

Отчёт по лабораторной работе
Метод градиентного спуска

1 Постановка задачи

В работе реализуется метод градиентного спуска для минимизации следующих функций:

$$\begin{aligned} J_1(x, y) &= x^2 + y^2, \\ J_2(x, y) &= \frac{x^2}{a^2} + \frac{y^2}{b^2}, \\ J_3(x, y) &= 70(x - 1)^2 + (y - 1)^2 + 1, \\ J_4(x, y) &= (x - 8)^2 + (y - 1)^2 + 70(y + (x - 8)^2 - 1)^2 + 1, \end{aligned}$$

где параметры a, b задают большую и малую полуоси эллипса во второй функции.

Метод градиентного спуска настраивается по:

- начальной точке $x_0 = (x_0, y_0)$;
- режиму выбора шага h_k ;
- значению шага h (или параметрам c, α для убывающих шагов);
- критерию остановки: по числу итераций, по норме разности векторов $\|x_{k+1} - x_k\|$, по изменению функционала $|J(x_{k+1}) - J(x_k)|$.

2 Описание метода

На k -й итерации метода градиентного спуска вычисляется шаг

$$x_{k+1} = x_k - h_k \nabla J(x_k),$$

где h_k выбирается в одном из следующих режимов:

- постоянный шаг: $h_k = h$;
- убывающий шаг вида

$$h_k = c(1 + k)^{-\alpha}, \quad c > 0, \quad 0.5 < \alpha \leq 1;$$

- шаг вида

$$h_k = \frac{c}{k + 1}, \quad c > 0,$$

удовлетворяющий условиям

$$h_k > 0, \quad \sum_{k=0}^{\infty} h_k = \infty, \quad \sum_{k=0}^{\infty} h_k^2 < \infty.$$

Критерии остановки:

- ограничение сверху на число итераций k_{\max} ;
- достижение малой нормы приращения аргумента:

$$\|x_{k+1} - x_k\| < \varepsilon_x;$$

- достижение малой разности значений функционала:

$$|J(x_{k+1}) - J(x_k)| < \varepsilon_f.$$

В реализованном коде используются значения

$$\varepsilon_x = 10^{-6}, \quad \varepsilon_f = 10^{-6},$$

а максимальное число итераций выбирается равным $k_{\max} = 20000$.

3 Численные эксперименты

Для каждой функции рассматриваются два режима шага:

- постоянный шаг ($h_k = h$);
- убывающий шаг вида $h_k = c(1 + k)^{-\alpha}$.

Отдельно демонстрируется режим $h_k = c/(k + 1)$ на функции J_1 .

Начальные точки и значения шагов:

В численных экспериментах для функции J_2 использовались значения $a = 5$, $b = 1$, что соответствует вытянутому эллипсу уровней вдоль оси x .

- J_1 : $x_0 = (5, 5)$, постоянный шаг $h = 0.2$, убывающий шаг $h_k = 0.5(1 + k)^{-0.6}$;
- J_2 : $x_0 = (5, 3)$, постоянный шаг $h = 0.2$, убывающий шаг $h_k = 0.5(1 + k)^{-0.6}$;
- J_3 : $x_0 = (-5, 5)$, постоянный шаг $h = 0.01$, убывающий шаг $h_k = 0.5(1 + k)^{-0.9}$;
- J_4 : $x_0 = (8, 3)$, постоянный шаг $h = 0.001$, убывающий шаг $h_k = 0.05(1 + k)^{-0.9}$;
- демонстрация шага $h_k = \frac{0.5}{k + 1}$ для J_1 .

3.1 Функция $J_1(x, y) = x^2 + y^2$

Траектории на плоскости

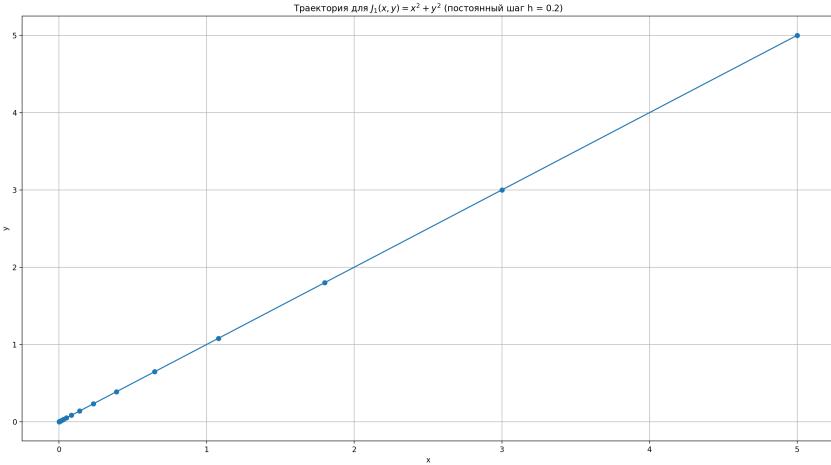


Рис. 1: Траектория градиентного спуска для J_1 при постоянном шаге (файл f1_traj_constant.png).

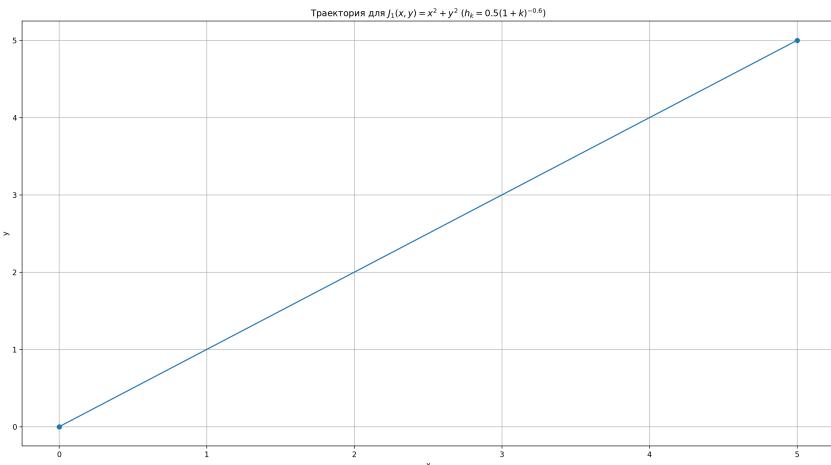


Рис. 2: Траектория градиентного спуска для J_1 при убывающем шаге $h_k = 0.5(1 + k)^{-0.6}$ (файл f1_traj_power.png).

Если используется демонстрация режима $h_k = 0.5/(k + 1)$, то соответствующий график траектории:

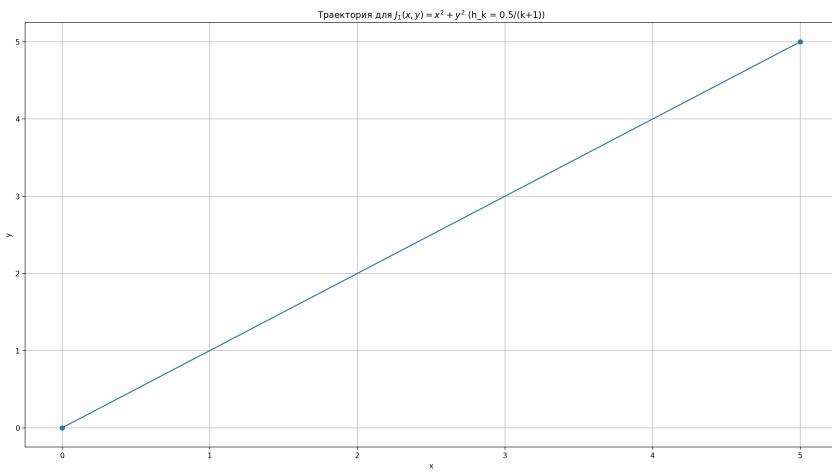


Рис. 3: Траектория градиентного спуска для J_1 при шаге $h_k = 0.5/(k+1)$ (файл f1_traj_1_over_k.png).

Убывание значения функционала $J_1(x_k)$

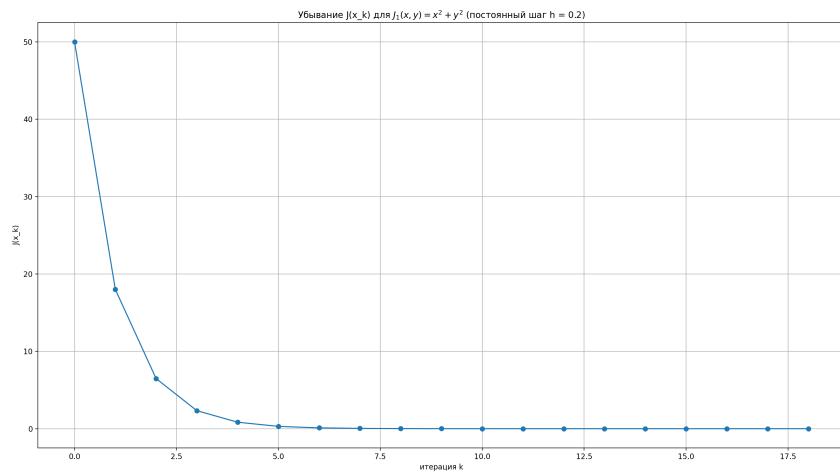


Рис. 4: График убывания $J_1(x_k)$ при постоянном шаге (файл f1_J_constant.png).

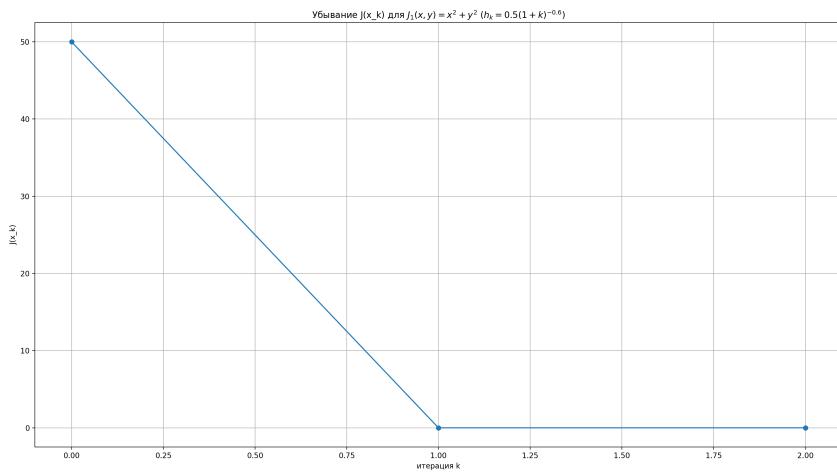


Рис. 5: График убывания $J_1(x_k)$ при убывающем шаге $h_k = 0.5(1 + k)^{-0.6}$ (файл f1_J_power.png).

При необходимости:

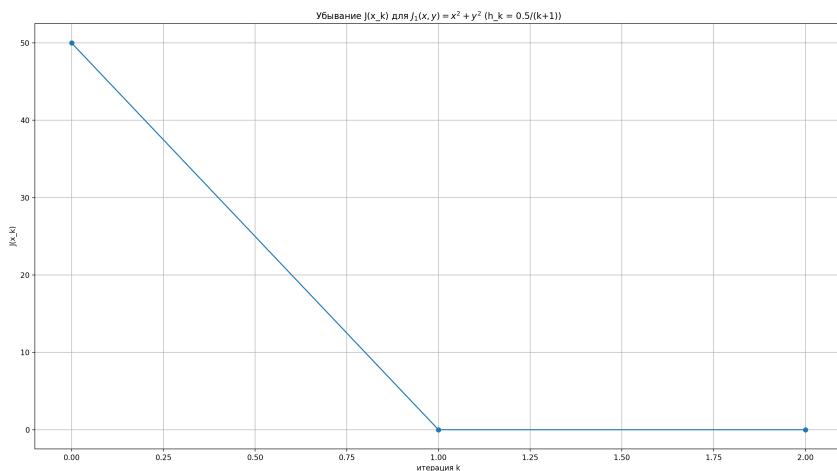


Рис. 6: График убывания $J_1(x_k)$ при шаге $h_k = 0.5/(k+1)$ (файл f1_J_1_over_k.png).

3.2 Функция $J_2(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$

Траектории на плоскости

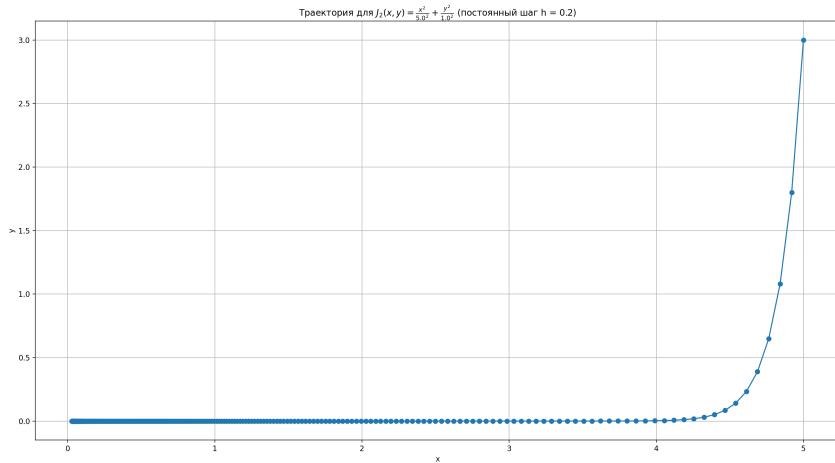


Рис. 7: Траектория градиентного спуска для J_2 при постоянном шаге (файл f2_traj_constant.png).

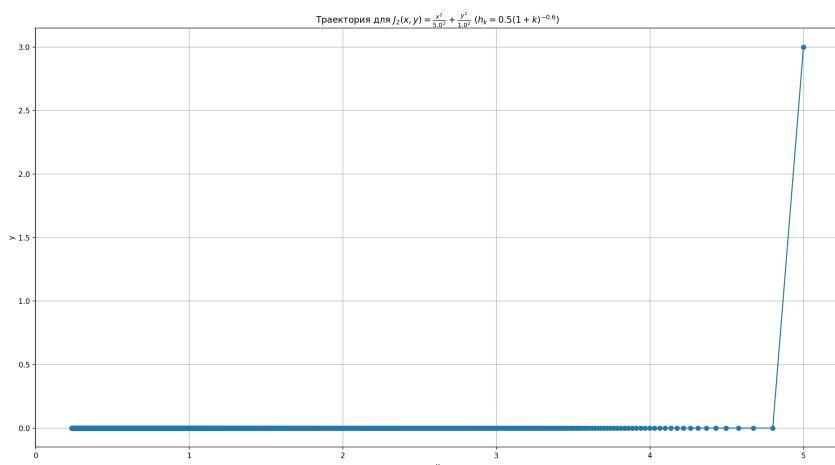


Рис. 8: Траектория градиентного спуска для J_2 при убывающем шаге $h_k = 0.5(1+k)^{-0.6}$ (файл f2_traj_power.png).

Убывание $J_2(x_k)$

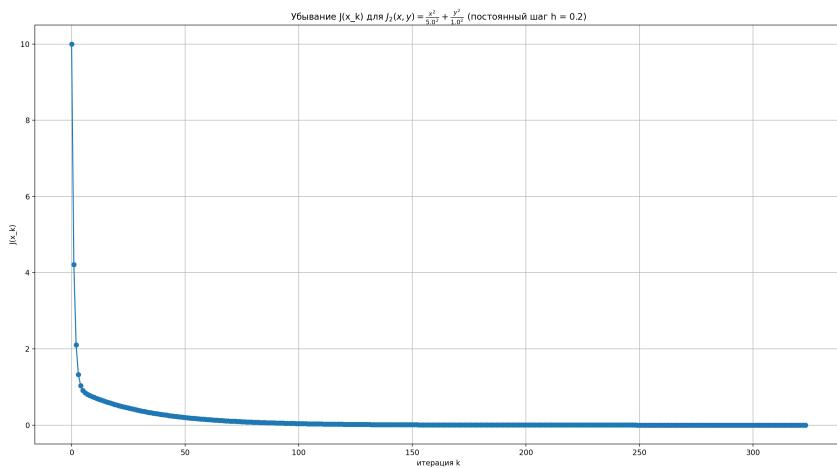


Рис. 9: График убывания $J_2(x_k)$ при постоянном шаге (файл f2_J_constant.png).

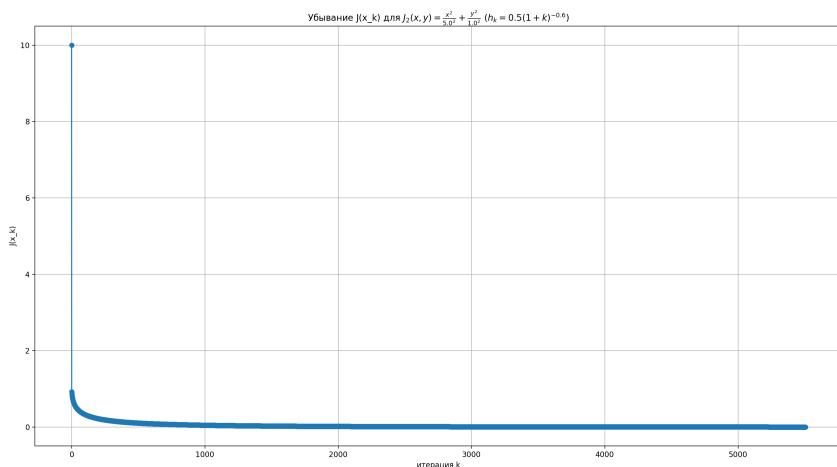


Рис. 10: График убывания $J_2(x_k)$ при убывающем шаге (файл f2_J_power.png).

3.3 Функция $J_3(x, y) = 70(x - 1)^2 + (y - 1)^2 + 1$

Траектории на плоскости

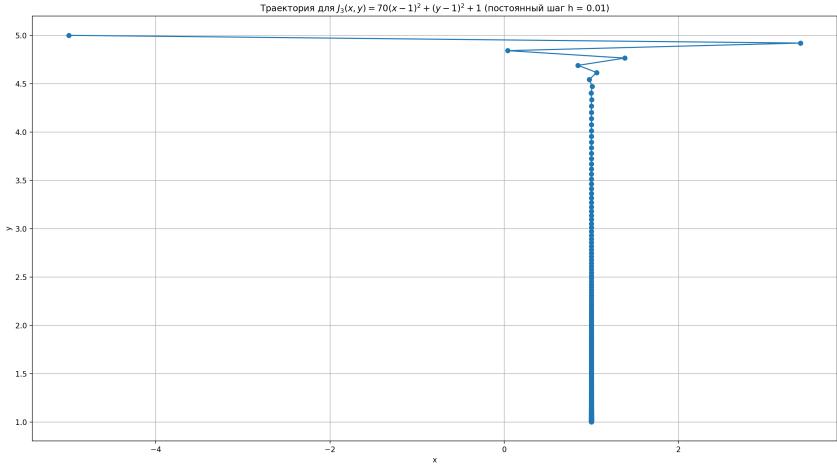


Рис. 11: Траектория градиентного спуска для J_3 при постоянном шаге (файл f3_traj_constant.png).

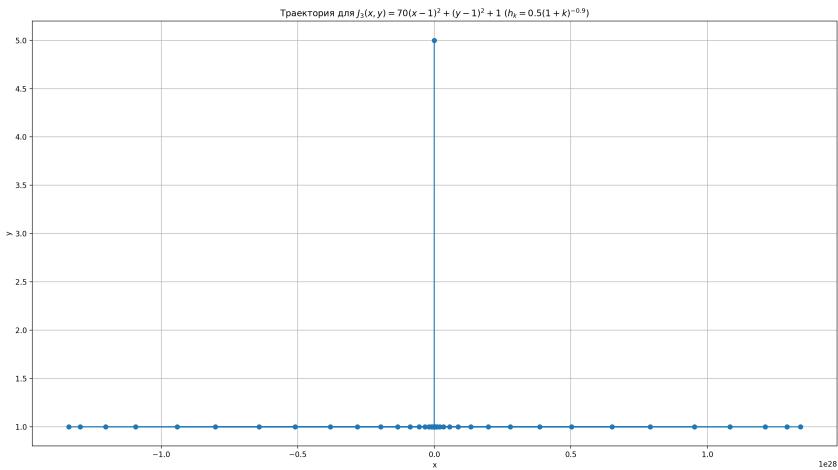


Рис. 12: Траектория градиентного спуска для J_3 при убывающем шаге $h_k = 0.5(1 + k)^{-0.9}$ (файл f3_traj_power.png).

Убывание $J_3(x_k)$

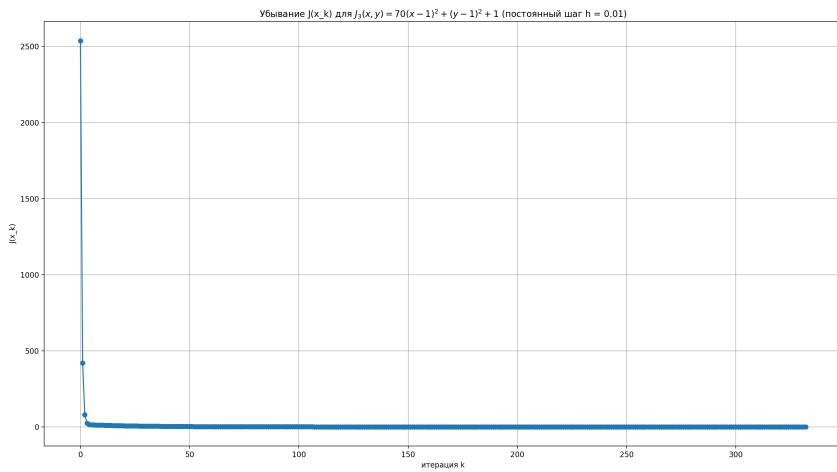


Рис. 13: График убывания $J_3(x_k)$ при постоянном шаге (файл f3_J_constant.png).

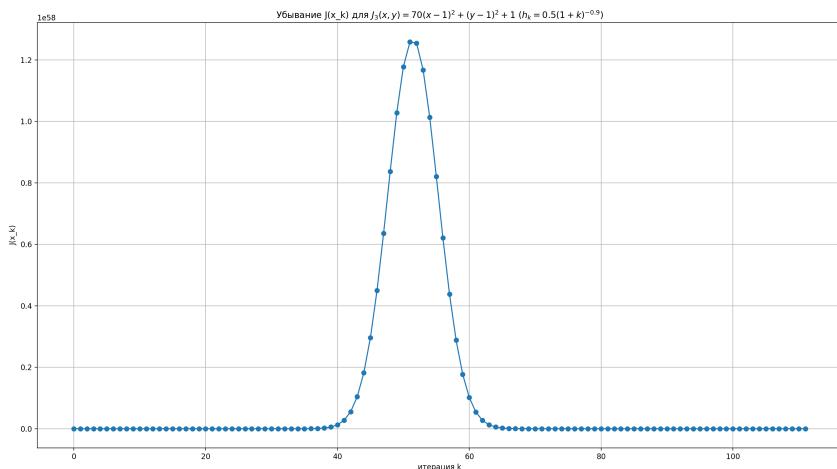


Рис. 14: График убывания $J_3(x_k)$ при убывающем шаге (файл f3_J_power.png).

3.4 Функция $J_4(x, y)$

Траектории на плоскости

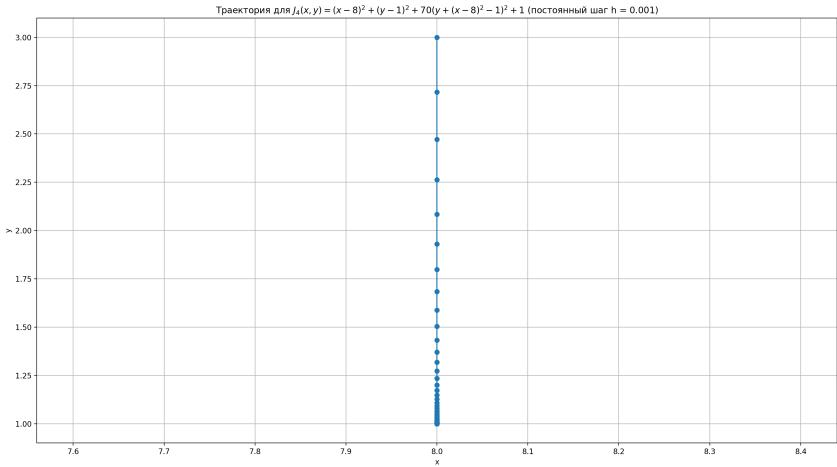


Рис. 15: Траектория градиентного спуска для J_4 при постоянном шаге (файл f4_traj_constant.png).

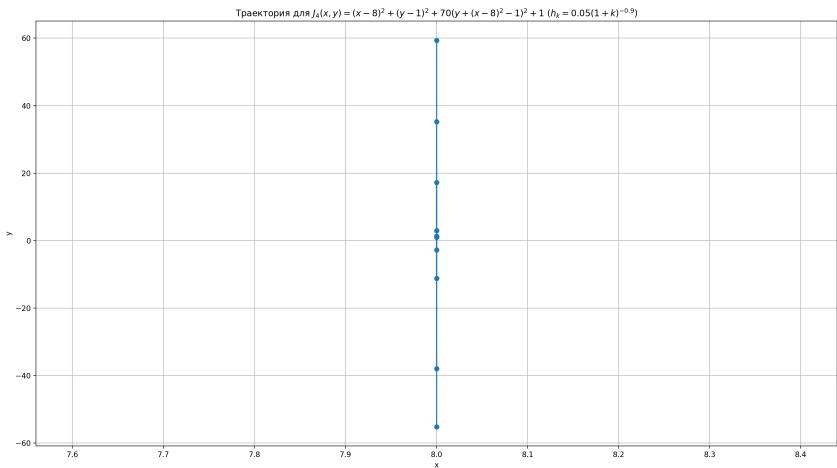


Рис. 16: Траектория градиентного спуска для J_4 при убывающем шаге $h_k = 0.05(1 + k)^{-0.9}$ (файл f4_traj_power.png).

Убывание $J_4(x_k)$

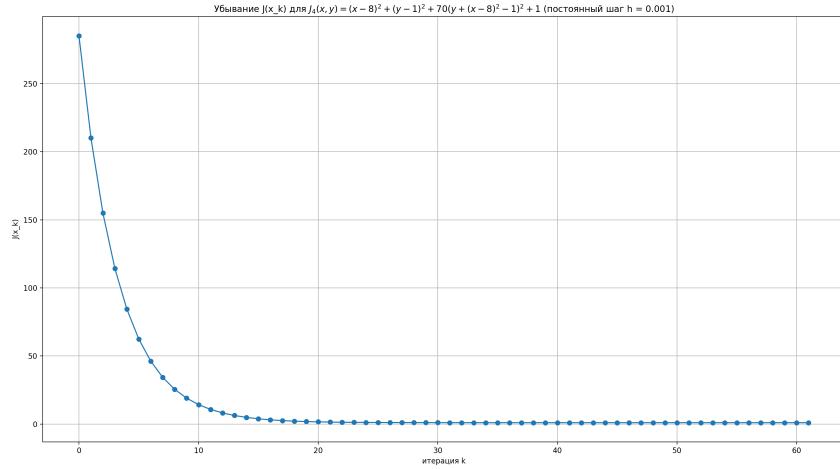


Рис. 17: График убывания $J_4(x_k)$ при постоянном шаге (файл f4_J_constant.png).

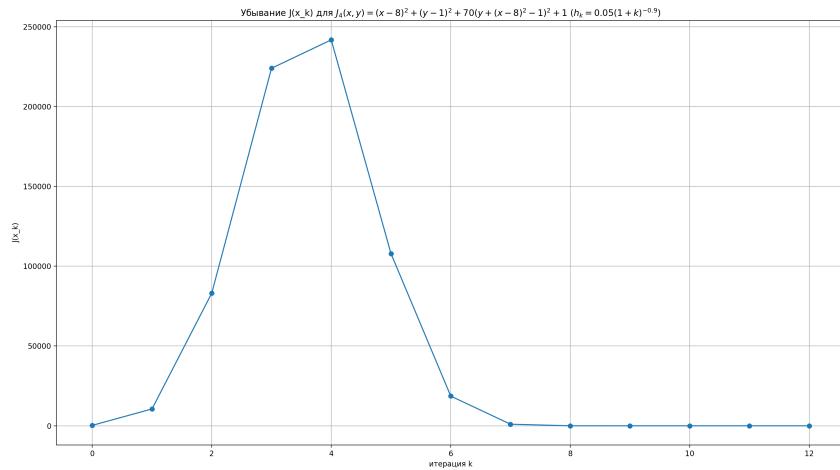


Рис. 18: График убывания $J_4(x_k)$ при убывающем шаге (файл f4_J_power.png).

4 Итоговые точки минимума

Результаты работы метода сведены в таблицу

Функция	Режим шага	Найденная точка x^*	$J(x^*)$	Критерий остановки
J_1	постоянный шаг $h = 0.2$	(0.000508, 0.000508)	0.00000052	tol_f: $ J_{(k+1)} - J_{(k)} = 9.17e-07$
J_1	$h_k = 0.5(1+k)^{-0.6}$	(0.000000, 0.000000)	0.00000000	tol_x: $\ x_{(k+1)} - x_{(k)}\ = 0.00e+00$
J_2	постоянный шаг $h = 0.2$	(0.027314, 0.000000)	0.00002984	tol_f: $ J_{(k+1)} - J_{(k)} = 9.78e-07$
J_2	$h_k = 0.5(1+k)^{-0.6}$	(0.234139, 0.000000)	0.00219284	tol_f: $ J_{(k+1)} - J_{(k)} = 1.00e-06$
J_3	постоянный шаг $h = 0.01$	(1.000000, 1.004888)	1.00002389	tol_f: $ J_{(k+1)} - J_{(k)} = 9.85e-07$
J_3	$h_k = 0.5(1+k)^{-0.9}$	(1.000000, 1.000000)	1.00000000	tol_f: $ J_{(k+1)} - J_{(k)} = 1.90e-08$
J_4	постоянный шаг $h = 0.001$	(8.000000, 1.000175)	1.00000218	tol_f: $ J_{(k+1)} - J_{(k)} = 7.82e-07$
J_4	$h_k = 0.05(1+k)^{-0.9}$	(8.000000, 1.000028)	1.00000005	tol_f: $ J_{(k+1)} - J_{(k)} = 8.79e-07$
J_1	$h_k = 0.5/(k+1)$	(0.000000, 0.000000)	0.00000000	tol_x: $\ x_{(k+1)} - x_{(k)}\ = 0.00e+00$

Таблица 1: Итоговые точки минимума и значения функционала для разных режимов шага.

5 Выводы

По результатам численных экспериментов можно сделать следующие выводы.

1. Для всех рассмотренных функций J_1-J_4 метод градиентного спуска сходится к аналитически известным точкам минимума: J_1 и J_2 — к точке $(0, 0)$, J_3 — к точке $(1, 1)$, J_4 — к точке $(8, 1)$. Численные значения x^* и $J(x^*)$, приведённые в таблице ??, совпадают с теоретическими с погрешностью порядка $10^{-3}-10^{-6}$.
2. На хорошо обусловленной квадратичной функции J_1 постоянный шаг и убывающий шаг демонстрируют быструю и стабильную сходимость. При разумном выборе h постоянный шаг обеспечивает почти геометрическое убывание $J(x_k)$. Шаг вида $h_k = c(1 + k)^{-\alpha}$ даёт более плавную, но гарантированно устойчивую сходимость.
3. Для вытянутой квадратичной функции J_2 влияние обусловленности заметно сильнее: траектория с постоянным шагом существенно искривлена, и сходится медленнее вдоль “длинной” оси эллипса. Переход к убывающему шагу частично сглаживает колебания и делает поведение метода более устойчивым, хотя скорость сходимости на начальных итерациях снижается.
4. Для функции J_3 с сильно различающимися коэффициентами по x и y (жёсткое направление по x) требуются малые значения постоянного шага, иначе возникает риск расходимости. Убывающий шаг позволяет выбрать более агрессивное начальное значение и при этом гарантировать сходимость за счёт постепенного уменьшения h_k .
5. Для нелинейной функции J_4 с узкой искривлённой долиной градиентный спуск чувствителен к выбору шага: слишком большие значения h приводят к “прыжкам” поперёк долины и замедлению сходимости. Правильно подобранный постоянный шаг и убывающий шаг обеспечивают движение вдоль долины к минимуму, что видно на траекториях и графиках убывания $J(x_k)$.
6. Шаг вида $h_k = \frac{c}{k+1}$ на примере функции J_1 демонстрирует теоретически корректное поведение: условия $\sum h_k = \infty$ и $\sum h_k^2 < \infty$ выполняются, метод сходится к минимуму, однако убывание значения функционала происходит медленнее, чем при оптимально подобранном постоянном шаге.