# Лабораторная №1 по предмету методы оптимизации университет ИТМО.

Группа: М3237

Команда: пацаны на отSOSе

Участники: Курдюков Кирилл Алексеевич, Харёв Павел Андреевич, Стрельников Илья Денисович.

1 Постановка задания:

1. Реализовать алгоритмы:

* метод градиентного спуска
* метод наискорейшего спуска
* метод сопряженных градиентов

Оценить, как меняется скорость сходимости, если для поиска величины шага использовать различные методы одномерного поиска:

* метод дихотомии
* метод золотого сечения
* метод Фиббоначи
* метод парабол
* комбинированный метод Брента

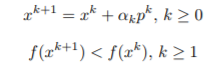
1. Проанализировать траектории методов для нескольких квадратичных функциий.
2. Исследовать, как зависит число итераций, необходимой методам для сходимости, от следующих двух параметров:

* числа обусловленности k >= 1 оптимизируемой функции;
* размерности пространства n оптимизируемых переменных

Для задачи параметров n и k сгенерировать случайным образом квадратичную задачу размера n с числом обусловленности k и запустить на ней методы с некоторой заданной точностью. Замерить число итераций T(n, k), которое потребовалось сделать методу до сходимости.

2 Вычислительная схема алгоритмов.

В данной лабораторной работе рассматриваются итерационные методы, имеющий вид:



**2.1 Метод градиентного спуска**

Метод градиентного спуска – метод нахождения минимума многомерной функции, основная идея которого – осуществлять оптимизацию функции в направлении наискорейшего спуска, а именно антиградиента .

На каждой итерации метода полагается . Если , то условие , очевидно, выполнено. Следовательно, направление вектора является направлением убывания функции , причем в малой окрестности точки  направление обеспечивает наискорейшее убывание этой функции. Поэтому можно найти такое , что будет выполняться условие .

* Начальная данные: 
* Итерационный процесс:

Шаг 1. Задать параметр точности , начальный шаг , выбрать  и вычислить f(x). Перейти к шагу 2.

Шаг 2. Вычислить  и проверить условие достижения точности .

Если оно выполнено, то вычисления завершить, полагая . Иначе перейти к шагу 3.

Шаг 3. Найти  и . Если , то положить x = y, f(x) = f(y) и перейти к шагу 2, иначе – к шагу 4.

Шаг 4. Положить a = a / 2 и перейти к шагу 3.

Необходимо отметить, что вблизи стационарной точки функции f(x) величина  становится малой. Это часто приводит к замедлению сходимости последовательности . Поэтому в основной формуле  иногда полагают , используя вместо антиградиента вектор единичной длины в этом же направлении.

**2.2 Метод наискорейшего спуска**

Метод наискорейшего спуска – модификация метода градиентного спуска. Движение происходит в направлении антиградиента до тех пор, пока не достигнем минимума функции f на этом направлении. Для каждого антиградиента ищется самая оптимальная величина шага  с помощью метода одномерной оптимизации задачи: т. е. на каждой итерации в направлении антиградиента  совершается исчерпывающий спуска.

* Итерационный процесс:

Шаг 1. Задать параметр точности , выбрать . Вычислить f(x). Перейти к шагу 2.

Шаг 2. Вычислить  и проверить условие достижения точности: . Если оно выполнено, вычисления завершить, полагая . Иначе - перейти к шагу 3.

Шаг 3. Решить задачу одномерной оптимизации выше. Положить  и перейти к шагу 2.

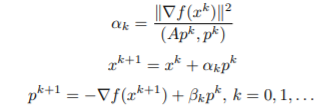
**2.3 Метод сопряженных градиентов**

В первых двух методах в качестве направления убывания функции использовался только вектор антиградиента. Однако такой выбор направления убывания не всегда бывает удачным. В частности, для плохо обусловленных задач минимизации направление антиградиента в точке 

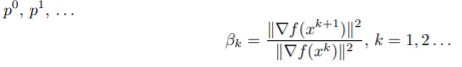
может значительно отличаться от направления к точке минимума . В результате траектория приближения к точке минимума имеет зигзагообразный характер.

В отличие от предыдущих двух методов, в методе сопряженных градиентов направления спуска представляют из себя А – ортогональные вектора, из – за чего метод решает квадратичную задачу оптимизации не более чем за n итераций, где n – размерность пространства, что является преимуществом метода. Метод сопряженных градиентов хорошо работает в случае, если f не является квадратичной функцией, но при этом необязательно сходится к минимуму за предсказуемое число шагов.

* Итерационный процесс:



Значения  выбираются так, чтобы при минимизации квадратичной функции с положительно определенной матрицей получалась последовательность ортогональных векторов:



Одной их важных оптимизаций метода является “перезапуск” метода  через каждые n шагов, где n - размерность пространства. Выполнять обновление необходимо из-за того, что возникает погрешность при вычислении коэффициентов , из – за того, что вектора  могут не образовывать ортогональную систему относительно матрицы А. В итоге вектора  перестанут указывать направление убывания функции.

Можно заметить, что при вычислении  в явном виде не используется матрица А, поэтому метод может применяться для минимизации не только квадратичных функций.

3 Результаты исследования

3.1 Исследование тракторий

При исследовании критерием останова принималось условие = 10^(-5), а за начальное приближение бралась точка (0, 0). Начальное  в градиентном спуске (скорость спуска) бралось равным 10.

* f(x, y) = 64 \* x \* x + 126 \* x \* y + 64 \* y \* y - 10 \* x + 30 \* y + 13

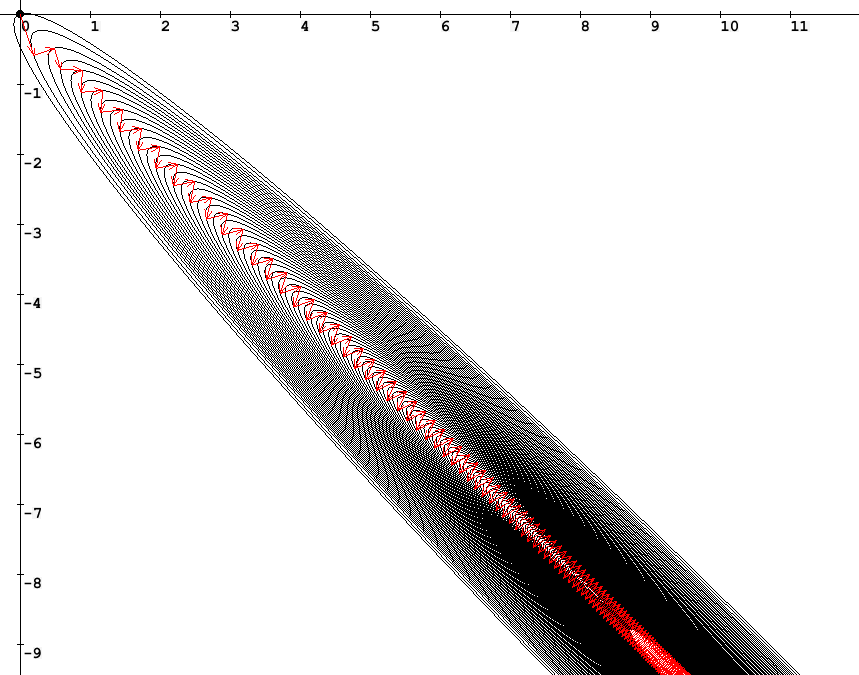


Рис 1.1 Градиентный спуск. На этом рисунке видно, как градиентный спуск двигается очень медленно к точке минимума “зигзагами”. Можно предположить, что как раз таки он двигался по оврагам и из – за чего долго искал минимум. Видно, что длина шага долгое время не изменялось, поэтому зигзаги учащались. Количество итераций – 916.

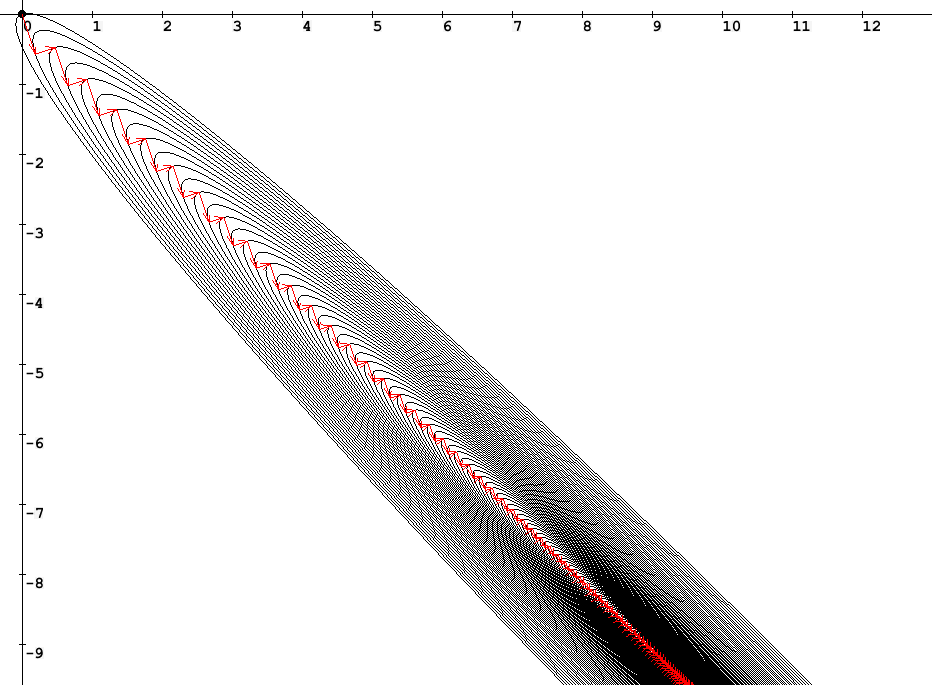


Рис 1.2 Наискорейший спуск. На этом видно, что наискорейший спуск соответственно тоже сходиться очень долго так как его направление – это все тот же антиградиент, что и наискорейшем спуске. Только длинна шага находиться с помощью одномерной минимизации, благодаря этому количество итераций сократилось почти в два раза. Зигзаги грустно :(

|  |  |
| --- | --- |
| Одномерный метод | Количество итераций |
| Дихотомия | 521 |
| Золотое сечение | 475 |
| Фибоначчи | 483 |
| Параболы | 475 |
| Брент | 638 |

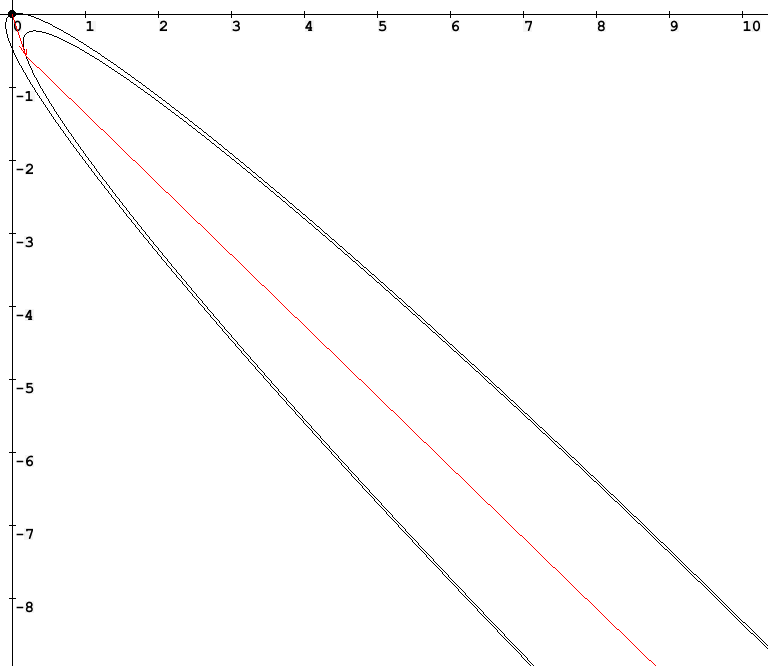


Рис 1.3 Сопряженные градиенты. За обещанные в теории два шага алгоритм находит точку минимума. Так как матрица А размера два на два. Для двумерных функций самый подходящий алгоритм. Количество итераций – 2.

* f(x, y) = 99 \* x \* x + 196 \* x \* y + 99 \* y \* y - 95 \* x - 9 \* y + 91

Функция похоже на первую. Такая же болезненная для первых двух методов.

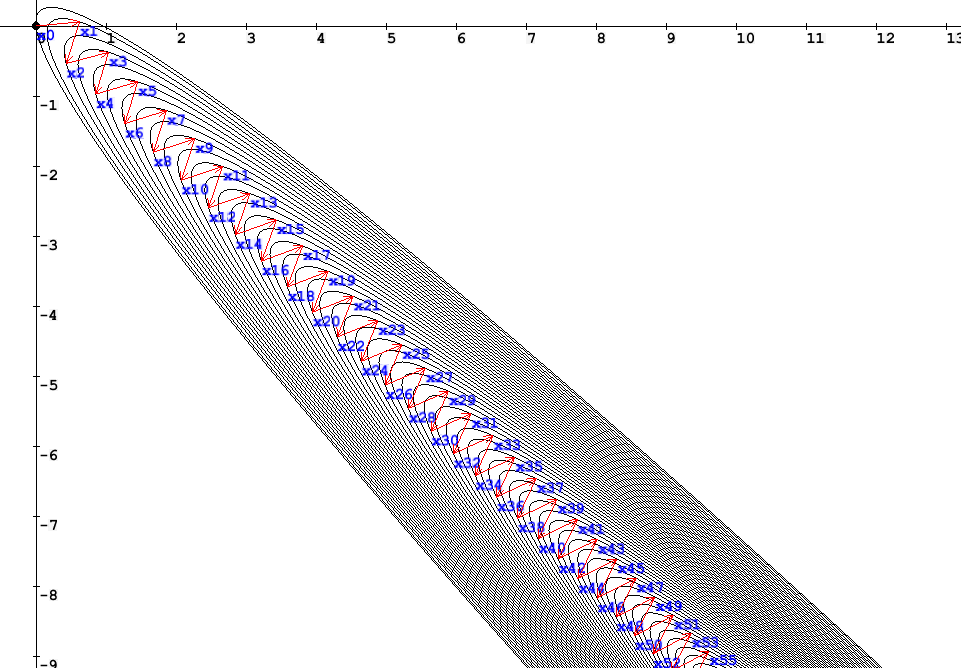


Рис 2.1 Градиентный спуск. Видно, что метод выбрал для себя альфа на очень долгое время из – за этого проскакивает верный маршрут к точке минимума. Количество итераций – 1358.

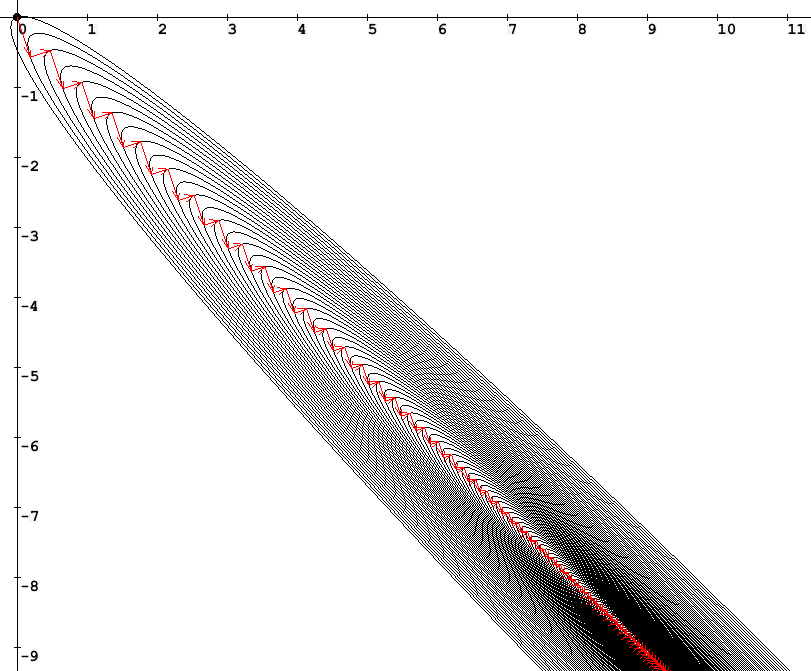


Рис 2.2 Наискорейший спуск. Видно, что с помощью одномерной минимизации мы к точке минимума идем гораздо быстрее.

|  |  |
| --- | --- |
| Одномерный метод | Количество итераций |
| Дихотомия | more than 1000 iterations! |
| Золотое сечение | more than 1000 iterations! |
| Фибоначчи | more than 1000 iterations! |
| Параболы | 995 |
| Брент | more than 1000 iterations! |

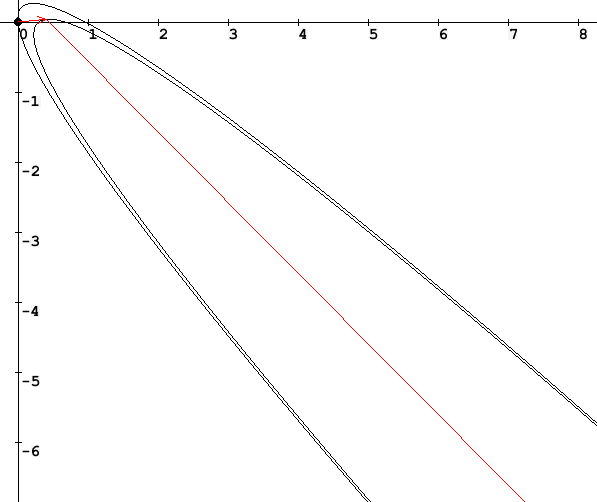


Рис 2.3 Метод сопряженных градиентов продолжает свое доминирование. Количество итераций – 2.

* f(x, y) = 10 \* x \* x + y \* y - 5 \* x + 3 \* y + 8

Более простая функция по сравнению с первыми двумя. И именно на такой функции шикарно прослеживается работа трех методов. Недостатки, преимущества.

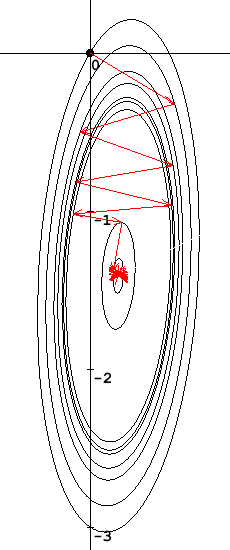
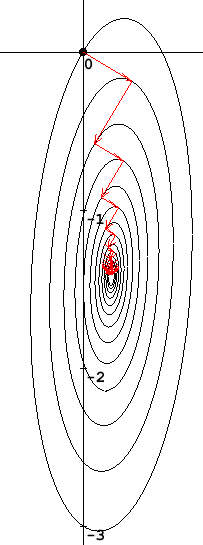
 

Рис 3.1 Градиентный спуск. Рис 3.2 Наискорейший спуск.

Количество итераций – 58. Количество итераций – 49.

|  |  |
| --- | --- |
| Одномерный метод | Количество итераций |
| Дихотомия | 49 |
| Золотое сечение | 49 |
| Фибоначчи | 49 |
| Параболы | 54 |
| Брент | 49 |

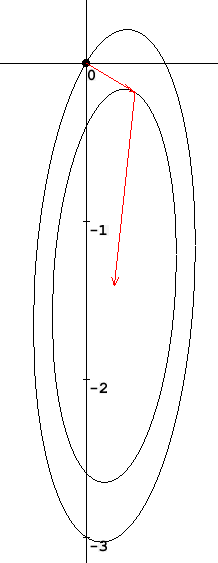


Рис 3.2 Сопряженные градиенты.

Стабильные 2 итерации. 😊

**Вывод**: метод сопряженных градиентов показал себя наилучшим, а метод градиентного спуска наихудшим. На плохо обусловленных функциях методы градиентного и наискорейшего спуска дольше сходятся.

|  |  |  |  |
| --- | --- | --- | --- |
| Функция | Собственное число 1 | Собственное число 2 | Обусловленность |
| F1 | 2 | 254 | 127 |
| F2 | 2 | 394 | 197 |
| F3 | -Sqrt(82) + 11 | Sqrt(82) + 11 | 1.56 |

Видно, что третья функция лучше всех обусловлена, и поэтому время работы первых двух методов приемлемо, гораздо лучше, чем для F1 и F2.

3.2 Исследование n – мерных функций.

Далее представлены таблицы, отражающие количество итераций методов на случайных квадратичных формах в зависимости от числа обусловленности и размерности пространства n.

Минимум всех квадратичных функций форм находился в начале координат, а стартовая точка – в (1, 1, …, 1). Значение в таблицах – количество итераций методов при заданных  и n.

1. F(x1, x2, … xn) = x1^2 + x2 ^ 2 + … + xn ^ 2

* = 1.

Метод градиентного спуска.

|  |  |  |
| --- | --- | --- |
| Размерность | Ответ | Количество итераций |
| 10 | 1.9227233284363255E-13 | 30 |
| 100 | 3.0814879110195774E-31 | 5 |
| 1000 | 1.9227232843264244E-11 | 39 |
| 10000 | 1.1209213202887523E-25 | 50 |

Метод наискорейшего спуска.

|  |  |  |
| --- | --- | --- |
| Размерность | Ответ | Количество итераций |
| 10 | 9.099541208320007E-12 | 20 |
| 100 | 2.274942744689507E-11 | 21 |
| 1000 | 1.4219110214124483E-11 | 23 |
| 10000 | 8.887392693876088E-12 | 25 |

Метод сопряженных градиентов.

|  |  |  |
| --- | --- | --- |
| Размерность | Ответ | Количество итераций |
| 10 | 4.0343434110195774E-21 | 10 |
| 100 | 2.0814879110195774E-31 | 19 |
| 1000 | 2.2342342334434474E-26 | 28 |
| 10000 | 1.0814823423423474E-11 | 54 |

Дальше генерировалась диагональная матрица А с произвольными диагональными элементами от 1 до . Таблицы представлены. С точностью eps = 10 ^ -4.

Градиентный спуск.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| n \ k | 5 | 10 | 20 | 100 | 500 | 1000 | 2000 |
| 10 | 33 | 69 | 142 | 587 | 2354 | 4712 | 9427 |
| 100 | 31 | 69 | 147 | 585 | 2354 | 4712 | 9427 |
| 500 | 35 | 72 | 147 | 588 | 2354 | 4714 | 9427 |
| 1000 | 35 | 76 | 152 | 612 | 2354 | 4712 | 9427 |
| 2000 | 36 | 78 | 155 | 589 | 2371 | 4808 | 9638 |
| 5000 | 38 | 805 | 163 | 621 | 2356 | 4734 | 9427 |
| 10000 | 39 | 83 | 169 | 639 | 2424 | 4715 | 9428 |

**Наискорейший спуск.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| n \ k | 5 | 10 | 20 | 100 | 500 | 1000 | 2000 |
| 10 | 25 | 48 | 97 | 467 | 2347 | 4759 | 9558 |
| 100 | 25 | 51 | 97 | 473 | 2381 | 4749 | 9445 |
| 500 | 26 | 51 | 100 | 491 | 2375 | 4745 | 9475 |
| 1000 | 27 | 53 | 103 | 489 | 2429 | 4731 | 9493 |
| 2000 | 28 | 54 | 104 | 503 | 2391 | 4749 | 9471 |
| 5000 | 29 | 57 | 110 | 503 | 2383 | 4743 | 9493 |
| 10000 | 30 | 58 | 113 | 519 | 2473 | 4921 | 9569 |

**Сопряженные градиенты.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| n \ k | 5 | 10 | 20 | 100 | 500 | 1000 | 2000 |
| 10 | 8 | 10 | 10 | 10 | 10 | 10 | 10 |
| 100 | 13 | 19 | 26 | 36 | 61 | 65 | 68 |
| 500 | 14 | 21 | 30 | 58 | 99 | 113 | 103 |
| 1000 | 14 | 21 | 31 | 69 | 111 | 157 | 186 |
| 2000 | 15 | 22 | 31 | 70 | 140 | 197 | 226 |
| 5000 | 15 | 22 | 32 | 72 | 158 | 223 | 261 |
| 10000 | 16 | 23 | 33 | 74 | 163 | 228 | 319 |

4 Сравнение методов

Теоретическая скорость сходимости метода наискорейшего спуска не выше скорости сходимости метода градиентного спуска. Однако на практике метод наискорейшего спуска требует меньше количество итераций, чем метод градиентного спуска. При этом недостатком метода наискорейшего спуска является необходимость решать одномерную задачу оптимизации.

Самым оптимальным методом является метод сопряженных градиентов. В отличие от метода наискорейшего спуска и градиентного спуска, которые для квадратичной функции сходятся лишь в пределе, метод сопряженных градиентов оптимизирует квадратичную функцию за конечное число итерации (не более, чем размер пространства n).

При оптимизации функций общего вида метод сопряженных градиентов сходится в несколько раз быстрее остальных методов.

На небольших n можно наглядно увидеть, что метод сопряженных градиентов совершает не более n итераций для квадратичных функций при любом числе обусловленности  (иногда метод все же может совершить более чем n итераций – это связано с погрешностью вычислений, которая является одной из главных проблем градиентных методов), тогда для остальных методов количество итераций увеличивается пропорционально . Можно пронаблюдать, что n – это верхняя граница на количество итераций методом сопряженных градиентов. Методы градиентного спуска и наискорейшего спуска такой границы не имеют.

Стоит заметить, что в отличие от методов второго порядка, для градиентных методов не требуется трудоемких вычислений вторых частных производных.

5 Выводы

В ходе работы были изучены градиентные методы многомерной оптимизации и проведено их сравнение.

Основные достоинства градиентных методов:

* Глобальная сходимость, нет серьезных ограничений на начальные данные, точка старта может быть далека от x\_min.
* Градиентные метода – метода первого порядка, необходима только первая производная функции, что уменьшает требования к функции и дает преимущества перед методами, использующими вторую производную.

Недостатки:

* Скорость сходимости градиентных методов существенно зависит от точности вычислений градиента.
* Потеря скорости сходимости в окрестности точки минимума.

Потеря точности обычно происходит в окрестности точек минимума или в овражной ситуации. Поэтому градиентные методы часто используют на начальной стадии решения в дополнение к более эффективным методам.

Были использованы различные оптимизации метода градиентного спуска:

1. Градиентный спуск – базовая реализация идея осуществления оптимизацию в направлении наискорейшего спуска, однако фиксированным сокращением шага не позволяет добиться оптимального количества итераций. Для маленького шага сходимость будет слишком медленной, а для большого возникнут проблемы в виде колебаний возле точки минимума.
2. Метод наискорейшего спуска выбирает шаг алгоритма из условия решения задачи одномерной оптимизации. Движение вдоль антиградиента происходит до тех пор, пока значение функции убывает. Однако у градиентных методов существует проблема овражных функций, для которых направление антиградиента отклоняется от направления в точку минимума, что приводит к замедлению скорости сходимости.
3. Метод сопряженных градиентов самый оптимальный из трех методов. Он использует оптимизации наискорейшего спуска, и добавляет оптимизацию для выбора метода направления спуска. Благодаря этому ему удается добиться сходимости за n(размерность пространства) итераций на квадратичных функциях и существенного улучшение на не квадратичных функциях (за счет того, что гладкие функции в окрестности точки минимума хорошо аппроксимируются квадратичной).