# Лабораторная №4 по предмету методы оптимизации университет ИТМО.

Группа: М3237

Команда: пацаны на отSOSе

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

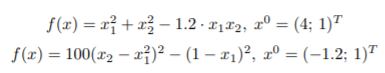
1 Постановка задания

1. Реализовать метод Ньютона: a) классический, б) с одномерным поиском (на выбор), в) с направлением спуска.

(а) Продемонстрировать работу методов на 2, 3 функциях, в том числе на не квадратичных.

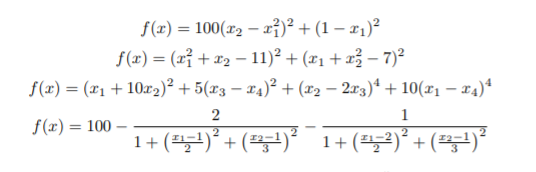
* Для поиска ньютоновского направления спуска необходимо использовать прямой или итерационный метод решения СЛАУ.
* Результаты иллюстрировать траекториями спуска.
* Указать количество итераций, необходимых для достижения заданной точности.
* В случае одномерного поиска указать найденные значения параметра.
* Провести исследование влияния выбора начального приближения на результат (не менее трех).

(b) Исследовать работу методов на двух функциях с заданными начальными приближениями:



* Для поиска ньютоновского направления спуска использовать прямой или итерационный метод решения СЛАУ.
* Сравнить результаты с минимизацией методов наискорейшего спуска из лабораторной №2.
* Построить таблицу или график зависимости “метод – количество итераций”.
* Привести траектории сходимости для каждого метода.

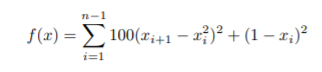
1. Реализовать квазиньютоновские методы (метод Давида – Флетчера – Пауэлла и метод Пауэлла). Сравнить работу квазиньютоновских методов с наилучшим методом Ньютона (по результатам 1b) на функциях:



* Для каждого метода привести иллюстрации траекторий сходимости.
* Привести исследование влияния выбора начального приближения на результат (не менее трех), оценить скорость сходимости.
* Построить таблицу или график зависимости “метод – количество итераций”.

Для отображения траектории поиска точки использовать ранее реализованный графический интерфейс.

1. Реализовать метод Марквардта двумя вариантами. Результаты работы продемонстрировать на минимизации многомерной функции. Розенброка(n = 100) в сравнении с наилучшим методом Ньютона (по результатам 1b):



Для обоих методов построить график зависимости “итерация – число разложения Холецкого”.

2 Вычислительная схема алгоритмов

2.1 Методы Ньютона

На вход методам подаются x0, eps и f – начальное приближение, ошибка и дважды дифференцируемая функция (которую и будем оптимизировать) соответственно. Пусть k – номер итерации.

**2.1.1 Классический метод Ньютона**

1. Решаем СЛАУ  где H – матрица Гессе функции f в точке , а  - градиент в этой точке. Из решения СЛАУ получим направление спуска (в случае если решения нет – берем антиградиент);

2. Положим , где ak = 1, pk = qk.

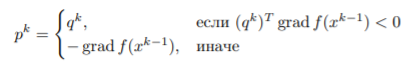
3. Если  оканчиваем итерационный процесс (ответ – xk), иначе переходим к шагу (1).

**2.1.2 Метод Ньютона с одномерным поиском**

Аналогичен классическому, но на шаге (2) полагаем, что 

**2.1.3 Метод Ньютона с направлением спуска**

Аналогичен методу Ньютона с одномерным поиском, но на шаге (2) полагаем, что



2.2 Квазиньютоновские методы

Следующие методы работают по алгоритму:

1. 
2. Нахождение положительно определенной матрицы Gk размерности n на n – определяется реализацией метода
3. 
4. 
5. Если , оканчиваем итерационный процесс (ответ – xk), иначе переходим к шагу (1)

Конкретные реализации квазиньютоновского метода, исследуемые в работе, отличаются только выбором Gk, поэтому в соответствующих секциях будет описано только это. Введем обозначения .

**2.2.1 Метод Давидона – Флетчера – Пауэлла**

* G1 = I – единичная матрица
* **** (1)

Стоит отметить, что для того, чтобы ослабить влияние накапливаемых погрешностей на сходимость итерационной последовательности и уменьшить вероятность появления после очередных n итераций линейно зависимых направлений спуск, в ДФП – методе применяют процедуру обновления алгоритма: через каждые n итераций в качества Gn используют Gn = I.

Свойства метода Давидона – Флетчера – Пауэлла:

* Gk сохраняет положительную определенность: Gk > 0 => Gk+1 (при условии  .
* При минимизации квадратичной функции с положительной матрицей А метод ДФП сводится к методу сопряженных направлений, т к первые n векторов pk (задающих направление спуска) является сопряженным относительно матрицы А. Следовательно ДФП – метод не более чем за n итераций дает точное решение квадратичной функции.
* Если в (1) Gk – симметрична, то Gk + 1 – симметрична
* Матрицы Gk вычисленные по ДФП – методу, связаны равенством:

. При k = n следует, что n векторов pi является собственными для симметричной матрицы GnA, а соответствующие собственные значения равны единице. Следовательно, GnA = I и , т.е Gn оказывается обратной к матрице Гессе H(x\*) = A квадратичной функции в точке x\*.

* Если f(x) не является квадратичной, то ДФП – метод не позволяет найти минимум за конечное число итераций.
* Если f(x) – квадратичная функция с H > 0, то метод ДФП (с точным одномерным поиском) после n итераций дает:



**2.2.2 Метод Пауэлла**

* G1 = I – единичная матрица
* ****

2.3 Метод Марквардта

**2.3.1 С уменьшением диагональной прибавки**

Данному методу дополнительно на вход подаются числа r0 и  - начальная диагональная прибавки и интенсивность ее уменьшения.

1. Инициируем метод , вычислим f(x);
2. Вычислим H – матрицу Гессе функции f в точке x и grad f(x) – градиент в этой точке;
3. Решаем СЛАУ . Из решения СЛАУ получим p;
4. Положим y = x + p, вычислим f(y);
5. Если f(y)> f(x), то и переходим к шагу (3);
6. Пусть ;
7. Если , оканчиваем итерационный процесс (ответ – x), иначе переходим к шагу (2).

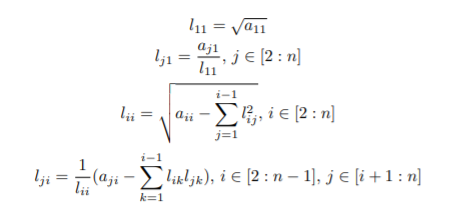
**2.3.2 С увеличением диагональной прибавки**

1. Инициируем метод: ;
2. Вычислим f(x), H – матрицу Гессе функции f в точке x и grad f(x) – градиент в этой точке;
3. Если (H + rl) не положительно определенная, то , возвращаемся в начало шага (3);
4. Решаем СЛАУ . Из решения СЛАУ получим p;
5. Положим x = x + p;
6. Если , оканчиваем итерационный процесс (ответ – x), иначе переходим к шагу (2).

Проверка матрицы на положительную определенность производится с помощью разложения Холецкого.

**2.3.3 Разложение Холецкого**

Разложение Холецкого , где L = (lij) – нижняя треугольная матрица, а А = (аij) – положительно определенная симметричная матрица, строиться по следующим формулам:



Если какое – либо подкоренное выражение отрицательно или lii = 0, то матрица А не является положительно определенной.

3 Результаты исследования

3.1 Методы Ньютона

**3.1.1 Демонстрация работы**

Далее будет продемонстрирована работа методов Ньютона на некоторых функциях: как квадратичных, так и не квадратичных. В качестве ошибки брался eps = 10 ^ -7. Для одномерной оптимизации был взят метод золотого сечения.

1. f(x, y) = x ^ 2 + y ^ 2

(а) Классический метод Ньютона.

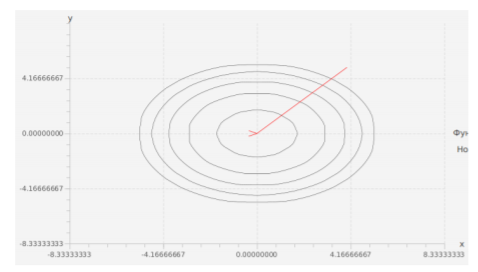
Известно, что метод Ньютона сходится за одну итерацию на квадратичных функциях. В данной ситуации метод сошелся к минимуму за 2 итерации: за одну итерацию был найдет минимум, за вторую метод сделал очень маленький шаг – так метод смог понять, что шаг достаточно мал, чтобы завершать работу. 

Рис. 1: Траектория поиска с началом в (4, 5)^T

(b) Метод Ньютона с одномерным поиском и метод Ньютона с направлением спуска.

Данные два метода имеют аналогичное прошлому методу поведение. Происходит это из-за того, что классический метод Ньютона – их основа. Так же ak примерно равен 1 (результат одномерной оптимизации) на каждом шагу, так как полученный вектор ведет ровно в минимум.

2. f(x, y) = 10x^ 2 + y ^ 2 + 5 \* x + y + 5

(а) Классический метод Ньютона.

Аналогично предыдущей функции, классический метод Ньютона завершается за 2 шага.

(b) Метод Ньютона с одномерным поиском и метод Ньютона с направлением спуска.

Данные методы имеют аналогичную траекторию поиска, но завершаются за 3 итерации. Результаты одномерной оптимизации: а1 1, a2  0.95, a3  0.13.

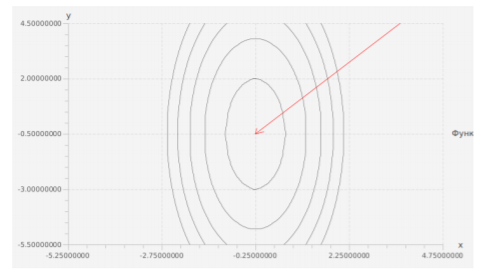


Рис. 2: Траектория поиска с началом в (4,5)^T

Данному явлению можно дать следующее объяснение: одномерная оптимизация дала ответ с ошибкой eps = 10^ -7, из-за чего погрешность ответа после первой итерации равна \* 10 ^ -7 – из-за этого может понадобиться еще как минимум одна итерация.

3. ln(x^2 + y^4 + 1)

(a) Классический метод Ньютона.

Метод был протестирован с тремя разными начальными точками: (10^-5; 10 ^ -5),

(0.01, 0.05), (0.6, 0.6). Для первых двух точек количество итераций соответсвенно: 2, 17. В третьем случае метод не сошелся.

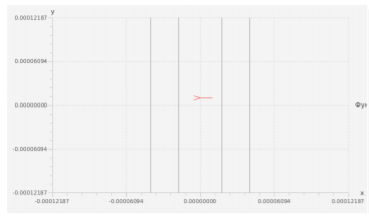


Рис. 3: Траектория поиска классического метода Ньютона с началом в (10 ^ -5, 10 ^ -5).

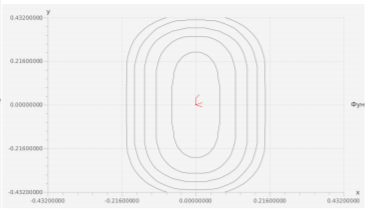


Рис 4: Траектория поиска классического метода Ньютона с началом в (0.01, 0,05).

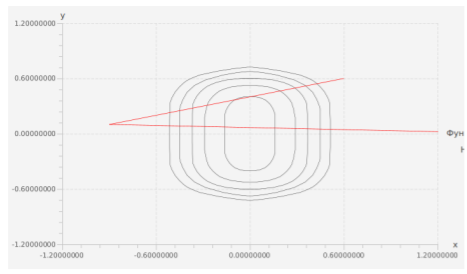
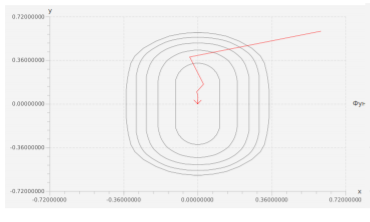


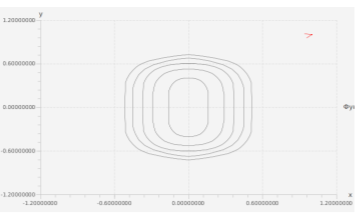
Рис. 5: Траектория поиска классического метода Ньютона с началом в (0.6, 0.6).

(b) Метод Ньютона с одномерным поиском.

Метод был протестирован с двумя разными начальными точками: (0.6, 0.6), (1, 1).

Понятно, что для точки (10^-5, 10^-5) результаты будут аналогичны прошлому методу. Начиная в первой точке, метод сошелся за 16 итераций. Для ak было получена последовательность (почти все числа с точностью до двух знаков после точки): 0.43, 1.68, 10 ^ -7, 1.13, 10 ^ -7, 1.53, … Во втором случае метод не сошелся ak примерно равен 10 ^ -7.

Рис. 6: Траектория поиска метода Ньютона с одномерным поиском, с началом в (0.6, 0.6).

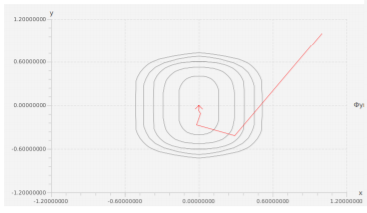
 Рис. 7: Траектория поиска метода Ньютона с одномерным поиском, с началом в (1, 1).

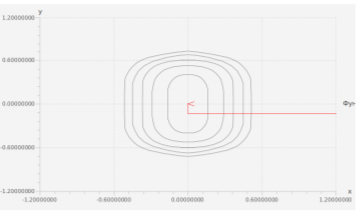
(c) Метод Ньютона с направлением спуска.

Метод был протестирован с тремя разными начальными точками: (0.6, 0.6), (1, 1), (100, 100).

Аналогично предыдущему методу, можно сделать вывод про начальную точку (10 ^ -5, 10 ^ -5).

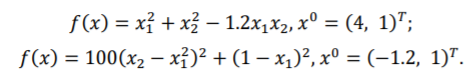
В первой точке результат был аналогичный прошлому методу. Для второй точки метод сошелся за 17 итераций, ak из (0.5, 2), a17 = 10^-7. Для третьей точки метод сошелся за 5 итераций, последовательность альф 2500, 5000, 3, 1, 10^-7. Улучшение сходимости в третьей точке можно обосновать тем, что спуск по антиградиенту оказался спуска по ньютоновскому направлению.

Рис. 8: Траектория поиска метода Ньютона с направлением, с началом в (1, 1).

Рис. 9: Траектория поиска метода Ньютона с направлением, с началом в (100, 100).

**3.1.2 Исследование работы на заданных функциях**

Далее будет исследована работа методов Ньютона и метода наискорейшего спуска на следующих функциях с заданными начальными приближениями:



1. 

Это квадратичная функция. Результаты, полученные в этом пункте аналогичны прежним результатам на квадратичных функциях. Для одномерного поиска ak примерно равна 1.

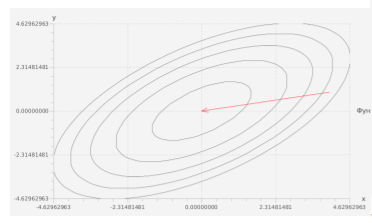


Рис. 10: Траектория поиска методов Ньютона

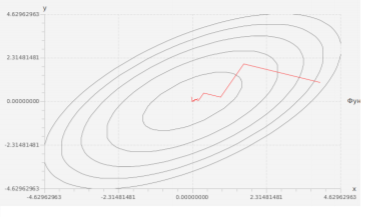


Рис. 11: Траектория поиска метода наискорейшего спуска

2. 

Далее будут представлены траектории спуска методов на этой функции. Можно заметить, что метод Ньютона с одномерным поиском неудачно сделал первую итерацию alpha примерно 719, из-за чего очень долго сходился к минимуму (ak примерно 10^-7). Метод Ньютона с направлением спуска сделал первую итерацию аналогично методу Ньютона с одномерным поиском, а далее начал более быстрый спуск к минимуму ak из (0.9, 3), k > 0. Метод наискорейшего спуска попал на аналогичную траекторию через 20 итераций и начал спускаться к минимуму.

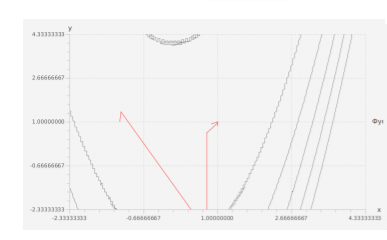


Рис. 12: Траектория поиска классического метода Ньютона

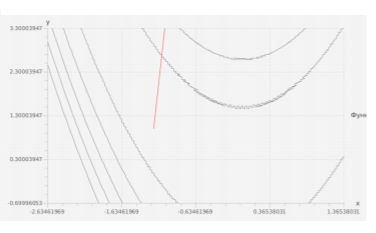


Рис. 13: Первая итерация метода Ньютона с одномерным поиском

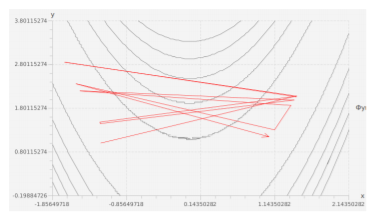


Рис. 14: Первые 20 итераций метода наискорейшего спуска

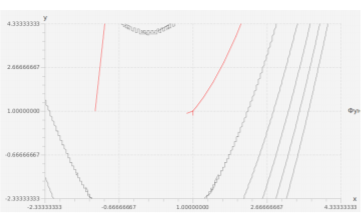


Рис.15: Первая итерация метода Ньютона с направлением спуска

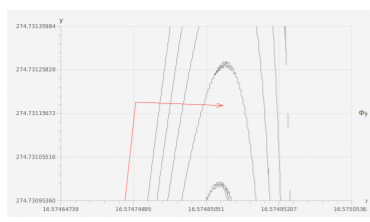


Рис. 16: Первые 2 итерации метода Ньютона с направлением спуска

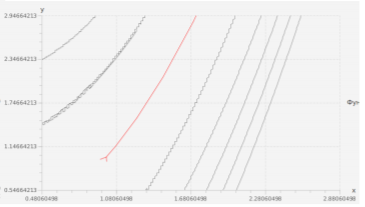


Рис. 17: Последние итерации метода Ньютона с направлением спуска

Далее представлена таблица с зависимостью количества итераций от метода.

|  |  |  |
| --- | --- | --- |
| Метод | Итераций на первой функции | Итерации на второй функции |
| Классический | 2 | 7 |
| С одномерным поиском | 3 | 4694163 |
| С направлением спуска | 3 | 40 |
| Наискорейшего спуска | 24 | 6701 |

Можно сделать вывод, что методы Ньютона значительно лучше на квадратичных функциях, чем метод наискорейшего спуска: количество итераций на второй функции отличается в 1000 раз. С другой стороны, методы Ньютона требуют от функции больше – необходимо знать ее гессиан. Также отсюда можно видеть плохой результат метода Ньютона с одномерным поиском на второй функции – гарантированной сходимости нет.

3.2 Квазиньютоновские методы

Далее будет продемонстрирована работа Квазиньютоновских методов на данных функциях. Здесь и далее eps = 10 ^ -7. Для одномерной оптимизации был взят метод золотого сечения.

1. 

(а) Для точки x0 = (-1.2, 1)

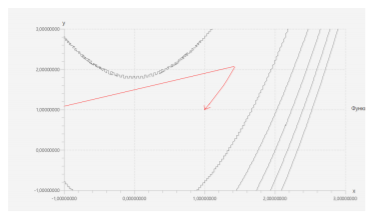


Рис. 18: Траектория метода ДФП

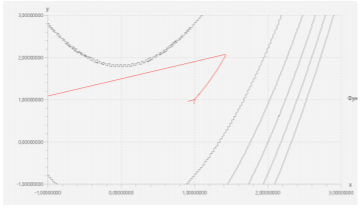


Рис. 19: Траектория метода Пауэлла

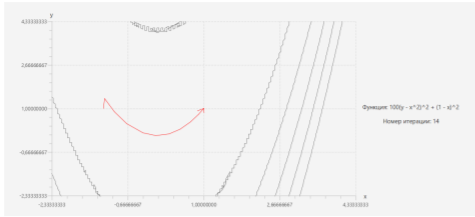


Рис. 20: Траектория поиска метода Ньютона с направлением спуска.

(b) Для точки x0 = (-1, 4)

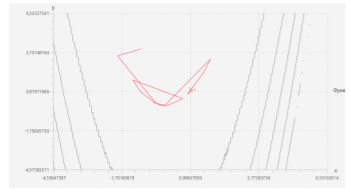


Рис. 21: Траектория метода ДФП

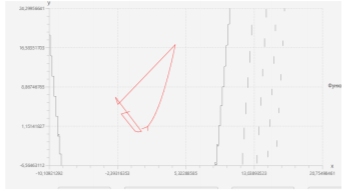


Рис. 22: Траектория метода Пауэлла

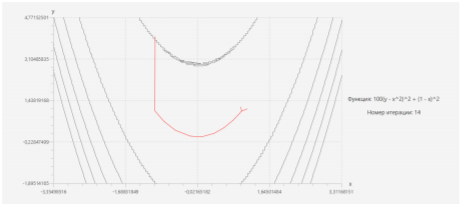


Рис. 23: Траектория поиска метода Ньютона с направлением спуска

(с) Для точки x0 = (-10, 60)

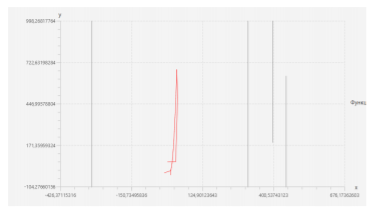


Рис. 24: Траектория метода ДФП

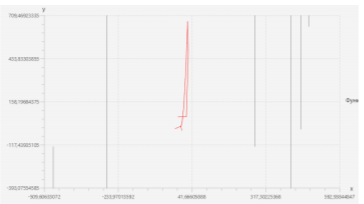


Рис. 25: Траектория метода Пауэлла

Можем заметить, что квазиньютоновские методы в пунктах (а) и (с) ведут себя одинаково и совершают одинаковое число итераций. Однако в пункте (b) траектория метода Пауэлла отличается от ДФП и кол-во итераций почти в два раза больше.

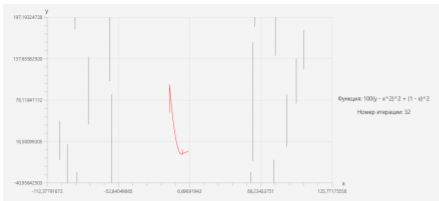


Рис. 26: Траектория поиска метода Ньютона с направлением спуска

Метод Ньютона с выбором направления работает стабильно и совершает примерно в 2 раза меньше итераций, чем квазиньютоновские методы.

2. 

(а) Для точки x0 = (-3, 5)

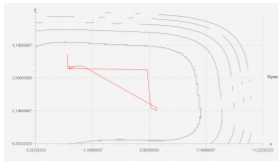


Рис. 27: Траектория метода ДФП

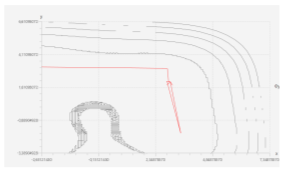


Рис. 28: Траектория метода Пауэлла

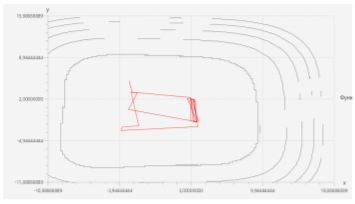


Рис. 29: Траектория поиска метода Ньютона с направлением спуска

У данной функции 4 минимума в пункте а) методы Пауэлла и Ньютона нашли один минимум, а метод ДФП в другой. Также в данном случае метод Ньютона отработал в два раза хуже, чем квазиньютоновские методы.

(b) Для точки х0 = (39; 65)

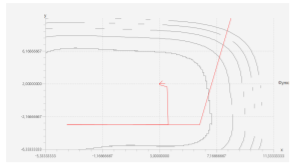


Рис. 30: Траектория метода ДФП

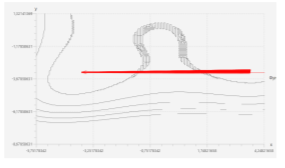


Рис. 31: Траектория метода Пауэлла

Метод Ньютона и метод ДФП сходятся за 13 и 14 операций соответственно, при этом метод Пауэлла совершает скачкообразные движения между двумя точками минимума, однако не сходится.

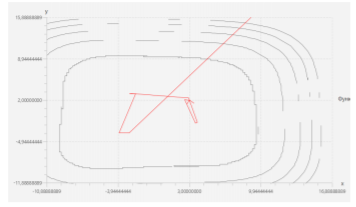


Рис. 32: Траектория поиска метода Ньютона с направлением спуска

3. 

Для данной функции отрисовка графиков оказалась сложной, поэтому будет предоставлена только таблица с зависимостью количества итераций от метода и начальной точки.

|  |  |  |  |
| --- | --- | --- | --- |
| Метод | (10^-5, 10^-5, 10^-5, 10^-5) | (1, 2, 3, 4) | (12.4, 25, 2.7841) |
| Пауэлла | 2 | 21 | 36 |
| Давидона-Флетчера-Пауэлла | 2 | 49 | 53 |
| Ньютона с направлением спуска | 3 | 25 | 24 |

4. 

(a) Для точки x0 = (1.29164, 0.99999)

Оба квазиньютоновских метода сошлись за одинаковое количество итераций и с похожими траекториями.

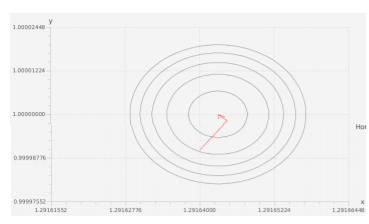


Рис. 36: Траектория поиска методов Пауэлла и ДФП

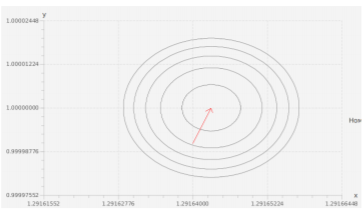


Рис. 37: Траектория поиска метода Ньютона с направлением спуска

(b) Для точки х0 = (2, 2)

Оба квазиньютоновских метода сошлись за одинаковое количество итераций и с похожими траекториями.

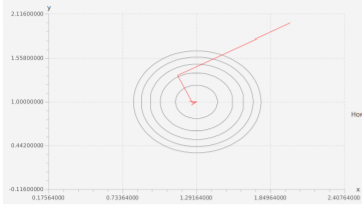


Рис. 38: Траектория поиска методов Пауэлла и ДФП

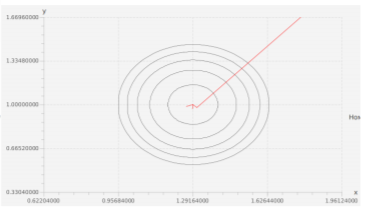


Рис. 39: Траектория поиска метода Ньютона с направлением спуска

(с) Для точки x0 = (300, 300)

С началом в данной точке метод Пауэлла не сошелся.

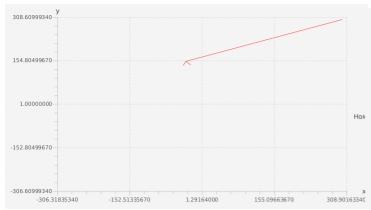


Рис. 40: Траектория поиска метода Пауэлла

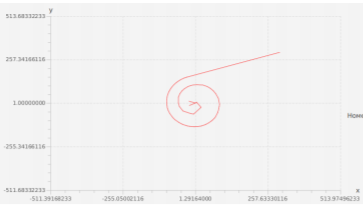


Рис. 41: Траектория поиска метода ДФП

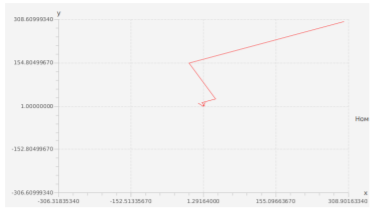


Рис. 42: Траектория поиска метода Ньютона с направлением спуска

Далее представлена таблица с зависимостью количества итераций от выбора начальной точки и метода.

|  |  |  |  |
| --- | --- | --- | --- |
| Метод | (1.29164, 0.99999) | (2, 2) | (300, 300) |
| Пауэлла | 3 | 5 | Неизвестно |
| Давидона – Флетчера – Пауэлла | 3 | 5 | 85 |
| Ньютона с направлением спуска | 2 | 4 | 9 |

3.3 Метод Марквардта

Далее будет исследована работа методов Марквадрта (обоих подходов) и метода Ньютона с направлением спуска (наилучшего, по нашему мнению, метода Ньютона) на многомерной функции Розенброка с n = 100:



Глобальный минимум находится в точке (1, 1, …, 1). Но существуют и другие локальные минимумы. В качестве eps бралось значения 10^-4.

**3.3.1 С увеличением диагональной прибавки**

Для первого подхода брались стартовые значения = 0.25,  = 1024.

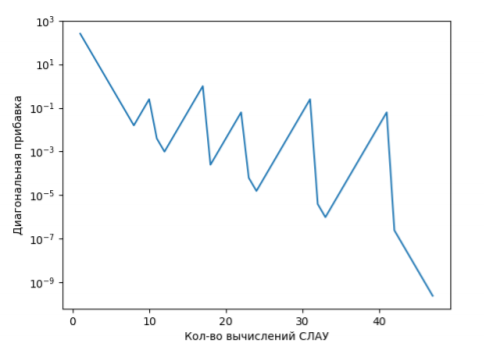


Рис. 43: Зависимость “итерация - ” с x0 = (4, 4, … , 4)

|  |  |
| --- | --- |
| Начальное приближение x0 | Кол-во вычислений СЛАУ |
| (4, 4, …, 4) | 47 |
| (2, 2, …, 2) | 15 |
| (0, 0, …, 0) | 2239 |
| -(2, 2, …, 2) | 2734 |
| -(4, 4, …, 4) | 2686 |
| -(6, 6, …, 6) | 2495 |

**3.3.2 С уменьшением диагональной прибавки**

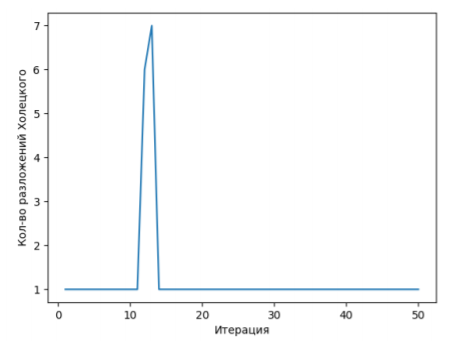
****

Рис. 44: Зависимость “итерация – число разложений Холецкого” с x0 = -(6, 6, …, 6) (показаны только первые 50 итераций, далее график остается константным)

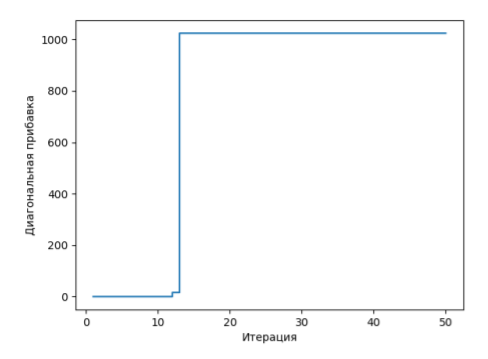


Рис. 45: Зависимость “итерация - ” с x0 = -(6, 6, …, 6) показаны только первые 50 итераций

|  |  |
| --- | --- |
| Начальное приближение x0 | Кол-во вычислений СЛАУ |
| (4, 4, …, 4) | 14 |
| (2, 2, …, 2) | 11 |
| (0, 0, …, 0) | 593164 |
| -(2, 2, …, 2) | 145 |
| -(4, 4, …, 4) | 37776 |
| -(6, 6, …, 6) | 19799 |

**3.3.3 Сравнение с методом Ньютона с направлением спуска**

|  |  |
| --- | --- |
| Начальное приближение x0 | Кол-во вычислений СЛАУ |
| (4, 4, …, 4) | 20 |
| (2, 2, …, 2) | 12 |
| (0, 0, …, 0) | 128 |
| -(2, 2, …, 2) | 141 |
| -(4, 4, …, 4) | 133 |
| -(6, 6, …, 6) | 138 |

Как видно из результатов, метод Марквардта с уменьшением диагональной прибавки делает стабильное количество итераций, но вычисляет СЛАУ на порядок раз больше, чем метод Ньютона с направлением спуска.

С другой стороны, метод Марквардта с увеличением диагональной прибавки очень сильно зависит от выбора начального приближения x0. Так как он на первых итерациях ведет себя также, как и классический метод Ньютона, то метод может начать расходиться от минимума, и из далеких точек начать делать маленькие шаги в сторону минимума.

4 Сравнение методов

**Глобальная сходимость**

Простая реализация метода Ньютона не обладает глобальной сходимостью. Т.е если выбирается начальная точка на слишком далеком расстоянии от x\*, то метод не сходится.

Модификация метода Ньютона с одномерным поиском и направлением спуска, а также метод Марквардта обладают глобальной сходимостью.

Квазиньютоновские методы, в частности метода Давидона-Флетчера-Пауэлла, обладают глобальной сходимостью в случае применения “рестарта” метода на каждой n итерации.

Из лабораторной по градиентным методам известно, что глобальная сходимость метода сопряженных градиентов обеспечивается при “рестарте” метода. При этом глобальная сходимость следует из глобальной сходимости метода наискорейшего спуска.

**Скорость сходимости**

* Методы сопряженных градиентов имеют сверх линейную скорость сходимости по n. В частности, если H(x) удовлетворяет условию Липшица, то эти методы имеют квадратичную скорость сходимости y = 2.
* Квазиньютоновские методы, в частности ДФП, имеют сверх линейную скорость сходимости. В частности, если H(x) удовлетворяет условию Липшица, то эти методы имеют квадратичную скорость сходимости y = 2.
* Метод Ньютона имеет квадратичную локальную сходимость, если H(x) удовлетворяет в окрестности т. x\* условию Липшица.
* Соответственно, можно пронаблюдать преимущество квазиньютоноских методов перед методами сопряженных градиентов, которые требуют в n раз больше итераци для одной и той же асимптотики. Однако нельзя забывать о промежуточных матричных вычислениях О(n^2) и использовании дополнительной памяти О(n^2)

Выводы

В ходе работы были изучены и реализованы методы Ньютона и его модификации, квазиньютоновские методы и метод Марквардта. Было проведено сравнение методов на различных матрицах.

Метод Ньютона и его модификации, а также метод Марквардта относятся к методам второго порядка. К основным преимуществам таких методов можно отнести:

* Квадратичная скорость сходимости

Недостатками методов второго порядка являются:

* Серьезные ограничения на функцию, она должна быть дважды непрерывно дифференцируема.
* Сложность вычисления матрицы вторых производных и решения СЛАУ с ее использованием (либо поиск обратной матрицы).

Стоит заметить, что метод Ньютона имеет квадратичную локальную сходимость, если H(x) удовлетворяет условию Липшица. В малой окрестность x м. Ньютона сходится быстрее остальных методов, т.е на практике возможно применение какого-то другого метода (например метода сопряженных градиентов, а при приближении к x\* использование метода Ньютона.