Московский государственный университет
имени М.В.Ломоносова
Механико-математический факультет

На правах рукописи

Букин Дмитрий Борисович

ЗАДАЧИ МОНЖА И КАНТОРОВИЧА В БЕСКОНЕЧНОМЕРНЫХ ПРОСТРАНСТВАХ

01.01.01 — вещественный, комплексный и функциональный анализ

ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата физико-математических наук

Научный руководитель
доктор физико-математических
наук, профессор В.И. Богачев

Москва, 2020

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Актуальность темы. Тематика диссертации находится на стыке функционального анализа, теории меры, стохастического анализа и теории экстремальных задач. Рассмотренные вопросы представляют интерес также и для уравнений с частными производными и дифференциальной геометрии (см. обзор В.И. Богачева и А.В. Колесникова [9]). Главными объектами исследования данной работы выступают функционалы в задачах Монжа и Канторовича оптимальной транспортировки мер, а также метрики типа Канторовича–Рубинштейна. Основные результаты работы относятся к теории меры и теории экстремальных задач, они могут быть полезны и для теории уравнений Фоккера–Планка–Колмогорова [10].

Задача Монжа возникла еще в XVIII веке и в первоначальной постановке заключалась в нахождении оптимальных путей переноса масс (скажем, грунта) с минимизацией произведенной работы. При этом считалось априори ясным, что оптимальная транспортировка есть, вопрос был в ее описании и исследовании различных свойств. Лишь почти через два столетия появилась точная постановка этой задачи, состоящая в следующем. Для вероятностных мер μ и ν на $\mathbb{R}^n$, заданных посредством плотностей относительно меры Лебега, надо найти борелевское отображение T пространства $\mathbb{R}^n$, которое переводит μ в ν, т. е. $\nu(B) = \mu(T^{-1}(B))$ для всех борелевских множеств B, и минимизирует интеграл

$$\int |x - T(x)|\, \mu(dx)$$

среди всевозможных отображений, переводящих μ в ν. Однако доказательство существования оптимального отображения оказалось удивительно сложным и было получено уже в XXI веке, спустя более двадцати лет после появления первого математического доказательства В.Н. Судакова [22] в 1970-х, в котором позже были найдены существенные пробелы. В настоящее время известно несколько доказательств, одно из них следует методу Судакова и обходит выявленные в его работе пробелы (это связано с тем, что одно промежуточное техническое утверждение Судакова оказалось неверным), но все эти доказательства чрезвычайно длинны. Подробные комментарии можно найти в [9], [5]. Также в 1970-х годах, начиная с А.М. Вершика [12] (см. также [13]), стала рассматриваться более общая задача минимизации интеграла от функции $c(x, T(x))$ для так называемой функции стоимости c на произведении множеств, на которых заданы меры μ и ν (скажем, метрики на произведении двух общих метрических пространств). В общей задаче Монжа в современной трактовке исследуются отображения T одного пространства с мерой (X, μ) в другое (Y, ν), которые переводят заданную меру μ в заданную меру ν. Эти меры называются маргинальными распределениями или маргиналами; обычно они считаются вероятностными. При этом на $X \times Y$ задана неотрицательная функция c, называемая функцией стоимости. Задача состоит в минимизации интеграла

$$\int_X c(x, T(x))\, \mu(dx)$$

по отображениям T, переводящим μ в ν. Значение указанного интеграла называют стоимостью транспортировки T.

Еще до всех этих событий второй половины XX века и начала XXI века в конце 1930-х – начале 1940-х годов Л.В. Канторович, тогда даже не знавший о задаче Монжа, поставил свою задачу оптимизации, вызвавшую весьма значительный поток исследований, не иссякающий и по сей день. В задаче Канторовича тоже даны вероятностные меры μ и ν на измеримых пространствах X и Y и измеримая функция стоимости c на $X \times Y$, но найти надо вероятностную меру π на $X \times Y$, проекции которой на X и Y есть μ и ν и которая минимизирует интеграл от функции стоимости c по всем мерам с данными проекциями. Меры с заданными проекциями называются планами Канторовича или транспортными планами. Минимизирующая мера (если она есть) называется оптимальным планом или решением задачи Канторовича; соответствующее минимальное значение интеграла называют стоимостью оптимальной транспортировки. Связь двух задач состоит в том, что всякое отображение T меры μ в меру ν приводит к мере σ на $X \times Y$ с проекциями μ и ν: в качестве такой меры берется образ μ при отображении $x \mapsto (x, T(x))$. Сам Л.В. Канторович рассмотрел случай, когда X и Y — метрические компакты и c — соответствующая метрика на их произведении, но вскоре стало понятно, что его задача имеет смысл в гораздо более широкой постановке. Узнав после выхода своей заметки о работах Г. Монжа, Л.В. Канторович даже указал, что решение его задачи позволяет решить и задачу Монжа, но затем выяснилось, что это не совсем так. Оказалось, что задача Канторовича разрешима при значительно более широких предположениях, чем задача Монжа. В частности, с ней не возникло никаких технических проблем типа тех, с которыми имели дело В.Н. Судаков и его последователи. Например, для существования решения в задаче Канторовича достаточно полунепрерывности снизу функции стоимости на вполне регулярных топологических пространствах с мерами Радона, в то время как задача Монжа может не иметь решения даже для очень хороших функций стоимости на плоскости. Однако глубокая связь между задачами Монжа и Канторовича есть, одно из ее проявлений состоит в том, что для непрерывной функции стоимости и радоновских мер μ и ν без атомов значения инфимумов в обеих задачах равны при условии сепарабельности обеих мер (что автоматически имеет место для многих пространств в приложениях, например для суслинских пространств), см. [5], а также [20] и [44]; в [5] указано, что равенство остается в силе и для так называемых виртуально непрерывных функций стоимости, введенных А.М. Вершиком, П.Б. Затицким, Ф.В. Петровым [14], [15]. По обеим задачам имеется обширная литература, ими занимались и занимаются многие известные математики, опубликовано множество обзоров и немало монографий, см. [23], [9], [36], [19], [21], [45], [51], [52]. В последнее десятилетие в этих исследованиях значительное внимание уделялось сингулярным функциям стоимости, не являющимся непрерывными и даже бесконечным на значительной части пространства, см. [24], [25], [26], [34], [35]. Часть основных результатов диссертации также лежит в этом русле, а другая часть относится, наоборот, к весьма регулярной ситуации, когда отображаются гауссовские меры, причем линейными операторами, но задача состоит в сравнении стоимостей транспортировок для двух важных классов отображений: собственно оптимальных и так называемых канонических треугольных.

Работу можно разделить на две основные части, соответствующие двум основным главам (помимо вспомогательной первой главы), причем эти части связаны как идейно, так и технически.

Во второй главе исследуются задачи Монжа и Канторовича об оптимальной транспортировке на пространстве функций, непрерывных на отрезке, для распределений диффузионных процессов. При этом рассматривается функция стоимости, заданная нормой пространства Камерона–Мартина меры Винера.

Вопросы, связанные с задачей об оптимальной транспортировке меры Винера с функцией стоимости, заданной нормой пространства Камерона–Мартина, как в постановке Монжа, так и в постановке Канторовича, изучались в работах многих авторов, в том числе Д. Фейеля, А.С. Устюнеля [34], [35], впервые поставивших такой вопрос, а также Ш. Фанга, Ж. Шао, К.-Т. Штурма [33], Ф. Кавалетти [31], В.И. Богачева и А.В. Колесникова [1], [3], [27], [8], [29]. В работе [9] приведена обширная библиография по современным исследованиям данных вопросов. Значительной особенностью такой функции стоимости является то, что пространство Камерона–Мартина имеет меру нуль относительно меры Винера, а для многих мер на произведении (включая квадрат меры Винера) функция стоимости почти всюду бесконечна. Это делает задачу абсолютно не похожей на случай, когда используется обычная норма пространства непрерывных на отрезке функций, и значительно более сложной, так как рассматриваемая функция стоимости конечна только на очень малой части всего пространства, так что класс планов транспортировки, для которых функционал стоимости конечен, очень узок. Выбор этой функции стоимости обусловлен тем, что мера, полученная сдвигом меры Винера на вектор из пространства Камерона–Мартина, эквивалентна мере Винера. Более того, как показано в [8], всякая вероятностная мера, абсолютно непрерывная относительно меры Винера, может быть представлена в виде указанного сдвига. В работе [31] было доказано, что для меры Винера существует оптимальное отображение такого вида, несмотря на указанную сингулярность функции стоимости. Аналогичный результат справедлив и для мер, абсолютно непрерывных относительно меры Винера.

Таким образом, возникает вопрос об обобщении этого результата на меры, являющиеся распределениями диффузионных процессов. Если коэффициент диффузии постоянный, то распределение диффузионного процесса будет абсолютно непрерывно относительно меры Винера, а задача Монжа в рассматриваемой постановке окажется разрешимой. Нетрудно построить пример, показывающий, что в случае постоянства коэффициента диффузии на некотором отрезке задача Монжа также разрешима: эта ситуация снова сводится к рассмотрению мер, абсолютно непрерывных относительно меры Винера.

Основной результат второй главы состоит в том, что для широкого класса распределений диффузионных процессов с непостоянным коэффициентом диффузии в задаче Монжа с функцией стоимости, заданной нормой пространства Камерона–Мартина, отсутствуют решения в случае несовпадающих исходных мер.

В третьей главе основным объектом исследования выступают некоторые специальные классы отображений пространства $\mathbb{R}^n$ и преобразования гауссовских мер при этих отображениях, а задача состоит в сравнении стоимостей соответствующих транспортировок. Кроме того, эти вопросы обсуждаются и в бесконечномерном случае. Особое внимание уделяется широко используемым классам треугольных отображений и оптимальных отображений, в частности проводится сравнение значений функционала стоимости в задаче Канторовича для этих классов отображений.

Метрика Канторовича–Рубинштейна впервые была введена Л.В. Канторовичем в работе [17]; вопросы, связанные с ней рассматриваются в книгах В.И. Богачева [1], [2] и работах А. Такацу [48], [49], М. Ловри, М. Мин-Оо, Е.А. Рух [41], К. Модин [43] и Л.Т. Сковгаард [47] (последние три работы посвящены римановой геометрии гауссовских распределений). Вопросы, связанные с оценкой метрики Канторовича–Рубинштейна и с неравенствами для гауссовских мер, изучались многими авторами, среди которых М. Талагран (см. [50]), М. Леду (см. [40]), В.И. Богачев и А.В. Колесников (см. [6], [18], [39]). Известные неравенства, полученные Талаграном в работе [50], являются основной предпосылкой к исследованию связи между оптимальными и треугольными отображениями. М. Талагран оценивает функционал стоимости как для оптимального, так и для треугольного отображений, переводящих стандартную гауссовскую меру γ в меру μ, абсолютно непрерывную относительно γ, с помощью энтропии меры μ относительно меры γ. Если эта энтропия достаточно мала, то логично было бы предположить, что значения функционала стоимости сравнимы для оптимального и для треугольного отображений, переводящих γ в μ. В третьей главе показано, что класс мер, для которых эти два значения сравнимы, весьма узок, но все же и в бесконечномерном случае имеются конструктивные условия, при которых стоимость транспортировки при треугольном отображении оценивается с универсальной постоянной через оптимальную стоимость.

Цель работы.

• Исследовать задачу Монжа об оптимальном отображении распределения диффузионного процесса μ на пространстве непрерывных функций в абсолютно непрерывную относительно μ меру с функцией стоимости, заданной нормой пространства Камерона–Мартина меры Винера.

• Исследовать задачу Канторовича для мер μ и ν, где μ — распределение диффузионного процесса в пространстве непрерывных функций, а мера ν абсолютно непрерывна относительно μ.

• Исследовать связь значений функционала стоимости в задаче Монжа для треугольных и для оптимальных отображений гауссовских мер. Описать класс мер, для которых эти значения сравнимы.

Научная новизна. Все результаты диссертации являются новыми и состоят в следующем.

1. Для распределений диффузионных процессов в пространстве непрерывных функций получены условия на коэффициент диффузии, необходимые для того, чтобы соответствующая задача Монжа с функцией стоимости, заданной нормой пространства Камерона–Мартина меры Винера, была разрешима. Это дает простые и легко проверяемые условия, при которых указанная задача неразрешима.

2. Для задачи Канторовича с функцией стоимости, заданной нормой пространства Камерона–Мартина меры Винера, показано, что она не имеет нетривиальных решений для широкого класса распределений диффузионных процессов. В частности, так обстоит дело для непостоянных аналитических коэффициентов диффузии.

3. Для гауссовских мер на пространстве $\mathbb{R}^n$ проведено сравнение функционалов стоимости в задаче Монжа для треугольных и для оптимальных отображений, показано, что эти значения асимптотически не сравнимы для весьма широкого класса гауссовских мер. Получены бесконечномерные аналоги этого утверждения. С другой стороны, указаны конструктивные условия, при которых даже в бесконечномерном случае стоимость транспортировки гауссовских мер при треугольном отображении оценивается с универсальной постоянной через оптимальную стоимость.

Положения, выносимые на защиту.

1. Конструктивные условия на коэффициент диффузии, необходимые для разрешимости соответствующей задачи Монжа для распределений диффузионных процессов в пространстве непрерывных функций с функцией стоимости, заданной нормой пространства Камерона–Мартина меры Винера.

2. Задача Канторовича с функцией стоимости, заданной нормой пространства Камерона–Мартина меры Винера, не имеет нетривиальных решений для широкого класса распределений диффузионных процессов.

3. Значения функционалов стоимости в задаче Монжа для треугольных

и для оптимальных отображений гауссовских мер на $\mathbb{R}^n$ асимптотически не сравнимы для широкого класса мер. Условия, при которых в бесконечномерном случае стоимость транспортировки при треугольном отображении оценивается с универсальной постоянной через оптимальную стоимость.

Методы исследования. В работе используются методы общей теории меры и нелинейного функционального анализа, а также некоторые конструкции автора.

Теоретическая и практическая ценность. Результаты диссертации носят теоретический характер и могут применяться в различных вопросах функционального анализа, в том числе бесконечномерного анализа, теории меры, теории экстремальных задач и стохастического анализа. Ее результаты и методы будут востребованы в исследованиях, проводимых в Московском государственном университете имени М.В. Ломоносова, Математическом институте имени В.А. Стеклова РАН, Институте проблем передачи информации имени А.А. Харкевича РАН, Санкт-Петербургском государственном университете, Новосибирском государственном университете и Национальном исследовательском университете «Высшая школа экономики».

Апробация диссертации.

Результаты диссертации докладывались автором на следующих научных конференциях.

1. Международная конференция «Бесконечномерный анализ, стохастика, математическое моделирование: новые задачи и методы. Проблемы математического и естественнонаучного образования» (Москва, РУДН, 2014 г.)

2. Международная конференция студентов, аспирантов и молодых ученых «Ломоносов» (Москва, МГУ, 2016 г.)

3. Международная конференция «Бесконечномерный анализ и теория управления» (Москва, МГУ, 2018 г.)

4. Международная конференция «Recent Advances in Mass Transporta-

tion». Poncelet Center and the Higher School of Economics (Москва, 2019 г.).

По теме диссертации были сделаны доклады на следующих научно-исследовательских семинарах.

1. Научно-исследовательский семинар «Бесконечномерный анализ и стохастика» под руководством В.И. Богачева, С.В. Шапошникова и Н.А. Толмачева (МГУ, многократно, 2013–2019 г.),

2. Международный научно-исследовательский семинар “Infinite-dimensional stochastic analysis” в университете г. Билефельда, Германия (2015 г.)

Публикации. Основные результаты диссертации опубликованы в 4 работах автора (см. [53], [54], [55], [56], последние две из которых в соавторстве) в рецензируемых научных журналах из списка ВАК, входящих в базы данных SCOPUS и Web of Science, а также представлены в тезисах 3 международных конференций (см. [57]–[59]).

Структура и объем работы. Диссертация состоит из введения, трех глав, заключения и списка литературы из 59 наименований. Общий объем диссертации составляет 61 страницу.

Краткое содержание диссертации. Нумерация приводимых здесь результатов соответствует нумерации в основном тексте диссертации.

Первая глава данной работы является вводной и содержит все необходимые определения, обозначения и уже известные результаты, которые используются в остальных главах работы. Проблемы, обсуждаемые во второй и третьей главе, имеют бесконечномерный характер.

Во второй главе исследуются задачи Монжа и Канторовича с функцией стоимости, заданной нормой пространства Камерона–Мартина меры Винера. Эти задачи рассмотрены для распределений диффузионных процессов в пространстве непрерывных функций на отрезке $[0, 1]$.

Мера μ — образ меры Винера P_W на пространстве $X = C[0, 1]$ при отображении $F(x(\cdot)) = f \circ x(\cdot)$, где f — диффеоморфизм вещественной прямой, π — произвольная борелевская вероятностная мера на $X \times X$ с заданными проекциями μ и ν на первый и второй сомножители.

Теорема 2.1.1. *Пусть мера ν абсолютно непрерывна относительно меры μ. Тогда для π-почти всех $(x, y) \in X \times X$ верно равенство*

$$f'[f^{-1}(x(t))] = f'[f^{-1}(y(t))], \quad t \in [0,1].$$

Смысл этого утверждения состоит в том, что указанное равенство накладывает весьма сильные ограничения на меру π, от которой даже не требуется никакой оптимальности. Например, в конечномерном случае в качестве π можно взять произведение мер μ и ν, здесь же произведение не подходит. Найденное условие позволяет получить ограничения на функцию f, чтобы задачи Монжа и Канторовича с функцией стоимости, равной норме Камерона–Мартина, имели решения с маргиналами μ и ν.

Теорема 2.2.1. *Пусть $f \in C^2(\mathbb{R})$ — возрастающий диффеоморфизм прямой, причем нули функции f'' изолированы. Предположим, что борелевская вероятностная мера π на $X \times X$, удовлетворяющая условию*

$$x - y \in H \quad \text{для } \pi\text{-почти всех} \quad (x, y) \in X \times X,$$

имеет проекции μ и ν такие, что ν абсолютно непрерывна относительно μ. Тогда $x = y$ для π-почти всех $(x, y) \in X \times X$ и $\mu = \nu$.

Следствие 2.2.2. *Пусть f — вещественно-аналитический, возрастающий, нелинейный диффеоморфизм прямой. Тогда мера π, удовлетворяющая условиям теоремы* 2.1.1*, сосредоточена на множестве*

$$\{(x, y) \colon x = y\}$$

и $\mu = \nu$.

Теорема 2.3.1. *Пусть $f \in C^2(\mathbb{R})$ — возрастающий диффеоморфизм вещественной прямой, причем множество нулей функции f'' имеет лебеговскую меру нуль. Предположим, что борелевская вероятностная мера π на $X \times X$ удовлетворяет условию*

$$x - y \in H \quad \text{для } \pi\text{-почти всех} \quad (x, y) \in X \times X,$$

причем имеет проекции μ и ν такие, что мера ν абсолютно непрерывна относительно μ. Тогда $x = y$ для π-почти всех $(x, y) \in X \times X$ и $\mu = \nu$.

Последние два утверждения показывают, что при условии $\mu \neq \nu$ не существует меры на $X \times X$ с проекциями μ и ν, относительно которой интегрируема функция $|x-y|_H$. Тем самым задачи Монжа и Канторовича не имеют решений.

В **третьей главе** проводится сравнение значений функционала стоимости в задаче Монжа для треугольных и оптимальных отображений гауссовских мер. Каноническим треугольным отображением борелевской вероятностной меры μ на $\mathbb{R}^n$ в борелевскую вероятностную меру ν называется такое борелевское отображение $T = (T_1, \ldots, T_n)$ пространства $\mathbb{R}^n$, переводящее μ в ν, что компонента T_1 есть функция только первой координаты, т. е. имеет вид $T_1(x_1)$, компонента T_2 имеет вид $T_2(x_1, x_2)$, компонента с номером k имеет вид $T_k(x_1, \ldots, x_k)$, причем функции

$$x_k \mapsto T_k(x_1, \ldots, x_k)$$

являются возрастающими при фиксированных значениях оставшихся переменных. В случае линейного отображения его матрица имеет треугольный вид. Известно (см. [2], [7], [8]), что такое отображение существует и единственно с точностью до переопределения на множестве меры нуль для всех пар абсолютно непрерывных мер. В одномерном случае каноническое отображение строится с помощью функций распределения и их обратных. В данной главе оценивается константа K в неравенстве

$$\int_{\mathbb{R}^n} |T(x)-x|^2\, \gamma(dx) \leqslant K \int_{\mathbb{R}^n} |T_0(x)-x|^2\, \gamma(dx),$$

где γ и μ — две центрированные гауссовские меры на $\mathbb{R}^n$, а T и T_0 — соответственно треугольное и оптимальное отображения, переводящие меру γ в меру μ.

Оказывается, что оба эти отображения в гауссовском случае линейны. Это вполне ожидаемое и кажущееся очевидным утверждение доказывается, как ни странно, довольно неэлементарно с привлечением глубоких результатов для общих оптимальных отображений.

Лемма 3.1.1. *Пусть γ и μ — две центрированные гауссовские меры на $\mathbb{R}^n$. Тогда оптимальное отображение T_0, переводящее меру γ в меру μ,* *линейно, а каноническое треугольное отображение T меры γ в меру μ также линейно.*

Несложные примеры показывают, что каноническое треугольное отображение гауссовских мер отнюдь не всегда оптимально. Это вполне естественно, так как выбор такого отображения тесно связан с выбором ортонормированного базиса, а оптимальное отображение не зависит от выбора ортонормированного базиса. Поэтому даже в том случае, когда оба отображения совпали, бывает несложно перейти к иному базису, где совпадения нет. Однако был открыт вопрос о том, сколь сильно могут отличаться стоимости транспортировок при этих двух отображениях. Этот вопрос естественно возник после того, как М. Талагран открыл свое замечательное неравенство, из которого вытекало, что обе стоимости при широких условиях допускают оценку через энтропию одного из маргиналов. Поэтому можно было рассчитывать, что они имеют одинаковый порядок. Однако, как показано в диссертации, такие ожидания не оправдываются, хотя при определенных дополнительных условиях сравнение все же оказывается возможным, причем в бесконечномерном случае.

Теорема 3.2.1. *Пусть γ — стандартная гауссовская мера на $\mathbb{R}^n$ и M — симметричная положительно определенная матрица с собственными значениями μ_i и определителем, равным единице. Тогда в некотором ортонормированном базисе пространства $\mathbb{R}^n$ центрированная гауссовская мера μ с ковариационной матрицей M обладает следующим свойством: для оптимального линейного отображения T_0, переводящего меру γ в меру μ (которое задается матрицей $\sqrt{M}$ в этом базисе) и для канонического треугольного отображения T меры γ в меру μ (в этом же базисе) справедлива оценка*

$$K \geqslant 1 + 2\frac{\sum_{i=1}^{n}(\lambda_i - 1)}{\sum_{i=1}^{n}(\lambda_i - 1)^2},$$

где λ_i — собственные числа матрицы A отображения T_0 в стандартном базисе, т. е. $\lambda_i = \sqrt{\mu_i}$.

Теорема 3.3.1. *Пусть γ — стандартная гауссовская мера на $\mathbb{R}^n$. Бу-*

дем брать в качестве μ всевозможные невырожденные центрированные гауссовские меры на $\mathbb{R}^n$. Если T — треугольное, а T_0 — оптимальное отображение, переводящее меру γ в меру μ, то для наименьшей возможной постоянной K в неравенстве

$$\int_{\mathbb{R}^n} |T(x)-x|^2\,\gamma(dx) \leqslant K\int_{\mathbb{R}^n} |T_0(x)-x|^2\,\gamma(dx)$$

верна оценка

$$K \geqslant n+\sqrt{n^2-n}.$$

В частности, если рассматривать меры μ с ковариационной матрицей, имеющей единичный определитель, то коэффициент K не может быть меньше $\sqrt{n}$.

Пусть мера γ на пространстве $\mathbb{R}^\infty$ всех последовательностей есть счетная степень стандартной гауссовской меры на прямой, тогда $H = l^2$ — ее пространство Камерона–Мартина. Рассмотрим измеримый линейный оператор B следующего вида:

$$Bx=\sum_{i=1}^{\infty} x_i B_0 e_i,$$

где $B_0\colon H \to H$ — линейный оператор, у которого норма Гильберта–Шмидта меньше 1.

Следствие 3.3.2. *Пусть норма Гильберта–Шмидта оператора B_0 не превосходит $1/2$. Тогда мера $\mu = \gamma\circ(I+B)^{-1}$ эквивалентна γ, имеет конечную энтропию и существует измеримое линейное треугольное отображение T, переводящее меру γ в меру μ, причем для оптимального отображения T_0 верна оценка*

$$\int |Tx-x|^2\,\gamma(dx) \leqslant 25\int |T_0x-x|^2\,\gamma(dx),$$

а также верны энтропийные неравенства Талаграна для интегралов от $|Tx-x|^2$ и $|T_0x-x|^2$.

Смысл этого утверждения состоит в том, что при указанных условиях каноническое треугольное отображение одной гауссовской меры в другую приводит к стоимости транспортировки, оцениваемой в фиксированное число раз через стоимость оптимальной транспортировки. Это отличает данную более специальную ситуацию от общей, в которой нет не зависящего от размерности коэффициента, позволяющего сравнить заведомо неоптимальную в типичных случаях стоимость с оптимальной.

Автор глубоко благодарен своему научному руководителю доктору физико-математических наук, профессору Владимиру Игоревичу Богачеву за постановку задач, их регулярное обсуждение и помощь в работе.

Глава 1

Определения, обозначения и вспомогательные сведения

В этой главе даются основные определения и обозначения, используемые в работе, а также приводятся известные результаты, применяемые при доказательствах основных теорем работы.

1.1 Определения и обозначения

Основным объектом, рассматриваемым в данной работе, являются преобразования вероятностных мер. Рассмотрим две вероятностные меры μ и ν на некоторых измеримых пространствах $(X, \mathcal{A})$ и $(Y, \mathcal{B})$ соответственно ($\mathcal{A}$ и $\mathcal{B}$ — σ-алгебры измеримых множеств на пространствах X и Y соответственно).

Напомним, что образом меры μ на пространстве X под действием измеримого отображения $T\colon X \to Y$ называется мера, задаваемая равенством

$$\mu \circ T^{-1}(A) = \mu\{x : T(x) \in A\} \quad A \in \mathcal{B}$$

В случае, если пространство X — линейное, через μ_h обозначим сдвиг меры μ на вектор $h \in X$, т. е. $\mu_h(A) = \mu(A - h)$ для всех $A \in \mathcal{B}$.

Пусть μ и ν — две вероятностные меры на некотором измеримом пространстве $(X, \mathcal{A})$. Мера μ называется абсолютно непрерывной относительно меры ν, если $\mu(A) = 0$ для всякого множества $A \in \mathcal{A}$ с $\nu(A) = 0$.

Обозначение: $\mu \ll \nu$. Меры μ и ν называются эквивалентными, если $\mu \ll \nu$ и $\nu \ll \mu$. Обозначение: $\mu \sim \nu$. Меры μ и ν называются сингулярными, если существует такое множество $\Omega \in \mathcal{A}$, что $\mu(\Omega) = 0$ и $\nu(X \setminus \Omega) = 0$. Обозначение: $\mu \perp \nu$.

Во второй главе мы будем рассматривать распределения диффузонных случайных процессов в пространстве траекторий. Напомним некоторые ключевые определения.

Случайный процесс $(\xi_t)_{t\in[0,1]}$ называется диффузионным, если он удовлетворяет стохастическому дифференциальному уравнению

$$d\xi_t = a(t, \xi_t)dt + B(t, \xi_t)dw_t,$$

где w_t — винеровский случайный процесс, ξ_t и w_t принимают значения в $\mathbb{R}$. Функции $a(t, \xi_t)dt$ и $B(t, \xi_t)$ определены на $[0, 1] \times C[0, 1]$, измеримы по совокупности переменных и для всех $t \in [0, 1]$ как функции от x_t измеримы относительно σ-алгебры, порожденной цилиндрическими множествами в $C[0, 1]$ с основаниями над $[0, t]$.

Пусть f — возрастающий диффеоморфизм прямой, тогда, по формуле Ито, процесс $\xi_t = f(w_t)$ удовлетворяет приведенному выше уравнению при $B(t, \xi_t) = f'(f^{-1}(\xi_t))$ и $a(t, \xi_t) = f''(f^{-1}(\xi_t))/2$. Поэтому, в большинстве случаев, распределение одномерной диффузии эквивалентно распределению процесса $f(w_t)$ для соответствующего f. Мера Винера P_W на $C[0, 1]$ задается как распределение винеровского процесса w_t на этом пространстве. Тогда распределение диффузионного процесса является образом P_W под действием отображения, заданного композицией каждой траектории с отображением f.

Напомним определение гауссовской меры в $\mathbb{R}^n$. Вероятностная мера γ на прямой называется гауссовской, если она либо является дираковской мерой в некоторой точке, либо имеет плотность вида

$$\left(2\pi\sigma^2\right)^{-1/2} \exp\left(-|x - a|^2/2\sigma^2\right)$$

для некоторых чисел a и σ. Гауссовская мера называется центрированной, если число a в формуле выше равно нулю. Вероятностная мера γ на $\mathbb{R}^n$

называется гауссовской, если для каждого линейного функционала l на $\mathbb{R}^n$ ее образ $\gamma \circ l^{-1}$ — гауссовская мера на прямой.

Известно, что (см. [1]) преобразование Фурье гауссовской меры на $\mathbb{R}^n$ имеет вид

$$\widetilde{\gamma}(y) = \exp\bigl(i(y,a) - 1/2(Ky,y)\bigr),$$

где a — некоторый вектор в $\mathbb{R}^n$ и K — неотрицательная матрица. Мера γ имеет плотность в том и только том случае, когда матрица K невырождена. При этом плотность меры γ имеет вид

$$(2\pi)^{-n/2}(\det K)^{-1/2}\exp\left\{-1/2\bigl(K^{-1}(x-a),(x-a)\bigr)\right\}.$$

Матрица K называется ковариационной матрицей меры γ. Стандартной гауссовской мерой на $\mathbb{R}^n$ будем называть меру, задаваемую плотностью $(2\pi)^{-n/2}\exp(-|x|^2/2)$, где $|\cdot|$ — стандартная евклидова норма.

Радоновская вероятностная мера γ на локально выпуклом пространстве X называется гауссовской, если для всякого $f \in X^*$ индуцированная мера $\gamma \circ f^{-1}$ — гауссовская на прямой (X^* — пространство всех линейных непрерывных функционалов на X).

Средним радоновской гауссовской меры γ называется такой вектор $m \in X$, что

$$f(m) = \int_X f(x)\gamma(dx)$$

для всякого $f \in X^*$. Если $m = 0$, то мера γ называется центрированной. Всякая радоновская гауссовская мера γ есть сдвиг центрированной гауссовской меры $\gamma_m(B) = \gamma(B+m)$.

Для центрированной радоновской гауссовской меры γ обозначим через X^*_γ замыкание X^* в $L^2(\gamma)$.

Существует оператор $R_\gamma \colon X^*_\gamma \mapsto X$, называемый ковариационным оператором меры γ, такой, что

$$f(R_\gamma g) = \int_X f(x)g(x)\gamma(dx)$$

для всяких $f \in X^*$, $g \in X^*_\gamma$. Положим $g = \hat{h}$, если $h = R_\gamma g$. Тогда $\hat{h}$ называется γ-измеримым линейным функционалом, порожденным h.

Пространство $H(\gamma) = R_\gamma(X_\gamma^*)$ называется пространством Камерона–Мартина меры γ. Оно оказывается гильбертовым относительно скалярного произведения

$$(h,k)_H = \int_X \hat{h}(x)\hat{k}(x)\gamma(dx)$$

с соответствующей нормой, заданной формулой

$$|h|_H = \|\hat{h}\|_{L^2(\gamma)}.$$

В случае, когда мера μ абсолютно непрерывна относительно меры γ, энтропия $\mathrm{Ent}_\gamma(\mu)$ вводится следующим образом:

$$\mathrm{Ent}_\gamma(\mu) = \mathrm{Ent}_\gamma(\rho) = \int_X \rho \ln \rho\, \gamma(dx),$$

где $\rho = \frac{d\mu}{d\gamma}$ — плотность меры μ относительно γ.

Отображение $T = (T_1, \ldots, T_n) : \mathbb{R}^n \to \mathbb{R}^n$ называется треугольным, если T_1 есть функция x_1, T_2 — функция (x_1, x_2), T_3 — функция (x_1, x_2, x_3) и т.д.: T_i есть функция от $(x_1, x_2, \ldots, x_i)$. Треугольное отображение называется возрастающим, если каждая его компонента T_i — возрастающая функция по переменной x_i; однако монотонности по другим переменным не требуется (см. [3]). Измеримое возрастающее треугольное отображение называют каноническим; при весьма широких условиях известны теоремы существования и единственности для таких (конечномерных) отображений (см. [2], [3], [7], [8]). Выбор термина “треугольное отображение” объясняется тем, что для таких дифференцируемых отображений матрица Якоби имеет треугольный вид. Несомненным достоинством треугольных отображений является их конструктивность, возможность получить явные (хотя и громоздкие) формулы.

Другой важный класс отображений возникает при рассмотрении задачи Канторовича об оптимальной транспортировке (см. [9], [51], [52]). Пусть X — измеримое пространство, на котором заданы меры γ и μ, а функция $c(x,y)$ неотрицательна и измерима на $X \times X$. Задача Канторовича состоит в минимизации функционала стоимости

$$K(\nu) = \int_{X\times X} c(x_1, x_2)\, \nu(dx_1 dx_2)$$

по всем мерам ν на $X\times X$ с проекциями γ и μ. В случае, когда X — метрическое пространство с метрикой d, расстояние Канторовича–Рубинштейна (см. [2], [4], [6]) между мерами γ и μ определяется с помощью функционала $K(\nu)$ с функцией стоимости вида $c(x,y)=d(x,y)^2$:

$$W_2(\gamma,\mu)=\inf\left[\int_{X\times X} d(x_1,x_2)^2\,\nu(dx_1dx_2)\right]^{1/2},$$

где инфимум берется по всем таким мерам ν, что γ и μ — проекции меры ν на первое и второе пространство соответственно.

Во многих частных случаях существует отображение (называемое оптимальным отображением мер) $T: X\to X$ такое, что $\mu=\gamma\circ T^{-1}$ и

$$W_2(\gamma,\mu)^2=\int_X d(x,T(x))^2\,\gamma(dx).$$

При широких условиях такие отображения существуют и единственны. В частности, для всяких двух вероятностных мер ρ_1dx и ρ_2dx на пространстве $X=\mathbb{R}^n$ существует оптимальное отображение T, переводящее ρ_1dx в ρ_2dx, причем оно ρ_1dx-единственно и имеет вид $T=\nabla\Psi$, где Ψ — выпуклая функция, удовлетворяющая уравнению Монжа–Ампера

$$\rho_2(\nabla\Psi)\det D^2\Psi=\rho_1.$$

Для нормы функции f в пространстве $L^p(\mu)$, равной

$$\left(\int |f|^p\,d\mu\right)^{1/p},$$

мы будем использовать обозначение $\|f\|_{L^p(\mu)}$ или $\|f\|_p$ для сокращения записи, если из контекста понятно, о какой мере идет речь.

Символами вида C, C_1, C_2 и c, c_1, c_2 обозначаются числовые абсолютные константы, а символами вида $C(d), C_1(d), C_2(d)$ и $c(d), c_1(d), c_2(d)$ обозначаются константы, зависящие только от числового параметра d.

1.2 Вспомогательные результаты

В этом параграфе мы приведем результаты, которыми будем пользоваться при доказательстве основных результатов работы.

Во второй главе нам понадобятся следующие теоремы.

Теорема 1.2.1. *Пространство H Камерона–Мартина меры Винера P_W совпадает с классом Соболева $W_0^{1,2}[0,1]$ таких функций f на $[0,1]$, что f абсолютно непрерывна, $f' \in L^2[0,1]$ и $f(0) = 0$. При этом*

$$|f|_H = \|f'\|_{L^2[0,1]}.$$

Доказательство этой теоремы можно найти в [1, с. 55].

Нам также понадобятся следующие две теоремы из теории случайных процессов. Первая из них приведена в третьем томе книги [16, с. 345].

Теорема 1.2.2. *Пусть ξ_t — диффузионный процесс, который является решением уравнения*

$$d\xi_t = a(t,\xi_t)dt + B(t,\xi_t)dw_t,$$

а функция $b(t,x_t)$ на $[0,1]\times C[0,1]$ со значениями из $\mathbb{R}$ является измеримой по совокупности переменных и для всех $t \in [0,1]$ как функция от x_t измерима относительно σ-алгебры, порожденной функционалами $s \mapsto x(s)$ с $[0,t]$.

Пусть плотность $\rho_1(x_t)$ определяется равенством

$$\rho_1(x_t) = \frac{d\mu}{d\mu_\xi}(x_t) = \exp\left\{\int_0^1 b(s,x_t)\,dz_s - \frac{1}{2}\int_0^1 b(s,x_t)^2\,ds\right\}.$$

В этом равенстве μ_ξ — распределение процесса ξ_t в $C[0,1]$, μ — некоторая мера, абсолютно непрерывная относительно μ_ξ, а z_t — случайный процесс, заданный формулой

$$z_t = \int_0^t B^{-1}(s,x_s)\,dy_s,$$

где

$$y_t = x_t - \int_0^t a(s,x_s)\,ds.$$

Тогда существует такое решение уравнения

$$d\eta_t = a_1(t,\eta_t)dt + B(t,\eta_t)dw_t,$$

где

$$a_1(t, x_t) = a(t, x_t) + B(t, x_t)b(t, x_t),$$

что мера μ_η, соответствующая решению η_t, абсолютно непрерывна относительно меры μ_ξ. При этом

$$\frac{d\mu_\eta}{d\mu_\xi}(x_t) = \rho_1(x_t).$$

Вторая теорема — классический закон повторного логарифма.

Теорема 1.2.3. *Пусть w_t — стандартный винеровский процесс, тогда*

$$P\left[\limsup_{t\to 0+} \frac{w_t}{(2t\ln\ln(1/t))^{1/2}} = 1\right] = 1.$$

Доказательство этого факта можно найти в [46, теорема 1.9].

Следующие теоремы применяются в третьей главе.

Теорема 1.2.4. *Пусть μ и ν — борелевские вероятностные меры на $\mathbb{R}^\infty$. Предположим, что возрастающие треугольные борелевские отображения $(T_n)_{n=1}^\infty$ и $S = (S_n)_{n=1}^\infty$ таковы, что $\mu\circ T^{-1} = \mu\circ S^{-1}$ и для каждого n отображение $(T_1, \ldots, T_n)$ инъективно на борелевском множестве полной меры относительно проекции μ на $\mathbb{R}^n$. Тогда $T(x) = S(x)$ для μ-почти всех x. В частности, если проекции мер μ и ν на пространства $\mathbb{R}^n$ абсолютно непрерывны, то существует каноническое треугольное отображение $T_{\mu,\nu}$, причем оно единственно с точностью до μ-эквивалентности в классе возрастающих борелевских треугольных отображений, переводящих μ в ν.*

Доказательство этой теоремы можно найти в [8, с. 9].

Теорема 1.2.5 ([42]). *Пусть μ и ν — вероятностные меры на $\mathbb{R}^n$, причем μ равна нулю на борелевских подмножествах $\mathbb{R}^n$ хаусдорфовой размерности не выше $n-1$. Тогда существует выпуклая функция ψ на $\mathbb{R}^n$ такая, что ее градиент $\nabla\psi$ переводит μ в ν. При этом $\nabla\psi$ определено единственным образом μ-почти всюду.*

Теорема 1.2.6 ([8]). *Пусть* μ *— стандартная гауссовская мера на* $\mathbb{R}^n$*, вероятностная мера* ν *на* $\mathbb{R}^n$ *абсолютно непрерывна относительно* μ*, причем для* $f = d\nu/d\mu$ *имеем* $f \ln f \in L^1(\mu)$*. Тогда для канонического треугольного отображения* $T_{\mu,\nu}$ *имеем*

$$\int_{\mathbb{R}^n} |x - T_{\mu,\nu}(x)|^2 d\mu \leqslant 2 \int_{\mathbb{R}^n} f(x) \ln f(x) d\mu.$$

Теорема 1.2.7 ([50]). *Пусть* μ *— стандартная гауссовская мера на* $\mathbb{R}^n$*, удовлетворяющая логарифмическому неравенству Соболева*

$$\mathrm{Ent}_\mu f^2 \leqslant 2 \int_{\mathbb{R}^n} |\nabla f|^2 \, d\mu,$$

где $f \in C_0^\infty(\mathbb{R}^n)$*. Тогда для всякой вероятностной меры* $g \cdot \mu$*, абсолютно непрерывной относительно* μ*, имеем*

$$W_2^2(\gamma, g \cdot \gamma) \leqslant 2 \int_{\mathbb{R}^n} g \ln g \, d\mu.$$

Глава 2

Задачи Монжа и Канторовича для распределений диффузионных процессов

Пусть $(X, \mathcal{B})$ — измеримое пространство с двумя заданными на нем вероятностными мерами μ и ν. Пусть $c(x, y)$ — неотрицательная измеримая функция на $X \times X$. Она называется функцией стоимости. Задача Монжа состоит в минимизации функционала стоимости

$$M(T) = \int_X c(x, T(x))\, \mu(dx)$$

по всем измеримым отображениям $T\colon X \to X$ таким, что $\mu \circ T^{-1} = \nu$. Если отображение T является решением задачи Монжа, то оно называется оптимальным отображением. В общем случае задача Монжа может не иметь решений. Если существует отображение T, переводящее меру μ в меру ν, такое, что $M(T)$ конечен, то можно, по крайней мере, рассматривать инфимум $M(T)$ по всем таким T.

Задача Канторовича (в некотором смысле, более общая) для функции стоимости c состоит в минимизации величины

$$K(\pi) = \int_{X \times X} c(x, y)\, \pi(dx, dy)$$

по всем вероятностным мерам π на $X \times X$ с проекциями μ и ν (такие меры называются планами транспортировки). Как и у задачи Монжа, в общем случае минимума может не существовать, однако задача Канторовича имеет решение при гораздо более общих условиях. Очевидно, что всякое преобразование T меры μ в меру ν порождает вероятностную меру π на $X \times X$ с проекциями μ и ν: достаточно рассмотреть образ μ при отображении $x \to (x, T(x))$. Поэтому инфимум $K(\pi)$ по всем планам транспортировки π не превосходит инфимума в задаче Монжа. В то же время бывает, что обе задачи разрешимы, но соответствующие минимумы различны. Функция стоимости c имеет довольно специальный вид (как правило, связанный с расстояниями), по крайней мере, она конечна.

Рассмотрим задачи Монжа и Канторовича на пространстве $X = C[0,1]$ с борелевской σ-алгеброй, соответствующей его естественной равномерной норме, с функциями стоимости $c(x,y) = |x-y|_H^2$ и $c(x,y) = |x-y|_H$, где

$$H = W_0^{2,1} = \{h\colon h \text{ абсолютно непрерывна на } [0,1], h' \in L^2[0,1], h(0) = 0\}$$

есть пространство Камерона–Мартина меры Винера P_W с нормой

$$|h|_H = \|h'\|_{L^2}$$

согласно теореме 1.2.1. Функционал стоимости принимает вид

$$M(T) = \int_X |T(x) - x|_H^2 \, \mu(dx)$$

для функции стоимости, заданной квадратом расстояния, или

$$M(T) = \int_X |T(x) - x|_H \, \mu(dx)$$

в случае $c(x,y) = |x-y|_H$.

В этой главе мы покажем, что при такой постановке эти задачи не имеют нетривиальных решений для весьма широкого класса диффузионных процессов. Более того, функционал стоимости оказывается конечен лишь для тождественного отображения T.

2.1 Некоторые свойства коэффициента диффузии

Известно (см. [28], [1, с. 309] или [3, с. 124]), что для нелинейного диффеоморфизма f распределение $\xi_t = f(w_t)$ может быть сингулярно с P_W, более того, во многих случаях, к примеру, для аналитической f с непостоянной f', сдвиг распределения μ процесса ξ_t на ненулевой вектор из $C[0,1]$ будет взаимно сингулярен с μ. Благодаря этому свойству возникает явление, изученное в этой главе.

С другой стороны, неверно, что T должно быть тождественно для всякого диффузионного процесса с непостоянным коэффициентом диффузии. К примеру, если $f(x) = x$ для всех $x \in [-1,1]$, но $f(x) \neq x$ во всех остальных точках, то распределение μ диффузии $\xi_t = f(w_t)$ совпадает с мерой Винера на единичном шаре U в пространстве $C[0,1]$. Если теперь в качестве ν взять вероятностную меру, которая совпадает с μ на дополнении к U, а на самом U эквивалентна ограничению меры Винера (и, следовательно, ограничению μ), но не совпадает с μ, то мы получаем, что существуют отображения $T = I + b$, где b принимает значения в H и тождественно равна нулю на дополнении к U, такие что $\nu = \mu \circ T^{-1}$ и $M(T) < \infty$ (на самом деле соответствующая задача Монжа разрешима). Таким образом, в отличие от описанной выше ситуации, когда ненулевые эквивалентные сдвиги отсутствуют, в общем случае понадобятся некоторые дополнительные ограничения на f, чтобы исключить нетривиальные решения задачи Монжа.

Отсутствие решений у задачи Монжа в случае $\nu \neq \mu$ будет следовать из более общего результата об отсутствии планов транспортировки.

Пусть π — вероятностная мера на $X \times X$ с проекциями μ и ν. Предположим, что интеграл

$$K(\pi) = \int_{X\times X} |x-y|_H \, \pi(dx,dy)$$

конечен. Тогда

$$x - y \in H \quad \text{для } \pi\text{-почти всех } (x,y) \in X \times X. \tag{2.1.1}$$

Производная функции f непрерывна и нигде не обращается в нуль на $\mathbb{R}$. Следовательно, она знакопостоянна. Будем считать, что $f'(x) > 0$. Случай $f'(x) < 0$ рассматривается аналогично.

Теорема 2.1.1. *Пусть мера ν абсолютно непрерывна относительно меры μ. Тогда для π-почти всех $(x, y) \in X \times X$ имеем*

$$f'[f^{-1}(x(t))] = f'[f^{-1}(y(t))], \quad t \in [0, 1],$$

Доказательство. В силу закона повторного логарифма (теорема 1.2.3), для всякого t из интервала $(0, 1)$ для P_W-почти всех траекторий x выполнено равенство

$$\limsup_{\delta \to 0+} \frac{|x(t+\delta) - x(t)|}{\delta^{\frac{1}{2}} |\ln|\ln\delta||^{\frac{1}{2}}} = 1.$$

Поскольку $\mu = P_W \circ f^{-1}$, это равенство можно переписать следующим образом:

$$Lx(t) = \limsup_{\delta \to 0+} \frac{|f^{-1}(x(t+\delta)) - f^{-1}(x(t))|}{\delta^{\frac{1}{2}} |\ln|\ln\delta||^{\frac{1}{2}}} = 1$$

для μ-почти всех $x \in X$. Пусть

$$A_t = \{x \in C[0, 1] : Lx(t) = 1\}.$$

Так как $\mu(A_t) = 1$, а мера ν абсолютно непрерывна относительно меры μ, то $\nu(A_t) = 1$. Следовательно, при каждом фиксированном t имеем

$$\pi(A_t \times A_t) = \pi((A_t \times X) \cap (X \times A_t)) = 1.$$

Таким образом, для π-почти всех $(x, y) \in X \times X$ получаем

$$Lx(t) = Ly(t) = 1, \quad h = y - x \in H.$$

Положим $g := f^{-1}$. Тогда, поскольку $y = x + h$ и $Ly(t) = 1$, верно равенство

$$\limsup_{\delta \to 0+} \frac{|g[x(t+\delta) + h(t+\delta)] - g[x(t) + h(t)]|}{\delta^{\frac{1}{2}} |\ln|\ln\delta||^{\frac{1}{2}}} = 1.$$

Применяя к функции g на отрезке с концами $x(t+\delta) + h(t+\delta)$ и $x(t) + h(t)$ теорему Лагранжа, получим:

$$|g[x(t+\delta) + h(t+\delta)] - g[x(t) + h(t)]| =$$
$$= g'[x(t) + h(t) + \theta_1]|x(t+\delta) - x(t) + h(t+\delta) - h(t)|,$$

где $|\theta_1| \leqslant |x(t+\delta)-x(t)+h(t+\delta)-h(t)| \to 0$ при $\delta \to 0+$ и, следовательно,

$$g'[x(t)+h(t)+\theta_1] \to g'[x(t)+h(t)] \quad \text{при } \delta \to 0+,$$

поскольку функция g' непрерывна. Отсюда

$$\limsup_{\delta\to 0+} g'[x(t)+h(t)]\frac{|x(t+\delta)-x(t)+h(t+\delta)-h(t)|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}} = 1. \tag{2.1.2}$$

Рассмотрим произвольную функцию $h \in W_0^{2,1}[0,1]$. Для нее в силу неравенства Коши–Буняковского в $L^2[t,t+\delta]$ выполняется оценка

$$|h(t+\delta)-h(t)| = \left|\int_t^{t+\delta} h'(t)\,dt\right| \leqslant \|h'\|_{L^2[t,t+\delta]}\cdot\|1\|_{L^2[t,t+\delta]} \leqslant \delta^{\frac{1}{2}}\|h'\|_{L^2}.$$

Из этой оценки для всякого $h \in W_0^{2,1}[0,1]$ имеем

$$Lh(t) \leqslant \limsup_{\delta\to 0+}\frac{\|h'\|_{L^2}}{|\ln|\ln\delta||^{\frac{1}{2}}} = 0,$$

откуда

$$Lh(t) = \lim_{\delta\to 0+}\frac{|h(t+\delta)-h(t)|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}} = 0. \tag{2.1.3}$$

Заметим, что

$$\begin{aligned}|x(t+\delta)-x(t)|-|h(t+\delta)-h(t)| &\leqslant \\ \leqslant |x(t+\delta)-x(t)+h(t+\delta)-h(t)| &\leqslant \\ \leqslant |x(t+\delta)-x(t)|+|h(t+\delta)-h(t)|&.\end{aligned}$$

Верхний предел суммы (произведения) равен сумме (соответственно произведению) верхних пределов, если один из верхних пределов является пределом в обычном смысле, поэтому

$$\begin{aligned}&\limsup_{\delta\to 0+}\frac{|x(t+\delta)-x(t)|-|h(t+\delta)-h(t)|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}} \leqslant \\ &\leqslant \limsup_{\delta\to 0+}\frac{|x(t+\delta)-x(t)+h(t+\delta)-h(t)|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}} \leqslant \\ &\leqslant \limsup_{\delta\to 0+}\frac{|x(t+\delta)-x(t)|+|h(t+\delta)-h(t)|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}}.\end{aligned}$$

Согласно равенству (2.1.3), это неравенство принимает вид

$$\limsup_{\delta\to 0+}\frac{|x(t+\delta)-x(t)|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}}\leqslant \limsup_{\delta\to 0+}\frac{|x(t+\delta)-x(t)+h(t+\delta)-h(t)|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}}\leqslant \limsup_{\delta\to 0+}\frac{|x(t+\delta)-x(t)|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}},$$

откуда

$$\limsup_{\delta\to 0+}\frac{|x(t+\delta)-x(t)+h(t+\delta)-h(t)|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}}=\limsup_{\delta\to 0+}\frac{|x(t+\delta)-x(t)|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}}.$$

Из полученного равенства и из равенства (2.1.2) следует, что

$$\limsup_{\delta\to 0+}\frac{|f[g(x(t+\delta))]-f[g(x(t))]|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}}g'[x(t)+h(t)]=1.$$

Теперь, применяя к функции f на отрезке с концами $g(x(t+\delta))$ и $g(x(t))$ теорему Лагранжа, получим равенство

$$|f[g(x(t+\delta))]-f[g(x(t))]|=f'[g(x(t))+\theta_2]|g(x(t+\delta))-g(x(t))|,$$

где

$$|\theta_2|\leqslant|g(x(t+\delta))-g(x(t))|\to 0\quad \text{при } \delta\to 0+\,.$$

Следовательно, $f'(x(t)+\theta_2)$ сходится к $f'(x(t))$ при $\delta\to 0+$, поскольку функция f' непрерывна. Таким образом,

$$\limsup_{\delta\to 0+}\frac{|g(x(t+\delta))-g(x(t))|}{\delta^{\frac{1}{2}}|\ln|\ln\delta||^{\frac{1}{2}}}f'[g(x(t))]g'[x(t)+h(t)]=1.$$

Это равенство можно переписать так:

$$Lx(t)f'[g(x(t))]g'[x(t)+h(t)]=1,$$

но $Lx(t)=1$ и $x(t)+h(t)=y(t)$, так что

$$f'[g(x(t))]g'[y(t)]=1.$$

По теореме о дифференцировании обратной функции

$$g'[y(t)]=\frac{1}{f'[g(y(t))]}.$$

Следовательно,

$$f'[g(x(t))] = f'[g(y(t))]$$

и

$$f'[f^{-1}(x(t))] = f'[f^{-1}(y(t))] \tag{2.1.4}$$

для π-почти всех (x, y) при фиксированном $t \in (0, 1)$.

Равенство (2.1.4) выполняется π-почти всюду сразу при всех рациональных $t \in (0, 1)$, поскольку в этом случае множество (x, y) таких, что (2.1.4) выполнено, является пересечением по всем рациональным $t \in (0, 1)$ множеств, для которых (2.1.4) выполнено при фиксированном t, т. е. пересечением счетного числа множеств меры 1, следовательно, имеет меру 1. Тогда из непрерывности функций f', f^{-1}, x и y следует, что равенство (2.1.4) выполняется тождественно на $[0, 1]$ для π-почти всех пар (x, y). $\square$

2.2 Случай изолированных нулей коэффициента диффузии

Теорема 2.1.1 позволяет показать, что для достаточно широкого класса функций f отображение T тождественно μ-почти всюду на X, если оно переводит меру μ в абсолютно непрерывную меру.

Теорема 2.2.1. *Пусть $f \in C^2(\mathbb{R})$ — возрастающий диффеоморфизм прямой, причем нули функции f'' изолированы. Предположим, что мера π на $X \times X$, удовлетворяющая условию* (2.1.1)*, имеет проекции μ и ν такие, что ν абсолютно непрерывна относительно μ. Тогда $x = y$ для π-почти всех $(x, y) \in X \times X$ и $\mu = \nu$.*

Доказательство. Зафиксируем пару траекторий $(x, y) \in X \times X$, для которых $(y - x) \in H$ и выполнено равенство (2.1.4). Положим также

$$v(t) := y(t) - x(t).$$

Необходимо доказать, что $v(t) = 0$ для всех $t \in [0, 1]$. Рассмотрим функцию

$$g(x) = f'[f^{-1}(x)].$$

Тогда

$$g'(x) = f''[f^{-1}(x)](f^{-1})'(x).$$

Так как f^{-1} — диффеоморфизм, а нули функции f'' изолированы, то нули функции $f''[f^{-1}(x)]$ также будут изолированы. Следовательно, и нули функции $g'(x)$ изолированы.

Пусть $x(t)$ является траекторией нашего диффузионного процесса. Известно, что у μ-почти каждой такой траектории нет односторонних экстремумов (это следует, например, из закона повторного логарифма). Будем считать, что траектория $x(t)$ лежит в этом множестве полной меры. Равенство (2.1.4) примет вид

$$g[x(t)] - g[x(t) + v(t)] = 0.$$

Применим к функции g на отрезке с концами $x(t)$ и $x(t) + v(t)$ теорему Лагранжа:

$$g'[x(t) + \theta(t)]v(t) = 0, \tag{2.2.1}$$

где $\theta(t)$ принимает значения в интервале с концами 0 и $v(t)$.

Предположим, что существует точка $t_1 \in [0, 1]$ такая, что $v(t_1) \neq 0$. Так как $v(0) = 0$, то найдется ближайшая к t_1 точка $t_0 \in [0, t_1]$ такая, что $v(t_0) = 0$. Это значит, что $v(t) \neq 0$ при $t_0 < t < t_1$.

Пусть $g'(x_0) \neq 0$, где $x_0 = x(t_0)$. Функция g' непрерывна, поэтому в некоторой окрестности U точки x_0 выполнено условие $g'(x) \neq 0$. Существует $\delta > 0$ такое, что $x(t)$ и $x(t)+v(t)$ лежат в U при всех $t_0 < t < t_0+\delta$. Тогда

$$x(t) + \theta(t) \in U \quad \text{и} \quad g'[x(t) + \theta(t)] \neq 0.$$

Следовательно, согласно равенству (2.2.1), $v(t) = 0$ при $t_0 \leqslant t < t_0 + \delta$. Получено противоречие

Пусть теперь $g'(x_0) = 0$. Поскольку нули функции g' изолированы, в некоторой окрестности U точки x_0 не будет нулей функции g', за исключением самой точки x_0. Функции g' и v непрерывны, поэтому существует $\delta > 0$ такое, что $x(t)$, $x(t) + v(t)$ и $x(t) + \theta(t)$ лежат в U при всех $t_0 < t < t_0 + \delta$. Мы можем выбрать δ таким образом, что $\delta < t_1 - t_0$. Тогда $v(t) \neq 0$ при всех $t_0 < t < t_0 + \delta$. Следовательно, из равенства (2.2.1) следует, что при всех $t_0 < t < t_0 + \delta$

$$g'[x(t) + \theta(t)] = 0,$$

поэтому $x(t) + \theta(t) = x_0$.

Пусть, для определенности, $v(t) > 0$ при $t_0 < t < t_1$, тогда $\theta(t) \geqslant 0$ и $x(t) \leqslant x(t_0)$ при $t_0 < t < t_0 + \delta$. Это значит, что $x(t)$ имеет односторонний максимум в точке t_0, чего не может быть благодаря выбору траектории $x(t)$. Следовательно, $v(t) = 0$ при всех $t_0 < t < t_0 + \delta$. Получено противоречие. Таким образом, $v(t) \equiv 0$ на $[0, 1]$, т.е. $x(\cdot) = y(\cdot)$.

Итак, равенство $x(\cdot) = y(\cdot)$ верно, если выполнено условие (2.1.4) и $x - y \in H$. Однако мера π множества таких траекторий равна 1 по теореме 2.1.1, т. е. $x(\cdot) = y(\cdot)$ для π-почти всех пар траекторий.

Для окончания доказательства заметим, что поскольку план π с проекциями μ и ν сосредоточен на диагонали, то $\mu = \nu$. □

Следствие 2.2.2. *Пусть f — вещественно-аналитический, возрастающий, нелинейный диффеоморфизм прямой. Тогда всякая мера π, удовлетворяющая условию* (2.1.1) *и имеющая проекции μ и ν такие, что мера ν абсолютно непрерывна относительно μ, сосредоточена на множестве $\{(x, y)\colon x = y\}$ и верно равенство $\mu = \nu$.*

Доказательство. Очевидно, что f удовлетворяет условиям теоремы 2.2.1, поскольку функция f'' также является вещественно-аналитической. □

2.3 Дальнейшее расширение класса диффузионных процессов

Результат теоремы 2.2.1 может быть улучшен при помощи несколько более технически сложного рассуждения. Покажем, что для отсутствия нетождественных планов транспортировки в задаче Канторовича достаточно потребовать, чтобы множество нулей функции f'' имело нулевую меру Лебега.

Как и прежде, мера μ является распределением диффузионного процесса $\xi_t = f(w_t)$.

Теорема 2.3.1. *Пусть $f \in C^2(\mathbb{R})$ — возрастающий диффеоморфизм вещественной прямой, причем множество нулей функции f'' имеет лебеговскую меру нуль. Предположим, что борелевская вероятностная мера π на $X \times X$ удовлетворяет условию* (2.1.1) *и имеет проекции μ и ν такие, что мера ν абсолютно непрерывна относительно μ. Тогда $x = y$ для π-почти всех $(x, y) \in X \times X$ и $\mu = \nu$.*

Доказательство. Пусть $A \subset \mathbb{R}$ — некоторое множество нулевой меры Лебега. Рассмотрим стандартный винеровский процесс $w(t, \omega)$ на отрезке $t \in [0, 1]$ и функцию

$$I\{w(t,\omega) \in A\} = \begin{cases} 1, & \text{если } w(t,\omega) \in A \\ 0, & \text{иначе.} \end{cases}$$

Стандартный винеровский процесс является измеримой функцией по паре аргументов (t, ω), поэтому по теореме Фубини

$$\int_\Omega \int_0^1 I\{w(t,\omega) \in A\}\, dt dP = \int_0^1 \int_\Omega I\{w(t,\omega) \in A\}\, dP dt.$$

Это равенство можно преобразовать следующим образом:

$$\int_\Omega \lambda\{t \in [0,1]\colon\ w(t,\omega) \in A\}\, dP = \int_0^1 P\{\omega\colon\ w(t,\omega) \in A\}\, dt,$$

где λ — мера Лебега на $[0, 1]$. Заметим, что

$$P\{\omega\colon\ w(t,\omega) \in A\} = \frac{1}{\sqrt{2\pi t}} \int_A \exp\left(-\frac{s^2}{2t}\right) ds = 0,$$

так как A имеет нулевую меру Лебега. Следовательно,

$$\int_{\Omega} \lambda\{t \in [0,1]\colon\ w(t,\omega) \in A\}\, dP = 0.$$

Таким образом, поскольку

$$\lambda\{t \in [0,1]\colon\ w(t,\omega) \in A\} \geqslant 0,$$

то выполнено равенство

$$\lambda\{t \in [0,1]\colon\ w(t,\omega) \in A\} = 0 \quad \text{п.н.}$$

или, эквивалентно,

$$\lambda\{t \in [0,1]\colon\ w(t) \in A\} = 0 \quad P_W\text{-почти всюду.}$$

Если f — диффеоморфизм, то $f^{-1}(A)$ также имеет нулевую меру Лебега. Следовательно,

$$\lambda\{t \in [0,1]\colon\ w(t) \in f^{-1}(A)\} = 0 \quad P_W\text{-почти всюду.}$$

Так как $\mu = P_W \circ f^{-1}$, а мера ν абсолютно непрерывна относительно μ, то имеем

$$\lambda\{t \in [0,1]\colon\ x(t) \in A\} = 0 \quad \nu\text{-почти всюду.}$$

Согласно теореме 2.2.1, для π-почти всех $(x,y) \in X \times X$ выполнено равенство

$$g(x(t)) = g(y(t)), \tag{2.3.1}$$

где $g(s) = f'[f^{-1}(s)]$. Пусть Z — множество нулей функции f'', которое по условию имеет нулевую меру Лебега. Множество нулей производной

$$g'(s) = f''(f^{-1}(s))(f^{-1})'(s)$$

есть в точности $f(Z)$, так как $(f^{-1})'(s) > 0$. Оно имеет нулевую меру Лебега, поскольку f — диффеоморфизм.

Далее, для ν-почти всех $y(\cdot)$ имеем

$$\lambda\{t \in [0,1]\colon\ y(t) \in f(Z)\} = 0, \tag{2.3.2}$$

поэтому для π-почти всех (x, y) выполнены условия (2.3.1) и (2.3.2). Зафиксируем пару траекторий (x, y) с этими условиями. В силу условия (2.3.2) для почти всех $t \in [0, 1]$

$$h'(t) \quad \text{определена и} \quad g'(y(t)) \neq 0. \tag{2.3.3}$$

Зафиксируем t_0 с условием (2.3.3). В некоторой окрестности V точки $s_0 = y(t_0)$ функция g обратима по теореме об обратной функции. Пусть g_1 — обратная к g в V, а U — такая окрестность точки t_0, что $y(t) \in V$ при $t \in U$. Тогда

$$y(t) = g_1[g(x(t))] \quad \text{при } t \in U.$$

Согласно условию (2.1.1), для некоторой функции $h \in H$

$$h(t) = G[x(t)] \quad \text{при } t \in U,$$

где $G(s) = g_1[g(s)] - s$ — непрерывно дифференцируемая функция в окрестности V точки $s_0 = y(t_0)$.

Пусть $G'[x(t_0)] \neq 0$, тогда по теореме об обратной функции в некоторой окрестности точки $s_1 = x(t_0)$ существует обратная к G дифференцируемая функция G_1. Таким образом,

$$x(t) = G_1(h(t))$$

в некоторой окрестности точки t_0. Функция в правой части равенства дифференцируема в точке t_0, в то время как $x(t)$ есть траектория диффузионного процесса и не является дифференцируемой ни в одной точке. Получено противоречие.

Пусть теперь $G'[x(t_0)] = 0$. Тогда $h'(t_0) = 0$. Поскольку t_0 выбиралось из множества полной меры, то получаем

$$h'(t) = 0 \quad \text{почти всюду на [0,1]}.$$

Таким образом,

$$|h|_H = 0 \quad \text{и} \quad h \equiv 0,$$

но это означает, что $x(\cdot) = y(\cdot)$.

Функции $x(\cdot)$ и $y(\cdot)$ совпадают, если $x-y \in H$ и справедливы условия (2.3.1) и (2.3.2). Все эти условия выполнены для π-почти всех пар (x, y), т. е. $x(\cdot) = y(\cdot)$ для π-почти всех $(x, y) \in X \times X$.

Для окончания доказательства заметим, что поскольку план π с проекциями μ и ν сосредоточен на диагонали, то $\mu = \nu$. $\square$

Итак, доказано, что при предположениях теоремы 2.3.1 не существует планов транспортировки конечной стоимости в том случае, когда меры μ и ν различны. В частности, не существует преобразований μ в ν конечной стоимости, следовательно, соответствующая задача Монжа тоже неразрешима.

Глава 3

Оценка значений функционала стоимости в задаче Монжа для оптимальных и треугольных отображений

В известной работе [50] М. Талагран доказал, что если γ — стандартная гауссовская мера на $\mathbb{R}^n$ и вероятностная мера $\mu \ll \gamma$ такова, что энтропия $\mathrm{Ent}_\gamma(\mu)$ конечна, то существует такое возрастающее треугольное отображение T, что $\mu = \gamma \circ T^{-1}$ и

$$\int_{\mathbb{R}^n} |T(x) - x|^2 \, \gamma(dx) \leqslant 2 \, \mathrm{Ent}_\gamma(\mu).$$

Аналогичное неравенство верно и для оптимального отображения, поэтому естественно сравнить значения функционала стоимости в задаче Канторовича для этих двух классов отображений.

3.1 Линейность оптимальных отображений гауссовских мер

В этом разделе мы докажем, что как оптимальное, так и каноническое треугольное отображения одной центрированной гауссовской меры в другую линейны.

Лемма 3.1.1. *Пусть γ и μ — две центрированные невырожденные гауссовские меры на $\mathbb{R}^n$. Тогда оптимальное отображение T_0, переводящее меру γ в меру μ линейно, а каноническое треугольное отображение T меры γ в меру μ также линейно.*

Доказательство. Сначала докажем, что отображение T_0 линейно. Известно (см. [51]), что оно γ-почти всюду единственно и имеет вид градиента $T_0 = \nabla\Psi$ некоторой выпуклой функции $\Psi\colon \mathbb{R}^n \to \mathbb{R}$. Следовательно, достаточно построить линейный оператор с симметричной матрицей, который переводит γ в μ (такой оператор является градиентом квадратичной формы).

Меры γ и μ задаются своими преобразованиями Фурье φ_γ и φ_μ соответственно (см. [1]):

$$\varphi_\gamma(x) = \exp\left(-\frac{(Mx,x)}{2}\right), \quad \varphi_\mu(x) = \exp\left(-\frac{(Nx,x)}{2}\right),$$

где симметричные положительно определенные матрицы M и N — ковариационные матрицы гауссовских мер γ и μ соответственно. Для всякой пары симметричных положительно определенных матриц M и N, существует симметричная положительно определенная матрица A такая, что $N = AMA$ (матричное уравнение Рикатти). Матрица A имеет вид:

$$A = M^{-1/2}\left(M^{1/2}NM^{1/2}\right)^{1/2}M^{-1/2}.$$

Действительно, подставляя A в требуемое равенство, получим:

$$\begin{aligned} AMA &= \left(M^{-1/2}\left(M^{1/2}NM^{1/2}\right)^{1/2}M^{-1/2}\right)\times \\ &\times\left(M^{1/2}M^{1/2}\right)\left(M^{-1/2}\left(M^{1/2}NM^{1/2}\right)^{1/2}M^{-1/2}\right) = \\ &= M^{-1/2}\left(M^{1/2}NM^{1/2}\right)^{1/2}\left(M^{1/2}NM^{1/2}\right)^{1/2}M^{-1/2} \\ &= M^{-1/2}M^{1/2}NM^{1/2}M^{-1/2} = N. \end{aligned}$$

В силу симметрии матрицы A соответствующее линейное отображение является градиентом квадратичной формы $2^{-1}(Ax,x)$. С другой стороны, мера γ переходит в меру μ под действием линейного оператора с матрицей A. Согласно теореме 1.2.5 градиент, задающий оптимальное отображение, единственен, поэтому полученное линейное отображение является оптимальным.

Построим теперь линейное треугольное отображение T. Найдется нижнетреугольная матрица D такая, что $N = DMD^T$. Применим разложение Холецкого для положительно определенных симметричных матриц M и N (см. [38]):

$$M = BB^T \quad \text{и} \quad N = CC^T,$$

где B и C — нижнетреугольные матрицы с положительными элементами на диагонали. Искомая матрица имеет вид $D = CB^{-1}$. Поскольку положительно определенные нижнетреугольные матрицы образуют группу, матрица D является нижнетреугольной. Следующее равенство показывает, что линейный оператор T с матрицей D переводит меру γ в меру μ:

$$DMD^T = \bigl(CB^{-1}\bigr)\bigl(BB^T\bigr)\bigl(CB^{-1}\bigr)^T = CB^{-1}BB^T\bigl(B^{-1}\bigr)^T C^T = CC^T = N.$$

Отображение T является каноническим треугольным отображением, переводящим меру γ в меру μ в силу единственности (теорема 1.2.4). $\square$

3.2 Случай ковариационной матрицы с единичным определителем

Нам предстоит исследовать, во сколько раз величина

$$\int_{\mathbb{R}^n} |T(x) - x|^2 \, \gamma(dx)$$

для треугольного отображения $T(x)$ может быть больше

$$\int_{\mathbb{R}^n} |T_0(x) - x|^2 \, \gamma(dx)$$

для оптимального $T_0(x)$. Оказывается, для сравнимости этих величин соответствующие меры и отображения должны удовлетворять достаточно жестким ограничениям; в частности, даже для произвольных линейных отображений гауссовских мер постоянная K в неравенстве

$$\int_{\mathbb{R}^n} |T(x) - x|^2 \gamma(dx) \leqslant K \int_{\mathbb{R}^n} |T_0(x) - x|^2 \gamma(dx) \qquad (3.2.1)$$

неограниченно возрастает с ростом размерности пространства n. В то же время, при некоторых ограничениях на меру, получаемую при отображении, справедлива оценка вида (3.2.1) с коэффициентом K, не зависящим от размерности.

Лемма 3.1.1 позволяет получить предварительную оценку для коэффициента K в неравенстве (3.2.1).

Теорема 3.2.1. *Пусть γ — стандартная гауссовская мера на $\mathbb{R}^n$ и M — симметричная положительно определенная матрица размера $n \times n$ с собственными значениями μ_i и определителем, равным единице. Тогда в некотором ортонормированном базисе пространства $\mathbb{R}^n$ центрированная гауссовская мера μ с ковариационной матрицей M обладает следующим свойством: для оптимального линейного отображения T_0, переводящего меру γ в меру μ (которое задается матрицей $\sqrt{M}$ в этом базисе) и для канонического треугольного отображения T меры γ в меру μ (в этом же базисе) справедлива оценка*

$$K \geqslant 1 + 2\frac{\sum_{i=1}^n (\lambda_i - 1)}{\sum_{i=1}^n (\lambda_i - 1)^2},$$

где λ_i — собственные числа матрицы A отображения T_0 в стандартном базисе, то есть $\lambda_i = \sqrt{\mu_i}$.

Доказательство. Оптимальное отображение T_0 линейно по лемме 3.1.1. Пусть $A = \sqrt{M}$ — матрица этого отображения. Она диагональна в некотором базисе, но не обязана быть треугольной в стандартном базисе. Заметим, что

$$\sum_{i,j=1}^n a_{ij}^2 = \operatorname{tr} M. \qquad (3.2.2)$$

Рассмотрим матрицу M. Покажем, как ортогональными преобразованиями привести ее к виду, в котором все угловые миноры будут равны 1.

С помощью ортогональной матрицы U можно привести матрицу M к диагональному виду:

$$U^{-1}\begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & \mu_n \end{pmatrix} U = M, \quad U^{-1} = U^T.$$

Без ограничения общности будем считать, что все μ_i отличны от 1 при всех $i = 1, \dots, n$, причем ни для какого конечного набора различных индексов из множества $\{1, \dots, n\}$ произведение чисел μ_i с этими индексами не равно 1, кроме набора $1, \dots, n$. В противном случае, перейдя к подпространствам меньшей размерности, проделаем в каждом из них замены базисов, описанные далее.

По индукции будем поворотами пар координатных векторов приводить диагональную матрицу к симметричному виду с единичными угловыми минорами. Пусть, без ограничения общности, $\mu_1 > 1 > \mu_n$ (этого всегда можно добиться перестановкой базисных векторов, выбрав при этом в качестве μ_1 и μ_n наибольшее и наименьшее среди μ_i соответственно). Рассмотрим поворот в плоскости (x_1, x_n) на угол φ:

$$\begin{pmatrix} \cos\varphi & 0 & \dots & 0 & \sin\varphi \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ -\sin\varphi & 0 & \dots & 0 & \cos\varphi \end{pmatrix} \begin{pmatrix} \mu_1 & 0 & \dots & 0 & 0 \\ 0 & \mu_2 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & \mu_{n-1} & 0 \\ 0 & 0 & \dots & 0 & \mu_n \end{pmatrix} \times$$

$$\times \begin{pmatrix} \cos\varphi & 0 & \dots & 0 & -\sin\varphi \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ \sin\varphi & 0 & \dots & 0 & \cos\varphi \end{pmatrix}.$$

После перемножения получим:

$$\begin{pmatrix} \mu_1\cos^2\varphi+\mu_n\sin^2\varphi & 0 & \dots & 0 & (\mu_n-\mu_1)\sin\varphi\cos\varphi \\ 0 & \mu_2 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & \mu_{n-1} & 0 \\ (\mu_n-\mu_1)\sin\varphi\cos\varphi & 0 & \dots & 0 & \mu_1\sin^2\varphi+\mu_n\cos^2\varphi \end{pmatrix}.$$

Для того чтобы $\mu_1\cos^2\varphi+\mu_n\sin^2\varphi=1$, достаточно положить

$$\operatorname{tg}\varphi=\sqrt{\frac{\mu_1-1}{1-\mu_n}}.$$

Пусть теперь первые $i-1$ угловых миноров ковариационной матрицы равны 1. Перестановкой базисных векторов добьемся того, чтобы $\mu_i>1>\mu_1\dots\mu_{i-1}\mu_n$ или, наоборот, $\mu_i<1<\mu_1\dots\mu_{i-1}\mu_n$. С помощью поворота в плоскости (x_i,x_n) получим следующее ($\widetilde{\mu}$ — элемент, полученный в правом нижнем углу матрицы после всех предыдущих поворотов):

$$\left(\begin{array}{ccc|cc} 1 & & 0 & 0 & 0 \\ & \ddots & & \vdots & \vdots \\ 0 & & 1 & 0 & 0 \\ \hline 0 & \dots & 0 & \cos\varphi & \sin\varphi \\ 0 & \dots & 0 & -\sin\varphi & \cos\varphi \end{array}\right)\left(\begin{array}{ccc|cc} m_{11} & \dots & m_{1,i-1} & 0 & m_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ m_{1,i-1} & \dots & m_{i-1,i-1} & 0 & m_{i-1,n} \\ \hline 0 & \dots & 0 & \mu_i & 0 \\ m_{1n} & \dots & m_{i-1,n} & 0 & \widetilde{\mu} \end{array}\right)\times$$

$$\times\left(\begin{array}{ccc|cc} 1 & & 0 & 0 & 0 \\ & \ddots & & \vdots & \vdots \\ 0 & & 1 & 0 & 0 \\ \hline 0 & \dots & 0 & \cos\varphi & -\sin\varphi \\ 0 & \dots & 0 & \sin\varphi & \cos\varphi \end{array}\right)=$$

$$=\left(\begin{array}{ccc|cc} m_{11} & \dots & m_{1,i-1} & m_{1n}\sin\varphi & m_{1n}\cos\varphi \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ m_{1,i-1} & \dots & m_{i-1,i-1} & m_{i-1,n}\sin\varphi & m_{i-1,n}\cos\varphi \\ \hline m_{1n}\sin\varphi & \dots & m_{i-1,n}\sin\varphi & \mu_i\cos^2\varphi+\widetilde{\mu}\sin^2\varphi & (\widetilde{\mu}-\mu_i)\sin\varphi\cos\varphi \\ m_{1n}\cos\varphi & \dots & m_{i-1,n}\cos\varphi & (\widetilde{\mu}-\mu_i)\sin\varphi\cos\varphi & \mu_i\sin^2\varphi+\widetilde{\mu}\cos^2\varphi \end{array}\right).$$

В этой формуле опущены направления $x_{i+1}, \ldots, x_{n-1}$, которые не участвовали ни в одном повороте.

Требуется найти угол φ такой, что следующий определитель равен единице:

$$D = \begin{vmatrix} m_{11} & \ldots & m_{1,i-1} & m_{1n}\sin\varphi \\ \vdots & \ddots & \vdots & \vdots \\ m_{1,i-1} & \ldots & m_{i-1,i-1} & m_{i-1,n}\sin\varphi \\ m_{1n}\sin\varphi & \ldots & m_{i-1,n}\sin\varphi & \mu_i\cos^2\varphi + \widetilde{\mu}\sin^2\varphi \end{vmatrix}.$$

По предположению индукции, угловой минор порядка $(i-1)\times(i-1)$ равен 1. Обозначим его M_{i-1}. При указанных заменах координат, величина углового минора, соответствующего подпространству $(x_1, \ldots, x_{i-1}, x_n)$, постоянна, поэтому верно равенство

$$\begin{vmatrix} m_{11} & \ldots & m_{1,i-1} & m_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ m_{1,i-1} & \ldots & m_{i-1,i-1} & m_{i-1,n} \\ m_{1n} & \ldots & m_{i-1,n} & \widetilde{\mu} \end{vmatrix} = \mu_1\cdots\mu_{i-1}\mu_n.$$

Разложим получившийся определитель по последней строке:

$$\mu_1\cdots\mu_{i-1}\mu_n = \widetilde{\mu}M_{i-1} + \sum_{k=1}^{i-1} m_{kn}M_{i-1,k} = \widetilde{\mu} + \sum_{k=1}^{i-1} m_{kn}M_{i-1,k},$$

где $M_{i-1,k}$ — дополнительные миноры порядка $(i-1)$ с соответствующими знаками. Определитель D также разложим по последней строке:

$$D = (\mu_i\cos^2\varphi + \widetilde{\mu}\sin^2\varphi)M_{i-1} + \sum_{k=1}^{i-1}(m_{kn}\sin\varphi)(M_{i-1,k}\sin\varphi) =$$

$$= \mu_i\cos^2\varphi + \left(\widetilde{\mu} + \sum_{k=1}^{i-1} m_{kn}M_{i-1,k}\right)\sin^2\varphi = \mu_i\cos^2\varphi + \mu_1\cdots\mu_{i-1}\mu_n\sin^2\varphi.$$

Как и в случае с самым первым поворотом, положим

$$\operatorname{tg}\varphi = \sqrt{\frac{\mu_i - 1}{1 - \mu_1\cdots\mu_{i-1}\mu_n}}.$$

Продолжая эту процедуру, получим матрицу M, у которой все угловые миноры равны единице, и ортогональный базис, в котором это имеет место.

По лемме 3.1.1 каноническое треугольное отображение стандартной гауссовской меры γ в центрированную гауссовскую меру μ (в построенном ортогональном базисе) является линейным оператором T, заданным нижнетреугольной матрицей B (в том же ортогональном базисе). Диагональные элементы матрицы B положительны, причем

$$M = BB^T \quad \text{и} \quad \operatorname{tr} M = \sum_{i,j=1}^{n} b_{ij}^2, \tag{3.2.3}$$

Докажем, что на диагонали матрицы B стоят единицы. Для этого выразим ее диагональные элементы через угловые миноры матрицы M. Пусть матрица B имеет вид

$$B = \begin{pmatrix} b_{11} & 0 & \dots & 0 \\ b_{12} & b_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ b_{1n} & b_{2n} & ... & b_{nn} \end{pmatrix}$$

Обозначим через B_i ее минор, образованный пересечением первых i строк и первых i столбцов. Тогда, согласно (3.2.3), угловой минор M_i матрицы M равен определителю следующего матричного произведения:

$$\begin{pmatrix} b_{11} & 0 & \dots & 0 \\ b_{12} & b_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ b_{1i} & b_{2i} & \dots & b_{ii} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1i} \\ 0 & b_{22} & \dots & b_{2i} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{ii} \end{pmatrix},$$

откуда $M_i = B_i^2$.

Поскольку матрица B — нижнетреугольная, всякий ее угловой минор равен произведению своих диагональных элементов, поэтому верны соотношения

$$M_1 = b_{11}^2, \quad M_2 = b_{11}^2 b_{22}^2, \ \dots\ , \quad M_n = |M| = b_{11}^2 \cdots b_{nn}^2.$$

Выразим отсюда диагональные элементы матрицы B:

$$b_{11} = \sqrt{M_1}, \quad b_{22} = \frac{\sqrt{M_2}}{\sqrt{M_1}}, \ \dots\ , \quad b_{nn} = \frac{\sqrt{M_n}}{\sqrt{M_{n-1}}}.$$

Все угловые миноры M_i матрицы M равны 1, следовательно,

$$b_{11} = b_{22} = \ldots = b_{nn} = 1.$$

Итак, мы выбрали базис, в котором матрицы линейных операторов T_0 и T имеют вид A и B соответственно и на диагонали матрицы B стоят единицы. Исследуемая оценка (3.2.1) примет вид

$$\int_{\mathbb{R}^n} |Bx - x|^2\, \gamma(dx) \leqslant K \int_{\mathbb{R}^n} |Ax - x|^2\, \gamma(dx). \tag{3.2.4}$$

Вычислим значение функционала стоимости для произвольного линейного отображения, заданного в некотором ортогональном базисе пространства $\mathbb{R}^n$ матрицей C:

$$\int_{\mathbb{R}^n} |Cx - x|^2\, \gamma(dx) = \operatorname{tr}(C^T C) - 2 \operatorname{tr} C + n = \sum_{i,j=1}^{n} c_{ij}^2 - 2 \sum_{i=1}^{n} c_{ii} + n.$$

Принимая во внимание равенства (3.2.2), (3.2.3), запишем значения функционалов стоимости для отображений T и T_0:

$$\int_{\mathbb{R}^n} |Ax - x|^2\, d\gamma = \operatorname{tr} M - 2 \operatorname{tr} A + n,$$
$$\int_{\mathbb{R}^n} |Bx - x|^2\, d\gamma = \operatorname{tr} M - 2 \operatorname{tr} B + n = \operatorname{tr} M - n.$$

Отметим, что первая из приведенных выше формул при $A = \sqrt{M}$ является частным случаем более общей формулы для квадратичного расстояния Канторовича–Рубинштейна между двумя произвольными гауссовскими мерами на $\mathbb{R}^n$ (см. [37] и [32]). Эта формула и первое равенство во второй формуле справедливы для произвольных линейных операторов, переводящих γ в μ, но оптимальное отображение имеет среди всех таких операторов максимальный след.

Подставляя полученные значения в (3.2.4), находим следующую оценку для константы K:

$$K \geqslant \frac{\operatorname{tr} M - 2\operatorname{tr} B + n}{\operatorname{tr} M - 2\operatorname{tr} A + n} = \frac{\operatorname{tr} M - n}{\operatorname{tr} M - 2\operatorname{tr} A + n} = \\ = 1 + 2\frac{\operatorname{tr} A - n}{\operatorname{tr} M - 2\operatorname{tr} A + n}. \quad (3.2.5)$$

След матрицы инвариантен при ортогональных заменах координат, поэтому перейдем в систему координат, где матрица A диагональна, а на диагонали стоят ее собственные числа $\lambda_1, \ldots, \lambda_n$. Тогда

$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i, \quad \operatorname{tr} M = \operatorname{tr}(A^2) = \sum_{i=1}^{n} \lambda_i^2,$$

и, подставляя в (3.2.5), получаем оценку

$$K \geqslant 1 + 2\frac{\sum_{i=1}^n \lambda_i - n}{\sum_{i=1}^n \lambda_i^2 - 2\sum_{i=1}^n \lambda_i + n} = 1 + 2\frac{\sum_{i=1}^n (\lambda_i - 1)}{\sum_{i=1}^n (\lambda_i - 1)^2},$$

что и требовалось доказать. □

3.3 Гауссовские меры с произвольной ковариационной матрицей

Рассмотрим теперь случай, когда μ — центрированная гауссовская мера с ковариационной матрицей M, определитель которой положителен. Такую меру можно получить из меры μ_1 с ковариационной матрицей M_1, определитель которой равен единице. Достаточно в качестве M_1 взять матрицу с единичным определителем, полученную из матрицы M умножением всех ее элементов на $|M|^{-1/n}$. Тогда μ будет получаться из μ_1 гомотетией, т.е. растяжением равномерно во всех направлениях с коэффициентом $\alpha = |M|^{1/2n}$. Рассмотрим ортогональный базис, построенный в теореме 3.2.1. Оптимальное отображение, переводящее меру γ в μ_1, линейно, причем его матрица A_1 симметрична, имеет единичный определитель и $A_1^2 = M_1$. Тогда матрица A, полученная домножением всех элементов матрицы A_1 на α, также симметрична, причем для нее верно равенство $A^2 = M$, т. е. этой матрицей задается оптимальное отображение γ в μ (доказательство буквально повторяет аналогичное рассуждение из теоремы 3.2.1). Каноническое треугольное отображение меры γ в меру μ_1 задается нижнетреугольной матрицей B_1 в рассматриваемом ортогональном базисе. Умножая все элементы B_1 на α, получим матрицу канонического треугольного отображения, переводящего меру γ в меру μ, причем ее след будет равен $n\alpha$.

Следы матриц A и M выражаются через собственные числа $\lambda_1, \dots, \lambda_n$ матрицы A_1 следующим образом:

$$\operatorname{tr} A = \alpha \sum_{i=1}^{n} \lambda_i, \quad \operatorname{tr} M = \alpha^2 \sum_{i=1}^{n} \lambda_i^2.$$

Подставим эти выражения в оценку (3.2.5) для коэффициента K:

$$K \geqslant 1 + 2\frac{\operatorname{tr} A - \operatorname{tr} B}{\operatorname{tr} M - 2\operatorname{tr} A + n} = 1 + 2\frac{\sum_{i=1}^{n} \alpha\lambda_i - n\alpha}{\sum_{i=1}^{n}(\alpha\lambda_i)^2 - 2\sum_{i=1}^{n} \alpha\lambda_i + n} = \\ = 1 + 2\frac{\alpha\sum_{i=1}^{n}(\lambda_i - 1)}{\sum_{i=1}^{n}(\alpha\lambda_i - 1)^2}. \quad (3.3.1)$$

В следующей теореме мы исследуем, какие значения может принимать выражение, полученное в (3.3.1), и получим оценку снизу для коэффициента K, а также покажем, что он неограниченно возрастает с ростом n. Из этого будет следовать отсутствие сравнимости треугольного и оптимального отображения в общем случае.

Прежде чем это доказать, заметим, что если для операторной нормы матрицы M мы имеем $\|M\| \geqslant 4$, то для произвольного линейного оператора L, переводящего стандартную гауссовскую меру γ на $\mathbb{R}^n$ в меру с ковариационной матрицей M, справедливо неравенство

$$\int_{\mathbb{R}^n} |Lx - x|^2\,\gamma(dx) \leqslant (5n + 4)\int_{\mathbb{R}^n} |T_0x - x|^2\,\gamma(dx),$$

где $T_0 = \sqrt{M}$ — оптимальное отображение. Действительно, мы доказали, что левая часть неравенства имеет вид $\operatorname{tr} M - 2\operatorname{tr} L + n$ и $L^T L = M$, поэтому $\|L\|^2 = \|M\|$. Тогда $2|\operatorname{tr} L| \leqslant 2n\|M\|^{1/2}$, поэтому левая часть указанного неравенства не превосходит $\operatorname{tr} M + 2n\|M\|^{1/2} + n$, что оценивается сверху правой частью, поскольку $M^{1/2} - I \geqslant 2^{-1}M^{1/2}$ и $\|M\|^{1/2} \leqslant \operatorname{tr} M/2$.

Теперь предположим, что γ — снова стандартная гауссовская мера на пространстве $\mathbb{R}^n$ и μ — центрированная гауссовская мера, у которой собственные значения ковариационной матрицы M равны $\mu_1, \dots \mu_n$. Тогда энтропия меры μ относительно меры γ имеет вид

$$\int \rho \ln \rho \, d\gamma = \int \ln \rho \, d\mu = 2^{-1} \sum_{i=1}^{n} (\mu_i - 1 - \ln \mu_i),$$

где $\rho = d\mu/d\gamma$. Действительно, предположим, что матрица M диагональна, тогда

$$\ln \rho = \sum_{i=1}^{n} (-2^{-1} \ln \mu_i + 2^{-1} x_i^2 (1 - 1/\mu_i)),$$

где интеграл от x_i^2 по мере μ равен μ_i. Если $|\mu_i - 1| \leqslant 1/2$, то

$$\int \rho \ln \rho \, d\gamma \leqslant 2 \sum_{i=1}^{n} (\mu_i - 1)^2 \leqslant \frac{25}{2} \sum_{i=1}^{n} (\mu_i^{1/2} - 1)^2,$$

поскольку $(\mu_i - 1)^2 \leqslant 25(\mu_i^{1/2} - 1)^2/4$. Справа записано значение функционала стоимости в задаче Монжа для оптимальной транспортировки $T_0 = M^{1/2}$, поэтому из неравенства Талаграна (теорема 1.2.7) следует, что

$$\int |Tx - x|^2 \, \gamma(dx) \leqslant 25 \int |T_0 x - x|^2 \, \gamma(dx), \text{ если } \|M - I\| \leqslant 1/2. \quad (3.3.2)$$

Как мы увидим далее, существенную роль здесь играет ограничение нормы M.

Теорема 3.3.1. *Пусть γ — стандартная гауссовская мера на $\mathbb{R}^n$. Будем рассматривать в качестве μ всевозможные невырожденные центрированные гауссовские меры на $\mathbb{R}^n$. Тогда, если T — треугольное, а T_0 — оптимальное отображение, переводящее меру γ в меру μ, то для наименьшей возможной постоянной K в неравенстве*

$$\int_{\mathbb{R}^n} |T(x) - x|^2 \, \gamma(dx) \leqslant K \int_{\mathbb{R}^n} |T_0(x) - x|^2 \, \gamma(dx)$$

верна оценка

$$K \geqslant n + \sqrt{n^2 - n}.$$

В частности, если рассматривать меры μ с ковариационной матрицей, имеющей единичный определитель, то коэффициент K не может быть меньше $\sqrt{n}$.

Доказательство. Второе утверждение является непосредственным следствием теоремы 3.2.1. Достаточно рассмотреть матрицу M такого вида, что $\lambda_i = \sqrt{\mu_i} = 1+n^{-1/2}$ при $i = 1,\dots,n-1$ и $\lambda_n = \sqrt{\mu_n} = (1+n^{-1/2})^{1-n}$. Заметим, что $(1+n^{-1/2})^n = \left((1+n^{-1/2})^{n^{1/2}}\right)^{n^{1/2}} = (e+o(1))^{n^{1/2}}$, следовательно,

$$\sum_{i=1}^{n}(\lambda_i - 1) = (n-1)n^{-1/2} - 1 + o(1), \quad \sum_{i=1}^{n}(\lambda_i - 1)^2 = (n-1)n^{-1} + 1 + o(1),$$

и отношение этих выражений имеет вид $n^{1/2} + o(1)$.

Покажем теперь, что в общем случае это отношение может быть еще больше. Как было доказано ранее, отображения T_0 и T — линейные, причем их матрицы A и B удовлетворяют равенствам:

$$AA^T = BB^T = M,$$

где M — ковариационная матрица меры μ. Пусть A_1 — матрица с определителем 1, гомотетичная с коэффициентом $\alpha = |M|^{1/2n}$ матрице A. Тогда A_1 симметрична, пусть $\lambda_1,\dots,\lambda_n$ — ее собственные числа. Найдем при фиксированных $\lambda_k,\ k = 1,\dots,n$ максимум величины

$$F_{\lambda_1,\dots,\lambda_n}(\alpha) = \frac{\alpha\sum_{i=1}^{n}(\lambda_i - 1)}{\sum_{i=1}^{n}(\alpha\lambda_i - 1)^2} \tag{3.3.3}$$

как функции от α.

Случай, когда все λ_i равны, не подходит по смыслу задачи, ведь если произведение λ_i равно 1 (это определитель матрицы A_1), то и $\lambda_i = 1$ при $i = 1,\dots,n$. Тогда при $\alpha = 1$ в выражении (3.3.3) возникает неопределенность, которая объясняется тем, что под действием линейного отображения с матрицей A мера γ переходит в себя, т. е. оптимальным является тождественное отображение, а функционал стоимости обнуляется.

Итак, не все λ_i равны между собой, поэтому функция $F_{\lambda_1,\dots,\lambda_n}(\alpha)$ определена и непрерывна при всех $\alpha > 0$. Отметим, что ее пределы при $\alpha \to 0$

и $\alpha \to \infty$ равны нулю, поэтому она ограничена на $(0, +\infty)$ и имеет на этом промежутке локальный экстремум, который мы найдем дифференцированием по α:

$$F'_{\lambda_1,\ldots,\lambda_n}(\alpha) = \frac{\Big(\sum_{i=1}^n(\alpha\lambda_i - 1)^2 - \alpha\sum_{i=1}^n 2\lambda_i(\alpha\lambda_i - 1)\Big)\sum_{i=1}^n(\lambda_i - 1)}{\Big(\sum_{i=1}^n(\alpha\lambda_i - 1)^2\Big)^2} = 0. \tag{3.3.4}$$

Поскольку не все λ_i равны между собой, то

$$\sum_{i=1}^n(\lambda_i - 1) = \sum_{i=1}^n \lambda_i - n > 0,$$

ибо по неравенству о средних

$$\frac{1}{n}\sum_{i=1}^n \lambda_i \geqslant \Big(\prod_{i=1}^n \lambda_i\Big)^{1/n} = 1,$$

причем равенство достигается лишь при условии, что все λ_i равны. Вернемся к уравнению (3.3.4). Упрощая, получим

$$\begin{aligned}\sum_{i=1}^n(\alpha\lambda_i - 1)^2 - \alpha\sum_{i=1}^n 2\lambda_i(\alpha\lambda_i - 1) &= \sum_{i=1}^n(\alpha^2\lambda_i^2 - 2\alpha\lambda_i + 1 - 2\alpha^2\lambda_i^2 + 2\alpha\lambda_i)\\ &= \sum_{i=1}^n(1 - \alpha^2\lambda_i^2) = n - \alpha^2\sum_{i=1}^n \lambda_i^2 = 0,\end{aligned}$$

откуда предполагаемая точка экстремума есть

$$\alpha_0 = \frac{\sqrt{n}}{\sqrt{\sum_{i=1}^n \lambda_i^2}}.$$

Заметим, что производная $F'_{\lambda_1,\ldots,\lambda_n}(\alpha)$ положительна при $0 < \alpha < \alpha_0$ и отрицательна при $\alpha > \alpha_0$, то есть в точке α_0 достигается максимум функции $F_{\lambda_1,\ldots,\lambda_n}(\alpha)$ при $\alpha > 0$ при фиксированных $\lambda_1, \ldots, \lambda_n$:

$$\begin{aligned}F_{\lambda_1,\ldots,\lambda_n}(\alpha_0) &= \frac{\sum_{i=1}^n(\lambda_i - 1)}{\alpha_0\sum_{i=1}^n \lambda_i^2 - 2\sum_{i=1}^n \lambda_i + n/\alpha_0} =\\ &= \frac{\sum_{i=1}^n(\lambda_i - 1)}{\sqrt{n\sum_{i=1}^n \lambda_i^2} - 2\sum_{i=1}^n \lambda_i + \sqrt{n\sum_{i=1}^n \lambda_i^2}} = \frac{\sum_{i=1}^n(\lambda_i - 1)}{2\Big(\sqrt{n\sum_{i=1}^n \lambda_i^2} - \sum_{i=1}^n \lambda_i\Big)}.\end{aligned}$$

Тогда оценка (3.3.1) принимает вид

$$K \geqslant 1 + 2 \cdot \frac{\sum_{i=1}^n \lambda_i - n}{2\Big(\sqrt{n\sum_{i=1}^n \lambda_i^2} - \sum_{i=1}^n \lambda_i\Big)} = \frac{\sqrt{n\sum_{i=1}^n \lambda_i^2} - n}{\sqrt{n\sum_{i=1}^n \lambda_i^2} - \sum_{i=1}^n \lambda_i}. \quad (3.3.5)$$

Рассмотрим следующий частный случай: пусть $\lambda_1 = \ldots = \lambda_{n-1} = \lambda$, $\lambda_n = \lambda^{1-n}$. Выражение в правой части оценки (3.3.5) принимает вид:

$$\frac{\sqrt{n\sum_{i=1}^n \lambda_i^2} - n}{\sqrt{n\sum_{i=1}^n \lambda_i^2} - \sum_{i=1}^n \lambda_i} = \frac{\sqrt{n(n-1)\lambda^2 + n\lambda^{2-2n}} - n}{\sqrt{n(n-1)\lambda^2 + n\lambda^{2-2n}} - (n-1)\lambda - \lambda^{1-n}} = \\ = \frac{\sqrt{n(n-1) + n\lambda^{-2n}} - n\lambda^{-1}}{\sqrt{n(n-1) + n\lambda^{-2n}} - (n-1) - \lambda^{-n}}.$$

Устремляя $\lambda \to \infty$, получим предел

$$\frac{\sqrt{n(n-1)}}{\sqrt{n(n-1)} - (n-1)} = \frac{\sqrt{n}}{\sqrt{n} - \sqrt{n-1}} = n + \sqrt{n^2 - n}.$$

Для всякого λ должна выполняться оценка (3.3.5), поэтому при переходе к пределу в неравенстве получим

$$K \geqslant n + \sqrt{n^2 - n},$$

что и требовалось доказать. □

Таким образом, для линейных образов меры γ коэффициент K в неравенстве

$$\int_{\mathbb{R}^n} |T(x) - x|^2 \, \gamma(dx) \leqslant K \int_{\mathbb{R}^n} |T_0(x) - x|^2 \, \gamma(dx)$$

ограничен снизу величиной $n + \sqrt{n^2 - n}$, так что он неограниченно растет с ростом размерности n. Кроме того, если образ меры γ — центрированная гауссовская мера, ковариационная матрица которой имеет единичный определитель, то коэффициент K не может быть меньше $n^{1/2}$.

Это означает, что для линейных преобразований гауссовских мер на бесконечномерных пространствах функционал стоимости в задаче Монжа, взятый для треугольного отображения, вообще говоря, нельзя оценить через его минимальное значение ни с какой константой (задача Монжа в бесконечномерных пространствах исследуется в [9], [53], [54], [34], [39]).

Однако некоторые оценки все же можно получить. Пусть мера γ на пространстве $\mathbb{R}^\infty$ всех последовательностей есть счетная степень стандартной гауссовской меры на прямой, тогда $H = l^2$ — ее пространство Камерона–Мартина. Рассмотрим задачи Монжа и Канторовича для гауссовских мер μ, эквивалентных мере γ с функцией стоимости

$$|x-y|_H^2 = \sum_{i=1}^{\infty}(x_i - y_i)^2.$$

Известно (см. [9]), что для меры μ, эквивалентной мере γ, функция стоимости конечна и существует измеримый линейный оператор T_0, переводящий γ в μ и минимизирующий величину

$$\int |T_0 x - x|_H^2 \, \gamma(dx).$$

Рассмотрим типичный пример меры μ, эквивалентной мере γ: положим $\mu = \gamma \circ A^{-1}$, где $A = I + B$ и B — измеримый линейный оператор на пространстве H, значения которого порождаются оператором Гильберта–Шмидта $B_0\colon H \to H$, причем норма Гильберта–Шмидта оператора B_0 меньше 1. Оператор B имеет вид

$$Bx = \sum_{i=1}^{\infty} x_i B_0 e_i,$$

причем ряд сходится γ-почти всюду (это следует из условия, что B_0 — оператор Гильберта–Шмидта). Из доказанного для конечномерного случая, а также из известного метода построения треугольных отображений в бесконечномерных пространствах (см. [8]) получаем такое следствие.

Следствие 3.3.2. *Пусть норма Гильберта–Шмидта оператора B_0 не превосходит $1/2$. Тогда мера $\mu = \gamma \circ (I+B)^{-1}$ эквивалентна γ, имеет конечную энтропию, существует измеримое линейное треугольное отображение T, переводящее меру γ в меру μ, причем для оптимального отображения T_0 верна оценка*

$$\int |Tx - x|^2 \, \gamma(dx) \leqslant 25 \int |T_0 x - x|^2 \, \gamma(dx),$$

а также верны энтропийные неравенства Талаграна для интегралов от $|Tx - x|^2$ *и* $|T_0x - x|^2$.

Отметим, что если $|\mu_i - 1| \leqslant q$ при некотором $q < 1$, то оценка (3.3.2) верна с другой константой, зависящей от q.

ЗАКЛЮЧЕНИЕ

В диссертации исследованы некоторые виды преобразований мер в бесконечномерных пространствах, а также задачи Монжа и Канторовича для этих мер. Показано, что для широкого класса распределений диффузионных процессов с непостоянным коэффициентом диффузии в пространстве непрерывных на отрезке $[0, 1]$ функций соответствующие задачи Монжа и Канторовича с функцией стоимости, заданной нормой пространства Камерона–Мартина меры Винера, не имеют нетривиальных решений. Доказано, что если распределение диффузионного процесса из построенного класса переходит в меру, не совпадающую с ним, то функционалы стоимости как в задаче Монжа, так и в задаче Канторовича не принимают конечных значений.

Также в диссертации исследованы некоторые свойства треугольных отображений гауссовских мер в бесконечномерных пространствах, а именно произведено сравнение значений функционала стоимости в задаче Монжа для оптимального и канонического треугольного отображений, переводящих одну гауссовскую меру в другую, и показано, что в общем случае стоимость транспортировки для треугольного отображения не оценивается через оптимальную стоимость. С другой стороны, найдены эффективные условия, при которых имеется оценка с универсальной постоянной.

Дальнейшие исследования по тематике диссертации могут проводиться в следующих направлениях.

1. Исследование задач Монжа и Канторовича для распределений диффузионных процессов на пространстве непрерывных на отрезке $[0, 1]$ траекторий с другими функциями стоимости.

2. Исследование класса гауссовских мер на пространстве $\mathbb{R}^\infty$, для которых значения функционала стоимости в задаче Монжа для оптимального и канонического треугольного отображения сравнимы.

Литература

[1] Богачев В.И. Гауссовские меры. М.: Наука. – 1997. – 352 с.

[2] Богачев В.И. Основы теории меры. Т.1, 2, 2-е изд. – М. – Ижевск: РХД. – 2006.

[3] Богачев В.И. Дифференцируемые меры и исчисление Маллявэна. М. – Ижевск: РХД. – 2008. – 544 с.

[4] Богачев В.И. Слабая сходимость мер. М. – Ижевск: ИКИ. – 2016. – 396 с.

[5] Богачев В.И., Калинин А.Н., Попова С.Н. О равенстве значений в задачах Монжа и Канторовича// Зап. науч. семин. ПОМИ. – 2017. – Т. 457. – С. 53–73.

[6] Богачев В.И., Колесников А.В. Нелинейные преобразования выпуклых мер и энтропия плотностей Радона–Никодима// ДАН. – 2004. – Т. 397. – №2. – С. 155–159.

[7] Богачев В.И., Колесников А.В., Медведев К.В. О треугольных преобразованиях мер// ДАН. – 2004. – Т. 396. – №6. – С. 438–442.

[8] Богачев В.И., Колесников А.В., Медведев К.В. Треугольные преобразования мер// Матем. сб. – 2005. – Т. 196. – №3. – С. 3–30.

[9] Богачев В.И., Колесников А.В. Задача Монжа–Канторовича: достижения, связи и перспективы// УМН. – 2012. – Т. 67. – №5. – С. 3–110.

[10] Богачев В.И., Крылов Н.В., Рёкнер М., Шапошников С.В. Уравнения Фоккера–Планка–Колмогорова. М. – Ижевск: ИКИ. – 2014. – 592 с.

[11] Богачев В.И., Смолянов О.Г. Действительный и функциональный анализ: университетский курс. 2-е изд. М. – Ижевск: РХД, 2011. – 728 с.

[12] Вершик А.М. Несколько замечаний о бесконечномерных задачах линейного программирования// Успехи матем. наук. – 1970. – Т. 25. – № 5. – С. 117–124.

[13] Вершик А.М. Метрика Канторовича: начальная история и малоизвестные применения// Зап. научн. сем. ПОМИ. – 2004. – Т. 312. – С. 69–85.

[14] Вершик А.М., Затицкий П.Б., Петров Ф.В. Виртуальная непрерывность измеримых функций многих переменных и ее приложения// Успехи матем наук. – 2014. – Т. 69. – № 6. – С. 81–114.

[15] Вершик А.М., Затицкий П.Б., Петров Ф.В. Интегрирование виртуально непрерывных функций по бистохастическим мерам и формула следа ядерных операторов// Алгебра и анализ. – 2015. – Т. 27. – № 3. – С. 66–74.

[16] Гихман И.И., Скороход А.В. Теория случайных процессов. М: Наука, 1971. – 664 с.

[17] Канторович Л.В. О перемещении масс// ДАН СССР. – 1942. – Т. 37. – №7–8. – С. 227–229.

[18] Колесников А.В. Неравенства выпуклости и нелинейные преобразования мер// ДАН. – 2004. – Т. 396. – №3. – Р. 300–304.

[19] Левин В.Л. Двойственность Монжа – Канторовича и ее применение в теории полезности// Экономика и матем. методы. – 2011. – Т. 47. – № 4. – С. 1–40.

[20] Липчюс А.А. Замечание о равенстве в задачах Монжа и Канторовича// Теория вероятн. и ее примен. – 2005. – Т. 50. – № 4. – С. 779–782.

[21] Рачев С.Т. Задача Монжа – Канторовича о перемещении масс и ее применения в стохастике// Теория вероятн. и ее примен. – 1984. – Т. 29. – № 4. – С. 625–653.

[22] Судаков В.Н. Геометрические проблемы теории бесконечномерных вероятностных распределений// Тр. Мат. ин-та АН СССР. – 1976. – Т. 140. – С. 1–190.

[23] Ambrosio L., Gigli N. A user's guide to optimal transport// Lecture Notes in Math. – 2013. – Vol. 2062. – P. 1–155.

[24] Beiglböck M., Goldstern M., Maresch G., Schachermayer W. Optimal and better transport plans// J. Funct. Anal. – 2009. – Vol. 256. – № 6. – P. 1907–1927.

[25] Beiglböck M., Leonard C., Schachermayer W. A general duality theorem for the Monge–Kantorovich transport problem// Studia Math. – 2012. – Vol. 209. – №2. – P. 151–167.

[26] Beiglböck M., Leonard C., Schachermayer W. On the duality theory for the Monge–Kantorovich transport problem. In: Optimal transportation, pp. 216–265. London Math. Soc. Lecture Note Ser., V. 413. Cambridge Univ. Press, Cambridge, 2014.

[27] Bogachev V.I. Gaussian measures on infinite-dimensional spaces// In: Real and Stochastic Analysis: Current Trends (M.M. Rao ed.), pp. 1–83, World Sci., Singapore, 2014.

[28] Bogachev V.I. Differential properties of measures on infinite dimensional spaces and the Malliavin calculus// Acta Math. Univ. Carolinae, Math. et Phys. – 1989. – Vol. 30. – №2. – P. 9–30.

[29] Bogachev. V.I., Kolesnikov A.V. On the Monge–Ampére equation in infinite dimensions// Infin. Dimen. Anal. Quantum Probab. Related Top. – 2005. – Vol. 8. – №4. – P. 547–572.

[30] Bogachev. V.I., Kolesnikov A.V. Sobolev regularity for the Monge–Ampère equation in the Wiener space// Kyoto J. Math. – 2013. – Vol. 53. – №4. – P. 713–738.

[31] Cavalletti F. The Monge problem in Wiener space// Calc. Var. Partial Diff. Equat. – 2012. – Vol. 45. – №1-2. – P. 101–124.

[32] Chevallier E., Kalunga E., Angulo J. Kernel density estimation on spaces of Gaussian distributions and symmetric positive definite matrices// SIAM J. Imaging Sciences. – 2017. – Vol. 10. – №1. – P. 191–215.

[33] Fang S., Shao J., Sturm K.-T. Wasserstein space over the Wiener space// Probab. Theory Related Fields. – 2010. – Vol. 146. – №3–4. – P. 535–565.

[34] Feyel D., Üstünel S. Monge–Kantorovitch measure transportation and Monge–Ampére equation on Wiener space// Probab. Theory Related Fields. – 2004. – Vol. 128. – №3. – P. 347–385.

[35] Feyel D., Üstünel S. Solution of the Monge–Ampere equation on Wiener space for general log-concave measures// J. Funct. Anal. – 2006. – Vol. 232. – №1. – P. 29–55.

[36] Gangbo W., McCann R.J. The geometry of optimal transportation// Acta Math. – 1996. – №177. – P. 113–161.

[37] Givens C.R., Shortt R.M. A class of Wasserstein metrics for probability distributions// Michigan Math. J. – 1984. – Vol. 31 – №2. – P. 231–240.

[38] Horn R., Johnson C. Matrix analysis. Cambridge University Press, Cambridge, 1985. – xiii+561 p.

[39] Kolesnikov A.V. Convexity inequalities and optimal transport of infinite-dimensional measures// J. Math. Pures Appl. – 2004. – Vol. 83. – №11. – P. 1373–1404.

[40] Ledoux M. The concentration of measure phenomenon. Amer. Math. Soc., Providence, Rhode Island, 2001. – x+181 p.

[41] Lovrić M., Min-Oo M., Ruh E.A. Multivariate normal distributions parametrized as a Riemannian symmetric space// J. Multivariate Anal. – 2000. – Vol. 74. – №1. – P. 36–48.

[42] McCann R.J. Existence and uniqueness of monotone measure-preserving maps// Duke Math. J. – 1995. – Vol. 80. – P. 309–323.

[43] Modin K. Geometry of matrix decompositions seen through optimal transport and information geometry// J. Geom. Mech. – 2017. – Vol. 9. – №3. – P. 335–390.

[44] Pratelli A. On the equality between Monge's infimum and Kantorovich's minimum in optimal mass transportation// Ann. Inst. H. Poincaré (B) Probab. Statist. – 2007. – Vol. 43, № 1. – P. 1–13.

[45] Rachev S.T., Rüschendorf L. Mass transportation problems. Vol. I, II. Springer, 1998.

[46] Revuz D., Yor M. Continuous martingales and Brownian motion. Springer, Berlin, 1999. – xiv+602 pp.

[47] Skovgaard L.T. A Riemannian geometry of the multivariate normal model// Scand. J. Statist. – 1984. – Vol. 11. – P. 211–223.

[48] Takatsu A. On Wasserstein geometry of Gaussian measures// In: Probabilistic approach to geometry. Adv. Stud. Pure Math. – 2010. – Vol. 57. – P. 463–472.

[49] Takatsu A. Wasserstein geometry of Gaussian measures// Osaka J. Math. – 2011. – Vol. 48. – №4. – P. 1005–1026.

[50] Talagrand M. Transportation cost for Gaussian and other product measures// Geom. Funct. Anal. – 1996. – Vol. 6. – №3. – P. 587–600.

[51] Villani C. Topics in optimal transportation. Amer. Math. Soc., Providence, Rhode Island, 2003. – xvi+370 p.

[52] Villani C. Optimal transport, old and new. Springer, New York, 2009. – xxii+973 p.

Работы автора по теме диссертации:

Статьи в научных журналах Web of Science, SCOPUS, RSCI

[53] Bukin D.B. On the Monge and Kantorovich problems for distributions of diffusion processes// Math. Notes. – 2014. – Vol. 96. – №5-6. – P. 864–870 (импакт-фактор WoS 0.612).

[54] Букин Д.Б. О задаче Канторовича для нелинейных образов меры Винера// Матем. заметки. – 2016. – Т. 100. – №5. – С. 682–688 (импакт-фактор WoS 0.612).

Bukin D.B. On the Kantorovich problem for nonlinear images of the Wiener measure// Mathematical Notes. – 2016. – Vol. 100. – №5-6. P. 660–665.

[55] Bukin D.B., Krugova E.P. Transportation costs for optimal and triangular transformations of Gaussian measures// Theory of Stochastic Processes. – 2018. – Vol. 23. – № 2. – P. 21–32 (импакт-фактор SJR 0.12).

[56] Bukin D.B., Krugova E.P. On triangular mappings of Gaussian measures// Mathematical Notes. – 2019. – Vol. 106. – №5. – P. 843–845 (импакт-фактор WoS 0.612).

Тезисы докладов на научных конференциях

[57] Букин Д.Б. О задаче Канторовича для распределений диффузионных процессов// Сборник тезисов докладов международной научной конференции «Бесконечномерный анализ, стохастика, математическое моделирование: новые задачи и методы. Проблемы математического и естественнонаучного образования». Москва, 2014. – С. 32–33.

[58] Букин Д.Б. О задаче Канторовича для нелинейных образов меры Винера// Сборник тезисов докладов на международной научной конференции студентов, аспирантов и молодых ученых «Ломоносов-2016». М.: МГУ, 2016. 1 с.

[59] Bukin D.B. On the Monge and Kantorovich problems for nonlinear images of the Wiener measure// Сборник тезисов докладов на международной научной конференции «Бесконечномерный анализ и теория управления». М.: МГУ, 2018. 1 с.