

Методы оптимизации. Семинар 1. Введение.

Александр Катруца

Московский физико-технический институт,
Факультет Управления и Прикладной Математики

2 сентября 2019 г.

Вопросы к студентам

- Имя
- Кафедра
- Знание $\text{T}_\text{E}^{\text{X}}/\text{L}^{\text{A}}\text{T}_\text{E}^{\text{X}}$
- Ожидания от курса

О чём этот курс?

Первый семестр (теоретический):

- Основы выпуклого анализа
- Условия оптимальности
- Теория двойственности
- Линейное программирование: симплекс-метод

О чём этот курс?

Первый семестр (теоретический):

- Основы выпуклого анализа
- Условия оптимальности
- Теория двойственности
- Линейное программирование: симплекс-метод

Второй семестр (практический):

- Методы безусловной минимизации первого и второго порядка
- Методы условной оптимизации
- Барьерные методы
- Оптимальные методы
- ...

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
- Домашние задания примерно раз в месяц

План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашние задания примерно раз в месяц
- Итоговая контрольная в конце семестра
(и промежуточная в середине семестра)

План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашние задания примерно раз в месяц
- Итоговая контрольная в конце семестра (и промежуточная в середине семестра)
- Экзамен в конце семестра. Оценка за курс — среднее арифметическое:

План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашние задания примерно раз в месяц
- Итоговая контрольная в конце семестра (и промежуточная в середине семестра)
- Экзамен в конце семестра. Оценка за курс — среднее арифметическое:
 - оценки за работу в семестре

План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашние задания примерно раз в месяц
- Итоговая контрольная в конце семестра (и промежуточная в середине семестра)
- Экзамен в конце семестра. Оценка за курс — среднее арифметическое:
 - оценки за работу в семестре
 - ответа на экзамене

План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашние задания примерно раз в месяц
- Итоговая контрольная в конце семестра (и промежуточная в середине семестра)
- Экзамен в конце семестра. Оценка за курс — среднее арифметическое:
 - оценки за работу в семестре
 - ответа на экзамене
- Piazza для Q&A

План на семестр

- Семинар и лекция раз в неделю
- Миниконтрольные в начале каждого семинара
- Домашние задания примерно раз в месяц
- Итоговая контрольная в конце семестра (и промежуточная в середине семестра)
- Экзамен в конце семестра. Оценка за курс — среднее арифметическое:
 - оценки за работу в семестре
 - ответа на экзамене
- Piazza для Q&A
- Gradescope для отправки домашних заданий

Основная книга

S. Boyd and L. Vandenberghe *Convex Optimization*

<https://web.stanford.edu/~boyd/cvxbook/>

- К.Ю. Осипенко Лекции
- В.Г. Жадан *Методы оптимизации. Часть 1. Введение в выпуклый анализ и теорию оптимизации*
- R.T. Rockafellar *Convex analysis*

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

Предварительные навыки

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

Основные этапы использования методов оптимизации при решении реальных задач:

1. Определение целевой функции

Основные этапы использования методов оптимизации при решении реальных задач:

1. Определение целевой функции
2. Определение допустимого множества решений

Основные этапы использования методов оптимизации при решении реальных задач:

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

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

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

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

Постановка задачи

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

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

Постановка задачи

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

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

Постановка задачи

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

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

Постановка задачи

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

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

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

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

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

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

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

- 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

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^T \mathbf{x} \\ \text{s.t. } \mathbf{a}_i^T \mathbf{x} \leq b_i, \quad i = 1, \dots, m \end{aligned}$$

- нет аналитического решения
- существуют эффективные алгоритмы
- разработанная технология
- симплекс-метод входит в Топ-10 алгоритмов XX века¹

¹<https://archive.siam.org/pdf/news/637.pdf>

Задача наименьших квадратов

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

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

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

Малоранговое приближение

$$\begin{aligned} \min_{\mathbf{X} \in \mathbb{R}^{m \times n}} \quad & \|\mathbf{A} - \mathbf{X}\|_F \\ \text{s.t.} \quad & \text{rank}(\mathbf{X}) \leq k \end{aligned}$$

Theorem (Eckart–Young, 1993)

Пусть $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ — сингулярное разложение матрицы \mathbf{A} , где $\mathbf{U} = [\mathbf{U}_k, \mathbf{U}_{r-k}] \in \mathbb{R}^{m \times r}$, $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_k, \dots, \sigma_r)$, $\mathbf{V} = [\mathbf{V}_k, \mathbf{V}_{r-k}] \in \mathbb{R}^{n \times r}$ и $r = \text{rank}(\mathbf{A})$. Тогда решение задачи можно записать в виде:

$$\mathbf{X} = \mathbf{U}_k \hat{\mathbf{\Sigma}} \mathbf{V}_k^\top,$$

где $\mathbf{U}_k \in \mathbb{R}^{m \times k}$, $\hat{\mathbf{\Sigma}} = \text{diag}(\sigma_1, \dots, \sigma_k)$, $\mathbf{V}_k \in \mathbb{R}^{n \times k}$.

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x}) \\ & \text{s.t. } f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\ & \mathbf{Ax} = \mathbf{b} \end{aligned}$$

- 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 $\alpha + \beta = 1$.

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x}) \\ & \text{s.t. } f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\ & \quad \mathbf{Ax} = \mathbf{b} \end{aligned}$$

- 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

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x}) \\ & \text{s.t. } f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\ & \mathbf{Ax} = \mathbf{b} \end{aligned}$$

- 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

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x}) \\ & \text{s.t. } f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\ & \mathbf{Ax} = \mathbf{b} \end{aligned}$$

- 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

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x}) \\ & \text{s.t. } f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\ & \mathbf{Ax} = \mathbf{b} \end{aligned}$$

- 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?

Ralph Tyrrell Rockafellar (1935 —)

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

Почему выпуклость так важна?

Ralph 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?

Ralph 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?

Ralph 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?

Ralph 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?

Ralph 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?

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