# Методы оптимизации. Семинар 1. Введение.

Александр Катруца

Московский физико-технический институт,
Факультет Управления и Прикладной Математики

18 августа 2018 г.

# Вопросы к студентам

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

# 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[1]

[1] http://www.cvxpy.org/en/latest/examples/index.html

# О чём этот курс?

Первый семестр (теоретический):

- Основы выпуклого анализа
- Теория двойственности
- Условия оптимальности

Второй семестр (практический):

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

# План на семестр

- Семинар и лекция раз в неделю
- 2 задания
- Итоговая контрольная в конце семестра
  (и промежуточная в середине семестра)
- Экзамен в конце семестра. Среднее арифметическое:
  - оценки за работу в семестре
  - ответа на экзамене
- Миниконтрольные в начале каждого семинара
- Домашнее задание на каждый семинар — $\LaTeX$

# 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
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

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

- $\mathbf{x} \in \mathbb{R}^n$ — 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

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

# Современнные 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

$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^\intercal \mathbf{x}$$
$$\text{s.t. } \mathbf{a}_i^\intercal \mathbf{x} \leq c_i, \;\; i = 1, \ldots, m$$

- нет аналитического решения
- существуют эффективные алгоритмы
- разработанная технология
- симплекс-метод входит в Top-10 алгоритмов XX века[2]

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

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

# Малоранговое приближение ранга $k$

$$\min_{\mathbf{X}\in\mathbb{R}^{m\times n}} \|\mathbf{A}-\mathbf{X}\|_F$$
$$\text{s.t. } \operatorname{rank}(\mathbf{X}) \le k$$

**Theorem (Eckart–Young, 1993)**

*Пусть* $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ *— сингулярное разложение матрицы* $\mathbf{A}$*, где* $\mathbf{U} = [\mathbf{U}_k, \mathbf{U}_{r-k}] \in \mathbb{R}^{m\times r}$, $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \dots, \sigma_k, \dots, \sigma_r)$, $\mathbf{V} = [\mathbf{V}_k, \mathbf{V}_{r-k}] \in \mathbb{R}^{n\times r}$ *и* $r = rank(\mathbf{A})$*. Тогда решение задачи* (14) *можно записать в виде:*

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

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

Алгоритм вычисления сингулярного разложения и быстрый, и устойчивый.

# Выпуклая оптимизация

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

- $f_0, f_i$ — 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 виду (15)

# Какая задача проще?

Поиск независимого под-алфавит максимальной мощности

$$\max_{\mathbf{x}\in\mathbb{R}^n} \sum_{i=1}^{n} x_i$$

$$\text{s.t. } x_i^2 - x_i = 0 \quad i = 1, \dots, n$$
$$x_i x_j = 0 \qquad \forall (i,j) \in \Gamma,$$

где $\Gamma$ — множество пар

Truss design

$$\min -2 \sum_{i=1}^{k} \sum_{j=1}^{m} c_{ij} x_{ij} + x_{00}$$

$$\text{s.t. } \sum_{i=1}^{k} x_i = 1$$
$$\lambda_{\min}(\mathbf{A}) \geq 0,$$

где $\mathbf{A} =$

$$\begin{bmatrix} x_1 & \dots & \dots & \sum\limits_{j=1}^{m} b_{pj} x_{1j} \\ \vdots & \ddots & & \vdots \\ & & x_k & \sum\limits_{j=1}^{m} b_{pj} x_{kj} \\ \sum\limits_{j=1}^{m} b_{pj} x_{1j} & \dots & \sum\limits_{j=1}^{m} b_{pj} x_{kj} & x_{00} \end{bmatrix}$$

# 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?

# Резюме

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