

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования

МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)
(МАИ)

На правах рукописи



Торишный Роман Олегович

**АППРОКСИМАЦИЯ ВЕРОЯТНОСТНЫХ КРИТЕРИЕВ И ИХ
ПРОИЗВОДНЫХ ПРИ НЕПРЕРЫВНЫХ РАСПРЕДЕЛЕНИЯХ
СЛУЧАЙНЫХ ПАРАМЕТРОВ**

Специальность 2.3.1. — Системный анализ, управление и обработка информации,
статистика

Диссертация на соискание ученой степени
кандидата физико-математических наук

Научный руководитель:
кандидат физико-математических наук
Соболь Виталий Романович

Москва, 2023 год

Оглавление

Введение	4
1 Исследование гладкой аппроксимации функции вероятности	15
1.1 Постановка задачи и проблема вычисления градиента	15
1.2 Представление градиента функции вероятности в форме поверхностного интеграла	16
1.3 Идея гладкой аппроксимации функции вероятности для одномерного случайного вектора	19
1.4 Сигмоидальная функция и функция вероятности. Погрешность аппроксимации . . .	20
1.5 Свойства дельта-функции Дирака и производной сигмоидальной функции	29
1.6 Сходимость производной функции вероятности по компонентам стратегии	34
1.7 Вычисление аппроксимированной производной функции вероятности по компоненте стратегии	44
1.8 Аппроксимированная производная функции вероятности по уровню потерь	47
1.9 Квантильный критерий. Аппроксимация	51
1.10 Обобщение на многомерный случай	52
1.11 Выводы по главе 1	56
2 Гладкая аппроксимация вероятностных функций в теоретических задачах стохастического программирования	58
2.1 Построение аппроксимации ядра вероятностной меры	58
2.1.1 Двумерный случай	59
2.1.2 Примеры построения аппроксимированной внешней границы ядра для двумерных распределений	61
2.1.3 Трехмерный случай	63
2.1.4 Пример построения аппроксимированной внешней границы ядра для трехмерного случайного вектора	65
2.2 Аппроксимация вторых производных	66
2.2.1 Анализ сходимости производных функции квантили	69
2.2.2 Численные примеры	71
2.3 Исследование задачи стохастического программирования с вероятностным критерием и полиэдральной функцией потерь	75
2.3.1 Аппроксимация полиэдральной функции потерь	77
2.3.2 Аппроксимация вероятностного критерия	79
2.3.3 Алгоритм оптимизации	80
2.3.4 Пример двумерной задачи с полиэдральной функцией потерь и полиэдральной функцией ограничений	81
2.4 Выводы по главе 2	83
3 Применение гладкой аппроксимации функции вероятности в прикладных задачах стохастического программирования	85
3.1 Задача оптимизации площади взлетно-посадочной полосы.	85

3.2	Определение множества допустимых скоростей ветра	89
3.3	Проектирование системы водоснабжения	91
3.4	Задача формирования портфеля ценных бумаг с логарифмической функцией потерь	94
3.5	Выводы по главе 3	97
4	Программный комплекс для анализа задач стохастического программирования с вероятностным критерием	99
4.1	Структура и функциональные блоки программы	99
4.2	Алгоритм работы с программой и ее функциональные возможности	100
4.3	Примеры использования программного комплекса	103
4.3.1	Построение функции распределения и плотности нормального распределения	103
4.3.2	Задача формирования портфеля акций с минимальным риском	104
4.4	Выводы по главе 4	107
	Заключение	107
	Список использованных источников	109

Введение

Актуальность темы исследования.

В прикладных задачах управления и оптимизации нередко возникают проблемы, требующие принятия решений в условиях неопределенности. При построении и исследовании математических моделей систем одним из основных способов учета неопределенности является подход на основе аппарата теории вероятностей. При этом качество функционирования системы зачастую описывается некоторой функцией потерь, зависящей от управления (стратегии) и случайных величин. Значения этой функции являются случайными, что делает невозможной непосредственную минимизацию (или максимизацию) данной функции. В связи с этим для систем с неопределенностями стохастического характера необходимо рассматривать такие критериальные функции, которые в свертке с функцией потерь продуцируют детерминированные функции. Для этого используются вероятностные критериальные функции, причем основными и наиболее распространенными являются функция вероятности и функция квантили. Функция вероятности, взятая в качестве критериальной функции, соответствует вероятности непревышения функцией потерь какого-либо фиксированного уровня. Функция квантили, взятая в качестве критериальной, соответствует минимальному уровню потерь, который не превышает функцию потерь с вероятностью не ниже заданной. Использование вероятностных критериальных функций в оптимизации порождает задачи стохастического программирования, представляющие большую сложность в сравнении с детерминированными оптимизационными задачами. Обычно под словосочетанием «вероятностный критерий» подразумевают максимизацию функции вероятности, а под словосочетанием «квантильный критерий» - минимизацию функции квантили.

Стохастическое программирование возникло как раздел теории оптимизации в 50-х годах XX века и изначально было связано с исследованием задач с вероятностными ограничениями на множество допустимых стратегий. Такие задачи были рассмотрены в работах А. Чарнса, В. Купера [40, 41], А. Чарнса, В. Купера и Г. Симмондса [42], А. Прекопы [70–73], С. Гартски [50], Дж. Сенгупты [79], П. Калла [55], П. Калла и С. Уоллеса [56], Дж. Вайды [86]. Дальнейшее расширение стохастического программирования было обусловлено исследованием задач с вероятностными ограничениями на выполнение системы неравенств, которым были посвящены работы В. Аоя [88], В. Аоя и Р. Хенриона [87], В. Аоя и К. Сагастисабал [89], Б. Миллера и Х. Вагнера [66]. Стоит отдельно отметить монографию А. Прекопы [69], которая посвящена в том числе теоретическим основам построения численных методов стохастической оптимизации на основе

выпуклого программирования, а также работы М. Лежена [60, 61], Д. Денчевой, А. Прекопы и А. Ружински [46], П. Беральди и А. Ружински [38], В. Акоя, В. Берже, В. Оливейра и К. Сагастисабал [90], которые посвящены алгоритмам решения задач стохастического программирования с дискретным распределением случайной величины на основе предложенного А. Прекопой механизма p -эффективных точек.

На основе впервые рассмотренного С. Катаокой [58] квантильного критерия возник целый класс задач с критерием оптимизации в форме квантили и новая область стохастического программирования, посвященная исследованию таких задач. Развитие этой области обеспечили работы Э. Райка [21–23], А.И. Кибзуна и В.В. Малышева [11], А.И. Кибзуна и В.Ю. Курбаковско-го [10, 59], Р. Лепша [18, 19], Э. Тамм [29, 30, 81], К. Марти [64, 65], Ю.М. Ермольева, В.И. Норкина и Р. Ветса [47], В.И. Норкина и Н.В. Роевко [20] и С.П. Урясьева [36, 82–84]. Здесь также стоит особо отметить монографии А.И. Кибзуна и Ю.С. Кана [8, 57], в которых достаточно полно отражены результаты этих и других работ, посвященных решению задач квантильной оптимизации.

Основная сложность в решении задач стохастического программирования с вероятностными критериальными функциями состоит в невозможности сравнения напрямую значений критериальной функции при различных векторах стратегии ввиду её стохастической природы. Суть существующих способов решения таких задач, достаточно полно описанных в [8, 57], заключается в различных способах перехода к детерминированным задачам, которые либо являются преобразованием исходной задачи, либо их решение может быть приближенным решением или хотя бы оценкой решения исходной задачи. Основным методом, используемым в таких задачах, является доверительный метод, суть которого заключается в переходе к минимаксной задаче оптимизации, где максимум берется по некоторому доверительному множеству, подлежащему оптимизации. Сложность метода заключается в поиске «хорошего» доверительного множества для каждой отдельной задачи; этому посвящены работы А.И. Кибзуна и В.В. Малышева [11, 12], Ю.С. Кана и Р.Т. Суринова [9], А.И. Кибзуна и А.В. Наумова [14], А.И. Кибзуна, В.В. Малышева и Д.Э. Чернова [13]. Менее популярные методы — например, метод детерминированного эквивалента и методы статистической аппроксимации — обычно накладывают существенные ограничения на условия исходных задач. Такие методы для разного рода задач рассматривались в работах С.В. Иванова и А.И. Кибзуна [3], А.Н. Игнатова и А.И. Кибзуна [6], А.И. Кибзуна, А.В. Наумова и В.И. Норкина [15], А.И. Кибзуна и В.Р. Соболя [16], Д. Денчевой и А. Ружински [43, 44].

Большим потенциалом обладает область исследования задач стохастического программирования, связанная с нахождением значения градиента критериальной вероятностной функции, что в теории может позволить использовать градиентные методы поиска оптимума критериальной функции. Впервые выражение градиента функции вероятности в виде поверхностного ин-

теграла было получено Э. Райком [24]. Дальнейшее исследование в области получения строгих формул для вычисления градиента вероятностных функций обеспечено трудами А.И. Кибзуна и Г.Л. Третьякова [17], которыми было получено выражение для градиента функции вероятности в форме интеграла Лебега по поверхности, являющейся границей области реализаций случайного вектора, и трудами С.П. Урясьева [84, 85], в которых показано, что для некоторых частных случаев этот интеграл сводится как минимум к сумме объемного и поверхностного интегралов. Ввиду большой вычислительной сложности поверхностного интегрирования и ограничений на распределение случайных функций и величин, методика использования строго вычисленных значений градиента функции вероятности не была широко распространена.

В более поздних работах, сопряженных с данной темой, рассматриваются различные методики аппроксимации или оценки значения градиента функции вероятности. Работы Р. Хенриона [51], Р. Хенриона и В. Аkoa [87] посвящены оценке нормы градиента функции вероятности и получению верхних оценок субдифференциалов Кларка и Мордуховича в случае гауссового распределения параметров. В работе Г. Пфлюга и Х. Вайсхаупта [68] производная функции вероятности была приближена с помощью слабых производных и оценена с помощью сэмплирования, а в работе Ч. Ю и Д. Зельтермана [91] получены аппроксимации квантили и производных плотности распределения с использованием рядов Тейлора для нормального, экспоненциального и хи-квадрат распределений. В работе Дж. Гарньера, А. Омране и Ю. Раучди [49] были получены асимптотические формулы для первой и второй производной, вычисляемые с помощью метода Монте-Карло, для случая гауссового распределения параметров с малой дисперсией. В общем, существует большое количество методов оценки или аппроксимации градиента функции вероятности, но применение этих методов может быть крайне ограничено типом распределения случайных параметров, методы могут быть неприменимы для разных типов задач и они могут опираться на другие случайные механизмы для вычисления значений градиента.

Одним из более универсальных и общих подходов к получению приближенных значений производных вероятностных функций является исследование субградиентов функции вероятности и его возможные приближения. Одними из первых работ, косвенно освещающих данную проблему, были работы Э. Парзена [67] и М. Розенблатта [74]. В этих работах была предложена процедура приближения функции вероятности путем построения последовательности аппроксимирующих функций, полученных заменой индикаторной функции в интегральном представлении функции вероятности на некоторую гладкую функцию, причем Э. Парзенoм было предложено несколько таких аппроксимирующих функций, а М. Розенблаттом были рассмотрены вопросы сходимости аппроксимаций к оригинальной функции. Данное направление было развито в трудах Р. Леппа [62], Э. Тамм [80], К. Марти [65], в которых были исследованы субградиенты получив-

шихся аппроксимированных функций и рассмотрены механизмы решения задач стохастического программирования, например, с вероятностными ограничениями. Тема минимизации вероятностной функции, заданной в виде аппроксимации Парзена-Розенблатта, была также затронута в работе Ю.М. Ермольева, В.И. Норкина и Р. Уэтса [48], хоть и не рассматривалась целенаправленно.

Также стоит отметить иное направление исследований, смежное с рассматриваемой областью. В работах С.П. Урясьева [63, 85, 92] рассматривается т.н. «буферная вероятность» (англ. *buffered probability of exceedance, bPOE*). Суть этого термина заключается в задании «безопасного буфера» как по вероятности превышения случайной величиной заранее заданного значения, так и по величине разницы между случайной величиной и заданным значением. Для буферной вероятности были получены формулы для вычисления субградиентов и градиентов в некоторых случаях, но, ввиду специфики рассматриваемой вероятности, эти формулы не применимы к вероятностям превышения или непревышения некоторого значения случайной функцией.

Исследование именно субградиентов вероятностных функций обусловлено тем, что при представлении функции вероятности в интегральном виде возникает разрыв, обусловленный свойствами индикаторной функции, который можно обойти, используя рассматриваемое обобщение. При аппроксимации методом Парзена-Розенблатта, последовательность функций, аппроксимирующих субградиент функции вероятности, на самом деле является последовательностью обыкновенных градиентов, что обуславливает возможность использования данной методики при приближенном решении задач стохастической оптимизации градиентными методами. В диссертационной работе де-факто рассматривается один из частных случаев метода аппроксимации Парзена-Розенблатта, заключающийся в замене индикаторной функции на сигмоидальную функцию. Принципиальными отличиями рассматриваемого метода от ранее указанных являются параметризация гладкой аппроксимации индикатора, менее строгие условия, предъявляемые к целевой функции и распределению случайных величин и параметров, а также общность метода аппроксимации при вычислении приближенных производных функции вероятности разных порядков, в том числе и нулевого, а также независимость от предполагаемой переменной дифференцирования. Рассматриваемая аппроксимация также представляет относительно невысокую вычислительную сложность, так как аппроксимации функции вероятности и ее производных представляются в виде объемных интегралов. Более того, в работе также рассматриваются вопросы вычисления и сходимости аппроксимированных производных функции вероятности с порядком выше первого, ранее практически не исследованные ввиду большой вычислительной сложности.

Цели и задачи исследования

Целью диссертации является разработка теоретических основ, методов и алгоритмов решения задач стохастического программирования с вероятностными критериальными функциями

на основе гладкой аппроксимации вероятностных функций.

Для достижения выбранной цели необходимо решить следующие задачи:

- 1) Исследовать гладкую аппроксимацию функции вероятности и ее производных для случая гладкой функции потерь и абсолютно непрерывных распределений случайных параметров, а также исследовать сходимость аппроксимаций к точным значениям;
- 2) Провести обобщение метода гладкой аппроксимации на случай многомерных случайных величин и для производных функции вероятности высших порядков;
- 3) Разработать алгоритм построения аппроксимации границы ядра вероятностной меры с помощью гладкой аппроксимации функции вероятности специального вида для решения задач стохастического программирования, в частности, задач квантильной оптимизации;
- 4) Разработать алгоритм решения задачи максимизации функции вероятности в случае полиэдральной функции потерь и полиэдральной функции ограничений;
- 5) Разработать численные процедуры, реализующие вычисление и применение аппроксимаций функции вероятности и ее производных к решению известных прикладных задач стохастического программирования.

Объект и предмет исследования

Объектом исследования являются задачи стохастического программирования с вероятностным и квантильным критериальными функциями, а также задачи с ограничениями в форме функции вероятности. Предметом исследования являются алгоритмы приближенного (численного) решения задач с использованием гладкой аппроксимации критериальной функции или функций ограничений.

Методы исследования

В диссертации используются современные методы теории вероятностей, математического моделирования, стохастического программирования, математического анализа, системного анализа, теории оптимизации, математического программирования.

Достоверность результатов

Достоверность результатов обеспечивается строгостью математических формулировок и доказательств утверждений, а также подтверждением полученных теоретических результатов численными экспериментами.

Научная новизна

В диссертации впервые подробно исследован метод аппроксимации вероятностных функций на основе гладкого приближения функции Хевисайда сигмоидальной функцией. Доказана сходимость аппроксимированной функции вероятности к оригинальной и сходимость аппроксимированных производных функции вероятности к оригинальным производным. Разработаны но-

вые алгоритмы решения задач стохастического программирования с вероятностной критериальной функцией и/или ограничениями в форме функции вероятности, основанные на применении гладкой аппроксимации. Разработан новый алгоритм построения аппроксимации ядра вероятностной меры на основе аппроксимации функции вероятности особого вида. Разработан программный комплекс для построения аппроксимированной функции вероятности, а также её производных в одномерном и двумерном случаях.

Практическая ценность

Практическая ценность работы состоит в том, что её теоретические результаты могут служить основой для разработки программно-алгоритмического обеспечения решения прикладных задач в областях авиационной и ракетно-космической техники, оптимизации функционирования транспортных и логистических систем, систем распределения ресурсов, оптимального инвестирования. Эффективность предложенных алгоритмов и методов продемонстрирована на примере ряда прикладных задач.

Соответствие диссертации паспорту научной специальности

В диссертации применены методы системного анализа для исследования сложных технических и экономических систем, проведено исследование задач оптимизации, разработаны методы и алгоритмы их решения, в том числе численные алгоритмы решения с использованием ЭВМ (области исследования 1, 4, 5 специальности 2.3.1.).

Структура и объем диссертации

Диссертация содержит введение, 4 главы, заключение и список используемой литературы. Работа состоит из 117 страниц, включая 24 рисунка, 5 таблиц и список литературы, содержащий 92 наименования.

Содержание диссертации

Во введении приведен подробный обзор работ по теме диссертационного исследования и смежным областям, сформулированы цели и задачи исследования, выделены предмет и объект исследования, описаны и аргументированы ее научная новизна и практическая ценность, кратко изложены содержание и структура диссертации.

В первой главе приводятся обобщенные постановки задач оптимизации с вероятностной и квантильной критериальной функцией, относящейся к функции потерь, зависящей от случайной переменной и вектора стратегии. Далее кратко описана методика получения градиента функции вероятности напрямую в форме поверхностного интеграла и изложены проблемы такого подхода.

В основной части первой главы предлагается аппроксимация функции вероятности, основанная на замене функции Хевисайда в интегральном представлении функции вероятности на параметрическую сигмоидальную функцию. Далее формулируется утверждение о сходимости этой

функции к функции Хевисайда почти наверное в случае одномерной случайной величины. Формулируется утверждение о сходимости аппроксимированной функции вероятности к оригинальной функции вероятности при стремлении параметра сигмоиды к бесконечности. Это утверждение доказывается на основе утверждения о сходимости сигмоиды к функции Хевисайда без введения дополнительных условий. Путем оценки модуля разности аппроксимированной и оригинальной функции вероятности определяется верхняя оценка погрешности аппроксимации функции вероятности, которая стремится к нулю при стремлении параметра сигмоиды к бесконечности.

Далее рассматриваются свойства производной сигмоидальной функции и свойства дельта-функции Дирака как производной функции Хевисайда в смысле обобщенных функций. Формулируется и доказывается утверждение о слабой сходимости последовательности производных сигмоиды с нарастающим параметром к дельта-функции Дирака в смысле обобщенных функций. Таким образом, производная сигмоиды обладает фильтрующим свойством, аналогичным таковому у дельта-функции. Формулируется и доказывается утверждение о сходимости производной аппроксимированной функции вероятности по компонентам вектора стратегии к производной оригинальной функции вероятности. Вычисление производной аппроксимированной функции вероятности сводится к дифференцированию математического ожидания сигмоидальной функции, однако в работе также получено более удобное выражение для этой производной, представляющее собой интеграл от производной сигмоидальной функции. Далее формулируется и доказывается утверждение о сходимости производной аппроксимированной функции по уровню потерь к производной оригинальной функции вероятности и аналогичным образом получается упрощенное выражение для вычисления аппроксимированной производной. Приводятся выражения для аппроксимации производных по компонентам вектора стратегии и уровню доверительной вероятности для функции квантили.

Далее следует обобщение аппроксимации на случай многомерной случайной величины. Определяются дополнительные условия, накладываемые на функцию потерь для применимости аппроксимации в многомерном случае. Доказываются аналоги утверждений о сходимости функции вероятности и ее производных по компонентам вектора стратегии и уровню потерь к соответствующим оригинальным функциям для многомерного случая.

Во второй главе рассматриваются возможные применения гладкой аппроксимации функции вероятности при исследовании теоретических задач стохастического программирования.

В первом разделе в общем виде рассматривается задача аппроксимации внешней границы α -ядра вероятностной меры для абсолютно непрерывного распределения. Вводятся определения α -доверительных множеств и α -ядра, а также рассматривается понятие регулярного α -ядра вместе с теоремами и утверждениями, определяющими необходимые и достаточные условия регулярно-

сти. Рассматривается построение внешней границы α -ядра в двумерном случае как построение изокванты функции вероятности особого вида. Построение изокванты, в свою очередь, сводится к решению дифференциального уравнения, что ведет к вычислению аппроксимированных производных функции вероятности особого вида. В предположении, что α -ядро содержит нулевую точку, совершается переход к полярной системе координат и определяется алгоритм построения внешней границы путем поворота радиус-вектора на каждой итерации на фиксированный шаг и приближенного вычисления значения вектора градиента функции вероятности с помощью описанной аппроксимации. В качестве иллюстрации приводятся результаты построения аппроксимации внешней границы ядра для равномерного распределения на квадрате и двумерного нормального распределения, совпадающие с известными результатами. Аналогично рассматривается трехмерный случай, предлагается алгоритм итерационного построения аппроксимированной поверхности внешней границы α -ядра и приводятся результаты построения границы α -ядра для трехмерного нормального распределения, соответствующие известным результатам.

Во втором разделе рассматриваются свойства и возможные способы применения производных второго порядка функции вероятности и функции квантили. Приводится краткий обзор работ, посвященных или сопряженных с темой выпуклости и квазивыпуклости функции вероятности и вопросом нахождения вторых производных. Далее приводятся выражения для аппроксимации смешанной второй производной по компонентам вектора стратегии для функции вероятности и функции квантили. Производится анализ сходимости гладкой аппроксимации вторых производных функции квантили к точным значениям для трех случаев, в которых для задачи стохастического программирования существует детерминированный эквивалент: случай сепарабельной функции потерь; функции потерь, представленной произведением зависящей только от вектора стратегии функции и зависящей только от случайного вектора функции; случай одномерного случайного вектора и монотонной функции потерь. Проведены численные эксперименты, подтверждающие близость аппроксимированной и исходной производных для двух случаев билинейной функции потерь и двух случаев квадратичной функции потерь.

В третьем разделе исследуется задача стохастического программирования с вероятностным критерием и полиэдральной функцией потерь. Приводится постановка задачи и описываются основные проблемы прямого применения описываемого метода аппроксимации. Для устранения неопределенности производной целевой функции в точках пересечения функций под знаком максимума рассматривается однопараметрическое преобразование гладкого максимума, представляющее собой гладкую дифференцируемую аппроксимацию функции максимума и основывающееся на применении экспоненциальной функции. Доказывается сходимость функции вероятности, полученной при помощи этого преобразования, к оригинальной функции при стремлении пара-

метра гладкого максимума к бесконечности. Далее рассматривается аппроксимация вероятностного критерия путем замены функций Хевисайда, относящихся к функции потерь и функции ограничений, на сигмоидальные функции. Выводятся выражения для вычисления приближенных значений производной функции вероятности. Приводится модифицированный алгоритм градиентного спуска для приближенного решения исходной задачи стохастического программирования. В качестве иллюстрации по описанному алгоритму решается задача стохастического программирования с известными полиэдральной функцией потерь и функцией ограничений, а также проводится сравнение с решениями, полученными другими методами.

В третьей главе рассматривается использование гладкой аппроксимации функции вероятности и функции квантили для решения некоторых известных практических задач стохастического программирования. Рассматривается задача оптимизации площади взлетно-посадочной полосы при ограничении на вероятность успешность посадки; решение задачи сравнивается с результатами, полученными с использованием доверительного и квазиградиентного методов. Рассматривается задача определения допустимого множества векторов скоростей ветра, определенного априори и гарантирующего безопасную посадку с некоторой гарантирующей вероятностью. Эта задача решается путем построения изокванты функции вероятности, что, в свою очередь, приводит к последовательному вычислению производных функции вероятности; построенное множество неотлично от выпуклого и обладает более простым построением в сравнении с известным результатом. Далее рассматривается задача проектирования системы водоснабжения в засушливом регионе с минимальной стоимостью и гарантированным с некоторой вероятностью удовлетворением спроса на пресную воду. Также рассматривается задача формирования портфеля ценных бумаг по вероятностному критерию с логарифмической функцией потерь.

В четвертой главе описывается программный комплекс, осуществляющий построение аппроксимированных функций вероятности и ее производных. Описаны основные модули программы и ее возможности, приведена краткая инструкция по взаимодействию с программой. В качестве иллюстрации работы программы рассмотрен тривиальный пример построения функции распределения и плотности нормального распределения, а также решена задача формирования портфеля с минимальным риском на основе реально наблюдаемых цен закрытия акций трех российских компаний.

В заключении приведены основные научные результаты, полученные автором работы.

Основные положения, выносимые на защиту

- 1) Доказана теорема о сходимости аппроксимированной функции вероятности к исходной и теоремы о сходимости производной аппроксимированной функции вероятности к производной функции вероятности по компонентам вектора стратегии и уровню потерь при воз-

растании параметра сигмоиды в случае одномерной случайной величины [25, 77, 78].

- 2) Доказаны аналоги теорем о сходимости аппроксимированной функции вероятности и ее производных к исходным функциям в случае многомерной случайной величины и получены аппроксимационные выражения в форме объемных интегралов. На их основе проведены численные решения 4 известных задач стохастического программирования [31, 75, 77, 78].
- 3) Разработан новый алгоритм решения задачи стохастического программирования с вероятностным критерием, полиэдральной функцией потерь и полиэдральной функцией ограничений на основе рассматриваемой аппроксимации и аппроксимации гладкого минимума [76].
- 4) Разработан новый алгоритм аппроксимации внешней границы альфа-ядра вероятностной меры на основе аппроксимации функции вероятности [77].
- 5) Разработан и зарегистрирован программный комплекс для построения графиков и поверхностей аппроксимированной функции вероятности, ее первых и вторых производных по компонентам вектора стратегии [28, 32].

Апробация работы

Основные результаты диссертации докладывались на следующих научных конференциях и семинарах:

- 1) 10-я Традиционная молодежная школа "Управление, информация и оптимизация" (Москва, 2018);
- 2) 17-я Международная конференция "Авиация и космонавтика" (Москва, 2018);
- 3) XLV Международная молодежная научная конференция "Гагаринские чтения" (Москва, 2019);
- 4) XXVI Международная научная конференция студентов, аспирантов и молодых ученых "Ломоносов" (Москва, 2019);
- 5) Международная научная конференция "Системный анализ, управление и навигация" (Евпатория, 2019);
- 6) 19-я Международная конференция "Авиация и космонавтика" (Москва, 2020);
- 7) International Conference "Mathematical Optimization Theory and Operations Research (MOTOR 2021)" (Irkutsk, Baikal, 2021);
- 8) International Conference "Mathematical Optimization Theory and Operations Research (MOTOR 2022)" (Petrozavodsk, Karelia, Russia, 2022);
- 9) Научный семинар кафедры теории вероятностей и компьютерного моделирования Московского авиационного института (под рук. проф. Кибзуна А.И.).

Работа поддержана грантом РФФИ № 20-31-90035.

Публикации

Основные результаты диссертационной работы опубликованы в 12 научных работах, из которых 3 опубликованы в периодических изданиях, входящих в перечень ВАК и цитируемых международными базами Web of Science и Scopus [25, 75, 78]; 2 опубликованы в периодических изданиях, входящих в перечень ВАК [31, 32]; 2 опубликованы в периодических изданиях, цитируемых международными базами Web of Science и Scopus [76, 77]; 5 опубликованы в качестве тезисов докладов в трудах российских и международных конференций [26, 27, 33–35]. На основе результатов диссертационной работы разработана и зарегистрирована программа для ЭВМ [28].

Личный вклад

В совместных публикациях с научным руководителем Соболев В.Р. руководителю принадлежат постановки некоторых задач и идеи их решения.

Благодарности

Автор выражает глубокую признательность научному руководителю Соболю В.Р., преподавателям и ассистентам кафедры «Теория вероятностей и компьютерное моделирование» Кибзуну А.И., Наумову А.В., Иванову С.В., Игнатову А.Н., Борисову А.В., Ибрагимову Д.Н. за ценные критические замечания, позволившие улучшить качество диссертационной работы.

1. Исследование гладкой аппроксимации функции вероятности

1.1 Постановка задачи и проблема вычисления градиента

Пусть $(\Omega, \mathcal{F}, \mathbb{P})$ — полное вероятностное пространство, где Ω — пространство элементарных событий, \mathcal{F} — σ -алгебра, определенная на этом пространстве, \mathbb{P} — вероятностная мера. Пусть X — случайный вектор размерности n с носителем распределения $G \subseteq \mathbb{R}^n$, заданный на этом пространстве и имеющий плотность распределения $f : \mathbb{R}^n \rightarrow \mathbb{R}^1$. Пусть также задана функция потерь $\Phi(u, x) : U \times \mathbb{R} \rightarrow \mathbb{R}$, зависящая от вектора стратегии u , принадлежащего непустому множеству допустимых стратегий $U \subseteq \mathbb{R}^m$, и от реализации x случайного вектора X . Также пусть задан произвольный максимально допустимый уровень потерь φ .

Целью задачи является поиск оптимальной стратегии, которая максимизирует вероятность того, что уровень потерь не превысит заранее заданный уровень φ . В такой задаче в качестве критериальной функции выступает функция вероятности, заданная следующим образом:

$$P_\varphi(u) \triangleq \mathbb{P} \{ \Phi(u, X) \leq \varphi \}, \quad (1.1.1)$$

а сама задача представляет собой максимизацию этой функции:

$$P_\varphi(u) \rightarrow \max_{u \in U}. \quad (1.1.2)$$

Существует также альтернативная постановка, заключающаяся в поиске оптимальной стратегии, доставляющей минимум функции квантили для некоторого фиксированного уровня доверия $\alpha \in (0; 1)$. Это означает, что оптимальной стратегии соответствует минимально возможный уровень потерь φ , и этот уровень не будет превышен с вероятностью не меньшей, чем α . В этом случае в качестве критериальной функции выбирается функция квантили, заданная следующим образом:

$$\varphi_\alpha(u) \triangleq \min \{ \varphi : P_\varphi(u) \geq \alpha \}, \quad (1.1.3)$$

а задача представляет собой минимизацию функции квантили:

$$\varphi_\alpha(u) \rightarrow \min_{u \in U}. \quad (1.1.4)$$

Представленные постановки задач описывают одношаговую модель принятия решения в условиях неопределенности, моделируемой случайным вектором X . При этом само принятие решения, то есть выбор стратегии u , производится априори до реализации случайного вектора, опи-

раясь исключительно на знание закона распределения случайного вектора, исходя из условия минимума (или максимума, в зависимости от формулировки) критерия оптимизации.

Также стоит отметить, что в любой из постановок задач могут быть дополнительные ограничения на значения переменных. Так как к вероятностным ограничениям также применима аппроксимация, описанная далее, а нестохастические ограничения лишь сужают область допустимых стратегий, и в обоих случаях структура целевой функции от ограничений не зависит, на данном шаге рассмотрения задачи дополнительные ограничения опускаются.

Как было отмечено ранее, для задач такого рода представляется интересной разработка унифицированного метода решения, не зависящего от вида целевой функции или распределения случайного вектора. Наиболее универсальные методы решения задач оптимизации - методы оптимизации, использующие градиент оптимизируемой функции. При этом на практике метод, например, градиентного спуска не применяется для решения задач вероятностной оптимизации по причине трудности прямого вычисления градиента функции вероятности. Известные методы прямого вычисления градиента функции вероятности описаны в пункте 1.2. Предлагаемая аппроксимация, в том числе упрощающая вычисление градиента функции вероятности, описана в пунктах 1.3-1.10, а также секции 2.2.

1.2 Представление градиента функции вероятности в форме поверхностного интеграла

Впервые формула для градиента функции вероятности была получена Райком [24] в форме поверхностного интеграла в смысле Римана. При этом функция потерь предполагалась векторной $\Phi(u, x) : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^k$.

Для целей данного раздела введем следующие обозначения:

$$\begin{aligned}
 P(u) &\triangleq P \{ \Phi(u, X) \leq 0 \}, \\
 V(u) &\triangleq \{ x \in \mathbb{R}^n : \Phi(u, x) \leq 0 \}, & S(u) &\triangleq \partial V(u), \\
 V_G(u) &\triangleq V(u) \cap G, & S_G(u) &\triangleq S(u) \cap G, \\
 V_U &\triangleq \text{cl} \left(\bigcup_{u \in U} V(u) \right), & \Delta V_U &\triangleq \text{cl} \left(\bigcup_{u \in U} \partial V(u) \right), \\
 \Delta V_{UG} &\triangleq \Delta V_U \cap G, & N_U &\triangleq \text{cl}(U) \times V_U, \\
 N_{UG} &\triangleq \text{cl}(U) \times V_{UG}, & \Delta N_{UG} &\triangleq \text{cl}(U) \times \Delta V_{UG},
 \end{aligned}$$

где $\text{cl}(\cdot)$ - замыкание множества. Далее предполагается, что $S(u)$ является $(n-1)$ -мерной поверхностью объемного множества $V(u)$. Поэтому множество ΔV_U содержит все возможные поверхности $S(u)$, а множество V_U — все возможные множества $V(u)$.

В [8] приведена модификация утверждения из [24] для случая $k = 1$, что соответствует рассмотренной ранее постановке задачи.

Утверждение 1. [8] Пусть выполнены следующие условия:

- 1) градиенты $\nabla_u \Phi(u, x)$, $\nabla_x \Phi(u, x)$ непрерывны на ΔN_{UG} ;
- 2) $\nabla_u \Phi(u, x) > 0$ покомпонентно на ΔN_{UG} ;
- 3) $\|\nabla_x \Phi(u, x)\| > 0$ на ΔN_{UG} .

Тогда функция вероятности $P(u)$ дифференцируема почти всюду на U , а ее градиент может быть представлен как поверхностный интеграл (в смысле Лебега):

$$\nabla_u P(u) = - \int_{S_G(u)} \frac{\nabla_u \Phi(u, x)}{\|\nabla_x \Phi(u, x)\|} f(x) dS. \quad (1.2.1)$$

На основании этой теоремы нельзя проверить дифференцируемость функции вероятности в конкретной точке $u_0 \in U$. Это можно сделать, наложив дополнительные ограничения. Сформулируем другое утверждение из [24].

Утверждение 2. [24] Пусть выполнены следующие условия:

- 1) множество $S(u)$ равномерно ограничено в некоторой окрестности точки $u_0 \in \mathbb{R}^m$;
- 2) вектор-функции $\nabla_u \Phi(u, x)$ и $\nabla_x \Phi(u, x)$ непрерывны на $\mathbb{R}^m \times \mathbb{R}^n$;
- 3) плотность $f(x)$ непрерывна на \mathbb{R}^n ;
- 4) вектор-функция $\nabla_x \Phi(u, x)$ удовлетворяет условию Липшица по x для всех $u \in U$;
- 5) $\|\nabla_x \Phi(u, x)\| > 0$ на поверхности $S(u_0)$;
- 6) существует поверхностный интеграл $\int_{S(u_0)} dS$.

Тогда функция вероятности $P(u)$ дифференцируема в точке u_0 , а ее градиент может быть представлен как поверхностный интеграл:

$$\nabla_u P(u) = - \int_{S(u_0)} \frac{\nabla_u \Phi(u, x)}{\|\nabla_x \Phi(u, x)\|} f(x) dS.$$

Существуют и другие утверждения о представлении градиента функции вероятности в форме поверхностного интеграла, отличающиеся набором условий.

Утверждение 3. [8] Пусть выполнены следующие условия:

- 1) множество ΔV_{UG} ограничено;
- 2) вектор-функции $\nabla_u \Phi(u, x)$ и $\nabla_x \Phi(u, x)$ непрерывны на ΔN_{UG} ;

3) плотность $f(x)$ непрерывна на ΔV_{UG} и для всех $u \in U$ выполняется условие:

$$\text{mes}_{n-1}(\{x \in \delta G : f(x) > 0\} \cap S(u)) = 0,$$

где $\text{mes}_{n-1}(\cdot)$ — поверхностная мера Лебега;

4) $\|\nabla_x \Phi(u, x)\| > 0$ на множестве $S_G(u)$ для всех $u \in U$.

Тогда функция вероятности $P(u)$ дифференцируема для всех $u \in U$ и ее градиент имеет вид (1.2.1)

В [8] также приводится утверждение о представлении градиента функции вероятности в форме объемного интеграла, изначально представленное в [84] для векторной функции потерь.

Утверждение 4. [8] Пусть выполнены следующие условия:

- 1) множество N_U ограничено;
- 2) вектор-функции $\nabla_u \Phi(u, x)$ и $\nabla_x \Phi(u, x)$ непрерывны на N_U ;
- 3) существует непрерывная функция $H(u, x)$, являющаяся решением следующего матричного уравнения:

$$H(u, x) \nabla_x \Phi(u, x) + \nabla_u \Phi(u, x) = 0, \quad (1.2.2)$$

для которой матричная производная $\nabla_x H(u, x)$ непрерывна на N_U ;

- 4) градиент плотности $\nabla_x f(x)$ непрерывен на V_U ;
- 5) $\|\nabla_x \Phi(u, x)\| \geq \Delta > 0$ на $G(u)$.

Тогда функция вероятности $P(u)$ дифференцируема в точке $u_0 \in U$ и

$$\nabla_u P(u_0) = \int_{V_G(u_0)} \text{div}_x \{H(u_0, x) f(x)\} dx.$$

Отметим, что проверка условий в утверждениях 1-4, а также определение множеств $V(u)$, $S(u)$ и иных связанных вспомогательных множеств, фигурирующих в утверждениях, является очень трудоемкой задачей, особенно в прикладных задачах со сложной функцией потерь. Отметим также, что в общем случае уравнение (1.2.2) имеет множество решений, то есть формула градиента в утверждении 4 порождает семейство градиентов в случае, когда функция потерь является вектор-функцией. Для исследуемого в работе случая в [84] было предложено одно аналитическое решение уравнения (1.2.2):

$$H(u, x) = -\|\nabla_x \Phi(u, x)\|^{-2} \nabla_u \Phi(u, x) |\nabla_x \Phi(u, x)|^T,$$

но практическое применение этого результата также будет ограничено необходимостью проверки условий 1-5 в утверждении 4.

1.3 Идея гладкой аппроксимации функции вероятности для одномерного случайного вектора

Для описания аппроксимации функции вероятности и ее производных рассмотрим функцию вероятности (1.1.1) в случае, когда вектор X является одномерным, т.е. X является случайной величиной. Согласно [37], функция вероятности может быть представлена в виде:

$$P_\varphi(u) = \mathbb{P} \{ \Phi(u, X) \leq \varphi \} = \mathbb{M} [I_\varphi(u, X)] = \int_{-\infty}^{+\infty} I_\varphi(u, x) f(x) dx, \quad (1.3.1)$$

где $I_\varphi(u, x)$ - индикаторная функция, заданная следующим образом:

$$I_\varphi(u, x) = \begin{cases} 1, & \Phi(u, x) \leq \varphi, \\ 0, & \Phi(u, x) > \varphi. \end{cases}$$

Отметим, что изначально область интегрирования в интеграле (1.3.1) есть носитель распределения, т.е. $G \subseteq \mathbb{R}$. В случае, если $G \neq \mathbb{R}$, то плотность распределения $f(x)$ доопределяется нулем на области $\mathbb{R} \setminus G$, что не оказывает влияния на дальнейшие рассуждения и позволяет интегрировать по всей области \mathbb{R} .

Известно, что индикаторную функцию можно записать как функцию Хевисайда:

$$\Theta(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Стоит отметить, что для полной взаимозаменяемости индикаторной функции и функции Хевисайда последняя доопределяется в нуле указанным выше образом, т.е. $\Theta(0) = 1$.

Таким образом, функция вероятности может быть записана как:

$$P_\varphi(u) = \int_{-\infty}^{+\infty} \Theta(\varphi - \Phi(u, x)) f(x) dx. \quad (1.3.2)$$

Для получения формулы дифференцирования правой части уравнения (1.3.2) по переменной φ или компонентам стратегии u необходима как минимум непрерывность подынтегральной функции вместе с ее производными по этим переменным. Функция Хевисайда же разрывна и не имеет производной в смысле стандартных функций в точке нуля своего аргумента. Если рассматривать функцию Хевисайда как обобщенную функцию, то ее производной в этом смысле является дельта-функция Дирака [2], но использование аппарата обобщенных функций в дальнейших аналитических преобразованиях, а именно дифференцировании и интегрировании в стандартном

смысле, невозможно. Поэтому логичным шагом видится замена функции Хевисайда в исходном определении функции вероятности на аппроксимирующую функцию, обладающую некоторыми свойствами:

- близость (сходимость) аппроксимирующей функции к функции Хевисайда, что в том числе значит стремление к нулю ошибки аппроксимации;
- непрерывность аппроксимирующей функции вместе с ее производными;
- близость производной аппроксимирующей функции к производной функции Хевисайда;
- близость аппроксимации функции вероятности к оригинальной функции;
- близость аппроксимации производной функции вероятности к производной оригинальной функции;
- выполнение этих же свойств для аппроксимации квантильной функции как критериальной функции.

В качестве аппроксимирующей функции, обладающей отмеченными свойствами, взята сигмоидальная функция того же аргумента, зависящая от параметра $\theta > 0$:

$$\Theta(\varphi - \Phi(u, x)) \approx S_{\theta}(\varphi - \Phi(u, x)) = \frac{1}{1 + e^{-\theta(\varphi - \Phi(u, x))}}. \quad (1.3.3)$$

Таким образом, в работе рассматриваются свойства этой аппроксимирующей функции, а также свойства полученной аппроксимации функции вероятности

$$P_{\varphi}^{\theta}(u) = \int_{-\infty}^{+\infty} S_{\theta}(\varphi - \Phi(u, x))f(x)dx, \quad (1.3.4)$$

вместе с ее производными.

1.4 Сигмоидальная функция и функция вероятности. Погрешность аппроксимации

Сигмоидальная функция определяется как:

$$S_{\theta}(t) = \frac{1}{1 + e^{-\theta t}}, \quad (1.4.1)$$

где θ — большое положительное число. График сигмоидальной функции с параметром $\theta = 5$ представлен на рисунке 1.1.

Сигмоидальная функция обладает непрерывностью и гладкостью, а параметр θ влияет на степень кривизны функции в окрестности нуля: чем выше параметр, тем круче график функции в окрестности нуля.

Рассмотрим некоторые утверждения и теоремы, необходимые для дальнейшего описания аппроксимации и дальнейших доказательств ее свойств.

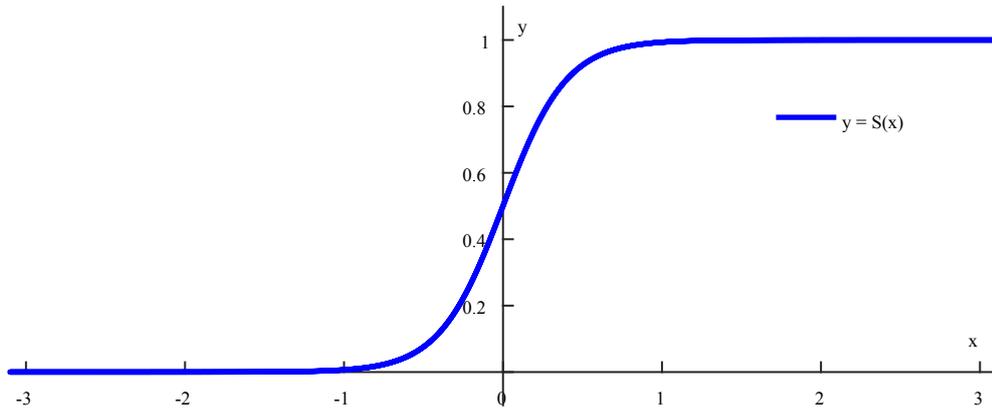


Рисунок 1.1. Сигмоидальная функция

Теорема 1. Пусть X — непрерывная случайная величина со значениями $x \in \mathbb{R}$ на полном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$. Пусть задано непустое множество $U \subseteq \mathbb{R}^m$ допустимых стратегий u и задана непрерывная строго кусочно-монотонная по x при любом $u \in U$ функция потерь $\Phi(u, x) : U \times \mathbb{R} \rightarrow \mathbb{R}$. Также пусть задан уровень потерь φ . Тогда для сигмоидальной функции $S_\theta(\cdot)$ и функции Хевисайда $\Theta(\cdot)$ верно следующее:

$$S_\theta(\varphi - \Phi(u, X)) \xrightarrow{\text{п.п.}} \Theta(\varphi - \Phi(u, X)) \quad \text{при } \theta \rightarrow +\infty \quad \forall u \in U. \quad (1.4.2)$$

Доказательство. Рассмотрим функции Хевисайда и сигмоиды при различных значениях x случайной величины X .

1) Пусть $x^* \in R^+$, $R^+ = \{x \mid \Phi(u, x) > \varphi\}$. Тогда

$$\Theta(\varphi - \Phi(u, x^*)) = 1, \quad (1.4.3)$$

$$S_\theta(\varphi - \Phi(u, x^*)) = \frac{1}{1 + e^{-\theta(\varphi - \Phi(u, x^*))}} \rightarrow 1 \quad \text{при } \theta \rightarrow +\infty. \quad (1.4.4)$$

Отсюда следует поточечная сходимость сигмоиды к функции Хевисайда на всей области R^+ :

$$S_\theta(\varphi - \Phi(u, x^*)) \rightarrow \Theta(\varphi - \Phi(u, x^*)) \quad \text{при } \theta \rightarrow +\infty \quad \forall x^* \in R^+. \quad (1.4.5)$$

2) Пусть $x^* \in R^-$, $R^- = \{x \mid \Phi(u, x) < \varphi\}$. Тогда по аналогии

$$\Theta(\varphi - \Phi(u, x^*)) = 0, \quad (1.4.6)$$

$$S_\theta(\varphi - \Phi(u, x^*)) = \frac{1}{1 + e^{-\theta(\varphi - \Phi(u, x^*))}} \rightarrow 0 \quad \text{при } \theta \rightarrow +\infty, \quad (1.4.7)$$

$$S_\theta(\varphi - \Phi(u, x^*)) \rightarrow \Theta(\varphi - \Phi(u, x^*)) \text{ при } \theta \rightarrow +\infty \quad \forall x^* \in R^-, \quad (1.4.8)$$

т.е. выполняется поточечная сходимость на R^- .

3) Пусть $x^0 \in R^0$, $R^0 = \{x \mid \Phi(u, x) = \varphi\}$. Тогда

$$\Theta(\varphi - \Phi(u, x^0)) = 1, \quad S_\theta(\varphi - \Phi(u, x^0)) = \frac{1}{2}. \quad (1.4.9)$$

Ввиду строгой кусочной монотонности $\Phi(u, x)$, любой заданный уровень потерь φ и значения функции $\Phi(u, x)$ равны друг другу максимум на множестве изолированных точек. Тогда множество R^0 — дискретное множество, состоящее из не более чем счетного числа точек. Поскольку мера \mathbb{P} порождена непрерывной величиной X , то:

$$\mathbb{P}\{R^0\} = 0. \quad (1.4.10)$$

Это означает, что поточечная сходимость функции сигмоиды к функции Хевисайда не выполняется только на множестве R^0 меры нуль. С учетом пунктов 1 и 2 доказательства, получаем, что имеет место сходимость функции сигмоиды к функции Хевисайда почти всюду на $R^+ \cup R^- \cup R^0$, что эквивалентно сходимости почти наверное на \mathbb{R} . \square

Далее рассмотрим функцию вероятности

$$P_\varphi(u) = \mathbf{M} [\Theta(\varphi - \Phi(u, X))] = \int_{-\infty}^{+\infty} \Theta(\varphi - \Phi(u, x)) f(x) dx \quad (1.4.11)$$

и ее аппроксимацию (1.3.4)

$$P_\varphi^\theta(u) = \mathbf{M} [S_\theta(\varphi - \Phi(u, X))] = \int_{-\infty}^{+\infty} S_\theta(\varphi - \Phi(u, x)) f(x) dx. \quad (1.4.12)$$

Теорема 2. Пусть X — непрерывная случайная величина со значениями $x \in \mathbb{R}$ и кусочно-непрерывной плотностью $f(x)$ на полном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$. Пусть задано непустое множество $U \subseteq \mathbb{R}^m$ допустимых стратегий u , задана непрерывная строго кусочно-монотонная по x при любом $u \in U$ функция потерь $\Phi(u, x) : U \times \mathbb{R} \rightarrow \mathbb{R}$, а также задан уровень потерь φ . Пусть также выражениями (1.4.11) и (1.3.4) заданы функции $P_\varphi^\theta(u)$ и $P_\varphi(u)$. Тогда верно

$$P_\varphi^\theta(u) \rightarrow P_\varphi(u) \text{ при } \theta \rightarrow +\infty \quad \forall u \in U. \quad (1.4.13)$$

Доказательство. В соответствии с Теоремой 1:

$$S_\theta(\varphi - \Phi(u, X)) \xrightarrow{п.п.} \Theta(\varphi - \Phi(u, X)) \quad \text{при } \theta \rightarrow +\infty. \quad (1.4.14)$$

Сигмоидальная функция мажорируется единицей, т.е. $|S_\theta(\varphi - \Phi(u, x))| \leq 1$ при любых значениях θ и аргументов. В силу этого на каждом из промежутков непрерывности $f(x)$ выполняется

$$S_\theta(\varphi - \Phi(u, x)) \cdot f(x) \leq f(x), \quad (1.4.15)$$

что в соответствии со свойствами интегралов влечет сходимость интеграла (1.4.12). Таким образом, математическое ожидание $\mathbb{M}[S_\theta(\varphi - \Phi(u, X))]$ существует, и по теореме Лебега о мажорируемой сходимости получим

$$\lim_{\theta \rightarrow +\infty} \mathbb{M}[S_\theta(\varphi - \Phi(u, X))] = \mathbb{M}[\Theta(\varphi - \Phi(u, X))], \quad (1.4.16)$$

что с учетом (1.4.11) и (1.3.4) эквивалентно

$$\lim_{\theta \rightarrow +\infty} P_\varphi^\theta(u) = P_\varphi(u), \quad (1.4.17)$$

что, в свою очередь, эквивалентно доказываемому утверждению. \square

Теперь получим выражение для верхней оценки погрешности аппроксимации. Для этого рассмотрим модуль разности функции вероятности $P_\varphi(u)$ и ее аппроксимации $P_\varphi^\theta(u)$. В соответствии со свойствами определенного интеграла:

$$\begin{aligned} |P_\varphi(u) - P_\varphi^\theta(u)| &= \left| \int_{-\infty}^{+\infty} (\Theta(\varphi - \Phi(u, x)) - S_\theta(\varphi - \Phi(u, x))) f(x) dx \right| \leq \\ &\leq \int_{-\infty}^{+\infty} |\Theta(\varphi - \Phi(u, x)) - S_\theta(\varphi - \Phi(u, x))| f(x) dx \quad (1.4.18) \end{aligned}$$

Рассмотрим подынтегральную функцию правой части этого выражения. Множитель $f(x)$ — кусочно-непрерывная положительная ограниченная функция, не влияющая на знаки неравенств или значения другого сомножителя. Поэтому основная содержательная часть состава погрешности зависит от модуля разности функций Хевисайда и сигмоиды. Таким образом, далее будет рассматриваться функция $g(\varphi - \Phi(u, x)) = |\Theta(\varphi - \Phi(u, x)) - S_\theta(\varphi - \Phi(u, x))|$.

Разобьем пространство \mathbb{R} аргумента x на четыре непересекающиеся области следующим

образом:

$$\chi_1^+(u) = \{x : 0 \leq \varphi - \Phi(u, x) < 1\}, \quad (1.4.19)$$

$$\chi_1^-(u) = \{x : -1 < \varphi - \Phi(u, x) < 0\}, \quad (1.4.20)$$

$$\chi_2^+(u) = \{x : \varphi - \Phi(u, x) \geq 1\}, \quad (1.4.21)$$

$$\chi_2^-(u) = \{x : \varphi - \Phi(u, x) \leq -1\}. \quad (1.4.22)$$

Рассматриваемая погрешность, представленная интегралом по всему пространству \mathbb{R} , оценивается по сумме интегралов на обозначенных выше областях. Стоит отметить, что границы ранее введенных областей зависят от значения вектора стратегии u , т.е. каждому u соответствуют свои области $\chi_1^+(u)$, $\chi_1^-(u)$, $\chi_2^+(u)$ и $\chi_2^-(u)$. Так как все рассуждения ниже справедливы при любом значении $u \in U$ и переносятся на любые удовлетворяющие определению области без изменений, то в целях упрощения записи зафиксируем некоторое $u \in U$ и опустим зависимости рассматриваемых областей от вектора стратегии.

Отметим, что ввиду непрерывности $\Phi(u, x)$ данные области представляют собой объединение непересекающихся интервалов и не более чем счетного числа точек. Пусть область χ состоит из не более чем счетного количества непересекающихся интервалов (x_j, x_{j+1}) ненулевой меры. С учетом этого введем обозначение:

$$\int_{\chi} (\cdot) dx = \sum_j \int_{x_j}^{x_{j+1}} (\cdot) dx, \quad (1.4.23)$$

Это означает, что в обозначении (1.4.23) «интеграл по области» χ представляет собой сумму не более чем счетного числа интегралов по не пересекающимся интервалам, составляющим эту область. Поскольку интеграл аддитивен, а все выкладки далее применимы сразу к любому интервалу (x_j, x_{j+1}) фиксированной области χ , то выкладки справедливы для суммы всех интегралов по всем областям, для которой и введено обозначение (1.4.23).

Перейдем к рассмотрению областей.

1) Рассмотрим область $\chi_2^+ = \{x : \varphi - \Phi(u, x) \geq 1\}$. Функция Хевисайда внутри функции $g(\varphi - \Phi(u, x))$ на этой области равна единице. Тогда

$$g(\varphi - \Phi(u, x)) = |1 - S_{\theta}(\varphi - \Phi(u, x))| = 1 - \frac{1}{1 + e^{-\theta \cdot (\varphi - \Phi(u, x))}}. \quad (1.4.24)$$

Из условия задания области следует цепочка неравенств:

$$\begin{aligned}\varphi - \Phi(u, x) &\geq 1, \\ 1 + e^{-\theta(\varphi - \Phi(u, x))} &\leq 1 + e^{-\theta}, \\ \frac{1}{1 + e^{-\theta(\varphi - \Phi(u, x))}} &\geq \frac{1}{1 + e^{-\theta}}, \\ 1 - \frac{1}{1 + e^{-\theta(\varphi - \Phi(u, x))}} &\leq 1 - \frac{1}{1 + e^{-\theta}}.\end{aligned}$$

Таким образом, на области χ_2^+ выполняется неравенство:

$$g(\varphi - \Phi(u, x)) \leq 1 - \frac{1}{1 + e^{-\theta}}. \quad (1.4.25)$$

Возвращаясь к интегралу (1.4.18), из положительности подынтегральной функции и свойств определенного интеграла следует неравенство:

$$\int_{\chi_2^+} g(\varphi - \Phi(u, x))f(x)dx \leq \int_{\chi_2^+} \left(1 - \frac{1}{1 + e^{-\theta}}\right) f(x)dx. \quad (1.4.26)$$

2) Рассмотрим область $\chi_2^- = \{x : \varphi - \Phi(u, x) \leq -1\}$. Функция Хевисайда внутри функции $g(\varphi - \Phi(u, x))$ на этой области равна нулю. Тогда

$$g(\varphi - \Phi(u, x)) = |0 - S_\theta(\varphi - \Phi(u, x))| = S_\theta(\varphi - \Phi(u, x)) = \frac{1}{1 + e^{-\theta \cdot (\varphi - \Phi(u, x))}} \quad (1.4.27)$$

Аналогично цепочке неравенств для области разбиения χ_2^+ , из неравенства условия:

$$\varphi - \Phi(u, x) \leq -1 \quad (1.4.28)$$

получаем неравенство:

$$\frac{1}{1 + e^{-\theta(\varphi - \Phi(u, x))}} \leq \frac{1}{1 + e^\theta}. \quad (1.4.29)$$

Аналогично рассуждениям для области разбиения χ_2^+ , получаем:

$$\int_{\chi_2^-} g(\varphi - \Phi(u, x))f(x)dx \leq \int_{\chi_2^-} \left(\frac{1}{1 + e^\theta}\right) f(x)dx. \quad (1.4.30)$$

3) Рассмотрим область $\chi_1^+ = \{x : 0 \leq \varphi - \Phi(u, x) < 1\}$. Нетрудно убедиться, что функция $g(\varphi - \Phi(u, x))$ убывает на области χ_1^+ при возрастании значения аргумента $\varphi - \Phi(u, x)$.

Возьмем произвольное значение $t^* \in (0, 1)$. Исходя из непрерывности и убывания $g(\varphi -$

$\Phi(u, x)$ следуют неравенства:

$$g(\varphi - \Phi(u, x)) \leq \frac{1}{2}, \quad \text{если } \varphi - \Phi(u, x) \in [0, t^*] \quad (1.4.31)$$

$$g(\varphi - \Phi(u, x)) \leq g(t^*), \quad \text{если } \varphi - \Phi(u, x) \in [t^*, 1] \quad (1.4.32)$$

Графическая иллюстрация полученных неравенств представлена на рисунке 1.2.

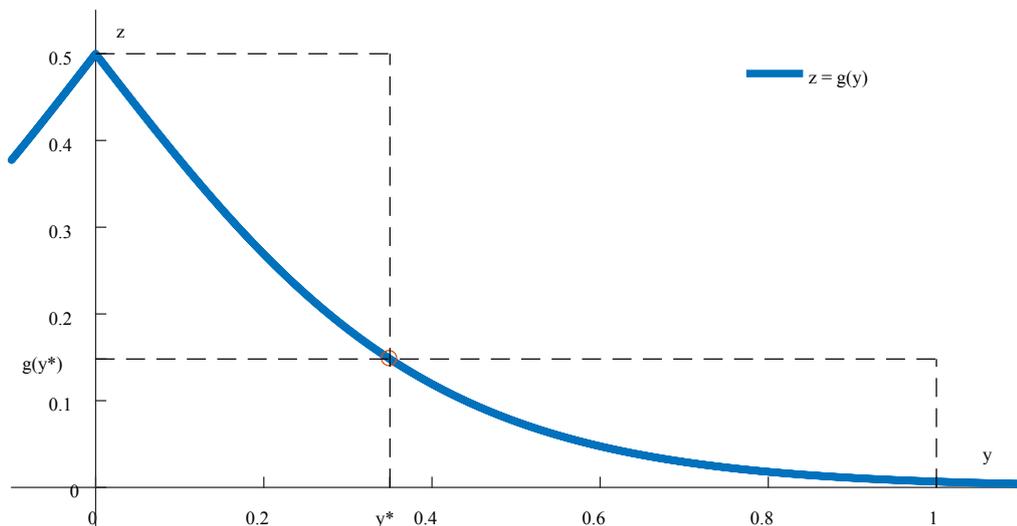


Рисунок 1.2. График функции $g(y)$ для $0 \leq y < 1$

Разобьем интеграл функции вероятности по области χ_1^+ на сумму двух интегралов:

$$\int_{\chi_1^+} g(\varphi - \Phi(u, x))f(x)dx = \int_{\chi_1^{++}} g(\varphi - \Phi(u, x))f(x)dx + \int_{\chi_1^{+-}} g(\varphi - \Phi(u, x))f(x)dx, \quad (1.4.33)$$

где

$$\chi_1^{++} = \{x : t^* < \varphi - \Phi(u, x) < 1\}, \quad (1.4.34)$$

$$\chi_1^{+-} = \{x : 0 < \varphi - \Phi(u, x) < t^*\}. \quad (1.4.35)$$

Учитывая неравенства (1.4.31) и (1.4.32) и свойства интегралов, получаем:

$$\int_{\chi_1^{++}} g(\varphi - \Phi(u, x))f(x)dx \leq \int_{\chi_1^{++}} g(t^*)f(x)dx, \quad (1.4.36)$$

$$\int_{\chi_1^{+-}} g(\varphi - \Phi(u, x))f(x)dx \leq \frac{1}{2} \int_{\chi_1^{+-}} f(x)dx. \quad (1.4.37)$$

Учитывая, что значение $g(t^*)$ константно и может быть вычислено непосредственно, получаем:

$$\int_{\chi_1^{++}} g(\varphi - \Phi(u, x))f(x)dx \leq \left(1 - \frac{1}{1 + e^{-\theta t^*}}\right) \int_{\chi_1^{++}} f(x)dx, \quad (1.4.38)$$

Так как неравенства (1.4.37), (1.4.38) действуют для любых $t^* \in (0, 1)$, возьмем в качестве t^* определенные значения, представляющие собой последовательность, зависящую от θ , причем

$$t_\theta^* = \frac{1}{\sqrt{\theta}}. \quad (1.4.39)$$

Тогда неравенство (1.4.38) записывается в виде:

$$\int_{\chi_1^{++}} g(\varphi - \Phi(u, x))f(x)dx \leq \left(1 - \frac{1}{1 + e^{-\sqrt{\theta}}}\right) \int_{\chi_1^{++}} f(x)dx,$$

а определения рассматриваемых областей χ_1^{++} , χ_1^{+-} могут быть переписаны с учетом замены значения t^* на t_θ^* . Таким образом, на области χ_1^+ оценка верхней границы погрешности выражается функцией, зависящей от параметра θ .

4) Рассмотрим область $\chi_1^- = \{x : -1 < \varphi - \Phi(u, x) < 0\}$. Ввиду симметричности функции $g(\varphi - \Phi(u, x))$ относительно аргумента $\varphi - \Phi(u, x)$, все рассуждения для области χ_1^+ переносятся на текущую область аналогичным образом с некоторыми отличиями.

Так же, как и при рассуждениях для области χ_1^+ , оценивается каждый из интегралов, получаемый при разбиении области χ_1^- на две подобласти χ_1^{-+} и χ_1^{--} точкой $t^* \in (-1; 0)$, т.е.

$$\chi_1^{-+} = \{x : t^* < \varphi - \Phi(u, x) < 0\}, \quad (1.4.40)$$

$$\chi_1^{--} = \{x : -1 < \varphi - \Phi(u, x) < t^*\}. \quad (1.4.41)$$

Функция $g(\varphi - \Phi(u, x))$ возрастает на этом промежутке относительно аргумента $\varphi - \Phi(u, x)$, поэтому

$$g(\varphi - \Phi(u, x)) \leq \frac{1}{2}, \quad \text{если } \varphi - \Phi(u, x) \in [t^*, 0] \quad (1.4.42)$$

$$g(\varphi - \Phi(u, x)) \leq g(t^*), \quad \text{если } \varphi - \Phi(u, x) \in [-1, t^*] \quad (1.4.43)$$

В качестве точки t^* берется значение, зависящее от θ :

$$t_\theta^* = -\frac{1}{\sqrt{\theta}}. \quad (1.4.44)$$

Аналогично предыдущему случаю находятся верхние оценки для интегралов по областям

χ_1^{-+} и χ_1^{--} :

$$\int_{\chi_1^{--}} g(\varphi - \Phi(u, x))f(x)dx \leq \left(\frac{1}{1 + e^{\sqrt{\theta}}}\right) \int_{\chi_1^{--}} f(x)dx, \quad (1.4.45)$$

$$\int_{\chi_1^{-+}} g(\varphi - \Phi(u, x))f(x)dx \leq \frac{1}{2} \int_{\chi_1^{-+}} f(x)dx. \quad (1.4.46)$$

В итоге после рассмотрения всех четырех областей получается следующее выражение для верхней оценки погрешности аппроксимации функции вероятности:

$$\begin{aligned} |P_\varphi(u) - P_\varphi^\theta(u)| &\leq \left(1 - \frac{1}{1 + e^{-\theta}}\right) \int_{\chi_2^+} f(x)dx + \left(\frac{1}{1 + e^\theta}\right) \int_{\chi_2^-} f(x)dx + \\ &+ \frac{1}{2} \int_{\chi_1^{+-}} f(x)dx + \left(1 - \frac{1}{1 + e^{-\sqrt{\theta}}}\right) \int_{\chi_1^{++}} f(x)dx + \\ &+ \frac{1}{2} \int_{\chi_1^{-+}} f(x)dx + \left(\frac{1}{1 + e^{\sqrt{\theta}}}\right) \int_{\chi_1^{--}} f(x)dx, \end{aligned} \quad (1.4.47)$$

где области χ_2^+ , χ_2^- задаются согласно (1.4.21), (1.4.22), а области χ_1^{+-} , χ_1^{++} , χ_1^{-+} , χ_1^{--} заданы следующим образом:

$$\chi_1^{++} = \left\{ x : \frac{1}{\sqrt{\theta}} < \varphi - \Phi(u, x) < 1 \right\}, \quad (1.4.48)$$

$$\chi_1^{+-} = \left\{ x : 0 < \varphi - \Phi(u, x) < \frac{1}{\sqrt{\theta}} \right\}. \quad (1.4.49)$$

$$\chi_1^{-+} = \left\{ x : -\frac{1}{\sqrt{\theta}} < \varphi - \Phi(u, x) < 0 \right\}, \quad (1.4.50)$$

$$\chi_1^{--} = \left\{ x : -1 < \varphi - \Phi(u, x) < -\frac{1}{\sqrt{\theta}} \right\}. \quad (1.4.51)$$

Полученная верхняя оценка погрешности, хоть и выглядит громоздкой, намного точнее простых оценок, например, оценок, полученных путем мажорирования значений функции $g(\varphi - \Phi(u, x))$ константой $1/2$ или набором линейных функций. Более того, данная оценка погрешности уменьшается с ростом θ . При $\theta \rightarrow +\infty$ каждое из слагаемых (1.4.47) стремится к нулю: коэффициенты перед интегралами по областям χ_2^+ , χ_2^- , χ_1^{++} , χ_1^{--} стремятся к нулю при ограниченном значении самого интеграла, а значения интегралов по областям χ_1^{+-} , χ_1^{-+} стремятся к нулю вследствие стремления области интегрирования к точке. Таким образом, данная верхняя оценка погрешности стремится к нулю при $\theta \rightarrow +\infty$.

1.5 Свойства дельта-функции Дирака и производной сигмоидальной функции

Для дальнейших выкладок необходимо показать некоторые свойства дельта-функции Дирака и производной сигмоидальной функции.

Обозначим пространство финитных бесконечно дифференцируемых на \mathbb{R} функций как C . Дельта-функция Дирака — обобщенная функция (непрерывный линейный функционал), определяемый для любой функции $g \in C$ следующим образом:

$$\langle \delta, g \rangle = \int_{-\infty}^{+\infty} g(t) d\Theta(t) = g(0), \quad (1.5.1)$$

где интеграл понимается как интеграл Стильеса. Дельта-функция обладает четностью в смысле обобщенных функций и для нее выполняется фильтрующее свойство:

$$\int_{-\infty}^{+\infty} \delta(t - t_0) g(t) dt = g(t_0). \quad (1.5.2)$$

Также известно, что при рассмотрении функции Хевисайда как обобщенной функции, т.е.

$$\langle \Theta, g \rangle = \int_{-\infty}^{+\infty} \Theta(t) g(t) dt, \quad (1.5.3)$$

ее производной является дельта-функция:

$$\Theta' = \delta. \quad (1.5.4)$$

Также для дельта-функции существует эвристическое описание:

$$\delta(t) = \begin{cases} +\infty, & t = 0, \\ 0, & t \neq 0, \end{cases} \quad (1.5.5)$$

$$\int_{-\infty}^{+\infty} \delta(t) dt = 1. \quad (1.5.6)$$

Дельта-функция не является классической функцией, однако при наличии классической функции, сходящейся к дельта-функции в смысле обобщенных функций, становится возможным использовать как фильтрующее свойство, так и определение (1.5.1), интерпретируя интеграл как обычный интеграл Римана-Стильеса.

Вернемся к сигмоидальной функции. Ее производная вычисляется непосредственно:

$$\begin{aligned} \frac{\partial S_\theta(t)}{\partial t} &= \left(\frac{1}{1 + e^{-\theta t}} \right)' = \frac{\theta e^{-\theta t}}{(1 + e^{-\theta t})^2} = \theta \cdot \frac{1}{1 + e^{-\theta t}} \cdot \frac{e^{-\theta t}}{1 + e^{-\theta t}} = \\ &= \theta \cdot \frac{1}{1 + e^{-\theta t}} \cdot \left(1 - \frac{1}{1 + e^{-\theta t}} \right) = \theta \cdot S_\theta(t) \cdot (1 - S_\theta(t)). \end{aligned} \quad (1.5.7)$$

График функции $S'_\theta(t) = \theta \cdot S_\theta(t) \cdot (1 - S_\theta(t))$ представлен на рисунке 1.3 (параметр $\theta = 5$).

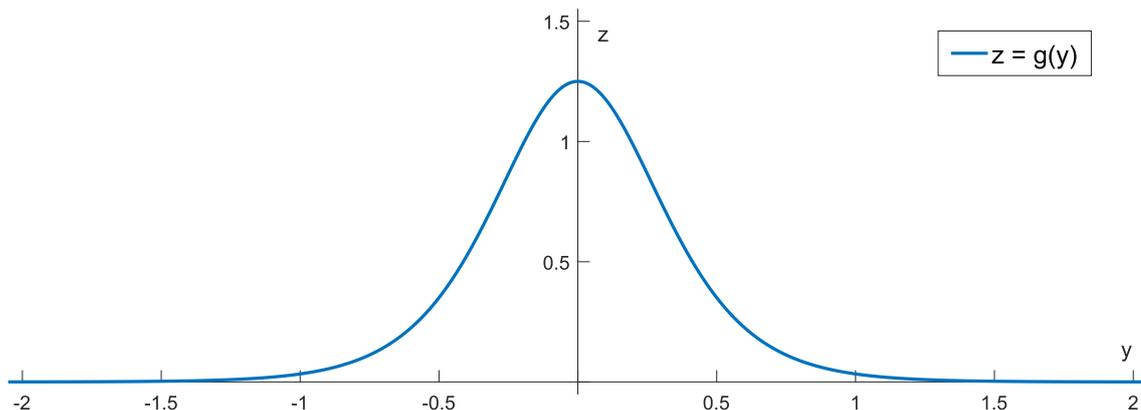


Рисунок 1.3. Производная сигмоидальной функции

Рассмотрим производную сигмоиды при фиксированных значениях аргумента t и при стремлении параметра сигмоиды θ в бесконечность. Значение производной при $t = 0$ вычисляется непосредственно:

$$S'_\theta(t) \Big|_{t=0} = \theta \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{\theta}{4} \rightarrow +\infty \quad \text{при } \theta \rightarrow +\infty. \quad (1.5.8)$$

При $t \neq 0$ рассматривается предел производной:

$$\lim_{\theta \rightarrow +\infty} S'_\theta(t) = \lim_{\theta \rightarrow +\infty} \frac{\theta e^{-\theta t}}{(1 + e^{-\theta t})^2} \quad (1.5.9)$$

Знаменатель выражения стремится к единице. Предел выражения в числителе вычисляется по правилу Лопиталя:

$$\lim_{\theta \rightarrow +\infty} \theta e^{-\theta t} = \lim_{\theta \rightarrow +\infty} \frac{\theta}{e^{\theta t}} = \lim_{\theta \rightarrow +\infty} \frac{1}{e^{\theta t} \cdot t} = 0 \quad (1.5.10)$$

Таким образом,

$$S'_\theta(t) \rightarrow 0 \quad \text{при } \theta \rightarrow +\infty, \quad t \neq 0, \quad (1.5.11)$$

Также непосредственно проверяется, что

$$\int_{-\infty}^{+\infty} S'_\theta(t) dt = S_\theta(t) \Big|_{-\infty}^{+\infty} = 1 - 0 = 1. \quad (1.5.12)$$

Таким образом, последовательность производных $\{S'_\theta(t)\}_{\theta=1}^{\infty}$ в пределе обладает свойствами, при-

сущими эвристическому описанию дельта-функции.

Теорема 3. Последовательность производных сигмоидальных функций $\{S'_\theta(t)\}_{\theta=1}^\infty$ сходится к дельта-функции Дирака в смысле обобщенных функций при стремлении параметра сигмоидальной функции к бесконечности, т.е.

$$\lim_{\theta \rightarrow +\infty} \int_{-\infty}^{+\infty} S'_\theta(t)g(t)dt = \int_{-\infty}^{+\infty} \delta(t)g(t)dt = g(0) \quad (1.5.13)$$

для любой финитной бесконечно дифференцируемой функции $g(t)$.

Доказательство. Исходя из определения дельта-функции, необходимо доказать, что

$$\lim_{\theta \rightarrow +\infty} \int_{-\infty}^{+\infty} S'_\theta(t)g(t)dt = g(0) \quad (1.5.14)$$

Введем обозначение:

$$I_\theta = \int_{-\infty}^{+\infty} S'_\theta(t)g(t)dt. \quad (1.5.15)$$

Рассмотрим разность $I_\theta - g(0)$. С учетом свойства (1.5.12) получим

$$I_\theta - g(0) = \int_{-\infty}^{+\infty} S'_\theta(t)g(t)dt - g(0) \int_{-\infty}^{+\infty} S'_\theta(t)dt = \int_{-\infty}^{+\infty} S'_\theta(t) [g(t) - g(0)] dt. \quad (1.5.16)$$

Далее оценим разность

$$I_\theta - g(0) \leq |I_\theta - g(0)| = \left| \int_{-\infty}^{+\infty} S'_\theta(t) [g(t) - g(0)] dt \right| \leq \int_{-\infty}^{+\infty} S'_\theta(t) |g(t) - g(0)| dt. \quad (1.5.17)$$

Пусть имеется произвольное число $T > 0$. Представим интеграл в виде суммы:

$$\int_{-\infty}^{+\infty} S'_\theta(t) |g(t) - g(0)| dt = I_\theta^1 + I_\theta^2 + I_\theta^3, \quad (1.5.18)$$

где

$$I_\theta^1 = \int_{-\infty}^{-T} S'_\theta(t) |g(t) - g(0)| dt, \quad (1.5.19)$$

$$I_{\theta}^2 = \int_{-T}^T S'_{\theta}(t) |g(t) - g(0)| dt, \quad (1.5.20)$$

$$I_{\theta}^3 = \int_T^{+\infty} S'_{\theta}(t) |g(t) - g(0)| dt. \quad (1.5.21)$$

Выберем произвольное $\varepsilon > 0$ и рассмотрим каждый из интегралов.

1) Рассмотрим интеграл I_{θ}^1 . Поскольку $g(t)$ финитна и непрерывна, то она является ограниченной на \mathbb{R} . Поэтому существует конечное положительное значение $G \in \mathbb{R}$, такое, что $g(t) \leq G \forall t \in \mathbb{R}$. Это означает, что

$$|g(t) - g(0)| \leq |g(t)| + |g(0)| \leq 2G. \quad (1.5.22)$$

Следовательно, значение I_{θ}^1 оценивается как

$$I_{\theta}^1 \leq \int_{-\infty}^{-T} 2G \cdot S'_{\theta}(t) dt = 2G \int_{-\infty}^{-T} S'_{\theta}(t) dt = 2G \cdot S_{\theta}(t) \Big|_{-\infty}^{-T} = 2G \cdot S_{\theta}(-T). \quad (1.5.23)$$

Тогда, поскольку значение сигмоиды при отрицательном аргументе положительно и стремится к нулю при увеличении параметра θ , то можно выбрать такое значение параметра $\theta_1 > 0$, что

$$\forall \theta > \theta_1, \quad \forall T > 0 : S_{\theta}(-T) < \frac{\varepsilon}{6G}. \quad (1.5.24)$$

Таким образом, получаем, что при параметре $\theta > \theta_1$ верно

$$I_{\theta}^1 < \frac{\varepsilon}{3}. \quad (1.5.25)$$

2) Аналогичным образом рассмотрим интеграл I_{θ}^3 . Из ограниченности функции $g(t)$ следует:

$$I_{\theta}^1 \leq \int_T^{+\infty} 2G \cdot S'_{\theta}(t) dt = 2G \cdot S_{\theta}(t) \Big|_T^{+\infty} = 2G \cdot (1 - S_{\theta}(T)). \quad (1.5.26)$$

Поскольку значение сигмоидальной функции стремится к единице при увеличении θ , то можно выбрать такое значение параметра $\theta_2 > 0$, что

$$\forall \theta > \theta_2, \quad \forall T > 0 : S_{\theta}(T) > 1 - \frac{\varepsilon}{6G}. \quad (1.5.27)$$

Таким образом, получаем, что при параметре $\theta > \theta_2$ верно

$$I_\theta^3 < \frac{\varepsilon}{3}. \quad (1.5.28)$$

3) Рассмотрим интеграл I_θ^2 . Исходя из непрерывности $g(t)$, существует непустая окрестность около нуля, т.е. интервал $(-R, R)$, для всех точек которого выполняется

$$\forall t \in (-R, R) : |g(t) - g(0)| < \frac{\varepsilon}{3}. \quad (1.5.29)$$

Так как значение предела интегрирования T произвольно и его выбор не влияет на рассуждения при оценке I_θ^1 и I_θ^3 , то выберем T таким образом, чтобы неравенство (1.5.29) выполнялось на интервале $(-T, T)$. Тогда значение интеграла I_θ^2 оценивается как

$$I_\theta^2 \leq \int_{-T}^T S'_\theta(t) \cdot \frac{\varepsilon}{3} dt = \frac{\varepsilon}{3} \int_{-T}^T S'_\theta(t) dt \leq \frac{\varepsilon}{3} \int_{-\infty}^{+\infty} S'_\theta(t) dt = \frac{\varepsilon}{3}. \quad (1.5.30)$$

Возвращаясь к исходным выражениям (1.5.17)-(1.5.18), получаем, что для произвольного $\varepsilon > 0$ и при параметре сигмоиды $\theta > \max(\theta_1, \theta_2)$ выполняется:

$$I_\theta - g(0) \leq |I_\theta - g(0)| \leq \int_{-\infty}^{+\infty} S'_\theta(t) |g(t) - g(0)| dt = I_\theta^1 + I_\theta^2 + I_\theta^3 \leq \varepsilon, \quad (1.5.31)$$

Поскольку ε может быть сколь угодно малым, то получаем, что

$$I_\theta - g(0) \rightarrow 0 \quad \text{при } \theta \rightarrow +\infty, \quad (1.5.32)$$

или, с учетом определения I_θ (1.5.14):

$$\int_{-\infty}^{+\infty} S'_\theta(t) g(t) dt \rightarrow g(0) \quad \text{при } \theta \rightarrow +\infty, \quad (1.5.33)$$

что, в свою очередь, эквивалентно доказываемому утверждению. \square

Таким образом, справедливы следующие два следствия.

Следствие 1. Для любой финитной бесконечно дифференцируемой функции $g(t)$ производная

сигмоидальной функции в пределе при $\theta \rightarrow +\infty$ обладает фильтрующим свойством:

$$\int_{-\infty}^{+\infty} S'_\theta(t - t_0)g(t)dt \rightarrow g(t_0) \text{ при } \theta \rightarrow +\infty. \quad (1.5.34)$$

Следствие 2. Для любой финитной бесконечно дифференцируемой функции $g(t)$ и на любом отрезке $[a, b]$, не включающем в себя нуль аргумента сигмоидальной функции, интеграл произведения производной сигмоидальной функции и $g(t)$ стремится к нулю при $\theta \rightarrow +\infty$:

$$\forall [a, b] : 0 \notin [a, b] \quad \int_a^b S'_\theta(t)g(t)dt \rightarrow 0 \text{ при } \theta \rightarrow +\infty. \quad (1.5.35)$$

1.6 Сходимость производной функции вероятности по компонентам стратегии

Далее рассматривается вопрос о сходимости производных аппроксимированной функции вероятности к производным исходной функции вероятности. Для начала рассмотрим теорему о сходимости частной производной по компонентам стратегии.

Теорема 4. Пусть X — абсолютно непрерывная случайная величина со значениями $x \in \mathbb{R}$ на полном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$. Пусть задано непустое множество $U \subseteq \mathbb{R}^m$ допустимых стратегий u . Пусть задана непрерывная дважды дифференцируемая по совокупности аргументов строго кусочно-монотонная по x для любого $u \in U$ функция потерь $\Phi(u, x) : U \times \mathbb{R} \rightarrow \mathbb{R}$, с ограниченными непрерывными производными по компонентам стратегии u для любого x , а также задан уровень потерь φ . Пусть также выражениями (1.4.11) и (1.3.4) заданы функции вероятности $P_\varphi(u)$ и ее аппроксимация $P_\varphi^\theta(u)$. Тогда последовательность частных производных по компонентам u_k стратегии u аппроксимированной вероятностной функции стремится к частной производной по компонентам u_k стратегии u исходной вероятностной функции везде, где производная функции вероятности определена, т.е.

$$\frac{\partial P_\varphi^\theta(u)}{\partial u_k} \rightarrow \frac{\partial P_\varphi(u)}{\partial u_k} \text{ при } \theta \rightarrow +\infty, \quad k = \overline{1, m}, \quad \forall u \in U. \quad (1.6.1)$$

Доказательство. Поскольку $\Phi(u, x)$ является гладкой кусочно-монотонной по x функцией, то выражение для функции вероятности $P_\varphi(u)$ можно представить в виде суммы интегралов особого вида:

$$P_\varphi(u) = \sum_i \int_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x))f(x)dx, \quad (1.6.2)$$

где $x_i^-(u)$ и $x_i^+(u)$ – границы i -го промежутка монотонности функции $\Phi(u, x)$ при заданной стратегии u . Эти точки определяются как точки перегибов и локальных экстремумов функции $\Phi(u, x)$ при заданном u , т.е. являются решениями уравнения:

$$\frac{\partial}{\partial x}\Phi(u, x) = 0. \quad (1.6.3)$$

Сначала рассмотрим случаи, когда точек перегибов и локальных экстремумов у функции $\Phi(u, x)$ по переменной x нет, т.е. уравнение (1.6.3) не дает решения.

Случай отсутствия перегибов и локальных экстремумов

1) Пусть $\forall u \in U$ уравнение $\Phi(u, x) = \varphi$ не имеет решений. В этом случае функция вероятности $P_\varphi(u)$ при любых $u \in U$ постоянна: либо везде равна единице, либо везде равна нулю, — и ее производная равняется нулю.

Рассмотрим аппроксимацию функции вероятности $P_\varphi^\theta(u)$. По правилу Лейбница:

$$\frac{\partial P_\varphi^\theta(u)}{\partial u_k} = \frac{\partial}{\partial u_k} \int_{-\infty}^{+\infty} S_\theta(\varphi - \Phi(u, x))f(x)dx = \int_{-\infty}^{+\infty} S'_\theta(\varphi - \Phi(u, x))(-\Phi'_{u_k}(u, x))f(x)dx. \quad (1.6.4)$$

В рассматриваемом случае аргумент производной сигмоиды $\varphi - \Phi(u, x)$ не обращается в ноль на всей области интегрирования. Отсюда в соответствии со Следствием 2:

$$\int_{-\infty}^{+\infty} S'_\theta(\varphi - \Phi(u, x))(-\Phi'_{u_k}(u, x))f(x)dx \rightarrow 0 \text{ при } \theta \rightarrow +\infty, \quad (1.6.5)$$

что влечет выполнение (1.6.1) в этом случае.

2) Пусть в некоторой окрестности $u \in U$ существует единственное решение уравнения $\Phi(u, x) = \varphi$ по x ; обозначим его $x^*(u)$. В таком случае в зависимости от вида функции $\Phi(u, X)$, условие $\Phi(u, X) \leq \varphi$ эквивалентно либо условию $X \leq x^*(u)$, либо условию $X \geq x^*(u)$. Без потери общности упростим функцию $P_\varphi(u)$:

$$P_\varphi(u) = \mathbb{P}\{\Phi(u, X) \leq \varphi\} = \mathbb{P}\{X \leq x^*(u)\} = \mathbb{M}[\Theta(x^*(u) - X)] \quad (1.6.6)$$

Пусть Δu_k - приращение по компоненте u_k вектора u . Обозначим вектор размерности m , имеющий единицу на k -том месте и нули на остальных, как E_k . Тогда у уравнения $\Phi(u, x) = \varphi$ будет корень $x^*(u + E_k \Delta u_k)$. Не теряя общности, положим $x^*(u + E_k \Delta u_k) > x^*(u)$. Производная функции

вероятности может быть выражена как:

$$\frac{\partial P_\varphi(u)}{\partial u_k} = \lim_{\Delta u_k \rightarrow 0} \frac{\mathbb{M}[\Theta(x^*(u + E_k \Delta u_k) - X)] - \mathbb{M}[\Theta(x^*(u) - X)]}{\Delta u_k}. \quad (1.6.7)$$

Руководствуясь тем, что условие $\Phi(u, X) \leq \varphi$ эквивалентно либо условию $X \leq x^*(u)$, либо $X \geq x^*(u)$, без потери общности упростим аппроксимацию $P_\varphi^\theta(u)$:

$$P_\varphi^\theta(u) = \mathbb{M}[S_\theta(\varphi - \Phi(u, x))] = \mathbb{M}[S_\theta(x^*(u) - X)], \quad (1.6.8)$$

и аналогично найдем производную:

$$\frac{\partial P_\varphi^\theta(u)}{\partial u_k} = \lim_{\Delta u_k \rightarrow 0} \frac{\mathbb{M}[S_\theta(x^*(u + E_k \Delta u_k) - X)] - \mathbb{M}[S_\theta(x^*(u) - X)]}{\Delta u_k}. \quad (1.6.9)$$

Составим разность этих производных:

$$\frac{\partial P_\varphi(u)}{\partial u_k} - \frac{\partial P_\varphi^\theta(u)}{\partial u_k} = \lim_{\Delta u_k \rightarrow 0} \left(\frac{\mathbb{M}[\Theta(x^*(u + E_k \Delta u_k) - X) - \Theta(x^*(u) - X)]}{\Delta u_k} - \frac{\mathbb{M}[S_\theta(x^*(u + E_k \Delta u_k) - X) - S_\theta(x^*(u) - X)]}{\Delta u_k} \right) \quad (1.6.10)$$

Рассмотрим пары разностей функции Хевисайда и сигмоиды при одинаковых аргументах. В соответствии с Теоремой 1:

$$S_\theta(t) \xrightarrow{\text{п.п.}} \Theta(t) \quad \text{при } \theta \rightarrow +\infty, \quad (1.6.11)$$

следовательно,

$$\forall \varepsilon > 0 \exists \theta_1^* > 0 : \forall \theta > \theta_1^* \quad | -\Theta(x^*(u) - X) + S_\theta(x^*(u) - X) | < \frac{\varepsilon}{2}. \quad (1.6.12)$$

С учетом этого факта получается цепочка неравенств:

$$\begin{aligned} \mathbb{M}[-\Theta(x^*(u) - X) + S_\theta(x^*(u) - X)] &\leq \left| \int_{-\infty}^{+\infty} (S_\theta(x^*(u) - X) - \Theta(x^*(u) - X)) f(x) dx \right| \leq \\ &\leq \int_{-\infty}^{+\infty} |S_\theta(x^*(u) - X) - \Theta(x^*(u) - X)| f(x) dx \leq \int_{-\infty}^{+\infty} \frac{\varepsilon}{2} f(x) dx \leq \frac{\varepsilon}{2} \end{aligned} \quad (1.6.13)$$

для всех $\theta > \theta_1^*$. Аналогичные выкладки проводятся и для второй пары:

$$\forall \varepsilon > 0 \exists \theta_2^* > 0 : \forall \theta > \theta_2^* \mathbb{M}[-\Theta(x^*(u + E_k \Delta u_k) - X) + S_\theta(x^*(u + E_k \Delta u_k) - X)] \leq \frac{\varepsilon}{2}. \quad (1.6.14)$$

Таким образом, для любого $\varepsilon > 0$ и для всех $\theta > \max(\theta_1^*, \theta_2^*)$ числитель отношения в выражении (1.6.10) меньше ε , что означает его стремление к нулю при $\theta \rightarrow +\infty$. Это означает, что в случае существования предела выражения (1.6.7) справедливо (1.6.1).

Случай наличия перегибов и локальных экстремумов

Итак, функция вероятности $P_\varphi(u)$ представляется в виде суммы интегралов функции Хевисайда в пределах от $x_i^-(u)$ и $x_i^+(u)$ – границ i -го промежутка монотонности функции $\Phi(u, x)$ при заданной стратегии u :

$$P_\varphi(u) = \sum_i \int_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx. \quad (1.6.15)$$

Эти точки являются решениями уравнения:

$$\frac{\partial}{\partial x} \Phi(u, x) = 0. \quad (1.6.16)$$

Здесь стоит отметить, что при наличии у функции конечного числа интервалов монотонности, концы которых являются решениями уравнения (1.6.15), крайний левый и крайний правый полуинтервалы монотонности содержат бесконечно удаленную точку. В силу гладкости и непрерывности функции $\Phi(u, x)$ небольшие изменения вектора u не влияют на изменение количества бесконечно удаленных корней уравнения, и дополнительно отметим, что производные границ интервалов в бесконечно удаленных точках можно доопределить нулем, что не оказывает влияния на рассуждения далее.

Уравнение (1.6.15) задает неявную функцию $\Phi'_x(u, x)$. Следовательно, по теореме о неявной функции возможно найти производные границ i -го интервала монотонности $x_i^-(u)$ и $x_i^+(u)$ по компоненте вектора стратегии u_k . Для краткости записей введем обозначения:

$$\frac{\partial}{\partial x} \Phi(u, x) \Big|_{x=t} = \Phi'_x(u, t), \quad \frac{\partial^2}{\partial x \partial u_k} \Phi(u, x) \Big|_{x=t} = \Phi''_{xu_k}(u, t). \quad (1.6.17)$$

Тогда производные задаются как:

$$\frac{\partial x_i^-(u)}{\partial u_k} = -\frac{\Phi''_{xu_k}(u, x_i^-(u))}{\Phi''_{xx}(u, x_i^-(u))}, \quad \frac{\partial x_i^+(u)}{\partial u_k} = -\frac{\Phi''_{xu_k}(u, x_i^+(u))}{\Phi''_{xx}(u, x_i^+(u))}. \quad (1.6.18)$$

Данные результаты будут необходимы далее. Отметим, что в случае, если $\Phi''_{xx}(u, x_i^+(u))$ или $\Phi''_{xx}(u, x_i^-(u))$ равняется нулю, то соответствующие им левые части выражений (1.6.18) доопределяются нулем в этих точках, т.к. фактически это означает, что корень уравнения (1.6.15) не изменяется при изменении значения вектора u .

Воспользуемся представлением (1.6.2) и запишем производную исходной функции вероятности $P_\varphi(u)$ по компоненте стратегии u_k :

$$\frac{\partial}{\partial u_k} P_\varphi(u) = \sum_i \frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx. \quad (1.6.19)$$

Рассмотрим правую часть этого равенства. На каждом из промежутков монотонности $[x_i^-(u), x_i^+(u)]$ возможен один из четырех вариантов взаимного расположения значений функции потерь $\Phi(u, x)$ и требуемого уровня потерь φ :

$$\text{Случай 1.} \quad \min_{x \in [x_i^-(u); x_i^+(u)]} \Phi(u, x) \geq \varphi.$$

В этом случае значение функции Хевисайда $\Theta(\varphi - \Phi(u, x))$ на всем промежутке монотонности равно нулю. Это значит, что интеграл (1.6.19) также равен нулю:

$$\int_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx = 0, \quad (1.6.20)$$

и его производная на всем промежутке тоже равна нулю:

$$\frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx = 0. \quad (1.6.21)$$

$$\text{Случай 2.} \quad \max_{x \in [x_i^-(u); x_i^+(u)]} \Phi(u, x) \leq \varphi.$$

В этом случае значение функции Хевисайда равно единице на всем промежутке. Это значит, что:

$$\int_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx = \int_{x_i^-(u)}^{x_i^+(u)} f(x) dx. \quad (1.6.22)$$

Согласно формуле Лейбница, с учетом соотношений (1.6.18) и учетом независимости $f(x)$ от u_k ,

производная интеграла по u_k представляется как:

$$\begin{aligned} \frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} f(x) dx &= f(x_i^+(u)) \frac{\partial x_i^+(u)}{\partial u_k} - f(x_i^-(u)) \frac{\partial x_i^-(u)}{\partial u_k} = \\ &= f(x_i^-(u)) \frac{\Phi''_{xu_k}(u, x_i^-(u))}{\Phi''_{xx}(u, x_i^-(u))} - f(x_i^+(u)) \frac{\Phi''_{xu_k}(u, x_i^+(u))}{\Phi''_{xx}(u, x_i^+(u))}. \end{aligned} \quad (1.6.23)$$

Случай 3. $\Phi(u, x_i^-(u)) \leq \varphi \leq \Phi(u, x_i^+(u))$.

В этом случае на отрезке $[x_i^-(u), x_i^+(u)]$ при заданном u существует корень уравнения $\Phi(u, x) = \varphi$, причем единственный; обозначим его как $x_i^*(u)$. Тогда функция Хевисайда равна единице на отрезке $[x_i^-(u), x_i^*(u)]$, и равна нулю вне его, что означает:

$$\int_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx = \int_{x_i^-(u)}^{x_i^*(u)} f(x) dx, \quad (1.6.24)$$

Согласно формуле Лейбница, производная этого интеграла по u_k равна:

$$\frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^*(u)} f(x) dx = f(x_i^*(u)) \frac{\partial x_i^*(u)}{\partial u_k} - f(x_i^-(u)) \frac{\partial x_i^-(u)}{\partial u_k}. \quad (1.6.25)$$

Производная $x_i^*(u)$ по u_k определяется аналогично (1.6.18) с помощью теоремы о неявной функции и равна:

$$\frac{\partial x_i^*(u)}{\partial u_k} = - \frac{\Phi'_{u_k}(u, x_i^*(u))}{\Phi'_x(u, x_i^*(u))}. \quad (1.6.26)$$

Следовательно,

$$\int_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx = f(x_i^-(u)) \frac{\Phi''_{xu_k}(u, x_i^-(u))}{\Phi''_{xx}(u, x_i^-(u))} - f(x_i^*(u)) \frac{\Phi'_{u_k}(u, x_i^*(u))}{\Phi'_x(u, x_i^*(u))}. \quad (1.6.27)$$

Случай 4. $\Phi(u, x_i^-(u)) \geq \varphi \geq \Phi(u, x_i^+(u))$.

По аналогии со случаем 3, получаем, что на отрезке $[x_i^-(u), x_i^+(u)]$ при заданном u существует единственный корень $x_i^*(u)$ уравнения $\Phi(u, x) = \varphi$; функция Хевисайда при этом равна

единице на отрезке $[x_i^*(u), x_i^+(u)]$ и равна нулю вне его, что означает:

$$\int_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx = \int_{x_i^*(u)}^{x_i^+(u)} f(x) dx, \quad (1.6.28)$$

Производная этого интеграла по u_k равна:

$$\begin{aligned} \frac{\partial}{\partial u_k} \int_{x_i^*(u)}^{x_i^+(u)} f(x) dx &= f(x_i^+(u)) \frac{\partial x_i^+(u)}{\partial u_k} - f(x_i^*(u)) \frac{\partial x_i^*(u)}{\partial u_k} = \\ &= f(x_i^*(u)) \frac{\Phi'_{u_k}(u, x_i^*(u))}{\Phi'_x(u, x_i^*(u))} - f(x_i^+(u)) \frac{\Phi''_{u_k x}(u, x_i^+(u))}{\Phi''_{xx}(u, x_i^+(u))}. \end{aligned} \quad (1.6.29)$$

Таким образом, частная производная функции вероятности по компоненте стратегии u_k определена на всем множестве значений x .

Проведем аналогичные выкладки для аппроксимированной функции вероятности $P_\varphi^\theta(u)$. Рассматриваемая производная определяется аналогично производной исходной функции. В итоге получается аналог формулы (1.6.19):

$$\frac{\partial}{\partial u_k} P_\varphi^\theta(u) = \sum_i \frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u, x)) f(x) dx. \quad (1.6.30)$$

Рассмотрим аналогичные варианты взаимного расположения значений функции потерь $\Phi(u, x)$ и требуемого уровня потерь φ для интеграла (1.6.30):

Случай 1. $\min_{x \in [x_i^-(u); x_i^+(u)]} \Phi(u, x) \geq \varphi$. В этом случае, согласно формуле Лейбница, производная интеграла по u_k равна:

$$\begin{aligned} \frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u, x)) f(x) dx &= S_\theta(\varphi - \Phi(u, x_i^+(u))) f(x_i^+(u)) \frac{\partial x_i^+(u)}{\partial u_k} - \\ &- S_\theta(\varphi - \Phi(u, x_i^-(u))) f(x_i^-(u)) \frac{\partial x_i^-(u)}{\partial u_k} - \int_{x_i^-(u)}^{x_i^+(u)} S'_\theta(\varphi - \Phi(u, x)) \Phi'_{u_k}(u, x) f(x) dx. \end{aligned} \quad (1.6.31)$$

На этом промежутке монотонности согласно свойствам сигмоиды выполняются следующие соотношения:

$$S_\theta(\varphi - \Phi(u, x_i^+(u))) \rightarrow 0 \quad \text{при } \theta \rightarrow +\infty, \quad (1.6.32)$$

$$S_\theta(\varphi - \Phi(u, x_i^-(u))) \rightarrow 0 \quad \text{при } \theta \rightarrow +\infty. \quad (1.6.33)$$

Таким образом, первые два слагаемых в (1.6.31) стремятся к нулю. Согласно Следствию 2, третье слагаемое стремится к нулю:

$$\int_{x_i^-(u)}^{x_i^+(u)} S'_\theta(\varphi - \Phi(u, x)) \Phi'_{u_k}(u, x) f(x) dx \rightarrow 0 \quad \text{при } \theta \rightarrow +\infty, \quad (1.6.34)$$

Это означает, что на данном промежутке монотонности с учетом (1.6.21) верно:

$$\frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u, x)) f(x) dx \rightarrow \frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx, \quad (1.6.35)$$

т.е. производная аппроксимированной функции сходится к производной исходной функции при $\theta \rightarrow +\infty$.

Случай 2. $\max_{x \in [x_i^-(u); x_i^+(u)]} \Phi(u, x) \leq \varphi$.

Аналогично случаю 1, производная интеграла по u_k равна:

$$\begin{aligned} \frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u, x)) f(x) dx &= S_\theta(\varphi - \Phi(u, x_i^+(u))) f(x_i^+(u)) \frac{\partial x_i^+(u)}{\partial u_k} - \\ &- S_\theta(\varphi - \Phi(u, x_i^-(u))) f(x_i^-(u)) \frac{\partial x_i^-(u)}{\partial u_k} - \int_{x_i^-(u)}^{x_i^+(u)} S'_\theta(\varphi - \Phi(u, x)) \Phi'_{u_k}(u, x) f(x) dx. \end{aligned} \quad (1.6.36)$$

Согласно свойствам сигмоиды, значения множителей $S_\theta(\varphi - \Phi(u, x_i^+(u)))$ и $S_\theta(\varphi - \Phi(u, x_i^-(u)))$ стремятся к единице при $\theta \rightarrow +\infty$. Последнее слагаемое, согласно Следствию 2, стремится к нулю при $\theta \rightarrow +\infty$. Следовательно,

$$\begin{aligned} \frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u, x)) f(x) dx &\rightarrow f(x_i^-(u)) \frac{\Phi''_{xu_k}(u, x_i^-(u))}{\Phi''_{xx}(u, x_i^-(u))} - \\ &- f(x_i^+(u)) \frac{\Phi''_{xu_k}(u, x_i^+(u))}{\Phi''_{xx}(u, x_i^+(u))} \quad \text{при } \theta \rightarrow +\infty. \end{aligned} \quad (1.6.37)$$

Сопоставляя полученное выражение и (1.6.23), получим, что на данном промежутке монотонности верно соотношение (1.6.35).

Случай 3. $\Phi(u, x_i^-(u)) \leq \varphi \leq \Phi(u, x_i^+(u))$.

В этом случае при заданном u существует единственный корень $x_i^*(u)$ уравнения $\Phi(u, x) = \varphi$ на отрезке $[x_i^-(u); x_i^+(u)]$. Производная интеграла по u_k определяется согласно формуле (1.6.36), и на этом промежутке монотонности согласно свойствам сигмоиды выполняются следующие соотношения:

$$S_\theta(\varphi - \Phi(u, x_i^+(u))) \rightarrow 0 \quad \text{при } \theta \rightarrow +\infty, \quad (1.6.38)$$

$$S_\theta(\varphi - \Phi(u, x_i^-(u))) \rightarrow 1 \quad \text{при } \theta \rightarrow +\infty. \quad (1.6.39)$$

В последнем слагаемом (1.6.36), а именно

$$\int_{x_i^-(u)}^{x_i^+(u)} f(x) S'_\theta(\varphi - \Phi(u, x)) \Phi'_{u_k}(u, x) dx \quad (1.6.40)$$

сделаем замену переменных

$$t = \Phi(u, x), \quad dt = \Phi'_x(u, x) dx, \quad (1.6.41)$$

получаем:

$$\int_{x_i^-(u)}^{x_i^+(u)} f(x) S'_\theta(\varphi - \Phi(u, x)) \Phi'_{u_k}(u, x) dx = \int_{t^-}^{t^+} f(x(t)) S'_\theta(\varphi - t) \frac{\Phi'_{u_k}(u, x(t))}{\Phi'_x(u, x(t))} dt, \quad (1.6.42)$$

где $t^- = t(x_i^-(u))$, $t^+ = t(x_i^+(u))$.

Исходя из Следствия 1 (фильтрующего свойства производной сигмоиды), значение этого интеграла стремится к значению функции $f(x(t)) \frac{\Phi'_{u_k}(u, x(t))}{\Phi'_x(u, x(t))}$ в точке нуля производной сигмоиды, т.е. к значению в точке $t = \varphi$, что эквивалентно $\Phi(u, x) = \varphi$. По аналогии с выкладками для исходной функции вероятности, обозначим корень уравнения $\Phi(u, x) = \varphi$ на этом промежутке монотонности как $x_i^*(u)$. Тогда точка $t = \varphi$ по определению эквивалентна точке $x = x_i^*(u)$. Таким образом, возвращаясь к исходным переменным, получаем:

$$\int_{x_i^-(u)}^{x_i^+(u)} f(x) S'_\theta(\varphi - \Phi(u, x)) \Phi'_{u_k}(u, x) dx \rightarrow f(x_i^*(u)) \frac{\Phi'_{u_k}(u, x_i^*(u))}{\Phi'_x(u, x_i^*(u))} \quad \text{при } \theta \rightarrow +\infty. \quad (1.6.43)$$

Рассматривая выражение (1.6.36) с учетом (1.6.38), (1.6.39), (1.6.43), получим, что

$$\begin{aligned} \frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u, x))f(x)dx &\rightarrow f(x_i^-(u)) \frac{\Phi''_{xu_k}(u, x_i^-(u))}{\Phi''_{xx}(u, x_i^-(u))} - \\ &- f(x_i^*(u)) \frac{\Phi'_{u_k}(u, x_i^*(u))}{\Phi'_x(u, x_i^*(u))} \quad \text{при } \theta \rightarrow +\infty, \end{aligned} \quad (1.6.44)$$

а отсюда следует выполнение соотношения (1.6.35) для этого промежутка монотонности.

$$\text{Случай 4. } \Phi(u, x_i^-(u)) \geq \varphi \geq \Phi(u, x_i^+(u)).$$

Данный случай рассматривается аналогично случаю 3. Производная интеграла определяется по формуле (1.6.36); для сигмоидальной функции в этом случае верны выражения:

$$S_\theta(\varphi - \Phi(u, x_i^+)) \rightarrow 1 \quad \text{при } \theta \rightarrow +\infty, \quad (1.6.45)$$

$$S_\theta(\varphi - \Phi(u, x_i^-)) \rightarrow 0 \quad \text{при } \theta \rightarrow +\infty. \quad (1.6.46)$$

На данном отрезке монотонности существует единственный корень $x_i^*(u)$ уравнения $\Phi(u, x) = \varphi$. В интегральном слагаемом (1.6.43) производной предполагается замена (1.6.41), и в соответствии со Следствием 1 получаем:

$$\int_{x_i^-(u)}^{x_i^+(u)} f(x)S'_\theta(\varphi - \Phi(u, x))\Phi'_{u_k}(u, x)dx \rightarrow f(x_i^*(u)) \frac{\Phi'_{u_k}(u, x_i^*(u))}{\Phi'_x(u, x_i^*(u))} \quad \text{при } \theta \rightarrow +\infty. \quad (1.6.47)$$

В итоге, рассматривая (1.6.36) с учетом (1.6.45), (1.6.46), (1.6.47), получим:

$$\begin{aligned} \frac{\partial}{\partial u_k} \int_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u, x))f(x)dx &\rightarrow -f(x_i^+(u)) \frac{\Phi''_{u_k x}(u, x_i^+(u))}{\Phi''_{xx}(u, x_i^+(u))} + \\ &+ f(x_i^*(u)) \frac{\Phi'_{u_k}(u, x_i^*(u))}{\Phi'_x(u, x_i^*(u))} \quad \text{при } \theta \rightarrow +\infty, \end{aligned}$$

что означает выполнение соотношения (1.6.35) для этого промежутка монотонности.

Таким образом, для каждого промежутка монотонности справедливо соотношение (1.6.35), т.е. производная аппроксимированной функции сходится к производной исходной функции при $\theta \rightarrow +\infty$ на всем \mathbb{R} . \square

Исходя из определения, вычисление производной аппроксимированной функции вероятности сводится к дифференцированию математического ожидания сигмоидальной функции. Однако,

данная аппроксимированная производная может быть представлена в более удобном для вычисления виде, который описан в следующем пункте работы.

1.7 Вычисление аппроксимированной производной функции вероятности по компоненте стратегии

Пусть условия удовлетворяют условиям Теоремы 4. Для получения удобной формулы для вычисления производной аппроксимации вероятностной функции перепишем функцию потерь $\Phi(u, x)$ в следующем виде:

$$\Phi(u, x) \equiv \Phi(u_k, \bar{u}_k, x), \quad (1.7.1)$$

где u_k – k -я компонента вектора u , $\bar{u}_k \in \mathbb{R}^{m-1}$ – вектор, полученный из вектора u исключением k -й координаты. Рассмотрим математическое ожидание функции $H(u, \Delta u_k, X)$, заданной как:

$$H(u, \Delta u_k, X) = \frac{S_\theta(\varphi - \Phi(u_k + \Delta u_k, \bar{u}_k, X)) - S_\theta(\varphi - \Phi(u_k, \bar{u}_k, X))}{\Delta u_k}, \quad (1.7.2)$$

где $\Delta u_k > 0$. Сигмоидальную функцию $S_\theta(\varphi - \Phi(u_k, \bar{u}_k, x))$ представим в виде интеграла от специально подобранной функции $h(\beta, \bar{u}_k, x)$:

$$h(\beta, \bar{u}_k, x) = -\theta S_\theta(\varphi - \Phi(\beta, \bar{u}_k, x)) (1 - S_\theta(\varphi - \Phi(\beta, \bar{u}_k, x))) \Phi'_\beta(\beta, \bar{u}_k, x); \quad (1.7.3)$$

функция подобрана так, что

$$S_\theta(\varphi - \Phi(u_k, \bar{u}_k, x)) = \int_{-\infty}^{u_k} h(\beta, \bar{u}_k, x) d\beta. \quad (1.7.4)$$

Из определений (1.7.2) и (1.7.4) следует:

$$H(u, \Delta u_k, X) = \frac{1}{\Delta u_k} \int_{u_k}^{u_k + \Delta u_k} h(\beta, \bar{u}_k, X) d\beta, \quad (1.7.5)$$

а также следует равенство математических ожиданий обеих частей равенства.

Возьмем произвольное $c \in \mathbb{R}$ и рассмотрим правую часть предыдущего равенства. Справедливо следующее тождество:

$$\begin{aligned} \int_{u_k}^{u_k + \Delta u_k} h(\beta, \bar{u}_k, X) d\beta &= \int_{u_k}^{u_k + \Delta u_k} h(\beta, \bar{u}_k, X) I \{ |h(\beta, \bar{u}_k, X)| < c \} d\beta + \\ &+ \int_{u_k}^{u_k + \Delta u_k} h(\beta, \bar{u}_k, X) I \{ |h(\beta, \bar{u}_k, X)| \geq c \} d\beta, \end{aligned} \quad (1.7.6)$$

где $I\{\cdot\}$ - индикаторная функция. Используя свойства индикаторной функции и интегралов, нетрудно установить, что

$$\left| \int_{u_k}^{u_k + \Delta u_k} h(\beta, \bar{u}_k, X) I \{ |h(\beta, \bar{u}_k, X)| < c \} d\beta \right| \leq c \cdot \Delta u_k. \quad (1.7.7)$$

Таким образом, возвращаясь к (1.7.5), верно:

$$|H(u, \Delta u_k, X)| \leq c + \frac{1}{\Delta u_k} \int_{u_k}^{u_k + \Delta u_k} h(\beta, \bar{u}_k, X) I \{ |h(\beta, \bar{u}_k, X)| \geq c \} d\beta. \quad (1.7.8)$$

Оценим значение следующего выражения с учетом предыдущего результата:

$$\begin{aligned} \mathbf{M} [|H(u, \Delta u_k, X)| I \{ |H(u, \Delta u_k, X)| > 2c \}] &\geq \mathbf{M} [c \cdot I \{ |H(u, \Delta u_k, X)| > 2c \}] + \\ &+ \mathbf{M} \left[\frac{1}{\Delta u_k} \int_{u_k}^{u_k + \Delta u_k} h(\beta, \bar{u}_k, X) I \{ |h(\beta, \bar{u}_k, X)| \geq c \} d\beta \cdot I \{ |H(u, \Delta u_k, X)| > 2c \} \right]. \end{aligned} \quad (1.7.9)$$

Первое слагаемое правой части может быть оценено как:

$$\begin{aligned} \mathbf{M} [c \cdot I \{ |H(u, \Delta u_k, X)| > 2c \}] &= c \cdot \mathbb{P} \{ |H(u, \Delta u_k, X)| > 2c \} \leq \\ &\leq c \cdot \mathbb{P} \left\{ \frac{1}{\Delta u_k} \int_{u_k}^{u_k + \Delta u_k} h(\beta, \bar{u}_k, X) I \{ |h(\beta, \bar{u}_k, X)| \geq c \} d\beta > c \right\} \leq \\ &\leq \mathbf{M} \left[\frac{1}{\Delta u_k} \int_{u_k}^{u_k + \Delta u_k} h(\beta, \bar{u}_k, X) I \{ |h(\beta, \bar{u}_k, X)| \geq c \} d\beta \right]. \end{aligned} \quad (1.7.10)$$

Второе слагаемое правой части неравенства (1.7.9) оценивается сверху выражением, аналогичным правой части выражения (1.7.10). Таким образом, оценка (1.7.9) представляется как:

$$\begin{aligned} \mathbf{M} [|H(u, \Delta u_k, X)| I \{ |H(u, \Delta u_k, X)| > 2c \}] &\leq \\ &\leq 2 \cdot \mathbf{M} \left[\frac{1}{\Delta u_k} \int_{u_k}^{u_k + \Delta u_k} h(\beta, \bar{u}_k, X) I \{ |h(\beta, \bar{u}_k, X)| \geq c \} d\beta \right]. \end{aligned} \quad (1.7.11)$$

Поскольку математическое ожидание может быть представлено в форме интеграла, то по теореме

Фубини и в соответствии со свойствами интеграла верно:

$$\begin{aligned}
\mathbf{M} [|H(u, \Delta u_k, X)| I \{ |H(u, \Delta u_k, X)| > 2c \}] &\leq \\
&\leq 2 \cdot \frac{1}{\Delta u_k} \int_{u_k}^{u_k + \Delta u_k} \mathbf{M} [h(\beta, \bar{u}_k, X) I \{ |h(\beta, \bar{u}_k, X)| \geq c \}] d\beta \leq \\
&\leq 2 \sup_{\beta} \mathbf{M} [h(\beta, \bar{u}_k, X) I \{ |h(\beta, \bar{u}_k, X)| \geq c \}] \leq \\
&\leq 2 \sup_{\beta} \int_{x: I \{ |h(\beta, \bar{u}_k, x)| \geq c \}} |h(\beta, \bar{u}_k, x)| f(x) dx. \quad (1.7.12)
\end{aligned}$$

Поскольку интеграл в правой части на всей области \mathbb{R} сходится, а также поскольку

$$\mathbb{P} \{ |h(\beta, \bar{u}_k, X)| \geq c \} \rightarrow 0 \text{ при } c \rightarrow +\infty, \quad (1.7.13)$$

то

$$\int_{x: I \{ |h(\beta, \bar{u}_k, x)| \geq c \}} |h(\beta, \bar{u}_k, x)| f(x) dx \rightarrow 0 \text{ при } c \rightarrow +\infty. \quad (1.7.14)$$

Таким образом,

$$\mathbf{M} [|H(u, \Delta u_k, X)| I \{ |H(u, \Delta u_k, X)| > 2c \}] \rightarrow 0 \text{ при } c \rightarrow +\infty, \quad (1.7.15)$$

что доказывает равномерную интегрируемость семейства функций $\{H(u, \Delta u_k, X) : u \in U\}$. Это, в свою очередь, означает, что

$$\lim_{\Delta u_k \rightarrow 0} \mathbf{M} [H(u, \Delta u_k, X)] = \mathbf{M} \left[\lim_{\Delta u_k \rightarrow 0} H(u, \Delta u_k, X) \right]. \quad (1.7.16)$$

Функция $H(u, \Delta u_k, X)$ по определению (1.7.2) представляет собой отношение приращения значения функции $S_\theta(\varphi - \Phi(u_k, \bar{u}_k, X))$ к приращению аргумента Δu_k ; таким образом, выражение (1.7.16) может быть преобразовано с помощью перехода от пределов к частным производным. В результате перехода с учетом свойств математического ожидания, получаем:

$$\frac{\partial}{\partial u_k} \mathbf{M} [S_\theta(\varphi - \Phi(u_k, \bar{u}_k, X))] = \mathbf{M} [h(u_k, \bar{u}_k, X)]. \quad (1.7.17)$$

Переходя к принятым ранее обозначениям, получим:

$$\frac{\partial}{\partial u_k} P_\varphi^\theta(u) = \mathbf{M} [-\theta S_\theta(\varphi - \Phi(u, X)) (1 - S_\theta(\varphi - \Phi(u, X))) \Phi'_{u_k}(u, X)], \quad (1.7.18)$$

или, в конечном виде:

$$\frac{\partial P_\varphi^\theta(u)}{\partial u_k} = \int_{-\infty}^{+\infty} \theta S_\theta(\varphi - \Phi(u, x)) (S_\theta(\varphi - \Phi(u, x)) - 1) \Phi'_{u_k}(u, x) f(x) dx. \quad (1.7.19)$$

Данную формулу можно использовать при решении задачи (1.1.2) с помощью методов оптимизации первого порядка.

1.8 Аппроксимированная производная функции вероятности по уровню потерь

Производная аппроксимации функции вероятности по уровню потерь рассматривается аналогично производной по компоненте стратегии. Эта производная, в первую очередь, необходимо для получения производной функции квантили.

Рассмотрим теорему о сходимости аппроксимированной производной по уровню потерь к производной оригинальной функции. Общий ход доказательства аналогичен таковому в доказательстве Теоремы 4, но имеет некоторые отличия.

Теорема 5. Пусть X — абсолютно непрерывная случайная величина со значениями $x \in \mathbb{R}$ на полном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$. Пусть задано непустое множество $U \subseteq \mathbb{R}^m$ допустимых стратегий u . Пусть задана непрерывная дифференцируемая по совокупности аргументов строго кусочно-монотонная по x для любого $u \in U$ функция потерь $\Phi(u, x) : U \times \mathbb{R} \rightarrow \mathbb{R}$, с ограниченными непрерывными производными по компонентам стратегии u для любого x , а также задан уровень потерь φ . Пусть также выражениями (1.4.11) и (1.3.4) заданы функции вероятности $P_\varphi(u)$ и ее аппроксимация $P_\varphi^\theta(u)$. Тогда последовательность частных производных по уровню потерь φ аппроксимированной вероятностной функции стремится к частной производной по уровню потерь φ исходной вероятностной функции везде, где производная функции вероятности определена, т.е.

$$\frac{\partial P_\varphi^\theta(u)}{\partial \varphi} \rightarrow \frac{\partial P_\varphi(u)}{\partial \varphi} \quad \text{при } \theta \rightarrow +\infty \quad \forall u \in U. \quad (1.8.1)$$

Доказательство. Зафиксируем некоторую стратегию $u \in U$. Поскольку $\Phi(u, x)$ является гладкой кусочно-монотонной по x функцией, то выражение для функции вероятности $P_\varphi(u)$ можно представить в виде суммы интегралов особого вида:

$$P_\varphi(u) = \sum_i \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} \Theta(\varphi - \Phi(u, x)) f(x) dx, \quad (1.8.2)$$

где $x_i^-(\varphi)$ и $x_i^+(\varphi)$ – пара соседних корней уравнения:

$$\varphi - \Phi(u, x) = 0. \quad (1.8.3)$$

Отдельно рассмотрим случай, когда уравнение $\Phi(u, x) = \varphi$ не имеет корней.

Случай отсутствия корней уравнения $\Phi(u, x) = \varphi$

Пусть $\forall u \in U$ уравнение $\Phi(u, x) = \varphi$ не имеет решений. В этом случае функция вероятности $P_\varphi(u)$ при любых $u \in U$ константна и ее производная равняется нулю.

Рассмотрим аппроксимацию функции вероятности $P_\varphi^\theta(u)$. По правилу Лейбница:

$$\frac{\partial P_\varphi^\theta(u)}{\partial \varphi} = \frac{\partial}{\partial \varphi} \int_{-\infty}^{+\infty} S_\theta(\varphi - \Phi(u, x)) f(x) dx = \int_{-\infty}^{+\infty} S'_\theta(\varphi - \Phi(u, x)) f(x) dx. \quad (1.8.4)$$

Аргумент производной сигмоиды $\varphi - \Phi(u, x)$ не обращается в ноль на всей области интегрирования, поэтому в соответствии со Следствием 2 интеграл в правой части стремится к нулю при $\theta \rightarrow +\infty$, что влечет выполнение (1.8.1).

Случай наличия не менее одного корня уравнения $\Phi(u, x) = \varphi$

Итак, функция вероятности $P_\varphi(u)$ представляется в виде суммы интегралов функции Хевисайда в пределах от $x_i^-(\varphi)$ до $x_i^+(\varphi)$ – пары соседних корней уравнения $\varphi - \Phi(u, x) = 0$. Это уравнение задает неявную функцию, соответственно, по теореме о неявной функции можно найти производную границ интегрирования $x_i^-(\varphi)$ и $x_i^+(\varphi)$ по уровню потерь φ :

$$\frac{\partial x_i^-(\varphi)}{\partial \varphi} = \frac{1}{\Phi'_x(u, x_i^-(\varphi))}, \quad \frac{\partial x_i^+(\varphi)}{\partial \varphi} = \frac{1}{\Phi'_x(u, x_i^+(\varphi))}, \quad (1.8.5)$$

где

$$\Phi'_x(u, t) = \left. \frac{\partial}{\partial x} \Phi(u, x) \right|_{x=t}. \quad (1.8.6)$$

Запишем производную исходной функции вероятности $P_\varphi(u)$ по уровню потерь φ :

$$\frac{\partial}{\partial \varphi} P_\varphi(u) = \sum_i \frac{\partial}{\partial \varphi} \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} \Theta(\varphi - \Phi(u, x)) f(x) dx. \quad (1.8.7)$$

На каждом из интервалов интегрирования $[x_i^-(\varphi), x_i^+(\varphi)]$ возможны два варианта взаимного расположения значений функции потерь $\Phi(u, x)$ и уровня потерь φ ; рассмотрим оба случая.

Случай 1. $\Phi(u, x) \geq \varphi$.

В этом случае значение функции Хевисайда $\Theta(\varphi - \Phi(u, x))$ равно нулю, следовательно,

интеграл (1.8.7) также равен нулю:

$$\int_{x_i^-(\varphi)}^{x_i^+(\varphi)} \Theta(\varphi - \Phi(u, x))f(x)dx = 0, \quad (1.8.8)$$

и его производная на всем промежутке тоже равна нулю:

$$\frac{\partial}{\partial \varphi} \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} \Theta(\varphi - \Phi(u, x))f(x)dx = 0. \quad (1.8.9)$$

Случай 2. $\Phi(u, x) < \varphi$.

В этом случае значение функции Хевисайда $\Theta(\varphi - \Phi(u, x))$ равно единице:

$$\int_{x_i^-(\varphi)}^{x_i^+(\varphi)} \Theta(\varphi - \Phi(u, x))f(x)dx = \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} f(x)dx. \quad (1.8.10)$$

Согласно формуле Лейбница и с учетом других соотношений производная интеграла по φ представляется как:

$$\begin{aligned} \frac{\partial}{\partial \varphi} \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} f(x)dx &= f(x_i^+(\varphi)) \frac{\partial x_i^+(\varphi)}{\partial \varphi} - f(x_i^-(\varphi)) \frac{\partial x_i^-(\varphi)}{\partial \varphi} = \\ &= \frac{f(x_i^+(\varphi))}{\Phi'_x(u, x_i^+(\varphi))} - \frac{f(x_i^-(\varphi))}{\Phi'_x(u, x_i^-(\varphi))}. \end{aligned} \quad (1.8.11)$$

Проведем аналогичные выкладки для аппроксимированной функции вероятности $P_\varphi^\theta(u)$. Рассматриваемая производная определяется аналогично производной исходной функции. В итоге получим:

$$\frac{\partial}{\partial \varphi} P_\varphi^\theta(u) = \sum_i \frac{\partial}{\partial \varphi} \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} S_\theta(\varphi - \Phi(u, x))f(x)dx. \quad (1.8.12)$$

Рассмотрим аналогичные варианты взаимного расположения значений функции потерь $\Phi(u, x)$ и уровня потерь φ для интеграла (1.8.12):

Случай 1. $\Phi(u, x) \geq \varphi$.

В этом случае, исходя из свойств сигмоиды, верно:

$$\int_{x_i^-(\varphi)}^{x_i^+(\varphi)} S_\theta(\varphi - \Phi(u, x))f(x)dx \rightarrow 0 \quad \text{при } \theta \rightarrow +\infty. \quad (1.8.13)$$

Производная этого интеграла стремится к нулю по Следствию 2:

$$\frac{\partial}{\partial \varphi} \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} S_\theta(\varphi - \Phi(u, x))f(x)dx = \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} S'_\theta(\varphi - \Phi(u, x))f(x)dx \rightarrow 0 \quad \text{при } \theta \rightarrow +\infty, \quad (1.8.14)$$

что сопоставимо с выражением (1.8.9).

Случай 2. $\Phi(u, x) < \varphi$.

В этом случае, согласно формуле Лейбница:

$$\begin{aligned} \frac{\partial}{\partial \varphi} \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} S_\theta(\varphi - \Phi(u, x))f(x)dx &= S_\theta(\varphi - \Phi(u, x_i^+(\varphi)))f(x_i^+(\varphi))\frac{\partial x_i^+(\varphi)}{\partial \varphi} - \\ &- S_\theta(\varphi - \Phi(u, x_i^-(\varphi)))f(x_i^-(\varphi))\frac{\partial x_i^-(\varphi)}{\partial \varphi} + \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} S'_\theta(\varphi - \Phi(u, x))f(x)dx. \end{aligned} \quad (1.8.15)$$

Согласно свойствам сигмоиды, значения множителей $S_\theta(\varphi - \Phi(u, x_i^+(\varphi)))$ и $S_\theta(\varphi - \Phi(u, x_i^-(\varphi)))$ стремятся к единице при $\theta \rightarrow +\infty$. Последнее слагаемое, согласно Следствию 2, стремится к нулю при $\theta \rightarrow +\infty$. Следовательно, итоговое выражение сопоставимо с выражением (1.8.11).

Таким образом, для каждого промежутка $[x_i^-(\varphi), x_i^+(\varphi)]$ выполняется:

$$\frac{\partial}{\partial \varphi} \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} S_\theta(\varphi - \Phi(u, x))f(x)dx \rightarrow \frac{\partial}{\partial \varphi} \int_{x_i^-(\varphi)}^{x_i^+(\varphi)} \Theta(\varphi - \Phi(u, x))f(x)dx, \quad (1.8.16)$$

что эквивалентно доказываемому утверждению. \square

Выкладки для получения удобной формулы вычисления производной аппроксимации функции вероятности по уровню потерь аналогичны таковым, представленным в пункте 1.7. Далее приведен сокращенный ход получения требуемой формулы.

Рассматривается математическое ожидание функции $H(\Delta\varphi, u, X)$:

$$H(\Delta\varphi, u, X) = \frac{S_\theta(\varphi + \Delta\varphi - \Phi(u, X)) - S_\theta(\varphi - \Phi(u, X))}{\Delta\varphi}, \quad (1.8.17)$$

где $\Delta\varphi > 0$. Сигмоидальная функция представляется как:

$$S_\theta(\varphi - \Phi(u, x)) = \int_{-\infty}^{\varphi} h(\beta, u, x) d\beta, \quad (1.8.18)$$

где

$$h(\beta, u, x) = \theta S_\theta(\beta - \Phi(u, x)) (1 - S_\theta(\beta - \Phi(u, x))). \quad (1.8.19)$$

Аналогично выкладкам пункта 1.7 доказывается равномерная интегрируемость семейства $\{H(\Delta\varphi, u, X)\}$, что означает

$$\frac{\partial}{\partial\varphi} \mathbf{M}[S_\theta(\varphi - \Phi(u, X))] = \mathbf{M}[h(\varphi, u, X)]. \quad (1.8.20)$$

Переходя к принятым ранее обозначениям, получим:

$$\frac{\partial}{\partial\varphi} P_\varphi^\theta(u) = \mathbf{M}[\theta S_\theta(\varphi - \Phi(u, X)) (1 - S_\theta(\varphi - \Phi(u, X)))], \quad (1.8.21)$$

или, в конечном виде:

$$\frac{\partial P_\varphi^\theta(u)}{\partial\varphi} = \int_{-\infty}^{+\infty} \theta S_\theta(\varphi - \Phi(u, x)) (1 - S_\theta(\varphi - \Phi(u, x))) f(x) dx, \quad (1.8.22)$$

1.9 Квантильный критерий. Аппроксимация

Согласно одномерному аналогу формулы представления градиента функции квантили из [8], справедливо равенство

$$\frac{\partial}{\partial u_j} \varphi_\alpha(u) = - \frac{\frac{\partial}{\partial u_j} P_\varphi(u) \Big|_{\varphi=\varphi_\alpha(u)}}{\frac{\partial}{\partial\varphi} P_\varphi(u) \Big|_{\varphi=\varphi_\alpha(u)}}. \quad (1.9.1)$$

Подставив в (1.9.1) выражения (1.7.19) и (1.8.22), получим выражение для аппроксимации производной квантили по компоненте стратегии:

$$\frac{\partial \varphi_\alpha^\theta(u)}{\partial u_j} = \frac{\int_{-\infty}^{+\infty} S_\theta(\varphi_\alpha(u) - \Phi(u, x)) (1 - S_\theta(\varphi_\alpha(u) - \Phi(u, x))) \Phi'_{u_j}(u, x) f(x) dx}{\int_{-\infty}^{+\infty} S_\theta(\varphi_\alpha(u) - \Phi(u, x)) (1 - S_\theta(\varphi_\alpha(u) - \Phi(u, x))) f(x) dx}. \quad (1.9.2)$$

Данное выражение может быть использовано при решении задачи (1.1.4) методами оптимизации первого порядка.

Согласно определению квантили и функции распределения, справедливо равенство:

$$F_X(\varphi_\alpha) = \alpha,$$

где φ_α — квантиль уровня α случайной величины X с функцией распределения $F_X(x)$ и плотностью $f_X(x)$. Продифференцируем равенство по α в обеих частях, получим:

$$f_X(\varphi_\alpha) \cdot \frac{d\varphi_\alpha}{d\alpha} = 1,$$

или

$$\frac{d\varphi_\alpha}{d\alpha} = \frac{1}{f_X(\varphi_\alpha)}.$$

Учитывая выражение (1.8.22), получим уравнение аппроксимации производной квантили по уровню для рассматриваемого случая

$$\frac{\partial \varphi_\alpha^\theta(u)}{\partial \alpha} = \frac{1}{\int_{-\infty}^{+\infty} \theta S_\theta(\varphi_\alpha(u) - \Phi(u, x)) (1 - S_\theta(\varphi_\alpha(u) - \Phi(u, x))) f(x) dx}, \quad (1.9.3)$$

Таким образом, оценив или определив значение квантили для заданных u и α , можно вычислить квантиль определенного уровня и использовать градиентные методы оптимизации для решения исходной задачи в квантильной постановке.

1.10 Обобщение на многомерный случай

Полученные в предыдущих пунктах соотношения могут быть обобщены на случай многомерной случайной величины. Рассмотрим многомерные аналоги некоторых теорем.

Теорема 6. Пусть X — случайный вектор со значениями $x \in \mathbb{R}^n$ на полном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$. Пусть задан уровень потерь φ , задано непустое множество $U \subseteq \mathbb{R}^m$ допустимых стратегий u и задана непрерывная по совокупности аргументов функция потерь $\Phi(u, x) : U \times \mathbb{R}^n \rightarrow \mathbb{R}$, такая, что

$$\mathbb{P}\{\Phi(u, X) = \varphi\} = 0 \quad \forall u \in U. \quad (1.10.1)$$

Тогда для сигмоидальной функции $S_\theta(\cdot)$ и функции Хевисайда $\Theta(\cdot)$ верно следующее:

$$S_\theta(\varphi - \Phi(u, X)) \xrightarrow{\text{п.п.}} \Theta(\varphi - \Phi(u, X)) \quad \text{при } \theta \rightarrow +\infty. \quad (1.10.2)$$

Доказательство. Требуемое утверждение доказывается аналогично Теореме 1 путем рассмотрения обеих функций на интервалах выполнения неравенств $\Phi(u, X) < \varphi$ и $\Phi(u, X) > \varphi$, и в точках

выполнения равенства $\Phi(u, X) = \varphi$. \square

Теорема 7. Пусть X — случайный вектор со значениями $x \in \mathbb{R}^n$ и кусочно-непрерывной по каждому из аргументов плотностью $f(x)$ и функцией распределения $F(x)$ на полном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$. Пусть задан уровень потерь φ , непустое множество $U \subseteq \mathbb{R}^m$ допустимых стратегий u . Пусть также задана непрерывная по совокупности аргументов функция потерь $\Phi(u, x) : U \times \mathbb{R} \rightarrow \mathbb{R}$, такая, что выполняется условие (1.10.1). Пусть заданы функция вероятности $F_\varphi^n(u)$ и ее аппроксимация $F_\varphi^{n,\theta}(u)$:

$$F_\varphi^n(u) = \mathbf{M}[\Theta(\varphi - \Phi(u, X))] = \int_{\mathbb{R}^n} \Theta(\varphi - \Phi(u, x)) dF(x) \quad (1.10.3)$$

$$F_\varphi^{n,\theta}(u) = \mathbf{M}[S_\theta(\varphi - \Phi(u, X))] = \int_{\mathbb{R}^n} S_\theta(\varphi - \Phi(u, x)) dF(x). \quad (1.10.4)$$

Тогда верно

$$F_\varphi^{n,\theta}(u) \rightarrow F_\varphi^n(u) \quad \text{при } \theta \rightarrow +\infty \quad \forall u \in U. \quad (1.10.5)$$

Доказательство. Требуемое утверждение доказывается аналогично Теореме 2 и следует из сходимости сигмоиды к функции Хевисайда по Теореме 6, мажорируемости сигмоиды единицей и теоремы Лебега о мажорируемой сходимости. \square

Далее стоит отметить, что доказательство пункта 1.7 о равномерной интегрируемости семейства функций $\{H(u, \Delta u_k, X), u \in U\}$, где

$$H(u, \Delta u_k, X) = \frac{S_\theta(\varphi - \Phi(u_k + \Delta u_k, \bar{u}_k, X)) - S_\theta(\varphi - \Phi(u_k, \bar{u}_k, X))}{\Delta u_k}, \quad (1.10.6)$$

для случая многомерной величины X полностью аналогично таковому для одномерной величины. Это значит, что производную аппроксимированной функции вероятности по компоненте u_j стратегии u можно представить в виде:

$$\frac{\partial F_\varphi^{n,\theta}(u)}{\partial u_j} = \int_{\mathbb{R}^n} \theta S_\theta(\varphi - \Phi(u, x)) (1 - S_\theta(\varphi - \Phi(u, x))) (-\Phi'_{u_j}(u, x)) dF(x). \quad (1.10.7)$$

То же самое справедливо и для производной аппроксимированной функции вероятности по уровню потерь φ :

$$\frac{\partial F_\varphi^{n,\theta}(u)}{\partial \varphi} = \int_{\mathbb{R}^n} \theta S_\theta(\varphi - \Phi(u, x)) (1 - S_\theta(\varphi - \Phi(u, x))) dF(x). \quad (1.10.8)$$

Теорема 8. Пусть X — случайный вектор со значениями $x \in \mathbb{R}^n$ и кусочно-непрерывной плотностью $f(x)$ и функцией распределения $F(x)$ на полном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$.

Пусть задан уровень потерь φ и непустое множество $U \subseteq \mathbb{R}^m$ допустимых стратегий u . Пусть также задана непрерывная и дважды дифференцируемая по совокупности аргументов функция потерь $\Phi(u, x) : U \times \mathbb{R} \rightarrow \mathbb{R}$, такая, что выполняется условие (1.10.1), а ее производные первого порядка по компонентам управления u ограничены и непрерывны. Пусть заданы функция вероятности $P_\varphi^n(u)$ и ее аппроксимация $P_\varphi^{n,\theta}(u)$ формулами (1.10.3) и (1.10.4). Тогда последовательность частных производных по компонентам u_k стратегии u аппроксимированной вероятностной функции стремится к частной производной по компонентам u_k стратегии u исходной вероятностной функции везде, где производная функции вероятности определена, т.е.

$$\frac{\partial P_\varphi^{n,\theta}(u)}{\partial u_k} \rightarrow \frac{\partial P_\varphi^n(u)}{\partial u_k} \quad \text{при } \theta \rightarrow +\infty, \quad k = \overline{1, m}, \quad \forall u \in U. \quad (1.10.9)$$

Доказательство. Доказательство проводится по принципу математической индукции при возрастании размерности n случайного вектора X .

База индукции, а именно доказательство (1.10.9) при $n = 1$, представлено в Теореме 2.

Предположим, что утверждение (1.10.9) верно для $n = \beta$. Рассмотрим условную плотность X при фиксированном значении $x_{\beta+1}^*$ компоненты $X_{\beta+1}$ случайного вектора X :

$$f_{(X|X_{\beta+1})}(x_1, \dots, x_\beta | x_{\beta+1}^*) = \frac{f(x_1, \dots, x_\beta, x_{\beta+1}^*)}{f_{X_{\beta+1}}(x_{\beta+1}^*)}, \quad (1.10.10)$$

где $f_{X_{\beta+1}}(\cdot)$ - безусловная плотность распределения компоненты $X_{\beta+1}$. Отметим, что в точках равенства знаменателя этого выражения нулю оно доопределяется нулем, что не оказывает влияния на дальнейшие рассуждения. Поскольку эта плотность является плотностью β -мерной случайной величины X , то справедлива следующая форма утверждения (1.10.9):

$$\begin{aligned} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} S'_\theta(\varphi - \Phi(u, \dots, x_\beta, x_{\beta+1}^*)) f_{(X|X_{\beta+1})}(x_1, \dots, x_\beta | x_{\beta+1}^*) dx_1 \dots dx_\beta \rightarrow \\ \rightarrow \frac{\partial \mathbb{P} \{ \Phi(u, X_1, \dots, X_\beta, X_{\beta+1}) < \varphi | X_{\beta+1} = x_{\beta+1}^* \}}{\partial u_k}. \end{aligned} \quad (1.10.11)$$

Применим формулу полной вероятности к обеим частям этого выражения для перехода от условных плотностей и условных вероятностей по компоненте $X_{\beta+1}$ случайного вектора X к

обычным плотности и вероятности. После применения формулы в левой части получим:

$$\int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} S'_\theta(\varphi - \Phi(u, x_1, \dots, x_\beta, x_{\beta+1}^*)) f_{(X|X_{\beta+1})}(x_1, \dots, x_\beta | x_{\beta+1}^*) dx_1 \dots dx_\beta \right\} \cdot f_{X_{\beta+1}}(x_{\beta+1}) dx_{\beta+1}. \quad (1.10.12)$$

что, в сущности, есть записанное в ином виде выражение (1.10.7) для $n = \beta + 1$. В правой части (1.10.11) после применения формулы полной вероятности получится, в сущности, определение производной $\frac{\partial P_\varphi^{\beta+1}(u)}{\partial u_k}$.

Отсюда, применяя к (1.10.12) теорему Лебега о мажорируемой сходимости и переходя обратно к кратким определениям, следует, что:

$$\frac{\partial P_\varphi^{\beta+1, \theta}(u)}{\partial u_k} \rightarrow \frac{\partial P_\varphi^{\beta+1}(u)}{\partial u_k}, \quad (1.10.13)$$

что завершает доказательство по принципу математической индукции. \square

Аналогичным образом доказывается многомерный аналог теоремы о сходимости аппроксимированной производной функции вероятности по уровню потерь к оригинальной производной.

Теорема 9. Пусть X — случайный вектор со значениями $x \in \mathbb{R}^n$ и кусочно-непрерывной плотностью $f(x)$ и функцией распределения $F(x)$ на полном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$. Пусть задан уровень потерь φ и непустое множество $U \subseteq \mathbb{R}^m$ допустимых стратегий u . Пусть также задана непрерывная по совокупности аргументов функция потерь $\Phi(u, x) : U \times \mathbb{R} \rightarrow \mathbb{R}$, такая, что выполняется условие (1.10.1), а ее производные первого порядка по компонентам управления u ограничены и непрерывны. Пусть заданы функция вероятности $P_\varphi^n(u)$ и ее аппроксимация $P_\varphi^{n, \theta}(u)$ формулами (1.10.3) и (1.10.4). Тогда последовательность частных производных по уровню потерь φ аппроксимированной вероятностной функции стремится к частной производной по уровню потерь φ исходной вероятностной функции везде, где производная функции вероятности определена, т.е.

$$\frac{\partial P_\varphi^{n, \theta}(u)}{\partial \varphi} \rightarrow \frac{\partial P_\varphi^n(u)}{\partial \varphi} \quad \text{при } \theta \rightarrow +\infty, \quad k = \overline{1, m}, \quad \forall u \in U. \quad (1.10.14)$$

Доказательство. Доказательство проводится по принципу математической индукции при возрастании размерности n случайного вектора X .

База индукции доказана в Теореме 5.

Доказательство шага индукции полностью аналогично доказательству Теоремы 7. \square

Таким образом, в многомерном случае выражения для аппроксимированной производной функции вероятности по компоненте вектора стратегии u_j и по уровню потерь φ также можно записать в виде объемных интегралов. Например, при наличии двумерной случайной величины $X = [X_1, X_2]$ и функции потерь $g(u, X_1, X_2)$, удовлетворяющих условиям Теорем 6-9, искомые производные можно записать как:

$$\frac{\partial \mathbb{P} \{g(u, X_1, X_2) \leq \varphi\}}{\partial u_j} \approx \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \theta S_\theta(\varphi - g(u, x_1, x_2)) \cdot (1 - S_\theta(\varphi - g(u, x_1, x_2))) \frac{\partial g(u, x_1, x_2)}{\partial u_j} f(x_1, x_2) dx_1 dx_2, \quad (1.10.15)$$

$$\frac{\partial \mathbb{P} \{g(u, X_1, X_2) \leq \varphi\}}{\partial \varphi} \approx \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \theta S_\theta(\varphi - g(u, x_1, x_2)) \cdot (S_\theta(\varphi - g(u, x_1, x_2)) - 1) f(x_1, x_2) dx_1 dx_2. \quad (1.10.16)$$

1.11 Выводы по главе 1

Основным объектом исследования является метод аппроксимации вероятностных критеральных функций в задачах стохастического программирования, основанный на сигмоидальной функции.

В ходе исследования аппроксимации были получены следующие основные результаты:

- Доказана теорема о сходимости аппроксимации функции вероятности к оригинальной функции в одномерном случае;
- Доказаны теоремы о сходимости производных аппроксимированной функции вероятности по компонентам вектора стратегии и уровню потерь к оригинальным производным в одномерном случае;
- На основе теорем о сходимости аппроксимированной функции вероятности и ее производных доказаны аналогичные теоремы о сходимости аппроксимированной функции вероятности и ее производных к оригинальным функциям в случае многомерной случайной величины;
- Получены новые выражения для аппроксимации функции вероятности, ее производных и производных функции квантили в форме объемных интегралов.
- Получена верхняя оценка погрешности аппроксимации функции вероятности и доказано ее стремление к нулю с ростом параметра сигмоидальной функции.

Полученные оценки значений функции вероятности и ее производной могут быть использованы при решении задач стохастического программирования при помощи численных методов оптимизации первого порядка, особенно в случаях, когда другие методы могут быть не применимы. Основным преимуществом такого способа решения является относительная вычислительная простота по сравнению, например, с доверительным методом, широко применяющимся при решении задач стохастического программирования. Вычисление градиента функции вероятности указанным способом также является менее трудоемким по сравнению с вычислением градиента в форме поверхностного интеграла.

Основные результаты главы 1 опубликованы в работах [25, 26, 33, 34, 77].

2. Гладкая аппроксимация вероятностных функций в теоретических задачах стохастического программирования

В главе 2 исследуются возможности применения гладкой аппроксимации функции вероятности при исследовании свойств вероятностных критериев и задач стохастического программирования. В разделе 2.1 предложен алгоритм построения внешней аппроксимации ядра вероятностной меры для случаев двумерного и трехмерного случайного вектора. В разделе 2.2 предложен подход к аппроксимации вторых производных функции вероятности и квантили, основанный на непосредственном дифференцировании выражений для гладкой аппроксимации производных первого порядка. Доказывается сходимость производных функции квантили первого и второго порядка по элементам вектора стратегии в отдельных случаях, когда для соответствующей задачи стохастического программирования существует детерминированный эквивалент. В разделе 2.3 предложен алгоритм решения задачи максимизации функции вероятности с полиэдральной функцией потерь и ограничений. В каждом из разделов приводятся численные примеры, подтверждающие корректность предложенных аппроксимаций и алгоритмов.

2.1 Построение аппроксимации ядра вероятностной меры

Введем необходимые определения и обозначения.

Определение 1. Множество $S \in \mathbb{R}^n$ называется α -доверительным множеством для случайного вектора X с реализациями $x \in \mathbb{R}^n$, если $\mathbb{P}\{X \in S\} \geq \alpha$, и α -супердоверительным, если $\mathbb{P}\{X \in S\} > \alpha$.

Определение 2. Пусть X — n -мерный случайный вектор. Пересечение всех выпуклых и замкнутых α -доверительных множеств в \mathbb{R}^n называется α -ядром вероятностной меры (распределения вектора X).

Очевидно, что α -ядро — выпуклое, замкнутое и ограниченное множество, то есть оно является выпуклым компактом. Установлено, что α -ядро K_α может быть определено как пересечение всех замкнутых α -доверительных полупространств:

$$K_\alpha = \bigcap_{\|c\|=1} \{x : c^T x \leq [c^T X]_\alpha\}.$$

Наиболее важным и принципиальным свойством α -ядра непрерывных распределений, является регулярность.

Определение 3. α -Ядро K_α называется регулярным, если всякое замкнутое полупространство, содержащее это ядро, является α -доверительным.

Достаточные условия регулярности α -ядра установлены в [8] в виде следующей теоремы.

Теорема 10. [8] Пусть выполнены следующие условия:

- случайный вектор X имеет плотность вероятности,
- граница α -ядра K_α для распределения вектора X является гладкой поверхностью в \mathbb{R}^n .

Тогда K_α является регулярным.

В частности, известно, что K_α регулярно для невырожденного многомерного нормального распределения вектора X при $\alpha \geq 1/2$.

Чтобы обобщить теорему 10 необходимо найти множество S_α всех точек x^* на границе α -ядра, в которых вектор внешней нормали к K_α в точках из S_α не является единственным.

Теорема 11. [8] Пусть выполнены следующие условия:

- случайный вектор X имеет плотность вероятности,
- $\mathbb{P}\{c^T X \leq c^T x^*\} \geq \alpha$ для любого $x^* \in S_\alpha$ и любого единичного нормального вектора c .

Тогда K_α является регулярным.

Теорема 12. [1] α -Ядро K_α n -мерного случайного вектора X непусто для любого $\alpha \in (\frac{n}{n+1}; 1)$

Теорема 13. [1] Пусть выполнены следующие условия:

- случайный вектор X имеет плотность вероятности и векторную медиану μ ,
- $\mathbb{P}\{\mu_k < X_k < \mu_k + \varepsilon\} > 0$ и $\mathbb{P}\{\mu_k - \varepsilon < X_k < \mu_k\} > 0$ для любого $\varepsilon > 0$ и любых компонент X_k вектора X и компонент μ_k вектора μ , $k = \overline{1, n}$,
- α -Ядро K_α непусто для любого $\alpha \in (1/2; 1)$.

Тогда при $\alpha = 0.5$ α -ядро K_α является регулярным и содержит векторную медиану μ в качестве единственной точки.

2.1.1 Двумерный случай

Для построения аппроксимации границы α -ядра составим и решим дифференциальное уравнение специального вида. Рассмотрим функцию, определяющую вероятность попадания случайного вектора X в полуплоскость, заданную вектором коэффициентов $c = [c_1, c_2]$:

$$P(c) = \mathbb{P}\{c_1 X_1 + c_2 X_2 \leq 1\}. \quad (2.1.1)$$

Рассмотрим задачу построения α -ядра для двумерного случайного вектора $X = (X_1, X_2)^T$ с

плотностью распределения $f(x_1, x_2)$. Значение функции $P(c)$ может быть вычислено как интеграл

$$P(c) = \iint_{c_1x_1+c_2x_2 \leq 1} f(x_1, x_2)dx_2dx_1 = \iint_{G_X} I(c, x_1, x_2)f(x_1, x_2)dx_2dx_1, \quad (2.1.2)$$

где $G_X = \text{supp}(X)$, а $I(c, x_1, x_2)$ – индикаторная функция:

$$I(c, x_1, x_2) = \begin{cases} 1, & c_1x_1 + c_2x_2 \leq 1, \\ 0, & c_1x_1 + c_2x_2 > 1. \end{cases} \quad (2.1.3)$$

Для того, чтобы определить границу ядра необходимо найти изокванту функции $P(c)$, то есть ее поверхность уровня. Изокванта будет содержать семейство векторов $c = [c_1, c_2]$, при которых значение функции вероятности постоянно. Множество таких векторов будет удовлетворять уравнению

$$\frac{\partial P(c)}{\partial c_1}dc_1 + \frac{\partial P(c)}{\partial c_2}dc_2 = 0. \quad (2.1.4)$$

Это уравнение эквивалентно следующему:

$$\frac{dc_2}{dc_1} = -\frac{\partial P(c)}{\partial c_1} \left(\frac{\partial P(c)}{\partial c_2} \right)^{-1}. \quad (2.1.5)$$

Чтобы упростить вычисление частных производных в правой части уравнения, предлагается перейти к гладкой аппроксимации функции вероятности, описанной ранее:

$$P_\theta(c) = \iint_{G_X} S_\theta(1 - c_1x_1 - c_2x_2)f(x_1, x_2)dx_2dx_1. \quad (2.1.6)$$

Предположим, что ядро вероятностной меры регулярно и содержит точку 0. Этого можно добиться, если рассмотреть смещенный случайный вектор, векторная медиана которого совпадает с точкой 0. Для построения границы множества перейдем к полярным координатам:

$$c_1 = r \cos t, \quad c_2 = r \sin t. \quad (2.1.7)$$

Запишем выражение для приближенной функции вероятности в полярных координатах:

$$P_\theta(r, t) = \iint_{G_X} S_\theta(1 - x_1r \cos t - x_2r \sin t)f(x_1, x_2)dx_2dx_1. \quad (2.1.8)$$

Частные производные по r и t равны соответственно:

$$\frac{\partial P_\theta(r, t)}{\partial r} = - \iint_{G_X} (x_1 \cos t + x_2 \sin t)S'_\theta(1 - x_1r \cos t - x_2r \sin t)f(x_1, x_2)dx_2dx_1, \quad (2.1.9)$$

$$\frac{\partial P_\theta(r, t)}{\partial t} = - \iint_{G_x} r(x_1 \cos t - x_2 \sin t) S'_\theta(1 - x_1 r \cos t - x_2 r \sin t) f(x_1, x_2) dx_2 dx_1. \quad (2.1.10)$$

Переформулируем уравнение (2.1.5) с учетом введенной аппроксимации и после перехода к новым координатам:

$$\frac{dr}{dt} = - \frac{\partial P_\theta(r, t)}{\partial t} \left(\frac{\partial P_\theta(r, t)}{\partial r} \right)^{-1}. \quad (2.1.11)$$

Опишем теперь алгоритм построения границы ядра.

Алгоритм 1. Алгоритм построения внешней аппроксимации границы α -ядра.

- 1) Задать начальные значения $t_0 = 0$, $r_0 = ([X_1]_\alpha)^{-1}$ и шаг Δt ;
- 2) Вычислить частные производные функции вероятности при текущих t_k и r_k при помощи формул (2.1.9) и (2.1.10);
- 3) Вычислить значение производной $\frac{dr}{dt}$ по формуле (2.1.11) и определить значение r_{k+1} каким-либо способом (например, методом Эйлера);
- 4) Задать $t_{k+1} = t_k + \Delta k$;
- 5) Повторять шаги 2-4 пока $t_{k+1} \leq 2\pi$.

Результатом работы алгоритма является заданная по точкам функция, аппроксимирующая внешнюю границу ядра. После обратного перехода от полярных координат к декартовым получаем набор векторов с компонентами c_1 и c_2 . Точность построения аппроксимации в данном случае зависит не только от параметров аппроксимации, но и от метода приближения следующего значения функции r_{k+1} , например, модификация Алгоритма 1, при которой значение r_{k+1} вычисляется на основе формул Адамса-Башфорта или аналогично методам Рунге-Кутты, дает меньшую кумулятивную ошибку по сравнению с методом Эйлера.

2.1.2 Примеры построения аппроксимированной внешней границы ядра для двумерных распределений

Зафиксируем параметры алгоритма:

$$\alpha = 0,8, \theta = 10, \Delta t = 0,01, t_0 = 0. \quad (2.1.12)$$

Значение функции r_{k+1} на каждом шаге алгоритма будем вычислять методом Эйлера.

Для случайного вектора X , равномерно распределенного на квадрате:

$$X \sim \mathbf{R}([-1, 1] \times [-1, 1]), r_0 = 1,667 \quad (2.1.13)$$

аппроксимация внешней границы ядра представлена на рисунке 2.1.

Для случайного вектора Y с нормальным распределением:

$$Y \sim \mathbf{N} \left((0, 0)^T, \begin{pmatrix} 2 & 0,7 \\ 0,7 & 1 \end{pmatrix} \right), \quad r_0 = 0,8402. \quad (2.1.14)$$

аппроксимация внешней границы ядра представлена на рисунке 2.2.

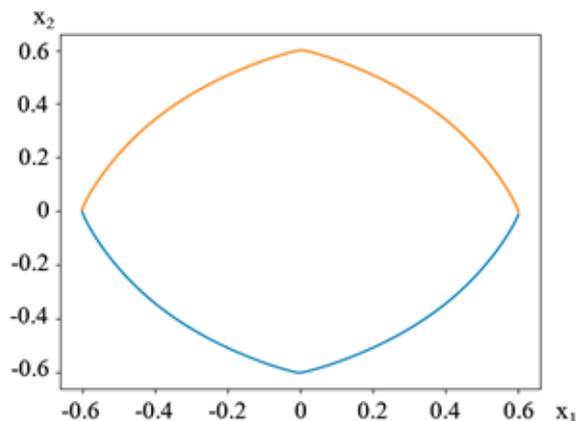


Рисунок 2.1. Внешняя аппроксимация границы α -ядра для равномерного распределения на квадрате

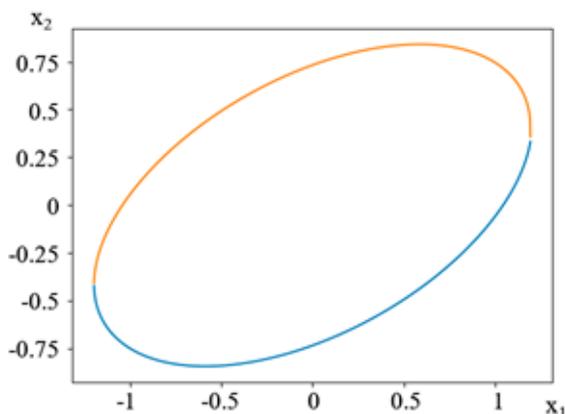


Рисунок 2.2. Внешняя аппроксимация границы α -ядра для двумерного нормального распределения

Стоит отметить, что в случае достаточно большого доверительного уровня α , алгоритм Эйлера для вычисления текущего значения функции r_{k+1} может давать достаточно большую кумулятивную ошибку, что может приводить к видимому «разрыву» границы аппроксимации. В этом случае можно доопределить границу ядра в точках «разрыва» каким-либо способом или использовать модификацию Алгоритма 1 с методом приближения значения функции большей точности.

В работах [1] и [8] были получены истинные виды границ ядер для рассматриваемых распределений: граница ядра для равномерного распределения состоит из четырех гиперболических сегментов, а граница ядра для двумерного гауссовского вектора представляет собой эллипсоид.

Полученные на рисунках 2.1 и 2.2 изображения внешних аппроксимация границ ядра неотличимы от известных истинных результатов.

2.1.3 Трехмерный случай

Рассмотрим задачу построения α -ядра для трехмерного случайного вектора $X = (X_1, X_2, X_3)^T$ с плотностью распределения $f(x_1, x_2, x_3)$. Также рассмотрим специальную функцию, определяющую вероятность попадания случайного вектора X в полуплоскость, заданную вектором коэффициентов $c = [c_1, c_2, c_3]$:

$$P(c) = \mathbb{P}\{c_1X_1 + c_2X_2 + c_3X_3 \leq 1\}. \quad (2.1.15)$$

Значение функции $P(c)$ может быть вычислено как интеграл

$$P(c) = \iiint_{c_1x_1 + c_2x_2 + c_3x_3 \leq 1} f(x_1, x_2, x_3) dx_3 dx_2 dx_1 = \iiint_{G_X} I(c, x_1, x_2, x_3) f(x_1, x_2, x_3) dx_3 dx_2 dx_1, \quad (2.1.16)$$

где $G_X = \text{supp}(X)$, а $I(c, x_1, x_2, x_3)$ – индикаторная функция:

$$I(c, x_1, x_2, x_3) = \begin{cases} 1, & c_1x_1 + c_2x_2 + c_3x_3 \leq 1, \\ 0, & c_1x_1 + c_2x_2 + c_3x_3 > 1. \end{cases} \quad (2.1.17)$$

Для того, чтобы определить границу ядра необходимо найти изокванту функции $P(c)$, то есть ее поверхность уровня. Изокванта будет содержать семейство векторов $c = [c_1, c_2, c_3]$, при которых значение функции вероятности постоянно. Множество таких векторов будет удовлетворять уравнению

$$\frac{\partial P(c)}{\partial c_1} dc_1 + \frac{\partial P(c)}{\partial c_2} dc_2 + \frac{\partial P(c)}{\partial c_3} dc_3 = 0. \quad (2.1.18)$$

Для упрощения вычисления производных перейдем к гладкой аппроксимации функции вероятности:

$$\begin{aligned} P_\theta(c) &= \iiint_{G_X} S_\theta(1 - c_1x_1 - c_2x_2 - c_3x_3) f(x_1, x_2, x_3) dx_3 dx_2 dx_1 = \\ &= \mathbf{M}[S_\theta(1 - c_1X_1 - c_2X_2 - c_3X_3)]. \end{aligned} \quad (2.1.19)$$

Предположим, что ядро вероятностной меры регулярно и содержит точку 0. Этого можно добиться, если рассмотреть смещенный случайный вектор, векторная медиана которого совпадает

с точкой 0. Для построения границы множества перейдем к сферическим координатам:

$$c_1 = r \sin \psi \cos t, \quad c_2 = r \sin \psi \sin t, \quad c_3 = r \cos \psi. \quad (2.1.20)$$

Запишем выражение для приближенной функции вероятности в полярных координатах:

$$P_\theta(r, \psi, t) = \mathbf{M} [S_\theta(1 - X_1 r \sin \psi \cos t - X_2 r \sin \psi \sin t - X_3 r \cos \psi)]. \quad (2.1.21)$$

Частные производные по r , ψ и t равны соответственно:

$$\frac{\partial P_\theta(r, \psi, t)}{\partial r} = -\mathbf{M} [(X_1 \sin \psi \cos t + X_2 \sin \psi \sin t + X_3 \cos \psi) S'_\theta(1 - X_1 r \sin \psi \cos t - X_2 r \sin \psi \sin t - X_3 r \cos \psi)] \quad (2.1.22)$$

$$\frac{\partial P_\theta(r, \psi, t)}{\partial t} = \mathbf{M} [r(X_1 \sin \psi \sin t - X_2 \sin \psi \cos t) S'_\theta(1 - X_1 r \sin \psi \cos t - X_2 r \sin \psi \sin t - X_3 r \cos \psi)], \quad (2.1.23)$$

$$\frac{\partial P_\theta(r, \psi, t)}{\partial \psi} = -\mathbf{M} [r(X_1 \cos \psi \cos t + X_2 \cos \psi \sin t - X_3 \sin \psi) S'_\theta(1 - X_1 r \sin \psi \cos t - X_2 r \sin \psi \sin t - X_3 r \cos \psi)], \quad (2.1.24)$$

Перепишем уравнение (2.1.18) с учетом введенной аппроксимации и перехода к новым координатам:

$$\frac{\partial P_\theta(r, \psi, t)}{\partial r} dr + \frac{\partial P_\theta(r, \psi, t)}{\partial \psi} d\psi + \frac{\partial P_\theta(r, \psi, t)}{\partial t} dt = 0. \quad (2.1.25)$$

Для построения аппроксимации ядра также необходимо определить набор векторов c , лежащих на границе множества уровня вспомогательной функции вероятности. Для направленного построения аппроксимирующего множества и его последующей визуализации построим спираль, проходящую по поверхности уровня функции вероятности. Для построения огибающей спирали составим алгоритм.

Алгоритм 2. Алгоритм построения спирали, огибающей набор векторов, определяющих аппроксимированную внешнюю границу α -ядра в трехмерном случае

- 1) Задать шаг $\Delta\psi > 0$ и $\Delta t > 0$. Для удобства можно положить $\Delta t = k\Delta\psi$, где k – натуральное число;
- 2) Задать начальные значения $t_0 = 0$, $\psi_0 = 0$, $r_0 = ([X_3]_\alpha)^{-1}$;
- 3) Вычислить частные производные функции вероятности (2.1.22-2.1.24) при текущих t_k , ψ_k

и r_k ;

4) Вычислить Δr в соответствии с уравнением (2.1.18) по следующей формуле:

$$\Delta r = -\frac{1}{\frac{\partial P_\theta(r, \psi, t)}{\partial r}} \left(\frac{\partial P_\theta(r, \psi, t)}{\partial \psi} \Delta \psi + \frac{\partial P_\theta(r, \psi, t)}{\partial t} \Delta t \right) \Big|_{r=r_k, t=t_k, \psi=\psi_k}. \quad (2.1.26)$$

5) Задать $r_{k+1} = r_k + \Delta r$, $t_{k+1} = t_k + \Delta k$, $\psi_{k+1} = \psi_k + \Delta \psi$;

6) Повторять шаги 2-5 пока $\psi_{k+1} \leq \pi$.

Выбор числа k в первом пункте алгоритма определяет количество витков спирали. Результатом работы алгоритма является заданная по точкам линия, огибающая витками поверхность векторов c , определяющих аппроксимацию ядра.

2.1.4 Пример построения аппроксимированной внешней границы ядра для трехмерного случайного вектора

Рассмотрим трехмерный гауссовский вектор $X = (X_1, X_2, X_3)^T$ с коррелированными компонентами:

$$X \sim \mathbf{N}((0, 0, 0)^T, K), \quad K = \begin{pmatrix} 2 & 0,5 & 0,1 \\ 0,5 & 1 & 0,3 \\ 0,1 & 0,3 & 1 \end{pmatrix} \quad (2.1.27)$$

Задав значения параметров Алгоритма 2 $\alpha = 0,9$, $\Delta \psi = \pi/180$, $k = 20$, в результате его работы получим огибающую спираль, изображенную на рисунке 2.3.

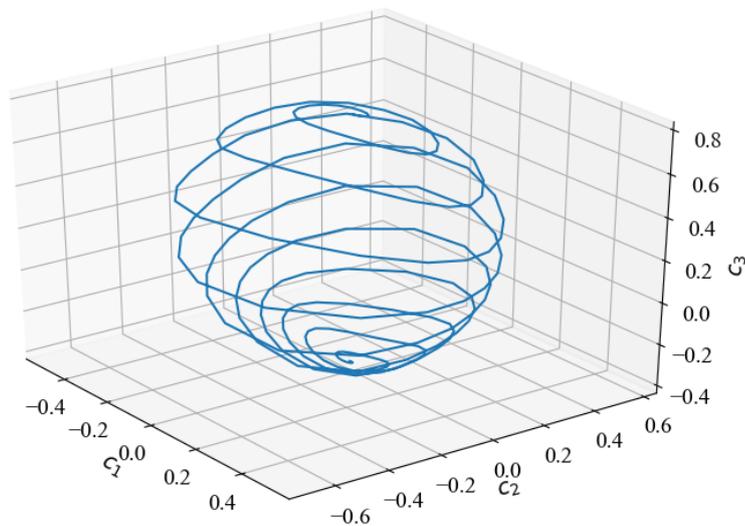


Рисунок 2.3. Огибающая поверхности векторов c

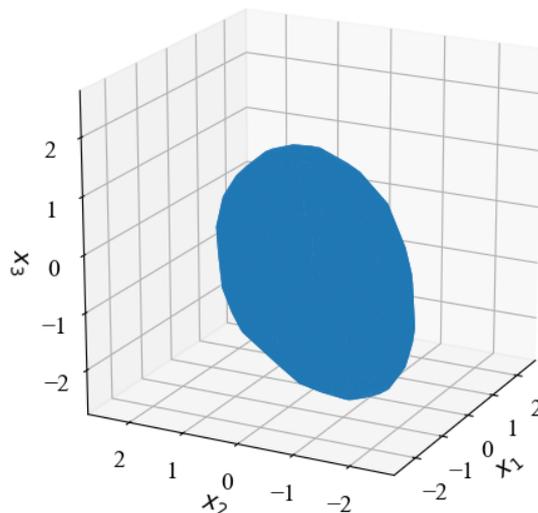


Рисунок 2.4. Аппроксимация α -ядра трехмерного нормального распределения

По множеству векторов c можно построить множество точек, принадлежащих α -ядру. Результат такого построения представлен на рисунке 2.4. В работах [1] и [8] было установлено, что границей ядра трехмерного нормального распределения является эллипсоид. Полученное изображение аппроксимированного ядра практически неотличимо от его истинного вида.

2.2 Аппроксимация вторых производных

Настоящий раздел посвящен вопросу применения численных методов оптимизации второго порядка в задачах стохастического программирования, включающих в себя функцию вероятности. Последняя может выступать в рассматриваемых задачах в качестве критерия оптимизации и/или в качестве функции ограничения. Известно, что функция вероятности не является выпуклой или вогнутой на всей области определения [8]. Как правило, функция вероятности исследуется на квазивыпуклость и используются свойства, следующие из её квазивыпуклости; данные свойства освещены в [7, 29, 45, 69]. Утверждения о свойствах квазивыпуклости функции вероятности опираются на понятия квазивогнутых и логарифмически вогнутых вероятностных мер, рассматриваемых в [20, 39, 70, 71]. Также, в других работах рассматриваются свойства множеств уровня функции вероятности, например, условия выпуклости множеств уровня для достаточно больших значений вероятности [88], условия связности данных множеств уровня [52].

В приведенных выше работах прямо не рассматриваются вопрос, смежный с анализом выпуклости функций или их оценкой – а именно вычисление или оценка значения производных функции вероятности. Как было отмечено ранее, основная сложность в вычислении производных функции вероятности заключается в функции Хевисайда, входящей в интегральное представление функции вероятности. Таким образом, при прямом дифференцировании функции вероятности

появляется дельта-функция Дирака, а производная в общем случае принимает вид поверхностного интеграла. В данном разделе получены выражения для вторых производных гладкой аппроксимации функции вероятности. Численные эксперименты показывают, что получаемые вторые производные гладкой аппроксимации сходятся ко вторым производным исходной функции вероятности при стремлении параметра в показателе экспоненты в сигмоиде к бесконечности. Вторые производные точной функции вероятности оценивались с помощью конечных разностей. Для сравнения близости аппроксимированного и точного значений производной рассмотрено три примера: с билинейной функцией потерь, квадратичной функцией потерь, а также с логарифмической функцией потерь. Доказательство сходимости аппроксимации вторых производных к точным значениям выходит за рамки представленной работы. В качестве примера применения алгоритма решения задачи с рассматриваемыми аппроксимациями приведено решение задачи портфельной оптимизации с логарифмической функцией потерь и вероятностным критерием.

Пусть условия и обозначения удовлетворяют таковым из Теоремы 2, и введено дополнительное условие на ограниченность смешанных частных производных второго порядка функции $\Phi(u, x)$ по компонентам вектора стратегии. Для компактной записи вторых производных функции вероятности введем обозначение для второй производной сигмоиды:

$$S''_{\theta}(x) = \theta^2 S_{\theta}(x)(1 - S_{\theta}(x))(1 - 2S_{\theta}(x)). \quad (2.2.1)$$

Тогда вторая смешанная частная производная функции вероятности по компонентам стратегии u_i и u_j , $i = \overline{1, m}$, $j = \overline{1, m}$ получается путем непосредственного дифференцирования формулы (1.7.19) и принимает вид:

$$\frac{\partial^2 P_{\varphi}^{\theta}(u)}{\partial u_i \partial u_j} = \int_G S''_{\theta}(\varphi - \Phi(u, x)) \Phi'_{u_i}(u, x) \Phi'_{u_j}(u, x) f(x) dx - \int_G S'_{\theta}(\varphi - \Phi(u, x)) \Phi''_{u_i u_j}(u, x) f(x) dx. \quad (2.2.2)$$

где $G = \text{supp}(X)$. Здесь стоит отметить, что, как и в случае с многомерными аналогами Теорем 1 и 2, доказательство пункта 1.7 о равномерной интегрируемости семейства функций - отношений приращения функции к аргументу - переносится на этот случай с минимальными изменениями.

Важным частным случаем является случай линейной или билинейной функции потерь. При этом вторые частные производные функции потерь будут равны нулю, и вторая смешанная частная производная примет более простой вид:

$$\frac{\partial^2 P_{\varphi}^{\theta}(u)}{\partial u_i \partial u_j} = \int_G S''_{\theta}(\varphi - \Phi(u, x)) \Phi'_{u_i}(u, x) \Phi'_{u_j}(u, x) f(x) dx. \quad (2.2.3)$$

Для аппроксимации производных функции квантили будет использоваться обозначение $\varphi_\alpha^\theta(u)$. В соответствии с формулой (1.9.3) аппроксимация производной функции квантили по уровню вероятности α принимает следующий вид:

$$\frac{\partial \varphi_\alpha^\theta(u)}{\partial \alpha} = \frac{1}{\mathbf{M}(S'_\theta(\varphi_\alpha(u) - \Phi(u, X)))}. \quad (2.2.4)$$

В формуле (1.9.1) ранее приводилось выражение производной функции квантили по элементам вектора стратегии через отношение частных производных функции вероятности:

$$\frac{\partial \varphi_\alpha(u)}{\partial u_i} = - \frac{\frac{\partial}{\partial u_i} P_\varphi(u) \Big|_{\varphi=\varphi_\alpha(u)}}{\frac{\partial}{\partial \varphi} P_\varphi(u) \Big|_{\varphi=\varphi_\alpha(u)}}. \quad (2.2.5)$$

Используя гладкую аппроксимацию функции вероятности и ее производных в выражении для производной функции квантили, получаем формулу (1.9.2):

$$\frac{\partial \varphi_\alpha^\theta(u)}{\partial u_i} = \frac{\mathbf{M}[S'_\theta(\varphi_\alpha(u) - \Phi(u, X))\Phi'_{u_i}(u, X)]}{\mathbf{M}[S'_\theta(\varphi_\alpha(u) - \Phi(u, X))]}, \quad (2.2.6)$$

записанную через отношение математических ожиданий.

Обозначим числитель и знаменатель как $V(u)$ и $W(u)$ соответственно:

$$V(u) \triangleq \mathbf{M}[S'_\theta(\varphi_\alpha(u) - \Phi(u, X))\Phi'_{u_i}(u, X)], \quad (2.2.7)$$

$$W(u) \triangleq \mathbf{M}[S'_\theta(\varphi_\alpha(u) - \Phi(u, X))]. \quad (2.2.8)$$

Аппроксимация второй частной производной функции квантили может быть определена следующим образом:

$$\frac{\partial^2 \varphi_\alpha^\theta(u)}{\partial u_i \partial u_j} = \frac{V'_{u_j}(u)W(u) - V(u)W'_{u_j}(u)}{W^2(u)}. \quad (2.2.9)$$

Производные функций $V(u)$ и $W(u)$ определяются непосредственно по правилам дифференцирования сложной функции:

$$V'_{u_j}(u) = \mathbf{M} \left[S'_\theta(\varphi_\alpha(u) - \Phi(u, X))\Phi''_{u_i u_j}(u, X) \right] + \mathbf{M} \left[S''_\theta(\varphi_\alpha(u) - \Phi(u, X))\Phi'_{u_i}(u, X) \left(\frac{\partial}{\partial u_j} \varphi_\alpha(u) - \Phi'_{u_j}(u, X) \right) \right], \quad (2.2.10)$$

$$W'_{u_j}(u) = \mathbf{M} \left[S''_\theta(\varphi_\alpha(u) - \Phi(u, X)) \left(\frac{\partial}{\partial u_j} \varphi_\alpha(u) - \Phi'_{u_j}(u, X) \right) \right]. \quad (2.2.11)$$

2.2.1 Анализ сходимости производных функции квантили

В данном разделе производится анализ сходимости гладкой аппроксимации производных функции квантили для 3 случаев, когда для соответствующей задачи минимизации функции квантили существует детерминированный эквивалент.

2.2.1.1 Сепарабельная функция потерь

Рассмотрим функцию потерь, представимую в виде суммы функции, зависящей только от вектора стратегии, и функции случайного вектора:

$$g(u, X) = g_1(u) + g_2(X). \quad (2.2.12)$$

Функция квантили в этом случае определяется следующим образом:

$$\varphi_\alpha(u) = g_1(u) + [g_2(X)]_\alpha, \quad (2.2.13)$$

где $[g_2(X)]_\alpha$ — квантиль уровня α распределения случайной величины $g_2(X)$. Производные функции квантили совпадают с производными функции $g_1(u)$:

$$\frac{\partial \varphi_\alpha(u)}{\partial u_i} = \frac{\partial g_1(u)}{\partial u_i}, \quad \frac{\partial^2 \varphi_\alpha(u)}{\partial u_i \partial u_j} = \frac{\partial^2 g_1(u)}{\partial u_i \partial u_j}. \quad (2.2.14)$$

Вспомогательные функции $V(u)$ и $W(u)$ определяются непосредственно:

$$V(u) = \mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))g'_{1u_i}(u)] = g'_{1u_i}(u)\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))], \quad (2.2.15)$$

$$W(u) = \mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))] \quad (2.2.16)$$

Частное функций $V(u)$ и $W(u)$ совпадает равно $g'_{1u_i}(u) = \frac{\partial}{\partial u_i} g_1(u)$. Соответственно, производная функции квантили и ее гладкая аппроксимация совпадают:

$$\frac{\partial \varphi_\alpha^\theta(u)}{\partial u_i} = \frac{V(u)}{W(u)} = \frac{\partial}{\partial u_i} g_1(u) = \frac{\partial \varphi_\alpha(u)}{\partial u_i}. \quad (2.2.17)$$

Аналогичным образом получаем, что гладкая аппроксимация второй производной функции квантили совпадает с исходной производной.

2.2.1.2 Произведение двух функций

Рассмотрим функцию потерь, представимую в виде произведения двух функций — функции, зависящей только от вектора стратегии, и функции случайного вектора:

$$g(u, X) = g_1(u)g_2(X). \quad (2.2.18)$$

Пусть $g_1(u) > 0$ для всех $u \in U$. Функция квантили определяется как

$$\varphi_\alpha(u) = g_1(u)[g_2(X)]_\alpha, \quad (2.2.19)$$

где $[g_2(X)]_\alpha$ — квантиль уровня α распределения случайной величины $g_2(X)$. В этом случае градиенты функции квантили и функции $g_1(u)$ сонаправлены:

$$\nabla \varphi_\alpha(u) = \nabla g_1(u)[g_2(X)]_\alpha. \quad (2.2.20)$$

Вспомогательные функции $V(u)$ и $W(u)$ определяются как:

$$\begin{aligned} V(u) &= \mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))g'_{1u_i}(u)g_2(X)] = g'_{1u_i}(u)\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))g_2(X)], \\ W(u) &= \mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))]. \end{aligned}$$

Аппроксимация производной функции квантили:

$$\frac{\partial \varphi_\alpha^\theta(u)}{\partial u_i} = \frac{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))g_2(X)]}{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))]} \frac{\partial}{\partial u_i} g_1(u). \quad (2.2.21)$$

Аппроксимация градиента функции квантили принимает вид:

$$\nabla \varphi_\alpha^\theta(u) = \nabla g_1(u) \cdot \frac{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))g_2(X)]}{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))]} \quad (2.2.22)$$

Аппроксимация градиента функции квантили и градиент исходной функции квантили сонаправлены. Следовательно, аппроксимация градиента функции квантили может использоваться при решении задачи квантильной минимизации с помощью численных методов первого порядка. Более того, согласно Следствию 2, при стремлении параметра θ к бесконечности будет выполняться предельный переход:

$$\lim_{\theta \rightarrow \infty} \frac{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))g_2(X)]}{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))]} = [g_2(X)]_\alpha, \quad (2.2.23)$$

так как аргумент производной сигмоиды в выражениях будет равен нулю только в точках, для которых $g(u, X) = \varphi_\alpha(u)$, т.е. в точках, в которых выполняется $g_2(X) = [g_2(X)]_\alpha$.

Выражение для аппроксимации второй смешанной частной производной функции квантили может быть получено непосредственно:

$$\begin{aligned} \frac{\partial^2 \varphi_\alpha^\theta(u)}{\partial u_i \partial u_j} &= \frac{\partial^2}{\partial u_i \partial u_j} g_1(u) \frac{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))g_2(X)]}{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))]} + \\ &+ \frac{\partial}{\partial u_i} g_1(u) \frac{\partial}{\partial u_j} g_1(u) \frac{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))g_2(X)] \mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))g_2(X)]}{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))]^2} - \\ &- \frac{\partial}{\partial u_i} g_1(u) \frac{\partial}{\partial u_j} g_1(u) \frac{\mathbf{M} [S''_\theta(\varphi_\alpha(u) - g(u, X))g_2^2(X)] \mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))]}{\mathbf{M} [S'_\theta(\varphi_\alpha(u) - g(u, X))]^2}. \end{aligned} \quad (2.2.24)$$

Рассмотрев поведение второй производной (2.2.1) отдельно для случаев $t = 0$ и $t \neq 0$, нетрудно убедиться, что для любого $t \in \mathbb{R}^1$ выполняется

$$\lim_{\theta \rightarrow \infty} S''_{\theta}(t) = 0.$$

Следовательно, второе и третье слагаемые в (2.2.24) будут стремиться к нулю при $\theta \rightarrow \infty$, а аппроксимация второй смешанной частной производной будет сходиться к точному значению:

$$\lim_{\theta \rightarrow \infty} \frac{\partial^2 \varphi_{\alpha}^{\theta}(u)}{\partial u_i \partial u_j} = \lim_{\theta \rightarrow \infty} \frac{\partial^2 g_1(u)}{\partial u_i \partial u_j} \frac{\mathbf{M} [S'_{\theta}(\varphi_{\alpha}(u) - g(u, X)) g_2(X)]}{\mathbf{M} [S'_{\theta}(\varphi_{\alpha}(u) - g(u, X))]} = \frac{\partial^2 g_1(u)}{\partial u_i \partial u_j} [g_2(X)]_{\alpha}.$$

2.2.1.3 Случай одномерного распределения и монотонной функции потерь

Пусть $X \in \mathbb{R}^1$, а функция $g(u, x)$ строго возрастающая и непрерывная слева по x для любого u . Этот случай описан в [8] в главе 4. Функция квантили определяется следующим образом:

$$\varphi_{\alpha}(u) = g(u, [X]_{\alpha}).$$

Используя сходимость сигмоиды к дельта-функции Дирака, получаем следующие выражения для вспомогательных функций $V(u)$ и $W(u)$:

$$\begin{aligned} V(u) &= \mathbf{M} [S'_{\theta}(\varphi_{\alpha}(u) - g(u, X)) g'_{u_i}(u, X)] \rightarrow \frac{g'_{u_i}(u, x)}{|g'_x(u, x)|} \Big|_{x=[X]_{\alpha}} \text{ при } \theta \rightarrow \infty, \\ W(u) &= \mathbf{M} [S'_{\theta}(\varphi_{\alpha}(u) - g(u, X))] \rightarrow \frac{1}{|g'_x(u, x)|} \Big|_{x=[X]_{\alpha}} \text{ при } \theta \rightarrow \infty, \\ V'_{u_j}(u) &\rightarrow \frac{g''_{u_i u_j}(u, x)}{|g'_x(u, x)|} \Big|_{x=[X]_{\alpha}} \text{ при } \theta \rightarrow \infty, \\ W'_{u_j}(u) &\rightarrow 0 \text{ при } \theta \rightarrow \infty. \end{aligned}$$

Таким образом, аппроксимация производной функции квантили сходится к производной функции потерь при $x = [X]_{\alpha}$.

$$\begin{aligned} \frac{\partial \varphi_{\alpha}^{\theta}(u)}{\partial u_i} &= \frac{V(u)}{W(u)} \rightarrow g'_{u_i}(u, [X]_{\alpha}) \text{ при } \theta \rightarrow \infty, \\ \frac{\partial^2 \varphi_{\alpha}^{\theta}(u)}{\partial u_i \partial u_j} &= \frac{V'_{u_j}(u) W(u) - V(u) W'_{u_j}(u)}{W^2(u)} \rightarrow g''_{u_i u_j}(u, [X]_{\alpha}) \text{ при } \theta \rightarrow \infty. \end{aligned}$$

2.2.2 Численные примеры

В данном разделе для функции вероятности приведены результаты расчетов ее гладкой аппроксимации, а также аппроксимации первых и вторых производных. Производные точной функции вероятности оценивались с помощью конечных разностей, а ее исходные значения – с помощью численного интегрирования средствами библиотеки SciPy языка Python. Значения аппрокси-

маций функции вероятности и ее производных, представляющие собой интегралы в соответствии с формулами (1.7.19), (2.2.2) и (2.2.3), вычислялись с помощью метода Монте-Карло на выборке из 10000 реализаций.

Для функции квантили и ее производных приведены два примера, в которых производится сравнение точной функции квантили и ее производных с оценками, полученными с помощью конечных разностей на основе выборочной квантили, а также с оценками, полученными с помощью гладкой аппроксимации. Оценки значений производных, получаемые на основе конечных разностей, вычислялись с шагом 0,01. Для оценки математических ожиданий в формулах (1.9.3) и (2.2.9) используется единая выборка из 10000 реализаций.

2.2.2.1 Пример 1

Рассмотрим пример с одной случайной величиной, одномерным управлением и билинейной функцией потерь:

$$\Phi(u, X) = 1 + u + X + uX. \quad (2.2.25)$$

Пусть значение целевого уровня потерь φ равно 2, а случайная величина X имеет нормальное распределение: $X \sim \mathcal{N}(1, 1)$. В примере проводится сравнение точных и аппроксимированных значений функции вероятности, её первой и второй производных, вычисленных соответствующими способами, при разных значениях параметра сигмоиды θ . Результаты сравнения представлены на рисунках 2.5 и 2.6.

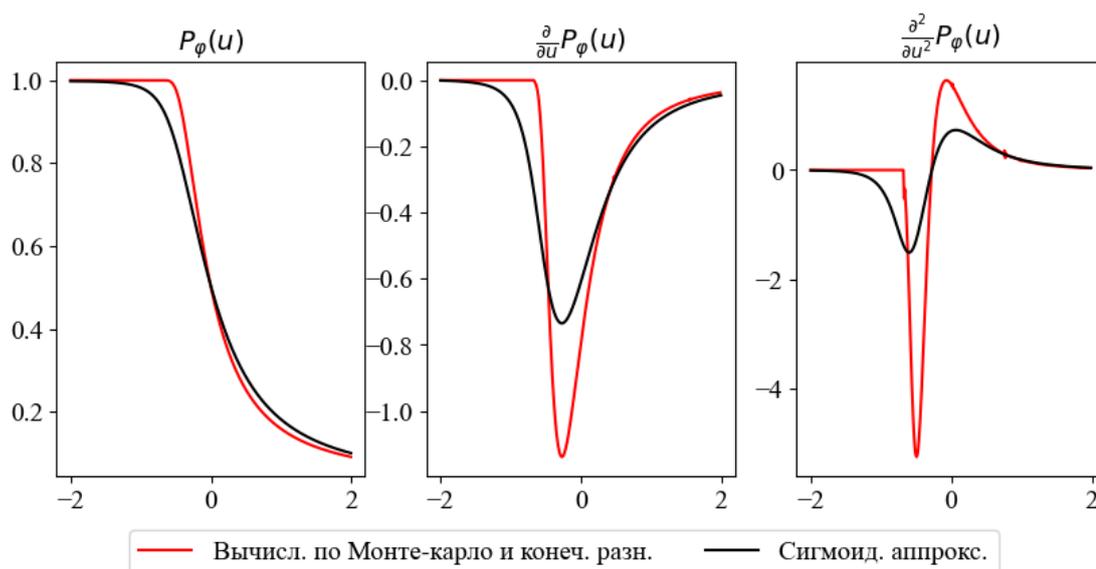


Рисунок 2.5. Сравнение точных (конечно-разностных) и аппроксимированных значений функций вероятности и ее производных при $\theta = 2$ и билинейной функции потерь

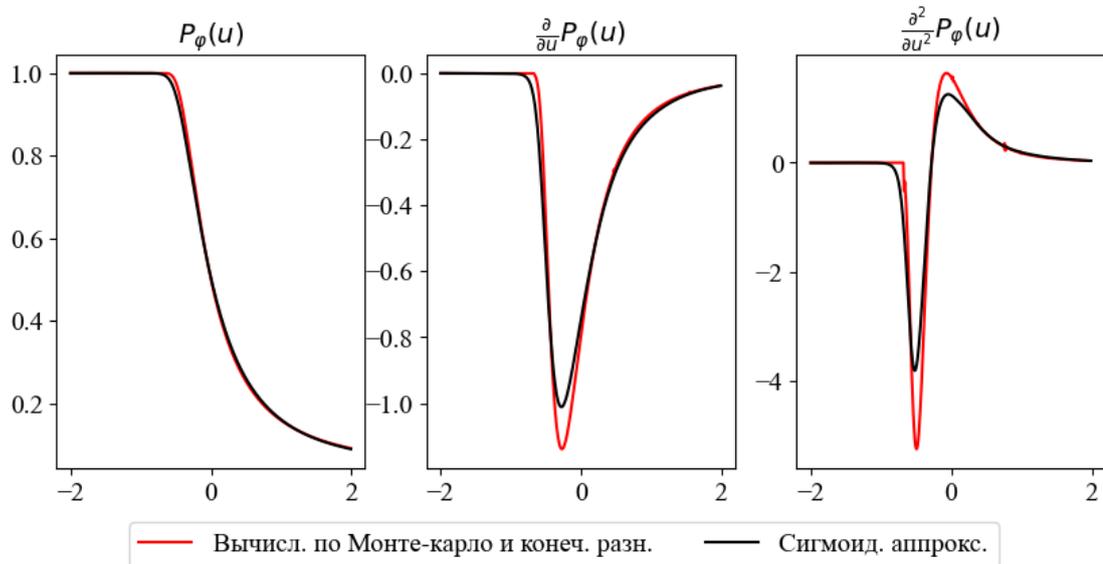


Рисунок 2.6. Сравнение точных (конечно-разностных) и аппроксимированных значений функций вероятности и ее производных при $\theta = 5$ и билинейной функции потерь

2.2.2.2 Пример 2

Рассмотрим пример с квадратичной функцией потерь. Пусть

$$\Phi(u, X) = 1 + u + X + (X - u)^2. \quad (2.2.26)$$

случайная величина имеет нормальное распределение $X \sim \mathcal{N}(1, 1)$, а допустимый уровень потерь φ равен 2. Как и в первом случае, сравним точные и аппроксимированные значения функции вероятности, её первой и второй производных при значении параметра сигмоиды $\theta = 10$. Результаты сравнения представлены на рисунке 2.7.

2.2.2.3 Пример 3

Далее рассмотрим примеры аппроксимации производных функции квантили. Пусть задана квадратичная функция потерь и гауссовская случайная величина:

$$g(u, X) = u^2 + 2uX + X^2 + 1, \quad X \sim \mathcal{N}(0, 1). \quad (2.2.27)$$

Параметр α равен 0,8, θ равна 30. Значение функции квантили для каждого u определяется как решение уравнения:

$$\Phi\left(-u + \sqrt{\varphi - 1}\right) - \Phi\left(-u - \sqrt{\varphi - 1}\right) = \alpha, \quad (2.2.28)$$

относительно φ . $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-s^2/2} ds$ – функция распределения стандартной гауссовской случайной величины.

На рисунке 2.8 представлены гладкая аппроксимация производных функций вероятности

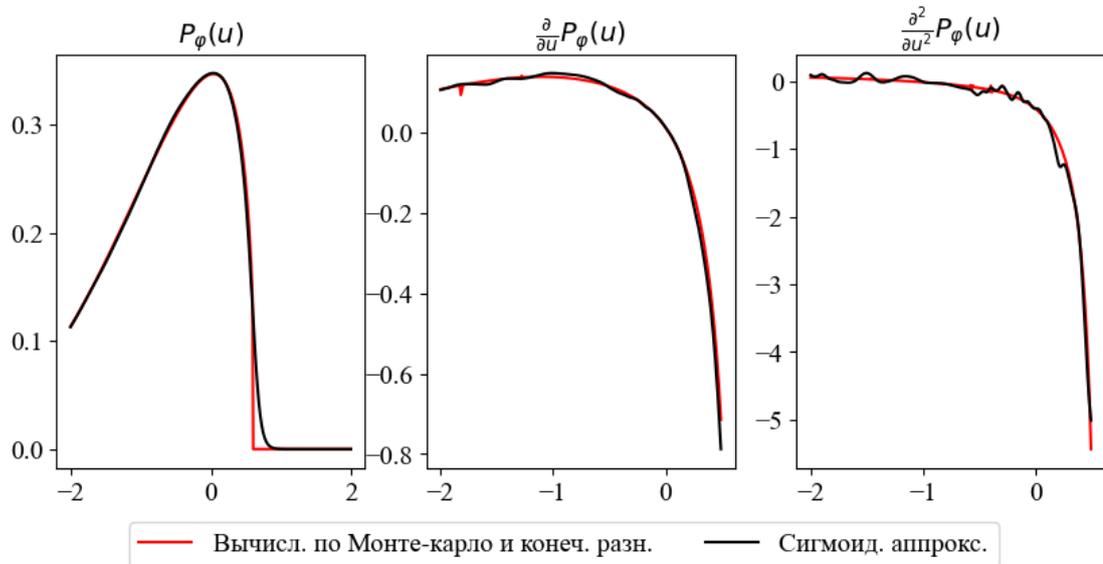


Рисунок 2.7. Сравнение точных (конечно-разностных) и аппроксимированных значений функций при $\theta = 10$ и квадратичной функции потерь

и конечно-разностная аппроксимация производных. Сплошная линия соответствует гладким ап-

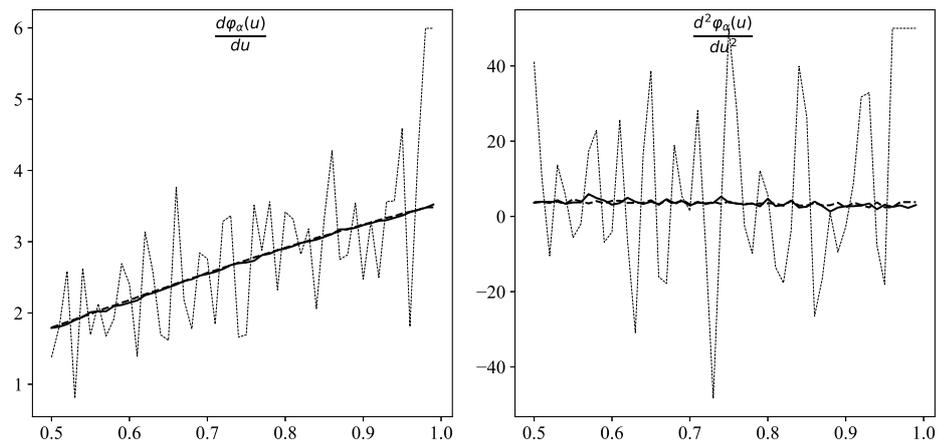


Рисунок 2.8. Оценки первой и второй производных функции квантили в примере 3.

проксимациям производных, пунктирная линия соответствует конечно-разностным оценкам на основе значений выборочной функции квантили, а прерывистая линия соответствует конечно-разностным оценкам производных, основанных на точных значениях функции квантили, определяемых для любого u как решение уравнения (2.2.28).

Как следует из рисунка 2.8, гладкая аппроксимация производных функции квантили близка к производным точной функции квантили, в то же время оценки производных с помощью выборочной квантили достаточно сильно флуктуируют.

2.2.2.4 Пример 4

Для функции квантили также рассмотрим пример с билинейной функцией потерь и двумерным нормальным распределением:

$$g(u, X) = u_1 X_1 + u_2 X_2, \quad X \sim \mathbb{N}(m, K), \quad (2.2.29)$$

где

$$m = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad K = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 2 \end{pmatrix}. \quad (2.2.30)$$

Функция квантили может быть найдена непосредственно:

$$\mathbf{M}(g(u, X)) = 0, \quad \mathbf{D}(g(u, X)) = u_1^2 + u_1 u_2 + 4u_2^2, \quad (2.2.31)$$

$$\varphi_\alpha(u) = x_\alpha \sqrt{u_1^2 + u_1 u_2 + 4u_2^2}, \quad (2.2.32)$$

где x_α — квантиль уровня α стандартного нормального распределения.

Сравнение точных производных функции квантили и их аппроксимаций представлено на рисунке 2.9. Гладкая аппроксимация первой производной функции квантили близка к точным значениям и обладает меньшей погрешностью в сравнении с оценками, полученными на основе выборочной квантили. Гладкая аппроксимация второй смешанной частной производной совпадает по знаку и наклону поверхности с точной производной, обладает большей погрешностью в сравнении с гладкой аппроксимацией первой производной.

2.3 Исследование задачи стохастического программирования с вероятностным критерием и полиэдральной функцией потерь

Рассмотрим полное вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$ и случайный вектор X на этом пространстве со значениями $x \in \mathbb{R}^m$. Пусть выпуклая полиэдральная функция потерь задана следующим образом:

$$\Phi(u, x) = \max_{i=\overline{1, k_1}} \{A_{1i}^T u + B_{1i}^T x + b_{1i}\}, \quad (2.3.1)$$

и пусть функции ограничений заданы следующим образом:

$$Q(u, x) = \max_{i=\overline{1, k_2}} \{A_{2i}^T u + B_{2i}^T x + b_{2i}\}, \quad (2.3.2)$$

где $u \in U$ - вектор стратегии; $U \subset \mathbb{R}^m$ - множество допустимых стратегий; $A_{1i}^T, A_{2i}^T, B_{1i}^T, B_{2i}^T$ - строки детерминированных матриц $A_1 \in \mathbb{R}^{k_1 \times n}, A_2 \in \mathbb{R}^{k_2 \times n}, B_1 \in \mathbb{R}^{k_1 \times m}, B_2 \in \mathbb{R}^{k_2 \times m}$ соответственно; $b_{1i}, i = \overline{1, k_1}$ и $b_{2i}, i = \overline{1, k_2}$ - компоненты числовых векторов $b_1 \in \mathbb{R}^{k_1}$ и $b_2 \in \mathbb{R}^{k_2}$ соответственно. Рассмотрим вероятность того, что функция потерь не превзойдет фиксированный заданный

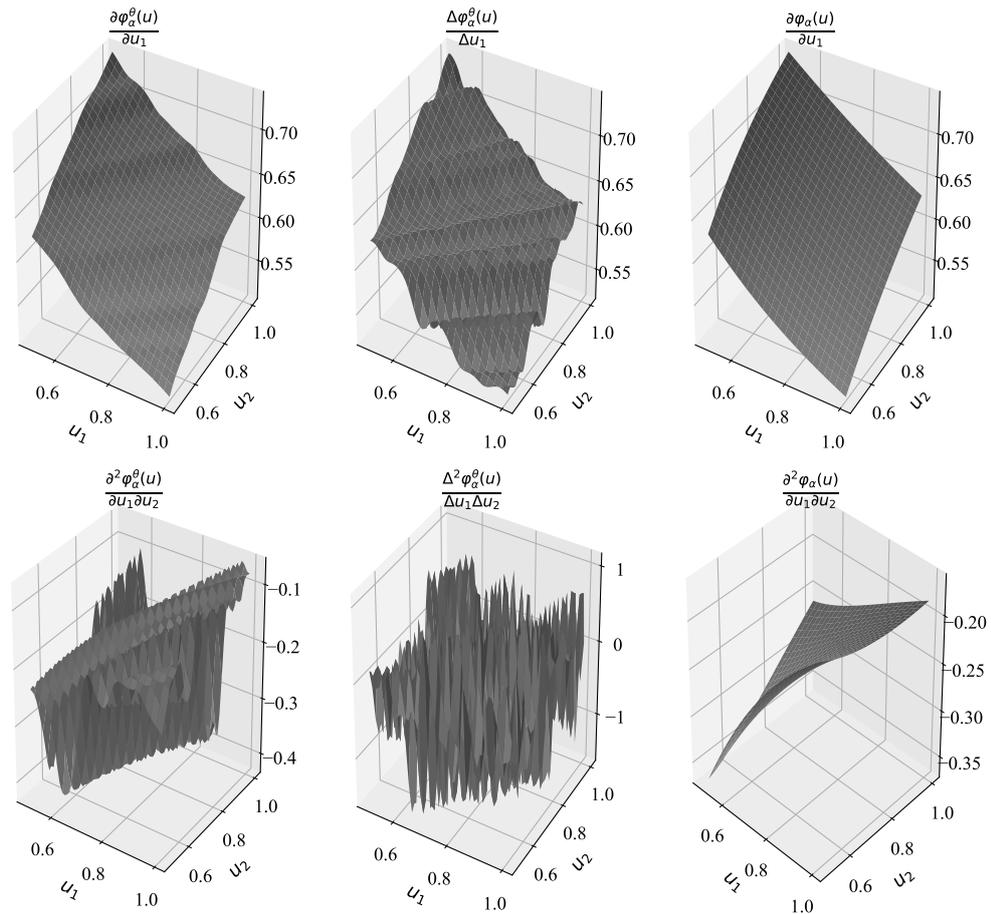


Рисунок 2.9. Оценки первой и второй производных функции квантили в примере 4

уровень потерь φ при выполняющихся ограничениях:

$$P_\varphi(u) = \mathbb{P} \{ \Phi(u, X) \leq \varphi, Q(u, X) \leq 0 \}. \quad (2.3.3)$$

Задача состоит в поиске оптимального вектора стратегии u_φ , который максимизирует рассматриваемую вероятность:

$$u_\varphi = \arg \max_{u \in U} P_\varphi(u). \quad (2.3.4)$$

Постановка задачи (2.3.4) ранее исследовалась в работах [53, 54]. Метод ее решения основывался на переходе к связанной задаче квантильной оптимизации и применении доверительного метода. Описанный далее подход предполагает решение исходной задачи максимизации функции вероятности без перехода к эквивалентной задаче квантильной оптимизации.

Для решения задачи (2.3.4) будет использоваться метод проекции градиента с использованием аппроксимации градиента функции вероятности. При этом необходимо учесть некоторые особенности рассматриваемой задачи:

- Для аппроксимации функции вероятности и ее производных необходимо существование градиента функции потерь и функции ограничений. Полиэдральная функция имеет произ-

водные во всех точках, кроме точек пересечения функций, находящихся под знаком максимума.

- Аппроксимация функции вероятности и ее градиента должна учитывать как целевую функцию, так и функцию ограничений

Первая проблема решается с помощью преобразования гладкого максимума, которое описано в пункте 2.3.1. Вторая проблема рассматривается в пункте 2.3.2.

2.3.1 Аппроксимация полиэдральной функции потерь

Полиэдральная функция потерь, представляющая собой максимум некоторых функций, в общем случае не имеет производной по аргументам в точках пересечения функций, находящихся под знаком максимума. При этом для дальнейшего построения аппроксимации градиента функции вероятности необходимо наличие градиента функции потерь во всех точках. Для решения этой проблемы к функции потерь применяется преобразование гладкого максимума.

Функция гладкого максимума набора линейных функций $f_1(x), \dots, f_z(x)$ задается следующим образом:

$$SM_\gamma(f_1(x), \dots, f_z(x)) = \frac{\sum_{i=1}^z f_i(x)e^{\gamma f_i(x)}}{\sum_{i=1}^z e^{\gamma f_i(x)}}, \quad (2.3.5)$$

где $\gamma > 0$ - параметр гладкости, представленный большим положительным числом. Для дальнейших выкладок необходимо показать, что

$$\lim_{\gamma \rightarrow +\infty} SM_\gamma(f_1(x), \dots, f_z(x)) = \max(f_1(x), \dots, f_z(x)), \quad (2.3.6)$$

$$\lim_{\gamma \rightarrow +\infty} \frac{d}{dx} SM_\gamma(f_1(x), \dots, f_z(x)) = \frac{d}{dx} \max(f_1(x), \dots, f_z(x)), \quad (2.3.7)$$

для каждой точки x , в которой определена производная правой части выражения (2.3.7).

Доказательство выражения (2.3.6) представляется тривиальным. Для доказательства (2.3.7) рассмотрим гладкий максимум двух функций:

$$\max(f_1(x), f_2(x)) \approx SM_\gamma(f_1(x), f_2(x)) = \frac{f_1(x)e^{\gamma f_1(x)} + f_2(x)e^{\gamma f_2(x)}}{e^{\gamma f_1(x)} + e^{\gamma f_2(x)}}.$$

Производная гладкого максимума может быть найдена непосредственно:

$$\frac{d}{dx} SM_\gamma(f_1(x), f_2(x)) = \frac{f_1'(x)e^{\gamma f_1(x)} + f_2'(x)e^{\gamma f_2(x)}}{e^{\gamma f_1(x)} + e^{\gamma f_2(x)}} + G(\gamma, f_1(x), f_2(x)), \quad (2.3.8)$$

где

$$G(\gamma, f_1(x), f_2(x)) = \gamma e^{\gamma f_1(x)} e^{\gamma f_2(x)} \frac{(f_1(x) - f_2(x))(f_1'(x) - f_2'(x))}{(e^{\gamma f_1(x)} + e^{\gamma f_2(x)})^2}. \quad (2.3.9)$$

Значение функции $G(\cdot)$ в точках пересечения функций под знаком максимума, т.е. в точках

$x^* : f_1(x^*) = f_2(x^*)$, равно нулю. В точках, отличных от x^* , рассмотрим поведение функции $G(\cdot)$ при $\gamma \rightarrow +\infty$.

Множитель $(f_1(x) - f_2(x))(f_1'(x) - f_2'(x))$ функции $G(\cdot)$ не зависит от параметра γ и может быть опущен. Без потери общности, предположим, что $f_2(x) > f_1(x)$ и разделим числитель и знаменатель выражения на $e^{2\gamma \min(f_1(x), f_2(x))}$, т.е. на $e^{2\gamma f_1(x)}$:

$$\lim_{\gamma \rightarrow +\infty} G(\cdot) = \lim_{\gamma \rightarrow +\infty} \frac{\gamma e^{\gamma f_1(x)} e^{\gamma f_2(x)}}{(e^{\gamma f_1(x)} + e^{\gamma f_2(x)})^2} = \lim_{\gamma \rightarrow +\infty} \frac{\gamma e^{\gamma(f_2(x) - f_1(x))}}{(1 + e^{\gamma(f_2(x) - f_1(x))})^2}. \quad (2.3.10)$$

Для упрощения выкладок сделаем замену $t = f_2(x) - f_1(x)$, $t > 0$. Единица в знаменателе выражения несущественна по сравнению с экспонентой, поэтому она может быть опущена. Далее, по правилу Лопиталья:

$$\lim_{\gamma \rightarrow +\infty} \frac{\gamma e^{\gamma(f_2(x) - f_1(x))}}{(1 + e^{\gamma(f_2(x) - f_1(x))})^2} = \lim_{\gamma \rightarrow +\infty} \frac{\gamma e^{\gamma t}}{(e^{\gamma t})^2} = \lim_{\gamma \rightarrow +\infty} \frac{\gamma}{e^{\gamma t}} = \lim_{\gamma \rightarrow +\infty} \frac{1}{te^{\gamma t}} = 0. \quad (2.3.11)$$

Получаем, что аппроксимация производной функции гладкого максимума может быть представлена как:

$$\frac{d}{dx} SM_\gamma(f_1(x), f_2(x)) \approx \frac{f_1'(x)e^{\gamma f_1(x)} + f_2'(x)e^{\gamma f_2(x)}}{e^{\gamma f_1(x)} + e^{\gamma f_2(x)}}. \quad (2.3.12)$$

Применяя подход, аналогичный таковому при рассмотрении функции $G(\cdot)$, показывается, что

$$\lim_{\gamma \rightarrow +\infty} \frac{f_1'(x)e^{\gamma f_1(x)} + f_2'(x)e^{\gamma f_2(x)}}{e^{\gamma f_1(x)} + e^{\gamma f_2(x)}} \rightarrow f_1'(x)$$

при условии $f_1(x) > f_2(x)$. Это означает, что производная функции гладкого максимума может быть использована как аппроксимация производной функции максимума.

Данный подход может быть обобщен на произвольное число функций под знаком максимума. В случае набора функций, выражение (2.3.12) записывается как:

$$\frac{d}{dx} SM_\gamma(f_1(x), \dots, f_z(x)) \approx \frac{\sum_{i=1}^z f_i'(x)e^{\gamma f_i(x)}}{\sum_{i=1}^z e^{\gamma f_i(x)}}.$$

Возвращаясь к первоначальной постановке задачи (2.3.4), функция потерь $\Phi(u, x)$ и функция ограничений $Q(u, x)$ аппроксимируются следующим образом:

$$\Phi_\gamma^*(u, x) = \frac{\sum_{i=1}^k [A_i^T u + B_i^T x + b_i] \exp \{ \gamma (A_i^T u + B_i^T x + b_i) \}}{\sum_{i=1}^k \exp \{ \gamma (A_i^T u + B_i^T x + b_i) \}}. \quad (2.3.13)$$

$$Q_\gamma^*(u, x) = \frac{\sum_{i=1}^{k_2} [A_{2i}^T u + B_{2i}^T x + b_{2i}] \exp \{ \gamma (A_{2i}^T u + B_{2i}^T x + b_{2i}) \}}{\sum_{i=1}^{k_2} \exp \{ \gamma (A_{2i}^T u + B_{2i}^T x + b_{2i}) \}}. \quad (2.3.14)$$

Частные производные этих аппроксимированных функций по компонентам вектора стратегии

$u_j, j = \overline{1, m}$ задаются как:

$$\frac{\partial \Phi_\gamma^*(u, x)}{\partial u_j} = \frac{\sum_{i=1}^k A_{ij} \exp \{ \gamma (A_i^T u + B_i^T x + b_i) \}}{\sum_{i=1}^k \exp \{ \gamma (A_i^T u + B_i^T x + b_i) \}}. \quad (2.3.15)$$

$$\frac{\partial Q_\gamma^*(u, x)}{\partial u_j} = \frac{\sum_{i=1}^{k_2} A_{2ij} \exp \{ \gamma (A_{2i}^T u + B_{2i}^T x + b_{2i}) \}}{\sum_{i=1}^{k_2} \exp \{ \gamma (A_{2i}^T u + B_{2i}^T x + b_{2i}) \}}. \quad (2.3.16)$$

Учитывая соотношения (2.3.6) и (2.3.7), верно утверждение:

$$\lim_{\gamma \rightarrow \infty} \mathbb{P} \{ \Phi_\gamma^*(u, X) \leq \varphi, Q_\gamma^*(u, X) \leq 0 \} = \mathbb{P} \{ \Phi(u, X) \leq \varphi, Q(u, X) \leq 0 \} = P_\varphi(u). \quad (2.3.17)$$

2.3.2 Аппроксимация вероятностного критерия

Возвращаясь к исходной проблеме (2.3.4), необходимо отметить, что функция вероятности (2.3.3) включает в себя два неравенства. Для применения описанной методики аппроксимации необходимо рассмотреть представление функции вероятности с индикаторной функцией двух неравенств:

$$\begin{aligned} P_\varphi(u) &= \mathbb{P} \{ \Phi(u, X) \leq \varphi, Q(u, X) \leq 0 \} = \\ &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} I \{ \Phi(u, x) \leq \varphi, Q(u, x) \leq 0 \} f(x) dx_1 \dots dx_m = \\ &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \Theta(\varphi - \Phi(u, x)) \Theta(-Q(u, x)) f(x) dx_1 \dots dx_m, \end{aligned} \quad (2.3.18)$$

где $f(x) = f(x_1, \dots, x_m)$ - плотность случайной величины X . Для использования аналогичного подхода к аппроксимации, функция Хевисайда заменяется сигмоидальной функцией. Дополнительно производится замена полиэдральных функции потерь и функции ограничений с помощью преобразования гладкого максимума. В результате этого получается дифференцируемая аппроксимация $P_\varphi^{\theta, \gamma}(u)$ функции вероятности $P_\varphi(u)$:

$$P_\varphi^{\theta, \gamma}(u) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} S_\theta(\varphi - \Phi_\gamma^*(u, x)) S_\theta(-Q_\gamma^*(u, x)) f(x) dx_1 \dots dx_m. \quad (2.3.19)$$

Частные производные функции $P_\varphi^{\theta,\gamma}(u)$ по компонентам стратегии u_j , $j = \overline{1, m}$, вычисляются аналогично (1.7.19):

$$\begin{aligned} \frac{\partial}{\partial u_j} P_\varphi^{\theta,\gamma}(u) &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \theta S_\theta(\varphi - \Phi_\gamma^*(u, x)) S_\theta(-Q_\gamma^*(u, x)) \cdot \\ &\quad \left([S_\theta(\varphi - \Phi_\gamma^*(u, x)) - 1] \Phi_{\gamma u_j}^{*'}(u, x) + [S_\theta(-Q_\gamma^*(u, x)) - 1] Q_{\gamma u_j}^{*'}(u, x) \right) \cdot \\ &\quad f(x) dx_1 \dots dx_m. \end{aligned} \quad (2.3.20)$$

Учитывая предыдущие соотношения, верны следующие утверждения:

$$\lim_{\theta \rightarrow +\infty, \gamma \rightarrow +\infty} P_\varphi^{\theta,\gamma}(u) = P_\varphi(u), \quad (2.3.21)$$

$$\lim_{\theta \rightarrow +\infty, \gamma \rightarrow +\infty} \frac{\partial}{\partial u_j} P_\varphi^{\theta,\gamma}(u) = \frac{\partial}{\partial u_j} P_\varphi(u). \quad (2.3.22)$$

В итоге, получается новая проблема оптимизации, аналогичная (2.3.4):

$$u_\varphi^{\theta,\gamma} = \arg \max_{u \in U} P_\varphi^{\theta,\gamma}(u). \quad (2.3.23)$$

Ввиду полученных соотношений о сходимости, решение данной приближенной задачи (2.3.23) может рассматриваться как приближенное решение исходной задачи (2.3.4).

2.3.3 Алгоритм оптимизации

В этом подпункте приведен используемый в задаче с полиэдральной функцией потерь алгоритм оптимизации. Данный алгоритм основан на методе градиентного спуска. Перед началом работы алгоритма задаются начальная точка u^0 , требуемый уровень потерь φ , базовое значение шага алгоритма h , максимальное количество шагов N , и границы u_i^{min} , u_i^{max} для каждой из компонент u_i вектора стратегии. Также перед началом работы алгоритма вычисляется значение функции вероятности p^0 в начальной точке u^0 по формуле (2.3.19).

Алгоритм 3. Модифицированный алгоритм градиентного спуска

- 1) Вычисляется аппроксимированное значение градиента $\Delta p = [\Delta p_1, \Delta p_2, \dots, \Delta p_n]^T$ функции вероятности при текущей стратегии u^k по формуле (2.3.20);
- 2) Значение градиента корректируется на основе значений u_i^k текущего вектора стратегии u^k .

Для каждого $i = \overline{1, n}$ проверяются условия:

- если $u_i^k < u_i^{min}$ и $\Delta p_i < 0$, то положим $\Delta p_i = 0$;
- если $u_i^k > u_i^{max}$ и $\Delta p_i > 0$, то положим $\Delta p_i = 0$;

3) Вычисляется текущее значение шага

$$h^k = h \cdot \left(\max_{i=\overline{1,n}} |\Delta p_i| \right)^{-1}; \quad (2.3.24)$$

4) Значение текущего шага корректируется на основе значений следующего предполагаемого вектора стратегии $u^{k+1} = u^k + h^k \cdot \Delta p$. Для каждого $i = \overline{1,n}$ проверяются условия:

- если $u_i^{k+1} < u_i^{min}$, то положим $h^k = \min\{h^k, (u_i^{min} - u_i^k)/\Delta p_i\}$;
- если $u_i^{k+1} > u_i^{max}$, то положим $h^k = \min\{h^k, (u_i^{max} - u_i^k)/\Delta p_i\}$;

5) Вычисляется следующий вектор стратегии $u^{k+1} = u^k + h^k \cdot \Delta p$;

6) Вычисляется значение функции вероятности p^{k+1} при стратегии u^{k+1} по формуле (2.3.19);

7) Проверяются условия остановки/итерации алгоритма:

- если $p^{k+1} \leq p^k$, то значение шага модифицируется $h = h \cdot L$, $L \in [0, 1]$, а алгоритм повторяется, положив $k = k + 1$;
- если $p^{k+1} > p^k$, алгоритм повторяется, положив $u^k = u^{k+1}$, $k = k + 1$;
- если $k \geq N$, алгоритм останавливается. Вектор оптимальной стратегии приравнивается к текущему вектору: $u^{opt} = u^k$.

Данный алгоритм был реализован на языке программирования Python.

2.3.4 Пример двумерной задачи с полиэдральной функцией потерь и полиэдральной функцией ограничений

Рассмотрим задачу (2.3.4) с функцией вероятности, заданной согласно (2.3.3), и функциями потерь, заданными согласно (2.3.1) и (2.3.2) соответственно со следующими матрицами:

$$A_1 = \begin{pmatrix} 6 & -6 \\ 1 & 1 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 & 2 \\ 10 & -10 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 1 \\ -5 \end{pmatrix}, \quad (2.3.25)$$

$$A_2 = \begin{pmatrix} 10 & -10 \end{pmatrix}, \quad B_2 = \begin{pmatrix} -3 & -1 \end{pmatrix}, \quad b_2 = \begin{pmatrix} -2 \end{pmatrix}. \quad (2.3.26)$$

Пусть значения компонент вектора стратегии ограничены следующим образом:

$$u_1^{min} = u_2^{min} = 0, \quad u_1^{max} = u_2^{max} = 10. \quad (2.3.27)$$

Пусть случайный двумерный вектор X со значениями $x \in \mathbb{R}^2$ имеет Гауссово распределение:

$$X \sim \mathbf{N}(0, I), \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (2.3.28)$$

Данная постановка задачи эквивалентна таковой в работе [53]. В работе [53] эта задача

приводится к эквивалентной задаче квантильной оптимизации с фиксированным уровнем вероятности, называемым доверительным уровнем. Полученная квантильная задача решается с использованием различных методов. При этом решением задачи в [53] является стратегия, доставляющее минимум функции квантили при заданных ограничениях на уровне доверительной вероятности. В качестве иллюстрации работы предложенного алгоритма, для каждого решения квантильной задачи из работы [53] полученное значение квантили при оптимальной стратегии интерпретируется как параметр φ приближенной задачи (2.3.23), которая затем решается с помощью представленного ранее алгоритма.

Зададим параметры алгоритма, сигмоидальной функции и функции гладкого максимума следующим образом:

$$u^0 = [5, 5], \quad h = 0,1, \quad N = 100, \quad \theta = \gamma = 12. \quad (2.3.29)$$

После нахождения оптимальной стратегии u^{opt} с использованием представленного алгоритма, дополнительно вычисляется оптимальный уровень вероятности $P_\varphi(u^{opt})$ без учета аппроксимации для оценки близости реальных и аппроксимированных значений. Решения, полученные в работе [53], представлены в таблицах 2.1 и 2.3, а результаты работы алгоритма при соответственно заданных уровнях потерь φ представлены в таблицах 2.2 и 2.4. Оптимальная стратегия, полученная при помощи других методов, обозначена в таблицах как u^* .

Таблица 2.1. Решение задачи (2.3.4) с $\alpha = 0.9$ [53]

Алгоритм	Опт. φ	u^*
Алгоритм 1 [53]	18,656	$\begin{pmatrix} 0 \\ 0,3218 \end{pmatrix}$
Алгоритм 2 [53]	14,674	$\begin{pmatrix} 0 \\ 1,44 \end{pmatrix}$
Квазиград.	13,906	$\begin{pmatrix} -0,0025 \\ 1,337 \end{pmatrix}$

Таблица 2.2. Решение приближенной задачи (2.3.23) с заданными уровнями потерь φ (при $\alpha = 0.9$)

φ	u^{opt}	$P_\varphi^{\theta,\gamma}(u^{opt})$	$P_\varphi(u^{opt})$
18,656	$\begin{pmatrix} 0 \\ 0,841 \end{pmatrix}$	0,9504	0,951
14,674	$\begin{pmatrix} 0 \\ 0,808 \end{pmatrix}$	0,9106	0,9105
13,906	$\begin{pmatrix} 0 \\ 0,751 \end{pmatrix}$	0,9017	0,9022

Таблица 2.3. Решение задачи (2.3.4) с $\alpha = 0.8$ [53]

Алгоритм	Опт. φ	u^*
Алгоритм 1 [53]	13,234	$\begin{pmatrix} 0 \\ 0,2032 \end{pmatrix}$
Алгоритм 2 [53]	8,511	$\begin{pmatrix} 0 \\ 1,48 \end{pmatrix}$
Квазиград.	7,943	$\begin{pmatrix} -0,005 \\ 1,084 \end{pmatrix}$

Таблица 2.4. Решение приближенной задачи (2.3.23) с заданными уровнями потерь φ (при $\alpha = 0.8$)

φ	u^{opt}	$P_\varphi^{\theta,\gamma}(u^{opt})$	$P_\varphi(u^{opt})$
13,234	$\begin{pmatrix} 0 \\ 0,7695 \end{pmatrix}$	0,892	0,8939
8,511	$\begin{pmatrix} 0 \\ 0,737 \end{pmatrix}$	0,8163	0,8165
7,943	$\begin{pmatrix} 0 \\ 0,6985 \end{pmatrix}$	0,8072	0,8076

Как видно из таблиц 2.1- 2.4, в каждом случае аппроксимированное значение функции ве-

роятности в точке оптимума выше доверительного уровня, использованного при решении задачи квантильной оптимизации, что показывает корректность полученного решения, так и алгоритма в целом для рассматриваемой постановки задачи. Уровень вероятности при оптимальной стратегии значительно выше, чем таковой при решении задачи алгоритмами 1 [53] или 2 [53], и это значит, что полученное представленным алгоритмом решение лучше решений, полученных этими методами. Значение функции вероятности в точке оптимума также очень близко к таковому значению, полученному при решении задачи квазиградиентным способом, который, в свою очередь, дает наиболее близкое к истинному решение [53]. Такой подход с использованием аппроксимации градиента функции вероятности может быть использован с другими методами решения, использующими градиент оптимизируемой функции. При решении этой задачи метод градиентного спуска выбран в основном из-за его простоты и понятности.

Также на основании данных таблиц 2.2 и 2.4 можно сделать выводы о точности аппроксимации функции вероятности. В случае решения задачи для доверительного уровня 0,9, относительная разница между аппроксимированным и истинным значением функции вероятности не превышает 0,07%. В случае решения задачи для доверительного уровня 0,8, относительная разница между аппроксимированным и истинным значением функции вероятности не превышает 0,22%. Ранее было показано, что выражения для абсолютного отклонения аппроксимированного значения функции вероятности от истинного значения представляются достаточно громоздкими даже в одномерном случае. Тем не менее, также было показано, что величина этого отклонения в одномерном случае стремится к нулю при стремлении параметра сигмоиды к бесконечности. Исходя из аналогичности доказательств сходимости в одномерном и двумерном случаях, а также отсутствии значимых изменений при построении аппроксимации для более высших размерностей, ожидается, что характер изменения рассматриваемого отклонения сохранится при переходе к более высоким размерностям.

2.4 Выводы по главе 2

С использованием гладкой аппроксимации функции вероятности и ее производных получены следующие результаты:

- Разработан алгоритм построения внешней аппроксимации границы ядра вероятностной меры для двумерного случайного вектора с абсолютно непрерывными распределениями компонент, основанный на рассматриваемой аппроксимации. Граница определяется через решение дифференциального уравнения первого порядка в полярных координатах. Аналогичный подход применен для построения внешней аппроксимации ядра для трехмерного случайного вектора, где граница множества векторов нормали, определяющих ядро, опре-

деляется как решение дифференциального уравнения первого порядка в сферических координатах.

- Разработан алгоритм решения задачи стохастического программирования с вероятностным критерием, полиэдральной функцией потерь и полиэдральной функцией ограничений. Численные примеры показывают, что полученное решение превосходит по значению критерия решение, полученное иными алгоритмами при переходе к эквивалентной задаче квантильной оптимизации.
- Получены выражения для вторых производных гладкой аппроксимации функции вероятности и функции квантили по компонентам вектора стратегии. Приведенные примеры демонстрируют, что эти производные хорошо аппроксимируют соответствующие производные точных функций, аналитическое вычисление которых сильно затруднено. Знание вторых производных позволяет применять численные методы второго порядка к решению задач стохастического программирования с вероятностными критериями или ограничением в форме функции вероятности.

Основные результаты главы 2 опубликованы в работах [27, 31, 76, 78].

3. Применение гладкой аппроксимации функции вероятности в прикладных задачах стохастического программирования

В третьей главе рассматривается использование гладкой аппроксимации функции вероятности и функции квантили для решения некоторых известных задач стохастического программирования. В разделе 3.1 рассматривается задача оптимизации площади взлетно-посадочной полосы при вероятностном ограничении на успешность посадки. В разделе 3.2 рассматривается задача определения допустимого множества векторов скоростей вектора, определенного априори и гарантирующего безопасную посадку с некоторой гарантирующей вероятностью. Далее в разделе 3.3 исследуется задача проектирования системы водоснабжения в засушливом регионе с минимальной стоимостью и гарантированным с некоторой вероятностью удовлетворением спроса на пресную воду. В разделе 3.4 рассматривается задача формирования портфеля ценных бумаг по вероятностному критерию с логарифмической функцией потерь.

3.1 Задача оптимизации площади взлетно-посадочной полосы.

В данном разделе рассматривается проблема оптимального определения длины и ширины взлетно-посадочной полосы (ВПП), предназначенной для посадки тяжелого летательного аппарата. Проблема оптимизации площади взлетно-посадочной полосы связана с двумя факторами:

- действием порывов ветра при посадке, что влияет на точку касания летательным аппаратом (ЛА) полотна, поэтому с точки зрения безопасности посадки полотно ВПП должно быть максимально обширным;
- дороговизной покрытия квадратного метра полотна ВПП, поэтому с этой точки зрения площадь ВПП должна быть минимально возможной.

В результате учета этих двух факторов и возникает задача оптимизации: найти минимально возможную площадь ВПП, при которой гарантируется некий установленный уровень безопасности при посадке ЛА. Суть задачи сводится в поиске оптимальных длины и ширины полотна взлетно-посадочной полосы, при которых выполняются вероятностные ограничения по безопасности, то есть превышает заданная вероятность успешной посадки.

Формализуем постановку задачи. Пусть O — расчетная точка касания самолетом ВПП; t_0 — номинальная длина пробега ЛА до полной остановки без учета внешнего воздействия; r_0 — половина расстояния между посадочными шасси. На случай недолета или перелета относительно расчетной точки касания вводятся величины запасов t_s на случай недолета и t_l на случай перелета. Ширина самой ВПП обозначается как $2r_1$. Все введенные величины представлены на рисунке 3.1.

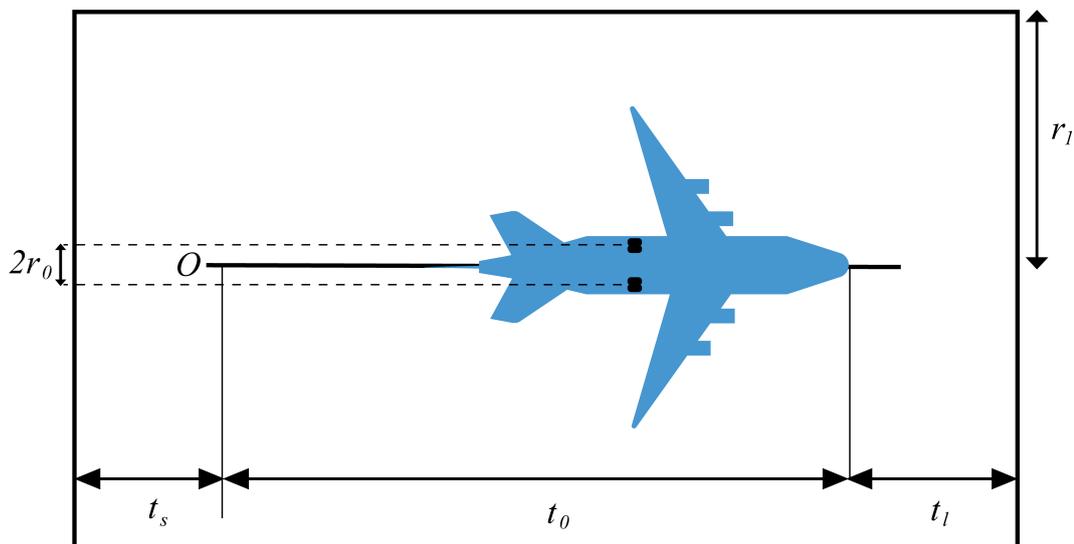


Рисунок 3.1. Параметры ВПП и ЛА при посадке

Площадь ВПП является целевой функцией, подлежащей оптимизации, и определяется выражением:

$$A(t_s, t_l, r_1) = (t_0 + t_s + t_l) \cdot 2r_1. \quad (3.1.1)$$

В целях упрощения модели предполагается, что $r_0 = 0$. Исходя из физического смысла введенных величин, накладываются следующие ограничения:

$$t_s \geq 0, \quad t_l \geq 0, \quad r_1 \geq 0. \quad (3.1.2)$$

Предположим, что известны распределения компонент скорости ветра на ВПП - продольной компоненты V_t и поперечной компоненты V_r , а также пусть эти компоненты независимы:

$$V_t \sim \mathcal{N}(m_t, \sigma_t^2), \quad V_r \sim \mathcal{N}(m_r, \sigma_r^2). \quad (3.1.3)$$

Если посадка управляется только путем изменения угла крена [12], то данные компоненты связаны со случайными продольным отклонением E_t и поперечным отклонением E_r от теоретической точки касания следующим образом:

$$E_t = c_{11}V_t + c_{12}|V_r|, \quad E_r = c_{22}V_r, \quad (3.1.4)$$

где c_{11}, c_{12}, c_{22} — эмпирически определяемые коэффициенты. Посадка самолета успешна, если фактическая точка касания самолетом полотно не выходит за рамки ВПП, т.е.:

$$-t_s \leq E_t \leq t_l, \quad |E_r| \leq r_1. \quad (3.1.5)$$

Таким образом требование по безопасности посадки с вероятностью не менее α задается как:

$$\mathbb{P} \{-t_s \leq E_t \leq t_l, |E_r| \leq r_1\} \geq \alpha. \quad (3.1.6)$$

В итоге задача оптимизации описывается как:

$$A(t_s, t_l, r_1) \rightarrow \min_{t_s, t_l, r_1} \quad (3.1.7)$$

при ограничениях (3.1.2), (3.1.6).

В книге [8] авторами приведена задача, эквивалентная задаче (3.1.7), но в постановке квантильной оптимизации, и при помощи аналитического алгоритма на основе доверительного метода определено оптимальное решение для различных значений параметра c_{12} . В частности, при $c_{12} = 0$ приближенное оптимальное решение задачи выглядит как:

$$t_s = t_l = k_2, \quad r_1 = \frac{\gamma}{2}, \quad (3.1.8)$$

где

$$k_2 = |c_{11}| \sigma_t g_\alpha, \quad \gamma = 2|c_{22}| \sigma_r g_\alpha, \quad g_\alpha = \sqrt{-2 \ln(1 - \alpha)}. \quad (3.1.9)$$

Авторами [8] также отмечено, что использование метода Монте-Карло или прямого метода оценки уровня квантили связано с огромными вычислительными затратами, а получение точного решения невозможно. Однако, в настоящий момент вычислительная сложность метода Монте-Карло даже для больших объемов выборок представляется приемлемой, поэтому он может применяться как для проверки выполнения ограничения на вероятность успешной посадки (3.1.6), так и для вычисления математических ожиданий в аппроксимации производных вероятностного ограничения.

Далее будем оптимизировать логарифм площади взлетно-посадочной полосы в целях уменьшения масштаба параметров. Таким образом, задача определяется как:

$$\ln r_1 + \ln(t_0 + t_s + t_l) \rightarrow \min \quad (3.1.10)$$

с вероятностным ограничением:

$$\mathbb{P} \{-t_s \leq E_t \leq t_l, |E_r| \leq r_1\} \geq \alpha. \quad (3.1.11)$$

Для решения задачи оптимизации с вероятностным ограничением будет использоваться метод проекции градиента с уменьшающимся шагом. Для этого перейдем от исходного вероят-

ностного ограничения (3.1.11) к его аппроксимации:

$$\mathbf{M} [S_{\theta}((t_s + E_t)(t_l - E_t)) S_{\theta}(r_1^2 - E_r^2)] \geq \alpha. \quad (3.1.12)$$

На каждом шаге метода оптимизации для вычисления вероятности в ограничении (3.1.12), а также ее производных, будем использовать метод Монте-Карло на выборке из 10000 реализаций.

Численные параметры для задачи заданы в соответствии с [8]:

$$c_{11} = c_{12} = -20 \text{ сек}, \quad c_{22} = 3 \text{ сек}, \quad t_0 = 1500 \text{ м}, \quad \sigma_t = \sigma_r = 5 \frac{\text{м}}{\text{сек}},$$

$$m_t = m_r = 0, \quad \alpha = 0,99.$$

Вычисления предлагаемым алгоритмом были выполнены за 7,59 секунд. В результате работы алгоритма получаем, что минимальная площадь ВПП, удовлетворяющая ограничению (3.1.12) составляет 0,185 км². Оптимальное значение параметра t_s (запас длины ВПП на случай недолета) равно 583,1м; оптимальное значение параметра t_l (запас длины ВПП на случай перелета) равно 283,1м; оптимальное значение параметра r_1 равно 39,2м.

В [8] представлены решения исследуемой задачи, полученные с помощью доверительного метода и метода стохастического квазиградиента. В таблице 3.1 представлено сравнение решений, полученных тремя методами, а также проведена проверка выполнения вероятностного ограничения (3.1.11) методом Монте-Карло на выборке из 1000000 реализаций без использования аппроксимации индикаторной функции сигмоидой.

Таблица 3.1. Результаты вычислений и другие решения

Метод решения	площадь ВПП A , км ²	t_s , м	t_l , м	r_1 , м	Вероятность успешной посадки
Аппроксимация функции вероятности	0,185	583,1	283,1	39,2	0,9903
Доверительный метод	0,202	430,2	290,1	45,54	0,9952
Метод квазиградиента	0,170	400,3	213,2	40,23	0,9849

Как видно из таблицы 3.1, предлагаемый метод решения дает меньшую оптимальную площадь в сравнении с результатами, полученными доверительным методом. При этом решение, полученное ранее методом стохастического квазиградиента, не удовлетворяет вероятностному ограничению.

3.2 Определение множества допустимых скоростей ветра

В данном разделе исследуется задача построения множества допустимых векторов скорости ветра, при которых диспетчерская служба может дать самолету разрешение на взлет. Рассмотрим ситуацию, когда самолет уже готов к вылету из аэропорта А в аэропорт Б. Перед вылетом мы можем измерить скорость и направление ветра в аэропорту прилета Б и принять решение будет ли посадка безопасной. Но необходимо учесть, что в течение полета как скорость, так и направление ветра могут измениться, и если условия посадки станут небезопасными, самолет будет перенаправлен на запасной аэродром, что ведет к дополнительным финансовым потерям. Задача заключается в том, чтобы определить набор допустимых векторов скорости ветра, которые гарантируют безопасную посадку с заданной вероятностью.

Задача исследуется в полярных координатах. Пусть s_0 и α_0 – начальные скорость и направление ветра в аэропорту прилета. Продольная составляющая V_l^0 и поперечная составляющая V_r^0 вектора скорости ветра в момент вылета равны:

$$V_l^0 = s_0 \cos \alpha_0, \quad V_r^0 = s_0 \sin \alpha_0. \quad (3.2.1)$$

Пусть χ и ψ – переменные, контролирующие изменения скорости и направление ветра. Предполагаем, что обе эти величины независимы и имеют нормальное распределение с заданными параметрами:

$$\chi \sim \mathbb{N}(0, \sigma_\chi^2), \quad \psi \sim \mathbb{N}(0, \sigma_\psi^2). \quad (3.2.2)$$

Продольная компонента V_l^t и поперечная компонента V_r^t вектора скорости ветра в момент вылета и в момент прилета соотносятся следующим образом:

$$V_l^t = (s_0 + \chi) \cos(\alpha_0 + \psi), \quad (3.2.3)$$

$$V_r^t = (s_0 + \chi) \sin(\alpha_0 + \psi). \quad (3.2.4)$$

Пусть v_l^{\max} и v_l^{\min} – максимальная и минимальная допустимые скорости для продольного ветра (или попутного и встречного ветра), v_r^{\max} – максимально допустимая скорость поперечного ветра. Самолет при подлете к аэропорту назначения получит разрешение на посадку только при непревышении текущими параметрами ветра установленных максимальных границ, т.е. при выполнении следующего условия:

$$|V_r^t| \leq v_r^{\max}, \quad v_l^{\min} \leq V_l^t \leq v_l^{\max}. \quad (3.2.5)$$

Определим функцию вероятности, как функцию от s_0 и α_0 :

$$P(s_0, \alpha_0) = \mathbb{P} (|V_r^t| \leq v_r^{\max}, v_l^{\min} \leq V_l^t \leq v_l^{\max}) . \quad (3.2.6)$$

Определим уровень доверительной вероятности (уровень надежности) β . Суть задачи сводится к нахождению множества точек $W = \{(s_0, \alpha_0)\}$, для которых выполняется вероятностное ограничение:

$$P(s_0, \alpha_0) \geq \beta \quad \forall (s_0, \alpha_0) \in W. \quad (3.2.7)$$

Набор допустимых векторов скорости ветра должен быть замкнутым выпуклым множеством, содержащим нулевую точку, чтобы любая аффинная комбинация этих векторов давала также допустимый вектор скорости ветра.

Для нахождения множества W будет строиться аппроксимация внешней границы этого множества методом, аналогичным описанному в пункте 2.1.1. Для этого мы будем аппроксимировать линию уровня функции вероятности $P(s_0, \alpha_0)$. В начале необходимо найти начальную точку с координатами (s_0, α_0) . Например, положим $\alpha_0 = 0$ и численно решим уравнение:

$$P(s_0, 0) = \beta. \quad (3.2.8)$$

После нахождения стартовой точки поворачиваем вектор скорости ветра с заданным шагом $\Delta\alpha$ и находим изменение скорости ветра Δs , для которой вероятность остается неизменной, то есть выполняется условие:

$$\frac{\partial}{\partial \alpha_0} P(s_0, \alpha_0) \Delta \alpha + \frac{\partial}{\partial s_0} P(s_0, \alpha_0) \Delta s = 0. \quad (3.2.9)$$

Таким образом, граница множества допустимых векторов скорости ветра является решением следующего уравнения:

$$\frac{ds}{d\alpha} = - \frac{\frac{\partial}{\partial \alpha} P(s, \alpha)}{\frac{\partial}{\partial s} P(s, \alpha)}, \quad (3.2.10)$$

с начальным условием:

$$s(\alpha_0) = s_0. \quad (3.2.11)$$

Начальные данные, отражающие изменение ветра в течение 3-х часов в районе аэропорта в Московской области, были взяты в соответствии с [8]:

$$\sigma_\chi = 1,9 \frac{\text{М}}{\text{сек}}, \quad \sigma_\psi = 27^\circ, \quad (3.2.12)$$

$$\beta = 0,99, \quad v_r^{\max} = 15 \frac{\text{М}}{\text{сек}}, \quad v_l^{\min} = -25 \frac{\text{М}}{\text{сек}}, \quad v_l^{\max} = 10 \frac{\text{М}}{\text{сек}}. \quad (3.2.13)$$

Найденная начальная точка: $s_0 = 6,101$, $\alpha_0 = 0$. Вычисления выполнены за 36,8 секунд. Получен-

ная приближенная граница множества представлена на рисунке 3.2.

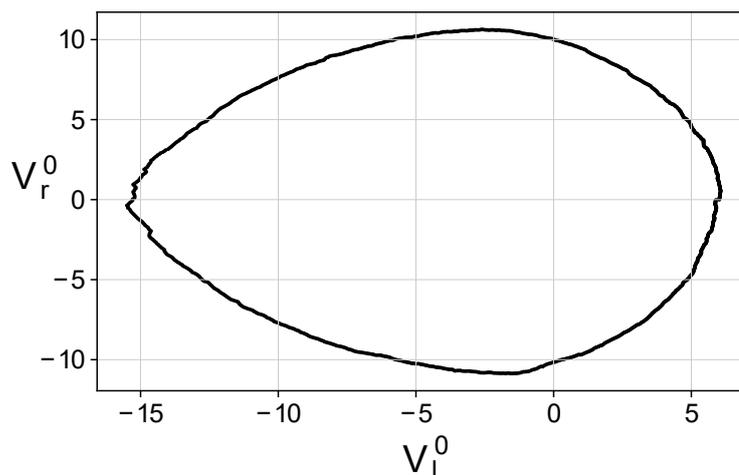


Рисунок 3.2. Аппроксимация границы множества допустимых скоростей ветра

В отличие от результатов из [8], в результате работы предлагаемого в настоящей работе метода было получено множество, близкое к выпуклому (с учетом погрешности вычислений и алгоритма построения). Множество, полученное в [8] было построено с помощью доверительного метода объединением двух конкретных доверительных множеств: эллипса и прямоугольника.

Эта же задача рассматривалась в качестве примера в работе [4], посвященной задаче построения доверительного множества для анализа стационарных стохастических систем. Результаты двух подходов очень близки друг к другу, но подход, предложенный в настоящей работе, оказывается проще в применении.

3.3 Проектирование системы водоснабжения

Рассмотрим систему водоснабжения в засушливом регионе. Предполагаем, что рядом находится источник с соленой водой, поэтому основной частью системы будет опреснительная установка. Это установка потребляет электроэнергию, вырабатываемую солнечной батареей, а полученная пресная вода хранится в резервуаре. Цена такой системы зависит от площади солнечной батареи и объема резервуара. В случае, когда пресной воды, выработанной опреснительной установкой, недостаточно, предполагается возможность доставить пресную воду из удаленного внешнего источника, но цена такой опции относительно высока. Также предполагаем, что в каждом месяце объем потребления воды известен заранее, но солнечная активность случайна – этот факт влияет на производительность солнечной батареи, которая, в свою очередь, влияет на объем пресной воды, выработанной опреснительной установкой. Задача состоит в том, чтобы спроектировать наиболее дешевую систему, которая с достаточно высокой вероятностью будет удовлетворять потребность в пресной воде.

Первый блок рассматриваемой системы - солнечная батарея площадью a м² со стоимостью установки m_1 долларов за 1 м², вырабатывающая электроэнергию для питания опреснительной установки. Пресная вода хранится в резервуаре объема b м³ со стоимостью постройки m_2 долларов за 1 м³. Функционирование системы рассматривается на горизонте t месяцев. Солнечная активность Y_i случайна для каждого i -ого месяца наблюдения, а объем опресненной в i -ом месяце воды выражается как aY_i . Предполагается, что величины Y_i , $i \in \overline{1, t}$ независимы, и каждая из них имеет гауссово распределение с известными параметрами

$$Y_i \sim \mathbb{N}(e_i, \sigma_i^2) \quad \forall i \in \overline{1, t}. \quad (3.3.1)$$

Потребление воды w_i м³ в i -й месяц полагается известным. В случае дефицита пресной воды возможна ее доставка из внешнего источника по цене m_0 долларов за 1 м³ (доставленный в i -й месяц объем воды обозначим как d_i м³).

Итоговая стоимость системы $C(h)$ может быть записана как:

$$C(h) = am_1 + bm_2 + m_0 \sum_{i=1}^t d_i. \quad (3.3.2)$$

где $h \triangleq (a, b, d_1, \dots, d_t)$ - набор управляемых параметров системы. Целью задачи, как было сказано ранее, является минимизация стоимости системы.

Определим ограничения задачи. Основное ограничение системы - удовлетворение спроса на пресную воду в каждом месяце. Этот спрос покрывается как за счет опресненной воды в текущем месяце, так и за счет излишков пресной воды, оставшихся с предыдущего месяца и хранящихся в резервуаре. Обозначим количество оставшейся в конце i -го месяца пресной воды как X_i , тогда излишек с предыдущего месяца выражается как $\min\{X_{i-1}, b\}$, а сама величина X_i выражается как:

$$X_i = \min\{X_{i-1}, b\} + aY_i + d_i - w_i, \quad X_0 = 0, \quad i = \overline{1, t}. \quad (3.3.3)$$

Система должна удовлетворять спрос в пресной воде с заданной вероятностью α , что эквивалентно следующему ограничению:

$$\mathbb{P}(X_i(h, Y_i) \geq 0) \geq \alpha \quad \forall i = \overline{1, t}. \quad (3.3.4)$$

Таким образом, задача формулируется как минимизация функции (3.3.2) по вектору управляемых параметров h при вероятностных ограничениях (3.3.4).

Для решения задачи перейдем к аппроксимации вероятностных ограничений задачи. Функ-

ция вероятности (3.3.4) может быть записана в следующем виде:

$$\mathbb{P}(X_i(h, Y_i) \geq 0, i = \overline{1, t}) = \mathbb{P}\left(\min_{i=\overline{1, t}} X_i(h, Y_i) \geq 0\right) \quad (3.3.5)$$

Для получения дифференцируемой аппроксимации ограничения внешний и внутренний минимума заменяются дифференцируемой аппроксимацией с помощью преобразования гладкого минимума (2.3.5) (аналогично подходу, описанному в пункте 2.3.1). Затем функция вероятности в ограничении задачи заменяется гладкой аппроксимацией (1.4.13).

Итоговая приближенная задача минимизации стоимости системы:

$$am_1 + bm_2 + m_0 \sum_{i=1}^t d_i \rightarrow \min_h \quad (3.3.6)$$

с вероятностным ограничением:

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} S_{\theta}(SM_{\gamma}[X_1(h, y), \dots, X_t(h, y)]) f_Y(y) dy_t \dots dy_1 \geq \alpha, \quad (3.3.7)$$

где $SM_{\gamma}(\cdot)$ – гладкий минимум t функций с параметром сглаживания γ .

В качестве метода решения этой задачи используется метод проекции градиента, в котором значения ограничений в точках и значения производной ограничений в точках находятся численно с помощью метода Монте-Карло.

Для численного решения возьмем параметры, представленные в [8]:

$$\alpha = 0,99, t = 6, m_0 = 250 \text{ \$/м}^3, m_1 = 37,5 \text{ \$/м}^2, m_2 = 100 \text{ \$/м}^3, \quad (3.3.8)$$

$$w_1 = 29,6 \text{ м}^3, w_2 = 23,9 \text{ м}^3, w_3 = 36,2 \text{ м}^3, w_4 = 82,1 \text{ м}^3, w_5 = 96,5 \text{ м}^3, w_6 = 173,4 \text{ м}^3, \quad (3.3.9)$$

$$e_1 = 0,0084, e_2 = 0,0083, e_3 = 0,0185, e_4 = 0,0063, e_5 = 0,123, e_6 = 0,137, \quad (3.3.10)$$

$$\sigma_1^2 = 5,8 \cdot 10^{-4}, \sigma_2^2 = 5,5 \cdot 10^{-4}, \sigma_3^2 = 12,3 \cdot 10^{-4}, \sigma_4^2 = 42,1 \cdot 10^{-4}, \sigma_5^2 = 81,8 \cdot 10^{-4}, \sigma_6^2 = 91,6 \cdot 10^{-4}. \quad (3.3.11)$$

Все вероятности и математические ожидания были вычислены с использованием метода Монте-Карло на выборке из 10000 наблюдений. Время работы алгоритма составило 292,69 секунды и в результате было получено следующее численное решение.

Оптимальная стоимость системы $C = 100613,5 \text{ \$}$; площадь солнечной батареи $a = 2488,1 \text{ м}^2$; объем резервуара $b = 31,4 \text{ м}^3$; объем доставленной воды в каждый из месяцев $d_1 = 15,3 \text{ м}^3$; $d_2 = 1,3 \text{ м}^3$; $d_3 = d_4 = d_5 = d_6 = 0 \text{ м}^3$. Таким образом, доставка воды из внешнего источника требуется только в первые два месяца. Полученное решение подразумевает большую пло-

щадь устанавливаемой солнечной батареи, но меньший совокупный объем доставляемой извне пресной воды по сравнению с ранее полученным решением [8].

3.4 Задача формирования портфеля ценных бумаг с логарифмической функцией потерь

В [5] рассматривалась задача формирования портфеля ценных бумаг с логарифмической функцией потерь, критерием в форме математического ожидания и равномерным распределением доходностей активов. Предполагается, что инвестор может вложить средства в безрисковый актив с доходностью b_0 и два рискованных актива с доходностями X_1 и X_2 . Доли капитала, вложенные в эти активы, обозначим u_1 и u_2 соответственно. В качестве функции потерь рассмотрим логарифм будущей стоимости портфеля, отражающий прирост капитала инвестора:

$$\Phi(u, X_1, X_2) = \ln(1 + (1 - u_1 - u_2)b_0 + u_1X_1 + u_2X_2) \quad (3.4.1)$$

Вероятностная критериальная функция примет вид:

$$P_\varphi(u) = \mathbb{P}\{\Phi(u, X_1, X_2) \geq \varphi^*\}. \quad (3.4.2)$$

где φ^* - логарифм желаемого уровня капитала. Задача состоит в максимизации этой вероятности:

$$P_\varphi(u) \rightarrow \max_u. \quad (3.4.3)$$

Доходности рискованных активов распределены равномерно:

$$X_1 \sim \mathbf{R}(-1, 1 + 2m_1), \quad X_2 \sim \mathbf{R}(-1, 1 + 2m_2). \quad (3.4.4)$$

Исходя из распределений доходностей рискованных активов и логики задачи, можно ввести следующее ограничение на уровень φ^* (имеется в виду, что при невыполнении этого условия задача становится тривиальной):

$$\ln b_0 < \varphi^* < \ln(\max(2m_1, 2m_2)). \quad (3.4.5)$$

Значения остальных параметров зададим следующим образом:

$$\varphi^* = 0,25, \quad b_0 = 0,05, \quad m_1 = 0,1, \quad m_2 = 0,2. \quad (3.4.6)$$

Результаты построения поверхностей точной вероятностной функции $P_\varphi(u)$ и аппроксимированной вероятностной функции $P_\varphi^\theta(u)$ при $\theta = 50$ представлены на рисунке 3.3. Для расчетов исходная поверхность для точной функции вероятностей строилась по сетке с разделением отрезка $[0, 1]$ по каждой переменной на 15 частей. При расчетах предполагалось, что у инвестора нет

возможности приобретать активы на заемные средства (т.е. существует запрет на короткие продажи); по этой причине расчеты ограничены областью, где $u_1 + u_2 \leq 1$. Также были построены поверхности второй смешанной частной производной по элементам вектора стратегии для точной и аппроксимированной функций вероятности (фрагменты полученных поверхностей представлены на рисунке 3.4).

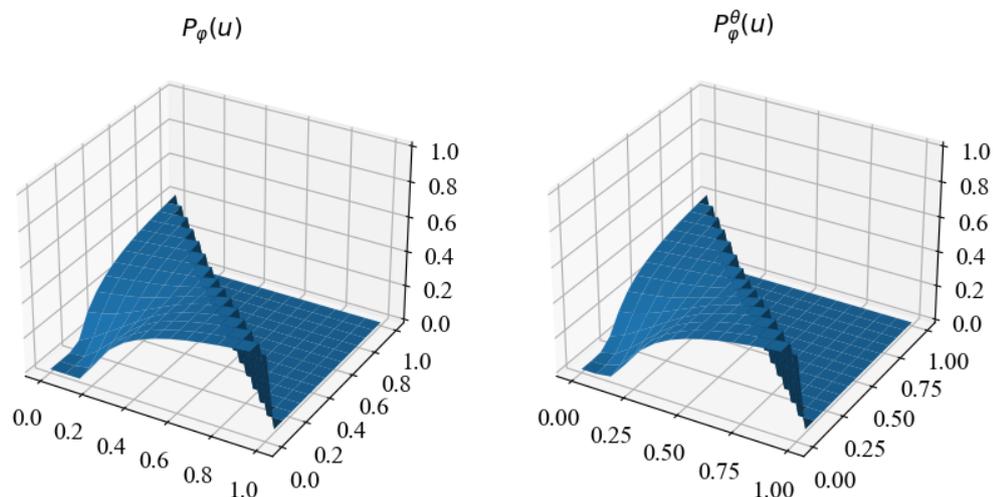


Рисунок 3.3. Сравнение поверхностей точной (слева) и аппроксимированной (справа) функций вероятности

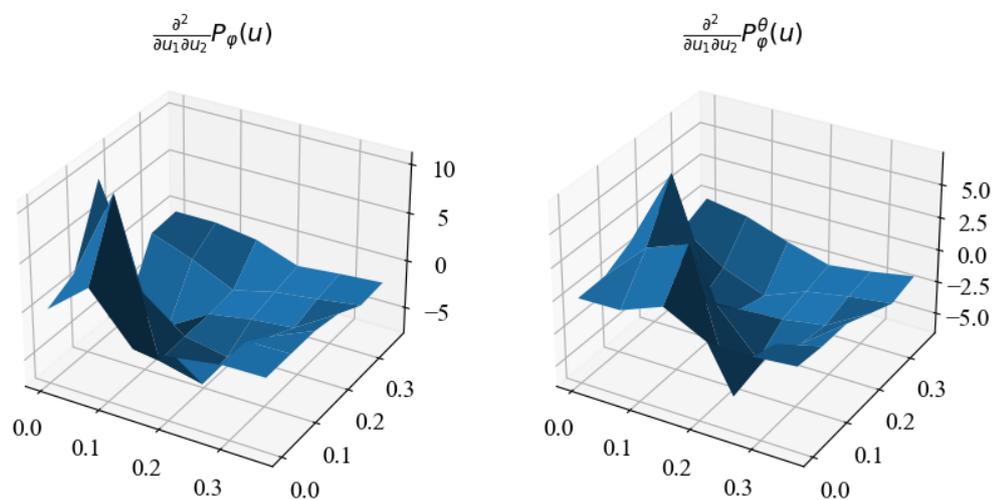


Рисунок 3.4. Сравнение фрагментов поверхностей вторых смешанных частных производных точной (слева) и аппроксимированной (справа) функций вероятности

Вторая смешанная частная производная гладкой аппроксимации при параметре $\theta = 50$ адекватно отражает форму исходной поверхности. Дополнительно было проведено сравнение поверхностей второй смешанной производной аппроксимированной функции вероятности для значений параметра сигмоиды 25 и 50. Получившиеся поверхности представлены на рисунке 3.5. Исходя из этих результатов, в данном случае глубина «оврага» (или величина «пиков» в окрестно-

сти точки $(0, 0)$) растет по мере увеличения параметра сигмоиды. В общем можно сделать вывод, что с уменьшением параметра сигмоиды увеличивается гладкость полученной поверхности, но теряется точность аппроксимации.

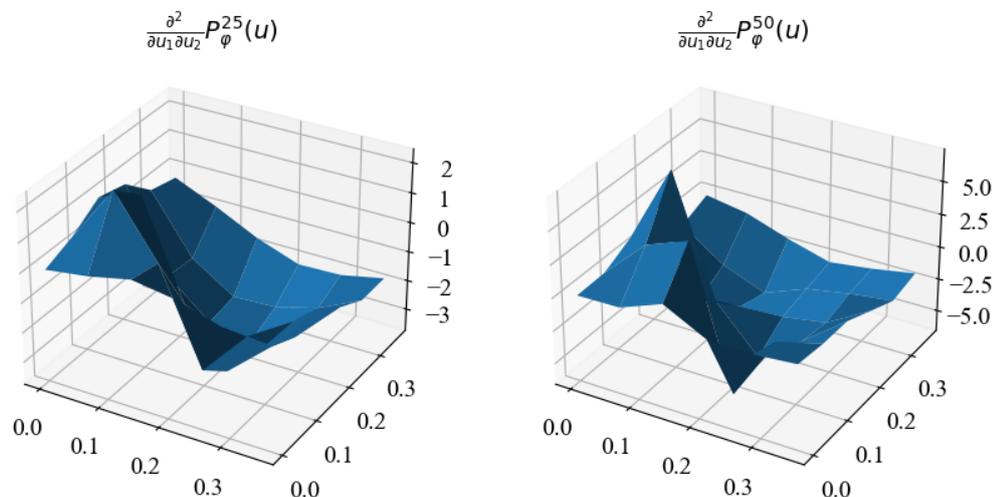


Рисунок 3.5. Сравнение фрагментов поверхностей вторых смешанных частных производных гладкой аппроксимации для параметров сигмоиды 25 (слева) и 50 (справа)

Предлагается решить приближенную к (3.4.3) задачу, в которой точная функция вероятности заменена на гладкую аппроксимацию. Зададим значение параметра сигмоиды $\theta = 50$ и количество реализаций для метода Монте-Карло равным 15000. Для решения задачи используется модификация метода Ньютона. В качестве начального приближения возьмем точку $u^{[0]} = (0,25; 0,25)$. Согласно классическому методу Ньютона, очередное приближение точки оптимума определяется через приближенное нахождение нуля градиента целевой функции, т.е. направление возрастания или убывания целевой функции не учитывается. В модификации новое приближение оптимума на каждом шаге выбирается как наилучшее по значению критерия из двух приближений: полученного по методу Ньютона и полученного по методу градиентного спуска. Новое приближение оптимума по методу Ньютона вычисляется как:

$$u^{[i+1]} = u^{[i]} - H^{-1}(u^{[i]}) \nabla P_{\varphi}^{\theta}(u^{[i]}), \quad (3.4.7)$$

где $H^{-1}(u^{[i]})$ - матрица, обратная матрице Гессе функции $P_{\varphi}^{\theta}(\cdot)$, вычисленная в текущей точке $u^{[i]}$. При решении учитываем ограничения на неотрицательность переменных и ограничение на сумму компонент вектора стратегии $u_1 + u_2 \leq 1$. При приближении текущего решения к границе области допустимых решений ограничивается шаг алгоритма до достижения соответствующей границы. При достижении границы области допустимых решений шаг алгоритма заменяется на проекцию полученного шага алгоритма на границу достигнутого ограничения. Результаты работы алгоритма решения представлены на рисунке 3.6.

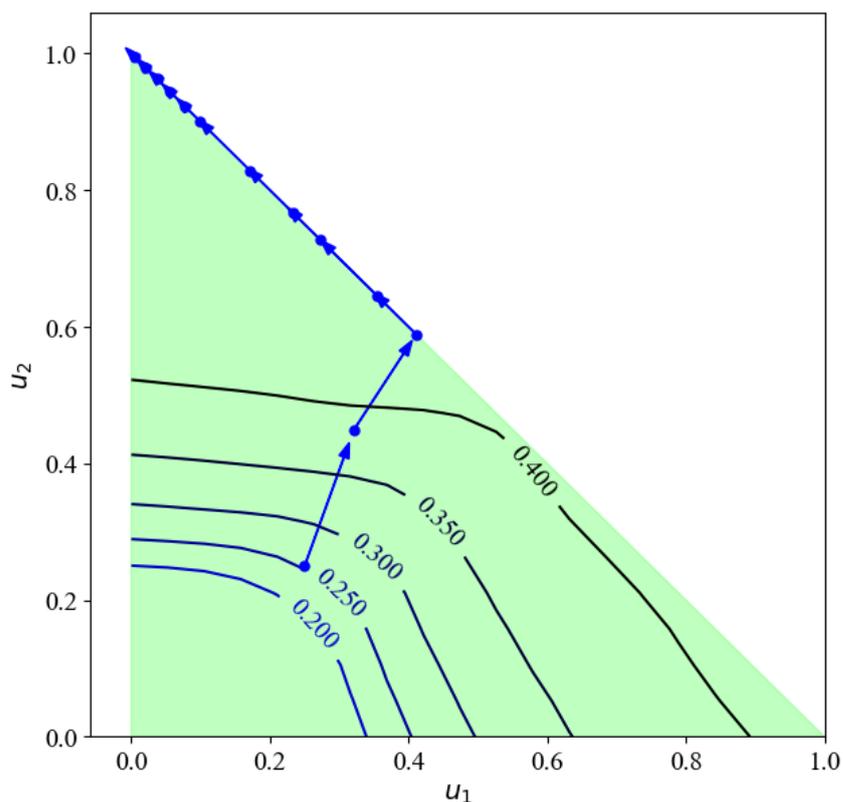


Рисунок 3.6. Иллюстрация пошаговой работы алгоритма по достижению оптимальной стратегии

На рисунке 3.6 черными и синими линиями показаны уровни функции вероятности, а стрелками — переходы от стратегии на текущем шаге к стратегии на следующем шаге. Алгоритм сходится в данном случае за 12 шагов, а оптимальной стратегией является точка $u = [0, 1]$. Таким образом, оптимальной стратегией является вложение всех средств в актив с более высокой ожидаемой доходностью.

3.5 Выводы по главе 3

С использованием гладкой аппроксимации функции вероятности и ее производных получены решения ряда теоретических и прикладных задач стохастического программирования:

- Получено решение задачи оптимизации площади взлетно-посадочной полосы при ограничении на вероятность успешной посадки методом проекции градиента. Площадь ВПП при найденном решении оказывается меньше площади, полученной при решении аналогичной задачи доверительным методом.
- Построено множество допустимых скоростей ветра, гарантирующее безопасную посадку летательного аппарата с учетом случайных изменений скорости и направления ветра за время перелета. Аппроксимация множества получена как решение дифференциального уравнения специального вида. Она соответствует множеству, полученному при решении той же задачи методами построения доверительных множеств поглощения, а также содержит

в себе множество, полученное при решении задачи доверительным методом.

- Получено решение задачи проектирования системы водоснабжения, состоящей из солнечной батареи, опреснительной установки и резервуара для хранения пресной воды. Эта задача с вероятностным ограничением решается методом проекции градиента. Полученное решение отличается от известного меньшим совокупным объемом воды, доставленной из внешнего источника, при большей площади солнечной батареи.
- С использованием модификации метода Ньютона и аппроксимации вторых производных функции вероятности решена задача формирования портфеля ценных бумаг с вероятностным критерием, логарифмической функцией полезности и равномерным распределением доходностей случайных параметров.

Основные результаты главы 3 опубликованы в работах [25, 26, 31, 35, 75].

4. Программный комплекс для анализа задач стохастического программирования с вероятностным критерием

В четвертой главе описывается программный комплекс PFDACC (Probability Function and Derivatives: Approximation, Calculation, Construction) для построения аппроксимированной функции вероятности и её производных. В разделе 4.1 приведена структура программы и описаны ее функциональные блоки. В разделе 4.2 приведен общий алгоритм работы с программой вкуче с указанием возможного функционала программы. В разделе 4.3 рассмотрены два случая использования программы: построение аппроксимированных функций распределения и плотности и решение задачи формирования портфеля ценных бумаг с минимальным риском.

4.1 Структура и функциональные блоки программы

Программный комплекс PFDACC реализован на языке программирования Python и состоит из набора модулей – файлов расширения .py, реализующих те или иные блоки программы. В модулях используются следующие библиотеки языка Python: NumPy, PyQt5, Pandas, Matplotlib и SciPy.

Входной точкой программы является модуль MainGUI, в котором содержится описание графического интерфейса программы. Данные, полученные через интерфейс программы, обрабатываются, хранятся и пересылаются в другие функциональные блоки с помощью объекта, описанного в модуле GUI_DO.

Интерфейс программы предполагает ввод пользователем целевой функции (функции потерь) в формате LaTeX наряду с выбором типов переменных и их значений. Строковая целевая функция автоматически преобразуется в объект специального класса LabFunc, описанный в модуле Labos_flow_v2. Этот класс позволяет хранить в древовидной структуре иерархию стандартных нелинейных функций, а также переменных, унарных и бинарных операций, составляющих итоговую функцию. Стандартные функции представляют собой объекты дочерних к LabFunc классов и включают в себя синус, косинус, тангенс, арктангенс, натуральный логарифм, экспоненту, минимум нескольких аргументов и максимум нескольких аргументов. Каждая функция, записанная в объекте LabFunc, поддерживает операцию дифференцирования: производная определяется либо по правилам дифференцирования сложной функции, либо напрямую в случае дифференцирования стандартной функции. Процедура представления введенной пользователем целевой функции в виде описанной древовидной структуры описана в модуле Latex_parser.

В модуле labos_point описан класс LabPoint, объекты которого являются аргументами

функций класса LabFunc. Этот класс является расширением стандартного типа данных-словаря языка Python и поддерживает операции над векторами и работу с массивами чисел в качестве значений. При этом вычисления, проводимые с объектами этого класса, реализованы при помощи векторизуемых процедур библиотеки NumPy, что повышает эффективность вычислений в сравнении с возможными нативными реализациями.

При наличии случайных величин среди переменных целевой функции, программой вычисляются значения и строится график аппроксимированной функции вероятности для заданного уровня потерь на сетке значений возможных стратегий. Значения подлежащих вычислению интегралов находятся с помощью метода Монте-Карло, а сама процедура вычисления находится в модуле `stochastic_model`. Также в этом модуле находится описание генераторов поддерживаемых программой типов случайных величин. На основе полученных значений программой строится график кривой или график поверхности функции при помощи методов, описанных в модуле `PainterBackend`.

4.2 Алгоритм работы с программой и ее функциональные возможности

Для начала работы с программой необходимо запустить модуль `MainGUI`, после чего перед пользователем появится первоначальный интерфейс программы, изображенный на рисунке 4.1

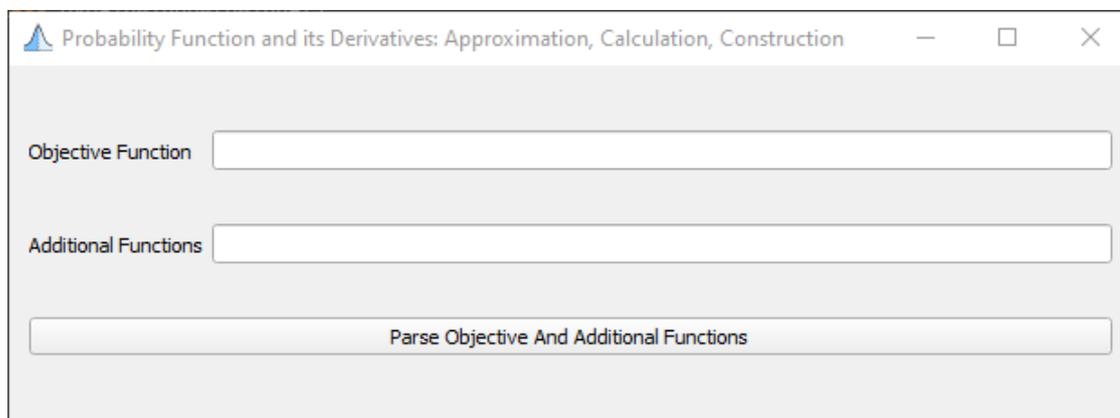


Рисунок 4.1. Начальный интерфейс программы

В поле “Objective Function” в формате LaTeX вводится целевая функция, состоящая из стандартных функций, арифметических операций, скобок и переменных. Переменные должны состоять из букв латинского алфавита и цифр, не должны начинаться с цифры, не должны совпадать с названием стандартной или вспомогательной функции и допускают задание нижнего индекса по правилам записи в формате LaTeX. Программой поддерживаются круглые, квадратные и фигурные скобки, в том числе с префиксами LaTeX `\left` и `\right`. Для упрощения записи целевой функции пользователем могут быть введены вспомогательные функции, задающиеся в поле

“Additional Functions” следующим образом: по правилам задания названия переменной выбирается обозначение функции, не совпадающее с названиями других переменных, и записывается в поле вспомогательных функций; затем ставится знак ”равно”и записывается тело вспомогательной функции по правилам записи целевой функции (например, $\text{func1}=\sin(x)+\cos(x)$). Несколько вспомогательных функций задаются в одном поле через разделитель - точку с запятой. После задания вспомогательных функций их обозначения можно использовать при задании целевой функции в качестве переменных.

При задании валидных целевой и вспомогательных функций по нажатию кнопки “Parse Objective and Additional Functions” программа автоматически создаст объект с функциональной интерпретацией целевой функции, определит переменные, входящие в эту функцию, и предоставит обновленное окно интерфейса, представленное на рисунке 4.2.

Рисунок 4.2. Интерфейс программы при задании целевой и вспомогательных функций

В обновленном окне находятся два блока: блок “Function Variables”, в котором задаются типы и значения переменных, входящих в целевую функцию, и блок “Static Variables”, в котором задаются значения служебных переменных - параметров алгоритмов вычисления.

Для каждой переменной в блоке “Function Variables” пользователю необходимо задать ее тип путем выбора в селекторе и ввести значения в соответствующее поле или поля. Программа поддерживает четыре типа переменных:

- 1) Константа (значение селектора “Constant”). Ее значение задается пользователем вручную и не требует повторного разбора формулы.
- 2) Компонента стратегии (значение селектора “Control”). Для переменной этого типа в двух полях указываются наибольшее и наименьшее значение для построения сетки, на которой будут вычисляться приближенные значения функции вероятности. Данный тип можно присвоить одной или двум переменным.
- 3) Случайная величина (значение селектора “Stochastic”). Случайная величина должна иметь один из стандартных законов распределения, который выбирается в отдельном селекторе. В поле значения указывается параметр (или несколько параметров через точку с запятой) выбранного распределения. Программой поддерживаются следующие непрерывные распределения: равномерное, нормальное, экспоненциальное, логистическое, логнормальное, Лапласа, Хи-квадрат, Стьюдента и Фишера. Все случайные величины предполагаются независимыми; для использования зависимых величин предлагается либо преобразовать исходное выражение, либо использовать выборку в качестве переменной.
- 4) Выборка (значение селектора “From File”). В этом случае переменная трактуется как случайная величина, для которой известна выборка значений. Выборка должна храниться в табличном формате в файле формата .xls, .xlsx или .csv. Файл должен быть предварительно загружен в программу по нажатию кнопки “Add Stochastic Variable Realization From File”. В одном файле могут храниться выборки нескольких переменных; при этом требуемая переменная выбирается в отдельном селекторе на основании названий столбцов переменных в файле.

В блоке “Static Variables” находятся параметры алгоритмов вычисления, причем для них заданы значения по умолчанию, которые могут быть изменены пользователем. В случае наличия в целевой функции переменной типа “случайная величина” или “выборка”, вместо исходной целевой функции будет строиться функция вероятности. В этом случае необходимо задать значение служебной переменной “Phi”, являющейся значением уровня потерь. По желанию можно изменить значения других служебных переменных. Переменная “Theta” является параметром гладкости сигмоиды. Переменная “Step” соответствует размеру шага при построении значений функции на сетке. Переменная “M-C” - количество реализаций, используемых для расчета функций по методу Монте-Карло. Если целевая функция содержит минимум или максимум, то при расчетах он будет заменен с помощью преобразования гладкого минимума (максимума) с отдельным парамет-

ром (см. раздел 2.3.1); этот параметр задается в поле служебной переменной “Min/Max”.

Далее, под блоками переменных в интерфейсе программы предлагается задать иные параметры вычисления. Для вычисления значений производных функции вместо значений самой функции вероятности, пользователю требуется отметить пункт “Calculate first-order derivatives” или “Calculate second-order derivatives” для вычисления аппроксимированных значений производной первого или второго порядка соответственно. Переменные дифференцирования выбираются пользователем в соответствующем селекторе; в случае вычисления второй производной необходимо выбрать две переменных отдельно в каждом из селекторов. В программе также имеется возможность вычисления значений функции вероятности и ее производных без учета гладкой аппроксимации функции вероятности на основе метода Монте-Карло; производные в этом случае приближаются конечными разностями с фиксированным шагом, указанным в служебном параметре “Step”. Для включения опции прямого расчета требуется отметить пункт “Calculate true values”.

При нажатии на кнопку “Calculate” производится расчет значений установленной функции на сетке ее аргументов. График искомой функции или поверхности строится автоматически при успешном расчете значений функции. Этот график откроется в отдельном окне, поддерживающем масштабирование контента, перемещение с помощью мыши, сохранение графика в качестве изображения в файл и другие операции. Результаты расчетов могут быть сохранены отдельно в табличный файл в формате, совместимом с ПО Microsoft Excel, по нажатию на кнопку “Dump Results Of Previous Calculation Into Excel File”. При последующим запуске программы данные табличные файлы можно использовать повторно для построения графиков кривой или поверхности; для этого необходимо нажать на кнопку “Plot Results Of Previous Calculation From Dumped File” и выбрать в открывшемся диалоговом окне необходимый файл.

4.3 Примеры использования программного комплекса

4.3.1 Построение функции распределения и плотности нормального распределения

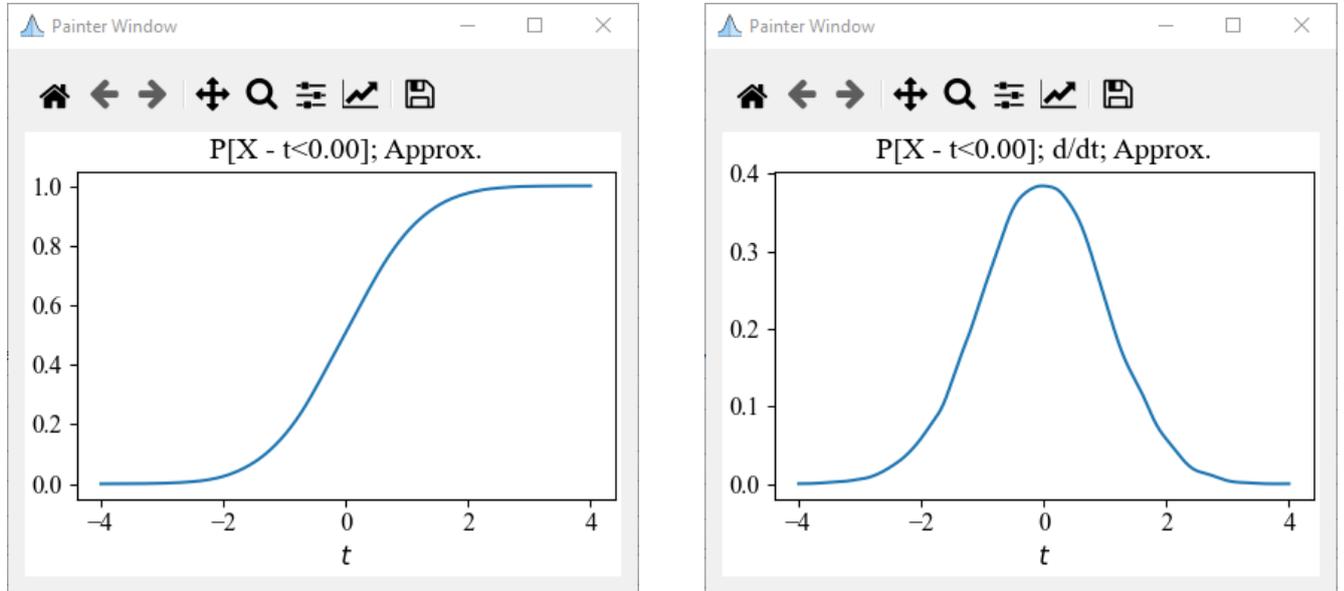
В данном подразделе рассматривается тривиальный пример построения функций распределения и плотности нормального распределения с помощью описанного программного комплекса. Согласно определению функции распределения $F_X(a)$ случайной величины X верно:

$$F_X(t) = \mathbb{P}\{X \leq t\} = \mathbb{P}\{X - t \leq 0\}. \quad (4.3.1)$$

Соотнеся полученное выражение и определение оптимизируемой функции $P_\varphi(u)$ (1.1.1), получим, что построение функции распределения переменной X эквивалентно заданию в программе целевой функции $\Phi(u, X) = X - t$, заданию уровня потерь $\varphi = 0$ и указания типов двух переменных: X - случайная переменная с нормальным законом распределения, t - переменная типа

”стратегии”.

Построение функции плотности переменной X сводится к построению производной указанной функции по переменной t . Окна программы с результатами расчетов для нормально распределенной переменной $X \sim \mathbf{N}(0; 1)$ и для диапазона значений переменной $t \in [-4, 4]$ приведены на рисунке 4.3.



(a) Функция распределения

(b) Функция плотности

Рисунок 4.3. Окна программы с результатами построения

4.3.2 Задача формирования портфеля акций с минимальным риском

Рассмотрим пример формирования портфеля из акций трех компаний из разных секторов экономики с ежедневной ребалансировкой портфеля в течение срока инвестирования: Yandex (биржевой тикер YNDX), ФосАгро (PHOR) и НЛМК (NLMK). Обозначим цены акций этих трех компаний в момент закрытия торгов как C_Y , C_P и C_N соответственно, а момент открытия торгов (покупки портфеля) как C_Y^0 , C_P^0 и C_N^0 . Первоначальный капитал обозначим как $K > 0$; долю капитала, вложенную в акции Yandex как u_Y , и долю капитала, вложенную в акции ФосАгро как u_P , тогда цена портфеля ценных бумаг в момент закрытия торгов задается выражением:

$$K \left(u_Y \frac{C_Y}{C_Y^0} + u_P \frac{C_P}{C_P^0} + (1 - u_Y - u_P) \frac{C_N}{C_N^0} \right). \quad (4.3.2)$$

Под портфелем с минимальным риском подразумевается портфель, составленный из указанных ценных бумаг, для которого вероятность снижения его рыночной стоимости по результатам минимальна. Задача формирования такого портфеля может быть записана как:

$$\mathbb{P} \left(u_Y \frac{C_Y - C_Y^0}{C_Y^0} + u_P \frac{C_P - C_P^0}{C_P^0} + (1 - u_Y - u_P) \frac{C_N - C_N^0}{C_N^0} \geq 0 \right) \rightarrow \max_{u_Y, u_P} \quad (4.3.3)$$

при ограничениях

$$u_Y \geq 0, u_P \geq 0, u_Y + u_P \leq 1, \quad (4.3.4)$$

Предполагается, что акции обладают бесконечной делимостью, а короткие продажи запрещены. Обозначим дневные приращения цен акций как $\Delta C_Y, \Delta C_P, \Delta C_N$ соответственно.

График цен закрытия торгов по акциям обозначенных компаний за период с 04.01.2021 по 30.12.2021 представлен на рисунке 4.4 (для сохранения масштаба график цены акций NLMK приведен за 20 акций компании).

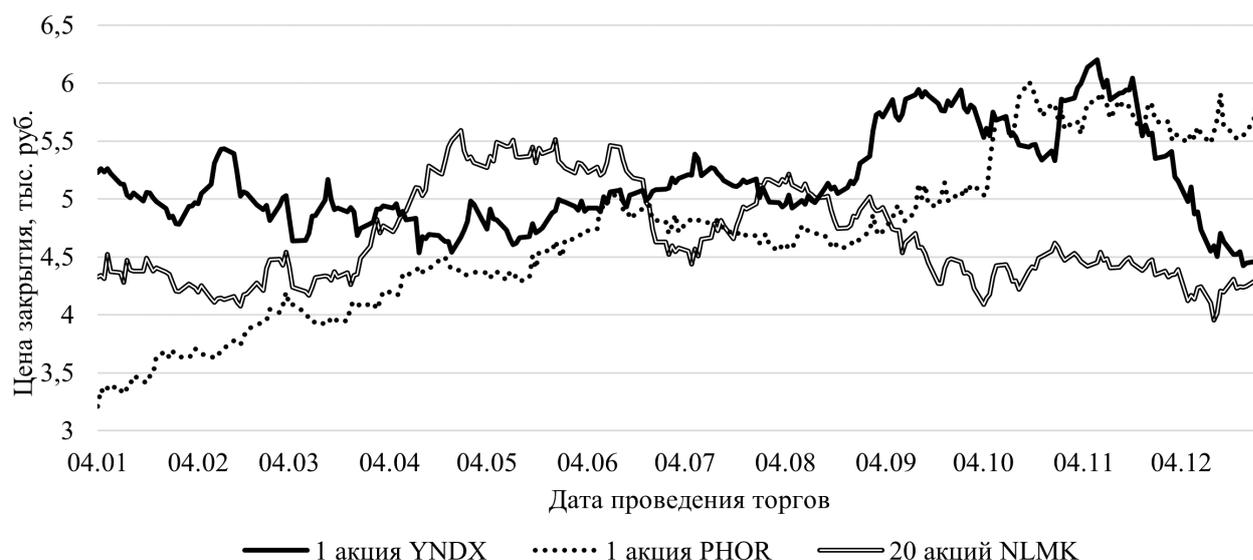


Рисунок 4.4. Графики цен закрытия акций Яндекс, ФосАгро и НЛМК за 2021 год

На основе этих данных строится выборка относительных изменений цен закрытия акций для каждой компании, т.е. выборки переменных $\Delta C_Y, \Delta C_P, \Delta C_N$. При построении аппроксимированной функции вероятности они используются как выборки случайной величины вместо случайных величин с заданным теоретическим распределением. Аппроксимированная функция вероятности строится путем задания в интерфейсе программы следующей целевой функции и набора параметров:

$$\Phi(u, \Delta C) = -u_Y \Delta C_Y - u_P \Delta C_P - (1 - u_Y - u_P) \Delta C_N, \quad (4.3.5)$$

$$\varphi = 0, u_Y \in [0, 1], u_P \in [0, 1]. \quad (4.3.6)$$

Дополнительно устанавливается значение параметра сигмоиды $\theta = 2$ для большего сглаживания получаемой поверхности (меньшие значения параметры сигмоиды позволяют уменьшить чувствительность вычисляемых значений функции вероятности к значениям входящих в целевую функцию переменных). Полученная в результате работы программы поверхность представлена

на рисунке 4.5.

$$-u_Y * c_Y - u_P * c_P - c_N + u_Y * c_N + u_P * c_N; \text{Approx.}$$

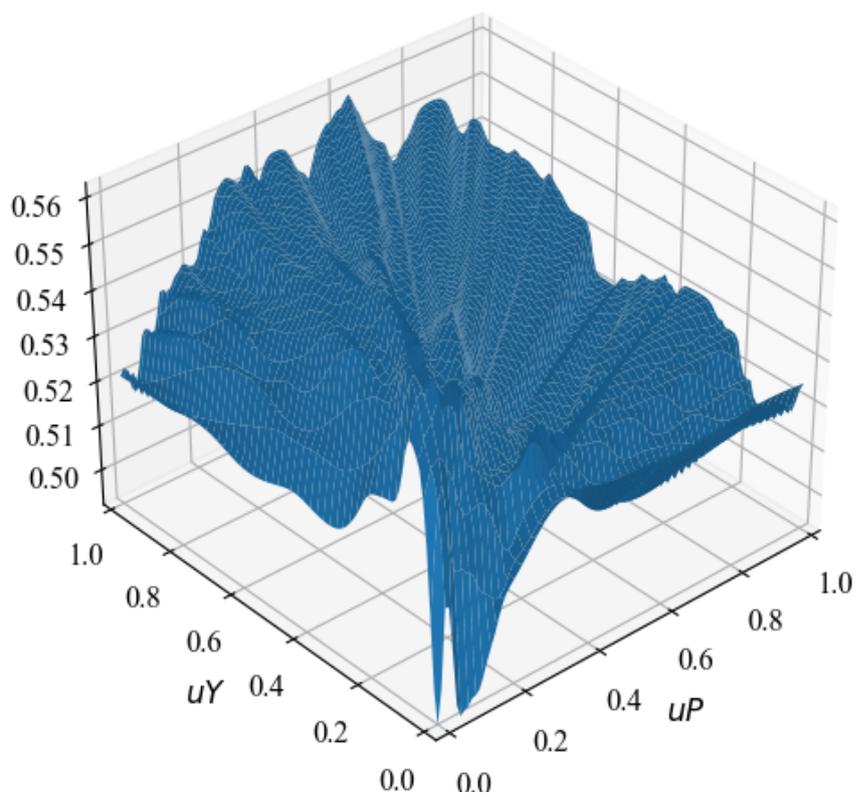


Рисунок 4.5. Аппроксимированная поверхность функции вероятности в задаче формирования портфеля

Построенные точки поверхности аппроксимированной функции вероятности выгружаются из программы в табличном виде и используются для непосредственного нахождения оптимальной стратегии (с учетом ограничения $u_Y + u_P \leq 1$). Исходя из полученных данных, оптимальная стратегия имеет вид:

$$u_Y = 0.14, u_P = 0.31, u_N = 1 - u_Y - u_P = 0.55. \quad (4.3.7)$$

Для оценки полученного решения строится график стоимости оптимального портфеля при первоначальном капитале в 1000 рублей в предположении, что цены закрытия торгов совпадают с таковыми в рассматриваемом периоде (т.е. соответствуют ценам, изображенным на рисунке 4.4). Данный график представлен на рисунке 4.6.

График стоимости оптимального портфеля имеет явно выраженный тренд и умеренную волатильность, но стоимость портфеля снижается в период с 01.12.2021 по 30.12.2021, поскольку ни один актив из выбранных для рассмотрения не рос в цене в этот период. Эта проблема может быть решена дальнейшей диверсификацией. Включение в портфель большего числа активов выходит за



Рисунок 4.6. Динамика стоимости оптимального портфеля за 2021 год

рамки возможностей программного комплекса, но остается реализуемым с использованием описанных ранее подходов и программных модулей.

4.4 Выводы по главе 4

На основе полученных теоретических положений о аппроксимации вероятностных критериев разработан программный комплекс PFDACC для построения аппроксимации функции вероятности и ее производных первого и второго порядков. Комплекс обладает достаточно широкими возможностями, может использоваться как дополнительное средство исследования задач стохастического программирования, и его работа продемонстрирована на примере решения задачи формирования портфеля акций с минимальным риском.

Основные результаты главы 4 опубликованы в работах [28, 32].

Заключение

Основным итогом диссертационной работы являются новые алгоритмы решения задач стохастического программирования с вероятностными критериальными функциями на основе аппроксимации функций вероятности и ее производных с помощью сигмоидальной функции. В ходе работы были получены следующие результаты:

- 1) Доказана теорема о сходимости аппроксимированной функции вероятности к исходной и теоремы о сходимости производной аппроксимированной функции вероятности к производной функции вероятности по компонентам вектора стратегии и уровню потерь при возрастании параметра сигмоиды в случае одномерной случайной величины [25, 77].
- 2) Доказаны аналоги теорем о сходимости аппроксимированной функции вероятности и ее производных к исходным функциям в случае многомерной случайной величины и получены аппроксимационные выражения в форме объемных интегралов [77].
- 3) Разработан новый алгоритм решения задачи стохастического программирования с вероятностным критерием, полиэдральной функцией потерь и полиэдральной функцией ограничений на основе рассматриваемой аппроксимации и аппроксимации гладкого минимума [76].
- 4) Разработан новый алгоритм аппроксимации внешней границы альфа-ядра вероятностной меры на основе аппроксимации функции вероятности [77].
- 5) Решены 4 известные прикладные задачи стохастического программирования с использованием рассматриваемой аппроксимации и обычных численных методов в качестве иллюстрации универсальности аппроксимации [75].
- 6) Разработан и зарегистрирован программный комплекс для построения графиков и поверхностей аппроксимированной функции вероятности, ее первых и вторых производных по компонентам вектора стратегии [28, 32].

Теоретическая часть приведенного исследования может быть расширена за счет доказательств сходимости аппроксимации функции квантили вместе с ее производными к оригинальным функциям. Помимо этого, исследование может быть продолжено в области построения методов решения теоретических задач и адаптация аппроксимации к задачам иного типа (например, к задачам с VAR- и CVAR-критериями). Большой перспективой также обладает развитие исследования в области прикладных задач стохастического программирования, а именно в разработке численных методов решения задач, для которых неприменимы существующие методы преобразования

задач, или существующие методы решения являются достаточно трудоемкими.

Список литературы

1. *Васильева С.Н.* Алгоритмы анализа и оптимизации квантильного критерия в задачах стохастического программирования с билинейными и квазилинейными функциями потерь.: дис. канд. ф.-м. наук: 05.13.01, 05.13.18: защищена 16.11.2018.
2. *Зорич В.А.* Математический анализ : учебник для студентов математических и физико-математических факультетов и специальностей высших учебных заведений // 6-е изд., доп.. – Москва : Изд-во МЦНМО, 2012. – 700 с. – ISBN 978-5-94057-892-5.
3. *Иванов С.В., Кибзун А.И.* Выборочная аппроксимация двухэтапной задачи стохастического линейного программирования с квантильным критерием // Тр. ИММ УрО РАН. 2017. Т. 23. № 3. С. 134–143.
4. *Иванов С.В., Кибзун А.И., Степанова А.С.* Построение доверительного множества поглощения в задачах анализа статических стохастических систем // *АиТ.* 2020. № 4. С. 21-36.
5. *Игнатов А.Н., Кибзун А.И.* О формировании портфеля ценных бумаг с равномерным распределением по логарифмическому критерию с приоритетной рисковой составляющей // *Автомат. и телемех.*, 2014. № 3. С. 87–105.
6. *Игнатов А.Н., Кибзун А.И.* Сведение двухшаговой задачи стохастического оптимального управления с билинейной моделью к задаче смешанного целочисленного линейного программирования // *Автоматика и телемеханика.* 2016. № 12. С. 89–111.
7. *Кан Ю.С., Кибзун А.И.* Свойства выпуклости функций вероятности и квантили в задачах оптимизации // *АиТ.* 1996. № 3. С. 82–102.
8. *Кан Ю.С., Кибзун А.И.* Задачи стохастического программирования с вероятностными критериями // М.: Физматлит. 2009. 372 с.
9. *Кан Ю.С., Суринов Р.Т.* О неравенстве треугольника для критерия VaR // *Моделирование и анализ безопасности и риска в сложных системах: Тр. Международной научной школы МАБР-2004.* – СПб.: СПбГУАП, 2004. С. 243–248.
10. *Кибзун А.И., Курбаковский В.Ю.* Численные алгоритмы квантильной оптимизации и их применение к решению задач с вероятностными ограничениями // *Изв. РАН, Техническая кибернетика,* 1992. № 1. С. 75–81.

11. *Кибзун А.И., Малышев В.В.* Обобщенный минимаксный подход к решению задач с вероятностными ограничениями // Изв. АН СССР, Техническая кибернетика, 1984. № 1. С. 20–29.
12. *Малышев В.В., Кибзун А.И.* Анализ и синтез высокоточного управления летательными аппаратами //– М.: Машиностроение, 1987. 303 С.
13. *Кибзун А.И., Малышев В.В., Чернов Д.Э.* Два подхода к решению задач стохастической оптимизации // Изв. АН СССР, Техническая кибернетика, 1988. № 3. С. 20–25.
14. *Кибзун А.И., Наумов А.В.* Гарантирующий алгоритм решения задачи квантильной оптимизации // Космические исследования, 1995. № 2. С. 160–165.
15. *Кибзун А.И., Наумов А.В., Норкин В.И.* О сведении задачи квантильной оптимизации с дискретным распределением к задаче смешанного целочисленного программирования // АиТ. 2013. № 6. С. 66–86.
16. *Кибзун А.И., Соболев В.Р.* Двухшаговая задача хеджирования европейского колл-опциона при случайной длительности транзакций // Тр. ИММ УрО РАН. 2015. Т. 21. № 3. С. 164–174.
17. *Кибзун А.И., Третьяков Г.Л.* Дифференцируемость функции вероятности. // Доклады Академии наук. 1997. № 2. С. 159–161.
18. *Ленн Р.* Максимизация функции вероятности при простых ограничениях // Изв. АН ЭССР, физ.-мат., 1979, 28, № 4, С. 303–309.
19. *Ленн Р.* Минимизация гладкой функции при вероятностных ограничениях // Изв. АН ЭССР, физ.-мат., 1980, 29, № 2, С. 140–144.
20. *Норкин В.И., Роевко Н.В.* α -вогнутые функции и меры их приложения. // Кибернетика и системный анализ, 1991, 6, С. 77–88.
21. *Райк Э.* О функции квантиля в стохастическом нелинейном программировании // Изв. АН ЭССР, физ.-мат., 1971, 20, № 2, С. 229–231.
22. *Райк Э.* Качественные исследования в задачах стохастического нелинейного программирования // Изв. АН ЭССР, физ.-мат., 1971, 20, № 1, С. 8–14.
23. *Райк Э.* О задачах стохастического программирования с функционалами вероятности и квантиля // Изв. АН ЭССР, физ.-мат., 1972, 21, № 2, С. 142–148.

24. Райк Э. Дифференцируемость по параметру функции вероятности и стохастический псевдо-градиентный метод для ее оптимизации // Известия Академии наук Эстонской ССР: физика, математика. 1975. № 1. С. 3–8.
25. Соболев В.Р., Торишинный Р.О. О гладкой аппроксимации вероятностных критериев в задачах стохастического программирования // Тр. СПИИРАН. Т. 19. № 1. 2020. С. 180–217.
26. Соболев В.Р., Торишинный Р.О. Об аппроксимации функции вероятности как критерия в задачах оптимизации // XXIV Международная научная конференция «Системный анализ, управление и навигация», Евпатория, 30 июня-07 июля 2019. — Тезисы докладов. — М.: МАИ-Принт, 2019. С. 168-170.
27. Соболев В.Р., Торишинный Р.О. Новый алгоритм аппроксимации границ альфа-ядер // 19-я Международная конференция «Авиация и космонавтика», Москва, 23-27 ноября 2019. — Тезисы докладов. — М.: Издательство «Перо», 2020. С. 497-498.
28. Соболев В.Р., Торишинный Р.О. Программно-алгоритмический комплекс для построения гладких аппроксимаций функции вероятности и её производных. // Федеральная служба по интелект. собственности. Св-во о гос. регистрации программы для ЭВМ № 2022613132. 2022.
29. Тамм Э. О квазивыпуклости функций вероятности и квантили // Изв. АН ЭССР, физ.-мат. 1976. Т. 25. № 2. С. 141–144.
30. Тамм Э. О минимизации функции вероятности // Изв. АН ЭССР, физ.-мат., 1979, 28, № 1, С. 17–24.
31. Торишинный Р.О. О применении численных методов второго порядка к задачам стохастического программирования с функцией вероятности // Тр. МАИ. 2021. № 121. С. 1-27.
32. Торишинный Р.О. Программный комплекс для анализа задач стохастического программирования с вероятностным критерием // ВКит. 2022. Т. 19. № 5(215). С. 3-12. DOI 10.14489/vkit.2022.05.pp.003-012.
33. Торишинный Р. О. Приближенное вычисление градиента функции вероятности и функции квантили // 17-ая Международная конференция «Авиация и космонавтика», Москва, 19-23 ноября 2018. — Тезисы докладов. — М.: Типография «Люксор», 2018. С. 473-474.
34. Торишинный Р.О. О сходимости и погрешности аппроксимации индикаторной функции в вероятностном критерии // XLV Международная молодежная научная конференция «Гагаринские

- чтения», Москва, 28 марта-19 апреля 2019. — Тезисы докладов. — М.: Изд-во МАИ, 2019. С. 1264-1265.
35. *Торшин П.О.* Об аппроксимации функции вероятности как критерия в задачах оптимизации // XXVI Международная научная конференция студентов, аспирантов и молодых ученых «Ломоносов-2019», Москва, 8-12 апреля 2019. — Тезисы докладов. — М.: ООО «МАКС Пресс», 2019. С. 140-142.
 36. *Урясьев С.П.* Адаптивные алгоритмы стохастической оптимизации и теории игр. // М.:Наука, 1979.
 37. *Шуряев А.Н.* Вероятность // М.:Наука. 2007. 552 с.
 38. *Beraldi P., Ruszczyński A.* A branch and bound method for stochastic integer problems under probabilistic constraints // Optimization Methods & Software. 2002. V. 17. No. 3. P. 359–382.
 39. *Borell C.* Convex Set Functions in d-Space // Period. Math. Hung. 1975. Vol. 6. No. 2. P. 111–136.
 40. *Charnes A., Cooper W.* Chance-Constrained Programming. // Management Sci., 1959, 6, P.73–79.
 41. *Charnes A., Cooper W.* Deterministic Equivalents for Optimizing and Satisficing under Chance-Constraints. // Oper. Res., 1963, 11, P.18–39.
 42. *Charnes A., Cooper W., Symonds G.* Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil. // Management. Sci., 1958, 4, P.235–263.
 43. *Dentcheva D., Penev S., Ruszczyński A.* Statistical estimation of composite risk functionals and risk optimization problems // Annals of the Institute of Statistical Mathematics. 2017. vol. 69. no 4. pp. 737–760.
 44. *Dentcheva D., Ruszczyński A.* Risk forms: representation, disintegration, and application to partially observable two-stage systems // Mathematical Programming. 2019. pp. 1–21.
 45. *Dentcheva D., Shapiro A., Ruszczyński A.* Lectures on Stochastic Programming. Modeling and Theory // Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 2009. 450 p.
 46. *Dentcheva D., Prekopa A., Ruszczyński A.* Concavity and efficient points of discrete distributions in probabilistic programming // Mathematical Programming. 2000. No. 89. P. 55–77.
 47. *Ermoliev Yu., Norkin V., Wets R.* The Minimization of Discontinuous Functions: Mollifier Subgradients // Working Paper WP-92-73, IIASA, Laxenburg, Austria, 1992.

48. *Ermoliev Yu., Norikin I., Wets R.* The Minimization of Semicontinuous Functions: Mollifier Subgradients // *Siam Journal on Control and Optimization* 33 (1995): 149-167.
49. *Garnier J., Omrane A., Rouchdy Y.* Asymptotic formulas for the derivatives of probability functions and their Monte Carlo estimations // *European Journal of Operational Research*. 2009. vol. 198. no. 3. pp. 848–858.
50. *Gartska S.* The Economic Equivalence of Several Stochastic Programming Models. In: *Stochastic Programming*, ed. M.A.H. Dempster, Academic Press, New York, 1980, P. 83–91.
51. *Henrion R.* Gradient estimates for Gaussian distribution functions: application to probabilistically constrained optimization problems // *Numerical Algebra, Control and Optimization*. 2012. vol. 2. no. 4. pp. 655–668.
52. *Henrion R.* On the Connectedness of Probabilistic Constraint Sets // *J. Optim. Theory Appl.* 2002. Vol. 112. No. 3. P. 657–663.
53. *Ivanov, S., Naumov, A.* On stochastic linear programming problems with the quantile criterion. // *Autom. Remote Control*. 72, 353–369 (2011).
54. *Ivanov, S., Naumov, A.* Algorithm to optimize the quantile criterion for the polyhedral loss function and discrete distribution of random parameters. // *Autom. Remote Control*. 73, 105–117 (2012).
55. *Kall P.* *Stochastic Linear Programming*. // Berlin: Springer-Verlag, 1976.
56. *Kall P., Wallace S.* *Stochastic Programming* // Chichester, Wiley, 1994.
57. *Kan Yu., Kibzun A.* *Stochastic Programming Problems with Probability and Quantile Functions*. Chichester, Wiley, 1996.
58. *Kataoka S.* Stochastic Programming Model // *Econometrica*. 1963. V. 31. No. 1-2. pp. 181-196.
59. *Kibzun A.I., Kurbakovskiy V.Yu.* Guaranteeing Approach to Solving Quantile Optimization Problem // *Annals of Operation Research*, 1991. Vol. 30. P. 81–93.
60. *Lejeune M.A.* Pattern definition of the p-efficiency concept // *Ann. Oper. Res.* 2012. No. 200 P. 23–36.
61. *Lejeune M.A.* Pattern-based modeling and solution of probabilistically constrained optimization problems // *Oper. Res.* 2012. No. 60. P. 1356–1372.
62. *Lepp R.* Minimization of Smooth Function for Probabilistic Constraints // *Izv. Eston. Akad. Nauk, Phys. Math.* – 1980. – T. 29. – №. 2. – C. 140-144.

63. *Mafusalov A., Uryasev S.* Buffered Probability of Exceedance: Mathematical Properties and Optimization // SIAM Journal on Optimization. 2018. vol. 28. no. 2. pp. 1077–1103.
64. *Marti K.* Stochastic Optimization Methods in Structural Mechanics. // ZAMM: Applied Mathematics and Mechanics, 1990, 70, pp.742–745.
65. *Marti K.* Approximations and Derivatives of Probability Functions. // In: Approximation, Probability and Related Fields, eds. G.Anastassiou and S.T.Rachev, Plenum Press, New York, 1994.
66. *Miller B.L., Wagner H.M.* Chance constrained programming with joint constraints // Oper. Res. 1965. No. 13. P. 930–945.
67. *Parzen E.* On Estimation of a Probability Density Function and Mode. The Annals of Mathematical Statistics, vol. 33, no. 3, 1962, pp. 1065–76.
68. *Pflug G.C., Weisshaupt H.* Probability gradient estimation by set-valued calculus and applications in network design // SIAM Journal on Optimization. 2005. vol. 15. pp. 898–914.
69. *Prekopa A.* Stochastic Programming. Springer Netherlands, 1995. 600 p.
70. *Prekopa A.* Logarithmic Concave Measures with Application to Stochastic Programming // Acta Sci. Math. (Szeged). 1971. Vol. 32. P. 301–316.
71. *Prekopa A.* On Logarithmic Concave Measures and Functions // Acta Sci. Math. (Szeged). 1973. Vol. 34. P. 335–343.
72. *Prekopa A.* Logarithmic Concave Measures and Related Topics. In: Stochastic Programming, ed. M.A.H.Dempster. – London: Academic Press, P.63–82.
73. *Prekopa A.* Numerical Solution of Probabilistic Constrained Programming Problems. In: Numerical Techniques for Stochastic Optimization, eds. Yu.Ermoliev and R.J.B.Wets, Springer-Verlag, Berlin, 1980, P.123–139.
74. *Rosenblatt M.* Remarks on Some Nonparametric Estimates of a Density Function. The Annals of Mathematical Statistics, vol. 27, no. 3, 1956, pp. 832–37.
75. *Sobol V.R., Torishnyy R.O., Pokhvalenskaya A.M.* Application of the smooth approximation of the probability function in some applied stochastic programming problems // Вестн. ЮУрГУ. Сер. Матем. моделирование и программирование. 14:3. 2021. С.33–45.

76. *Sobol V., Torishnyi R.* Application of Smooth Approximation in Stochastic Optimization Problems with a Polyhedral Loss Function and Probability Criterion // In: Strekalovsky A., Kochetov Y., Gruzdeva T., Orlov A. (eds) *Mathematical Optimization Theory and Operations Research: Recent Trends. MOTOR 2021. Communications in Computer and Information Science. vol 1476.* Springer, Cham. DOI:10.1007/978-3-030-86433-0_7
77. *Sobol V., Torishnyi R.* Smooth approximation of probability and quantile functions: vector generalization and its applications // *J. Phys.: Conf. Ser.* 1925 012034. 2021. pp 1-10.
78. *Sobol V.R., Torishnyy R.O.* Smooth Approximation of the Quantile Function Derivatives // *Вестн. ЮУрГУ. Сер. Матем. моделирование и программирование.* 15:4. 2022. С. 115–122.
79. *Sengupta J.K.* *Stochastic Programming Methods and Applications* // North-Holland, Amsterdam, 1972.
80. *Tamm E.* On a probability function optimization, *Izvestia Akademii Nauk Estonskoy SSR. Physics and Mathematics* 28 (1979), no. 1, 17-24.
81. *Tamm E.* On Minimization of a Function under an Equality Chance Constraint. // *Math. Operationsforsch. Statist., Ser. Optimization*, 1981, 12, No.2, P. 253–262.
82. *Uryasev S.* Differentiability of an Integral over a Set Defined by Inclusion. // *Cybernetics*, 1988, v.24, 5, P. 638–642.
83. *Uryasev S.* A Differentiation Formula for Integrals over Sets Given by Inclusion. // *Numerical Functional Analysis and Optimization*, 1989, 10(7 & 8), P. 827–841.
84. *Uryasev S.* Derivatives of Probability Functions and Integrals over Sets Given by Inequalities. // *J. Computational and Applied Mathematics*, 1995, 56. no. 1. pp. 197–223.
85. *Uryasev S.* Derivatives of probability functions and some applications. // *Annals of Operations Research*. 1995, Vol. 56. № 1. pp. 287–311.
86. *Vajda J.* *Probabilistic Programming.* // Acad. Press, New York, London, 1972.
87. *Van Ackooij W., Henrion R.* (Sub-)Gradient Formulae for Probability Functions of Random Inequality Systems under Gaussian Distribution // *SIAM/ASA Journal on Uncertainty Quantification*. 2017. vol. 5. no. 1. pp. 63–87.
88. *Van Ackooij W.* Eventual Convexity of Chance Constrained Feasible Sets // *J. Math. Programm. Oper. Res.* 2015. Vol. 64. No. 5. P. 1263-1284.

89. *Van Ackooij W., Sagastiz'abal C.* Constrained bundle methods for upper inexact oracles with application to joint chance constrained energy problems. *SIAM Journal on Optimization*, 2014, no. 24(2), P. 733–765.
90. *Van Ackooij W., Sagastiz'abal C., Oliveira W., Berge V.* Probabilistic optimization via approximate p-efficient points and bundle methods, *Computers & Operations Research*, Vol. 77, 2017, P. 177-193.
91. *Zel'terman D., Yu C.* A general approximation to quantiles. // *Communications in Statistics-Theory and Methods*. 2017. vol. 46(19). pp. 9834–9841.
92. *Zhang T., Uryasev S., Guan Y.* Derivatives and subderivatives of buffered probability of exceedance // *Operations Research Letters*. 2019. vol. 47. no. 2. pp 130–132.