4. ЛЕКЦИЯ. Безусловная оптимизация многих переменных. Методы первого порядка.

Методы первого порядка используют информацию о значениях целевой функции $f(\overline{x})$ и её первых производных. Предполагается, что функция $f(\overline{x})$ и её первые производные существуют и непрерывны. Направление смещения от точки $\overline{x^{(k)}}$ к точке $\overline{x^{(k+1)}}$ описывается представленной на первой лекции интеграционной процедурой (2.1)

$$\overline{x^{(k+1)}} = \overline{x^{(k)}} + h_k\overline{p^{(k)}}, \tag{2.1}$$

где k – номер итерации $k = 0, 1, \ldots$;

$\overline{x^{(k)}}$ – текущее приближение;

h_k – величина шага;

$\overline{p^k}$ – вектор, определяющий направление убывания функции $f(\overline{x})$.

и совпадает с направлением вектора антиградиента целевой функции $\overline{p^{(k)}} = -\nabla f(\overline{x^{(k)}})$. Все итерационные процессы, в которых направление движения на каждом шаге совпадает с антиградиентом функции, называются градиентными методами. Существует несколько модификаций градиентных методов, различающихся правилом выбора длины шага в направлении антиградиента функции.

Мы рассмотрим наиболее распространённые на практике следующие методы: *метод градиентного спуска с постоянным шагом, метод наискорейшего градиентного спуска, метод покоординатного спуска, метод сопряжённых направлений.*

Метод градиентного спуска с постоянным шагом

Сущность метода градиентного спуска с постоянным шагом заключается в следующем. Выбирается начальная точка $\overline{x^{(0)}}$ из области определения функции $f(\overline{x})$. Координаты новой точки вычисляются по формуле

$$\overline{x^{(k+1)}} = \overline{x^{(k)}} - h_k \nabla f(\overline{x^{(k)}}), \tag{2.2}$$

где k − номер итерации $k = 0, 1, \ldots$,

h_k − величина шага,

$\nabla f\,(\overline{x^{(k)}})$ – *градиент функции* $f(\,\overline{x}\,)$ в точке $\overline{x^{(k)}}$,

$$\nabla f\,(\overline{x^{(k)}}) \;=\left(\frac{\partial f\,(\overline{x^{(k)}}}{\partial x_1},\frac{\partial f\,(\overline{x^{(k)}}}{\partial x_2},\;\dots,\frac{\partial f\,(\overline{x^{(k)}})}{\partial x_n}\right).$$

Начальная величина шага h_0 задаётся пользователем. В каждом новой точке поиска $\overline{x^{(k+1)}}$ проверяется условие убывания функции $f\,(\overline{x^{(k+1)}}) < f\,(\overline{x^{(k)}})$. Если условие нарушается, то постепенно уменьшается величина шага h_k, т.е. точка $\overline{x^{(k+1)}}$ приближается к точке $\overline{x^{(k)}}$ до тех пор, пока условие не выполнится. В полученной точке $\overline{x^{(k+1)}}$ определяется новое направление градиента и осуществляется новый спуск. Процесс продолжается пока не будет выполнено условие окончания поиска. В качестве условия окончания поиска используется близость к нулю нормы градиента $\left\|\nabla f\,(\overline{x^{(k+1)}})\right\| \le$ ε.

Геометрическая иллюстрация поиска минимума целевой функции методом градиентного спуска с постоянным шагом для случая *n* = 2 представлена на рис. 1.

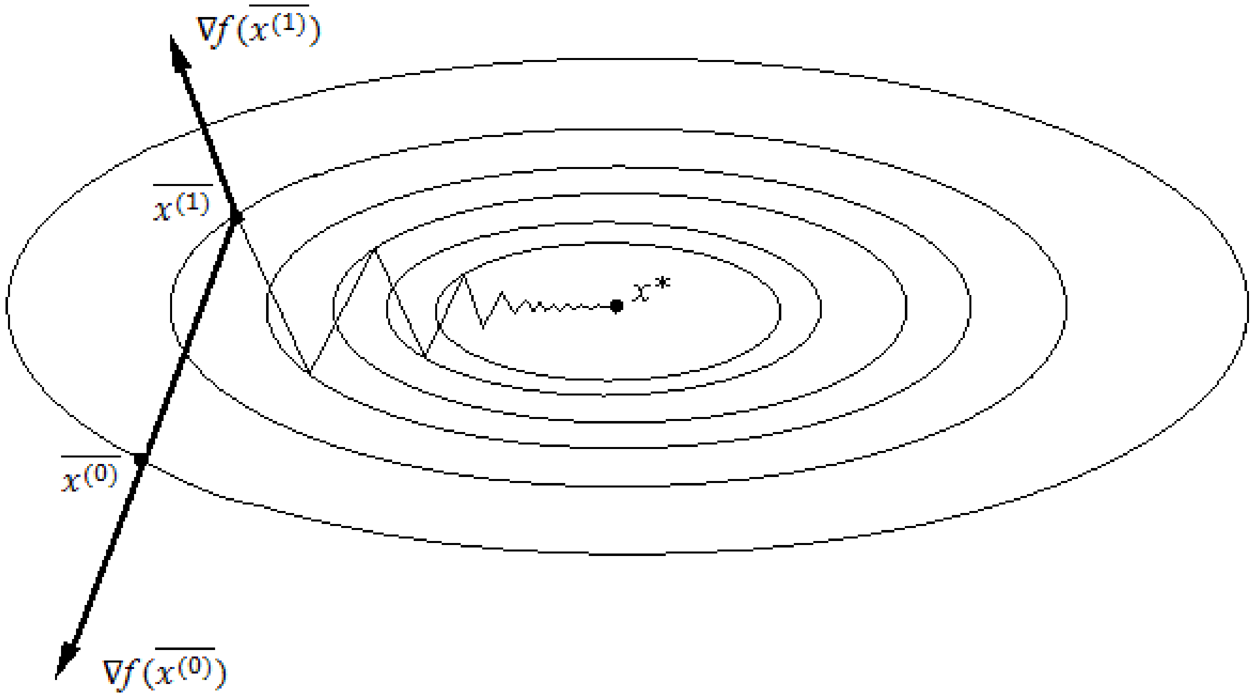


Рис. 1. Графическая иллюстрация поиска точки минимума методом градиентного спуска с постоянным шагом

Алгоритм метода минимизации целевой функции $f\ (\overline{x})$ методом градиентного спуска с постоянным шагом заключается в следующем:

1. Задать размерность задачи оптимизации n, координаты начальной точки $\overline{x^{(0)}} = ({x_1}^{(0)}, {x_2}^{(0)}, \dots, {x_n}^{(0)})$, начальную величину шага h_0, точность поиска ε.
2. Положить счётчик числа итерация k = 0
3. Вычислить значение функции $f(\overline{x^{(k)}})$ в точке $\overline{x^{(k)}}$.
4. Определить координаты вектора градиента функции $f(\ \overline{x}\ )$ в точке

$$\overline{x^{(k)}} = ({x_1}^{(k)}, {x_2}^{(k)}, \dots, {x_n}^{(k)}).$$

5. Проверить условие окончания поиска

$$\left\|\nabla f\ (\overline{x^{(k)}})\right\| = \sqrt{\sum_{i=1}^{n}\left(\frac{\partial f\ (\overline{x^{(k)}})}{\partial x_i}\right)^2} \leq \varepsilon.$$

Если условие выполнено, то перейти к пункту 8, иначе – к пункту 6.

6. Определить координаты точки $\overline{x^{(k+1)}} = \overline{x^{(k)}} - h_k \nabla \mathrm{f}\ (\overline{x^{(k)}})$ и значение целевой функции f $(\overline{x^{(k+1)}})$.
7. Проверить условие убывания функции f $(\overline{x^{(k+1)}})$ < f $(\overline{x^{(k)}})$.

Если условие выполнено, то положить k = k +1, f $(\overline{x^{(k)}})$ = f $(\overline{x^{(k+1)}})$ и перейти к пункту 4, иначе – положить $h_k = \frac{h_k}{2}$ и перейти к пункту 6.

8. Расчёт окончен. Полагаем $\overline{x^*} = \overline{x^{(k)}}$.

Пример. Найти минимум целевой функции

$$f\,(\,\bar{x}\,) = x_1^2 - \mathrm{x}_1\mathrm{x}_2 + 3x_2^2 - \mathrm{x}_1$$

методом градиентного спуска с постоянным шагом с точностью ε = 0,1.

Решение. Зададим начальную точку $\overline{x^{(0)}} = ({x_1}^{(0)}, {x_2}^{(0)})^T = (0;\ 0)^T$ и начальную величину шага h = 0,4.

Найдем градиент функции в произвольной точке $\overline{x^{(k)}} = ({x_1}^{(k)}, {x_2}^{(k)})^T$

$$\nabla f\left(\overline{x^{(k)}}\right) = \left(\frac{\partial f\left(\overline{x^{(k)}}\right)}{\partial x_1}; \frac{\partial f\left(\overline{x^{(k)}}\right)}{\partial x_2}\right)^T = \left(2x_1^{(k)} - x_2^{(k)} - 1; -x_1^{(k)} + 6x_2^{(k)}\right)^T.$$

Итерация k = 0. Вычислим значение целевой функции $f\left(\overline{x^{(0)}}\right)$ и градиент $\nabla f\left(\overline{x^{(0)}}\right)$ в начальной точке $x^{(0)}$: $f\left(\overline{x^{(0)}}\right) = 0$; $\nabla f\left(\overline{x^{(0)}}\right) = (-1;\ 0)^T$.

		Координаты вершины		Значение функции
		1	2	
Номер вершины	0	$x_1^{(0)} = 0$	$x_2^{(0)} = 0$	$f\left(\overline{x^{(0)}}\right) = 0$

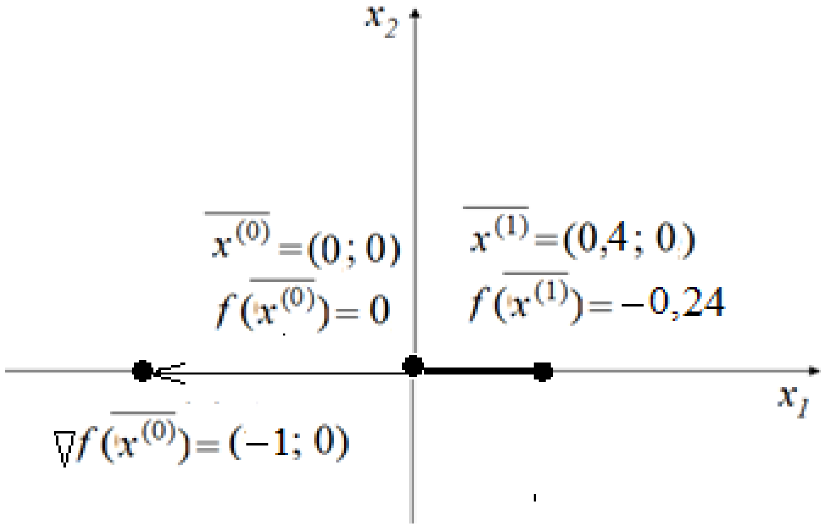
Определим координаты точки $\overline{x^{(1)}}$

$$\overline{x^{(1)}} = \overline{x^{(0)}} - h\,\nabla f\left(\overline{x^{(0)}}\right) = (0;\ 0)^T - 0{,}4 * (-1;\ 0)^T = (0{,}4;\ 0)^T$$

и значение целевой функции в этой точке

$$f\left(\overline{x^{(1)}}\right) = -0{,}240.$$

Тогда



		Координаты вершины		Значение функции
		1	2	
Номер вершины	0	$x_1^{(0)} = 0$	$x_2^{(0)} = 0$	$f\left(\overline{x^{(0)}}\right) = 0$
	1	$x_1^{(1)} = 0{,}4$	$x_2^{(1)} = 0$	$f\left(\overline{x^{(1)}}\right) = -0{,}24$

Сравним $f\,(\overline{x^{(0)}})$ и $f\,(\overline{x^{(1)}})$. Поскольку $f\,(\overline{x^{(1)}}) < f\,(\overline{x^{(0)}})$, то условие убывания функции выполнено.

Проверим условие окончания поиска. Для этого вычислим вектор градиента $\nabla f\,(\overline{x^{(1)}})$ в точке $\overline{x^{(1)}}$:

$$\nabla f\,(\overline{x^{(1)}}) = (-0{,}2;\ -0{,}4)^T.$$

Найдем норму вектора градиента $\overline{x^{(1)}}= (0{,}4;\ 0)^T$

$$\nabla f\,(\overline{x^{(1)}})\ == (2x_1^{(1)} - x_2^{(1)} - 1;\ -x_1^{(1)} + 6x_2^{(1)})^T = (-0{,}2; -0{,}4)$$

$$\left\|\nabla f\ (\overline{x^{(1)}})\right\| = \sqrt{(-0{,}2)^2 + (-0{,}4)^2} = 0{,}447 > \varepsilon.$$

Итерации продолжаются.

***Итерация k* = 1.** Определим координаты точки $\overline{x^{(2)}}$

$$\overline{x^{(2)}} = \overline{x^{(1)}} - h\ \nabla f\,(\overline{x^{(1)}}) = (0{,}4;\ 0)^T - 0{,}4 * (-0{,}2;\ -0{,}4)^T = (0{,}48;\ 0{,}16)^T$$

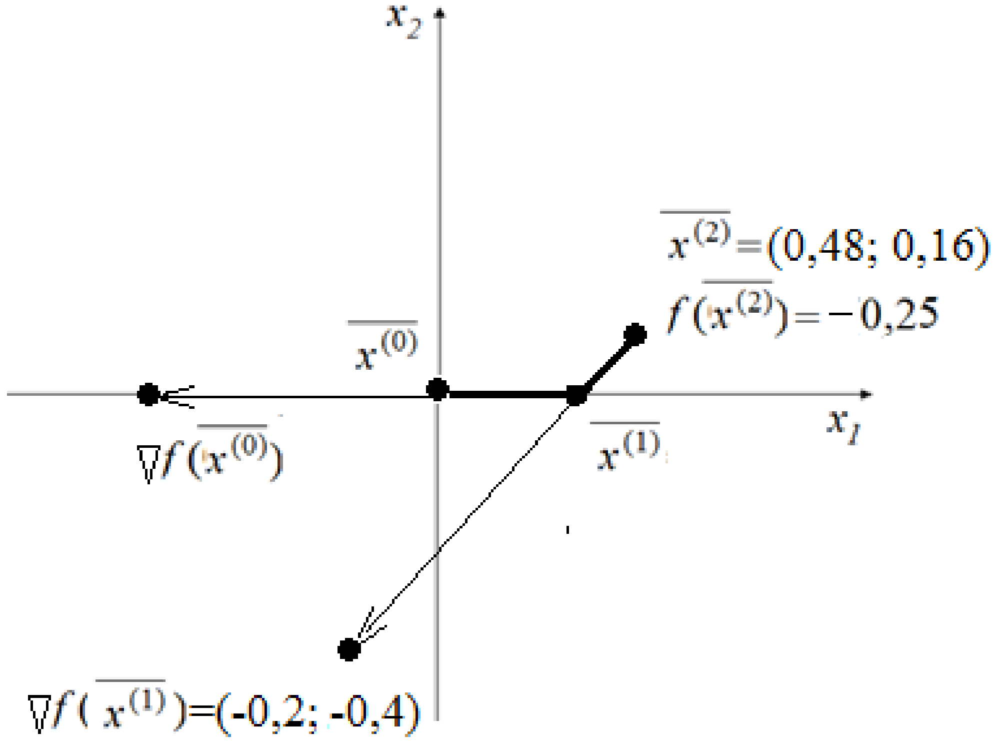
и значение целевой функции в этой точке

$$f\,(\overline{x^{(2)}}) = -0{,}250.$$

Запишем значения в таблицу

		Координаты вершины		Значение функции
		1	2	
Номер вершины	0	$x_1^{(0)} = 0$	$x_2^{(0)} = 0$	$f\left(\overline{x^{(0)}}\right) = 0$
	1	$x_1^{(1)} = 0{,}4$	$x_2^{(1)} = 0$	$f\left(\overline{x^{(1)}}\right) = -0{,}24$
	2	$x_1^{(2)} = 0{,}48$	$x_2^{(2)} = 0{,}16$	$f\left(\overline{x^{(2)}}\right) = -0{,}25$

Построим график



Сравним $f(\overline{x^{(1)}})$ и $f(\overline{x^{(2)}})$. Поскольку $f(\overline{x^{(2)}}) < f(\overline{x^{(1)}})$, то условие убывания функции выполнено.

Проверим условие окончания поиска. Для этого вычислим вектор градиента $\nabla f(\overline{x^{(2)}})$ в точке $\overline{x^{(2)}}$:

$$\nabla f(\overline{x^{(2)}}) = (-0{,}200;\ 0{,}481)^T.$$

Найдем норму вектора градиента

$$\left\|\nabla f\ (\overline{x^{(2)}})\right\| = \sqrt{(-0{,}200)^2 + (0{,}481)^2} = 0{,}521 > \varepsilon.$$

Итерации продолжаются.

***Итерация k* = 2.** Определим координаты точки $\overline{x^{(3)}}$

$$\overline{x^{(3)}} = \overline{x^{(2)}} - h\ \nabla f(\overline{x^{(2)}}) = (0{,}48;\ 0{,}16)^T - 0{,}4 * (-0{,}200;\ 0{,}481)^T = (0{,}560;\ -0{,}032)^T$$

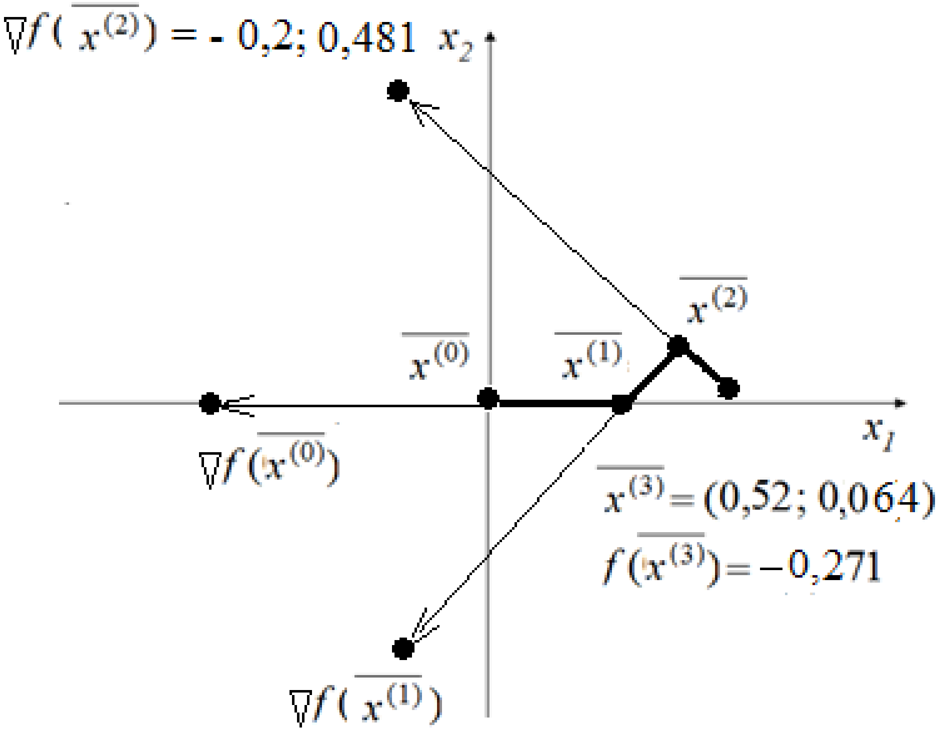
и значение целевой функции в этой точке

$$f(\overline{x^{(3)}}) = -0{,}225.$$

Сравним $f\,(\overline{x^{(2)}})$ и $f\,(\overline{x^{(3)}})$. Поскольку $f\,(\overline{x^{(3)}}) > f\,(\overline{x^{(2)}})$, то условие убывания функции не выполнено. Уменьшим величину шага $h = \frac{h}{2} = 0{,}2$ и повторим вычисления координат точки $\overline{x^{(3)}}$

$$\overline{x^{(3)}} = \overline{x^{(2)}} - h\,\nabla f\,(\overline{x^{(2)}}) = (0{,}48;\ 0{,}16)^T - 0{,}2 * (-0{,}2;\ 0{,}481)^T = (0{,}520;\ 0{,}064)^T$$

Значение целевой функции $f\,(\overline{x^{(3)}}) = -0{,}271$.



Сравним $f\,(\overline{x^{(2)}})$ и $f\,(\overline{x^{(3)}})$. Поскольку $f\,(\overline{x^{(3)}}) < f\,(\overline{x^{(2)}})$, то условие убывания функции выполнено.

		Координаты вершины		Значение функции
		1	2	
Номер вершины	0	$x_1^{(0)} = 0$	$x_2^{(0)} = 0$	$f\left(\overline{x^{(0)}}\right) = 0$
	1	$x_1^{(1)} = 0{,}4$	$x_2^{(1)} = 0$	$f\left(\overline{x^{(1)}}\right) = -0{,}24$
	2	$x_1^{(2)} = 0{,}48$	$x_2^{(2)} = 0{,}16$	$f\left(\overline{x^{(2)}}\right) = -0{,}25$
	3	$x_1^{(3)} = 0{,}52$	$x_2^{(3)} = 0{,}064$	$f(\overline{x^{(3)}}) = -0{,}271.$

Проверим условие окончания поиска. Для этого вычислим вектор градиента $\nabla f\,(\overline{x^{(3)}})$ в точке $\overline{x^{(3)}}$:

$$\nabla f\,(\overline{x^{(3)}}) = (-0{,}024;\ -0{,}136)^T.$$

Найдем норм вектора градиента

$$\left\|\nabla f\ (\overline{x^{(3)}})\right\| = \sqrt{(-0{,}024)^2 + (-0{,}136)^2} = 0{,}138 > \varepsilon.$$

Итерации продолжаются.

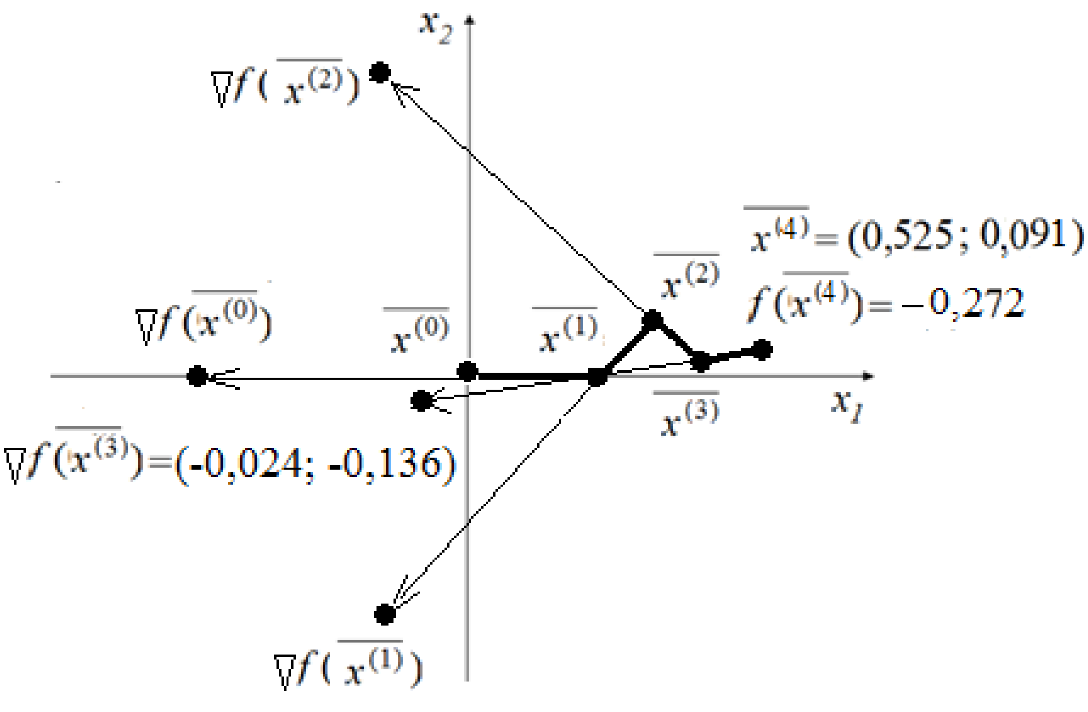
***Итерация k* = 3.** Определим координаты точки $\overline{x^{(4)}}$

$$\overline{x^{(4)}} = \overline{x^{(3)}} - h\,\nabla f\,(\overline{x^{(3)}}) = (0{,}520;\ 0{,}064)^T - 0{,}2 * (-0{,}024;\ -0{,}136)^T =$$
$$= (0{,}525;\ 0{,}091)^T$$

и значение целевой функции в этой точке

$$f\,(\overline{x^{(4)}}) = -0{,}272.$$

Сравним $f\ (\overline{x^{(3)}})$ и $f\ (\overline{x^{(4)}})$. Поскольку $f\ (\overline{x^{(3)}}) < f\ (\overline{x^{(4)}})$, то условие убывания функции выполнено



		Координаты вершины		Значение функции
		1	2	
Номер вершины	0	$x_1^{(0)} = 0$	$x_2^{(0)} = 0$	$f\left(\overline{x^{(0)}}\right) = 0$
	1	$x_1^{(1)} = 0{,}4$	$x_2^{(1)} = 0$	$f\left(\overline{x^{(1)}}\right) = -\,0{,}24$
	2	$x_1^{(2)} = 0{,}48$	$x_2^{(2)} = 0{,}16$	$f\left(\overline{x^{(2)}}\right) = -0{,}25$
	3	$x_1^{(3)} = 0{,}52$	$x_2^{(3)} = 0{,}064$	$f(\overline{x^{(3)}}) = -0{,}271.$
	4	$x_1^{(4)} = 0{,}525$	$x_2^{(4)} = 0{,}091$	$f\,(\overline{x^{(4)}}) = -0{,}272$

Проверим условие окончания поиска. Для этого вычислим вектор градиента ∇f $(\overline{x^{(4)}})$ в точке $\overline{x^{(4)}}$:

$$\nabla f\,(\overline{x^{(4)}}) = (-0{,}042;\ 0{,}022)^T$$

Найдем норму вектора градиента

$$\left\|\nabla f\ (\overline{x^{(4)}})\right\| = \sqrt{(-0{,}042)^2 + (0{,}022)^2} = 0{,}047 < \varepsilon.$$

Расчет окончен. Найдена точка

$$x^{(*)} = x^{(4)} = (0{,}525; 0{,}091)^T,\ f\ (\overline{x^{*}}) = -0{,}272.$$