Методы оптимизации
Лекция 1: Введение. Выпуклые множества

Александр Катруца

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

7 сентября 2020 г.

О чём этот семестр?

Основы выпуклого анализа

- Выпуклые множества и функции
- Субградиенты и субдифференциалы
- Конусы

Условия оптимальности и двойственность

- Условия Каруша-Куна-Таккера
- Основы теории двойственности и её применение
- Линейное программирование

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-fupm

Литература

Основная книга

S. Boyd and L. Vandenberghe *Convex Optimization*

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

- A. Nemirovski *Lecture notes on Modern Convex Optimization*
- R. T. Rockafellar *Convex analysis*
- В. Г. Жадан *Методы оптимизации. Часть 1. Введение в выпуклый анализ и теорию оптимизации: учебное пособие*

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

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

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

Постановка задачи

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

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

Постановка задачи

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

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

Постановка задачи

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

- $\mathbf{x} \in \mathbb{R}^n$ — 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

Пример: построение линейного классификатора

- Дана выборка: $(\mathbf{x}_i, y_i)$,
  $\mathbf{x}_i \in \mathbb{R}^n,\ y_i = \{+1, -1\},\ i = 1, \ldots, m$

Пример: построение линейного классификатора

- Дана выборка: $(\mathbf{x}_i, y_i)$,
  $\mathbf{x}_i \in \mathbb{R}^n,\ y_i = \{+1, -1\},\ i = 1, \ldots, m$
- Линейный классификатор $\hat{y} = \mathrm{sign}(\mathbf{w}^\top \mathbf{x} + b)$

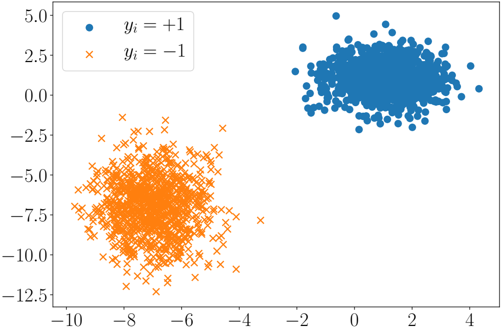
Пример: построение линейного классификатора

- Дана выборка: $(\mathbf{x}_i, y_i)$,
  $\mathbf{x}_i \in \mathbb{R}^n,\ y_i = \{+1, -1\},\ i = 1, \dots, m$
- Линейный классификатор $\hat{y} = \operatorname{sign}(\mathbf{w}^\top \mathbf{x} + b)$
- $\begin{cases} \mathbf{w}^\top \mathbf{x}_i + b > 1, & y_i = +1 \\ \mathbf{w}^\top \mathbf{x}_i + b < -1, & y_i = -1 \end{cases}$

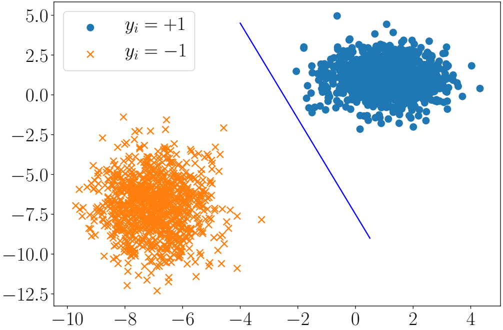
Пример: построение линейного классификатора

- Дана выборка: $(\mathbf{x}_i, y_i)$,
  $\mathbf{x}_i \in \mathbb{R}^n,\ y_i = \{+1, -1\},\ i = 1, \dots, m$
- Линейный классификатор $\hat{y} = \mathrm{sign}(\mathbf{w}^\top \mathbf{x} + b)$
- $\begin{cases} \mathbf{w}^\top \mathbf{x}_i + b > 1, & y_i = +1 \\ \mathbf{w}^\top \mathbf{x}_i + b < -1, & y_i = -1 \end{cases}$
- $y_i(\mathbf{w}^\top \mathbf{x}_i + b) > 1$

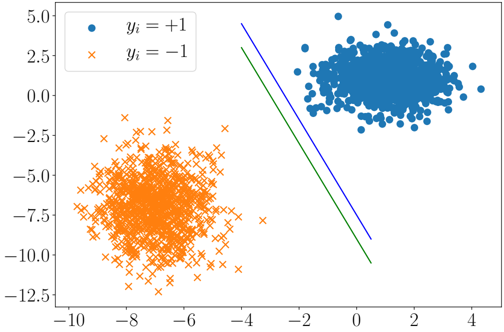
Возможные решения



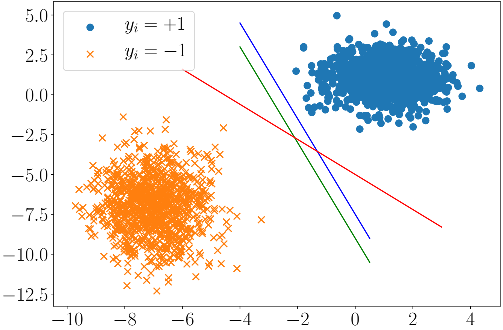
Возможные решения



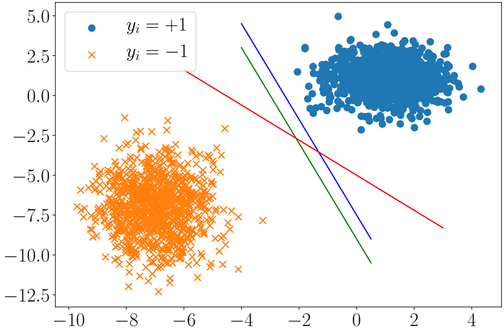
Возможные решения



Возможные решения



Возможные решения



Q: Как однозначно задать разделяющую гиперплоскость?

Максимизация зазора

Максимизация зазора

- Для опорных объектов каждого класса выполнено

$$\begin{cases} \mathbf{w}^\top \mathbf{x}_k + b = 1, & y_k = +1 \\ \mathbf{w}^\top \mathbf{x}_j + b = -1, & y_j = -1 \end{cases}$$

Максимизация зазора

- Для опорных объектов каждого класса выполнено

$$\begin{cases} \mathbf{w}^\top \mathbf{x}_k + b = 1, & y_k = +1 \\ \mathbf{w}^\top \mathbf{x}_j + b = -1, & y_j = -1 \end{cases}$$

- Расстояние между гиперплоскостями

$$d = \frac{|c_1 - c_2|}{\|\mathbf{w}\|_2} = \frac{2}{\|\mathbf{w}\|_2}$$

Максимизация зазора

- Для опорных объектов каждого класса выполнено

$$\begin{cases} \mathbf{w}^\top \mathbf{x}_k + b = 1, & y_k = +1 \\ \mathbf{w}^\top \mathbf{x}_j + b = -1, & y_j = -1 \end{cases}$$

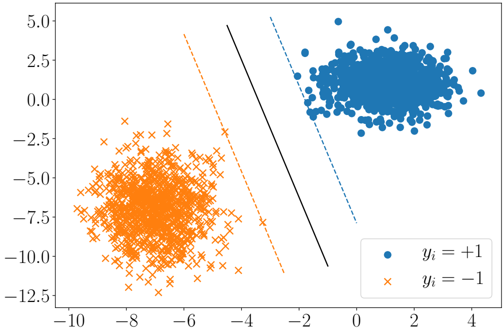
- Расстояние между гиперплоскостями

$$d = \frac{|c_1 - c_2|}{\|\mathbf{w}\|_2} = \frac{2}{\|\mathbf{w}\|_2}$$

Финальная задача

$$\min_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|_2^2$$
$$\text{s.t. } y_i(\mathbf{w}^\top \mathbf{x}_i + b) > 1,\ i = 1, \ldots, m$$

Оптимальная гиперплоскость



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

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

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

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 тензорные методы

Современные направления

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

Современные направления

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

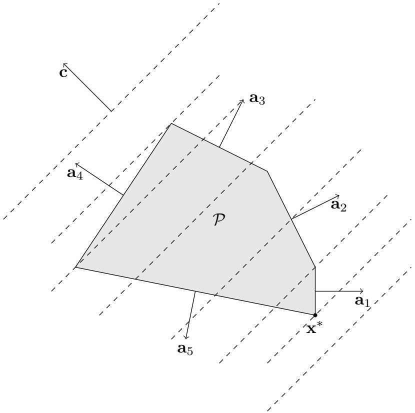
Современные направления

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

Линейное программирование (linear programming)

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

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



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

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

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)

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

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

Малоранговое приближение (low-rank approximation)

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

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

Теорема (Eckart–Young, 1993)

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

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

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

Сжатие

- 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 $\frac{3\times(493\times50+50+50\times700)}{493\times700\times3} = 0.173$

Сжатие: $k = 100$

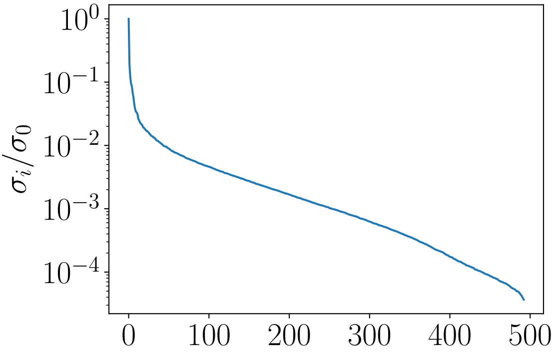
- Коэффициент сжатия $\frac{3\times(493\times100+100+100\times700)}{493\times700\times3} = 0.346$

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 σ_k



- Некоторые авторы предлагают брать порог равным $4/\sqrt{3}$

Выпуклая оптимизация (convex optimization)

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

Выпуклая оптимизация (convex optimization)

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

- 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 $\alpha + \beta = 1$.

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

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)

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

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

Выпуклая оптимизация (convex optimization)

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

- f_0, f_i — выпуклые функции:

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

  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 (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?

Выпуклые множества (convex sets)

Определение

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

$$\alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \in \mathcal{X}.$$

Выпуклые множества (convex sets)

Определение

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

$$\alpha \mathbf{x} + (1 - \alpha)\mathbf{y} \in \mathcal{X}.$$

Примеры

- Многоугольники

Выпуклые множества (convex sets)

Определение

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

$$\alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \in \mathcal{X}.$$

Примеры

- Многоугольники
- Гиперплоскости

Выпуклые множества (convex sets)

Определение

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

$$\alpha \mathbf{x} + (1 - \alpha)\mathbf{y} \in \mathcal{X}.$$

Примеры

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

Выпуклые множества (convex sets)

Определение

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

$$\alpha \mathbf{x} + (1 - \alpha)\mathbf{y} \in \mathcal{X}.$$

Примеры

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

Свойство выпуклых множеств

Утверждение

Если множество $\mathcal{X}$ выпукло, то все точки вида $\sum_{i=1}^{k} \alpha_i \mathbf{x}_i$, где $\mathbf{x}_i \in \mathcal{X}$ и $\alpha \in \Delta_k = \{\alpha \mid \alpha_i \geq 0, \; \sum_{i=1}^{k} \alpha_i = 1\}$, также лежат в этом множестве. Точки такого вида называются выпуклыми комбинациями точек $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

Доказательство по индукции

- База индукции: при $k = 1$ получаем сами точки множества.

Свойство выпуклых множеств

Утверждение

Если множество $\mathcal{X}$ выпукло, то все точки вида $\sum_{i=1}^{k} \alpha_i \mathbf{x}_i$, где $\mathbf{x}_i \in \mathcal{X}$ и $\alpha \in \Delta_k = \{\alpha \mid \alpha_i \geq 0, \ \sum_{i=1}^{k} \alpha_i = 1\}$, также лежат в этом множестве. Точки такого вида называются выпуклыми комбинациями точек $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

Доказательство по индукции

- База индукции: при $k = 1$ получаем сами точки множества.
- Предположение: пусть выполнено $\sum_{i=1}^{m-1} \alpha_i \mathbf{x}_i \in \mathcal{X}$, где $\alpha \in \Delta_{m-1}$

Свойство выпуклых множеств

Утверждение

Если множество $\mathcal{X}$ выпукло, то все точки вида $\sum_{i=1}^k \alpha_i \mathbf{x}_i$, где $\mathbf{x}_i \in \mathcal{X}$ и $\alpha \in \Delta_k = \{\alpha \mid \alpha_i \geq 0, \ \sum_{i=1}^k \alpha_i = 1\}$, также лежат в этом множестве. Точки такого вида называются выпуклыми комбинациями точек $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

Доказательство по индукции

- База индукции: при $k = 1$ получаем сами точки множества.
- Предположение: пусть выполнено $\sum_{i=1}^{m-1} \alpha_i \mathbf{x}_i \in \mathcal{X}$, где $\alpha \in \Delta_{m-1}$
- Рассмотрим точку вида $\sum_{i=1}^m \hat{\alpha}_i \mathbf{x}_i$, где $\hat{\alpha} \in \Delta_m$

Свойство выпуклых множеств

Утверждение

Если множество $\mathcal{X}$ выпукло, то все точки вида $\sum_{i=1}^k \alpha_i \mathbf{x}_i$, где $\mathbf{x}_i \in \mathcal{X}$ и $\alpha \in \Delta_k = \{\alpha \mid \alpha_i \geq 0, \ \sum_{i=1}^k \alpha_i = 1\}$, также лежат в этом множестве. Точки такого вида называются выпуклыми комбинациями точек $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

Доказательство по индукции

- База индукции: при $k = 1$ получаем сами точки множества.
- Предположение: пусть выполнено $\sum_{i=1}^{m-1} \alpha_i \mathbf{x}_i \in \mathcal{X}$, где $\alpha \in \Delta_{m-1}$
- Рассмотрим точку вида $\sum_{i=1}^m \hat{\alpha}_i \mathbf{x}_i$, где $\hat{\alpha} \in \Delta_m$
- Тогда найдётся $\hat{\alpha}_k < 1$ и $1 - \hat{\alpha}_k = \sum_{i \neq k} \hat{\alpha}_i$

Свойство выпуклых множеств

Утверждение

Если множество $\mathcal{X}$ выпукло, то все точки вида $\sum_{i=1}^k \alpha_i \mathbf{x}_i$, где $\mathbf{x}_i \in \mathcal{X}$ и $\alpha \in \Delta_k = \{\alpha \mid \alpha_i \geq 0,\ \sum_{i=1}^k \alpha_i = 1\}$, также лежат в этом множестве. Точки такого вида называются выпуклыми комбинациями точек $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

Доказательство по индукции

- База индукции: при $k = 1$ получаем сами точки множества.
- Предположение: пусть выполнено $\sum_{i=1}^{m-1} \alpha_i \mathbf{x}_i \in \mathcal{X}$, где $\alpha \in \Delta_{m-1}$
- Рассмотрим точку вида $\sum_{i=1}^m \hat{\alpha}_i \mathbf{x}_i$, где $\hat{\alpha} \in \Delta_m$
- Тогда найдётся $\hat{\alpha}_k < 1$ и $1 - \hat{\alpha}_k = \sum_{i \neq k} \hat{\alpha}_i$
- $\sum_{i=1}^m \hat{\alpha}_i \mathbf{x}_i = \sum_{i \neq k} \hat{\alpha}_i \mathbf{x}_i + \hat{\alpha}_k \mathbf{x}_k = (1 - \hat{\alpha}_k) \sum_{i \neq k} \frac{\hat{\alpha}_i}{1 - \hat{\alpha}_k} \mathbf{x}_i + \hat{\alpha}_k \mathbf{x}_k = (1 - \hat{\alpha}_k)\mathbf{y} + \hat{\alpha}_k \mathbf{x}_k \in \mathcal{X}$ так как множество выпукло

Пересечение выпуклых множеств

Теорема

Пересечение конечного или бесконечного числа выпуклых множеств $\mathcal{X}_i$ является выпуклым множеством:

$$\mathcal{X} = \bigcap_{i \in \mathcal{I}} \mathcal{X}_i.$$

Пересечение выпуклых множеств

Теорема

Пересечение конечного или бесконечного числа выпуклых множеств $\mathcal{X}_i$ является выпуклым множеством:

$$\mathcal{X} = \bigcap_{i \in \mathcal{I}} \mathcal{X}_i.$$

Доказательство

- Рассмотрим $\mathbf{x}, \mathbf{y} \in \mathcal{X} \to \mathbf{x}, \mathbf{y} \in \mathcal{X}_i, \forall i \in \mathcal{I}$

Пересечение выпуклых множеств

Теорема

Пересечение конечного или бесконечного числа выпуклых множеств $\mathcal{X}_i$ является выпуклым множеством:

$$\mathcal{X} = \bigcap_{i \in \mathcal{I}} \mathcal{X}_i.$$

Доказательство

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

Пересечение выпуклых множеств

Теорема

Пересечение конечного или бесконечного числа выпуклых множеств $\mathcal{X}_i$ *является выпуклым множеством:*

$$\mathcal{X} = \bigcap_{i \in \mathcal{I}} \mathcal{X}_i.$$

Доказательство

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

Пересечение выпуклых множеств

Теорема

Пересечение конечного или бесконечного числа выпуклых множеств $\mathcal{X}_i$ является выпуклым множеством:

$$\mathcal{X} = \bigcap_{i \in \mathcal{I}} \mathcal{X}_i.$$

Доказательство

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

Линейное отображение выпуклых множеств

Теорема

Образ выпуклого множества при линейном отображении является выпуклым множеством.

Доказательство

- Пусть $\mathcal{X}$ — выпуклое множество и $\mathbf{x}, \mathbf{y} \in \mathcal{X}$

Линейное отображение выпуклых множеств

Теорема

Образ выпуклого множества при линейном отображении является выпуклым множеством.

Доказательство

- Пусть $\mathcal{X}$ — выпуклое множество и $\mathbf{x}, \mathbf{y} \in \mathcal{X}$
- Пусть f — линейное отображение вида $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$

Линейное отображение выпуклых множеств

Теорема

Образ выпуклого множества при линейном отображении является выпуклым множеством.

Доказательство

- Пусть $\mathcal{X}$ — выпуклое множество и $\mathbf{x}, \mathbf{y} \in \mathcal{X}$
- Пусть f — линейное отображение вида $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$
- Покажем, что $\alpha f(\mathbf{x}) + (1 - \alpha) f(\mathbf{y}) \in f(\mathcal{X})$, где $\alpha \in [0, 1]$

Линейное отображение выпуклых множеств

Теорема

Образ выпуклого множества при линейном отображении является выпуклым множеством.

Доказательство

- Пусть $\mathcal{X}$ — выпуклое множество и $\mathbf{x}, \mathbf{y} \in \mathcal{X}$
- Пусть f — линейное отображение вида $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$
- Покажем, что $\alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y}) \in f(\mathcal{X})$, где $\alpha \in [0, 1]$
- Действительно,

$$\alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y}) = \alpha(\mathbf{A}\mathbf{x} + \mathbf{b}) + (1-\alpha)(\mathbf{A}\mathbf{y} + \mathbf{b}) =$$
$$\mathbf{A}(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) + \mathbf{b} = \mathbf{A}\mathbf{z} + \mathbf{b} = f(\mathbf{z}),$$

где $\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in \mathcal{X}$

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 $\mathcal{X}_1, \mathcal{X}_2$ — выпуклые множества. Рассмотрим $\mathcal{X} = \mathcal{X}_1 + \mathcal{X}_2 = \{\mathbf{x}_1 + \mathbf{x}_2 \mid \mathbf{x}_1 \in \mathcal{X}_1,\ \mathbf{x}_2 \in \mathcal{X}_2\}$

Арифметические операции над выпуклыми множествами

Теорема

Сумма Минковского выпуклых множеств является выпуклым множеством.

Доказательство

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

Арифметические операции над выпуклыми множествами

Теорема

Сумма Минковского выпуклых множеств является выпуклым множеством.

Доказательство

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

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

  где $\mathbf{y}_1 \in \mathcal{X}_1$ и $\mathbf{y}_2 \in \mathcal{X}_2$ в силу выпуклости множеств $\mathcal{X}_1, \mathcal{X}_2$.

Арифметические операции над выпуклыми множествами

Теорема

Сумма Минковского выпуклых множеств является выпуклым множеством.

Доказательство

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

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

  где $\mathbf{y}_1 \in \mathcal{X}_1$ и $\mathbf{y}_2 \in \mathcal{X}_2$ в силу выпуклости множеств $\mathcal{X}_1, \mathcal{X}_2$.

Следствие

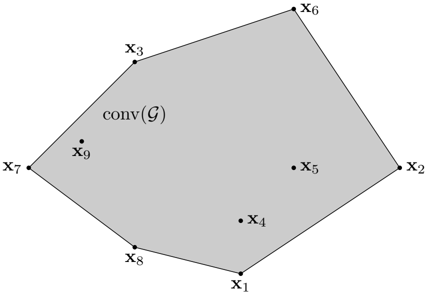
Линейная комбинация выпуклых множеств — выпуклое множество

Выпуклая оболочка (convex hull)

Определение

Выпуклой оболочкой множества $\mathcal{G}$ называется следующее множество

$$\mathrm{conv}\,(\mathcal{G}) = \left\{ \sum_{i=1}^{k} \theta_i \mathbf{x}_i \mid \mathbf{x}_i \in \mathcal{G}, \sum_{i=1}^{k} \theta_i = 1, \theta_i \geq 0 \right\}$$



Альтернативные интерпретации

Эквивалентные формулировки

Выпуклая оболочка множества $\mathcal{G}$ — это

1) минимальное по включению выпуклое множество, содержащее $\mathcal{G}$, то есть если $\mathcal{X}$ выпуклое множество и $\mathcal{G} \subseteq \mathcal{X}$, то $\operatorname{conv}(\mathcal{G}) \subseteq \mathcal{X}$
2) пересечение всех выпуклых множеств, содержащих $\mathcal{G}$

Доказательство

- $\operatorname{conv}(\mathcal{G})$ выпуклое множество (проверьте по определению!)
- $\mathcal{G} \in \operatorname{conv}(\mathcal{G})$
- $\mathcal{X}$ содержит все выпуклые комбинации своих точек
- Если $\mathcal{G} \subseteq \mathcal{X}$, то $\mathcal{X}$ содержит все выпуклые комбинации точек из $\mathcal{G}$
- А значит $\operatorname{conv}(\mathcal{G}) \subseteq \mathcal{X}$

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

Конусы (cones)

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* $\theta_1 \mathbf{x}_1 + \theta_2 \mathbf{x}_2 \in \mathcal{K}$.

Конусы (cones)

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

Конусы (cones)

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
  $\mathbb{R}^n_+ = \{\mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, \; i = 1, \dots, n\} \to$ LP

Конусы (cones)

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

Конусы (cones)

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

Коническая оболочка (conic hull)

Определение

Конической оболочкой множества $\mathcal{G}$ называется множество точек вида $\sum_{i=1}^{k} \theta_i \mathbf{x}_i$, где $\mathbf{x}_i \in \mathcal{G}$ и $\theta_i \geq 0$ и обозначается $\mathrm{cone}\,(\mathcal{G})$.

Теорема

Выпуклый конус содержит все конические оболочки своих элементов

Теорема

Коническая оболочка множества $\mathcal{G}$ — это минимальный выпуклый конус, который содержит $\mathcal{G}$, то есть если $\mathcal{K}$ выпуклый конус и $\mathcal{G} \subseteq \mathcal{K}$, то $\mathrm{cone}\,(\mathcal{G}) \subseteq \mathcal{K}$.

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