

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«Московский физико-технический институт
(национальный исследовательский университет)»
(МФТИ, Физтех)**

На правах рукописи

Плетнев Никита Вячеславович

**Применение градиентных методов оптимизации для решения
некоторых обратных задач математической физики**

**Специальность: 1.2.2. Математическое моделирование,
численные методы и комплексы программ**

**ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата физико-математических наук**

Научный руководитель:
Гасников Александр Владимирович,
доктор физико-математических наук,
профессор, член-корреспондент РАН

Долгопрудный – 2025

Оглавление

Введение	5
1 Общая характеристика и примеры задач	16
1.1 Постановка и общие свойства	16
1.2 Задачи, решаемые в работе	17
2 Используемые методы оптимизации	20
2.1 Подходы к выбору длины шага	20
2.1.1 Выбор длины шага по принципу минимизации функционала	20
2.1.2 Выбор длины шага по принципу минимизации расстояния до точки минимума	21
2.2 Существующие методы	22
2.2.1 Метод градиентного спуска	22
2.2.2 Метод сопряжённых градиентов	24
2.2.3 Метод подобных треугольников	25
2.2.4 Метод минорант Поляка	26
2.2.5 Адаптивный тяжёлый шарик Поляка	26
2.3 Новые предлагаемые методы	27
2.3.1 Метод покомпонентного спуска	27
2.3.2 Моментный метод минимальных ошибок	28
2.3.3 Многошаговый метод минимальных ошибок	31
2.3.4 Применение при составном функционале	35
2.3.5 Теоремы о сходимости построенных методов	36
2.3.6 О применении построенных методов в условиях неточных вычислений	45
3 Применение к обратным задачам математической физики	47
3.1 Начально-краевая задача для уравнения Гельмгольца	47

3.1.1	Происхождение, постановка и некорректность задачи	47
3.1.2	Операторы задачи	48
3.1.3	Спектр соответствующего задаче самосопряжённого оператора	53
3.1.4	Приближённые вычисления	54
3.1.5	Эксперименты	57
3.1.6	Выводы	60
3.2	Ретроспективная задача Коши для уравнения теплопроводности в одномерном пространстве	61
3.2.1	Происхождение, постановка и некорректность задачи	61
3.2.2	Операторы задач	62
3.2.3	Спектр соответствующего задаче самосопряжённого оператора	67
3.2.4	Приближённые вычисления	68
3.2.5	Эксперименты	70
3.2.6	Выводы	80
3.3	Ретроспективная задача Коши для уравнения теплопроводности в трёхмерном пространстве	81
3.3.1	Происхождение, постановка и некорректность задачи	81
3.3.2	Операторы задач	82
3.3.3	Спектр соответствующего задаче самосопряжённого оператора	84
3.3.4	Приближённые вычисления	84
3.3.5	Эксперименты	86
3.3.6	Выводы	94
3.4	Обратная задача термоакустики	95
3.4.1	Происхождение, постановка и некорректность задачи	96
3.4.2	Операторы задач	97
3.4.3	Приближённые вычисления	99
3.4.4	Эксперименты	100
3.4.5	Выводы	116
3.5	Интегральное уравнение Фредгольма 1-го рода	117
3.5.1	Происхождение, постановка и некорректность задачи	117
3.5.2	Операторы задачи	118
3.5.3	Эксперименты	119
3.5.4	Выводы	130
3.6	Об использовании построенных методов при неизвестном точном решении	132

Заключение	134
Список литературы	137
Список рисунков	143
Список таблиц	146
Список сокращений	148

Введение

Актуальность темы исследования.

Теория обратных и некорректных задач — очень обширная и бурно развивающаяся область математической физики. Понятие некорректной задачи было введено в 1900 году Ж. Адамаром, и он считал, что физический смысл может иметь лишь корректно поставленная задача. Однако оказалось, что некорректно поставленные задачи часто встречаются в приложениях: физике, технике, астрономии, геофизике.

В частности, пример некорректной задачи, придуманный Ж. Адамаром, возник в работах А. Н. Тихонова при изучении электромагнитного поля земли (пример такой работы — статья [41]).

Интересна взаимосвязь между решением некорректных задач и регуляризацией расходимостей различного происхождения, описанная в работе [18].

В монографии [53] описаны разнообразные обратные и некорректно поставленные задачи, известные в настоящий момент. Среди них есть интегральные уравнения, а также задачи для дифференциальных уравнений в частных производных эллиптического, параболического и гиперболического типов. Некоторые из таких задач решаются в данной диссертации.

Часто некорректные задачи можно свести к операторному уравнению. Первоначально их решали методом подбора (согласно [44]): решали прямые задачи для определённого подкласса возможных решений и выбирали среди них наилучшее по величине некоторой метрики.

Идея решения уравнений путём минимизации квадрата невязки восходит ещё к работам А. М. Лежандра и К. Ф. Гаусса. Простейший метод градиентного спуска для решения систем линейных алгебраических уравнений был предложен О. Л. Коши, и эта идея стала основой для построения многообразных методов оптимизации.

Преимущество методов первого порядка, использующих только значение функции и её градиент, заключается в том, что вычислительная сложность их реализации относительно невысока. Их применение к решению обратных задач

оказывается возможным, благодаря существованию производной Фреше у соответствующего функционала. Методы вычисления градиента изложены, например, в книгах [4] и [12].

В настоящее время большой популярностью пользуются ускоренные (то есть, моментные) методы. При одинаковой сложности одной итерации, они сходятся намного быстрее метода градиентного спуска. Это позволяет эффективно применять их к решению обратных задач, что и выполнено в данной работе.

Чаще всего обсуждается вопрос о сходимости методов по функционалу. Однако сходимость по аргументу важнее, исходя из следующих соображений. Для сильно выпуклых задач все три вида сходимости (по аргументу, по функционалу, по норме градиента) эквивалентны. Для выпуклых, но не сильно выпуклых задач в бесконечномерном пространстве это не так: из сходимости по аргументу следует сходимость по функционалу (а из неё — по норме градиента), но не наоборот. Напротив, обязательно существует последовательность точек, вдоль которой функционал стремится к нулю, а расстояние до точного решения бесконечно возрастает. Поэтому чрезвычайно важно доказательство, что методы оптимизации порождают не такую последовательность, а сходящуюся к точному решению (хотя бы при точных вычислениях). Это доказательство также получено в диссертации.

Подобная теорема доказывалась и в книге [53], но с использованием других идей. Метод доказательства, применённый в работе, позволил доказать и другую важную теорему — об отсутствии равномерной сходимости по аргументу для широкого класса градиентных методов.

Возникающие при решении обратных задач функционалы часто являются квадратичными. Этим знанием структуры можно воспользоваться для построения более эффективных (пусть и не допускающих обобщения на другие задачи оптимизации) методов первого порядка. В работе выполняется такое построение с использованием вспомогательной минимизации расстояния до точного решения. Полученные методы применяются к решению обратных задач и сравниваются с классическими методами.

Степень разработанности темы исследования.

Теория некорректно поставленных и обратных задач восходит к работам А. Н. Тихонова (например, [41], [42], [43]) и М. М. Лаврентьева (например, [24], [25], [26], [27], [28], [29]). В них исследуются разнообразные обратные задачи, возникающие из практических приложений — например, теории оптимального управления.

Очень важно для исследований понятие условной корректности (корректности по Тихонову). Она обозначает непрерывную зависимость решения от начальных данных не на всём пространстве, а на некотором его подмножестве.

Огромное количество интересных работ посвящены решению некорректных задач с помощью градиентных методов оптимизации. Например, в статье [50] А. В. Гасниковым, С. И. Кабанихиным, М. А. Шишлениным и А. Mohammed предложен следующий подход к решению обратных задач: ставится эквивалентная задача оптимизации; вместо построения аппроксимирующей задачи, допускающей точное решение, выполняется минимизация в гильбертовом пространстве, при этом градиент вычисляется приближённо.

Как уже было сказано, некорректно поставленные и обратные задачи часто встречаются в приложениях. Приведём примеры работ, в которых с помощью методов оптимизации решаются разнообразные обратные задачи из техники, медицины, астрофизики, геофизики.

В книге [12] Ю. Г. Евтушенко приводит пример применения численной оптимизации в решении конкретной технической задачи выбора наилучшего контура проектируемого самолёта (по заказу КБ Сухого). Была составлена математическая модель динамики самолёта, вычислены градиенты и проведена процедура минимизации функции Понтрягина. Интересно, что это оказалось возможным, несмотря на отсутствие гарантii выпуклости задачи оптимизации.

Статья М. А. Шишленина, Н. А. Савченко, Н. С. Новикова и Д. В. Ключинского [56] содержит решение обратной коэффициентной задачи гиперболического типа первого порядка. В ней с помощью оптимизационного алгоритма восстанавливается коэффициент акустического затухания по дополнительной информации о волновом поле в ряде приёмников.

Статья С. И. Кабанихина и М. А. Шишленина [21] посвящена решению обратной задачи определения старшего коэффициента в параболическом уравнении диффузии с использованием нелокальной дополнительной информации — интеграла решения уравнения по пространственной координате. В ней проведены численные эксперименты, показывающие преимущество оптимизационного подхода к решению обратных задач перед использовавшимся ранее эволюционным алгоритмом.

В статьях [1], [9], [14] (А. Ф. Албу, А. Ю. Горчаков, В. И. Зубов), [20] (С. И. Кабанихин, А. Х. Хасанов, А. В. Пененко) решаются разные варианты обратной задачи для уравнения теплопроводности, имеющего параболический тип. В них

восстанавливается входящий в уравнение переменный коэффициент теплопроводности. Целевой функционал — среднеквадратичное отклонение рассчитываемых значений теплового потока (или температурного поля) от наблюдаемых. Градиент вычисляется с помощью техники быстрого автоматического дифференцирования, изложенной в [12], что позволяет применить методы оптимизации первого порядка. Результаты экспериментов, отражённые в этих статьях, показывают эффективность данного подхода к решению обратной коэффициентной задачи. При этом в статьях [1], [9] и [14] искомый коэффициент зависит от температуры, поэтому задача является нелинейной, и нет гарантий выпуклости функционала.

Ретроспективные обратные задачи имеют большое значение для самых разных областей науки. Например, в статье [19] С. И. Кабанихиным, И. М. Куликовым и М. А. Шишлениным построен алгоритм восстановления начального состояния сверхновой звезды путём решения обратной задачи с помощью градиентного метода оптимизации. А В. Г. Романов и П. С. Мошканев в статье [36] используют решение обратной задачи для определения источника цунами. Статья [5] (В. В. Васин и др.) посвящена решению обратных задач гравиметрии и магнитометрии.

С другой стороны, интенсивно развиваются градиентные методы оптимизации. Подробному изложению современного состояния этой области математики посвящено, например, пособие А. В. Гасникова [7]. К примеру, в статье [57] А. Васиным и соавторами предложен ускоренный метод подобных треугольников (STM) с фиксированными шагами, для которого получены оптимальные оценки сходимости по функционалу в условиях неточного вычисления градиента. Решение обратных задач, как указано в этой статье, является естественным приложением градиентных методов, поскольку связанные с такими задачами функционалы часто являются выпуклыми и квадратичными, а приближённое вычисление градиента оказывается возможным путём решения прямых задач.

В статьях [2] (Г. М. Вайникко) и [31] (А. С. Немировский) проведён анализ сходимости методов градиентного спуска и сопряжённых градиентов в применении к решению операторных уравнений при неточном вычислении оператора, сформулированы правила останова.

Статьи [35] (Б. Т. Поляк) и [49] (N. Devanathan, S. Boyd) посвящены методу минорант Поляка. Его смысл заключается в использовании известного значения минимума и выполнении каждого шага путём приведения к целевому значению

локального нижнего приближения («миноранты») оптимизируемого функционала. Формулы, получаемые при реализации такого подхода, похожи на те, которые возникают при вспомогательной минимизации расстояния до точного решения; похожи и некоторые свойства. В данных работах доказана сходимость метода, теоретические результаты подтверждены численными экспериментами.

Статья [51] (B. Goujaud, A. Taylor, A. Dieuleveut) содержит попытку построить метод, соединяющий в себе метод тяжёлого шарика и шаг Поляка. Он должен минимизировать расстояние до точного решения на каждом шаге. Однако построение в статье не привело к успеху, что подтверждают и эксперименты, и теоретические выкладки.

Вклад настоящей работы заключается в построении метода первого порядка, выполняющего поставленную задачу — достижение на каждом шаге минимально возможного расстояния до точного решения. Под расстоянием понимается обычна для гильбертовых функциональных пространств метрика — среднеквадратичное отклонение. Также доказаны его свойства и проведена экспериментальная проверка полученных результатов.

Для проверки построенного метода и его сравнения с ранее существовавшими методами используются следующие некорректные и обратные задачи: начально-краевая задача для уравнения Гельмгольца, ретроспективная задача Коши для уравнения теплопроводности (в одномерном и трёхмерном пространстве), обратная задача термоакустики, интегральное уравнение Фредгольма 1-го рода.

Рассмотрим работы, посвящённые решению этих и близких к ним задач.

Например, в статье [52] (С. И. Кабанихин, М. А. Шишленин и соавторы) решалась начально-краевая задача для уравнения Гельмгольца. Это уравнение эллиптического типа. В статье вычислен градиент целевого функционала, проведены эксперименты по его минимизации с помощью метода градиентного спуска. Полученный алгоритм сравнивается с регуляризациями А. Н. Тихонова и С. К. Годунова. Выяснилось, что методы регуляризации, применённые к дискретизованной задаче, более эффективны.

В статье [39] (К. Б. Сабиров, Н. В. Мартемьянова) исследуются обратные задачи для эллиптических уравнений Лапласа и Гельмгольца. При их решении использован метод Фурье. Полученные представления в виде тригонометрических рядов позволили доказать существование и единственность решений при условии достаточной гладкости.

Статья [38] (К. Б. Сабиров) посвящена постановке нескольких обратных

задач для уравнения Гельмгольца с интегральным дополнительным условием. Также применены ряды Фурье для доказательства существования и единственности решения.

В данной диссертации метод Фурье использован для доказательства разрешимости и условной корректности начально-краевой задачи для уравнения Гельмгольца, а также для вычисления спектра связанного с задачей самосопряжённого оператора. Проведены численные эксперименты, показавшие преимущество нового построенного метода в применении к решению данной задачи. При интерпретации полученных результатов использован вычисленный спектр.

Статья [15] (В. К. Иванов) посвящена теоретическому решению ретроспективной задачи Коши для уравнения теплопроводности на числовой оси. Использованы регуляризация и преобразование Фурье.

В статье [13] (М. С. Еремеева) проведено экспериментальное сравнение различных итерационных методов решения ретроспективной задачи Коши для уравнения теплопроводности на отрезке. Из всех рассмотренных там методов наилучшим оказался метод сопряжённых градиентов. В статье [40] (А. А. Самарский, П. Н. Вабищевич, В. И. Васильев) кратко освещён тот же вопрос.

В статье [3] ретроспективная задача Коши для уравнения теплопроводности сводится к интегральному уравнению Фредгольма 1-ого рода с использованием интеграла Пуассона, а интегральное уравнение решается приближённо путём дискретизации и сведения к системе линейных алгебраических уравнений.

В данной диссертации ретроспективная задача Коши для уравнения теплопроводности в одномерном и трёхмерном случаях используется для проверки сходимости градиентных методов оптимизации. Оказывается, что при высокой точности вычислений построенный в работе t -моментный метод минимальных ошибок позволяет достичь лучших результатов; при ухудшении точности он достигает примерно такой же невязки, как и метод сопряжённых градиентов, но требует приблизительно вдвое меньших вычислительных затрат.

Статья [22] (С. И. Кабанихин, М. А. Шишленин, О. И. Криворотко) посвящена решению обратной задачи термоакустики, уравнение в которой имеет гиперболический тип. Поставлены три варианта задачи, отличающиеся количеством дополнительных условий. Для всех сформулированных задач выписаны функционалы и вычислены их градиенты. Для решения использованы метод градиентного спуска и метод сопряжённых градиентов. Показано, что ме-

тод сопряжённых градиентов значительно быстрее приближается к точному решению. Качество полученного решения оказалось достаточным для обнаружения неоднородностей в практических приложениях.

В диссертации применяется новый построенный метод для минимизации функционалов, возникающих при решении обратной задачи термоакустики. Это позволяет улучшить качество решения задачи.

В работе В. К. Иванова [16] представлен метод решения обратной задачи потенциала, которая сводится к интегральному уравнению Фредгольма 1-ого рода. Строится последовательность приближённых решений, лежащих в конечномерных подпространствах частичных сумм рядов Фурье, и доказывается её равномерная сходимость.

В статье [11] (Т. Ф. Долгополова, В. К. Иванов) исследуется задача численного дифференцирования и строится алгоритм её приближённого решения, основанный на взаимосвязи с обратной задачей для интегрального уравнения. Решению интегральных уравнений Фредгольма 1-ого рода с помощью регуляризации посвящена также статья В. К. Иванова [17]. В статье [6] (В. В. Васин, Т. И. Сережникова) решается похожее уравнение Фредгольма-Стилтьеса.

Статья [30] (Д. В. Лукьяненко, А. Г. Ягола) посвящена решению многомерного уравнения Фредгольма 1-ого рода, возникающего при решении задачи определения намагниченности по внешним измерениям. Используются регуляризация Тихонова и метод сопряжённых градиентов. Отмечено, что вычисление функционала и градиента допускает распараллеливание, что позволяет уменьшить время работы.

Данная диссертация содержит численные эксперименты, в которых интегральное уравнение Фредгольма 1-ого рода численно решается с помощью градиентных методов оптимизации. Показано, что новый построенный метод позволяет достичь наилучшего качества решения, и притом быстрее метода сопряжённых градиентов.

Цели исследования.

1. Получить эффективные методы оптимизации для решения некорректных и обратных задач.
2. Проверить работу методов оптимизации в применении к разнообразным задачам.
3. Выяснить зависимость качества решения от свойств задачи, например спектра связанного с задачей оператора.

Задачи исследования.

1. Построить методы оптимизации первого порядка, основанные на вспомогательной минимизации расстояния до точного решения.
2. Сформулировать и доказать оценки сходимости новых методов в бесконечномерном пространстве.
3. Экспериментально проверить построенные методы: применить их к решению обратных задач, сравнить с ранее существующими методами.
4. Определить влияние характеристик задачи на эффективность применения методов оптимизации.

Научная новизна.

1. Получен новый метод оптимизации для решения квадратичных задач, возникающих при решении уравнений (в том числе операторных), — это t -моментный метод минимальных ошибок.
2. Доказаны сходимость по аргументу и оптимальность построенного метода с $t = \infty$ при условии точности всех вычислений. Также доказано, что для некорректной задачи сходимость по аргументу может быть сколь угодно медленной в зависимости от начального приближения.
3. Проведены эксперименты, сравнивающие t -моментный метод минимальных ошибок с ранее существующими методами при решении разнообразных некорректных задач математической физики.

Теоретическая и практическая значимость работы.

1. Теоретические результаты: построен t -моментный метод минимальных ошибок, пригодный для решения квадратичных оптимизационных задач с известным минимальным значением, и показано, что при $t = \infty$ он является оптимальным среди методов, работающих в подпространствах Крылова. Доказана его сходимость по аргументу в любом пространстве (в n -мерном за $\leq n$ шагов; для сильно выпуклой задачи — со скоростью не медленнее геометрической прогрессии; для не сильно выпуклой — сам факт сходимости). При этом установлено, что для некорректной задачи сходимость по аргументу может оказаться сколь угодно медленной.

2. Экспериментальные результаты: для таких задач, как начально-краевая задача для уравнения Гельмгольца, ретроспективная задача Коши для уравнения теплопроводности (в одномерном и трёхмерном пространстве), обратная задача термоакустики, интегральное уравнение Фредгольма 1-го рода, проведено сравнение t -моментного метода минимальных ошибок, разных вариантов метода сопряжённых градиентов, градиентного спуска с шагом Поляка, адаптивного тяжёлого шарика и метода подобных треугольников. Теоретические результаты подтверждены на практике. В задачах, для которых достаточная точность вычислений обеспечивается, заметно превосходство нового метода. В других задачах достигаемое им качество решения сопоставимо с результатами работы метода сопряжённых градиентов.

Методология и методы исследования.

Используется общая постановка задачи в виде операторного уравнения из [50] и [53]. Сначала проводятся вычисления длины шага из соображений одномерной минимизации расстояния до точного решения для произвольного направления спуска. Потом выполняется аналогичная вспомогательная двумерная минимизация, результаты которой обобщаются на случай, когда рассматриваются t предыдущих шагов, с использованием метода Гаусса для решения системы линейных алгебраических уравнений. Корректность такого построения доказывается при помощи метода математической индукции.

Анализ сходимости полученного метода проводится с использованием разложения элементов гильбертова пространства по базису из собственных функций связанного с задачей самосопряжённого оператора. Существование такого базиса обусловлено общими свойствами компактных операторов, изложенными, в частности, в учебном пособии А. Н. Колмогорова и С. В. Фомина [23].

Для экспериментального сравнения все рассматриваемые методы оптимизации были реализованы на Python. Также реализованы приближённые оракулы первого порядка для всех задач оптимизации, к которым сводятся изучаемые некорректные задачи.

Градиенты функционалов вычисляются как решения двойственных задач с использованием техники, которая описана, например, в пособии Ф. П. Васильева [4].

Возникающие при вычислении функционалов и градиентов корректные задачи для дифференциальных уравнений в частных производных решаются приближённо

с помощью устойчивых разностных схем первого и второго порядка аппроксимации, теория которых изложена, например, в учебном пособии В. С. Рябенького [37].

Положения, выносимые на защиту.

1. *t*-моментный метод минимальных ошибок для решения квадратичных задач оптимизации с известным минимальным значением.
2. Теоремы об оптимальности и сходимости метода; теорема о существовании начального приближения, обеспечивающего сколь угодно медленную сходимость градиентных методов.
3. Многочисленные эксперименты, показывающие эффективность применения градиентных методов оптимизации (и, в особенности, построенного *t*-моментного метода минимальных ошибок) к решению некорректных задач.

Степень достоверности и апробация результатов.

Достоверность результатов работы обеспечивается математическими доказательствами теорем, корректным проведением математического и компьютерного моделирования.

Результаты диссертации опубликованы в 3 статьях: [32], [33] и [47]. 3 из них — в рецензируемых журналах «Компьютерные исследования и моделирование» (работы [32], [33]) и «Журнал вычислительной математики и математической физики» (работа [47]), рекомендованных для публикации результатов диссертаций в МФТИ. Статья [34] принята к публикации в «Журнал вычислительной математики и математической физики».

Результаты работы были обсуждены и получили одобрение специалистов на следующих конференциях и семинарах:

1. «Quasilinear Equations, Inverse Problems and their Applications» (2022, Сириус, Россия). «Application of gradient optimization methods to solve ill-posed Cauchy problems».
2. 65 Всероссийская научная конференция МФТИ (2023, Долгопрудный, Россия). «О модификации метода покомпонентного спуска для решения некоторых обратных задач математической физики»
3. «Quasilinear Equations, Inverse Problems and their Applications» (2023, Долгопрудный, Россия). «Application of first-order optimization methods to solving the retrospective Cauchy problem for the three-dimensional heat equation».

4. Семинар Лаборатории математических методов оптимизации МФТИ (2023, Долгопрудный, Россия). «Применение градиентных методов оптимизации к решению ретроспективной задачи Коши для уравнения теплопроводности в трёхмерном пространстве и к решению интегрального уравнения Фредгольма 1-го рода».
5. Постерная сессия в рамках «ASCOMP» (2024, Иннополис, Россия). «Application of gradient methods to solving ill-posed problems of mathematical physics».
6. «Quasilinear Equations, Inverse Problems and their Applications» (2024, Сириус, Россия). «On modification of the conjugate gradient method with minimization of the distance to the exact solution when choosing the step length».

Структура диссертации.

Работа состоит из введения, 3 глав (постановка задач, построение методов оптимизации, применение методов к задачам), заключения, списков литературы, рисунков, таблиц и сокращений. Список литературы включает 57 работ. Диссертация содержит 58 рисунков и 23 таблицы.

Глава 1.

Общая характеристика и примеры задач

Данная глава содержит формулировки некорректных задач, к решению которых применяются методы оптимизации, их общую постановку в виде операторного уравнения, а также некоторые общие свойства.

1.1. Постановка и общие свойства

Задача является прямой, если в ней по начальным условиям и параметрам среды требуется найти описание явления. Если наоборот, то задача обратная. Именно обратные задачи чаще всего являются некорректными.

С другой стороны, часто некорректная задача для дифференциального уравнения в частных производных может быть представлена в виде обратной, то есть как задача поиска неизвестной функции q , удовлетворяющей операторному уравнению

$$Aq = f. \quad (1.1)$$

Здесь A — оператор, вычисление которого является корректно поставленной задачей.

Идея такого представления очень проста. Пусть есть две задачи для дифференциального уравнения в частных производных, отличающиеся только одним из граничных условий. Первая (исходная) задача имеет условие на одной границе (назовём её «наблюдаемой»), выражющееся функцией f , и является некорректно поставленной. Вторая задача имеет условие на другой («ненаблюдаемой») границе, выражющееся функцией q , при этом поставлена корректно. Эти задачи должны иметь одно и то же единственное решение.

Оператор задачи A сопоставляет условию q на ненаблюдаемой границе значение решения второй задачи на наблюдаемой границе. Его вычисление — прямая задача.

Поиск такого q , при котором решение на наблюдаемой границе равно f , — обратная задача.

Как будет видно из рассмотренных примеров, оператор A не всегда является линейным из-за неоднородных граничных условий. При исследовании подобных задач вводится оператор A_0 , который определяется аналогично, но с нулями вместо всех начальных и граничных условий, кроме заданного функцией q . Важное свойство: $Aq - Aq' = A_0(q - q')$.

Все задачи, перечисленные в этой главе и рассматриваемые в работе, объединяет сводимость к операторному уравнению. В свою очередь, решение такого уравнения является точкой глобального минимума выпуклого функционала

$$J(q) = \frac{1}{2} \|Aq - f\|^2 \quad (1.2)$$

в гильбертовом пространстве функций со стандартным скалярным произведением и определяемой им нормой. Это показывает возможность применения к решению обратных задач методов выпуклой оптимизации.

Более того, градиент указанного функционала (производная Фреше) вычисляется по общей формуле

$$\nabla J(q) = A^*(Aq - f), \quad (1.3)$$

где A^* — сопряжённый оператор, вычисление которого сводится к решению корректной задачи, похожей на задачу вычисления A .

Можно заметить, что градиент удовлетворяет условию Липшица с константой $L = \|A_0\|^2$. Поскольку задача вычисления оператора A является корректной, A_0 ограничен.

$$\|\nabla J(q) - \nabla J(q')\| = \|A^*(Aq - f) - A^*(Aq' - f)\| = \|A^*A_0(q - q')\| \leq L\|q - q'\|. \quad (1.4)$$

1.2. Задачи, решаемые в работе

Как известно, задача называется корректной по Адамару, если её решение существует, единственно и непрерывно зависит от начальных (граничных) условий. Если хотя бы одно из этих свойств не выполняется, задача некорректна.

Определение непрерывной зависимости решения задачи u от её начальных условий f , согласно [46]:

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall f \in H \text{ т. ч. } \|f\| < \delta, \|g\| < \delta \rightarrow \|u\| < \varepsilon.$$

Здесь H — линейное пространство функций, удовлетворяющих необходимым условиям гладкости. Используется C -норма на тех множествах, на которых заданы соответствующие функции.

В работе решаются разнообразные задачи, некорректность которых будет показана в соответствующих разделах:

- Начально-краевая задача для уравнения Гельмгольца — дифференциального уравнения в частных производных эллиптического типа:

$$\begin{cases} u_{xx} + u_{yy} + \kappa^2 u = r(x, y), & (x, y) \in \Omega = (0, 1) \times (0, 1) \\ u|_{x=0} = f(y), & y \in [0, 1] \\ u_x|_{x=0} = g(y), & y \in [0, 1] \\ u|_{y=0} = u|_{y=1} = 0, & x \in [0, 1] \end{cases} \quad (1.5)$$

Здесь коэффициент κ и функции $r(x, y)$, $f(y)$, $g(y)$ известны, требуется найти $u(x, y)$. Мы ограничимся восстановлением $u(1, y)$, поскольку уравнение Гельмгольца с условиями на $u|_{x=1}$, $u_x|_{x=0}$, $u|_{y=0}$, $u|_{y=1}$ — корректная (краевая) задача, решение которой не представляет сложности.

- Ретроспективная задача Коши для уравнения теплопроводности — дифференциального уравнения в частных производных параболического типа (в одномерном и трёхмерном пространствах):

— с граничными условиями на значение функции:

$$\begin{cases} u_t - \kappa^2 u_{xx} = 0, & (x, t) \in \Omega = (0, 1) \times (0, 1) \\ u|_{x=0} = a(t), & t \in [0, 1] \\ u|_{x=1} = b(t), & t \in [0, 1] \\ u|_{t=1} = f(x), & x \in [0, 1] \end{cases} \quad (1.6)$$

$$\begin{cases} u_t - \kappa^2 \Delta_x u = 0, & (x, t) \in \Omega = \Pi \times (0, 1) \\ u|_{x \in \partial\Pi} = 0, & t \in [0, 1] \\ u|_{t=1} = f(x), & x \in \Pi \end{cases} \quad (1.7)$$

— с граничными условиями на производную по пространственной

координате:

$$\begin{cases} u_t - \kappa^2 u_{xx} = 0, & (x, t) \in \Omega = (0, 1) \times (0, 1) \\ u_x|_{x=0} = a(t), & t \in [0, 1] \\ u_x|_{x=1} = b(t), & t \in [0, 1] \\ u|_{t=1} = f(x), & x \in [0, 1] \end{cases} \quad (1.8)$$

$$\begin{cases} u_t - \kappa^2 \Delta_x u = 0, & (x, t) \in \Omega = \Pi \times (0, 1) \\ \frac{\partial u}{\partial n}|_{x \in \partial \Pi} = 0, & t \in [0, 1] \\ u|_{t=1} = f(x), & x \in \Pi \end{cases} \quad (1.9)$$

Коэффициент κ может быть постоянным или зависеть от координат и времени. $\kappa, f(x)$, граничные условия $a(t)$ и $b(t)$ известны, требуется найти $u(x, t)$. Опять же ограничимся восстановлением $u(x, 0)$, потому что это условие (вместо $u(x, 1)$) превратит задачу в корректную, допускающую решение стандартными способами, например, с помощью разностных схем.

- обратная задача термоакустики — краевая задача для дифференциального уравнения в частных производных гиперболического типа;

$$\begin{cases} u_{tt} = u_{xx} + u_{yy}, & (x, y) \in \Omega = (0, L) \times (0, L), t \in (0, T) \\ u_t|_{t=0} = 0, & (x, y) \in \Omega \\ u_x|_{x=0} = u_x|_{x=L} = 0, & y \in (0, L), t \in (0, T) \\ u_y|_{y=0} = u_y|_{y=L} = 0, & x \in (0, L), t \in (0, T) \\ u|_{x=0} = f_1(y, t), & y \in (0, L), t \in (0, T) \\ u|_{x=L} = f_2(y, t), & y \in (0, L), t \in (0, T) \\ u|_{y=L} = f_3(x, t), & x \in (0, L), t \in (0, T) \end{cases} \quad (1.10)$$

Границные условия f_1, f_2, f_3 могут задаваться в любых комбинациях: одно из них, два или все три. Требуется найти начальное условие $u(x, y, 0)$.

- Интегральное уравнение Фредгольма 1-го рода:

$$\int_0^1 K(x, s)q(s)ds = f(x), \quad x \in [0, 1] \quad (1.11)$$

Функции $K(x, s)$ и $f(x)$ известны, нужно найти $q(s)$.

Глава 2.

Используемые методы оптимизации

В этой главе рассматриваются методы оптимизации, применяемые к решению некорректных задач. Приведены описания уже существующих методов, с которыми производится сравнение. Построены новые методы, а также доказаны теоремы об их свойствах.

2.1. Подходы к выбору длины шага

Специальный вид функционала и градиента позволяет детально рассмотреть (согласно [53], раздел 2.6) разные подходы к выбору длины шага. Пусть шаг k делается в направлении s_k :

$$q_{k+1} = q_k + \alpha_k s_k. \quad (2.1.0.1)$$

2.1.1. Выбор длины шага по принципу минимизации функционала

$$\alpha_k(J) = \arg \min_{\alpha \geq 0} J(q_k + \alpha s_k).$$

Элементарные преобразования минимизируемого выражения позволяют получить искомое значение коэффициента в общем случае:

$$\begin{aligned} J(q_k + \alpha s_k) &= \frac{1}{2} \|A(q_k + \alpha s_k) - f\|^2 = \frac{1}{2} \|Aq_k - f + \alpha A_0 s_k\|^2 = \\ &= J(q_k) + \alpha \langle Aq_k - f, A_0 s_k \rangle + \frac{\alpha^2}{2} \|A_0 s_k\|^2 = \\ &= J(q_k) + \alpha \langle \nabla J(q_k), s_k \rangle + \frac{\alpha^2}{2} \|A_0 s_k\|^2 \Rightarrow \\ \Rightarrow \alpha_k(J) &= -\frac{\langle \nabla J(q_k), s_k \rangle}{\|A_0 s_k\|^2}. \end{aligned} \quad (2.1.1.1)$$

Очевидно, эта формула применима во всех случаях, когда доступно вычисление оператора и градиента.

При этом достигается уменьшение функционала:

$$J(q_{k+1}) = J(q_k) - \frac{\langle \nabla J(q_k), s_k \rangle^2}{2\|A_0 s_k\|^2}.$$

Суммируя от 0 до n , получаем:

$$J(q_0) - J(q_{n+1}) = \sum_{k=0}^n \frac{\langle \nabla J(q_k), s_k \rangle^2}{2\|A_0 s_k\|^2}.$$

Убывающая и ограниченная снизу нулём последовательность $J(q_{n+1})$ имеет предел, перейдём к нему:

$$J(q_0) - \lim_{n \rightarrow \infty} J(q_n) = \sum_{k=0}^{\infty} \frac{\langle \nabla J(q_k), s_k \rangle^2}{2\|A_0 s_k\|^2}. \quad (2.1.1.2)$$

Следовательно:

- ряд в правой части всегда сходится;
- его сумма не превосходит $J(q_0)$;
- равенство суммы $J(q_0)$ — необходимое и достаточное условие сходимости по функционалу метода с шагами s_k и длиной каждого шага $\alpha_k(J)$.

2.1.2. Выбор длины шага по принципу минимизации расстояния до точки минимума

$$\alpha_k(\rho) = \arg \min_{\alpha \geq 0} \|q_k + \alpha s_k - q^*\|^2.$$

Элементарные преобразования минимизируемого выражения также позволяют получить искомое значение коэффициента:

$$\begin{aligned} \|q_k + \alpha s_k - q^*\|^2 &= \|q_k - q^*\|^2 + 2\alpha \langle q_k - q^*, s_k \rangle + \alpha^2 \|s_k\|^2 \Rightarrow \\ &\Rightarrow \alpha_k(\rho) = -\frac{\langle q_k - q^*, s_k \rangle}{\|s_k\|^2}. \end{aligned} \quad (2.1.2.1)$$

Поскольку точное решение q^* неизвестно, эта формула применима лишь тогда, когда скалярное произведение в числителе может быть преобразовано к вычислимому виду. Соответственно, минимизация расстояния до точного решения возможна не при любом выборе направления спуска.

Если она возможна, то на шаге k достигается уменьшение расстояния до точки минимума:

$$\|q_{k+1} - q^*\|^2 = \|q_k - q^*\|^2 - \frac{\langle q_k - q^*, s_k \rangle^2}{\|s_k\|^2} = \|q_k - q^*\|^2 - \|q_{k+1} - q_k\|^2.$$

Суммируя от 0 до n , получаем:

$$\|q_0 - q^*\|^2 - \|q_{n+1} - q^*\|^2 = \sum_{k=0}^n \|q_{k+1} - q_k\|^2.$$

Убывающая и ограниченная снизу нулём последовательность $\|q_n - q^*\|^2$ имеет предел, перейдём к нему:

$$\|q_0 - q^*\|^2 - \lim_{n \rightarrow \infty} \|q_n - q^*\|^2 = \sum_{k=0}^{\infty} \|q_{k+1} - q_k\|^2. \quad (2.1.2.2)$$

Следовательно:

- ряд в правой части всегда сходится;
- его сумма не превосходит $\|q_0 - q^*\|^2$;
- равенство суммы $\|q_0 - q^*\|^2$ — необходимое и достаточное условие сильной сходимости метода с шагами s_k и длиной каждого шага $\alpha_k(\rho)$;
- $\lim_{k \rightarrow \infty} \|q_{k+1} - q_k\| = 0$ (необходимое условие сходимости ряда). Из признака сравнения рядов следует, что для любой числовой последовательности x_k такой, что ряд $\sum_{k=1}^{\infty} x_k$ расходится, выполнено условие $\|q_{k+1} - q_k\| = o(\sqrt{x_k})$.

В частности, $\|q_{k+1} - q_k\| = o\left(\frac{1}{\sqrt{k}}\right)$.

2.2. Существующие методы

2.2.1. Метод градиентного спуска

Самый распространённый метод первого порядка — простой градиентный спуск, во многих работах называемый итерацией Ландвебера (например, в статье [52] он

вводится под таким именем):

$$s_k = -\nabla J(q_k). \quad (2.2.1.1)$$

Часто данный метод используется с постоянным шагом $\alpha_k \equiv \alpha \in \left(0, \frac{1}{\|A_0\|^2}\right)$. Для этого случая известна теорема о сходимости по функционалу с линейным порядком ([53], 2.7.2).

Для сильно выпуклых функционалов все виды сходимости (по аргументу, функционалу и норме градиента) эквивалентны. Однако при решении некорректных задач функционал оказывается выпуклым, но не сильно выпуклым. Поэтому эквивалентности нет. Более того, возможно существование такой «плохой» последовательности точек q_m , что $J(q_m) \rightarrow 0$, и при этом $\|q_m - q^*\| \rightarrow \infty$.

В разделе 2.3.5 доказана теорема 2, гарантирующая сходимость по аргументу в достаточно общем случае. То есть, метод градиентного спуска (и другие методы, для которых доказана сходимость не хуже) порождает «хорошую» последовательность точек.

Вычисление нормы оператора и, соответственно, выбор α представляет определённую сложность. Однако применимы формулы (2.1.1.1) и (2.1.2.1), полученные в разделе 2.1:

$$\alpha_k(J) = -\frac{\langle \nabla J(q_k), s_k \rangle}{\|A_0 s_k\|^2} = \frac{\|\nabla J(q_k)\|^2}{\|A_0 \nabla J(q_k)\|^2}; \quad (2.2.1.2)$$

$$\begin{aligned} \alpha_k(\rho) &= -\frac{\langle q_k - q^*, s_k \rangle}{\|s_k\|^2} = \frac{\langle q_k - q^*, \nabla J(q_k) \rangle}{\|\nabla J(q_k)\|^2} = \frac{\langle q_k - q^*, A^*(Aq_k - f) \rangle}{\|\nabla J(q_k)\|^2} = \\ &= \frac{\langle A_0(q_k - q^*), Aq_k - f \rangle}{\|\nabla J(q_k)\|^2} = \frac{\|Aq - f\|^2}{\|\nabla J(q_k)\|^2} \Rightarrow \\ &\Rightarrow \alpha_k(\rho) = \frac{2J(q_k)}{\|\nabla J(q_k)\|^2}. \end{aligned} \quad (2.2.1.3)$$

Один шаг градиентного спуска с выбором длины шага по принципу минимизации функционала (метод наискорейшего спуска) требует трёх вычислений оператора; с выбором длины шага по принципу минимизации расстояния до точки минимума (метод минимальных ошибок) — двух.

По определению операторной нормы $\forall x \in H \|A_0 x\| \leq \|A_0\| \cdot \|x\|$, поэтому $\alpha_k(J) \geq \frac{1}{\|A_0\|^2}$ и $\alpha_k(\rho) \geq \frac{1}{\|A_0\|^2}$.

Метод градиентного спуска — это основа для построения более сложных и быстрых методов, некоторые из которых описаны в последующих разделах.

2.2.2. Метод сопряжённых градиентов

Классический метод второго порядка — это метод сопряжённых градиентов. Как известно ([54]), в конечномерном квадратичном случае он гарантирует сходимость за конечное число шагов.

$$s_0 = -\nabla J(q_0), \quad s_k = -\nabla J(q_k) + \beta_k s_{k-1}. \quad (2.2.2.1)$$

Существуют разные варианты вычисления параметра β_k :

- алгоритм Флетчера-Ривза $\beta_k = \frac{\|\nabla J(q_k)\|^2}{\|\nabla J(q_{k-1})\|^2}$;
- алгоритм Полака-Рибьера $\beta_k = \max \left(0, \frac{\langle \nabla J(q_k), \nabla J(q_k) - \nabla J(q_{k-1}) \rangle}{\|\nabla J(q_{k-1})\|^2} \right)$;
- $\beta_k = \frac{\langle \nabla J(q_k), s_{k-1} \rangle}{\|s_{k-1}\|^2}$: выбор из соображений ортогональности соседних шагов $\langle s_k, s_{k-1} \rangle = 0$.

Длина шага выбирается из соображений минимизации функционала, то есть используется формула (2.1.1.1).

Она не поддаётся преобразованию, в отличие от (2.1.2.1):

$$\alpha_k(\rho) = -\frac{\langle q_k - q^*, s_k \rangle}{\|s_k\|^2} = -\frac{\langle q_k - q^*, -\nabla J(q_k) + \beta_k s_{k-1} \rangle}{\|s_k\|^2} = \frac{2J(q_k)}{\|s_k\|^2}, \quad (2.2.2.2)$$

поскольку, независимо от способа выбора множителя β_k ,

$$\begin{aligned} \langle q_k - q^*, s_{k-1} \rangle &= \langle q_{k-1} - q^*, s_{k-1} \rangle + \langle q_k - q_{k-1}, s_{k-1} \rangle = \\ &= -\alpha_{k-1}(\rho) \|s_{k-1}\|^2 + \langle \alpha_{k-1}(\rho) s_{k-1}, s_{k-1} \rangle = 0. \end{aligned}$$

Это вычисление опирается на предположение, что предыдущий шаг тоже сделан по принципу минимизации расстояния до точного решения.

Количество вычислений оператора на одном шаге — такое же, как и для градиентного спуска.

Сочетание выбора β_k из соображений ортогональности соседних шагов с использованием длины шага $\alpha_k(\rho)$ образует новый метод оптимизации, строгому построению которого посвящён раздел 2.3.2.

2.2.3. Метод подобных треугольников

В статье [57] вводится ускоренный метод подобных треугольников (алгоритм 1), обычно называемый STM, который может применяться для решения выпуклых (не обязательно сильно выпуклых) задач с известными константами Липшица градиента и сильной выпуклости.

Algorithm 1 STM

Input: $f : Q \rightarrow \mathbb{R}$, $x_{start} \in Q$, L , N .

```

1:  $\tilde{x}_0 := x_{start}$ ;  $A_0 := \frac{1}{L}$ ;  $\alpha_0 := \frac{1}{L}$ ;  $z_0 := \tilde{x}_0 - \alpha_0 \tilde{\nabla} f(\tilde{x}_0)$ ;  $z_0 := x_0$ ;
2: for  $k = 1 \dots N$  do
3:    $\alpha_k := \frac{1}{2L} + \sqrt{\frac{1}{4L^2} + A_{k-1}}$ ;
4:    $A_k := A_{k-1} + \alpha_k$ ;
5:    $\tilde{x}_k := \frac{A_{k-1}x_{k-1} + \alpha_k z_{k-1}}{A_k}$ ;
6:    $z_k := z_{k-1} - \alpha_k \tilde{\nabla} f(\tilde{x}_k)$ ;
7:    $x_k := \frac{A_{k-1}x_{k-1} + \alpha_k z_k}{A_k}$ ;
8: end for
```

Output: x_N .

Достоинством такого метода с постоянными коэффициентами является отсутствие необходимости решать задачи вспомогательной минимизации и связанной с ними погрешности.

К недостаткам относится необходимость предварительной оценки констант Липшица и сильной выпуклости. Неправильная оценка может привести к потере сходимости, а слишком осторожная — к её замедлению. В простых случаях константа Липшица градиента может быть вычислена.

Для не сильно выпуклой функции в статье [57] доказана теорема о сходимости данного метода в случае аддитивного шума в градиенте:

$$f(x_N) - f(x^*) \leq \frac{4LR^2}{N^2} + 3\tilde{R}\delta + \frac{N\delta^2}{2L},$$

где $R = \|x_0 - x^*\|$, $\tilde{R} = \max_{0 \leq k \leq N} \{\|x_k - x^*\|, \|z_k - x^*\|, \|\tilde{x}_k - x^*\|\}$, δ — верхняя оценка шума.

2.2.4. Метод минорант Поляка

В статье [35] был предложен, а в статье [49] развит развит так называемый РММ — метод минорант Поляка. Он основан на следующей идее: на каждом шаге k подбирается «миноранта» $m_k(q)$ — функция простого вида, для которой выполняются соотношения $m_k(q) \leq J(q)$ и $m_k(q_k) = J(q_k)$. Подбирается точка q_{k+1} , для которой миноранта принимает целевое значение: $m_k(q_{k+1}) = J(q^*) = 0$. При этом $J(q_{k+1}) < J(q_k)$.

Поскольку «естественная» миноранта для гладкого выпуклого функционала — это его линейное приближение, метод принимает вид:

$$q_{k+1} = q_k - \frac{J(q_k)}{\|\nabla J(q_k)\|^2} \nabla J(q_k) \quad (2.2.4.1)$$

Эта формула отличается от (2.2.1.3), задающей метод минимальных ошибок, лишь отсутствием коэффициента 2. Выполняется доказанное в [49] и [35] неравенство

$$\|q_{k+1} - q^*\|^2 \leq \|q_k - q^*\|^2 - \|q_{k+1} - q_k\|^2,$$

откуда, при достаточно мягких условиях, следует сходимость не только по функционалу, но и по аргументу.

Эксперименты показывают, что метод Поляка эффективнее одномерной минимизации расстояния. Как указывает сам Б. Т. Поляк, его скорость сходимости примерно соответствует характеристикам обычных градиентных методов. Однако, оказывается возможным достижение лучших результатов при использовании многомерной вспомогательной минимизации.

2.2.5. Адаптивный тяжёлый шарик Поляка

В статье [51] предпринята попытка построить метод квадратичной минимизации, оптимальный в смысле минимизации расстояния до точного решения. Метод, описанный там, назван в честь Б. Т. Поляка, поскольку развивает его идеи: с одной стороны, он похож на метод тяжёлого шарика, а с другой — на шаг Поляка, описанный в предыдущем разделе.

Авторы статьи утверждают, что этой оптимальностью обладает следующий алгоритм (в оригинальных обозначениях, с учётом $f^* = 0$):

$$x_{t+1} = x_t - (1 + m_t)h_t \nabla f(x_t) + m_t(x_t - x_{t-1});$$

$$h_t = \frac{2f(x_t)}{\|\nabla f(x_t)\|^2};$$

$$m_0 = 0; m_t = \frac{-f(x_t)\langle \nabla f(x_t), \nabla f(x_{t-1}) \rangle}{f(x_{t-1})\|\nabla f(x_t)\|^2 + f(x_t)\langle \nabla f(x_t), \nabla f(x_{t-1}) \rangle}.$$

Однако при работе метода, удовлетворяющего требуемому свойству, расстояние до точного решения должно монотонно убывать. Эксперименты показывают, что построенный в указанной статье метод этим свойством не обладает.

В данной работе проблема решается: оптимальный метод строится, его свойства строго доказываются.

2.3. Новые предлагаемые методы

Ключевые результаты данного раздела, посвящённые построению m -моментного метода минимальных ошибок и доказательству его свойств, были изложены автором в статье [34].

2.3.1. Метод покомпонентного спуска

В статье [33] с целью решения ретроспективной задачи Коши для уравнения теплопроводности был введён метод покомпонентного спуска — алгоритм 2. Он применим в тех случаях, когда известны собственные функции w_m самосопряжённого оператора A^*A_0 и соответствующие им собственные значения λ_m .

Algorithm 2 Restarted coordinate descent

Input: $q^{(start)} \in Q$, $\{\lambda_m\}_{m=0}^{\infty}$, $\{w_m\}_{m=0}^{\infty}$, N_{steps} , $N_{restarts}$.

- 1: $q := q^{(start)}$;
- 2: **for** $n = 1 \dots N_{restarts}$ **do**
- 3: $q^{(n,0)} := q$
- 4: $g := \nabla J(q)$
- 5: **for** $k = 0 \dots N_{steps} - 1$ **do**
- 6: $q^{(n,k+1)} := q^{(n,k)} - \frac{\langle g, w_k \rangle}{\lambda_k} w_k$;
- 7: **if** $J(q^{(n,k+1)}) > J(q^{(n,k)})$ **then**
- 8: $q := q^{(n,k)}$
- 9: **break**;
- 10: **else**
- 11: $q := q^{(n,k+1)}$
- 12: **end if**
- 13: **end for**
- 14: **end for**

Output: q .

Коэффициент на шаге 6 — это $\alpha_k(J)$ при $s_k = w_k$, если w_k является собственной функцией оператора A_0 .

Собственные значения и функции оператора A^*A_0 известны лишь в редких частных случаях: постоянный коэффициент в эллиптическом или параболическом дифференциальном уравнении, вырожденное ядро в интегральном уравнении. Поэтому большого практического значения данный метод не имеет.

2.3.2. Моментный метод минимальных ошибок

Моментными называют методы следующего вида:

$$q_{k+1} = q_k - \alpha_k \nabla J(q_k) + \gamma_k (q_k - q_{k-1}) \quad (2.3.2.1)$$

или

$$s_k = -\nabla J(q_k) + \beta_k s_{k-1}; \quad q_{k+1} = q_k + \alpha_k s_k \quad (2.3.2.2)$$

Поскольку векторы s_{k-1} и $q_k - q_{k-1}$ коллинеарны, эти записи эквивалентны, и одни коэффициенты можно выразить через другие: $\gamma_k \alpha_{k-1} = \alpha_k \beta_k$.

Можно заметить, что к данному классу относятся различные уже рассмотренные методы, такие как метод сопряжённых градиентов, метод подобных треугольников или метод тяжёлого шарика. Идея заключается в том, чтобы оптимально подобрать коэффициенты.

Выполним вспомогательную двумерную минимизацию расстояния до точного решения при вычислении q_{k+1} в форме (2.3.2.1). Пусть $\gamma_0 = 0$, $\alpha_0 = \alpha_0(\rho)$. $q_k \neq q^*$ (иначе новый шаг не нужен).

$$\begin{aligned} \|q_{k+1} - q^*\|^2 &= \|q_k - q^* - \alpha_k \nabla J(q_k) + \gamma_k(q_k - q_{k-1})\|^2 = \\ &= \|q_k - q^*\|^2 + \alpha_k^2 \|\nabla J(q_k)\|^2 + \gamma_k^2 \|q_k - q_{k-1}\|^2 - 2\alpha_k \langle q_k - q^*, \nabla J(q_k) \rangle + \\ &\quad + 2\gamma_k \langle q_k - q^*, q_k - q_{k-1} \rangle - 2\alpha_k \gamma_k \langle \nabla J(q_k), q_k - q_{k-1} \rangle. \end{aligned}$$

Вычислим частные производные. Их равенство нулю является необходимым условием экстремума.

$$\begin{aligned} \frac{1}{2} \frac{\partial}{\partial \alpha_k} \|q_{k+1} - q^*\|^2 &= \alpha_k \|\nabla J(q_k)\|^2 - \gamma_k \langle \nabla J(q_k), q_k - q_{k-1} \rangle - \langle q_k - q^*, \nabla J(q_k) \rangle; \\ \frac{1}{2} \frac{\partial}{\partial \gamma_k} \|q_{k+1} - q^*\|^2 &= -\alpha_k \langle \nabla J(q_k), q_k - q_{k-1} \rangle + \gamma_k \|q_k - q_{k-1}\|^2 + \langle q_k - q^*, q_k - q_{k-1} \rangle. \end{aligned}$$

Преобразуем слагаемые, содержащие q^* :

$$\begin{aligned} \langle q_k - q^*, \nabla J(q_k) \rangle &= \langle q_k - q^*, A^*(Aq_k - f) \rangle = \langle A_0(q_k - q^*), Aq_k - f \rangle = \\ &= \|Aq_k - f\|^2 = 2J(q_k); \\ \langle q_k - q^*, q_k - q_{k-1} \rangle &= 0, \end{aligned} \tag{2.3.2.3}$$

потому что q_k выбрано на предыдущем шаге, как ближайший к q^* элемент линейного многообразия $q_{k-1} + \text{Lin} \{ \nabla J(q_{k-1}), q_{k-1} - q_{k-2} \}$, а это проекция на гиперплоскость.

Итак, получаем систему линейных уравнений:

$$\begin{cases} \alpha_k \|\nabla J(q_k)\|^2 - \gamma_k \langle \nabla J(q_k), q_k - q_{k-1} \rangle = 2J(q_k); \\ -\alpha_k \langle \nabla J(q_k), q_k - q_{k-1} \rangle + \gamma_k \|q_k - q_{k-1}\|^2 = 0. \end{cases}$$

Для её решения необходимо проверить, что она не является вырожденной. Для этого вычислим определитель:

$$\Delta = \|\nabla J(q_k)\|^2 \|q_k - q_{k-1}\|^2 - \langle \nabla J(q_k), q_k - q_{k-1} \rangle^2.$$

Неравенство Коши — Буняковского показывает, что всегда $\Delta \geq 0$, и равенство эквивалентно коллинеарности $\nabla J(q_k)$ и $q_k - q_{k-1}$. Поскольку $\langle q_k - q^*, q_k - q_{k-1} \rangle = 0$, одновременно выполняется и следующее равенство:

$$2J(q_k) = \langle q_k - q^*, \nabla J(q_k) \rangle = \langle q_k - q^*, \lambda(q_k - q_{k-1}) \rangle = 0.$$

Но это возможно только в том случае, когда q_k — точное решение задачи. По исходному предположению, это не так.

Близость Δ к нулю обозначает плохую обусловленность системы и низкую точность очередного шага в условиях реальных вычислений. Этот признак может быть использован для формулировки критерия остановки метода.

Если решение задачи оптимизации не достигнуто, то $\Delta > 0$. Тогда система имеет единственное решение (α_k, γ_k) . Оно является точкой минимума для $\|q_{k+1} - q^*\|^2$, поскольку матрица вторых производных — положительно определённая по критерию Сильвестра: её главные миноры $\|\nabla J(q_k)\|^2$ и Δ положительны.

Осталось вычислить это решение. Из второго уравнения:

$$\gamma_k = \alpha_k \cdot \frac{\langle \nabla J(q_k), q_k - q_{k-1} \rangle}{\|q_k - q_{k-1}\|^2}.$$

Подставляя в первое уравнение, получаем:

$$\alpha_k = \frac{2J(q_k)}{\|\nabla J(q_k)\|^2 - \frac{\langle \nabla J(q_k), q_k - q_{k-1} \rangle^2}{\|q_k - q_{k-1}\|^2}}.$$

Полученный метод является оптимальным среди моментных методов по признаку уменьшения расстояния до точного решения после одного шага.

Введём обозначение:

$$s_k = -\nabla J(q_k) + \frac{\langle \nabla J(q_k), q_k - q_{k-1} \rangle}{\|q_k - q_{k-1}\|^2} (q_k - q_{k-1}).$$

Поскольку $q_k - q_{k-1} = \alpha_{k-1} s_{k-1}$, выполняется равенство:

$$s_k = -\nabla J(q_k) + \beta_k s_{k-1},$$

где

$$\beta_k = \frac{\langle \nabla J(q_k), s_{k-1} \rangle}{\|s_{k-1}\|^2}.$$

Очевидно, $\langle s_k, q_k - q_{k-1} \rangle = 0$. Поэтому

$$\|s_k\|^2 + \frac{\langle \nabla J(q_k), q_k - q_{k-1} \rangle^2}{\|q_k - q_{k-1}\|^2} = \|\nabla J(q_k)\|^2,$$

откуда

$$\alpha_k = \frac{2J(q_k)}{\|s_k\|^2}.$$

Вычисление $\alpha_k(\rho)$ для введённого шага s_k приводит к тому же результату:

$$\alpha_k(\rho) = -\frac{\langle q_k - q^*, s_k \rangle}{\|s_k\|^2} = \frac{\langle q_k - q^*, \nabla J(q_k) \rangle}{\|s_k\|^2} - \beta_k \frac{\langle q_k - q^*, s_{k-1} \rangle}{\|s_k\|^2} = \frac{2J(q_k)}{\|s_k\|^2},$$

так как $\langle q_k - q^*, s_{k-1} \rangle = 0$ по доказанному в ходе преобразования частных производных.

Итак, построенный локально оптимальный моментный метод может быть представлен как моментный метод (2.3.2.2) с коэффициентом сопряжённости

$$\beta_k = \frac{\langle \nabla J(q_k), s_{k-1} \rangle}{\|s_{k-1}\|^2}, \quad (2.3.2.4)$$

вычисленным из соображения ортогональности соседних шагов, и выбором длины шага

$$\alpha_k = \alpha_k(\rho) = \frac{2J(q_k)}{\|s_k\|^2} \quad (2.3.2.5)$$

по принципу минимизации расстояния до точного решения.

2.3.3. Многошаговый метод минимальных ошибок

Рассмотрим следующее обобщение построенного метода, зависящее от натурального параметра m . Назовём его m -моментным методом минимальных ошибок. При $m = 1$ получится метод из раздела 2.3.2.

$$q_{k+1} = \arg \min_{q \in q_k + \text{Lin}\{-\nabla J(q_k), q_k - q_{k-1}, \dots, q_{k-m+1} - q_{k-m}\}} \|q - q^*\|^2. \quad (2.3.3.1)$$

То есть, на каждом шаге выбирается ближайшая к решению точка из линейного многообразия, проходящего через последнюю полученную точку и натянутого на векторы антиградиента и m предыдущих шагов.

Если $k < m$, то используются все k шагов, в этом случае размерность пространства равна $k + 1$.

Для краткости обозначим $h_j := q_{j+1} - q_j$ при всех j .

$$q_{k+1} = q_k - \gamma_k^0 \nabla J(q_k) + \sum_{i=1}^m \gamma_k^i h_{k-i}. \quad (2.3.3.2)$$

Коэффициенты γ_k^j в (2.3.3.2) получаются путём решения задачи минимизации (2.3.3.1).

Выпишем минимизируемое выражение:

$$\begin{aligned} \|q_{k+1} - q^*\|^2 &= \|q_k - q^*\|^2 + (\gamma_k^0)^2 \|\nabla J(q_k)\|^2 + \sum_{i=1}^m (\gamma_k^i)^2 \|h_{k-i}\|^2 - \\ &- 2\gamma_k^0 \langle q_k - q^*, \nabla J(q_k) \rangle + 2 \sum_{1 \leq i < j \leq m} \gamma_k^i \gamma_k^j \langle h_{k-i}, h_{k-j} \rangle - \\ &- 2 \sum_{i=1}^m \gamma_k^0 \gamma_k^i \langle \nabla J(q_k), h_{k-i} \rangle + 2 \sum_{i=1}^m \gamma_k^i \langle q_k - q^*, h_{k-i} \rangle. \end{aligned}$$

Преобразуем слагаемые, содержащие неизвестное q^* :

$$\langle q_k - q^*, \nabla J(q_k) \rangle = \langle A_0(q_k - q^*), Aq_k - f \rangle = \|Aq_k - f\|^2 = 2J(q_k); \quad (2.3.3.3)$$

$$\langle q_k - q^*, h_{k-i} \rangle = \langle q_k - q_{k-i+1}, h_{k-i} \rangle + \langle q_{k-i+1} - q^*, h_{k-i} \rangle. \quad (2.3.3.4)$$

Второе слагаемое в (2.3.3.4) равно 0, поскольку точки q^*, q_{k-i}, q_{k-i+1} образуют прямоугольный треугольник с прямым углом q_{k-i+1} . Это связано с тем, что выбранная на шаге $k - i$ точка является ближайшей к q^* в своём подпространстве — значит и на принадлежащей ему прямой.

Теперь (2.3.3.4) можно преобразовать:

$$\langle q_k - q^*, h_{k-i} \rangle = \langle q_k - q_{k-i+1}, h_{k-i} \rangle = \left\langle \sum_{j=1}^{i-1} h_{k-j}, h_{k-i} \right\rangle = \sum_{j=1}^{i-1} \langle h_{k-j}, h_{k-i} \rangle. \quad (2.3.3.5)$$

Все слагаемые, кроме $\|q_k - q^*\|^2$, не зависящего от вектора γ_k , удалось выразить в вычислимом виде. Следовательно, минимум может быть найден с помощью необходимого (для выпуклой квадратичной функции — достаточного) условия экстремума. Для этого вычислим частные производные по γ_k^i :

$$\frac{1}{2} \frac{\partial}{\partial \gamma_k^0} \|q_{k+1} - q^*\|^2 = \gamma_k^0 \|\nabla J(q_k)\|^2 - \sum_{i=1}^m \gamma_k^i \langle \nabla J(q_k), h_{k-i} \rangle - 2J(q_k);$$

при $1 \leq i \leq m$

$$\begin{aligned} \frac{1}{2} \frac{\partial}{\partial \gamma_k^i} \|q_{k+1} - q^*\|^2 &= \gamma_k^i \|h_{k-i}\|^2 + \sum_{j=1, j \neq i}^m \gamma_k^j \langle h_{k-j}, h_{k-i} \rangle - \\ &- \gamma_k^0 \langle \nabla J(q_k), h_{k-i} \rangle + \sum_{j=1}^{i-1} \langle h_{k-j}, h_{k-i} \rangle. \end{aligned}$$

Таким образом, условие экстремума можно представить в матричном виде:

$$G\gamma_k = b,$$

где

$$G_{00} = \|\nabla J(q_k)\|^2, G_{0i} = G_{i0} = -\langle \nabla J(q_k), h_{k-i} \rangle,$$

$$G_{ii} = \|h_{k-i}\|^2, G_{ij} = G_{ji} = \langle h_{k-i}, h_{k-j} \rangle,$$

$$\gamma_k = (\gamma_k^0, \gamma_k^1, \dots, \gamma_k^m)^T,$$

$$b = \left(2J(q_k), 0, -\langle h_{k-1}, h_{k-2} \rangle, \dots, -\sum_{j=1}^{i-1} \langle h_{k-i}, h_{k-j} \rangle, \dots, -\sum_{j=1}^{m-1} \langle h_{k-m}, h_{k-j} \rangle \right)^T.$$

Это представление вызывает следующие вопросы. Может ли матрица G оказаться вырожденной? Возможно ли упрощение вычислений (решения системы линейных уравнений порядка $m+1$)? Ответ на эти вопросы даёт следующая лемма.

Лемма 1. *Каждый шаг ортогонален предыдущим (или всем, если его номер меньше m), то есть $\langle h_k, h_{k-i} \rangle = 0$ при $1 \leq i \leq m$.*

Доказательство. Применим метод математической индукции по номеру шага k . База очевидна: при $k = 0$ или $k = 1$ и произвольном $m \geq 1$ метод совпадает с моментным методом из раздела (2.3.2), для которого ортогональность соседних шагов уже была доказана.

Пусть лемма верна для всех шагов с номерами меньше k . Тогда на шаге k все компоненты вектора b , кроме имеющей нулевой номер, равны нулю по предположению индукции. Также равны нулю все элементы матрицы G , кроме нулевой строки, нулевого столбца и главной диагонали.

Поскольку $h_k = -\gamma_k^0 \nabla J(q_k) + \sum_{i=1}^m \gamma_k^i h_{k-i}$, при $1 \leq i \leq m$. можно вычислить скалярное произведение:

$$\langle h_k, h_{k-i} \rangle = -\gamma_k^0 \langle \nabla J(q_k), h_{k-i} \rangle + \gamma_k^i \|h_{k-i}\|^2 = \sum_{j=0}^m G_{ji} \gamma_k^j = b_i = 0,$$

что и требовалось доказать. □

Из леммы 1 и формулы (2.3.3.5) сразу следует, что при $1 \leq i \leq m$

$$\langle q_k - q^*, h_{k-i} \rangle = 0.$$

Предположим теперь, что система векторов, из линейной оболочки которых выбирается вектор очередного шага, является линейно зависимой, то есть

$$\nabla J(q_k) = \sum_{i=1}^m \lambda_i h_{k-i}.$$

Умножим скалярно на $q_k - q^*$, получим равенство

$$2J(q_k) = \sum_{i=1}^m \lambda_i \langle q_k - q^*, h_{k-i} \rangle = 0,$$

которое выполняется только при достижении точного решения.

Соответственно, при $q_k \neq q^*$ эта система векторов линейно независима. Для решения системы уравнений $G\gamma_k = b$ выпишем и преобразуем (методом Гаусса) её расширенную матрицу:

$$\begin{array}{c|c} \left(\begin{array}{ccccc} \|\nabla J(q_k)\|^2 & -\langle \nabla J(q_k), h_{k-1} \rangle & \cdots & -\langle \nabla J(q_k), h_{k-m} \rangle & 2J(q_k) \\ -\langle \nabla J(q_k), h_{k-1} \rangle & \|\hbar_{k-1}\|^2 & \cdots & 0 & 0 \\ -\langle \nabla J(q_k), h_{k-2} \rangle & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\langle \nabla J(q_k), h_{k-m} \rangle & 0 & \cdots & \|\hbar_{k-m}\|^2 & 0 \end{array} \right) & \sim \\ \sim \left(\begin{array}{ccccc} \|\nabla J(q_k)\|^2 - \sum_{i=1}^m \frac{\langle \nabla J(q_k), h_{k-i} \rangle^2}{\|h_{k-i}\|^2} & 0 & \cdots & 0 & 2J(q_k) \\ -\langle \nabla J(q_k), h_{k-1} \rangle & \|\hbar_{k-1}\|^2 & \cdots & 0 & 0 \\ -\langle \nabla J(q_k), h_{k-2} \rangle & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\langle \nabla J(q_k), h_{k-m} \rangle & 0 & \cdots & \|\hbar_{k-m}\|^2 & 0 \end{array} \right) & \end{array}$$

Можно заметить, что вычитаемая сумма квадратов в левом верхнем углу — это квадрат модуля проекции $-\nabla J(q_k)$ на $\text{Lin}\{h_{k-1}, \dots, h_{k-m}\}$. Обозначим угол между антиградиентом и этой проекцией — φ_k . Тогда левый верхний элемент матрицы равен $\|\nabla J(q_k)\|^2 \sin^2 \varphi_k$, и можно вычислить компоненты вектора γ_k :

$$\gamma_k^0 = \frac{2J(q_k)}{\|\nabla J(q_k)\|^2 \sin^2 \varphi_k}; \quad \gamma_k^i = \gamma_k^0 \cdot \frac{\langle \nabla J(q_k), h_{k-i} \rangle}{\|h_{k-i}\|^2} \text{ при } 1 \leq i \leq m.$$

Теперь шаг метода может быть представлен в следующем виде:

$$s_k = -\nabla J(q_k) + \sum_{i=1}^m \frac{\langle \nabla J(q_k), h_{k-i} \rangle}{\|h_{k-i}\|^2} h_{k-i} \quad (2.3.3.6)$$

— проекция антиградиента на ортогональное дополнение линейной оболочки m предыдущих шагов; длина шага α_k вычисляется по формуле (2.3.2.5).

Формула для s_k может быть переписана рекуррентно:

$$s_k = -\nabla J(q_k) + \sum_{i=1}^m \frac{\langle \nabla J(q_k), s_{k-i} \rangle}{\|s_{k-i}\|^2} s_{k-i}. \quad (2.3.3.7)$$

При $k < m$ суммирование выполняется до индекса $i = k$. Формально положив $m = \infty$, получаем метод, каждый шаг которого ортогонален всем предыдущим. Его применение с n шагами требует n вычислений градиента и порядка n^2 вычислений скалярного произведения. Важно заметить, что, в отличие от вычисления градиента, скалярное произведение — «дешёвая» операция, поэтому на практике квадратичная сложность не приводит к заметному увеличению времени работы.

2.3.4. Применение при составном функционале

Все вычисления проведены в предположении, что функционал имеет простой квадратичный вид (1.2). Но в некоторых задачах функционал более сложен: он состоит из нескольких квадратичных слагаемых с общей точкой минимума. Это, например, обратная задача термоакустики или задача восстановления источников для уравнения теплопроводности.

Такой функционал имеет следующий вид:

$$J(q) = \frac{1}{2} \sum_{l=1}^s \|A_l q - f_l\|^2, \quad (2.3.4.1)$$

причём $\forall l \in \{1, \dots, m\}$ $A_l q^* = f_l$, поэтому $J(q^*) = 0$.

Его градиент вычисляется аналогично простому случаю:

$$\nabla J(q) = \sum_{l=1}^s A_l^* (A_l q - f_l). \quad (2.3.4.2)$$

В этом случае все преобразования этого раздела повторяются дословно, за исключением (2.3.2.3) и (2.3.3.3):

$$\begin{aligned} \langle q_k - q^*, \nabla J(q_k) \rangle &= \left\langle q_k - q^*, \sum_{l=1}^s A_l^* (A_l q_k - f_l) \right\rangle = \\ &= \sum_{l=1}^s \langle A_l (q_k - q^*), A_l q_k - f_l \rangle = \sum_{l=1}^s \|A_l q_k - f_l\|^2 = 2J(q_k). \end{aligned}$$

Результат аналогичный, поэтому все полученные формулы сохраняются. Оценки сходимости также будут общими.

2.3.5. Теоремы о сходимости построенных методов

Теорема 1 (Об оптимальности ∞ -моментного ММО). *∞ -моментный метод минимальных ошибок, полученный в разделе 2.3.3, является глобально оптимальным для решения задачи минимизации квадратичного функционала в гильбертовом пространстве среди методов, работающих в подпространствах Крылова. То есть, никакой метод из указанного класса не позволяет достичь меньшего расстояния до точного решения при одинаковом количестве шагов.*

Доказательство. Обозначим $B = A^*A_0$ — самосопряжённый оператор, связанный с задачей. Пусть q_0 — фиксированный элемент множества, на котором определён оператор A , (если $A = A_0$, то $q_0 = 0$). Тогда градиент функционала представляется так:

$$\nabla J(q) = B(q - q^*) = B(q - q_0) + A^*(Aq_0 - f) = B(q - q_0) + \nabla J(q_0).$$

Функционал можно представить в виде, типичном для применения классического метода сопряжённых градиентов:

$$J(q) = \frac{1}{2} \langle q - q_0, B(q - q_0) \rangle - \langle A^*(f - Aq_0), q - q_0 \rangle + \|Aq_0 - f\|^2.$$

То есть, обозначая градиент в начальной точке $\nabla J(q_0) = g$, получаем подпространства Крылова для задачи:

$$K_n = \text{Lin}\{g, Bg, \dots, B^{n-1}g\}.$$

Докажем по индукции следующее утверждение о последовательности q_n , порождаемой ∞ -моментным минимальных ошибок:

если векторы $g, Bg, B^2g, \dots, B^n g$ линейно независимы,

то $\nabla J(q_n) \in K_{n+1} \setminus K_n$, $h_n \in K_{n+1} \setminus K_n$.

При $n = 0$ оно очевидно: по определению $K_0 = \{0\}$, $K_1 = \text{Lin}\{\nabla J(q_0)\}$ — одномерное подпространство ($g \neq 0$, поскольку система, содержащая нулевой вектор, была бы линейно зависимой).

Пусть для всех шагов с номерами меньше n утверждение выполнено. Докажем его для номера n .

$$\nabla J(q_n) = \nabla J(q_{n-1}) + B(q_n - q_{n-1}) = \nabla J(q_{n-1}) + Bh_{n-1}.$$

Первое слагаемое, по предположению индукции, лежит в $K_n \subset K_{n+1}$. Посмотрим на второе слагаемое. По предположению индукции $h_{n-1} \in K_n$ и $h_{n-1} \notin K_{n-1}$, то есть h_{n-1} раскладывается по системе $\{g, Bg, \dots, B^{n-1}g\}$, причём последний коэффициент ненулевой, поэтому Bh_{n-1} раскладывается по системе $\{Bg, B^2g, \dots, B^n g\}$, и последний коэффициент — снова ненулевой (разложение по линейно независимой системе единственno), то есть $Bh_{n-1} \in K_{n+1}$ и $Bh_{n-1} \notin K_n$. Следовательно, $\nabla J(q_n) \in K_{n+1}$ (как сумма двух его элементов) и $\nabla J(q_n) \notin K_n$ (иначе Bh_{n-1} лежало бы в K_n , как разность его элементов).

$$h_n = -\gamma_n^0 \nabla J(q_n) + \sum_{i=1}^n \gamma_n^i h_{n-i}.$$

Здесь $\gamma_n^0 \neq 0$, и все слагаемые принадлежат K_{n+1} (по доказанному), а все, кроме первого, — K_n (по предположению индукции). Отсюда сразу получается требуемое.

Соответственно, на каждом шаге очередная точка q_{n+1} выбирается из $q_n + K_{n+1}$. Поскольку $q_n - q_0 \in K_{n+1}$, это множество совпадает с $q_0 + K_{n+1}$.

В рассматриваемом множестве ∞ -моментный метод выбирает точку, ближайшую к решению. Поэтому совпадение множеств

$$q_n + \text{Lin}\{-\nabla J(q_n), h_{n-1}, \dots, h_1, h_0\} = q_0 + K_{n+1}$$

и обозначает справедливость утверждения теоремы в случае, если для любого номера шага n векторы $g, Bg, B^2g, \dots, B^n g$ образуют линейно независимую систему.

Если же для некоторого n эта система впервые оказалась линейно зависимой, то максимальная размерность подпространств Крылова равна n . Соответственно, n ортогональных векторов h_0, h_1, \dots, h_{n-1} образуют в K_n базис. Поскольку градиент в любой порождаемой методом первого порядка точке принадлежит соответствующему подпространству Крылова, $\nabla J(q_n)$ раскладывается по этому базису. По следствию из леммы 1, это обозначает, что $q_n = q^*$. Но в таком случае утверждение теоремы тоже выполняется. \square

Важность доказанной теоремы следует из того, что все моментные методы, включая различные варианты метода сопряжённых градиентов, лежат в классе методов, работающих в подпространствах Крылова. Теорема показывает, что ∞ -моментный метод минимальных ошибок обеспечивает наилучшую сходимость по аргументу в этом классе.

Теорема 2 (О сходимости по аргументу метода градиентного спуска). *Если некорректно поставленная задача (1.1) в гильбертовом пространстве H имеет единственное решение q^* и оператор A_0 компактен, то простой градиентный спуск*

$$q_{k+1} = q_k - \alpha \nabla J(q_k)$$

(при $\alpha \leq \frac{1}{L}$) сходится по аргументу к точному решению, то есть

$$\lim_{k \rightarrow \infty} \|q_k - q^*\|^2 = 0.$$

Доказательство. Если оператор A_0 компактен, то сопряжённый ему A^* тоже компактен ([23], гл.4, §6, т. 3). Тогда оператор $B = A^*A_0$ — самосопряжённый компактный, и по теореме Гильберта—Шмидта ([23], гл.4, §6, т. 5) существуют монотонно стремящаяся к нулю последовательность его собственных значений $\{\lambda_n\}_{n=1}^\infty$ и последовательность соответствующих им собственных векторов $\{w_n\}_{n=1}^\infty$ такие, что произвольный вектор $x \in H$ может быть представлен в виде $x = \sum_{j=1}^\infty \alpha_j w_j + x'$, где $x' \in \text{Ker } B$, и $Bx = \sum_{j=1}^\infty \alpha_j \lambda_j w_j$.

Поскольку уравнение $Aq = f$ имеет единственное решение, оператор A_0 (а значит и A^*) инъективен, и ядро тривиально. Следовательно $x' = 0$, и $\{w_n\}_{n=1}^\infty$ — базис в H .

Пусть $q \in H$ — произвольный вектор из H . Разложим $q - q^*$ по базису:

$$q - q^* = \sum_{n=1}^\infty \xi_n w_n.$$

Тогда можно вычислить градиент

$$\nabla J(q) = B(q - q^*) = \sum_{n=1}^\infty \lambda_n \xi_n w_n$$

и функционал

$$J(q) = \frac{1}{2} \|A_0(q - q^*)\|^2 = \frac{1}{2} \langle q - q^*, B(q - q^*) \rangle = \frac{1}{2} \sum_{n=1}^\infty \lambda_n \xi_n^2.$$

Применим в точке q градиентный спуск с шагом $\alpha \leq \frac{1}{L} = \frac{1}{\lambda_1}$:

$$q_{new} = q - \alpha \nabla J(q) = q^* + \sum_{n=1}^\infty \xi_n (1 - \alpha \lambda_n) w_n.$$

То есть, каждая компонента просто умножается на своё число из полуинтервала $[0; 1)$, потому что $1 > 1 - \alpha \lambda_n \geq 1 - \frac{\lambda_n}{\lambda_1} \geq 0$.

Пусть $q_0 = q^* + \sum_{n=1}^{\infty} \xi_n w_n$ — это начальное приближение. Тогда после k шагов получается приближение

$$q_k = q^* + \sum_{n=1}^{\infty} \xi_n (1 - \alpha \lambda_n)^k w_n.$$

Расстояние от него до точного решения вычисляется стандартно:

$$\|q_k - q^*\|^2 = \sum_{n=1}^{\infty} \xi_n^2 (1 - \alpha \lambda_n)^{2k}.$$

Этот ряд сходится, в том числе и при $k = 0$, поскольку ξ_k — коэффициенты Фурье некоторого элемента H .

Из сходимости ряда следует, что $\forall \varepsilon > 0 \exists N = N(\varepsilon) \in \mathbb{N} : \sum_{n=N+1}^{\infty} \xi_n^2 < \frac{\varepsilon}{2}$.

Для всех $n \in \{1, \dots, N\}$ существуют такие $K_n \in \mathbb{N}$, что при $k \geq K_n$ $\xi_n^2 (1 - \alpha \lambda_n)^{2k} < \frac{\varepsilon}{2N}$. Обозначим $K = \max\{K_1, \dots, K_N\}$. По определению, при фиксированной последовательности ξ_n (коэффициентов разложения начального приближения) оно зависит только от выбора ε .

Соответственно, $\forall \varepsilon > 0 \exists K = K(\varepsilon) \in \mathbb{N} : \forall k \geq K \sum_{n=1}^{\infty} \xi_n^2 (1 - \alpha \lambda_n)^{2k} = \sum_{n=1}^N \xi_n^2 (1 - \alpha \lambda_n)^{2k} + \sum_{n=N+1}^{\infty} \xi_n^2 (1 - \alpha \lambda_n)^{2k} < N \cdot \frac{\varepsilon}{2N} + \frac{\varepsilon}{2} = \varepsilon$.

А это и есть определение нулевого предела. \square

Теорема 3 (О сходимости по аргументу ∞ -моментного метода минимальных ошибок). *В условиях теоремы 2, ∞ -моментный метод минимальных ошибок сходится по аргументу к точному решению, то есть*

$$\lim_{k \rightarrow \infty} \|q_k - q^*\| = 0.$$

Доказательство. Пусть q_0 — начальное приближение; $\{q_k\}_{k=0}^{\infty}$ — последовательность точек, порождаемых ∞ -моментным методом минимальных ошибок; $\{\tilde{q}_k\}_{k=0}^{\infty}$ — последовательность точек, порождаемых простейшим градиентным спуском. По теореме 1, $\|q_k - q^*\| \leq \|\tilde{q}_k - q^*\|$ при всех натуральных k . По теореме 2, $\|\tilde{q}_k - q^*\| \rightarrow 0$ при $k \rightarrow \infty$. Из теоремы о двух милиционерах получаем требуемое. \square

Строгое доказательство этого простого следствия необходимо. Ведь в конечномерном пространстве максимальная по включению линейно независимая система образует базис (что гарантирует достижение решения рассматриваемым

методом), а в бесконечномерном пространстве гипотетически возможна ситуация, когда $q_0 - q^*$ не раскладывается по системе $\{g, Bg, \dots\}$, и достижение q^* оказывается невозможным. Данная теорема гарантирует сходимость к решению, если оно единственное.

Если же решений больше одного (для квадратичной задачи это обозначает существование линейного многообразия решений), то метод будет сходиться к одному из них.

Теорема 4 (О скорости сходимости по функционалу m -моментного метода минимальных ошибок). *m -моментный метод минимальных ошибок ($m \geq 0$) сходится по функционалу со сверхлинейной скоростью, то есть*

$$J(q_k) = o\left(\frac{1}{k}\right), k \rightarrow \infty.$$

Доказательство. Сначала проведём вычисления для метода минимальных ошибок ($m = 0$).

$$\begin{aligned} J(q_k) &= \frac{1}{2}\alpha_k\|s_k\|^2 = \frac{1}{2\alpha_k}\|q_{k+1} - q_k\|^2 = \frac{\|\nabla J(q_k)\|^2}{4J(q_k)}\|q_{k+1} - q_k\|^2 \Rightarrow \\ &\Rightarrow J(q_k) = \frac{1}{2}\|\nabla J(q_k)\| \cdot \|q_{k+1} - q_k\|. \end{aligned}$$

Поскольку для функционала с липшицевым градиентом выполняется неравенство

$$J(q) \geq \frac{\|\nabla J(q)\|^2}{2L},$$

можно получить следующую оценку:

$$J(q_k) \leq \frac{1}{2}\sqrt{2LJ(q_k)} \cdot \|q_{k+1} - q_k\| \Rightarrow J(q_k) \leq \frac{L}{2}\|q_{k+1} - q_k\|^2,$$

откуда и вытекает требуемая оценка скорости сходимости: из сходимости ряда (2.1.2.2) следует стремление общего члена к нулю быстрее $\frac{1}{k}$.

При $m \geq 1$ вычисления очень похожи:

$$\begin{aligned} J(q_k) &= \frac{1}{2}\alpha_k\|s_k\|^2 = \frac{1}{2\alpha_k}\|q_{k+1} - q_k\|^2 = \frac{\|\nabla J(q_k)\|^2 \sin^2 \varphi_k}{4J(q_k)}\|q_{k+1} - q_k\|^2 \Rightarrow \\ &\Rightarrow J(q_k) = \frac{1}{2}\|\nabla J(q_k)\| \cdot \|q_{k+1} - q_k\| \sin \varphi_k. \end{aligned}$$

Подставляя оценку нормы градиента, получаем:

$$J(q_k) \leq \frac{L}{2}\|q_{k+1} - q_k\|^2 \sin^2 \varphi_k.$$

Это обозначает сходимость не хуже метода минимальных ошибок. \square

Теорема 5 (О скорости сходимости по аргументу t -моментного метода минимальных ошибок). *Метод минимальных ошибок для сильно выпуклого квадратичного функционала сходится по аргументу со скоростью геометрической прогрессии, то есть*

$$\|q_{k+1} - q^*\|^2 \leq \left(1 - \frac{\mu}{L}\right) \cdot \|q_k - q^*\|^2.$$

Для t -моментного метода минимальных ошибок оценка сходимости улучшается:

$$\|q_{k+1} - q^*\|^2 \leq \left(1 - \frac{\mu}{L \sin^2 \varphi_k}\right) \cdot \|q_k - q^*\|^2,$$

где φ_k — угол между $-\nabla J(q_k)$ и проекцией $-\nabla J(q_k)$ на $\text{Lin}\{h_{k-1}, \dots, h_{k-m}\}$.

Доказательство. Воспользуемся оценками, полученными при доказательстве прошлой теоремы, и свойством сильно выпуклого функционала

$$J(q) \geq \frac{\mu\|q - q^*\|^2}{2}.$$

Для простого метода минимальных ошибок:

$$\begin{aligned} \|q_{k+1} - q_k\|^2 &\geq \frac{2J(q_k)}{L} \geq \frac{\mu\|q_k - q^*\|^2}{L} \Rightarrow \\ &\Rightarrow \|q_{k+1} - q^*\|^2 = \|q_k - q^*\|^2 - \|q_{k+1} - q_k\|^2 \leq \left(1 - \frac{\mu}{L}\right) \cdot \|q_k - q^*\|^2, \end{aligned}$$

что и требовалось доказать.

Эта оценка соответствует обычным оценкам сходимости для градиентных методов.

Для t -моментного метода минимальных ошибок:

$$\begin{aligned} \|q_{k+1} - q_k\|^2 &\geq \frac{2J(q_k)}{L \sin^2 \varphi_k} \geq \frac{\mu\|q_k - q^*\|^2}{L \sin^2 \varphi_k} \Rightarrow \\ &\Rightarrow \|q_{k+1} - q^*\|^2 = \|q_k - q^*\|^2 - \|q_{k+1} - q_k\|^2 \leq \left(1 - \frac{\mu}{L \sin^2 \varphi_k}\right) \cdot \|q_k - q^*\|^2, \end{aligned}$$

что и требовалось доказать.

Эта оценка лучше предыдущей. Убывание расстояния до точного решения зависит от текущей величины угла между антиградиентом и его проекцией на линейную оболочку векторов t предыдущих шагов. Если этот угол близок к $\frac{\pi}{2}$, то для данного шага оценка соответствует обычной оценке для градиентных методов. Если же угол мал, то наблюдается очень быстрое убывание. \square

Получены оценки сходимости в сильно выпуклом случае; в конечномерном пространстве доказана сходимость за конечное число шагов. Возникает вопрос, нет ли хороших глобальных оценок скорости сходимости по аргументу и для не сильно выпуклых задач оптимизации, которые возникают при решении некорректных задач? Отрицательный ответ на этот вопрос даёт следующая теорема.

Теорема 6 (О неравномерной сходимости по аргументу; или о существовании начального приближения, при котором сходимость по аргументу оказывается сколь угодно медленной). *В условиях теоремы 2, для любых наперёд заданных $\varepsilon \in (0, 1)$ и $N \in \mathbb{N}$ существует такое начальное приближение q_0 , что $\|q_0 - q^*\| = 1$, а для любого метода, работающего в подпространствах Крылова, в том числе и для ∞ -моментного метода минимальных ошибок, $\|q_N - q^*\|^2 > \varepsilon$.*

Доказательство. Зафиксируем произвольное натуральное число N и $\varepsilon \in (0, 1)$.

$\nabla J(q_0) = B(q_0 - q^*)$, поэтому $K_N = \text{Lin}(\{B^k(q_0 - q^*)\}_{k=1}^N)$. При доказательстве теоремы 2 получен ортонормированный базис из собственных векторов B , будем использовать разложение по нему:

$$q_0 - q^* = \sum_{n=1}^{\infty} \xi_n w_n; \quad B^k(q_0 - q^*) = \sum_{n=1}^{\infty} \lambda_n^k \xi_n w_n.$$

Будем считать, что каждому λ_n соответствует только один собственный вектор, то есть последовательность собственных значений — строго убывающая. Если это не так, и каким-то собственным значениям соответствует несколько собственных векторов, то будем рассматривать лишь такие начальные приближения q_0 , для которых не более одного коэффициента ξ_n при векторе из такого набора отличается от нуля. Или просто (аналогично методу вспомогательного угла из школьной тригонометрии) представим линейную комбинацию собственных векторов, соответствующих одному и тому же собственному значению, в виде одного такого вектора, умноженного на число: $\sum_{i=1}^k \xi_{n_i} w_{n_i} = r \sum_{i=1}^k \frac{\xi_{n_i}}{r} w_{n_i}$, где

$$r = \sqrt{\sum_{i=1}^k \xi_{n_i}^2}.$$

Любой метод, работающий в подпространствах Крылова, порождает на шаге N точку $q_N = q_0 + \sum_{i=1}^N \eta_i B^i(q_0 - q^*)$ для некоторого набора чисел η_1, \dots, η_N . Преобразуем расстояние до точного решения:

$$\|q_N - q^*\|^2 = \|q_0 - q^* + \sum_{i=1}^N \eta_i B^i(q_0 - q^*)\|^2 = \left\| \sum_{n=1}^{\infty} \xi_n \left(1 + \sum_{i=1}^N \eta_i \lambda_n^i \right) w_n \right\|^2 =$$

$$= \sum_{n=1}^{\infty} \xi_n^2 \left(1 + \sum_{i=1}^N \eta_i \lambda_n^i \right)^2 =: \Psi(\eta; q_0).$$

$\Psi(\eta; q_0)$ — квадратичная функция от N -мерного вектора $\eta = (\eta_1, \dots, \eta_N)^T$, зависящая от параметра q_0 . Наша цель — подобрать такие ξ_n , что $\sum_{n=1}^{\infty} \xi_n^2 = 1$, а

$$\min_{\eta \in \mathbb{R}^n} \Psi \left(\eta; \sum_{n=1}^{\infty} \xi_n w_n \right) > \varepsilon. \text{ Тогда утверждение теоремы будет выполнено.}$$

Заметим: если количество ненулевых ξ_n не превосходит N , то $\min_{\eta \in \mathbb{R}^n} \Psi \left(\eta; \sum_{n=1}^{\infty} \xi_n w_n \right) = 0$. Действительно, для обнуления $\Psi \left(\eta; \sum_{n=1}^{\infty} \xi_n w_n \right)$ необходимо и достаточно выполнение равенства $1 + \sum_{i=1}^N \eta_i \lambda_n^i = 0$ для всех таких n , что $\xi_n \neq 0$ (обычное условие стационарности через частные производные, естественно, при этом выполняется). Если количество уравнений в этой системе оказалось меньше N , дополним её до этого количества и другими номерами. Определитель получившейся системы — это известный определитель Вандермонда (помноженный на произведение λ_n), а он не равен нулю при различных положительных λ_n . Поэтому решение существует.

Возьмём $\xi = \xi_1 = \xi_2 = \dots = \xi_N = \sqrt{\frac{1-\varepsilon}{2N}}$, $\xi_M = \sqrt{\frac{1+\varepsilon}{2}}$, где $M > N$ — число, которое будет выбрано позднее; все остальные ξ_n равны нулю. Условие нормировки выполняется. Обозначим $\tilde{\eta}$ — набор, при котором $\Psi \left(\tilde{\eta}; \sum_{n=1}^N \xi w_n \right) = 0$. По доказанному в предыдущем абзаце, он существует.

Запишем необходимое условие минимальности $\Psi \left(\eta; \sum_{n=1}^N \xi w_n + \xi_M w_M \right)$ в частных производных:

$$\begin{aligned} \frac{1}{2} \frac{\partial}{\partial \eta_j} \Psi \left(\eta; \sum_{n=1}^N \xi w_n + \xi_M w_M \right) &= \frac{1}{2} \frac{\partial}{\partial \eta_j} \left(\xi^2 \sum_{n=1}^N \left(1 + \sum_{i=1}^N \eta_i \lambda_n^i \right)^2 + \xi_M^2 \left(1 + \sum_{i=1}^N \eta_i \lambda_M^i \right)^2 \right) = \\ &= \xi^2 \sum_{n=1}^N \left(1 + \sum_{i=1}^N \eta_i \lambda_n^i \right) \lambda_n^j + \xi_M^2 \left(1 + \sum_{i=1}^N \eta_i \lambda_M^i \right) \lambda_M^j = 0 \text{ при } 1 \leq j \leq N. \end{aligned} \tag{2.3.5.1}$$

Обозначим вектор $(1, \lambda_n, \dots, \lambda_n^{N-1})^T =: x_n$. Тогда система уравнений (2.3.5.1) принимает вид

$$\left(\sum_{n=1}^N \xi^2 \lambda_n^2 x_n x_n^T + \xi_M^2 \lambda_M^2 x_M x_M^T \right) \eta = -\xi^2 \sum_{n=1}^N \lambda_n x_n - \xi_M^2 \lambda_M x_M$$

Подставим значения ξ и ξ_M , а затем домножим левую и правую части на одно и то же число:

$$\left(\sum_{n=1}^N \lambda_n^2 x_n x_n^T + N \frac{1+\varepsilon}{1-\varepsilon} \lambda_M^2 x_M x_M^T \right) \eta = - \sum_{n=1}^N \lambda_n x_n - N \frac{1+\varepsilon}{1-\varepsilon} \lambda_M x_M. \quad (2.3.5.2)$$

Матрица $L = \sum_{n=1}^N \lambda_n^2 x_n x_n^T$ — положительно определённая. Действительно:

предположим, что для некоторого $y \neq \bar{0}$ $y^T L y = 0$. Тогда $\sum_{n=1}^N \lambda_n^2 \langle x_n, y \rangle^2 = 0$.

Поскольку все $\lambda_n > 0$, это возможно, только если $\langle x_n, y \rangle = 0$ при всех n . Но векторы x_n , как уже было замечено, — это N линейно независимых векторов в \mathbb{R}^N . Поэтому ненулевой $y \in \mathbb{R}^N$ не может быть ортогонален им всем.

Значит, система (2.3.5.2) имеет единственное решение, поскольку её матрица — это сумма положительно определённой L и неотрицательно определённой матрицы, заданной последним слагаемым. Обозначим это решение $\hat{\eta}(M)$.

Поскольку набор $\tilde{\eta}$ обнуляет все остальные слагаемые в определении,

$$\Psi \left(\tilde{\eta}; \sum_{n=1}^N \xi w_n + \xi_M w_M \right) = \left\| \xi_M \left(1 + \sum_{i=1}^n \tilde{\eta}_i \lambda_M^i \right) w_i \right\|^2 = \frac{1+\varepsilon}{2} \left(1 + \sum_{i=1}^n \tilde{\eta}_i \lambda_M^i \right)^2.$$

При $M \rightarrow \infty \lambda_M \rightarrow 0$, поэтому выражение в скобке стремится к 1. Соответственно, при достаточно больших M выполняется оценка

$$\Psi \left(\tilde{\eta}; \sum_{n=1}^N \xi w_n + \xi_M w_M \right) > \frac{1+3\varepsilon}{4} \quad (2.3.5.3)$$

Известно, что при положительно определённой матрице решение системы линейных алгебраических условий непрерывно зависит от левой и правой частей: если $\delta L \rightarrow 0$ и $\delta b \rightarrow 0$ по некоторой норме (например, по евклидовой; хотя в конечномерном пространстве, в котором решается система (2.3.5.2), они все эквивалентны), то решение системы $(L+\delta L)\eta = b+\delta b$ стремится к решению $L\eta = b$ по той же норме.

Обозначим слагаемые в (2.3.5.2), содержащие номер M , за δL и δb . Заметим также, что при $M \rightarrow \infty \lambda_M \rightarrow 0$, $\|x_M\| \rightarrow 1$. Значит, $\|\hat{\eta}(M) - \tilde{\eta}\| \rightarrow 0$, и, поскольку $\Psi(\eta; q)$ — непрерывная функция, $\left| \Psi \left(\hat{\eta}(M); \sum_{n=1}^N \xi w_n \right) - \Psi \left(\tilde{\eta}; \sum_{n=1}^N \xi w_n \right) \right| \rightarrow 0$.

$$\left| \Psi \left(\eta; \sum_{n=1}^N \xi w_n + \xi_M w_M \right) - \Psi \left(\eta; \sum_{n=1}^N \xi w_n \right) \right| = \frac{1+\varepsilon}{2} \left(1 + \sum_{i=1}^n \eta_i \lambda_M^i \right)^2.$$

При $M \rightarrow \infty$ эта величина стремится к $\frac{1+\varepsilon}{2}$, даже если η не фиксирован, а зависит от номера M и лишь ограничен по норме.

Значит, $\Psi\left(\hat{\eta}(M); \sum_{n=1}^N \xi w_n + \xi_M w_M\right) - \Psi\left(\tilde{\eta}; \sum_{n=1}^N \xi w_n + \xi_M w_M\right) = \Delta_1 + \Delta_2 - \Delta_3$,
где

$$\begin{aligned}\Delta_1 &= \Psi\left(\hat{\eta}(M); \sum_{n=1}^N \xi w_n\right) - \Psi\left(\tilde{\eta}; \sum_{n=1}^N \xi w_n\right) \rightarrow 0 \text{ при } M \rightarrow \infty; \\ \Delta_2 &= \Psi\left(\hat{\eta}(M); \sum_{n=1}^N \xi w_n + \xi_M w_M\right) - \Psi\left(\hat{\eta}(M); \sum_{n=1}^N \xi w_n\right) \rightarrow \frac{1+\varepsilon}{2} \text{ при } M \rightarrow \infty; \\ \Delta_3 &= \Psi\left(\tilde{\eta}; \sum_{n=1}^N \xi w_n + \xi_M w_M\right) - \Psi\left(\tilde{\eta}; \sum_{n=1}^N \xi w_n\right) \rightarrow \frac{1+\varepsilon}{2} \text{ при } M \rightarrow \infty.\end{aligned}$$

Соответственно, $\Psi\left(\hat{\eta}(M); \sum_{n=1}^N \xi w_n + \xi_M w_M\right) - \Psi\left(\tilde{\eta}; \sum_{n=1}^N \xi w_n + \xi_M w_M\right) \rightarrow 0$
при $M \rightarrow \infty$. Поэтому, начиная с некоторого M , выполняется оценка

$$\left| \Psi\left(\hat{\eta}(M); \sum_{n=1}^N \xi w_n + \xi_M w_M\right) - \Psi\left(\tilde{\eta}; \sum_{n=1}^N \xi w_n + \xi_M w_M\right) \right| < \frac{1-\varepsilon}{4} \quad (2.3.5.4)$$

Выберем такое большое M , чтобы выполнялись обе оценки (2.3.5.3) и (2.3.5.4).

Тогда для $\hat{\eta}(M)$ — точки минимума функции Ψ — выполняется следующая оценка:
 $\Psi\left(\hat{\eta}(M); \sum_{n=1}^N \xi w_n + \xi_M w_M\right) > \frac{1+3\varepsilon}{4} - \frac{1-\varepsilon}{4} = \varepsilon$, что и требовалось доказать. \square

Эта теорема показывает, что, хотя для любого начального приближения и достигается сходимость ∞ -моментного метода минимальных ошибок (то есть, стремление к нулю расстояния до точного решения), равномерной для всех точек оценки этой сходимости не существует. За фиксированное количество шагов сокращение расстояния может быть сколь угодно слабым.

Важно отметить, что доказанное свойство сколь угодно медленной сходимости относится не к ∞ -моментному методу минимальных ошибок: на нём используемый в доказательстве минимум реализуется, а для любого другого метода первого порядка сходимость будет ещё медленнее (теорема 1). Оно относится именно к бесконечномерному пространству, в котором решаются задачи оптимизации.

2.3.6. О применении построенных методов в условиях неточных вычислений

Все рассматриваемые задачи оптимизации поставлены в бесконечномерных функциональных пространствах. Решение вспомогательных прямых задач требует

решения дифференциальных или интегральных уравнений. Они чаще всего не могут быть решены точно, поэтому используются приближённые вычисления. Соответственно, и функционал, и градиент вычисляются с некоторой погрешностью.

Также погрешность возникает из-за неточности измерений на наблюдаемой границе, результатом которых является правая часть операторного уравнения.

Проблема влияния погрешности вычисления градиента обсуждается, например, в статье [57]. Там проводится исследование влияния шума на сходимость метода подобных треугольников. Наличие шума может приводить даже к расходимости градиентных методов, если шум имеет аддитивный характер (это показано Б. Т. Поляком в статье [55]). При этом относительный шум в градиенте приводит лишь к линейному замедлению сходимости метода градиентного спуска ([54]).

Построенный t -моментный метод минимальных ошибок делает каждый шаг в направлении проекции антиградиента на линейную оболочку предыдущих шагов. Длина шага вычисляется по общей формуле, вывод которой существенно опирается на такой выбор направления. В этой формуле значение функционала находится в числителе, а норма проекции антиградиента — в знаменателе. Соответственно, влияние погрешности вычислений может быть значительным — как из-за отклонения направления шага от оптимального, так и из-за деления малых чисел.

Оценить это влияние теоретически не представляется возможным, поэтому применимость метода в реальных условиях можно проверить только экспериментально. Эта проверка выполнена в главе 3.

Глава 3.

Применение к обратным задачам математической физики

Данная глава посвящена решению конкретных обратных задач математической физики с применением градиентных методов оптимизации. Для каждой задачи вычислен градиент; проведены эксперименты, в которых сравнивается работа различных методов оптимизации.

По результатам всех экспериментов построены графики в логарифмическом масштабе по вертикальной оси, вдоль которой откладываются расстояние до точного решения и значение функционала. Вдоль горизонтальной оси откладывается номер итерации.

3.1. Начально-краевая задача для уравнения Гельмгольца

3.1.1. Происхождение, постановка и некорректность задачи

Результаты, в основном, были изложены автором в статье [32]; в части, относящейся к применению метода покомпонентного спуска и m -моментного метода минимальных ошибок, — в статьях [33] и [34].

Задача (1.5), которой посвящён данный раздел, предложена доктором физико-математических наук М. А. Шишлениным. Похожая задача решалась в статье [52].

Покажем некорректность задачи (1.5). Взяв $\varepsilon = 1$ и зафиксировав $\delta > 0$, рассмотрим такие начальные условия: $f(y) = \frac{\delta}{\pi n} \sin \pi n y$ и $g(y) = \sqrt{\pi^2 n^2 - \kappa^2} f(y)$ (n — натуральное число, которое будет выбрано позднее). Для них решение задачи Коши с $r(x, y) = 0$ — функция $u(x, y) = \frac{\delta}{\pi n} e^{\sqrt{\pi^2 n^2 - \kappa^2} x} \sin \pi n y$, что легко проверяется по определению.

$\|f\| = \frac{\delta}{\pi n} < \delta$; $\|g\| = \frac{\delta}{\pi n} \sqrt{\pi^2 n^2 - \kappa^2} < \delta$; $\|u\| = \frac{\delta}{\pi n} e^{\sqrt{\pi^2 n^2 - \kappa^2}} \rightarrow +\infty$ при $n \rightarrow +\infty$, поэтому найдётся n такое, что $\|u\| > \varepsilon$. Это и обозначает отсутствие непрерывности решения по начальным условиям, то есть некорректность поставленной задачи.

3.1.2. Операторы задачи

Начально-краевая задача для уравнения Гельмгольца некорректна, но корректна краевая задача:

$$\begin{cases} \Delta u + \kappa^2 u = r(x, y), & (x, y) \in \Omega = (0, 1) \times (0, 1) \\ u_x|_{x=0} = g(y), & y \in [0, 1] \\ u|_{y=0} = u|_{y=1} = 0, & x \in [0, 1] \\ u|_{x=1} = q(y), & y \in [0, 1] \end{cases}. \quad (3.1.2.1)$$

Эта задача отличается от исходной лишь заменой условия $u|_{x=0} = f(y)$ на $u|_{x=1} = q(y)$. Граница $x = 0$ — наблюдаемая, $x = 1$ — ненаблюдаемая.

Уточним условия гладкости. Для этого воспользуемся представлением решения задачи в виде ряда по собственным функциям стационарного оператора « $-\Delta$ » (подробнее о базисе — в разделе 3.1.3). Будем рассматривать случай $r(x, y) = 0$ и, учитывая граничные условия при $y \in \{0, 1\}$, представим функции в виде рядов Фурье: $g(y) = \sum_{m=1}^{\infty} g_m \sin \pi m y$, $q(y) = \sum_{m=1}^{\infty} q_m \sin \pi m y$. Для непрерывно дифференцируемых функций g, q коэффициенты Фурье необходимо удовлетворяют соотношениям: $g_m = O(m^{-2})$, $q_m = O(m^{-2})$.

Ищем решение в виде $u(x, y) = \sum_{m=1}^{\infty} X_m(x) \sin \pi m y$, как это сделано в статье [32].

После несложных преобразований получаем (обозначив $\gamma_m = \sqrt{\pi^2 m^2 - \kappa^2}$ для краткости), что $X_m(x) = \left(\frac{q_m}{\operatorname{ch} \gamma_m} - \frac{g_m}{\gamma_m} \operatorname{th} \gamma_m \right) \operatorname{ch} \gamma_m x + \frac{g_m}{\gamma_m} \operatorname{sh} \gamma_m x$.

Данный ряд сходится, причём $|X_m(x)| \leq |q_m| + |g_m|$.

В частности, $u(0, y) = \sum_{m=1}^{\infty} \left(\frac{q_m}{\operatorname{ch} \gamma_m} - \frac{g_m}{\gamma_m} \operatorname{th} \gamma_m \right) \sin \pi m y$.

При $m \rightarrow \infty$ $\operatorname{th} \gamma_m \rightarrow 1$, $\operatorname{ch} \gamma_m \sim \frac{e^{\pi m}}{2}$. Поэтому коэффициенты полученного ряда сходятся к нулю, как минимум, со скоростью $O(m^{-3})$. Следовательно, его сумма — непрерывно дифференцируемая функция, принимающая нулевые значения на концах отрезка $[0, 1]$.

Отметим, что данный способ решения уравнений является стандартным. Например, похожее рассуждение используется при доказательстве существования и единственности решения в статье [39].

Итак, на пространстве непрерывно дифференцируемых функций

$$H = \{q \in C^1[0, 1] : q(0) = q(1) = 0\} \subset L_2[0, 1]$$

со стандартным скалярным произведением можно определить оператор $A : H \rightarrow H$ следующим образом:

$$(Aq)(y) = u(0, y) \quad \forall y \in (0, 1), \quad (3.1.2.2)$$

где u — решение корректной краевой задачи (3.1.2.1).

Из изложенного вытекает, что для существования решения задачи $Aq = f$ необходимо (но не достаточно), чтобы коэффициенты Фурье функции $f(y)$ убывали со скоростью $O(m^{-3})$. Докажем достаточное условие.

Теорема 7 (О разрешимости и условной корректности однородной начально-краевой задачи для уравнения Гельмгольца). *Если $f(y) \in H_c$, где*

$$H_c = \left\{ \sum_{m=1}^{\infty} f_m \sin \pi m y \mid f_m = O(m^{-3} e^{-\pi m}) \text{ при } m \rightarrow \infty \right\},$$

то задача 1.5 с $r(x, y) = g(y) = 0$ имеет решение и является условно корректной на множестве

$$M = \left\{ \sum_{m=1}^{\infty} q_m \sin \pi m y \mid q_m = O(m^{-3}) \text{ при } m \rightarrow \infty \right\}.$$

Доказательство. Формула для $u(0, y)$, полученная выше, связывает коэффициенты Фурье функций $q(y)$ и $f(y)$ при $g(y) = 0$ следующим образом: $q_m = f_m \cdot \operatorname{ch} \gamma_m$.

Поскольку $\operatorname{ch} \gamma_m \cdot e^{-\pi m} \rightarrow \frac{1}{2}$ при $m \rightarrow \infty$, $\exists G > 0$: $\operatorname{ch} \gamma_m \cdot e^{-\pi m} < G$ для всех m .

Если $f(y) \in H_c$, то $q_m = f_m \cdot \operatorname{ch} \gamma_m = O(m^{-3})$. Тогда ряд Фурье $\sum_{m=1}^{\infty} q_m \sin \pi m y$, а также его формальная производная, сходится по признаку Вейерштрасса абсолютно и равномерно к непрерывно дифференцируемой функции. Следовательно, он является решением задачи $Aq = f$.

Учитывая связь между коэффициентами Фурье для $q(y)$ и $f(y)$, можно заметить, что $H_c = A(M)$.

Введём на H_c естественную норму: $\|x\|_{H_c} = \sup_{m \in \mathbb{N}} |x_m| m^3 e^{\pi m}$.

Рассмотрим две обратных задачи $Aq = f^{(1)}$, $Aq = f^{(2)}$ с решениями $q^{(1)}$ и $q^{(2)}$. Пусть $\tilde{f} = f^{(1)} - f^{(2)}$, $\tilde{q} = q^{(1)} - q^{(2)}$.

Если $\|\tilde{f}\|_{H_c} < \delta$, то $|\tilde{f}_m| < \delta m^{-3} e^{-\pi m}$ для всех m .

Тогда $|\tilde{q}_m| < C\delta m^{-3}$, поэтому ряд $\sum_{m=1}^{\infty} \tilde{q}_m \sin \pi m y$ равномерно сходится, причём его сумма не превышает по модулю $2C\delta$ и является непрерывно дифференцируемой.

Взяв для произвольного $\varepsilon > 0$ $\delta = \frac{\varepsilon}{2C}$, получаем: если $\|\tilde{f}\|_{H_c} < \delta$, то $\|\tilde{q}\| < \varepsilon$. А это и есть определение условной устойчивости, то есть непрерывности решения по начальным данным на M . \square

Следуя методу из пособия [4], вычислим сопряжённый оператор и, соответственно, градиент функционала задачи. Эти вычисления были проведены в статье [32].

Итак,

$$J(q) = \frac{1}{2} \int_0^1 |u(0, y) - f(y)|^2 dy,$$

где u — решение задачи (3.1.2.1).

Запишем функцию Лагранжа, взяв множители для первых двух строк, а две последние считая задающими «простое» множество:

$$\begin{aligned} L(q, u; \psi = (\psi(x, y), \lambda(y))) &= \frac{1}{2} \int_0^1 |u(0, y) - f(y)|^2 dy - \\ &- \int_0^1 \int_0^1 (u_{xx} + u_{yy} + \kappa^2 u - r(x, y)) \psi(x, y) dx dy + \int_0^1 (q(y) - u(1, y)) \lambda(y) dy. \end{aligned}$$

Для получения сопряжённой задачи на ψ варьируем по $\delta u(x, y)$:

$$\begin{aligned} \delta_u L(q, u; \psi) &= \int_0^1 (u(0, y) - f(y)) \delta u(0, y) dy - \\ &- \int_0^1 \int_0^1 (\delta u_{xx} + \delta u_{yy} + \kappa^2 \delta u) \psi(x, y) dx dy - \int_0^1 \delta u(1, y) \lambda(y) dy. \end{aligned}$$

При этом

$$\delta u_x(0, y) = \delta u(x, 0) = \delta u(x, 1) = 0.$$

Интегрируем по частям слагаемые с вторыми производными:

$$\begin{aligned}
& \int_0^1 \int_0^1 \psi(x, y) \delta u_{xx} dx dy = \int_0^1 \psi(1, y) \delta u_x(1, y) dy - \\
& - \int_0^1 (\psi_x(1, y) \delta u(1, y) - \psi_x(0, y) \delta u(0, y)) dy + \int_0^1 \int_0^1 \psi_{xx} \delta u(x, y) dx dy; \\
& \int_0^1 \int_0^1 \psi(x, y) \delta u_{yy} dx dy = \int_0^1 \psi(x, 1) \delta u_y(x, 1) dx - \\
& - \int_0^1 \psi(x, 0) \delta u_y(x, 0) dx + \int_0^1 \int_0^1 \psi_{yy} \delta u(x, y) dx dy.
\end{aligned}$$

Подставляем и вычисляем вариацию функции Лагранжа:

$$\begin{aligned}
\delta_u L(q, u; \psi) = & - \int_0^1 \int_0^1 (\psi_{xx} + \psi_{yy} + \kappa^2 \psi) \delta u(x, y) dx dy + \\
& + \int_0^1 (u(0, y) - f(y) - \psi_x(0, y)) \delta u(0, y) dy + \int_0^1 (\psi_x(1, y) - \lambda(y)) \delta u(1, y) dy - \\
& - \int_0^1 \psi(1, y) \delta u_x(1, y) dy + \int_0^1 \psi(x, 0) \delta u_y(x, 0) dx - \int_0^1 \psi(x, 1) \delta u_y(x, 1) dx.
\end{aligned}$$

Из необходимого условия экстремума получаем систему уравнений:

$$\begin{cases} \psi_{xx} + \psi_{yy} + \kappa^2 \psi = 0, & (x, y) \in (0, 1) \times (0, 1) \\ \psi|_{y=0} = \psi|_{y=1} = 0, & x \in [0, 1] \\ \psi|_{x=1} = 0, & y \in [0, 1] \\ \psi_x|_{x=0} = u(0, y) - f(y), & y \in [0, 1] \\ \psi_x|_{x=1} = \lambda(y), & y \in [0, 1] \end{cases}$$

Последнее уравнение — связь между множителями Лагранжа. Поскольку

$$\lambda(y) = L_q(y) = \nabla J(q)(y)$$

(формула Демьянова—Данскина—Рубинова [10], [48]), вычисление градиента функционала свелось к решению двух краевых задач.

Оператор $A^* : H \rightarrow H$ задаётся следующим образом:

$$(A^*p)(y) = \psi_x(1, y) \quad \forall y \in (0, 1), \quad (3.1.2.3)$$

где ψ — решение задачи

$$\begin{cases} \psi_{xx} + \psi_{yy} + \kappa^2\psi = 0, & (x, y) \in (0, 1) \times (0, 1) \\ \psi|_{y=0} = \psi|_{y=1} = 0, & x \in [0, 1] \\ \psi|_{x=1} = 0, & y \in [0, 1] \\ \psi_x|_{x=0} = p(y), & y \in [0, 1] \end{cases}. \quad (3.1.2.4)$$

Как видно из определения, оператор A (в отличие от сопряжённого к нему A^*) не является линейным из-за неоднородных краевых условий в определяющей его задаче. Однако линейным является следующий оператор $A_0 : H \rightarrow H$:

$$(A_0q)(y) = u(0, y) \quad \forall y \in (0, 1), \quad (3.1.2.5)$$

где u — решение краевой задачи, в которой все краевые условия, кроме q , равны нулю:

$$\begin{cases} u_{xx} + u_{yy} + \kappa^2u = 0, & (x, y) \in (0, 1) \times (0, 1) \\ u|_{x=1} = q(y), & y \in [0, 1] \\ u_x|_{x=0} = 0, & y \in [0, 1] \\ u|_{y=0} = u|_{y=1} = 0, & x \in [0, 1] \end{cases}. \quad (3.1.2.6)$$

Очевидны из определения свойства:

$$Aq = A_0 + A_0q, \quad Aq_1 - Aq_2 = A_0(q_1 - q_2).$$

Удобно ввести общее обозначение, которое позволяет кратко записать алгоритм точного вычисления градиента.

Обозначим за $\Gamma[a(y), b(y), c(x, y)](x, y)$ решение $v(x, y)$ следующей краевой задачи:

$$\begin{cases} v_{xx} + v_{yy} + \kappa^2v = c(x, y), & (x, y) \in (0, 1) \times (0, 1) \\ v|_{x=1} = a(y), & y \in [0, 1] \\ v_x|_{x=0} = b(y), & y \in [0, 1] \\ v|_{y=0} = v|_{y=1} = 0, & x \in [0, 1] \end{cases}.$$

Очевидно, что

$$(Aq)(y) = \Gamma[q(y), g(y), r(x, y)](0, y),$$

$$(A_0 q)(y) = \Gamma[q(y), 0, 0](0, y),$$

$$(A^* p)(y) = \frac{\partial}{\partial x} \Gamma[0, p(y), 0](1, y).$$

Тогда градиент вычисляется алгоритмом 3.

Algorithm 3 Gradient Helmholtz Precise Oracle

Input: $q(y) \in H; f(y), g(y), r(x, y), \kappa$ — параметры задачи

- 1: вычислить $u(x, y) := \Gamma[q(y), g(y), r(x, y)](x, y);$
- 2: вычислить $\psi(x, y) := \Gamma[0, u(0, y) - f(y), 0](x, y);$
- 3: вычислить $\nabla J(q)(y) := \psi_x(1, y).$

Output: $\nabla J(q)(y).$

3.1.3. Спектр соответствующего задаче самосопряжённого оператора

Очевидно, $A^* \neq A_0$, поэтому оператор задачи не является самосопряжённым. Однако самосопряжённым, как и в общем случае, является оператор $A^* A_0$, причём градиент липшицев с константой, равной максимальному собственному значению оператора $A^* A_0$.

Для нахождения собственных чисел ищем

$$u(x, y) = \Gamma[q(y), 0, 0](x, y) \text{ и } (A^* A_0 q)(y) = \Gamma[0, u(0, y), 0]_x(1, y)$$

в виде ряда по собственным функциям стационарного оператора, то есть в виде

$$u(x, y) = \sum_{m=1}^{\infty} X_m(x) Y_m(y).$$

Каждое слагаемое $X_m(x) Y_m(y)$ удовлетворяет уравнению Гельмгольца и условиям

$$Y_m(0) = Y_m(1) = 0.$$

Подставляя в уравнение, получаем:

$$X_m''(x) Y_m(y) + X_m(x) Y_m''(y) + \kappa^2 X_m(x) Y_m(y) = 0,$$

откуда

$$\frac{X_m''(x)}{X_m(x)} + \kappa^2 = -\frac{Y_m''(y)}{Y_m(y)} = l.$$

Краевая задача

$$Y''(y) = -l Y(y), Y(0) = Y(1) = 0$$

имеет нетривиальные решения только при

$$l_m = \pi^2 m^2, \quad Y_m(y) = \sin \pi m y, \quad m = 1, 2, \dots$$

Как известно из курса математического анализа, система функций $\{\sin \pi m y\}_{m=1}^\infty$ является базисом в пространстве функций H со стандартным скалярным произведением.

Решая соответствующие уравнения, находим собственные значения оператора $A^* A_0$:

$$\lambda_m = \begin{cases} \frac{1}{\cos^2 \sqrt{\kappa^2 - \pi^2 m^2}}, & \pi m < \kappa, \\ 1, & \pi m = \kappa, \\ \frac{1}{\operatorname{ch}^2 \sqrt{\pi^2 m^2 - \kappa^2}}, & \pi m > \kappa. \end{cases} \quad (3.1.3.1)$$

Ограничимся случаем $\kappa \leq \pi$: тогда градиент липшицев с $L \leq 1$. Поскольку $\lim_{m \rightarrow \infty} \lambda_m = 0$, функционал не является сильно выпуклым ($\mu = 0$).

Эти оценки объединяет следующая лемма.

Лемма 2. При $\kappa \leq \pi$ задача минимизации функционала $J(q)$, соответствующего задаче Коши для уравнения Гельмгольца, является выпуклой с липшицевым градиентом, спектр $\lambda_m = \frac{1}{\operatorname{ch}^2 \sqrt{\pi^2 m^2 - \kappa^2}}$ ($m \geq 1$), константа Липшица градиента $L = \frac{1}{\operatorname{ch}^2 \sqrt{\pi^2 - \kappa^2}} \leq 1$, константа сильной выпуклости $\mu = 0$.

3.1.4. Приближённые вычисления

Стандартный подход к решению краевых задач — аппроксимация уравнений в частных производных разностными схемами. Так действовали авторы статьи [52], и мы последуем их примеру.

Разностная схема для краевой задачи

Пусть на $[0, 1] \times [0, 1]$ задана сетка с шагом h , то есть

$$x = ih, \quad y = jh, \quad 0 \leq i, j \leq P \quad (Ph = 1).$$

Тогда $\Gamma_h[a_j, b_j, c_{i,j}]_{i,j}$ вводится как решение разностной схемы, аналогичной

применённой в работе [52]

$$\begin{cases} \frac{v_{i+1,j} - 2v_{i,j} + v_{i-1,j}}{h^2} + \frac{v_{i,j+1} - 2v_{i,j} + v_{i,j-1}}{h^2} + \kappa^2 v_{i,j} = c_{i,j}, & 1 \leq i, j \leq P-1 \\ v_{P,j} = a_j, & 1 \leq j \leq P-1 \\ \frac{v_{1,j} - v_{0,j}}{h} = b_j, & 1 \leq j \leq P-1 \\ v_{i,0} = v_{i,P} = 0, & 0 \leq i \leq P \end{cases} \quad (3.1.4.1)$$

При этом

$$a_j = a(jh), \quad b_j = b(jh), \quad c_{i,j} = c(ih, jh), \quad v_{i,j} = v(ih, jh).$$

Поскольку

$$\begin{aligned} v_{i\pm 1,j} &= v(ih, jh) \pm \frac{\partial v}{\partial x} h + \frac{\partial^2 v}{\partial x^2} \frac{h^2}{2} \pm \frac{\partial^3 v}{\partial x^3} \frac{h^3}{6} + \frac{\partial^4 v}{\partial x^4} \frac{h^4}{24} + O(h^5), \quad h \rightarrow 0, \\ \frac{v_{i+1,j} - 2v_{i,j} + v_{i-1,j}}{h^2} &= \frac{\partial^2 v}{\partial x^2} + O(h^2), \end{aligned}$$

аналогично

$$\frac{v_{i,j+1} - 2v_{i,j} + v_{i,j-1}}{h^2} = \frac{\partial^2 v}{\partial y^2} + O(h^2)$$

(из формулы Тейлора; все производные берутся в точке (ih, jh)), поэтому первое уравнение системы аппроксимирует уравнение Гельмгольца с невязкой второго порядка.

Аналогично получается, что сеточное приближение для граничного условия имеет первый порядок аппроксимации. Поэтому схема имеет первый порядок аппроксимации.

Поскольку схема устойчива, из теоремы Рябенького—Лакса [37] следует сходимость первого порядка к решению дифференциальной краевой задачи при стремлении шага сетки к нулю.

Приближённое вычисление градиента на сетке

При применении градиентного метода в качестве оракула первого порядка используется алгоритм 4:

Algorithm 4 Gradient Helmholtz Approximate Oracle

Input: $q(y) \in H; f_j = f(jh), g_j = g(jh), r_{i,j} = r(ih, jh), \kappa$ — параметры задачи, $h = \frac{1}{P}$ — шаг сетки

- 1: вычислить $u_{i,j} := \Gamma_h[q_j, g_j, r_{i,j}]_{i,j};$
- 2: вычислить $\psi_{i,j} := \Gamma_h[0, u_{0,j} - f_j, 0]_{i,j};$
- 3: вычислить $\tilde{\nabla} J(q)_j := \frac{\psi_{P,j} - \psi_{P-1,j}}{h}.$

Output: $\tilde{\nabla} J(q)_j.$

Теоретическая оценка погрешности

Поскольку на каждом шаге алгоритма 4 результат отличается от точного на величину, пропорциональную шагу сетки, то и

$$\max_j |\tilde{\nabla} J(q)_j - \nabla J(q)(jh)| = O(h), h \rightarrow 0.$$

Введём обозначения: $\delta\varphi$ — погрешность вычисления функции φ (это может быть u, ψ);

$$\delta\Gamma[a, b, c]_{i,j} := \Gamma_h[a(ih), b(jh), c(ih, jh)]_{i,j} - \Gamma[a(y), b(y), c(x, y)](ih, jh)$$

— погрешность решения краевой задачи Γ при её аппроксимации задачей Γ_h на равномерной сетке. Считаем, что $\varphi_{i,j}$ и $\varphi(ih, jh)$ — это одно и то же. Тогда можно вычислить:

$$\begin{aligned} \delta u_{i,j} &= \delta\Gamma[q, g, r]_{i,j}; \\ \delta\psi_{i,j} &= \delta\Gamma[0, u(0, y) - f(y), 0]_{i,j} + \Gamma_h[0, \delta u_{0,j}, 0]_{i,j} = \\ &= \delta\Gamma[0, \Gamma[q, g, r](0, y) - f(y), 0]_{i,j} + \Gamma_h[0, \delta\Gamma[q, g, r]_{0,j}, 0]_{i,j} = \\ &= (\delta\Gamma[0, \Gamma[0, g, r](0, y) - f(y), 0]_{i,j} + \Gamma_h[0, \delta\Gamma[0, g, r]_{0,j}]_{i,j}, 0) + \\ &\quad + (\delta\Gamma[0, \Gamma[q, 0, 0](0, y), 0]_{i,j} + \Gamma_h[0, \delta\Gamma[q, 0, 0]_{0,j}, 0]_{i,j}). \end{aligned}$$

Погрешность — это сумма двух слагаемых, одно из которых зависит от q , а другое — нет. При этом оно не равно тождественно нулю, и зависит от мелкости сетки и начальных условий.

Соответственно, градиент не удовлетворяет концепции относительного шума, потому что погрешность такого рода не допускает оценки сверху величиной, пропорциональной норме истинного значения градиента с малым коэффициентом пропорциональности.

С другой стороны, погрешность пропорциональна мелкости разбиения, поэтому концепция аддитивного шума корректно описывает данный способ вычисления.

Обозначая количество точек отрезка $[0, 1]$ на сетке за P , получим, что краевая задача сводится к решению линейной системы порядка P^2 . Это требует $O(P^6)$ арифметических операций, поскольку стандартные методы решения систем линейных уравнений имеют сложность кубического порядка. При этом погрешность вычисления градиента составит $O(P^{-1})$.

3.1.5. Эксперименты

В качестве иллюстрации для сравнения методов использована модельная задача (1.5) с параметрами $r(x, y) = -x(2 - y + y^2)$, $f(y) = 0$, $g(y) = y - y^2$. Известно её точное решение: $q(y) = y - y^2$.

Результаты экспериментов (достигнутые наименьшие расстояния до точного решения в нормах L_2 и C , значения функционала) представлены в таблице 1.

Таблица 1 – Результаты работы методов. Задача (1.5).

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	6.14×10^{-4}	1.34×10^{-3}	7.38×10^{-24}
ММО, $m = 2$	6.14×10^{-4}	1.33×10^{-3}	3.32×10^{-25}
ММО, $m = 5$	6.14×10^{-4}	1.22×10^{-3}	1.53×10^{-25}
ММО, $m = \infty$	6.14×10^{-4}	1.23×10^{-3}	2.19×10^{-25}
Сопряжённые градиенты	1.58×10^{-3}	2.96×10^{-3}	1.29×10^{-19}
Тяжёлый шарик	1.58×10^{-3}	2.87×10^{-3}	1.36×10^{-19}
Шаг Поляка	6.94×10^{-3}	1.06×10^{-2}	8.30×10^{-13}
STM	1.56×10^{-3}	2.97×10^{-3}	1.24×10^{-19}

Использованы сетка с шагом $h = 0.01$ и нулевое начальное приближение. Начальное значение функционала $J(q_0) = 1.77 \cdot 10^{-4}$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.183$.

Рисунок 1 показывает, что 1-моментный метод минимальных ошибок позволяет за небольшое число шагов достичь меньшей невязки, чем адаптивный метод тяжёлого шарика из статьи [51], градиентный спуск с шагом Поляка и метод подобных треугольников, и по аргументу, и по функционалу. При этом сходимость по функционалу не является монотонной.

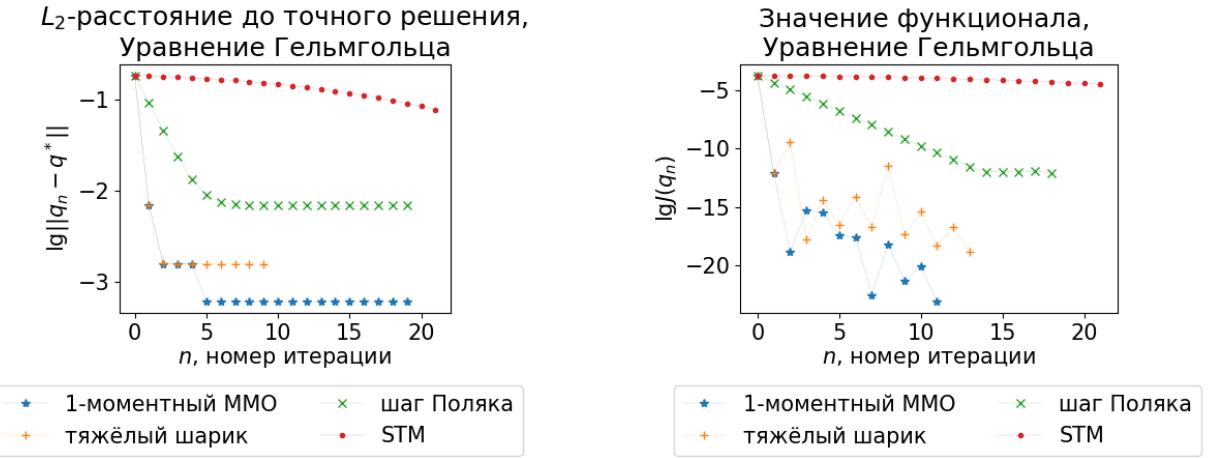


Рисунок 1 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.5).

Увеличение количества итераций не приводит к заметным изменениям результатов, кроме метода подобных треугольников. Для него достигнутые результаты оказались близки к методу сопряжённых градиентов.

Рисунок 2 показывает, что различные варианты метода сопряжённых градиентов дают на рассматриваемой задаче примерно одинаковые результаты, а 1-моментный метод минимальных ошибок их превосходит.

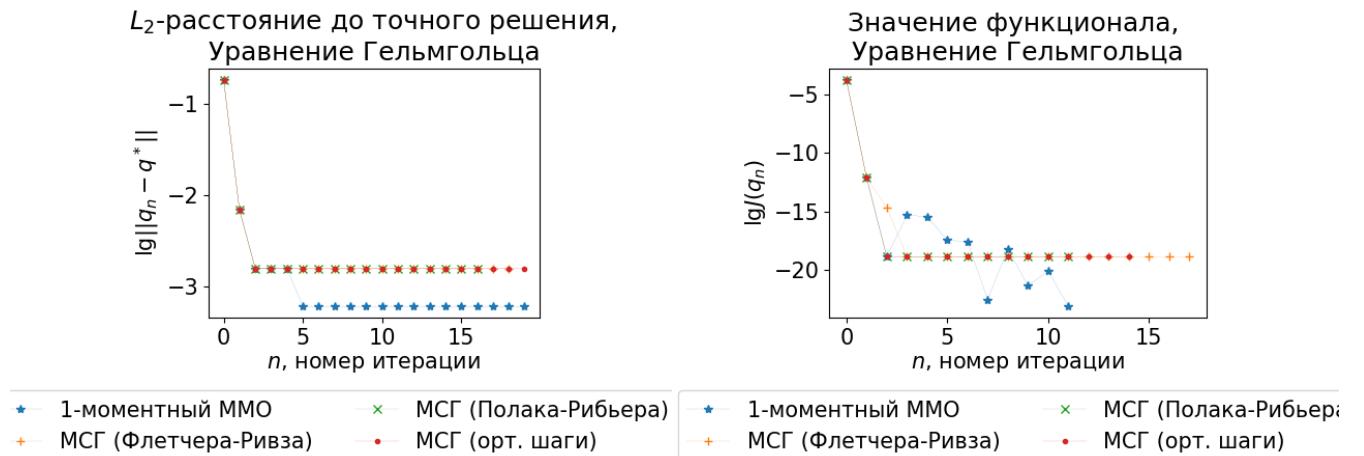


Рисунок 2 – Сходимость: 1-моментный ММО и МСГ. Задача (1.5).

На рисунке 3 сравниваются t -моментные методы минимальных ошибок при разных t . Теоретически, чем больше t , тем лучше должна быть сходимость по аргументу. Однако на данной задаче такой эффект не наблюдается: все методы дают очень близкие результаты.

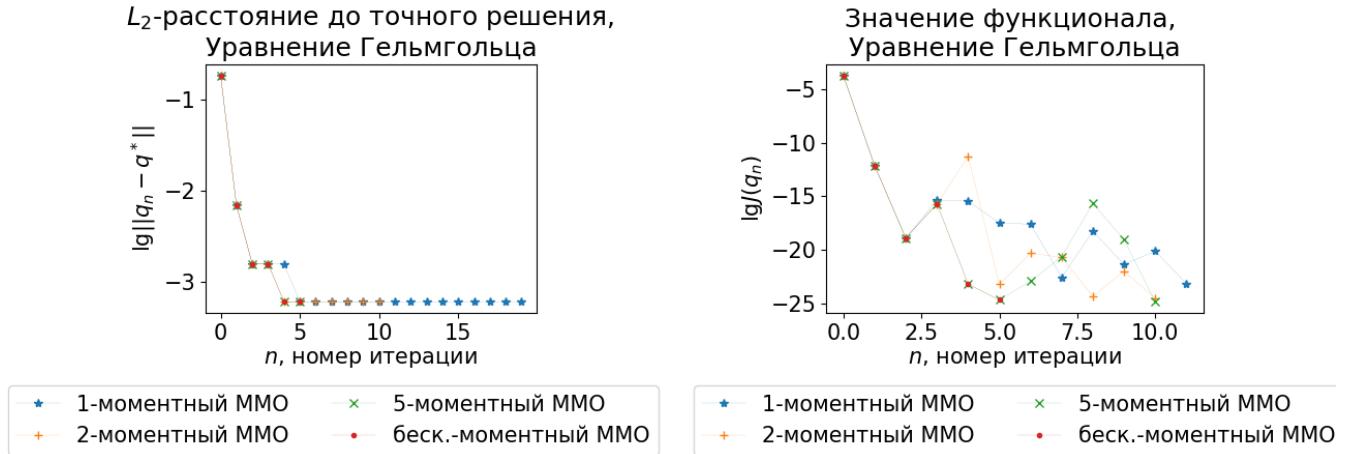


Рисунок 3 – Сходимость: t -моментный ММО с $t \in \{1, 2, 5, \infty\}$. Задача (1.5).

Сравним полученные решения с известной функцией $q^*(y)$. Изобразим наилучшие по функционалу и по расстоянию приближения, полученные с помощью 5-моментного ММО и с помощью метода сопряжённых градиентов, на рисунке 4.

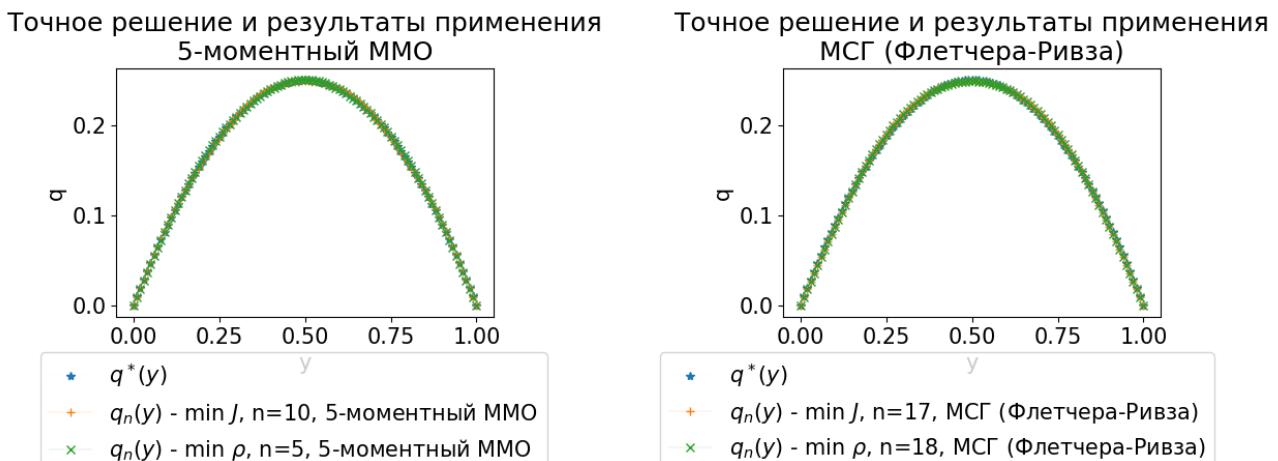


Рисунок 4 – Сравнение с точным решением: 5-моментный ММО, МСГ. Задача (1.5).

Функция восстановлена: точки на графике почти неразличимы.

Проверим теперь, насколько методы оптимизации чувствительны к шуму. Для этого рассмотрим модельную задачу с $r(x, y) = g(y) = 0$, точным решением $q^*(y) = y - y^2$ и условием $f(y) = (Aq^*)(y) \cdot (1 + \delta \cdot \mathcal{U}[-1, 1])$.

Достигнутые расстояния до точного решения для t -моментных ММО при $t \in \{1, 2, 5\}$ и МСГ в форме Флетчера-Ривза с разными δ представлены в таблице 2. Также проверена модификация 1-моментного ММО, в которой $\beta_{2k} \equiv 0$.

Таблица 2 – Расстояние до точного решения при наличии шума δ в правой части. Задача (1.5).

$\delta, \%$	0	0.5	1	2	5
ММО, $m = 1$	6.14×10^{-4}	6.94×10^{-3}	6.94×10^{-3}	6.94×10^{-3}	6.96×10^{-3}
ММО, $m = 1$ (2)	1.58×10^{-3}	2.09×10^{-3}	3.71×10^{-3}	6.94×10^{-3}	6.96×10^{-3}
ММО, $m = 2$	6.14×10^{-4}	6.94×10^{-3}	6.94×10^{-3}	6.94×10^{-3}	6.96×10^{-3}
ММО, $m = 5$	3.33×10^{-4}	6.94×10^{-3}	6.94×10^{-3}	6.94×10^{-3}	6.96×10^{-3}
МСГ	1.58×10^{-3}	2.98×10^{-3}	5.05×10^{-3}	6.94×10^{-3}	6.95×10^{-3}

Видно, что преимущество m -моментного ММО, наблюдаемое при решении точных задач, теряется при добавлении шума. При этом качество решения всё ещё сопоставимо с МСГ.

3.1.6. Выводы

Эксперименты показывают, что градиентные методы оптимизации позволяют эффективно решать начально-краевую задачу для уравнения Гельмгольца. Если значения $f(y)$ заданы точно, то получаемые приближённые решения достаточно близки к известным точным решениям и «в среднем» (по L_2 -норме), и равномерно (по C -норме). Максимальное отклонение полученных решений от точного — на два порядка меньше, чем разброс между минимальным и максимальным значениями $q^*(y)$. Это свидетельствует о хорошем восстановлении формы графика, что можно наблюдать и непосредственно.

Практические результаты согласуются с теоретическими: оператор задачи компактен, поэтому по теореме 2 сходится метод простейшего градиентного спуска; соответственно, рассмотренные методы тоже должны сходиться. В частности, сходимость m -моментного метода минимальных ошибок установлена теоремой 3.

Теорема 1 об оптимальности ∞ -моментного метода минимальных ошибок нашла своё подтверждение: новые методы сходятся значительно лучше ранее существовавших. Ускоренные методы достигли лучших результатов, чем неускоренный (градиентный спуск с шагом Поляка).

Как показано в разделе 3.1.3, собственные значения связанного с задачей самосопряжённого оператора A^*A_0 имеют вид $\lambda_n = \frac{1}{\operatorname{ch}^2 \sqrt{\pi^2 n^2 - \kappa^2}}$. Поскольку при

$t \rightarrow \infty$ $\operatorname{ch} t \sim \frac{e^t}{2}$, а $\sqrt{\pi^2 n^2 - \kappa^2} = \pi n - \frac{\kappa^2}{2\pi n} + o\left(\frac{1}{n}\right)$, $n \rightarrow \infty$, собственные значения стремятся к нулю эквивалентно геометрической прогрессии со знаменателем $e^{-2\pi}$.

Тот факт, что последовательность собственных значений связанного с задачей самосопряжённого оператора стремится к нулю не слишком быстро, может быть причиной быстрого достижения градиентными методами оптимизации точек, близких к точному решению. Это связано с тем, что такая задача хорошо приближается конечномерной, если в рядах Фурье для $f(y)$ и $g(y)$ отбросить слагаемые, начиная с некоторого номера. С одной стороны, остаток ряда получается достаточно маленьким; с другой стороны, константа сильной выпуклости (минимальное собственное значение, соответствующее оставшимся слагаемым) оказывается не слишком близкой к нулю.

При добавлении шума работа m -моментного метода минимальных ошибок ухудшается, что подтверждает его чувствительность к точности измерений и вычислений. Качество становится примерно таким же, как у метода сопряжённых градиентов. До некоторой степени, ситуацию улучшают рестарты, препятствующие накоплению погрешности при увеличении количества шагов.

3.2. Ретроспективная задача Коши для уравнения теплопроводности в одномерном пространстве

Результаты были изложены автором в статьях [33] и [34].

3.2.1. Происхождение, постановка и некорректность задачи

Постановка задач взята из книги [53]. Их физический смысл очень прост. Известно распределение температуры стержня в конечный момент времени, а также температура (или тепловой поток) на концах стержня в течение всего отрезка времени. Вне концов стержня теплообмена с окружающей средой нет. Требуется найти распределение температуры в начальный момент времени.

Физический смысл задачи с граничными условиями на значение функции (1.6): в каждый момент времени измеряется температура на концах стержня.

Физический смысл задачи с граничными условиями на производную по

пространственной координате (1.8): в каждый момент времени измеряется тепловой поток через концы стержня.

Докажем некорректность обеих задач с постоянным κ и нулевыми граничными условиями: $a(t) = b(t) = 0$. Взяв $\varepsilon = 1$ и зафиксировав $\delta > 0$, рассмотрим следующие начальные условия: $f(x) = \frac{\delta}{\pi n} \sin \pi n x$ для задачи с граничными условиями на значение функции и $f(x) = \frac{\delta}{\pi n} \cos \pi n x$ для задачи с граничными условиями на производную по пространственной координате (в обоих случаях n — натуральное число, которое будет выбрано позднее). Решением задачи Коши в каждом случае будет $u(x, t) = e^{\pi^2 n^2 \kappa^2 (1-t)} f(x)$, что легко проверяется по определению.

$\|f\| = \frac{\delta}{\pi n} < \delta$; $\|u\| = \frac{\delta}{\pi n} e^{\pi^2 n^2 \kappa^2} \rightarrow +\infty$ при $n \rightarrow +\infty$, поэтому найдётся n такое, что $\|u\| > \varepsilon$. Это и обозначает отсутствие непрерывности решения по начальным условиям, то есть некорректность поставленной задачи.

3.2.2. Операторы задач

Ретроспективные задачи Коши для уравнения теплопроводности некорректны. Однако корректны задачи с условием в начальный момент времени:

$$\begin{cases} u_t - \kappa^2 u_{xx} = 0, & (x, t) \in \Omega = (0, 1) \times (0, 1) \\ u|_{x=0} = a(t), & t \in [0, 1] \\ u|_{x=1} = b(t), & t \in [0, 1] \\ u|_{t=0} = q(x), & x \in [0, 1] \end{cases} \quad (3.2.2.1)$$

или

$$\begin{cases} u_t - \kappa^2 u_{xx} = 0, & (x, t) \in \Omega = (0, 1) \times (0, 1) \\ u_x|_{x=0} = a(t), & t \in [0, 1] \\ u_x|_{x=1} = b(t), & t \in [0, 1] \\ u|_{t=0} = q(x), & x \in [0, 1] \end{cases} \quad (3.2.2.2)$$

Эти задачи отличаются от исходных лишь заменой условия $u|_{t=1} = f(x)$ на $u|_{t=0} = q(x)$. Граница $t = 1$ — наблюдаемая, $t = 0$ — ненаблюдаемая.

Рассмотрим задачу (3.2.2.1) с однородными граничными условиями $a(t) = b(t) = 0$ и найдём её решение в виде $u(x, t) = \sum_{m=1}^{\infty} X_m(t) \sin \pi m x$. Пусть

$q(x) = \sum_{m=1}^{\infty} q_m \sin \pi mx$. Функции $X_m(x)$ являются решениями задач Коши:

$$\begin{cases} X'_m(t) + \kappa^2 \pi^2 m^2 X_m(t) = 0, & t \in (0, 1) \\ X_m(0) = q_m \end{cases}.$$

Поэтому $u(x, t) = \sum_{m=1}^{\infty} q_m e^{-\kappa^2 \pi^2 m^2 t} \sin \pi mx$.

В частности, $u(x, 1) = \sum_{m=1}^{\infty} q_m e^{-\kappa^2 \pi^2 m^2} \sin \pi mx$. Коэффициенты этого ряда стремятся к нулю быстрее, чем q_m , поэтому условия гладкости, которым удовлетворяет $q(x)$, заведомо окажутся выполнены для $u(x, 1)$.

Более того, поскольку для любого натурального l $m^l e^{-\kappa^2 \pi^2 m^2} = o(m^{-2})$, всегда $u(x, 1) \in C^\infty[0, 1]$.

Выкладки для задачи (3.2.2.2) аналогичны, только синусы заменяются на косинусы.

Для каждой из рассматриваемых задач на пространстве непрерывных функций со стандартным скалярным произведением

$$H = C[0, 1] \subset L^2[0, 1]$$

определим оператор $A : H \rightarrow H$ следующим образом:

$$(Aq)(x) = u(x, 1) \quad \forall x \in (0, 1), \quad (3.2.2.3)$$

где u — решение соответствующей корректной задачи Коши (3.2.2.1) или (3.2.2.2).

Как показано выше, $f \in C^\infty[0, 1]$ — необходимое условие существования решения уравнения $Aq = f$.

Важно заметить, что областью определения является не всё пространство. Ограничения возникают из условий непрерывности для существования классического решения задачи Коши.

Для оператора задачи с граничными условиями на значение функции: $q(0) = a(0), q(1) = b(0)$. При этом $(Aq)(0) = a(1), (Aq)(1) = b(1)$.

Для оператора задачи с граничными условиями на производную по пространственной координате: $q'(0) = a(0), q'(1) = b(0)$. При этом $(Aq)'(0) = a(1), (Aq)'(1) = b(1)$.

Полученные ограничения задают линейные многообразия — множества простой структуры. Как будет видно из формул (3.2.2.6) и (3.2.2.7), градиент функционала в каждой точке такого многообразия принадлежит параллельному ему

многообразию, заданному нулевыми граничными условиями. Поэтому проекция градиента совпадает с самим градиентом, и можно применять методы оптимизации, как для задач без ограничений.

Следуя методу из пособия [4], вычислим сопряжённые операторы и, соответственно, градиенты функционалов задач. Сначала проведём вычисления для задачи с граничными условиями на значение функции.

Итак,

$$J(q) = \frac{1}{2} \int_0^1 |u(x, 1) - f(x)|^2 dx,$$

где u — решение задачи (3.2.2.1).

Запишем функцию Лагранжа, взяв множители для первой и последней строк, а две другие считая задающими «простое» множество:

$$\begin{aligned} L(q, u; \psi = (\psi(x, t), \lambda(x))) &= \frac{1}{2} \int_0^1 |u(x, 1) - f(x)|^2 dx - \\ &- \int_0^1 \int_0^1 (u_t - \kappa^2 u_{xx}) \psi(x, t) dx dt + \int_0^1 (q(x) - u(x, 0)) \lambda(x) dx. \end{aligned}$$

Для получения сопряжённой задачи на ψ варьируем по $\delta u(x, t)$:

$$\begin{aligned} \delta_u L(q, u; \psi) &= \int_0^1 (u(x, 1) - f(x)) \delta u(x, 1) dx - \\ &- \int_0^1 \int_0^1 (\delta u_t - \kappa^2 \delta u_{xx}) \psi(x, t) dx dt - \int_0^1 \delta u(x, 0) \lambda(x) dx. \end{aligned}$$

При этом

$$\delta u(0, t) = \delta u(1, t) = 0.$$

Интегрируем по частям слагаемые с производными u :

$$\begin{aligned} \int_0^1 \int_0^1 \psi(x, t) \delta u_t dx dt &= \int_0^1 \psi(x, 1) \delta u(x, 1) dx - \\ &- \int_0^1 \psi(x, 0) \delta u(x, 0) dx - \int_0^1 \int_0^1 \psi_t \delta u(x, t) dx dt; \end{aligned}$$

$$\int_0^1 \int_0^1 \kappa^2 \psi(x, t) \delta u_{xx} dx dt = \int_0^1 \kappa^2 \psi(1, t) \delta u_x(1, t) dt - \\ - \int_0^1 \kappa^2 \psi(0, t) \delta u_x(0, t) dt + \int_0^1 \int_0^1 (\kappa^2 \psi)_{xx}(x, t) \delta u(x, t) dx dt.$$

Подставляем и вычисляем вариацию функции Лагранжа:

$$\delta_u L(q, u; \psi) = \int_0^1 (u(x, 1) - f(x) - \psi(x, 1)) \delta u(x, 1) dx - \\ - \int_0^1 (\lambda(x) - \psi(x, 0)) \delta u(x, 0) dx + \int_0^1 \kappa^2 \psi(1, t) \delta u_x(1, t) dt - \\ - \int_0^1 \kappa^2 \psi(0, t) \delta u_x(0, t) dt + \int_0^1 \int_0^1 (\psi_t(x, t) + (\kappa^2 \psi)_{xx}(x, t)) \delta u(x, t) dx dt.$$

Из необходимого условия экстремума получаем систему уравнений:

$$\begin{cases} \psi_t + (\kappa^2 \psi)_{xx} = 0, & (x, t) \in (0, 1) \times (0, 1) \\ \psi|_{x=0} = \psi|_{x=1} = 0, & t \in [0, 1] \\ \psi|_{t=1} = u(x, 1) - f(x), & x \in [0, 1] \\ \psi|_{t=0} = \lambda(x), & x \in [0, 1] \end{cases}$$

Последнее уравнение — связь между множителями Лагранжа. Поскольку

$$\lambda(x) = L_q(x) = \nabla J(q)(x)$$

(формула Демьянова—Данскина—Рубинова [10], [48]), вычисление градиента функционала свелось к решению двух корректных задач.

Оператор $A^* : H \rightarrow H$ задаётся следующим образом:

$$(A^* p)(x) = \psi(x, 0) \quad \forall x \in (0, 1), \tag{3.2.2.4}$$

где ψ — решение задачи

$$\begin{cases} \psi_t + (\kappa^2 \psi)_{xx} = 0, & (x, t) \in (0, 1) \times (0, 1) \\ \psi|_{x=0} = \psi|_{x=1} = 0, & t \in [0, 1] \\ \psi|_{t=1} = p(x), & x \in [0, 1] \end{cases}. \tag{3.2.2.5}$$

Замена переменных $\psi(x, t) = \phi(x, 1 - t)$ позволяет привести эту задачу к виду, похожему на исходную. При этом в случае κ , зависящего от времени, выполняется равенство $\bar{\kappa}(x, t) = \kappa(x, 1 - t)$.

$$(A^* p)(x) = \phi(x, 1) \quad \forall x \in (0, 1), \quad (3.2.2.6)$$

где ϕ — решение задачи

$$\begin{cases} \phi_t - (\bar{\kappa}^2 \phi)_{xx} = 0, & (x, t) \in (0, 1) \times (0, 1) \\ \phi|_{x=0} = \phi|_{x=1} = 0, & t \in [0, 1] \\ \phi|_{t=0} = p(x), & x \in [0, 1] \end{cases}. \quad (3.2.2.7)$$

Как видно из определения, оператор A (в отличие от сопряжённого к нему A^*) не является линейным из-за неоднородных краевых условий в определяющей его задаче. Однако линейным является следующий оператор $A_0 : H \rightarrow H$:

$$(A_0 q)(x) = u(x, 1) \quad \forall x \in (0, 1), \quad (3.2.2.8)$$

где u — решение краевой задачи, в которой все краевые условия, кроме q , равны нулю:

$$\begin{cases} u_t - \kappa^2 u_{xx} = 0, & (x, t) \in \Omega = (0, 1) \times (0, 1) \\ u|_{x=0} = u|_{x=1} = 0, & t \in [0, 1] \\ u|_{t=0} = q(x), & x \in [0, 1] \end{cases}. \quad (3.2.2.9)$$

Свойство $Aq_1 - Aq_2 = A_0(q_1 - q_2)$ выполняется, как и для оператора, определяемого уравнением Гельмгольца.

Важно заметить, что $A^* = A_0$ тогда и только тогда, когда κ не зависит от координаты. В противном случае задача, определяющая оператор A^* , отличается от исходной, хотя остаётся параболической и корректной. При этом необходимо $\kappa^2 \in C^2(0, 1)$.

Для задачи (1.8) с граничными условиями на производную по пространственной координате вычисления практически полностью аналогичны. Отличие лишь в одном слагаемом:

$$\begin{aligned} \int_0^1 \int_0^1 \kappa^2 \psi(x, t) \delta u_{xx} dx dt &= - \int_0^1 (\kappa^2 \psi)_x(1, t) \delta u(1, t) dt + \\ &+ \int_0^1 (\kappa^2 \psi)_x(0, t) \delta u(0, t) dt + \int_0^1 \int_0^1 (\kappa^2 \psi)_{xx}(x, t) \delta u(x, t) dx dt. \end{aligned}$$

Соответственно, в данном случае оператор $A^* : H \rightarrow H$ задаётся так:

$$(A^*p)(x) = \phi(x, 1) \quad \forall x \in (0, 1), \quad (3.2.2.10)$$

где ϕ — решение задачи

$$\begin{cases} \phi_t - (\bar{\kappa}^2 \phi)_{xx} = 0, & (x, t) \in (0, 1) \times (0, 1) \\ (\bar{\kappa}^2 \phi)_x|_{x=0} = (\bar{\kappa}^2 \phi)_x|_{x=1} = 0, & t \in [0, 1] \\ \phi|_{t=0} = p(x), & x \in [0, 1] \end{cases}. \quad (3.2.2.11)$$

Остальные выкладки совпадают.

3.2.3. Спектр соответствующего задаче самосопряжённого оператора

Оператор $A^* A_0$ является самосопряжённым. Константа Липшица градиента равна его максимальному собственному значению.

Аналитическое вычисление A_0 представляется возможным только при постоянном κ . В этом случае $A^* = A_0$, и $\lambda_m(A^* A_0) = \lambda_m^2(A_0)$.

В области определения оператора задачи с однородными граничными условиями на значение функции существует базис — это система функций $\{w_m\}_{m=1}^\infty = \{\sin \pi m x\}_{m=1}^\infty$. Непосредственная проверка показывает, что $A_0 w_m = e^{-\pi^2 \kappa^2 m^2} w_m$.

Соответственно, $\lambda_m(A^* A_0) = e^{-2\pi^2 \kappa^2 m^2}$, и константа Липшица градиента $L = e^{-2\pi^2 \kappa^2}$.

В области определения оператора задачи с однородными граничными условиями на производную по пространственной координате также существует базис — это система функций $\{w_m\}_{m=0}^\infty = \{\cos \pi m x\}_{m=0}^\infty$. Непосредственная проверка показывает, что $A_0 w_m = e^{-\pi^2 \kappa^2 m^2} w_m$.

Соответственно, $\lambda_m(A^* A_0) = e^{-2\pi^2 \kappa^2 m^2}$, и константа Липшица градиента $L = 1$.

Эти оценки объединяет следующая лемма.

Лемма 3. При постоянном $\kappa > 0$ задача минимизации функционала $J(q)$, соответствующего ретроспективной задаче Коши для уравнения теплопроводности, является выпуклой с липшицевым градиентом, спектр $\lambda_m = e^{-2\pi^2 \kappa^2 m^2}$. Если задача имеет граничные условия на значение функции, то $m \geq 1$, константа Липшица $L = e^{-2\pi^2 \kappa^2} \leq 1$, константа сильной выпуклости

$\mu = 0$. Если задача имеет граничные условия на производную по пространственной координате, то $m \geq 0$, константа Липшица градиента $L = 1$, константа сильной выпуклости $\mu = 0$.

3.2.4. Приближённые вычисления

Корректные задачи Коши для уравнения теплопроводности удобно решать с помощью явных разностных схем. Рассмотрим их.

Разностные схемы для задач Коши

На $[0, 1] \times [0, 1]$ задаётся сетка с шагами h по пространству и τ по времени, то есть

$$\{x = ih, t = j\tau \mid 0 \leq i \leq P, 0 \leq j \leq T\} \quad (Ph = T\tau = 1).$$

Прямая задача с граничными условиями на значение функции аппроксимируется схемой первого порядка по τ и второго порядка по h :

$$\begin{cases} \frac{u_{i,j+1} - u_{i,j}}{\tau} - \kappa_{i,j}^2 \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} = 0, & 1 \leq i \leq P-1, 0 \leq j \leq T-1 \\ u_{0,j} = a_j, & 1 \leq j \leq T \\ u_{M,j} = b_j, & 1 \leq j \leq T \\ u_{i,0} = q_i, & 0 \leq i \leq P \end{cases} \quad (3.2.4.1)$$

Разностная схема для задачи с граничными условиями на производную по пространственной координате выглядит очень похоже:

$$\begin{cases} \frac{u_{i,j+1} - u_{i,j}}{\tau} - \kappa_{i,j}^2 \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} = 0, & 1 \leq i \leq P-1, 0 \leq j \leq T-1 \\ \frac{-3u_{0,j} + 4u_{1,j} - u_{2,j}}{2h} = a_j, & 1 \leq j \leq T \\ \frac{-3u_{M,j} + 4u_{M-1,j} - u_{M-2,j}}{2h} = b_j, & 1 \leq j \leq T \\ u_{i,0} = q_i, & 0 \leq i \leq P \end{cases} \quad (3.2.4.2)$$

Здесь использована аппроксимация второго порядка для частной производной. При этом

$$\kappa_{i,j} = \kappa(ih, j\tau), \quad a_j = a(j\tau), \quad b_j = b(j\tau), \quad q_i = q(ih), \quad u_{i,j} = u(ih, j\tau).$$

Условие сходимости (спектральный признак Неймана) для явных схем даёт условие устойчивости $\tau \leq \frac{h^2}{2\kappa_{max}^2}$.

Сопряжённые задачи аппроксимируются другими схемами:

$$\begin{cases} \frac{\phi_{i,j+1} - \phi_{i,j}}{\tau} - \frac{\bar{\kappa}_{i+1,j}^2 \phi_{i+1,j} - 2\bar{\kappa}_{i,j}^2 \phi_{i,j} + \bar{\kappa}_{i-1,j}^2 \phi_{i-1,j}}{h^2} = 0, & 1 \leq i \leq P-1, 1 \leq j \leq T-1 \\ \phi_{0,j} = 0, & 1 \leq j \leq T-1 \\ \phi_{M,j} = 0, & 1 \leq j \leq T-1 \\ \phi_{i,0} = p_i, & 0 \leq i \leq P \end{cases} \quad (3.2.4.3)$$

$$\begin{cases} \frac{\phi_{i,j+1} - \phi_{i,j}}{\tau} - \frac{\bar{\kappa}_{i+1,j}^2 \phi_{i+1,j} - 2\bar{\kappa}_{i,j}^2 \phi_{i,j} + \bar{\kappa}_{i-1,j}^2 \phi_{i-1,j}}{h^2} = 0, & 1 \leq i \leq P-1, 1 \leq j \leq T-1 \\ \frac{-3\phi_{0,j} + 4\phi_{1,j} - \phi_{2,j}}{2h} = 0, & 1 \leq j \leq T-1 \\ \frac{-3\phi_{M,j} + 4\phi_{M-1,j} - \phi_{M-2,j}}{2h} = 0, & 1 \leq j \leq T-1 \\ \phi_{i,0} = p_i, & 0 \leq i \leq P \end{cases} \quad (3.2.4.4)$$

При постоянном κ эти схемы совпадают с точностью до обозначений со схемами для прямых задач.

Поскольку $\tau = O(h^2)$, схемы имеют второй порядок аппроксимации по h . Из теоремы Рябенького—Лакса [37] следует сходимость второго порядка к решению дифференциальной задачи при согласованном стремлении шагов сетки к нулю.

Приближённое вычисление градиента на сетке

При применении градиентного метода в качестве оракула первого порядка используется приближённый алгоритм 5.

Algorithm 5 Gradient Heat Approximate Oracle

Input: $q(x) \in H$; $f_i = f(ih)$, $a_j = a(j\tau)$, $b_j = b(j\tau)$, $\kappa_{i,j} = \kappa(ih, j\tau)$, κ — параметры задачи, $h = \frac{1}{P}$, $\tau = \frac{1}{T}$ — шаги сетки

- 1: вычислить $u_{i,j}$ как решение задачи (3.2.4.1) для задачи с граничными условиями на значение функции или (3.2.4.2) для задачи с граничными условиями на производную по пространственной координате;
- 2: вычислить $\phi_{i,j}$ как решение задачи (3.2.4.3) для задачи с граничными условиями на значение функции или (3.2.4.4) для задачи с граничными условиями на производную по пространственной координате с начальным условием $p_i = u_{i,T} - f_i$.

Output: $\tilde{\nabla} J(q)_i = \phi_{i,T}$.

Теоретическая оценка погрешности

Поскольку на каждом шаге алгоритма 5 результат отличается от точного на величину, пропорциональную квадрату пространственного шага сетки, то и

$$\max_j |\tilde{\nabla} J(q)_j - \nabla J(q)(jh)| = O(h^2), h \rightarrow 0.$$

Погрешность пропорциональна квадрату мелкости разбиения, поэтому концепция аддитивного шума корректно описывает данный способ вычисления.

Обозначая количество точек отрезка $[0, 1]$ на сетке за P , получим, что решение краевой задачи требует $O(PT) = O(P^3)$ арифметических операций. При этом погрешность вычисления градиента составит $O(P^{-2})$.

3.2.5. Эксперименты

Задача с граничными условиями на значение функции

В качестве иллюстрации для сравнения методов использована модельная задача (1.6) с параметрами $a(t) = b(t) = 0$, $\kappa_{max} \in \{0.4, 0.6\}$,

$$\kappa(x) = \kappa_{max} \cdot \left(0.2 + 0.8 \cdot [|x - 0.5| < 0.2] \cdot \left(1 - \frac{|x - 0.5|}{0.2} \right)^2 \right),$$

$$q(x) = x - x^2, f(x) = (Aq)(x).$$

Квадратные скобки в формуле, как обычно, обозначают функцию-индикатор, которая принимает значение 1, если утверждение внутри скобок истинно, и 0 в противоположном случае.

Выбор переменного коэффициента теплопроводности обусловлен факторами:

- возможные технические приложения обычно связаны с неоднородными средами. Здесь моделируется стержень с неоднородностью в середине;
- неоднородность сглажена для соблюдения условия дифференцируемости;
- из физических соображений, чем больше коэффициент теплопроводности, тем быстрее выравнивается температура. Поэтому предполагается, что увеличение κ_{max} должно приводить к ухудшению качества решения задачи.

Результаты экспериментов (достигнутые наименьшие расстояния до точного решения в нормах L_2 и C , значения функционала) представлены в таблице 3.

Таблица 3 – Результаты работы методов. Задача (1.6), $\kappa_{max} = 0.4$.

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	1.48×10^{-3}	4.49×10^{-3}	2.13×10^{-15}
ММО, $m \geq 2$	7.45×10^{-3}	1.75×10^{-2}	2.61×10^{-7}
Сопряжённые градиенты	3.31×10^{-3}	9.16×10^{-3}	2.56×10^{-11}
Тяжёлый шарик	3.28×10^{-3}	5.94×10^{-3}	4.95×10^{-11}
Шаг Поляка	2.72×10^{-3}	7.64×10^{-3}	9.89×10^{-14}
STM	2.64×10^{-3}	7.38×10^{-3}	9.90×10^{-14}

Использованы сетка с шагом $h = 0.01$ и нулевое начальное приближение. Начальное значение функционала $J(q_0) = 0.013$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.183$.

Рисунок 5 показывает, что 1-моментный метод минимальных ошибок позволяет достичь меньшей невязки, чем адаптивный метод тяжёлого шарика из статьи [51], градиентный спуск с шагом Поляка и метод подобных треугольников, и по аргументу, и по функционалу. При этом сходимость по функционалу не является монотонной.

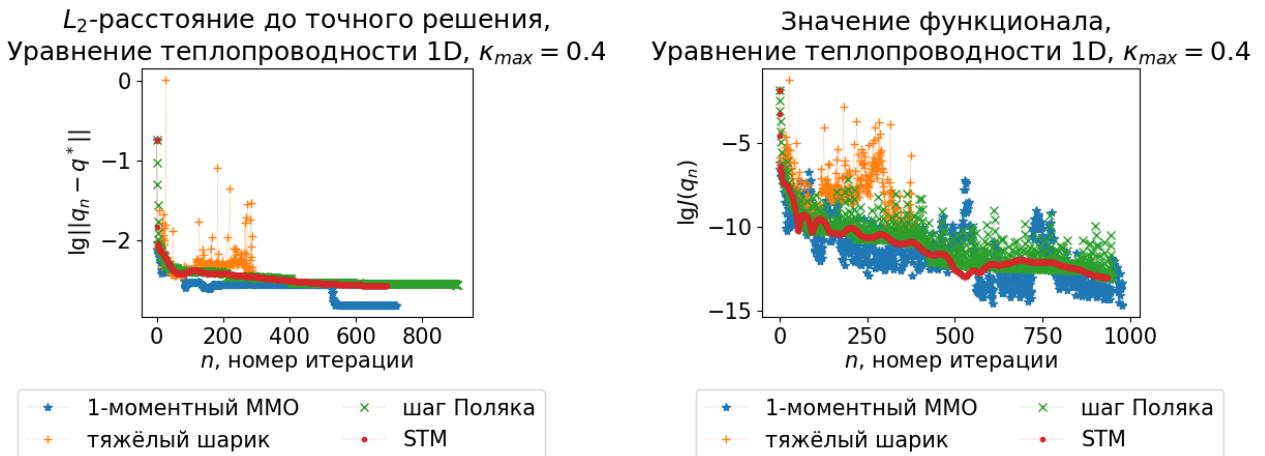


Рисунок 5 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.6), $\kappa_{max} = 0.4$.

Графики построены только до точки минимума. Можно заметить, что с некоторого номера итерации расстояние до точного решения перестаёт уменьшаться. Это связано, в частности, с неточностью вычислений.

Рисунок 6 показывает, что классические варианты метода сопряжённых градиентов сначала превосходят метод с ортогональными шагами и вспомогательной минимизацией функционала, а потом перестают работать. При

в этом вспомогательная минимизация расстояния до точного решения позволяет улучшить качество решения по обоим параметрам.

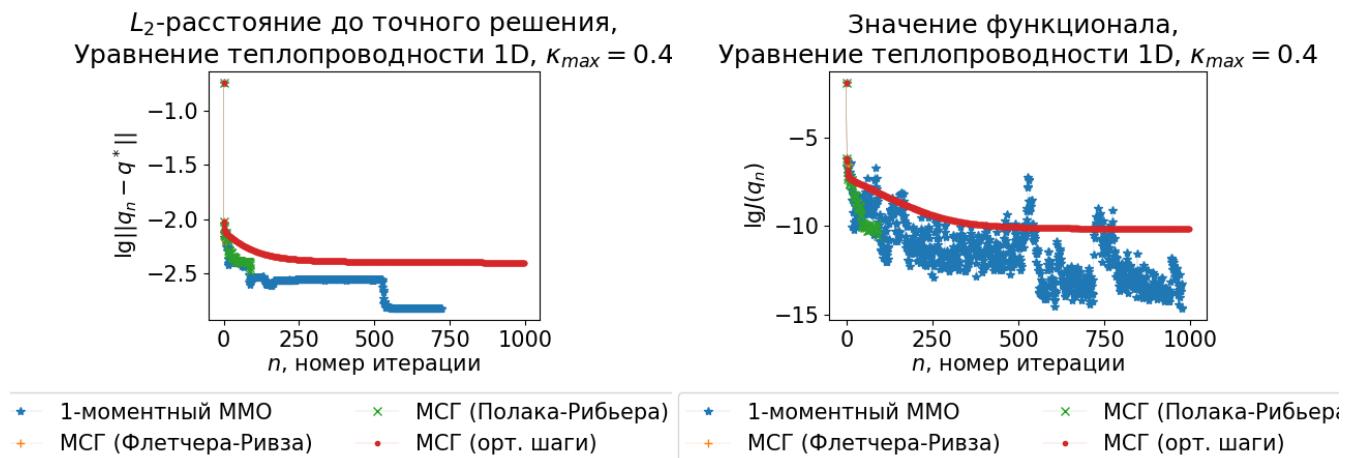


Рисунок 6 – Сходимость: 1-моментный ММО и МСГ. Задача (1.6), $\kappa_{max} = 0.4$.

Увеличение параметра t в многошаговом методе минимальных ошибок приводит к тому, что после второй итерации расстояние до точного решения перестаёт уменьшаться. Это связано с тем, что длина проекции градиента на ортогональное дополнение линейной оболочки t предыдущих шагов оказывается в знаменателе, и её малость приводит к потере точности вычислений.

Сравним полученные решения с известной функцией $q^*(x)$. Изобразим наилучшие по функционалу и по расстоянию приближения, полученные с помощью 1-моментного ММО и с помощью метода сопряжённых градиентов, на рисунке 7.

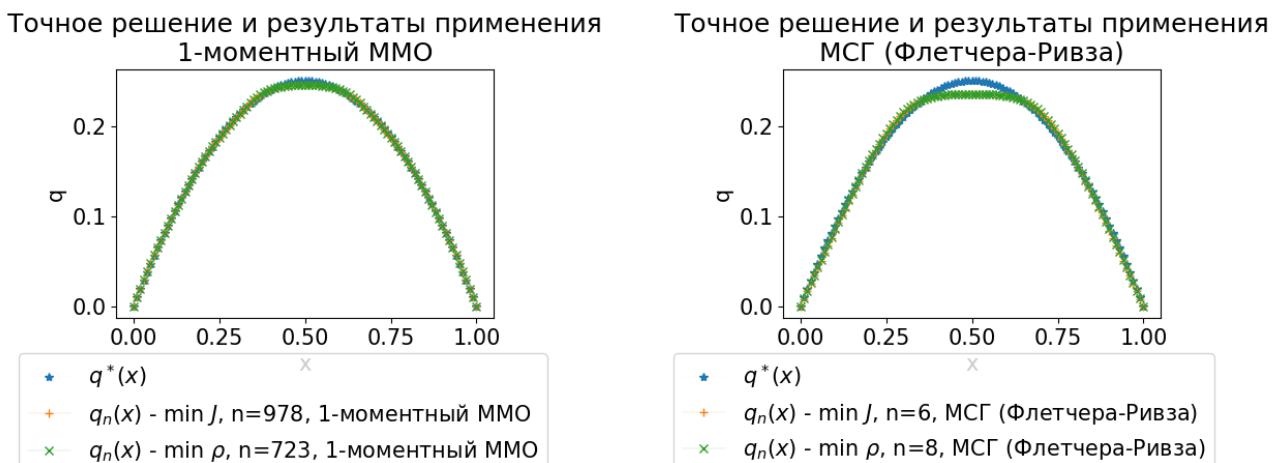


Рисунок 7 – Сравнение с точным решением: 1-моментный ММО, МСГ. Задача (1.6), $\kappa_{max} = 0.4$.

1-моментный метод минимальных ошибок справился с восстановлением формы

графика лучше: в обоих случаях есть погрешность в окрестности экстремума, но у 1-моментного ММО она меньше.

Результаты экспериментов для аналогичной задачи с $\kappa_{max} = 0.6$ представлены в таблице 4.

Таблица 4 – Результаты работы методов. Задача (1.6), $\kappa_{max} = 0.6$.

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	3.33×10^{-3}	8.44×10^{-3}	3.62×10^{-16}
ММО, $m \geq 2$	8.29×10^{-3}	1.94×10^{-2}	6.13×10^{-8}
Сопряжённые градиенты (ФР)	7.49×10^{-3}	1.71×10^{-2}	6.12×10^{-8}
Сопряжённые градиенты (ПР)	4.15×10^{-3}	1.05×10^{-2}	4.82×10^{-11}
Сопряжённые градиенты (\perp)	7.76×10^{-3}	1.83×10^{-2}	7.42×10^{-11}
Тяжёлый шарик	6.35×10^{-3}	1.10×10^{-2}	5.62×10^{-11}
Шаг Поляка	4.30×10^{-3}	1.15×10^{-2}	4.43×10^{-13}
STM	5.17×10^{-3}	1.34×10^{-2}	8.54×10^{-12}

Начальное значение функционала $J(q_0) = 9.48 \times 10^{-3}$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.183$.

На рисунках 8 и 9 видны те же эффекты, которые наблюдались и при решении задачи с $\kappa_{max} = 0.4$, только ещё более явно. 1-моментный метод минимальных ошибок превосходит остальные методы и по расстоянию до точного решения, и по функционалу.

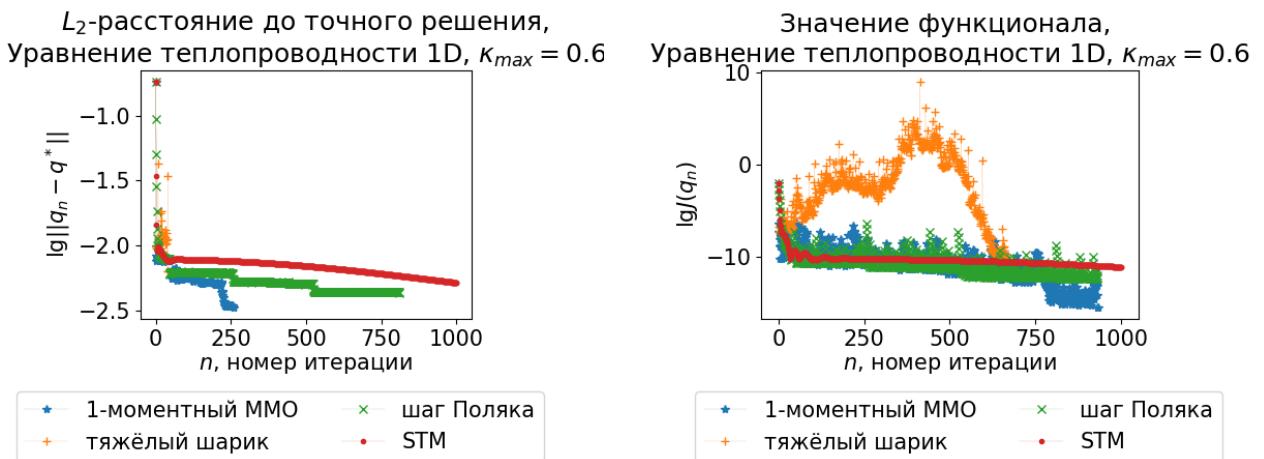


Рисунок 8 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.6), $\kappa_{max} = 0.6$.

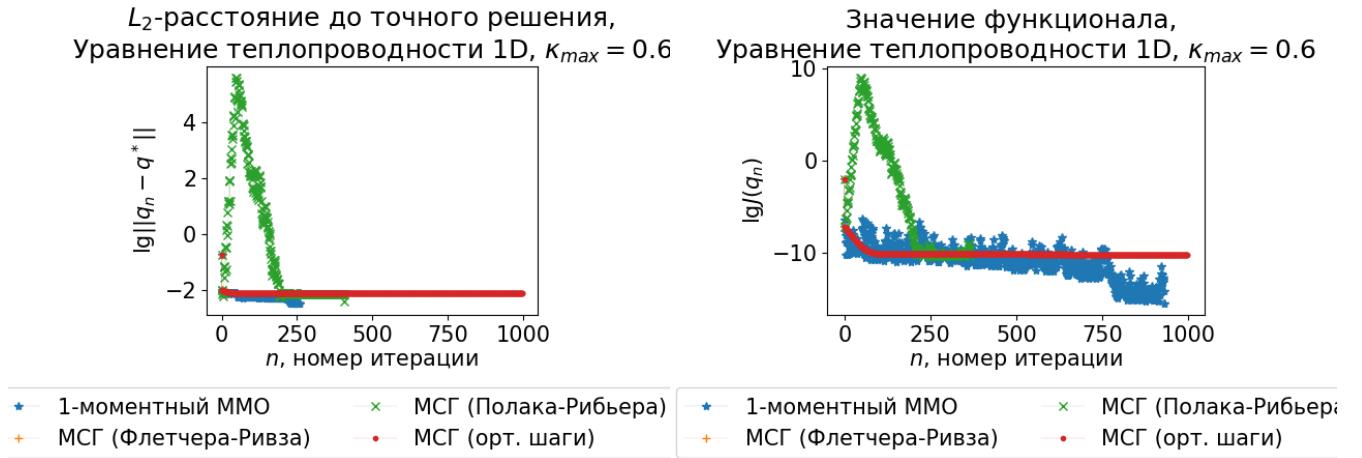


Рисунок 9 – Сходимость: 1-моментный ММО и МСГ. Задача (1.6), $\kappa_{max} = 0.6$.

Сравним полученные решения с известной функцией $q^*(x)$. Изобразим наилучшие по функционалу и по расстоянию приближения, полученные с помощью 1-моментного ММО и с помощью метода сопряжённых градиентов, на рисунке 10.

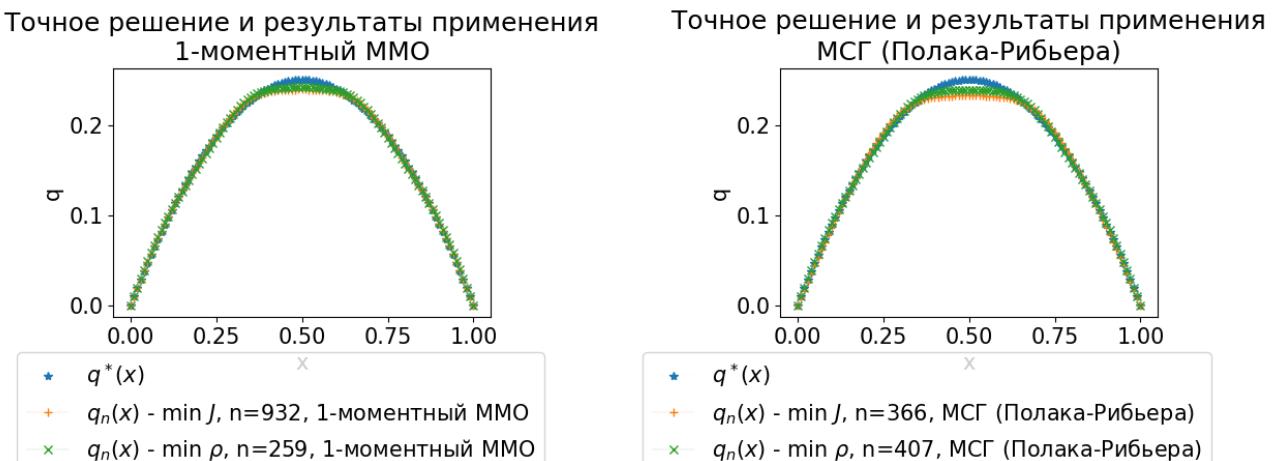


Рисунок 10 – Сравнение с точным решением: 1-моментный ММО, МСГ. Задача (1.6), $\kappa_{max} = 0.6$.

Результаты очень похожи на те, которые были получены при $\kappa_{max} = 0.4$.

Проверим теперь, насколько методы оптимизации чувствительны к шуму. Для этого добавим к рассматриваемой задаче (при $\kappa_{max} = 0.6$) шум в правой части: теперь $f(x) = (Aq^*)(x) \cdot (1 + \delta \cdot \mathcal{U}[-1, 1])$.

Достигнутые расстояния до точного решения для m -моментных ММО при $m \in \{1, 2, 5\}$ и МСГ в форме Флетчера-Ривза с разными δ представлены в таблице 5. Также проверена модификация 1-моментного ММО, в которой $\beta_{2k} \equiv 0$.

Таблица 5 – Расстояние до точного решения при наличии шума δ в правой части. Задача (1.6), $\kappa_{max} = 0.4$.

$\delta, \%$	0.5	1	2	5
ММО, $m = 1$	6.78×10^{-3}	6.50×10^{-3}	9.86×10^{-3}	9.77×10^{-3}
ММО, $m = 1$ (2)	6.63×10^{-3}	6.52×10^{-3}	6.36×10^{-3}	8.23×10^{-3}
ММО, $m = 2$	6.78×10^{-3}	8.29×10^{-3}	9.86×10^{-3}	9.77×10^{-3}
ММО, $m = 5$	8.34×10^{-3}	9.27×10^{-3}	9.86×10^{-3}	9.77×10^{-3}
МСГ	6.86×10^{-3}	7.10×10^{-3}	7.61×10^{-3}	7.53×10^{-3}

Видно, что преимущество m -моментного ММО, наблюдаемое при решении точных задач, теряется при добавлении шума. Чем больше m , тем чувствительнее к шуму m -моментный метод минимальных ошибок. Однако расстояние до точного решения всё ещё не слишком велико.

Задача с граничными условиями на производную по пространственной координате

В качестве иллюстрации для сравнения методов использована модельная задача (1.8) с параметрами $a(t) = b(t) = 0$, $\kappa_{max} \in \{0.4, 0.6\}$,

$$\kappa(x) = \kappa_{max} \cdot \left(0.2 + 0.8 \cdot [|x - 0.5| < 0.2] \cdot \left(1 - \frac{|x - 0.5|}{0.2} \right)^2 \right),$$

$$q(x) = 2x^2 - x^4, f(x) = (Aq)(x).$$

Переменный коэффициент теплопроводности выбран по тем же причинам, что и при исследовании задачи с граничными условиями на значение функции.

Результаты экспериментов представлены в таблице 6.

Таблица 6 – Результаты работы методов. Задача (1.8), $\kappa_{max} = 0.4$.

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	1.27×10^{-3}	3.74×10^{-3}	4.72×10^{-14}
ММО, $m = 2$	1.42×10^{-3}	3.01×10^{-3}	1.05×10^{-10}
ММО, $m = 5$	4.45×10^{-3}	1.09×10^{-2}	3.56×10^{-7}
ММО, $m = \infty$	4.45×10^{-3}	1.09×10^{-2}	3.56×10^{-7}
Сопряжённые градиенты (ФР)	2.02×10^{-3}	3.77×10^{-3}	6.96×10^{-10}
Сопряжённые градиенты (ПР)	2.00×10^{-3}	4.76×10^{-3}	3.86×10^{-10}
Сопряжённые градиенты (\perp)	2.23×10^{-3}	6.24×10^{-3}	1.16×10^{-10}
Тяжёлый шарик	2.09×10^{-3}	4.52×10^{-3}	3.41×10^{-10}
Шаг Поляка	1.73×10^{-3}	4.95×10^{-3}	7.07×10^{-13}
STM	1.57×10^{-3}	4.56×10^{-3}	3.04×10^{-13}

Использованы сетка с шагом $h = 0.01$ и нулевое начальное приближение. Начальное значение функционала $J(q_0) = 0.164$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.583$.

На рисунках 11 и 12 видны те же эффекты, которые наблюдались и при исследовании задачи (1.6). 1-моментный метод минимальных ошибок по-прежнему превосходит остальные методы и по расстоянию до точного решения, и по функционалу.

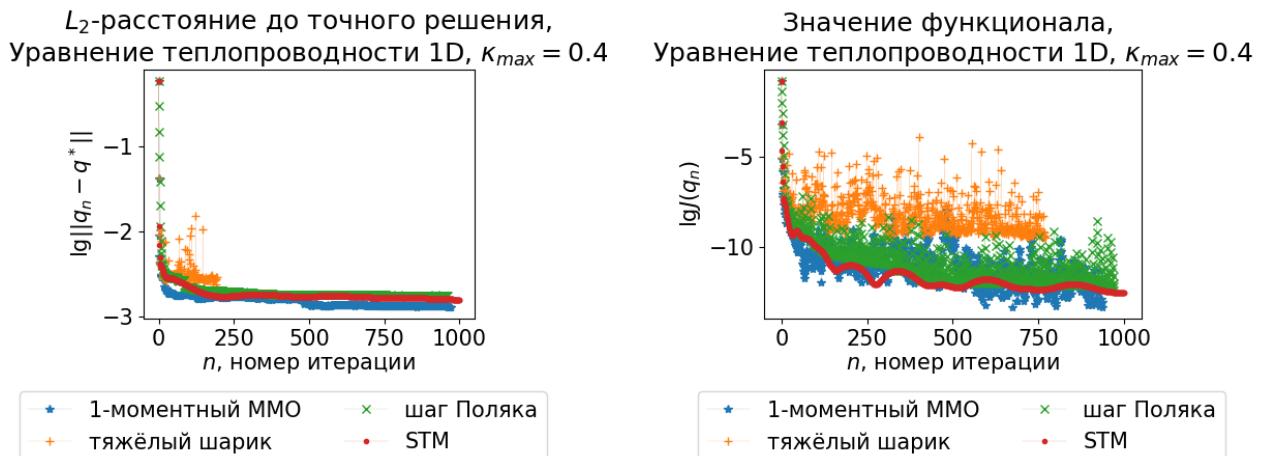


Рисунок 11 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.8), $\kappa_{max} = 0.4$.

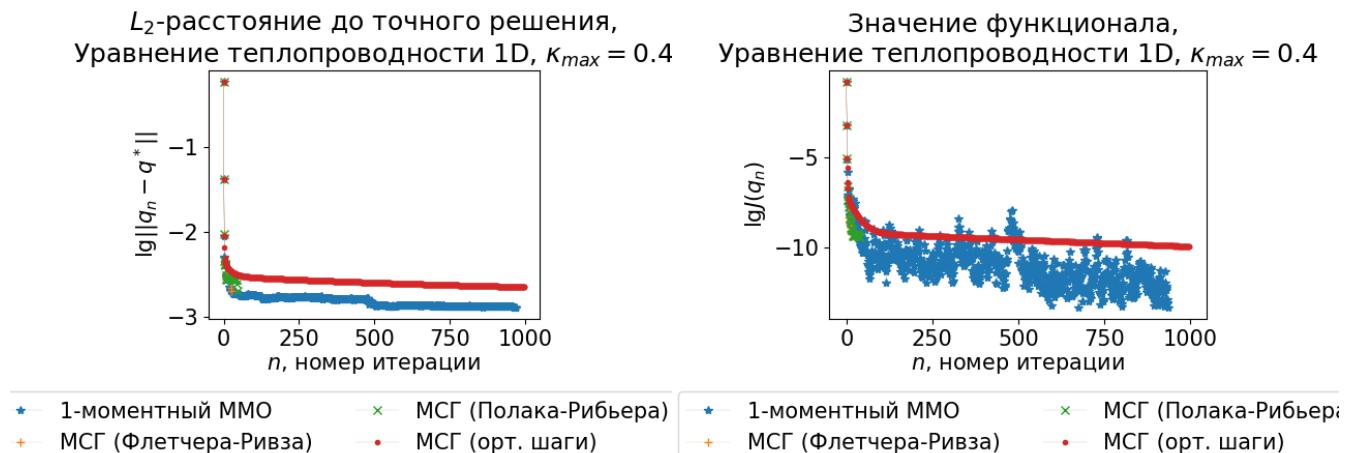


Рисунок 12 – Сходимость: 1-моментный ММО и МСГ. Задача (1.8), $\kappa_{max} = 0.4$.

Увеличение параметра t в многошаговом методе минимальных ошибок не приводит к улучшению результатов.

Сравним полученные решения с известной функцией $q^*(x)$. Изобразим наилучшие по функционалу и по расстоянию приближения, полученные с помощью

1-моментного ММО и с помощью STM, показавшего в данном эксперименте наилучшие результаты среди ранее существовавших методов, на рисунке 13.

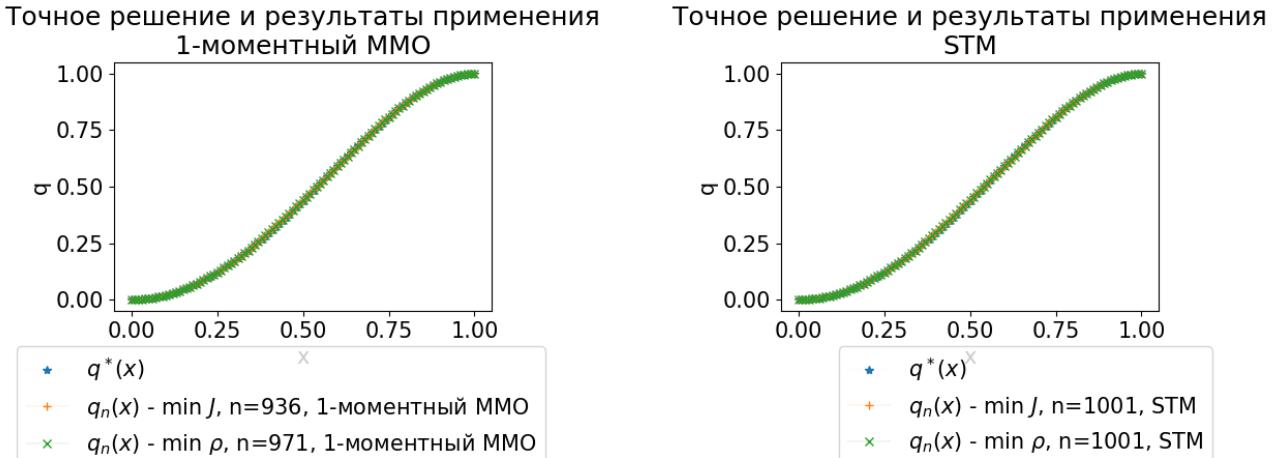


Рисунок 13 – Сравнение с точным решением: 1-моментный ММО, STM. Задача (1.8), $\kappa_{max} = 0.4$.

Функция восстановлена: точки на графике почти неразличимы.

Результаты экспериментов с решением задачи (1.8) с $\kappa_{max} = 0.6$ представлены в таблице 7.

Таблица 7 – Результаты работы методов. Задача (1.8), $\kappa_{max} = 0.6$.

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	2.61×10^{-3}	6.89×10^{-3}	4.76×10^{-14}
ММО, $m = 2$	4.23×10^{-3}	1.01×10^{-2}	1.69×10^{-7}
ММО, $m = 5$	5.95×10^{-3}	1.32×10^{-2}	3.45×10^{-6}
ММО, $m = \infty$	5.95×10^{-3}	1.32×10^{-2}	3.45×10^{-6}
Сопряжённые градиенты (ФР)	2.97×10^{-3}	6.70×10^{-3}	6.83×10^{-10}
Сопряжённые градиенты (ПР)	2.88×10^{-3}	6.69×10^{-3}	6.28×10^{-10}
Сопряжённые градиенты (\perp)	3.73×10^{-3}	9.55×10^{-3}	1.55×10^{-10}
Тяжёлый шарик	3.90×10^{-3}	8.17×10^{-3}	6.08×10^{-10}
Шаг Поляка	3.00×10^{-3}	7.90×10^{-3}	9.70×10^{-14}
STM	2.95×10^{-3}	7.73×10^{-3}	9.15×10^{-14}

Начальное значение функционала $J(q_0) = 0.157$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.583$.

Увеличение κ_{max} с 0.4 до 0.6 не приводит к качественному изменению результатов, что и наблюдаем на рисунках 14 и 15.

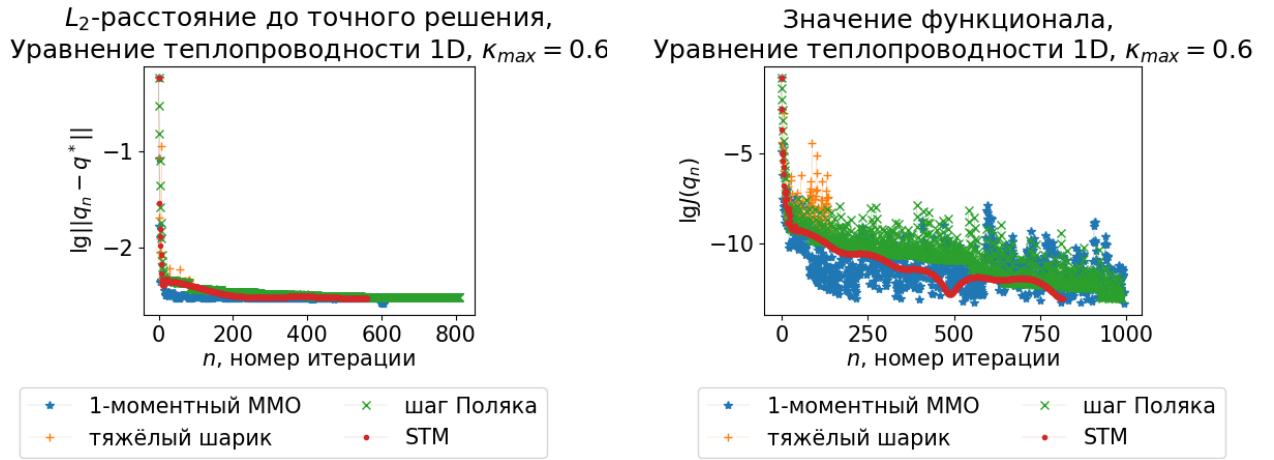


Рисунок 14 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM.
Задача (1.8), $\kappa_{max} = 0.6$.

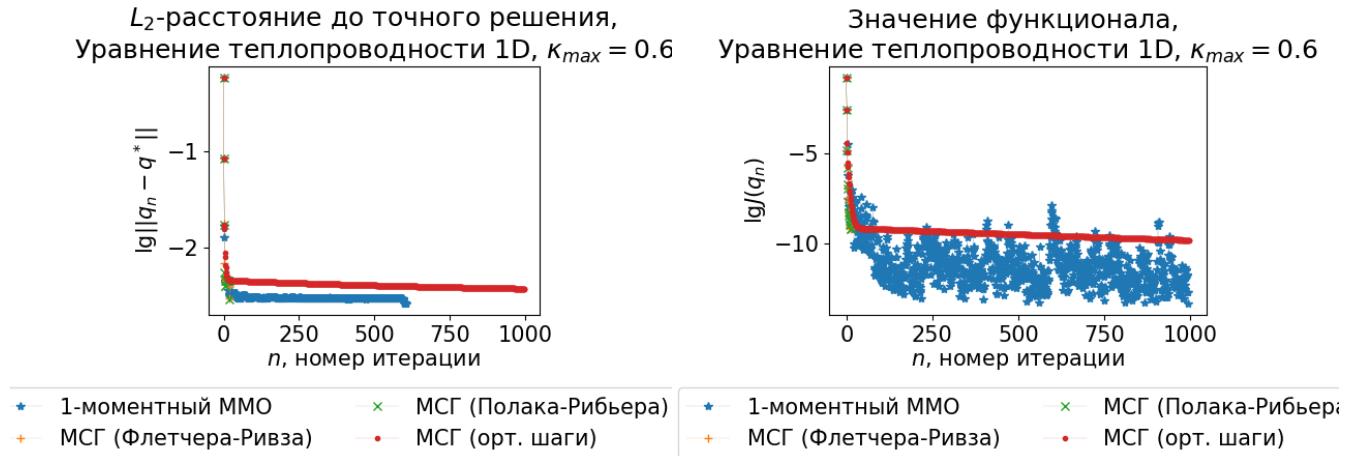


Рисунок 15 – Сходимость: 1-моментный ММО и МСГ. Задача (1.8), $\kappa_{max} = 0.6$.

Сравним полученные решения с известной функцией $q^*(x)$. Изобразим наилучшие по функционалу и по расстоянию приближения, полученные с помощью 1-моментного ММО и с помощью метода сопряжённых градиентов, на рисунке 16.

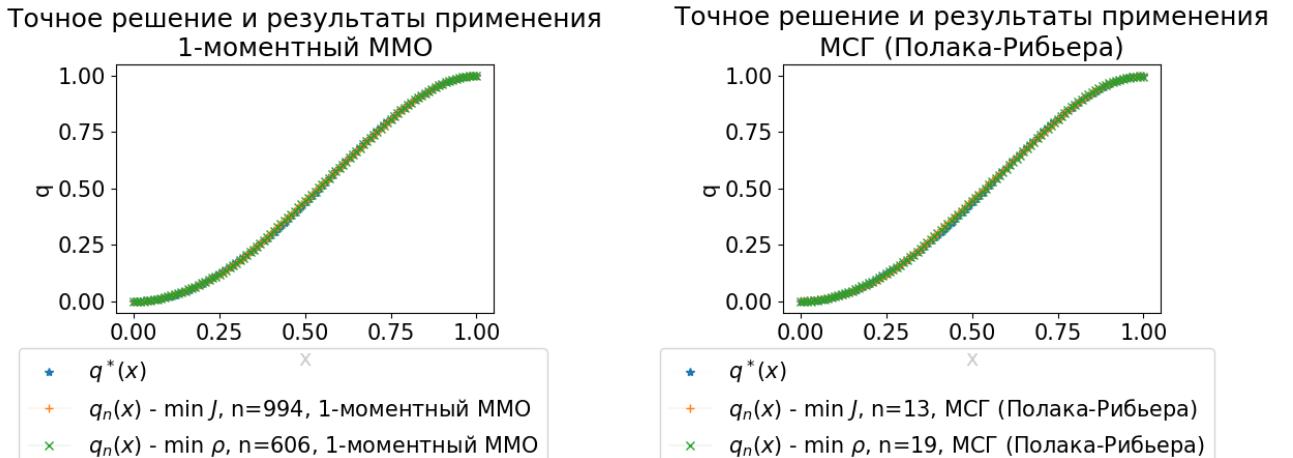


Рисунок 16 – Сравнение с точным решением: 1-моментный ММО, МСГ. Задача (1.8), $\kappa_{max} = 0.6$.

Функция восстановлена: точки на графике почти неразличимы.

Проверим теперь, насколько методы оптимизации чувствительны к шуму. Для этого добавим к рассматриваемой задаче (при $\kappa_{max} = 0.6$) шум в правой части: теперь $f(x) = (Aq^*)(x) \cdot (1 + \delta \cdot \mathcal{U}[-1, 1])$.

Достигнутые расстояния до точного решения для m -моментных ММО при $m \in \{1, 2, 5\}$ и МСГ в форме Флетчера-Ривза с разными δ представлены в таблице 8. Также проверена модификация 1-моментного ММО, в которой $\beta_{2k} \equiv 0$.

Таблица 8 – Расстояние до точного решения при наличии шума δ в правой части. Задача (1.8), $\kappa_{max} = 0.6$.

$\delta, \%$	0.5	1	2	5
MMO, $m = 1$	4.48×10^{-3}	7.71×10^{-3}	1.60×10^{-2}	1.67×10^{-2}
MMO, $m = 1$ (2)	3.87×10^{-3}	5.66×10^{-3}	8.94×10^{-3}	1.49×10^{-2}
MMO, $m = 2$	4.25×10^{-3}	6.93×10^{-3}	1.66×10^{-2}	1.67×10^{-2}
MMO, $m = 5$	5.76×10^{-3}	7.12×10^{-3}	1.66×10^{-2}	1.67×10^{-2}
МСГ	4.50×10^{-3}	4.56×10^{-3}	4.71×10^{-3}	5.45×10^{-3}

Видно, что преимущество m -моментного ММО перед МСГ, наблюдаемое при решении точных задач, теряется при добавлении шума. Чем больше m , тем чувствительнее к шуму m -моментный метод минимальных ошибок. Рестарты позволяют уменьшить невязку, поскольку препятствуют накоплению погрешности.

3.2.6. Выводы

Эксперименты показывают, что градиентные методы оптимизации позволяют эффективно решать ретроспективную задачу Коши для одномерного уравнения теплопроводности, если правая часть операторного уравнения задана точно. Получаемые приближённые решения достаточно близки к известным точным решениям и «в среднем» (по L_2 -норме), и равномерно (по C -норме). Максимальное отклонение полученных решений от точного — на два порядка меньше, чем разброс между минимальным и максимальным значениями $q^*(y)$. Форма графика восстановлена точно.

Однако для получения решений такого же качества, как для задачи (1.5), потребовалось на порядок больше операций. Величина коэффициента теплопроводности также влияет на качество решения: чем он больше, тем хуже работают методы оптимизации.

Это неудивительно, ведь собственные значения связанного с задачей самосопряжённого оператора имеют вид $\lambda_n = e^{-2\pi^2\kappa^2n^2}$, то есть стремятся к нулю значительно быстрее геометрической прогрессии. Чем больше κ , тем быстрее это стремление.

В доказательстве теорем 2 и 3 используется представление расстояния до точного решения в зависимости от номера шага в виде суммы бесконечного числа геометрических прогрессий. Знаменатель n -ой прогрессии равен $1 - \alpha\lambda_n$, поэтому с ростом номера знаменатель стремится к 1.

Коэффициенты, с которыми эти прогрессии входят в сумму, являются коэффициентами Фурье некоторого элемента гильбертова пространства, поэтому они стремятся к нулю. Однако чем быстрее стремятся к единице знаменатели прогрессий, тем больше оказывается вклад медленно убывающих слагаемых и, соответственно, хуже сходимость.

Таким образом, результаты экспериментов согласуются с теоретическими выводами.

Причиной плохой сходимости m -моментного метода минимальных ошибок с $m \geq 2$ может являться низкая точность вычисления длины шага в том случае, когда угол между антиградиентом и линейной оболочкой m предыдущих шагов слишком мал, ведь синус этого угла находится в знаменателе. А метод существенно опирается

на предположение о точности выполнения предыдущих шагов. Данная проблема может быть решена с помощью рестартов.

Однако ММО с $m = 1$ продемонстрировал лучшее качество решения точной задачи, чем все другие методы — как классические, так и более новые. Это верно для обоих вариантов граничных условий.

Если же правая часть операторного уравнения содержит шум, то сходимость ухудшается. Когда шум достаточно мал (до 0.5%), m -моментный ММО с $m \leq 2$ превосходит метод сопряжённых градиентов, но при увеличении шума он теряет преимущество.

3.3. Ретроспективная задача Коши для уравнения теплопроводности в трёхмерном пространстве

Экспериментальные результаты были изложены автором в статье [34].

3.3.1. Происхождение, постановка и некорректность задачи

Задачи являются обобщением задач, рассмотренных в предыдущем разделе, на случай трёхмерного пространства. Их физический смысл очень похож, только вместо изолированного стержня — часть пространства, например, холодильная камера. Известно распределение температуры в области в конечный момент времени. Температура (или тепловой поток) на её границе в течение всего отрезка времени постоянна и равна нулю. Требуется найти распределение температуры в начальный момент времени.

Физический смысл задачи с граничными условиями на значение функции (1.7): область окружена средой с большой теплоёмкостью, поэтому температура на границе не меняется.

Физический смысл задачи с граничными условиями на производную по пространственной координате (1.9): область теплоизолирована, поэтому тепловой поток через границу равен нулю.

Рассмотрим эти задачи в области $\Pi = (0, 1)^3$.

Докажем некорректность обеих задач с постоянным κ и нулевыми граничными

условиями: $a(t) = b(t) = 0$. Взяв $\varepsilon = 1$ и зафиксировав $\delta > 0$, рассмотрим следующие начальные условия:

$f(x) = \frac{\delta}{\pi n} \sin \pi n x_1 \sin \pi n x_2 \sin \pi n x_3$ для задачи с граничными условиями на значение функции и $f(x) = \frac{\delta}{\pi n} \cos \pi n x_1 \cos \pi n x_2 \cos \pi n x_3$ для задачи с граничными условиями на производную по пространственной координате (в обоих случаях n — натуральное число, которое будет выбрано позднее). Решением задачи Коши в каждом случае будет $u(x, t) = e^{3\pi^2 n^2 \kappa^2(1-t)} f(x)$, что легко проверяется по определению.

$\|f\| = \frac{\delta}{\pi n} < \delta$; $\|u\| = \frac{\delta}{\pi n} e^{3\pi^2 n^2 \kappa^2} \rightarrow +\infty$ при $n \rightarrow +\infty$, поэтому найдётся n такое, что $\|u\| > \varepsilon$. Это и обозначает отсутствие непрерывности решения по начальным условиям, то есть некорректность поставленной задачи.

3.3.2. Операторы задач

Ретроспективные задачи Коши для уравнения теплопроводности некорректны. Однако корректны задачи с условием в начальный момент времени:

$$\begin{cases} u_t - \kappa^2 \Delta_x u = 0, & (x, t) \in \Omega = \Pi \times (0, 1) \\ u|_{x \in \partial\Pi} = 0, & t \in [0, 1] \\ u|_{t=0} = q(x), & x \in \Pi \end{cases} \quad (3.3.2.1)$$

или

$$\begin{cases} u_t - \kappa^2 \Delta_x u = 0, & (x, t) \in \Omega = \Pi \times (0, 1) \\ \frac{\partial u}{\partial \bar{n}}|_{x \in \partial\Pi} = 0, & t \in [0, 1] \\ u|_{t=0} = q(x), & x \in \Pi \end{cases} \quad (3.3.2.2)$$

Эти задача отличаются от исходных лишь заменой условия $u|_{t=1} = f(x)$ на $u|_{t=0} = q(x)$. Граница $t = 1$ — наблюдаемая, $t = 0$ — ненаблюдаемая.

Рассуждения, проведённые в одномерном случае, переносятся и на рассматриваемые задачи. Поэтому для каждой из них на пространстве непрерывных функций со стандартным скалярным произведением

$$H = C(\Pi) \subset L^2(\Pi)$$

определим оператор $A : H \rightarrow H$ следующим образом:

$$(Aq)(x) = u(x, 1) \quad \forall x \in \Pi, \quad (3.3.2.3)$$

где u — решение соответствующей корректной задачи Коши (3.3.2.1) или (3.3.2.2).

Важно заметить, что областью определения, как и в одномерном случае, является не всё пространство. Ограничения возникают из условий непрерывности для существования классического решения задачи Коши.

Для оператора задачи с граничными условиями на значение функции: $q(x) = 0$ при $x \in \partial\Pi$. При этом $(Aq)(x) = 0$ для $x \in \partial\Pi$.

Для оператора задачи с граничными условиями на производную по пространственной координате: $\frac{\partial}{\partial n}q(x) = 0$ при $x \in \partial\Pi$. При этом $\frac{\partial}{\partial n}(Aq)(x) = 0$ для $x \in \partial\Pi$.

Вычисление сопряжённого оператора и градиента аналогично одномерному случаю, только вместо одного слагаемого с второй производной по x — три однотипных, для каждой пространственной переменной своё.

Оператор $A^* : H \rightarrow H$ задаётся следующим образом:

$$(A^*p)(x) = \phi(x, 1) \quad \forall x \in \Pi, \quad (3.3.2.4)$$

где ϕ — решение задачи

$$\begin{cases} \phi_t - \Delta_x(\bar{\kappa}^2\phi) = 0, & (x, t) \in \Pi \times (0, 1) \\ \phi|_{x \in \partial\Pi} = 0, & t \in [0, 1] \\ \phi|_{t=0} = p(x), & x \in \Pi \end{cases}. \quad (3.3.2.5)$$

Как видно из определения, оператор A , как и A^* , является линейным.

Важно заметить, что, аналогично одномерному случаю, $A^* = A$ тогда и только тогда, когда κ не зависит от координаты. В противном случае задача, определяющая оператор A^* , отличается от исходной, хотя остаётся параболической и корректной.

Для задачи (1.9) с граничными условиями на производную по пространственной координате оператор $A^* : H \rightarrow H$ задаётся так:

$$(A^*p)(x) = \phi(x, 1) \quad \forall x \in \Pi, \quad (3.3.2.6)$$

где ϕ — решение задачи

$$\begin{cases} \phi_t - \Delta_x(\bar{\kappa}^2\phi)_{xx} = 0, & (x, t) \in \Pi \times (0, 1) \\ (\bar{\kappa}^2\phi)_x|_{x \in \partial\Pi} = 0, & t \in [0, 1] \\ \phi|_{t=0} = p(x), & x \in [0, 1] \end{cases}. \quad (3.3.2.7)$$

3.3.3. Спектр соответствующего задаче самосопряжённого оператора

Оператор A^*A является самосопряжённым. Константа Липшица градиента равна его максимальному собственному значению.

Точное аналитическое вычисление A представляется возможным лишь при $\kappa = \text{const}$. В этом случае $A^* = A$, и $\lambda_m(A^*A) = \lambda_m^2(A)$.

В области определения оператора задачи с граничными условиями на значение функции существует базис — это система функций $\{w_{k,l,m}\}_{k,l,m=1}^\infty = \{\sin \pi kx_1 \sin \pi lx_2 \sin \pi mx_3\}_{k,l,m=1}^\infty$. Непосредственная проверка показывает, что $Aw_{k,l,m} = e^{-\pi^2 \kappa^2 (k^2 + l^2 + m^2)} w_{k,l,m}$.

Соответственно, $\lambda_{k,l,m}(A^*A) = e^{-2\pi^2 \kappa^2 (k^2 + l^2 + m^2)}$, и градиент липшицев с константой $L = e^{-6\pi^2 \kappa^2}$.

В области определения оператора задачи с граничными условиями на нормальную производную по пространственной координате также существует базис — это система функций $\{w_{k,l,m}\}_{k,l,m=0}^\infty = \{\cos \pi kx_1 \cos \pi lx_2 \cos \pi mx_3\}_{k,l,m=0}^\infty$. Непосредственная проверка показывает, что $Aw_{k,l,m} = e^{-\pi^2 \kappa^2 (k^2 + l^2 + m^2)} w_{k,l,m}$.

Соответственно, $\lambda_{k,l,m}(A^*A) = e^{-2\pi^2 \kappa^2 (k^2 + l^2 + m^2)}$, и градиент липшицев с константой $L = 1$.

Эти оценки объединяет следующая лемма.

Лемма 4. При $\kappa > 0$ задача минимизации функционала $J(q)$, соответствующего ретроспективной задаче Коши для уравнения теплопроводности, является выпуклой с липшицевым градиентом, спектр $\lambda_{k,l,m} = e^{-2\pi^2 \kappa^2 (k^2 + l^2 + m^2)}$. Если задача имеет граничные условия на значение функции, то $k, l, m \geq 1$, константа Липшица градиента $L = e^{-6\pi^2 \kappa^2} \leq 1$, константа сильной выпуклости $\mu = 0$. Если задача имеет граничные условия на производную по пространственной координате, то $k, l, m \geq 0$, константа Липшица градиента $L = 1$, константа сильной выпуклости $\mu = 0$.

3.3.4. Приближённые вычисления

Корректные задачи Коши для уравнения теплопроводности в трёхмерном пространстве решаются подобно одномерным аналогам.

Разностные схемы для задач Коши

На $[0, 1]^3 \times [0, 1]$ задаётся сетка с шагами h по пространству и τ по времени, то есть

$$\{x_1 = i_1 h, x_2 = i_2 h, x_3 = i_3 h, t = j\tau \mid 0 \leq i_1, i_2, i_3 \leq P, 0 \leq j \leq T\}$$
$$(Ph = T\tau = 1).$$

Прямая задача с граничными условиями на значение функции аппроксимируется схемой первого порядка по τ и второго порядка по h , аналогичной схеме (3.2.4.1) с однотипными слагаемыми для второй производной по каждой пространственной координате

Разностная схема для задачи с граничными условиями на производную по пространственной координате аналогична схеме (3.2.4.2) с такими же слагаемыми.

Условие сходимости (спектральный признак Неймана) для явных схем даёт условие устойчивости $\tau \leq \frac{h^2}{6\kappa_{max}^2}$.

Сопряжённые задачи аппроксимируются схемами, аналогичными (3.2.4.3) и (3.2.4.4). Используется аппроксимация лапласиана $\bar{\kappa}^2 \phi$.

При постоянном κ эти схемы совпадают с точностью до обозначений со схемами для прямых задач.

Поскольку $\tau = O(h^2)$, схемы имеют второй порядок аппроксимации по h . Из теоремы Рябенького—Лакса [37] следует сходимость второго порядка к решению дифференциальной задачи при согласованном стремлении шагов сетки к нулю.

Приближённое вычисление градиента на сетке

При применении градиентного метода в качестве оракула первого порядка используется алгоритм 5 с соответствующей заменой схем на их трёхмерные аналоги.

Теоретическая оценка погрешности

Поскольку на каждом шаге алгоритма 5 результат отличается от точного на величину, пропорциональную квадрату пространственного шага сетки, то и

$$\max_j |\tilde{\nabla} J(q)_j - \nabla J(q)(jh)| = O(h^2), h \rightarrow 0.$$

Погрешность пропорциональна квадрату мелкости разбиения, поэтому концепция аддитивного шума корректно описывает данный способ вычисления.

Обозначая количество точек отрезка $[0, 1]$ на сетке за P , получим, что решение краевой задачи требует $O(P^3 T) = O(P^5)$ арифметических операций. При этом погрешность вычисления градиента составит $O(P^{-2})$.

3.3.5. Эксперименты

Задача с граничными условиями на значение функции

В качестве иллюстрации для сравнения методов использована модельная задача (1.7) с переменным коэффициентом теплопроводности

$$\kappa(x) = \begin{cases} \kappa_{max}, & \text{если } 0.4 < x_1, x_2, x_3 < 0.6; \\ \frac{\kappa_{max}}{5}, & \text{иначе.} \end{cases},$$

где $\kappa_{max} \in \{0.4, 0.6\}$, точным решением $q(x) = \sin 2\pi x_1 \cdot \sin^2 2\pi x_2 \cdot \sin^3 2\pi x_3$ и условием на наблюдаемой границе $f(x) = (Aq)(x)$.

Результаты экспериментов с $\kappa_{max} = 0.4$ представлены в таблице 9.

Таблица 9 – Результаты работы методов. Задача (1.7), $\kappa_{max} = 0.4$.

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	1.65×10^{-3}	3.91×10^{-2}	1.70×10^{-14}
ММО, $m = 2$	1.65×10^{-3}	3.91×10^{-2}	1.38×10^{-14}
ММО, $m = 5$	1.64×10^{-3}	3.91×10^{-2}	7.27×10^{-15}
ММО, $m = \infty$	1.18×10^{-3}	3.91×10^{-2}	5.03×10^{-26}
Сопряжённые градиенты (ФР)	1.66×10^{-3}	3.92×10^{-2}	7.51×10^{-16}
Сопряжённые градиенты (ПР)	1.66×10^{-3}	3.92×10^{-2}	9.67×10^{-16}
Сопряжённые градиенты (\perp)	6.59×10^{-3}	4.35×10^{-2}	1.46×10^{-9}
Тяжёлый шарик	6.73×10^{-3}	4.19×10^{-2}	2.75×10^{-9}
Шаг Поляка	3.09×10^{-3}	3.92×10^{-2}	1.76×10^{-12}
STM	2.89×10^{-3}	3.92×10^{-2}	2.14×10^{-12}

Использованы сетка с шагом $h = 0.04$ и нулевое начальное приближение. Начальное значение функционала $J(q_0) = 6.27 \times 10^{-3}$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.242$.

Рисунок 17 показывает, что 1-моментный метод минимальных ошибок позволяет достичь меньшей невязки, чем адаптивный метод тяжёлого шарика из статьи [51], градиентный спуск с шагом Поляка и метод подобных треугольников, и по аргументу, и по функционалу. При этом сходимость по функционалу не является монотонной.

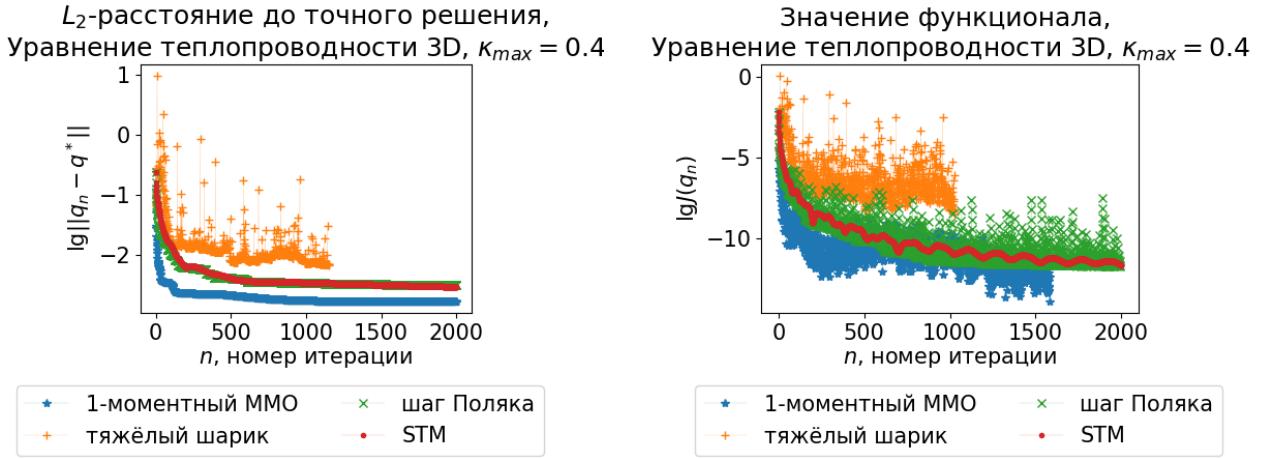


Рисунок 17 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.7), $\kappa_{max} = 0.4$.

Рисунок 18 показывает, что классические варианты метода сопряжённых градиентов и новый 1-моментный метод минимальных ошибок дают близкую невязку.

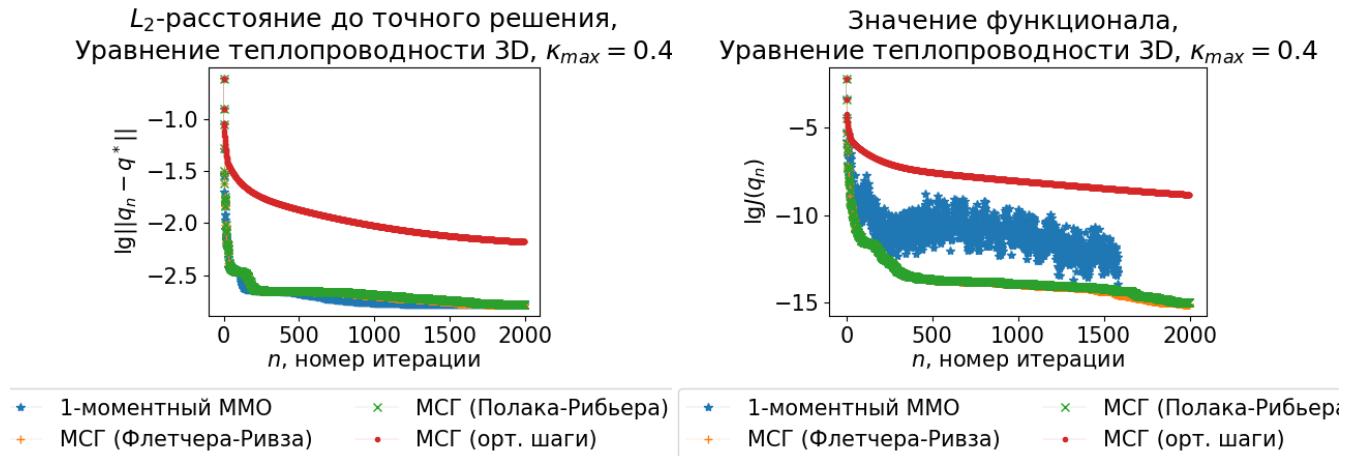


Рисунок 18 – Сходимость: 1-моментный ММО и МСГ. Задача (1.7), $\kappa_{max} = 0.4$.

Рисунок 19 показывает, что ∞ -моментный метод минимальных ошибок превосходит остальные рассмотренные методы: он достигает меньшей невязки и по аргументу, и по функционалу.

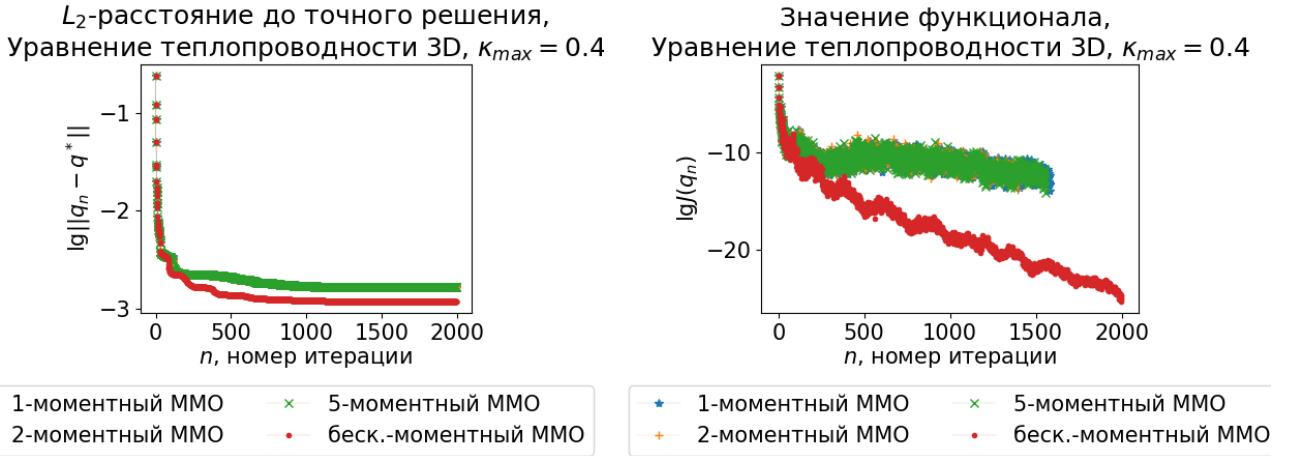


Рисунок 19 – Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.7), $\kappa_{max} = 0.4$.

Результаты экспериментов с $\kappa_{max} = 0.6$ представлены в таблице 10.

Таблица 10 – Результаты работы методов. Задача (1.7), $\kappa_{max} = 0.6$.

	$ q_n - q^* $	$ q_n - q^* _C$	$J(q_n)$
MMO, $m = 1$	3.74×10^{-3}	3.92×10^{-2}	9.93×10^{-16}
MMO, $m = 2$	3.72×10^{-3}	3.92×10^{-2}	1.35×10^{-15}
MMO, $m = 5$	3.72×10^{-3}	3.92×10^{-2}	2.36×10^{-15}
MMO, $m = \infty$	1.66×10^{-3}	3.92×10^{-2}	3.41×10^{-25}
Сопряжённые градиенты (ФР)	4.76×10^{-3}	3.93×10^{-2}	3.89×10^{-16}
Сопряжённые градиенты (ПР)	6.00×10^{-3}	3.94×10^{-2}	9.26×10^{-16}
Сопряжённые градиенты (\perp)	74.3×10^{-3}	0.280	4.14×10^{-8}
Тяжёлый шарик	63.4×10^{-3}	0.216	3.26×10^{-8}
Шаг Поляка	26.0×10^{-3}	9.41×10^{-2}	1.19×10^{-10}
STM	29.6×10^{-3}	0.110	2.99×10^{-10}

Начальное значение функционала $J(q_0) = 1.34 \times 10^{-3}$, начальное расстояние до точного решения $||q_0 - q^*|| = 0.242$.

Эксперименты с $\kappa_{max} = 0.6$ дают очень похожие результаты. Моментный метод минимальных ошибок достигает меньшего расстояния до точного решения, чем все остальные рассмотренные методы. $m = \infty$ позволяет достичь наименьшей невязки как по аргументу, так и по функционалу. Сходимость различных методов оптимизации показана на рисунках 20, 21 и 22.

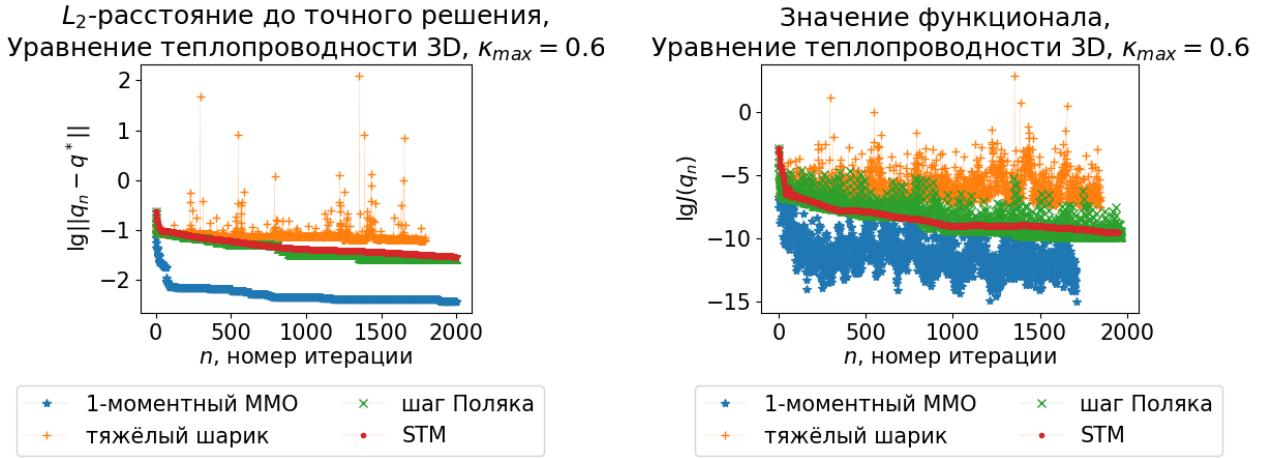


Рисунок 20 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM.
Задача (1.7), $\kappa_{max} = 0.6$.

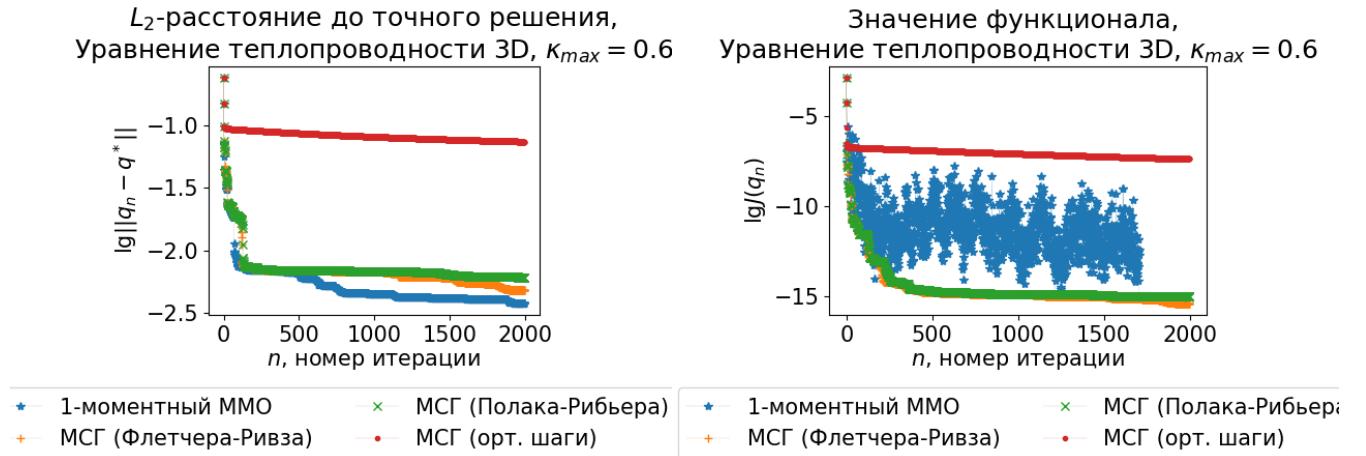


Рисунок 21 – Сходимость: 1-моментный ММО и МСГ. Задача (1.7), $\kappa_{max} = 0.6$.

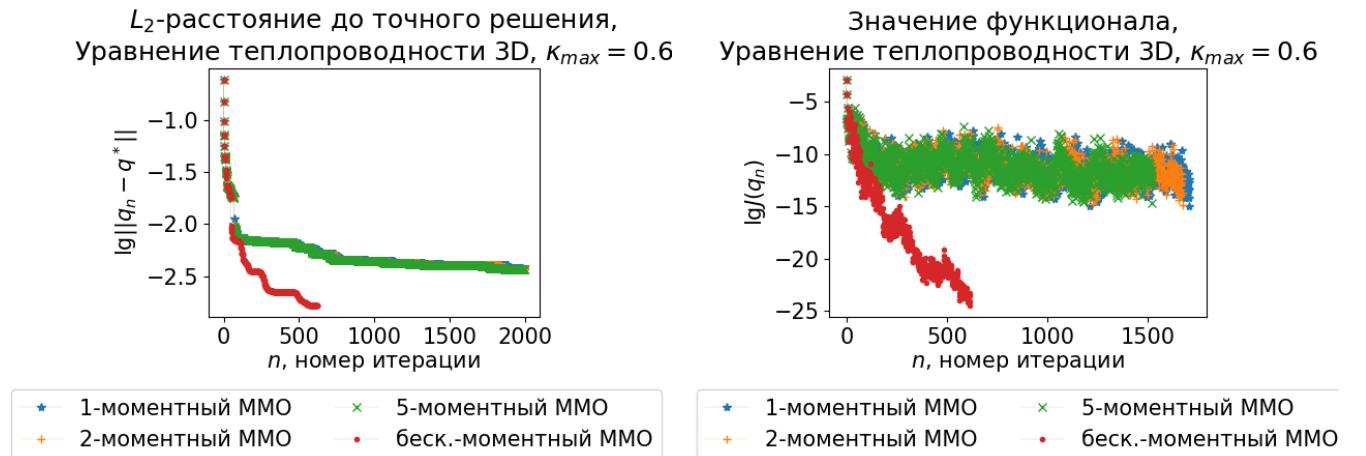


Рисунок 22 – Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.7), $\kappa_{max} = 0.6$.

Задача с граничными условиями на производную по пространственной координате

В качестве иллюстрации для сравнения методов использована модельная задача (1.9) с переменным коэффициентом теплопроводности

$$\kappa(x) = \begin{cases} \kappa_{max}, & \text{если } 0.4 < x_1, x_2, x_3 < 0.6; \\ \frac{\kappa_{max}}{5}, & \text{иначе.} \end{cases},$$

где $\kappa_{max} \in \{0.1, 0.2\}$, точным решением $q(x) = \cos 2\pi x_1 \cdot \cos^2 2\pi x_2 \cdot \cos^3 2\pi x_3$ и условием на наблюдаемой границе $f(x) = (Aq)(x)$.

Результаты экспериментов с $\kappa_{max} = 0.1$ представлены в таблице 11.

Таблица 11 – Результаты работы методов. Задача (1.9), $\kappa_{max} = 0.1$.

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	7.12211×10^{-3}	0.111	4.82×10^{-19}
ММО, $m = 2$	7.12209×10^{-3}	0.116	5.57×10^{-20}
ММО, $m = 5$	7.12172×10^{-3}	0.120	5.19×10^{-21}
ММО, $m = \infty$	18.3×10^{-3}	0.269	1.04×10^{-8}
Сопряжённые градиенты (ФР)	7.12211×10^{-3}	0.113	2.26×10^{-25}
Сопряжённые градиенты (ПР)	7.12260×10^{-3}	0.112	8.99×10^{-21}
Сопряжённые градиенты (\perp)	38.1×10^{-3}	0.488	1.27×10^{-7}
Тяжёлый шарик	18.3×10^{-3}	0.269	2.45×10^{-11}
Шаг Поляка	18.3×10^{-3}	0.269	3.71×10^{-16}
STM	18.3×10^{-3}	0.269	4.11×10^{-16}

Использованы сетка с шагом $h = 0.04$ и нулевое начальное приближение. Начальное значение функционала $J(q_0) = 0.0240$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.242$.

Рисунок 23 показывает, что 1-моментный метод минимальных ошибок позволяет достичь меньшей невязки, чем адаптивный метод тяжёлого шарика из статьи [51], градиентный спуск с шагом Поляка и метод подобных треугольников, и по аргументу, и по функционалу. При этом сходимость по функционалу не является монотонной.

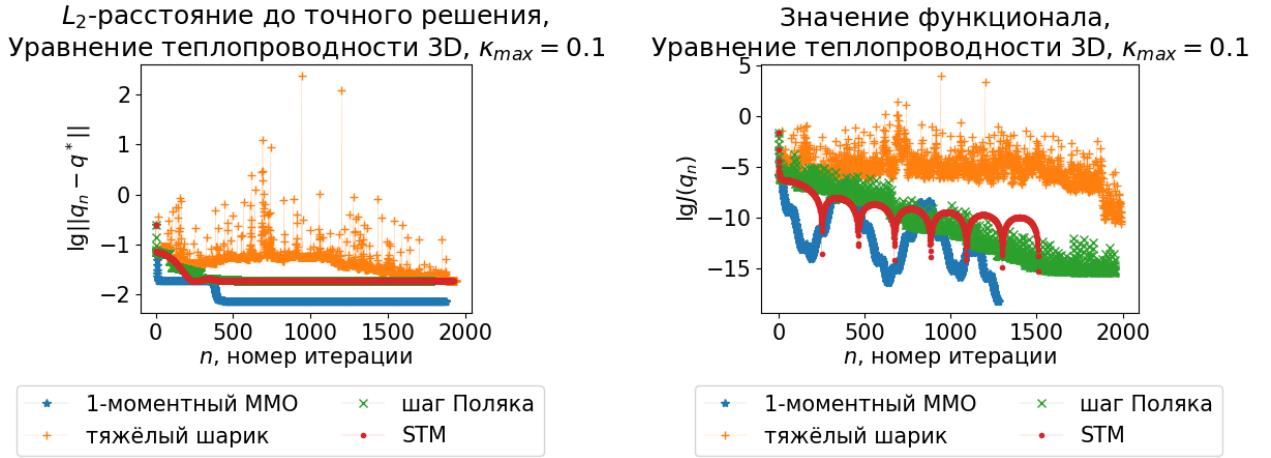


Рисунок 23 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.9), $\kappa_{max} = 0.1$.

Рисунок 24 показывает, что классические варианты метода сопряжённых градиентов и новый 1-моментный метод минимальных ошибок дают близкую невязку по аргументу: $\varepsilon \approx 7.12 \times 10^{-3}$. Невязка по функционалу у метода Флетчера-Ривза оказалась меньше. Однако 1-моментный метод минимальных ошибок достиг ε -окрестности точного решения примерно на 100 итераций быстрее.

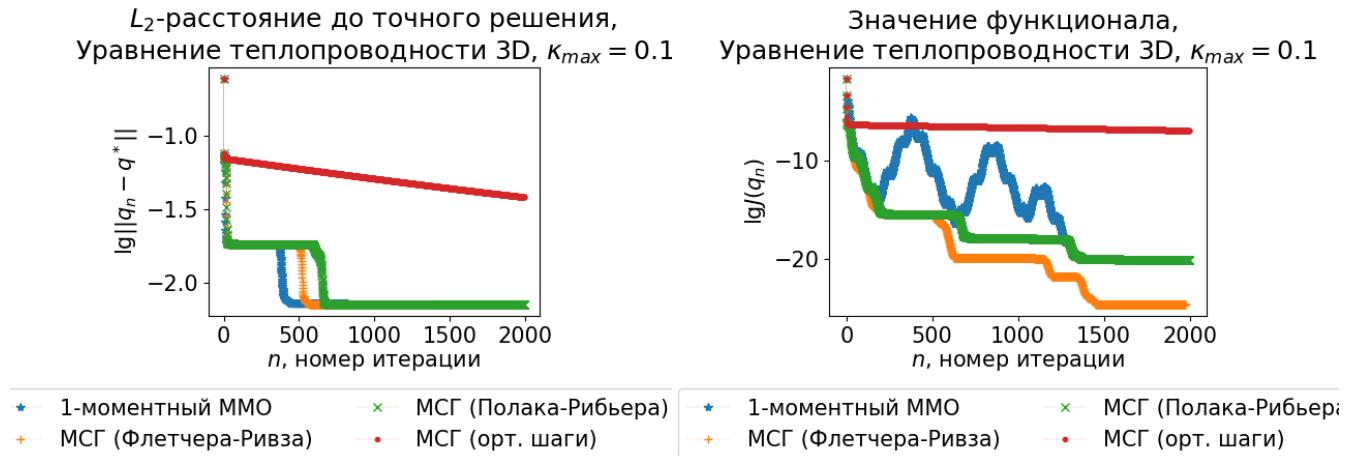


Рисунок 24 – Сходимость: 1-моментный ММО и МСГ. Задача (1.9), $\kappa_{max} = 0.1$.

Рисунок 25 показывает, что при увеличении m от 1 до 5 m -моментный метод минимальных ошибок достигает ε -окрестности точного решения раньше. Однако при $m = \infty$ работа метода ухудшается.

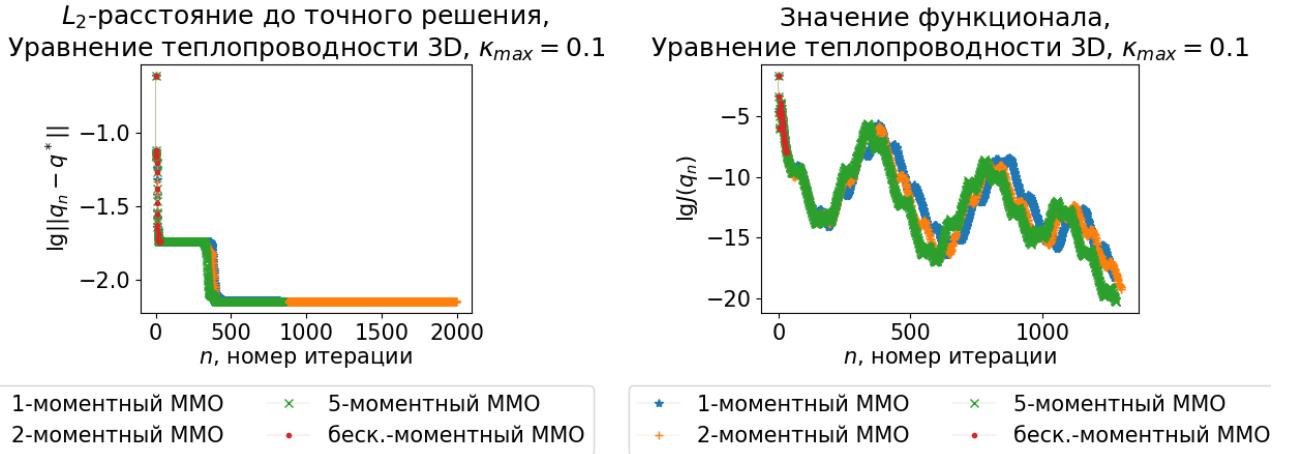


Рисунок 25 – Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.9), $\kappa_{max} = 0.1$.

Эксперименты с $\kappa_{max} = 0.2$ показывают, что ни один из методов не даёт приемлемого качества решения. Расстояние до точного решения уменьшается лишь в 3 раза за первые несколько итераций, после чего практически перестаёт меняться. Результаты представлены в таблице 12.

Таблица 12 – Результаты работы методов. Задача (1.9), $\kappa_{max} = 0.2$.

	$ q_n - q^* $	$ q_n - q^* _C$	$J(q_n)$
ММО, $m = 1$	7.03×10^{-2}	0.937	1.85×10^{-12}
ММО, $m = 2$	7.03×10^{-2}	0.937	1.80×10^{-12}
ММО, $m = 5$	7.03×10^{-2}	0.937	5.10×10^{-11}
ММО, $m = \infty$	7.04×10^{-2}	0.927	2.62×10^{-7}
Сопряжённые градиенты (ФР)	7.03×10^{-2}	0.937	8.86×10^{-15}
Сопряжённые градиенты (ПР)	7.03×10^{-2}	0.937	5.64×10^{-15}
Сопряжённые градиенты (\perp)	7.03×10^{-2}	0.938	3.95×10^{-11}
Тяжёлый шарик	7.03×10^{-2}	0.924	7.98×10^{-12}
Шаг Поляка	7.03×10^{-2}	0.937	1.72×10^{-13}
STM	7.03×10^{-2}	0.937	3.92×10^{-14}

Начальное значение функционала $J(q_0) = 0.0168$, начальное расстояние до точного решения $||q_0 - q^*|| = 0.242$.

Работа методов — уменьшение функционала и расстояния до точного решения — показана на рисунках 26, 27 и 28.

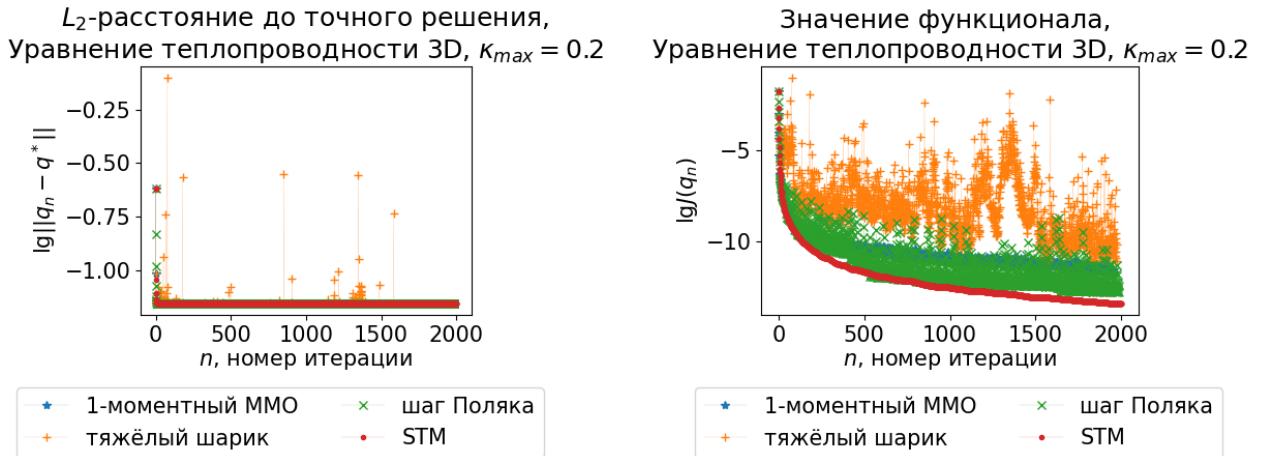


Рисунок 26 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM.
Задача (1.9), $\kappa_{max} = 0.2$.

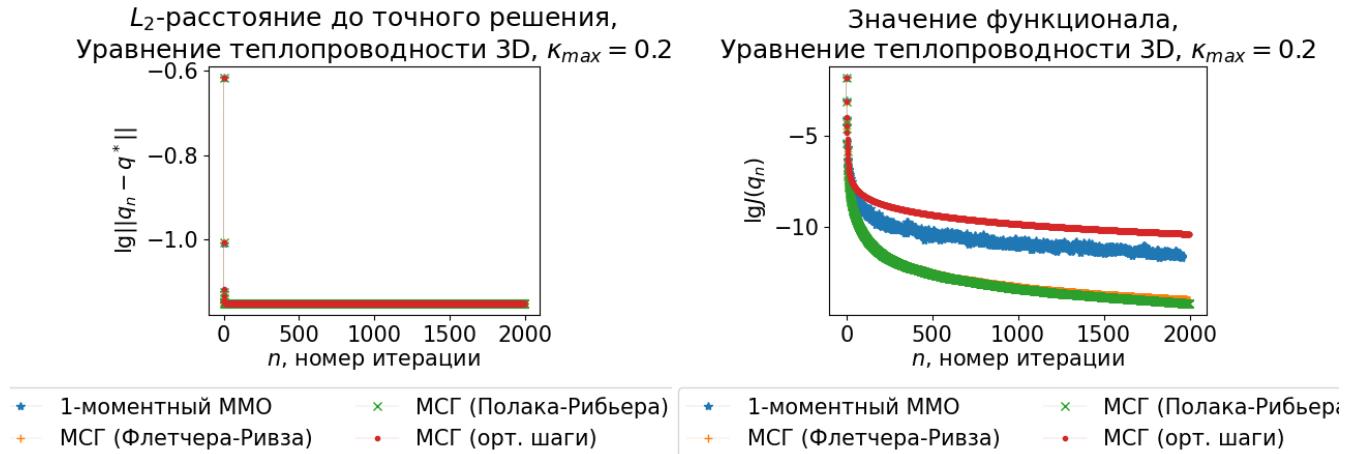


Рисунок 27 – Сходимость: 1-моментный ММО и МСГ. Задача (1.9), $\kappa_{max} = 0.2$.

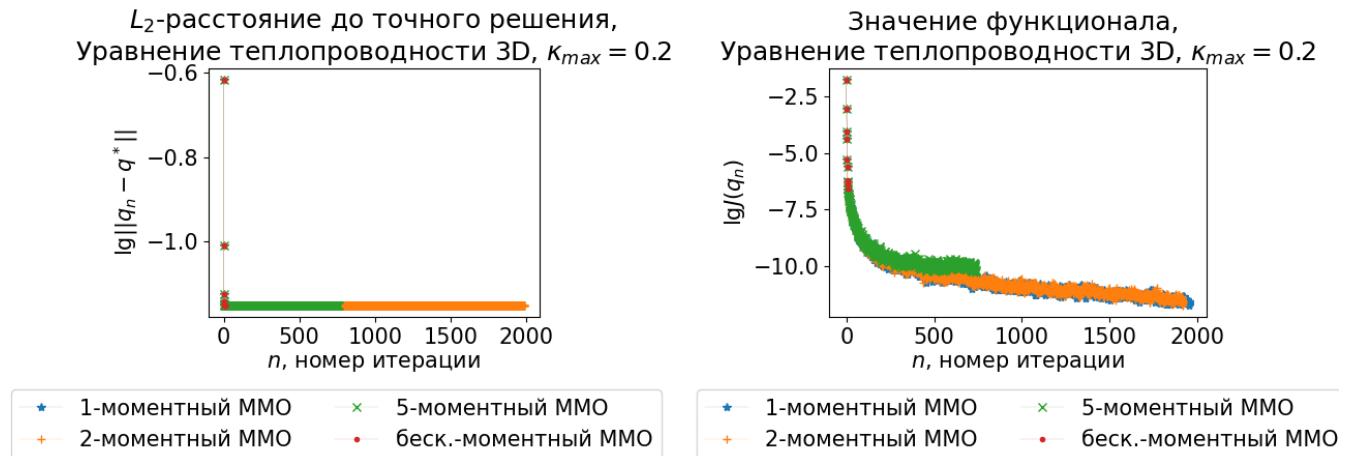


Рисунок 28 – Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.9), $\kappa_{max} = 0.2$.

3.3.6. Выводы

Эксперименты показывают, что градиентные методы оптимизации позволяют эффективно решать ретроспективную задачу Коши для трёхмерного уравнения теплопроводности с граничными условиями на значение функции. Получаемые приближённые решения достаточно близки к известным точным решениям и «в среднем» (по L_2 -норме), и равномерно (по C -норме).

Максимальное отклонение полученных решений от точного — почти на два порядка меньше, чем разброс между минимальным и максимальным значениями $q^*(y)$. Это свидетельствует о хорошем восстановлении формы графика.

Как и в одномерном случае, увеличение коэффициента теплопроводности ухудшает качество решения. Это не вызывает удивления, ведь свойство «жёсткости» спектра сохраняется при переходе в пространство более высокой размерности.

Моментный метод минимальных ошибок позволяет получить лучшее качество решения по сравнению с другими методами. Особенно хорошие результаты даёт метод с $m = \infty$.

Этот результат вполне согласуется с теоретическими выкладками —теоремами 1 и 3.

Качество решения задачи с граничными условиями на производную по пространственной координате оказалось значительно хуже. Получаемые приближённые решения достаточно близки к известным точным решениям «в среднем» (по L_2 -норме), но отклонение по C -норме оказалось значительно больше.

Для m -моментного ММО и метода сопряжённых градиентов максимальное отклонение полученных решений от точного составило примерно 5% разброса между минимальным и максимальным значениями $q^*(y)$. Это свидетельствует о восстановлении формы графика «в общем», но локальные свойства решения могут утрачиваться.

Такой результат связан с особенностями аппроксимации задачи разностными схемами. Значение функции при переходе к дискретной задаче сохраняется, а производная заменяется на её разностное приближение. В окрестностях рёбер куба, где граница области, на которой задано уравнение, теряет гладкость, это приближение оказывается особенно плохим.

Если для точной задачи рёбрами (как множеством меры нуль) можно пренебречь, то в дискретной задаче с делением отрезка на P частей они содержат $12P$ точек,

то есть $\frac{12P}{P^3} = \frac{12}{P^2}$ от всех точек. При $P = 25$ это почти 2% — достаточно много. Неудивительно, что погрешность накапливается.

Эта погрешность сильно сказывается на работе m -моментных методов минимальных ошибок, которые по своему устройству весьма чувствительны к точности вычислений. Такое влияние погрешностей вполне согласуется с наблюдениями, сделанными для одномерных задач с искусственным шумом.

Достижение приемлемого качества решения задачи с граничными условиями на производную по пространственной координате и сравнение методов оказались возможными лишь при достаточно малом коэффициенте теплопроводности: $\kappa_{max} = 0.1$. В этом случае моментный метод минимальных ошибок и классические варианты метода сопряжённых градиентов позволяют получить примерно одинаковое качество решения.

Сравним вычислительную сложность указанных решений. Как было отмечено в описании графиков, 1-моментный ММО затрачивает для достижения наилучшего решения примерно на 100 итераций меньше, чем метод сопряжённых градиентов. Это около 20% от количества итераций метода сопряжённых градиентов. Самая затратная операция — вычисление As (или A_0s) по элементу $s \in H$. Каждая итерация m -моментного ММО требует двух вычислений оператора, а каждая итерация метода сопряжённых градиентов — трёх. Поэтому время работы 1-моментного ММО составляет примерно $0.8 \cdot \frac{2}{3} \approx \frac{1}{2}$ времени работы метода сопряжённых градиентов при одинаковом качестве решения.

Соответственно, ретроспективная задача Коши для трёхмерного уравнения теплопроводности с граничными условиями на производную по пространственной координате допускает эффективное решение с помощью нового моментного ММО тогда же, когда хорошего решения можно достичь методом сопряжённых градиентов; в таком случае вычислительная сложность нового метода оказывается примерно вдвое ниже.

3.4. Обратная задача термоакустики

Экспериментальные результаты этого раздела изложены в статье [34].

3.4.1. Происхождение, постановка и некорректность задачи

Задача 1.10 взята из статьи [22] С. И. Кабанихина, М. А. Шишленина и О. И. Кривортько, по которой излагается необходимая теоретическая часть данного раздела. Указано, что задача связана с томографией и имеет прямое применение в медицине. В той же статье обоснована её некорректность.

Эксперименты, как и в статье [22], проводятся с тремя задачами, которые отличаются количеством дополнительной информации. Здесь $L = T = 1$, хотя это необязательно.

Обратная задача 1:

$$\begin{cases} u_{tt} = u_{xx} + u_{yy}, & (x, y) \in \Omega = (0, 1) \times (0, 1), t \in (0, 1) \\ u_t|_{t=0} = 0, & (x, y) \in \Omega \\ u_x|_{x=0} = u_x|_{x=1} = 0, & y \in (0, 1), t \in (0, 1) \\ u_y|_{y=0} = u_y|_{y=1} = 0, & x \in (0, 1), t \in (0, 1) \\ u|_{x=0} = f_1(y, t), & y \in (0, 1), t \in (0, 1) \end{cases} \quad (3.4.1.1)$$

Обратная задача 2:

$$\begin{cases} u_{tt} = u_{xx} + u_{yy}, & (x, y) \in \Omega = (0, 1) \times (0, 1), t \in (0, 1) \\ u_t|_{t=0} = 0, & (x, y) \in \Omega \\ u_x|_{x=0} = u_x|_{x=1} = 0, & y \in (0, 1), t \in (0, 1) \\ u_y|_{y=0} = u_y|_{y=1} = 0, & x \in (0, 1), t \in (0, 1) \\ u|_{x=0} = f_1(y, t), & y \in (0, 1), t \in (0, 1) \\ u|_{x=1} = f_2(y, t), & y \in (0, 1), t \in (0, 1) \end{cases} \quad (3.4.1.2)$$

Обратная задача 3:

$$\begin{cases} u_{tt} = u_{xx} + u_{yy}, & (x, y) \in \Omega = (0, 1) \times (0, 1), t \in (0, 1) \\ u_t|_{t=0} = 0, & (x, y) \in \Omega \\ u_x|_{x=0} = u_x|_{x=1} = 0, & y \in (0, 1), t \in (0, 1) \\ u_y|_{y=0} = u_y|_{y=1} = 0, & x \in (0, 1), t \in (0, 1) \\ u|_{x=0} = f_1(y, t), & y \in (0, 1), t \in (0, 1) \\ u|_{x=1} = f_2(y, t), & y \in (0, 1), t \in (0, 1) \\ u|_{y=1} = f_3(x, t), & x \in (0, 1), t \in (0, 1) \end{cases} \quad (3.4.1.3)$$

3.4.2. Операторы задач

Прямая задача термоакустики:

$$\begin{cases} u_{tt} = u_{xx} + u_{yy}, & (x, y) \in \Omega = (0, 1) \times (0, 1), t \in (0, 1) \\ u|_{t=0} = q(x, y), & (x, y) \in \Omega \\ u_t|_{t=0} = 0, & (x, y) \in \Omega \\ u_x|_{x=0} = u_x|_{x=1} = 0, & y \in (0, 1), t \in (0, 1) \\ u_y|_{y=0} = u_y|_{y=1} = 0, & x \in (0, 1), t \in (0, 1) \end{cases} \quad (3.4.2.1)$$

Эта корректно поставленная задача получается заменой граничных условий на функцию начальным условием. По смыслу задачи, именно начальное условие нужно найти, зная граничные. Функция $u(x, y, t)$ определена на параллелепипеде. Его граница $t = 0$ — ненаблюдаемая; границы $x = 0, x = 1, y = 1$ — наблюдаемые.

Для каждой наблюдаемой границы можно определить свой оператор. Важно, что все они действуют из $C^1([0, L] \times [0, L])$ в $C^1([0, L] \times [0, T])$. В общем случае, области определения функций q и $A_i q$ могут оказаться различными.

$$\begin{aligned} (A_1 q)(y, t) &= u(0, y, t) \quad \forall (y, t) \in [0, 1] \times [0, 1] \\ (A_2 q)(y, t) &= u(1, y, t) \quad \forall (y, t) \in [0, 1] \times [0, 1] \\ (A_3 q)(x, t) &= u(x, 1, t) \quad \forall (x, t) \in [0, 1] \times [0, 1] \end{aligned} \quad (3.4.2.2)$$

Во всех случаях $u(x, y, t)$ — решение корректной задачи (3.4.2.1).

Для каждой из поставленных обратных задач определим функционал:

$$J(q) = \frac{1}{2} \sum_{l=1}^s \|A_l q - f_l\|^2 \quad (3.4.2.3)$$

Здесь s — номер задачи. При $s \in \{2, 3\}$ то тот самый составной функционал (2.3.4.1), вычисления с которым проводились в разделе 2.3.4.

Градиент, как обычно, вычисляется с помощью сопряжённых операторов. Приведём их формулы согласно [22]:

$$A_l^* : C^1([0, 1] \times [0, 1]) \rightarrow C^1([0, 1] \times [0, 1])$$

$$(A_1^* p)(x, y) = \psi_{1t}(x, y, 0), \text{ где } \psi_1(x, y, t) — \text{решение задачи}$$

$$\begin{cases} \psi_{1tt} = \psi_{1xx} + \psi_{1yy}, & (x, y) \in \Omega = (0, 1) \times (0, 1), t \in (0, 1) \\ \psi_1|_{t=T} = \psi_{1t}|_{t=T} = 0, & (x, y) \in \Omega \\ \psi_{1x}|_{x=0} = p(y, t), & y \in (0, 1), t \in (0, 1) \\ \psi_{1x}|_{x=1} = 0, & y \in (0, 1), t \in (0, 1) \\ \psi_{1y}|_{y=0} = \psi_{1y}|_{y=1} = 0, & x \in (0, 1), t \in (0, 1) \end{cases}$$

$(A_2^*p)(x, y) = \psi_{2t}(x, y, 0)$, где $\psi_2(x, y, t)$ — решение задачи

$$\begin{cases} \psi_{2tt} = \psi_{2xx} + \psi_{2yy}, & (x, y) \in \Omega = (0, 1) \times (0, 1), t \in (0, 1) \\ \psi_2|_{t=T} = \psi_{2t}|_{t=T} = 0, & (x, y) \in \Omega \\ \psi_{2x}|_{x=0} = 0, & y \in (0, 1), t \in (0, 1) \\ \psi_{2x}|_{x=1} = -p(y, t), & y \in (0, 1), t \in (0, 1) \\ \psi_{2y}|_{y=0} = \psi_{2y}|_{y=1} = 0, & x \in (0, 1), t \in (0, 1) \end{cases}$$

$(A_3^*p)(x, y) = \psi_{3t}(x, y, 0)$, где $\psi_2(x, y, t)$ — решение задачи

$$\begin{cases} \psi_{3tt} = \psi_{3xx} + \psi_{3yy}, & (x, y) \in \Omega = (0, 1) \times (0, 1), t \in (0, 1) \\ \psi_3|_{t=T} = \psi_{3t}|_{t=T} = 0, & (x, y) \in \Omega \\ \psi_{3x}|_{x=0} = \psi_{3x}|_{x=1} = 0, & y \in (0, 1), t \in (0, 1) \\ \psi_{3y}|_{y=0} = 0, & x \in (0, 1), t \in (0, 1) \\ \psi_{3y}|_{y=1} = -p(x, t), & x \in (0, 1), t \in (0, 1) \end{cases}$$

Точный оракул первого порядка задаётся алгоритмом 6.

Algorithm 6 Gradient Thermoacoustics Precise Oracle

Input: $q(x, y) \in C^1([0, L] \times [0, L])$; s — номер решаемой обратной задачи; $f_l(y, t)$ ($1 \leq l \leq s$) — параметры задачи.

- 1: вычислить $u(x, y, t)$ — решение задачи (3.4.2.1);
- 2: положить $J = 0$, $\mathbf{g} = \mathbf{0}$;
- 3: **for** $l = 1 \dots s$ **do**
- 4: вычислить $A_l q$ согласно формуле (3.4.2.2);
- 5: вычислить $A_l q - f_l$;
- 6: вычислить S_l — двойной интеграл $A_l q - f_l$ по $[0, L] \times [0, T]$;
- 7: $J := J + \frac{S_l}{2}$;
- 8: вычислить ψ_l — решить соответствующую сопряжённую задачу;
- 9: $\mathbf{g} := \mathbf{g} + \frac{\partial}{\partial t} \psi_l(x, y, t)|_{t=0}$.
- 10: **end for**

Output: $J(q) = J$, $\nabla J(q)(x, y) = \mathbf{g}$.

3.4.3. Приближённые вычисления

Разностная схема для задач Коши

На $[0, 1] \times [0, 1] \times [0, 1]$ задаётся равномерная сетка с шагами h по пространственным координатам и τ по времени. То есть,

$$x = ih, y = jh, t = k\tau, 0 \leq i, j \leq P \quad (Ph = 1), \quad 0 \leq k \leq T \quad (T\tau = 1).$$

Разностная схема для решения прямой задачи:

$$\left\{ \begin{array}{l} \frac{u_{i,j,k+1} - 2u_{i,j,k} + u_{i,j,k-1}}{\tau^2} = \frac{u_{i+1,j,k} - 2u_{i,j,k} + u_{i-1,j,k}}{h^2} + \frac{u_{i,j+1,k} - 2u_{i,j,k} + u_{i,j-1,k}}{h^2} \\ u_{i,j,0} = q_{i,j} \\ \frac{u_{i,j,1} - u_{i,j,0}}{\tau} = 0 \\ \frac{u_{1,j,k} - u_{0,j,k}}{h} = 0 \\ \frac{u_{P,j,k} - u_{P-1,j,k}}{h} = 0 \\ \frac{u_{i,1,k} - u_{i,0,k}}{h} = 0 \\ \frac{u_{i,P,k} - u_{i,P-1,k}}{h} = 0 \end{array} \right. \quad (3.4.3.1)$$

Разностная схема для решения сопряжённых задач:

$$\left\{ \begin{array}{l} \frac{\psi_{i,j,k+1} - 2\psi_{i,j,k} + \psi_{i,j,k-1}}{\tau^2} = \frac{\psi_{i+1,j,k} - 2\psi_{i,j,k} + \psi_{i-1,j,k}}{h^2} + \frac{\psi_{i,j+1,k} - 2\psi_{i,j,k} + \psi_{i,j-1,k}}{h^2} \\ \psi_{i,j,T} = 0 \\ \frac{\psi_{i,j,T} - \psi_{i,j,T-1}}{\tau} = 0 \\ \frac{\psi_{1,j,k} - \psi_{0,j,k}}{h} = p_{j,k}^{(1)} \\ \frac{\psi_{P,j,k} - \psi_{P-1,j,k}}{h} = -p_{j,k}^{(2)} \\ \frac{\psi_{i,1,k} - \psi_{i,0,k}}{h} = 0 \\ \frac{\psi_{i,P,k} - \psi_{i,P-1,k}}{h} = -p_{i,k}^{(3)} \end{array} \right. \quad (3.4.3.2)$$

Ненулевым является одно из граничных условий $p^{(i)}$ в зависимости от номера задачи.

Обе схемы имеют первый порядок аппроксимации по h, τ . При согласованном стремлении шагов сетки к нулю решение разностной задачи сходится к решению исходной дифференциальной задачи.

Приближённое вычисление градиента на сетке

В качестве оракула первого порядка используется алгоритм 7:

Algorithm 7 Gradient Thermoacoustics Approximate Oracle

Input: $q_{i,j} = q(ih, jh)$; s — номер решаемой обратной задачи; $f_{j,k}^{(l)} = f^{(l)}(jh, k\tau)$ ($1 \leq l \leq s$) — параметры задачи.

- 1: вычислить $u_{i,j,k}$ — решение задачи (3.4.3.1);
- 2: положить $J = 0$, $\mathbf{g} = \mathbf{0}$;
- 3: **for** $l = 1 \dots s$ **do**
- 4: вычислить $A_l q$: $(A_1 q)_{j,k} = u_{0,j,k}$, $(A_2 q)_{j,k} = u_{P,j,k}$ или $(A_3 q)_{i,k} = u_{i,P,k}$;
- 5: вычислить $A_l q - f^{(l)}$;
- 6: приближённо вычислить S_l — двойной интеграл $A_l q - f_l$ по $[0, L] \times [0, T]$;
- 7: $J := J + \frac{S_l}{2}$;
- 8: вычислить $\psi^{(l)}$ — решить соответствующую сопряжённую задачу с помощью сетки (3.4.3.2);
- 9: вычислить $\psi_t^{(l)}|_{t=0} := \frac{\psi_{i,j,1}^{(l)} - \psi_{i,j,0}^{(l)}}{\tau}$
- 10: $\mathbf{g} := \mathbf{g} + \psi_t^{(l)}|_{t=0}$.
- 11: **end for**

Output: $J(q) = J$, $\tilde{\nabla} J(q)(x, y) = \mathbf{g}$.

Теоретическая оценка погрешности

Поскольку на каждом шаге алгоритма 7 результат отличается от точного на величину, пропорциональную пространственному шагу сетки, то и

$$\max_j |\tilde{\nabla} J(q)_j - \nabla J(q)(jh)| = O(h), h \rightarrow 0.$$

Погрешность пропорциональна мелкости разбиения, поэтому концепция аддитивного шума корректно описывает данный способ вычисления.

Обозначая количество точек отрезка $[0, 1]$ на сетке за P , получим, что решение задачи Коши требует $O(P^2T)$ арифметических операций. При этом погрешность вычисления градиента составит $O(P^{-1})$.

3.4.4. Эксперименты

Для проверки методов используются задачи с

$$q^*(x, y) = 0.1 + \frac{(1 + \cos 8\pi x)(1 + \cos 8\pi y)}{32} \left[x, y \in \left[\frac{1}{8}, \frac{3}{8} \right] \cup \left[\frac{5}{8}, \frac{7}{8} \right] \right].$$

На таких же задачах проводились эксперименты в статье [22].

$q^*(x, y)$ моделирует область с четырьмя неоднородностями.

Для приближённого решения задач применяются сетки с $h = 0.02$, $\tau = 0.0001$.

Эксперименты с обратной задачей 1

Результаты представлены в таблице 13.

Таблица 13 – Результаты работы методов. Задача (3.4.1.1).

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	8.94×10^{-3}	4.79×10^{-2}	1.99×10^{-8}
ММО, $m = 2$	8.88×10^{-3}	4.77×10^{-2}	6.80×10^{-9}
ММО, $m = 5$	7.35×10^{-3}	3.42×10^{-2}	1.31×10^{-10}
ММО, $m = \infty$	12.2×10^{-3}	5.76×10^{-2}	3.28×10^{-6}
Сопряжённые градиенты (ФР)	14.8×10^{-3}	6.85×10^{-2}	1.89×10^{-5}
Сопряжённые градиенты (ПР)	14.8×10^{-3}	6.69×10^{-2}	1.84×10^{-5}
Сопряжённые градиенты (\perp)	10.3×10^{-3}	5.67×10^{-2}	1.20×10^{-8}
Тяжёлый шарик	10.4×10^{-3}	5.36×10^{-2}	4.05×10^{-5}
Шаг Поляка	8.50×10^{-3}	4.58×10^{-2}	1.90×10^{-9}
STM	13.6×10^{-3}	7.06×10^{-2}	1.02×10^{-5}

Использованы сетка с шагами $h = 0.02$, $\tau = 0.0001$ и нулевое начальное приближение. Начальное значение функционала $J(q_0) = 6.07 \times 10^{-3}$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.11$.

Рисунок 29 показывает, что только 5-моментный метод минимальных ошибок и градиентный спуск с шагом Поляка демонстрируют монотонную сходимость; два других метода быстро перестают уменьшать невязку.

5-моментный метод минимальных ошибок показывает значительно лучшие результаты.

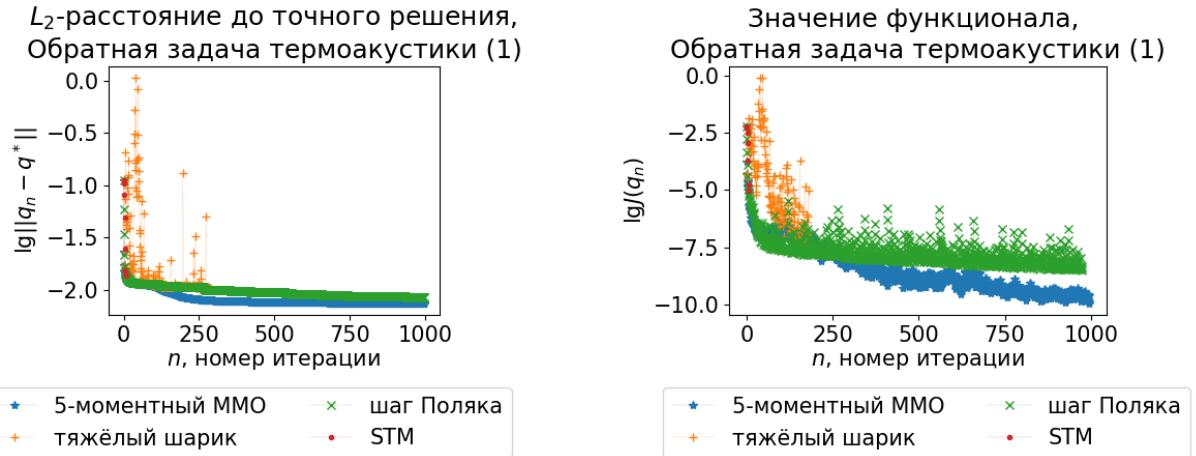


Рисунок 29 – Сходимость: 5-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (3.4.1.1).

Классические варианты метода сопряжённых градиентов перестают уменьшать невязку после нескольких итераций (рисунок 30). Метод с ортогональными шагами и вспомогательной минимизацией функционала продолжает работать, но его результаты хуже, чем у 5-моментного ММО.

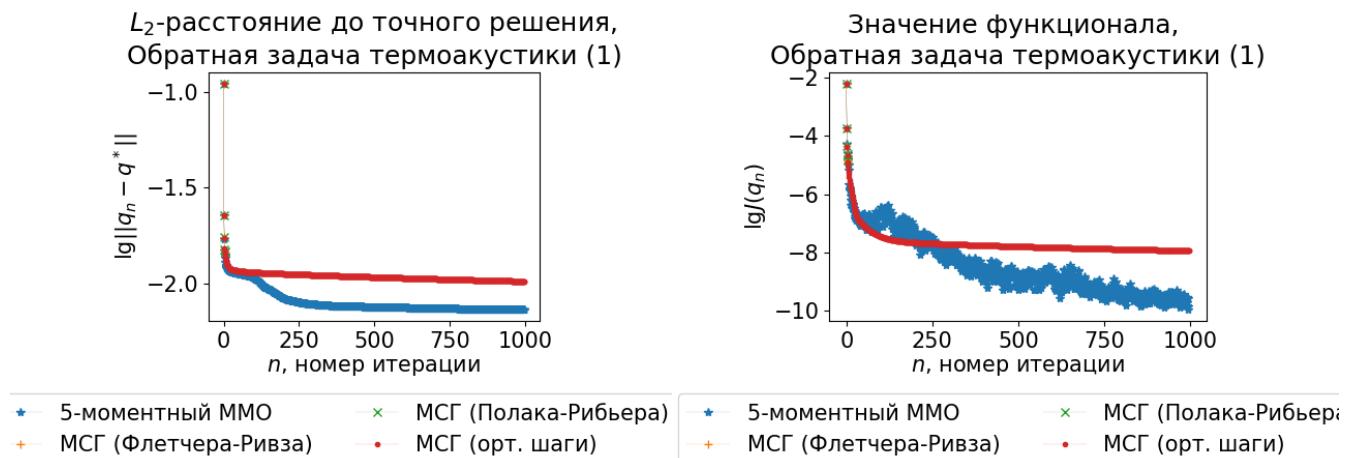
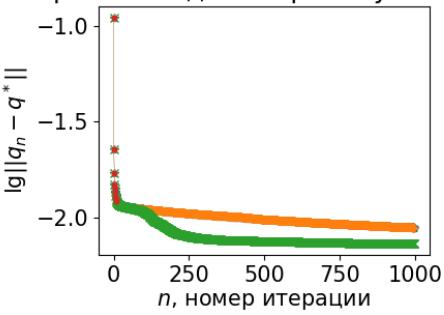


Рисунок 30 – Сходимость: 5-моментный ММО и МСГ. Задача (3.4.1.1).

Среди m -моментных методов минимальных ошибок наилучшие результаты достигнуты при $m = 5$, что видно на рисунке 31. Метод с $m = \infty$, наилучший в теории, слишком чувствителен к погрешностям вычислений.

L_2 -расстояние до точного решения,
Обратная задача термоакустики (1)



Значение функционала,
Обратная задача термоакустики (1)

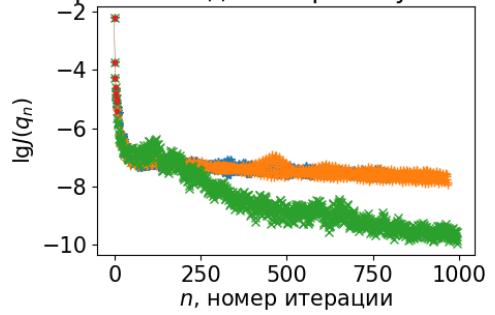


Рисунок 31 – Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (3.4.1.1).

Сравним полученные решения с известной функцией $q^*(x, y)$. Изобразим наилучшие по функционалу приближения, полученные с помощью 5-моментного ММО, МСГ и градиентного спуска с шагом Поляка, на рисунке 32.

Точное решение

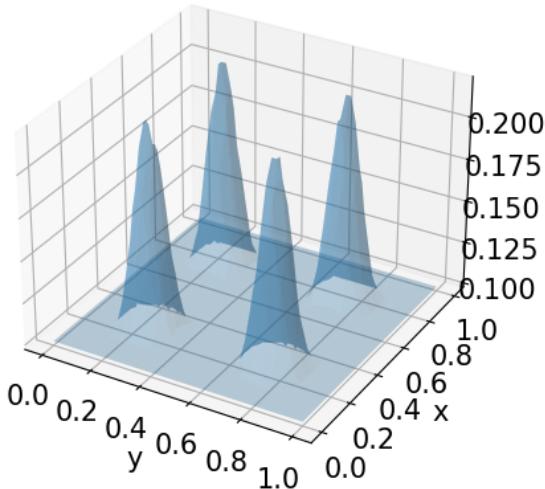
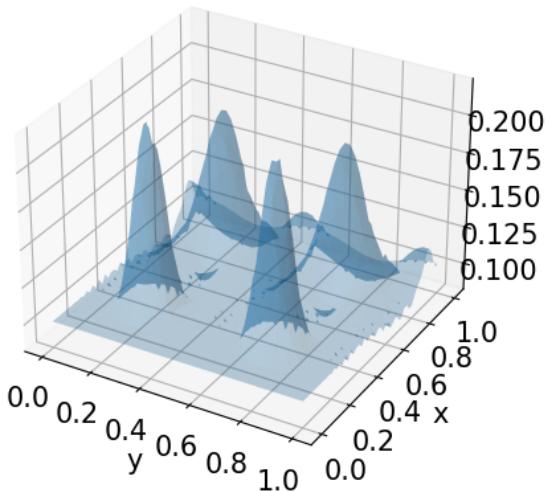
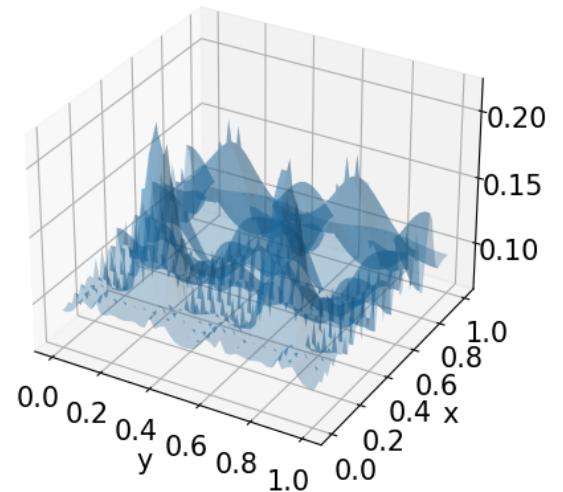
шаг Поляка
 $q_n(x, y) - \min J, n=975$ 5-моментный ММО
 $q_n(x, y) - \min J, n=996$ МСГ (Полака-Рибьера)
 $q_n(x, y) - \min J, n=5$ 

Рисунок 32 – Сравнение с точным решением: 5-моментный ММО, МСГ, шаг Поляка. Задача (3.4.1.1).

Как и в статье [22], лучше всего восстанавливаются неоднородности, близкие к границе $x = 0$, на которой задано условие. С восстановлением других неоднородностей методыправляются по-разному. Самый лучший результат у 5-моментного ММО, что подтверждается как вычисленными метриками, так и видом восстановленного графика: высота неоднородностей ближе к истинной, чем у других методов.

Оценим теперь ухудшение качества решения при наличии погрешности

измерений правой части интегрального уравнения. Для этого добавим к рассматриваемой задаче шум в правой части: $f(x) = (Aq^*)(x) \cdot (1 + \delta \cdot \mathcal{U}[-1, 1])$.

Достигнутые расстояния до точного решения для m -моментных ММО при $m \in \{1, 2, 5\}$ и МСГ в форме Флетчера-Ривза с разными δ представлены в таблице 14. Также проверена модификация 1-моментного ММО, в которой $\beta_{2k} \equiv 0$.

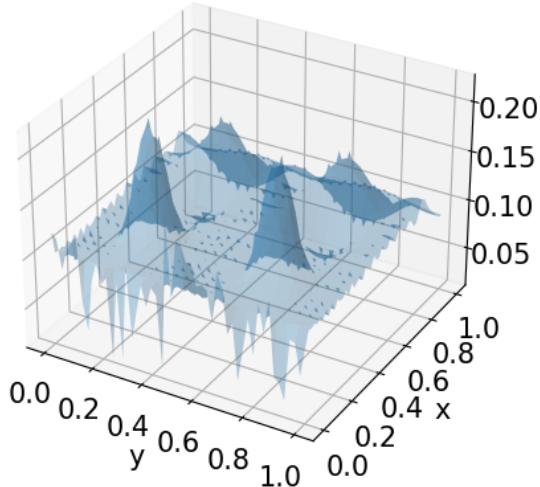
Таблица 14 – Расстояние до точного решения при наличии шума δ в правой части. Задача (3.4.1.1).

δ	2×10^{-4}	5×10^{-4}	10^{-3}
ММО, $m = 1$	1.26×10^{-2}	1.40×10^{-2}	1.53×10^{-2}
ММО, $m = 1$ (2)	1.25×10^{-2}	1.38×10^{-2}	1.52×10^{-2}
ММО, $m = 2$	1.25×10^{-2}	1.37×10^{-2}	1.52×10^{-2}
ММО, $m = 5$	1.24×10^{-2}	1.33×10^{-2}	1.53×10^{-2}
МСГ	1.49×10^{-2}	1.53×10^{-2}	1.61×10^{-2}

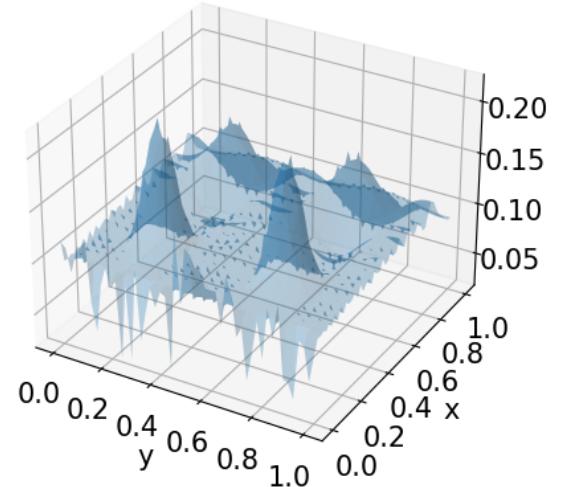
Эти значения невязки примерно соответствуют результатам работы МСГ при отсутствии шума.

На рисунке 33 построены графики функций, порождаемых различными методами при величине шума $\delta = 2 \times 10^{-4}$.

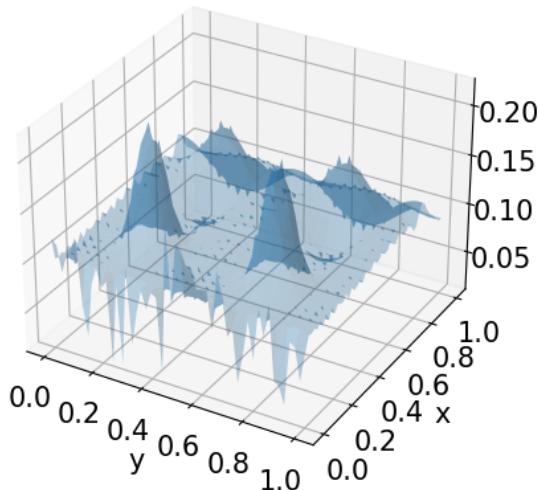
1-моментный ММО, рестарт 2
 $q_n(x, y) - \min J, n=21$



5-моментный ММО
 $q_n(x, y) - \min J, n=14$



2-моментный ММО
 $q_n(x, y) - \min J, n=17$



МСГ (Флетчера-Ривза)
 $q_n(x, y) - \min J, n=5$

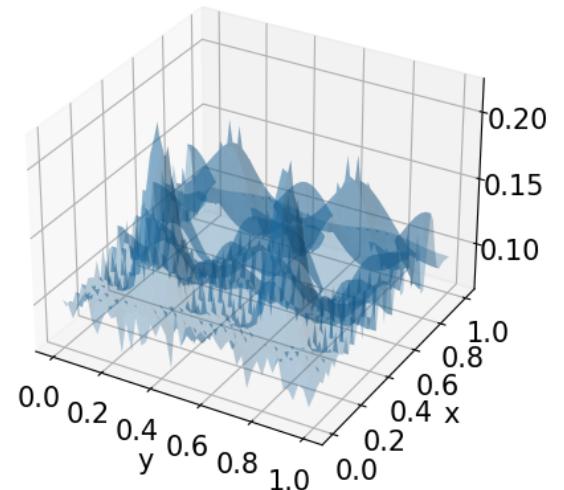
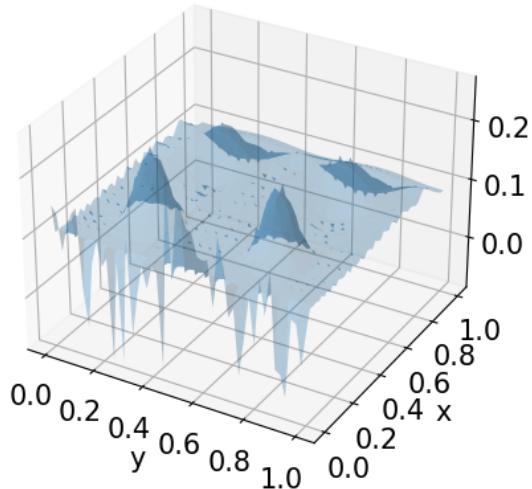


Рисунок 33 – Сравнение с точным решением: m -моментный ММО с $m \in \{1, 2, 5\}$, МСГ. Шум $\delta = 2 \times 10^{-4}$. Задача (3.4.1.1).

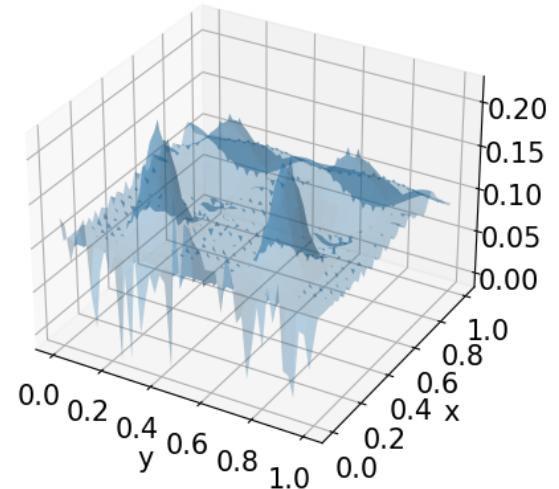
Разные варианты m -моментного ММО восстанавливают две неоднородности, хотя и с большими отклонениями в других областях. Если допускается использование априорной информации (например, $q^*(x, y) \geq 0.1$), то эти погрешности устранимы. Результат работы МСГ значительно хуже.

На рисунке 34 построены графики функций, порождаемых различными методами при величине шума $\delta = 5 \times 10^{-4}$.

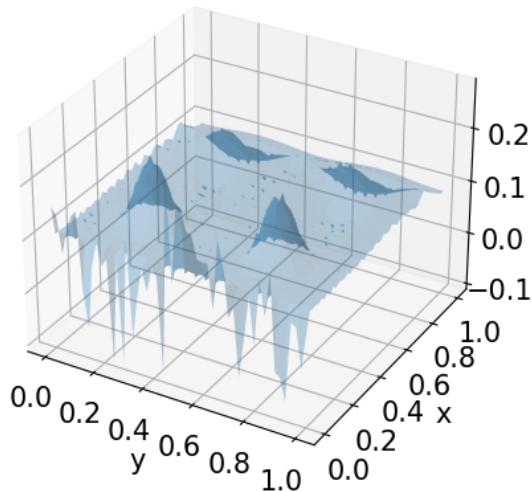
1-моментный ММО, рестарт 2
 $q_n(x, y) - \min J$, n=18



5-моментный ММО
 $q_n(x, y) - \min J$, n=18



2-моментный ММО
 $q_n(x, y) - \min J$, n=17



МСГ (Флетчера-Ривза)
 $q_n(x, y) - \min J$, n=5

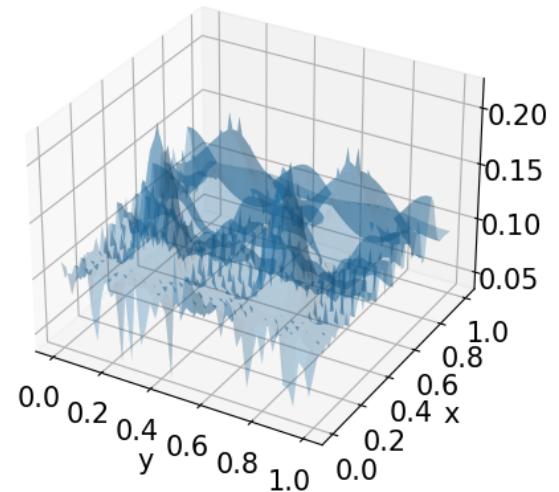
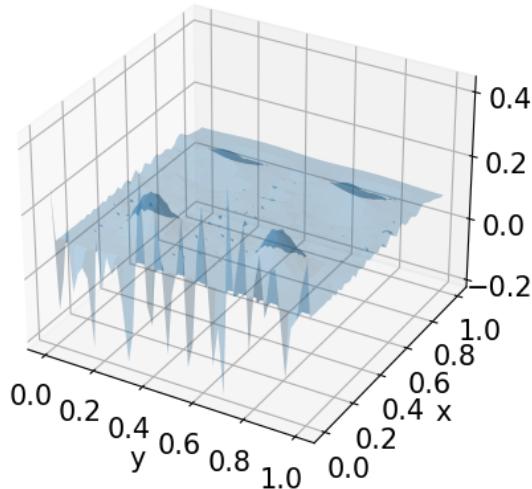


Рисунок 34 – Сравнение с точным решением: m -моментный ММО с $m \in \{1, 2, 5\}$, МСГ. Шум $\delta = 5 \times 10^{-4}$. Задача (3.4.1.1).

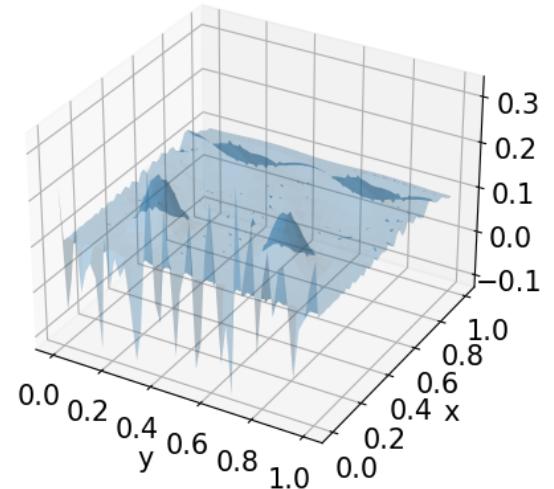
5-моментный ММО справился с восстановлением ближайших к поверхности измерения неоднородностей лучше остальных методов.

На рисунке 35 построены графики функций, порождаемых различными методами при величине шума $\delta = 10^{-3}$.

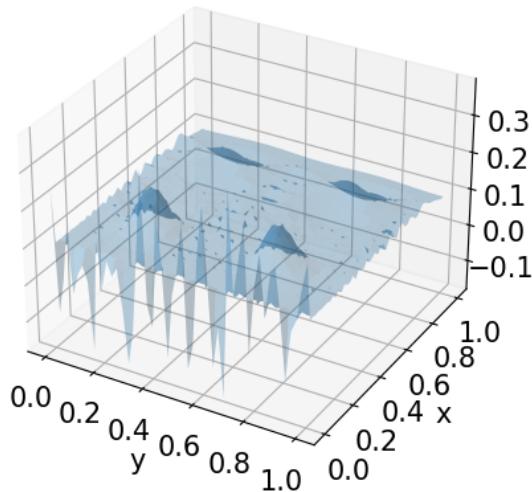
1-моментный ММО, рестарт 2
 $q_n(x, y) - \min J$, n=14



5-моментный ММО
 $q_n(x, y) - \min J$, n=10



2-моментный ММО
 $q_n(x, y) - \min J$, n=13



МСГ (Флетчера-Ривза)
 $q_n(x, y) - \min J$, n=4

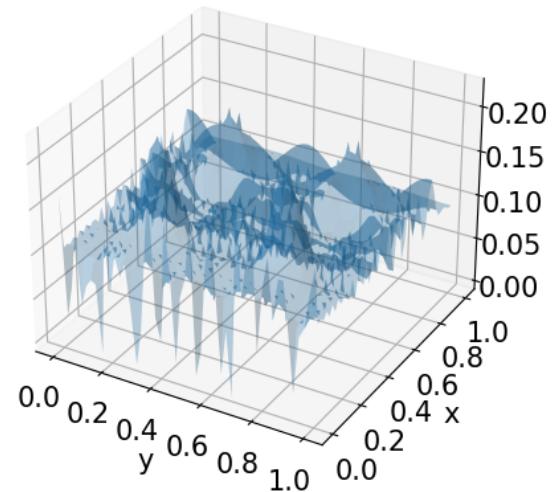


Рисунок 35 – Сравнение с точным решением: m -моментный ММО с $m \in \{1, 2, 5\}$, МСГ. Шум $\delta = 10^{-3}$. Задача (3.4.1.1).

При шуме $\delta = 10^{-3}$ неоднородности восстанавливаются очень плохо. Следовательно, обратная задача 1 в условиях неточных измерений на наблюдаемой границе решается только частично и лишь при малых погрешностях: $\delta \leq 5 \times 10^{-4}$.

Эксперименты с обратной задачей 2

Результаты представлены в таблице 15.

Таблица 15 – Результаты работы методов. Задача (3.4.1.2).

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	1.15×10^{-4}	5.33×10^{-4}	7.88×10^{-12}
ММО, $m = 2$	9.66×10^{-5}	4.16×10^{-4}	1.07×10^{-12}
ММО, $m = 5$	4.31×10^{-5}	1.67×10^{-4}	1.41×10^{-13}
ММО, $m = \infty$	3.95×10^{-3}	1.08×10^{-2}	5.60×10^{-6}
Сопряжённые градиенты (ФР)	1.13×10^{-3}	3.75×10^{-3}	2.69×10^{-7}
Сопряжённые градиенты (ПР)	5.93×10^{-4}	1.90×10^{-3}	2.38×10^{-9}
Сопряжённые градиенты (\perp)	1.91×10^{-4}	8.24×10^{-4}	2.96×10^{-11}
Тяжёлый шарик	6.35×10^{-4}	2.20×10^{-3}	1.00×10^{-10}
Шаг Поляка	9.90×10^{-5}	4.29×10^{-4}	9.25×10^{-13}
STM	3.70×10^{-3}	1.42×10^{-2}	2.75×10^{-6}

Использованы сетка с шагами $h = 0.02$, $\tau = 0.0001$ и нулевое начальное приближение. Начальное значение функционала $J(q_0) = 0.012$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.11$.

Рисунок 36 показывает, что только 5-моментный метод минимальных ошибок и градиентный спуск с шагом Поляка демонстрируют монотонную сходимость.

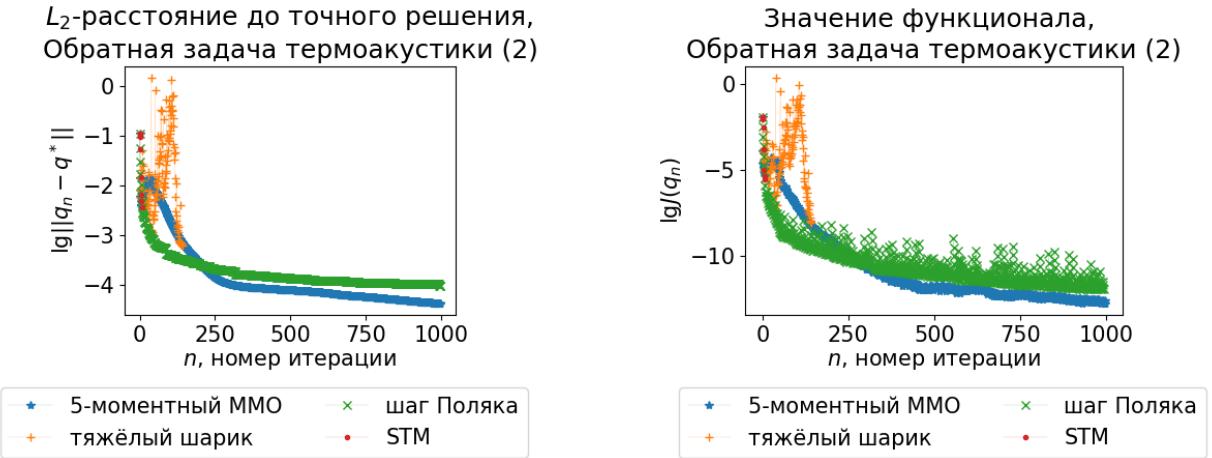


Рисунок 36 – Сходимость: 5-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (3.4.1.2).

5-моментный метод минимальных ошибок показывает значительно лучшие результаты.

Классические варианты метода сопряжённых градиентов перестают уменьшать невязку после нескольких итераций (рисунок 37). Метод с ортогональными шагами и вспомогательной минимизацией функционала продолжает работать, но его результаты хуже, чем у 5-моментного ММО.

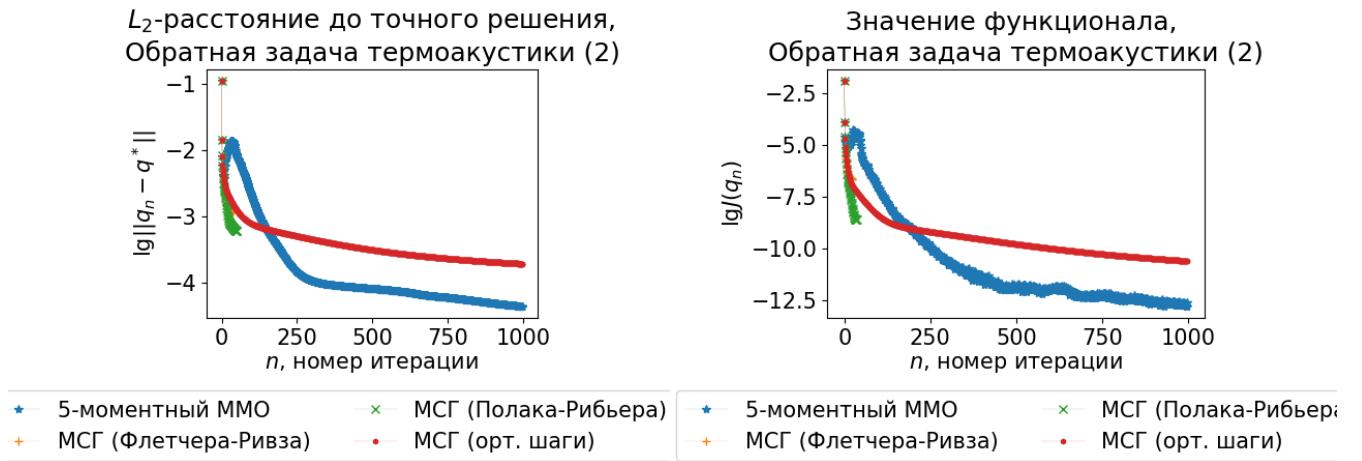


Рисунок 37 – Сходимость: 5-моментный ММО и МСГ. Задача (3.4.1.2).

Среди m -моментных методов минимальных ошибок наилучшие результаты достигнуты при $m = 5$ (рисунок 38). Метод с $m = \infty$, наилучший в теории, слишком чувствителен к погрешностям вычислений.

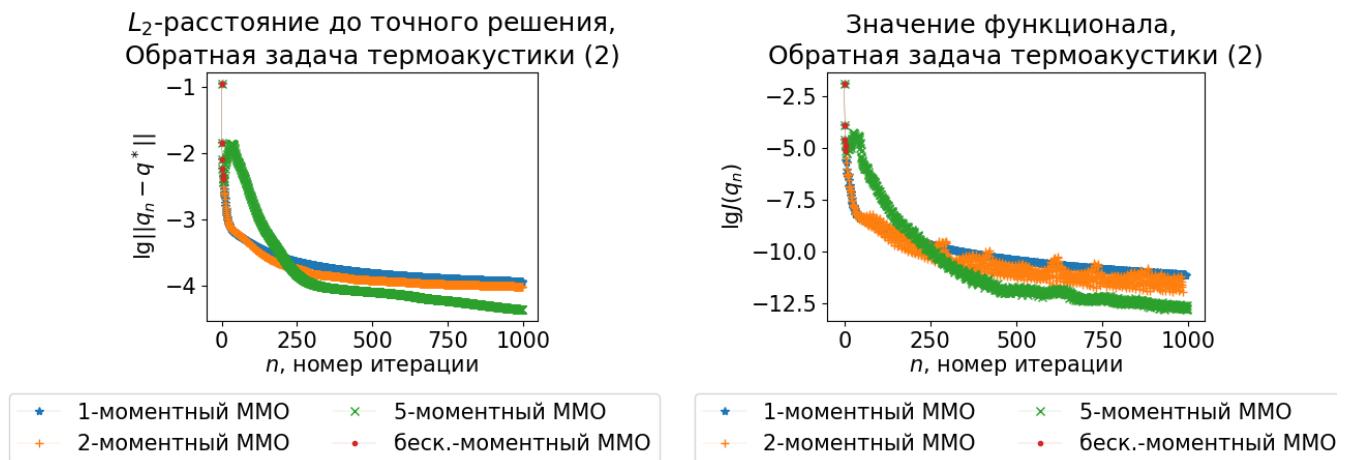


Рисунок 38 – Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (3.4.1.2).

Сравним полученные решения с известной функцией $q^*(x, y)$. Изобразим наилучшие по функционалу приближения, полученные с помощью 5-моментного ММО, МСГ и градиентного спуска с шагом Поляка, на рисунке 39.

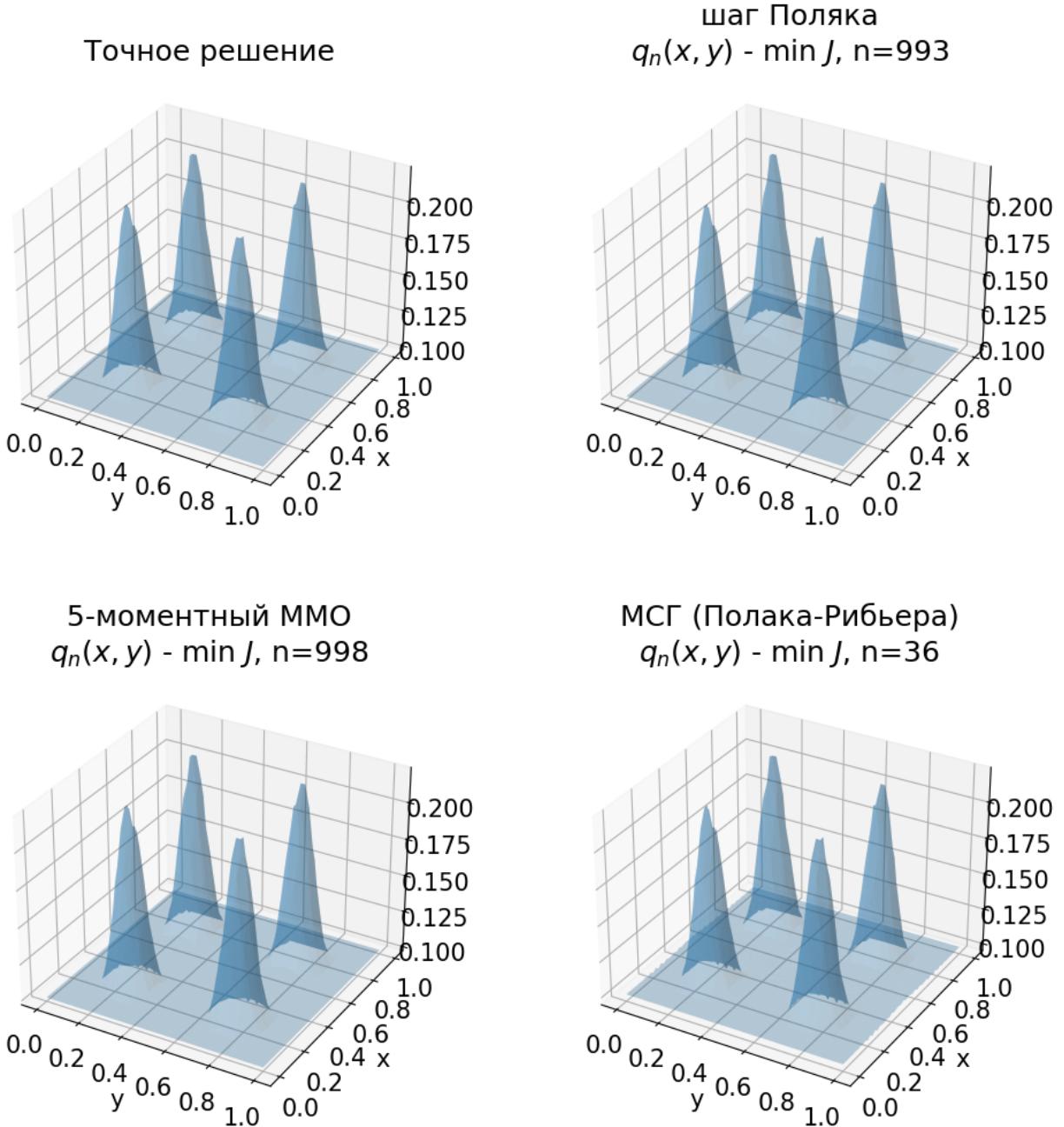


Рисунок 39 – Сравнение с точным решением: 5-моментный ММО, МСГ, шаг Поляка. Задача (3.4.1.2).

Все методы достигли достаточно хорошего качества восстановления неоднородностей. Различие в качестве нельзя заметить на графике.

Оценим теперь ухудшение качества решения при наличии погрешности измерений правой части интегрального уравнения. Для этого добавим к рассматриваемой задаче шум в правой части: $f(x) = (Aq^*)(x) \cdot (1 + \delta \cdot \mathcal{U}[-1, 1])$.

Достигнутые расстояния до точного решения для m -моментных ММО при $m \in \{1, 2, 5\}$ и МСГ в форме Флетчера-Ривза с разными δ представлены в таблице 16. Также проверена модификация 1-моментного ММО, в которой $\beta_{2k} \equiv 0$.

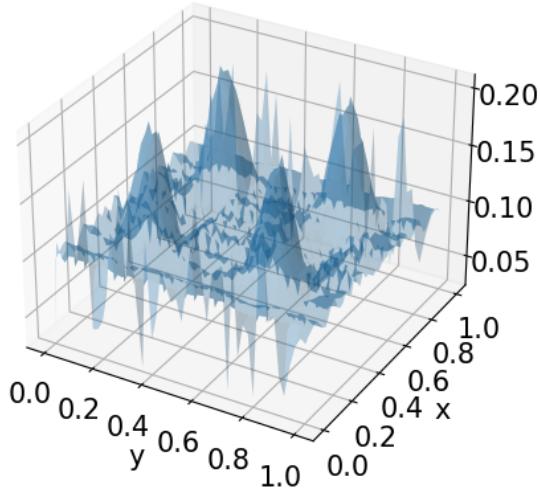
Таблица 16 – Расстояние до точного решения при наличии шума δ в правой части. Задача (3.4.1.2).

δ	2×10^{-4}	5×10^{-4}	10^{-3}	2×10^{-3}	5×10^{-3}
MMO, $m = 1$	3.62×10^{-3}	5.34×10^{-3}	7.29×10^{-3}	9.80×10^{-3}	1.54×10^{-2}
MMO, $m = 1$ (2)	3.63×10^{-3}	5.27×10^{-3}	7.32×10^{-3}	9.80×10^{-3}	1.54×10^{-2}
MMO, $m = 2$	3.64×10^{-3}	5.40×10^{-3}	7.50×10^{-3}	9.80×10^{-3}	1.54×10^{-2}
MMO, $m = 5$	4.44×10^{-3}	5.82×10^{-3}	7.50×10^{-3}	9.80×10^{-3}	1.54×10^{-2}
MСГ	3.85×10^{-3}	5.61×10^{-3}	7.27×10^{-3}	9.77×10^{-3}	1.61×10^{-2}

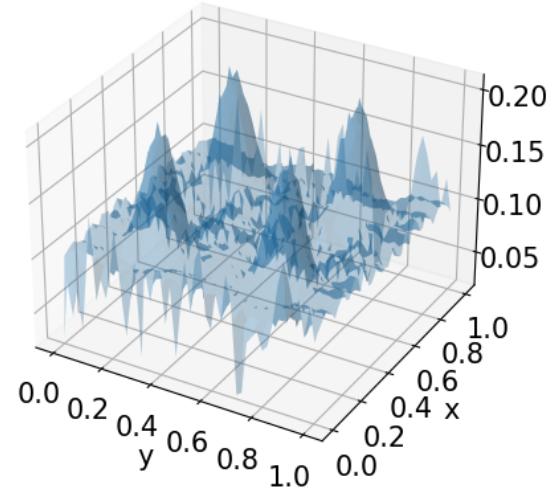
Эти результаты значительно хуже получаемых при отсутствии шума. Теряется преимущество 5-моментного ММО.

На рисунке 40 построены графики функций, порождаемых различными методами при величине шума $\delta = 2 \times 10^{-3}$.

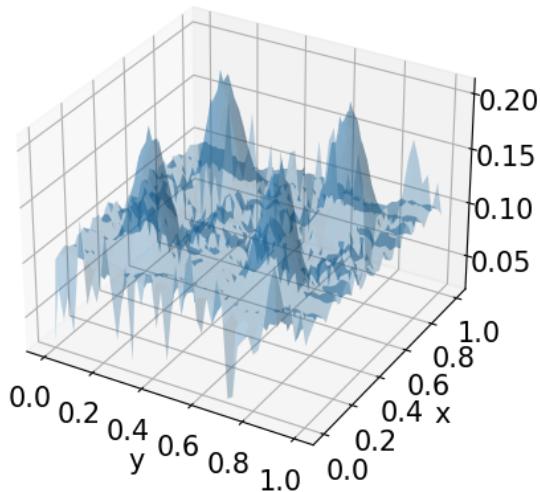
шаг Поляка
 $q_n(x, y) - \min \rho, n=5$



5-моментный ММО
 $q_n(x, y) - \min \rho, n=2$



2-моментный ММО
 $q_n(x, y) - \min \rho, n=2$



МСГ (Флетчера-Ривза)
 $q_n(x, y) - \min \rho, n=2$

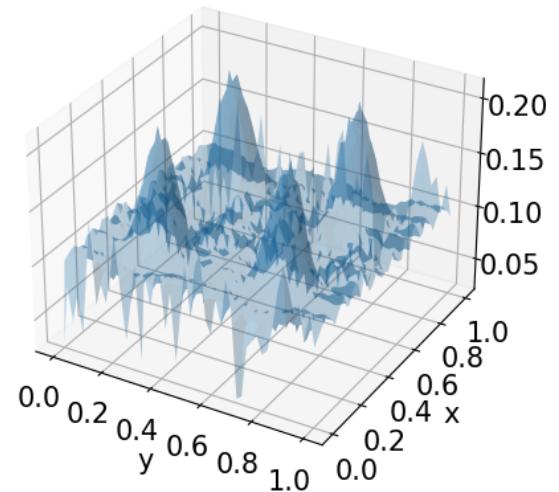
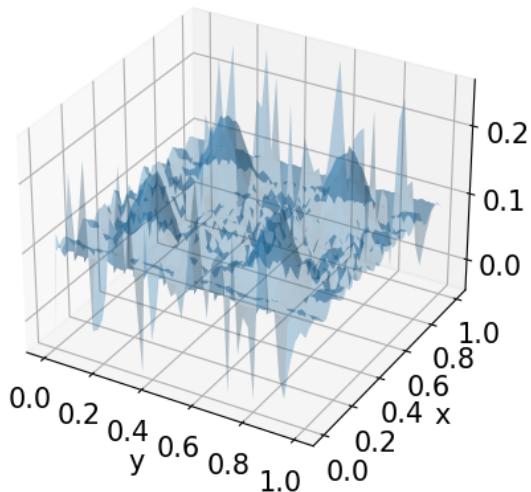


Рисунок 40 – Сравнение решений: шаг Поляка, m -моментный ММО с $m \in \{2, 5\}$, МСГ. Шум $\delta = 2 \times 10^{-3}$. Задача (3.4.1.2).

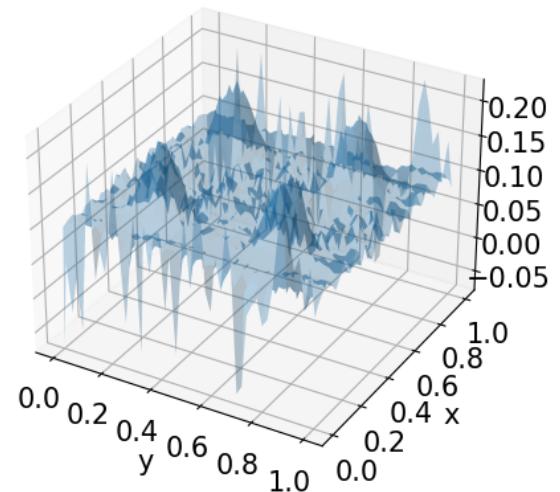
Рассмотренные методы удовлетворительно восстанавливают неоднородности. Значит, для обратной задачи 2 допустимая величина шума выше, чем для обратной задачи 1.

На рисунке 41 построены графики функций, порождаемых различными методами при величине шума $\delta = 5 \times 10^{-3}$.

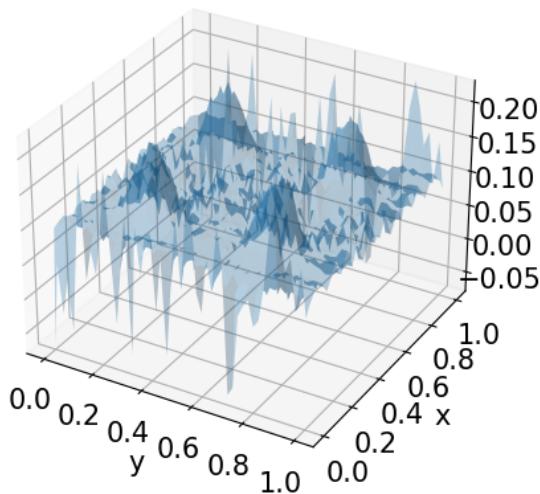
шаг Поляка
 $q_n(x, y) - \min \rho, n=4$



5-моментный ММО
 $q_n(x, y) - \min \rho, n=2$



2-моментный ММО
 $q_n(x, y) - \min \rho, n=2$



МСГ (Флетчера-Ривза)
 $q_n(x, y) - \min \rho, n=1$

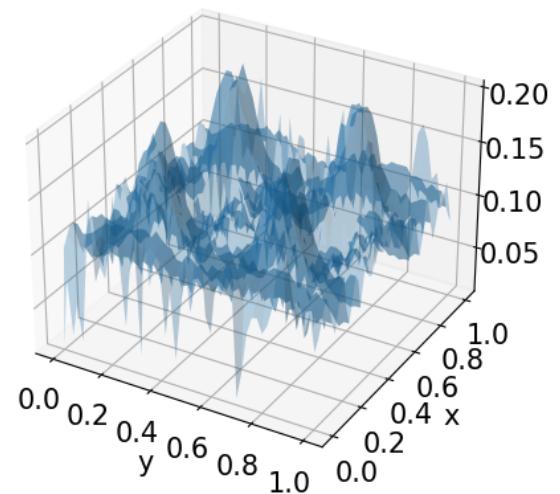


Рисунок 41 – Сравнение решений: шаг Поляка, m -моментный ММО с $m \in \{2, 5\}$, МСГ. Шум $\delta = 5 \times 10^{-3}$. Задача (3.4.1.2).

Настоящие неоднородности почти не выделяются среди множества максимумов и минимумов.

Эксперименты с обратной задачей 3

Результаты представлены в таблице 17.

Таблица 17 – Результаты работы методов. Задача (3.4.1.3).

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	8.50×10^{-5}	4.90×10^{-4}	6.93×10^{-12}
ММО, $m = 2$	5.68×10^{-5}	2.58×10^{-4}	9.47×10^{-13}
ММО, $m = 5$	3.77×10^{-5}	1.89×10^{-4}	2.29×10^{-13}
ММО, $m = \infty$	3.67×10^{-3}	1.69×10^{-2}	8.65×10^{-6}
Сопряжённые градиенты (ФР)	2.90×10^{-3}	1.37×10^{-2}	3.90×10^{-6}
Сопряжённые градиенты (ПР)	8.56×10^{-5}	4.97×10^{-4}	1.47×10^{-12}
Сопряжённые градиенты (\perp)	1.91×10^{-4}	1.24×10^{-3}	3.63×10^{-11}
Тяжёлый шарик	3.69×10^{-3}	1.70×10^{-2}	1.18×10^{-5}
Шаг Поляка	7.20×10^{-5}	3.55×10^{-4}	2.77×10^{-13}
STM	7.30×10^{-3}	2.48×10^{-2}	3.91×10^{-5}

Использованы сетка с шагами $h = 0.02$, $\tau = 0.0001$ и нулевое начальное приближение. Начальное значение функционала $J(q_0) = 0.018$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.11$.

Рисунок 42 показывает, что только 5-моментный метод минимальных ошибок и градиентный спуск с шагом Поляка демонстрируют монотонную сходимость по аргументу.

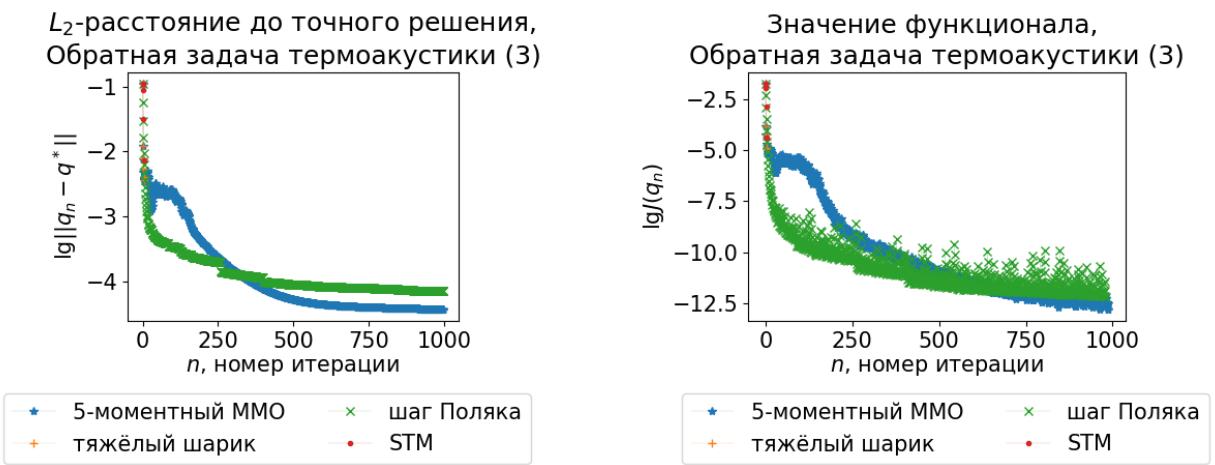


Рисунок 42 – Сходимость: 5-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (3.4.1.3).

5-моментный метод минимальных ошибок показывает значительно лучшие результаты.

Метод сопряжённых градиентов в форме Флетчера-Ривза перестаёт уменьшать невязку после нескольких итераций. Метод сопряжённых градиентов в форме Полака-Рибьера, как и метод с ортогональными шагами и вспомогательной

минимизацией функционала, продолжает работать, но его результаты хуже, чем у 5-моментного ММО. Это показывает рисунок 43.

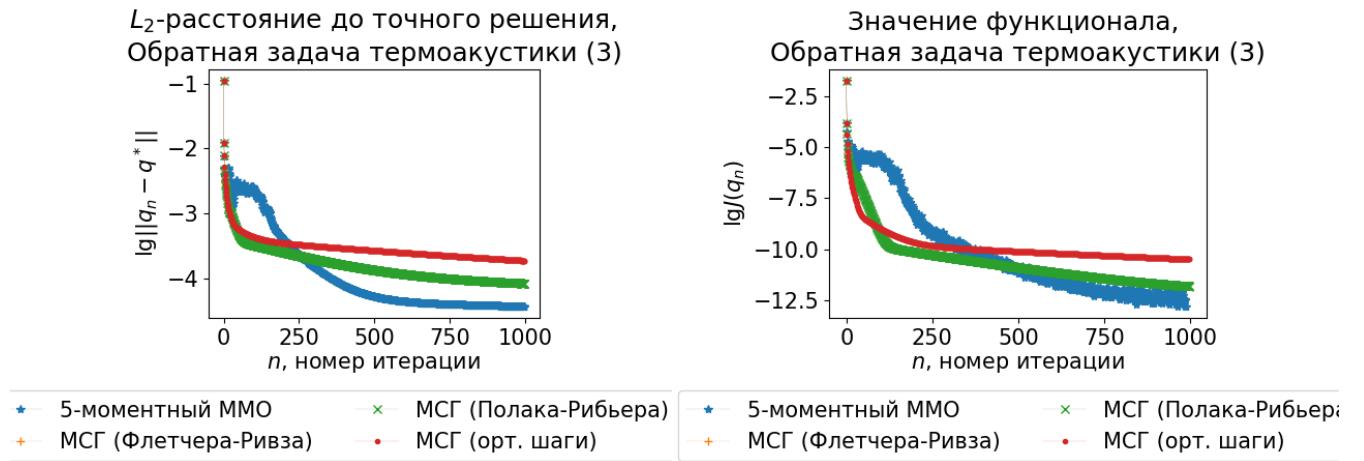


Рисунок 43 – Сходимость: 5-моментный ММО и МСГ. Задача (3.4.1.3).

Среди m -моментных методов минимальных ошибок наилучшие результаты достигнуты при $m = 5$ (рисунок 44). Метод с $m = \infty$, наилучший в теории, слишком чувствителен к погрешностям вычислений.

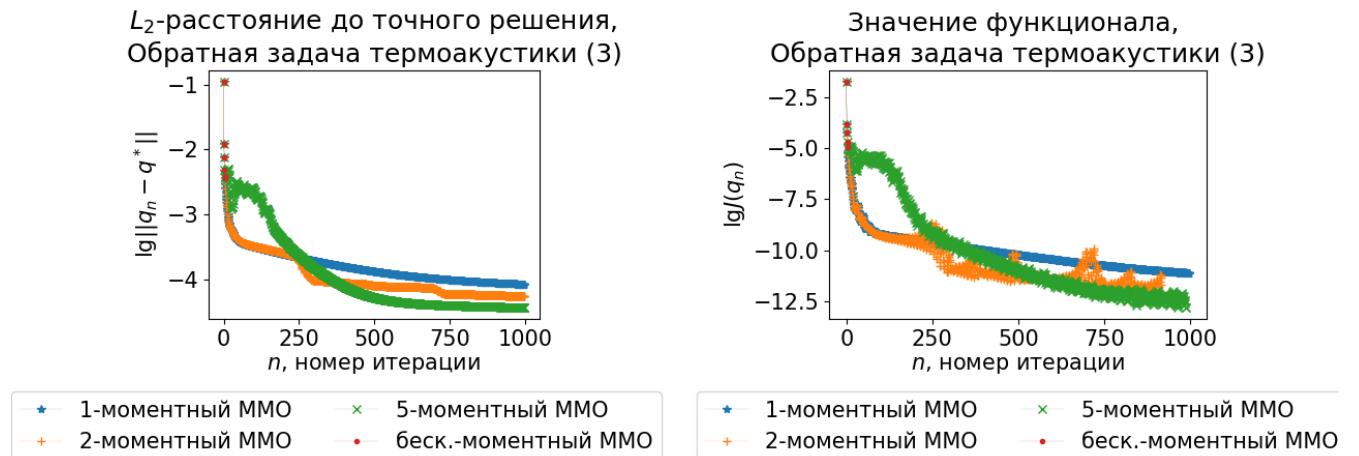


Рисунок 44 – Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (3.4.1.3).

3.4.5. Выводы

Для обратной задачи 1 ни один из методов не позволил уменьшить расстояние до точного решения в 15 или более раз. Наилучшее качество решения — у m -моментных ММО и градиентного спуска с шагом Поляка. Максимальное

отклонение полученных с их помощью решений от точного составило примерно 30% разброса между минимальным и максимальным значениями $q^*(x, y)$. Форма графика восстановлена, но не полностью: величины экстремумов отличаются от настоящих.

Увеличение номера обратной задачи (то есть, количества дополнительной информации) приводит к улучшению качества решения. Это вполне естественно. 5-моментный ММО во всех случаях показал наилучшие результаты. Это подтверждает выводы раздела 2.3.4 о возможности применения построенных методов к составному функционалу — сумме квадратичных функционалов, имеющих общую точку минимума.

Для обратных задач 2 и 3 максимальное отклонение полученных с помощью 5-моментного ММО решений от точного на три порядка меньше разброса между минимальным и максимальным значениями $q^*(x, y)$. Это свидетельствует о хорошем восстановлении формы графика, что подтверждается его построением.

Теоретические ожидания лучшей сходимости при $m = \infty$ не оправдались. Это объясняется чувствительностью такого метода к погрешностям и их накоплением.

Эксперименты показывают возможность решения обратной задачи термоакустики с помощью градиентных методов оптимизации, в особенности — нового m -моментного метода минимальных ошибок.

Обнаружено также, что построенные алгоритмы численного решения обратной задачи термоакустики устойчивы к сравнительно небольшому шуму. Обратная задача 1 решается с помощью 5-моментного ММО при величине относительного шума $\delta \leq 5 \times 10^{-4}$. Увеличение количества дополнительной информации (то есть, переход к обратной задаче 2) повышает приемлемый уровень шума примерно до $\delta = 2 \times 10^{-3}$.

3.5. Интегральное уравнение Фредгольма 1-го рода

3.5.1. Происхождение, постановка и некорректность задачи

Постановка задачи (1.11) взята из книги [53]. Подобные интегральные уравнения возникают, например, при обращении «усечённого» преобразования Фурье, которое использует интегрирование по отрезку вместо всей числовой прямой.

Некорректность обоснована в разделе 4.1 [53]. Например, уравнение (1.11) с непрерывно дифференцируемым ядром $K(x, s)$ и недифференцируемой правой частью $f(x)$ не может иметь решений среди непрерывных функций из-за теоремы о дифференировании интеграла по параметру. С другой стороны, при вырожденном ядре нарушается единственность решения.

3.5.2. Операторы задачи

Пусть $K(x, s)$ — непрерывная на $[0, 1] \times [0, 1]$ функция двух переменных. Определим на пространстве непрерывных функций вещественной переменной со стандартным скалярным произведением

$$H = C[0, 1] \subset L^2[0, 1]$$

оператор $A : H \rightarrow H$:

$$(Aq)(x) = \int_0^1 K(x, s)q(s)ds \quad \forall x \in [0, 1]. \quad (3.5.2.1)$$

Правая часть — непрерывная функция параметра, поэтому определение корректно. Тогда

$$J(q) = \frac{1}{2} \int_0^1 \left(\int_0^1 K(x, s)q(s)ds - f(x) \right)^2 dx. \quad (3.5.2.2)$$

Для вычисления градиента функционала (и, соответственно, сопряжённого оператора) рассмотрим его приращение при изменении q на δq :

$$\begin{aligned} J(q + \delta q) - J(q) &\simeq \int_0^1 \left(\int_0^1 K(x, s)q(s)ds - f(x) \right) \left(\int_0^1 K(x, s)\delta q(s)ds \right) dx = \\ &= \int_0^1 (Aq - f)(x) \left(\int_0^1 K(x, s)\delta q(s)ds \right) dx = \int_0^1 \int_0^1 (Aq - f)(x)K(x, s)\delta q(s)dsdx. \end{aligned}$$

Подынтегральная функция непрерывна, поэтому можно поменять порядок интегрирования:

$$\int_0^1 \int_0^1 (Aq - f)(x)K(x, s)\delta q(s)dxds = \left\langle \int_0^1 (Aq - f)(x)K(x, s)dx, \delta q(s) \right\rangle.$$

Соответственно,

$$(\nabla J(q))(x) = \int_0^1 (Aq - f)(x)K(x, s)dx = (A^*(Aq - f))(x), \quad (3.5.2.3)$$

где

$$(A^*p)(x) = \int_0^1 K(s, x)p(s)ds. \quad (3.5.2.4)$$

Очевидно, что $A = A^*$ тогда и только тогда, когда ядро симметричное, то есть при $0 \leq x, s \leq 1$ $K(x, s) = K(s, x)$.

3.5.3. Эксперименты

Для проверки работы методов выбрана модельная задача восстановления тождественной функции $q^*(x) = x$ по результату действия на неё интегрального оператора (3.5.2.1) с различными ядрами. Соответственно, в каждом примере $f(x) = (Aq^*)(x)$.

Ядро — сумма синусов

$$K_1(x, s) = 1 + \sum_{k=1}^{10} \frac{1}{2^k} \sin \pi kx \sin \pi ks$$

Результаты представлены в таблице 18.

Таблица 18 – Результаты работы методов. Задача (1.11), ядро K_1 .

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
MMO, $m = 1$	9.59×10^{-2}	0.500	3.51×10^{-32}
MMO, $m = 2$	9.59×10^{-2}	0.500	1.32×10^{-32}
MMO, $m = 5$	9.59×10^{-2}	0.500	5.13×10^{-33}
MMO, $m = \infty$	9.59×10^{-2}	0.500	1.73×10^{-32}
Сопряжённые градиенты (ФР)	9.59×10^{-2}	0.500	1.58×10^{-31}
Сопряжённые градиенты (ПР)	9.59×10^{-2}	0.500	3.60×10^{-27}
Сопряжённые градиенты (\perp)	0.115	0.509	4.16×10^{-9}
Тяжёлый шарик	0.125	0.500	1.05×10^{-7}
Шаг Поляка	0.106	0.500	7.73×10^{-10}
STM	0.142	0.506	8.68×10^{-7}

Задача решается с интегрированием по сетке с шагом $h = 0.001$. Начальное значение функционала $J(q_0) = 0.184$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.577$.

Рисунок 45 показывает, что 1-моментный метод минимальных ошибок позволяет достичь меньшей невязки, чем адаптивный метод тяжёлого шарика из статьи [51], градиентный спуск с шагом Поляка и метод подобных треугольников, и по аргументу, и по функционалу. При этом сходимость по функционалу не является монотонной.

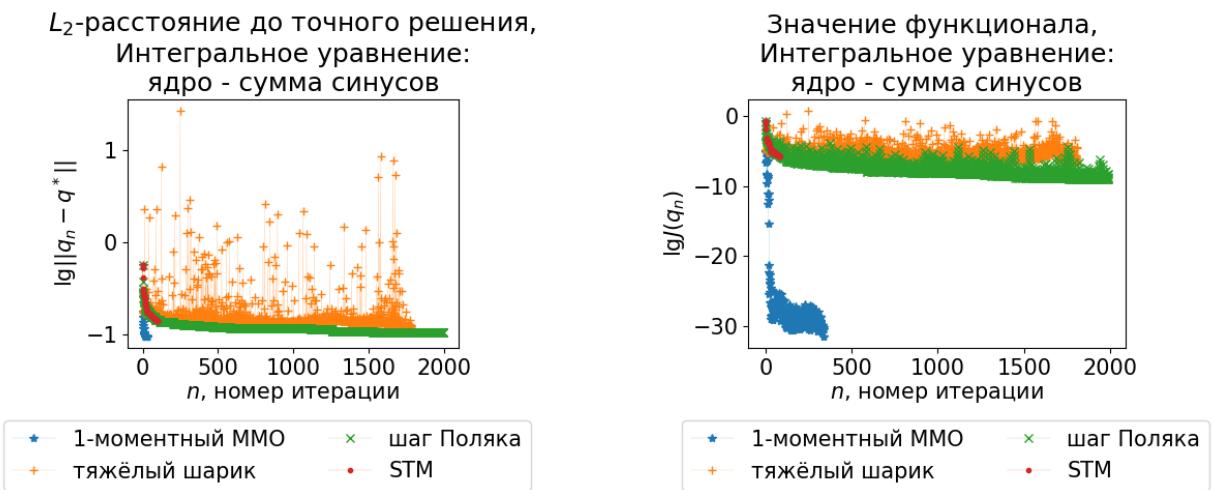


Рисунок 45 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.11), ядро K_1 .

Начиная с определённого момента, расстояние до точного решения перестаёт убывать. Это связано с вырожденностью ядра: достигается проекция точного решения на линейную оболочку базисных функций, составляющих ядро.

Рисунок 46 показывает, что классические варианты метода сопряжённых градиентов и новый 1-моментный метод минимальных ошибок достигают минимально возможной невязки по аргументу.

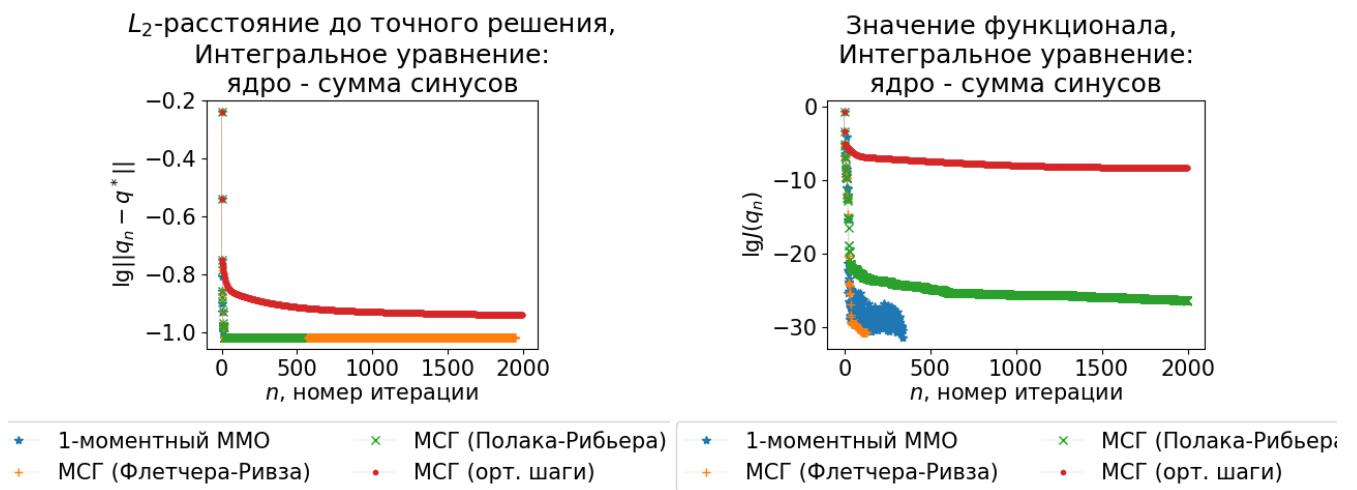


Рисунок 46 – Сходимость: 1-моментный ММО и МСГ. Задача (1.11), ядро K_1 .

Рисунок 47 показывает, что при всех m m -моментный ММО достигает минимально возможной невязки по аргументу. Невязка по функционалу — в пределах погрешности вычислений.

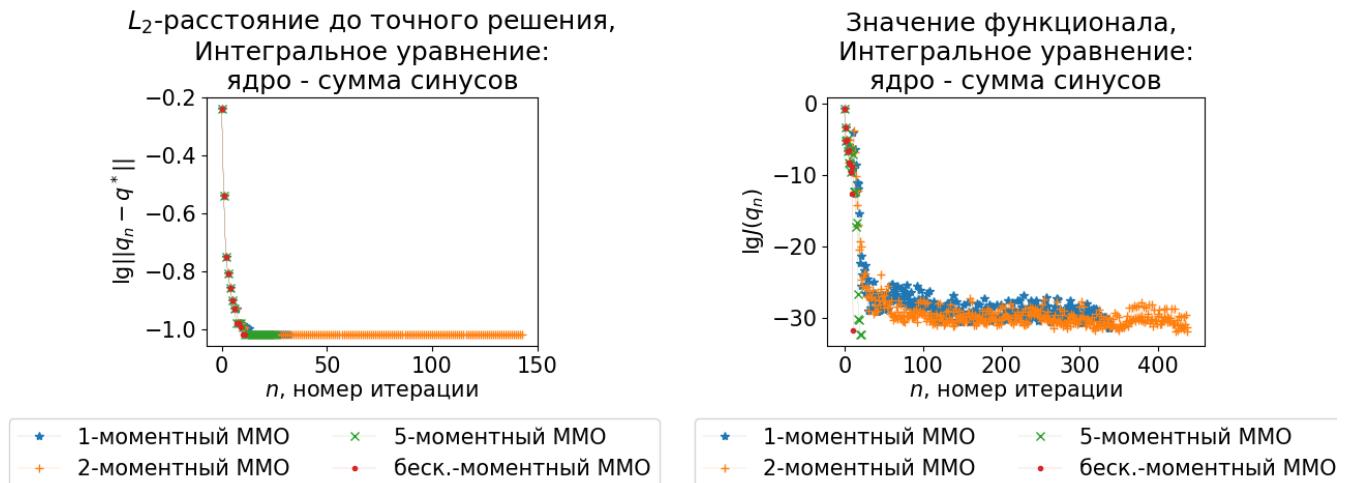


Рисунок 47 – Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.11), ядро K_1 .

Сравним полученные решения с известной функцией $q^*(x)$. Изобразим наилучшие по функционалу и по расстоянию приближения, полученные с помощью 5-моментного ММО и с помощью метода сопряжённых градиентов, на рисунке 48.

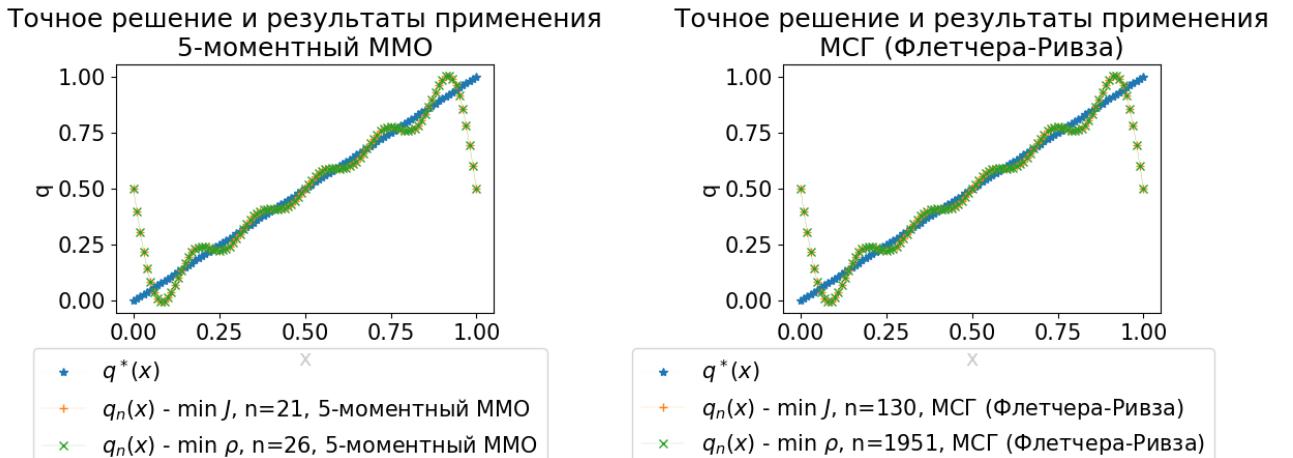


Рисунок 48 – Сравнение с точным решением: 5-моментный ММО, МСГ. Задача (1.11), ядро K_1 .

Функция восстановлена плохо, что объясняется невозможностью её разложения по системе собственных функций ядра. Методы показали близкие результаты, но для 5-моментного ММО количество вызовов оракула на порядок меньше.

Оценим теперь ухудшение качества решения при наличии погрешности измерений правой части интегрального уравнения. Для этого добавим к рассматриваемой задаче шум в правой части: $f(x) = (Aq^*)(x) \cdot (1 + \delta \cdot \mathcal{U}[-1, 1])$.

Достигнутые расстояния до точного решения для m -моментных ММО при $m \in \{1, 2, 5\}$ и МСГ в форме Флетчера-Ривза с разными δ представлены в таблице 19. Также проверена модификация 1-моментного ММО, в которой $\beta_{2k} \equiv 0$.

Таблица 19 – Расстояние до точного решения при наличии шума δ в правой части. Задача (1.11), ядро K_1 .

$\delta, \%$	0.5	1	2	5
ММО, $m = 1$	0.143	0.175	0.179	0.198
ММО, $m = 1$ (2)	0.118	0.121	0.131	0.164
ММО, $m = 2$	0.143	0.175	0.179	0.198
ММО, $m = 5$	0.143	0.175	0.179	0.198
МСГ	0.108	0.117	0.125	0.139

Видно, что для m -моментного ММО качество решения ухудшается сильнее, чем для МСГ. Рестарты позволяют уменьшить невязку, поскольку препятствуют накоплению погрешности. Изобразим наилучшие по функционалу и по расстоянию приближения, полученные при $\delta = 1\%$ с помощью 1-моментного ММО с рестартами и с помощью метода сопряжённых градиентов, на рисунке 49.

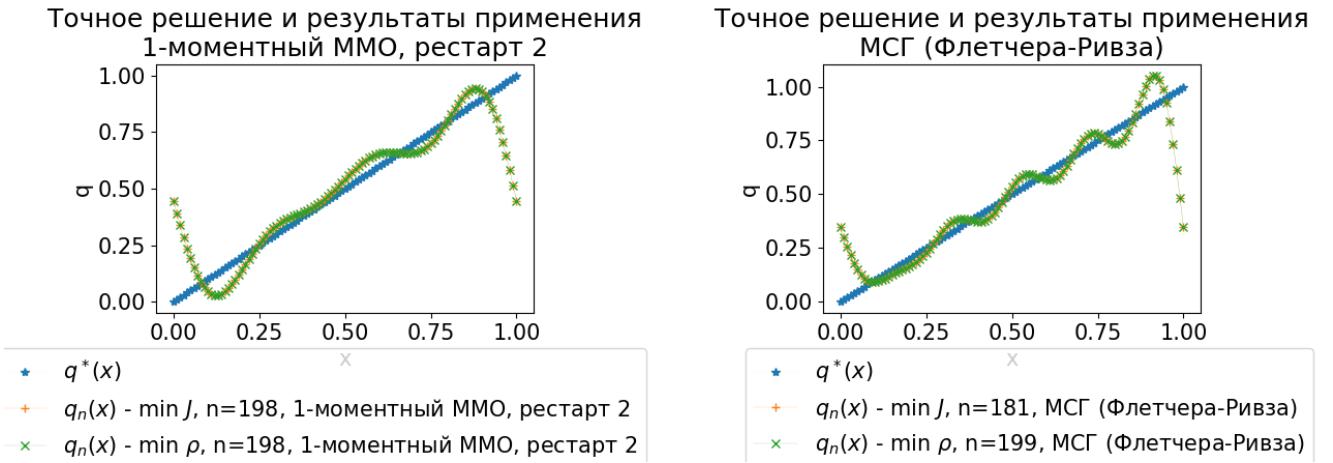


Рисунок 49 – Сравнение с точным решением: 1-моментный ММО с рестартами, МСГ. Шум $\delta = 1\%$. Задача (1.11), ядро K_1 .

Качественно графики похожи на те, которые были получены при решении незашумлённой задачи с таким же ядром.

Ядро — сумма синусов и косинусов

$$K_2(x, s) = 1 + \sum_{k=1}^{10} \frac{1}{2^k} (\cos \pi kx \cos \pi ks + \sin \pi kx \sin \pi ks) = 1 + \sum_{k=1}^{10} \frac{1}{2^k} \cos \pi k(x - s)$$

Результаты представлены в таблице 20.

Таблица 20 – Результаты работы методов. Задача (1.11), ядро K_2 .

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
MMO, $m = 1$	4.86×10^{-6}	3.00×10^{-5}	3.06×10^{-29}
MMO, $m = 2$	4.86×10^{-6}	3.12×10^{-5}	8.33×10^{-31}
MMO, $m = 5$	4.60×10^{-6}	2.85×10^{-5}	8.78×10^{-31}
MMO, $m = \infty$	3.54×10^{-6}	2.31×10^{-5}	3.34×10^{-31}
Сопряжённые градиенты (ФР)	9.57×10^{-5}	5.90×10^{-4}	2.94×10^{-23}
Сопряжённые градиенты (ПР)	1.27×10^{-4}	1.03×10^{-3}	6.13×10^{-21}
Сопряжённые градиенты (\perp)	0.013	0.072	3.57×10^{-9}
Тяжёлый шарик	9.87×10^{-3}	3.15×10^{-2}	9.46×10^{-9}
Шаг Поляка	6.12×10^{-3}	3.74×10^{-2}	5.37×10^{-11}
STM	0.022	0.086	1.76×10^{-7}

Задача решается с интегрированием по сетке с шагом $h = 0.001$. Начальное значение функционала $J(q_0) = 0.188$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.577$.

Рисунок 50 показывает, что 1-моментный метод минимальных ошибок позволяет достичь меньшей невязки, чем адаптивный метод тяжёлого шарика из статьи [51], градиентный спуск с шагом Поляка и метод подобных треугольников, и по аргументу, и по функционалу. При этом сходимость по функционалу не является монотонной.

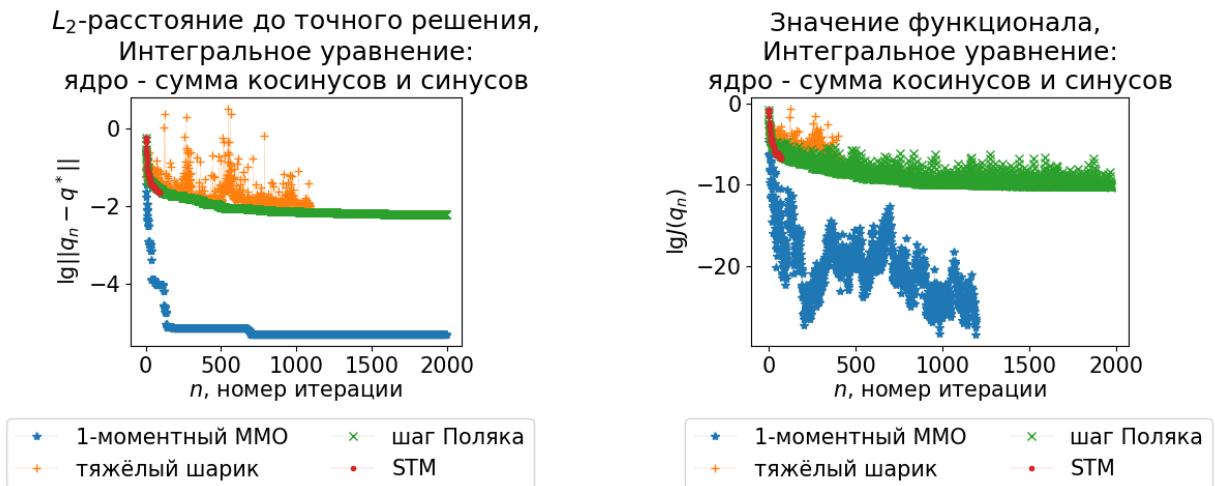


Рисунок 50 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.11), ядро K_2 .

Начиная с определённого момента, расстояние до точного решения перестаёт убывать, не доходя до нуля. При этом функционал очень близок к нулю — в пределах погрешности вычислений. Это связано с вырожденностью ядра: достигается проекция точного решения на линейную оболочку базисных функций, составляющих ядро. При этом невязка оказывается меньше, чем для ядра — суммы синусов. Это естественно, поскольку разложение $q(x) = x$ в ряд Фурье содержит косинусы с ненулевыми коэффициентами.

Рисунок 51 показывает, что новый 1-моментный метод минимальных ошибок достигает меньшей невязки по аргументу, чем метод сопряжённых градиентов.

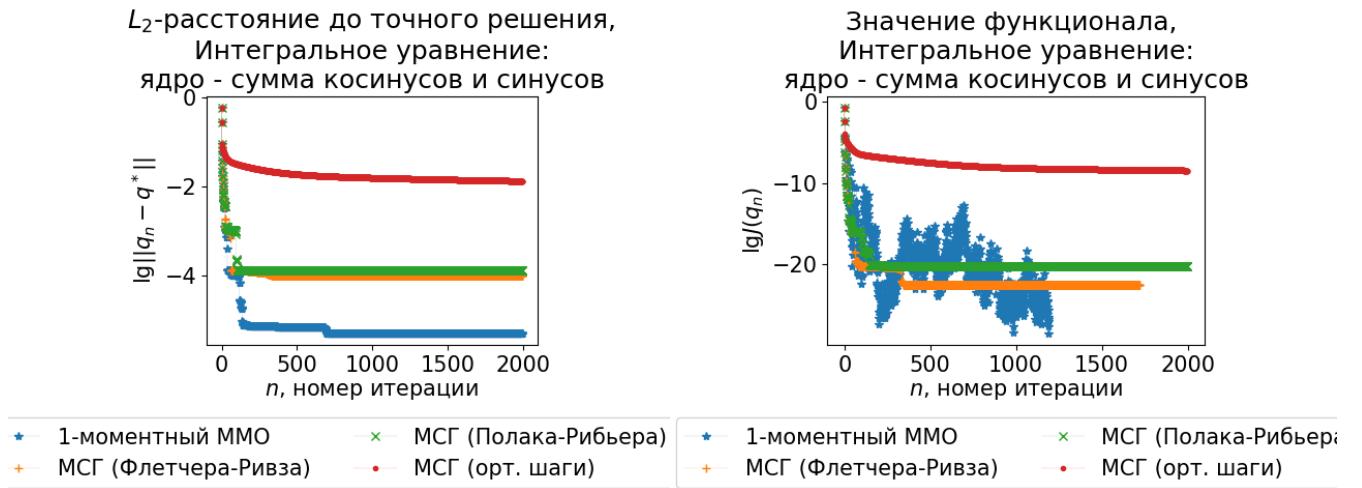


Рисунок 51 – Сходимость: 1-моментный ММО и МСГ. Задача (1.11), ядро K_2 .

Рисунок 52 показывает, что чем больше m , тем меньшей невязки достигает m -моментный метод минимальных ошибок и меньше итераций на это требуется. При $m = \infty$ требуется лишь 18 итераций.

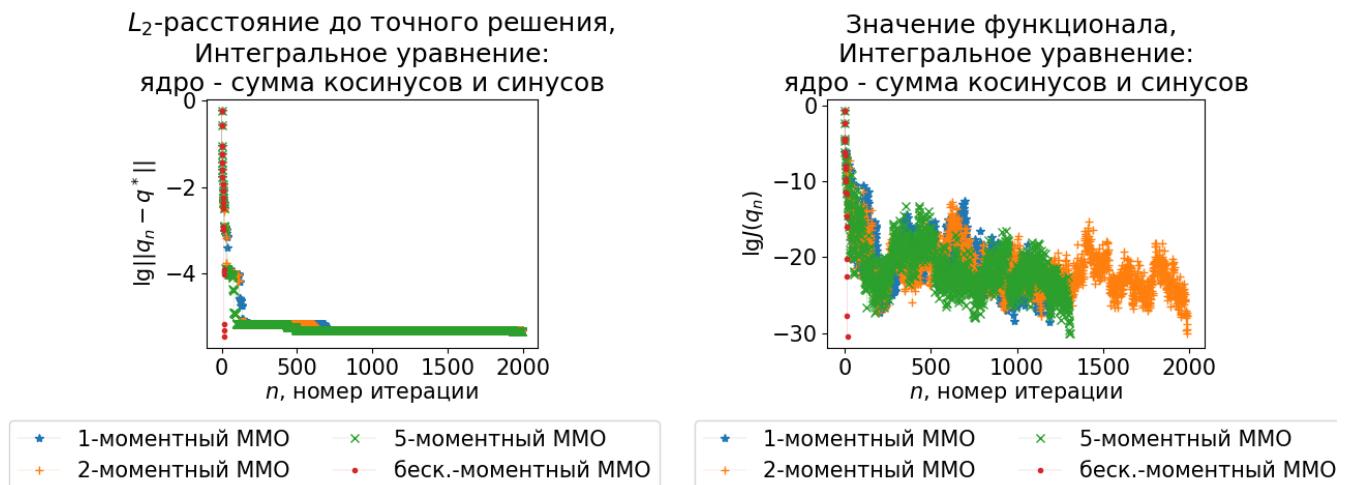


Рисунок 52 – Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.11), ядро K_2 .

Сравним полученные решения с известной функцией $q^*(x)$. Изобразим наилучшие по функционалу и по расстоянию приближения, полученные с помощью 5-моментного ММО и с помощью метода сопряжённых градиентов, на рисунке 53.

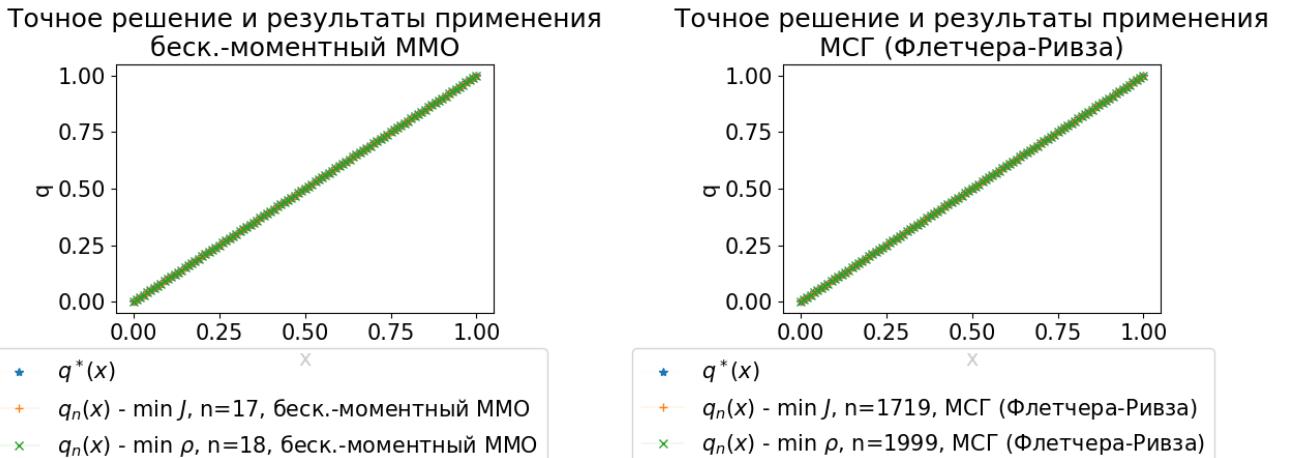


Рисунок 53 – Сравнение с точным решением: 5-моментный ММО, МСГ. Задача (1.11), ядро K_2 .

Функция восстановлена. Результаты очень близки, но ∞ -моментный ММО позволил достичь их в 100 раз быстрее.

Оценим теперь ухудшение качества решения при наличии погрешности измерений правой части интегрального уравнения. Для этого добавим к рассматриваемой задаче шум в правой части: $f(x) = (Aq^*)(x) \cdot (1 + \delta \cdot \mathcal{U}[-1, 1])$.

Достигнутые расстояния до точного решения для m -моментных ММО при $m \in \{1, 2, 5\}$ и МСГ в форме Флетчера-Ривза с разными δ представлены в таблице 21. Также проверена модификация 1-моментного ММО, в которой $\beta_{2k} \equiv 0$.

Таблица 21 – Расстояние до точного решения при наличии шума δ в правой части. Задача (1.11), ядро K_2 .

$\delta, \%$	0.5	1	2	5
ММО, $m = 1$	3.72×10^{-2}	4.23×10^{-2}	6.91×10^{-2}	8.89×10^{-2}
ММО, $m = 1$ (2)	1.75×10^{-2}	2.02×10^{-2}	2.71×10^{-2}	4.33×10^{-2}
ММО, $m = 2$	3.72×10^{-2}	4.23×10^{-2}	6.91×10^{-2}	8.89×10^{-2}
ММО, $m = 5$	3.72×10^{-2}	4.23×10^{-2}	6.91×10^{-2}	8.89×10^{-2}
МСГ	1.66×10^{-2}	1.84×10^{-2}	2.06×10^{-2}	2.68×10^{-2}

Видно, что для m -моментного ММО качество решения ухудшается сильнее, чем для МСГ. Рестарты позволяют уменьшить невязку, поскольку препятствуют накоплению погрешности. Изобразим наилучшие по функционалу и по расстоянию приближения, полученные при $\delta = 1\%$ с помощью 1-моментного ММО с рестартами и с помощью метода сопряжённых градиентов, на рисунке 54.

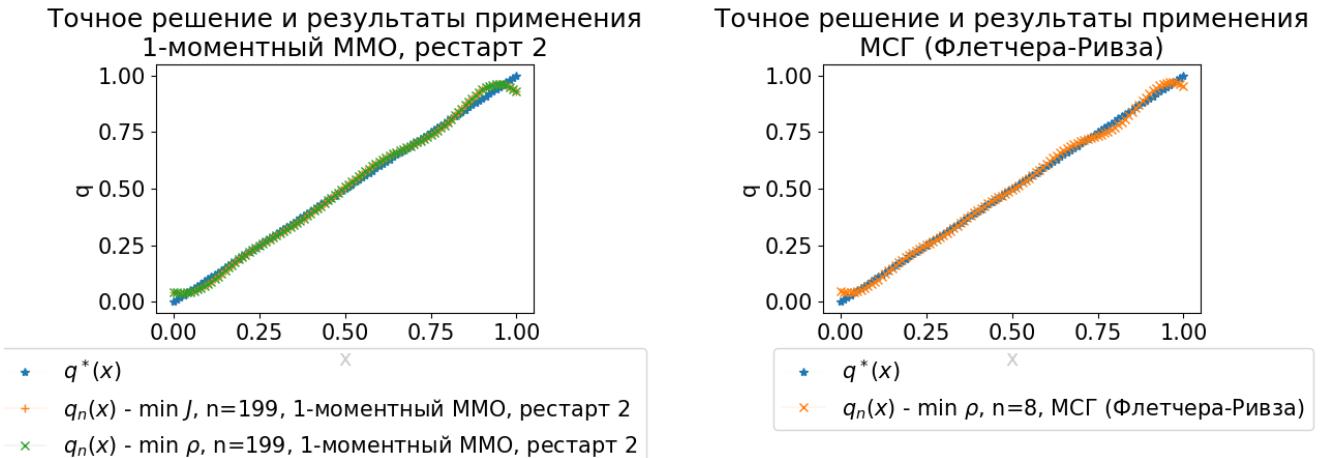


Рисунок 54 – Сравнение с точным решением: 1-моментный ММО с рестартами, МСГ. Шум $\delta = 1\%$. Задача (1.11), ядро K_2 .

Решение восстанавливается, но хуже, чем при точно заданной правой части. Функция $q_{194}(x)$, на которой функционал достиг минимального значения при работе метода сопряжённых градиентов, не изображена, поскольку принимает очень большие по модулю значения. Этого можно было бы избежать при использовании регуляризации.

Для методов без рестартов наблюдается следующий известный эффект: невязка по функционалу и аргументу сначала убывает, потом возрастает. 1-моментный ММО с рестартами достиг точки с минимальной невязкой на последней итерации, что может свидетельствовать о возможности получения решения с лучшим качеством при увеличении числа итераций.

Ядро Фурье

$$K_3(x, s) = \sin \pi x s$$

Результаты представлены в таблице 22.

Таблица 22 – Результаты работы методов. Задача (1.11), ядро K_3 .

	$\ q_n - q^*\ $	$\ q_n - q^*\ _C$	$J(q_n)$
ММО, $m = 1$	1.13×10^{-8}	2.36×10^{-8}	1.91×10^{-33}
ММО, $m = 2$	1.07×10^{-8}	4.49×10^{-8}	1.43×10^{-33}
ММО, $m = 5$	1.13×10^{-8}	5.42×10^{-8}	1.72×10^{-33}
ММО, $m = \infty$	1.13×10^{-8}	5.42×10^{-8}	1.72×10^{-33}
Сопряжённые градиенты (ФР)	1.66×10^{-6}	7.27×10^{-6}	3.14×10^{-25}
Сопряжённые градиенты (ПР)	1.66×10^{-6}	7.27×10^{-6}	3.14×10^{-25}
Сопряжённые градиенты (\perp)	1.69×10^{-3}	6.16×10^{-3}	1.94×10^{-11}
Тяжёлый шарик	1.52×10^{-4}	2.79×10^{-4}	3.99×10^{-17}
Шаг Поляка	1.52×10^{-4}	6.07×10^{-4}	1.24×10^{-16}
STM	1.52×10^{-4}	2.71×10^{-4}	3.83×10^{-17}

Задача решается с интегрированием по сетке с шагом $h = 0.001$. Начальное значение функционала $J(q_0) = 5.83 \times 10^{-2}$, начальное расстояние до точного решения $\|q_0 - q^*\| = 0.577$.

Рисунок 55 показывает, что 1-моментный метод минимальных ошибок всего за 16 итераций позволяет достичь меньшей невязки, чем адаптивный метод тяжёлого шарика из статьи [51], градиентный спуск с шагом Поляка и метод подобных треугольников. При этом сходимость по функционалу не является монотонной.

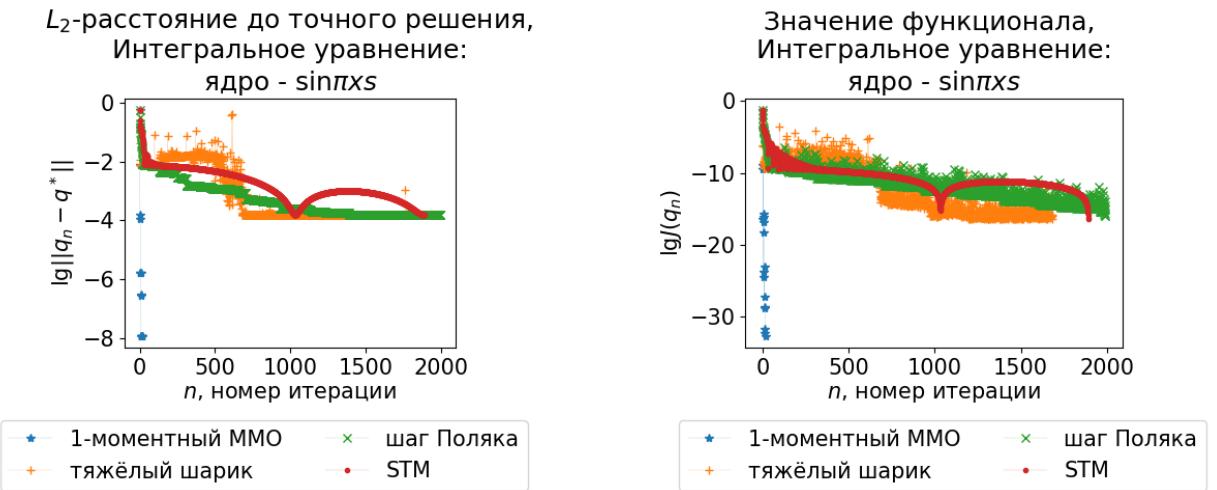


Рисунок 55 – Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.11), ядро K_3 .

Рисунок 56 показывает, что новый 1-моментный метод минимальных ошибок достигает меньшей невязки, чем метод сопряжённых градиентов. Классические варианты метода сопряжённых градиентов перестают уменьшать невязку после небольшого (14 и 65) числа итераций.

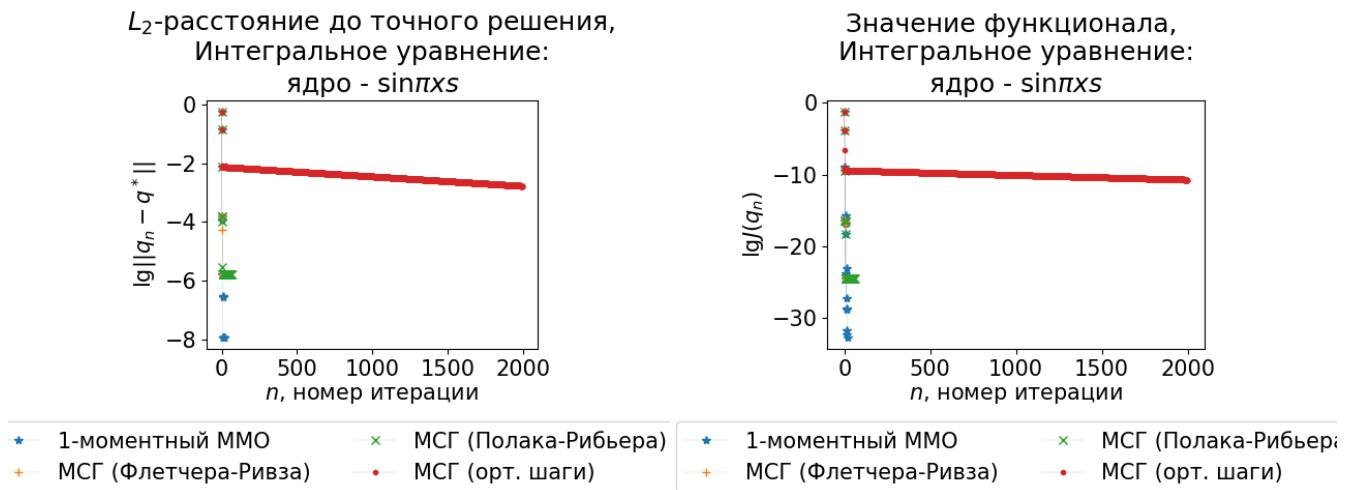


Рисунок 56 – Сходимость: 1-моментный ММО и МСГ. Задача (1.11), ядро K_3 .

Рисунок 57 показывает, что при $m \in \{5, \infty\}$ точки, генерируемые m -моментным методом минимальных ошибок, перестают приближаться к точному решению, начиная с 5 итерации. При $m = 2$ метод достигает несколько меньшей невязки.

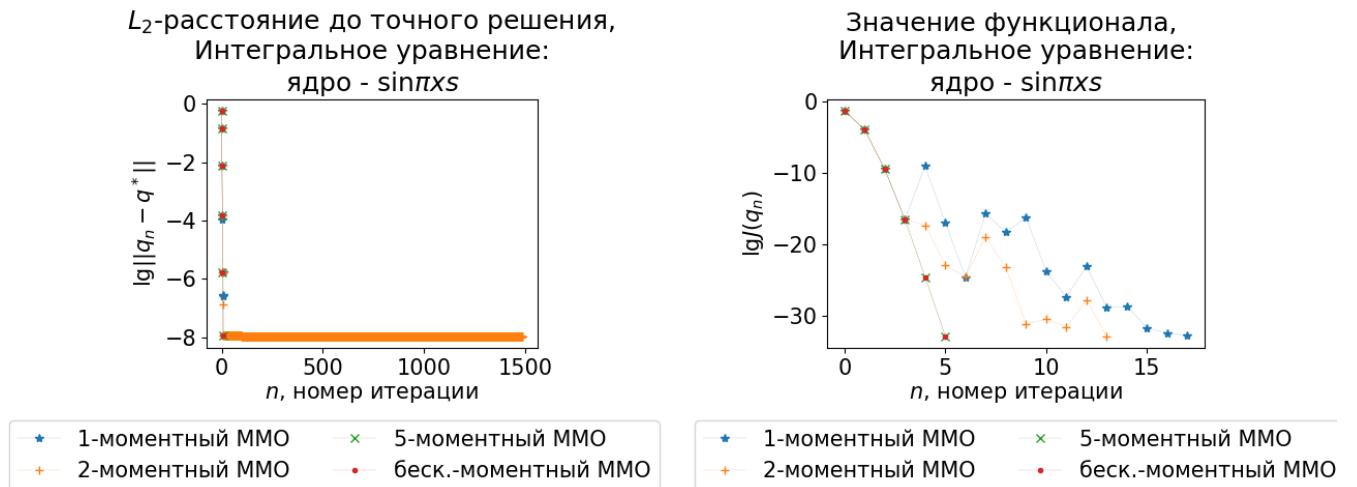


Рисунок 57 – Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.11), ядро K_3 .

Оценим теперь ухудшение качества решения при наличии погрешности измерений правой части интегрального уравнения. Для этого добавим к рассматриваемой задаче шум в правой части: $f(x) = (Aq^*)(x) \cdot (1 + \delta \cdot \mathcal{U}[-1, 1])$.

Достигнутые расстояния до точного решения для m -моментных ММО при $m \in \{1, 2, 5\}$ и МСГ в форме Флетчера-Ривза с разными δ представлены в таблице 23. Также проверена модификация 1-моментного ММО, в которой $\beta_{2k} \equiv 0$.

Таблица 23 – Расстояние до точного решения при наличии шума δ в правой части. Задача (1.11), ядро K_3 .

$\delta, \%$	0.5	1	2	5
ММО, $m = 1$	7.67×10^{-3}	8.24×10^{-3}	1.31×10^{-2}	5.78×10^{-2}
ММО, $m = 1$ (2)	1.72×10^{-3}	8.24×10^{-3}	1.07×10^{-2}	5.73×10^{-2}
ММО, $m = 2$	7.67×10^{-3}	8.24×10^{-3}	1.31×10^{-2}	5.78×10^{-2}
ММО, $m = 5$	7.67×10^{-3}	8.24×10^{-3}	1.31×10^{-2}	5.78×10^{-2}
МСГ	4.11×10^{-3}	7.67×10^{-3}	7.88×10^{-3}	9.28×10^{-3}

Видно, что для m -моментного ММО качество решения ухудшается сильнее, чем для МСГ. Рестарты позволяют уменьшить невязку, поскольку препятствуют накоплению погрешности. При небольшом шуме ($\delta \leq 0.5\%$) метод с рестартами позволяет достичь наилучшего качества решения. Изобразим наилучшие по функционалу и по расстоянию приближения, полученные при $\delta = 1\%$ с помощью 1-моментного ММО с рестартами и без них, на рисунке 58.

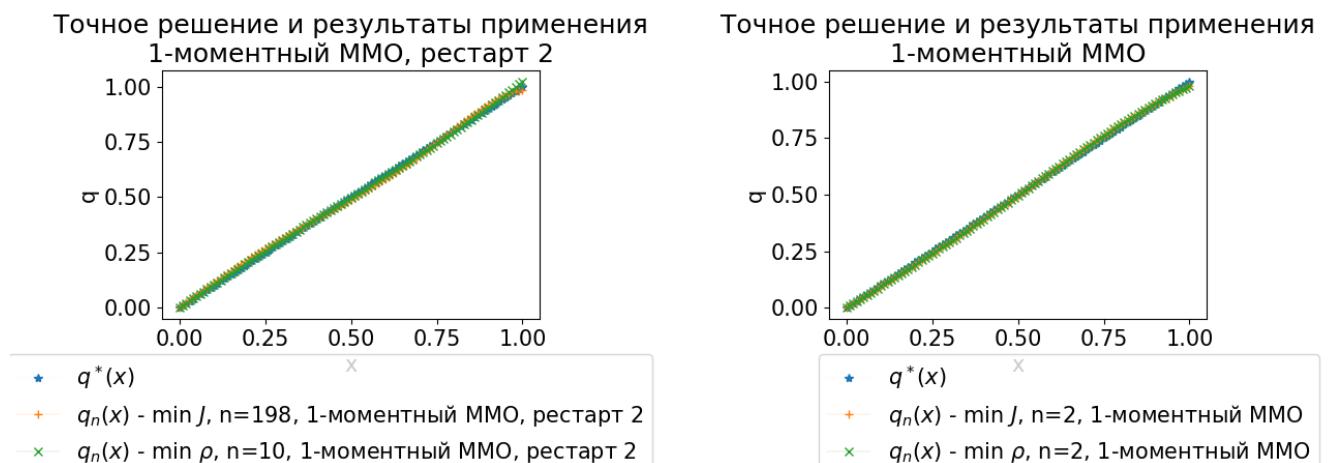


Рисунок 58 – Сравнение с точным решением: 1-моментный ММО с рестартами и без них. Шум $\delta = 1\%$. Задача (1.11), ядро K_3 .

Решение восстанавливается почти так же хорошо, как при точно заданной правой части.

3.5.4. Выводы

Эксперименты показывают, что градиентные методы оптимизации позволяют эффективно решать интегральные уравнения Фредгольма 1-ого рода.

Это согласуется с теорией: интегральный оператор с непрерывным ядром является компактным, поэтому при точном вычислении градиентные методы сходились бы к точному решению по теореме 2. При наличии неизбежной погрешности качество решения ограничивается лишь её величиной.

Если ядро вырожденное ($K(x, s) = \sum_{k=1}^n \lambda_k \varphi_k(x) \varphi_k(s)$, где $\|\varphi_k\| = 1$), то вместо $q^*(x)$ восстанавливается проекция $q^*(x)$ на $\text{Lin}\{\varphi_1, \dots, \varphi_n\}$. Минимально возможное расстояние — «зазор» в неравенстве Бесселя:

$$\|q - q^*\|^2 \geq \|q^*\|^2 - \sum_{k=1}^n \langle q^*, \varphi_k \rangle^2.$$

Отсутствие сходимости к q^* никак не опровергает теорему 3, ведь среди её условий — инъективность оператора A . Интегральный оператор с вырожденным ядром инъективным не является.

Это подтверждает эмпирический результат: чем больше слагаемых содержит ядро, тем меньше достижимая невязка по аргументу при решении соответствующего интегрального уравнения.

На всех трёх задачах m -моментный метод минимальных ошибок показал наилучшие результаты по сравнению с другими методами.

Для первого ядра (сумма синусов) решение восстанавливается плохо: отклонение достигает 0.1 в L_2 -норме и 0.5 в C -норме. Это ожидаемо: в разложении x в ряд Фурье по системе $\{1, \cos \pi n x, \sin \pi n x\}_{n=1}^\infty$ коэффициенты при нечётных косинусах не равны нулю. Максимальное значение соответствующей части ряда Фурье составляет как раз 0.5.

Для второго и третьего ядер (сумма синусов и косинусов, ядро Фурье) решение восстанавливается хорошо. Максимальное отклонение найденных с помощью m -моментного ММО приближённых решений от известного точного решения $q(x)$ — минимум на 4 порядка меньше диапазона изменения $q(x)$. Форма графика восстанавливается.

Построенные алгоритмы численного решения интегрального уравнения Фредгольма 1-ого рода оказались менее устойчивыми, чем основанные на применении метода сопряжённых градиентов, при достаточно большом уровне шума. При малом шуме они позволяют достаточно хорошо решать задачу, в том числе и с использованием рестартов.

3.6. Об использовании построенных методов при неизвестном точном решении

На модельных задачах, для которых известно точное решение и, соответственно, можно вычислить расстояние до него после каждого шага, t -моментный метод минимальных ошибок показал хорошие результаты. Возникает вопрос, как применять его и использовать оценки сходимости при неизвестном точном решении. Кратко напомним доказанные теоретически и полученные экспериментально свойства метода.

- При решении обратной задачи невозможна общая оценка сходимости по аргументу, не зависящая от конкретного начального приближения.
- Оценка сходимости по функционалу не позволяет сделать вывод о приближении к точному решению.
- Сходимость по функционалу не является монотонной.
- При отсутствии шума часто сначала достигается минимальное значение функционала, потом минимальное расстояние до точного решения; затем погрешность вычислений приводит к прекращению приближения к точному решению. Это видно, например, на графиках 11, 17, 21, 24, 29.
- При достаточно большом шуме t -моментный ММО без рестартов работает хуже, чем классические методы.
- При отсутствии шума приближённые решения на шагах t -моментного ММО, соответствующих минимумам функционала и расстояния до точного решения, почти не отличаются.

Исходя из этих фактов, можно предложить следующий способ применения построенного t -моментного метода минимальных ошибок к реальным задачам с небольшой погрешностью определения правой части операторного уравнения.

- Выбирается число T , например $T = 100$ — «задержка останова».
- На каждом шаге сохраняется текущее минимальное достигнутое значение функционала J_{min} и номер шага t , на котором оно достигнуто.

- Если в течение T шагов J_{min} не обновлено, работа m -моментного ММО прекращается.
- В качестве приближённого решения выбирается q_t .

Чем меньше T , тем осторожнее алгоритм. Соответственно, увеличение T может быть уместно при высокой точности вычислений. Возможен также рестарт метода при достижении условия остановки: тогда при линейном увеличении числа операций качество решения заведомо не ухудшится.

Заключение

В данной диссертации построен m -моментный метод минимальных ошибок для решения задач оптимизации квадратичных функционалов. Стого доказаны его теоретические свойства, проведена их экспериментальная проверка. Проверена численная устойчивость при решении разнообразных обратных задач. Также предложен способ применения построенного метода к реальным задачам, когда точное решение неизвестно, а погрешность вычислений может влиять на качество решения.

В работе доказан ряд теорем о построенном методе.

Теорема 1 показывает, что при $m = \infty$ построенный метод оптимальен: никакой метод первого порядка не может сходиться по аргументу быстрее него. Это верно при точных вычислениях, поэтому не всегда результаты экспериментов согласуются с теорией. Теорема подтверждается на практике, если применение m -моментного метода минимальных ошибок сочетается с высокой точностью вычислений на всех промежуточных шагах.

Теорема 3 устанавливает сам факт сходимости по аргументу даже для не сильно выпуклого функционала (в предположении точных вычислений). Однако теорема 6 свидетельствует о невозможности общих оценок сверху для скорости сходимости по аргументу при решении некорректных задач. Этот отрицательный результат относится не к самому построенному методу, а к области его применения.

Проведены многочисленные эксперименты на некорректных задачах различных типов (эллиптическая, параболическая, гиперболическая, интегральное уравнение), показывающие эффективность их решения с применением градиентных методов оптимизации и преимущество построенного m -моментного метода минимальных ошибок. Его вычислительная сложность — на треть меньше, чем у метода сопряжённых градиентов при одинаковом числе итераций.

Например, начально-краевая задача для уравнения Гельмгольца требует всего нескольких итераций нового метода для нахождения решения с погрешностью не более 10^{-3} . А при добавлении к дополнительному условию относительного шума

качество решения соответствует качеству решения с помощью метода сопряжённых градиентов.

Ретроспективную задачу Коши для уравнения теплопроводности характеризует значительно более жёсткий, чем у предыдущей задачи, спектр связанного с ней самосопряжённого оператора. Это замедляет сходимость методов оптимизации. t -моментный метод минимальных ошибок превосходит другие методы при точных вычислениях. Он оказался достаточно чувствительным к погрешностям, возникающим из-за шума или ошибок дискретизации, но и в таких случаях он достигает примерно такой же точности, как и метод сопряжённых градиентов, при меньшей вычислительной сложности. А при уровне шума $\delta \leq 5 \times 10^{-3}$ преимущество нового метода сохраняется.

Обратная задача термоакустики интересна тем, что её функционал имеет более сложный вид — сумма простых квадратичных функционалов. t -моментный метод минимальных ошибок показал высокую эффективность и на ней. Он обеспечил наилучшее восстановление неоднородностей по сравнению с другими методами, даже при минимальном количестве дополнительной информации (задача 1). Эффективность решения задачи 1 сохраняется при добавлении относительного шума с $\delta \leq 5 \times 10^{-4}$. Задача 2, в которой больше дополнительной информации, удовлетворительно решается при $\delta \leq 2 \times 10^{-3}$.

Интегральные уравнения Фредгольма 1-ого рода также допускают эффективное численное решение с помощью t -моментного метода минимальных ошибок. При уровне шума $\delta \leq 5 \times 10^{-3}$ оказывается выгодным его применение с рестартами.

Итак, построенный в работе метод позволяет лучше решать разнообразные обратные задачи.

Благодарности

Автор выражает огромную благодарность своему научному руководителю, доктору физико-математических наук Александру Владимировичу Гасникову за поддержку, помошь и ценные рекомендации на протяжении всей исследовательской работы.

Также автор благодарит доктора физико-математических наук Максима Александровича Шишленина за предложенную им начально-краевую задачу для уравнения Гельмгольца, решение которой стало началом работы в области применения методов оптимизации к решению некорректных и обратных задач.

Автор выражает признательность доктору технических наук Александру Юрьевичу Горнову за ряд ценных замечаний, в частности, за обнаружение общих свойств построенного метода с методом Поляка.

Автор благодарит Наталью Николаевну Клесову и Александра Савельевича Штерна — своих наставников в школьные годы — за переданные знания, сохранение и развитие интереса к математике.

Список литературы

- [1] Албу А.Ф., Зубов В.И. О восстановлении коэффициента теплопроводности вещества по температурному полю // Журнал вычислительной математики и математической физики. — 2018. — Т. 58, № 10. — С. 1640-1655.
- [2] Вайникко Г.М. Оценки погрешности метода последовательных приближений для некорректных задач // Автоматика и телемеханика. — 1980. — № 3. — С. 84–92.
- [3] Васильев В.И., Кардашевский А.М. Численное решение ретроспективной обратной задачи теплопроводности с помощью интеграла Пуассона // Сибирский журнал индустриальной математики. — 2018. — Т. 21, № 3. — С. 26–36.
- [4] Васильев Ф.П. Методы оптимизации, часть вторая: Оптимизация в функциональных пространствах. Регуляризация. Аппроксимация // МЦНМО. — Москва. — 2011. — 433 с. — ISBN 978-5-94057-708-9
- [5] Васин В.В., Пересторонина Г.Я., Пруткин И.Л., Тимерханова Л.Ю. Решение трехмерных обратных задач гравиметрии и магнитометрии для трехслойной среды // Математическое моделирование. — 2003. — Т. 15, № 2. — С. 69-76.
- [6] Васин В.В., Сережникова Т.И. Об одном алгоритме решения уравнения Фредгольма–Стильтьеса // Известия вузов. Математика. — 2001. — № 4. — С. 3-10.
- [7] Гасников А.В. Современные численные методы оптимизации. Метод универсального градиентного спуска: Учебное пособие. – 2-е изд., испр. // МЦНМО. — Москва. — 2021. — 272 с. — ISBN 978-5-4439-4199-8
- [8] Гласко В.Б., Кулик Н.И., Тихонов А.Н. Об определении геоэлектрического разреза на основе метода регуляризации // Журнал вычислительной математики и математической физики. — 1972. — Т. 12, № 1. — С. 139-149.

- [9] Горчаков А.Ю., Зубов В.И. Об одновременном определении коэффициента теплопроводности и объемной теплоемкости вещества // Журнал вычислительной математики и математической физики. — 2023. — Т. 63, № 8. — С. 1279–1295.
- [10] Демьянов В.Ф. Минимакс: дифференцируемость по направлениям // Издательство Ленинградского университета. — Ленинград. — 1974. — 112 с.
- [11] Долгополова Т.Ф., Иванов В.К. О численном дифференцировании // Журнал вычислительной математики и математической физики. — 1966. — Т. 6, № 3. — С. 223–232.
- [12] Евтушенко Ю.Г. Оптимизация и быстрое автоматическое дифференцирование // Вычислительный центр им. А.А. Дородницына Российской академии наук. — Москва. — 2013. — 144 с.
- [13] Еремеева М.С. Сравнение итерационных методов решения обратной ретроспективной задачи теплопроводности // Вестник СВФУ. — 2015. — Т. 12, № 1. — С. 15–24.
- [14] Зубов В.И., Албу А.Ф. Восстановление коэффициента теплопроводности вещества по поверхностному тепловому потоку // Журнал вычислительной математики и математической физики. — 2018. — Т. 58, № 12. — С. 2112–2126.
- [15] Иванов В.К. Задача квазиобращения для уравнения теплопроводности в равномерной метрике // Дифференциальные уравнения. — 1972. — Т. 8, № 4. — С. 652-658.
- [16] Иванов В.К. Интегральные уравнения первого рода и приближенное решение обратной задачи потенциала // Доклады Академии наук СССР. — 1962. — Т. 142, № 5. — С. 998-1000.
- [17] Иванов В.К. Об интегральных уравнениях Фредгольма первого рода // Дифференциальные уравнения. — 1967. — Т. 3, № 3. — С. 410-421.
- [18] Иванов В.К., Мельникова И.В. Общая схема устранения расходимостей разного рода // Сибирский математический журнал. — 1988. — Т. 29, № 6. — С. 66–73.
- [19] Кабанихин С.И., Куликов И.М. Шишленин М.А. Алгоритм восстановления характеристик начального состояния сверхновой звезды // Журнал

вычислительной математики и математической физики. — 2020. — Т. 60, № 6. — С. 1035–1034.

- [20] Кабанихин С. И., Хасанов А. Х. Пененко А. В. Метод градиентного спуска для решения обратной коэффициентной задачи теплопроводности // Сибирский журнал вычислительной математики. — 2008. — Т. 11, № 1. — С. 41–51.
- [21] Кабанихин С. И., Шишленин М. А. Восстановление коэффициента диффузии, зависящего от времени, по нелокальным данным // Сибирский журнал вычислительной математики. — 2018. — Т. 21, № 1. — С. 55–63.
- [22] Кабанихин С. И., Шишленин М. А., Криворотко О. И. Оптимизационный метод решения обратной задачи термоакустики // Сибирские электронные математические известия. — 2011. — Т. 8. — С. 263–292.
- [23] Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа // ФИЗМАТЛИТ. — Москва. — 2004. — 572 с. — ISBN 5-9221-0266-4
- [24] Лаврентьев М. М. О задаче Коши для уравнения Лапласа // Известия Академии наук СССР. Серия математическая. — 1956. — Т. 20, № 6. — С. 819–842.
- [25] Лаврентьев М. М. Математические задачи томографии и гиперболические отображения // Сибирский математический журнал. — 2001. — Т. 42, № 5. — С. 1094–1105.
- [26] Лаврентьев М. М., Васильев В. Г. О постановке некоторых некорректных задач математической физики // Сибирский математический журнал. — 1966. — Т. 7, № 3. — с. 559–576.
- [27] Лаврентьев М. М., Клибанов М. В. Об одной обратной задаче для уравнения параболического типа // Дифференциальные уравнения. — 1975. — Т. 11, № 9. — С. 1647–1651.
- [28] Лаврентьев М. М., Максимов В. И. О восстановлении правой части параболического уравнения // Журнал вычислительной математики и математической физики. — 2008. — Т. 48, № 4. — С. 674–680.
- [29] Лаврентьев М. М., Федотов А. М. О постановке некоторых некорректных задач математической физики со случайными исходными данными // Журнал вычислительной математики и математической физики. — 1982. — Т. 2, № 1. — С. 133–143.

- [30] *Лукьяненко Д. В., Ягола А. Г.* Использование многопроцессорных систем для решения обратных задач, сводящихся к интегральным уравнениям Фредгольма 1-го рода // Труды Института математики и механики УрО РАН. — 2012. — Т. 18, № 1. — С. 222–234.
- [31] *Немировский А. С.* О регуляризующих свойствах метода сопряжённых градиентов на некорректных задачах // Журнал вычислительной математики и математической физики. — 1986. — Т. 26, № 3. — С. 332-347.
- [32] *Плетнев Н. В., Двуреченский П. Е., Гасников А. В.* Применение градиентных методов оптимизации для решения задачи Коши для уравнения Гельмгольца // Компьютерные исследования и моделирование. — 2022. — Т. 14, № 2. — С. 417-444
- [33] *Плетнев Н. В., Матюхин В. В.* О модификации метода покомпонентного спуска для решения некоторых обратных задач математической физики // Компьютерные исследования и моделирование. — 2023. — Т. 15, № 2. — С. 301-316.
- [34] *Плетнев Н. В.* О построении градиентного метода квадратичной оптимизации, оптимального с точки зрения минимизации расстояния до точного решения // Журнал вычислительной математики и математической физики. — 2025. — Т. 65, № 10. — С. 1625-1648.
- [35] *Поляк Б. Т.* Минимизация негладких функционалов // Журнал вычислительной математики и математической физики. — 1969. — Т. 9, № 3. — С. 509-521.
- [36] *Романов В. Г., Мошков П. С.* Одномерная обратная задача об определении источника цунами // Сибирский журнал индустриальной математики. — 2011. — Т. 14, № 3. — С. 87–99.
- [37] *Рябенький В. С.* Введение в вычислительную математику: Учебное пособие. — 2-е изд., испр. // ФИЗМАТЛИТ. — Москва. — 2000. — 296 с. — ISBN 5-9221-0047-5
- [38] *Сабиров К. Б.* Обратные задачи для уравнения Гельмгольца по отысканию правой части с нелокальным интегральным наблюдением // Журнал вычислительной математики и математической физики. — 2023. — Т. 63, № 7. — С. 1145-1155.

- [39] Сабиров К. Б., Мартемьянова Н. В. К вопросу о корректности обратных задач для неоднородного уравнения Гельмгольца // Вестник СГТУ. Серия: Физико-математические науки. — 2018. — Т. 22, № 2. — С. 269-292.
- [40] Самарский А. А., Вабищевич П. Н., Васильев В. И. Итерационное решение ретроспективной обратной задачи теплопроводности // Математическое моделирование. — 1997. — Т. 9, № 5. — С. 119-127.
- [41] Тихонов А. Н. К математическому обоснованию теории электромагнитных зондирований // Журнал вычислительной математики и математической физики. — 1965. — Т. 5, № 3. — С. 545–548.
- [42] Тихонов А. Н. Об устойчивости алгоритмов для решения вырожденных систем линейных алгебраических уравнений // Журнал вычислительной математики и математической физики. — 1965. — Т. 5, № 4. — С. 718–722.
- [43] Тихонов А. Н. О некорректных задачах оптимального планирования // Журнал вычислительной математики и математической физики. — 1966. — Т. 6, № 1. — С. 81-89.
- [44] Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач // Наука. — Москва. — 1979. — 283 с.
- [45] Тихонов А. Н., Гончарский А. В., Степанов В. В., Ягола А. Г. Численные методы решения некорректных задач // Наука. — Москва. — 1990. — 232 с. — ISBN 5-02-014135-6
- [46] Тихонов А. Н., Самарский А. А. Уравнения математической физики: Учебное пособие. – 6-е изд., испр, и доп. // Издательство МГУ. — Москва. — 1999. — 798 с. — ISBN: 5-211-04138-0
- [47] Akindinov G. D., Gasnikov A. V., Krivorotko O. I., Matyukhin V. V., Pletnev N. V. Gradient-type Approaches to Inverse and Ill-Posed Problems of Mathematical Physics // Computational Mathematics and Mathematical Physics. — 2024. — Vol. 64. — P. 1974–1990. — <https://doi.org/10.1134/S0965542524701136>
- [48] Danskin J. M. The theory of Max Min // Springer. — Berlin. — 1967. — 126 p.
- [49] Devanathan N., Boyd S. Polyak Minorant Method for Convex Optimization // arXiv preprint arXiv:2310.07922. — 2024. — <https://arxiv.org/abs/2310.07922>; (Accessed on October 14, 2025).

- [50] *Gasnikov A., Kabanikhin S., Mohammed A., Shishlenin M.* Convex optimization in Hilbert space with applications to inverse problems // arXiv preprint arXiv:1703.00267. — 2017. — <https://arxiv.org/abs/1703.00267>; (Accessed on October 14, 2025).
- [51] *Goujaud B., Taylor A., Dieuleveut A.* Quadratic minimization: from conjugate gradient to an adaptive Heavy-ball method with Polyak step-sizes // arXiv preprint arXiv:2210.06367. — 2022. — <https://arxiv.org/abs/2210.06367>; (Accessed on October 14, 2025).
- [52] *Kabanikhin S., Shishlenin M. et al* Comparative Analysis of Methods for Regularizing an Initial Boundary Value Problem for the Helmholtz Equation // Journal of Applied Mathematics. — 2014. — Vol. 2014. — P. 786326 (7 pages). <https://doi.org/10.1155/2014/786326>
- [53] *Kabanikhin S. I.* Inverse and ill-posed problems: theory and applications // Walter de Gruyter GmbH & Co. — Berlin. — 2012. — 459 p. — ISBN 978-3-11-022400-9
- [54] *Polyak B. T.* Introduction to Optimization // Optimization Software. — New York. — 1987. — 438 p.
- [55] *Polyak B. T.* Iterative algorithms for singular minimization problems // Nonlinear Programming 4, Academic Press. — 1981. — P. 147–166.
- [56] *Shishlenin M. A., Savchenko N. A., Novikov N. S., Klyuchinskiy D. V.* On the reconstruction of the absorption coefficient for the 2D acoustic system // Сибирские электронные математические известия. — 2023. — Vol. 20, no. 2. — P. 1474–1489.
- [57] *Vasin A., Gasnikov A., Dvurechensky P., Spokoiny V.* Stopping rules for accelerated gradient methods with additive noise in gradient // arXiv preprint arXiv:2102.02921. — 2021. — <https://arxiv.org/abs/2102.02921>; (Accessed on October 14, 2025).

Список рисунков

1. Рисунок 1 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.5).	58
2. Рисунок 2 — Сходимость: 1-моментный ММО и МСГ. Задача (1.5).	58
3. Рисунок 3 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.5).	59
4. Рисунок 4 — Сравнение с точным решением: 5-моментный ММО, МСГ. Задача (1.5).	59
5. Рисунок 5 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.6), $\kappa_{max} = 0.4$.	71
6. Рисунок 6 — Сходимость: 1-моментный ММО и МСГ. Задача (1.6), $\kappa_{max} = 0.4$.	72
7. Рисунок 7 — Сравнение с точным решением: 1-моментный ММО, МСГ. Задача (1.6), $\kappa_{max} = 0.4$.	72
8. Рисунок 8 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.6), $\kappa_{max} = 0.6$.	73
9. Рисунок 9 — Сходимость: 1-моментный ММО и МСГ. Задача (1.6), $\kappa_{max} = 0.6$.	74
10. Рисунок 10 — Сравнение с точным решением: 1-моментный ММО, МСГ. Задача (1.6), $\kappa_{max} = 0.6$.	74
11. Рисунок 11 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.8), $\kappa_{max} = 0.4$.	76
12. Рисунок 12 — Сходимость: 1-моментный ММО и МСГ. Задача (1.8), $\kappa_{max} = 0.4$.	76
13. Рисунок 13 — Сравнение с точным решением: 1-моментный ММО, STM. Задача (1.8), $\kappa_{max} = 0.4$.	77
14. Рисунок 14 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.8), $\kappa_{max} = 0.6$.	78
15. Рисунок 15 — Сходимость: 1-моментный ММО и МСГ. Задача (1.8), $\kappa_{max} = 0.6$.	78
16. Рисунок 16 — Сравнение с точным решением: 1-моментный ММО, МСГ. Задача (1.8), $\kappa_{max} = 0.6$.	79
17. Рисунок 17 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.7), $\kappa_{max} = 0.4$.	87
18. Рисунок 18 — Сходимость: 1-моментный ММО и МСГ. Задача (1.7), $\kappa_{max} = 0.4$.	87

19. Рисунок 19 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.7), $\kappa_{max} = 0.4$.	88
20. Рисунок 20 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.7), $\kappa_{max} = 0.6$.	89
21. Рисунок 21 — Сходимость: 1-моментный ММО и МСГ. Задача (1.7), $\kappa_{max} = 0.6$.	89
22. Рисунок 22 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.7), $\kappa_{max} = 0.6$.	89
23. Рисунок 23 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.9), $\kappa_{max} = 0.1$.	91
24. Рисунок 24 — Сходимость: 1-моментный ММО и МСГ. Задача (1.9), $\kappa_{max} = 0.1$.	91
25. Рисунок 25 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.9), $\kappa_{max} = 0.1$.	92
26. Рисунок 26 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.9), $\kappa_{max} = 0.2$.	93
27. Рисунок 27 — Сходимость: 1-моментный ММО и МСГ. Задача (1.9), $\kappa_{max} = 0.2$.	93
28. Рисунок 28 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.9), $\kappa_{max} = 0.2$.	93
29. Рисунок 29 — Сходимость: 5-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (3.4.1.1).	102
30. Рисунок 30 — Сходимость: 5-моментный ММО и МСГ. Задача (3.4.1.1).	102
31. Рисунок 31 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (3.4.1.1).	103
32. Рисунок 32 — Сравнение с точным решением: 5-моментный ММО, МСГ, шаг Поляка. Задача (3.4.1.1).	104
33. Рисунок 33 — Сравнение с точным решением: m -моментный ММО с $m \in \{1, 2, 5\}$, МСГ. Шум $\delta = 2 \times 10^{-4}$. Задача (3.4.1.1).	106
34. Рисунок 34 — Сравнение с точным решением: m -моментный ММО с $m \in \{1, 2, 5\}$, МСГ. Шум $\delta = 5 \times 10^{-4}$. Задача (3.4.1.1).	107
35. Рисунок 35 — Сравнение с точным решением: m -моментный ММО с $m \in \{1, 2, 5\}$, МСГ. Шум $\delta = 10^{-3}$. Задача (3.4.1.1).	108
36. Рисунок 36 — Сходимость: 5-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (3.4.1.2).	109
37. Рисунок 37 — Сходимость: 5-моментный ММО и МСГ. Задача (3.4.1.2).	110
38. Рисунок 38 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (3.4.1.2).	110
39. Рисунок 39 — Сравнение с точным решением: 5-моментный ММО, МСГ, шаг Поляка. Задача (3.4.1.2).	111

40. Рисунок 40 — Сравнение решений: шаг Поляка, m -моментный ММО с $m \in \{2, 5\}$, МСГ. Шум $\delta = 2 \times 10^{-3}$. Задача (3.4.1.2).	113
41. Рисунок 41 — Сравнение решений: шаг Поляка, m -моментный ММО с $m \in \{2, 5\}$, МСГ. Шум $\delta = 5 \times 10^{-3}$. Задача (3.4.1.2).	114
42. Рисунок 42 — Сходимость: 5-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (3.4.1.3).	115
43. Рисунок 43 — Сходимость: 5-моментный ММО и МСГ. Задача (3.4.1.3).	116
44. Рисунок 44 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (3.4.1.3).	116
45. Рисунок 45 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.11), ядро K_1 .	120
46. Рисунок 46 — Сходимость: 1-моментный ММО и МСГ. Задача (1.11), ядро K_1 .	121
47. Рисунок 47 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.11), ядро K_1 .	121
48. Рисунок 48 — Сравнение с точным решением: 5-моментный ММО, МСГ. Задача (1.11), ядро K_1 .	122
49. Рисунок 49 — Сравнение с точным решением: 1-моментный ММО с рестартами, МСГ. Шум $\delta = 1\%$. Задача (1.11), ядро K_1 .	123
50. Рисунок 50 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.11), ядро K_2 .	124
51. Рисунок 51 — Сходимость: 1-моментный ММО и МСГ. Задача (1.11), ядро K_2 .	125
52. Рисунок 52 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.11), ядро K_2 .	125
53. Рисунок 53 — Сравнение с точным решением: 5-моментный ММО, МСГ. Задача (1.11), ядро K_2 .	126
54. Рисунок 54 — Сравнение с точным решением: 1-моментный ММО с рестартами, МСГ. Шум $\delta = 1\%$. Задача (1.11), ядро K_2 .	127
55. Рисунок 55 — Сходимость: 1-моментный ММО, адаптивный тяжёлый шарик, шаг Поляка и STM. Задача (1.11), ядро K_3 .	128
56. Рисунок 56 — Сходимость: 1-моментный ММО и МСГ. Задача (1.11), ядро K_3 .	129
57. Рисунок 57 — Сходимость: m -моментный ММО с $m \in \{1, 2, 5, \infty\}$. Задача (1.11), ядро K_3 .	129
58. Рисунок 58 — Сравнение с точным решением: 1-моментный ММО с рестартами и без них. Шум $\delta = 1\%$. Задача (1.11), ядро K_3 .	130

Список таблиц

1. Таблица 1 — Результаты работы методов. Задача (1.5).	57
2. Таблица 2 — Расстояние до точного решения при наличии шума δ в правой части. Задача (1.5).	60
3. Таблица 3 — Результаты работы методов. Задача (1.6), $\kappa_{max} = 0.4$	71
4. Таблица 4 — Результаты работы методов. Задача (1.6), $\kappa_{max} = 0.6$	73
5. Таблица 5 — Расстояние до точного решения при наличии шума δ в правой части. Задача (1.6), $\kappa_{max} = 0.4$	75
6. Таблица 6 — Результаты работы методов. Задача (1.8), $\kappa_{max} = 0.4$	75
7. Таблица 7 — Результаты работы методов. Задача (1.8), $\kappa_{max} = 0.6$	77
8. Таблица 8 — Расстояние до точного решения при наличии шума δ в правой части. Задача (1.8), $\kappa_{max} = 0.6$	79
9. Таблица 9 — Результаты работы методов. Задача (1.7), $\kappa_{max} = 0.4$	86
10. Таблица 10 — Результаты работы методов. Задача (1.7), $\kappa_{max} = 0.6$	88
11. Таблица 11 — Результаты работы методов. Задача (1.9), $\kappa_{max} = 0.1$	90
12. Таблица 12 — Результаты работы методов. Задача (1.9), $\kappa_{max} = 0.2$	92
13. Таблица 13 — Результаты работы методов. Задача (3.4.1.1).	101
14. Таблица 14 — Расстояние до точного решения при наличии шума δ в правой части. Задача (3.4.1.1).	105
15. Таблица 15 — Результаты работы методов. Задача (3.4.1.2).	109
16. Таблица 16 — Расстояние до точного решения при наличии шума δ в правой части. Задача (3.4.1.2).	112
17. Таблица 17 — Результаты работы методов. Задача (3.4.1.3).	115
18. Таблица 18 — Результаты работы методов. Задача (1.11), ядро K_1	119
19. Таблица 19 — Расстояние до точного решения при наличии шума δ в правой части. Задача (1.11), ядро K_1	122
20. Таблица 20 — Результаты работы методов. Задача (1.11), ядро K_2	123
21. Таблица 21 — Расстояние до точного решения при наличии шума δ в правой части. Задача (1.11), ядро K_2	126

22. Таблица 22 — Результаты работы методов. Задача (1.11), ядро K_3 .	128
23. Таблица 23 — Расстояние до точного решения при наличии шума δ в правой части. Задача (1.11), ядро K_3 .	130

Список сокращений

PMM	Polyak Minorant Method
STM	Similar Triangle Method
ММО	Метод минимальных ошибок
МСГ	Метод сопряжённых градиентов
ПР	Полака-Рибьера
ФР	Флэтчера-Ривза