

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 первого порядка в
конечномерных пространствах**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента 2 курса 217 группы

направление **01.04.02 – Прикладная математика и информатика**
механико-математического факультета

Захарова Марка Александровича

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. В работе исследуются численные методы поиска экстремума в конечномерных пространствах.

Данная работа посвящена методам первого порядка. Одним из самых известных таких методов современной оптимизации является градиентный метод. Данный метод лежит в основе многих численных методов оптимизации. Широкому распространению этого метода способствовала его сравнительная простота и возможность применения для минимизации широкого класса функций. Существует несколько вариантов этого метода, отличающиеся способом выбора длины шага: это градиентный метод с дроблением шага, градиентный метод с постоянным шагом, градиентный спуск с заранее заданным шагом, метод наискорейшего градиентного спуска. В градиентных методах с постоянным и заранее заданным шагом расстояние, проходимое на каждой итерации вдоль антиградиента мало, поэтому они сходятся медленно. В методе же наискорейшего градиентного спуска с помощью решения задачи одномерной минимизации, находится оптимальная длина шага, что ведет к более быстрой сходимости.

Сначала введем обозначения, напомним некоторые определения из линейной алгебры и математического анализа. Через $\mathbf{R}^n$ будем обозначать n -мерное вещественное линейное пространство, состоящее из вектор-столбцов

$$u = \begin{bmatrix} u^1 \\ \dots \\ u^n \end{bmatrix}, \quad v = \begin{bmatrix} v^1 \\ \dots \\ v^n \end{bmatrix}, \quad w = \begin{bmatrix} w^1 \\ \dots \\ w^n \end{bmatrix}$$

с действительными координатами $u^i, v^i, w^i, \dots (i = 1, \dots, n)$; сумма $u + v$ двух вектор-столбцов и произведение αu векторстолбца u на действительное число α в $\mathbf{R}^n$ определяется так:

$$u + v = \begin{bmatrix} u^1 + v^1 \\ \dots \\ u^n + v^n \end{bmatrix}, \quad \alpha u = \begin{bmatrix} \alpha u^1 \\ \dots \\ \alpha u^n \end{bmatrix};$$

вектор-столбец

$$0 = \begin{bmatrix} 0 \\ \dots \\ 0 \end{bmatrix}$$

называется нулевым. Вектор-строку, полученную транспонированием вектор-столбца u , обозначим через $u^T = (u^1, \dots, u^n)$. Там, где не могут возникнуть недоразумения, вектор-столбец или вектор-строку из $\mathbf{R}^n$ для краткости мы часто будем называть просто вектором или точкой, а знак транспонирования «*» будем опускать. Если в $\mathbf{R}^n$ ввести скалярное произведение двух векторов

$$\langle u, v \rangle = \sum_{i=1}^n u^i v^i, \quad u, v \in \mathbf{R}^n$$

то $\mathbf{R}^n$ превращается в n -мерное евклидово пространство, которое будем обозначать через E^n . Длина вектора или норма вектора в E^n определяется так:

$$|u| = \langle u, u \rangle^{1/2} = \left(\sum_{i=1}^n |u^i|^2 \right)^{1/2}$$

Величину

$$\rho(u, v) = |u - v| = \left(\sum_{i=1}^n |u^i - v^i|^2 \right)^{1/2}$$

называют евклидовым расстоянием между точками $u, v \in E^n$. Для любых точек $u, v, w \in E^n$ справедливо неравенство

$$|u - v| \leq |u - w| + |w - v|$$

называемое неравенством треугольника. Когда важно подчеркнуть, что скалярное произведение, норма, расстояние взяты именно в E^n , мы будем писать $\langle u, v \rangle_{E^n}$, $|u|_{E^n}$, $|u - v|_{E^n}$.

Основное содержание работы. Перейдем к постановке задачи минимизации. Пусть U - некоторое непустое множество из E^n , а $J(u)$ - функция, определенная на этом множестве. Всюду ниже, если не оговорено противное, мы будем рассматривать лишь функции, принимающие во всех точках

$u \in U$ конечные вещественные значения. Определения таких понятий, как точка минимума и максимума, наименьшее и наибольшее значение, ограниченность снизу и сверху, нижняя и верхняя грань функции $J(u)$ на множестве U , минимизирующая и максимизирующая последовательность, точка локального и строгого локального минимума и максимума функции, сходимости последовательности к заданному множеству в пространстве E^n получаются из определений, нужно лишь под u понимать точку $u = (u^1, \dots, u^n)$ из E^n , под $|u|$ - норму u в E^n . Поэтому здесь мы не будем воспроизводить определения перечисленных понятий. Примеры могут служить иллюстрацией к этим понятиям и в E^n , так как функция одной переменной является частным случаем функции n переменных.

Нижнюю грань функции $J(u)$ на множестве U по-прежнему будем обозначать через

$$\inf_U J(u) = J_*,$$

а множество точек минимума $J(u)$ на U - через

$$U_* = \{u : u \in U, \quad J(u) = J_*\}$$

Для обозначения задачи минимизации функции $J(u)$ на множестве U часто будем пользоваться следующей краткой стандартной записью:

$$J(u) \rightarrow \inf; \quad u \in U.$$

Будем различать задачи минимизации двух типов. В задачах первого типа ищется точное или приближенное значение величины J_* , и здесь неважно, будет ли множество U_* пустым или непустым. В задачах второго типа наряду с величиной J_* ищется точка $u \in U$, которая достаточно близка к множеству U_* или даже принадлежит U_* , - здесь естественно требовать, чтобы $J_* > -\infty, U_* \neq \emptyset$.

Для приближенного решения задач обоих типов на практике обычно строят какую-либо минимизирующую последовательность $\{u_k\}$:

$$u_k \in U, \quad k = 1, 2, \dots, \quad \lim_{k \rightarrow \infty} J(u_k) = J_*$$

(при $U_* \neq \emptyset$ возможно, например, $u_k = u_* \in U_*, k = 1, 2, \dots$). Тогда, как нетрудно видеть, в качестве приближения для J_* можно взять величину $J(u_k)$ при достаточно большом k . В том случае, если $\{u_k\}$ сходится к множеству U_* , т. е. $\rho(u_k, U_*) = \inf_{U_*} |u_k - u| \rightarrow 0$ при $k \rightarrow \infty$, точку u_k и соответствующее значение функции $J(u_k)$ при достаточно большом k можно принять за приближенное решение задачи второго типа. Однако, как мы видели в примере 1.1.5, условие $\lim_{k \rightarrow \infty} \rho(u_k, U_*) = 0$ имеет место не всегда. Поэтому в задачах второго типа построение минимизирующих последовательностей, сходящихся к U_* , в общем случае требует привлечения специальных методов.

В то же время имеются классы задач второго типа, у которых любая минимизирующая последовательность $\{u_h\}$ сходится к U_* . Эти классы задач хороши тем, что для их приближенного решения достаточно построить произвольную минимизирующую последовательность $\{u_k\}$ и затем пару $(u_k, J(u_k))$ при достаточно большом k принять за приближенное решение. Один такой класс задач будет описан ниже в теореме 1. Эту теорему будем называть теоремой Вейерштрасса, поскольку она является некоторым обобщением хорошо известной из учебников по математическому анализу теоремы Вейерштрасса о достижении нижней грани непрерывной функции на замкнутом ограниченном множестве из E^n [10, 160, 165, 233].

Перейдем к формулировке теоремы Вейерштрасса.

В первой главе рассматривается метод покоординатного спуска и его модификации.

Сначала опишем этот метод для задачи

$$J(u) \rightarrow \inf; \quad u \in U = E^n. \quad (3.1.1)$$

Обозначим $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ - единичный координатный (базис) вектор, у которого i -я координата равна 1, остальные равны нулю $i = 1, \dots, n$. Пусть u_0 - некоторое начальное приближение, а α_0 - некоторое положительное число, являющееся параметром метода. Допустим, что нам уже известны точка $u_k \in E^n$ и число $\alpha_k > 0$ при каком-либо $k \geq 0$. Положим

$$p_k = e_{i_k}, \quad i_k = k - n \left[\frac{k}{n} \right] + 1, \quad (3.1.2)$$

где $\left[\frac{k}{n} \right]$ означает целую часть числа k/n .

Условие (3.1.2) обеспечивает циклический перебор координатных векторов $e_1, e_2, \dots, e_n$, т. е.

$$p_0 = e_1, \dots, p_{n-1} = e_n, \quad p_n = e_1, \dots, p_{2n-1} = e_n, \quad p_{2n} = e_1, \dots$$

Вычислим значение функции $J(u)$ в точке $u = u_k + \alpha_k p_k$ и проверим неравенство

$$J(u_k + \alpha_k p_k) < J(u_k). \quad (3.1.3)$$

Если (3.1.3) выполняется, то примем

$$u_{k+1} = u_k + \alpha_k p_k, \quad \alpha_{k+1} = \alpha_k. \quad (3.1.4)$$

В том случае, если (3.1.3) не выполняется, то вычисляем значение функции $J(u)$ в точке $u = u_k - \alpha_k p_k$ и проверяем неравенство

$$J(u_k - \alpha_k p_k) < J(u_k) \quad (3.1.5)$$

В случае выполнения (3.1.5) положим

$$u_{k+1} = u_k - \alpha_k p_k, \quad \alpha_{k+1} = \alpha_k \quad (3.6)$$

Назовем $(k+1)$ -ю итерацию удачной, если справедливо хотя бы одно из неравенств (3.1.3) или (3.1.5). Если $(k+1)$ -я итерация неудачная, т. е. не выполняются оба неравенства (3.1.3) и (3.1.5), то полагаем

$$u_{k+1} = u_k, \quad \alpha_{k+1} = \begin{cases} \lambda \alpha_k, & i_k = n, u_k = u_{k-n+1} \\ \alpha_k, & i_k \neq n \text{ или } u_k \neq u_{k-n+1} \\ & \text{или } 0 \leq k \leq n-1 \end{cases} \quad (3.7)$$

здесь $\lambda(0 < \lambda < 1)$ - фиксированное число, являющееся параметром метода.

Условия (3.1.7) означают, что если за один цикл из n итераций при переборе направлений всех координатных осей $e_1, \dots, e_n$ с шагом α_k реализовалась хотя бы одна удачная итерация, то длина шага α_k не дробится и сохраняется на протяжении по крайней мере следующего цикла из n итераций. Если же среди последних n итераций не оказалось ни одной удачной итерации, то шаг α_k дробится. Таким образом, если на итерации с номером $k = k_m$ произошло дробление α_k , то

$$J(u_{k_m} + \alpha_{k_m} e_i) \geq J(u_{k_m}), \quad J(u_{k_m} - \alpha_{k_m} e_i) \geq J(u_{k_m}) \quad (3.1.8)$$

при всех $i = 1, 2, \dots, n$. Метод покоординатного спуска для задачи (3.1.1) описан.

Описанный выше метод покоординатного спуска нетрудно модифицировать применительно к задаче минимизации функции на параллелепипеде:

$$J(u) \rightarrow \inf; \quad u \in U = \{ (u^1, \dots, u^n) : a_i \leq u^i \leq b_i, i = 1, \dots, n \}, \quad (3.2.1)$$

где a_i, b_i - заданные числа, $a_i < b_i$ ($i = 1, \dots, n$). А именно, пусть k -е приближение $u_k \in U$ и число $\alpha_k > 0$ при некотором $k \geq 0$ уже найдены. Выберем вектор $p_k = e_{i_k}$ согласно формуле (2), составим точку $u_k + \alpha_k p_k$ и проверим условия

$$u_k + \alpha_k p_k \in U, \quad J(u_k + \alpha_k p_k) < J(u_k). \quad (3.2.2)$$

Если оба условия (3.2.2) выполняются, то следующее приближение u_{k+1}, α_{k+1} определяем по формулам (3.1.4). Если же хотя бы одно условие (3.2.2) не выполняется, то составляем точку $u_k - \alpha_k p_k$ и проверяем условия

$$u_k - \alpha_k p_k \in U, \quad J(u_k - \alpha_k p_k) < J(u_k). \quad (3.2.3)$$

В случае выполнения обоих условий (3.2.3) следующее приближение определяем по формулам (3.1.6), а если хотя бы одно из условий (3.2.3) не выполняется, то следующее приближение находится из неравенств (3.1.7).

Существуют также и другие варианты метода покоординатного спуска.

Можно, например, строить последовательность $\{u_k\}$ по правилу

$$u_{k+1} = u_k + \alpha_k p_k, \quad (3.2.4)$$

где p_k определяется согласно (3.1.2), а α_k - условиями

$$\alpha_k \geq 0, \quad f_k(\alpha_k) = \min_{-\infty < \alpha < +\infty} f_k(\alpha), \quad f_k(\alpha) = J(u_k + \alpha p_k). \quad (3.2.5)$$

Метод (3.2.4), (3.2.5) имеет смысл применять в том случае, когда величина α_k из (3.13) находится в явном виде. Так будет, если функция $J(u)$ - квадратичная, т. е.

$$J(u) = \frac{1}{2} \langle Au, u \rangle - \langle b, u \rangle, \quad u \in E^n, \quad (3.2.6)$$

где A - симметричная положительно определенная матрица $b \in E^n$. Нетрудно убедиться, что для функции (3.2.6) метод (3.2.3), (3.2.4) приводит к хорошо известному методу Зейделя из линейной алгебры.

Хотя и скорость сходимости метода покоординатного спуска, вообще говоря, невысокая, благодаря простоте каждой итерации, скромным требованиям к гладкости минимизируемой функции этот метод довольно широко применяется на практике.

Существуют и другие методы минимизации, использующие лишь значения функции и не требующие для своей реализации вычисления производных. Например, используя вместо производных их разностные аппроксимации, можно построить модификации рассматривавшихся в предыдущих параграфах методов, требующие вычисления лишь значений функции в подходящим образом выбранных точках.

Во втором разделе будут рассматриваться радиентные методы поиска экстремума функции. Для наглядности рассматривать будем метод скорейшего спуска. Поставим задачу и определим метод.

Будем рассматривать задачу

$$J(u) \rightarrow \inf; \quad u \in U \equiv E^n, \quad (4.1.1)$$

предполагая, что функция $J(u)$ непрерывно дифференцируема на E^n , т. е. $J(u) \in C^1(E^n)$. Согласно дифференцируемой функции

$$J(u+h) - J(u) = \langle J'(u), h \rangle + o(h; u), \quad (4.1.2)$$

где $\lim_{|h| \rightarrow 0} o(h; u)|h|^{-1} = 0$. Если $J'(u) \neq 0$, то при достаточно малых $|h|$ главная часть приращения (4.1.2) будет определяться дифференциалом функции $dJ(u) = \langle J'(u), h \rangle$. Справедливо неравенство Коши - Буняковского

$$-|J'(u)| \cdot |h| \leq \langle J'(u), h \rangle \leq |J'(u)| \cdot |h|,$$

$$u_{k+1} = u_k - \alpha_k J'(u_k), \quad \alpha_k > 0, \quad k = 0, 1, \dots \quad (4.1.3)$$

Число α_k из (3) часто называют длиной шага или просто шагом градиентного метода. Если $J'(u_k) \neq 0$, то шаг $\alpha_k > 0$ можно выбрать так, чтобы $J(u_{k+1}) < J(u_k)$. В самом деле, из равенства (4.1.2) имеем

$$J(u_{k+1}) - J(u_k) = \alpha_k \left[-|J'(u_k)|^2 + o(\alpha_k) \alpha_k^{-1} \right] < 0$$

Существуют различные способы выбора величины α_k в методе (4.1.3). В зависимости от способа выбора α_k можно получить различные варианты градиентного метода. Укажем несколько наиболее употребительных на практике способов выбора α_k .

1) На луче $\{u \in E^n : u = u_k - \alpha J'(u_k), \alpha \geq 0\}$, направленном по антиградиенту, введем функцию одной переменной

$$f_k(\alpha) = J(u_k - \alpha J'(u_k)), \quad \alpha \geq 0,$$

и определим α_k из условий

$$f_k(\alpha_k) = \inf_{\alpha \geq 0} f_k(\alpha) = f_{k*}, \quad \alpha_k > 0. \quad (4.1.4)$$

Метод (4.1.3), (4.1.4) принято называть методом скорейшего спуска. При $J'(u_k) \neq 0$ согласно формуле (2.3.1) $f'_k(0) = -|J'(u_k)|^2 < 0$, поэтому

нижняя грань в (4.1.4) может достигаться лишь при $\alpha_k > 0$.

Далее рассматриваются четыре метода: Рассматривается такая же задача (5.1.1), предполагая, что функция $J(u)$ непрерывно дифференцируема на E^n

Если функция $J(u) \in C^{1,1}(E^n)$, т. е. $J(u) \in C^1(E^n)$, и градиент $J'(u)$ удовлетворяет условию

$$|J'(u) - J'(v)| \leq L|u - v|, \quad u, v \in E^n$$

причем константа L известна, то в (5.1.3) в качестве α_k может быть взято любое число, удовлетворяющее условиям

$$0 < \varepsilon_0 \leq \alpha_k \leq 2/(L + 2\varepsilon), \quad (5.2.1)$$

где $\varepsilon_0, \varepsilon$ - положительные числа, являющиеся параметрами метода. В частности, при $\varepsilon = L/2, \varepsilon_0 = 1/L$ получим метод (5.1.3) с постоянным шагом $\alpha_k = 1/L$. Отсюда ясно, что если константа L большая или получена с помощью слишком грубых оценок, то шаг α_k в (4.1.3) будет маленьким.

Далее рассмотрим методы релаксации, суть которых заключается в том, что последующий шаг итерации не превышает предыдущий, т.е. $U_{k_{j+1}} \leq U_{k_j}$.

Существуют достаточно простые классы функций, которыми можно аппроксимировать минимизируемую функцию и для которых задача минимизации легко решается. Выбирая, например, в качестве такого аппроксимирующего класса квадратичные функции, график которых представляет собой параболу, мы придем к методу минимизации, который естественно назвать методом парабол.

В последней главе рассмотрим численные реализации методов поиска локального экстремума. Для наглядности будем рассматривать градиентные методы, а точнее, метод наискорейшего спуска и его модификацию. В данном разделе рассматривается применение градиентного метода для нахождения минимума заданной функции $f(x)$ при заданных начальной точке, параметрах $\alpha, \lambda, \varepsilon$.

Здесь и далее обозначим

$$x := (x_1, x_2), f(x) := f(x_1, x_2),$$

$$x_k := (x_{1_k}, x_{2_k}), f(x_k) := f(x_{1_k}, x_{2_k}).$$

Рассмотрим функцию $f(x)$ вида

$$f(x) = f(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2 + \left(\frac{3}{4}x_1^3 - x_2 + 0,9\right)^2. \quad (6.1.1)$$

Требуется найти точку x_* , которая является минимумом $f(x)$ на множестве R^2 . Для нахождения x_* воспользуемся градиентным методом.

По градиентному методу требуется построить итерационный процесс вида (6.1.1).

Градиент рассматриваемой функции $f'(x)$ имеет вид

$$f'(x) = \left(\frac{\partial f(x_1, x_2)}{\partial x_1}, \frac{\partial f(x_1, x_2)}{\partial x_2} \right), \quad (6.1.2)$$

где

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 4x_1^3 + 4x_1 \cdot x_2^2 - 4x_1 + \frac{9}{16} \cdot 6x_1^5 - \frac{9}{2}x_1^2x_2 + \frac{9}{2} \cdot 0,9x_1^2,$$

$$\frac{\partial f(x_1, x_2)}{\partial x_2} = 4x_1^2 \cdot x_2 + 4x_2^3 - 4x_2 - \frac{3}{2}x_1^3 + 2x_2 - 2 \cdot 0,9$$

Таким образом получаем итерационный процесс вида

$$x_{k+1} = x - \alpha_k f'(x_k), \alpha_k > 0, k = 0, 1, \dots \quad (6.1.3)$$

где $f'(x_k)$ вычисляется по (6.1.2), $\alpha_k = \alpha \cdot \lambda^k$. Так как $f(x) \in R, x \in R^2$, будем использовать условие в виде

$$|f(x_{k+1}) - f(x_k)| \leq \varepsilon \alpha_k \|f'(x_k)\|^2. \quad (6.1.4)$$

В итоге имеем следующий алгоритм:

1. Задаём начальное приближение, иначе x_0 , и значения параметров $\alpha, \lambda, \varepsilon$.

2. Вычисляем x_{k+1} по (6.1.2), пока не выполнится (6.1.3).

3. Когда выполнилось (6.1.3), фиксируем полученную x_{k+1} . Это и есть искомая точка.

Будем рассматривать функцию $f(x)$ вида (6.1.1), градиент которой имеет вид (6.1.2). Требуется найти точку x_m , которая будет минимумом функции $f(x)$ вида (6.1.1) на множестве R^2 . Воспользуемся упомянутым в предыдущем разделе методом для отыскания x_m .

Возьмём

$$\hat{x}_{k+1} = x_k - \alpha f'(x_k) \quad (6.2.1)$$

где $\alpha > 0$ задаётся с клавиатуры. Искомые x_{k+1} будем вычислять по

$$x_{k+1} = (1 - \xi_M) x_k + \xi_M \hat{x}_{k+1} \quad (6.2.2)$$

где

$$\xi_M = -\frac{(f'(x_k), \hat{x}_{k+1} - x_k)}{2 \cdot (f(\hat{x}_{k+1}) - f(x_k) - (f'(x_k), \hat{x}_{k+1} - x_k))} \quad (6.2.3)$$

Помним, что

$$x_k = (x_{k,1}, x_{k,2})$$

По (6.2.3) строим приближения до тех пор, пока не выполнится условие

$$(|x_{k+1,1} - x_{k,1}| < \varepsilon) \&\& (|x_{k+1,2} - x_{k,2}| < \varepsilon). \quad (6.2.4)$$

Когда (6.2.4) выполнилось, фиксируем полученную x_{k+1} . Это и будет точка x_M - искомый минимум функции $f(x)$ на R^2 .

На языке Python разработана программа, в которой реализован указанный выше итерационный процесс. Через консоль осуществляется ввод входных данных. Результаты работы программы представлены на рисунках ниже. Зелёным цветом показан график значений функции $f(x)$, синим цветом показаны x_k , красным цветом показано значение $f(x)$ в найденной точке минимума. Будем брать $\lambda = 0,95$. В качестве начальной точки возьмём $x_0 := (-0,5; -0,5)$.

Получим минимум $f(x)$ путём подбора параметров ε и α

На рисунках 6.3.5 - 6.3.6 продемонстрировано полученное за 540 ите-

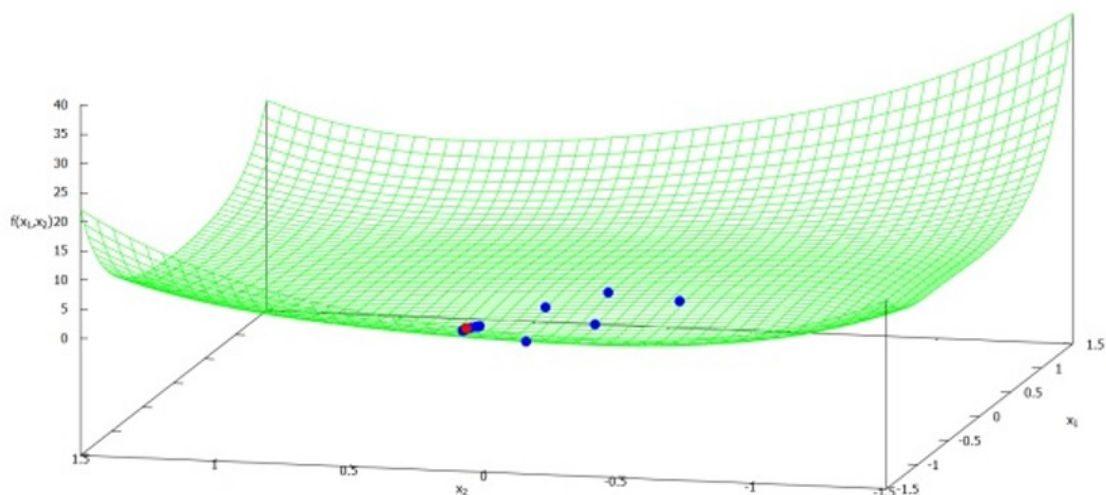


Рисунок 6.3.5 $\varepsilon = 0.0001, \alpha = 0.145$

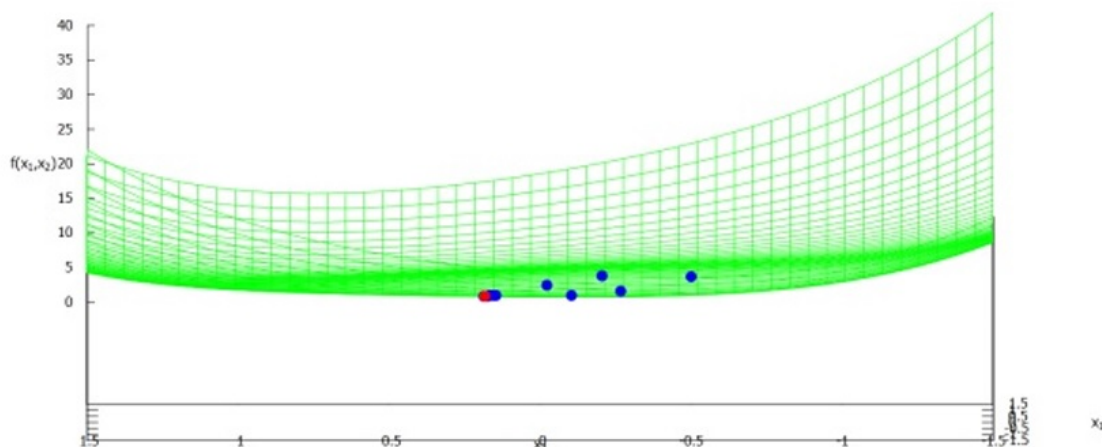


Рисунок 6.3.6 $\varepsilon = 0.0001, \alpha = 0.145$

раций приближенное решение $x^* = (-0,9846633; 0,1818098)$. Это решение является ближайшим к минимуму $f(x)$ на рассматриваемом множестве $[-1,5; 1,5] \times [-1,5; 1,5]$.

$$f(x^*) = 1,1560956324190181E - 5.$$

Далее результаты работы программы модификации метода.

Через консоль осуществляется ввод координат стартовой точки x_k и параметра α . Будем брать $\varepsilon = 10^{-6}$. Результаты работы программы представлены на рисунках ниже. Зелёным цветом показан график функции $f(x)$, оранжевым цветом показаны значения x_k , черным цветом показано значение $f(x)$ в найденной точке минимума функции $f(x)$.

В качестве начальной точки также возьмём $x_0 = (-0,5; -0,5)$. Подбором параметра α получим значение x_M . На рисунках 6.3.7 - 6.3.8 показано по-

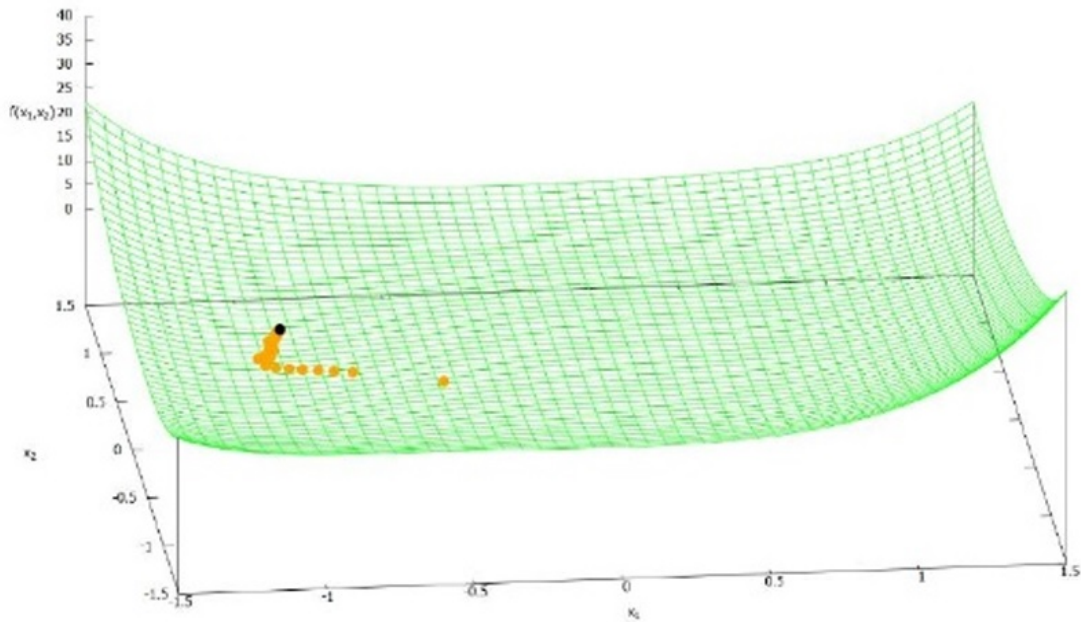


Рисунок 6.3.7 $\varepsilon = 10^{-6}, \alpha = 0.1$

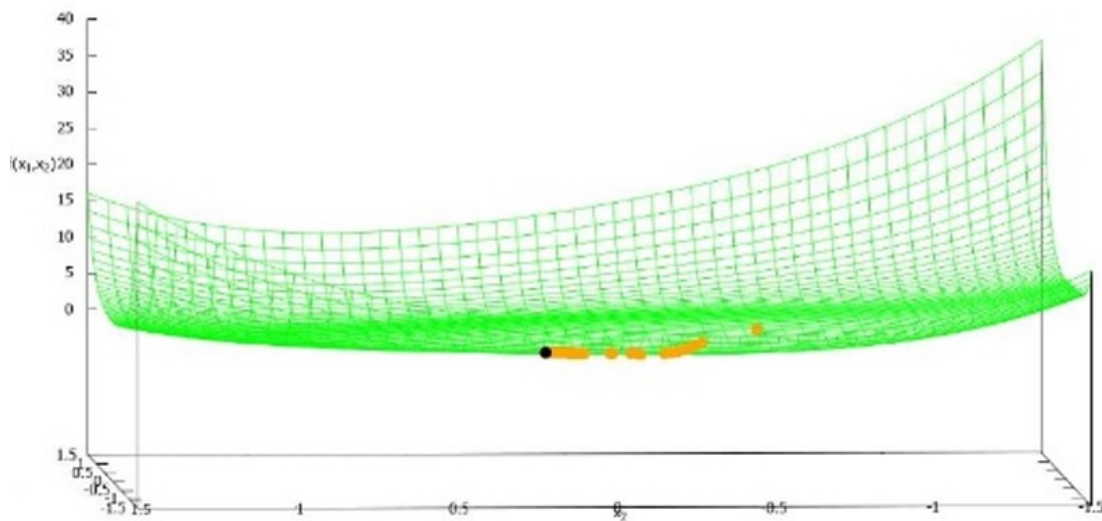


Рисунок 6.3.8 $\varepsilon = 10^{-6}, \alpha = 0.1$

лученное за 62 итераций приближенное решение $x_M = (-0.98125; 0.192469)$, а значение функции в этой точке $f(x_M) = 1.2395905E - 13$.

Как видно на рисунках 6.3.5 - 6.3.6 и при сравнении полученных значений функций среди наилучших результатов градиентного спуска и моди-

фикации градиентного метода, модификация градиентного метода позволяет найти лучшее решение за меньшее количество итераций и при меньшем количестве параметров, от которых зависят как скорость нахождения приближённого решения, так и близость к точному решению найденного

Заключение. В данной работе была проведена всесторонняя исследовательская работа по методам поиска экстремума первого порядка в конечно-мерных пространствах. Мы рассмотрели основные определения и принципы, лежащие в основе задач минимизации, и проанализировали важные теоремы, такие как теорема Вейерштрасса, которая служит теоретической основой для существования оптимальных решений. В заключение, можно отметить, что методы поиска экстремума первого порядка остаются актуальной и активно развивающейся областью исследования. Будущие работы могут быть направлены на дальнейшую оптимизацию существующих алгоритмов, а также на разработку новых подходов, учитывающих особенности конкретных задач. Применение этих методов в различных областях, таких как экономика, инженерия и машинное обучение, подчеркивает их важность и универсальность.