

# Содержание

Кликните по названию лекции, чтобы перейти к ней в документе.

Лекция 2. Сведения из математического анализа .....	2
Лекция 3. Теория квадратичных форм и основы оптимизации .....	24
Лекция 4. Методы одномерной оптимизации .....	38
Лекция 5. Методы спуска .....	52
Лекция 6. Градиентный метод и правила выбора шага .....	58
Лекция 7. Градиентные методы. Сильная выпуклость и эффект оврагов .....	73
Лекция 8. Градиентный спуск. Эвристики, покоординатный метод .... и метод сопряжённых направлений	76
Лекция 9. Метод сопряжённых направлений (продолжение). .... Метод сопряжённых градиентов	83
Лекция 10. Метод Ньютона. Квазиньютоновские методы. .... Условная оптимизация	92

## Лекция 2. Сведения из математического анализа

### 1. Дифференцирование скалярных функций векторного аргумента

#### Определение 2.1. Дифференцируемость и градиент

Пусть задана скалярная функция

$$f(x): \mathbb{R}^n \rightarrow \mathbb{R}.$$

**1)** Функция  $f(x)$  называется *дифференцируемой в точке*  $x \in \mathbb{R}^n$ , если существует вектор  $a \in \mathbb{R}^n$  такой, что для всех  $y \in \mathbb{R}^n$  выполняется

$$f(x + y) = f(x) + \langle a, y \rangle + o(y), \quad y \rightarrow 0. \quad (1)$$

Здесь  $\langle \cdot, \cdot \rangle$  — стандартное скалярное произведение в  $\mathbb{R}^n$ , а  $o(y)$  — величина, удовлетворяющая

$$\frac{o(y)}{\|y\|} \rightarrow 0 \quad \text{при } y \rightarrow 0.$$

**2)** Вектор  $a$  называется *градиентом функции*  $f(x)$  в точке  $x$  и обозначается

$$a = \nabla f(x).$$

Таким образом, дифференцируемость эквивалентна существованию линейной аппроксимации первого порядка функции в данной точке (л. 2, стр. 1–2).

Лекция 2. Сведения из математич...

Из (1) следует эквивалентная форма записи:

$$f(x + y) = f(x) + \langle \nabla f(x), y \rangle + o(y). \quad (2)$$

Следовательно, функция  $f(x)$  дифференцируема в точке  $x$  тогда и только тогда, когда она допускает в этой точке линейную аппроксимацию первого порядка (л. 2, стр. 2).

#### Связь градиента с частными производными

Если функция  $f(x)$  дифференцируема в точке  $x$ , то её градиент выражается через частные производные:

$$\nabla f(x) = \left( \frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right).$$

Градиент можно вычислять:

- непосредственно из определения дифференцируемости;
- с помощью координатной записи;
- используя правила дифференцирования сложных функций (л. 2, стр. 2–3).

## Пример: квадратичная функция

Пусть

$$f(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle,$$

где:

- $A = (a_{ij})_{i,j=1}^n$  — симметричная матрица размера  $n \times n$ ,
- $b \in \mathbb{R}^n$ ,
- $x \in \mathbb{R}^n$ .

Рассмотрим приращение:

$$\begin{aligned} f(x+y) &= \frac{1}{2} \langle A(x+y), x+y \rangle - \langle b, x+y \rangle \\ &= \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle + \langle Ax - b, y \rangle + \frac{1}{2} \langle Ay, y \rangle. \end{aligned}$$

Отсюда:

$$f(x+y) = f(x) + \langle Ax - b, y \rangle + \frac{1}{2} \langle Ay, y \rangle.$$

Так как

$$|\langle Ay, y \rangle| \leq \|A\| \|y\|^2,$$

то  $\frac{1}{2} \langle Ay, y \rangle = o(\|y\|)$  при  $y \rightarrow 0$ . Следовательно,

$$\nabla f(x) = Ax - b. \quad (3)$$

(л. 2, стр. 3-5)

## Замечание о нормах матриц

Норма квадратной матрицы  $A = (a_{ij})$  определяется как

$$\|A\| = \left( \sum_{i,j=1}^n |a_{ij}|^2 \right)^{1/2},$$

и называется *евклидовой нормой* или *нормой Фробениуса*.

Также вводится *операторная норма*:

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|,$$

для которой всегда выполняется оценка

$$\|Ax\| \leq \|A\| \|x\|.$$

(л. 2, стр. 5–6)

## Дифференцируемость на множестве

### Определение 2.2

Функция  $f(x): Q \subset \mathbb{R}^n \rightarrow \mathbb{R}$  называется *дифференцируемой на множестве  $Q$* , если она дифференцируема в каждой точке  $x \in Q$ .

Если множество  $Q = \mathbb{R}^n$ , то говорят, что функция дифференцируема всюду (л. 2, стр. 6).

## Связь с функцией одной переменной

Рассмотрим функцию  $\varphi(t): \mathbb{R} \rightarrow \mathbb{R}$ , такую что:

- $\varphi(t)$  непрерывна и дифференцируема,
- $\varphi'(t)$  интегрируема.

Тогда имеет место формула Ньютона–Лейбница:

$$\varphi(1) = \varphi(0) + \int_0^1 \varphi'(z) dz.$$

Пусть теперь  $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$  дифференцируема на отрезке

$$Q = [x, x + y] = \{x + ty \mid 0 \leq t \leq 1\}.$$

Введём функцию одной переменной:

$$\varphi(t) = f(x + ty).$$

Тогда

$$\frac{\varphi(t + \Delta t) - \varphi(t)}{\Delta t} = \frac{f(x + (t + \Delta t)y) - f(x + ty)}{\Delta t} = \langle \nabla f(x + ty), y \rangle + o(\|y\|).$$

Переходя к пределу при  $\Delta t \rightarrow 0$ , получаем:

$$\varphi'(t) = \langle \nabla f(x + ty), y \rangle. \quad (4)$$

Следовательно, функция  $\varphi(t)$  дифференцируема на  $[0, 1]$  (л. 2, стр. 7–9).

## Производная по направлению

### Определение 2.3

Величина

$$f'(x; y) = \lim_{\varepsilon \rightarrow 0} \frac{f(x + \varepsilon y) - f(x)}{\varepsilon} \quad (5)$$

называется *производной функции  $f(x)$  по направлению  $y$  в точке  $x$* .

Производная по направлению может существовать даже для недифференцируемых функций. Например, для функции

$$f(x) = \|x\|, \quad x \in \mathbb{R}^n,$$

имеем

$$f'(0; y) = \|y\|.$$

Если функция  $f(x)$  имеет в точке  $x$  производную по всем направлениям, причём

$$f'(x; y) = \langle a, y \rangle, \quad a \in \mathbb{R}^n,$$

то функция  $f(x)$  дифференцируема по Гато в точке  $x$ , и

$$a = \left( \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right).$$

Из формулы (4) следует, что если  $f(x)$  дифференцируема в точке  $x$ , то она дифференцируема и по Гато, причём

$$f'(x; y) = \langle \nabla f(x), y \rangle. \quad (6)$$

Обратное утверждение в общем случае неверно (л. 2, стр. 9–10).

Источники

## Обратимость дифференцируемости по Гато: контрпример

Рассмотрим функцию  $f(x): \mathbb{R}^n \rightarrow \mathbb{R}, n \geq 2$ , заданную формулой

$$f(x) = \begin{cases} 1, & \text{если } \|x - a\| = \|x\|, x \neq 0, \\ 0, & \text{в остальных точках,} \end{cases}$$

где  $a \in \mathbb{R}^n, a \neq 0$ .

Эта функция:

- имеет производную по направлению в нуле по любому направлению  $y$ , причём

$$f'(0; y) = 0 \quad \forall y \in \mathbb{R}^n,$$

- следовательно, дифференцируема по Гато в точке 0,

однако:

- не является дифференцируемой в смысле Фреше в точке 0,
- более того, она не является непрерывной в нуле.

Этот пример показывает, что существование производных по всем направлениям недостаточно для дифференцируемости в смысле (1) (л. 2, стр. 11).

## Интегральная форма остаточного члена

Пусть функция  $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$  дифференцируема на отрезке

$$Q = [x, x + y].$$

Как и ранее, введём функцию

$$\varphi(t) = f(x + ty).$$

Используя формулу для производной (4) и формулу Ньютона–Лейбница, получаем:

$$\begin{aligned} f(x + y) &= \varphi(1) = \varphi(0) + \int_0^1 \varphi'(z) dz \\ &= f(x) + \int_0^1 \langle \nabla f(x + zy), y \rangle dz. \end{aligned} \tag{7}$$

Выделяя линейную часть, перепишем:

$$\begin{aligned} f(x + y) &= f(x) + \langle \nabla f(x), y \rangle \\ &\quad + \int_0^1 \langle \nabla f(x + zy) - \nabla f(x), y \rangle dz. \end{aligned} \tag{8}$$

Формула (8) является *интегральной формой остаточного члена* разложения функции в окрестности точки  $x$  (л. 2, стр. 12–13).

## Теорема о среднем для скалярных функций

Из формулы (7) следует теорема о среднем.

Действительно, так как функция

$$z \mapsto \langle \nabla f(x + zy), y \rangle$$

непрерывна на  $[0, 1]$ , то существует  $\theta \in [0, 1]$  такое, что

$$\int_0^1 \langle \nabla f(x + zy), y \rangle dz = \langle \nabla f(x + \theta y), y \rangle.$$

Следовательно,

$$f(x + y) = f(x) + \langle \nabla f(x + \theta y), y \rangle, \quad 0 \leq \theta \leq 1. \tag{9}$$

(л. 2, стр. 13)

## Задачи

1. Доказать, что

$$\nabla \|x\| = \frac{x}{\|x\|}, \quad x \neq 0.$$

2. Доказать, что функция  $\|x\|$  не дифференцируема в точке  $x = 0$ .

3. Доказать, что непрерывность по  $x$  производной по Гато влечёт дифференцируемость функции  $f(x)$ .

(л. 2, стр. 13)

## 2. Вторые производные скалярных функций векторного аргумента

### Определение 2.4

Скалярная функция

$$f(x): \mathbb{R}^n \rightarrow \mathbb{R}$$

называется *дважды дифференцируемой* в точке  $x \in \mathbb{R}^n$ , если:

- она дифференцируема в точке  $x$ ,
- существует симметричная матрица  $H \in \mathbb{R}^{n \times n}$  такая, что

$$f(x + y) = f(x) + \langle \nabla f(x), y \rangle + \frac{1}{2} \langle Hy, y \rangle + o(\|y\|^2).$$

Матрица  $H$  называется *матрицей вторых производных* (матрицей Гессе, гессианом) и обозначается:

$$H = f''(x) = \nabla^2 f(x). \quad (10)$$

(л. 2, стр. 14)

### Эквивалентное определение

Функция  $f(x)$  дважды дифференцируема в точке  $x$ , если она допускает *квадратичную аппроксимацию* в окрестности этой точки, то есть существует квадратичная функция

$$\tilde{f}(y) = f(x) + \langle \nabla f(x), y \rangle + \frac{1}{2} \langle \nabla^2 f(x) y, y \rangle,$$

такая, что

$$|f(x + y) - \tilde{f}(y)| = o(\|y\|^2).$$

(л. 2, стр. 15)

## Оценки для дважды дифференцируемых функций

Пусть  $f(x)$  дважды дифференцируема на множестве

$$Q = \{x + ty \mid 0 \leq t \leq 1\}.$$

Рассмотрим функцию

$$\varphi(t) = f(x + ty).$$

Тогда:

- $\varphi(t)$  дважды дифференцируема,
- её вторая производная имеет вид

$$\varphi''(t) = \langle \nabla^2 f(x + ty)y, y \rangle. \quad (11)$$

Кроме того, справедливо равенство:

$$\varphi(1) = \varphi(0) + \varphi'(0) + \int_0^1 \int_0^t \varphi''(z) dz dt. \quad (12)$$

(л. 2, стр. 16–17)

Подставляя выражения для  $\varphi(0)$ ,  $\varphi'(0)$  и  $\varphi''(t)$ , получаем интегральную формулу:

$$f(x + y) = f(x) + \langle \nabla f(x), y \rangle + \int_0^1 \int_0^t \langle \nabla^2 f(x + zy)y, y \rangle dz dt. \quad (13)$$

### Квадратичная оценка остаточного члена

Предположим, что на отрезке  $Q$  выполнено ограничение

$$\|\nabla^2 f(x + ty)\| \leq L. \quad (14)$$

Тогда из (13) следует оценка:

$$|f(x + y) - f(x) - \langle \nabla f(x), y \rangle| \leq \frac{L}{2} \|y\|^2. \quad (15)$$

(л. 2, стр. 18–19)

### Кубическая оценка при липшицевости гессиана

Если дополнительно выполнено условие Липшица:

$$\|\nabla^2 f(x + ty) - \nabla^2 f(x)\| \leq L t \|y\|,$$

то имеет место более точная оценка:



$$|f(x+y) - f(x) - \langle \nabla f(x), y \rangle - \frac{1}{2} \langle \nabla^2 f(x) y, y \rangle| \leq \frac{L}{6} \|y\|^3. \quad (16)$$

(л. 2, стр. 19)

## Формула Тейлора с остаточным членом в форме Лагранжа

Существует  $\theta \in [0, 1]$  такое, что

$$f(x+y) = f(x) + \langle \nabla f(x), y \rangle + \frac{1}{2} \langle \nabla^2 f(x + \theta y) y, y \rangle. \quad (17)$$

(л. 2, стр. 20)

## Задачи

1. Показать, что элементы матрицы Гессе равны соответствующим вторым частным производным:

$$(\nabla^2 f(x))_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}.$$

2. Доказать, что

$$\nabla^2 \langle Ax, x \rangle = A + A^\top,$$

а в случае симметричной матрицы  $A$  выполняется  $\nabla^2 \langle Ax, x \rangle = 2A$ .

(л. 2, стр. 20)

**Задание: проверить аксиомы нормы для матричной нормы**  $\|A\| = \sup_{\|x\|=1} \|Ax\|$

(л. 2, стр. 5-6)

Пусть  $\|\cdot\|$  — евклидова норма в  $\mathbb{R}^n$ .

1. **Неотрицательность:**  $\|A\| \geq 0$ , так как  $\|Ax\| \geq 0$  для всех  $x$ , а супремум неотрицательных чисел неотрицателен.
2. **Невырожденность:** если  $\|A\| = 0$ , то  $\|Ax\| = 0$  для всех  $\|x\| = 1$ , значит  $Ax = 0$  для всех таких  $x$ . Тогда для любого  $x \neq 0$ :  $x = \|x\| \cdot \frac{x}{\|x\|}$ , и

$$Ax = \|x\| A \left( \frac{x}{\|x\|} \right) = \|x\| \cdot 0 = 0,$$

то есть  $A = 0$ . Обратно: если  $A = 0$ , то  $\|A\| = 0$ .

3. **Однородность:** для  $\alpha \in \mathbb{R}$

$$\|\alpha A\| = \sup_{\|x\|=1} \|\alpha Ax\| = |\alpha| \sup_{\|x\|=1} \|Ax\| = |\alpha| \|A\|.$$

#### 4. Неравенство треугольника:

$$\|A + B\| = \sup_{\|x\|=1} \|(A + B)x\| \leq \sup_{\|x\|=1} (\|Ax\| + \|Bx\|) \leq \sup_{\|x\|=1} \|Ax\| + \sup_{\|x\|=1} \|Bx\| = \|A\| + \|B\|.$$

**Задача: вывести оценку  $\|Ax\| \leq \|A\| \|x\|$  из определения подчинённой нормы (л. 2, стр. 6)**

Если  $x = 0$ , то верно тривиально. Пусть  $x \neq 0$ . Тогда  $\left\|\frac{x}{\|x\|}\right\| = 1$ , поэтому по определению  $\|A\|$ :

$$\left\|A \frac{x}{\|x\|}\right\| \leq \sup_{\|u\|=1} \|Au\| = \|A\|.$$

Умножая на  $\|x\|$ , получаем

$$\|Ax\| \leq \|A\| \|x\|.$$

**Задача 1: доказать  $\nabla\|x\| = \frac{x}{\|x\|}$ ,  $x \neq 0$  (л. 2, стр. 13)**

Пусть  $f(x) = \|x\| = (\langle x, x \rangle)^{1/2}$ . Для  $x \neq 0$  функция гладкая, и по правилу цепочки:

$$\frac{\partial}{\partial x_i} \|x\| = \frac{\partial}{\partial x_i} (\langle x, x \rangle^{1/2}) = \frac{1}{2} \langle x, x \rangle^{-1/2} \cdot 2x_i = \frac{x_i}{\|x\|}.$$

Значит,

$$\nabla\|x\| = \left( \frac{x_1}{\|x\|}, \dots, \frac{x_n}{\|x\|} \right) = \frac{x}{\|x\|}.$$

**Задача 2: доказать, что  $\|x\|$  не дифференцируема в  $x = 0$  (л. 2, стр. 13)**

По определению производной по направлению (5):

$$f'(0; y) = \lim_{\varepsilon \rightarrow 0} \frac{\|\varepsilon y\| - \|0\|}{\varepsilon} = \lim_{\varepsilon \rightarrow 0} \frac{|\varepsilon| \|y\|}{\varepsilon}.$$

Предел справа не существует как двусторонний (слева даёт  $-\|y\|$ , справа  $+\|y\|$ ); в лекции отмечено, что для  $\|x\|$  в нуле производная по направлению рассматривается как  $f'(0; y) = \|y\|$  (стр. 9), но главное здесь: отображение  $y \mapsto \|y\|$  **не линейно**.

Если бы  $f$  была дифференцируема в нуле, то по формуле (6) из дифференцируемости следовало бы

$$f'(0; y) = \langle \nabla f(0), y \rangle,$$

то есть зависимость от  $y$  линейная. Противоречие с нелинейностью  $y \mapsto \|y\|$ . Следовательно,  $f(x) = \|x\|$  не дифференцируема в 0.

### Задача 3: доказать, что непрерывность по $x$ производной Гато влечёт дифференцируемость (л. 2, стр. 13)

Пусть в окрестности точки  $x$  существует производная по Гато  $f'(u; \cdot)$ , причём:

- для каждого  $u$  отображение  $y \mapsto f'(u; y)$  линейно;
- как функция от  $u$  (в норме линейных функционалов) она непрерывна:  $f'(u; \cdot) \rightarrow f'(x; \cdot)$  при  $u \rightarrow x$ .

Рассмотрим  $\varphi(t) = f(x + ty)$ . Тогда по смыслу производной по направлению:

$$\varphi'(t) = f'(x + ty; y).$$

Непрерывность по  $x$  производной Гато даёт непрерывность  $\varphi'(t)$  по  $t$  на  $[0, 1]$ , поэтому по формуле Ньютона–Лейбница:

$$f(x + y) - f(x) = \varphi(1) - \varphi(0) = \int_0^1 f'(x + ty; y) dt.$$

Вычтем линейную часть  $f'(x; y)$ :

$$f(x + y) - f(x) - f'(x; y) = \int_0^1 (f'(x + ty; y) - f'(x; y)) dt.$$

Оценим модуль, используя норму линейного функционала:

$$|f'(x + ty; y) - f'(x; y)| \leq \|f'(x + ty; \cdot) - f'(x; \cdot)\| \|y\|.$$

Так как  $x + ty \rightarrow x$  при  $y \rightarrow 0$  равномерно по  $t \in [0, 1]$ , по непрерывности  $f'(x + ty; \cdot) \rightarrow f'(x; \cdot)$  и, значит,

$$\sup_{t \in [0, 1]} \|f'(x + ty; \cdot) - f'(x; \cdot)\| \xrightarrow{y \rightarrow 0} 0.$$

Следовательно,

$$\frac{|f(x + y) - f(x) - f'(x; y)|}{\|y\|} \leq \sup_{t \in [0, 1]} \|f'(x + ty; \cdot) - f'(x; \cdot)\| \xrightarrow{y \rightarrow 0} 0,$$

то есть

$$f(x + y) = f(x) + f'(x; y) + o(\|y\|),$$

а значит  $f$  дифференцируема в смысле (1), причём  $\langle \nabla f(x), y \rangle = f'(x; y)$ .

### Задача: получить оценку (15) из (13) при условии (14) (л. 2, стр. 19)

Из (13):

$$f(x+y) - f(x) - \langle \nabla f(x), y \rangle = \int_0^1 \int_0^t \langle \nabla^2 f(x+zy)y, y \rangle dz dt.$$

Берём модуль и используем  $|\langle Hy, y \rangle| \leq \|H\| \|y\|^2$  и (14):

$$|f(x+y) - f(x) - \langle \nabla f(x), y \rangle| \leq \int_0^1 \int_0^t \|\nabla^2 f(x+zy)\| \|y\|^2 dz dt \leq L \|y\|^2 \int_0^1 \int_0^t dz dt.$$

Но  $\int_0^1 \int_0^t dz dt = \int_0^1 t dt = \frac{1}{2}$ . Поэтому

$$|f(x+y) - f(x) - \langle \nabla f(x), y \rangle| \leq \frac{L}{2} \|y\|^2,$$

то есть (15).

**Задача: доказать оценку (16) при условии Липшица для  $\nabla^2 f$  (л. 2, стр. 19)**

Рассмотрим разность между (13) и квадратичным членом:

$$R := f(x+y) - f(x) - \langle \nabla f(x), y \rangle - \frac{1}{2} \langle \nabla^2 f(x)y, y \rangle.$$

Заметим, что

$$\frac{1}{2} \langle \nabla^2 f(x)y, y \rangle = \left( \int_0^1 \int_0^t dz dt \right) \langle \nabla^2 f(x)y, y \rangle = \int_0^1 \int_0^t \langle \nabla^2 f(x)y, y \rangle dz dt.$$

Тогда из (13):

$$R = \int_0^1 \int_0^t \langle (\nabla^2 f(x+zy) - \nabla^2 f(x))y, y \rangle dz dt.$$

Берём модуль и используем  $|\langle Hy, y \rangle| \leq \|H\| \|y\|^2$ :

$$|R| \leq \int_0^1 \int_0^t \|\nabla^2 f(x+zy) - \nabla^2 f(x)\| \|y\|^2 dz dt.$$

По условию Липшица (в записи лекции):

$$\|\nabla^2 f(x+zy) - \nabla^2 f(x)\| \leq L z \|y\|.$$

Подставляем:

$$|R| \leq L \|y\|^3 \int_0^1 \int_0^t z dz dt.$$

Считаем интеграл:

$$\int_0^1 \int_0^t z dz dt = \int_0^1 \frac{t^2}{2} dt = \frac{1}{6}.$$

Итого:

$$|R| \leq \frac{L}{6} \|y\|^3,$$

что и есть (16).

**Задача 1: показать, что  $(\nabla^2 f(x))_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$  (л. 2, стр. 20)**

Из определения дважды дифференцируемости (10) получаем асимптотику при  $y \rightarrow 0$ :

$$f(x+y) = f(x) + \langle \nabla f(x), y \rangle + \frac{1}{2} \langle \nabla^2 f(x) y, y \rangle + o(\|y\|^2).$$

Пусть  $e_i$  — стандартный базис. Подставим  $y = h e_i$ :

$$f(x + h e_i) = f(x) + h \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} h^2 \langle \nabla^2 f(x) e_i, e_i \rangle + o(h^2).$$

Сравнивая с разложением функции одной переменной по  $h$ , получаем

$$\frac{\partial^2 f}{\partial x_i^2}(x) = \langle \nabla^2 f(x) e_i, e_i \rangle = (\nabla^2 f(x))_{ii}.$$

Для смешанных производных используем поляризацию квадратичной формы  $q(y) = \langle \nabla^2 f(x) y, y \rangle$ :

$$\langle \nabla^2 f(x) e_i, e_j \rangle = \frac{1}{4} (q(e_i + e_j) - q(e_i - e_j)).$$

А коэффициенты при  $h k$  в разложении  $f(x + h e_i + k e_j)$  совпадают со смешанной производной  $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ . Отсюда

$$(\nabla^2 f(x))_{ij} = \langle \nabla^2 f(x) e_j, e_i \rangle = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}.$$

**Задача 2: найти  $\nabla^2 f(x)$  для  $f(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle$  (л. 2, стр. 20; ср. пример на стр. 3-5)**

Из примера получено

$$\nabla f(x) = Ax - b \quad (\text{формула (3)}).$$

Градиент — аффинная функция, поэтому её производная по  $x$  постоянна и равна  $A$ . То есть

$$\nabla^2 f(x) = A \quad (\text{при симметричной } A),$$

а в координатах:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = a_{ij}.$$

Если  $A$  не предполагать симметричной, то  $\frac{1}{2}\langle Ax, x \rangle = \frac{1}{2}x^\top Ax = \frac{1}{2}x^\top \frac{A+A^\top}{2}x$ , и тогда  $\nabla^2 f(x) = \frac{A+A^\top}{2}$ ; в лекции в этом месте  $A$  уже симметрична.

### 3. Дифференцирование вектор-функций

Под *вектор-функцией* понимается отображение вида

$$g(x): \mathbb{R}^n \rightarrow \mathbb{R}^m,$$

в общем случае  $m \neq n$  (л. 2, стр. 21).

#### Определение 2.5

Функция

$$g(x): \mathbb{R}^n \rightarrow \mathbb{R}^m$$

называется *дифференцируемой в точке*  $x \in \mathbb{R}^n$ , если существует матрица

$$A \in \mathbb{R}^{m \times n}$$

такая, что для всех  $y \in \mathbb{R}^n$

$$g(x+y) = g(x) + Ay + o(y), \quad y \rightarrow 0. \quad (18)$$

Матрица  $A$  называется *производной* (или *матрицей Якоби*) отображения  $g$  в точке  $x$  и обозначается

$$A = g'(x).$$

Таким образом, дифференцируемость вектор-функции означает существование линейной аппроксимации первого порядка (л. 2, стр. 21–22).

#### Эквивалентное определение

Отображение  $g(x)$  дифференцируемо в точке  $x$  тогда и только тогда, когда оно допускает в этой точке линейную аппроксимацию вида (18).

Элементы матрицы Якоби выражаются через частные производные компонент:

$$(g'(x))_{ij} = \frac{\partial g_i(x)}{\partial x_j}. \quad (19)$$

(л. 2, стр. 22)

## Правило дифференцирования сложных функций (цепное правило)

Пусть:

- $g(x): \mathbb{R}^n \rightarrow \mathbb{R}^m$  — дифференцируемая в точке  $x$  функция,
- $h(u): \mathbb{R}^m \rightarrow \mathbb{R}^s$  — дифференцируемая в точке  $u = g(x)$  функция.

Тогда композиция

$$(h \circ g)(x) = h(g(x))$$

дифференцируема в точке  $x$ , причём выполняется *цепное правило*:

$$(h \circ g)'(x) = h'(g(x)) \cdot g'(x). \quad (19)$$

Здесь произведение — обычное произведение матриц размеров  $(s \times m)$  и  $(m \times n)$ .

Размерность матрицы производной композиции определяется как

$$(s \times m) \cdot (m \times n) = (s \times n).$$

(л. 2, стр. 23)

## Теорема о среднем для вектор-функций (неверность)

В отличие от скалярного случая, *теорема о среднем* для вектор-функций в общем случае **неверна**.

### Утверждение

Не существует такого  $\theta \in [0, 1]$ , что для произвольной дифференцируемой на отрезке

$$Q = [x, x + y]$$

функции

$$g(x): \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad m > 1,$$

выполнялось бы равенство

$$g(x + y) = g(x) + g'(x + \theta y) y.$$

(л. 2, стр. 24)

Вместо этого справедлива *интегральная формула* (сравнить с формулой (8) для скалярных функций):

$$g(x + y) = g(x) + g'(x)y + \int_0^1 (g'(x + ty) - g'(x))y dt. \quad (20)$$

(л. 2, стр. 24)

## Следствия и оценки

Из формулы (20) выводятся следующие полезные оценки.

### 1. Линейная оценка

Если

$$\|g'(x + ty)\| \leq L, \quad 0 \leq t \leq 1,$$

то

$$\|g(x + y) - g(x)\| \leq L \|y\|. \quad (21)$$

(л. 2, стр. 25)

---

### 2. Квадратичная оценка при липшицевости производной

Если производная  $g'(x)$  удовлетворяет условию Липшица на множестве

$$Q = [x, x + y],$$

то есть

$$\|g'(u) - g'(v)\| \leq L \|u - v\|, \quad u, v \in Q,$$

то выполняется оценка:

$$\|g(x + y) - g(x) - g'(x)y\| \leq \frac{L}{2} \|y\|^2. \quad (22)$$

(л. 2, стр. 25)

---

## Задача

Доказать, что

$$\nabla \|Ax - b\| = 2A^\top (Ax - b),$$

где  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ .

(л. 2, стр. 25)

---

## Элементы выпуклого анализа

### 1. Выпуклые множества и их свойства

#### Определение 2.5

Множество  $Q \subset \mathbb{R}^n$  называется *выпуклым*, если для любых  $x, y \in Q$  и любого  $\lambda \in [0, 1]$  выполняется:



$$\lambda x + (1 - \lambda)y \in Q. \quad (21)$$

(л. 2, стр. 26)

## Выпуклые комбинации

Точка

$$\sum_{i=1}^m \lambda_i x_i \quad (22)$$

называется *выпуклой комбинацией* точек  $x_1, \dots, x_m \in Q$ , если

$$\lambda_i \geq 0, \quad \sum_{i=1}^m \lambda_i = 1.$$

По индукции доказывается, что любая выпуклая комбинация точек из  $Q$  также принадлежит  $Q$ :

$$\sum_{i=1}^m \lambda_i x_i \in Q. \quad (23)$$

(л. 2, стр. 26)

## Примеры

- Выпуклое множество: отрезок, шар, многогранник.
- Невыпуклое множество: множество с «дыркой» или разрывом (геометрические иллюстрации приведены в лекции).

(л. 2, стр. 27)

## Определение 2.6. Выпуклая функция

Функция

$$f(x): Q \subset \mathbb{R}^n \rightarrow \mathbb{R}$$

называется *выпуклой на выпуклом множестве  $Q$* , если для всех  $x, y \in Q$  и  $\lambda \in [0, 1]$

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y). \quad (24)$$

Если неравенство строгое при  $0 < \lambda < 1$  и  $x \neq y$ , то функция называется *строго выпуклой*.

(л. 2, стр. 27)

## Геометрическая интерпретация выпуклости

График выпуклой функции лежит *не выше* отрезка, соединяющего точки графика над  $x$  и  $y$ ; для строго выпуклой — *строго ниже* (иллюстрация приведена в лекции).

(л. 2, стр. 27)

## Определение 2.7

Функция  $f(x)$ , определённая на выпуклом множестве  $Q \subset \mathbb{R}^n$ , называется *вогнутой* на  $Q$ , если функция  $-f(x)$  является выпуклой на  $Q$ .

Из определения 2.6 непосредственно следует, что *линейная (аффинная)* функция вида

$$f(x) = a^\top x + b, \quad a \in \mathbb{R}^n, b \in \mathbb{R},$$

является одновременно и выпуклой, и вогнутой.

Это единственный класс функций, обладающих одновременно этими свойствами.

(л. 2, стр. 28–29)

## Подуровни выпуклой функции

Для выпуклой функции  $f(x)$  множество

$$Q_d = \{x \in \mathbb{R}^n \mid f(x) \leq d\} \tag{25}$$

является выпуклым для любого  $d \in \mathbb{R}$ .

### Доказательство.

Пусть  $x, y \in Q_d$ , то есть  $f(x) \leq d, f(y) \leq d$ . По выпуклости функции:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \leq d,$$

следовательно,  $\lambda x + (1 - \lambda)y \in Q_d$ .  $\square$

(л. 2, стр. 29)

## Обратное утверждение (в общем случае неверно)

Из выпуклости множества

$$Q_d = \{x \mid f(x) \leq d\}$$

**не следует**, что функция  $f(x)$  является выпуклой.

## Пример

Рассмотрим функцию

$$f(x) = \sqrt{\|x\|}.$$

Она **не является выпуклой**, однако множество

$$Q = \{x \in \mathbb{R}^n \mid f(x) \leq 2\} = \{x \in \mathbb{R}^n \mid \|x\| \leq 4\}$$

является выпуклым (шар в  $\mathbb{R}^n$ ).

Такие функции называются *квазивыпуклыми*.

(л. 2, стр. 30)

### Замечание о выпуклых множествах

Если множество  $Q$  является невыпуклым, то его можно *выпуклить*, то есть рассмотреть наименьшее выпуклое множество, содержащее  $Q$ .

### Определение 2.8

Выпуклой оболочкой множества  $Q \subset \mathbb{R}^n$  называется наименьшее выпуклое множество, содержащее  $Q$ , и обозначается

$$\text{conv } Q.$$

Эквивалентно,

$$\text{conv } Q = \left\{ \sum_{i=1}^m \lambda_i x_i \mid x_i \in Q, \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1, m \in \mathbb{N} \right\}.$$

(л. 2, стр. 30)

### Итог раздела «Элементы выпуклого анализа»

1. Выпуклость — геометрическое свойство множества, сохраняющееся при взятии выпуклых комбинаций.
2. Выпуклые функции обобщают линейные и играют ключевую роль в задачах оптимизации.
3. Подуровни выпуклых функций всегда выпуклы, но обратное утверждение неверно.
4. Выпуклая оболочка — фундаментальное понятие, позволяющее работать с произвольными множествами через их «выпуклое замыкание».

(л. 2, стр. 27–30)

**Задача (л. 2, стр. 25).** Показать, что

$$\nabla \|Ax - b\| = 2A^\top(Ax - b), \quad [A] = m \times n, \quad x \in \mathbb{R}^n, \quad b \in \mathbb{R}^m.$$

**Замечание об обозначениях.** Правая часть  $2A^\top(Ax - b)$  соответствует функции

$$F(x) = \|Ax - b\|^2 = \langle Ax - b, Ax - b \rangle,$$

а не самой норме  $\|Ax - b\|$  (для  $\|Ax - b\|$  градиент равен  $A^\top(Ax - b)/\|Ax - b\|$  при  $Ax \neq b$ ).  
Далее решаем в корректной форме: для  $F(x) = \|Ax - b\|^2$ .

Лекция 2. Сведения из математики...

## Решение

Обозначим

$$g(x) = Ax - b \in \mathbb{R}^m.$$

Тогда  $g$  — вектор-функция, и её производная (матрица Якоби) постоянна:

$$g'(x) = A.$$

Рассмотрим скалярную функцию

$$F(x) = \|g(x)\|^2 = \langle g(x), g(x) \rangle.$$

Возьмём приращение при  $x \mapsto x + y$ :

$$\begin{aligned} F(x + y) - F(x) &= \langle g(x + y), g(x + y) \rangle - \langle g(x), g(x) \rangle \\ &= \langle g(x) + Ay, g(x) + Ay \rangle - \langle g(x), g(x) \rangle \\ &= 2\langle g(x), Ay \rangle + \langle Ay, Ay \rangle. \end{aligned}$$

Последний член  $\langle Ay, Ay \rangle = O(\|y\|^2) = o(\|y\|)$ , поэтому линейная часть приращения равна

$$2\langle g(x), Ay \rangle = 2\langle A^\top g(x), y \rangle.$$

Сравнивая с формой дифференцируемости (2):  $F(x + y) = F(x) + \langle \nabla F(x), y \rangle + o(y)$ , получаем

$$\nabla F(x) = 2A^\top g(x) = 2A^\top(Ax - b).$$

**Ответ:**  $\boxed{\nabla \|Ax - b\|^2 = 2A^\top(Ax - b)}$  (л. 2, стр. 25).

Источники

(продолжаем без нового главного заголовка)

## Пример квазивыпуклой функции

Рассмотрим функцию

$$f(x) = \sqrt{\|x\|}.$$

Она **не является выпуклой** (и даже не является строго выпуклой), однако её подуровни

$$Q = \{x \in \mathbb{R}^n \mid f(x) \leq 2\} = \{x \in \mathbb{R}^n \mid \|x\| \leq 4\}$$

являются выпуклыми множествами (шарами).

Такие функции называются **квазивыпуклыми**: у них все множества подуровней выпуклы, но сама функция может не быть выпуклой.

Если множество  $Q$  является невыпуклым, то его можно «выпуклить», то есть заменить на выпуклое множество, связанное с ним.

(л. 2, стр. 30)

## Определение 2.8. Выпуклая оболочка

Выпуклой оболочкой множества  $Q \subset \mathbb{R}^n$  называется наименьшее выпуклое множество, содержащее  $Q$ . Обозначается:

$$\text{conv } Q.$$

По определению:

$$Q \subset \text{conv } Q.$$

Иными словами,  $\text{conv } Q$  — это пересечение всех выпуклых множеств, содержащих  $Q$ .

Множество  $\text{conv } Q$  существует и непусто для любого  $Q \neq \emptyset$ .

Пример:

- выпуклой оболочкой сферы является шар;
- выпуклой оболочкой двух точек — отрезок, соединяющий их.

(л. 2, стр. 30–31)

## Явное описание выпуклой оболочки

Выпуклую оболочку можно задать в виде:

$$\text{conv } Q = \left\{ x \mid x = \sum_{i=1}^m \lambda_i x_i, x_i \in Q, \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1 \right\}. \quad (26)$$

То есть  $\text{conv } Q$  состоит из всех выпуклых комбинаций конечного числа точек из  $Q$ .

(л. 2, стр. 31)

## Лемма 2.1 (Каратеодори)

Для любого множества  $Q \subset \mathbb{R}^n$  в формуле (26) достаточно брать

$$m \leq n + 1. \quad (27)$$

Иначе говоря, **любая точка выпуклой оболочки в  $\mathbb{R}^n$**  представима как выпуклая комбинация не более чем  $n + 1$  точек из  $Q$ .

(л. 2, стр. 32)

## Лемма 2.2

Если множество  $Q \subset \mathbb{R}^n$  замкнуто и ограничено, то его выпуклая оболочка  $\text{conv } Q$  также является замкнутым и ограниченным множеством.

(л. 2, стр. 32)

## Проекция на множество

В задачах оптимизации важную роль играет операция **проецирования**.

### Определение 2.9

Проекцией точки  $x \in \mathbb{R}^n$  на множество  $Q \subset \mathbb{R}^n$  называется точка

$$P_Q(x) = \arg \min_{y \in Q} \|x - y\|. \quad (28)$$

Из определения следует, что если  $x \in Q$ , то

$$P_Q(x) = x.$$

По теореме Вейерштрасса, если  $Q$  замкнуто и ограничено, то проекция существует.

Если множество  $Q$  **выпуклое**, то проекция **единственна** (доказывается отдельно).

(л. 2, стр. 33–34)

## Свойства проекции на выпуклое замкнутое множество

Пусть  $Q \subset \mathbb{R}^n$  — выпуклое и замкнутое множество. Тогда для проекции  $P_Q(x)$  выполняются следующие свойства:

### 1. Характеризационное неравенство

$$\langle x - P_Q(x), y - P_Q(x) \rangle \leq 0 \quad \forall y \in Q. \quad (29)$$

### 2. Неувеличение расстояния (неexpansive-свойство)

$$\|P_Q(x) - P_Q(y)\| \leq \|x - y\| \quad \forall x, y \in \mathbb{R}^n. \quad (30)$$

(л. 2, стр. 34)

## Теоремы отделимости. Теорема Хана–Банаха

### Определение 2.10

Два множества  $Q_1, Q_2 \subset \mathbb{R}^n$  называются **отделимыми**, если существует гиперплоскость, разделяющая их, то есть существуют  $a \in \mathbb{R}^n$ ,  $a \neq 0$ , и число  $d \in \mathbb{R}$  такие, что

$$\langle a, x \rangle \geq d \quad \forall x \in Q_1,$$

$$\langle a, x \rangle \leq d \quad \forall x \in Q_2.$$

Множества называются **строго отделимыми**, если существуют  $d_1 > d_2$  такие, что

$$\langle a, x \rangle \geq d_1 \quad \forall x \in Q_1,$$

$$\langle a, x \rangle \leq d_2 \quad \forall x \in Q_2.$$

(л. 2, стр. 35)

### Теорема 1 (об отделимости)

Пусть  $Q_1, Q_2 \subset \mathbb{R}^n$  — выпуклые, замкнутые, непересекающиеся множества, причём  $Q_2$  ограничено. Тогда множества  $Q_1$  и  $Q_2$  **строго отделимы**.

Геометрически это означает существование зазора между множествами, который можно отделить гиперплоскостью.

(л. 2, стр. 35–36)

### Определение 2.11. Опорная гиперплоскость

Гиперплоскость

$$L = \{x \in \mathbb{R}^n \mid \langle a, x \rangle = d\}$$

называется **опорной** к множеству  $Q$  в точке  $x_0$ , если:

- $x_0 \in L$ ,
- всё множество  $Q$  лежит в одном полупространстве:

$$\langle a, x \rangle \leq d \quad \forall x \in Q.$$

(л. 2, стр. 36)

### Теорема 2 (о существовании опорной гиперплоскости)

Пусть  $Q \subset \mathbb{R}^n$  — выпуклое множество, а  $x_0$  — граничная точка множества  $Q$ . Тогда существует гиперплоскость, **опорная** к  $Q$  в точке  $x_0$ .

Эта теорема является частным случаем теорем Хана–Банаха и лежит в основе многих методов выпуклой оптимизации и вариационного анализа.

(л. 2, стр. 36)

# Лекция 3. Теория квадратичных форм и основы оптимизации

## 3.1. Некоторые сведения из теории квадратичных форм

### Определение 3.1.

Функция  $f$  от  $n$  вещественных переменных

$$f(x_1, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$$

называется **квадратичной формой**, если она имеет вид

$$f(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j, \quad (3.1)$$

где  $a_{ij} \in \mathbb{R}$ , а  $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ .

Эквивалентно, квадратичную форму можно записать в матричном виде:

$$f(x) = x^T A x,$$

где  $A = (a_{ij})_{i,j=1}^n$  — квадратная матрица коэффициентов (л. 3, стр. 1).

Действительно,

$$x^T A x = (x_1, \dots, x_n) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

(л. 3, стр. 1-2).

**Замечание.** Матрицу  $A$ , задающую квадратичную форму, всегда можно считать **симметрической**.

Если исходная матрица  $A$  несимметрична, то квадратичная форма не изменится при замене  $A$  на

$$\tilde{A} = \frac{1}{2}(A + A^T), \quad (3.2)$$

так как

$$x^T A x = x^T \tilde{A} x$$

для всех  $x \in \mathbb{R}^n$  (л. 3, стр. 2).

### Определение 3.2.

1. Квадратичная форма называется **знакоположительной**, если

$$x^T A x \geq 0 \quad (3.3)$$



для всех  $x \in \mathbb{R}^n$ .

2. Квадратичная форма называется **положительно определённой**, если

$$x^T A x > 0 \quad (3.4)$$

для всех  $x \in \mathbb{R}^n$ ,  $x \neq 0$ .

Аналогично определяются **знакотрицательные** и **отрицательно определённые** квадратичные формы (л. 3, стр. 3).

**Замечание.** Для удобства далее часто говорят о знакоопределённости матрицы  $A$ , имея в виду соответствующую квадратичную форму  $x^T A x$ .

## Главные миноры и критерии знакоопределённости

Рассмотрим квадратную матрицу  $A = (a_{ij})_{i,j=1}^n$ .

### Определение.

Минором порядка  $p$  называется определитель подматрицы, полученной выбором строк  $i_1, \dots, i_p$  и столбцов  $j_1, \dots, j_p$ :

$$D \begin{pmatrix} i_1, \dots, i_p \\ j_1, \dots, j_p \end{pmatrix} = \det \begin{pmatrix} a_{i_1 j_1} & \dots & a_{i_1 j_p} \\ \vdots & \ddots & \vdots \\ a_{i_p j_1} & \dots & a_{i_p j_p} \end{pmatrix}. \quad (3.5)$$

Минор называется **главным**, если он составлен из одинаковых наборов строк и столбцов, то есть  $i_k = j_k$  (л. 3, стр. 5).

### Пример.

Для матрицы

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

главные миноры:

- порядка 1: 1, 5, 9;
- порядка 2:  $\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}, \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix}, \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix};$
- порядка 3:  $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$   
(л. 3, стр. 6).

Общее число главных миноров квадратной матрицы размера  $n \times n$  равно  $2^n - 1$  (л. 3, стр. 7).

Часто используют **ведущие главные миноры**:

$$D_1 = a_{11}, \quad D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \dots, \quad D_n = \det A. \quad (3.6)$$

## Критерии Сильвестра

Все дальнейшие критерии применимы **только к симметрическим матрицам**. Если матрица несимметрична, её следует предварительно заменить на  $\tilde{A} = \frac{1}{2}(A + A^T)$  (л. 3, стр. 8).

**Теорема (критерий Сильвестра).**

1. Матрица  $A$  (и соответствующая квадратичная форма) **положительно определена** тогда и только тогда, когда все её ведущие главные миноры положительны:

$$D_1 > 0, \quad D_2 > 0, \quad \dots, \quad D_n > 0. \quad (3.7)$$

2. Матрица  $A$  **знакоположительна** (неотрицательно определена), если и только если **все её главные миноры неотрицательны**:

$$D \begin{pmatrix} i_1, \dots, i_p \\ i_1, \dots, i_p \end{pmatrix} \geq 0, \quad 1 \leq p \leq n. \quad (3.8)$$

(л. 3, стр. 8–9)

**Замечание.** Условие неотрицательности всех главных миноров является **необходимым**, но не достаточным для строгой положительной определённости.

**Пример.**

Пусть

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad a_{11} = 0, \quad a_{12} = a_{21} = 0, \quad a_{22} < 0.$$

Тогда

$$D_1 = 0, \quad D_2 = 0,$$

однако

$$f(x) = x^T A x = a_{22} x_2^2 < 0$$

при  $x_2 \neq 0$ . Следовательно, форма не является знакоположительной, несмотря на неотрицательность ведущих миноров (л. 3, стр. 9–10).

## Проверка отрицательной определённости

Для установления **отрицательной определённости** матрицы  $A$  можно:

1. Умножить матрицу на  $-1$  и проверить положительную определённость матрицы  $-A$  по критериям выше.
2. Использовать специальные критерии:
  - критерий отрицательной определённости:

$$(-1)^p D_p > 0, \quad p = 1, \dots, n;$$

- критерий неотрицательной определённости:

$$(-1)^p D \begin{pmatrix} i_1, \dots, i_p \\ i_1, \dots, i_p \end{pmatrix} \geq 0.$$

(л. 3, стр. 11)

На этом месте изложение первой половины лекции 3 обрывается; далее темы будут продолжены в следующей части.

## 3.2. Локальные методы безусловной оптимизации

Рассматривается простейшая задача **безусловной оптимизации**:

$$f(x) \rightarrow \min, \quad x \in \mathbb{R}^n,$$

где целевая функция  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  предполагается дифференцируемой (л. 3, стр. 13).

Основная цель — поиск **локального минимума** функции. При решении таких задач необходимо опираться на итерационные принципы, гарантирующие сходимость соответствующих вычислительных процедур. Существенную роль играет:

- класс целевой функции  $f$ ;
- допустимое множество (в безусловной задаче  $X = \mathbb{R}^n$ );
- свойства функции (выпуклость, гладкость и т.п.).

Выделяют методы поиска **локальных** и **глобальных** решений (л. 3, стр. 13–14).

Любой численный метод оптимизации основан на вычислении значений целевой функции и/или её производных, а также на анализе этих величин и их приращений.

### Определение 4.1.

Метод оптимизации называется **пассивным**, если точки, в которых производятся вычисления, задаются заранее и не зависят от результатов предыдущих вычислений (л. 3, стр. 14).

### Определение 4.2.

Метод оптимизации называется **интерактивным (адаптивным)**, если точки, в которых вычисляются значения функции и её производных, выбираются в процессе вычислений на основе уже полученной информации (л. 3, стр. 14–15).

**Замечание.** Подавляющее большинство методов оптимизации являются итерационными.

## Итерационные методы

Итерационный метод порождает последовательность точек

$$x_0, x_1, x_2, \dots, x_k, x_{k+1}, \dots,$$

которая (в идеальном случае) сходится к решению задачи (л. 3, стр. 15).

### Определение 4.3.

Точка  $x_k$  называется  $k$ -м приближением к решению, а вся последовательность  $\{x_k\}_{k=0}^{\infty}$  — **траекторией метода** (л. 3, стр. 15).

Последовательность  $\{x_k\}$  называется **релаксационной**, если

$$f(x_{k+1}) \leq f(x_k), \quad \forall k \geq 0. \quad (4.1)$$

Переход от  $x_k$  к  $x_{k+1}$  называется **шагом** или **итерацией метода**.

Часто метод задаётся в виде **итерационной схемы**:

$$x_{k+1} = \Phi(x_k), \quad k = 0, 1, 2, \dots, \quad (4.2)$$

где отображение  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  называется **алгоритмическим отображением** (л. 3, стр. 16).

Схема (4.2) является **одномерной**, так как для вычисления  $x_{k+1}$  используется только предыдущее приближение  $x_k$ . В более общих методах алгоритмическое отображение может зависеть от нескольких предыдущих приближений:

$$x_{k+1} = \Phi(x_k, x_{k-1}), \quad (4.3)$$

при этом в начале процесса необходимо задать  $x_0$  и  $x_1$  (л. 3, стр. 17).

## Сходимость итерационных методов

### Определение 4.4.

Итерационный метод

$$x_{k+1} = \Phi(x_k)$$

называется **сходящимся** к точке  $x^*$  из начальной точки  $x_0$ , если

$$\lim_{k \rightarrow \infty} x_k = x^*.$$

(л. 3, стр. 18)

Возможна ситуация, когда вся последовательность  $\{x_k\}$  не сходится, но существует сходящаяся подпоследовательность. Тогда говорят, что метод **не сходится**, несмотря на наличие предельных точек.

В ряде случаев используется более слабое понятие сходимости.

### Определение 4.5.

Метод называется сходящимся к множеству  $X^*$ , если

$$\lim_{k \rightarrow \infty} \rho(x_k, X^*) = 0,$$

где  $\rho(x, X^*)$  — расстояние от точки  $x$  до множества  $X^*$  (л. 3, стр. 19).

Если  $X^* = \{x^*\}$ , то это определение эквивалентно обычной сходимости  $x_k \rightarrow x^*$ .

## Минимизирующие последовательности и сходимость по функционалу

#### Определение 4.6.

Последовательность  $\{x_k\}$  называется **минимизирующей**, если

$$\lim_{k \rightarrow \infty} f(x_k) = f^* = \inf_{x \in X} f(x). \quad (4.4)$$

(л. 3, стр. 19)

Если условие (4.4) выполнено, говорят, что метод **сходится по функционалу**.

**Замечание.** Сходимость по функционалу не гарантирует сходимости последовательности  $\{x_k\}$  по аргументу.

#### Пример.

Рассмотрим задачу

$$f(x) \rightarrow \inf, \quad x \in X = \mathbb{R}, x \geq 0,$$

где

$$f(x) = \frac{x}{1+x^2}.$$

В этой задаче  $f^* = 0$ ,  $X^* = \{0\}$ . Последовательность

$$x_k = k, \quad k \geq 0,$$

является минимизирующей, так как

$$\lim_{k \rightarrow \infty} f(x_k) = 0,$$

но при этом  $x_k$  не сходится к решению  $x^* = 0$  (л. 3, стр. 20).

### Скорость сходимости

Пусть последовательность  $\{x_k\}$  сходится к точке  $x^*$ , которая является решением задачи оптимизации:

$$x^* = \arg \min_{x \in X} f(x), \quad (4.5)$$

где  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ .

#### Определение 4.7.

Метод

$$x_{k+1} = \Phi(x_k)$$

сходится к  $x^*$  **с линейной скоростью**, если существует константа  $C > 0$ , число  $q \in (0, 1)$  и номер  $K \in \mathbb{N}$  такие, что

$$\|x_{k+1} - x^*\| \leq C \|x_k - x^*\|, \quad (4.6)$$

для всех  $k \geq K$  (л. 3, стр. 21).

Метод сходится **со сверхлинейной скоростью**, если

$$\|x_{k+1} - x^*\| \leq C_k \|x_k - x^*\|,$$

где  $C_k > 0$  и

$$\lim_{k \rightarrow \infty} C_k = 0$$

(л. 3, стр. 22).

Метод сходится **с квадратичной скоростью**, если существуют константа  $C > 0$  и номер  $K \in \mathbb{N}$  такие, что

$$\|x_{k+1} - x^*\| \leq C \|x_k - x^*\|^2, \quad (4.7)$$

для всех  $k \geq K$ .

Эквивалентная форма линейной сходимости: существует  $C > 0$ ,  $q \in (0, 1)$  и  $K \in \mathbb{N}$ , что

$$\|x_k - x^*\| \leq C q^k, \quad k \geq K. \quad (4.8)$$

Тогда говорят, что метод сходится к  $x^*$  **геометрически** (л. 3, стр. 22).

## 3.2. Локальные методы безусловной оптимизации (продолжение)

Ранее было введено понятие линейной (геометрической) скорости сходимости последовательности  $\{x_k\}$  к точке  $x^*$ :

$$\|x_k - x^*\| \leq C q^k, \quad 0 < q < 1.$$

Аналогично вводятся определения скорости сходимости **по целевой функции**.

**Замечание.** Аналогично определениям скорости сходимости по аргументу, вводятся понятия скорости сходимости и по значению целевой функции. В этом случае исследуется последовательность

$$f(x_k) \rightarrow f^*,$$

где  $f^* = f(x^*)$  (л. 3, стр. 23).

Например, линейная скорость сходимости по целевой функции означает существование констант  $C > 0$ ,  $q \in (0, 1)$  и номера  $K \in \mathbb{N}$  таких, что

$$|f(x_k) - f^*| \leq C q^k, \quad (4.9)$$

для всех  $k \geq K$ .

---

## Правила останова итерационного процесса

При практической реализации итерационных методов возникает вопрос: **когда следует прекращать вычисления**.

**Правила останова итерационного процесса** должны:

1. носить эвристический характер;
2. быть применимыми как к задачам без ограничений, так и с ограничениями;
3. учитывать класс целевой функции (линейность, квадратичность и т.п.);
4. учитывать, какие параметры вычисляются в ходе итерационного процесса (л. 3, стр. 23–24).

Пусть  $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0$  — заданные параметры точности решения задачи оптимизации.

Итерационный процесс обычно прекращают, если выполнено одно из условий:

$$\|x_{k+1} - x_k\| \leq \varepsilon_1, \quad |f(x_{k+1}) - f(x_k)| \leq \varepsilon_2. \quad (4.10)$$

Если целевая функция гладкая, то дополнительно проверяют условие

$$\|\nabla f(x_k)\| \leq \varepsilon_3, \quad (4.11)$$

которое указывает на приближённое выполнение необходимых условий оптимальности (л. 3, стр. 24–25).

**Замечание.** Условия (4.10) основаны на использовании абсолютных приращений аргумента и функции. Более устойчивыми на практике являются условия, основанные на **относительных приращениях**:

$$\begin{aligned} \|x_{k+1} - x_k\| &\leq \varepsilon_1 (1 + \|x_{k+1}\|), \\ |f(x_{k+1}) - f(x_k)| &\leq \varepsilon_2 (1 + |f(x_{k+1})|) \end{aligned} \quad (4.12)$$

(л. 3, стр. 25).

Условие (4.11) тесно связано с необходимыми условиями оптимальности в задачах оптимизации без ограничений.

## Проблемы существования решения

При решении любой оптимизационной задачи возникают три основные проблемы.

**Первая проблема** связана с существованием решения задачи оптимизации.

В ряде случаев вопрос о существовании минимума (локального или глобального) решается с помощью **теоремы Вейерштрасса**.

**Теорема (Вейерштрасса).**

Если функция  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$  непрерывна на компактном множестве  $X$ , то она достигает на  $X$  своего минимума и максимума (л. 3, стр. 26).

В этой теореме условие компактности множества  $X$  является ключевым, однако во многих задачах оптимизации допустимое множество не является компактным (например,  $X = \mathbb{R}^n$ ,  $X = \mathbb{R}_+^n$ ).

Если ослабить ограничение на множество  $X$  и наложить дополнительные условия на функцию  $f$ , то можно получить аналоги теоремы Вейерштрасса.

**Следствие 1.**

Если функция  $f(x)$  непрерывна на непустом и замкнутом множестве  $X$  и существует точка  $z \in X$  такая, что множество

$$\{x \in X \mid f(x) \leq f(z)\}$$

ограничено, то функция  $f$  достигает на  $X$  своей точки минимума (л. 3, стр. 26–27).

### Следствие 2.

Если функция  $f(x)$  непрерывна на непустом и замкнутом множестве  $X$  и для любой последовательности  $\{x_k\} \subset X$ , для которой

$$\|x_k\| \rightarrow \infty,$$

выполняется

$$f(x_k) \rightarrow +\infty,$$

то функция  $f$  достигает на  $X$  своей точки минимума (л. 3, стр. 27).

## Условия оптимальности. Направления убывания

**Вторая проблема** заключается в поиске условий, которым должно удовлетворять оптимальное решение задачи.

Рассмотрим множество уровней функции:

$$L_f(d) = \{x \in \mathbb{R}^n \mid f(x) \leq d\}.$$

Пусть  $\tilde{x}$  — фиксированная точка,  $d = f(\tilde{x})$ . Введём множество **направлений, касательных к множеству уровней** в точке  $\tilde{x}$ :

$$S_f(\tilde{x}) = \left\{ s \in \mathbb{R}^n \mid s = \lim_{k \rightarrow \infty} \frac{y_k - \tilde{x}}{\|y_k - \tilde{x}\|}, y_k \rightarrow \tilde{x}, y_k \in L_f(d) \right\}. \quad (4.13)$$

(л. 3, стр. 28)

Пусть  $s \in S_f(\tilde{x})$ ,  $\|s\| = 1$ . Рассмотрим локальное поведение функции в направлении  $s$  и введём величину

$$\Delta(s) = \lim_{d \rightarrow 0} \frac{f(\tilde{x} + ds) - f(\tilde{x})}{d}. \quad (4.14)$$

Используя линейную аппроксимацию функции,

$$f(\tilde{x} + ds) - f(\tilde{x}) = \langle \nabla f(\tilde{x}), ds \rangle + o(d),$$

получаем

$$\Delta(s) = \langle \nabla f(\tilde{x}), s \rangle. \quad (4.15)$$

(л. 3, стр. 29)

По неравенству Коши–Буняковского:

$$-\|\nabla f(\tilde{x})\| \|s\| \leq \langle \nabla f(\tilde{x}), s \rangle \leq \|\nabla f(\tilde{x})\| \|s\|.$$

Так как  $\|s\| = 1$ , имеем



$$\Delta(s) \geq -\|\nabla f(\tilde{x})\|.$$

Минимальное значение  $\Delta(s)$  достигается при

$$s = -\frac{\nabla f(\tilde{x})}{\|\nabla f(\tilde{x})\|}.$$

Следовательно, **антиградиент** является направлением наискорейшего убывания функции в точке  $\tilde{x}$  (л. 3, стр. 29–30).

## Условия оптимальности первого порядка

### Теорема 4.1 (необходимое условие оптимальности первого порядка).

Пусть  $x^* \in \mathbb{R}^n$  — точка локального минимума дифференцируемой функции

$$f : \mathbb{R}^n \rightarrow \mathbb{R}.$$

Тогда

$$\nabla f(x^*) = 0.$$

(л. 3, стр. 30)

Доказательство основано на разложении функции в окрестности точки  $x^*$  и анализе знака первого дифференциала.

#### Следствие 4.1.1.

Пусть  $x^*$  — точка локального минимума функции  $f(x)$  при линейных ограничениях

$$x \in X = \{x \in \mathbb{R}^n \mid Ax = b\},$$

где  $A$  — матрица размера  $m \times n$ ,  $b \in \mathbb{R}^m$ . Тогда существует вектор множителей  $\lambda^* \in \mathbb{R}^m$  такой, что

$$\nabla f(x^*) = A^T \lambda^*. \quad (4.17)$$

(л. 3, стр. 31)

#### Замечание.

1. Теорема 4.1 и следствие 4.1.1 дают **необходимые**, но не достаточные условия локального минимума.
2. Все точки  $x^* \in \mathbb{R}^n$ , удовлетворяющие условию

$$\nabla f(x^*) = 0,$$

называются **стационарными (критическими) точками** функции  $f$  (л. 3, стр. 31–32).

#### Пример.

Функция  $f(x) = x^3$  имеет критическую точку  $x^* = 0$ , так как  $f'(x) = 3x^2$ , но точка  $x^* = 0$  не является точкой минимума (л. 3, стр. 32).

## Условия оптимальности второго порядка

### Теорема 4.2 (необходимые условия оптимальности второго порядка).

Пусть  $x^* \in \mathbb{R}^n$  — точка локального минимума дважды дифференцируемой функции  $f$ . Тогда

$$\nabla f(x^*) = 0, \quad \nabla^2 f(x^*) \succeq 0, \quad (4.18)$$

то есть матрица Гессе  $\nabla^2 f(x^*)$  является неотрицательно определённой (л. 3, стр. 32–33).

Доказательство основано на разложении Тейлора второго порядка:

$$f(x^* + th) = f(x^*) + t\langle \nabla f(x^*), h \rangle + \frac{t^2}{2}\langle \nabla^2 f(x^*)h, h \rangle + o(t^2\|h\|^2),$$

и анализе знака квадратичной формы при  $t \rightarrow 0$ .

## Достаточные условия локального минимума

### Теорема 4.3 (достаточные условия оптимальности).

Пусть функция  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  дважды дифференцируема в окрестности точки  $x^*$  и выполнены условия:

$$\nabla f(x^*) = 0, \quad \nabla^2 f(x^*) \succ 0,$$

то есть матрица Гессе положительно определена. Тогда  $x^*$  является точкой локального минимума функции  $f$  (л. 3, стр. 33).

Таким образом, для многомерной задачи оптимизации без ограничений проверка оптимальности сводится к:

1. проверке условия  $\nabla f(x^*) = 0$ ;
2. исследованию знакоопределённости матрицы Гессе  $\nabla^2 f(x^*)$  с использованием критериев Сильвестра.

## 3.2. Локальные методы безусловной оптимизации (продолжение)

### Теорема 4.1 (необходимое условие оптимальности первого порядка).

Пусть  $x^* \in \mathbb{R}^n$  — точка локального минимума дифференцируемой функции  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Тогда

$$\nabla f(x^*) = 0.$$

(л. 3, стр. 33)

### Доказательство.

Так как  $x^*$  — точка локального минимума, существует  $\varepsilon > 0$  такое, что для всех  $y \in \mathbb{R}^n$ ,  $\|y - x^*\| \leq \varepsilon$ , выполняется

$$f(y) \geq f(x^*).$$

Поскольку функция  $f$  дифференцируема, имеем разложение:

$$f(y) = f(x^*) + \langle \nabla f(x^*), y - x^* \rangle + o(\|y - x^*\|).$$

Следовательно,

$$\langle \nabla f(x^*), y - x^* \rangle + o(\|y - x^*\|) \geq 0.$$

Положим  $y - x^* = \beta s$ , где  $\|s\| = 1$ . Делим неравенство на  $\beta > 0$  и переходим к пределу при  $\beta \rightarrow 0$ :

$$\langle \nabla f(x^*), s \rangle \geq 0$$

для всех единичных векторов  $s$ . Выбирая

$$s = -\frac{\nabla f(x^*)}{\|\nabla f(x^*)\|},$$

получаем

$$-\|\nabla f(x^*)\| \geq 0,$$

откуда следует  $\nabla f(x^*) = 0$ .  $\square$

(л. 3, стр. 33–34)

### Следствие 4.1.1.

Пусть  $x^*$  — точка локального минимума функции  $f(x)$  при линейных ограничениях

$$x \in X = \{x \in \mathbb{R}^n \mid Ax = b\},$$

где  $A$  — матрица размера  $m \times n$ ,  $b \in \mathbb{R}^m$ . Тогда существует вектор  $\lambda^* \in \mathbb{R}^m$  такой, что

$$\nabla f(x^*) = A^T \lambda^*. \quad (4.17)$$

(л. 3, стр. 34)

### Замечания.

1. Теорема 4.1 и следствие 4.1.1 дают только **необходимые** условия локального минимума.
2. Все точки  $x^* \in \mathbb{R}^n$ , удовлетворяющие условию

$$\nabla f(x^*) = 0,$$

называются **стационарными (критическими) точками** функции  $f$ .

3. Не каждая стационарная точка является точкой локального минимума (л. 3, стр. 34–35).

### Пример.

Функция  $f(x) = x^3$ ,  $x \in \mathbb{R}$ , имеет стационарную точку  $x^* = 0$ , поскольку

$$f'(x) = 3x^2, \quad f'(0) = 0,$$

однако  $x^* = 0$  не является точкой минимума (л. 3, стр. 35).

## Условия оптимальности второго порядка

### Теорема 4.2 (необходимые условия оптимальности второго порядка).

Пусть  $x^*$  — точка локального минимума дважды дифференцируемой функции  $f$ . Тогда

$$\nabla f(x^*) = 0, \quad \nabla^2 f(x^*) \succeq 0, \quad (4.18)$$

то есть матрица Гессе  $\nabla^2 f(x^*)$  является неотрицательно определённой (л. 3, стр. 35).

#### Доказательство.

Из теоремы 4.1 следует, что  $\nabla f(x^*) = 0$ . Рассмотрим разложение Тейлора второго порядка:

$$f(x^* + th) = f(x^*) + t\langle \nabla f(x^*), h \rangle + \frac{t^2}{2}\langle \nabla^2 f(x^*)h, h \rangle + o(t^2\|h\|^2).$$

Так как  $f(x^* + th) \geq f(x^*)$  при малых  $t$ , получаем:

$$\frac{t^2}{2}\langle \nabla^2 f(x^*)h, h \rangle + o(t^2\|h\|^2) \geq 0.$$

Делим на  $t^2$  и переходим к пределу при  $t \rightarrow 0$ :

$$\langle \nabla^2 f(x^*)h, h \rangle \geq 0$$

для всех  $h \in \mathbb{R}^n$ . Следовательно, квадратичная форма, задаваемая матрицей Гессе, неотрицательно определена.  $\square$

(л. 3, стр. 35–37)

## Достаточные условия локального минимума

### Теорема 4.3 (достаточные условия оптимальности).

Пусть функция  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  дважды дифференцируема в окрестности точки  $x^*$  и выполнены условия:

$$\nabla f(x^*) = 0, \quad \nabla^2 f(x^*) \succ 0,$$

то есть матрица Гессе положительно определена. Тогда  $x^*$  является точкой локального минимума функции  $f$ .

(л. 3, стр. 37)

#### Доказательство.

Предположим противное. Тогда существует последовательность  $\{x_k\}$ ,  $x_k \rightarrow x^*$ , такая что

$$f(x_k) \leq f(x^*).$$

Представим

$$x_k = x^* + \alpha_k h_k, \quad \alpha_k = \|x_k - x^*\|, \quad h_k = \frac{x_k - x^*}{\|x_k - x^*\|}, \quad \|h_k\| = 1.$$

Из компактности единичной сферы можно считать, что  $h_k \rightarrow h$ ,  $\|h\| = 1$ .

Используя разложение Тейлора второго порядка и условие  $\nabla f(x^*) = 0$ , получаем:

$$f(x_k) - f(x^*) = \frac{\alpha_k^2}{2} \langle \nabla^2 f(x^*) h_k, h_k \rangle + o(\alpha_k^2).$$

Из неравенства  $f(x_k) - f(x^*) \leq 0$  следует

$$\langle \nabla^2 f(x^*) h, h \rangle \leq 0,$$

что противоречит положительной определённости матрицы Гессе. Следовательно, точка  $x^*$  является точкой локального минимума.  $\square$

(л. 3, стр. 37–41)

Таким образом, для многомерной задачи безусловной оптимизации:

- проверка условия  $\nabla f(x^*) = 0$  даёт кандидатов на экстремум;
- проверка условия  $\nabla^2 f(x^*) \succeq 0$  — необходимое условие локального минимума;
- проверка условия  $\nabla^2 f(x^*) \succ 0$  — достаточное условие локального минимума;

а исследование знакоопределённости матрицы Гессе проводится с помощью критериев Сильвестра, изложенных ранее (л. 3, стр. 41–42).

# Лекция 4. Методы одномерной оптимизации

## Постановка задачи

Часто встречающаяся в приложениях задача — задача минимизации скалярной функции одного аргумента. Рассматривается функция

$f : X \rightarrow \mathbb{R}$ , где  $X \subset \mathbb{R}^1$ , обычно  $X = [a, b]$ ,  $a < b$ . Предполагается, что функция  $f(x)$  определена и непрерывна на множестве  $X$  (л. 4, стр. 1).

Задача одномерной оптимизации формулируется следующим образом:

$$f^* = \min_{x \in X} f(x), \quad x^* = \arg \min_{x \in X} f(x), \quad (4.1)$$

где  $f^* = f(x^*)$  — минимальное значение функции, а  $x^*$  — точка минимума (л. 4, стр. 1).

**Замечание.** Задача (4.1) часто используется как вспомогательная при численном решении более сложных оптимизационных задач. Поэтому построение и анализ эффективных методов её решения имеет самостоятельную ценность (л. 4, стр. 2).

## Унимодальные функции

Большинство алгоритмов, рассматриваемых далее, опираются на свойство унимодальности целевой функции.

**Определение 4.1.** Функция  $f(x)$  называется *унимодальной* на отрезке  $X = [a, b]$ , если существует точка  $x^* \in [a, b]$  такая, что:

- для любых  $a \leq x_1 < x_2 \leq x^*$  выполняется  $f(x_1) > f(x_2)$ ;
  - для любых  $x^* \leq x_1 < x_2 \leq b$  выполняется  $f(x_1) < f(x_2)$
- (л. 4, стр. 2–3).

Таким образом, унимодальная функция строго убывает слева от точки минимума и строго возрастает справа от неё. При этом унимодальная функция может быть как непрерывной, так и разрывной. На рисунке 4.1.а показан пример непрерывной унимодальной функции, на рисунке 4.1.б — разрывной (л. 4, стр. 3).

## Лемма о единственности минимума

**Лемма 4.1.** Если на отрезке  $X = [a, b] \subset \mathbb{R}^1$  унимодальная функция  $f(x)$  достигает своего минимума, то этот минимум достигается в единственной точке  $x^*$  (л. 4, стр. 4).

Из леммы следует, что для непрерывных целевых функций унимодальность эквивалентна наличию на отрезке  $X$  единственного локального минимума, который одновременно является и глобальным минимумом (л. 4, стр. 4).

## Свойства унимодальных функций

**Лемма 4.2.** Пусть функция  $f(x)$  унимодальна на отрезке  $[a, b]$  и достигает минимума в точке  $x^* \in (a, b)$ . Пусть выбраны точки  $x_1, x_2 \in [a, b]$ ,  $x_1 < x_2$ . Тогда:

1. если  $f(x_1) \leq f(x_2)$ , то  $x^* \in [a, x_2]$ ;
  2. если  $f(x_1) \geq f(x_2)$ , то  $x^* \in [x_1, b]$
- (л. 4, стр. 5).

*Идея доказательства.* Докажем, например, первое утверждение. Предположим противное:  $x^* > x_2$ . Тогда из унимодальности следует, что при  $x_1 < x_2 < x^*$  должно выполняться  $f(x_1) > f(x_2)$ , что противоречит условию  $f(x_1) \leq f(x_2)$ . Противоречие доказывает утверждение (л. 4, стр. 5–6).

Лемма 4.2 лежит в основе методов последовательной локализации решения задачи (4.1).

## Общая схема методов последовательной локализации

Для непрерывных унимодальных функций достаточно эффективными оказываются методы последовательной локализации решения.

**Шаг 1.** Строится последовательность вложенных отрезков локализации решения:

$$[a, b] \supset [a_0, b_0] \supset [a_1, b_1] \supset \dots \supset [a_k, b_k] \supset \dots$$

так, чтобы каждый отрезок  $[a_k, b_k]$  содержал точку минимума  $x^*$  (л. 4, стр. 6–7).

Обозначим длины отрезков:

$$\Delta_k = b_k - a_k, \quad k = 0, 1, 2, \dots$$

Обычно строится монотонно убывающая последовательность длин:

$$\Delta_0 > \Delta_1 > \dots > \Delta_k > \dots, \quad \lim_{k \rightarrow \infty} \Delta_k = 0$$

(л. 4, стр. 7).

**Шаг 2.** После построения очередного отрезка  $[a_k, b_k]$  в качестве приближённого решения задачи выбирается точка  $x_k \in [a_k, b_k]$ . Чаще всего берётся середина отрезка:

$$x_k = \frac{a_k + b_k}{2}. \quad (4.3)$$

Тогда автоматически выполняется оценка погрешности:

$$|x_k - x^*| \leq \frac{\Delta_k}{2}. \quad (4.4)$$

В общем случае для произвольной точки  $x_k \in [a_k, b_k]$  имеет место более грубая оценка:

$$|x_k - x^*| \leq \Delta_k. \quad (4.5)$$

(л. 4, стр. 7–8).

## Критерий остановки

**Шаг 3.** Итерационный процесс локализации продолжается до тех пор, пока не будет достигнута заданная точность  $\varepsilon$ .

Под точностью решения в данном классе методов понимается оценка расстояния между точкой минимума и найденным приближением:

$$|x_k - x^*|.$$

Процесс прекращается, когда выполнено условие:

$$|x_k - x^*| \leq \varepsilon. \quad (4.6)$$

(л. 4, стр. 8–9).

Если в качестве  $x_k$  берётся середина отрезка  $[a_k, b_k]$ , то достаточно выбрать минимальный номер итерации  $k_0$ , для которого

$$\Delta_{k_0} \leq 2\varepsilon. \quad (4.7)$$

(л. 4, стр. 9).

**Замечание.** Длина отрезка локализации решения в два раза превышает точность решения. Такая оценка справедлива независимо от того, с какой стороны от середины отрезка находится точка минимума  $x^*$  (л. 4, стр. 9–10).

## Классификация методов

Различные алгоритмы решения задачи (4.1) отличаются друг от друга конкретными способами перехода от отрезка  $[a_k, b_k]$  к следующему отрезку  $[a_{k+1}, b_{k+1}]$ . При этом во всех случаях используется унимодальность функции  $f(x)$  (л. 4, стр. 10).

## Метод равномерного поиска

Метод равномерного поиска относится к пассивным методам оптимизации.

На каждом шаге значения целевой функции вычисляются в точках, определяемых равномерным разбиением текущего отрезка  $[a, b]$  на  $N$  одинаковых промежутков. Из всех вычисленных значений выбирается наименьшее.

Пусть минимум достигается в точке  $x_j$ . Тогда из унимодальности функции следует, что отрезки

$$[a, x_{j-1}] \quad \text{и} \quad [x_{j+1}, b]$$

можно исключить из дальнейшего рассмотрения (л. 4, стр. 11).

В результате множество поиска сужается до интервала

$$(x_{j-1}, x_{j+1}),$$

который и используется на следующем шаге метода (л. 4, стр. 11).



Ограничение на выбор числа узлов разбиения:

$$N \geq 3,$$

так как при  $N \leq 2$  исключение промежутков невозможно (л. 4, стр. 12).

Алгоритм равномерного поиска минимума относится к классу *пассивных методов поиска*, поскольку схема вычисления значений целевой функции заранее фиксирована и не зависит от результатов предыдущих вычислений (л. 4, стр. 12).

### Шаг 1. Нулевой шаг.

Полагаем:

$$k = 0, \quad a_0 = a, \quad b_0 = b, \quad \Delta_0 = b_0 - a_0,$$

и в качестве текущего приближения берём середину отрезка:

$$\bar{x}_0 = \frac{a_0 + b_0}{2}.$$

(л. 4, стр. 12).

**Шаг 2.** Фиксируем число  $N$  и строим равномерную сетку с  $N + 1$  узлами на текущем отрезке  $[a_k, b_k]$ :

$$x_k^{(i)} = a_k + \frac{\Delta_k}{N} i, \quad i = 0, 1, \dots, N,$$

где  $\Delta_k = b_k - a_k$  (л. 4, стр. 12–13). Геометрическая схема построения сетки показана на рис. 4.2 (л. 4, стр. 13).

**Шаг 3.** Вычисляем значения целевой функции  $f(x)$  во всех узлах построенной сетки, кроме крайних точек:

$$f(x_k^{(1)}), f(x_k^{(2)}), \dots, f(x_k^{(N-1)}).$$

(л. 4, стр. 13).

**Шаг 4.** Из набора вычисленных значений выбираем минимальное, то есть находим индекс  $j$ , такой что

$$f(x_k^{(j)}) = \min_{i=1, \dots, N-1} f(x_k^{(i)}).$$

(л. 4, стр. 13).

**Шаг 5.** Обновляем отрезок локализации решения:

$$a_{k+1} = x_k^{(j-1)}, \quad b_{k+1} = x_k^{(j+1)},$$

и новую середину:

$$\bar{x}_{k+1} = \frac{a_{k+1} + b_{k+1}}{2}.$$

При этом длина нового отрезка равна

$$\Delta_{k+1} = b_{k+1} - a_{k+1} = \frac{2}{N} \Delta_k.$$

(л. 4, стр. 14).

**Шаг 6.** Если выполнены условия остановки:

$$\frac{\Delta_{k+1}}{2} \leq \varepsilon_x, \quad |f(\bar{x}_{k+1}) - f(\bar{x}_k)| \leq \varepsilon_f,$$

то полагаем  $x^* = \bar{x}_{k+1}$  и вычислительный процесс завершается. В противном случае полагаем  $k := k + 1$  и возвращаемся к шагу 2 (л. 4, стр. 14).

**Замечание.** В качестве приближения к точке минимума на  $k$ -й итерации принимается середина текущего отрезка  $[a_k, b_k]$ . Число итераций, необходимое для достижения точности  $\varepsilon_x$ , определяется из условия:

$$\left(\frac{2}{N}\right)^k \frac{b-a}{2} \leq \varepsilon_x. \quad (4.8)$$

Оно не зависит от вида целевой функции  $f(x)$  (л. 4, стр. 15).

**Задача.** Найти пару  $\{x^*, f(x^*)\}$  для функции

$$f(x) = \frac{\sqrt{x^2 + 3}}{5} + \frac{5-x}{5},$$

на отрезке  $X = [-3, 8]$ , используя метод равномерного поиска при параметрах  $N = 10$ ,  $\varepsilon_x = 0,05$ ,  $\varepsilon_f = 0,001$  (л. 4, стр. 15).

## Метод дихотомии (метод деления отрезка пополам)

Метод дихотомии является наиболее простым и интуитивно понятным методом одномерной оптимизации (л. 4, стр. 16).

**Идея метода.** Решение задачи оптимизации на каждом шаге локализуется внутри некоторого отрезка

$$[a_{k+1}, b_{k+1}] \subset [a_k, b_k],$$

длина которого уменьшается вдвое:

$$\Delta_{k+1} = \frac{\Delta_k}{2}, \quad \Delta_k = b_k - a_k.$$

Этот отрезок гарантированно содержит точку минимума  $x^*$  (л. 4, стр. 16).

**Начальная итерация.**

Полагаем:

$$k = 0, \quad a_0 = a, \quad b_0 = b,$$

и вычисляем:

$$c_0 = \frac{a_0 + b_0}{2}, \quad f(c_0).$$

(л. 4, стр. 16).

**Общая  $k$ -я итерация.**

**Шаг 1.** Берём точку

$$y_k = \frac{a_k + c_k}{2}$$

и вычисляем значение  $f(y_k)$ .

Рассматриваются два случая (л. 4, стр. 17):

а) если  $f(y_k) \leq f(c_k)$ , то полагаем

$$a_{k+1} = a_k, \quad b_{k+1} = c_k, \quad c_{k+1} = y_k,$$

и переходим к следующей итерации;

б) если  $f(y_k) > f(c_k)$ , то вводим новую точку

$$z_k = \frac{c_k + b_k}{2}$$

и вычисляем  $f(z_k)$ .

**Шаг 2.** Если  $f(c_k) \leq f(z_k)$ , то полагаем

$$a_{k+1} = y_k, \quad b_{k+1} = b_k, \quad c_{k+1} = c_k;$$

если же  $f(c_k) > f(z_k)$ , то

$$a_{k+1} = c_k, \quad b_{k+1} = b_k, \quad c_{k+1} = z_k.$$

(л. 4, стр. 17).

**Шаг 3.** Полагаем  $k := k + 1$  и переходим к шагу 1 (л. 4, стр. 18).

Геометрическая интерпретация возможных случаев сравнения значений функции приведена на рис. 4.3–4.5 (л. 4, стр. 18–19).

Из леммы 4.2 следует, что на каждом шаге точка минимума удовлетворяет включению

$$x^* \in [a_{k+1}, b_{k+1}],$$

а длина отрезка локализации равна

$$\Delta_{k+1} = b_{k+1} - a_{k+1} = \frac{1}{2^k}(b - a). \quad (4.9)$$

(л. 4, стр. 19).

В качестве очередного приближения к точке минимума  $x^*$  обычно выбирают середину текущего отрезка:

$$c_{k+1} = \frac{a_{k+1} + b_{k+1}}{2}.$$

(л. 4, стр. 20).

На каждой итерации алгоритма вычисляются не более двух значений целевой функции  $f(x)$ , а на начальной итерации — только одно (л. 4, стр. 20).

Метод дихотомии особенно эффективен в задачах, где вычисление значений функции является трудоёмким (л. 4, стр. 20).

## Оценка скорости сходимости метода дихотомии

Пусть допускается выполнение не более  $N$  вычислений значений функции  $f(x)$ . Тогда метод дихотомии можно выполнить как минимум

$$M = \frac{N - 1}{2}$$

итераций и получить приближение

$$x_{M+1} = c_{M+1}$$

(л. 4, стр. 21).

Для погрешности имеет место оценка:

$$|x_{M+1} - x^*| \leq \frac{1}{2}(b_{M+1} - a_{M+1}) = \frac{1}{2^{M+1}}(b - a) \approx 0,707^{N-1} \frac{b - a}{2}. \quad (4.10)$$

(л. 4, стр. 21–22).

Таким образом, погрешность убывает не медленнее, чем по геометрической прогрессии со знаменателем

$$q = 0,707.$$

Из оценки (4.10) легко получить число итераций  $M$ , необходимое для достижения заданной точности  $\varepsilon$ :

$$|x_{M+1} - x^*| \leq \frac{1}{2}(b_{M+1} - a_{M+1}) \leq \varepsilon.$$

Отсюда следует оценка:

$$M \geq \left\lceil \log_2 \frac{b - a}{\varepsilon} - 1 \right\rceil, \quad (4.11)$$

где  $\lceil \cdot \rceil$  — оператор округления до ближайшего большего целого числа (л. 4, стр. 22).

## Метод золотого сечения

**Определение 4.2.** *Золотым сечением* называется такое деление отрезка на две неравные части, при котором отношение длины всего отрезка к большей части равно отношению большей части к меньшей (л. 4, стр. 23).

Рассмотрим отрезок  $[a, b]$  и точку  $y$ , делящую его в пропорции золотого сечения и расположенную ближе к левому концу  $a$ . Тогда по определению выполняется соотношение:

$$\frac{b-a}{b-y} = \frac{b-y}{y-a}.$$

Отсюда получаем:

$$y(a, b) = a + \frac{2}{3 + \sqrt{5}}(b - a) \approx a + 0,382(b - a). \quad (4.12)$$

(л. 4, стр. 23–24).

Если точка  $z$  расположена ближе к правому концу отрезка  $b$ , то из определения золотого сечения имеем:

$$\frac{b-a}{z-a} = \frac{z-a}{b-z},$$

откуда следует:

$$z(a, b) = a + \frac{\sqrt{5}-1}{2}(b - a) \approx a + 0,618(b - a). \quad (4.13)$$

(л. 4, стр. 24).

## Малая и большая золотые точки

**Определение 4.3.** Точка  $y(a, b)$  называется *меньшей золотой точкой* отрезка  $[a, b]$ , а точка  $z(a, b)$  — *большой золотой точкой* этого отрезка (л. 4, стр. 25).

На рис. 4.6 показано взаимное расположение меньшей и большей золотых точек на отрезке  $[a, b]$  (л. 4, стр. 25).

Введём обозначение:

$$\tau = \frac{1 + \sqrt{5}}{2},$$

— число золотого сечения.

Тогда координаты золотых точек удобно записываются в виде:

$$y(a, b) = a + \frac{1}{\tau^2}(b - a), \quad z(a, b) = a + \frac{1}{\tau}(b - a). \quad (4.14)$$

(л. 4, стр. 25).

## Свойства золотых точек

**Лемма 4.3.** Пусть  $y$  и  $z$  — соответственно меньшая и большая золотые точки отрезка  $[a, b]$ . Тогда выполняются следующие равенства:

1.

$$z - a = b - y = \frac{\sqrt{5} - 1}{2}(b - a) = \frac{b - a}{\tau};$$

2. точка  $y$  является большей золотой точкой отрезка  $[a, z]$ , а точка  $z$  — меньшей золотой точкой отрезка  $[y, b]$ .  
(л. 4, стр. 26).

Эти свойства лежат в основе экономичности метода золотого сечения.

## Алгоритм метода золотого сечения

### Начальная итерация.

Полагаем:

$$k = 1, \quad a_1 = a, \quad b_1 = b.$$

По формулам (4.14) вычисляем:

$$y_1 = y(a_1, b_1), \quad z_1 = z(a_1, b_1),$$

и на первом шаге вычисляем одно значение функции:

$$f_1 = f(y_1).$$

(л. 4, стр. 26).

## Общая $k$ -я итерация

**Шаг 1.** Вычисляем то из значений  $f(y_k)$  или  $f(z_k)$ , которое ещё не было вычислено.

Если

$$f(y_k) \leq f(z_k),$$

то полагаем:

$$a_{k+1} = a_k, \quad b_{k+1} = z_k, \quad z_{k+1} = y_k, \quad y_{k+1} = y(a_{k+1}, b_{k+1}).$$

Если же

$$f(y_k) > f(z_k),$$

то полагаем:

$$a_{k+1} = y_k, \quad b_{k+1} = b_k, \quad y_{k+1} = z_k, \quad z_{k+1} = z(a_{k+1}, b_{k+1}).$$

(л. 4, стр. 27).

**Шаг 2.** Полагаем  $k := k + 1$  и переходим к шагу 1 (л. 4, стр. 27).

В соответствии с леммой 4.3 на каждой итерации выполняется:

$$y_{k+1} = y(a_{k+1}, b_{k+1}), \quad z_{k+1} = z(a_{k+1}, b_{k+1}),$$

а из унимодальности функции следует включение:

$$x^* \in [a_{k+1}, b_{k+1}].$$

(л. 4, стр. 27–28).

В качестве приближения к решению задачи можно брать любую точку текущего отрезка  $[a_{k+1}, b_{k+1}]$ .

## Оценка длины отрезка локализации

В методе золотого сечения на каждой итерации требуется вычисление значения целевой функции только в *одной* новой точке.

Длина отрезка локализации удовлетворяет соотношению:

$$\Delta_{k+1} = b_{k+1} - a_{k+1} = \frac{1}{\tau}(b_k - a_k),$$

и, следовательно,

$$\Delta_{k+1} = \left(\frac{1}{\tau}\right)^k (b - a). \quad (4.15)$$

(л. 4, стр. 28).

## Оценка точности метода золотого сечения

Если допускается вычисление значений функции в  $N$  точках, то можно выполнить

$$k = N - 1$$

итераций метода золотого сечения. Для погрешности приближения  $x_N$  к точке минимума  $x^*$  имеет место оценка:

$$|x_N - x^*| \leq \Delta_{k+1} = \left(\frac{1}{\tau}\right)^{N-1} (b - a) \approx 0,618^{N-1}(b - a). \quad (4.16)$$

(л. 4, стр. 29).

При достаточно больших  $N$  данная оценка существенно лучше, чем для метода дихотомии.

Для достижения заданной точности  $\varepsilon$  необходимо выполнить число итераций, удовлетворяющее условию:

$$k \geq \left\lceil \frac{\ln \frac{b-a}{\varepsilon}}{\ln \tau} \right\rceil + 1. \quad (4.17)$$

(л. 4, стр. 29–30).

## Метод Фибоначчи

Метод Фибоначчи близок по идее к методу золотого сечения, однако в нём вместо постоянного коэффициента золотого сечения используется последовательность чисел Фибоначчи (л. 4, стр. 30).

Последовательность чисел Фибоначчи определяется рекуррентным соотношением:

$$F_{i+1} = F_i + F_{i-1}, \quad i \geq 1, \quad (4.18)$$

при начальных условиях

$$F_0 = F_1 = 1.$$

Отсюда последовательно получаем:

$$F_2 = 2, F_3 = 3, F_4 = 5, F_5 = 8, \dots$$

(л. 4, стр. 30).

## Связь чисел Фибоначчи с золотым сечением

Рассмотрим рекуррентное уравнение (4.18) и будем искать его решение в классе геометрических прогрессий вида

$$F_i = t^i.$$

Подстановка в (4.18) даёт:

$$t^{i+1} = t^i + t^{i-1}, \quad t \neq 0,$$

или, после деления на  $t^{i-1}$ ,

$$t^2 - t - 1 = 0. \quad (4.20)$$

Корни этого квадратного уравнения равны:

$$t_1 = \frac{1 + \sqrt{5}}{2}, \quad t_2 = \frac{1 - \sqrt{5}}{2}.$$

Первый корень совпадает с числом золотого сечения:

$$\tau = \frac{1 + \sqrt{5}}{2}.$$

(л. 4, стр. 31).

Общее решение рекуррентного соотношения (4.18) представляется в виде линейной комбинации:

$$F_i = C_1 \tau^i + C_2 \left(-\frac{1}{\tau}\right)^i.$$

Коэффициенты  $C_1, C_2$  находятся из начальных условий  $F_0 = F_1 = 1$ , в результате чего получается *формула Бине*:



$$F_i = \frac{\tau^{i+1} - \left(-\frac{1}{\tau}\right)^{i+1}}{\sqrt{5}}.$$

При  $i \rightarrow \infty$  второе слагаемое стремится к нулю, и имеет место асимптотическая формула:

$$F_i \approx \frac{\tau^{i+1}}{\sqrt{5}}. \quad (4.21)$$

(л. 4, стр. 31–32).

## Алгоритм метода Фибоначчи

Для простоты дальнейшего изложения считаем, что исходный отрезок приведён к виду

$$[a, b] = [0, 1].$$

В общем случае переход осуществляется заменой переменной

$$x = a + y(b - a), \quad y \in [0, 1].$$

(л. 4, стр. 32).

Пусть задано целое число  $N > 2$ , равное максимально допустимому числу вычислений значений функции  $f(x)$ .

Полагаем:

$$a_1 = 0, \quad b_1 = 1, \quad \Delta_1 = b_1 - a_1 = 1.$$

Начальные точки выбираются по формулам:

$$y_1 = a_1 + \frac{F_{N-2}}{F_N} = b_1 - \frac{F_{N-1}}{F_N},$$

$$z_1 = a_1 + \frac{F_{N-1}}{F_N} = b_1 - \frac{F_{N-2}}{F_N},$$

причём точки  $y_1$  и  $z_1$  расположены симметрично относительно середины отрезка  $[a_1, b_1]$  (л. 4, стр. 32).

Вычисляем значения  $f(y_1)$  и  $f(z_1)$ .

## Итерационный процесс

Рассматриваются два случая.

**Случай 1.** Если

$$f(y_1) \leq f(z_1),$$

то полагаем:

$$a_2 = a_1, \quad b_2 = z_1, \quad z_2 = y_1,$$

и определяем новую точку:

$$y_2 = a_2 + \frac{F_{N-3}}{F_N} = b_2 - \frac{F_{N-2}}{F_N}.$$

**Случай 2.** Если

$$f(y_1) > f(z_1),$$

то полагаем:

$$a_2 = y_1, \quad b_2 = b_1, \quad y_2 = z_1,$$

и определяем новую точку:

$$z_2 = a_2 + \frac{F_{N-2}}{F_N} = b_2 - \frac{F_{N-3}}{F_N}.$$

(л. 4, стр. 33).

В обоих случаях длина нового отрезка локализации равна:

$$\Delta_2 = b_2 - a_2 = \frac{F_{N-1}}{F_N} \Delta_1 < \Delta_1.$$

Точки  $y_2$  и  $z_2$  вновь располагаются симметрично относительно середины отрезка  $[a_2, b_2]$  (л. 4, стр. 33–34).

## Общая схема

Аналогичным образом на  $k$ -м шаге получается отрезок локализации  $[a_k, b_k]$  длины

$$\Delta_k = \frac{F_{N-k+1}}{F_N} \Delta_1,$$

а всего можно выполнить  $N - 2$  итерации метода. На последнем шаге точки  $y_{N-1}$  и  $z_{N-1}$  совпадают и равны середине отрезка  $[a_{N-2}, b_{N-2}]$  (л. 4, стр. 34–35).

Эта точка и принимается в качестве приближённого решения задачи оптимизации:

$$x_{N-1}.$$

## Оценка точности метода Фибоначчи

Для погрешности имеет место оценка:

$$|x_{N-1} - x^*| \leq \frac{\Delta_{N-1}}{2} = \frac{b - a}{F_N}. \quad (4.22)$$

Однако, поскольку точка минимума может располагаться как слева, так и справа от найденного приближения, длина итогового отрезка локализации равна:

$$\Delta_{N-1} = \frac{2}{F_N} (b - a). \quad (4.23)$$

(л. 4, стр. 35–36).

Используя асимптотическую формулу (4.21), из (4.22) можно получить оценку на число допустимых вычислений функции, необходимых для достижения точности  $\varepsilon$ :

$$N \geq \left\lceil \frac{\ln \left( \sqrt{5} \frac{b-a}{2\varepsilon} \right)}{\ln \tau} \right\rceil - 1. \quad (4.24)$$

(л. 4, стр. 36).

## Сравнение с методом золотого сечения

Промежуток локализации в методе Фибоначчи примерно на 17% меньше, чем в методе золотого сечения при одном и том же числе вычислений функции (л. 4, стр. 37).

Однако метод Фибоначчи, в отличие от метода золотого сечения, нельзя прервать произвольно: число вычислений  $N$  должно быть задано заранее. Если процесс остановить раньше, то полученный отрезок локализации не будет оптимальным.

# Лекция 5. Методы спуска

## 1. Задача безусловной оптимизации

Рассматривается задача безусловной оптимизации

$$f(x) \rightarrow \inf, \quad x \in X, \quad (5.1)$$

где целевая функция  $f(\cdot) : X \rightarrow \mathbb{R}$ , а допустимое множество  $X = \mathbb{R}^n$  (л. 5, стр. 1).

**Замечание.**

1. Относительно функции  $x \mapsto f(x)$  предполагается, что она непрерывна и дифференцируема достаточное число раз.
2. Рассматриваются методы поиска локального минимума, однако если целевая функция  $f(x)$  является выпуклой, то найденный локальный минимум одновременно является и глобальным (л. 5, стр. 1).

## 2. Общая схема методов спуска

Рассмотрим общую итерационную схему:

$$x_{k+1} = \Phi(x_k), \quad (5.2)$$

где строится последовательность точек  $\{x_k\} \subset \mathbb{R}^n, k = 0, 1, \dots$ , такая что

$$f(x_{k+1}) < f(x_k). \quad (5.3)$$

Таким образом, последовательность значений целевой функции  $\{f(x_k)\}$  является убывающей:

$$f(x_{k+1}) < f(x_k). \quad (5.4)$$

Такая последовательность называется **релаксационной** (л. 5, стр. 2).

Начальная точка  $x_0$  задаётся, а на каждой последующей итерации точка  $x_k$  выбирается так, чтобы выполнялось условие (5.3). Этого добиваются выбором направления убывания функции  $f(x)$  и шага, на который осуществляется перемещение вдоль этого направления (л. 5, стр. 2–3).

При этом шаг, с которым движутся по направлению убывания функции  $f(x)$ , подбирают таким образом, чтобы для следующей точки выполнялось

$$f(x_{k+1}) < f(x_k).$$

Различные методы спуска отличаются друг от друга:

- способом выбора направления убывания;
- способом выбора шага;
- правилами остановки итерационного процесса (условия сходимости, критерии завершения) (л. 5, стр. 3).

### 3. Направление убывания. Определение и свойства

#### Определение 5.1.

Вектор  $s \in \mathbb{R}^n$ ,  $s \neq 0$ , называется **направлением убывания** целевой функции  $f(x)$  в точке  $x \in \mathbb{R}^n$ , если

$$f(x + \alpha s) < f(x) \quad (5.5)$$

для всех достаточно малых  $\alpha > 0$  (л. 5, стр. 3–4).

Множество всех направлений убывания функции  $f(x)$  в точке  $x$  обозначается через

$$T(x, f).$$

Это множество образует конус в  $\mathbb{R}^n$  (л. 5, стр. 4).

### 4. Связь направления убывания с градиентом

#### Лемма 5.1.

Пусть целевая функция  $f(x)$  дифференцируема в точке  $x \in \mathbb{R}^n$ . Тогда:

1. Если вектор  $s \in \mathbb{R}^n$  удовлетворяет условию

$$\langle f'(x), s \rangle < 0, \quad (5.6)$$

то  $s \in T(x, f)$ .

2. Если  $s \in T(x, f)$ , то

$$\langle f'(x), s \rangle \leq 0. \quad (5.7)$$

(л. 5, стр. 4).

#### Доказательство (п. 1).

Пусть выполнено условие (5.6). Рассмотрим приращение функции:

$$f(x + \alpha s) = f(x) + \langle f'(x), \alpha s \rangle + o(\alpha).$$

Тогда

$$f(x + \alpha s) - f(x) = \alpha \langle f'(x), s \rangle + o(\alpha).$$

Делим обе части на  $\alpha > 0$  и переходим к пределу при  $\alpha \rightarrow 0$ :

$$\lim_{\alpha \rightarrow 0} \frac{f(x + \alpha s) - f(x)}{\alpha} = \langle f'(x), s \rangle < 0.$$

Отсюда следует, что при достаточно малых  $\alpha > 0$  выполняется  $f(x + \alpha s) < f(x)$ , то есть  $s$  является направлением убывания.

Доказательство п. 2 проводится аналогично (л. 5, стр. 5–6).

#### Замечание.

Геометрически условие (5.6) означает, что вектор  $s$  образует тупой угол с градиентом целевой

функции  $f'(x) = \nabla f(x)$  в точке  $x$  (л. 5, стр. 6).

## 5. Итерационная схема метода спуска

Зададим отображение  $\Phi(\cdot)$  в итерационной формуле (5.2) следующим образом:

$$\Phi(x_k) = x_k + \alpha_k s_k. \quad (5.8)$$

Тогда

$$x_{k+1} = x_k + \alpha_k s_k. \quad (5.9)$$

### Определение 5.2.

Итерационный метод (5.9) называется **методом спуска**, если:

- вектор  $s_k$  задаёт направление убывания целевой функции  $f(x)$  в точке  $x_k$ , то есть

$$s_k \in T(x_k, f), \quad k = 0, 1, \dots, \quad (5.10)$$

- число  $\alpha_k > 0$  (шаг метода спуска) выбрано так, что

$$f(x_{k+1}) < f(x_k)$$

для всех  $k = 0, 1, \dots$  (л. 5, стр. 7).

Реализация метода (5.9) зависит от выбора векторов  $s_k$  и шага  $\alpha_k$ .

## 6. Простейший метод спуска: градиентный метод

Простейшим примером метода спуска является **градиентный метод**, для которого

$$s_k = -f'(x_k) = -\nabla f(x_k). \quad (5.11)$$

Тогда итерационная формула принимает вид:

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k). \quad (5.12)$$

При условии  $f'(x_k) \neq 0$  вектор  $s_k = -f'(x_k)$  действительно является направлением убывания, поскольку

$$\langle f'(x_k), s_k \rangle = -\|f'(x_k)\|^2 < 0,$$

то есть  $s_k \in T(x_k, f)$  (л. 5, стр. 8).

## 7. Принципы выбора шага метода спуска

Один из основных принципов выбора шага  $\alpha_k$  — минимизация целевой функции вдоль заданного направления. При фиксированном направлении  $s_k$  параметр  $\alpha_k$  в методе (5.9) определяется как решение одномерной задачи оптимизации:

$$f(x_k + \alpha s_k) \rightarrow \min_{\alpha \geq 0}. \quad (5.13)$$

В этом случае минимум берётся по  $\alpha \geq 0$ , поскольку  $s_k$  — направление убывания функции  $f(x)$  в точке  $x_k$ . Такой способ выбора шага называется **точным линейным поиском** (л. 5, стр. 8).

Рассмотренный способ выбора длины шага  $\alpha_k$  в методе спуска (5.9) является **точным линейным поиском**, поскольку он обеспечивает достижение наименьшего значения целевой функции вдоль заданного направления  $s_k \in T(x_k, f)$ . Однако для его реализации требуется решение вспомогательной задачи одномерной оптимизации (5.13), что, как правило, связано с существенными вычислительными затратами (л. 5, стр. 9).

В практических алгоритмах во многих случаях величину  $\alpha_k$  удаётся находить в явном виде.

### Пример 5.1. Квадратичная целевая функция

Пусть целевая функция имеет вид

$$f(x) = \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle, \quad (5.14)$$

где  $A$  — симметричная матрица,  $A = (a_{ij})_{i,j=1}^n$ ,  $b \in \mathbb{R}^n$ ,  $x \in \mathbb{R}^n$  (л. 5, стр. 10).

Рассмотрим схему метода спуска (5.9) для функции (5.14). Подставляя  $x_k + \alpha s_k$  в  $f(x)$ , получаем:

$$\begin{aligned} f(x_k + \alpha s_k) &= \frac{1}{2} \langle A(x_k + \alpha s_k), x_k + \alpha s_k \rangle + \langle b, x_k + \alpha s_k \rangle \\ &= \frac{1}{2} \langle A s_k, s_k \rangle \alpha^2 + \langle A x_k + b, s_k \rangle \alpha + \left\langle \frac{1}{2} A x_k + b, x_k \right\rangle. \end{aligned}$$

Обозначим

$$\varphi(\alpha) = f(x_k + \alpha s_k).$$

Тогда  $\varphi(\alpha)$  — квадратичная функция вида

$$\varphi(\alpha) = M_0 \alpha^2 + M_1 \alpha + M_2,$$

где

$$M_0 = \frac{1}{2} \langle A s_k, s_k \rangle, \quad M_1 = \langle A x_k + b, s_k \rangle, \quad M_2 = \left\langle \frac{1}{2} A x_k + b, x_k \right\rangle$$

(л. 5, стр. 10).

Найдём точку минимума:

$$\frac{d\varphi(\alpha)}{d\alpha} = 2M_0\alpha + M_1.$$

Из условия  $\varphi'(\alpha) = 0$  получаем

$$\alpha^* = -\frac{M_1}{2M_0} = -\frac{\langle A x_k + b, s_k \rangle}{\langle A s_k, s_k \rangle}.$$

Заметим, что

$$f'(x_k) = Ax_k + b,$$

поэтому оптимальный шаг можно записать в виде

$$\alpha^* = -\frac{\langle f'(x_k), s_k \rangle}{\langle As_k, s_k \rangle}. \quad (5.15)$$

Если  $s_k$  — направление убывания функции  $f(x)$ , то по лемме 5.1 выполняется

$$\langle f'(x_k), s_k \rangle \leq 0,$$

следовательно,

$$\alpha^* \geq 0.$$

Таким образом, в рассматриваемом случае

$$\varphi(\alpha_k) = \min_{\alpha \geq 0} \varphi(\alpha) = \min_{\alpha \in \mathbb{R}} \varphi(\alpha)$$

(л. 5, стр. 11–12).

## 2. Адаптивный способ определения шага

Адаптивный способ выбора шага  $\alpha_k$  не требует дополнительных вычислений производных или решения вспомогательных задач оптимизации. В отличие от точного линейного поиска, он использует лишь значения градиента в текущей точке и некоторые априорные оценки, характеризующие поведение целевой функции (л. 5, стр. 12).

Шаг  $\alpha_k$  выбирается так, чтобы при выполнении соответствующих предположений выполнялось неравенство

$$f(x_k + \alpha_k s_k) - f(x_k) \leq \varepsilon \alpha_k \langle f'(x_k), s_k \rangle, \quad (5.16)$$

где  $\varepsilon \in (0, 1)$ , а  $s_k$  — направление спуска, для которого

$$\langle f'(x_k), s_k \rangle < 0.$$

### Замечание 1.

Из неравенства (5.16) непосредственно следует

$$f(x_{k+1}) < f(x_k),$$

то есть соответствующий метод действительно является методом спуска. Действительно, так как  $\langle f'(x_k), s_k \rangle < 0$ ,  $\alpha_k > 0$  и  $\varepsilon > 0$ , правая часть (5.16) отрицательна, а значит

$$f(x_k + \alpha_k s_k) - f(x_k) < 0$$

(л. 5, стр. 13–14).

### Замечание 2.

Неравенство (5.16) играет важную роль при обосновании сходимости методов оптимизации,



относящихся к классу методов спуска. Ниже приводятся утверждения, дающие оценки для допустимого шага  $\alpha_k$  (л. 5, стр. 14).

### Лемма 5.2 (условие Липшица для градиента)

Пусть целевая функция  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$  дифференцируема на  $\mathbb{R}^n$ , а её градиент удовлетворяет условию Липшица:

$$\|f'(x) - f'(\tilde{x})\| \leq M\|x - \tilde{x}\|, \quad (5.17)$$

где  $x, \tilde{x} \in \mathbb{R}^n$ ,  $M > 0$ .

Тогда для любых  $\varepsilon \in (0, 1)$ ,  $x_k \in \mathbb{R}^n$  и направлений  $s_k$ , удовлетворяющих условию

$$\langle f'(x_k), s_k \rangle < 0,$$

неравенство (5.16) выполняется, если

$$0 < \alpha_k \leq -\frac{(1 - \varepsilon)\langle f'(x_k), s_k \rangle}{M\|s_k\|^2}. \quad (5.18)$$

(л. 5, стр. 15).

### Лемма 5.3 (оценка через вторую производную)

Пусть целевая функция  $f(x)$ ,  $x \in \mathbb{R}^n$ , дважды дифференцируема, и её матрица вторых производных (гессиан) удовлетворяет условию:

$$\langle f''(x)s, s \rangle \leq D\|s\|^2, \quad (5.19)$$

где  $D > 0$ .

Тогда для любых  $\varepsilon \in (0, 1)$ ,  $x_k \in \mathbb{R}^n$  и направлений  $s_k$ , удовлетворяющих условию

$$\langle f'(x_k), s_k \rangle < 0,$$

неравенство (5.16) выполняется при

$$0 < \alpha_k \leq -\frac{2(1 - \varepsilon)\langle f'(x_k), s_k \rangle}{D\|s_k\|^2}. \quad (5.20)$$

(л. 5, стр. 15–16).

#### Замечание.

Постоянные  $M$  и  $D$ , как правило, неизвестны, и условия для выбора шага  $\alpha_k$  в формулах (5.18) и (5.20) на практике не реализуются напрямую. Однако эти оценки подсказывают разумный способ адаптивного выбора шага  $\alpha_k$ , при котором неравенство (5.16) гарантированно выполняется (л. 5, стр. 16).

# Лекция 6. Градиентный метод. Сводка по выбору шага в методах спуска

## 1. Правило одномерной оптимизации

В методах спуска шаг  $\alpha_k$  на  $k$ -й итерации можно выбирать из условия минимизации целевой функции вдоль заданного направления спуска  $s_k$ :

$$f(x_k + \alpha_k s_k) = \min_{\alpha \geq 0} f(x_k + \alpha s_k).$$

Тем самым задача выбора шага сводится к задаче одномерной оптимизации

$$\varphi_k(\alpha) \rightarrow \min, \quad \alpha \in \mathbb{R}_{\geq 0}, \quad (6.1)$$

где введена вспомогательная функция

$$\varphi_k(\alpha) := f(x_k + \alpha s_k), \quad \varphi_k : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}.$$

(л. 6, стр. 1)

Предположим, что функция  $f(x)$  дифференцируема в точках вида  $x_k + \alpha s_k$ ,  $\alpha \geq 0$ . Тогда необходимое условие стационарности для задачи (6.1) имеет вид

$$\varphi'_k(\alpha) = \frac{d}{d\alpha} f(x_k + \alpha s_k) = 0.$$

Отсюда следует условие

$$\langle f'(x_k + \alpha s_k), s_k \rangle = 0. \quad (6.2)$$

Геометрически это означает ортогональность направления спуска  $s_k$  градиенту целевой функции в новой точке  $x_{k+1} = x_k + \alpha_k s_k$ : градиент в точке  $x_{k+1}$  ортогонален траектории движения вдоль  $s_k$  (л. 6, стр. 2).

Такой способ выбора шага называется **точным линейным поиском**. Он позволяет получить на очередной итерации наименьшее возможное значение целевой функции вдоль выбранного направления.

На практике задачу (6.1) часто заменяют более простой:

$$\varphi_k(\alpha) \rightarrow \min, \quad \alpha \in [0, \bar{\alpha}], \quad (6.3)$$

где  $\bar{\alpha} > 0$  — фиксированный параметр. В методах градиентного спуска (методах спуска по антиградиенту) именно эта постановка используется наиболее часто (л. 6, стр. 3).

## 2. Правило Армихо

Пусть функция  $f(x)$  дифференцируема в точке  $x_k$ . Зафиксируем параметры  $\varepsilon \in (0, 1)$  и  $\bar{\alpha} > 0$ . Положим первоначально  $\alpha = \bar{\alpha}$ .

**Шаг 1.** Проверяется выполнение неравенства Армихо:

$$f(x_k + \alpha s_k) \leq f(x_k) + \varepsilon \alpha \langle f'(x_k), s_k \rangle. \quad (6.4)$$

**Шаг 2.** Если неравенство (6.4) не выполняется, то шаг уменьшают, например, по правилу

$$\alpha := \frac{1}{2} \alpha,$$

и снова переходят к шагу 1. Процесс продолжается до тех пор, пока неравенство Армихо не станет верным (л. 6, стр. 4).

Итоговый шаг  $\alpha_k$  выбирается как первое значение  $\alpha$ , полученное в результате дробления начального шага  $\bar{\alpha}$ , для которого выполняется (6.4). Отсюда сразу видно, что направление  $s_k$  обязательно должно быть направлением убывания:

$$\langle f'(x_k), s_k \rangle < 0. \quad (6.5)$$

## Лемма 6.1

Пусть  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  — дифференцируемая функция,  $x_k \in \mathbb{R}^n$ . Если направление  $s_k \in \mathbb{R}^n$  удовлетворяет условию

$$\langle f'(x_k), s_k \rangle < 0, \quad (6.5)$$

то неравенство Армихо (6.4) выполняется для любого достаточно малого шага  $\alpha > 0$  (л. 6, стр. 5).

**Замечание.** Условие (6.5) гарантирует, что при выборе шага по правилу Армихо выполняется строгое убывание целевой функции:

$$f(x_{k+1}) < f(x_k). \quad (6.6)$$

Кроме того, неравенство Армихо даёт количественную оценку уменьшения значения функции, что важно для анализа скорости сходимости метода спуска.

## Лемма 6.2

Пусть  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  дифференцируема на  $\mathbb{R}^n$ , и её градиент липшицев непрерывен с константой  $L > 0$ , то есть

$$\|f'(x) - f'(y)\| \leq L \|x - y\| \quad \forall x, y \in \mathbb{R}^n.$$

Если для некоторых  $x_k, s_k \in \mathbb{R}^n$  выполнено условие

$$\langle f'(x_k), s_k \rangle < 0,$$

то неравенство Армихо выполняется для всех

$$\alpha \in (0, \bar{\alpha}_k], \quad (6.7)$$

где

$$\bar{\alpha}_k = -\frac{2(1-\varepsilon)\langle f'(x_k), s_k \rangle}{L\|s_k\|^2} > 0. \quad (6.8)$$

(л. 6, стр. 6–7)

### Лемма 6.3

При выполнении условий леммы 6.2 для любых  $x, z \in \mathbb{R}^n$  справедлива оценка

$$|f(x+z) - f(x) - \langle f'(x), z \rangle| \leq \frac{L}{2}\|z\|^2.$$

**Доказательство.** По формуле Ньютона–Лейбница:

$$f(x+z) = f(x) + \int_0^1 \langle f'(x+tz), z \rangle dt.$$

Вычитая  $\langle f'(x), z \rangle$  и применяя неравенство Коши–Буняковского и липшицевость градиента, получаем требуемую оценку (л. 6, стр. 7–9).

### Следствия для правила Армихо

Из лемм 6.2 и 6.3 следует, что при выборе

$$\alpha_k = -\frac{2(1-\varepsilon)\langle f'(x_k), s_k \rangle}{L\|s_k\|^2}$$

неравенство Армихо гарантированно выполняется. Более того, существует число  $\delta > 0$ , не зависящее от номера итерации  $k$ , такое что

$$\frac{\langle f'(x_k), s_k \rangle}{\|s_k\|^2} < \delta. \quad (6.10)$$

Отсюда следует важный практический вывод: если параметры  $\varepsilon$  и  $\bar{\alpha}$  фиксированы одинаково на всех итерациях, то число дроблений шага в правиле Армихо будет равномерно ограничено по  $k$ , а шаги  $\alpha_k$  не будут стремиться к нулю слишком быстро (л. 6, стр. 10–12).

На практике это означает, что начиная с некоторой итерации шаг перестаёт уменьшаться, и алгоритм фактически переходит к использованию почти постоянного шага.

Из неравенства

$$\frac{\langle f'(x_k), s_k \rangle}{\|s_k\|^2} < \delta, \quad (6.10)$$

где  $\delta > 0$  не зависит от номера итерации  $k$ , а число  $\varepsilon$  выбрано фиксированным и одинаковым для всех шагов, следует, что количество дроблений шага в правиле Армихо будет конечно и

равномерно ограничено по  $k$ . Иными словами, шаги  $\alpha_k$  не стремятся к нулю слишком быстро и остаются отделёнными от нуля при больших  $k$  (л. 6, стр. 12).

Практический смысл этого факта состоит в том, что начиная с некоторой итерации длина шага перестаёт уменьшаться, и метод фактически переходит к использованию почти постоянного шага. Это сближает правило Армихо с методом постоянного шага, но с автоматическим подбором подходящего значения.

## Правило постоянного шага (постоянного параметра)

Фиксируется число  $\alpha > 0$ , не зависящее от номера итерации  $k$ , и полагается

$$\alpha_k \equiv \alpha.$$

Это наиболее простой способ выбора шага, который обычно применяется тогда, когда вычисление целевой функции является сравнительно дешёвым и заранее известны оценки на параметры задачи.

Если в условиях леммы 6.2 выполнено условие

$$\frac{\langle f'(x_k), s_k \rangle}{\|s_k\|^2} < \delta,$$

то при достаточно малом фиксированном  $\alpha$  неравенство Армихо будет выполняться на каждом шаге при  $\alpha_k = \alpha$ . В этом случае анализ метода спуска с постоянным шагом полностью аналогичен анализу метода с выбором шага по правилу Армихо (л. 6, стр. 13).

Если известна константа Липшица  $L$  (или верхняя оценка для неё), то длину шага можно оценить непосредственно по формуле

$$\alpha_k = -\frac{2(1 - \varepsilon)\langle f'(x_k), s_k \rangle}{L\|s_k\|^2}.$$

Из этой формулы следует, что чем больше  $L$ , тем меньше допустимый шаг, и наоборот. Отсюда видно, что скорость сходимости метода спуска существенно зависит от величины константы Липшица градиента (л. 6, стр. 14).

Основные правила выбора шага:

- правило одномерной оптимизации;
- правило Армихо;
- правило постоянного шага.

## Правило Голдстейна

Правило Голдстейна задаётся двойным неравенством

$$\varepsilon_1 \leq \frac{f(x_k + \alpha s_k) - f(x_k)}{\alpha \langle f'(x_k), s_k \rangle} \leq \varepsilon_2, \quad \varepsilon_1, \varepsilon_2 \in (0, 1), \quad \varepsilon_1 < \varepsilon_2. \quad (6.11)$$

(л. 6, стр. 14)

## Интерпретация.

- Левое неравенство в (6.11) совпадает с неравенством Армихо при  $\varepsilon = \varepsilon_1$  и обеспечивает убывание целевой функции вдоль траектории метода.
- Правое неравенство запрещает выбирать слишком малый шаг, который не даёт заметного продвижения к минимуму.

Согласно лемме 6.1, неравенство Армихо выполняется для достаточно малого шага  $\alpha$ , однако это не гарантирует выполнения правого неравенства в (6.11). Поэтому правило Голдстейна, в отличие от правила Армихо, не допускает чрезмерного уменьшения шага и тем самым предотвращает излишне медленную сходимость (л. 6, стр. 15).

## Правило Вульфа

Правило Вульфа задаётся системой неравенств

$$f(x_k + \alpha s_k) \leq f(x_k) + \varepsilon_1 \alpha \langle f'(x_k), s_k \rangle, \quad (6.12)$$

$$\langle f'(x_k + \alpha s_k), s_k \rangle \geq \varepsilon_2 \langle f'(x_k), s_k \rangle, \quad 0 < \varepsilon_1 < \varepsilon_2 < 1. \quad (6.13)$$

(л. 6, стр. 16)

Первое неравенство совпадает с условием Армихо и гарантирует убывание функции. Второе неравенство ограничивает уменьшение производной вдоль направления  $s_k$  и, тем самым, предотвращает выбор слишком малого шага.

Если вычисление градиента не является чрезмерно дорогим, правило Вульфа считается одним из наиболее эффективных правил выбора шага в методах спуска.

## Алгоритм реализации правила Вульфа

1. Фиксируются параметры  $\varepsilon_1, \varepsilon_2 \in (0, 1)$ ,  $\varepsilon_1 < \varepsilon_2$ .
  2. Полагается  $\underline{\alpha} = 0$ .
  3. Выбирается начальное пробное значение шага  $\alpha > 0$ .
  4. Проверяются неравенства (6.12) и (6.13).
    - Если оба выполняются, переходят к шагу 6.
    - Если нарушено (6.12), полагают  $\bar{\alpha} = \alpha$  и переходят к шагу 5.
    - Если нарушено (6.13), полагают  $\underline{\alpha} = \alpha$  и переходят к шагу 5.
  5. Если  $\bar{\alpha} = 0$ , выбирается новое пробное значение  $\alpha > \underline{\alpha}$  (экстраполяция).  
Если  $\bar{\alpha} > 0$ , выбирается новое  $\alpha \in (\underline{\alpha}, \bar{\alpha})$  (интерполяция).
  6. Полагается  $\alpha_k = \alpha$ .
- (л. 6, стр. 16–17)

## Интерпретация правила Вульфа.

- Нарушение (6.12) означает, что выбранный шаг слишком велик.

- Нарушение (6.13) означает, что шаг слишком мал.

В процессе работы алгоритма шаг сначала может увеличиваться, затем уменьшается, причём всегда сохраняется условие  $\underline{\alpha} < \alpha < \bar{\alpha}$  (л. 6, стр. 18).

## Схема градиентного метода

**Предположение.** Целевая функция  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  дифференцируема.

В градиентных методах спуска направление убывания выбирается в виде антиградиента:

$$s_k = -\nabla f(x_k) = -f'(x_k). \quad (6.14)$$

Если  $f'(x_k) = 0$ , то точка  $x_k$  является стационарной точкой задачи минимизации, и работа метода прекращается (л. 6, стр. 19).

Тогда итерационная формула градиентного метода имеет вид

$$x_{k+1} = x_k - \alpha_k f'(x_k), \quad (6.15)$$

где  $k = 0, 1, 2, \dots$ . Из формулы (6.15) следует, что градиентные методы относятся к методам первого порядка (л. 6, стр. 20).

## Алгоритм градиентного метода

1. Выбирается начальная точка  $x_0 \in \mathbb{R}^n$ , полагается  $k = 0$ .
2. Выбирается правило определения шага  $\alpha_k$  и соответствующие параметры:
  - $\bar{\alpha} > 0$  — при правиле одномерной оптимизации;
  - $\bar{\alpha} > 0, \varepsilon \in (0, 1)$  — при правиле Армихо;
  - $\alpha > 0$  — при правиле постоянного шага.
3. Вычисляется шаг  $\alpha_k$  в соответствии с выбранным правилом.
4. Если  $f'(x_k) \neq 0$ , вычисляется новая точка

$$x_{k+1} = x_k - \alpha_k f'(x_k).$$

5. Полагается  $k := k + 1$  и осуществляется переход к шагу 3.  
(л. 6, стр. 21)

**Определение 6.2.** Градиентный метод, в котором шаг  $\alpha_k$  выбирается по правилу одномерной оптимизации, называется **методом скорейшего спуска** (л. 6, стр. 21).

## Особенности метода скорейшего спуска

Важным отличительным свойством метода скорейшего спуска является выполнение равенства

$$\langle f'(x_{k+1}), f'(x_k) \rangle = 0. \quad (6.16)$$

(л. 6, стр. 22)

Действительно, шаг  $\alpha_k$  выбирается как решение задачи одномерной оптимизации

$$\alpha_k = \arg \min_{\alpha \geq 0} \varphi_k(\alpha), \quad \varphi_k(\alpha) = f(x_k - \alpha f'(x_k)).$$

Условие стационарности даёт

$$\varphi'_k(\alpha_k) = -\langle f'(x_{k+1}), f'(x_k) \rangle = 0,$$

что и эквивалентно (6.16). Геометрически это означает, что градиенты в соседних точках ортогональны (л. 6, стр. 22–23).

## Условия эффективности градиентного метода

Рассматривается класс функций  $f \in C^{1,1}(\mathbb{R}^n)$ , то есть функций, непрерывно дифференцируемых на  $\mathbb{R}^n$  и имеющих липшицев непрерывный градиент:

$$\|f'(x) - f'(y)\| \leq L\|x - y\|.$$

Для одной итерации градиентного метода

$$x_{k+1} = x_k - \alpha_k f'(x_k)$$

условие направления убывания выполняется автоматически, так как при

$$s_k = -f'(x_k), \quad f'(x_k) \neq 0,$$

имеем

$$\langle f'(x_k), s_k \rangle = -\|f'(x_k)\|^2 < 0.$$

(л. 6, стр. 24)

Теперь воспользуемся леммой 6.3. Для любых  $x, z \in \mathbb{R}^n$  справедлива оценка

$$|f(x+z) - f(x) - \langle f'(x), z \rangle| \leq \frac{L}{2} \|z\|^2.$$

Положим

$$z = x_{k+1} - x_k, \quad x = x_k.$$

Тогда

$$|f(x_{k+1}) - f(x_k) - \langle f'(x_k), x_{k+1} - x_k \rangle| \leq \frac{L}{2} \|x_{k+1} - x_k\|^2.$$

(л. 6, стр. 25)

Подставим формулу градиентного шага  $x_{k+1} = x_k - \alpha_k f'(x_k)$ . Получаем



$$f(x_{k+1}) - f(x_k) - \langle f'(x_k), -\alpha_k f'(x_k) \rangle \leq \frac{L}{2} \|\alpha_k f'(x_k)\|^2.$$

Отсюда

$$f(x_{k+1}) \leq f(x_k) - \alpha_k \|f'(x_k)\|^2 + \frac{L}{2} \alpha_k^2 \|f'(x_k)\|^2.$$

Или, что эквивалентно,

$$f(x_{k+1}) = f(x_k) - \alpha_k \left(1 - \frac{L\alpha_k}{2}\right) \|f'(x_k)\|^2. \quad (6.17)$$

(л. 6, стр. 26)

Следовательно, уменьшение целевой функции за одну итерацию при фиксированном шаге  $\alpha_k$  равно

$$\delta f(\alpha_k) := f(x_k) - f(x_{k+1}) = \alpha_k \left(1 - \frac{L\alpha_k}{2}\right) \|f'(x_k)\|^2. \quad (6.18)$$

Чтобы за одну итерацию убывание целевой функции было максимальным, необходимо максимизировать величину  $\delta f(\alpha_k)$  по  $\alpha_k > 0$ . Представим её в виде

$$\delta f(\alpha_k) = \Delta(\alpha_k) \|f'(x_k)\|^2, \quad (6.19)$$

где

$$\Delta(\alpha_k) = \alpha_k \left(1 - \frac{L\alpha_k}{2}\right). \quad (6.20)$$

Тогда задача максимизации  $\delta f(\alpha_k)$  эквивалентна задаче

$$\Delta(\alpha_k) \rightarrow \max, \quad \alpha_k > 0. \quad (6.21)$$

(л. 6, стр. 27)

Условие стационарности:

$$\Delta'(\alpha_k) = 1 - L\alpha_k = 0,$$

откуда

$$\alpha_k^* = \frac{1}{L}, \quad L > 0. \quad (6.22)$$

Вторая производная

$$\Delta''(\alpha_k) = -L < 0,$$

поэтому  $\alpha_k^*$  действительно соответствует максимуму.

Следовательно, за один шаг градиентного метода с постоянным шагом  $\alpha_k^* = \frac{1}{L}$  целевая функция уменьшается на величину

$$\delta f(\alpha_k^*) = \Delta(\alpha_k^*) \|f'(x_k)\|^2 = \frac{1}{L} \left(1 - \frac{1}{2}\right) \|f'(x_k)\|^2 = \frac{1}{2L} \|f'(x_k)\|^2. \quad (6.23)$$

## Постоянный шаг

Рассмотрим выбор постоянного шага

$$\alpha_k = \alpha = \text{const}, \quad k = 0, 1, \dots$$

Тогда схема градиентного метода имеет вид

$$x_{k+1} = x_k - \alpha f'(x_k). \quad (6.24)$$

Неравенство (6.17) принимает вид

$$f(x_k) - f(x_{k+1}) \geq \alpha \left(1 - \frac{L\alpha}{2}\right) \|f'(x_k)\|^2. \quad (6.25)$$

Если положить

$$\alpha = \frac{2h}{L}, \quad h \in (0, 1), \quad (6.26)$$

то

$$f(x_k) - f(x_{k+1}) \geq \frac{2h}{L} (1 - h) \|f'(x_k)\|^2.$$

Введём функцию

$$\Delta(h) = \frac{2}{L} h(1 - h).$$

Её производная

$$\Delta'(h) = \frac{2}{L} (1 - 2h) = 0$$

даёт стационарную точку

$$h^* = \frac{1}{2}.$$

Так как

$$\Delta''(h) = -\frac{4}{L} < 0,$$

эта точка соответствует максимуму. При  $h = h^*$  получаем

$$\alpha = \frac{1}{L}, \quad (6.27)$$

и неравенство (6.25) принимает вид

$$f(x_k) - f(x_{k+1}) \geq \frac{1}{2L} \|f'(x_k)\|^2. \quad (6.28)$$

(л. 6, стр. 29)

Из (6.28) следует, что гарантированная оценка уменьшения целевой функции равна

$$\delta f(\alpha) = \frac{1}{2L} \|f'(x_k)\|^2, \quad (6.29)$$

что совпадает с оценкой (6.23). Таким образом, шаг

$$\alpha_k = \frac{1}{L} \quad (6.30)$$

является оптимальным для градиентного метода с постоянным шагом.

## Правило Армихо для градиентного метода

Для градиентного метода условие Армихо имеет вид

$$f(x_{k+1}) \leq f(x_k) + \varepsilon \alpha_k \langle f'(x_k), s_k \rangle = f(x_k) - \varepsilon \alpha_k \|f'(x_k)\|^2,$$

то есть

$$f(x_k) - f(x_{k+1}) \geq \varepsilon \alpha_k \|f'(x_k)\|^2.$$

С учётом (6.17) получаем систему неравенств:

$$f(x_k) - f(x_{k+1}) \geq \varepsilon \alpha_k \|f'(x_k)\|^2,$$

$$f(x_k) - f(x_{k+1}) \geq \alpha_k \left(1 - \frac{L\alpha_k}{2}\right) \|f'(x_k)\|^2.$$

Отсюда следует условие

$$\varepsilon \alpha_k \leq \alpha_k \left(1 - \frac{L\alpha_k}{2}\right),$$

то есть

$$\alpha_k \leq \frac{2}{L} (1 - \varepsilon), \quad \varepsilon \in (0, 1). \quad (6.31)$$

(л. 6, стр. 31)

Если это условие выполнено, то уменьшение целевой функции при правиле Армихо равно

$$\delta f(\alpha_k) = \Delta(\alpha_k) \|f'(x_k)\|^2, \quad (6.32)$$

где  $\Delta(\alpha_k) = \alpha_k \left(1 - \frac{L\alpha_k}{2}\right)$ .

Сравнивая оценки (6.23), (6.29) и (6.32), приходим к выводу, что при использовании всех трёх основных правил выбора шага в методе градиентного спуска выполняется неравенство

$$f(x_k) - f(x_{k+1}) \geq \frac{\omega}{L} \|f'(x_k)\|^2, \quad (6.33)$$

где  $\omega > 0$  — некоторая постоянная (л. 6, стр. 32).

Просуммируем неравенство (6.33) по  $k = 0, 1, \dots, N$ :

$$\frac{\omega}{L} \sum_{k=0}^N \|f'(x_k)\|^2 \leq \sum_{k=0}^N (f(x_k) - f(x_{k+1})) = f(x_0) - f(x_{N+1}) \leq f(x_0) - f^*,$$

где  $f^*$  — решение оптимизационной задачи. Тогда

$$\frac{\omega}{L} \sum_{k=0}^{\infty} \|f'(x_k)\|^2 \leq f(x_0) - f^*. \quad (6.34)$$

Отсюда следует

$$\lim_{k \rightarrow \infty} \|f'(x_k)\| = 0. \quad (6.35)$$

То есть градиентный метод спуска сходится (л. 6, стр. 33).

## Оценка скорости сходимости

Введём обозначение

$$g_k := \|f'(x_k)\|, \quad g_N^* := \min_{0 \leq k \leq N} g_k.$$

Из (6.34) следует оценка

$$g_N^* \leq \frac{1}{\sqrt{N+1}} \left[ \frac{L}{\omega} (f(x_0) - f^*) \right]^{1/2}. \quad (6.36)$$

(л. 6, стр. 34)

Правая часть неравенства (6.36) даёт явную скорость сходимости последовательности  $\{g_N^*\}$  к нулю. При этом о глобальной сходимости последовательностей  $\{x_k\}$  и  $\{f(x_k)\}$  в общем случае утверждать нельзя.

## Локальная сходимость градиентного метода

Рассмотрим задачу оптимизации

$$f(x) \rightarrow \min, \quad x \in \mathbb{R}^n,$$

при следующих допущениях:

1.  $f \in C^{2,2}(\mathbb{R}^n)$ ;
2. существует точка локального минимума  $x^* \in \mathbb{R}^n$ ,

$$x^* = \arg \min f(x),$$

в которой гессиан положительно определён:

$$f''(x^*) > 0;$$

3. гессиан удовлетворяет оценке

$$\ell E_n \leq f''(x^*) \leq L E_n, \quad (6.37)$$

где  $0 < \ell \leq L < \infty$ ;

4. начальная точка  $x_0$  достаточно близка к  $x^*$ .

(л. 6, стр. 35)

Рассмотрим градиентную схему

$$x_{k+1} = x_k - \alpha_k f'(x_k).$$

При её реализации выполняется условие стационарности в точке минимума:

$$f'(x^*) = 0.$$

Для градиента в точке  $x_k$  справедливо точное интегральное представление:

$$f'(x_k) = f'(x_k) - f'(x^*) = \int_0^1 f''(x^* + \tau(x_k - x^*)) (x_k - x^*) d\tau. \quad (6.38)$$

(л. 6, стр. 36)

Обозначим

$$G_k := \int_0^1 f''(x^* + \tau(x_k - x^*)) d\tau.$$

Тогда формула (6.38) переписывается в виде

$$f'(x_k) = G_k(x_k - x^*). \quad (6.39)$$

Подставляя это выражение в градиентную схему, получаем

$$x_{k+1} = x_k - \alpha_k G_k(x_k - x^*),$$

или, что эквивалентно,

$$x_{k+1} - x^* = (E_n - \alpha_k G_k)(x_k - x^*). \quad (6.40)$$

Обозначим ошибку итерации

$$e_k := x_k - x^*.$$

Тогда рекуррентное соотношение для ошибок имеет вид

$$e_{k+1} = (E_n - \alpha_k G_k)e_k. \quad (6.41)$$

Для анализа сходимости используется принцип сжимающих отображений. Пусть последовательность  $\{a_k\}$  определяется рекуррентно:

$$a_{k+1} = A_k a_k, \quad k = 0, 1, \dots,$$

где  $A_k$  — квадратные матрицы такие, что

$$\|A_k\| \leq 1 - q, \quad q \in (0, 1).$$

Тогда

$$\|a_{k+1}\| \leq (1 - q)\|a_k\| \leq (1 - q)^{k+1}\|a_0\|,$$

и, следовательно,

$$(1 - q)^{k+1}\|a_0\| \xrightarrow{k \rightarrow \infty} 0. \quad (6.43)$$

(л. 6, стр. 37)

В нашем случае необходимо оценить норму оператора

$$E_n - \alpha_k G_k.$$

Положим

$$r_k := \|x_k - x^*\| = \|e_k\|.$$

## Лемма 6.4

Пусть  $f \in C^{2,2}(\mathbb{R}^n)$  и  $\|y - x\| = r$ . Тогда

$$f''(x) - Mr E_n \leq f''(y) \leq f''(x) + Mr E_n,$$

где  $M > 0$  — константа Липшица для гессиана, а знак  $\leq$  понимается в смысле неравенства для симметричных матриц (л. 6, стр. 38).

**Доказательство.** Матрица Гессе симметрична. Рассмотрим

$$G = f''(y) - f''(x).$$

Так как  $f \in C^{2,2}$ , имеем

$$\|G\| = \|f''(y) - f''(x)\| \leq M\|y - x\| = Mr.$$

Так как  $G$  симметрична, её собственные значения  $\lambda_i(G)$  удовлетворяют

$$|\lambda_i(G)| \leq Mr,$$

откуда

$$-Mr E_n \leq G \leq Mr E_n.$$

□

Применим лемму 6.4 к матрице

$$y = x^* + \tau(x_k - x^*), \quad x = x^*.$$

Тогда

$$\|y - x^*\| = \tau \|x_k - x^*\| = \tau r_k.$$

Из условий (6.37)

$$\ell E_n \leq f''(x^*) \leq L E_n$$

и леммы 6.4 следует:

$$\ell E_n - \tau r_k M E_n \leq f''(x^* + \tau(x_k - x^*)) \leq L E_n + \tau r_k M E_n.$$

Интегрируя по  $\tau \in [0, 1]$ , получаем оценку для  $G_k$ :

$$\left(\ell - \frac{M r_k}{2}\right) E_n \leq G_k \leq \left(L + \frac{M r_k}{2}\right) E_n.$$

(л. 6, стр. 39–40)

Домножим обе части неравенства на  $\alpha_k$  и вычтем из  $E_n$ :

$$\left[1 - \alpha_k \left(L + \frac{M r_k}{2}\right)\right] E_n \leq E_n - \alpha_k G_k \leq \left[1 - \alpha_k \left(\ell - \frac{M r_k}{2}\right)\right] E_n.$$

Введём обозначения

$$a_k(\alpha_k) = 1 - \alpha_k \left(\ell - \frac{M r_k}{2}\right), \quad (6.44)$$

$$b_k(\alpha_k) = 1 - \alpha_k \left(L + \frac{M r_k}{2}\right).$$

Тогда

$$\|E_n - \alpha_k G_k\| \leq \max\{|a_k(\alpha_k)|, |b_k(\alpha_k)|\}. \quad (6.45)$$

Заметим, что

$$a_k(0) = 1, \quad b_k(0) = 1.$$

Кроме того,

$$a'_k(\alpha_k) = -\left(\ell - \frac{M r_k}{2}\right) < 0 \quad \text{при } r_k < \frac{2\ell}{M}.$$

Следовательно, функция  $a_k(\alpha_k)$  строго убывает при

$$r_k < \bar{r} := \frac{2\ell}{M}.$$

Если же  $r_k > \bar{r}$ , то  $a'_k(\alpha_k) > 0$ , и функция становится возрастающей. Из (6.45) видно, что при достаточно малых  $\alpha_k$  всегда выполняется

$$\|E_n - \alpha_k G_k\| < 1,$$

и, следовательно,

$$r_{k+1} < r_k.$$

(л. 6, стр. 41–42)

Рассмотрим теперь оптимальный выбор шага в смысле принципа сжимающих отображений, то есть задачу

$$\max\{|a_k(\alpha_k)|, |b_k(\alpha_k)|\} \rightarrow \min, \quad \alpha_k \geq 0. \quad (6.46)$$

Пусть  $r_k < \bar{r}$ . Тогда решение задачи (6.46) достигается в точке, где

$$a_k(\alpha_k) = -b_k(\alpha_k).$$

Это приводит к уравнению

$$1 - \alpha_k \left( \ell - \frac{Mr_k}{2} \right) = - \left[ 1 - \alpha_k \left( L + \frac{Mr_k}{2} \right) \right].$$

Отсюда находим оптимальный шаг:

$$\alpha_k^* = \frac{2}{\ell + L}. \quad (6.47)$$

Заметим, что  $\alpha_k^*$  **не зависит** от константы Липшица  $M$  для гессиана и от номера итерации  $k$  (л. 6, стр. 43).

## Теорема 6.1

Пусть целевая функция  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  удовлетворяет условиям 1)–4), и

$$r_0 = \|x_0 - x^*\| < \bar{r} = \frac{2\ell}{M}.$$

Тогда градиентный метод с постоянным шагом

$$\alpha_k = \frac{2}{\ell + L}$$

сходится линейно, причём

$$\|x_k - x^*\| \leq \frac{\bar{r} r_0}{\bar{r} - r_0} \left( 1 - \frac{2\ell}{L + 3\ell} \right)^k.$$

Таким образом, при достаточно близком к минимуму начальном приближении градиентный метод обладает линейной (геометрической) скоростью локальной сходимости (л. 6, стр. 44–48).



# Лекция 7. Градиентные методы. Сильная выпуклость.

## Релаксационные схемы

### Сильно выпуклые функции

**Определение.** Функция  $f(x)$  называется *сильно выпуклой*, если она дважды непрерывно дифференцируема и её матрица Гессе  $f''(x)$  при всех  $x \in \mathbb{R}^n$  удовлетворяет неравенству

$$m\|y\|^2 \leq \langle f''(x)y, y \rangle \leq M\|y\|^2 \quad \forall y \in \mathbb{R}^n,$$

где  $M \geq m > 0$  — некоторые постоянные (л. 7, стр. 1).

Иными словами, все собственные значения матрицы Гессе равномерно ограничены снизу и сверху положительными константами. Это означает, что функция имеет «одинаковую» кривизну во всех направлениях и не содержит плоских направлений.

### Сходимость градиентного метода для сильно выпуклых функций

**Теорема.** Пусть  $f(x)$  — сильно выпуклая функция, а последовательность  $\{x_k\}_{k=1}^\infty$  строится по методу градиентного спуска

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k),$$

где шаг  $\alpha_k$  выбирается по правилу точного одномерного поиска:

$$\alpha_k = \arg \min_{\alpha > 0} f(x_k - \alpha \nabla f(x_k)).$$

Тогда последовательность  $\{x_k\}$  сходится к точке минимума  $x^*$  со скоростью геометрической прогрессии (л. 7, стр. 1).

Более точно, для достаточно больших  $k$  выполняется оценка

$$\|x_{k+1} - x^*\| \leq q \|x_k - x^*\|, \quad q = \frac{M - m}{M + m}.$$

Здесь  $q \in (0, 1)$  — знаменатель геометрической прогрессии, определяющий скорость сходимости (л. 7, стр. 2).

### Эффект оврагов и релаксационные методы

Скорость сходимости градиентного метода существенно зависит от соотношения  $M/m$ , то есть от числа обусловленности матрицы Гессе.

Если  $M \approx m$ , то  $q \approx 0$ , и метод сходится быстро. Если же

$$\frac{m}{M} \ll 1,$$

то  $q \approx 1$ , и градиентные методы сходятся медленно (л. 7, стр. 3).

Этот эффект имеет наглядную геометрическую интерпретацию и называется *эффектом оврагов*.

## Геометрическая интерпретация эффекта оврагов

Поверхности уровня функции  $f(x) = \text{const}$  в плохо обусловленном случае имеют форму сильно вытянутых эллипсов. Траектория градиентного спуска быстро «падает» на дно оврага, но затем медленно движется вдоль него к точке минимума, совершая зигзагообразные шаги (л. 7, стр. 4).

На рисунке показана типичная траектория градиентного метода, сходящаяся к минимуму в точке  $(0, 0)^T$ .

## Масштабирование переменных

Один из способов борьбы с эффектом оврагов — изменение масштаба переменных.

Рассмотрим пример. Пусть

$$f(x) = \sum_{i=1}^n \lambda_i x_i^2, \quad \lambda_i > 0,$$

и коэффициенты  $\lambda_i$  сильно различаются между собой (л. 7, стр. 5).

Поверхности уровня такой функции вытянуты вдоль осей  $x_i$ , соответствующих малым значениям  $\lambda_i$ .

Введём замену переменных

$$x_i = \mu_i y_i,$$

и подберём коэффициенты  $\mu_i$  так, чтобы поверхности уровня в новых переменных стали сферами. Достаточно положить

$$\mu_i = \lambda_i^{-1/2}.$$

Тогда

$$x_i = \lambda_i^{-1/2} y_i,$$

и после подстановки в функцию получаем

$$f(x) = \sum_{i=1}^n y_i^2,$$

то есть в новых координатах поверхности уровня являются сферами (л. 7, стр. 5–6).

## Общий случай: использование матрицы Гессе

Если функция  $f(x)$  не является квадратичной, но достаточно гладкая, коэффициенты масштабирования выбирают на основе локальной кривизны:

$$\mu_i = \left( \frac{\partial^2 f(x)}{\partial x_i^2} \right)^{-1/2},$$

где производные вычисляются в точке одномерного минимума вдоль соответствующей оси (л. 7, стр. 6).

В более общем виде для устранения эффекта оврагов используют все элементы матрицы Гессе и применяют преобразование

$$y = (f''(x))^{1/2} x,$$

где под матрицей  $(f''(x))^{1/2}$  понимается симметричная матрица  $A$ , такая что

$$f''(x) = AA = A^2, \quad f''(x) > 0$$

(л. 7, стр. 7).

Такое преобразование приводит (локально) к выравниванию кривизны поверхности уровня.

## Релаксационные матричные градиентные методы

В результате масштабирования переменных естественным образом возникает итерационный процесс вида

$$x_{k+1} = x_k - \alpha_k B_k \nabla f(x_k),$$

где матрица  $B_k$  может зависеть от номера итерации  $k$  (л. 7, стр. 8).

Методы такого вида называются *релаксационными матричными градиентными методами*.

Частный случай:

- при  $B_k = E$  (единичная матрица) получаем обычный градиентный метод;
- при специальном выборе  $B_k$  удаётся существенно ускорить сходимость, уменьшая эффект оврагов (л. 7, стр. 8).

# Лекция 8. Эвристические схемы метода градиентного спуска

## 1. Общая идея эвристических схем градиентного спуска

Эвристические схемы метода градиентного спуска применяются для минимизации функций со **сложной топологической структурой**, когда классические схемы градиентного метода оказываются неэффективными или сходятся слишком медленно. Основная идея состоит в **комбинировании градиентного шага с релаксационными приёмами**, позволяющими адаптировать направление и величину шага в процессе итераций (л. 8, стр. 1).

Общая итерационная схема имеет вид

$$x_{k+1} = x_k - \alpha_k B_k \nabla f(x_k),$$

где

- $x_k \in \mathbb{R}^n$  — текущая итерация,
- $\alpha_k > 0$  — шаг метода,
- $B_k$  — некоторая матрица (в простейшем случае единичная),
- $\nabla f(x_k)$  — градиент минимизируемой функции  $f$  в точке  $x_k$  (л. 8, стр. 1).

## 2. Релаксационные процедуры

### Процедура 1

Пусть в текущей точке  $x$  вычислены все частные производные

$$\frac{\partial f}{\partial x_i}, \quad i = 1, \dots, n.$$

Задаётся малое число  $\varepsilon_1 > 0$ . Если

$$\left| \frac{\partial f}{\partial x_i} \right| \leq \varepsilon_1,$$

то считается, что по координате  $x_i$  функция практически не убывает, и движение вдоль этой координаты временно не производится (л. 8, стр. 1).

### Интерпретация.

Фактически исключаются направления, в которых функция убывает слишком слабо. Это позволяет сосредоточить вычислительные ресурсы на «существенных» направлениях убывания.

### Процедура 2

Задаётся число  $\varepsilon_2 > \varepsilon_1$ . Если

$$\left| \frac{\partial f}{\partial x_i} \right| \geq \varepsilon_2,$$

то по координате  $x_i$  выполняется градиентный шаг (л. 8, стр. 2).

В этом случае движение производится **в первую очередь** по тем координатам, вдоль которых убывание функции наиболее интенсивно.

### 3. Комбинированная эвристическая схема

Комбинируя процедуры 1 и 2, получают следующую схему (л. 8, стр. 2–3):

1. В каждой итерации анализируются значения всех частных производных.
2. Формируется подпространство координат, по которым убывание функции существенно.
3. Градиентный шаг выполняется **только в этом подпространстве**.
4. Процесс продолжается до тех пор, пока очередная итерация не приведёт к точке, где значения всех частных производных малы.

Если выполнены условия сходимости, то после некоторого числа итераций происходит переход к обычному градиентному методу в полном пространстве.

#### Замечание.

Подобные методы особенно эффективны для целевых функций со сложной топологией, содержащих «плато» или узкие долины (л. 8, стр. 3).

### 4. Геометрическая интерпретация

На плоскости  $x_1, x_2$  траектория итераций может выглядеть следующим образом: точки  $x_0, x_1, x_2, \dots$  последовательно приближаются к линии уровня, после чего движение продолжается вдоль неё с последующим спуском к минимуму (л. 8, стр. 4).

Идея состоит в том, что алгоритм:

- сначала быстро «падает» к области минимума,
- затем уточняет положение минимума, двигаясь вдоль линий уровня.

### 5. Метод покоординатного спуска

#### Цель метода

Минимизация функции  $f(x)$  осуществляется **последовательно по отдельным координатам**, а не по всему вектору  $x$  сразу (л. 8, стр. 5).

Пусть

$$x_0 = (x_0^{(1)}, \dots, x_0^{(n)})^T$$

— начальное приближение.

## Первая итерация

Вычисляется частная производная по первой координате:

$$\left. \frac{\partial f(x)}{\partial x^{(1)}} \right|_{x=x_0}.$$

Далее выполняется шаг

$$x_1 = x_0 - \alpha_0 \frac{\partial f(x_0)}{\partial x^{(1)}} e_1,$$

где

- $e_1 = (1, 0, \dots, 0)^T$  — единичный вектор вдоль первой координаты,
- $\alpha_0 > 0$  — шаг метода (л. 8, стр. 5).

## Общая итерация

На следующем шаге:

$$x_2 = x_1 - \alpha_1 \frac{\partial f(x_1)}{\partial x^{(2)}} e_2,$$

и т. д. В общем виде

$$x_{k+1} = x_k - \alpha_k \frac{\partial f(x_k)}{\partial x^{(s)}} e_s,$$

где  $s$  — номер координаты, по которой осуществляется спуск на данной итерации (л. 8, стр. 6).

## 6. Координатная форма метода

Если использовать координатную запись, то

$$x_{k+1}^{(i)} = \begin{cases} x_k^{(i)}, & i \neq s, \\ x_k^{(s)} - \alpha_k \frac{\partial f(x_k)}{\partial x^{(s)}}, & i = s. \end{cases}$$

Здесь на каждом шаге изменяется **ровно одна координата**, остальные остаются фиксированными (л. 8, стр. 7).

## 7. Выбор шага

Если шаг  $\alpha_k = \text{const}$ , то получается классический метод покоординатного спуска.

Если же  $\alpha_k$  выбирается из условия одномерной минимизации

$$\varphi(\alpha) = f\left(x_{k,s} - \alpha \frac{\partial f(x_{k,s})}{\partial x^{(s)}} e_s\right) \rightarrow \min,$$

то метод называется **покоординатным спуском с оптимальным шагом**. В этом случае он эквивалентен частному случаю метода Гаусса–Зейделя (л. 8, стр. 8).

## 8. Связь с методом Гаусса–Зейделя

Метод Гаусса–Зейделя можно рассматривать как частный случай релаксационного процесса, когда на каждой итерации корректируется **одна координата** решения системы, причём используется уже обновлённая информация (л. 8, стр. 8–9).

## 9. Основные свойства покоординатных методов

1. Метод относится к методам **первого порядка**, так как использует только первые производные функции.
2. По сравнению с классическим градиентным методом он может сходиться быстрее для задач большой размерности.
3. Реализация алгоритма существенно проще, так как на каждом шаге решается одномерная задача минимизации.
4. Метод применим для поиска

$$x^* = \arg \min_{x \in \mathbb{R}^n} f(x)$$

при сравнительно небольших вычислительных затратах (л. 8, стр. 9).

4. Так как каждая итерация в покоординатном спуске выполняется в окрестности текущей точки, то метод хорошо подходит для минимизации функций многих переменных, особенно в задачах большой размерности (л. 8, стр. 9).
5. Метод первоначально был разработан для решения систем линейных алгебраических уравнений

$$Ax = b, \quad A = (a_{ij})_{i,j=1}^n, \quad x \in \mathbb{R}^n, \quad b \in \mathbb{R}^n,$$

где матрица  $A$  является симметричной и положительно определённой (л. 8, стр. 10).

В этом случае задача решения системы  $Ax = b$  эквивалентна задаче минимизации квадратичной функции

$$\varphi(x) = \langle Ax, x \rangle - \langle b, x \rangle,$$

и точка минимума

$$x^* = \arg \min_{x \in \mathbb{R}^n} \varphi(x)$$

совпадает с решением системы  $Ax = b$  (л. 8, стр. 10).

## Сопряжённость и сопряжённые направления

### Определение.

Два вектора  $x, y \in \mathbb{R}^n$  называются  **$H$ -сопряжёнными** (или сопряжёнными относительно матрицы  $H$ ), если матрица  $H$  симметрична и выполнено условие

$$\langle x, Hy \rangle = 0. \quad (1)$$

Если  $H = E$  — единичная матрица, то условие сопряжённости переходит в обычную ортогональность:

$$\langle x, y \rangle = 0.$$

Таким образом, сопряжённость является обобщением понятия ортогональности (л. 8, стр. 11).

### Замечание.

В задачах минимизации квадратичных функций роль матрицы  $H$  обычно играет матрица Гессе  $f''(x)$ . В этом случае направления, сопряжённые относительно  $f''(x)$ , обладают важными оптимизационными свойствами.

## Пример (построение сопряжённых направлений)

Рассмотрим задачу минимизации

$$f(x) \rightarrow \min, \quad x \in \mathbb{R}^n,$$

где

$$f(x) = (x^{(1)})^2 + 4(x^{(2)})^2 - 1, \quad x = (x^{(1)}, x^{(2)})^T, \quad n = 2$$

(л. 8, стр. 12).

Пусть начальная точка

$$x_0 = (1, 1)^T,$$

а  $s_0$  — произвольное начальное направление поиска, например

$$s_0 = (1, 2)^T.$$

Тогда следующая точка определяется формулой

$$x_1 = x_0 + \lambda^{(0)} s_0, \quad (2)$$

где величина шага  $\lambda^{(0)}$  находится из условия одномерной минимизации

$$\left. \frac{d}{d\lambda} f(x_0 + \lambda s_0) \right|_{\lambda=\lambda^{(0)}} = 0. \quad (3)$$

## Нахождение шага $\lambda^{(0)}$

Вычислим градиент:



$$\nabla f(x) = \begin{pmatrix} 2x^{(1)} \\ 8x^{(2)} \end{pmatrix}.$$

Подставляя  $x = x_0 + \lambda s_0$ , получаем условие (3) в виде

$$\langle \nabla f(x_0 + \lambda s_0), s_0 \rangle = 0.$$

Так как

$$x_0 + \lambda s_0 = (1 + \lambda, 1 + 2\lambda),$$

то

$$\frac{\partial f}{\partial x^{(1)}} = 2(1 + \lambda), \quad \frac{\partial f}{\partial x^{(2)}} = 8(1 + 2\lambda).$$

Следовательно,

$$2(1 + \lambda) \cdot 1 + 8(1 + 2\lambda) \cdot 2 = 0,$$

откуда

$$\begin{aligned} 2(1 + \lambda) + 16(1 + 2\lambda) &= 0, \\ 18 + 34\lambda &= 0, \quad \lambda^{(0)} = -\frac{9}{17}. \end{aligned}$$

## Первая итерация

Подставляя  $\lambda^{(0)}$  в (2), получаем

$$x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{9}{17} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{8}{17} \\ -\frac{1}{17} \end{pmatrix}$$

(л. 8, стр. 15).

## Выбор сопряжённого направления

Далее требуется выбрать направление  $s_1$ , сопряжённое с  $s_0$  относительно матрицы Гессе

$$f''(x) = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}.$$

Условие сопряжённости:

$$\langle s_1, f''(x_1) s_0 \rangle = 0. \tag{4}$$

Пусть

$$s_1 = (s_1^{(1)}, s_1^{(2)})^T.$$

Тогда

$$f''(x_1)s_0 = (2, 16)^T,$$

и условие (4) принимает вид

$$2s_1^{(1)} + 16s_1^{(2)} = 0.$$

Выбирая, например,

$$s_1^{(1)} = 1, \quad s_1^{(2)} = -\frac{1}{8},$$

получаем сопряжённое направление (л. 8, стр. 16).

## Вторая итерация

Следующая точка вычисляется по формуле

$$x_2 = x_1 + \lambda^{(1)}s_1,$$

где шаг  $\lambda^{(1)}$  снова находится из условия

$$\left. \frac{d}{d\lambda} f(x_1 + \lambda s_1) \right|_{\lambda=\lambda^{(1)}} = 0.$$

В координатной форме:

$$x_2^{(1)} = \frac{8}{17} + \lambda^{(1)}, \quad x_2^{(2)} = -\frac{1}{17} - \frac{1}{8}\lambda^{(1)}.$$

Подстановка в условие оптимальности даёт

$$2\left(\frac{8}{17} + \lambda^{(1)}\right) + 8\left(-\frac{1}{17} - \frac{1}{8}\lambda^{(1)}\right) \cdot \left(-\frac{1}{8}\right) = 0,$$

откуда

$$\lambda^{(1)} = -\frac{8}{17}.$$

В результате

$$x_2 = (0, 0)^T,$$

что является точкой минимума функции  $f(x)$  (л. 8, стр. 17).

Таким образом, для квадратичной функции метод сопряжённых направлений приводит к точке минимума за конечное число шагов (не более  $n$ ), что иллюстрирует его высокую эффективность.

## Лекция 9. Метод сопряжённых направлений (продолжение)

Рассматривается задача минимизации квадратичной целевой функции

$$f(x) = a + \langle x, b \rangle + \frac{1}{2} \langle x, Hx \rangle, \quad x \in \mathbb{R}^n, \quad (9.1)$$

где  $a \in \mathbb{R}$ ,  $b \in \mathbb{R}^n$ ,  $H \in \mathbb{R}^{n \times n}$  — симметричная положительно определённая матрица (л. 9, стр. 1). Требуется найти решение задачи безусловной оптимизации

$$f(x) \rightarrow \min, \quad x \in \mathbb{R}^n,$$

причём предполагается, что функция имеет вид (9.1).

Пусть заданы  $n$  векторов  $s_0, \dots, s_{n-1} \in \mathbb{R}^n$ , образующих систему  $H$ -сопряжённых направлений, то есть

$$\langle s_i, Hs_j \rangle = 0, \quad i \neq j, \quad i, j = 0, \dots, n-1$$

(л. 9, стр. 1).

### Итерационный процесс метода сопряжённых направлений

Вводится итерационная схема

$$x_{k+1} = x_k + \alpha_k s_k, \quad (9.2)$$

где  $s_k$  — направление спуска на  $k$ -м шаге, а  $\alpha_k$  — величина шага, выбираемая из условия точного одномерного минимума:

$$\alpha_k = \arg \min_{\alpha \geq 0} \varphi(\alpha), \quad \varphi(\alpha) = f(x_k + \alpha s_k). \quad (9.3)$$

Из (9.1)–(9.3) непосредственно следует явная формула для шага:

$$\alpha_k = - \frac{\langle f'(x_k), s_k \rangle}{\langle s_k, Hs_k \rangle}. \quad (9.4)$$

Здесь градиент квадратичной функции равен

$$f'(x) = b + Hx.$$

Задав начальное приближение  $x_0 \in \mathbb{R}^n$ , из (9.2) и (9.4) получаем выражение для результата после  $n$  итераций:

$$x_n = x_0 + \sum_{k=0}^{n-1} \alpha_k s_k = x_0 - \sum_{k=0}^{n-1} \frac{\langle f'(x_k), s_k \rangle}{\langle s_k, Hs_k \rangle} s_k. \quad (9.5)$$

Покажем, что вектор  $x_n$ , определённый формулой (9.5), совпадает с точкой глобального минимума функции  $f(x)$  (л. 9, стр. 2–3).

## Точная точка минимума квадратичной функции

Стационарная (критическая) точка функции (9.1) определяется условием

$$f'(x) = 0.$$

Так как  $f'(x) = b + Hx$ , то единственная точка минимума имеет вид

$$x^* = -H^{-1}b. \quad (9.6)$$

Из положительной определённости  $H$  следует, что эта точка действительно является глобальным минимумом.

## Линейная независимость и разложение решения

**Лемма 9.1.** Векторы  $s_0, \dots, s_{n-1}$ , попарно  $H$ -сопряжённые, линейно независимы (л. 9, стр. 3).

Так как система  $\{s_i\}$  линейно независима, любой вектор, в частности разность  $x^* - x_0$ , допускает разложение

$$x^* - x_0 = \sum_{i=0}^{n-1} d_i s_i.$$

Отсюда

$$x^* = x_0 + \sum_{i=0}^{n-1} d_i s_i. \quad (9.7)$$

Введём обозначения:

$$\ell = (d_0, \dots, d_{n-1})^T \in \mathbb{R}^n,$$

$S = (s_0, \dots, s_{n-1}) \in \mathbb{R}^{n \times n}$  — матрица, столбцами которой являются векторы  $s_i$ . Тогда (9.7) можно переписать в матричном виде:

$$x^* = x_0 + S\ell. \quad (9.8)$$

## Вычисление коэффициентов разложения

Подставим (9.8) в условие стационарности  $f'(x^*) = 0$ :

$$f'(x^*) = b + H(x_0 + S\ell) = 0.$$

Отсюда

$$H(x_0 + S\ell) = -b, \quad HS\ell = -(Hx_0 + b).$$

Так как  $f'(x_0) = Hx_0 + b$ , получаем

$$HS\ell = -f'(x_0). \quad (9.9)$$

Домножая (9.9) слева на  $S^T$ , приходим к системе

$$(S^T HS)\ell = -S^T f'(x_0). \quad (9.10)$$

Матрица  $S^T H S$  симметрична и невырождена, поскольку  $H$  положительно определена, а векторы  $s_i$  линейно независимы. Следовательно, существует обратная матрица

$$(S^T H S)^{-1}. \quad (9.11)$$

Из (9.10) получаем явную формулу

$$\ell = -(S^T H S)^{-1} S^T f'(x_0). \quad (9.12)$$

## Диагональный вид матрицы и явные формулы

Так как векторы  $s_i$   $H$ -сопряжены, матрица  $S^T H S$  имеет диагональный вид:

$$S^T H S = \begin{pmatrix} d_0 & & 0 \\ & \ddots & \\ 0 & & d_{n-1} \end{pmatrix}, \quad d_i = \langle s_i, H s_i \rangle > 0. \quad (9.13)$$

Следовательно, её обратная матрица также диагональна:

$$d_i^{-1} = \frac{1}{\langle s_i, H s_i \rangle}. \quad (9.14)$$

Из (9.12) и (9.14) находим коэффициенты разложения:

$$d_i = -\frac{\langle f'(x_0), s_i \rangle}{\langle s_i, H s_i \rangle}, \quad i = 0, \dots, n-1. \quad (9.15)$$

Сравнивая формулы (9.15) и (9.4), видим, что при выборе  $x_0 = x_k$  коэффициенты  $d_i$  совпадают с шагами  $\alpha_k$ . Тем самым подтверждается, что итерационный процесс (9.2)–(9.4) приводит точно в точку минимума  $x^*$  за не более чем  $n$  шагов для квадратичной функции (9.1) (л. 9, стр. 7).

## Метод сопряжённых градиентов

Рассмотрим теперь квадратичную целевую функцию специального вида

$$f(x) = \langle b, x \rangle + \frac{1}{2} \langle A x, x \rangle, \quad (9.16)$$

где  $A \in \mathbb{R}^{n \times n}$  — симметричная положительно определённая матрица,  $b \in \mathbb{R}^n$ ,  $x \in \mathbb{R}^n$  (л. 9, стр. 8).

В методе сопряжённых направлений требуется заранее построить систему  $H$ -сопряжённых направлений. Метод сопряжённых градиентов позволяет формировать такие направления итерационно, используя информацию о градиентах, так что последовательная минимизация вдоль этих направлений приводит к решению задачи

$$f(x) \rightarrow \min, \quad x \in \mathbb{R}^n,$$

при любом начальном приближении  $x_0$ .

Итерационная схема записывается в виде

$$x_{k+1} = x_k + \alpha_k s_k, \quad (9.17)$$

где шаг  $\alpha_k$  выбирается из условия точного одномерного минимума:

$$f(x_k + \alpha_k s_k) = \min_{\alpha \in \mathbb{R}} f(x_k + \alpha s_k), \quad k = 0, 1, 2, \dots$$

Направления строятся по формулам

$$s_0 = -f'(x_0), \quad s_k = -f'(x_k) + \beta_{k-1} s_{k-1}, \quad k \geq 1. \quad (9.18)$$

Коэффициент  $\beta_{k-1}$  определяется из условия сопряжённости направлений  $s_{k-1}$  и  $s_k$ :

$$\langle s_k, A s_{k-1} \rangle = 0.$$

Подставляя (9.18), получаем

$$0 = \langle -f'(x_k) + \beta_{k-1} s_{k-1}, A s_{k-1} \rangle = -\langle f'(x_k), A s_{k-1} \rangle + \beta_{k-1} \langle s_{k-1}, A s_{k-1} \rangle,$$

откуда

$$\beta_{k-1} = \frac{\langle f'(x_k), A s_{k-1} \rangle}{\langle s_{k-1}, A s_{k-1} \rangle}. \quad (9.19)$$

Таким образом, метод (9.17)–(9.19) относится к классу методов сопряжённых направлений. Далее будут показаны его важные свойства, в частности ортогональность градиентов на последовательных шагах (л. 9, стр. 9–10).

Из (9.17)–(9.19) следует, что метод сопряжённых градиентов обладает рядом важных свойств.

**Лемма 9.2.** Векторы градиентов  $f'(x_k)$  и  $f'(x_{k+1})$  ортогональны:

$$\langle f'(x_k), f'(x_{k+1}) \rangle = 0, \quad k = 0, 1, 2, \dots$$

(л. 9, стр. 10).

**Задача.** Доказать лемму 9.2.

**Лемма 9.3.** Пусть  $x_0 \in \mathbb{R}^n$ , и последовательности  $x_0, \dots, x_n$  и  $s_0, \dots, s_{n-1}$  получены по формулам (9.17)–(9.19), причём  $f'(x_k) \neq 0$  при  $k = 0, 1, \dots, n-1$ . Тогда векторы  $s_0, \dots, s_{n-1}$  являются взаимно  $A$ -сопряжёнными, а градиенты  $f'(x_0), \dots, f'(x_{n-1})$  — взаимно ортогональными (л. 9, стр. 10).

Из лемм 9.2 и 9.3 следует, что метод сопряжённых градиентов действительно принадлежит к классу методов сопряжённых направлений. В частности, он обеспечивает нахождение минимума квадратичной функции (9.16) не более чем за  $n$  шагов.

## Переход к неквадратичным функциям

Построим теперь метод сопряжённых градиентов для минимизации неквадратичных гладких функций. Формально идея заключается в использовании той же схемы, что и для квадратичного случая, с заменой матрицы  $A$  на некоторую аппроксимацию, связанную с локальной кривизной функции.

Итерационная схема остаётся прежней:

$$x_{k+1} = x_k + \alpha_k s_k,$$

где шаг  $\alpha_k$  находится из условия одномерной минимизации:

$$\alpha_k = \arg \min_{\alpha \geq 0} f(x_k + \alpha s_k), \quad k = 0, 1, 2, \dots \quad (9.24)$$

Направления выбираются по формулам:

$$s_0 = -f'(x_0), \quad s_k = -f'(x_k) + \beta_{k-1} s_{k-1}, \quad k \geq 1,$$

где коэффициент  $\beta_{k-1}$  определяется следующим образом:

$$\beta_{k-1} = \begin{cases} \frac{\langle f'(x_k), f'(x_k) - f'(x_{k-1}) \rangle}{\|f'(x_{k-1})\|^2}, & k \notin \{n, 2n, 3n, \dots\}, \\ 0, & k \in \{n, 2n, 3n, \dots\}. \end{cases}$$

(л. 9, стр. 15–16).

При таком выборе периодически происходит «перезапуск» метода, то есть после каждых  $n$  итераций направление снова полагается равным антиградиенту. Это делается для повышения устойчивости алгоритма.

Можно показать, что при реализации метода (9.24) выполняется условие

$$\lim_{k \rightarrow \infty} \|f'(x_k)\| = 0, \quad (9.25)$$

если целевая функция  $f(x)$  ограничена снизу и её градиент удовлетворяет условию Липшица (л. 9, стр. 16).

Если, кроме того, функция сильно выпукла, то есть её гессиан ограничен и положительно определён, то метод (9.24) сходится к точке минимума  $x^*$  со сверхлинейной скоростью. В частности, имеет место оценка:

$$\|x_{k+n} - x^*\| \leq C \|x_k - x^*\|^2, \quad (9.26)$$

где  $k = 0, n, 2n, \dots$ , а константа  $C > 0$  не зависит от  $k$  (л. 9, стр. 17).

## Метод тяжёлого шарика

Отдельно рассматривается метод тяжёлого шарика. В градиентном методе на каждом шаге используется только информация о текущем градиенте. В методах с «памятью» новое приближение зависит от нескольких предыдущих шагов.

**Определение 9.1.** Методы оптимизации, в которых новое приближение  $x_{k+1}$  зависит от  $s$  предыдущих приближений,

$$x_{k+1} = \Phi(x_k, \dots, x_{k-s+1}), \quad s \geq 1, \quad (9.27)$$

называются  $s$ -шаговыми (л. 9, стр. 17).

Одним из двухшаговых методов оптимизации при  $s = 2$  является метод тяжёлого шарика. Его схема имеет вид:

$$x_{k+1} = x_k - \alpha f'(x_k) + \beta(x_k - x_{k-1}), \quad (9.28)$$

где  $\alpha > 0$  и  $\beta \geq 0$  — параметры метода. При  $\beta = 0$  схема (9.28) вырождается в обычный градиентный метод.

## Физическая интерпретация

Пусть  $f(x)$  — потенциальная энергия материальной точки («шарика»), движущейся по вязкой поверхности. Тогда уравнение движения с линейным трением имеет вид

$$\ddot{x}(t) = -f'(x(t)) - \mu \dot{x}(t), \quad (9.29)$$

где  $\mu > 0$  — коэффициент вязкости (л. 9, стр. 18).

Дискретный аналог уравнения (9.29) при соответствующем выборе шага по времени приводит именно к схеме метода тяжёлого шарика (9.28). Устойчивость этого уравнения соответствует наличию минимума потенциальной энергии, то есть решению задачи оптимизации.

## Сходимость метода тяжёлого шарика

Введение в схему (9.28) инерционного члена может существенно ускорить сходимость по сравнению с чисто градиентным методом, однако при неудачном выборе параметров возможны колебания и потеря устойчивости.

**Теорема 9.1.** Пусть  $x^*$  — невырожденная точка минимума функции  $f(x)$ , то есть

$$f''(x^*) > 0.$$

Пусть при некоторых  $l > 0, L > 0$  выполняется

$$lI \leq f''(x) \leq LI.$$

Тогда при параметрах

$$0 < \alpha < \frac{2(1+\beta)}{L}, \quad 0 \leq \beta < 1,$$

найдётся такое  $\varepsilon > 0$ , что для всех начальных  $x_0, x_1$  из  $\varepsilon$ -окрестности точки  $x^*$  метод (9.28) сходится к  $x^*$  с геометрической скоростью:

$$\|x_k - x^*\| \leq C(\delta)(q + \delta)^k, \quad q < 1, \delta > 0, \quad (9.30)$$

где  $C(\delta)$  — константа, не зависящая от  $k$  (л. 9, стр. 20).

Оптимальная величина коэффициента сжатия равна

$$q_0 = \frac{\sqrt{L} - \sqrt{l}}{\sqrt{L} + \sqrt{l}}, \quad (9.31)$$

и достигается при выборе

$$\alpha_0 = \frac{4}{(\sqrt{L} + \sqrt{l})^2}, \quad \beta_0 = q_0^2.$$



## Сравнение скоростей сходимости

Для одномерного градиентного метода оценка скорости сходимости имеет вид

$$q_1 \approx 1 - \frac{l}{2L},$$

где  $\frac{1}{2}l$  — параметр сильной выпуклости, а  $L$  — константа Липшица градиента.

Для двухшагового метода тяжёлого шарика получаем существенно лучшую оценку:

$$q_2 \approx 1 - 2\sqrt{\frac{l}{L}}.$$

Тем самым двухшаговый метод по порядку величины быстрее одношагового примерно в  $\sqrt{L/l}$  раз (л. 9, стр. 21).

### Задача (л. 9, стр. 10). Доказать лемму 9.2: $\langle f'(x_k), f'(x_{k+1}) \rangle = 0$

Рассматривается квадратичная функция

$$f(x) = \langle b, x \rangle + \frac{1}{2} \langle Ax, x \rangle, \quad A = A^T > 0, \quad (9.16)$$

и метод

$$x_{k+1} = x_k + \alpha_k s_k, \quad s_0 = -f'(x_0), \quad s_k = -f'(x_k) + \beta_{k-1} s_{k-1}, \quad (9.17), (9.18)$$

где  $\beta_{k-1}$  выбрано из условия  $A$ -сопряжённости  $\langle s_k, A s_{k-1} \rangle = 0$ , что даёт формулу (9.19) (л. 9, стр. 9–10).

Обозначим  $g_k := f'(x_k)$ . Для квадратичной  $f$  имеем

$$g(x) = Ax + b \Rightarrow g_{k+1} = g_k + \alpha_k A s_k. \quad (1)$$

#### Шаг 1. Условие точного одномерного минимума.

Так как  $\alpha_k$  выбирается из условия точной минимизации вдоль  $s_k$ ,

$$\alpha_k = \arg \min_{\alpha \in \mathbb{R}} f(x_k + \alpha s_k),$$

то производная по  $\alpha$  в точке  $\alpha = \alpha_k$  равна нулю:

$$0 = \frac{d}{d\alpha} f(x_k + \alpha s_k) \Big|_{\alpha=\alpha_k} = \langle f'(x_k + \alpha_k s_k), s_k \rangle = \langle g_{k+1}, s_k \rangle. \quad (2)$$

#### Шаг 2. Покажем, что $\langle g_{k+1}, s_{k-1} \rangle = 0$ .

Из (1):

$$\langle g_{k+1}, s_{k-1} \rangle = \langle g_k, s_{k-1} \rangle + \alpha_k \langle A s_k, s_{k-1} \rangle. \quad (3)$$

Первое слагаемое равно нулю по аналогичному условию точного минимума на предыдущем шаге:

$$0 = \frac{d}{d\alpha} f(x_{k-1} + \alpha s_{k-1}) \Big|_{\alpha=\alpha_{k-1}} = \langle g_k, s_{k-1} \rangle. \quad (4)$$

Второе слагаемое равно нулю по  $A$ -сопряжённости направлений  $s_k$  и  $s_{k-1}$ :

$$\langle As_k, s_{k-1} \rangle = \langle s_k, As_{k-1} \rangle = 0. \quad (5)$$

Из (3)–(5) следует

$$\langle g_{k+1}, s_{k-1} \rangle = 0. \quad (6)$$

### Шаг 3. Получим ортогональность градиентов.

Из (2) и формулы (9.18)  $s_k = -g_k + \beta_{k-1}s_{k-1}$  имеем:

$$0 = \langle g_{k+1}, s_k \rangle = \langle g_{k+1}, -g_k + \beta_{k-1}s_{k-1} \rangle = -\langle g_{k+1}, g_k \rangle + \beta_{k-1} \langle g_{k+1}, s_{k-1} \rangle.$$

По (6) второе слагаемое равно нулю, значит

$$\langle g_{k+1}, g_k \rangle = 0,$$

то есть

$$\langle f'(x_k), f'(x_{k+1}) \rangle = 0, \quad k = 0, 1, 2, \dots$$

Что и требовалось (л. 9, стр. 10).

### Задача (л. 9, стр. 21). Показать, что двухшаговый метод «в $\sqrt{L/l}$ раз быстрее» одномшагового (по порядку величины)

Пусть ошибка убывает геометрически:  $\|e_k\| \leq Cq^k$ , где  $q \in (0, 1)$ . Тогда для достижения  $\|e_k\| \leq \varepsilon$  достаточно

$$Cq^k \leq \varepsilon \iff k \geq \frac{\ln(\varepsilon/C)}{\ln q}. \quad (1)$$

Если  $q \approx 1$ , то  $\ln q = \ln(1 - (1 - q)) \approx -(1 - q)$ . Поэтому асимптотически

$$k \approx \frac{\ln(C/\varepsilon)}{1 - q}. \quad (2)$$

В конспекте приведены оценки (л. 9, стр. 21):

$$q_1 \approx 1 - \frac{l}{2L}, \quad q_2 \approx 1 - 2\sqrt{\frac{l}{L}}. \quad (3)$$

Отсюда

$$1 - q_1 \approx \frac{l}{2L}, \quad 1 - q_2 \approx 2\sqrt{\frac{l}{L}}. \quad (4)$$

Подставляя (4) в (2), получаем оценки числа итераций:

$$k_1 \approx \frac{\ln(C/\varepsilon)}{l/(2L)} = \frac{2L}{l} \ln \frac{C}{\varepsilon}, \quad k_2 \approx \frac{\ln(C/\varepsilon)}{2\sqrt{l/L}} = \frac{1}{2} \sqrt{\frac{L}{l}} \ln \frac{C}{\varepsilon}. \quad (5)$$

Следовательно, отношение требуемых итераций

$$\frac{k_1}{k_2} \approx \frac{\frac{2L}{l} \ln(C/\varepsilon)}{\frac{1}{2} \sqrt{L/l} \ln(C/\varepsilon)} = 4 \sqrt{\frac{L}{l}}. \quad (6)$$

То есть с точностью до константы метод с коэффициентом  $q_2$  требует примерно в  $\sqrt{L/l}$  раз меньше итераций, чем метод с  $q_1$ ; это и означает «быстрее по порядку величины в  $\sqrt{L/l}$  раз» (л. 9, стр. 21).

# Лекция 10. Метод Ньютона. Квазиньютоновские методы. Условная оптимизация

## Метод Ньютона

### Схема метода Ньютона.

Рассматривается задача минимизации гладкой функции  $f(x)$ ,  $x \in \mathbb{R}^n$ . Итерационная схема метода Ньютона имеет вид

$$x_{k+1} = x_k - (f''(x_k))^{-1} f'(x_k). \quad (10.1)$$

(л. 10, стр. 1)

Если ввести направление

$$s_k = -(f''(x_k))^{-1} f'(x_k), \quad (10.2)$$

и положить шаг  $\alpha_k = 1$ , то схема Ньютона переписывается в виде

$$x_{k+1} = x_k + \alpha_k s_k. \quad (10.3)$$

(л. 10, стр. 1)

Таким образом, метод Ньютона формально относится к классу **методов спуска**. Действительно, проверим условие спуска:

$$\langle f'(x_k), s_k \rangle < 0.$$

Имеем

$$\langle f'(x), s \rangle = - \langle f'(x), (f''(x))^{-1} f'(x) \rangle.$$

Если матрица Гессе  $f''(x)$  положительно определена, то

$$\langle f'(x), (f''(x))^{-1} f'(x) \rangle > 0,$$

следовательно,

$$\langle f'(x), s \rangle < 0,$$

и схема (10.1) действительно является методом спуска.

(л. 10, стр. 1-2)

## Ньютоновская длина

### Определение 10.1. Величину

$$\langle f'(x), (f''(x))^{-1} f'(x) \rangle^{1/2}$$

называют **ньютоновской длиной** функции  $f(x)$  в точке  $x \in \mathbb{R}^n$ .

(л. 10, стр. 2)

## Сильная выпуклость

Пусть функция  $f(x)$  **сильно выпукла**, то есть для любых  $x, y \in \mathbb{R}^n$  выполняются неравенства

$$f(y) \geq f(x) + \langle f'(x), y - x \rangle + \frac{G_2}{2} \|y - x\|^2, \quad (10.4)$$

$$\langle f'(x) - f'(y), x - y \rangle \geq G_2 \|x - y\|^2,$$

где  $G_2 > 0$  — параметр сильной выпуклости.

(л. 10, стр. 2)

Для дважды дифференцируемой функции условие сильной выпуклости эквивалентно неравенству

$$\langle f''(x)h, h \rangle \geq G_2 \|h\|^2, \quad \forall h \in \mathbb{R}^n. \quad (10.5)$$

(л. 10, стр. 3)

**Замечание.** Условия (10.4) и (10.5) являются необходимыми и достаточными. Кроме того, часто дополнительно предполагают ограниченность матрицы Гессе сверху:

$$m \|h\|^2 \leq \langle f''(x)h, h \rangle \leq M \|h\|^2, \quad (10.6)$$

где  $0 < m \leq M < +\infty$ .

(л. 10, стр. 3)

## Свойства метода Ньютона при сильной выпуклости

Из неравенства (10.6) для обратной матрицы Гессе следует:

$$\frac{1}{M} \|s\|^2 \leq \langle s, (f''(x))^{-1} s \rangle \leq \frac{1}{m} \|s\|^2. \quad (10.7)$$

(л. 10, стр. 4)

При сделанных предположениях:

- решение задачи оптимизации  $x^*$  существует и единственно;
- последовательность  $\{x_k\}$ , построенная по схеме Ньютона (10.1), является ограниченной и имеет предельные точки.

(л. 10, стр. 4)

**Теорема 10.1.** Пусть  $f(x)$  — сильно выпуклая дважды дифференцируемая функция, для которой выполнены условия (10.6) и (10.7). Тогда метод Ньютона (10.1) при  $\alpha_k = 1$  локально сходится к  $x^*$  со **сверхлинейной скоростью**.

(л. 10, стр. 4)

Если, кроме того, для любых  $x, y \in \mathbb{R}^n$  выполнено условие липшицевости второй производной

$$\|f''(x) - f''(y)\| \leq L\|x - y\|, \quad (10.8)$$

где  $L > 0$ , то метод Ньютона сходится к  $x^*$  с **квадратичной скоростью**.

(л. 10, стр. 5)

**Замечание.** Основной недостаток метода Ньютона — его локальная сходимость: для сходимости требуется, чтобы начальная точка  $x_0$  была достаточно близка к  $x^*$ . В частности, должно выполняться

$$\|x_0 - x^*\| < \bar{r},$$

где  $\bar{r}$  — некоторая константа (в конспекте указано приближённо  $\bar{r} \approx 1,392$ ).

(л. 10, стр. 5)

## Демпфированный метод Ньютона

**Схема метода.**

$$x_{k+1} = x_k - \alpha_k (f''(x_k))^{-1} f'(x_k). \quad (10.9)$$

(л. 10, стр. 5)

Здесь вводится **переменный шаг**  $\alpha_k$ , который выбирается по правилу Армихо:

$$0 < \alpha_k < \frac{1}{2}.$$

(л. 10, стр. 6)

**Теорема 10.2.** В условиях теоремы 10.1 демпфированный метод Ньютона с выбором шага по правилу Армихо сходится из любой начальной точки к единственному решению  $x^*$  со **сверхлинейной скоростью**. Если дополнительно выполнено условие липшицевости (10.8), то скорость сходимости становится **квадратичной**.

(л. 10, стр. 6)

## Квазиньютоновские методы

Квазиньютоновские методы близки к методу Ньютона, но **не требуют явного вычисления матрицы Гессе**  $f''(x)$ .

(л. 10, стр. 6–7)

**Основная идея.** Используются матрицы  $H_k$ , аппроксимирующие обратную матрицу Гессе:

$$H_k \approx (f''(x_k))^{-1}.$$

Рассматривается итерационный процесс

$$\begin{cases} x_{k+1} = x_k + \alpha_k s_k, \\ s_k = -H_k f'(x_k), \end{cases} \quad (10.10)$$

где  $H_k$  — симметричная положительно определённая матрица.

(л. 10, стр. 7)

Процесс (10.10) относится к методам спуска, поскольку

$$\langle f'(x_k), s_k \rangle = -\langle f'(x_k), H_k f'(x_k) \rangle < 0,$$

если  $f'(x_k) \neq 0$ .

(л. 10, стр. 7)

### Замечания.

1. Если  $H_k = E$ , то схема (10.10) совпадает с градиентным спуском с малым шагом.
2. Если  $f(x)$  дважды дифференцируема и выпукла, а матрица Гессе положительно определена, то при выборе

$$H_k = (f''(x_k))^{-1}$$

получаем метод Ньютона с переменным шагом.

(л. 10, стр. 7-8)

Особый интерес представляет класс методов, для которых

$$\lim_{k \rightarrow \infty} H_k = (f''(x^*))^{-1}. \quad (10.11)$$

(л. 10, стр. 8)

## Построение аппроксимации матриц $H_k$

Пусть  $H_0$  — произвольная симметричная положительно определённая матрица. Линейный процесс для матриц имеет вид

$$H_{k+1} = H_k + \Delta H_k, \quad (10.12)$$

где  $\Delta H_k$  — некоторая симметричная матрица.

(л. 10, стр. 9)

Рассмотрим разложение градиента:

$$f'(x_k) - f'(x_{k+1}) = f''(x_{k+1})(x_k - x_{k+1}) + O(\|x_k - x_{k+1}\|).$$

Положив

$$\Delta x_k = x_{k+1} - x_k, \quad \Delta y_k = f'(x_{k+1}) - f'(x_k),$$

получаем приближённое соотношение

$$\Delta x_k \approx (f''(x_{k+1}))^{-1} \Delta y_k.$$

Требуя, чтобы матрица  $H_{k+1}$  аппроксимировала  $(f''(x_{k+1}))^{-1}$ , приходим к **секущему условию**

$$\Delta x_k = H_{k+1} \Delta y_k. \quad (10.13)$$

(л. 10, стр. 9–10)

Подставляя (10.12) в (10.13), получаем

$$\Delta H_k \Delta y_k = \Delta x_k - H_k \Delta y_k. \quad (10.14)$$

(л. 10, стр. 10)

Дополнительно обычно требуют, чтобы:

- $\Delta H_k$  была симметричной;
- матрица  $H_{k+1}$  оставалась положительно определённой.

Система (10.14) является недоопределённой, поэтому существует множество квазиньютоновских методов, удовлетворяющих текущему условию.

(л. 10, стр. 10–11)

## Общая схема квазиньютоновских методов

Методы вида

$$\begin{cases} x_{k+1} = x_k + \alpha_k s_k, \\ s_k = -H_k f'(x_k), \\ H_{k+1} = H_k + \Delta H_k, \\ \Delta H_k \Delta y_k = \Delta x_k - H_k \Delta y_k, \end{cases} \quad (10.15)$$

называются **квазиньютоновскими методами**.

(л. 10, стр. 11)

## Метод Бройдена

В методе Бройдена матрица  $\Delta H_k$  выбирается так, что

$$\text{rank } \Delta H_k = 1.$$

Формула обновления имеет вид

$$\Delta H_k = \frac{(\Delta x_k - H_k \Delta y_k)(\Delta x_k - H_k \Delta y_k)^T}{\langle \Delta x_k - H_k \Delta y_k, \Delta y_k \rangle}, \quad (10.16)$$

где

$$\langle \Delta x_k - H_k \Delta y_k, \Delta y_k \rangle \neq 0.$$

(л. 10, стр. 11)

## Метод Дэвидона–Флетчера–Пауэлла (DFP)

В этом методе



$$\text{rank } \Delta H_k = 2.$$

Формула обновления:

$$\Delta H_k = \frac{\Delta x_k \Delta x_k^\top}{\langle \Delta x_k, \Delta y_k \rangle} - \frac{H_k \Delta y_k (H_k \Delta y_k)^\top}{\langle H_k \Delta y_k, \Delta y_k \rangle}. \quad (10.17)$$

(л. 10, стр. 12)

## Метод Бroyдена-Флетчера-Гольдфарба-Шанно (BFGS)

В рамках этого метода матрица  $\Delta H_k$  имеет представление

$$\Delta H_k = Q \bar{D} Q^\top, \quad (10.18)$$

где

$$\text{rank } \Delta H_k = 2, \quad Q = [q_1, q_2], \quad q_1, q_2 \in \mathbb{R}^n,$$

$$\bar{D} = \begin{pmatrix} a & c \\ c & b \end{pmatrix}.$$

(л. 10, стр. 12)

На этом месте изложение темы квазиньютоновских методов в рамках рассматриваемых страниц обрывается.

Формула для  $\Delta H_k$  в этом случае имеет вид

$$\Delta H_k = \frac{1}{\langle \Delta x_k, \Delta y_k \rangle} \left[ \beta \Delta x_k (\Delta x_k)^\top - H_k \Delta y_k (\Delta x_k)^\top - \Delta x_k (H_k \Delta y_k)^\top \right], \quad (10.19)$$

где

$$\beta = 1 + \frac{\langle H_k \Delta y_k, \Delta y_k \rangle}{\langle \Delta x_k, \Delta y_k \rangle}.$$

(л. 10, стр. 13)

Если ввести вектор

$$p_k = \sqrt{\langle \Delta y_k, H_k \Delta y_k \rangle} \left( \frac{\Delta x_k}{\langle \Delta x_k, \Delta y_k \rangle} - \frac{H_k \Delta y_k}{\langle \Delta y_k, H_k \Delta y_k \rangle} \right),$$

то формула (10.19) принимает более компактный вид:

$$\Delta H_k = \frac{\Delta x_k (\Delta x_k)^\top}{\langle \Delta x_k, \Delta y_k \rangle} - \frac{H_k \Delta y_k (H_k \Delta y_k)^\top}{\langle H_k \Delta y_k, \Delta y_k \rangle} + p_k p_k^\top. \quad (10.20)$$

(л. 10, стр. 13–14)

## Выводы о квазиньютоновских методах

1. Квазиньютоновские методы являются методами **первого порядка**, но обладают **сверхлинейной**, а при дополнительных условиях — **квадратичной скоростью сходимости**.
2. Доказано, что если шаг  $\alpha_k$  выбирается из условия достаточного уменьшения (по правилу Армико), то любой из рассмотренных квазиньютоновских методов находит минимум выпуклой квадратичной функции за конечное число шагов, не превосходящее размерность задачи.
3. Квазиньютоновские методы тесно связаны с методами переменной метрики.  
(л. 10, стр. 14)

## Связь с методами переменной метрики

Пусть задана симметричная положительно определённая матрица  $A$ . Тогда можно ввести скалярное произведение

$$\langle x, y \rangle_A = \langle x, Ay \rangle = \langle Ax, y \rangle, \quad (10.21)$$

где  $x, y \in \mathbb{R}^n$ ,  $A = A^T$ .

(л. 10, стр. 15)

Соответствующая метрика в пространстве  $\mathbb{R}^n$  имеет вид

$$\rho(x, y) = \|x - y\|_A = \sqrt{\langle x - y, x - y \rangle_A}. \quad (10.22)$$

(л. 10, стр. 15)

Линейную часть разложения функции  $f(x)$  в окрестности точки  $x$  можно записать как

$$\langle f'(x), \Delta x \rangle = \langle AA^{-1}f'(x), \Delta x \rangle = \langle A^{-1}f'(x), \Delta x \rangle_A.$$

(л. 10, стр. 15–16)

Следовательно, вектор  $A^{-1}f'(x)$  является градиентом функции  $f(x)$  в пространстве со скалярным произведением  $\langle \cdot, \cdot \rangle_A$ .

Итерационный процесс

$$x_{k+1} = x_k + \alpha_k s_k, \quad s_k = -H_k f'(x_k),$$

можно интерпретировать как **градиентный метод** в пространстве со скалярным произведением, задаваемым матрицей  $H_k^{-1}$ .

(л. 10, стр. 16)

Основная задача состоит в том, чтобы заменить матрицу Гессе на некоторую симметричную положительно определённую матрицу, которая:

- проще вычисляется;
- требует меньших затрат памяти;
- аппроксимирует оператор второго порядка так же, как и матрица Гессе.

(л. 10, стр. 16–17)

# Условная оптимизация

## Постановка задачи условной оптимизации

Рассматривается задача условной оптимизации

$$f_0(x) \rightarrow \min, \quad x \in Q, \quad (10.23)$$

где

$$Q = \{x \in \mathbb{R}^n \mid f_i(x) \leq 0, i = 1, \dots, m\}.$$

Предполагается, что все функции  $f_i(x)$  гладкие на множестве  $Q$ .

(л. 10, стр. 17)

**Идея.** Приближённое решение задачи (10.23) строится путём последовательного решения задач безусловной оптимизации.

Выделяют две основные группы методов:

- методы штрафных функций;
  - барьерные методы.
- (л. 10, стр. 18)

## Методы штрафных функций

**Определение 10.2.** Непрерывная функция  $\Phi(x)$  называется **штрафной функцией** для замкнутого множества  $Q \subset \mathbb{R}^n$ , если:

$$\Phi(x) = 0 \quad \forall x \in Q,$$

$$\Phi(x) > 0 \quad \forall x \notin Q.$$

(л. 10, стр. 18)

### Свойства штрафных функций.

Пусть:

- $\Phi_1(x)$  — штрафная функция для множества  $Q_1$ ;
- $\Phi_2(x)$  — штрафная функция для множества  $Q_2$ .

Тогда функция

$$\Phi_1(x) + \Phi_2(x)$$

является штрафной функцией для множества  $Q_1 \cap Q_2$ .

(л. 10, стр. 18–19)

Примером штрафной функции для множества  $Q$  из (10.23) является **квадратичная штрафная функция**

$$\Phi(x) = \sum_{i=1}^m (f_i(x))^2. \quad (10.24)$$

(л. 10, стр. 19)

## Общая схема метода штрафных функций

**Шаг 0.** Выбирается начальная точка  $x_0 \in \mathbb{R}^n$  и возрастающая последовательность штрафных коэффициентов  $\{t_k\}$ :

$$0 < t_1 < \dots < t_k < t_{k+1} < \dots$$

**Шаг  $k$ .** Находится

$$x_{k+1} = \arg \min_{x \in \mathbb{R}^n} \{f_0(x) + t_k \Phi(x)\},$$

используя точку  $x_k$  в качестве начального приближения.

(л. 10, стр. 19)

## Теорема о сходимости метода штрафных функций

**Теорема 10.3.** Пусть существует число  $\bar{t}$  такое, что множество

$$S_{\bar{t}} = \{x \in \mathbb{R}^n \mid f_0(x) + \bar{t} \Phi(x) \leq f_0(x^*)\}$$

ограничено, где  $x^*$  — решение задачи (10.23). Тогда:

$$\lim_{k \rightarrow \infty} f_0(x_k) = f_0(x^*), \quad \lim_{k \rightarrow \infty} \Phi(x_k) = 0.$$

(л. 10, стр. 20)

Данный результат носит **теоретический характер**. На практике возникают следующие трудности:

1. выбор конкретного вида штрафной функции;
2. выбор последовательности штрафных коэффициентов;
3. точность решения вспомогательных задач безусловной оптимизации.

(л. 10, стр. 20)

## Барьерные методы

**Определение 10.3.** Пусть  $Q$  — замкнутое множество,  $Q \neq \emptyset$ . Непрерывная функция  $F(x)$  называется **барьерной функцией** (барьером) для множества  $Q$ , если

$$F(x) \rightarrow +\infty \quad \text{при } x \rightarrow \partial Q.$$

(л. 10, стр. 21)

## Свойства барьеров.

Если:

- $F_1(x)$  — барьер для  $Q_1$ ;
- $F_2(x)$  — барьер для  $Q_2$ ,

то функция

$$F_1(x) + F_2(x)$$

является барьером для множества  $Q_1 \cap Q_2$ .

(л. 10, стр. 21)

Для применения барьерных методов к задаче (10.23) требуется выполнение условия существования **строго допустимой точки**:

$$\exists \bar{x} \in \mathbb{R}^n : f_i(\bar{x}) < 0, \quad i = 1, \dots, m. \quad (10.25)$$

(л. 10, стр. 21)

## Примеры барьерных функций

Пусть

$$Q = \{x \in \mathbb{R}^n \mid f_i(x) \leq 0, \quad i = 1, \dots, m\}.$$

Тогда следующие функции являются барьерами:

### 1. Степенной барьер

$$F(x) = \sum_{i=1}^m \frac{1}{(-f_i(x))^p}, \quad p \geq 1.$$

### 2. Логарифмический барьер

$$F(x) = - \sum_{i=1}^m \ln(-f_i(x)).$$

### 3. Экспоненциальный барьер

$$F(x) = \sum_{i=1}^m \exp\left(\frac{1}{-f_i(x)}\right).$$

(л. 10, стр. 22)

## Схема барьерного метода

**Шаг 0.** Выбирается начальная точка

$$x_0 \in \text{int } Q,$$

и убывающая последовательность барьерных коэффициентов:

$$0 < t_1 < \dots < t_k < t_{k+1} < \dots$$

**Шаг  $k$ .** Находится

$$x_{k+1} = \arg \min_{x \in Q} \left\{ f_0(x) + \frac{1}{t_k} F(x) \right\},$$

используя  $x_k$  в качестве начальной точки.

(л. 10, стр. 23)

## Теорема о сходимости барьерного метода

**Теорема 10.4.** Пусть барьерная функция  $F(x)$  ограничена снизу на множестве  $Q$ . Тогда

$$\lim_{k \rightarrow \infty} x_k = x^*, \quad \lim_{k \rightarrow \infty} H_k^* = f_0^*,$$

где

$$H_k(x) = f_0(x) + \frac{1}{t_k} F(x), \quad H_k^* = \min_{x \in Q} H_k(x), \quad f_0^* = \min_{x \in Q} f_0(x).$$

(л. 10, стр. 23)