Санкт-Петербургский
Политехнический университет Петра Великого

Отчет по лабораторной работе №4
по дисциплине
"Методы оптимизации"

Решение задачи многомерной минимизации

Выполнили:

Марков Михаил Денисович
Аптуков Михаил Ильдусович
группа: 5030102/10201

Преподаватель:

Родионова Елена Александровна

Санкт-Петербург
2024

Содержание

1 Постановка задачи

Пусть имеется задача многомерной минимизации :

$$f(x) = x_1^2 + x_1 + 3x_2 + 4 + 1.5x_2^2 \longrightarrow \min_x$$

Необходимо:

1. Решить данную задачу градиентным методом 1-го порядка наискорейшего спуска с постоянным шагом.
2. Решить данную задачу методом 2-го порядка с выбором шага по методу Пшеничного.
3. Обосновать выбор условия окончания вычислений для метода одномерной минимизации по соответствующему условию градиентного метода. Показать справедливость своего вывода в ходе вычислительного эксперимента (для точности градиентного метода 0,01).
4. Решить данную задачу методом 2-го порядка с выбором шага по методу Пшеничного.
5. Решить данную задачу градиентным методом Девидона-Флетчера–Пауэлла второго порядка.
6. Выполнить сравнительный анализ алгоритмов методов.
7. Нарисовать линии уровня функции цели .

2 Описание алгоритмов

2.1 Градиентный метод 1-го порядка с постоянным шагом

Начальные параметры:

- f — заданная функция.
- $\varepsilon > 0$ — заданная точность.
- x_0 — начальное приближение.
- $0 < \alpha \leq 1$ — шаг.

Алгоритм k-го шага:

1. Вычисляется значение градиента в текущей точке $\nabla f(x_k)$.
2. Вычисляется следующее значение $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$.
3. Если $\|\nabla f(x_k)\| < \varepsilon$ — останавливаем алгоритм, иначе возвращаемся к шагу 1.

2.2 Градиентный метод 2-го порядка с выбором шага по методу Пшеничного

1. **Инициализация**: Задается начальное приближение x_0 и требуемая точность ε.
2. **Основной цикл**: Повторяются следующие шаги, пока не будет достигнут критерий остановки:
   (a) **Вычисление градиента**: Вычисляется градиент целевой функции $\nabla f(x_k)$ в текущей точке x_k.
   (b) **Решение системы линейных уравнений**: Решается система линейных уравнений методом LU-разложения $H(x_k)p_k = -\nabla f(x_k)$, где p_k - направление спуска.
   (c) **Определение оптимального шага**: Определяется оптимальный шаг α_k по методу Пшеничного, который минимизирует функцию на следующем шаге.
   (d) **Обновление переменных**: Обновляются переменные по формуле $x_{k+1} = x_k + \alpha_k p_k$.
3. **Критерий остановки**: Когда условие остановки выполнено, то есть $||\nabla f(x)|| < \epsilon$, возвращаем текущую точку x.

2.3 Метод Девидона - Флетчера - Пауэлла

1. **Начальный этап**:
   (a) Выбрать параметр $\varepsilon > 0$ - параметр окончания вычислений
   (b) Выбрать начальное приближение x_0
   (c) Положить $k = 1$
   (d) Положить для матрицы $A_1 = E$, E - соответствующая единичная матрица
   (e) Построить множество моментов обновления алгоритма $I_0 = \{n, 2n, ...\}$
   (f) Вычислить $\omega_1 = -gradf(x_0)$
2. **Основной этап**:
   (a) Найти направление спуска $p_k = A_k\omega_k$
   (b) Подобрать шаг $\alpha_k : min\{\phi_k(\alpha)\} = f(x_{k-1} + \alpha_k p_k)$ решив задачу одномерной минимизации.
   (c) Вычислить точку $x_k = x_{k-1} + \alpha_k p_k$ и $\omega_{k+1} = -gradf(x_k)$
   (d) Если $k \in I_0$, то положить $A_{k+1} = E$ и вернуться к шагу 2.(a)
   (e) Положить $\Delta x_k = x_k - x_{k-1}$ и $\Delta\omega_k = \omega_{k+1} - \omega_k$
   (f) Построить матрицу A_{k+1} по формуле и вернуться к шагу 2.(a):

$$A_{k+1} = A_k - \frac{\Delta x_k(\Delta x_k)^T}{\Delta\omega_k^T \Delta x_k} - \frac{A_k\Delta\omega_k(\Delta\omega_k)^T A_k^T}{\Delta\omega_k^T A_k \Delta\omega_k} \tag{1}$$

3. **Условие окончания вычислений:** $||\omega_k|| < \varepsilon$

3 Исследование применимости методов

Оба метода гарантированно приводят к решению, если функция имеет котловинный рельеф.

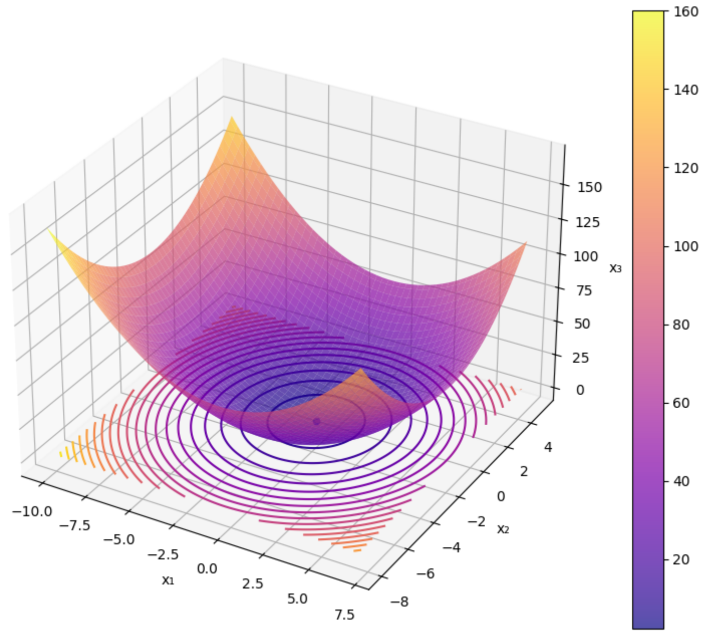


Рис. 1: График функции

Из графика видно, что заданная функция имеет котловинный рельеф, так как линии уровня представляют собой эллипсы. Кроме того, функция имеет единственный минимум.

Теорема. Пусть $f(x)$ — дважды непрерывно дифференцируемая функция, удовлетворяющая условию:

$$\exists 0 < m \leq M : m\|x\|^2 \leq x^T H(y)x \leq M\|x\|^2 \ \forall x, y. \tag{2}$$

Тогда последовательность $\{x_k\} \xrightarrow[k\to\infty]{} x^*$, $f(x_k) \xrightarrow[k\to\infty]{} f(x^*)$ при любом начальном приближении x_0.

Исследуем заданную функцию $f(x) = x_1^2 + x_1 + 3x_2 + 4 + 1.5x_2^2$:

$$\frac{\partial f}{\partial x_1} = 2x_1 + 1 \tag{3}$$

$$\frac{\partial f}{\partial x_2} = 3x_2 + 3 \tag{4}$$

$$\frac{\partial^2 f}{\partial x_1^2} = 2 \tag{5}$$

$$\frac{\partial^2 f}{\partial x_2^2} = 3 \tag{6}$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = 0 \tag{7}$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = 0 \tag{8}$$

$$H = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{vmatrix} \tag{9}$$

$$\mathbf{det} H = 2 * 3 - 0 * 0 = 6 \tag{10}$$

$$\begin{cases} \frac{\partial^2 f}{\partial x_1^2} > 0 \\ \mathbf{det} H > 0 \end{cases} \tag{11}$$

Следовательно, по критерию Сильвестра, матрица Гессе положительно определена.

Теперь найдем m, M как наименьшее и наибольшее собственные числа матрицы гессиана.

$$|H - \lambda E| = 0 \tag{12}$$

$$\begin{vmatrix} 2-\lambda & 0 \\ 0 & 3-\lambda \end{vmatrix} = 0 \tag{13}$$

$$|H - \lambda E| = (2-\lambda)(3-\lambda) = 0 \tag{14}$$

$$m = \lambda_1 = 2; \ M = \lambda_2 = 3 \tag{15}$$

Положительная определенность Гессиана и то, что m, M – его собственные числа, дают нам выполнение неравенства (1). Следовательно, условия теоремы выполнены,

методы применимы к данной задаче и сходятся к оптимальному решению.
Теперь найдем константу q для оценки линейной сходимости метода наискорейшего спуска.

$$\alpha = 1 - \frac{m}{2M}\left(1 + \frac{m}{M}\right) = 1 - \frac{2}{6}\left(1 + \frac{2}{3}\right) = \frac{5}{9} = 0.444 \tag{16}$$

4 Решение задачи

Для оценки точности полученных решений, найдем точное решение исходной задачи минимизации:
$f(x) = f(x) = x_1^2 + x_1 + 3x_2 + 4 + 1.5x_2^2$
$f(x) = (x_1 + 0.5)^2 + 1.5(x_2 + 1)^2 + 2$
$\nabla f(x^*) = 0$
$x^* = (-0.5, -1)$
$f(x^*) = 2$
Продемонстрируем результаты вычислительного эксперимента в виде графиков и таблиц:

Заданная точность	Ошибка	Точка минимума	Итерации
10^{-1}	$4.1*10^{-2}$	(-0.4999915 -1.0041070)	5
10^{-2}	$4.6*10^{-3}$	(-0.4999998 -1.0004559)	7
10^{-3}	$7.21*10^{-8}$	(-0.4999999 -1.0000506)	9

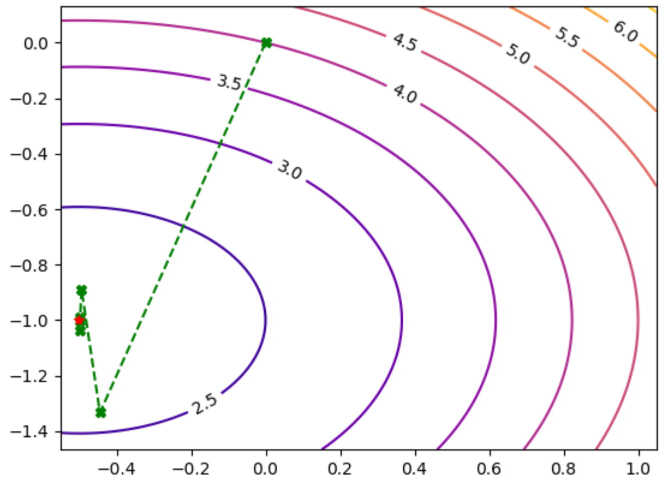
Таблица 1: Результат градиентного метода 1-го порядка с постоянным шагом по методу Фибоначчи

Заданная точность	Ошибка	Точка минимума	Итерации
10^{-1}	0	(0.5000000 -1.0000000)	1
10^{-2}	0	(0.5000000 -1.0000000)	1
10^{-3}	0	(0.5000000 -1.0000000)	1

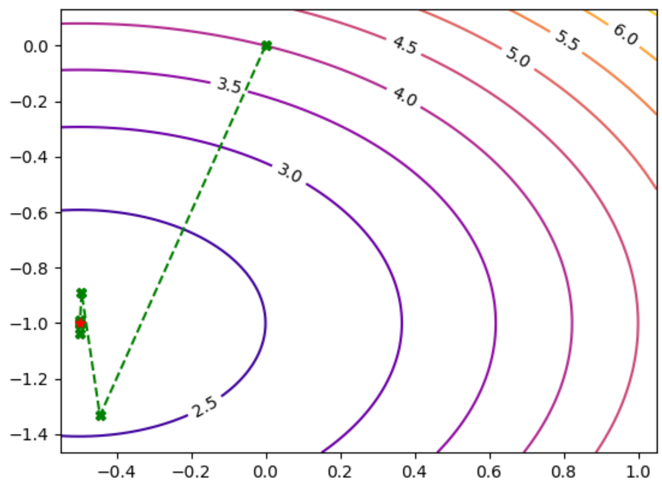
Таблица 2: Результат градиентного метода 2-го порядка с выбором шага по методу Пшеничного

Заданная точность	Ошибка	Точка минимума	Итерации
10^{-1}	$4.5*10^{-6}$	(-0.50000075 -0.99999557)	3
10^{-2}	$1.7*10^{-6}$	(-0.50000169 -0.99999972)	4
10^{-3}	$1.7*10^{-6}$	(-0.50000169 -0.99999972)	4

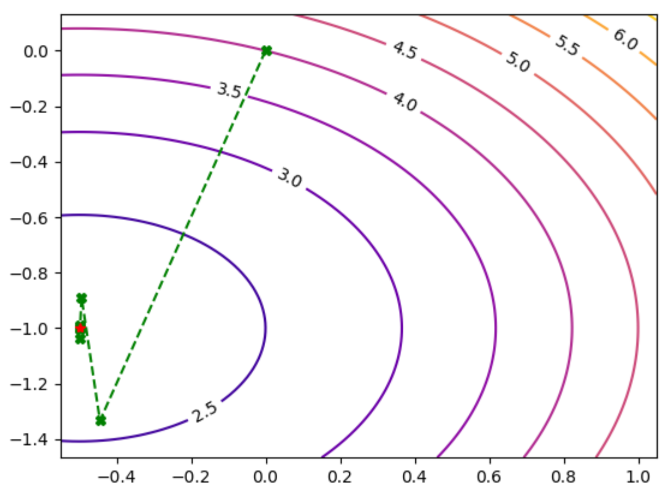
Таблица 3: Результат градиентного метода ДФП с выбором шага по методу Фибоначчи



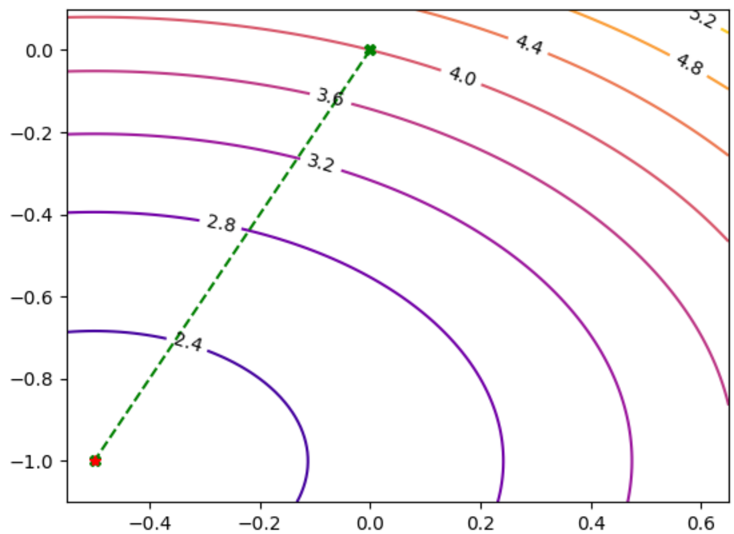
Метод 1 порядка,$\epsilon = 10^{-1}$



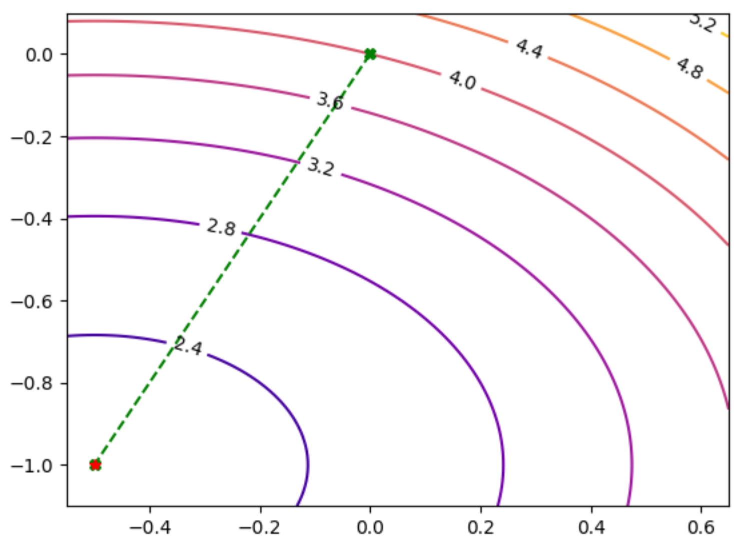
Метод 1 порядка,$\epsilon = 10^{-2}$



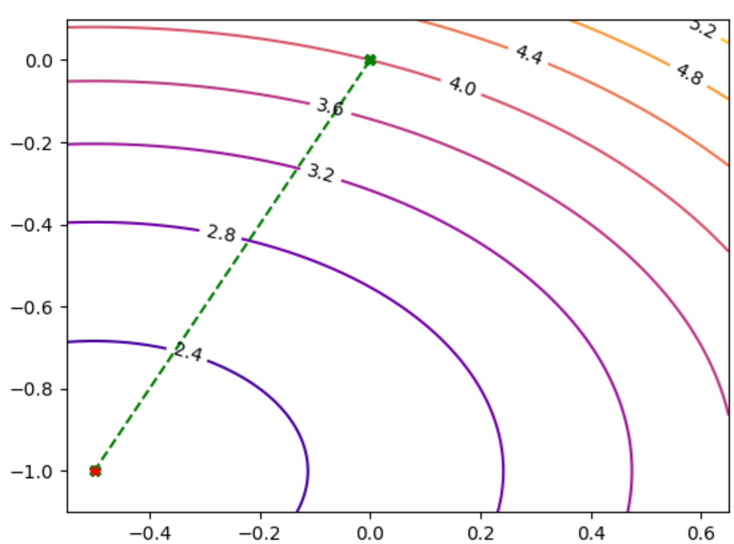
Метод 1 порядка,$\epsilon = 10^{-3}$



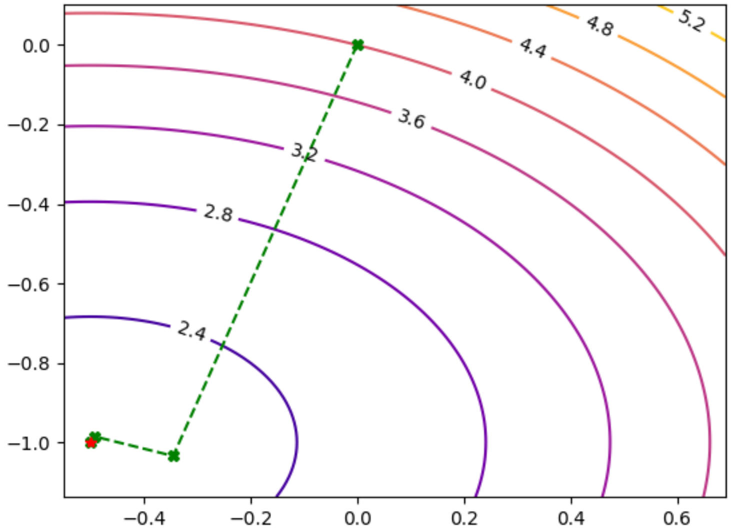
Метод 2 порядка,$\epsilon = 10^{-1}$



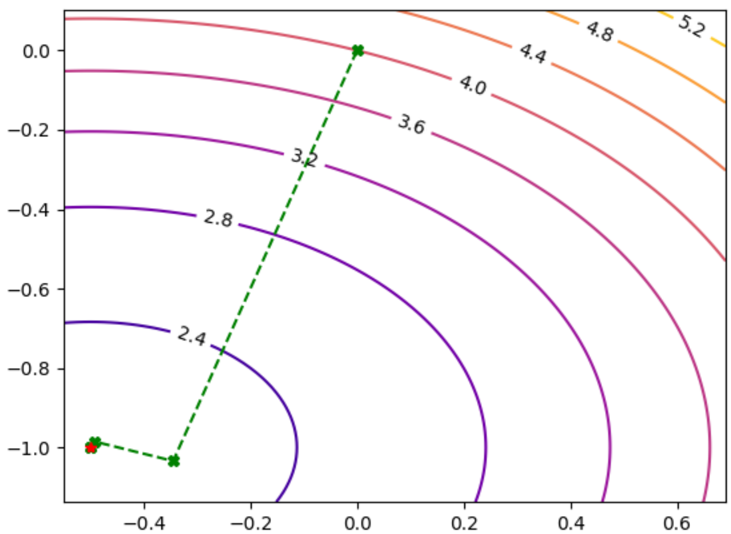
Метод 2 порядка,$\epsilon = 10^{-2}$



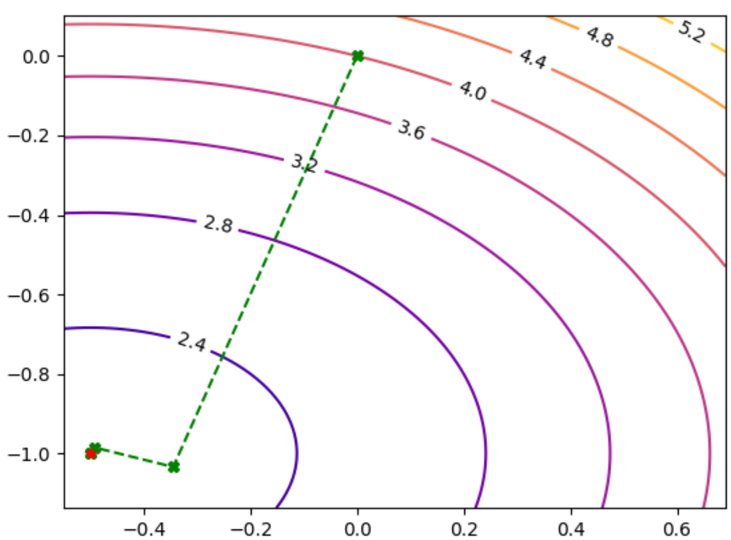
Метод 2 порядка,$\epsilon = 10^{-3}$



ДФП, $\epsilon = 10^{-1}$



ДФП, $\epsilon = 10^{-2}$



ДФП, $\epsilon = 10^{-3}$

5 Сравнительный анализ методов

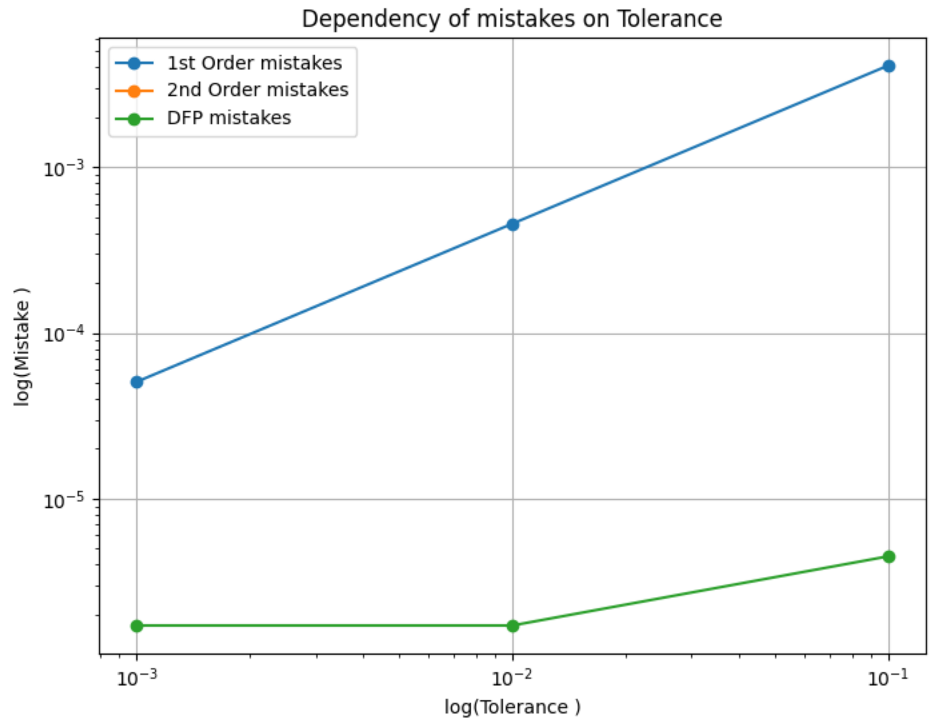
Метод наискорейшего спуска и метод Девидона-Флетчера–Пауэлла второго порядка относятся к методам оптимизации без ограничений. Они оба основаны на градиентном спуске, но метод Девидона-Флетчера–Пауэлла второго порядка учитывает информацию о кривизне функции с помощью матрицы Гессе.

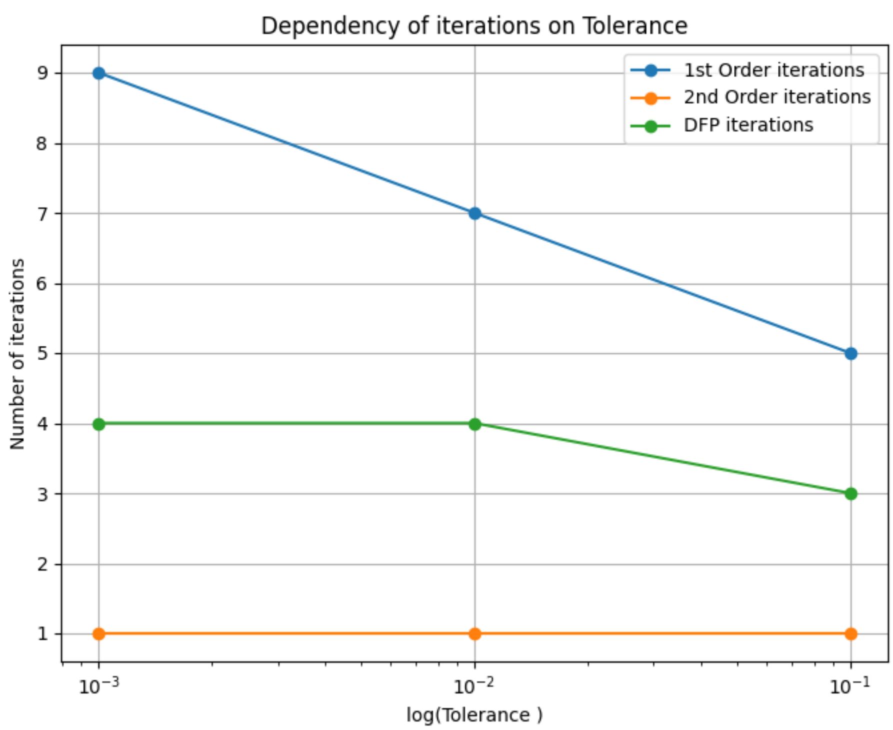
Основное различие между методами заключается в том, что метод наискорейшего спуска использует только первые производные функции (градиент), а метод Девидона-Флетчера–Пауэлла второго порядка использует еще и вторые производные (матрицу Гессе). Это позволяет методу Девидона-Флетчера–Пауэлла второго порядка более эффективно и быстро сходиться к оптимальному решению.

Однако, метод Девидона-Флетчера–Пауэлла второго порядка имеет некоторые ограничения. В частности, вычисление и хранение матрицы Гессе может быть затруднительным для больших размерностей задачи оптимизации. Кроме того, приближение матрицы Гессе может быть неточным, что может приводить к неустойчивости метода.

Таким образом, метод наискорейшего спуска может быть предпочтительнее для задач с большим количеством переменных или при ограниченной вычислительной мощности. Метод Девидона-Флетчера–Пауэлла второго порядка может быть эффективнее для задач с небольшим количеством переменных и точным приближением матрицы Гессе.

Метод градиентного спуска второго порядка может сходиться за одну итерацию, потому что исходная функция является квадратичной функцией и ее гессиан постоянен.





6 Дополнительные исследования

По графику линий уровня функции и градиентной ломаной для метода наискорейшего спуска видим общую закономерность в том, что шаг уменьшается по мере приближения к оптимальной точке.
Из данных графиком мы видим, что и метод наискорейшего спуска, и ДФП-метод достигают точности на порядок выше заданной, при этом количество итераций для метода

наискорейшего спуска растет линейно, в то время как для ДФП-метода зависимость более пологая. Метод градиентного спуска второго порядка сходится за одну итерацию, потому что исходная функция является квадратичной функцией и ее гессиан постоянен.

7 Выводы

Важнейшими характеристиками градиентных методов различных порядков используемых производных являются скорость и порядок сходимости метода, а также их эффективность, то есть количество обращений к функции цели при достижении заданной точности вычислений.

Выбор между этими методами обычно зависит от того, что более важно для конкретной задачи: скорость сходимости, требования к памяти, возможность хранения и вычисления гессиана, и так далее. В общем случае, если решается задача многомерной минимизации, где вычисления не слишком затратны, и хранение и вычисление гессиана не представляют проблему, метод второго порядка может быть предпочтительным из-за своей быстрой сходимости. Однако, если есть ограничения на вычислительные ресурсы или возможность хранения гессиана, квазиньютоновский метод может быть более подходящим выбором.

Стоит сказать, что методы первого порядка дают приближение к окрестности оптимальной точки, а методы второго порядка - уточняют ее.

8 Вычисленная оценка

8.1 Градиентный метод наискорейшего спуска

Произведем оценку вычислительного эксперимента по методу 1-го порядка.

Будем оценивать следующие значения, получаемые в ходе итерационного процесса оптимизации с помощью градиентного метода наискорейшего спуска для точности 0.01:

$m(1 + m/M)(f(x_k) - f(x_*))$

$m||x_k - x_*||$

$||\nabla f(x_k)||$

Согласно теореме о скорости сходимости градиентного метода выполним оценку, а именно выполнение следующих неравенств:

$m(1 + m/M)(f(x_k) - f(x_*) < m||x_k - x_*|| < ||\nabla f(x_k)||,$

где $f(x) = x_1^2 + x_1 + 3x_2 + 4 + 1.5x_2^2,$

$\nabla f(x) = (2x_1 + 1, 3x_2 + 3),$

$m = 2$, $M = 3$

Полученные результаты представим в виде таблицы:

Итерация	$m(1 + m/M)(f(x_k) - f(x_*)$	$m\|\|x_k - x_*\|\|$	$\|\|\nabla f(x_k)\|\|$	Выполнение неравенств
1	0.37	0.66	6	Да
2	2.1*10^{-2}	1.63*10^{-1}	1.63*10^{-1}	Да
3	1.85*10^{-3}	4.56*10^{-2}	1.63*10^{-1}	Да
4	1.46*10^{-4}	1.35*10^{-2}	1.63*10^{-1}	Да
5	1.46*10^{-4}	1.35*10^{-2}	1.63*10^{-1}	Да
6	1.46*10^{-4}	1.35*10^{-2}	1.63*10^{-1}	Да
7	1.46*10^{-4}	1.35*10^{-2}	1.63*10^{-1}	Да

Таблица 4: Оценка вычислительного эксперимента

В ходе вычислительного эксперимента, теоретическая оценка выполняется на каждой итерации, что свидетельствует о корректной работе алгоритма

8.2 Градиентный метод Девидона-Флетчера-Пауэлла

Произведем оценку вычислительного эксперимента по методу 2-го порядка.

Будем оценивать значения норм разности значения на текущей итерации и точного решения, получаемые в ходе итерационного процесса оптимизации с помощью градиентного метода ДФП для точности 0.001:

Согласно теореме о сходимости, для градиентного метода второго порядка выполняется следующая оценка:

$||x_k - x_*|| < ||x_{k-1} - x_*||^2$

Произведем данную оценку для итерационного процесса градиентного метода ДФП для нашей задачи, вычисляя следующие значения на каждой итерации и оценивая единицей сверху:
$||x_k - x_*||\,||\mathrm{x}_{k-1} - x_*||^2 < 1$ Полученные результаты представим в виде таблицы:

Итерация	$\|\|x_k - x_*\|\|\,\|\|\mathrm{x}_{k-1} - x_*\|\|^2$
1	$2.2*10^{-2}$
2	$1.7*10^{-3}$
3	$1.2*10^{-9}$
4	$6.6*10^{-7}$

Таблица 5: Оценка вычислительного эксперимента

В ходе вычислительного эксперимента был получен корректный результат, а именно - на каждой итерации исследуемое соотношение меньше единицы, что согласуется с теоремой и свидетельствует о корректной работе алгоритма