

# Методы оптимизации

## Лекция 2: Сопряжённые конусы. Отделимость. Выпуклые функции

Александр Катруца

Физтех-школа прикладной математики и информатики  
Московский физико-технический институт



14 сентября 2020 г.

## На прошлой лекции

- ▶ О чём этот курс и почему он нужен
- ▶ Примеры постановок задач оптимизации
- ▶ Выпуклые множества и их свойства

# Напоминание: конусы

## Определение

*Множество  $K$  называется конусом, если для любого  $x \in K$  и произвольного числа  $\theta \geq 0$  выполнено  $\theta x \in K$ .*

# Напоминание: конусы

## Определение

Множество  $K$  называется конусом, если для любого  $x \in K$  и произвольного числа  $\theta \geq 0$  выполнено  $\theta x \in K$ .

## Определение

Множество  $K$  называется **выпуклым** конусом, если для любых точек  $x_1, x_2 \in K$  и любых чисел  $\theta_1 \geq 0, \theta_2 \geq 0$  выполнено  $\theta_1 x_1 + \theta_2 x_2 \in K$ .

# Напоминание: конусы

## Определение

Множество  $K$  называется конусом, если для любого  $\mathbf{x} \in K$  и произвольного числа  $\theta \geq 0$  выполнено  $\theta\mathbf{x} \in K$ .

## Определение

Множество  $K$  называется **выпуклым** конусом, если для любых точек  $\mathbf{x}_1, \mathbf{x}_2 \in K$  и любых чисел  $\theta_1 \geq 0, \theta_2 \geq 0$  выполнено  $\theta_1\mathbf{x}_1 + \theta_2\mathbf{x}_2 \in K$ .

## Важные конусы

- ▶ Неотрицательный октант  
 $\mathbb{R}_+^n = \{\mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, i = 1, \dots, n\} \rightarrow \text{LP}$
- ▶ Конус второго порядка  $\{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq t\} \rightarrow \text{SOCP}$
- ▶ Конус симметричных положительно полуопределённых матриц  $\mathbf{S}_+^n \rightarrow \text{SDP}$

# Сопряжённый конус (dual cone)

## Определение

Пусть  $K$  — конус. Тогда множество

$$K^* = \{\mathbf{y} \mid \langle \mathbf{y}, \mathbf{x} \rangle \geq 0, \mathbf{x} \in K\}$$

называется сопряжённым конусом.

# Сопряжённый конус (dual cone)

## Определение

Пусть  $K$  — конус. Тогда множество

$$K^* = \{\mathbf{y} \mid \langle \mathbf{y}, \mathbf{x} \rangle \geq 0, \mathbf{x} \in K\}$$

называется сопряжённым конусом.

## Свойства

- ▶  $K^*$  — конус
- ▶  $K^*$  — выпуклый конус для любого конуса  $K$
- ▶ Если  $K_1 \subseteq K_2$ , то  $K_2^* \subseteq K_1^*$

# Сопряжённый конус (dual cone)

## Определение

Пусть  $K$  — конус. Тогда множество

$$K^* = \{\mathbf{y} \mid \langle \mathbf{y}, \mathbf{x} \rangle \geq 0, \mathbf{x} \in K\}$$

называется сопряжённым конусом.

## Свойства

- ▶  $K^*$  — конус
- ▶  $K^*$  — выпуклый конус для любого конуса  $K$
- ▶ Если  $K_1 \subseteq K_2$ , то  $K_2^* \subseteq K_1^*$

## Определение

Если  $K = K^*$ , то конус называется самосопряжённым (self-dual)



# Самосопряжённые конусы

# Самосопряжённые конусы

## Определение

Сопряжённой нормой относительно  $\|\cdot\|$  называется

$$\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \leq 1} \mathbf{z}^\top \mathbf{x}.$$

# Самосопряжённые конусы

## Определение

Сопряжённой нормой относительно  $\|\cdot\|$  называется

$$\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \leq 1} \mathbf{z}^\top \mathbf{x}.$$

## Примеры

- ▶  $\|\cdot\|_1 \rightarrow \|\cdot\|_* = \|\cdot\|_\infty$
- ▶  $\|\cdot\|_2 \rightarrow \|\cdot\|_* = \|\cdot\|_2$
- ▶  $\|\cdot\|_\infty \rightarrow \|\cdot\|_* = \|\cdot\|_1$

# Самосопряжённые конусы

## Определение

Сопряжённой нормой относительно  $\|\cdot\|$  называется

$$\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \leq 1} \mathbf{z}^\top \mathbf{x}.$$

## Примеры

- ▶  $\|\cdot\|_1 \rightarrow \|\cdot\|_* = \|\cdot\|_\infty$
- ▶  $\|\cdot\|_2 \rightarrow \|\cdot\|_* = \|\cdot\|_2$
- ▶  $\|\cdot\|_\infty \rightarrow \|\cdot\|_* = \|\cdot\|_1$

## Самосопряжённые конусы

- ▶  $\mathbb{R}_+^n$

# Самосопряжённые конусы

## Определение

Сопряжённой нормой относительно  $\|\cdot\|$  называется

$$\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \leq 1} \mathbf{z}^\top \mathbf{x}.$$

## Примеры

- ▶  $\|\cdot\|_1 \rightarrow \|\cdot\|_* = \|\cdot\|_\infty$
- ▶  $\|\cdot\|_2 \rightarrow \|\cdot\|_* = \|\cdot\|_2$
- ▶  $\|\cdot\|_\infty \rightarrow \|\cdot\|_* = \|\cdot\|_1$

## Самосопряжённые конусы

- ▶  $\mathbb{R}_+^n$
- ▶ Конус второго порядка  $\{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq t\}$

# Самосопряжённые конусы

## Определение

Сопряжённой нормой относительно  $\|\cdot\|$  называется

$$\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \leq 1} \mathbf{z}^\top \mathbf{x}.$$

## Примеры

- ▶  $\|\cdot\|_1 \rightarrow \|\cdot\|_* = \|\cdot\|_\infty$
- ▶  $\|\cdot\|_2 \rightarrow \|\cdot\|_* = \|\cdot\|_2$
- ▶  $\|\cdot\|_\infty \rightarrow \|\cdot\|_* = \|\cdot\|_1$

## Самосопряжённые конусы

- ▶  $\mathbb{R}_+^n$
- ▶ Конус второго порядка  $\{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq t\}$
- ▶  $\mathbf{S}_+^n$

Зачем нужны конусы?

## Зачем нужны конусы?

Обобщённое отношение частичного порядка

Пусть  $K$  — конус. Тогда  $x \leq_K y \Leftrightarrow y - x \in K$ .



## Зачем нужны конусы?

Обобщённое отношение частичного порядка

Пусть  $K$  — конус. Тогда  $\mathbf{x} \leq_K \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in K$ .

Пример: конус  $\mathbf{S}_+^n$

Пусть  $\mathbf{X}, \mathbf{Y} \in \mathbf{S}^n$ . Тогда  $\mathbf{X} \leq_K \mathbf{Y}$  означает, что  $\mathbf{Y} - \mathbf{X} \in \mathbf{S}_+^n$

# Зачем нужны конусы?

Обобщённое отношение частичного порядка

Пусть  $K$  — конус. Тогда  $\mathbf{x} \leq_K \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in K$ .

Пример: конус  $\mathbf{S}_+^n$

Пусть  $\mathbf{X}, \mathbf{Y} \in \mathbf{S}^n$ . Тогда  $\mathbf{X} \leq_K \mathbf{Y}$  означает, что  $\mathbf{Y} - \mathbf{X} \in \mathbf{S}_+^n$

Задача линейного программирования

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \\ & x_i \geq 0 \end{aligned}$$

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{x} \in \mathbb{R}_+^n \end{aligned}$$

# Зачем нужны конусы?

## Обобщённое отношение частичного порядка

Пусть  $K$  — конус. Тогда  $\mathbf{x} \leq_K \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in K$ .

Пример: конус  $\mathbf{S}_+^n$

Пусть  $\mathbf{X}, \mathbf{Y} \in \mathbf{S}^n$ . Тогда  $\mathbf{X} \leq_K \mathbf{Y}$  означает, что  $\mathbf{Y} - \mathbf{X} \in \mathbf{S}_+^n$

## Задача линейного программирования

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \\ & x_i \geq 0 \end{aligned}$$

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{x} \in \mathbb{R}_+^n \end{aligned}$$

## Введение нелинейности

Использование декартового произведения трёх самосопряжённых конусов позволяет записать многие практически важные выпуклые задачи

# Отделимость выпуклых множеств

## Определение

Множества  $A, B$  называются *отделимыми*, если существует вектор  $\mathbf{a} \neq 0$  и число  $b$  такие что

- ▶  $\mathbf{a}^\top \mathbf{x} + b \geq 0$  для всех  $\mathbf{x} \in A$
- ▶  $\mathbf{a}^\top \mathbf{y} + b \leq 0$  для всех  $\mathbf{y} \in B$ .

# Отделимость выпуклых множеств

## Определение

Множества  $A, B$  называются *отделимыми*, если существует вектор  $\mathbf{a} \neq 0$  и число  $b$  такие что

- ▶  $\mathbf{a}^\top \mathbf{x} + b \geq 0$  для всех  $\mathbf{x} \in A$
- ▶  $\mathbf{a}^\top \mathbf{y} + b \leq 0$  для всех  $\mathbf{y} \in B$ .

## Теорема

Пусть  $A$  и  $B$  — выпуклые и непересекающиеся множества. Тогда существует разделяющая их гиперплоскость.

## Доказательство

# Отделимость выпуклых множеств

## Определение

Множества  $A, B$  называются *отделимыми*, если существует вектор  $\mathbf{a} \neq 0$  и число  $b$  такие что

- ▶  $\mathbf{a}^\top \mathbf{x} + b \geq 0$  для всех  $\mathbf{x} \in A$
- ▶  $\mathbf{a}^\top \mathbf{y} + b \leq 0$  для всех  $\mathbf{y} \in B$ .

## Теорема

Пусть  $A$  и  $B$  — выпуклые и непересекающиеся множества. Тогда существует разделяющая их гиперплоскость.

## Доказательство

- ▶ Предположим, что расстояние между  $A$  и  $B$  положительно:  $\inf_{\mathbf{x} \in A, \mathbf{y} \in B} \|\mathbf{x} - \mathbf{y}\|_2 > 0$

# Отделимость выпуклых множеств

## Определение

Множества  $A, B$  называются *отделимыми*, если существует вектор  $\mathbf{a} \neq 0$  и число  $b$  такие что

- ▶  $\mathbf{a}^\top \mathbf{x} + b \geq 0$  для всех  $\mathbf{x} \in A$
- ▶  $\mathbf{a}^\top \mathbf{y} + b \leq 0$  для всех  $\mathbf{y} \in B$ .

## Теорема

Пусть  $A$  и  $B$  — выпуклые и непересекающиеся множества. Тогда существует разделяющая их гиперплоскость.

## Доказательство

- ▶ Предположим, что расстояние между  $A$  и  $B$  положительно:  $\inf_{\mathbf{x} \in A, \mathbf{y} \in B} \|\mathbf{x} - \mathbf{y}\|_2 > 0$
- ▶ Пусть  $\mathbf{c} \in A$  и  $\mathbf{d} \in B$  — точки, в которых этот минимум достигается

# Отделимость выпуклых множеств

## Определение

Множества  $A, B$  называются *отделимыми*, если существует вектор  $\mathbf{a} \neq 0$  и число  $b$  такие что

- ▶  $\mathbf{a}^\top \mathbf{x} + b \geq 0$  для всех  $\mathbf{x} \in A$
- ▶  $\mathbf{a}^\top \mathbf{y} + b \leq 0$  для всех  $\mathbf{y} \in B$ .

## Теорема

Пусть  $A$  и  $B$  — выпуклые и непересекающиеся множества. Тогда существует разделяющая их гиперплоскость.

## Доказательство

- ▶ Предположим, что расстояние между  $A$  и  $B$  положительно:  $\inf_{\mathbf{x} \in A, \mathbf{y} \in B} \|\mathbf{x} - \mathbf{y}\|_2 > 0$
- ▶ Пусть  $\mathbf{c} \in A$  и  $\mathbf{d} \in B$  — точки, в которых этот минимум достигается
- ▶ Рассмотрим  $f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} + b$ , где  $\mathbf{a} = \mathbf{d} - \mathbf{c}$  и  $b = \frac{\|\mathbf{d}\|_2^2 - \|\mathbf{c}\|_2^2}{2}$



- Покажем, что  $f(\mathbf{y}) \geq 0$  для всех  $\mathbf{y} \in B$

- ▶ Покажем, что  $f(\mathbf{y}) \geq 0$  для всех  $\mathbf{y} \in B$
- ▶ Пусть найдётся  $\mathbf{u} \in B$  такая что  $f(\mathbf{u}) < 0$

- ▶ Покажем, что  $f(\mathbf{y}) \geq 0$  для всех  $\mathbf{y} \in B$
- ▶ Пусть найдётся  $\mathbf{u} \in B$  такая что  $f(\mathbf{u}) < 0$
- ▶  $f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \frac{1}{2}(\mathbf{d} + \mathbf{c})) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) + \frac{1}{2}\|\mathbf{d} - \mathbf{c}\|_2^2$

- ▶ Покажем, что  $f(\mathbf{y}) \geq 0$  для всех  $\mathbf{y} \in B$
- ▶ Пусть найдётся  $\mathbf{u} \in B$  такая что  $f(\mathbf{u}) < 0$
- ▶  $f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \frac{1}{2}(\mathbf{d} + \mathbf{c})) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) + \frac{1}{2}\|\mathbf{d} - \mathbf{c}\|_2^2$
- ▶  $(\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) < 0$

- ▶ Покажем, что  $f(\mathbf{y}) \geq 0$  для всех  $\mathbf{y} \in B$
- ▶ Пусть найдётся  $\mathbf{u} \in B$  такая что  $f(\mathbf{u}) < 0$
- ▶  $f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \frac{1}{2}(\mathbf{d} + \mathbf{c})) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) + \frac{1}{2}\|\mathbf{d} - \mathbf{c}\|_2^2$
- ▶  $(\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) < 0$
- ▶ Заметим, что

$$\left. \frac{d}{dt} \|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2^2 \right|_{t=0} = 2(\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) < 0$$

а значит для  $t \in (0, 1]$

$$\|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2 \leq \|\mathbf{d} - \mathbf{c}\|_2.$$

- ▶ Покажем, что  $f(\mathbf{y}) \geq 0$  для всех  $\mathbf{y} \in B$
- ▶ Пусть найдётся  $\mathbf{u} \in B$  такая что  $f(\mathbf{u}) < 0$
- ▶  $f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \frac{1}{2}(\mathbf{d} + \mathbf{c})) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) + \frac{1}{2}\|\mathbf{d} - \mathbf{c}\|_2^2$
- ▶  $(\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) < 0$
- ▶ Заметим, что

$$\left. \frac{d}{dt} \|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2^2 \right|_{t=0} = 2(\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) < 0$$

а значит для  $t \in (0, 1]$

$$\|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2 \leq \|\mathbf{d} - \mathbf{c}\|_2.$$

- ▶ Точка  $\mathbf{d} + t(\mathbf{u} - \mathbf{d}) \in B$  ближе к  $\mathbf{c}$ , чем  $\mathbf{d}$ , противоречие.

- ▶ Покажем, что  $f(\mathbf{y}) \geq 0$  для всех  $\mathbf{y} \in B$
- ▶ Пусть найдётся  $\mathbf{u} \in B$  такая что  $f(\mathbf{u}) < 0$
- ▶  $f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \frac{1}{2}(\mathbf{d} + \mathbf{c})) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) + \frac{1}{2}\|\mathbf{d} - \mathbf{c}\|_2^2$
- ▶  $(\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) < 0$
- ▶ Заметим, что

$$\left. \frac{d}{dt} \|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2^2 \right|_{t=0} = 2(\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) < 0$$

а значит для  $t \in (0, 1]$

$$\|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2 \leq \|\mathbf{d} - \mathbf{c}\|_2.$$

- ▶ Точка  $\mathbf{d} + t(\mathbf{u} - \mathbf{d}) \in B$  ближе к  $\mathbf{c}$ , чем  $\mathbf{d}$ , противоречие.
- ▶ Случай  $f(\mathbf{v}) \leq 0$  для всех  $\mathbf{v} \in A$  аналогичен для  $-f$

- ▶ Покажем, что  $f(\mathbf{y}) \geq 0$  для всех  $\mathbf{y} \in B$
- ▶ Пусть найдётся  $\mathbf{u} \in B$  такая что  $f(\mathbf{u}) < 0$
- ▶  $f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \frac{1}{2}(\mathbf{d} + \mathbf{c})) = (\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) + \frac{1}{2}\|\mathbf{d} - \mathbf{c}\|_2^2$
- ▶  $(\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) < 0$
- ▶ Заметим, что

$$\left. \frac{d}{dt} \|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2^2 \right|_{t=0} = 2(\mathbf{d} - \mathbf{c})^\top (\mathbf{u} - \mathbf{d}) < 0$$

а значит для  $t \in (0, 1]$

$$\|\mathbf{d} - \mathbf{c} + t(\mathbf{u} - \mathbf{d})\|_2 \leq \|\mathbf{d} - \mathbf{c}\|_2.$$

- ▶ Точка  $\mathbf{d} + t(\mathbf{u} - \mathbf{d}) \in B$  ближе к  $\mathbf{c}$ , чем  $\mathbf{d}$ , противоречие.
- ▶ Случай  $f(\mathbf{v}) \leq 0$  для всех  $\mathbf{v} \in A$  аналогичен для  $-f$

**Q:** верно ли обратное утверждение?



# Критерий отделимости для выпуклых множеств

## Теорема

*Два выпуклых множества, **одно из которых открыто**, не пересекаются тогда и только когда, когда они отделимы.*

# Критерий отделимости для выпуклых множеств

## Теорема

*Два выпуклых множества, **одно из которых открыто**, не пересекаются тогда и только когда, когда они отделимы.*

## Доказательство

- ▶ Пусть два выпуклых множества  $A, B$  отделимы и  $B$  является открытым

# Критерий делимости для выпуклых множеств

## Теорема

*Два выпуклых множества, **одно из которых открыто**, не пересекаются тогда и только когда, когда они делимы.*

## Доказательство

- ▶ Пусть два выпуклых множества  $A, B$  делимы и  $B$  является открытым
- ▶ Тогда  $\mathbf{a}^\top \mathbf{y} + b < 0$  на всех  $\mathbf{y} \in B$

# Критерий делимости для выпуклых множеств

## Теорема

*Два выпуклых множества, **одно из которых открыто**, не пересекаются тогда и только когда, когда они делимы.*

## Доказательство

- ▶ Пусть два выпуклых множества  $A, B$  делимы и  $B$  является открытым
- ▶ Тогда  $\mathbf{a}^\top \mathbf{y} + b < 0$  на всех  $\mathbf{y} \in B$
- ▶ Если  $\mathbf{u} \in B$  и  $\mathbf{a}^\top \mathbf{u} + b = 0$ , то в окрестности  $\mathbf{u}$  нашлась бы точка  $\mathbf{u}_+$ , в которой  $\mathbf{a}^\top \mathbf{u}_+ + b > 0$

# Критерий отделимости для выпуклых множеств

## Теорема

*Два выпуклых множества, **одно из которых открыто**, не пересекаются тогда и только когда, когда они отделимы.*

## Доказательство

- ▶ Пусть два выпуклых множества  $A, B$  отделимы и  $B$  является открытым
- ▶ Тогда  $\mathbf{a}^\top \mathbf{y} + b < 0$  на всех  $\mathbf{y} \in B$
- ▶ Если  $\mathbf{u} \in B$  и  $\mathbf{a}^\top \mathbf{u} + b = 0$ , то в окрестности  $\mathbf{u}$  нашлась бы точка  $\mathbf{u}_+$ , в которой  $\mathbf{a}^\top \mathbf{u}_+ + b > 0$
- ▶ Тогда  $A$  и  $B$  не пересекаются, так как на элементах  $A$  выполнено  $\geq$ , а на элементах  $B$  выполнено  $<$ .

# Лемма Фаркаша (Farkas' lemma)

## Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество  $\{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$  непусто
- ▶ существует вектор  $\mathbf{p}$  такой что  $\mathbf{p}^\top \mathbf{A} \geq 0$  и  $\mathbf{p}^\top \mathbf{b} < 0$

# Лемма Фаркаша (Farkas' lemma)

## Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество  $\{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$  непусто
- ▶ существует вектор  $\mathbf{p}$  такой что  $\mathbf{p}^\top \mathbf{A} \geq 0$  и  $\mathbf{p}^\top \mathbf{b} < 0$

## Доказательство

# Лемма Фаркаша (Farkas' lemma)

## Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество  $\{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$  непусто
- ▶ существует вектор  $\mathbf{p}$  такой что  $\mathbf{p}^\top \mathbf{A} \geq 0$  и  $\mathbf{p}^\top \mathbf{b} < 0$

## Доказательство

- ▶ Первое условие означает, что  $\mathbf{b}$  лежит в конусе  $C$ , образованном столбцами матрицы  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_m]$



# Лемма Фаркаша (Farkas' lemma)

## Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество  $\{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$  непусто
- ▶ существует вектор  $\mathbf{p}$  такой что  $\mathbf{p}^\top \mathbf{A} \geq 0$  и  $\mathbf{p}^\top \mathbf{b} < 0$

## Доказательство

- ▶ Первое условие означает, что  $\mathbf{b}$  лежит в конусе  $C$ , образованном столбцами матрицы  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_m]$
- ▶ Если это не так, то существует гиперплоскость, которая *строго* отделяет конус от точки  $\mathbf{b}$ :

$$\mathbf{c}^\top \mathbf{y} < d, \mathbf{y} \in C \quad \mathbf{c}^\top \mathbf{b} > d.$$

# Лемма Фаркаша (Farkas' lemma)

## Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество  $\{x \mid Ax = b, x \geq 0\}$  непусто
- ▶ существует вектор  $p$  такой что  $p^T A \geq 0$  и  $p^T b < 0$

## Доказательство

- ▶ Первое условие означает, что  $b$  лежит в конусе  $C$ , образованном столбцами матрицы  $A = [a_1, \dots, a_m]$
- ▶ Если это не так, то существует гиперплоскость, которая *строго* отделяет конус от точки  $b$ :

$$c^T y < d, y \in C \quad c^T b > d.$$

- ▶ Поскольку  $0 \in C$ , то  $d > 0$ . Также  $a_i \in C \Rightarrow \alpha a_i \in C \quad \alpha > 0$

# Лемма Фаркаша (Farkas' lemma)

## Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество  $\{x \mid Ax = b, x \geq 0\}$  непусто
- ▶ существует вектор  $p$  такой что  $p^T A \geq 0$  и  $p^T b < 0$

## Доказательство

- ▶ Первое условие означает, что  $b$  лежит в конусе  $C$ , образованном столбцами матрицы  $A = [a_1, \dots, a_m]$
- ▶ Если это не так, то существует гиперплоскость, которая *строго* отделяет конус от точки  $b$ :

$$c^T y < d, y \in C \quad c^T b > d.$$

- ▶ Поскольку  $0 \in C$ , то  $d > 0$ . Также  $a_i \in C \Rightarrow \alpha a_i \in C \quad \alpha > 0$
- ▶ Значит  $c^T \alpha a_i < d \Rightarrow c^T a_i < d/\alpha$ . При  $\alpha \rightarrow \infty$ ,  $c^T a_i \leq 0$

# Лемма Фаркаша (Farkas' lemma)

## Лемма Фаркаша

Выполнено одно и только одно из следующих условий

- ▶ множество  $\{x \mid Ax = b, x \geq 0\}$  непусто
- ▶ существует вектор  $p$  такой что  $p^T A \geq 0$  и  $p^T b < 0$

## Доказательство

- ▶ Первое условие означает, что  $b$  лежит в конусе  $C$ , образованном столбцами матрицы  $A = [a_1, \dots, a_m]$
- ▶ Если это не так, то существует гиперплоскость, которая *строго* отделяет конус от точки  $b$ :

$$c^T y < d, y \in C \quad c^T b > d.$$

- ▶ Поскольку  $0 \in C$ , то  $d > 0$ . Также  $a_i \in C \Rightarrow \alpha a_i \in C \quad \alpha > 0$
- ▶ Значит  $c^T \alpha a_i < d \Rightarrow c^T a_i < d/\alpha$ . При  $\alpha \rightarrow \infty$ ,  $c^T a_i \leq 0$
- ▶ Таким образом,  $p = -c$  и выполнено второе условие

# Главное в первой части

- ▶ Сопряжённые конусы и геометрическая интерпретация

# Главное в первой части

- ▶ Сопряжённые конусы и геометрическая интерпретация
- ▶ Самосопряжённые конусы

# Главное в первой части

- ▶ Сопряжённые конусы и геометрическая интерпретация
- ▶ Самосопряжённые конусы
- ▶ Отделимость выпуклых множеств

# Выпуклая функция (convex function)

## Определение

Функция  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$  называется выпуклой (*строго выпуклой*), если  $X$  — *выпуклое множество* и для

$\forall \mathbf{x}_1, \mathbf{x}_2 \in X$  и  $\alpha \in [0, 1]$  ( $\alpha \in (0, 1)$ ) выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$



# Выпуклая функция (convex function)

## Определение

Функция  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$  называется выпуклой (*строго выпуклой*), если  $X$  — *выпуклое множество* и для

$\forall \mathbf{x}_1, \mathbf{x}_2 \in X$  и  $\alpha \in [0, 1]$  ( $\alpha \in (0, 1)$ ) выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

## Определение

Функция  $f$  вогнута (concave), если функция  $-f$  выпукла.

# Выпуклая функция (convex function)

## Определение

Функция  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$  называется выпуклой (*строго выпуклой*), если  $X$  — *выпуклое множество* и для

$\forall \mathbf{x}_1, \mathbf{x}_2 \in X$  и  $\alpha \in [0, 1]$  ( $\alpha \in (0, 1)$ ) выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

## Определение

Функция  $f$  вогнута (concave), если функция  $-f$  выпукла.

## Примеры выпуклых функций

- ▶  $x^p$  для  $x \geq 0$  и  $p \geq 1$
- ▶  $x \log x$ , где  $x > 0$
- ▶  $\max\{x_1, \dots, x_n\}$
- ▶  $\|\mathbf{x}\|$
- ▶  $\log \left( \sum_{i=1}^n e^{x_i} \right)$
- ▶  $-\log \det \mathbf{X}$  для  $\mathbf{X} \in \mathbf{S}_{++}^n$

# Надграфик и выпуклость

## Определение

Множество  $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$  называется надграфиком (эпиграфом) функции  $f$ .

# Надграфик и выпуклость

## Определение

Множество  $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$  называется надграфиком (эпиграфом) функции  $f$ .

## Теорема

Функция  $f$  выпукла  $\Leftrightarrow \text{epi } f$  выпуклое множество.

# Надграфик и выпуклость

## Определение

Множество  $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$  называется надграфиком (эпиграфом) функции  $f$ .

## Теорема

Функция  $f$  выпукла  $\Leftrightarrow \text{epi } f$  выпуклое множество.

## Доказательство

# Надграфик и выпуклость

## Определение

Множество  $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$  называется надграфиком (эпиграфом) функции  $f$ .

## Теорема

Функция  $f$  выпукла  $\Leftrightarrow \text{epi } f$  выпуклое множество.

## Доказательство

1. Пусть  $f$  выпуклая функция

# Надграфик и выпуклость

## Определение

Множество  $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$  называется надграфиком (эпиграфом) функции  $f$ .

## Теорема

Функция  $f$  выпукла  $\Leftrightarrow \text{epi } f$  выпуклое множество.

## Доказательство

1. Пусть  $f$  выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа  $(\mathbf{x}_1, t_1)$  и  $(\mathbf{x}_2, t_2)$ , где  $t_1 \geq f(\mathbf{x}_1)$  и  $t_2 \geq f(\mathbf{x}_2)$

# Надграфик и выпуклость

## Определение

Множество  $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$  называется надграфиком (эпиграфом) функции  $f$ .

## Теорема

Функция  $f$  выпукла  $\Leftrightarrow \text{epi } f$  выпуклое множество.

## Доказательство

1. Пусть  $f$  выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа  $(\mathbf{x}_1, t_1)$  и  $(\mathbf{x}_2, t_2)$ , где  $t_1 \geq f(\mathbf{x}_1)$  и  $t_2 \geq f(\mathbf{x}_2)$
- ▶ Проверим принадлежность эпиграфу точки  $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha t_1 + (1 - \alpha) t_2)$



# Надграфик и выпуклость

## Определение

Множество  $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$  называется надграфиком (эпиграфом) функции  $f$ .

## Теорема

Функция  $f$  выпукла  $\Leftrightarrow \text{epi } f$  выпуклое множество.

## Доказательство

1. Пусть  $f$  выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа  $(\mathbf{x}_1, t_1)$  и  $(\mathbf{x}_2, t_2)$ , где  $t_1 \geq f(\mathbf{x}_1)$  и  $t_2 \geq f(\mathbf{x}_2)$
- ▶ Проверим принадлежность эпиграфу точки  $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha t_1 + (1 - \alpha) t_2)$
- ▶ В силу выпуклости функции  $\alpha t_1 + (1 - \alpha) t_2 \geq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \geq f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2)$ .

# Надграфик и выпуклость

## Определение

Множество  $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$  называется надграфиком (эпиграфом) функции  $f$ .

## Теорема

Функция  $f$  выпукла  $\Leftrightarrow \text{epi } f$  выпуклое множество.

## Доказательство

1. Пусть  $f$  выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа  $(\mathbf{x}_1, t_1)$  и  $(\mathbf{x}_2, t_2)$ , где  $t_1 \geq f(\mathbf{x}_1)$  и  $t_2 \geq f(\mathbf{x}_2)$
- ▶ Проверим принадлежность эпиграфу точки  $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha t_1 + (1 - \alpha) t_2)$
- ▶ В силу выпуклости функции 
$$\alpha t_1 + (1 - \alpha) t_2 \geq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \geq f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2).$$

2. Пусть надграфик выпуклое множество

# Надграфик и выпуклость

## Определение

Множество  $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$  называется надграфиком (эпиграфом) функции  $f$ .

## Теорема

Функция  $f$  выпукла  $\Leftrightarrow \text{epi } f$  выпуклое множество.

## Доказательство

1. Пусть  $f$  выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа  $(\mathbf{x}_1, t_1)$  и  $(\mathbf{x}_2, t_2)$ , где  $t_1 \geq f(\mathbf{x}_1)$  и  $t_2 \geq f(\mathbf{x}_2)$
- ▶ Проверим принадлежность эпиграфу точки  $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha t_1 + (1 - \alpha) t_2)$
- ▶ В силу выпуклости функции 
$$\alpha t_1 + (1 - \alpha) t_2 \geq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \geq f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2).$$

2. Пусть надграфик выпуклое множество

- ▶  $(\mathbf{x}_1, f(\mathbf{x}_1))$  и  $(\mathbf{x}_2, f(\mathbf{x}_2)) \in \text{epi } f$ , то  $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)) \in \text{epi } f$

# Надграфик и выпуклость

## Определение

Множество  $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$  называется надграфиком (эпиграфом) функции  $f$ .

## Теорема

Функция  $f$  выпукла  $\Leftrightarrow \text{epi } f$  выпуклое множество.

## Доказательство

1. Пусть  $f$  выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа  $(\mathbf{x}_1, t_1)$  и  $(\mathbf{x}_2, t_2)$ , где  $t_1 \geq f(\mathbf{x}_1)$  и  $t_2 \geq f(\mathbf{x}_2)$
- ▶ Проверим принадлежность эпиграфу точки  $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha t_1 + (1 - \alpha) t_2)$
- ▶ В силу выпуклости функции 
$$\alpha t_1 + (1 - \alpha) t_2 \geq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \geq f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2).$$

2. Пусть надграфик выпуклое множество

- ▶  $(\mathbf{x}_1, f(\mathbf{x}_1))$  и  $(\mathbf{x}_2, f(\mathbf{x}_2)) \in \text{epi } f$ , то  $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)) \in \text{epi } f$
- ▶ Из определения надграфика следует выпуклость  $f$

# Сильно выпуклая функция (strongly convex function)

## Определение

Функция  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$  называется сильно выпуклой с константой  $m > 0$ , если  $X$  — выпуклое множество и для  $\forall \mathbf{x}_1, \mathbf{x}_2 \in X$  и  $\alpha \in [0, 1]$  выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) - \frac{m}{2} \alpha (1 - \alpha) \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$$

# Сильно выпуклая функция (strongly convex function)

## Определение

Функция  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$  называется сильно выпуклой с константой  $m > 0$ , если  $X$  — выпуклое множество и для  $\forall \mathbf{x}_1, \mathbf{x}_2 \in X$  и  $\alpha \in [0, 1]$  выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) - \frac{m}{2} \alpha (1 - \alpha) \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$$

- ▶ Выпуклость  $\supset$  строгая выпуклость  $\supset$  сильная выпуклость

# Сильно выпуклая функция (strongly convex function)

## Определение

Функция  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$  называется сильно выпуклой с константой  $m > 0$ , если  $X$  — выпуклое множество и для  $\forall \mathbf{x}_1, \mathbf{x}_2 \in X$  и  $\alpha \in [0, 1]$  выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) - \frac{m}{2} \alpha (1 - \alpha) \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$$

- ▶ Выпуклость  $\supset$  строгая выпуклость  $\supset$  сильная выпуклость
- ▶ Для сильно выпуклых функций многие утверждения о методах оказываются более сильными, чем просто для выпуклых функций: пример о сходимости градиентного спуска.

## Дифференциальные критерии

Будем считать выпуклую функцию сильно выпуклой с  $m = 0$ .



## Дифференциальные критерии

Будем считать выпуклую функцию сильно выпуклой с  $m = 0$ .

### Теорема

*Пусть функция  $f(\mathbf{x})$  дифференцируема и определена на выпуклом множестве  $X \subseteq \mathbb{R}^n$ . Тогда  $f(\mathbf{x})$  сильно выпукла с константой  $m \geq 0$  в том и только том случае, если*

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

## Дифференциальные критерии

Будем считать выпуклую функцию сильно выпуклой с  $m = 0$ .

### Теорема

*Пусть функция  $f(\mathbf{x})$  дифференцируема и определена на выпуклом множестве  $X \subseteq \mathbb{R}^n$ . Тогда  $f(\mathbf{x})$  сильно выпукла с константой  $m \geq 0$  в том и только том случае, если*

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

Доказательство: пусть  $f$  выпукла

- По определению:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

## Дифференциальные критерии

Будем считать выпуклую функцию сильно выпуклой с  $m = 0$ .

### Теорема

Пусть функция  $f(\mathbf{x})$  дифференцируема и определена на выпуклом множестве  $X \subseteq \mathbb{R}^n$ . Тогда  $f(\mathbf{x})$  сильно выпукла с константой  $m \geq 0$  в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

Доказательство: пусть  $f$  выпукла

- По определению:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

- Перепишем в виде

$$f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) \leq f(\mathbf{x}_2) + \alpha(f(\mathbf{x}_1) - f(\mathbf{x}_2)) \text{ или}$$

$$\frac{f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2)}{\alpha} \leq f(\mathbf{x}_1) - f(\mathbf{x}_2)$$

## Дифференциальные критерии

Будем считать выпуклую функцию сильно выпуклой с  $m = 0$ .

### Теорема

Пусть функция  $f(\mathbf{x})$  дифференцируема и определена на выпуклом множестве  $X \subseteq \mathbb{R}^n$ . Тогда  $f(\mathbf{x})$  сильно выпукла с константой  $m \geq 0$  в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

Доказательство: пусть  $f$  выпукла

- ▶ По определению:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

- ▶ Перепишем в виде

$$f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) \leq f(\mathbf{x}_2) + \alpha(f(\mathbf{x}_1) - f(\mathbf{x}_2)) \text{ или}$$

$$\frac{f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2)}{\alpha} \leq f(\mathbf{x}_1) - f(\mathbf{x}_2)$$

- ▶ При  $\alpha \rightarrow 0$  получим

$$\langle f'(\mathbf{x}_2), \mathbf{x}_1 - \mathbf{x}_2 \rangle \leq f(\mathbf{x}_1) - f(\mathbf{x}_2)$$

Доказательство: пусть выполнено неравенство

- Рассмотрим  $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$

Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим  $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$
- ▶ Запишем два неравенства для  $\mathbf{z}, \mathbf{x}_1$  и  $\mathbf{z}, \mathbf{x}_2$

$$f(\mathbf{x}_1) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим  $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$
- ▶ Запишем два неравенства для  $\mathbf{z}, \mathbf{x}_1$  и  $\mathbf{z}, \mathbf{x}_2$

$$f(\mathbf{x}_1) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

- ▶ Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим  $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$
- ▶ Запишем два неравенства для  $\mathbf{z}, \mathbf{x}_1$  и  $\mathbf{z}, \mathbf{x}_2$

$$f(\mathbf{x}_1) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

- ▶ Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$



Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим  $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$
- ▶ Запишем два неравенства для  $\mathbf{z}, \mathbf{x}_1$  и  $\mathbf{z}, \mathbf{x}_2$

$$f(\mathbf{x}_1) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

- ▶ Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Сильно выпуклый случай

Для случая сильной выпуклости необходимо применить аналогичные выкладки к функции  $f(\mathbf{x}) - \frac{m}{2} \|\mathbf{x}\|_2^2$ .

Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим  $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$
- ▶ Запишем два неравенства для  $\mathbf{z}, \mathbf{x}_1$  и  $\mathbf{z}, \mathbf{x}_2$

$$f(\mathbf{x}_1) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

- ▶ Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

## Сильно выпуклый случай

Для случая сильной выпуклости необходимо применить аналогичные выкладки к функции  $f(\mathbf{x}) - \frac{m}{2} \|\mathbf{x}\|_2^2$ .

## Упражнение

Покажите, что  $f$  сильно выпукла  $\Leftrightarrow f(\mathbf{x}) - \frac{m}{2} \|\mathbf{x}\|_2^2$  выпукла.

## Теорема

Дважды непрерывно дифференцируемая функция  $f$  выпукла

$$\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$$

## Теорема

Дважды непрерывно дифференцируемая функция  $f$  выпукла

$$\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$$

## Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

## Теорема

Дважды непрерывно дифференцируемая функция  $f$  выпукла

$$\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$$

## Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

- Если  $f''(\mathbf{x}) \succeq m\mathbf{I}$ , то  $\frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$ ,  
и по критерию первого порядка  $f$  выпукла

## Теорема

Дважды непрерывно дифференцируемая функция  $f$  выпукла  
 $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

## Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

- Если  $f''(\mathbf{x}) \succeq m\mathbf{I}$ , то  $\frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$ ,  
и по критерию первого порядка  $f$  выпукла
- Если найдётся точка  $\mathbf{z}$  такая, что  $f''(\mathbf{z}) \not\succeq m\mathbf{I}$ , тогда  
найдётся направление  $\mathbf{d}$  такое, что  $\mathbf{d}^\top f''(\mathbf{z})\mathbf{d} \leq m$

## Теорема

Дважды непрерывно дифференцируемая функция  $f$  выпукла  
 $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

## Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

- Если  $f''(\mathbf{x}) \succeq m\mathbf{I}$ , то  $\frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$ ,  
и по критерию первого порядка  $f$  выпукла
- Если найдётся точка  $\mathbf{z}$  такая, что  $f''(\mathbf{z}) \not\succeq m\mathbf{I}$ , тогда найдётся направление  $\mathbf{d}$  такое, что  $\mathbf{d}^\top f''(\mathbf{z}) \mathbf{d} \leq m$
- В таком случае в силу критерия первого порядка  $f$  невыпукла — противоречие

## Теорема

Дважды непрерывно дифференцируемая функция  $f$  выпукла  
 $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

## Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

- Если  $f''(\mathbf{x}) \succeq m\mathbf{I}$ , то  $\frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$ ,  
и по критерию первого порядка  $f$  выпукла
- Если найдётся точка  $\mathbf{z}$  такая, что  $f''(\mathbf{z}) \not\succeq m\mathbf{I}$ , тогда найдётся направление  $\mathbf{d}$  такое, что  $\mathbf{d}^\top f''(\mathbf{z}) \mathbf{d} \leq m$
- В таком случае в силу критерия первого порядка  $f$  невыпукла — противоречие



## Теорема

Дважды непрерывно дифференцируемая функция  $f$  выпукла  
 $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

## Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

- Если  $f''(\mathbf{x}) \succeq m\mathbf{I}$ , то  $\frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$ ,  
и по критерию первого порядка  $f$  выпукла
- Если найдётся точка  $\mathbf{z}$  такая, что  $f''(\mathbf{z}) \not\succeq m\mathbf{I}$ , тогда найдётся направление  $\mathbf{d}$  такое, что  $\mathbf{d}^\top f''(\mathbf{z}) \mathbf{d} \leq m$
- В таком случае в силу критерия первого порядка  $f$  невыпукла — противоречие

## Напоминание

Для проверки определённости матрицы нужно использовать определение или критерий Сильвестра

# Выпуклые композиции функций

- ▶ Если  $f(\mathbf{x})$  — выпукла, то  $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$  также выпукла

# Выпуклые композиции функций

- ▶ Если  $f(\mathbf{x})$  — выпукла, то  $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$  также выпукла
- ▶ Если  $f(\mathbf{x})$  — выпукла, то  $g(t) = f(\mathbf{x} + t\mathbf{y})$  также выпукла

# Выпуклые композиции функций

- ▶ Если  $f(\mathbf{x})$  — выпукла, то  $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$  также выпукла
- ▶ Если  $f(\mathbf{x})$  — выпукла, то  $g(t) = f(\mathbf{x} + t\mathbf{y})$  также выпукла
- ▶ Если  $f_i$  — выпуклы, то  $\max_{i=1,\dots,m} f_i$  также выпукла

# Выпуклые композиции функций

- ▶ Если  $f(\mathbf{x})$  — выпукла, то  $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$  также выпукла
- ▶ Если  $f(\mathbf{x})$  — выпукла, то  $g(t) = f(\mathbf{x} + t\mathbf{y})$  также выпукла
- ▶ Если  $f_i$  — выпуклы, то  $\max_{i=1,\dots,m} f_i$  также выпукла
- ▶ Сумма выпуклых функций с неотрицательными коэффициентами — выпуклая функция

# Выпуклые композиции функций

- ▶ Если  $f(\mathbf{x})$  — выпукла, то  $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$  также выпукла
- ▶ Если  $f(\mathbf{x})$  — выпукла, то  $g(t) = f(\mathbf{x} + t\mathbf{y})$  также выпукла
- ▶ Если  $f_i$  — выпуклы, то  $\max_{i=1,\dots,m} f_i$  также выпукла
- ▶ Сумма выпуклых функций с неотрицательными коэффициентами — выпуклая функция
- ▶ Скалярная композиция  $h(f(\mathbf{x}))$

# Локальный минимум является глобальным минимумом

## Теорема

Если  $f$  выпуклая функция и  $x^*$  *локальный* минимум, то  $x^*$  — *глобальный* минимум.

# Локальный минимум является глобальным минимумом

## Теорема

Если  $f$  выпуклая функция и  $x^*$  *локальный* минимум, то  $x^*$  — *глобальный* минимум.

## Доказательство от противного

- ▶ Пусть  $y^* \neq x^*$  — глобальный минимум:  $f(y^*) < f(x^*)$



# Локальный минимум является глобальным минимумом

## Теорема

Если  $f$  выпуклая функция и  $\mathbf{x}^*$  **локальный** минимум, то  $\mathbf{x}^*$  — **глобальный** минимум.

## Доказательство от противного

- ▶ Пусть  $\mathbf{y}^* \neq \mathbf{x}^*$  — глобальный минимум:  $f(\mathbf{y}^*) < f(\mathbf{x}^*)$
- ▶ По определению локального минимума:  $f(\mathbf{x}^*) \leq f(\mathbf{x})$ , где  $\|\mathbf{x}^* - \mathbf{x}\|_2 \leq \delta$

# Локальный минимум является глобальным минимумом

## Теорема

Если  $f$  выпуклая функция и  $\mathbf{x}^*$  **локальный** минимум, то  $\mathbf{x}^*$  — **глобальный** минимум.

## Доказательство от противного

- ▶ Пусть  $\mathbf{y}^* \neq \mathbf{x}^*$  — глобальный минимум:  $f(\mathbf{y}^*) < f(\mathbf{x}^*)$
- ▶ По определению локального минимума:  $f(\mathbf{x}^*) \leq f(\mathbf{x})$ , где  $\|\mathbf{x}^* - \mathbf{x}\|_2 \leq \delta$
- ▶ Выберем достаточно малое  $\alpha \in (0, 1)$  и рассмотрим точку  $\mathbf{z} = (1 - \alpha)\mathbf{x}^* + \alpha\mathbf{y}^*$  такую что  $\|\mathbf{z} - \mathbf{x}^*\|_2 \leq \delta$

# Локальный минимум является глобальным минимумом

## Теорема

Если  $f$  выпуклая функция и  $\mathbf{x}^*$  **локальный** минимум, то  $\mathbf{x}^*$  — **глобальный** минимум.

## Доказательство от противного

- ▶ Пусть  $\mathbf{y}^* \neq \mathbf{x}^*$  — глобальный минимум:  $f(\mathbf{y}^*) < f(\mathbf{x}^*)$
- ▶ По определению локального минимума:  $f(\mathbf{x}^*) \leq f(\mathbf{x})$ , где  $\|\mathbf{x}^* - \mathbf{x}\|_2 \leq \delta$
- ▶ Выберем достаточно малое  $\alpha \in (0, 1)$  и рассмотрим точку  $\mathbf{z} = (1 - \alpha)\mathbf{x}^* + \alpha\mathbf{y}^*$  такую что  $\|\mathbf{z} - \mathbf{x}^*\|_2 \leq \delta$
- ▶  $f(\mathbf{x}^*) \leq f(\mathbf{z}) \leq \alpha f(\mathbf{y}^*) + (1 - \alpha)f(\mathbf{x}^*) < f(\mathbf{x}^*)$

# Пример сложной выпуклой задачи

## Определение

Множество  $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$  называется *copositive cone*.

# Пример сложной выпуклой задачи

## Определение

Множество  $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$  называется *copositive cone*.

- $\mathcal{C}^n$  выпукло

# Пример сложной выпуклой задачи

## Определение

Множество  $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$  называется *copositive cone*.

- ▶  $\mathcal{C}^n$  выпукло
- ▶  $\mathbf{S}_+^n \subset \mathcal{C}^n$

# Пример сложной выпуклой задачи

## Определение

Множество  $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$  называется *copositive cone*.

- ▶  $\mathcal{C}^n$  выпукло
- ▶  $\mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки  $\mathbf{X} \notin \mathcal{C}^n$  является co-NP полной!

# Пример сложной выпуклой задачи

## Определение

Множество  $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$  называется *copositive cone*.

- ▶  $\mathcal{C}^n$  выпукло
- ▶  $\mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки  $\mathbf{X} \notin \mathcal{C}^n$  является co-NP полной!
- ▶ Задача конической оптимизации с конусом  $\mathcal{C}^n$  является NP-трудной



# Пример сложной выпуклой задачи

## Определение

Множество  $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$  называется *copositive cone*.

- ▶  $\mathcal{C}^n$  выпукло
- ▶  $\mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки  $\mathbf{X} \notin \mathcal{C}^n$  является co-NP полной!
- ▶ Задача конической оптимизации с конусом  $\mathcal{C}^n$  является NP-трудной

# Пример сложной выпуклой задачи

## Определение

Множество  $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$  называется *copositive cone*.

- ▶  $\mathcal{C}^n$  выпукло
- ▶  $\mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки  $\mathbf{X} \notin \mathcal{C}^n$  является co-NP полной!
- ▶ Задача конической оптимизации с конусом  $\mathcal{C}^n$  является NP-трудной

## Пример

Задача определения максимального независимого множества вершин графа сводится к задаче оптимизации на множестве  $\mathcal{C}^n$ . Подробности [тут](#)

## Пример простой невыпуклой задачи

Дана матрица  $\mathbf{Q} \in \mathbf{S}_{++}^n$ .

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

## Пример простой невыпуклой задачи

Дана матрица  $\mathbf{Q} \in \mathbf{S}_{++}^n$ .

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

- ▶ Допустимое множество невыпукло

## Пример простой невыпуклой задачи

Дана матрица  $\mathbf{Q} \in \mathbf{S}_{++}^n$ .

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

- ▶ Допустимое множество невыпукло
- ▶ Целевая функция выпукла

## Пример простой невыпуклой задачи

Дана матрица  $\mathbf{Q} \in \mathbf{S}_{++}^n$ .

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

- ▶ Допустимое множество невыпукло
- ▶ Целевая функция выпукла

**Q:** какая интерпретация у  $\mathbf{x}^*$  и  $f(\mathbf{x}^*)$ ?

# Неравенство Йенсена

## Теорема

Если функция  $f$  выпукла, то  $f\left(\sum_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$ , где

$$\sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0.$$

# Неравенство Йенсена

## Теорема

Если функция  $f$  выпукла, то  $f\left(\sum_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$ , где

$$\sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0.$$

## Доказательство по индукции

- База  $k = 2$  выполнена в силу определения



# Неравенство Йенсена

## Теорема

Если функция  $f$  выпукла, то  $f\left(\sum_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$ , где

$$\sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0.$$

## Доказательство по индукции

- ▶ База  $k = 2$  выполнена в силу определения
- ▶ Пусть неравенство выполнено для  $k = m - 1$ :

$$f\left(\sum_{i=1}^{m-1} \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^{m-1} \alpha_i f(\mathbf{x}_i) \text{ и } \sum_{i=1}^{m-1} \alpha_i = 1, \alpha_i \geq 0$$

# Неравенство Йенсена

## Теорема

Если функция  $f$  выпукла, то  $f\left(\sum_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$ , где

$$\sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0.$$

## Доказательство по индукции

► База  $k = 2$  выполнена в силу определения

► Пусть неравенство выполнено для  $k = m - 1$ :

$$f\left(\sum_{i=1}^{m-1} \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^{m-1} \alpha_i f(\mathbf{x}_i) \text{ и } \sum_{i=1}^{m-1} \alpha_i = 1, \alpha_i \geq 0$$

► Рассмотрим  $k = m$ :  $f\left(\sum_{i=1}^m \hat{\alpha}_i \mathbf{x}_i\right) = f\left(\sum_{i=1}^{m-1} \hat{\alpha} \mathbf{x}_i + \hat{\alpha}_m \mathbf{x}_m\right) =$

$$f\left((1 - \hat{\alpha}_m) \sum_{i=1}^{m-1} \frac{\hat{\alpha}_i}{1 - \hat{\alpha}_m} \mathbf{x}_i + \hat{\alpha}_m \mathbf{x}_m\right) \leq$$

$$(1 - \hat{\alpha}_m) f\left(\sum_{i=1}^{m-1} \frac{\hat{\alpha}_i}{1 - \hat{\alpha}_m} \mathbf{x}_i\right) + \hat{\alpha}_m f(\mathbf{x}_m) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$$

## Следствия и обобщения

- Запись неравенства Йенсена для функции  $-\log x$

$$\frac{1}{m} \sum_{i=1}^m x_i \geq \sqrt[m]{x_1 \cdot \dots \cdot x_m}$$

## Следствия и обобщения

- ▶ Запись неравенства Йенсена для функции  $-\log x$

$$\frac{1}{m} \sum_{i=1}^m x_i \geq \sqrt[m]{x_1 \cdot \dots \cdot x_m}$$

- ▶ Неравенство Гёльдера

$$\sum_{i=1}^n x_i y_i \leq \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} \left( \sum_{i=1}^n |y_i|^q \right)^{1/q}$$

## Следствия и обобщения

- ▶ Запись неравенства Йенсена для функции  $-\log x$

$$\frac{1}{m} \sum_{i=1}^m x_i \geq \sqrt[m]{x_1 \cdot \dots \cdot x_m}$$

- ▶ Неравенство Гёльдера

$$\sum_{i=1}^n x_i y_i \leq \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} \left( \sum_{i=1}^n |y_i|^q \right)^{1/q}$$

- ▶ Обобщение на непрерывный случай даёт неравенство для выпуклой функции от математического ожидания

$$f(\mathbb{E}(\mathbf{x})) \leq \mathbb{E}(f(\mathbf{x}))$$

# Главное во второй части

- ▶ Выпуклые функции и их свойства

# Главное во второй части

- ▶ Выпуклые функции и их свойства
- ▶ Операции, сохраняющие выпуклость

# Главное во второй части

- ▶ Выпуклые функции и их свойства
- ▶ Операции, сохраняющие выпуклость
- ▶ Неравенство Йенсена