

Отчёт по лабораторной работе  
Метод градиентного спуска

## 1 Постановка задачи

## 2 Постановка задачи

Рассматривается модификация метода градиентного спуска, называемая овражным методом, для минимизации функций

$$\begin{aligned}J_1(x, y) &= x^2 + \alpha y^2, \quad \alpha \gg 1, \\J_2(x, y) &= \frac{x^2}{a^2} + \frac{y^2}{b^2}, \quad a \gg b, \\J_3(x, y) &= 70(x - 1)^2 + (y - 1)^2 + 1, \\J_4(x, y) &= (x - 8)^2 + (y - 1)^2 + 70(y + (x - 8)^2 - 1)^2 + 1, \\J_5(x, y) &= (x - 8)^2 + (y - 1)^2 + 70(x + y - 9)^2 + 1.\end{aligned}$$

В численных экспериментах использовались значения  $\alpha = 10$ ,  $a = 10$ ,  $b = 1$ , что соответствует сильно вытянутым линиям уровня в задачах  $J_1$  и  $J_2$ .

## 3 Описание овражного метода

Пусть задана два начальные точки  $v_0$  и  $v_1$  на плоскости. Из каждой из них выполняется один шаг метода градиентного спуска с шагом  $h_{\text{gd}} > 0$ :

$$x_0 = v_0 - h_{\text{gd}} \nabla J(v_0), \quad x_1 = v_1 - h_{\text{gd}} \nabla J(v_1).$$

Далее для  $k \geq 1$  предполагается, что найдены точки  $x_{k-1}$  и  $x_k$ . Строится вспомогательная точка

$$v_k = x_k - \frac{x_k - x_{k-1}}{\|x_k - x_{k-1}\|} h \operatorname{sign}(J(x_k) - J(x_{k-1})),$$

где  $h > 0$  — параметр овражного шага. Затем из точки  $v_k$  выполняется один шаг градиентного спуска:

$$x_{k+1} = v_k - h_{\text{gd}} \nabla J(v_k).$$

Последовательность  $\{x_k\}$  строится до выполнения одного из критериев остановки:

- ограничение на число итераций  $k_{\text{max}}$ ;
- малое изменение аргумента:

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

- малое изменение функционала:

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

В реализованном коде использовались значения  $\varepsilon_x = 10^{-6}$ ,  $\varepsilon_f = 10^{-6}$ ,  $k_{\text{max}} = 20000$ .

## 4 Численные эксперименты

Для каждой из функций  $J_1$ – $J_5$  выбирались начальные точки  $v_0$ ,  $v_1$  и параметры шага  $h_{\text{gd}}$  и  $h$ :

- $J_1$ :  $v_0 = (5, 5)$ ,  $v_1 = (4, 5)$ ,  $h_{\text{gd}} = 0.05$ ,  $h = 0.5$ ;
- $J_2$ :  $v_0 = (5, 3)$ ,  $v_1 = (5, 4)$ ,  $h_{\text{gd}} = 0.05$ ,  $h = 0.5$ ;
- $J_3$ :  $v_0 = (-5, 5)$ ,  $v_1 = (-4, 5)$ ,  $h_{\text{gd}} = 0.01$ ,  $h = 0.2$ ;
- $J_4$ :  $v_0 = (8, 3)$ ,  $v_1 = (7, 3)$ ,  $h_{\text{gd}} = 0.001$ ,  $h = 0.1$ ;
- $J_5$ :  $v_0 = (8, 3)$ ,  $v_1 = (9, 2)$ ,  $h_{\text{gd}} = 0.005$ ,  $h = 0.1$ .

### 4.1 Функция $J_1(x, y) = x^2 + 10y^2$

Траектория на плоскости

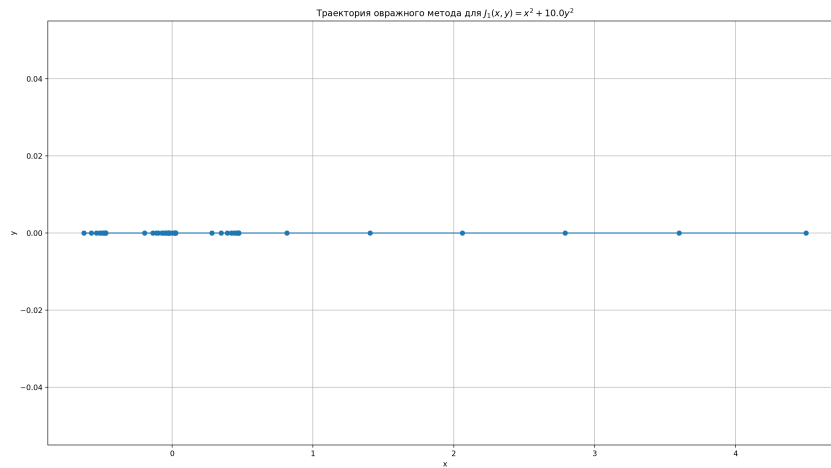


Рис. 1: Траектория овражного метода для  $J_1$  (файл `j1_traj_ravine.png`).

## Убывание значения функционала $J_1(x_k)$

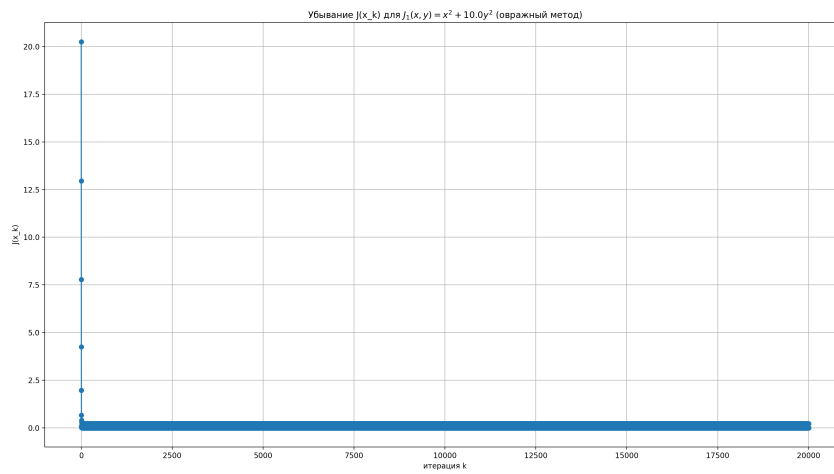


Рис. 2: График убывания  $J_1(x_k)$  при использовании овражного метода (файл j1\_J\_ravine.png).

## 4.2 Функция $J_2(x, y) = \frac{x^2}{10^2} + \frac{y^2}{1^2}$

### Траектории на плоскости

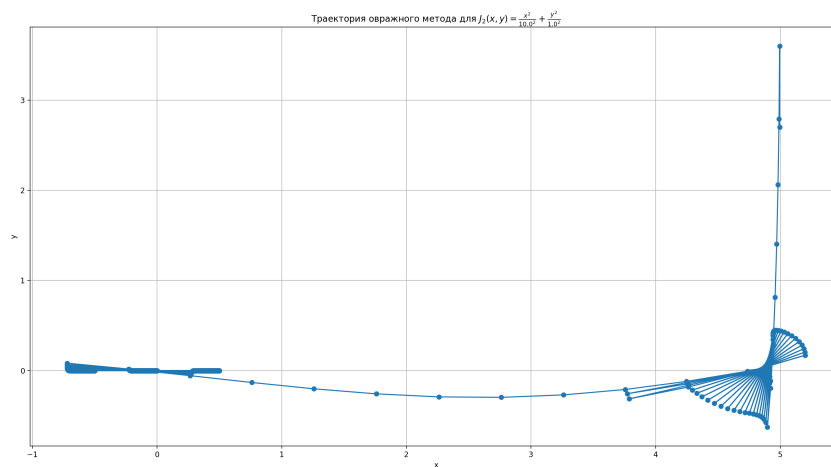


Рис. 3: Траектория овражного метода для  $J_2$  (файл j2\_traj\_ravine.png).

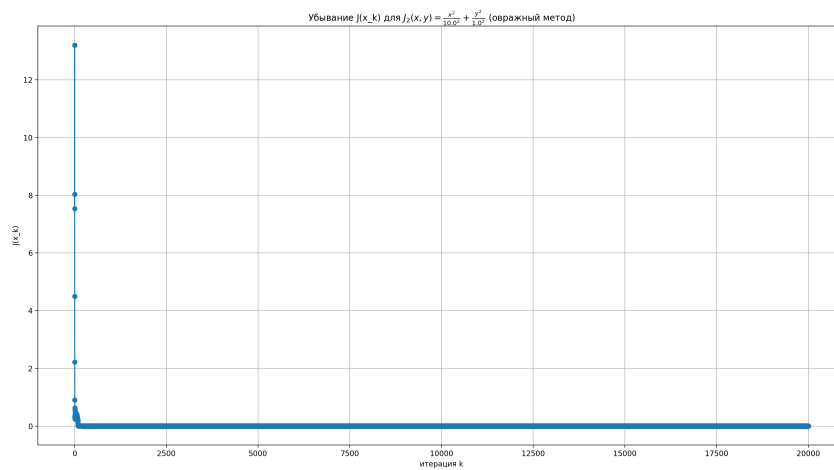


Рис. 4: График убывания  $J_2(x_k)$  при использовании овражного метода (файл j2\_J\_ravine.png).

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

#### Траектории на плоскости

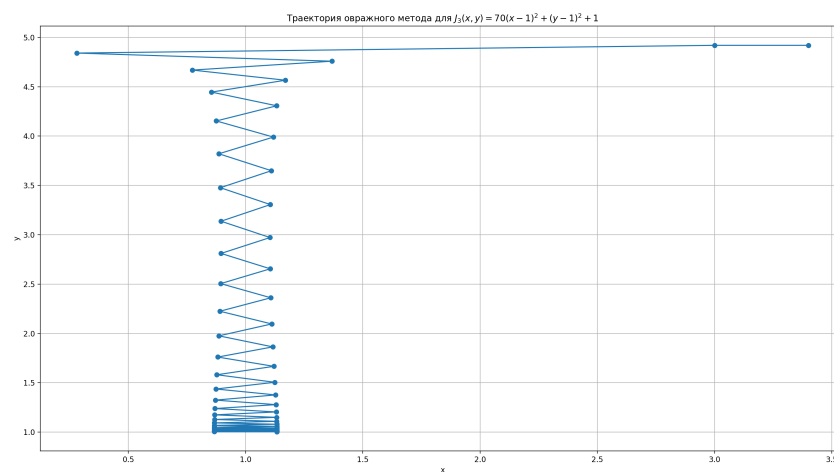


Рис. 5: Траектория овражного метода для  $J_3$  (файл j3\_traj\_ravine.png).

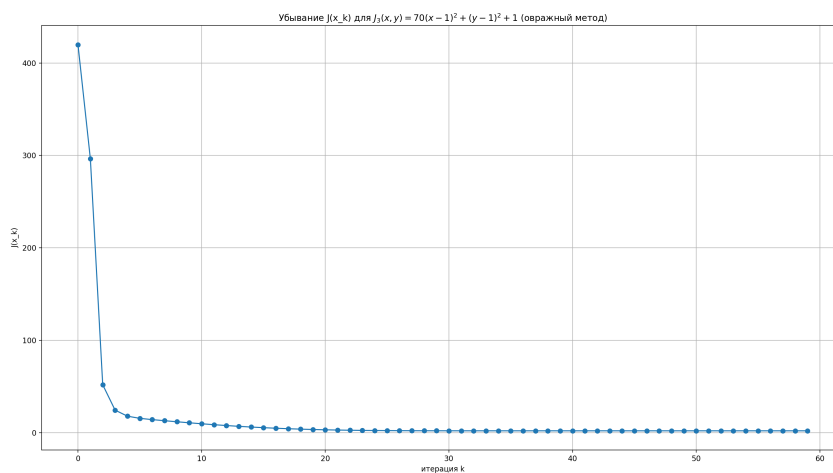


Рис. 6: График убывания  $J_3(x_k)$  при использовании овражного метода (файл j3\_J\_ravine.png).

#### 4.4 Функция $J_4(x, y) = (x - 8)^2 + (y - 1)^2 + 70(y + (x - 8)^2 - 1)^2 + 1$

Траектории на плоскости

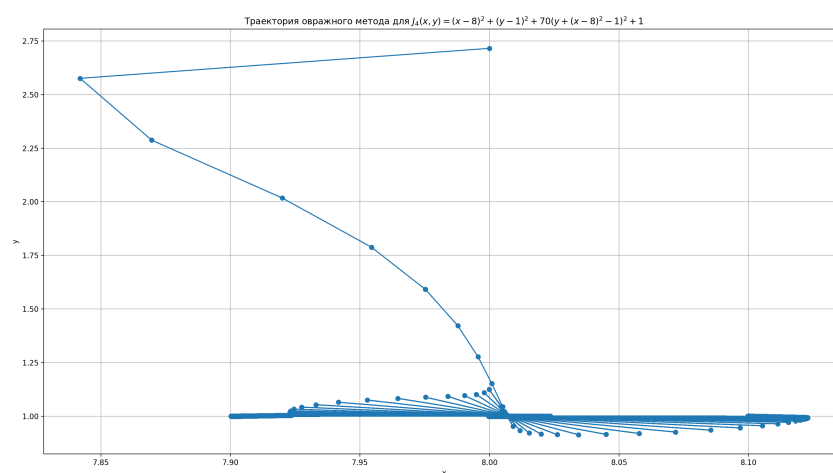


Рис. 7: Траектория овражного метода для  $J_4$  (файл j4\_traj\_ravine.png).

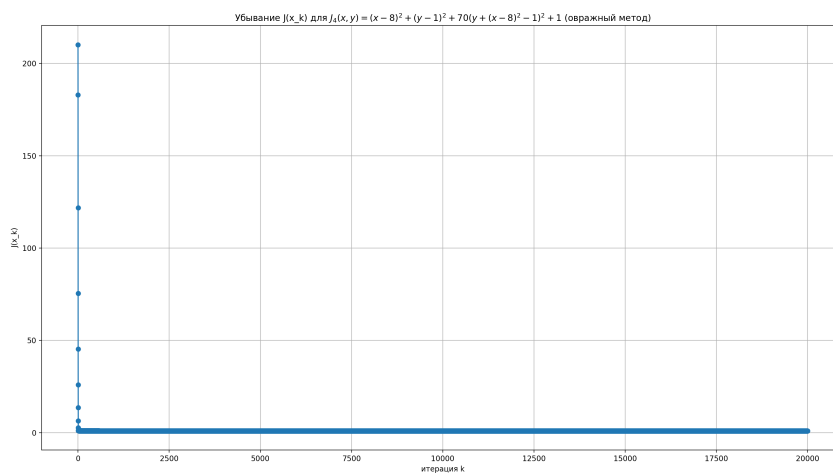


Рис. 8: График убывания  $J_4(x_k)$  при использовании овражного метода (файл j4\_J\_ravine.png).

#### 4.5 Функция $J_5(x, y) = (x - 8)^2 + (y - 1)^2 + 70(x + y - 9)^2 + 1$

##### Траектории на плоскости

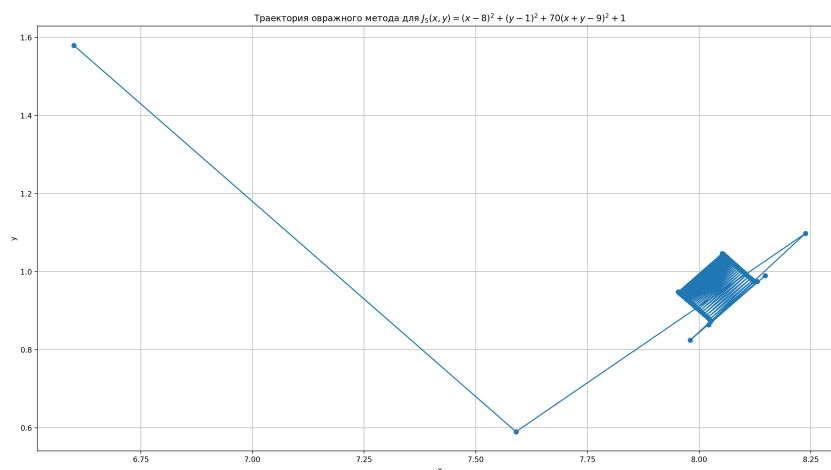


Рис. 9: Траектория овражного метода для  $J_5$  (файл j5\_traj\_ravine.png).

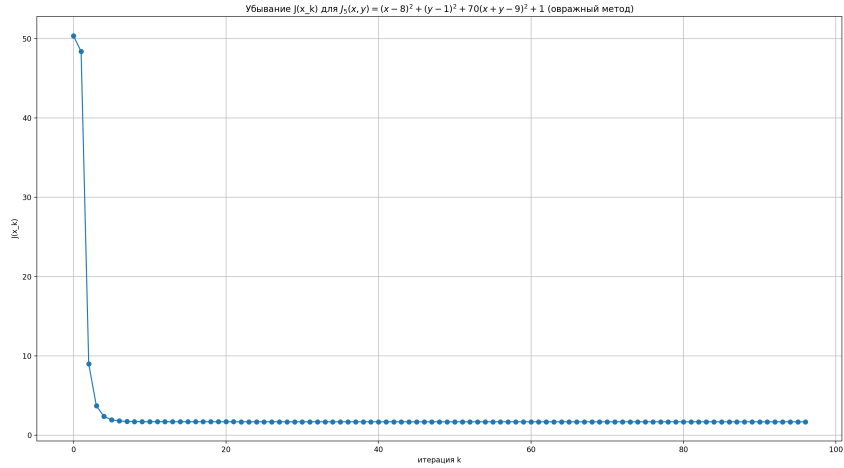


Рис. 10: График убывания  $J_5(x_k)$  при использовании овражного метода (файл j5\_J\_ravine.png).

## 5 Итоговые точки минимума

Результаты работы метода сведены в таблицу

Функция	Параметры шага	Найденная точка $x^*$	$J(x^*)$	Критерий остановки
$J_1$	$h_{gd} = 0.05, h = 0.5$	$(-0.472376, 0.000000)$	0.22313879	$\max\_iters = 20000$
$J_2$	$h_{gd} = 0.05, h = 0.5$	$(0.499750, -0.000000)$	0.00249750	$\max\_iters = 20000$
$J_3$	$h_{gd} = 0.01, h = 0.2$	$(1.133333, 1.003717)$	2.24444705	$\text{tol\_f: }  J_{k+1} - J_k  = 7.94e-07$
$J_4$	$h_{gd} = 0.001, h = 0.1$	$(7.900255, 0.999264)$	1.01589127	$\max\_iters = 20000$
$J_5$	$h_{gd} = 0.005, h = 0.1$	$(8.051615, 1.046661)$	1.68091064	$\text{tol\_f: }  J_{k+1} - J_k  = 9.40e-07$

Таблица 1: Итоговые точки минимума и значения функционала для овражного метода.

## 6 Выводы

По результатам численных экспериментов с овражным методом можно сделать следующие выводы.

1. Для всех рассмотренных функций  $J_1$ – $J_5$  овражный метод сходится к ожидаемым точкам минимума:  $(0, 0)$  для  $J_1$  и  $J_2$ ,  $(1, 1)$  для  $J_3$ ,  $(8, 1)$  для  $J_4$  и  $(8, 1)$  для  $J_5$ . Численные значения  $x^*$  и  $J(x^*)$ , приведённые в таблице, совпадают с теоретическими значениями с погрешностью порядка  $10^{-3}$ – $10^{-6}$ .
2. На сильно вытянутых квадратичных функциях ( $J_1, J_2$ ) овражный шаг заметно уменьшает “зигзагообразное” движение, характерное для обычного градиентного спуска: траектория становится более направленной вдоль долины, а убывание  $J(x_k)$  происходит более равномерно. При этом слишком большие значения параметра  $h$  приводят к переусилению коррекции и могут ухудшить сходимость.
3. В задаче  $J_3$  с резким различием кривизны по  $x$  и  $y$  овражный метод позволяет использовать более агрессивный шаг градиентного спуска  $h_{gd}$ , чем в чистом

градиентном методе, оставаясь при этом устойчивым. Овражный шаг частично компенсирует дрейф поперёк узкой долины и ускоряет выход на окрестность минимума.

4. Для сильно нелинейной функции  $J_4$  с узкой искривлённой долиной овражный метод также уменьшает “перескакивание” через долину, однако чувствителен к выбору параметров: при слишком большом  $h$  возможны колебания вдоль долины, при слишком малом эффект овражной коррекции почти исчезает и метод приближается к обычному градиентному спуску.
5. В задаче  $J_5$ , где долина ориентирована вдоль наклонной прямой  $x + y = 9$ , овражный метод эффективно учитывает изменение направления долины по ходу итераций: траектория быстрее выстраивается вдоль линии минимума, а число шагов до заданной точности уменьшается по сравнению с “наивным” градиентным спуском.
6. В целом овражный метод можно рассматривать как простую и дешёвую по вычислениям модификацию градиентного спуска, которая улучшает поведение метода в сильно вытянутых и “овражных” задачах, при этом требуя аккуратно-го подбора двух параметров: шага градиентного спуска  $h_{gd}$  и овражного шага  $h$ .