

# Лекция 6

Цуя Yaroshevskiy

13 мая 2023 г.

## Содержание

<b>1</b>	<b>Постановка задачи</b>	<b>1</b>
1.1	Свойства квадратичных форм . . . . .	2
1.2	Свойства выпуклых множеств и выпуклых функций . . . . .	2
1.3	Необходимое и достаточное условие безусловного экстремума . . . . .	3
1.3.1	Проверка выполнения условий . . . . .	3

## 1 Постановка задачи

- $x^* = (x_1, x_2, \dots, x_n)^T$ ,  $x_i \in U \subset E_n$ , где  $U$  — множество допустимых значений,  $E_n$  — евклидово пространство размера  $n$ .  $f(x^*) = \min_{x \in U} f(x)$ . Если ставится задача найти максимум, то можно перейти к поиску минимума:  $f(x^*) = \max_{x \in U} f(x) = -\min_{x \in U} (-f(x))$
- $f(x^*) = \text{extr}_{x \in U} f(x)$
- Если  $U$  задается ограничением на вектор  $x$ , то задача поиска условного экстремума. Если  $U = E_n$  — не имеет ограничений, то задача поиска безусловного экстремума
- Решение задачи поиска экстремума — пара  $(x^*, f(x^*))$

---

Если  $\forall x \in U f(x^*) \leq f(x)$  — то  $x^*$  — глобальный минимум. Локальный минимум  $x^* \in U$ : если  $\exists \varepsilon > 0$ , что  $\forall x \in U$  и  $\|x - x^*\| < \varepsilon$ , то  $f(x^*) \leq f(x)$

**Определение. Поверхностью уровня** функции  $f(x)$  называется множество точек, в которых функция принимает постоянные значения, т.е.  $f(x) = \text{const}$

**Определение. Градиентом**  $\nabla f(x)$  непрерывно дифференцируемой функции  $f(x)$  в  $x$ :

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}$$

Градиент направлен по нормали к поверхности уровня, т.е. перпендикулярно к касательной плоскости в точке  $x$ , проведенной в сторону наибольшего возрастания функции

**Определение. Матрицей Гессе**  $H(x)$  дважды непрерывно дифференцируемой в точке  $x$  функции  $f(x)$  называется матрица частных производных второго порядка, вычисленных в данной точке.

$$H(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n x_1} & \frac{\partial^2 f(x)}{\partial x_n x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$

- $H(x)$  — симметричная, размер  $nn$
- Антиградиент: вектор, равный по модулю вектору градиента, но противоположный по направлению. Указывает в сторону наибольшего убывания функции  $f(x)$
- 

$$\nabla f(x) = f(x + \Delta x) - f(x) = \nabla f(x)^T \Delta x + \frac{1}{2} \Delta x^T H(x) \Delta x + o(\|\Delta x\|^2)$$

$o(\|\Delta x\|^2)$  — сумма всех членов разложения, имеющих порядок выше второго,  $\Delta x^T H(x) \Delta x$  — квадратичная форма

## 1.1 Свойства квадратичных форм

Квадратичная форма  $\Delta x^T H(x) \Delta x$  (и соответствующая матрица  $H(x)$ ) называется:

- положительно определенной  $H(x) > 0$ , если  $\forall \Delta x \neq 0 \Delta x^T H(x) \Delta x > 0$
- отрицательно определенной  $H(x) < 0$ , если  $\forall \Delta x \neq 0 \Delta x^T H(x) \Delta x < 0$
- положительно полуопределенной  $H(x) \geq 0$ , если  $\forall \Delta x \neq 0 \Delta x^T H(x) \Delta x \geq 0$  и имеется  $\Delta x \neq 0 : \Delta x^T H(x) \Delta x = 0$
- отрицательно полуопределенной  $H(x) \leq 0$ , если  $\forall \Delta x \neq 0 \Delta x^T H(x) \Delta x \leq 0$  и имеется  $\Delta x \neq 0 : \Delta x^T H(x) \Delta x = 0$
- неопределенной, если  $\exists \Delta x, \Delta \tilde{x} : \Delta x^T H(x) \Delta x > 0, \Delta \tilde{x}^T H(\tilde{x}) \Delta \tilde{x} < 0$
- тождественно равной нулю  $H(x) \equiv 0$ , если  $\forall \Delta x \Delta x^T H(x) \Delta x = 0$

## 1.2 Свойства выпуклых множеств и выпуклых функций

**Определение.** Пусть  $x, y \in E_n$ . Множество точек вида  $\{z\} \subset E_n : z = \alpha x + (1 - \alpha)y, \alpha \in [0, 1]$ ,  $z$  — отрезок, соединяющий  $x$  и  $y$ .

*Пример.*  $E_n : n \leq 3$ :  $z$  — отрезок (обычный)

**Определение.**  $U \subset E_n$  выпуклое, если вместе с точками  $x$  и  $(y \in U)$  оно содержит и весь отрезок  $z = \alpha x + (1 - \alpha)y, \alpha \in [0, 1]$

**Определение.** Функция  $f(x)$ , заданная на выпуклом  $U \subset E_n$  называется:

- выпуклой, если  $\forall x, y \in U$  и  $\forall \alpha \in [0, 1]$  выполняется  $f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$
- строго выпуклой, если  $\forall \alpha \in (0, 1)$  выполняется  $f(\alpha x + (1 - \alpha)y) < \alpha f(x) + (1 - \alpha)f(y)$
- сильно выпуклой с константой  $l > 0$ , если  $\forall x, y \in U$  и  $\forall \alpha \in [0, 1]$  выполняется  $f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) - \frac{l}{2} \alpha(1 - \alpha) \|x - y\|^2$

*Свойства:*

1. Функция  $f(x)$  выпуклая, если ее график целиком лежит не выше отрезка, соединяющего две ее произвольные точки  
Функция  $f(x)$  строго выпуклая, если ее график лежит целиком ниже отрезка, соединяющего две ее произвольные, но не совпадающие точки
2. Если функция  $f(x)$  сильно выпуклая, то она одновременно строго выпуклая и выпуклая  
Если функция  $f(x)$  строго выпуклая, то она одновременно выпуклая
3. Выпуклость функции можно определить по матрице Гессе  $H(x)$ 
  - Если  $H(x) \geq 0 \forall x \in E_n$ , то  $f(x)$  выпуклая
  - Если  $H(x) > 0 \forall x \in E_n$ , то  $f(x)$  строго выпуклая
  - Если  $H(x) \geq lE \forall x \in E_n$ , где  $E$  — единичная матрица, то  $f(x)$  сильно выпуклая

*Свойства выпуклых функций:*

1. Если  $f(x)$  выпуклая функция на выпуклом множестве  $U$ , то всякая точка локального минимума есть точка глобального минимума на  $U$
2. Если выпуклая функция достигает своего минимума в двух различных точках, то она достигает минимума во всех точках отрезка, соединяющих эти точки.
3. Если  $f(x)$  строго выпуклая функция множества  $U$ , то она может достигать своего глобального минимума на  $U$  не более чем в одной точке

### 1.3 Необходимое и достаточное условие безусловного экстремума

**Теорема 1.1** (Необходимое условие экстремума первого порядка). Пусть  $x^* \in E_n$  — локальный минимум или максимум  $f(x)$  на  $E_n$  и  $f(x)$  — дифференцируема в точке  $x^*$

Тогда  $\nabla f(x)$  в точке  $x^*$  равен нулю  $\nabla f(x^*) = 0$ , т.е.

$$\frac{\partial f(x^*)}{\partial x_i} = 0, \quad i = \overline{1, n}$$

**Определение.** Точки  $x^* : \nabla f(x^*) = 0$  — **стационарные**

**Теорема 1.2** (Необходимое условие экстремума второго порядка). Пусть  $x^* \in E_n$  — точка локального минимума или максимума  $f(x)$  на  $E_n$  и  $f(x)$  — дважды дифференцируемая в точке.

Тогда  $H(x^*)$  — является положительно или отрицательно (если максимум) полуопределенной, т.е.  $H(x^*) \geq 0$  или  $H(x^*) \leq 0$  (если максимум)

**Теорема 1.3** (Достаточное условие экстремума). Пусть  $f(x)$  в  $x^* \in E_n$  дважды дифференцируема, ее  $\nabla f(x) = 0$ , а  $H(x^*) > 0$  или  $H(x^*) < 0$  (для максимума).

Тогда  $x^*$  — точка локального минимума (максимума)  $f(x)$  на  $E_n$

#### 1.3.1 Проверка выполнения условий

- вычисление угловых миноров  $H(x)$
- вычисление главных миноров  $H(x)$

1. Исследование положительной или отрицательной определенности угловых и главных миноров
2. Анализ собственных значений матрицы  $H(x)$