# Методы оптимизации
# Лекция 1: Введение. Выпуклые множества

Александр Катруца

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

5 сентября 2022 г.

# О чём этот курс?

## Теория: сентябрь — середина октября

- Выпуклые множества и функции
- Условия оптимальности
- Основы теории двойственности

## Методы и приложения: середина октября — середина декабря

- Постановки задач оптимизации
- Методы решения задач без ограничений
- Методы решения задач с простыми ограничениями
- Линейное программирование
- Задачи конической оптимизации и SDP

# Организационные вопросы

- Лекция и семинар каждую неделю

# Организационные вопросы

- Лекция и семинар каждую неделю
- Отчётность

# Организационные вопросы

- Лекция и семинар каждую неделю
- Отчётность
- Репозиторий со слайдами лекций:
  https://github.com/amkatrutsa/optimization-fivt

# Организационные вопросы

- Лекция и семинар каждую неделю
- Отчётность
- Репозиторий со слайдами лекций:
  https://github.com/amkatrutsa/optimization-fivt
- Репозиторий со старыми семинарами:
  https://github.com/amkatrutsa/seminars-fivt

# Отзывы о курсе год назад

| Как Вы оцениваете лекции по курсу Методы оптимизации? | Что Вам больше всего понравилось в лекциях? | Что Вам больше всего НЕ понравилось в лекциях? |
|---|---|---|
| 7 | Мнгомерное дифференцирование | С презентации доказательства не заходят, хочется на доске расписывать |
| 8 | Структурированность изложения | Хотелось бы побольше доказательств  и теор. фактов |
| 9 | Всё прикреплялось примерами, что способствовало лёгкому восприятию темы. | Большая часть лекций было онлайн. |
| 3 | Не смотрел | Не смотрел |
| 9 | Лектор хорошо взаимодействует с аудиторией и добивается понимания материала | Ничего критичного (мб, что были ошибки в слайдах, но их исправили и всё супер👍 ) |
| 9 | Примеры задач | Все понравилось |

# Что нового будет в этом году?

- Больше примеров задач: оптимизация на многообразиях, робастная оптимизация

# Что нового будет в этом году?

- Больше примеров задач: оптимизация на многообразиях, робастная оптимизация
- Больше методов решения: специфика использования смешанной точности

# Что нового будет в этом году?

- Больше примеров задач: оптимизация на многообразиях, робастная оптимизация
- Больше методов решения: специфика использования смешанной точности
- Меньше времени уделим долгим доказательствам

# Что нового будет в этом году?

- Больше примеров задач: оптимизация на многообразиях, робастная оптимизация
- Больше методов решения: специфика использования смешанной точности
- Меньше времени уделим долгим доказательствам
- Больше примеров комбинации теоретических результатов и пакетов для решения задач: cvxpy, pymanopt, apex, etc

# Литература

Основная книга
S. Boyd and L. Vandenberghe *Convex Optimization*
https://web.stanford.edu/~boyd/cvxbook/

## Теория

- Ю.Е. Нестеров *Введение в выпуклую оптимизацию*
- A. Nemirovski *Lecture notes on Modern Convex Optimization*
- S. Bubeck *Convex Optimization: Algorithms and Complexity*
- R. T. Rockafellar *Convex analysis*

## Методы и приложения

- J. Nocedal, S. J. Wright *Numerical Optimization*
- P. E. Gill, W. Murray, M. H. Wright *Practical optimization*
- Б.Т. Поляк *Введение в оптимизацию*

# 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
  - 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
  - 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
  - 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
  - выбор активов (portfolio optimization)
  - оптимальное управление
  - обработка сигналов
  - робототехника

# Зачем этот курс?

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

# 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

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

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

# 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

# Определения решений

# Определения решений

## Определение

*Точка* $\mathbf{x}^*$ *называется точкой* ***глобального*** *минимума, если* $f(\mathbf{x}) \geq f(\mathbf{x}^*)$ *для всех* $\mathbf{x}$ *из допустимого множества.*

# Определения решений

## Определение

*Точка* $\mathbf{x}^*$ *называется точкой* ***глобального*** *минимума, если* $f(\mathbf{x}) \geq f(\mathbf{x}^*)$ *для всех* $\mathbf{x}$ *из допустимого множества.*

## Определение

*Точка* $\mathbf{x}^*$ *называется точкой* ***локального*** *минимума, если* $f(\mathbf{x}) \geq f(\mathbf{x}^*)$ *для всех* $\mathbf{x}$ *из окрестности точки* $\mathbf{x}^*$ *и допустимого множества.*

# Определения решений

## Определение

*Точка* $\mathbf{x}^*$ *называется точкой* ***глобального*** *минимума, если* $f(\mathbf{x}) \geq f(\mathbf{x}^*)$ *для всех* $\mathbf{x}$ *из допустимого множества.*

## Определение

*Точка* $\mathbf{x}^*$ *называется точкой* ***локального*** *минимума, если* $f(\mathbf{x}) \geq f(\mathbf{x}^*)$ *для всех* $\mathbf{x}$ *из окрестности точки* $\mathbf{x}^*$ *и допустимого множества.*

## Альтернативная запись задачи

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

# Как решать?

В общем случае:

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

# Как решать?

В общем случае:

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

НО определённые классы задач могут быть решены быстро!

# Как решать?

В общем случае:

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

НО определённые классы задач могут быть решены быстро!

- Линейное программирование
- Задача наименьших квадратов
- Задача о малоранговом приближении матрицы
- Выпуклая оптимизация

# История развития

- 1940-ые — линейное программирование
- 1950-ые — квадратичное программирование
- 1960-ые — геометрическое программирование
- 1990-ые — полиномиальные методы внутренней точки для задач конической оптимизации
- 2000-ые — релаксации комбинаторных задач и робастная оптимизация
- 2010-ые — стохастические методы, невыпуклые задачи и глобальная оптимизация

# 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 высокого порядка

## Современные направления

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

# Современные направления

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

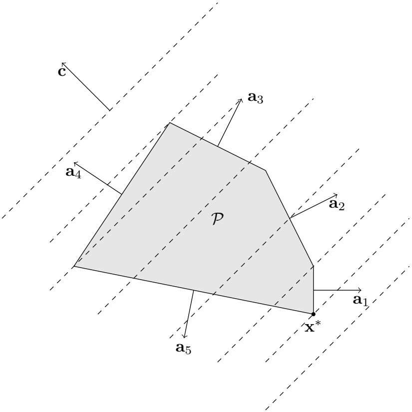
# Современные направления

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

# Линейное программирование (linear programming)

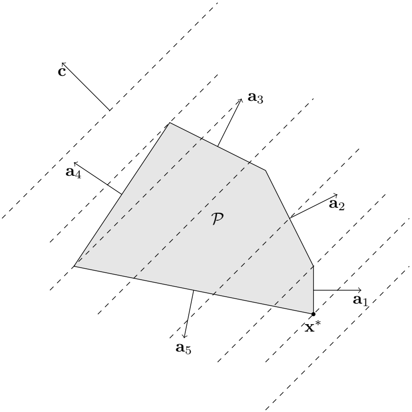
$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^\top \mathbf{x}$$

$$\text{s.t. } \mathbf{a}_i^\top \mathbf{x} \le b_i,\ i = 1, \dots, m$$



# Линейное программирование (linear programming)

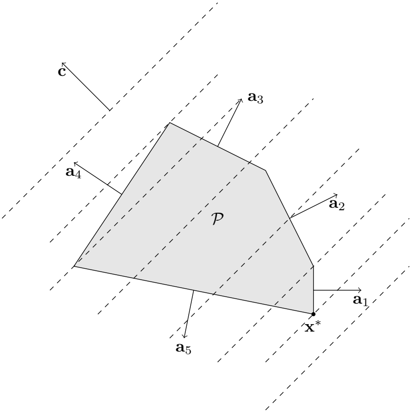
$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^\top \mathbf{x}$$
$$\text{s.t. } \mathbf{a}_i^\top \mathbf{x} \le b_i,\ i = 1, \ldots, m$$




- нет аналитического решения

# Линейное программирование (linear programming)

$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^\top \mathbf{x}$$
$$\text{s.t. } \mathbf{a}_i^\top \mathbf{x} \le b_i,\ i = 1, \dots, m$$

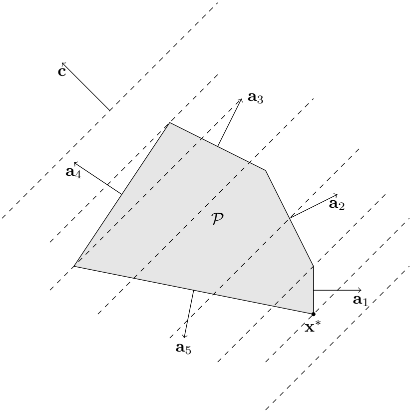



- нет аналитического решения
- существуют эффективные алгоритмы

# Линейное программирование (linear programming)

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

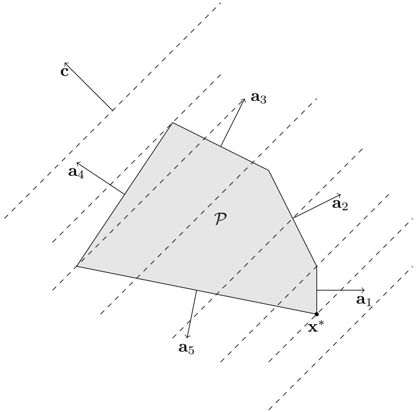
$$\text{s.t. } \mathbf{a}_i^\top \mathbf{x} \leq b_i, \; i = 1, \ldots, m$$




- нет аналитического решения
- существуют эффективные алгоритмы
- разработанная технология

# Линейное программирование (linear programming)

$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^\top \mathbf{x}$$
$$\text{s.t. } \mathbf{a}_i^\top \mathbf{x} \le b_i,\ i = 1, \dots, m$$

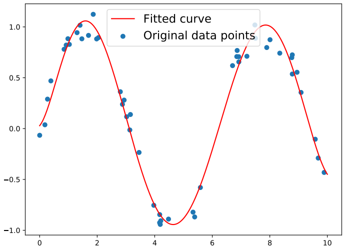



- нет аналитического решения
- существуют эффективные алгоритмы
- разработанная технология
- симплекс-метод для решения таких задач входит в Top-10 алгоритмов XX века

# Задача наименьших квадратов (linear least squares problem)

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

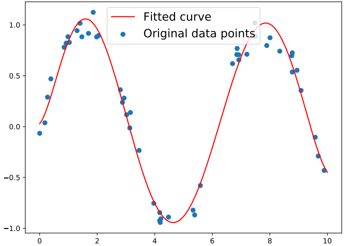
где $\mathbf{A} \in \mathbb{R}^{m \times n}$ и $\mathbf{b} \in \mathbb{R}^m$.



# Задача наименьших квадратов (linear least squares problem)

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

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

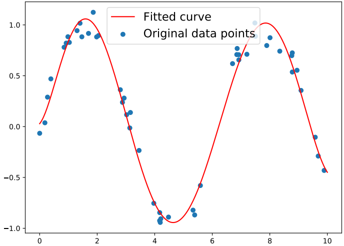



- имеет аналитическое решение: $\mathbf{x}^* = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{b}$

# Задача наименьших квадратов (linear least squares problem)

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

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

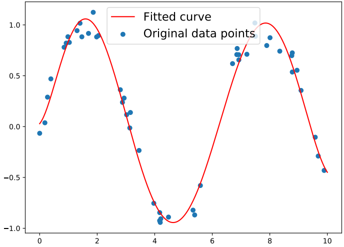



- имеет аналитическое решение: $\mathbf{x}^* = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}$
- существуют эффективные алгоритмы

# Задача наименьших квадратов (linear least squares problem)

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

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

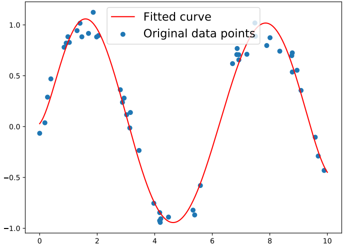



- имеет аналитическое решение: $\mathbf{x}^* = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}$
- существуют эффективные алгоритмы
- разработанная технология

# Задача наименьших квадратов (linear least squares problem)

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

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

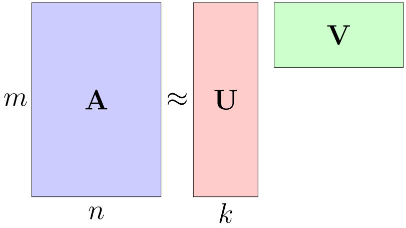



- имеет аналитическое решение: $\mathbf{x}^* = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{b}$
- существуют эффективные алгоритмы
- разработанная технология
- имеет статистическую интерпретацию

# Малоранговое приближение (low-rank approximation)

$$\min_{\mathbf{X}\in\mathbb{R}^{m\times n}} \|\mathbf{A}-\mathbf{X}\|_F$$

$$\text{s.t. rank}(\mathbf{X}) \le k$$



# Эта задача имеет аналитическое решение

**Теорема (Eckart–Young, 1993)**

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

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

*где* $\hat{\mathbf{\Sigma}} = \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$*.*

# Сжатие: toy problem


- Изображение $493 \times 700 \times 3$
- Каков эффективный ранг матрицы для каждого цвета?

# Сжатие: $k = 10$

- Коэффициент сжатия $\frac{3\times(493\times10+10+10\times700)}{493\times700\times3} = 0.035$

# Сжатие: $k = 50$


- Коэффициент сжатия $\frac{3\times(493\times 50+50+50\times 700)}{493\times 700\times 3} = 0.173$

# Сжатие: $k = 100$

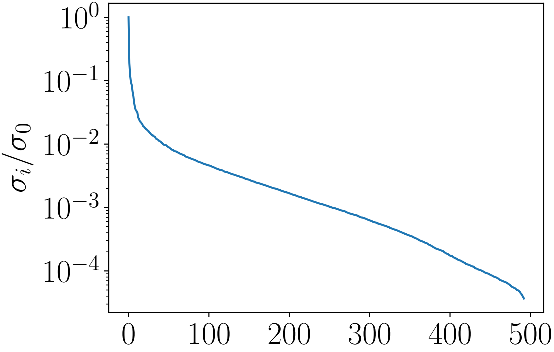

- Коэффициент сжатия $\frac{3\times(493\times100+100+100\times700)}{493\times700\times3} = 0.346$

# Сжатие: $k = 150$


- Коэффициент сжатия $\frac{3\times(493\times150+150+150\times700)}{493\times700\times3} = 0.519$

# Определение ранга

- Убывание сингулярных чисел связано с ошибкой аппроксимации
- Выбор ранга по величине сингулярного числа $\sigma_k$




Код доступен тут

# Сжатие: real problem

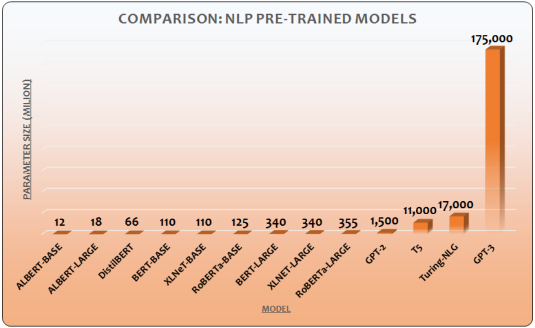



График взят отсюда

- Современные нейросети содержат миллиарды оптимизируемых параметров

# Сжатие: real problem

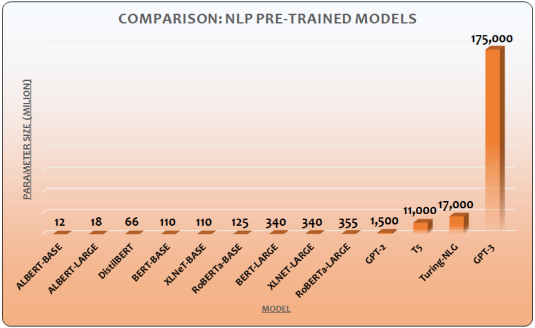



График взят отсюда

- Современные нейросети содержат миллиарды оптимизируемых параметров
- Для масштабирования моделей необходимо снижать внутреннюю размерность слоёв

# Сжатие: real problem

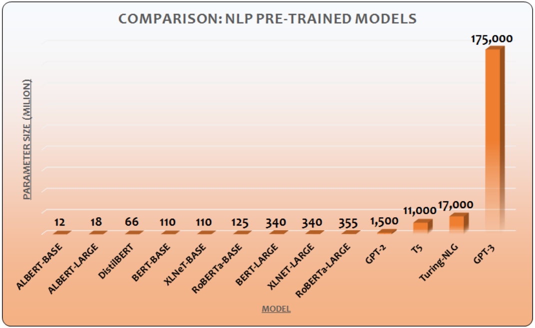



График взят отсюда

- Современные нейросети содержат миллиарды оптимизируемых параметров
- Для масштабирования моделей необходимо снижать внутреннюю размерность слоёв
- Многие методы сжатия основаны на малоранговой аппроксимации, см обзор

# Выпуклая оптимизация (convex optimization)

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

# Выпуклая оптимизация (convex optimization)

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

- $f_0, f_i$ — 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 $\alpha + \beta = 1$.

# Выпуклая оптимизация (convex optimization)

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

- $f_0, f_i$ — 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

# Выпуклая оптимизация (convex optimization)

$$\min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x})$$
$$\text{s.t. } f_i(\mathbf{x}) \le 0, \ i = 1, \dots, m$$
$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- $f_0, f_i$ — 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
- существуют эффективные алгоритмы

# Выпуклая оптимизация (convex optimization)

$$\min_{\mathbf{x}\in\mathbb{R}^n} f_0(\mathbf{x})$$
$$\text{s.t. } f_i(\mathbf{x}) \le 0,\ i = 1,\dots,m$$
$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- $f_0, f_i$ — 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
- существуют эффективные алгоритмы
- часто сложно «увидеть» задачу выпуклой оптимизации

# Выпуклая оптимизация (convex optimization)

$$\min_{\mathbf{x}\in\mathbb{R}^n} f_0(\mathbf{x})$$
$$\text{s.t. } f_i(\mathbf{x}) \le 0,\ i = 1, \dots, m$$
$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

- $f_0, f_i$ — 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
- существуют эффективные алгоритмы
- часто сложно «увидеть» задачу выпуклой оптимизации
- существуют приёмы для преобразования задачи к стандартному виду

# Почему выпуклость так важна?

### Ralph Tyrrell Rockafellar (born 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 (born 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 (born 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 (born 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 (born 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 (born 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?

# Структура задачи

- От простого к сложному

## Структура задачи

- От простого к сложному
- От общего к частному

# Структура задачи

- От простого к сложному
- От общего к частному
- Чем больше информации о задаче вам известно, тем быстрее вы можете её решить

# Структура задачи

- От простого к сложному
- От общего к частному
- Чем больше информации о задаче вам известно, тем быстрее вы можете её решить
- Распараллеленные модификации методов могут существенно ускорить процесс решения задачи

# Выпуклые множества (convex sets)

**Определение**

*Множество* $C \subseteq \mathbb{R}^n$ *называется выпуклым, если для всех* $\alpha \in [0, 1]$ *и любых* $x, y \in C$ *выполнено*

$$\alpha x + (1 - \alpha) y \in C.$$

# Выпуклые множества (convex sets)

## Определение

*Множество* $C \subseteq \mathbb{R}^n$ *называется выпуклым, если для всех* $\alpha \in [0, 1]$ *и любых* $x, y \in C$ *выполнено*

$$\alpha x + (1 - \alpha) y \in C.$$

## Примеры

- Многоугольники

# Выпуклые множества (convex sets)

## Определение

*Множество $C \subseteq \mathbb{R}^n$ называется выпуклым, если для всех $\alpha \in [0, 1]$ и любых $x, y \in C$ выполнено*

$$\alpha x + (1 - \alpha) y \in C.$$

## Примеры

- Многоугольники
- Гиперплоскости

# Выпуклые множества (convex sets)

## Определение

*Множество* $C \subseteq \mathbb{R}^n$ *называется выпуклым, если для всех* $\alpha \in [0, 1]$ *и любых* $x, y \in C$ *выполнено*

$$\alpha x + (1 - \alpha) y \in C.$$

## Примеры

- Многоугольники
- Гиперплоскости
- Шары в *любой* норме и эллипсоиды

# Выпуклые множества (convex sets)

## Определение

*Множество* $C \subseteq \mathbb{R}^n$ *называется выпуклым, если для всех* $\alpha \in [0, 1]$ *и любых* $x, y \in C$ *выполнено*

$$\alpha x + (1 - \alpha) y \in C.$$

## Примеры

- Многоугольники
- Гиперплоскости
- Шары в *любой* норме и эллипсоиды
- Симметричные положительно определённые матрицы

# Пересечение выпуклых множеств

## Теорема

*Пересечение конечного или бесконечного числа выпуклых множеств $C_i$ является выпуклым множеством:*

$$C = \bigcap_{i \in \mathcal{I}} C_i.$$

# Пересечение выпуклых множеств

## Теорема

*Пересечение конечного или бесконечного числа выпуклых множеств $C_i$ является выпуклым множеством:*

$$C = \bigcap_{i \in \mathcal{I}} C_i.$$

## Доказательство

- Рассмотрим $x, y \in C \to x, y \in C_i, \forall i \in \mathcal{I}$

# Пересечение выпуклых множеств

### Теорема

*Пересечение конечного или бесконечного числа выпуклых множеств $C_i$ является выпуклым множеством:*

$$C = \bigcap_{i \in \mathcal{I}} C_i.$$

### Доказательство

- Рассмотрим $x, y \in C \to x, y \in C_i, \forall i \in \mathcal{I}$
- Построим точку $z = \alpha x + (1 - \alpha) y$, $\alpha \in [0, 1]$

# Пересечение выпуклых множеств

## Теорема

*Пересечение конечного или бесконечного числа выпуклых множеств $C_i$ является выпуклым множеством:*

$$C = \bigcap_{i \in \mathcal{I}} C_i.$$

## Доказательство

- Рассмотрим $x, y \in C \to x, y \in C_i, \forall i \in \mathcal{I}$
- Построим точку $z = \alpha x + (1 - \alpha) y$, $\alpha \in [0, 1]$
- Так как все $C_i$ выпуклы, то $z \in C_i$, $\forall i \in \mathcal{I}$

# Пересечение выпуклых множеств

## Теорема

*Пересечение конечного или бесконечного числа выпуклых множеств $C_i$ является выпуклым множеством:*

$$C = \bigcap_{i \in \mathcal{I}} C_i.$$

## Доказательство

- Рассмотрим $x, y \in C \to x, y \in C_i, \forall i \in \mathcal{I}$
- Построим точку $z = \alpha x + (1 - \alpha) y$, $\alpha \in [0, 1]$
- Так как все $C_i$ выпуклы, то $z \in C_i$, $\forall i \in \mathcal{I}$
- Следовательно, $z \in C$ и $C$ — выпукло

# Линейное отображение выпуклых множеств

### Теорема

*Образ выпуклого множества при линейном отображении является выпуклым множеством.*

### Доказательство

- Пусть $C$ — выпуклое множество и $x, y \in C$

# Линейное отображение выпуклых множеств

## Теорема

*Образ выпуклого множества при линейном отображении является выпуклым множеством.*

## Доказательство

- Пусть $C$ — выпуклое множество и $x, y \in C$
- Пусть $f$ — линейное отображение вида $f(x) = Ax + b$

# Линейное отображение выпуклых множеств

### Теорема

*Образ выпуклого множества при линейном отображении является выпуклым множеством.*

### Доказательство

- Пусть $C$ — выпуклое множество и $x, y \in C$
- Пусть $f$ — линейное отображение вида $f(x) = Ax + b$
- Покажем, что $\alpha f(x) + (1 - \alpha) f(y) \in f(C)$, где $\alpha \in [0, 1]$

# Линейное отображение выпуклых множеств

### Теорема

*Образ выпуклого множества при линейном отображении является выпуклым множеством.*

### Доказательство

- Пусть $C$ — выпуклое множество и $x, y \in C$
- Пусть $f$ — линейное отображение вида $f(x) = Ax + b$
- Покажем, что $\alpha f(x) + (1-\alpha) f(y) \in f(C)$, где $\alpha \in [0, 1]$
- Действительно,

$$\alpha f(x) + (1-\alpha) f(y) = \alpha(Ax + b) + (1-\alpha)(Ay + b) =$$
$$A(\alpha x + (1-\alpha)y) + b = Az + b = f(z),$$

где $z = \alpha x + (1-\alpha) y \in C$

# Арифметические операции над выпуклыми множествами

Теорема

*Сумма Минковского выпуклых множеств является выпуклым множеством.*

# Арифметические операции над выпуклыми множествами

### Теорема

*Сумма Минковского выпуклых множеств является выпуклым множеством.*

### Доказательство

- Пусть $C_1, C_2$ — выпуклые множества. Рассмотрим $C = C_1 + C_2 = \{x_1 + x_2 \mid x_1 \in C_1,\ x_2 \in C_2\}$

# Арифметические операции над выпуклыми множествами

## Теорема

*Сумма Минковского выпуклых множеств является выпуклым множеством.*

## Доказательство

- Пусть $C_1, C_2$ — выпуклые множества. Рассмотрим $C = C_1 + C_2 = \{x_1 + x_2 \mid x_1 \in C_1,\ x_2 \in C_2\}$
- Пусть $\hat{x} = \hat{x}_1 + \hat{x}_2$ и $\tilde{x} = \tilde{x}_1 + \tilde{x}_2$ лежат в $C$. Покажем, что в $C$ лежит точка $\alpha\hat{x} + (1-\alpha)\tilde{x}$

# Арифметические операции над выпуклыми множествами

## Теорема

*Сумма Минковского выпуклых множеств является выпуклым множеством.*

## Доказательство

- Пусть $C_1, C_2$ — выпуклые множества. Рассмотрим $C = C_1 + C_2 = \{x_1 + x_2 \mid x_1 \in C_1,\ x_2 \in C_2\}$
- Пусть $\hat{x} = \hat{x}_1 + \hat{x}_2$ и $\tilde{x} = \tilde{x}_1 + \tilde{x}_2$ лежат в $C$. Покажем, что в $C$ лежит точка $\alpha\hat{x} + (1-\alpha)\tilde{x}$
- Действительно,

  $\alpha\hat{x}+(1-\alpha)\tilde{x} = [\alpha\hat{x}_1+(1-\alpha)\tilde{x}_1]+[\alpha\hat{x}_2+(1-\alpha)\tilde{x}_2] = y_1+y_2,$

  где $y_1 \in C_1$ и $y_2 \in C_2$ в силу выпуклости множеств $C_1, C_2$.

# Арифметические операции над выпуклыми множествами

## Теорема

*Сумма Минковского выпуклых множеств является выпуклым множеством.*

## Доказательство

- Пусть $C_1, C_2$ — выпуклые множества. Рассмотрим $C = C_1 + C_2 = \{x_1 + x_2 \mid x_1 \in C_1,\ x_2 \in C_2\}$
- Пусть $\hat{x} = \hat{x}_1 + \hat{x}_2$ и $\tilde{x} = \tilde{x}_1 + \tilde{x}_2$ лежат в $C$. Покажем, что в $C$ лежит точка $\alpha\hat{x} + (1-\alpha)\tilde{x}$
- Действительно,
  $\alpha\hat{x}+(1-\alpha)\tilde{x} = [\alpha\hat{x}_1+(1-\alpha)\tilde{x}_1]+[\alpha\hat{x}_2+(1-\alpha)\tilde{x}_2] = y_1+y_2,$
  где $y_1 \in C_1$ и $y_2 \in C_2$ в силу выпуклости множеств $C_1, C_2$.

## Следствие

Линейная комбинация выпуклых множеств — выпуклое множество

# Конусы (cones)

### Определение

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

### Определение

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

# Конусы (cones)

### Определение

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

### Определение

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

### Важные конусы

# Конусы (cones)

### Определение

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

### Определение

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

### Важные конусы

- Неотрицательный октант
  $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0,\ i = 1, \ldots, n\} \rightarrow$ LP

# Конусы (cones)

### Определение

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

### Определение

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

### Важные конусы

- Неотрицательный октант
  $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0,\ i = 1, \dots, n\} \to$ LP
- Конус второго порядка $\{(x, t) \in \mathbb{R}^{n+1} \mid \|x\|_2 \leq t\} \to$ SOCP

# Конусы (cones)

## Определение

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

## Определение

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

## Важные конусы

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

# Выпуклая оболочка (convex hull)

## Определение

*Выпуклой оболочкой множества* $X$ *называется такое множество* $\mathrm{conv}(X)$, *что*

- *оно является пересечением всех выпуклых множеств, содержащих* $X$
- *оно содержит все выпуклые комбинации точек из* $X$

$$\mathrm{conv}(X) = \left\{ \sum_{i=1}^{k} \theta_i x_i \mid x_i \in X, \sum_{i=1}^{k} \theta_i = 1, \theta_i \geq 0 \right\}$$

- *оно является минимальным по включению выпуклым множеством, содержащим* $X$

# Использование выпуклых оболочек

- При постановке задачи допустимое множество получилось **невыпуклым**

# Использование выпуклых оболочек

- При постановке задачи допустимое множество получилось **невыпуклым**
- Можно заменить само множество его выпуклой оболочкой

# Использование выпуклых оболочек

- При постановке задачи допустимое множество получилось **невыпуклым**
- Можно заменить само множество его выпуклой оболочкой
- Решить задачу на этом множестве

# Использование выпуклых оболочек

- При постановке задачи допустимое множество получилось **невыпуклым**
- Можно заменить само множество его выпуклой оболочкой
- Решить задачу на этом множестве
- Восстановить некоторым образом приближённое решение из исходной области

# Главное

- Постановки задач оптимизации: целевая функция, допустимое множество, ограничения

# Главное

- Постановки задач оптимизации: целевая функция, допустимое множество, ограничения
- Примеры задач оптимизации и приложения

# Главное

- Постановки задач оптимизации: целевая функция, допустимое множество, ограничения
- Примеры задач оптимизации и приложения
- Выпуклые множества

# Главное

- Постановки задач оптимизации: целевая функция, допустимое множество, ограничения
- Примеры задач оптимизации и приложения
- Выпуклые множества
- Способы определения является ли множество выпуклым

# Литература


Suvrit Sra, Sebastian Nowozin, and Stephen J Wright.
*Optimization for machine learning*.
MIT Press, 2012.

Tamás Terlaky, Miguel F Anjos, and Shabbir Ahmed.
*Advances and trends in optimization with engineering applications*.
SIAM, 2017.

Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski.
*Robust optimization*, volume 28.
Princeton university press, 2009.

Amir Beck.
*First-order methods in optimization*.
SIAM, 2017.

P-A Absil, Robert Mahony, and Rodolphe Sepulchre.
Optimization algorithms on matrix manifolds.
In *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2009.