Методы оптимизации, Лабораторная работа №1

Кирилл Кадомцев и Андрей Крюков

Апрель 2025

Содержание

1. Описание методов

Для градиентного спуска были реализованы следующие методы выбора шага:

- Постоянный
- Кусочно-постоянный
- Наискорейший спуск (золотое сечение)
- Наискорейший спуск (дихотомия)

Для удобства и расширяемости, методы реализовывались так, чтоб можно было в дальнейшем использовать их для любых размерностей (исходя из предположения, что это может стать объектом исследования в дальнейших лабораторных работах). Для стратегий был создан специальный интерфейс с методом *step*, позволяющий в дальнейшем расширить список реализованных методов.

2. Тестирование

Для тестирования были выбраны несколько функций с различными точками минимума. Также, была использованна функция Химмельблау, поскольку она мультимодальная.

1. **Параболоид**
    - $f(x, y) = x^2 + y^2$
    - $\nabla f(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$
2. **Эллипс**
    - $f(x, y) = 4x + y$
    - $\nabla f(x, y) = \begin{pmatrix} 8x \\ 2y \end{pmatrix}$
3. **Функция Розенброка**
    - $f(x, y) = (1 - x)^2 + 100(y - x^2)^2$
    - $\nabla f(x, y) = \begin{pmatrix} -2(1 - x) - 400x(y - x^2) \\ 200(y - x^2) \end{pmatrix}$
4. **Квадратичная форма (3, -2)**
    - $f(x, y) = (x - 3)^2 + (y + 2)^2$
    - $\nabla f(x, y) = \begin{pmatrix} 2(x - 3) \\ 2(y + 2) \end{pmatrix}$
5. **Квадратичная форма (2, -1)**
    - $f(x, y) = x + y$
    - $\nabla f(x, y) = \begin{pmatrix} 10(x - 2) \\ 6(y + 1) \end{pmatrix}$
6. **Функция Химмельблау**
    - $f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2$
    - $\nabla f(x, y) = \begin{pmatrix} 4x(x^2 + y - 11) + 2(x + y^2 - 7) \\ 2(x^2 + y - 11) + 4y(x + y^2 - 7) \end{pmatrix}$
    - глобальные минимумы: $\begin{cases} (3.0,\ 2.0) \\ (-2.805118,\ 3.131312) \\ (-3.779310,\ -3.283186) \\ (3.584428,\ -1.848126) \end{cases}$

Ниже приведены результаты оптимизации в зависимости от стратегии выбора шага. В данном случае *(nan, nan)* является результатом переполнения, то есть провалом поиска минимума.

Функция	**argmin**	**Пост. шаг**	**Кус.-пост. шаг**	**Наиск. спуск**
paraboloid	$(0,\ 0)$	$(0,\ 0)$	$(0,\ 0)$	$(0,\ 0)$
ellipse	$(0,\ 0)$	$(0,\ 0)$	$(0,\ 0)$	$(0,\ 0)$
rosenbrock	$(1,\ 1)$	(nan, nan)	(nan, nan)	$(0.9991,\ 0.9982)$
min3m2	$(3,\ -2)$	$(2.9996,\ -1.9997)$	$(2.9996,\ -1.9997)$	$(3.0000,\ -2.0000)$
min2m1	$(2,\ -1)$	$(2.0000,\ -0.9999)$	$(2.0000,\ -0.9999)$	$(0.5932,\ -0.1903)$
himmelblau	см. список	(nan, nan)	(nan, nan)	$(3.5844,\ -1.8481)$
noisy rosenbrock	$(1,\ 1)$	(nan, nan)	(nan, nan)	(nan, nan)

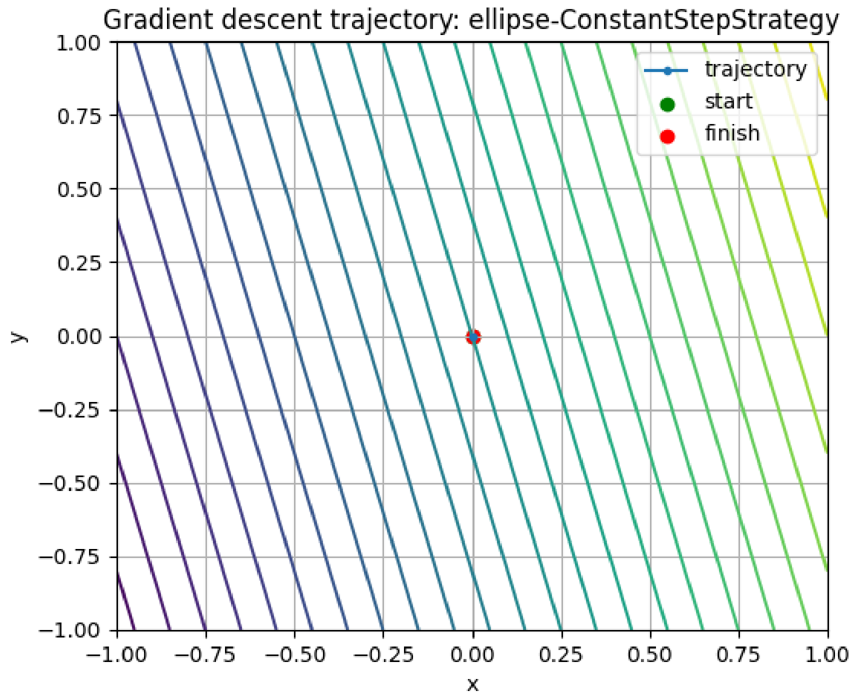
Таблица 1: Сравнение точек минимума для разных стратегий шага

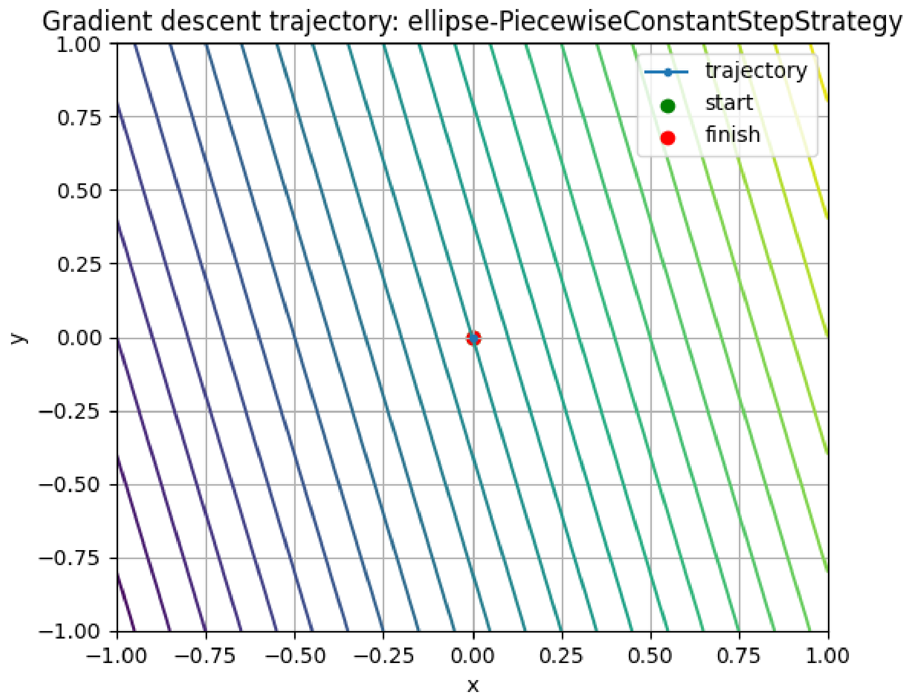
Функция	**argmin**	**Пост. шаг**	**Кус.-пост. шаг**	**Наиск. спуск**
paraboloid	$(0,\ 0)$	1	1	1
ellipse	$(0,\ 0)$	1	1	1
rosenbrock	$(1,\ 1)$	10000	10000	5723
min3m2	$(3,\ -2)$	45	85	2
min2m1	$(2,\ -1)$	10	25	10000
himmelblau	см. список	10000	10000	12
noisy rosenbrock	$(1,\ 1)$	10000	10000	10000

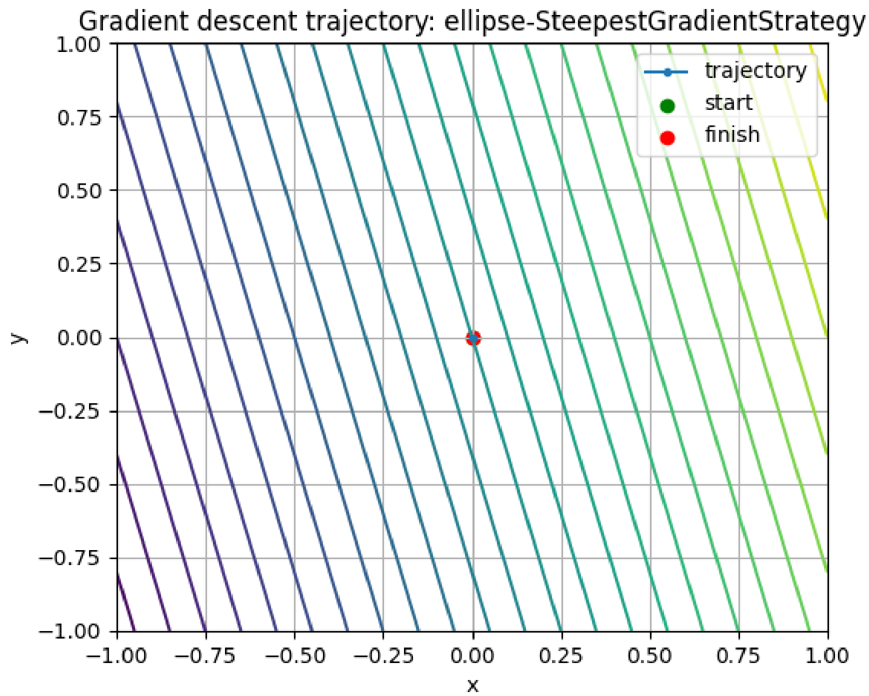
Таблица 2: Сравнение количества итераций для разных стратегий шага

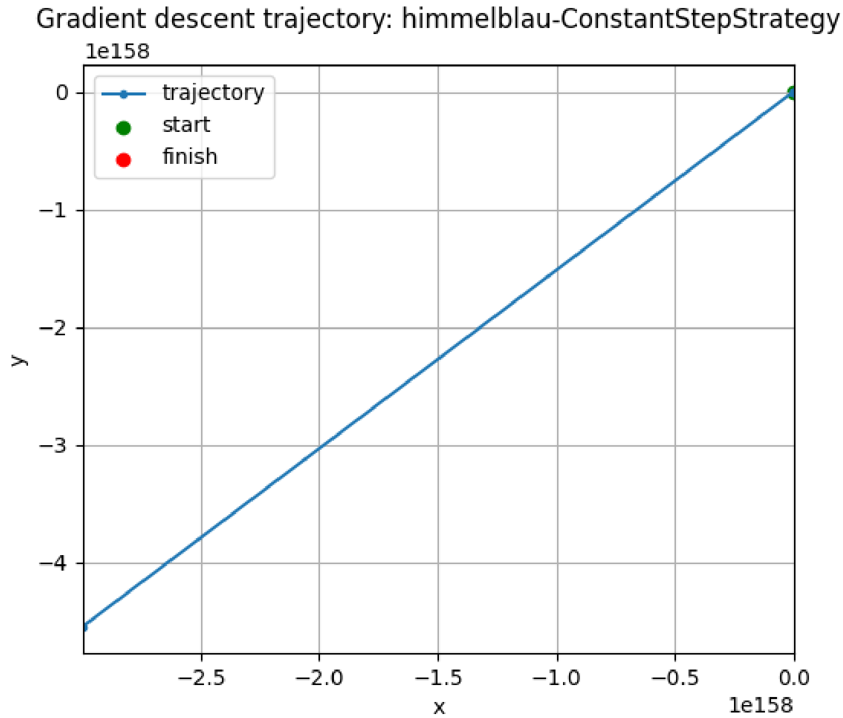
3. Графики

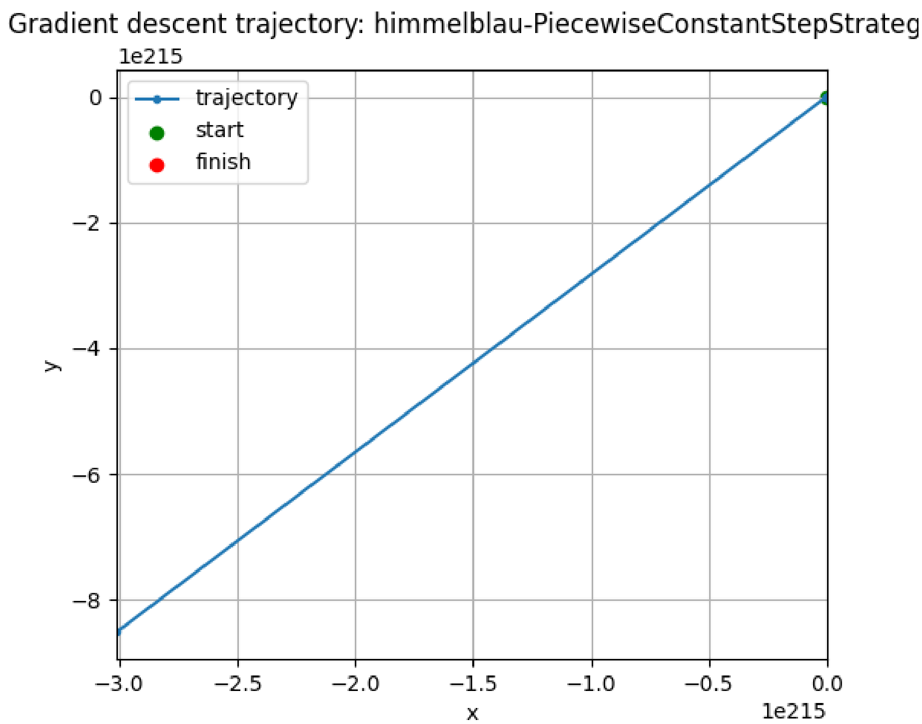
Не будем приводить все графики ввиду избыточности. Графики для эллипса демонстрируют, что точка минимума $(0, 0)$ находится за 1 итерацию. Графики для функции Химмельблау показывают несовершенство постоянного и кусочно-постоянного методов для мультимодальной функции. Для квадратичной функции графики приведены для демонстрации работы корректности работы в случае, если минимум отличен от нуля.

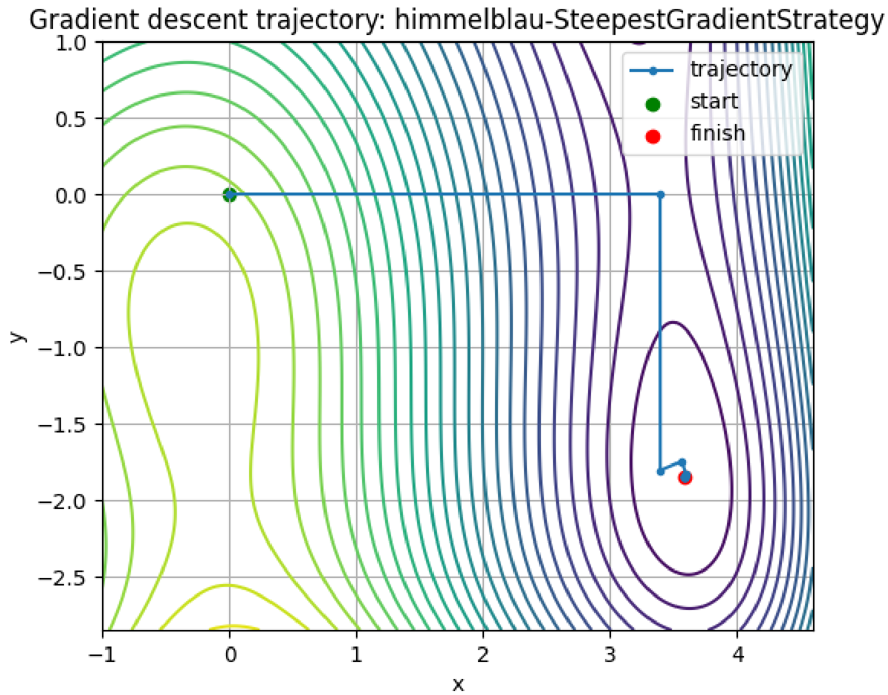

Gradient descent trajectory: ellipse-ConstantStepStrategy
trajectory
start
finish
x
y
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00


Gradient descent trajectory: ellipse-PiecewiseConstantStepStrategy
trajectory
start
finish
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
y
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
x

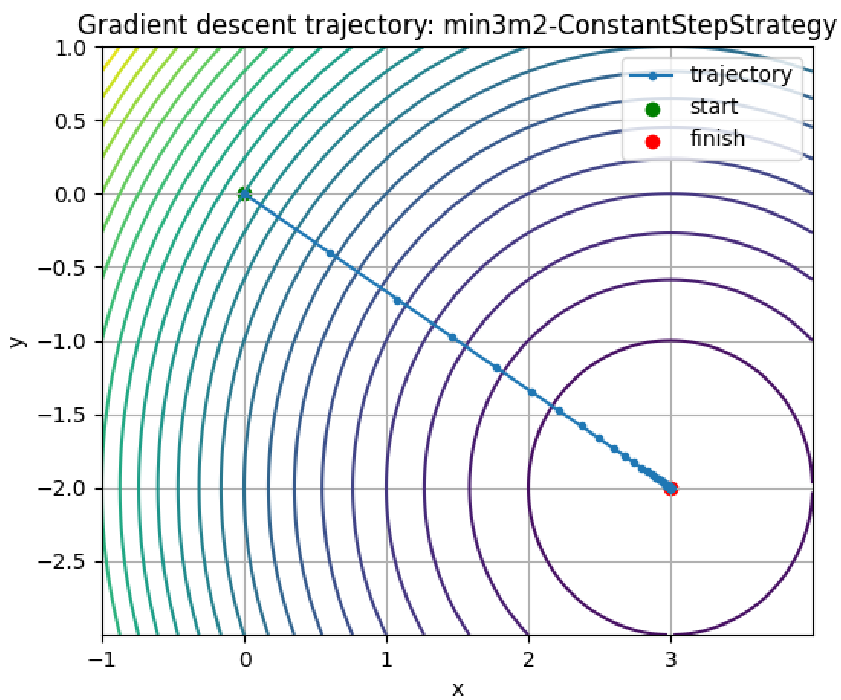

Gradient descent trajectory: ellipse-SteepestGradientStrategy
trajectory
start
finish
x
y
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00

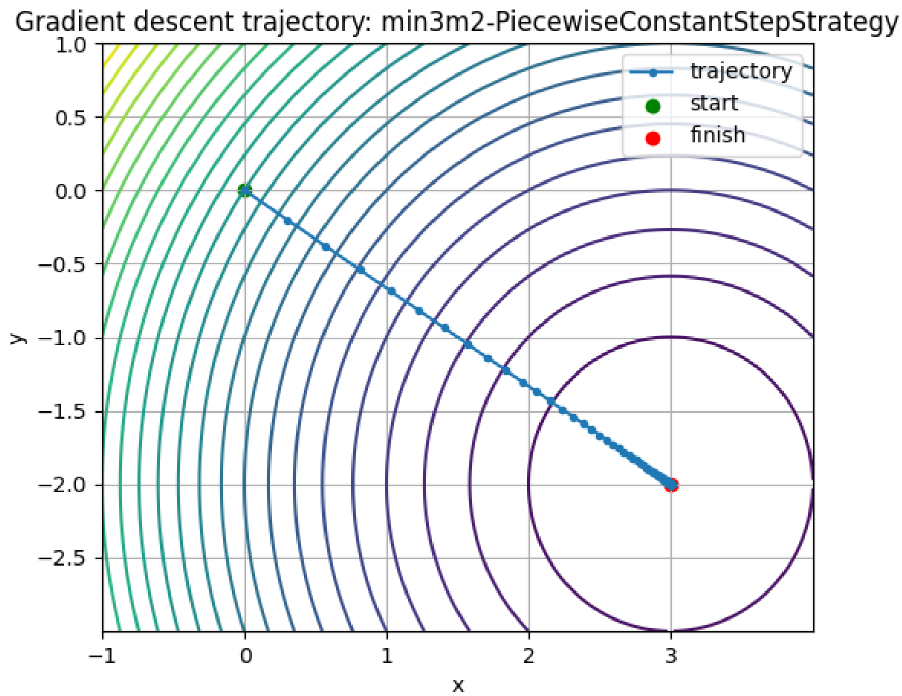

Gradient descent trajectory: himmelblau-ConstantStepStrategy
1e158
trajectory
start
finish
0
−1
−2
−3
−4
y
−2.5
−2.0
−1.5
−1.0
−0.5
0.0
x
1e158

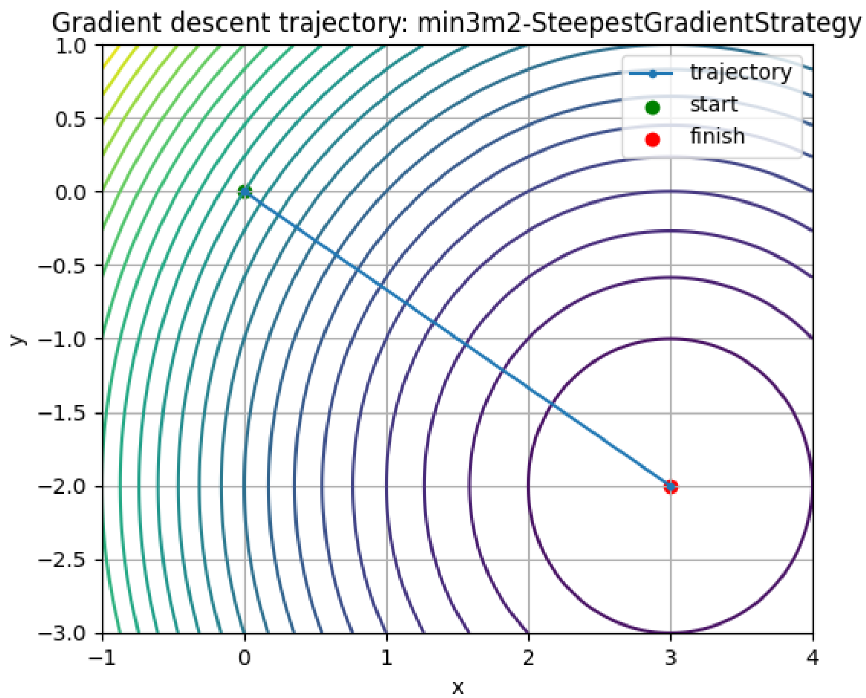

Gradient descent trajectory: himmelblau-PiecewiseConstantStepStrateg
1e215
trajectory
start
finish
0
−2
−4
−6
−8
y
−3.0
−2.5
−2.0
−1.5
−1.0
−0.5
0.0
x
1e215



Gradient descent trajectory: himmelblau-SteepestGradientStrategy
trajectory
start
finish
x
y
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−1
0
1
2
3
4


Gradient descent trajectory: min3m2-ConstantStepStrategy
trajectory
start
finish
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
y
−1
0
1
2
3
x


Gradient descent trajectory: min3m2-PiecewiseConstantStepStrategy
trajectory
start
finish
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
y
−1
0
1
2
3
x



Gradient descent trajectory: min3m2-SteepestGradientStrategy
trajectory
start
finish
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
y
−1
0
1
2
3
4
x