МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра математической физики и вычислительной математики

Методы поиска экстремума нулевого порядка в кнечномерных пространствах

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента 2 курса 217 группы

направление **01.04.02 – Прикладная математика и информатика**

механико-математического факультета

Леднова Александра Петровича

Научный руководитель

зав.кафедрой, д.ф.-м.н., профессор _______________ В.А. Юрко

Заведующий кафедрой

зав.кафедрой, д.ф.-м.н., профессор _______________ В.А. Юрко

Саратов 2025

Введение Поиск экстремума целевой функции является одной из ключевых задач прикладной математики и вычислительной науки, находящей широкое применение в инженерии, экономике, медицине и других областях. Минимизация или максимизация функций позволяет формализовать процессы принятия решений в условиях многопараметрических ограничений. Современные задачи оптимизации характеризуются высокой размерностью, особенно в областях машинного обучения, физического моделирования и инженерного проектирования. Это делает поиск экстремума не только теоретически значимой, но и вычислительно сложной задачей, требующей применение специализированных и высокоэффективных алгоритмов для анализа поведения функций в многомерных пространствах.

Методы поиска экстремума классифицируются в зависимости от используемой информации: методы нулевого порядка опираются только на значения функции; методы первого порядка используют также первые производные; методы второго порядка дополнительно учитывают вторые производные. В практических задачах часто отсутствует аналитическое выражение целевой функции, и можно лишь вычислять ее значения в отдельных точках. В таких условиях особое значение приобретают методы нулевого порядка, также называемые методами прямого поиска, которые не требуют информации о производных и потому обладают широкой применимостью. Эти методы подразделяются на алгоритмы с фиксированной стратегией поиска, к которым относятся, например, метод покоординатного спуска и метод Хука–Дживса, и адаптивные алгоритмы, такие как методы Розенброка и Пауэлла, где направления поиска изменяются в ходе оптимизации на основе анализа поведения целевой функции.

Актуальность настоящей работы определяется возросшей потребностью в эффективных численных методах оптимизации, способных решать задачи высокой размерности в условиях ограниченной информации о целевой функции. Такие условия характерны для многих прикладных задач, в которых невозможно получить аналитическое выражение функции, а её значения доступны лишь в результате экспериментов или трудоемких вычислений. Несмотря на широкое распространение методов нулевого порядка, многие из них основаны на эвристических подходах и не имеют строгого теоретического обоснования, что затрудняет оценку их эффективности. Кроме того, выбор подходящего алгоритма во многом зависит от конкретных условий задачи, таких как рельеф целевой функции, требования к точности и вычислительные ограничения, что делает исследование и сопоставление различных подходов особенно важным.

Целью магистерской работы является исследование методов нулевого порядка, анализ их алгоритмических особенностей, а также проведение сравнительного анализа эффективности этих методов для различных типов рельефа целевой функции.

Для достижения данной цели можно выделить ряд задач:

— изучение теоретических основ релаксационных последовательностей и условий их сходимости;
— анализ общей схемы методов спуска, включая выбор направления и длины шага;
— рассмотрение наиболее распространенных методов прямого поиска;
— проведение вычислительного эксперимента по сравнению эффективности изученных методов нулевого порядка.

Работа включает введение, четыре раздела, заключение, список использованных источников из 20 наименований, а также приложение с программным кодом. В первом разделе изложены теоретические основы релаксационных последовательностей и условия их сходимости. Второй раздел посвящен общей схеме методов спуска. Здесь рассматриваются принципы выбора направления спуска и длины шага. Третий раздел содержит описание алгоритмов наиболее распространенных методов прямого поиска: метода покоординатного спуска, метода Хука-Дживса, метода Розенброка, метода Пауэлла. В четвертом разделе представлены результаты вычислительных экспериментов, позволяющие сравнить эффективность рассмотренных алгоритмов в зависимости от рельефа целевой функции.

Основное содержание работы В первом разделе магистерской работы рассматриваются теоретические основы релаксационных последовательностей, а также условия их сходимости. Численные методы решения задачи поиска экстремума целевой функции базируются на построении подобных последовательностей.

Пусть требуется найти точку $x^* \in \mathbb{R}^n$, в которой ограниченная снизу целевая функция $f(x)$, определенная в $\mathbb{R}^n$, достигает своего наименьшего значения (считается, что такая точка существует). Тогда задачу поиска минимума функции $f(x)$ можно представить в виде

$$f(x) \to \min, \quad x \in \mathbb{R}^n.$$

Общей чертой всех численных методов решения этой задачи является последовательный переход от точки x^{k-1} к точке x^k, $k \in \mathbb{N}$, начиная с некоторой начальной точки $x^0 \in \mathbb{R}^n$, причем на каждой итерации с номером k выполняется условие

$$f_k = f(x^k) \leqslant f(x^{k-1}) = f_{k-1}, \quad k \in \mathbb{N}. \tag{1}$$

Опредление 1. *Последовательность $\{x^k\}$, элементы которой удовлетворяют неравенству (1), называют релаксационной.*

Численные методы, направленные на построение релаксационных последовательностей, относятся к классу методов спуска.

Второй раздел магистерской работы посвящен изучению методов спуска, включая принципы выбора направления спуска и длины шага. Поскольку различные методы спуска отличаются либо выбором направления, либо способом движения вдоль выбранного направления, для них можно сформулировать единую общую схему.

Пусть существует точка $x^* \in \mathbb{R}^n$, в которой целевая функция $f(x)$ достигает своего минимума. В качестве начального приближения выбирается, вообще говоря, произвольная точка $x^0 \in \mathbb{R}^n$. Тогда процесс поиска решения x^* в рамках методов спуска может быть описан с помощью рекуррентного соотношения:

$$x^k = x^{k-1} + \alpha_k u^k, \quad k \in \mathbb{N},$$

где $u^k \in \mathbb{R}^n$ — единичный вектор, задающий направление спуска на k-й итерации, а $\alpha_k \geqslant 0$ — длина шага, определяющая расстояние, на которое осуществляется переход от точки x^{k-1} к новой точке x^k вдоль направления u^k.

Различия между конкретными численными методами заключаются, как правило, в способах выбора направлений u^k и шагов α_k. Выбор этих параметров определяет как эффективность метода, так и его сходимость.

Критерии завершения итерационного процесса могут базироваться на достижении заданной точности. Обычно процесс минимизации завершается при выполнении одного или обоих следующих условий:

$$\left|x^k - x^{k-1}\right| = \alpha_k < \varepsilon_1, \quad \left|f(x^k) - f(x^{k-1})\right| < \varepsilon_2, \tag{2}$$

где $\varepsilon_1 > 0$ и $\varepsilon_2 > 0$ — заданные допустимые уровни точности по аргументу и по значению функции соответственно. Эти условия позволяют гарантировать, что итерационный процесс прекращается либо при малом изменении координат между последовательными приближениями, либо при незначительном изменении значений целевой функции.

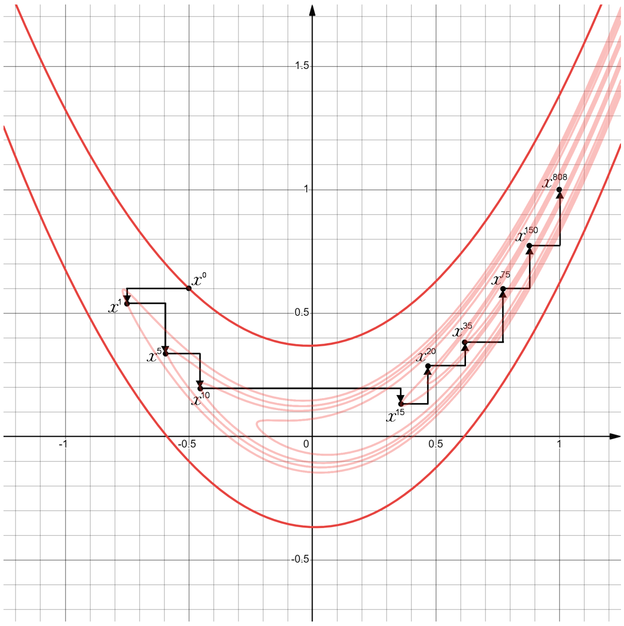
Третий раздел содержит описание алгоритмов наиболее распространенных методов прямого поиска: метода покоординатного спуска, метода Хука-Дживса, метода Розенброка, метода Пауэлла. Графическая иллюстрация работы этих алгоритмов приведена на рисунке 1.

Метод циклического покоординатного спуска представляет собой один из алгоритмов прямого поиска, основанный на идее градиентных методов. Его суть заключается в последовательной одномерной минимизации целевой функции по каждой из координат вектора аргумента $x \in \mathbb{R}^n$.

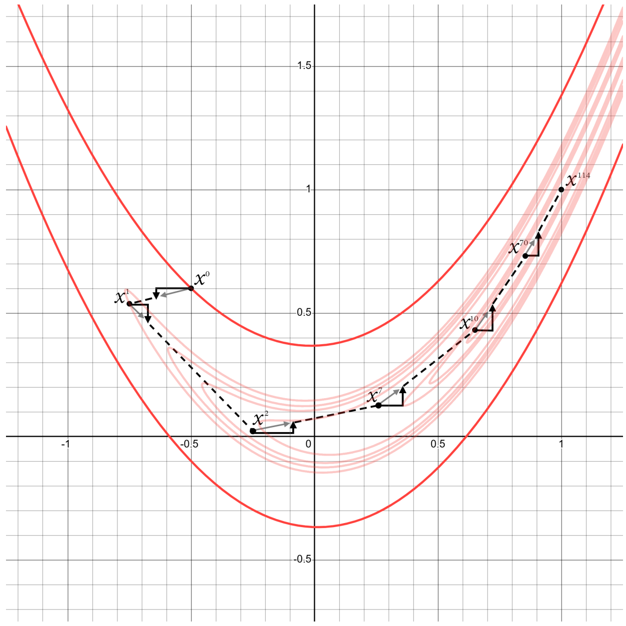
Как иллюстрируется на рисунке 1a, данный метод осуществляет поочередный одномерный поиск вдоль направлений координат стандартного базиса в пространстве $\mathbb{R}^2$. Такой подход прост в реализации, но характеризуется медленной сходимостью на участках с криволинейной геометрией, поскольку поиск осуществляется только вдоль координатных осей.

Эффективность прямого поиска минимума ограниченной снизу функции можно существенно повысить, если на каждом шаге алгоритма целенаправленно выбирать направление спуска. В частности, в методе Хука–Дживса для этой цели на каждом шаге k реализуется предварительный этап исследующего поиска, направленный на определение локального направления убывания целевой функции.

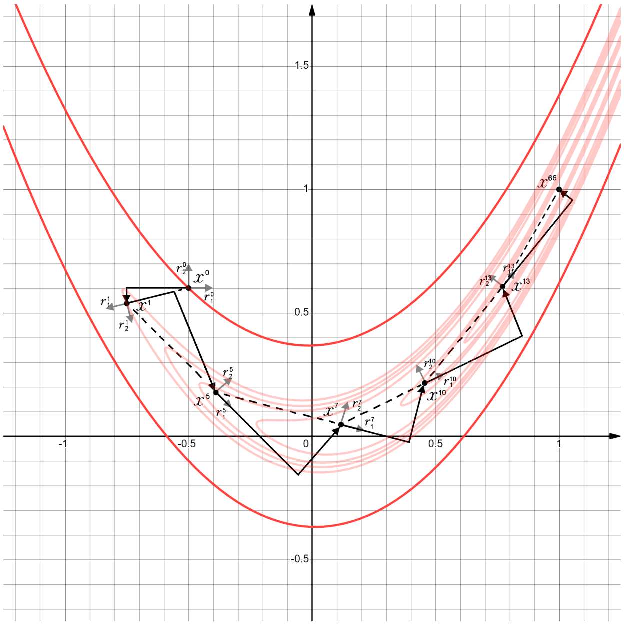
На данном этапе анализируется поведение функции $f(x)$ в окрестности



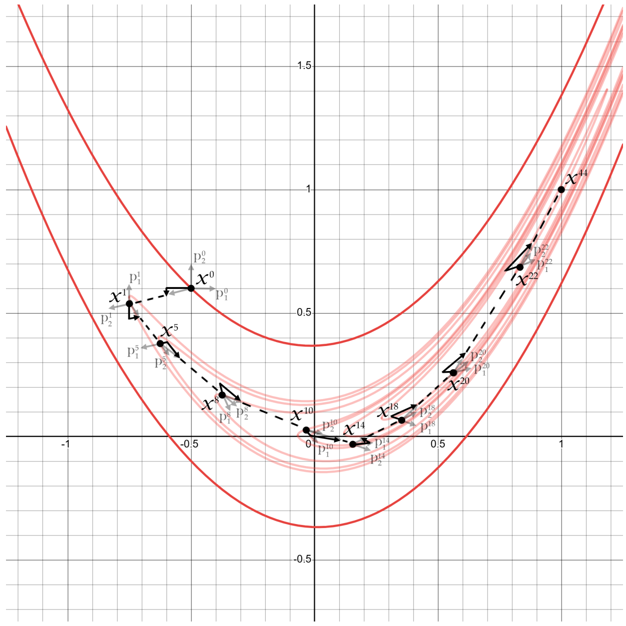
(a) Метод покоординатного спуска



(b) Метод Хука-Дживса



(c) Метод Розенброка



(d) Метод Пауэлла

Рисунок 1 – Графическая иллюстрация методов нулевого порядка

текущей точки x^{k-1}, полученной на предыдущей итерации. В результате формируется новая точка $\widetilde{x}^k$, для которой выполняется условие $f(\widetilde{x}^k) < f(x^{k-1})$, а направление спуска определяется вектором $\widetilde{x}^k - x^{k-1}$.

Как показано на рисунке 1b, после этапа исследующего поиска метод выполняет шаг ускорения вдоль найденного направления убывания. Такой прием позволяет существенно повысить скорость сходимости алгоритма на участках с относительно прямолинейным рельефом целевой функции.

Особенно эффективно это проявляется в случаях, когда направление спуска изменяется незначительно, поскольку один шаг ускорения может заменить несколько последовательных исследующих шагов. В результате общее количество вычислений значений функции значительно сокращается.

Еще одна стратегия поиска точки минимума ограниченной снизу функции $f(x)$, $x \in \mathbb{R}^n$, получила название метода Розенброка. Метод, реализующий эту стратегию, также как и метод Хука–Дживса, предусматривает проведение исследующего поиска на каждом k-ом шаге. Исследование проводится в окрестности текущей точки x^{k-1}, полученной на предыдущем шаге, и направлено на выявление направления убывания значения целевой функции.

Ключевая особенность метода заключается в динамическом построении нового ортонормированного базиса в пространстве $\mathbb{R}^n$ на каждом шаге. Это позволяет адаптировать направления исследующего поиска к локальной структуре функции. В результате определяется точка $\widetilde{x}^k$, удовлетворяющая условию $f(\widetilde{x}^k) < f(x^{k-1})$, а соответствующий вектор $\widetilde{x}^k - x^{k-1}$ определяет направление спуска на текущем шаге алгоритма.

На рисунке 1c иллюстрируется основное преимущество метода – динамическое обновление ортогонального базиса в ходе итераций. В отличие от методов со статическими направлениями поиска, такая адаптивность позволяет более эффективно учитывать геометрию целевой функции. Особенно это проявляется на узких вытянутых участках, где оптимальный подбор направлений способствует сокращению числа итераций. Следует отметить, однако, что высокая адаптивность сопровождается увеличением вычислительной сложности метода.

Кроме метода Розенброка существует еще один алгоритм прямого поиска, называемый методом Пауэлла, направленный на адаптацию направлений движения к локальной структуре целевой функции. На начальных этапах алгоритм во многом аналогичен методу Розенброка: реализуется исследующий поиск в окрестности текущей точки, определяются направления убывания и соответствующие точки перехода. Однако ключевое отличие метода Пауэлла проявляется на следующем этапе формирования направлений спуска.

В отличие от статических схем и процедур ортонормирования, метод Пауэлла обновляет направления на основе линейных комбинаций ранее использованных векторов. Этот механизм позволяет учитывать накопленную информацию о поведении функции и более точно адаптироваться к ее локальной геометрии.

Как иллюстрируется на рисунке 1d, динамическое обновление направлений способствует ускорению сходимости, особенно на участках с выраженной кривизной. Адаптивность метода обеспечивает более эффективный выбор направлений движения без необходимости прямого использования градиентной информации. Однако данное преимущество сопровождается дополнительными вычислительными затратами, связанными с регулярным пересчетом и переопределением направлений.

Целью четвертого раздела магистерской работы является практическая проверка эффективности и особенностей работы изученных ранее алгоритмов нулевого порядка для поиска экстремума, а также их сравнительный анализ на примере тестовой функции. Для проведения сравнительного анализа методов прямого поиска была разработана программа на языке C#, реализующая программные модели методов циклического покоординатного спуска, Хука-Дживса, Розенброка и Пауэлла.

В качестве тестовой функции выбрана двумерная функция Розенброка:

$$f(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2, \quad x^* = (1, 1),$$

которая традиционно используется для оценки эффективности алгоритмов оптимизации. Функция Розенброка обладает узкой, извилистой долиной, ведущей к глобальному минимуму, и характеризуется выраженной нелинейностью. Это обеспечивает благоприятные условия для тестирования алгоритмов, особенно тех, которые не используют информацию о значениях производных, поскольку правильное направление поиска зачастую не совпадает с координатными осями, а рельеф функции затрудняет быстрое достижение экстремума.

В качестве критерия остановки используются оба условия из выражения (2), где заданы значения точности $\varepsilon_1 = \varepsilon_2 = 10^{-9}$. Начальное значение ускоряющего множителя в методах Хука-Дживса и Пауэлла установлено равным 1. Начальная длина шага во всех методах установлена равной 0.5 и в дальнейшем изменяется в соответствии с методом удвоения. Метод удвоения представляет собой стратегию адаптивного выбора шага, при которой шаг сначала увеличивается до тех пор, пока обеспечивается уменьшение функции. При прекращении убывания шаг уменьшается, если это не приводит к снижению значения функции, шаг дополнительно сокращается.

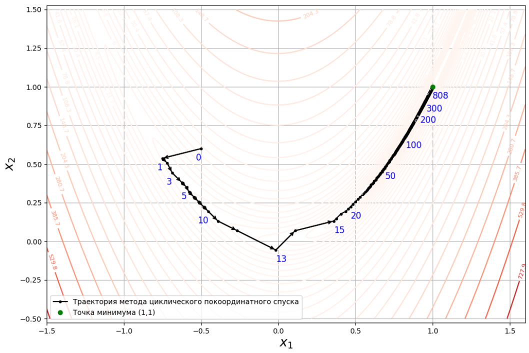
Для анализа поведения изученных методов нулевого порядка в условиях сложного рельефа целесообразно рассмотреть их работу при выборе начальной точки $x^0 = (-0.5, 0.6)$, расположенной вблизи узкой изогнутой долины функции Розенброка. Этот участок часто используется для тестирования алгоритмов из-за высокой кривизны и отклонения направления наискорейшего спуска от координатных осей. Анализ поведения алгоритмов на данном участке позволяет оценить их способность эффективно преодолевать овражную структуру рельефа и дает представление о характере формируемых траекторий в условиях искаженного пространства поиска.

Сравнительный анализ эффективности методов нулевого порядка в условиях овражного рельефа функции Розенброка, проведенный при стартовой точке $x^0 = (-0.5, 0.6)$, выявил существенные различия в работе рассматриваемых алгоритмов. Как показывают данные таблицы 1 и визуализация траекторий на рисунке 2, поведение методов существенно зависит от их способности адаптироваться к сложной геометрии целевой функции.

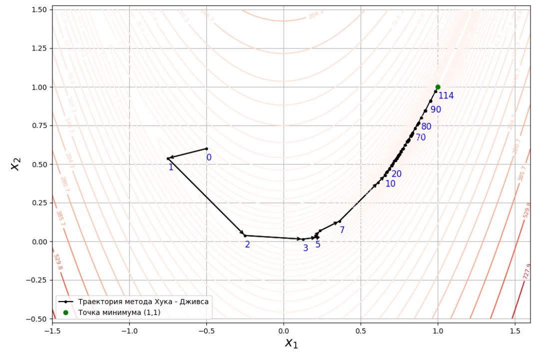
Таблица 1 – Сравнение методов нулевого порядка в овражной области функции

	Метод покоординатного спуска	Метод Хука-Дживса	Метод Розенброка	Метод Пауэлла
Колличество шагов	808	114	66	44
Колличество вызовов функции	5732	512	194	384

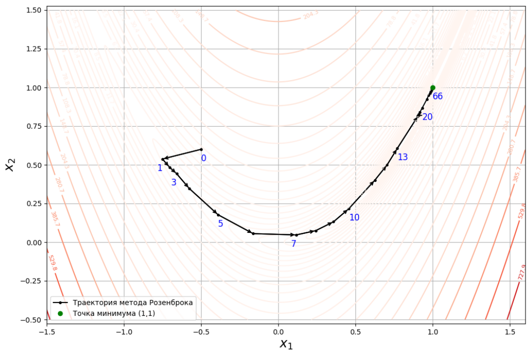
Сравнительный анализ методов нулевого порядка показал, что наименьшую эффективность продемонстрировал метод покоординатного спуска



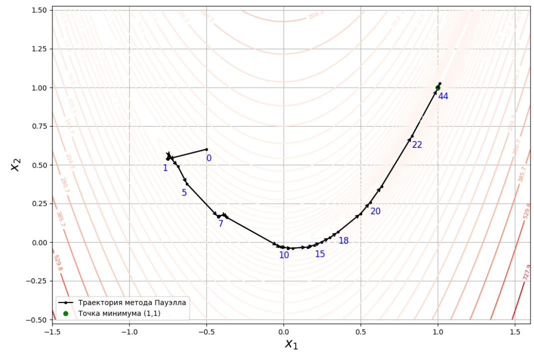
(a) Метод покоординатного спуска



(b) Метод Хука-Дживса



(c) Метод Розенброка



(d) Метод Пауэлла

Рисунок 2 – Визуализация траектории методов нулевого порядка для овражной области целевой функции

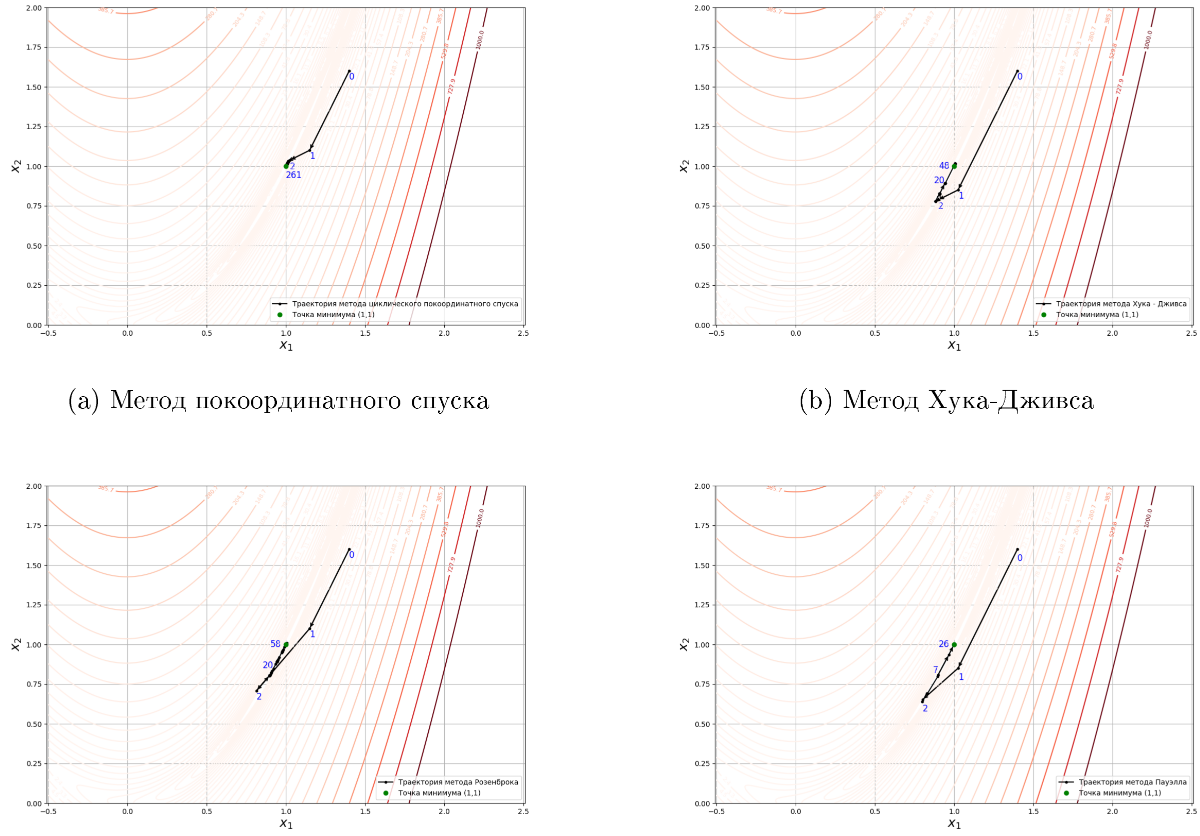
из-за игнорирования взаимосвязей между переменными и слабой приспособленности к криволинейному рельефу, что привело к высокой вычислительной нагрузке. Метод Хука–Дживса показал более высокую эффективность за счет чередования исследовательских и ускоряющих шагов. Наиболее эффективными оказались методы с адаптивным выбором направлений: метод Розенброка (лучший по числу вычислений) и метод Пауэлла (наилучший по числу итераций). Первый эффективно адаптируется к геометрии рельефа за счет обновления базиса, второй – благодаря построению сопряженных направлений. Таким образом, при оптимизации функций с овражной структурой целесообразно применять адаптивные алгоритмы, обеспечивающие лучшее сочетание сходимости и вычислительной эффективности.

Рассмотрим теперь поведение исследуемых алгоритмов в условиях, когда направление к экстремуму характеризуется меньшим отклонением от координатных осей. Выбор начальной точки $x^0 = (1.4, 1.6)$ позволяет проанализировать работу методов на относительно прямолинейном участке долины Розенброка. Согласно данным таблицы 2 и визуализации траекторий на рисунке 3, все методы демонстрируют улучшение показателей сходимости по сравнению с овражной областью, что обусловлено упрощением геометрической структуры рельефа. Однако сравнительный анализ выявляет существенные изменения в относительной эффективности алгоритмов по сравнению с овражной областью.

Таблица 2 – Сравнение методов нулевого порядка в прямолинейной области функции

	Метод покоординатного спуска	Метод Хука-Дживса	Метод Розенброка	Метод Пауэлла
Колличество шагов	261	48	58	26
Колличество вызовов функции	2148	234	176	278

На прямолинейном участке функции эффективность всех методов возрастает за счет упрощенной геометрии рельефа. Метод покоординатного спуска демонстрирует сокращение числа итераций (261 итерация против 808 в овражной области), сохраняя при этом наибольшее количество вычислений целевой функции. Метод Хука–Дживса показывает изменение относительной эффективности метода по сравнению с овражной областью - в данных условиях он превосходит метод Розенброка по количеству итераций (48 против 58), хотя и требует больше вычислений функции (234 против 176). Такое поведение объясняется повышенной эффективностью ускоряющих шагов на участках с малой кривизной, где направление спуска остается практически постоянным. Метод Розенброка остается наиболее экономичным по числу вычислений функции (176), демонстрируя при этом большее число итераций (58 шагов) по сравнению с методами Хука-Дживса и Пауэлла из-за необходимости настройки направлений. Метод Пауэлла вновь лидирует по скорости схо-



(a) Метод покоординатного спуска

(b) Метод Хука-Дживса

(c) Метод Розенброка

(d) Метод Пауэлла

Рисунок 3 – Визуализация траектории методов нулевого порядка для прямолинейной области целевой функции

димости (26 шагов), но его преимущество в вычислительной эффективности снижается. Сравнительный анализ результатов для овражного и прямолинейного участков функции Розенброка показывает, что методы с адаптивными направлениями сохраняют преимущество, однако при упрощении рельефа разница между алгоритмами уменьшается, и простые стратегии могут оказаться предпочтительнее.

Заключение

В данной работе была рассмотрена задача поиска экстремума целевой функции. Численные методы ее решения относятся к классу методов спуска и основаны на построении релаксационной последовательности. В ходе работы были подробно изучены основные понятия, связанные с релаксационными последовательностями. Кроме того, была изучена общая схема методов спуска и принципы, в соответствии с которыми осуществляется выбор направления и длины шага на каждом этапе итерационного процесса.

Среди методов спуска, применяемых для численного решения поставленной задачи, были рассмотрены алгоритмы покоординатного спуска, Хука—Дживса, Розенброка и Пауэлла. Все они относятся к методам нулевого порядка, то есть не используют производные целевой функции. Стратегия этих методов основана исключительно на информации о значениях функции, что делает их особенно полезными в ситуациях, когда аналитическое выражение целевой функции отсутствует, а имеется лишь возможность определить ее значение в любой точке рассматриваемой области с помощью расчетов.

Для оценки эффективности алгоритмов была разработана программа на языке C#, реализующая все рассматриваемые методы. В качестве тестовой функции использовалась функция Розенброка — классический пример функции с узкой извилистой долиной и высокой нелинейностью, усложняющей поиск экстремума. Визуализация траекторий оптимизации позволила проанализировать характер движения каждого алгоритма и их способность адаптироваться к различным участкам рельефа.

Проведенный сравнительный анализ показал, что методы с адаптивным выбором направлений (Розенброка и Пауэлла) обладают значительными преимуществами на рельефах с выраженной кривизной. Метод Пауэлла продемонстрировал наилучший баланс между числом итераций и вычислительными затратами, в то время как метод Розенброка оказался предпочтительнее при необходимости минимизировать количество вызовов целевой функции. Методы с фиксированными направлениями, такие как покоординатный спуск и Хука–Дживса, оказались менее эффективными в условиях сложного рельефа целевой функции.

На прямолинейных участках эффективность всех алгоритмов значительно возрастает, при этом относительное преимущество адаптивных методов снижается. Метод Розенброка продолжает демонстрировать низкие вычислительные затраты, но уступает методам Хука-Дживса и Пауэлла по числу итераций. Метод Пауэлла, несмотря на лидерство в скорости сходимости, уступает методам Хука-Дживса и Розенброка по общим вычислительным затратам. Эти результаты подчеркивают важность выбора метода с учетом локальных особенностей рельефа целевой функции.