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 минимизации нулевого порядка в конечномерном пространстве
наименование темы курсовой работы полужирным шрифтом

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента __4___ курса _____411_____ группы

направления_01.03.02 - Прикладная математика и информатика___________________
код и наименование направления

_________________механико-математического факультета_______________________
наименование факультета, института, колледжа

______________________Щербакова Ильи Алексеевича___________________________
фамилия. имя, отчество

Научный руководитель
зав. кафедрой, д.ф-м.н., профессор _________________ ___В.А. Юрко____
должность, уч. степень, уч.звание подпись, дата инициалы, фамилия

Зав. кафедрой
д.ф-м.н., профессор________ __________________ ____В.А. Юрко___
должность, уч. степень, уч.звание подпись, дата инициалы, фамилия

Саратов 2020

Введение. Целью данной выпускной квалификационной работы бакалавра является рассмотрение численных методов задачи поиска минимума функции в конечномерном векторном пространстве.

Эти методы особенно актуальны, так как нашли свое применение в обучении нейронных сетей, где саму математическую задачу можно, как правило, сформулировать как задачу минимизации. Среди более классических сфер применения этих методов можно выделить технику, экономику и управление процессами.

Кроме численных методов существуют аналитические, графические и экспериментальные, но эти способы часто оказываются неприемлимыми по той или иной причине. Поэтому большой практический интерес представляют численные методы, использующие предшествующую информацию для построения решения задачи при помощи итерационых процедур.

Эти методы являются итерационными, т.е. в них строится последовательность приближения $x^0, ..., x^n, ...$ сходящихся к точке минимума, а также релаксационными (или монотонными), т.е. значение функции монотонно убывает от итерации к итерации:

$$f(x^0) > f(x^1) > ... > f(x^k) > ...$$

Алгоритмы безусловной минимизации подразделяются на три группы в зависимости от максимального порядка производных минимизирующей функции. Наибольшей популярностью, при решении задач такого рода на компьютере, пользуются методы 0-го порядка (или прямые) для нахождения минимума функции, которые используют лишь значения этой функции.

В этих методах для определения направления спуска не требуется вычислять производные целевой функции. Направление минимизации в данном случае полностью определяется последовательными вычислениями значений функции. Следует отметить, что при решении задач безусловной минимизации методы первого и второго порядков обладают, как правило, более высокой скоростью сходимости, чем методы нулевого порядка. Однако на практике вычисление первых и вторых производных функции большого количества переменных весьма трудоемко. В ряде случаев они не могут быть получены в виде аналитических функций. Определение производных с помощью различных численных методов осуществляется с ошибками, которые могут ограничить применение таких методов.

Для достижения поставленной цели в работе необходимо решить следующие **задачи**:

- Рассмотреть оценки скорости сходимости численных методов общего вида, независимо от конкретной реализации этих процессов.
- Изучить вопросы сходимости релаксационных методов при решении задач поиска безусловного минимума выпуклой дифференцируемой функции.
- Рассмотреть методы минимизации функций нулевого порядка.
- Реализовать программы минимизации функций нелевого порядка.

Структура работы содержит 4 раздела:

1. Общая схема методов спуска
2. Общая теорема сходимости релаксационных методов минимизации
3. Метод покоординатного спуска
4. Описание программ и численные результаты

Основное содержание работы.

Во **введении** формулируется цель работы и решаемые задачи.

В **первой** разделе приводится общая схема методов спуска изучается выбор направления длины шага.

Методы спуска состоят в следующей процедуре построения последовательности $\{x_k\}$. В качестве начального приближения выбирается, вообще говоря, любая точка $x_0 \in E_n$.

Последовательные приблежения $x_1, x_2, ...$ строятся по следующей схеме:

1. в точке x_k выирают направление спуска: $-s_k$
2. находят $(k+1)$ - е приближение по формуле

$$x_{k+1} = x_k - \beta_k s_k$$

Всегда в релаксационных методах **направление спуска** $-s_k$ выбирают таким образом, чтобы обеспечить неравенство $f(x_{k+1}) < f(x_k)$ по крайней мере для малых значений величины β_k. Поскольку в реальных задачах, как правило, имеется информация лишь о поведении функции $f(x)$ в локальной окрестности точки x_k, то гарантирует убывание функции при перемещении из точки вдоль направления $-s_k$ соотношение

$$\langle f'(x_k), s_k \rangle > 0$$

Число β_k определяет расстояние от точки x_k до точки x_{k+1}. Будем называть это число **длиной шага** или просто шагом. Основная задача при выборе величины β_k в релаксационных процессах - это обеспечить выполнение неравенства $f(x_{k+1}) < f(x_k)$. Одним из элементарных способов является способ удвоения (или способ Розенброка).

Выбирают $\beta_k = \beta_{k-1}$. Если при этом $f(x_{k+1}) < f(x_k)$, то либо переходят к следующей $(k+2)$ - ой итерации, либо выбирают $\beta_k = 2\beta_{k-1}$. Если значение $f(x)$ меньше предыдущего значения, то процесс удвоения можно продолжать до тех пор пока убывание не прекратится. Если $f(x_{k+1}) \geq f(x_k)$ то выбирают $\beta_k = \frac{1}{2}\beta_{k-1}$. Если $f(x_k - \frac{1}{2}\beta_{k-1}s_k) < f(x_k)$, то полагают $x_{k+1} = x_k - \frac{1}{2}\beta_{k-1}s_k$ и переходят к следующей $(k+2)$ - й итерации. Если же $f(x_k) - \frac{1}{2}\beta_{k-1}s_k \geq f(x_k)$ то выбирают $\beta_k = \frac{1}{4}\beta_{k-1}$ и т.д.

Во **втором** разделе доказывается общая теорема сходимости для релаксационных методов.

Процесс построения последовательности точек x_k будем называть релаксационным, если $x_k \in X$ и $f(x_{k+1}) \leq f(x_k)$, $k = 0, 1, ...$

Будем предполагать, что множество $X^* = x^* \in X : f(x^*) = \min\limits_{x \in X} f(x)$ не пусто и при этом $x_k \notin X^*$.

Обозначим $\mu_k := f(x_k) - f(x^*)$. Из выпуклости функции $f(x)$ следует, что

$$0 < \mu_k \leq \langle f'(x_k), x_k - x^* \rangle$$

Это соотношение справедливо $\forall\ x_k$, ввиду предположения $f(x_k) \neq f(x^*)$.

В основе большинства исследований сходимости релаксационных методов минимизации лежат следующие теоремы об оценках.

Теорема 1. Если

1. $X = E_n$;
2. выпуклая функция $f(x)$ дифференцируема;
3. множество $X_0 = x : \ f(x) \leq f(x_0)$ ограниченно: $diam X_0 = \eta(x_0) = \eta < \infty$;

4. последовательность x_k релаксационная,

то справедливо равенство:

$$f(x_m) - f(x^*) \le \mu_0 \left[1 + \mu_0 \frac{1}{\eta^2} \sum_{k=0}^{m-1} \frac{f(x_k) - f(x_{k+1})}{\|f'(x_k)\|^2}\right]^{-1}, m = 1, 2, ...,$$

где $\mu_0 := f(x_0) - f(x^*)$.

Теорема 2. Если

1. выпуклая функция $f(x) \in C^{1,1}(E_n)$;
2. $diam X_0 = \eta < \infty$;
3. последовательность x_k определяется соотношениями:

$$x_{k+1} = x_k - \beta_k s_k,$$

$$f(x_k) - f(x_k - \beta_k \Delta_k s_k) \ge \beta_k \Delta_k q_k \langle f'(x_k), s_k \rangle \beta_k \ge 0,$$

то

$$f(x_m) - f(x^*) \le \mu_0 \left[1 + \frac{1}{L\eta^2} \mu_0 \sum_{k=0}^{m-1} q_k \alpha_k^2\right]^{-1}, m = 1, 2, ...$$

, где α_k - величина косинуса угла между направлениями антиградиента $-f'(x)$ в точке x_k (т.е. направлением наискорейшего убывания функции $f(x)$ в этой точке) и направлением спуска s_k из этой точки: $\alpha_k := \frac{\langle f'(x_k), s_k \rangle}{\|f'(x_k)\| \|s_k\|}$.

Теорема 3. Если

1. выпуклая функция $f(x) \in C^{1,1}(E_n)$;
2. $diam X_0 = \eta < \infty$;
3. последовательность x_k строится по формулам:

$$x_{k+1} = x_k - \beta_k s_k,$$

$$f(x_k - \beta_k s_k) \le (1 - \lambda_k) f(x_k) + \lambda_k \omega_k.$$

то справедлива оценка

$$f(x_m) - f(x^*) \le \mu_0 \left[1 + C\mu_0 \sum_{k-0}^{m-1} \lambda_k \alpha_k^2\right]^{-1} \forall 0 < C < \frac{1}{2L\eta^2}, m = 1, 2, ...$$

В **третьем** разделе рассматриваются различные методы минимизации функции нулевого порядка, метод покоординатного спуска, метод прямого поиска, метод поиска по деформируемому многограннику, метод Розенброка и метод Пауэлла.

В **четвертом** разделе даётся описание программы, реализующей методы, так же приводятся численные результаты.

В данной разделе реализована программы поиска минимума функции методом покоординатного спуска и методом Розенброка, выполненая на языке C++. Программа составлена для двумерного случая. В качестве необходимой точности взята стандратная константа DBL_EPSILON = 2.2204460492503131e−016. Для методов покоординатного спуска с обратной и последовательной параболическими интерполяциями использовалась точность 1e−03, т.к. при стандартной в их алгоритмах происходило деление на ноль.

Метод покоординатного спуска реализован в виде функции которая принимает начальное приближение, необходимая точность, целевую функцию и метод одномерной минимизации. На выходе получаем точку минимума и вывод количества итераций. После этого выводим значение функции в ней. Критерием окончания поиска является невязка функции на соседних итерациях.

В качестве методов одномерной минимизации были использованы обратная параболическая интерполяция, последовательная параболическая интерполяция и метод золотого сечения.

Метод Розенброка также реализован в виде функции, принимающей те же параметры, что и метод покоординатного спуска, но без метода одномерной минимизации. Критерием окончания поиска также является невязка функции. Направления осей представлены в виде массивов. Новая система координат строится не с помощью ортогонализации Грама-Шмидта, а просто строится вектор, перпендикулярный уже имеющемуся, так как это легче в двумерном пространстве.

Таблица 1 — Минимизация функции $(x1^2 + x2^2 - 1)^2 + (0.75x1^3 - 0.9x2^2 - 1)^2$

Начальная точка	$(0.5, 0.5)$	$(10, -10)$
Значение функции в начальной точке	1.52973	473882
Метод покоординатного спуска с обратной параболической интерполяцией при точности 1e−03		
Количество итераций	6	4
Точка минимума	$(-0.0640646, 0.0152112)$	(0.108954, −0.081725)
Минимум функции	1.99216	1.97335
Метод покоординатного спуска с последовательной параболической интерполяцией при точности 1e−03		
Количество итераций	2	5
Точка минимума	$(8.52471e{-}05, 0.235504)$	$(1.06089, -1.23878e - 05)$
Минимум функции	1.99448	0.0266645
Метод покоординатного спуска с методом золотого сечения		
Количество итераций	20	25
Точка минимума	(1.06073, −3.76712e−11)	(1.06073, −1.80813e−11)
Минимум функции	0.0266642	0.0266642
Метод Розенброка		
Количество итераций	110	161
Точка минимума	$(1.06073, 8.85359\text{e}{-}08)$	$(1.06073, 1.18919\text{e}{-}07)$
Минимум функции	0.0266642	0.0266642

Как видно из таблицы 1, метод покоординатного спуска с обратной параболической интерполяцией нашел локальный минимум, а с последовательной нашел локальный минимум, при точке близкой к нему, и глобальный, при более удаленной точке. Метод покоординатного спуска с методом золотого сечения и метод Розенброка нашли глобальный минимум, но методу Розерброка понадобилось на это немного больше итераций.

Таблица 2 — Минимизация функции Розенброка $100(x2 - x1^2)^2 + (1 - x1)^2$, которая имеет глобальный минимум 0 в точке $(1, 1)$

Начальная точка	$(0.5, 0.5)$	$(10, -10)$
Значение функции в начальной точке	6.5	1.21008e+06
Метод покоординатного спуска с обратной параболической интерполяцией при точности 1e−03		
Количество итераций	107	13
Точка минимума	$(0.908644, 0.825496))$	$(1.03828, 1.07772)$
Минимум функции	0.00834786	0.00147507
Метод покоординатного спуска с последовательной параболической интерполяцией при точности 1e−03		
Количество итераций	2	101
Точка минимума	$(0.709882, 0.503932)$	$(0.69309, 0.480374)$
Минимум функции	0.0841684	0.0941937
Метод покоординатного спуска с методом золотого сечения		
Количество итераций	4736	4514
Точка минимума	$(0.999998, 0.999997)$	$(0.999997, 0.999995)$
Минимум функции	2.52532e−12	7.03485e−12
Метод Розенброка		
Количество итераций	4123	3894
Точка минимума	$(1.00001, 1.00001)$	$(0.999992, 0.999984)$
Минимум функции	3.56695e−11	6.68589e−11

Как видно из таблицы 2, методы покоординатного спуска с обратной и последовательной интерполяциями нашли глобальный минимум, но не точный, за счет пониженной точности. Метод покоординатного спуска с методом золотого сечения и метод Розенброка нашли глобальный минимум, но методу Розерброка потребовалось на это немного меньше итераций, но минимум получился на порядок менее точный.

Таблица 3 — Минимизация функции $(x1^2 + x2 - 11)^2 + (x1 + x2^2 - 7)^2$

Начальная точка	$(0.5, 0.5)$	$(10, -10)$
Значение функции в начальной точке	144.125	16850
Метод покоординатного спуска с обратной параболической интерполяцией при точности 1e−03		
Количество итераций	59	3
Точка минимума	$(3.5834, -1.84823)$	$(-2.80533, 3.13031)$
Минимум функции	$5.68098e{-05}$	4.20548e−05
Метод покоординатного спуска с последовательной параболической интерполяцией при точности 1e−03		
Количество итераций	3	5
Точка минимума	$(-0.270908, -0.922984)$	$(-3.77941, -3.28325)$
Минимум функции	181.617	5.45718e−07
Метод покоординатного спуска с методом золотого сечения		
Количество итераций	671	10000
Точка минимума	$(3, 2)$	$(-3.77837, -1.63466)$
Минимум функции	2.87314e−14	68.4058
Метод Розенброка		
Количество итераций	88	81
Точка минимума	$(3.58443, -1.84813)$	$(3.58443, -1.84813)$
Минимум функции	6.98159e−15	7.63765e−15

Как видно из таблицы 3, метод покоординатного спуска с обратной и последовательной параболической интерполяцией нашел три разных минимума при разных начальных точках, и не нашел при последовательной параболической интерполяции с начальном точкой $(0.5, 0.5)$. Метод покоординатного спуска с методом золотого сечения нашел минимум при начальной точке $(0.5, 0.5)$ и не справился при $(10, -10)$. Метод Розенброка нашел минимум при обеих начальных точках, причем за меньшее количество шагов, чем метод покоординатного спуска с методом золотого сечения.

Заключение. В данной бакалаврской работе были рассмотрены методы нулевого порядка минимизации в конечномерных пространствах: метод покоординатного спуска, метод прямого поиска, метод поиска по деформируемому многограннику, метод Розенброка и метод Пауэлла. Рассмотрены условия обеспечивающие сходимость алгоритмов и доказана теорема сходимости релаксационных процессов. Также были реализованы методы покоординатного спуска и Розенброка.