

# Методы оптимизации

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

Александр Катруца

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



12 сентября 2022 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  $\mathcal{K}$  называется конусом, если для любого  $\mathbf{x} \in \mathcal{K}$  и произвольного числа  $\theta \geq 0$  выполнено  $\theta\mathbf{x} \in \mathcal{K}$ .*

# Напоминание: конусы

## Определение

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

## Определение

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

# Напоминание: конусы

## Определение

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

## Определение

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

## 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  $\{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}_+ \mid \|\mathbf{x}\|_2 \leq t\} \rightarrow \text{SOCP}$
- ▶ Конус симметричных положительно полуопределённых матриц  $\mathbf{S}_+^n \rightarrow \text{SDP}$

# Сопряжённый конус (dual cone)

## Определение

Пусть  $\mathcal{K}$  — конус. Тогда множество

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

называется сопряжённым конусом.

# Сопряжённый конус (dual cone)

## Определение

Пусть  $\mathcal{K}$  — конус. Тогда множество

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

называется сопряжённым конусом.

## Свойства

- ▶  $\mathcal{K}^*$  — конус
- ▶  $\mathcal{K}^*$  — выпуклый конус для любого конуса  $\mathcal{K}$
- ▶ Если  $\mathcal{K}_1 \subseteq \mathcal{K}_2$ , то  $\mathcal{K}_2^* \subseteq \mathcal{K}_1^*$

# Сопряжённый конус (dual cone)

## Определение

Пусть  $\mathcal{K}$  — конус. Тогда множество

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

называется сопряжённым конусом.

## Свойства

- ▶  $\mathcal{K}^*$  — конус
- ▶  $\mathcal{K}^*$  — выпуклый конус для любого конуса  $\mathcal{K}$
- ▶ Если  $\mathcal{K}_1 \subseteq \mathcal{K}_2$ , то  $\mathcal{K}_2^* \subseteq \mathcal{K}_1^*$

## Определение

Если  $\mathcal{K} = \mathcal{K}^*$ , 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  $\{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}_+ \mid \|\mathbf{x}\|_2 \leq t\}$

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

# Правильный конус (proper cone)

## Определение

Конус  $\mathcal{K}$  называется правильным (proper), если

- ▶  $\mathcal{K}$  выпуклый
- ▶  $\mathcal{K}$  замкнутый
- ▶  $\mathcal{K}$  не содержит прямых
- ▶ внутренность  $\mathcal{K}$  непуста

# Правильный конус (proper cone)

## Определение

Конус  $\mathcal{K}$  называется правильным (proper), если

- ▶  $\mathcal{K}$  выпуклый
- ▶  $\mathcal{K}$  замкнутый
- ▶  $\mathcal{K}$  не содержит прямых
- ▶ внутренность  $\mathcal{K}$  непуста

## Упражнение

Покажите, что самосопряжённые конусы, перечисленные выше, являются правильными.



Зачем нужны конусы?

# Зачем нужны конусы?

Обобщённое отношение частичного порядка

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

# Зачем нужны конусы?

## Обобщённое отношение частичного порядка

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

## Пример: конус $\mathbf{S}_+^n$

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

# Зачем нужны конусы?

## Обобщённое отношение частичного порядка

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

## Пример: конус $\mathbf{S}_+^n$

Пусть  $\mathbf{X}, \mathbf{Y} \in \mathbf{S}^n$ . Тогда  $\mathbf{X} \leq_{\mathcal{K}} \mathbf{Y}$  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  $\mathcal{K}$  — правильный конус. Тогда  $\mathbf{x} \leq_{\mathcal{K}} \mathbf{y} \Leftrightarrow \mathbf{y} - \mathbf{x} \in \mathcal{K}$ .

## Пример: конус $\mathbf{S}_+^n$

Пусть  $\mathbf{X}, \mathbf{Y} \in \mathbf{S}^n$ . Тогда  $\mathbf{X} \leq_{\mathcal{K}} \mathbf{Y}$  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  $\mathcal{A}, \mathcal{B}$  называются *отделимыми*, если существует вектор  $\mathbf{a} \neq 0$  и число  $b$  такие что

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

# Отделимость выпуклых множеств

## Определение

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

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

## Теорема

Пусть  $\mathcal{A}$  и  $\mathcal{B}$  — выпуклые и непересекающиеся множества.  
Тогда существует разделяющая их гиперплоскость.

# Отделимость выпуклых множеств

## Определение

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

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

## Теорема

Пусть  $\mathcal{A}$  и  $\mathcal{B}$  — выпуклые и непересекающиеся множества. Тогда существует разделяющая их гиперплоскость.

## Теорема

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



# 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  $\{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$  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  $\mathbf{b}$ :

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

- ▶ Поскольку  $0 \in C$ , то  $d > 0$ . Также  $\mathbf{a}_i \in C \Rightarrow \alpha \mathbf{a}_i \in C \quad \alpha > 0$

# 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

- ▶ 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  $\mathbf{b}$ :

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

- ▶ Поскольку  $0 \in C$ , то  $d > 0$ . Также  $\mathbf{a}_i \in C \Rightarrow \alpha \mathbf{a}_i \in C \quad \alpha > 0$
- ▶ Значит  $\mathbf{c}^\top \alpha \mathbf{a}_i < d \Rightarrow \mathbf{c}^\top \mathbf{a}_i < d/\alpha$ . При  $\alpha \rightarrow \infty$ ,  $\mathbf{c}^\top \mathbf{a}_i \leq 0$

# 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

- ▶ 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  $\mathbf{b}$ :

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

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

## Приложение: теорема об арбитраже

- ▶ Пусть есть  $n$  активов с ценами  $p_1, \dots, p_n$  до и  $v_1, \dots, v_n$  в конце периода инвестирования
- ▶ Пусть  $\mathbf{x} \in \mathbb{R}^n$  размер инвестиций в каждый актив
- ▶ Значения для цен  $v_i$  неизвестны, но пусть возможно  $K$  наборов таких цен, которые известны
- ▶ Если  $\langle \mathbf{p}, \mathbf{x} \rangle < 0$  и  $\langle \mathbf{v}^{(k)}, \mathbf{x} \rangle \geq 0$  для всех  $k = 1, \dots, K$ , то такая стратегия гарантировано принесёт прибыль!
- ▶ Ситуация на рынке, при которой существует гарантированно прибыльная стратегия называется *арбитражем*
- ▶ Такая ситуация в общем случае не обязана выполняться, то есть система  $\mathbf{V}\mathbf{x} \geq 0$ ,  $\langle \mathbf{p}, \mathbf{x} \rangle < 0$  несовместна
- ▶ По лемме Фаркаша это равносильно существованию  $\mathbf{y} \geq 0$  такому, что  $\mathbf{V}^\top \mathbf{y} = \mathbf{p}$



## Полный рынок (complete market)

- ▶ Пусть известна вся матрица  $\mathbf{V}$  и все  $p_i$  кроме  $p_n$
- ▶ Тогда можно поставить задачу поиска интервала для  $p_n$

$$\begin{aligned} & \max_{p_n, \mathbf{y}} / \min_{p_n, \mathbf{y}} p_n \\ \text{s.t. } & \mathbf{V}^\top \mathbf{y} = \mathbf{p} \\ & \mathbf{y} \geq 0 \end{aligned}$$

- ▶ Если условие арбитража приводит к единственным ценам, то такой рынок называется *полным*.

# 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  $\mathbf{d}^\top f''(\mathbf{z})\mathbf{d} < m\|\mathbf{d}\|_2^2$

## 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  $\mathbf{d}^\top f''(\mathbf{z})\mathbf{d} < m\|\mathbf{d}\|_2^2$
- Пусть  $\mathbf{y} = \mathbf{x} + \varepsilon\mathbf{d}$  для некоторого  $\varepsilon > 0$

## 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  $\mathbf{d}^\top f''(\mathbf{z})\mathbf{d} < m\|\mathbf{d}\|_2^2$
- ▶ Пусть  $\mathbf{y} = \mathbf{x} + \varepsilon\mathbf{d}$  для некоторого  $\varepsilon > 0$
- ▶ При  $\varepsilon$  достаточно малом,  $\mathbf{x}_\alpha$  и  $\mathbf{y}$  так близки к  $\mathbf{x}$ , что  
 $\mathbf{d}^\top f''(\mathbf{x}_\alpha)\mathbf{d} < m\|\mathbf{d}\|_2^2$  в силу непрерывности гессиана

## 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  $\mathbf{d}^\top f''(\mathbf{z})\mathbf{d} < m\|\mathbf{d}\|_2^2$
- ▶ Пусть  $\mathbf{y} = \mathbf{x} + \varepsilon\mathbf{d}$  для некоторого  $\varepsilon > 0$
- ▶ При  $\varepsilon$  достаточно малом,  $\mathbf{x}_\alpha$  и  $\mathbf{y}$  так близки к  $\mathbf{x}$ , что  
 $\mathbf{d}^\top f''(\mathbf{x}_\alpha)\mathbf{d} < m\|\mathbf{d}\|_2^2$  в силу непрерывности гессиана
- ▶ В таком случае в силу критерия первого порядка  $f$   
невыпукла — противоречие

# 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  $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b})$  также выпукла
- ▶  $f(\mathbf{x})$  — выпукла iff  $g(t) = f(\mathbf{x} + t\mathbf{y})$  выпукла как функция скалярного аргумента при условии что  $\mathbf{x} + t\mathbf{y} \in \text{dom } f$



# Выпуклые композиции функций

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

# Выпуклые композиции функций

- ▶ Если  $f(\mathbf{x})$  — выпукла, то  $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b})$  также выпукла
- ▶  $f(\mathbf{x})$  — выпукла iff  $g(t) = f(\mathbf{x} + t\mathbf{y})$  выпукла как функция скалярного аргумента при условии что  $\mathbf{x} + t\mathbf{y} \in \text{dom } f$
- ▶ Если  $f_i$  — выпуклы, то  $\max_{i=1, \dots, m} f_i$  также выпукла
- ▶ Сумма выпуклых функций с неотрицательными коэффициентами — выпуклая функция
- ▶ Скалярная композиция  $h(f(\mathbf{x}))$
- ▶ Перспективное преобразование: если  $f(\mathbf{x})$  выпукла, то  $g(\mathbf{x}, t) = tf(\mathbf{x}/t)$ , где  $t > 0$  и  $\mathbf{x}/t \in \text{dom } f$  также выпукла

# 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  $f(\mathbf{x}^*)$ ?

Почему решением является  $\lambda_{\min}(\mathbf{Q})$ ?

- ▶ Так как  $\mathbf{Q} \in \mathbf{S}_{++}^n$ , то существует ортогональный базис из собственных векторов  $\mathbf{v}_1, \dots, \mathbf{v}_n$  таких что  $\|\mathbf{v}_i\|_2 = 1$  и  $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$  для  $i \neq j$

## Почему решением является $\lambda_{\min}(\mathbf{Q})$ ?

- ▶ Так как  $\mathbf{Q} \in \mathbf{S}_{++}^n$ , то существует ортогональный базис из собственных векторов  $\mathbf{v}_1, \dots, \mathbf{v}_n$  таких что  $\|\mathbf{v}_i\|_2 = 1$  и  $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$  для  $i \neq j$
- ▶ Будем искать решение в виде  $\mathbf{x}_* = \sum_{i=1}^n \alpha_i \mathbf{v}_i$

## Почему решением является $\lambda_{\min}(\mathbf{Q})$ ?

- ▶ Так как  $\mathbf{Q} \in \mathbf{S}_{++}^n$ , то существует ортогональный базис из собственных векторов  $\mathbf{v}_1, \dots, \mathbf{v}_n$  таких что  $\|\mathbf{v}_i\|_2 = 1$  и  $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$  для  $i \neq j$
- ▶ Будем искать решение в виде  $\mathbf{x}_* = \sum_{i=1}^n \alpha_i \mathbf{v}_i$
- ▶ Тогда  $\mathbf{x}_*^\top \mathbf{Q} \mathbf{x}_* = \sum_{i=1}^n \alpha_i^2 \lambda_i$ , где  $\lambda_1 \geq \dots \geq \lambda_n > 0$

## Почему решением является $\lambda_{\min}(\mathbf{Q})$ ?

- ▶ Так как  $\mathbf{Q} \in \mathbf{S}_{++}^n$ , то существует ортогональный базис из собственных векторов  $\mathbf{v}_1, \dots, \mathbf{v}_n$  таких что  $\|\mathbf{v}_i\|_2 = 1$  и  $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$  для  $i \neq j$
- ▶ Будем искать решение в виде  $\mathbf{x}_* = \sum_{i=1}^n \alpha_i \mathbf{v}_i$
- ▶ Тогда  $\mathbf{x}_*^\top \mathbf{Q} \mathbf{x}_* = \sum_{i=1}^n \alpha_i^2 \lambda_i$ , где  $\lambda_1 \geq \dots \geq \lambda_n > 0$
- ▶ Ограничение  $\|\mathbf{x}_*\|_2 = 1$  даёт ограничения на  $\alpha_i$ :  
$$\sum_{i=1}^n \alpha_i^2 = 1$$



## Почему решением является $\lambda_{\min}(\mathbf{Q})$ ?

- ▶ Так как  $\mathbf{Q} \in \mathbf{S}_{++}^n$ , то существует ортогональный базис из собственных векторов  $\mathbf{v}_1, \dots, \mathbf{v}_n$  таких что  $\|\mathbf{v}_i\|_2 = 1$  и  $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$  для  $i \neq j$
- ▶ Будем искать решение в виде  $\mathbf{x}_* = \sum_{i=1}^n \alpha_i \mathbf{v}_i$
- ▶ Тогда  $\mathbf{x}_*^\top \mathbf{Q} \mathbf{x}_* = \sum_{i=1}^n \alpha_i^2 \lambda_i$ , где  $\lambda_1 \geq \dots \geq \lambda_n > 0$
- ▶ Ограничение  $\|\mathbf{x}_*\|_2 = 1$  даёт ограничения на  $\alpha_i$ :  
$$\sum_{i=1}^n \alpha_i^2 = 1$$
- ▶ Получим оценку снизу  $\sum_{i=1}^n \alpha_i^2 \lambda_i \geq \lambda_{\min} \sum_{i=1}^n \alpha_i^2 = \lambda_{\min}$

## Почему решением является $\lambda_{\min}(\mathbf{Q})$ ?

- ▶ Так как  $\mathbf{Q} \in \mathbf{S}_{++}^n$ , то существует ортогональный базис из собственных векторов  $\mathbf{v}_1, \dots, \mathbf{v}_n$  таких что  $\|\mathbf{v}_i\|_2 = 1$  и  $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$  для  $i \neq j$
- ▶ Будем искать решение в виде  $\mathbf{x}_* = \sum_{i=1}^n \alpha_i \mathbf{v}_i$
- ▶ Тогда  $\mathbf{x}_*^\top \mathbf{Q} \mathbf{x}_* = \sum_{i=1}^n \alpha_i^2 \lambda_i$ , где  $\lambda_1 \geq \dots \geq \lambda_n > 0$
- ▶ Ограничение  $\|\mathbf{x}_*\|_2 = 1$  даёт ограничения на  $\alpha_i$ :  
$$\sum_{i=1}^n \alpha_i^2 = 1$$
- ▶ Получим оценку снизу  $\sum_{i=1}^n \alpha_i^2 \lambda_i \geq \lambda_{\min} \sum_{i=1}^n \alpha_i^2 = \lambda_{\min}$
- ▶ Эта оценка достигается на коэффициентах

$$\alpha_i = \begin{cases} 0, & i \neq n \\ 1, & i = n \end{cases}$$

# 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

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

где  $\frac{1}{p} + \frac{1}{q} = 1$  и  $p, q \geq 1$



## 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

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

где  $\frac{1}{p} + \frac{1}{q} = 1$  и  $p, q \geq 1$

- ▶ 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
- ▶ Сложная задача выпуклой оптимизации и простая невыпуклая задача

# Главное во второй части

- ▶ Выпуклые функции и их свойства
- ▶ Операции, сохраняющие выпуклость
- ▶ Сложная задача выпуклой оптимизации и простая невыпуклая задача
- ▶ Неравенство Йенсена