# Методы оптимизации. Семинар 1. Введение. Выпуклые множества. Конусы

Александр Катруца

Московский физико-технический институт
Факультет Инноваций и Высоких Технологий

4 сентября 2018 г.

# Вопросы к студентам

- Имя
- Кафедра
- Знание TEX/LATEX
- Ожидания от курса

# План на семестр

- Семинар и лекция раз в неделю

# План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара

# План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашнее задание после каждого семинара

# План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашнее задание после каждого семинара
- Итоговая контрольная в конце семестра и промежуточная в середине семестра

# План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашнее задание после каждого семинара
- Итоговая контрольная в конце семестра и промежуточная в середине семестра
- Piazza для Q& A

# План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашнее задание после каждого семинара
- Итоговая контрольная в конце семестра и промежуточная в середине семестра
- Piazza для Q& A

? Проект

# План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашнее задание после каждого семинара
- Итоговая контрольная в конце семестра и промежуточная в середине семестра
- Piazza для Q& A

? Проект

- Итоговая оценка – взвешенная сумма баллов за все активности

# Зачем этот курс?

- Формализация задачи выбора элемента из множества

# Зачем этот курс?

- Формализация задачи выбора элемента из множества
- Обоснование правильности принятия решения

# Зачем этот курс?

- Формализация задачи выбора элемента из множества
- Обоснование правильности принятия решения
- Разнообразные приложения:

# Зачем этот курс?

- Формализация задачи выбора элемента из множества
- Обоснование правильности принятия решения
- Разнообразные приложения:
  - машинное обучение: классификация, кластеризация, регрессия

# Зачем этот курс?

- Формализация задачи выбора элемента из множества
- Обоснование правильности принятия решения
- Разнообразные приложения:
    - машинное обучение: классификация, кластеризация, регрессия
    - молекулярное моделирование

# Зачем этот курс?

- Формализация задачи выбора элемента из множества
- Обоснование правильности принятия решения
- Разнообразные приложения:
  - машинное обучение: классификация, кластеризация, регрессия
  - молекулярное моделирование
  - анализ рисков

# Зачем этот курс?

- Формализация задачи выбора элемента из множества
- Обоснование правильности принятия решения
- Разнообразные приложения:
  - машинное обучение: классификация, кластеризация, регрессия
  - молекулярное моделирование
  - анализ рисков
  - выбор активов (portfolio optimization)

# Зачем этот курс?

- Формализация задачи выбора элемента из множества
- Обоснование правильности принятия решения
- Разнообразные приложения:
  - машинное обучение: классификация, кластеризация, регрессия
  - молекулярное моделирование
  - анализ рисков
  - выбор активов (portfolio optimization)
  - оптимальное управление

# Зачем этот курс?

- Формализация задачи выбора элемента из множества
- Обоснование правильности принятия решения
- Разнообразные приложения:
  - машинное обучение: классификация, кластеризация, регрессия
  - молекулярное моделирование
  - анализ рисков
  - выбор активов (portfolio optimization)
  - оптимальное управление
  - обработка сигналов

# Зачем этот курс?

- Формализация задачи выбора элемента из множества
- Обоснование правильности принятия решения
- Разнообразные приложения:
  - машинное обучение: классификация, кластеризация, регрессия
  - молекулярное моделирование
  - анализ рисков
  - выбор активов (portfolio optimization)
  - оптимальное управление
  - обработка сигналов
  - оценка параметров в статистике

# Зачем этот курс?

- Формализация задачи выбора элемента из множества
- Обоснование правильности принятия решения
- Разнообразные приложения:
    - машинное обучение: классификация, кластеризация, регрессия
    - молекулярное моделирование
    - анализ рисков
    - выбор активов (portfolio optimization)
    - оптимальное управление
    - обработка сигналов
    - оценка параметров в статистике
    - и другие

# О чём этот курс?

Теоретическая половина (сентябрь-октябрь):

- Основы выпуклого анализа
- Условия оптимальности
- Теория двойственности

# О чём этот курс?

Теоретическая половина (сентябрь-октябрь):

- Основы выпуклого анализа
- Условия оптимальности
- Теория двойственности

Практическая половина (ноябрь-декабрь):

- Методы безусловной минимизации первого и второго порядка
- Методы условной оптимизации
- Линейное программирование: симплекс-метод и пр.
- Оптимальные методы
- ...

# Предварительные навыки

- Линейная алгебра
- Математический анализ
- Программирование: Python (NumPy, SciPy, CVXPY) или MATLAB
- Элементы вычислительной математики

# Методология

Основные этапы использования методов оптимизации при решении реальных задач:

1. Определение целевой функции

# Методология

Основные этапы использования методов оптимизации при решении реальных задач:

1. Определение целевой функции
2. Определение допустимого множества решений

# Методология

Основные этапы использования методов оптимизации при решении реальных задач:

1. Определение целевой функции
2. Определение допустимого множества решений
3. Постановка и анализ оптимизационной задачи

# Методология

Основные этапы использования методов оптимизации при решении реальных задач:

1. Определение целевой функции
2. Определение допустимого множества решений
3. Постановка и анализ оптимизационной задачи
4. Выбор наилучшего алгоритма для решения поставленной задачи

# Методология

Основные этапы использования методов оптимизации при решении реальных задач:

1. Определение целевой функции
2. Определение допустимого множества решений
3. Постановка и анализ оптимизационной задачи
4. Выбор наилучшего алгоритма для решения поставленной задачи
5. Реализация алгоритма и проверка его корректности

# Постановка задачи

$$
\begin{aligned}
&\min_{\mathbf{x} \in X} f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) = 0, \; i = 1, \ldots, p \\
& f_j(\mathbf{x}) \leq 0, \; j = p+1, \ldots, m,
\end{aligned}
$$

# Постановка задачи

$$
\begin{aligned}
& \min_{\mathbf{x} \in X} f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) = 0, \; i = 1, \ldots, p \\
& f_j(\mathbf{x}) \leq 0, \; j = p+1, \ldots, m,
\end{aligned}
$$

- $\mathbf{x} \in \mathbb{R}^n$ — искомый вектор

# Постановка задачи

$$
\begin{aligned}
& \min_{\mathbf{x} \in X} f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) = 0, \ i = 1, \dots, p \\
& f_j(\mathbf{x}) \leq 0, \ j = p+1, \dots, m,
\end{aligned}
$$

- $\mathbf{x} \in \mathbb{R}^n$ — искомый вектор
- $f_0(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ — целевая функция

# Постановка задачи

$$\begin{aligned} &\min_{\mathbf{x} \in X} f_0(\mathbf{x}) \\ \text{s.t. } & f_i(\mathbf{x}) = 0, \; i = 1, \dots, p \\ & f_j(\mathbf{x}) \leq 0, \; j = p+1, \dots, m, \end{aligned}$$

- $\mathbf{x} \in \mathbb{R}^n$ — искомый вектор
- $f_0(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ — целевая функция
- $f_k(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ — функции ограничений

# Постановка задачи

$$\begin{aligned} &\min_{\mathbf{x} \in X} f_0(\mathbf{x}) \\ \text{s.t. } & f_i(\mathbf{x}) = 0, \ i = 1, \ldots, p \\ & f_j(\mathbf{x}) \leq 0, \ j = p+1, \ldots, m, \end{aligned}$$

- $\mathbf{x} \in \mathbb{R}^n$ — искомый вектор
- $f_0(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ — целевая функция
- $f_k(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ — функции ограничений

Пример: выбор объектов для вложения денег и определение в какой объект сколько вкладывать

- $\mathbf{x}$ — размер инвестиций в каждый актив
- $f_0$ — суммарный риск или вариация прибыли
- $f_k$ — бюджетные ограничения, min/max вложения в актив, минимально допустимая прибыль

# Как решать?

В общем случае:

- NP-полные
- рандомизированные алгоритмы: время vs. стабильность

НО определённые классы задач могут быть решены быстро!

- Линейное программирование
- Метод наименьших квадратов
- Малоранговое приближение порядка $k$
- Выпуклая оптимизация

# История развития

- 1940-ые — линейное программирование
- 1950-ые — квадратичное программирование
- 1960-ые — геометрическое программирование
- 1990-ые — полиномиальные методы внутренней точки для произвольной задачи выпуклой оптимизации

# Современные направления

- Решение задач огромной размерности $(\sim 10^8 - 10^{12})$

# Современные направления

- Решение задач огромной размерности $(\sim 10^8 - 10^{12})$
- Распределённая оптимизация

# Современные направления

- Решение задач огромной размерности $(\sim 10^8 - 10^{12})$
- Распределённая оптимизация
- Быстрые методы первого порядка

# Современные направления

- Решение задач огромной размерности $(\sim 10^8 - 10^{12})$
- Распределённая оптимизация
- Быстрые методы первого порядка
- Стохастические алгоритмы: масштабируемость vs. точности

# Современные направления

- Решение задач огромной размерности ($\sim 10^8 - 10^{12}$)
- Распределённая оптимизация
- Быстрые методы первого порядка
- Стохастические алгоритмы: масштабируемость vs. точности
- Невыпуклые задачи определённой структуры

# Современные направления

- Решение задач огромной размерности $(\sim 10^8 - 10^{12})$
- Распределённая оптимизация
- Быстрые методы первого порядка
- Стохастические алгоритмы: масштабируемость vs. точности
- Невыпуклые задачи определённой структуры
- Приложения выпуклой оптимизации

# Линейное программирование

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^\mathsf{T} \mathbf{x}$$

$$\text{s.t. } \mathbf{a}_i^\mathsf{T} \mathbf{x} \leq b_i, \; i = 1, \ldots, m$$

- нет аналитического решения
- существуют эффективные алгоритмы
- разработанная технология
- симплекс-метод входит в Top-10 алгоритмов XX века[1]

[1] https://www.siam.org/pdf/news/637.pdf

# Задача наименьших квадратов

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2,$$

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

- имеет аналитическое решение: $\mathbf{x}^* = (\mathbf{A}^\intercal\mathbf{A})^{-1}\mathbf{A}^\intercal\mathbf{b}$
- существуют эффективные алгоритмы
- разработанная технология
- имеет статистическую интерпретацию

# Выпуклая оптимизация

$$\min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x})$$
$$\text{s.t. } f_i(\mathbf{x}) \leq b_i, \; i = 1, \ldots, m$$

- $f_0, f_i$ — выпуклые функции:

$$f(\alpha \mathbf{x}_1 + \beta \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + \beta f(\mathbf{x}_2),$$

где $\alpha, \beta \geq 0$ и $\alpha + \beta = 1$.

# Выпуклая оптимизация

$$\min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x})$$
$$\text{s.t. } f_i(\mathbf{x}) \le b_i, \; i = 1, \ldots, m$$

- $f_0, f_i$ — выпуклые функции:
$$f(\alpha \mathbf{x}_1 + \beta \mathbf{x}_2) \le \alpha f(\mathbf{x}_1) + \beta f(\mathbf{x}_2),$$
где $\alpha, \beta \ge 0$ и $\alpha + \beta = 1$.
- нет аналитического решения

# Выпуклая оптимизация

$$\min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x})$$

$$\text{s.t. } f_i(\mathbf{x}) \leq b_i, \; i = 1, \ldots, m$$

- $f_0, f_i$ — выпуклые функции:

  $$f(\alpha \mathbf{x}_1 + \beta \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + \beta f(\mathbf{x}_2),$$

  где $\alpha, \beta \geq 0$ и $\alpha + \beta = 1$.
- нет аналитического решения
- существуют эффективные алгоритмы

# Выпуклая оптимизация

$$\min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x})$$
$$\text{s.t. } f_i(\mathbf{x}) \leq b_i, \; i = 1, \ldots, m$$

- $f_0, f_i$ — выпуклые функции:

$$f(\alpha \mathbf{x}_1 + \beta \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + \beta f(\mathbf{x}_2),$$

  где $\alpha, \beta \geq 0$ и $\alpha + \beta = 1$.
- нет аналитического решения
- существуют эффективные алгоритмы
- часто сложно «увидеть» задачу выпуклой оптимизации

# Выпуклая оптимизация

$$\min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x})$$
$$\text{s.t. } f_i(\mathbf{x}) \leq b_i, \; i = 1, \ldots, m$$

- $f_0, f_i$ — выпуклые функции:
$$f(\alpha \mathbf{x}_1 + \beta \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + \beta f(\mathbf{x}_2),$$
где $\alpha, \beta \geq 0$ и $\alpha + \beta = 1$.
- нет аналитического решения
- существуют эффективные алгоритмы
- часто сложно «увидеть» задачу выпуклой оптимизации
- существуют приёмы для преобразования задачи к стандартному виду

# Почему выпуклость так важна?

**R. Tyrrell Rockafellar (1935 —)**

The great watershed in optimization is not between linearity and non-linearity, but convexity and non-convexity.

# Почему выпуклость так важна?

**R. Tyrrell Rockafellar (1935 —)**

The great watershed in optimization is not between linearity and non-linearity, but convexity and non-convexity.

- Локальный оптимум является глобальным

# Почему выпуклость так важна?

**R. Tyrrell Rockafellar (1935 —)**

The great watershed in optimization is not between linearity and non-linearity, but convexity and non-convexity.

- Локальный оптимум является глобальным
- Необходимое условие оптимальности является достаточным

# Почему выпуклость так важна?

**R. Tyrrell Rockafellar (1935 —)**

The great watershed in optimization is not between linearity and non-linearity, but convexity and non-convexity.

- Локальный оптимум является глобальным
- Необходимое условие оптимальности является достаточным

Вопросы:

# Почему выпуклость так важна?

**R. Tyrrell Rockafellar (1935 —)**

The great watershed in optimization is not between linearity and non-linearity, but convexity and non-convexity.

- Локальный оптимум является глобальным
- Необходимое условие оптимальности является достаточным

Вопросы:

- Любую ли задачу выпуклой оптимизации можно эффективно решить?

# Почему выпуклость так важна?

**R. Tyrrell Rockafellar (1935 —)**

The great watershed in optimization is not between linearity and non-linearity, but convexity and non-convexity.

- Локальный оптимум является глобальным
- Необходимое условие оптимальности является достаточным

Вопросы:

- Любую ли задачу выпуклой оптимизации можно эффективно решить?
- Можно ли эффективно решить невыпуклые задачи оптимизации?

# Выпуклое множество

### Выпуклое множество

Множество $C$ называется выпуклым, если
$$\forall x_1,\ x_2 \in C, \theta \in [0,1] \to \theta x_1 + (1-\theta)x_2 \in C.$$
$\emptyset$ и $\{x_0\}$ также считаются выпуклыми.

Примеры: $\mathbb{R}^n$, аффинное множество, луч, отрезок.

### Выпуклая комбинация точек

Пусть $x_1, \ldots, x_k \in G$, тогда точка $\theta_1 x_1 + \ldots + \theta_k x_k$ при $\sum\limits_{i=1}^{k} \theta_i = 1,\ \theta_i \geq 0$ называется выпуклой комбинацией точек $x_1, \ldots, x_k$.

### Выпуклая оболочка точек

Множество $\left\{ \sum\limits_{i=1}^{k} \theta_i x_i \mid x_i \in G, \sum\limits_{i=1}^{k} \theta_i = 1, \theta_i \geq 0 \right\}$ называется выпуклой оболочкой множества $G$ и обозначается **conv**(G).

# Операции, сохраняющие выпуклость

- Пересечение любого (конечного или бесконечного) числа выпуклых множеств — выпуклое множество
- Образ аффинного отображения выпуклого множества — выпуклое множество
- Линейная комбинация выпуклых множеств — выпуклое множество
- Декартово произведение выпуклых множеств — выпуклое множество

# Примеры

Проверьте на аффинность и выпуклость следующие множества:

1. Полупространство: $\{\mathbf{x} | \mathbf{a}^\intercal \mathbf{x} \leq c\}$
2. Многоугольник: $\{\mathbf{x} | \mathbf{A}\mathbf{x} \preceq \mathbf{b},\ \mathbf{C}\mathbf{x} = 0\}$
3. Шар по норме в $\mathbb{R}^n$: $B(r, x_c) = \{x \mid \|x - x_c\| \leq r\}$
4. Эллипсоид: $\mathcal{E}(x_c, \mathbf{P}, r) = \{x \mid (x - x_c)^\intercal \mathbf{P}^{-1}(x - x_c) \leq r^2\}$
5. Множество симметричных положительно-определённых матриц: $\mathbf{S}^n_+ = \{\mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathbf{X}^\intercal = \mathbf{X},\ \mathbf{X} \succeq 0\}$
6. $\{\mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathrm{Tr}(\mathbf{X}) = const\}$

# Конус

## Конус (выпуклый)

Множество $C$ называется конусом (выпуклым конусом), если

$$\forall x \in C, \theta \geq 0 \to \theta x \in C$$

$$(\forall x_1, x_2 \in C, \theta_1, \theta_2 \geq 0 \to \theta_1 x_1 + \theta_2 x_2 \in C)$$

Примеры: $\mathbb{R}^n$, аффинное множество, проходящее через 0, луч.

## Коническая (неотрицательная) комбинация точек

Пусть $x_1, \ldots, x_k \in G$, тогда точка $\theta_1 x_1 + \ldots + \theta_k x_k$ при $\theta_i \geq 0$ называется конической (неотрицательной) комбинацией точек $x_1, \ldots, x_k$.

## Коническая оболочка точек

Множество $\left\{ \sum\limits_{i=1}^{k} \theta_i x_i \mid x_i \in G, \theta_i \geq 0 \right\}$ называется конической оболочкой множества $G$ и обозначается **cone**(G).

# Примеры конусов

1. $\mathbf{S}^n_+$
2. Нормальный конус: $\{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\| \leq t\}$
   Для $\ell_2$-нормы называется конусом второго порядка или Лоренцевым конусом

# Резюме

- Организация работы
- Предмет курса по оптимизации
- Общая формулировка оптимизационной задачи
- Классические оптимизационные задачи
- Выпуклые множества
- Конусы