

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертацию КОРОТИНА Александра Андреевича
«Параметрические методы вычисления оптимальных транспортных отображений,
расстояний и барицентров», представленную на соискание
ученой степени кандидата физико-математических наук по специальности 1.2.2 -
«Математическое моделирование, численные методы и комплексы программ»

Актуальность работы

Многие методы математического моделирования для решения задач машинного обучения и обработки изображений зачастую имеют эвристический характер. Среди них можно выделить некоторые методы генеративного моделирования (генеративные состязательные сети и пр.), которые, несмотря на популярность и хорошую производительность в практических задачах до сих пор имеют ограниченное теоретическое обоснование; результат их работы зачастую сложно интерпретировать и контролировать, так как механизмы обучения таких моделей состоят из ряда базовых элементов (генератор, дискриминатор, функция потерь), настройка которых требует существенных трудозатрат. При этом высокие вычислительные затраты имеет и сам процесс обучения таких моделей.

Тематика описываемой работы — численные методы оптимального транспорта и их приложения в задачах математического моделирования. Оптимальный транспорт (ОТ) — это общий термин для описания набора взаимосвязанных математических задач поиска эффективного перемещения массы между вероятностными распределениями данных. Ключевой смысл применения инструментария ОТ к задачам моделирования заключается в том, что решения этих задач, полученные с помощью ОТ, хорошо интерпретируемы. Например, оптимальные транспортные отображения могут быть применены в задачах синтеза искусственных данных из других данных, так как имеют интуитивно понятную структуру, поскольку стараются максимально сохранить входной объект ввиду самой постановки задачи ОТ. Центры масс ОТ представляют собой геометрический осмысленный способ усреднения распределений данных и т.д. Иными словами, применяя ОТ на вопрос «что ожидать от модели» зачастую можно получить простой и понятный ответ.

Теоретические свойства ОТ (существование и характеристика решений, двойственные постановки и т.д.) детально изучены в релевантной теоретической литературе (начиная с ранних работ Л. Канторовича в 20 веке). Благодаря этому, существует видимый потенциал для разработки эффективных численных методов ОТ, которые в применении к задачам моделирования могут превзойти существующие подходы для этих задач. Методы ОТ уже сейчас уже находят применение в задачах обработки изображений, доменной адаптации, синтеза данных и пр. Большая часть существующих методов ОТ рассматривают

маломерные дискретные распределения данных и плохо масштабируются, например, с ростом размера выборки или размерности данных. Применение нейронных сетей - естественный способ исправить этот недостаток, масштабировать численные методы ОТ, и, в частности, расширить класс практических задач к которым инструментарий ОТ может быть применим.

Подводя итог можно уверенно утверждать, что рассматриваемая тема диссертации «Параметрические методы вычисления оптимальных транспортных отображений, расстояний и барицентров» является актуальной и перспективной. Разработка новых нейросетевых методов ОТ позволит строить новые математические модели данных и разрабатывать более эффективные и теоретические обоснованные численные методы для решения прикладных задач с помощью инструментария ОТ.

Характеристика работы

Введение содержит общую характеристику работы и описывает актуальность темы, цели и задачи, методологию исследования, значимость, новизну, достоверность, публикации, личный вклад автора и положения, выносимые на защиту.

Первая глава представляет собой краткое, но достаточное для понимания описание постановок рассматриваемых в работе математических задач и недостатков существующих численных методов для их решения. Вводятся задачи ОТ (с квадратичной функцией стоимости), поиска среднего по ОТ, градиентных потоков функционалов на метрическом пространстве Вассерштейн-2, бенчмаркинга (сравнения) методов ОТ. Детально каждая из задач рассматривается и решается в последующих главах работы.

Вторая глава рассматривает основную задачу вычисления ОТ отображений между непрерывными вероятностными распределениями, доступными через случайные эмпирические выборки. Рассматривается двойственная задача и описываются проблемы существующих подходов к её численному решению на практике. Для устранения этих проблем, предлагается новая регуляризация («циклическая состоятельность»), позволяющая сформулировать неминимаксный целевой функционал для оптимизации. Приводится теоретическое обоснование предложенного функционала (связывающее ошибку решения регуляризованной двойственной задачи и восстановления решения прямой задачи), а затем предлагается оптимизировать полученный функционал с помощью стохастического градиентного спуска с использованием выпуклых по входу нейросетей. Предложенный подход тестируется на ряде как модельных задач с данными малой размерности, так и в многомерных задачах обработки изображений.

В третьей главе предлагается нейросетевой алгоритм вычисления барицентров. Глава структурирована аналогично главе 2 — приводится двойственная задача барицентра, предлагается подход (на основе регуляризации), позволяющий сформулировать неминимаксный целевой функционал для оптимизации. Для функционала доказывается теоретический результат, по смыслу аналогичный результату главы 2, но для задачи вычисления барицентра. Предложенный подход эмпирически тестируется на ряде примеров и показывается превосходство над существующими методами.

Четвертая глава посвящена моделированию Васерштейн-2 градиентных потоков функционалов на пространстве вероятностных распределений. Предлагается и тестируется алгоритм для их моделирования на основе выпуклых нейронных сетей. В главе преимущественно рассматриваются градиентные потоки функционала свободной энергии Фоккера-Планка, но указано, что предлагаемый алгоритм может быть адаптирован и на случай других функционалов. Эмпирически показывается заметное превосходство над существующими алгоритмами моделирования градиентных потоков.

В пятой главе рассматривается вопрос тестирования качества параметрических численных методов ОТ. Объясняется, почему его сложно объективно оценивать и предлагается методология того, как это можно делать. Для этого в главе разрабатывается новый теоретически обоснованный подход к синтезу искусственных данных (эталонных пар непрерывных распределений), которые могут быть использованы для тестирования методов ОТ. Синтезируется ряд датасетов, которые далее используются для тестирования существующих методов ОТ. Один из выводов главы заключается в том, что некоторые из существующих методов ОТ по факту вообще сложно назвать «методами ОТ», т.к. они показывают колоссальную ошибку в вычислении ОТ.

Заключение подводит итоги полученных результатов и формулирует направления дальнейших исследований.

Характеристика результатов (новизна, достоверность и обоснованность)

Новизна работы заключается в том, что предложенные новые алгоритмы ОТ используют нейронные сети, имеют неминимаксный целевой критерий оптимизации, который не вносит смещения в решение задачи. Это позволяет масштабировать применение методов ОТ на пространства большой размерности с большими выборками данных.

Важно отметить, что в работе нейронные сети применяются не потому что «так можно», а потому, что это действительно (как показано в работе) позволяет масштабировать

численные методы ОТ и применить их к реальным задачам моделирования, решать которые без применения нейронных сетей затруднительно. Примером такой задачи, например, может служить задача непарного переноса стиля между наборами изображений в разделе 2.6.7. С другой стороны, в нескольких местах работы явно показано превосходство разработанных нейросетевых методов над существующими методами, не использующими нейросети, например, в разделах 3.6.1 (вычисления барицентров) или 4.4.2 (моделирование градиентных потоков).

В то же время, предложенные подходы для регуляризации двойственных задач ОТ (главы 2-3), являются не эвристикой, а представляют собой строго математически обоснованную автором конструкцию. Данная конструкция позволяет избавиться от необходимости решать минимаксную задачу поиска седловой точки и свести решение задач ОТ к чистой минимизации, что упрощает использование методов на практике. В частности, предложенный подход (за счет избавления от минимаксности) упрощает теоретический анализ качества решений, полученных в результате оптимизации. В общем, разработанные численные методы ОТ, в отличие от многих существующих нейросетевых методов математического моделирования, позволяют не просто «применять нейросети для эффективного решения задач на практике», но и анализировать их теоретические свойства решения, полученных с их помощью.

Отдельно стоит отметить новизну и значимость результатов главы 5. В любой научной области для прозрачности, воспроизводимости и достоверности получаемых результатов, требуется наличие бенчмарков (наборов тестов, датасетов) и метрик для проверки функционирования разрабатываемых методов. Поскольку нейросетевые численные методы ОТ и их приложения являются относительно новой областью исследования, таких бенчмарков в области пока нет. Фактически глава 5 данной работы решает эту проблему и предлагает первый такой бенчмарк и метрики. Это позволит внести ясность и прозрачность в исследования области, а также осуществлять сравнения будущих численных методов в области.

Утверждения и результаты, приводимые в данной работе достаточно обоснованы. Теоретические результаты (предложенные регуляризации, алгоритмы и соответствующие теоремы) математически корректны и к ним приведены строгие доказательства. Экспериментальные результаты (тестирование методов на модельных задачах) детально описаны и содержат необходимые подробности для их воспроизведения; произведено экспериментальное сравнение с многими существующими методами в области; программный код для воспроизведения экспериментов выложен автором в общий доступ.

Подводя итог, предложенные в данной диссертации численные методы ОТ и математические подходы, на которых они основаны, новы, достоверны и обоснованы. Они

позволяют масштабировать применения методов ОТ на пространства большей размерности и большие выборки данных. В то же время, предлагаемые методы имеют потенциал расширить класс задач, которым применимы методы ОТ. Так, например, метод вычисления градиентных потоков (глава 4 работы) потенциально может быть адаптирован для применения к задачам генеративного моделирования, аналогично тому, как это делают современные модели на основе диффузий. Одно из ограничений данной работы заключается в том, что построенные алгоритмы работают только для квадратичной функции стоимости транспорта. Расширение предлагаемой автором методологии решения задач ОТ на случай общей функции цены является перспективным направлением исследования, которое, я надеюсь, будет рассмотрено автором в будущем как логичное продолжение данной работы.

Замечания по работе

1. Понятие Beta-гладкости функции впервые используется в теоремах главы 2 работы, а формально вводится только в главе 3 в теореме 3.6.3.
2. В работе автор оперирует с градиентами выпуклых функций, неявно предполагая, что они существуют. Для прозрачности автору следовало сослаться на классические результаты выпуклого анализа, объясняющие почти всюду дифференцируемость выпуклых функций.
3. При определении градиентных потоков (главы 1, 4) в работе не упоминаются вопросы их существования или единственности. Хотя эти вопросы обсуждаются в приведенных в работе литературных источниках, автору стоило бы обсудить их непосредственно в работе.
4. В предложенных алгоритмах (главы 2, 3, 4) требуются специфические архитектуры нейронных сетей (выпуклые по входу). Такие архитектуры за счет встроенных ограничений могут быть субоптимальными при решении задач моделирования. Это обсуждается в работе в главе 5, но было бы осмысленно упомянуть об этом ранее, например, в главе 2, где представлены первые эксперименты с такими архитектурами.
5. В разделе 2.6.7 (перенос стилей между наборами изображений) при использовании предложенного метода ОТ фактически используется попиксельная L2 функция как метрика близости входного и выходного изображения. Такая метрика которая может не точно показывать семантическую близость изображений. Здесь интересно было бы осуществлять ОТ в скрытом пространстве автокодировщика (аналогично разделу 2.6.6), т.к. L2 между латентными кодами представляется более осмысленной метрикой близости изображений.
6. В разделах 2.6.4 и 3.6.3 (перенос цвета между двумя/тремя изображениями) ОТ осуществляется на цветовых RGB палитрах 3D изображений. Хотелось бы увидеть, как метод будет работать в случае, когда рассматриваются не изображения, а

глубокие карты признаков, например, полученные предобученной VGG-16 нейронной сетью.

Заключение

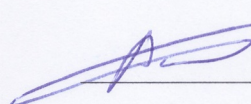
Приведенные выше замечания не влияют на общую положительную оценку работы и скорее носят методический характер. Диссертация А.А. Коротина выполнена на высоком уровне и является законченной научно-квалификационной работой, в которой предложены и теоретически обоснованы эффективные численные методы оптимального транспорта на основе нейронных сетей и приведены экспериментальные результаты подтверждающие их функционирование в ряде модельных задач, включая ряд задач обработки изображений.

С учетом вышесказанного, диссертация А.А. Коротина «Параметрические методы вычисления оптимальных транспортных отображений, расстояний и барицентров» полностью соответствует требованиям «Положения о присуждении ученых степеней», предъявляемых к кандидатским диссертациям. Считаю, что А.А. Коротин заслуживает присуждения ему ученой степени кандидата физико-математических наук по специальности 1.2.2 «Математическое моделирование, численные методы и комплексы программ».

Официальный оппонент

кандидат физико-математических наук
(специальность 05.13.11 - «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей»),
старший научный сотрудник, руководитель лаборатории
«Компьютерной графики и мультимедиа»
Факультета Вычислительной Математики и Кибернетики
Московского Государственного Университета им. Ломоносова
Ватолин Дмитрий Сергеевич

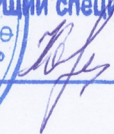
« 28 » декабря 2022 г.

 / Ватолин Д. С.

Адрес: 119991, Российская Федерация, Москва, Ленинские горы, д. 1
Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет имени М.В.Ломоносова»
Тел. +7 (495) 939-10-00.
E-mail: info@rector.msu.ru
Сайт <https://www.msu.ru>



Подпись удостоверяю
Ведущий специалист по кадрам

 Т.Г. Коваленко