Математика для Data Science. Математический анализ.
Условия задач

Содержание

Замечание. Вот этим цветом отмечены ссылки на страницы внутри этого файла.

4.1 Одномерный градиентный спуск

Задача 1

Дана функция $f(x) = x^2$. Мы начинаем в точке $r_1 = 1024$.

1. В какой точке мы окажемся через 1 шаг градиентного спуска с learning rate $\frac{1}{4}$?
2. В какой точке мы окажемся через 2 шага градиентного спуска с learning rate $\frac{1}{4}$?
3. В какой точке мы окажемся через 13 шагов градиентного спуска с learning rate $\frac{1}{4}$?

Задача 2

В алгоритме градиентного спуска мы надеемся, что для всех i будет выполнено $f(r_i) > f(r_{i+1})$. То есть, что на каждом шаге значение функции уменьшается.

Обязательно ли это будет так? Другими словами, может ли так случиться, что мы делаем шаг градиентного спуска и попадаем в точку, в которой значение функции больше, чем в предыдущей?

Если может – объясните, почему, и приведите пример. Если не может – докажите, что не может.

Задача 3

Пусть про функцию $f : \mathbb{R} \to \mathbb{R}$ и точку r_1 известно, что

- градиентный спуск с learning rate $= 1$ и стартовой точкой r_1 достигает точки глобального минимума за 1 шаг, и
- градиентный спуск с learning rate $= \frac{1}{2}$ и стартовой точкой r_1 не достигает 1-окрестности точки глобального минимума ни на каком шаге (пояснение: 1-окрестность это ε-окрестность с $\varepsilon = 1$:))

Как может выглядеть график такой функции и где должна располагаться на нём точка r_1?

Задача 4

Существует ли дифференцируемая функция $f : \mathbb{R} \to \mathbb{R}$ и точка r_1, такие что:

- функция f достигает глобального минимума,
- $f'(r_1) \neq 0$,
- градиентный спуск с начальной точкой r_1 и любым положительным learning rate не достигает 1-окрестности глобального минимума.

Если да — приведите пример функции f и точки r_1. Если нет — докажите, что такой функции не существует.

Если ответ «да», то достаточно объяснить, как выглядит график такой функции f и где должна располагаться на нём точка r_1. Формулу для $f(x)$ находить не обязательно.

Задача 5

Существует ли функция $f : \mathbb{R} \to [0, +\infty)$, у которой есть бесконечно много локальных минимумов, но нет глобального минимума?

Если да — приведите пример функции f. Если нет — докажите.

Если ответ «да», то достаточно объяснить, как выглядит график такой функции f. Формулу для $f(x)$ находить не обязательно.

4.3 $\mathbb{R}^n$: расстояния и векторы.

Задача 1

По аналогии с определением ε-окрестности точки для случая $\mathbb{R}^1$, мы определяем ε-окрестность точки $a \in \mathbb{R}^n$.

Определение. ε-окрестностью точки $a \in \mathbb{R}^n$ называется множество всех таких точек $x \in \mathbb{R}^n$, что расстояние от x до a меньше ε.

1. Как выглядит ε-окрестность точки $a \in \mathbb{R}^2$?
2. Как выглядит ε-окрестность точки $a \in \mathbb{R}^3$?
3. Как записать условие "x лежит в ε-окрестности точки a"в виде условия на координаты точки $x = (x_1, \dots, x_n)$?

Пример. Проверим, лежит ли точка $(5, -4)$ в 6-окрестности точки $(3, 1)$ в $\mathbb{R}^2$. Для этого вычислим расстояние $d((5, -4), (3, 1))$. Оно равно $||(5, -4) - (3, 1)||$ по предыдущей задаче, что, в свою очередь, равно $||(2, -5)|| = \sqrt{2^2 + (-5)^2} = \sqrt{29}$. А $\sqrt{29}$ меньше 6, потому что оба числа положительны и $\sqrt{29}^2 = 29 < 6^2 = 36$. Следовательно, $(5, -4)$ находится внутри 6-окрестности $(3, 1)$.

Задача 2

Дана последовательность $\{z_i\}$ точек в $\mathbb{R}^n$. Обозначим k-ую координату точки z_i за z_{ik}. То есть по координатам точки нашей последовательности $\{z_i\}$ запишутся так: $z_1 = (z_{11}, z_{12}, \ldots, z_{1n})$, $z_2 = (z_{21}, z_{22}, \ldots, z_{2n})$, ... $z_i = (z_{i1}, z_{i2}, \ldots, z_{in})$, ... Заметим, что $z_i \in \mathbb{R}^n$ это точки из многомерного пространства, а вот координаты $z_{ik} \in \mathbb{R}$ это просто обычные действительные числа.

Докажите, что эта последовательность сходится к точке $a = (a_1, \ldots, a_n)$ если и только если одновременно выполнены следующие n условий:

- последовательность первых координат точек z_i сходится к a_1,
- последовательность вторых координат точек z_i сходится к a_2,
- ...
- последовательность n-ых координат точек z_i сходится к a_n.

Для понимания может быть полезно сначала решить эту задачу в случае $n = 2$.

Пример. Дана последовательность $\{z_i\} = \left\{\left(\frac{100i^5}{i^5+i^3}, \frac{2^i+16}{2^i}\right)\right\}$.

- Последовательность первых координат точек z_i это последовательность $\left\{\frac{100i^5}{i^5+i^3}\right\}$. Она сходится к числу 100.
- Последовательность вторых координат точек z_i это последовательность $\left\{\frac{2^i+16}{2^i}\right\}$. Она сходится к числу 1.

Поэтому $\lim\limits_{i\to\infty}(z_i)$ это точка $(100, 1)$.

Комментарий. В этой задаче вы можете пользоваться следующим утверждением. Последовательность неотрицательных чисел стремится к нулю тогда и только тогда, когда последовательность корней из этих чисел стремится нулю. Другими словами, если для любого i выполнено $y_i > 0$, то

$$\lim_{i\to\infty}(y_i) = 0 \Leftrightarrow \lim_{i\to\infty}(\sqrt{y_i}) = 0.$$

Задача 3

Дайте определения:

1. Точки локального минимума функции $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$,
2. Точки глобального минимума функции $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$,
3. Предела функции $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ в точке x_0,
4. Непрерывности функции $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ в точке x_0.

Для понимания может быть полезно сначала дать эти определения в случае $\mathbb{R}^2$, то есть $n = 2$.

По сути, в этой задаче мы просим вас обобщить понятия, которые мы уже определяли для случая $\mathbb{R}^1$ в первых двух неделях. Вот что может пригодиться:

1. Определение локального минимума функции для случая $\mathbb{R}^1$. В определении для случая $\mathbb{R}^1$ мы использовали ε-окрестность в $\mathbb{R}^1$. На прошлых шагах мы научились определять ε-окрестность в $\mathbb{R}^n$.
2. Определение глобального минимума функции для случая $\mathbb{R}^1$.
3. Определение предела функции для случая $\mathbb{R}^1$. Предел функции для случая $\mathbb{R}^1$ определялся через предел последовательности в $\mathbb{R}^1$. На прошлых шагах мы научились определять предел последовательности в $\mathbb{R}^n$.
4. Определение непрерывности функции в точке для случая $\mathbb{R}^1$.

Пример 1. Точка $(0,0)$ является точкой локального минимума функции $f(x) = x_1^2 + x_2^2 + 3$, потому что

$$f(0,0) = 3 \leq x_1^2 + x_2^2 + 3 = f(x_1, x_2)$$

для любого $x = (x_1, x_2)$ из 5-окрестности точки $(0,0)$.

Пример 2. Точка $(0,0)$ также является и точкой глобального минимума функции $f(x) = x_1^2 + x_2^2 + 3$, потому что

$$f(0,0) = 3 \leq x_1^2 + x_2^2 + 3 = f(x_1, x_2)$$

для любого $x = (x_1, x_2)$. Тем самым, глобальный минимум функции f это $f(0,0) = 3$

Пример 3. Рассмотрим функцию $f : \mathbb{R}^2 \to \mathbb{R}$, заданную условием $f(x_1, x_2) = x_1^2 + 3x_2$. Докажем, что предел этой функции в точке $(4,5)$ равен $4^2 + 3 \cdot 5 = 31$.

Для этого рассмотрим любую последовательность $\{z_i\}$, сходящуюся к точке $(4,5)$, такую что $z_i \neq (4,5)$ для всех z_i. Обозначим первую координату z_i за z_{i1}, а вторую за z_{i2}, то есть z_i это точка (z_{i1}, z_{i2}). Тогда

$$\lim_{i\to\infty}(f(z_i)) = \lim_{i\to\infty}(z_{i1}^2 + 3z_{i2}) = (\lim_{i\to\infty}(z_{i1}))^2 + 3\lim_{i\to\infty}(z_{i2}) = 4^2 + 3 \cdot 5 = 31,$$

где предпоследнее равенство выполнено по предыдущей задаче. Так как ответ 31 не зависит от выбора последовательности, мы доказали, что $\lim\limits_{x \to (4,5)} f(x) = 31$.

Пример 4. Дана функция f, заданная так $f(x_1, x_2) = 4$. Ясно, что $\lim\limits_{x \to (7,5)} f(x) = 4$ и $f(7,5) = 4$. Поэтому функция f непрерывна в точке $(7,5)$. Неформально: в окрестности точки $(7,5)$ функция f ведёт себя так же, как и в самой точке $(7,5)$.

Задача 4

Пусть $x, y \in \mathbb{R}^n$. Докажите, что $d(x,y) = ||x - y||$.

Попробуйте понять, какая геометрическая интерпретация есть у этого равенства.

Заметьте, что в левой части равенства мы используем расстояние между точками x и y, а в правой — длину вектора $x - y$.

4.4 Дифференциал

Задача 1

Пример. Докажем, что для любой точки $x = (x_1, x_2)$ дифференциал функции $f(x_1, x_2) = 3x_1$ равен

$$d_x f(\Delta x_1, \Delta x_2) = 3\ \Delta x_1 + 0\ \Delta x_2 = 3\Delta x_1.$$

Подставим f и $d_x f$ в наше определение дифференциала:

$$\lim_{(\Delta x_1, \Delta x_2)\to(0,0)} \frac{f(x_1 + \Delta x_1, x_2 + \Delta x_2) - (f(x) + a_1 \Delta x_1 + a_2 \Delta x_2)}{||(\Delta x_1, \Delta x_2)||} =$$
$$= \lim_{(\Delta x_1, \Delta x_2)\to(0,0)} \frac{3(x_1 + \Delta x_1) - (3x_1 + 3\Delta x_1 + 0\Delta x_2)}{||(\Delta x_1, \Delta x_2)||} =$$
$$= \lim_{(\Delta x_1, \Delta x_2)\to(0,0)} \frac{0}{||(\Delta x_1, \Delta x_2)||} = 0.$$

Что и требовалось доказать.

Задача.

1. В начале этого урока мы нестрого показали, что дифференциал функции $f(x_1, x_2) = 4 + 3x_1 + 5x_2$ в каждой точке равен $3\ \Delta x_1 + 5\ \Delta x_2$. Докажите это утверждение строго, используя определение дифференциала.

2. Найдите дифференциал функции $f(x_1, x_2) = p + qx_1 + rx_2$ в каждой точке.

Задача 2

В первой половине урока мы нестрого нашли дифференциалы функций $x_1^2 + x_2^2$ и x_1x_2. На этом шаге мы сделаем это строго.

Пример. Докажем, что для любой точки $x = (x_1, x_2)$ дифференциал функции $f(x_1, x_2) = x_1^2$ равен

$$d_x f(\Delta x_1, \Delta x_2) = 2x_1\Delta x_1 + 0\ \Delta x_2 = 2x_1\Delta x_1.$$

Подставим f и $d_x f$ в наше определение дифференциала:

$$\lim_{(\Delta x_1,\Delta x_2)\to(0,0)} \frac{f(x_1+\Delta x_1, x_2+\Delta x_2) - (f(x) + a_1\Delta x_1 + a_2\Delta x_2)}{||(\Delta x_1, \Delta x_2)||} =$$
$$= \lim_{(\Delta x_1,\Delta x_2)\to(0,0)} \frac{(x_1+\Delta x_1)^2 - (x_1^2 + 2x_1\Delta x_1 + 0\Delta x_2)}{||(\Delta x_1, \Delta x_2)||} =$$
$$= \lim_{(\Delta x_1,\Delta x_2)\to(0,0)} \frac{\cancel{x_1^2} + \cancel{2x_1\Delta x_1} + \Delta x_1^2 - \cancel{x_1^2} - \cancel{2x_1\Delta x_1}}{||(\Delta x_1, \Delta x_2)||} = \lim_{(\Delta x_1,\Delta x_2)\to(0,0)} \frac{\Delta x_1^2}{||(\Delta x_1, \Delta x_2)||}.$$

А этот предел равен нулю. Докажем это.

Неформальное доказательство. Числитель по модулю не больше, чем квадрат длины вектора $(\Delta x_1, \Delta x_2)$. Значит, сама дробь по модулю не больше длины вектора $(\Delta x_1, \Delta x_2)$. Когда $(\Delta x_1, \Delta x_2)$ стремится к нулю, длина вектора $(\Delta x_1, \Delta x_2)$ тоже стремится к нулю.

Формальное доказательство. Рассмотрим любую последовательность векторов $\{(\Delta x_{1n}, \Delta x_{2n})\}$ сходящуюся к точке $(0, 0)$. Нам нужно доказать, что тогда последовательность $\{\frac{\Delta x_{1n}^2}{||(\Delta x_{1n}, \Delta x_{2n})||}\}$ сходится к нулю.

1. Так как выполнено $|\Delta x_{1n}| \le ||(\Delta x_{1n}, \Delta x_{2n})||$, имеем $\frac{\Delta x_{1n}^2}{||(\Delta x_{1n}, \Delta x_{2n})||} \le \frac{||(\Delta x_{1n}, \Delta x_{2n})||^2}{||(\Delta x_{1n}, \Delta x_{2n})||} = ||(\Delta x_{1n}, \Delta x_{2n})||$

2. Так как последовательность векторов $\{(\Delta x_{1n}, \Delta x_{2n})\}$ сходится к точке $(0, 0)$, последовательность $||(\Delta x_{1n}, \Delta x_{2n})||$ сходится к нулю.

3. Так как $0 \le \frac{\Delta x_{1n}^2}{||(\Delta x_{1n}, \Delta x_{2n})||} \le ||(\Delta x_{1n}, \Delta x_{2n})||$ и последовательность $||(\Delta x_{1n}, \Delta x_{2n})||$ сходится к нулю, последовательность $\{\frac{\Delta x_{1n}^2}{||(\Delta x_{1n}, \Delta x_{2n})||}\}$ тоже сходится к нулю.

Что и требовалось доказать.

Комментарий. Несколько шагов назад мы нестрого говорили, что член со второй степенью Δx_1 можно игнорировать. Рассуждение выше формализует это утверждение.

Задача.

1. Докажите, что дифференциал функции $f(x_1, x_2) = x_1^2 + x_2^2$ в точке (x_1, x_2) равен $2x_1\Delta x_1 + 2x_2\Delta x_2$.

2. Докажите, что дифференциал функции $f(x_1, x_2) = x_1x_2$ в точке (x_1, x_2) равен $x_2\ \Delta x_1 + x_1\Delta x_2$.

Задача 3

Пример.

- Пусть функция $f : \mathbb{R}^n \to \mathbb{R}$ дифференцируема в точке x, её дифференциал в точке x это $d_x f(\Delta x) := a_1\Delta x_1 + \cdots + a_n\ \Delta x_n$.

- Зафиксируем любое число $c \in \mathbb{R}$.

Докажем, что функция (cf) дифференцируема в точке x, и её дифференциал в точке x это

$$d_x(cf)(\Delta x) := ca_1\Delta x_1 + \cdots + ca_n\ \Delta x_n.$$

Доказательство. Подставим функцию $ca_1\Delta x_1 + \cdots + ca_n\ \Delta x_n$ в определение дифференциала функции (cf). Если полученный предел окажется равным нулю, то это и будет значить, что $ca_1\Delta x_1 + \cdots + ca_n\ \Delta x_n$ – дифференциал функции (cf) в точке x (в частности, мы получим, что дифференциал cf в точке x существует).

$$\lim_{\Delta x\ \to 0} \frac{(cf)(x+\Delta x) - ((cf)(x) + ca_1 \Delta x_1 + \dots + ca_n\ \Delta x_n)}{||\Delta x||} =$$
$$= \lim_{\Delta x\ \to 0} \frac{c \cdot f(x+\Delta x) - (c \cdot f(x) + c \cdot d_x f(\Delta x))}{||\Delta x||} =$$
$$= c \cdot \lim_{\Delta x\ \to 0} \frac{f(x+\Delta x) - (f(x) + d_x f(\Delta x))}{||\Delta x||} = c \cdot 0 = 0.$$

Что и требовалось доказать. Предпоследнее равенство выполнено, потому что $d_x f(\Delta x)$ это дифференциал функции f.

Комментарий. Тем самым, дифференциал ведёт себя как производная – при умножении функции на константу c, дифференциал умножается на c.

Задача.

- Пусть функция $f : \mathbb{R}^n \to \mathbb{R}$ дифференцируема в точке x, её дифференциал в точке x это $d_x f(\Delta x) := a_1 \Delta x_1 + \dots + a_n \Delta x_n$.
- Пусть функция $g : \mathbb{R}^n \to \mathbb{R}$ дифференцируема в точке x, её дифференциал в точке x это $d_x g(\Delta x) = b_1 \Delta x_1 + \dots + b_n \Delta x_n$.

Докажите, что сумма этих двух функций тоже дифференцируема в точке x, и её дифференциал в точке x равен

$$d_x f(\Delta x) + d_x g(\Delta x) = (a_1 + b_1)\Delta x_1 + \dots + (a_n + b_n)\Delta x_n.$$

Другими словами, докажите, что $d_x(f+g)(\Delta x) = d_x f(\Delta x) + d_x g(\Delta x)$.

Комментарий. Тем самым, дифференциал ведёт себя как производная – если f и g дифференцируемы в точке x, то $(f+g)$ тоже дифференцируема в точке x, и её дифференциал равен сумме дифференциала f и дифференциала g.

4.5 Частная производная

Задача 1

1. Чему равняется частная производная функции $f(x_1, x_2, x_3) = x_1^2 x_2 - x_1 x_2 x_3 + x_2^6 x_3^3 + x_3 + 10$ по x_2 в точке $(4, 1, -3)$?
2. Чему равняется частная производная функции $g(x_1, x_2) = x_1^4 - x_1^{x_2+2} - \operatorname{tg} x_2$ по x_1 в точке $(x_1, x_2) = (2, 3)$?
3. Чему равняется частная производная функции $h(x_1, x_2, x_3) = \ln(\cos(\frac{\sqrt{\frac{x_1^2 + x_3^2 - x_1 x_3}{2\sin(\operatorname{arctg}(x_3 - x_1))}} + x_1 - x_3^{sin x_1}}{e^{x_3}}) + \sin(\frac{2^{\operatorname{tg}(x_1 x_3)}}{\ln(x_1 x_3)})) + 2x_2$ по второй координате в точке $(x_1, x_2, x_3) = (e^{\sqrt{75}}, \cos 197, \operatorname{arctg} 1453)$?

Задача 2

1. Чему равняется частная производная функции $f(x_1, x_2, \dots, x_n) = 7 + x_1 + 2x_2 + \dots + nx_n$ по x_5 в точке $(10, 10, \dots, 10)$?
2. Чему равняется частная производная функции $g(x_1, x_2, \dots, x_9) = x_1 x_2 \dots x_9$ по x_4 в точке $(2, 2, \dots, 2)$?

Задача 3

Теорема. Дана функция f от n переменных. Пусть у f в точке x существует дифференциал $d_x f(\Delta x_1, \dots, \Delta x_n) = a_1 \Delta x_1 + \dots + a_n\ \Delta x_n$ и частные производные $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$. Тогда

$$a_1 = \frac{\partial f}{\partial x_1}, \dots, a_n\ = \frac{\partial f}{\partial x_n}.$$

То есть для любого $j = 1, \dots, n$ число a_j равно частной производной функции f по j-ой координате, вычисленной в точке x. Другими словами:

$$d_x f(\Delta x_1, \ldots, \Delta x_n) = \frac{\partial f}{\partial x_1}\Delta x_1 + \frac{\partial f}{\partial x_2}\Delta x_2 \ + \cdots + \frac{\partial f}{\partial x_n}\Delta x_n,$$

где все частные производные вычислены в точке x.

Задача. Докажите теорему.

Задача 4

Пример.

Вычислим дифференциал функции

$$f(x_1, x_2, x_3) = x_1 x_2^3 + x_2 x_3^2$$

в точке $(4, -2, 1)$. Воспользуемся теоремой с предыдущего шага и вычислим сначала частные производные:

- $\frac{\partial(x_1 x_2^3 + x_2 x_3^2)}{\partial x_1}\Big|_{(4,-2,1)} = x_2^3 = -8$
- $\frac{\partial(x_1 x_2^3 + x_2 x_3^2)}{\partial x_2}\Big|_{(4,-2,1)} = 3x_1 x_2^2 + x_3^2 = 49.$
- $\frac{\partial(x_1 x_2^3 + x_2 x_3^2)}{\partial x_3}\Big|_{(4,-2,1)} = 2x_2 x_3 = -4.$

Следовательно, дифференциал f в этой точке будет равен

$$d_x f(\Delta x_1, \Delta x_2) = -8\Delta x_1 + 49\Delta x_2 - 4\Delta x_3.$$

Задача.

1. Найдите дифференциал функции $g(x_1, x_2, x_3) = -x_1^4 x_3 + x_1 x_2^3 x_3 + x_1 x_3^2$ в точке $(2, 5, 4)$
2. Найдите дифференциал функции $h(x_1, x_2, \ldots, x_n) = x_1^2 x_2 + x_2^2 x_3 + \cdots + x_{n-1}^2 x_n + x_n^2 x_1$ в точке $(2, 2, \ldots, 2)$

Задача 5

Как вы помните, в случае функции одной переменной верно следующее утверждение:

Утверждение. Если функция f непрерывна и дифференцируема на отрезке $[a, b]$, то локальный минимум или максимум не может достигаться в точках $x \in (a, b)$, в которых $f'(x) \neq 0$.

Сформулируем аналогичное утверждение для функции от нескольких переменных.

Теорема. Дана функция f от n переменных. Пусть f определена в некоторой окрестности точки x, и в точке x у f существуют частные производные по всем координатам. Тогда x может быть точкой локального минимума или максимума только если все частные производные равны нулю.

Другими словами, если для какого-то k частная производная f по k-ой координате не равна нулю, то x не может быть точкой локального минимума или максимума.

Следствие. Пусть в точке x так же существует дифференциал $d_x f$. Точка x может быть точкой локального минимума или максимума, только если $d_x f = 0$ (то есть $d_x f(\Delta x_1, \ldots, \Delta x_n) = 0$ для любых $\Delta x_1, \ldots, \Delta x_n$).

Геометрически думать про это можно так: точка x может быть локальным минимумом или максимумом, если касательная плоскость к графику f в точке x будет горизонтальной (см геометрическую интерпретацию дифференциала с этого шага)

Задача. Докажите теорему и следствие.

Комментарий. Уровень строгости доказательства остаётся на ваше усмотрение (можно ограничиться уровнем строгости рассуждения с этого шага)

4.6 Направление и градиент

Задача 1

Дан произвольный вектор $a := (a_1, \ldots, a_n) \in \mathbb{R}^n$, не равный нулевому вектору. Докажите, что вектор $(\frac{-a_1}{||a||}, \ \frac{-a_2}{||a||}, \ldots \ , \frac{-a_n}{||a||}) = -\frac{a}{||a||}$ действительно является направлением – то есть, что его длина действительно равна 1.

Задача 2

Как мы уже видели, у любого ненулевого вектора есть длина и направление. Докажите, что они однозначно задают вектор.

Другими словами, если дано положительное число l и вектор e длины 1, то существует и единственен вектор a, такой что длина a это l, а направление a это e.

Задача для проверки. Известно, что вектор a имеет направление $\left(\frac{3}{5}, \frac{-4}{5}\right)$ и длину 10. Восстановите вектор a по этой информации.

Задача 3

Пример. Дана функция $f(x_1, x_2) = x_1 x_2$, мы находимся в точке $x = (5, 12)$. Найдём значение функции в точке, в которой мы окажемся, совершив из начальной точки шаг длины 2 по направлению, противоположному направлению градиента. Будет ли значение f в этой точке меньше, чем значение f в точке x?

- Сначала вычислим градиент:
$$\nabla f(x) = (x_2, x_1) = (12, 5).$$

- Следовательно, шаг мы будет делать в направлении
$$\frac{-\nabla f(x)}{||\nabla f(x)||} = (-\frac{12}{\sqrt{5^2+12^2}}, -\frac{5}{\sqrt{5^2+12^2}}) = (-\frac{12}{13}, -\frac{5}{13}).$$

- Так как длина шага должна быть равна 2, мы сместимся на вектор $2 \cdot (-\frac{12}{13}, -\frac{5}{13}) = (-\frac{24}{13}, -\frac{10}{13})$, и попадём в точку $(5 - \frac{24}{13}, 12 - \frac{10}{13})$.

- Значение функции в получившейся точке тогда будет равно
$$f(5 - \frac{24}{13}, 12 - \frac{10}{13}) = \frac{41}{13} \cdot \frac{146}{13} \approx 35.42,$$
что действительно меньше, чем $f(5, 12) = 5 \cdot 12 = 60$.

Задача. Дана функция $f(x_1, x_2) = (x_1 + x_2)^2$, мы находимся в точке $(3, 4)$. Значение в этой точке равно $f(3, 4) = (3 + 4)^2 = 49$.

Мы делаем шаг длины $2\sqrt{2}$ в направлении, противоположном направлению градиента.

Каково значение функции f в точке, в которой мы оказались?