

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»**

Кафедра математической физики и вычислительной математики

**О методах 1-го порядка поиска
экстремума в конечномерных пространствах**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента 2 курса 217 группы

направление 01.04.02 — Прикладная математика и информатика

механико-математического факультета

Кашкалда Владислава Денисовича

Научный руководитель

Профессор, д. ф.-м. н., профессор

В.А. Юрко

Зав. кафедрой

Профессор, д. ф.-м. н., профессор

В.А. Юрко

Саратов 2022

Введение. В работе исследуются численные методы поиска экстремума в конечномерных пространствах. Данная задача часто встречается как в математике так и в различных областях естествознания для нахождения как точных решений, если это возможно, так и приближенных решений. Численные методы поиска экстремума делятся на: 1) нулевого порядка, в них используется информация только о значениях минимизируемой функции; 2) первого порядка, использующие информацию также и о значениях первых производных; 3) второго порядка, использующие, кроме того, информацию о вторых производных и так далее. Известно, что чем больше информации об исследуемой функции в процессе минимизации используется, тем быстрее достигается результат. Таким образом, скорость сходимости процесса минимизации увеличивается, но попутно возрастает количество необходимых вычислений. Так, медленнее всего сходятся методы нулевого порядка, далее идут методы первого, а затем второго порядка. 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. На основе градиентного метода была рассмотрена модификация градиентного метода, проведен анализ результатов и выявлены преимущества модификации градиентного метода над градиентным методом. **Основное содержание работы. Постановка задачи.** Пусть задана выпуклая непрерывно дифференцируемая функция $f(x)$, $x \in E^n$, определенная при всех $x \in E^n$, и выпуклое множество X . Требуется найти минимум $f(x)$ на множестве X , т.е. найти такую точку

x_* , что $f(x_*) \leq f(x)$, $x \in X$. Для решения задачи используются итерационные процессы типа: $x_{k+1} = x_k + d_k p_k$, где p_k - вектор, определяющий направление движения из точки x_k , d_k - шаг в направлении p_k . Будем использовать следующие оценки скорости сходимости. **Обозначения.** Говорят, что последовательность $\{x_k\}$ сходится к точке x_* с линейной скоростью или со скоростью геометрической прогрессии (со знаменателем q), если, начиная с некоторого k , выполняется неравенство: $\|x_{k+1} - x_k\| \leq q \|x_k - x_*\|$, $0 < q < 1$. При выполнении неравенства $\|x_{k+1} - x_k\| \leq q_k \|x_k - x_*\|$, $q_k \rightarrow 0$ при $k \rightarrow \infty$, говорят, что скорость сходимости последовательности $\{x_k\}$ сверхлинейная, или выше скорости сходимости любой геометрической прогрессии. Если $q_k \leq C \|x_k - x_*\| \rightarrow 0$, то будет $\|x_{k+1} - x_*\| \leq C \|x_k - x_*\|^2$. Эта оценка характеризует квадратичную скорость сходимости. **Основные определения.** Введем векторно-матричные обозначения. Будем рассматривать n -мерное пространство векторов, обозначаемое через E^n . Сами векторы будем обозначать латинскими буквами x, y и т.п., а их компоненты - индексами сверху, так что x^i - это i -я компонента вектора x . Нижними индексами обозначим элементы некоторой последовательности.

Матрицы обозначим большими буквами A, B и т.п. Звездочка сверху при этом будет означать транспонирование, т.е. A^* - это транспонированная матрица A . Под вектором x будем понимать вектор-столбец, так что x^* обозначает вектор-строку. Скалярное произведение двух векторов обозначим как (x, y) , т.е. $(x, y) = \sum_{j=1}^n x^j y^j$. При этом под нормой вектора будем понимать его евклидову норму: $\|x\| = \sqrt{(x, x)}$; $f(x)$ - минимизируемая функция n переменных; $f'(x)$ - ее градиент, т.е. вектор с координатами $\frac{\partial f}{\partial x^1}, \dots, \frac{\partial f}{\partial x^n}$. $f''(x)$ - матрица $n \times n$ с элементами $\frac{\partial^2 f}{\partial x^i \partial x^j}$. M и m - соответственно наибольшее и наименьшее собственные значения матрицы $f''(x)$.

Через $y = Ax$ обозначим умножение матрицы A на вектор x , т.е. $y^i = \sum_{j=1}^n a^{ij} x^j$, $i = 1, 2, \dots, n$. Греческие буквы α, β и т.п. будут применяться для обозначения числовых величин. Пусть x_* - точка минимума функции $f(x)$; $f(x_*) = f^*$. Для краткости обозначим $f'_k = f'(x_k)$, $f''_{kc} = f''(x_{kc})$, $x_{kc} = x_k + \theta(x - x_k)$, $\theta \in [0, 1]$.

Градиентный метод: численное нахождение минимума функции

Для того, чтобы приблизиться к точке x_* (в общем случае x_* - точка, где выполняются - возможно, с определенной точностью - необходимые условия экстремума функции $f(x)$), естественно двигаться из точки x_k в направлении убывания функции, т.е. в направлении спуска.

Итерационный процесс

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

получающийся при таком выборе направления движения, называется градиентным методом. В координатной форме процесс (1) записывается следующим образом:

$$x_{k+1}^i = x_k^i - \alpha_k \frac{\partial f(x_k)}{\partial x^i}, \quad i = 1, 2, \dots, n.$$

В настоящее время градиентный метод является одним из наиболее известных методов минимизации. Широкому распространению этого метода способствовала его сравнительная простота и возможность применения для минимизации весьма широкого класса функций.

Изучим свойства алгоритма (1).

Шаг α_k выбирается методом наискорейшего спуска.

- 1) Выбираем некоторое произвольное значение α (одно и то же на всех итерациях) и определяем точку $x = x_k - \alpha f'_k$.
- 2) Вычисляем $f(x) = f(x_k - \alpha f'_k)$.
- 3) Производим проверку неравенства

$$f(x) - f(x_k) \leq \varepsilon \alpha (f'_{k+1} - f'_k), \quad (2)$$

где $0 < \varepsilon < 1$ - произвольно выбранная константа.

- 4) Если неравенство (2) выполнится, то значение α и берем в качестве искомого: $\alpha_k = \alpha$. Если же неравенство не выполнится, производим дробление α (путем умножения α на произвольное число $\lambda < 1$) до тех пор, пока неравенство (2) не окажется справедливым.

Этот способ выбора α_k нуждается в обосновании: необходимо выяснить, при каких условиях существуют ненулевые значения параметра α , при

которых неравенство (2) выполняется. Такое обоснование способа выбора α_k приводится в следующей теореме.

Теорема 1. Если функция $f(x)$ ограничена снизу, ее градиент $f'(x)$ удовлетворяет условию Липшица:

$$\| f'(x) - f'(y) \| \leq R \| x - y \| \quad (3)$$

при любых $x, y \in E^n$, а выбор значения α_k производится описанным выше способом, то для процесса (1) будет $\| f'_k \| \rightarrow 0$ при $k \rightarrow \infty$, какова бы ни была начальная точка x_0 .

Лемма 1. Выпуклая непрерывная функция $f(x)$ достигает своего минимума на компактном выпуклом множестве X .

Лемма 2. Пусть множество X замкнуто, а $f(x)$ - дважды непрерывно дифференцируемая сильно выпуклая функция. Тогда $f(x)$ достигает своего минимума на X .

Рассмотрим теперь пересечение $X \cap Y$. Очевидно, что если x_* - точка минимума $f(x)$ на множестве $X \cap Y$, то эта точка будет и точкой минимума $f(x)$ на X . Но множество $X \cap Y$ ограничено и замкнуто как пересечение двух замкнутых множеств, одно из которых ограничено. Поэтому $f(x)$ достигает своего минимума на $X \cap Y$, а значит, и на всем X .

Лемма 3. Строго выпуклая функция достигает своего минимума на выпуклом множестве X в единственной точке.

Теорема 2. Пусть $f(x)$ - дважды непрерывно дифференцируемая функция, причем ее матрица вторых производных удовлетворяет условиям

$$m \| y \|^2 \leq (f''(x)y, y) \leq M \| y \|^2, M \geq m > 0, \quad (4)$$

при любых $x, y \in E^n$, а последовательность $\{x_k\}$ строится по формуле (1), где α_k выбирается методом наискорейшего спуска. Тогда при любой начальной точке x_0 будет $x_k \rightarrow x_*$, $f(x_k) \rightarrow f(x_*)$, где x_* - точка минимума (единственная) $f(x)$.

Для скорости сходимости справедливы оценки:

$$\begin{aligned} f_k - f_* &\leq q^k (f_0 - f_*), \\ \|x_k - x_*\| &\leq cq^{\frac{k}{2}}, \\ c < \infty, 0 < q < 1. \end{aligned} \tag{5}$$

Неравенство (2) заведомо выполняется при значениях $\alpha \leq \frac{1-\varepsilon}{R}$ (теорема 1), либо при $\alpha \leq \frac{2(1-\varepsilon)}{M}$ (теорема 2). Если известны константы R или M , характеризующие минимизируемую функцию $f(x)$, то в методе (1) можно заранее выбрать $\alpha_k \equiv \bar{\alpha}$, где $0 < \bar{\alpha} \leq \frac{1-\varepsilon}{R}$, либо $0 < \bar{\alpha} \leq \frac{2(1-\varepsilon)}{M}$, и при этом теоремы 1 и 2 сохраняют силу. В таком варианте градиентного метода можно уточнить величину знаменателя q в оценках скорости сходимости (5).

Теорема 3. Если функция $f(x)$ удовлетворяет условиям теоремы 2 и в методе (1) $\alpha_k = \alpha$, $0 < \alpha < \frac{2}{M}$, то для скорости сходимости последовательности $\{x_k\}$ справедлива оценка

$$\begin{aligned} \|x_k - x_*\| &\leq q^k \|x_0 - x_*\|, \\ q &= \max\{|1 - \alpha m|, |1 - \alpha M|\}, \end{aligned}$$

причем минимальное значение $q_{\min} = \frac{M-m}{M+m}$ достигается при $\alpha = \frac{2}{M+m}$.

Отметим, что при $\alpha = \frac{2}{M+m}$ первая из оценок (5) уточняется следующим образом:

$$f_{k+1} - f_* \leq \left(\frac{M-m}{M+m}\right)^2 (f_k - f_*). \tag{6}$$

Численное нахождение минимума функции градиентным методом

В данном разделе рассматривается применение градиентного метода для нахождения минимума заданной функции $f(x)$ при заданных начальной точке, параметрах α , λ , ε .

Здесь и далее обозначим

$$\begin{aligned} x &:= (x_1, x_2), f(x) := f(x_1, x_2), \\ x_k &:= (x_{1_k}, x_{2_k}), f(x_k) := f(x_{1_k}, x_{2_k}). \end{aligned}$$

Рассмотрим функцию $f(x)$ вида

$$f(x) = f(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2 + \left(\frac{3}{4}x_1^3 - x_2 + 0,9\right)^2. \quad (7)$$

Требуется найти точку x_* , которая является минимумом $f(x)$ вида (7) на множестве R^2 . Для нахождения x_* воспользуемся градиентным методом.

По градиентному методу требуется построить итерационный процесс вида (1).

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

$$f'(x) = \left(\frac{\partial f(x_1, x_2)}{\partial x_1}, \frac{\partial f(x_1, x_2)}{\partial x_2} \right), \quad (8)$$

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

$$x_{k+1} = x - \alpha_k f'(x_k), \alpha_k > 0, k = 0, 1, \dots, \quad (9)$$

где $f'(x_k)$ вычисляется по (8), $\alpha_k = \alpha \cdot \lambda^k$.

Так как $f(x) \in R$, $x \in R^2$, заменяя условие (2), которое будем использовать в виде

$$|f(x_{k+1}) - f(x_k)| \leq \varepsilon \alpha_k \|f'(x_k)\|^2. \quad (10)$$

В итоге имеем следующий алгоритм:

1. Задаём начальное приближение, иначе x_0 , и значения параметров α , λ , ε .
2. Вычисляем x_{k+1} по (9), пока не выполнится (10).
3. Когда выполнилось (10), фиксируем полученную x_{k+1} . Это и есть искомая точка.

На языке Java 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 .

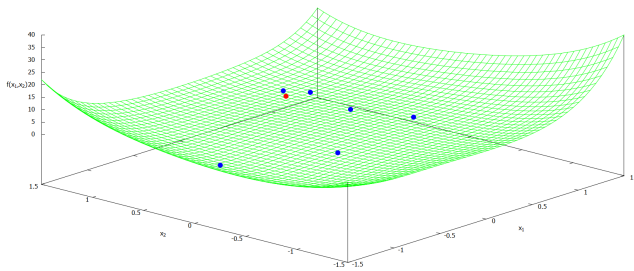


Рисунок 1 — $\varepsilon = 0,1$; $\alpha = 0,3$.

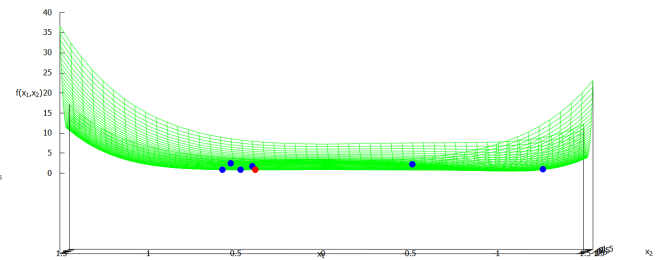


Рисунок 2 — $\varepsilon = 0,1$; $\alpha = 0,3$.

На рисунках 1 — 2 показано полученное за 6 итераций приближенное решение $x_* = (0,422975; 0,83745425)$, $f(x_*) = 0,028575675162065883$. Как видно на рисунках, полученное приближенное решение далеко от настоящего минимума.

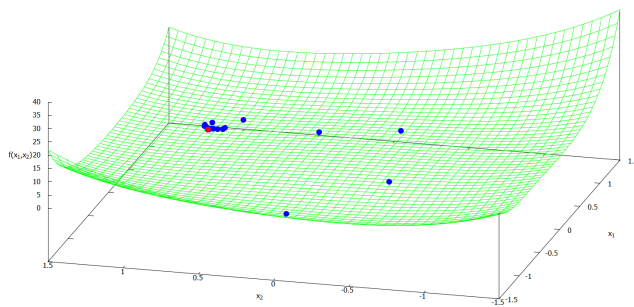


Рисунок 3 — $\varepsilon = 0,001$; $\alpha = 0,3$.

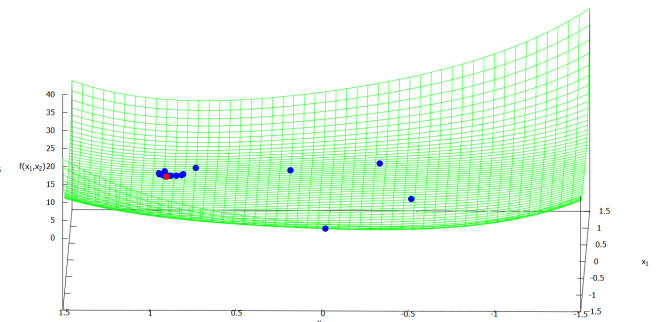


Рисунок 4 — $\varepsilon = 0,001$; $\alpha = 0,3$.

На рисунках 3 — 4 показано полученное за 568 итераций приближенное решение $x_* = (0,359696; 0,933426887)$, $f(x_*) = 2,625845738691603E-6$. Как видно на рисунках, полученное приближенное решение располагается не близко к настоящему минимуму на рассматриваемом множестве.

На рисунках 5 — 6 показано полученное за 599 итераций приближенное решение $x_* = (0,322289; 0,94224)$. $f(x_*) = 3,5159373008077793E-4$.

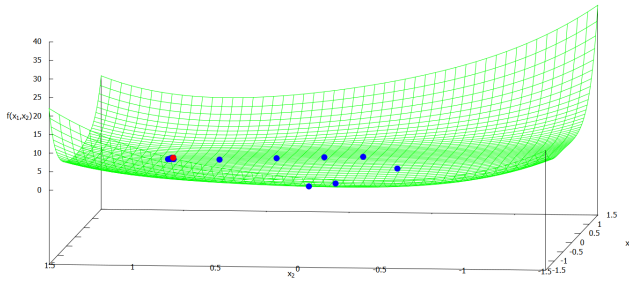


Рисунок 5 — $\varepsilon = 0,01$; $\alpha = 0,2$.

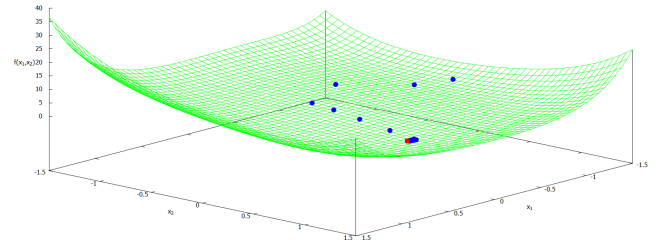


Рисунок 6 — $\varepsilon = 0,01$; $\alpha = 0,2$.

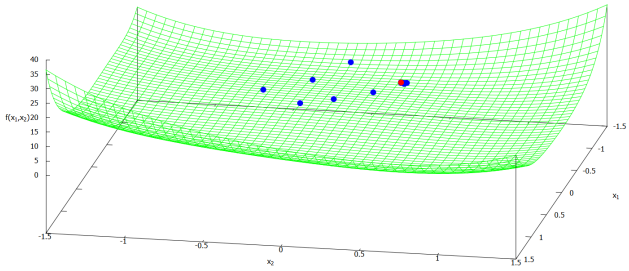


Рисунок 7 — $\varepsilon = 0,01$; $\alpha = 0,15$.

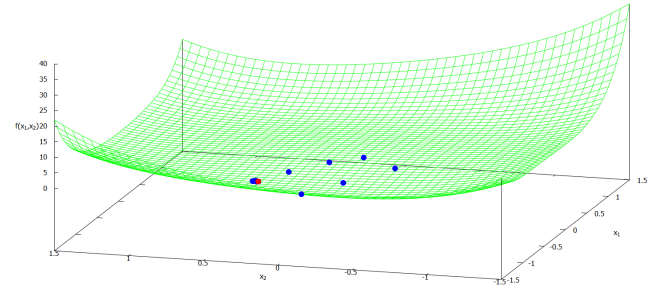


Рисунок 8 — $\varepsilon = 0,01$; $\alpha = 0,15$.

На рисунках 7 — 8 показано приближенное решение $x_* = (-0,9437684; 0,2912)$, полученное за 603 итерации. $f(x_*) = 0,0010688685961931649$.

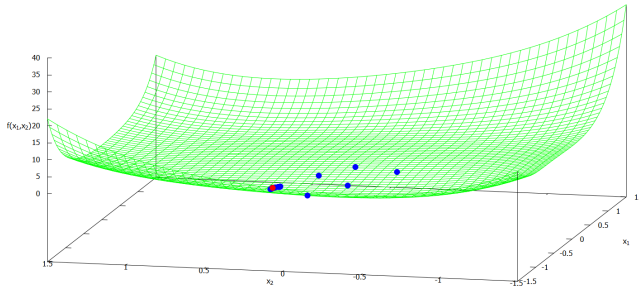


Рисунок 9 — $\varepsilon = 0,0001$; $\alpha = 0,145$.

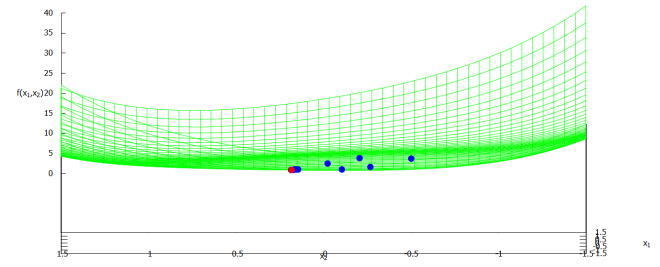


Рисунок 10 — $\varepsilon = 0,0001$; $\alpha = 0,145$.

На рисунках 9 — 10 продемонстрировано полученное за 540 итераций приближенное решение $x_* = (-0,9846633; 0,1818098)$. Это решение является ближайшим к минимуму $f(x)$ на рассматриваемом множестве $[-1,5; 1,5] \times [-1,5; 1,5]$.

$$f(x_*) = 1,1560956324190181\text{E-}5.$$

Лучший результат был получен при $\alpha = 0,145$, $\varepsilon = 0,0001$, $\lambda = 0,95$.

Модификация градиентного метода: численное нахождение минимума функции

Рассмотрим вариант метода наискорейшего спуска для минимизации выпуклой функции. Алгоритм обладает достаточно хорошей скоростью сходимости и не требует специального способа выбора шага в направлении спуска.

1) Определяется точка

$$\hat{x}_{k+1} = x_k - \alpha \frac{\partial f(x_k)}{\partial x_k}, \quad \alpha > 0, \quad x_k \in E^n, \quad k = 0, 1, 2, \dots \quad (11)$$

α - некоторое произвольно выбранное значение шага в направлении спуска.

2) Вычисляется $f(\hat{x}_{k+1}) = f(x_k - \alpha f'_k)$.

3) Вводится функция $\varphi(\xi)$ следующим образом:

$$\begin{aligned} \varphi(\xi) &= A\xi^2 + B\xi + C, \\ \varphi(0) &= 0, \\ \varphi(1) &= f(\hat{x}_{k+1}) - f(x_k), \\ \varphi'(0) &= (f'(x_k), \hat{x}_{k+1} - x_k). \end{aligned}$$

Тогда коэффициенты

$$A = \varphi(1) - B; \quad B = (f'_k, \hat{x}_{k+1} - x_k); \quad C = 0;$$

а точка минимума функции $\varphi(\xi)$

$$\xi_M = -\frac{B}{2A} \quad (12)$$

4) Последующее приближение строится в виде

$$x_{k+1} = (1 - \xi_M)x_k + \xi_M \hat{x}_{k+1} \quad (13)$$

Подставив (11) в (13), получим

$$x_{k+1} = x_k - \bar{\alpha} f'_k, \quad (14)$$

а из (12), разложив в знаменателе $f(\hat{x}_{k+1})$ в ряд Тейлора, - значение

$$\bar{\alpha} = \frac{\|f'_k\|^2}{(f''_{kc}f'_k, f'_k)}, \quad (15)$$

где $\bar{\alpha} = \alpha\xi_M$,

$$f''_{kc} = f''((1 - \theta)x_k + \theta\hat{x}_{k+1}), \quad \theta \in [0, 1]$$

Лемма 4. Пусть $f(x)$ - дважды непрерывно дифференцируемая функция, причем ее матрица вторых производных удовлетворяет условиям

$$m\|y\|^2 \leq (f'(x)y, y) \leq M\|y\|^2, \quad M \geq m > 0 \quad (16)$$

при любых $x, y \in E^n$, а последовательность $\{x_k\}$ строится по описанному выше алгоритму.

Тогда справедливы оценки

$$\frac{1}{M} \leq \bar{\alpha} \leq \frac{1}{m} \quad (17)$$

$$\|x_{k+1} - x_k\| \leq \frac{1}{m}\|f'_k\| \quad (18)$$

$$-\frac{2M^2 - m^2}{2m^2M}\|f'_K\|^2 \leq f_{k+1} - f_k \leq \frac{M^2 - 2m^2}{2m^2M}\|f'_k\|^2 \quad (19)$$

Для описанного алгоритма при выполнении условий леммы 4.1 справедлива теорема 3.2.

При любой начальной точке x_0 будет $x_k \rightarrow x_*$, $f(x_k) \rightarrow f(x_*)$.

Для скорости сходимости справедливы оценки

$$f_k - f_* \leq q^k(f_0 - f_*), \quad \|x_k - x_*\| \leq cq^{\frac{k}{2}}, \\ c < \infty, 0 < q < 1.$$

В отличие от указанной теоремы, в доказательстве которой неравенство $1 - \frac{\alpha M}{2} \geq \varepsilon$ применяется для оценки сверху шага α в направлении спуска, здесь это неравенство с учетом левой части (17) дает оценку для $\varepsilon : 0 < \varepsilon \leq \frac{1}{2}$

и минимальное значение знаменателя прогрессии

$$q_{\min} = 1 - \frac{m}{2M} \left(1 + \frac{m}{M}\right)$$

Выражение для $q_{\min}$ то же, что и в теореме 3.3, однако получено для произвольного положительного значения α в (11).

Ввиду произвольности выбора шага в направлении спуска описанный алгоритм обладает преимуществом перед обычным градиентным методом, особенно в том случае, когда для выпуклой функции $f(x)$ неизвестны оценки M и m для матрицы вторых производных.

Численное нахождение минимума функции модификацией градиентного метода.

В данном разделе рассматривается применение модификации градиентного метода для отыскания минимума заданной функции $f(x)$ при заданных начальной точке и параметре α . Его целью является ускорение сходимости.

Будем рассматривать функцию $f(x)$ вида (7), градиент которой имеет вид (8). Требуется найти точку x_m , которая будет минимумом функции $f(x)$ вида (7) на множестве R^2 . Воспользуемся упомянутым в предыдущем разделе методом для отыскания x_m .

Возьмём

$$\hat{x}_{k+1} = x_k - \alpha f'(x_k), \quad (20)$$

где $\alpha > 0$ задаётся с клавиатуры.

Искомые x_{k+1} будем вычислять по

$$x_{k+1} = (1 - \xi_M)x_k + \xi_M \hat{x}_{k+1}, \quad (21)$$

где

$$\xi_M = -\frac{(f'(x_k), \hat{x}_{k+1} - x_k)}{2 \cdot (f(\hat{x}_{k+1}) - f(x_k) - (f'(x_k), \hat{x}_{k+1} - x_k))}. \quad (22)$$

Помним, что

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

По (22) строим приближения до тех пор, пока не выполнится условие

$$(|x_{k+1,1} - x_{k,1}| < \varepsilon) \&\& (|x_{k+1,2} - x_{k,2}| < \varepsilon). \quad (23)$$

Когда (23) выполнилось, фиксируем полученную x_{k+1} . Это и будет точка x_M - искомый минимум функции $f(x)$ на R^2 .

Вышеуказанный итерационный процесс реализован в программе, разработанной на языке программирования Java. 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 так же возьмём $x_0 = (-0, 5; -0, 5)$. Подбором параметра α получим значение x_M .

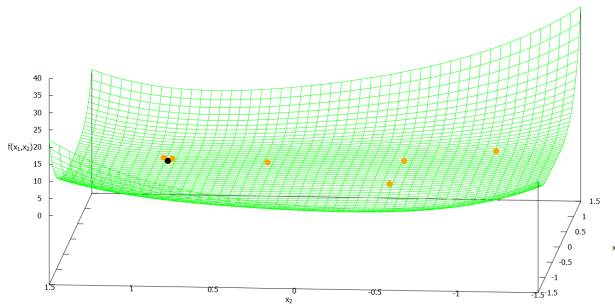


Рисунок 11 — $\alpha = 0, 1$.

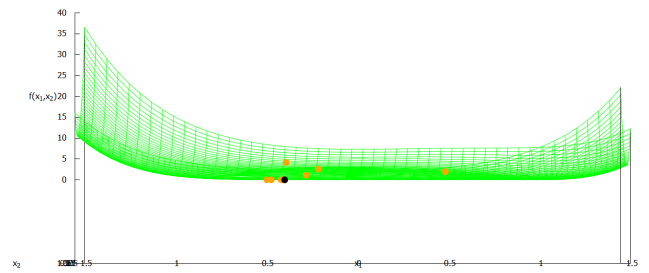


Рисунок 12 — $\alpha = 0, 1$.

На рисунках 11 — 12 показано полученное за 535 итераций приближенное решение $x_M = (0.3569747; 0.934114575)$, а значение функции в этой точке $f(x_M) = 8.01931591E - 12$.

На рисунках 13 — 14 показано полученное за 99 итераций приближенное решение $x_M = (0.356965535; 0.934117149)$, а значение функции в этой точке $f(x_M) = 7.1367558792E - 12$.

На рисунках 15 — 16 показано полученное за 139 итераций приближенное решение $x_M = (-0.98170995; 0.193988025)$, а значение функции в этой точке $f(x_M) = 7.02374005E - 11$.

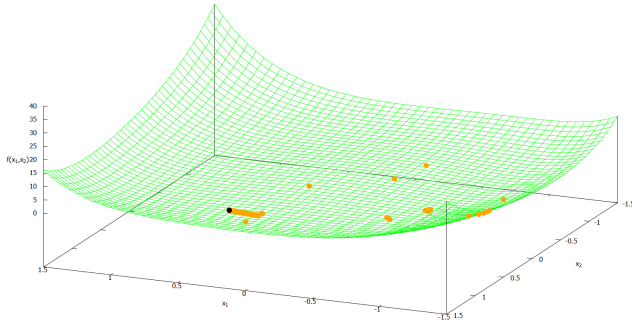


Рисунок 13 — $\alpha = 0,3$.

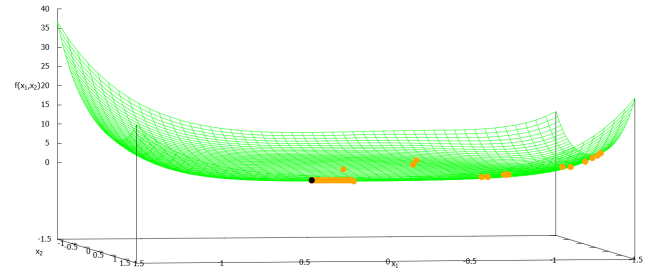


Рисунок 14 — $\alpha = 0,3$.

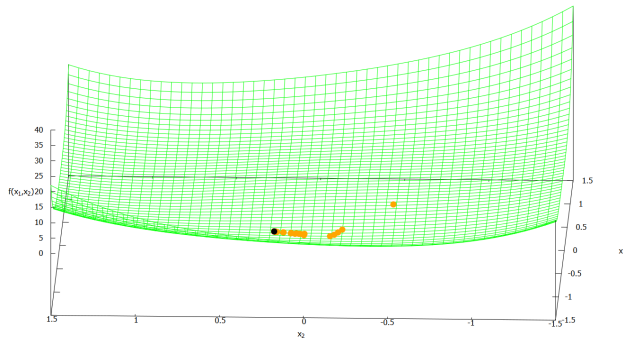


Рисунок 15 — $\alpha = 0,35$.

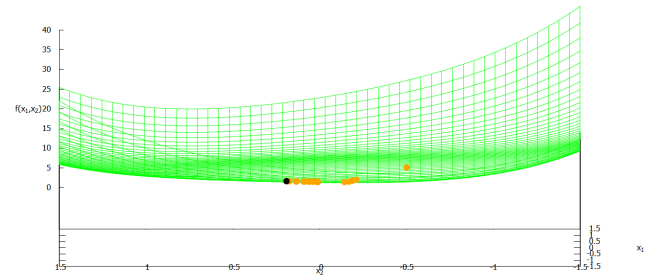


Рисунок 16 — $\alpha = 0,35$.

На рисунках 17 — 18 показано полученное за 57 итераций приближенное решение $x_M = (-0.98170; 0.190419)$, а значение функции в этой точке $f(x_M) = 1.2402905E - 13$.

Как видно на рисунках 9–10 и при сравнении полученных значений функций среди наилучших результатов градиентного спуска и модификации 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 приближенного решения.

В качестве подтверждения вышесказанного отобразим на одной картинке лучший результат градиентного метода и лучший результат модификации градиентного метода. Зелёным цветом показан график функции $f(x)$, синим — x_k метода градиентного спуска, красным — решение методом градиентного спуска, оранжевым — x_k модификация градиентного метода, черным — решение модификации градиентного метода.

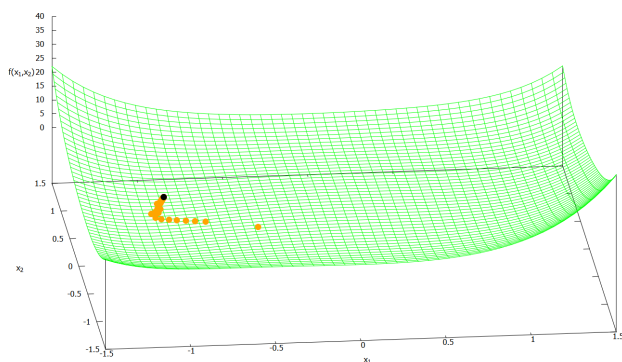


Рисунок 17 — $\alpha = 0,4$.

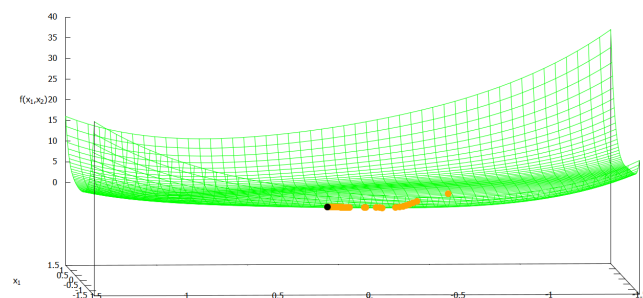


Рисунок 18 — $\alpha = 0,4$.

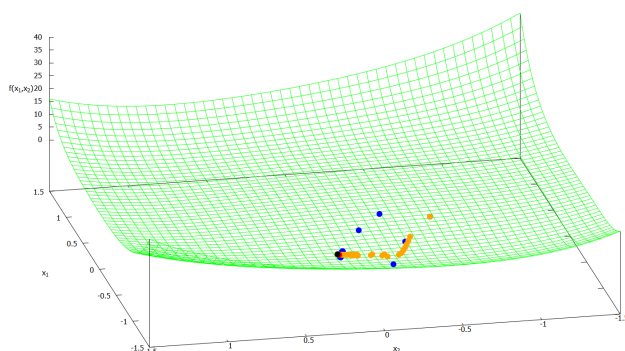


Рисунок 19

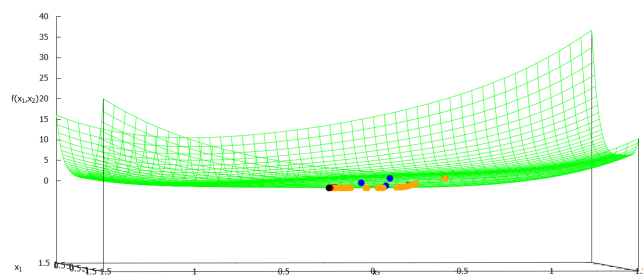


Рисунок 20

Заключение. В отличие от градиентного метода, в котором градиент определяют на каждом шаге, в модификации градиентного метода, градиент находят в начальной точке и движение в найденном направлении продолжают одинаковыми шагами до тех пор, пока значение функции уменьшается (увеличивается). Если на каком-либо шаге $f(x)$ возросло (уменьшилось), то движение в данном направлении прекращается, последний шаг снимается полностью или наполовину и вычисляется новое значение градиента и новое направление. Модификация градиентного метода требует решения на каждом шаге задачи одномерной оптимизации. Практика показывает, что этот метод часто требует меньшего числа операций, чем градиентный метод с постоянным шагом.