Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение высшего образования

«Национальный исследовательский университет ИТМО»

Факультет информационных технологий и программирования

Прикладная математика

Лабораторная работа №3

Выполнили студенты группы № M32091

Фисенко Никита Данилович

Рустамов Марк Самирович

Санкт-Петербург 2023

Постановка задачи:

- Реализовать алгоритм градиентного спуска с постоянным шагом.
- Реализовать алгоритм спуска с дроблением шага, используя условие Армихо.
- Реализовать метод наискорейшего спуска.
- Реализовать метод сопряженных градиентов.
- Проанализировать траектории предложенных алгоритмов на примерах квадратичных функций.
- Исследовать сходимость градиентного спуска с постоянным шагом, а также сравнить полученные результаты для всех функций; сравнить эффективность методов, а также исследовать работу методов в зависимости от выбора начальной точки.
- Реализовать генератор случайных квадратичных функций и исследовать зависимость числа итераций, необходимых градиентному спуску для сходимости от размерности пространства и числа обусловленности оптимизируемой функции.

Цель работы:

Изучение градиентных методов, их реализация, анализ сходимости, а также сравнение эффективности и изображение на графиках.

Теория:

1. ***Метод градиентного спуска с постоянным шагом.***

$$x^{[k+1]}=x^{[k]}-\lambda^{[k]}\nabla f(x^{[k]})$$

Метод градиентного спуска с постоянным шагом является одним из наиболее простых и эффективных методов оптимизации функций. Он используется для минимизации функций, имеющих множество локальных минимумов. Основная идея метода заключается в том, чтобы двигаться в направлении антиградиента функции с фиксированным шагом. Антиградиент функции — это вектор, направленный в сторону наиболее быстрого убывания функции. Шаг — это размер шага, который определяет, насколько далеко нужно переместиться в направлении антиградиента. При выборе значения шага нужно учитывать, что слишком большой шаг может привести к расходимости алгоритма, а слишком маленький - к слишком медленной сходимости.

2. ***Метод градиентного спуска с дроблением шага, с использованием условия Армихо.***

В этом варианте градиентного метода величина шага на каждой итерации выбирается из условия выполнения неравенства (условие Армихо):

$$f(x^{[k+1]}) \leq f(x^{[k]}) - \epsilon\lambda^{[k]} \| f'(x^{[k]}) \|^2$$

В данном алгоритме условие Армихо позволяет контролировать выбор шага на каждой итерации. Оно гарантирует, что новая точка будет достаточно близко к предыдущей точке и что изменение функции f(x) будет не менее значимым, чем ожидаемое. Это помогает избежать слишком малых или слишком больших шагов, которые могут замедлить сходимость алгоритма.

3. ***Метод наискорейшего спуска (на основе метода Брента).***

Этот вариант градиентного метода основывается на выборе шага из следующего соображения. Из точки x(k) будем двигаться в направлении антиградиента до тех пор, пока не достигнем минимума функции f на этом направлении:

$$\lambda^{[k]} = \arg \min_{\lambda \in [0,\infty)} f(x^{[k]} - \lambda f'(x^{[k]}))$$

Другими словами, шаг выбирается так, чтобы следующая итерация была точкой минимума функции f. Метод наискорейшего спуска требует решения на каждом шаге задачи одномерной оптимизации. В нашей работе для её решения используется метод Брента.

4. ***Метод сопряженных градиентов.***

Метод сопряжённых градиентов — итерационный метод для безусловной оптимизации в многомерном пространстве. Он основывается на использовании градиентов и ортогональности направлений спуска. Основная идея метода заключается в том, чтобы выбрать направления спуска таким образом, чтобы они были ортогональны друг другу. Это позволяет избежать зацикливания в локальных минимумах и быстрее достигать глобального минимума. Основным достоинством метода является то, что он решает квадратичную задачу оптимизации за конечное число шагов.

Реализация методов:

https://github.com/russianZAK/applied-mathematics/blob/main/Lab%203/lab3.ipynb

Траектории алгоритмов на примере квадратичных функций f, g, h;

Все описанные алгоритмы хорошо сходились на примере функций f, g, h.

Сходимость градиентного спуска с постоянным шагом:

Сходимость градиентного спуска с постоянным шагом зависит от выбора шага и свойств функции, которую мы оптимизируем. Если шаг выбран правильно, то градиентный спуск с постоянным шагом сходится к оптимальному решению. Однако, если шаг выбран слишком большим, то градиентный спуск может расходиться и не достигать оптимального решения. Если же шаг выбран слишком маленьким, то градиентный спуск будет сходиться очень медленно и может затянуться на несколько итераций. Для того чтобы гарантировать сходимость градиентного спуска с постоянным шагом, необходимо выбирать шаг таким образом, чтобы он был достаточно маленьким для обеспечения сходимости, но при этом не слишком маленьким, чтобы обеспечить быструю сходимость. В ходе работы данные выводы подтвердились.

Сравнение эффективности методов с точки зрения количества вычислений минимизируемой функции и её градиентов:

В ходе работы было установлено, что наиболее эффективным на всех 3 примерах квадратичных функций f, g, h с точки зрения количества вычислений минимизируемой функции и её градиентов стал метод сопряженных градиентов. Метод наискорейшего спуска на основе метода Брента работал практически одинаково с методом сопряженных градиентов на функциях f и g, однако стал заметно менее эффективным на функции h. Самым же наименее эффективным показал себя метод градиентного спуска с постоянным шагом.

Работа методов в зависимости от выбора начальной точки:

Выбор начальной точки может сильно влиять на эффективность методов. Если начальная точка выбрана неправильно, то метод может зациклиться в локальных минимумах или сходиться медленно. Однако ходе работы было установлено, что все описанные методы были достаточно устойчивыми в плане выбора начальной точки. Только для функции h метод наискорейшего спуска на основе метода Брента сходился плохо.

Зависимость числа итераций, необходимых градиентному спуску для сходимости в зависимости от размерности пространства и числа обусловленности оптимизируемой функции:

В ходе работы было установлено, что на малых (<50) размерностях пространства число итераций не превышает 200 вне зависимости от числа обусловленности минимизируемой функции. В остальных случаях при одинаковом числе обусловленности число итераций одинаковое вне зависимости от размерности и увеличивается при увеличении числа обусловленности.

Выводы:

- Все описанные алгоритмы хорошо сходились на примере функций f, g, h.
- Сходимость градиентного спуска с постоянным шагом зависит от выбора шага и свойств функции, которую мы оптимизируем. Если шаг выбран правильно, то градиентный спуск с постоянным шагом сходится к оптимальному решению.
- Наиболее эффективным на всех 3 примерах квадратичных функций f, g, h с точки зрения количества вычислений минимизируемой функции и её градиентов стал метод сопряженных градиентов. Самым же наименее эффективным показал себя метод градиентного спуска с постоянным шагом.
- Выбор начальной точки может сильно влиять на эффективность методов. Однако все описанные методы были достаточно устойчивыми в плане выбора начальной точки.
- На малых (<50) размерностях пространства число итераций не превышает 200 вне зависимости от числа обусловленности минимизируемой функции. В остальных случаях при одинаковом числе обусловленности число итераций одинаковое вне зависимости от размерности и увеличивается при увеличении числа обусловленности.