

Методы оптимизации

Лекция 3: Примеры задач выпуклой оптимизации

Александр Катруца

Факультет инноваций и высоких технологий
Физтех-школа прикладной математики и информатики



21 сентября 2020 г.

На прошлой лекции

- ▶ Самосопряжённые конусы
- ▶ Отделимость выпуклых множеств
- ▶ Выпуклые функции и способы проверки функции на выпуклость

Основные результаты по выпуклым функциям

Основные результаты по выпуклым функциям

Определение

Функция $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется выпуклой (*строго выпуклой*), если X — *выпуклое множество* и для

$\forall \mathbf{x}_1, \mathbf{x}_2 \in X$ и $\alpha \in [0, 1]$ ($\alpha \in (0, 1)$) выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Основные результаты по выпуклым функциям

Определение

Функция $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется выпуклой (*строго выпуклой*), если X — *выпуклое множество* и для

$\forall \mathbf{x}_1, \mathbf{x}_2 \in X$ и $\alpha \in [0, 1]$ ($\alpha \in (0, 1)$) выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Основные результаты по выпуклым функциям

Определение

Функция $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется выпуклой (*строго выпуклой*), если X — *выпуклое множество* и для

$\forall \mathbf{x}_1, \mathbf{x}_2 \in X$ и $\alpha \in [0, 1]$ ($\alpha \in (0, 1)$) выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Теорема

Пусть функция $f(\mathbf{x})$ дифференцируема и определена на выпуклом множестве $X \subseteq \mathbb{R}^n$. Тогда $f(\mathbf{x})$ сильно выпукла с константой $m \geq 0$ в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

Основные результаты по выпуклым функциям

Определение

Функция $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется выпуклой (*строго выпуклой*), если X — *выпуклое множество* и для

$\forall \mathbf{x}_1, \mathbf{x}_2 \in X$ и $\alpha \in [0, 1]$ ($\alpha \in (0, 1)$) выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Теорема

Пусть функция $f(\mathbf{x})$ дифференцируема и определена на выпуклом множестве $X \subseteq \mathbb{R}^n$. Тогда $f(\mathbf{x})$ сильно выпукла с константой $m \geq 0$ в том и только том случае, если

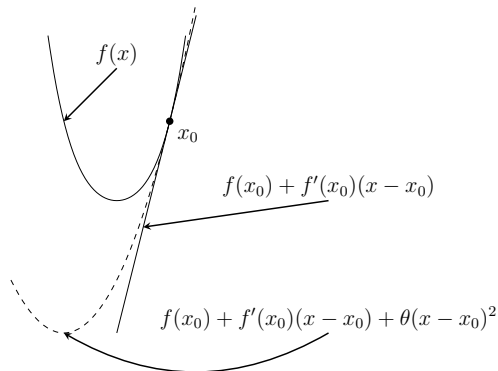
$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

Теорема

Дважды непрерывно дифференцируемая функция f выпукла $\Leftrightarrow f''(\mathbf{x}) \succeq m\mathbf{I}$

Иллюстрация дифференциальных критериев

Пусть $f(x) = x^2 + \frac{1}{2}x^4$



- ▶ Линейная глобальная оценка снизу для выпуклой функции
- ▶ Квадратичная глобальная оценка снизу для сильно выпуклой функции

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

- ▶ \mathcal{C}^n выпуклое множество

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

- ▶ \mathcal{C}^n выпуклое множество
- ▶ $\mathbf{S}_+^n \subset \mathcal{C}^n$

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

- ▶ \mathcal{C}^n выпуклое множество
- ▶ $\mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки $\mathbf{X} \notin \mathcal{C}^n$ является co-NP полной!

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

- ▶ \mathcal{C}^n выпуклое множество
- ▶ $\mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки $\mathbf{X} \notin \mathcal{C}^n$ является co-NP полной!
- ▶ Задача конической оптимизации с конусом \mathcal{C}^n является NP-трудной

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

- ▶ \mathcal{C}^n выпуклое множество
- ▶ $\mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки $\mathbf{X} \notin \mathcal{C}^n$ является co-NP полной!
- ▶ Задача конической оптимизации с конусом \mathcal{C}^n является NP-трудной

Пример

Задача определения максимального независимого множества вершин графа сводится к задаче оптимизации на множестве \mathcal{C}^n . Подробности [тут](#)

Пример простой невыпуклой задачи

Дана матрица $\mathbf{Q} \in \mathbf{S}^n$.

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

Пример простой невыпуклой задачи

Дана матрица $\mathbf{Q} \in \mathbf{S}^n$.

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

- ▶ Допустимое множество невыпукло

Пример простой невыпуклой задачи

Дана матрица $\mathbf{Q} \in \mathbf{S}^n$.

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

- ▶ Допустимое множество невыпукло
- ▶ Целевая функция невыпукла

Пример простой невыпуклой задачи

Дана матрица $\mathbf{Q} \in \mathbf{S}^n$.

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

- ▶ Допустимое множество невыпукло
- ▶ Целевая функция невыпукла

Пример простой невыпуклой задачи

Дана матрица $\mathbf{Q} \in \mathbf{S}^n$.

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

- ▶ Допустимое множество невыпукло
- ▶ Целевая функция невыпукла

Q: какая интерпретация у \mathbf{x}^* и $f(\mathbf{x}^*)$?

Степенной метод (Power method)

- ▶ Степенной метод для поиска собственного вектора для максимального по модулю собственного значения

$$\mathbf{x}_{k+1} = \frac{\mathbf{A}\mathbf{x}_k}{\|\mathbf{A}\mathbf{x}_k\|_2}, \quad \hat{\lambda}_{k+1} = \mathbf{x}_{k+1}^\top \mathbf{A}\mathbf{x}_{k+1}$$

- ▶ \mathbf{A} – диагонализуема, поэтому есть базис из собственных векторов $[\mathbf{v}_1, \dots, \mathbf{v}_n]$ причём предположим $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$
- ▶ $\|\mathbf{x}_k - \mathbf{v}_1\|_2 \leq \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k \|\mathbf{x}_0 - \mathbf{v}_1\|_2$
- ▶ Для поиска максимального и минимального собственных значений (не по модулю) существуют более продвинутые методы (см. [метод Ланцоша](#))
- ▶ Чтобы понять, как ими пользоваться см. [документацию в SciPy](#)

Стандартная форма записи задачи выпуклой оптимизации

$$\begin{aligned} \min f_0(\mathbf{x}) \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, p \end{aligned}$$

- ▶ f_0 – выпуклая целевая функция
- ▶ f_i – выпуклые функции для ограничений типа неравенств
- ▶ Ограничения типа равенств *только* линейные

Стандартная форма записи задачи выпуклой оптимизации

$$\begin{aligned} \min & f_0(\mathbf{x}) \\ \text{s.t. } & \mathbf{Ax} = \mathbf{b} \\ & f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, p \end{aligned}$$

- ▶ f_0 – выпуклая целевая функция
- ▶ f_i – выпуклые функции для ограничений типа неравенств
- ▶ Ограничения типа равенств *только* линейные

Ограничения стандартной формы записи

- ▶ Выпуклое множество может быть задано более общим образом

$$\begin{aligned} \min & x^2 \\ \text{s.t. } & (x - 2)^2 = 0 \\ & x^3 \geq 0 \end{aligned}$$

Линейное программирование (LP)

$$\begin{aligned} \min \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ x_i \geq 0, \ i = 1, \dots, n \end{aligned}$$

Линейное программирование (LP)

$$\begin{aligned} \min \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ x_i \geq 0, \ i = 1, \dots, n \end{aligned}$$

- ▶ Существуют различные формы задачи LP, но все они сводятся к указанной

Линейное программирование (LP)

$$\begin{aligned} \min \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ x_i \geq 0, \ i = 1, \dots, n \end{aligned}$$

- ▶ Существуют различные формы задачи LP, но все они сводятся к указанной
- ▶ Простейший пример задачи конической оптимизации

Линейное программирование (LP)

$$\begin{aligned} \min \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ x_i \geq 0, \ i = 1, \dots, n \end{aligned}$$

- ▶ Существуют различные формы задачи LP, но все они сводятся к указанной
- ▶ Простейший пример задачи конической оптимизации
- ▶ Замена \mathbb{R}_+^n на другие конусы даёт более богатое семейство задач, примеры далее

Линейное программирование (LP)

$$\begin{aligned} \min \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ x_i \geq 0, \ i = 1, \dots, n \end{aligned}$$

- ▶ Существуют различные формы задачи LP, но все они сводятся к указанной
- ▶ Простейший пример задачи конической оптимизации
- ▶ Замена \mathbb{R}_+^n на другие конусы даёт более богатое семейство задач, примеры далее

Пример: составление диеты минимальной стоимости

Линейное программирование (LP)

$$\begin{aligned} \min \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ x_i \geq 0, \quad i = 1, \dots, n \end{aligned}$$

- ▶ Существуют различные формы задачи LP, но все они сводятся к указанной
- ▶ Простейший пример задачи конической оптимизации
- ▶ Замена \mathbb{R}_+^n на другие конусы даёт более богатое семейство задач, примеры далее

Пример: составление диеты минимальной стоимости

- ▶ Дано n продуктов, цена единицы каждого c_i

Линейное программирование (LP)

$$\begin{aligned} \min \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ x_i \geq 0, \ i = 1, \dots, n \end{aligned}$$

- ▶ Существуют различные формы задачи LP, но все они сводятся к указанной
- ▶ Простейший пример задачи конической оптимизации
- ▶ Замена \mathbb{R}_+^n на другие конусы даёт более богатое семейство задач, примеры далее

Пример: составление диеты минимальной стоимости

- ▶ Дано n продуктов, цена единицы каждого c_i
- ▶ Необходимо, чтобы человек получил m питательных веществ в количествах не менее b_1, \dots, b_m

Линейное программирование (LP)

$$\begin{aligned} \min \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ x_i \geq 0, \ i = 1, \dots, n \end{aligned}$$

- ▶ Существуют различные формы задачи LP, но все они сводятся к указанной
- ▶ Простейший пример задачи конической оптимизации
- ▶ Замена \mathbb{R}_+^n на другие конусы даёт более богатое семейство задач, примеры далее

Пример: составление диеты минимальной стоимости

- ▶ Дано n продуктов, цена единицы каждого c_i
- ▶ Необходимо, чтобы человек получил m питательных веществ в количествах не менее b_1, \dots, b_m
- ▶ Известно, что в j -ом продукте содержится a_{ij} i -го питательного вещества

Линейное программирование (LP)

$$\begin{aligned} \min \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ x_i \geq 0, \ i = 1, \dots, n \end{aligned}$$

- ▶ Существуют различные формы задачи LP, но все они сводятся к указанной
- ▶ Простейший пример задачи конической оптимизации
- ▶ Замена \mathbb{R}_+^n на другие конусы даёт более богатое семейство задач, примеры далее

Пример: составление диеты минимальной стоимости

- ▶ Дано n продуктов, цена единицы каждого c_i
- ▶ Необходимо, чтобы человек получил m питательных веществ в количествах не менее b_1, \dots, b_m
- ▶ Известно, что в j -ом продукте содержится a_{ij} i -го питательного вещества
- ▶ Необходимо определить количество каждого продукта

Квадратичное программирование с квадратичными ограничениями (QCQP)

$$\begin{aligned} & \min \frac{1}{2} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}^\top \mathbf{x} + r_0 \\ & \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ & \frac{1}{2} \mathbf{x}^\top \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^\top \mathbf{x} + r_i \leq 0, \quad i = 1, \dots, n \end{aligned}$$

Квадратичное программирование с квадратичными ограничениями (QCQP)

$$\begin{aligned} \min \quad & \frac{1}{2} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}^\top \mathbf{x} + r_0 \\ \text{s.t.} \quad & \mathbf{A} \mathbf{x} = \mathbf{b} \\ & \frac{1}{2} \mathbf{x}^\top \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^\top \mathbf{x} + r_i \leq 0, \quad i = 1, \dots, n \end{aligned}$$

- Задача будет выпукла, если все $\mathbf{P}_i \in \mathbf{S}_+^n$

Квадратичное программирование с квадратичными ограничениями (QCQP)

$$\begin{aligned} \min \quad & \frac{1}{2} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}^\top \mathbf{x} + r_0 \\ \text{s.t.} \quad & \mathbf{A} \mathbf{x} = \mathbf{b} \\ & \frac{1}{2} \mathbf{x}^\top \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^\top \mathbf{x} + r_i \leq 0, \quad i = 1, \dots, n \end{aligned}$$

- ▶ Задача будет выпукла, если все $\mathbf{P}_i \in \mathbf{S}_+^n$
- ▶ Может быть сведена к оптимизации на конусе второго порядка $\mathcal{Q}^n = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq t\}$

Квадратичное программирование с квадратичными ограничениями (QCQP)

$$\begin{aligned} \min \quad & \frac{1}{2} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}^\top \mathbf{x} + r_0 \\ \text{s.t.} \quad & \mathbf{A} \mathbf{x} = \mathbf{b} \\ & \frac{1}{2} \mathbf{x}^\top \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^\top \mathbf{x} + r_i \leq 0, \quad i = 1, \dots, n \end{aligned}$$

- ▶ Задача будет выпукла, если все $\mathbf{P}_i \in \mathbf{S}_+^n$
- ▶ Может быть сведена к оптимизации на конусе второго порядка $\mathcal{Q}^n = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq t\}$
- ▶ При $\mathbf{P}_i = 0$ получим задачу LP

Линейные системы с неквадратными матрицами

- ▶ Линейная задача наименьших квадратов

$$\min \frac{1}{2} \|\mathbf{Ax} - \mathbf{b}\|_2^2$$

Возможно добавление линейных ограничений вида

$$l_i \leq x_i \leq u_i$$

Линейные системы с неквадратными матрицами

- ▶ Линейная задача наименьших квадратов

$$\min \frac{1}{2} \|\mathbf{Ax} - \mathbf{b}\|_2^2$$

Возможно добавление линейных ограничений вида

$$l_i \leq x_i \leq u_i$$

- ▶ Поиск решения минимальной нормы

$$\begin{aligned} \min \quad & \frac{1}{2} \|\mathbf{x}\|_2^2 \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \end{aligned}$$

Линейные системы с неквадратными матрицами

- ▶ Линейная задача наименьших квадратов

$$\min \frac{1}{2} \|\mathbf{Ax} - \mathbf{b}\|_2^2$$

Возможно добавление линейных ограничений вида

$$l_i \leq x_i \leq u_i$$

- ▶ Поиск решения минимальной нормы

$$\min \frac{1}{2} \|\mathbf{x}\|_2^2$$

$$\text{s.t. } \mathbf{Ax} = \mathbf{b}$$

Q: Каков геометрический смысл у решения?

Линейные системы с неквадратными матрицами

- ▶ Линейная задача наименьших квадратов

$$\min \frac{1}{2} \|\mathbf{Ax} - \mathbf{b}\|_2^2$$

Возможно добавление линейных ограничений вида

$$l_i \leq x_i \leq u_i$$

- ▶ Поиск решения минимальной нормы

$$\min \frac{1}{2} \|\mathbf{x}\|_2^2$$

$$\text{s.t. } \mathbf{Ax} = \mathbf{b}$$

Q: Каков геометрический смысл у решения?

Q: К какой задаче сводится похожая задача?

$$\min \|\mathbf{x}\|_\infty$$

$$\text{s.t. } \mathbf{Ax} = \mathbf{b}$$

Задача о составлении оптимального портфеля

Дано

Задача о составлении оптимального портфеля

Дано

- ▶ n АКТИВОВ

Задача о составлении оптимального портфеля

Дано

- ▶ n активов
- ▶ Изменение относительной цены активов — случайный вектор со средним $\bar{\mathbf{p}}$ и ковариационной матрицей Σ

Задача о составлении оптимального портфеля

Дано

- ▶ n активов
- ▶ Изменение относительной цены активов — случайный вектор со средним $\bar{\mathbf{p}}$ и ковариационной матрицей Σ
- ▶ Минимально допустимый средний доход \bar{r}

Задача о составлении оптимального портфеля

Дано

- ▶ n активов
- ▶ Изменение относительной цены активов — случайный вектор со средним $\bar{\mathbf{p}}$ и ковариационной матрицей Σ
- ▶ Минимально допустимый средний доход \bar{r}

Классическая задача составления оптимального портфеля

$$\begin{aligned} & \min \mathbf{x}^\top \Sigma \mathbf{x} \\ \text{s.t. } & \bar{\mathbf{p}}^\top \mathbf{x} \geq \bar{r} \\ & \sum_{i=1}^n x_i = 1, \quad x_i \geq 0, \quad i = 1, \dots, n \end{aligned}$$

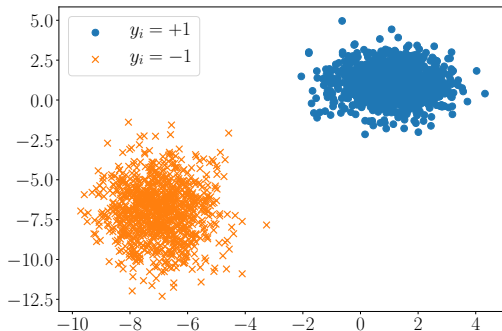
- ▶ Минимум риска при минимально допустимом доходе
- ▶ Существуют многочисленные вариации, которые не выводят задачу из класса QCQP или SOCP

Задача классификации

- ▶ Дана выборка (\mathbf{x}_i, y_i) , где $\mathbf{x}_i \in \mathbb{R}^n$, а $y_i \in \{-1, +1\}$
- ▶ Необходимо построить гиперплоскость так, чтобы $\mathbf{a}^\top \mathbf{x}_i + b > 0$, если $y_i = +1$ и $\mathbf{a}^\top \mathbf{x}_i + b < 0$, если $y_i < 0$

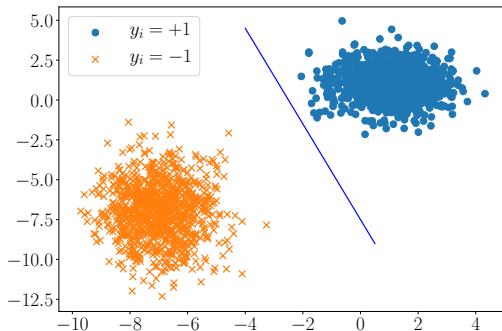
Задача классификации

- ▶ Дана выборка (\mathbf{x}_i, y_i) , где $\mathbf{x}_i \in \mathbb{R}^n$, а $y_i \in \{-1, +1\}$
- ▶ Необходимо построить гиперплоскость так, чтобы $\mathbf{a}^\top \mathbf{x}_i + b > 0$, если $y_i = +1$ и $\mathbf{a}^\top \mathbf{x}_i + b < 0$, если $y_i < 0$



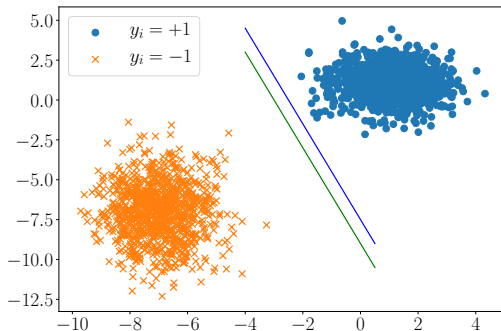
Задача классификации

- ▶ Дана выборка (\mathbf{x}_i, y_i) , где $\mathbf{x}_i \in \mathbb{R}^n$, а $y_i \in \{-1, +1\}$
- ▶ Необходимо построить гиперплоскость так, чтобы $\mathbf{a}^\top \mathbf{x}_i + b > 0$, если $y_i = +1$ и $\mathbf{a}^\top \mathbf{x}_i + b < 0$, если $y_i < 0$



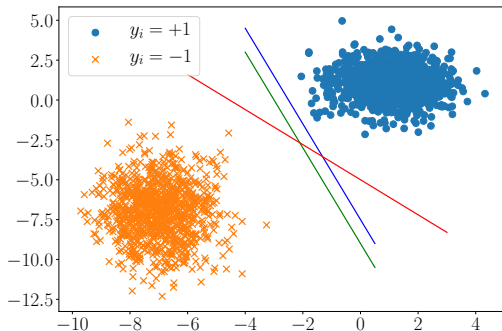
Задача классификации

- ▶ Дана выборка (\mathbf{x}_i, y_i) , где $\mathbf{x}_i \in \mathbb{R}^n$, а $y_i \in \{-1, +1\}$
- ▶ Необходимо построить гиперплоскость так, чтобы $\mathbf{a}^\top \mathbf{x}_i + b > 0$, если $y_i = +1$ и $\mathbf{a}^\top \mathbf{x}_i + b < 0$, если $y_i < 0$



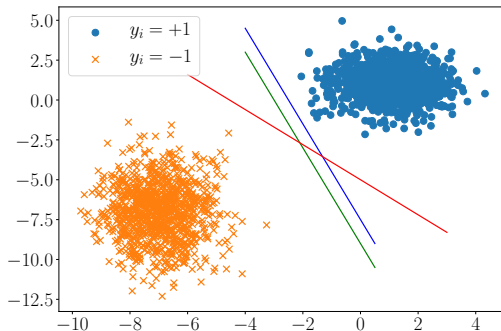
Задача классификации

- ▶ Дана выборка (\mathbf{x}_i, y_i) , где $\mathbf{x}_i \in \mathbb{R}^n$, а $y_i \in \{-1, +1\}$
- ▶ Необходимо построить гиперплоскость так, чтобы $\mathbf{a}^\top \mathbf{x}_i + b > 0$, если $y_i = +1$ и $\mathbf{a}^\top \mathbf{x}_i + b < 0$, если $y_i < 0$



Задача классификации

- ▶ Дана выборка (\mathbf{x}_i, y_i) , где $\mathbf{x}_i \in \mathbb{R}^n$, а $y_i \in \{-1, +1\}$
- ▶ Необходимо построить гиперплоскость так, чтобы $\mathbf{a}^\top \mathbf{x}_i + b > 0$, если $y_i = +1$ и $\mathbf{a}^\top \mathbf{x}_i + b < 0$, если $y_i < 0$



Q: Как однозначно задать разделяющую гиперплоскость?

Максимизация зазора: метод опорных векторов (SVM)

Максимизация зазора: метод опорных векторов (SVM)

- ▶ Для опорных объектов каждого класса выполнено

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

Максимизация зазора: метод опорных векторов (SVM)

- ▶ Для опорных объектов каждого класса выполнено

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

- ▶ Расстояние между гиперплоскостями

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

Максимизация зазора: метод опорных векторов (SVM)

- ▶ Для опорных объектов каждого класса выполнено

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

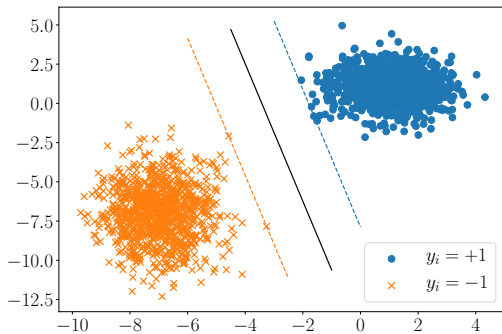
- ▶ Расстояние между гиперплоскостями

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

Финальная задача

$$\begin{aligned} \min_{\mathbf{a}, b} \quad & \frac{1}{2} \|\mathbf{a}\|_2^2 \\ \text{s.t.} \quad & y_i (\mathbf{a}^\top \mathbf{x}_i + b) > 1, \quad i = 1, \dots, m \end{aligned}$$

Оптимальная гиперплоскость



Оптимизация на конусе второго порядка (SOCP)

Коническая форма записи

$$\begin{aligned} \min \mathbf{f}^\top \mathbf{x} \\ \text{s.t. } \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}_i^\top \mathbf{x} + d_i \\ \mathbf{F} \mathbf{x} = \mathbf{g} \end{aligned}$$

$$\begin{aligned} \min \mathbf{f}^\top \mathbf{x} \\ \text{s.t. } (\mathbf{A}_i \mathbf{x} + \mathbf{b}_i, \mathbf{c}_i^\top \mathbf{x} + d_i) \succeq_{K_i} 0 \\ \mathbf{F} \mathbf{x} = \mathbf{g} \end{aligned}$$

QCQP \rightarrow SOCP

- ▶ Пусть $\mathbf{x}^\top \mathbf{P} \mathbf{x} + \mathbf{q}^\top \mathbf{x} + r \leq 0$ и $0 \prec \mathbf{P} = \mathbf{L} \mathbf{L}^\top$
- ▶ $(\mathbf{L}^\top \mathbf{x})^\top (\mathbf{L}^\top \mathbf{x}) + 2\tilde{\mathbf{q}}^\top \mathbf{L}^{-\top} \mathbf{L}^\top \mathbf{x} + \|\mathbf{L}^{-1} \tilde{\mathbf{q}}\|_2^2 \leq \|\mathbf{L}^{-1} \tilde{\mathbf{q}}\|_2^2 - r$
- ▶ Или $\|\mathbf{L}^\top \mathbf{x} + \mathbf{L}^{-1} \tilde{\mathbf{q}}\|_2 \leq \sqrt{\tilde{\mathbf{q}}^\top \mathbf{P}^{-1} \tilde{\mathbf{q}} - r}$

QCQP \rightarrow SOCP: преобразование целевой функции

$$\begin{array}{ll} \min_{\mathbf{x}} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^\top \mathbf{x} + r_0 & \Rightarrow \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} & \end{array} \quad \begin{array}{l} \min_{\mathbf{x}, t} t + \mathbf{q}_0^\top \mathbf{x} + r_0 \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t \end{array}$$

QCQP \rightarrow SOCP: преобразование целевой функции

$$\begin{array}{ll} \min_{\mathbf{x}} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^\top \mathbf{x} + r_0 & \implies \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} & \min_{\mathbf{x}, t} t + \mathbf{q}_0^\top \mathbf{x} + r_0 \\ & \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ & \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t \end{array}$$

► Так как $\mathbf{P}_0 \succ 0$, то $\mathbf{P} = \mathbf{LL}^\top$

QCQP \rightarrow SOCP: преобразование целевой функции

$$\begin{array}{ll} \min_{\mathbf{x}} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^\top \mathbf{x} + r_0 & \implies \min_{\mathbf{x}, t} t + \mathbf{q}_0^\top \mathbf{x} + r_0 \\ \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} & \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t \end{array}$$

- ▶ Так как $\mathbf{P}_0 \succ 0$, то $\mathbf{P} = \mathbf{L}\mathbf{L}^\top$
- ▶ Тогда $\mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t$ сводится к $\|\mathbf{L}^\top \mathbf{x}\|_2^2 \leq t$

QCQP \rightarrow SOCP: преобразование целевой функции

$$\begin{array}{ll} \min_{\mathbf{x}} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^\top \mathbf{x} + r_0 & \implies \min_{\mathbf{x}, t} t + \mathbf{q}_0^\top \mathbf{x} + r_0 \\ \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} & \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t \end{array}$$

- ▶ Так как $\mathbf{P}_0 \succ 0$, то $\mathbf{P} = \mathbf{L}\mathbf{L}^\top$
- ▶ Тогда $\mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t$ сводится к $\|\mathbf{L}^\top \mathbf{x}\|_2^2 \leq t$
- ▶ Повёрнутый конус второго порядка:
 $\mathcal{Q}_{\text{rot}}^n = \{(\mathbf{x}, y, z) \in \mathbb{R}^{n+2} \mid \|\mathbf{x}\|_2 \leq \sqrt{yz}, y \geq 0, z \geq 0\}$

QCQP \rightarrow SOCP: преобразование целевой функции

$$\begin{array}{ll} \min_{\mathbf{x}} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^\top \mathbf{x} + r_0 & \Rightarrow \min_{\mathbf{x}, t} t + \mathbf{q}_0^\top \mathbf{x} + r_0 \\ \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} & \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t \end{array}$$

- ▶ Так как $\mathbf{P}_0 \succ 0$, то $\mathbf{P} = \mathbf{L}\mathbf{L}^\top$
- ▶ Тогда $\mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t$ сводится к $\|\mathbf{L}^\top \mathbf{x}\|_2^2 \leq t$
- ▶ Повёрнутый конус второго порядка:
 $\mathcal{Q}_{\text{rot}}^n = \{(\mathbf{x}, y, z) \in \mathbb{R}^{n+2} \mid \|\mathbf{x}\|_2 \leq \sqrt{yz}, y \geq 0, z \geq 0\}$
- ▶ Условие $\|\mathbf{x}\|_2 \leq \sqrt{yz}, y, z \geq 0 \Leftrightarrow \|(2\mathbf{x}, y - z)\|_2 \leq y + z$

QCQP \rightarrow SOCP: преобразование целевой функции

$$\begin{array}{ll} \min_{\mathbf{x}} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^\top \mathbf{x} + r_0 & \Rightarrow \min_{\mathbf{x}, t} t + \mathbf{q}_0^\top \mathbf{x} + r_0 \\ \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} & \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t \end{array}$$

- ▶ Так как $\mathbf{P}_0 \succ 0$, то $\mathbf{P} = \mathbf{L}\mathbf{L}^\top$
- ▶ Тогда $\mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t$ сводится к $\|\mathbf{L}^\top \mathbf{x}\|_2^2 \leq t$
- ▶ Повёрнутый конус второго порядка:
 $\mathcal{Q}_{\text{rot}}^n = \{(\mathbf{x}, y, z) \in \mathbb{R}^{n+2} \mid \|\mathbf{x}\|_2 \leq \sqrt{yz}, y \geq 0, z \geq 0\}$
- ▶ Условие $\|\mathbf{x}\|_2 \leq \sqrt{yz}, y, z \geq 0 \Leftrightarrow \|(2\mathbf{x}, y - z)\|_2 \leq y + z$
- ▶ Тогда $(\mathbf{L}^\top \mathbf{x}, t, 1) \in \mathcal{Q}_{\text{rot}}^n$

QCQP \rightarrow SOCP: преобразование целевой функции

$$\begin{array}{ll} \min_{\mathbf{x}} \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^\top \mathbf{x} + r_0 & \Rightarrow \min_{\mathbf{x}, t} t + \mathbf{q}_0^\top \mathbf{x} + r_0 \\ \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} & \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t \end{array}$$

- ▶ Так как $\mathbf{P}_0 \succ 0$, то $\mathbf{P} = \mathbf{L}\mathbf{L}^\top$
- ▶ Тогда $\mathbf{x}^\top \mathbf{P}_0 \mathbf{x} \leq t$ сводится к $\|\mathbf{L}^\top \mathbf{x}\|_2^2 \leq t$
- ▶ Повёрнутый конус второго порядка:
 $\mathcal{Q}_{\text{rot}}^n = \{(\mathbf{x}, y, z) \in \mathbb{R}^{n+2} \mid \|\mathbf{x}\|_2 \leq \sqrt{yz}, y \geq 0, z \geq 0\}$
- ▶ Условие $\|\mathbf{x}\|_2 \leq \sqrt{yz}, y, z \geq 0 \Leftrightarrow \|(2\mathbf{x}, y - z)\|_2 \leq y + z$
- ▶ Тогда $(\mathbf{L}^\top \mathbf{x}, t, 1) \in \mathcal{Q}_{\text{rot}}^n$

Отношения между рассмотренными типами задач

$$\text{LP} \subset \text{QCQP} \subset \text{SOCP}$$

Оптимизация на конусе \mathbf{S}_+^n (SDP)

Коническая форма

$$\begin{aligned} \min_{\mathbf{x}} \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ \mathbf{G} + \sum_{i=1}^n x_i \mathbf{F}_i \preceq 0, \end{aligned}$$

где $\mathbf{G} \in \mathbf{S}^n$ и все $\mathbf{F}_i \in \mathbf{S}^n$.

Стандартная форма

$$\begin{aligned} \min_{\mathbf{X}} \text{trace}(\mathbf{CX}) \\ \text{s.t. } \text{trace}(\mathbf{A}_i \mathbf{X}) = b_i \\ \mathbf{X} \succeq 0, \end{aligned}$$

где $\mathbf{C} \in \mathbf{S}^n$ и все $\mathbf{A}_i \in \mathbf{S}^n$

Оптимизация на конусе \mathbf{S}_+^n (SDP)

Коническая форма

$$\begin{aligned} \min_{\mathbf{x}} \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ \mathbf{G} + \sum_{i=1}^n x_i \mathbf{F}_i \preceq 0, \end{aligned}$$

где $\mathbf{G} \in \mathbf{S}^n$ и все $\mathbf{F}_i \in \mathbf{S}^n$.

Стандартная форма

$$\begin{aligned} \min_{\mathbf{X}} \text{trace}(\mathbf{CX}) \\ \text{s.t. } \text{trace}(\mathbf{A}_i \mathbf{X}) = b_i \\ \mathbf{X} \succeq 0, \end{aligned}$$

где $\mathbf{C} \in \mathbf{S}^n$ и все $\mathbf{A}_i \in \mathbf{S}^n$

- ▶ Полная аналогия с LP с точностью до определения скалярного произведения и конуса

Оптимизация на конусе \mathbf{S}_+^n (SDP)

Коническая форма

$$\begin{aligned} \min_{\mathbf{x}} \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\ \mathbf{G} + \sum_{i=1}^n x_i \mathbf{F}_i \preceq 0, \end{aligned}$$

где $\mathbf{G} \in \mathbf{S}^n$ и все $\mathbf{F}_i \in \mathbf{S}^n$.

Стандартная форма

$$\begin{aligned} \min_{\mathbf{X}} \text{trace}(\mathbf{C}\mathbf{X}) \\ \text{s.t. } \text{trace}(\mathbf{A}_i \mathbf{X}) = b_i \\ \mathbf{X} \succeq 0, \end{aligned}$$

где $\mathbf{C} \in \mathbf{S}^n$ и все $\mathbf{A}_i \in \mathbf{S}^n$

- ▶ Полная аналогия с LP с точностью до определения скалярного произведения и конуса
- ▶ Геометрию таких задач рассмотрим ближе к концу курса, когда будем говорить о методах

Оптимизация на конусе \mathbf{S}_+^n (SDP)

Коническая форма

$$\begin{aligned} \min_{\mathbf{x}} \mathbf{c}^\top \mathbf{x} \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ \mathbf{G} + \sum_{i=1}^n x_i \mathbf{F}_i \preceq 0, \end{aligned}$$

где $\mathbf{G} \in \mathbf{S}^n$ и все $\mathbf{F}_i \in \mathbf{S}^n$.

Стандартная форма

$$\begin{aligned} \min_{\mathbf{X}} \text{trace}(\mathbf{CX}) \\ \text{s.t. } \text{trace}(\mathbf{A}_i \mathbf{X}) = b_i \\ \mathbf{X} \succeq 0, \end{aligned}$$

где $\mathbf{C} \in \mathbf{S}^n$ и все $\mathbf{A}_i \in \mathbf{S}^n$

- ▶ Полная аналогия с LP с точностью до определения скалярного произведения и конуса
- ▶ Геометрию таких задач рассмотрим ближе к концу курса, когда будем говорить о методах
- ▶ Из одной формы можно получить другую

LP и SOCP как задачи SDP

LP и SOCP как задачи SDP

LP

- ▶ $\mathbf{G} = 0$
- ▶ \mathbf{F}_i такие, что $\sum_{i=1}^n x_i \mathbf{F}_i = -\text{diag}(\mathbf{x})$

LP и SOCP как задачи SDP

LP

- ▶ $\mathbf{G} = 0$
- ▶ \mathbf{F}_i такие, что $\sum_{i=1}^n x_i \mathbf{F}_i = -\text{diag}(\mathbf{x})$

Дополнение по Шуру

Если $\mathbf{C} \succ 0$, то

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix} \succeq 0 \Leftrightarrow \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\top \succeq 0$$

LP и SOCP как задачи SDP

LP

- ▶ $\mathbf{G} = 0$
- ▶ \mathbf{F}_i такие, что $\sum_{i=1}^n x_i \mathbf{F}_i = -\text{diag}(\mathbf{x})$

Дополнение по Шуру

Если $\mathbf{C} \succ 0$, то

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix} \succeq 0 \Leftrightarrow \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\top \succeq 0$$

SOCP

- ▶ $\|\mathbf{x}\|_2 \leq t \Leftrightarrow \begin{bmatrix} t & \mathbf{x}^\top \\ \mathbf{x} & t\mathbf{I} \end{bmatrix} \succeq 0$
- ▶ Аналогично для \mathcal{Q}^n :
$$\|\mathbf{A}^\top \mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}^\top \mathbf{x} + d_i \Leftrightarrow \begin{bmatrix} \mathbf{c}^\top \mathbf{x} + d_i & (\mathbf{A}^\top \mathbf{x} + \mathbf{b}_i)^\top \\ \mathbf{A}^\top \mathbf{x} + \mathbf{b}_i & (\mathbf{c}^\top \mathbf{x} + d_i)\mathbf{I} \end{bmatrix} \succeq 0$$

Задача MAXCUT и её выпуклая релаксация

- ▶ Граф $G = (V, E)$ и матрица весов рёбер $\mathbf{W} \in \mathbf{S}^n$ и $\mathbf{W} \geq 0$

Задача MAXCUT и её выпуклая релаксация

- ▶ Граф $G = (V, E)$ и матрица весов рёбер $\mathbf{W} \in \mathbf{S}^n$ и $\mathbf{W} \geq 0$
- ▶ Задача MAXCUT

$$\max \frac{1}{2} \sum_{i < j} w_{ij} (1 - x_i x_j)$$

$$\text{s.t. } x_i \in \{-1, +1\}$$

Задача MAXCUT и её выпуклая релаксация

- ▶ Граф $G = (V, E)$ и матрица весов рёбер $\mathbf{W} \in \mathbf{S}^n$ и $\mathbf{W} \geq 0$
- ▶ Задача MAXCUT

$$\max \frac{1}{2} \sum_{i < j} w_{ij} (1 - x_i x_j)$$

$$\text{s.t. } x_i \in \{-1, +1\}$$

- ▶ В матрично-векторном виде

$$\min \frac{1}{2} \mathbf{x}^\top \mathbf{W} \mathbf{x}$$

$$\text{s.t. } x_i \in \{-1, +1\}$$

Задача MAXCUT и её выпуклая релаксация

- ▶ Граф $G = (V, E)$ и матрица весов рёбер $\mathbf{W} \in \mathbf{S}^n$ и $\mathbf{W} \geq 0$
- ▶ Задача MAXCUT

$$\max \frac{1}{2} \sum_{i < j} w_{ij} (1 - x_i x_j)$$

$$\text{s.t. } x_i \in \{-1, +1\}$$

- ▶ В матрично-векторном виде

$$\min \frac{1}{2} \mathbf{x}^\top \mathbf{W} \mathbf{x}$$

$$\text{s.t. } x_i \in \{-1, +1\}$$

- ▶ Эквивалентный вид

$$\min \frac{1}{2} \text{trace}(\mathbf{W} \mathbf{X})$$

$$\text{s.t. } \mathbf{X} \succeq 0, \text{rank}(\mathbf{X}) = 1, \text{diag}(\mathbf{X}) = \mathbf{1}$$

Задача MAXCUT и её выпуклая релаксация

- ▶ Граф $G = (V, E)$ и матрица весов рёбер $\mathbf{W} \in \mathbf{S}^n$ и $\mathbf{W} \succeq 0$
- ▶ Задача MAXCUT

$$\max \frac{1}{2} \sum_{i < j} w_{ij} (1 - x_i x_j)$$

$$\text{s.t. } x_i \in \{-1, +1\}$$

- ▶ В матрично-векторном виде

$$\min \frac{1}{2} \mathbf{x}^\top \mathbf{W} \mathbf{x}$$

$$\text{s.t. } x_i \in \{-1, +1\}$$

- ▶ Эквивалентный вид

$$\min \frac{1}{2} \text{trace}(\mathbf{W} \mathbf{X})$$

$$\text{s.t. } \mathbf{X} \succeq 0, \text{rank}(\mathbf{X}) = 1, \text{diag}(\mathbf{X}) = \mathbf{1}$$

- ▶ SDP релаксация

$$\min \text{trace}(\mathbf{W} \mathbf{X})$$

$$\text{s.t. } \mathbf{X} \succeq 0, \text{rank}(\mathbf{X}) \leq 1, \text{diag}(\mathbf{X}) = \mathbf{1}$$

Экстремальный эллипсоид для данного множества точек

- ▶ Дано k точек $\mathbf{x}_i \in \mathbb{R}^n$

Экстремальный эллипсоид для данного множества точек

- ▶ Дано k точек $\mathbf{x}_i \in \mathbb{R}^n$
- ▶ Необходимо найти эллипсоид минимальной площади, который покрывает все \mathbf{x}_i

Экстремальный эллипсоид для данного множества точек

- ▶ Дано k точек $\mathbf{x}_i \in \mathbb{R}^n$
- ▶ Необходимо найти эллипсоид минимальной площади, который покрывает все \mathbf{x}_i
- ▶ Эллипсоид можно задать аффинным преобразованием

$$\{\mathbf{x} \mid \|\mathbf{x}\|_2^2 \leq 1\} \rightarrow \{\mathbf{u} \mid \|\mathbf{u}\|_2^2 \leq 1, \mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$$

Экстремальный эллипсоид для данного множества точек

- ▶ Дано k точек $\mathbf{x}_i \in \mathbb{R}^n$
- ▶ Необходимо найти эллипсоид минимальной площади, который покрывает все \mathbf{x}_i
- ▶ Эллипсоид можно задать аффинным преобразованием

$$\{\mathbf{x} \mid \|\mathbf{x}\|_2^2 \leq 1\} \rightarrow \{\mathbf{u} \mid \|\mathbf{u}\|_2^2 \leq 1, \mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$$

- ▶ Тогда площадь увеличивается в $\det(\mathbf{A}^{-1})$ раз.

Экстремальный эллипсоид для данного множества точек

- ▶ Дано k точек $\mathbf{x}_i \in \mathbb{R}^n$
- ▶ Необходимо найти эллипсоид минимальной площади, который покрывает все \mathbf{x}_i
- ▶ Эллипсоид можно задать аффинным преобразованием

$$\{\mathbf{x} \mid \|\mathbf{x}\|_2^2 \leq 1\} \rightarrow \{\mathbf{u} \mid \|\mathbf{u}\|_2^2 \leq 1, \mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$$

- ▶ Тогда площадь увеличивается в $\det(\mathbf{A}^{-1})$ раз.
- ▶ Детерминант не является выпуклой/вогнутой функцией

Экстремальный эллипсоид для данного множества точек

- ▶ Дано k точек $\mathbf{x}_i \in \mathbb{R}^n$
- ▶ Необходимо найти эллипсоид минимальной площади, который покрывает все \mathbf{x}_i
- ▶ Эллипсоид можно задать аффинным преобразованием

$$\{\mathbf{x} \mid \|\mathbf{x}\|_2^2 \leq 1\} \rightarrow \{\mathbf{u} \mid \|\mathbf{u}\|_2^2 \leq 1, \mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$$

- ▶ Тогда площадь увеличивается в $\det(\mathbf{A}^{-1})$ раз.
- ▶ Детерминант не является выпуклой/вогнутой функцией
- ▶ $\log \det(\mathbf{A}^{-1}) = -\log \det(\mathbf{A})$ – выпуклая функция при $\mathbf{A} \succ 0$

Экстремальный эллипсоид для данного множества точек

- ▶ Дано k точек $\mathbf{x}_i \in \mathbb{R}^n$
- ▶ Необходимо найти эллипсоид минимальной площади, который покрывает все \mathbf{x}_i
- ▶ Эллипсоид можно задать аффинным преобразованием

$$\{\mathbf{x} \mid \|\mathbf{x}\|_2^2 \leq 1\} \rightarrow \{\mathbf{u} \mid \|\mathbf{u}\|_2^2 \leq 1, \mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$$

- ▶ Тогда площадь увеличивается в $\det(\mathbf{A}^{-1})$ раз.
- ▶ Детерминант не является выпуклой/вогнутой функцией
- ▶ $\log \det(\mathbf{A}^{-1}) = -\log \det(\mathbf{A})$ – выпуклая функция при $\mathbf{A} \succ 0$

$$\begin{aligned} & \min_{\mathbf{A}, \mathbf{b}} \log \det \mathbf{A}^{-1} \\ & \text{s.t. } \mathbf{A} \succ 0 \\ & \quad \|\mathbf{A}\mathbf{x}_i + \mathbf{b}\|_2 \leq 1 \end{aligned}$$

Экстремальный эллипсоид для данного множества точек

- ▶ Дано k точек $\mathbf{x}_i \in \mathbb{R}^n$
- ▶ Необходимо найти эллипсоид минимальной площади, который покрывает все \mathbf{x}_i
- ▶ Эллипсоид можно задать аффинным преобразованием

$$\{\mathbf{x} \mid \|\mathbf{x}\|_2^2 \leq 1\} \rightarrow \{\mathbf{u} \mid \|\mathbf{u}\|_2^2 \leq 1, \mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$$

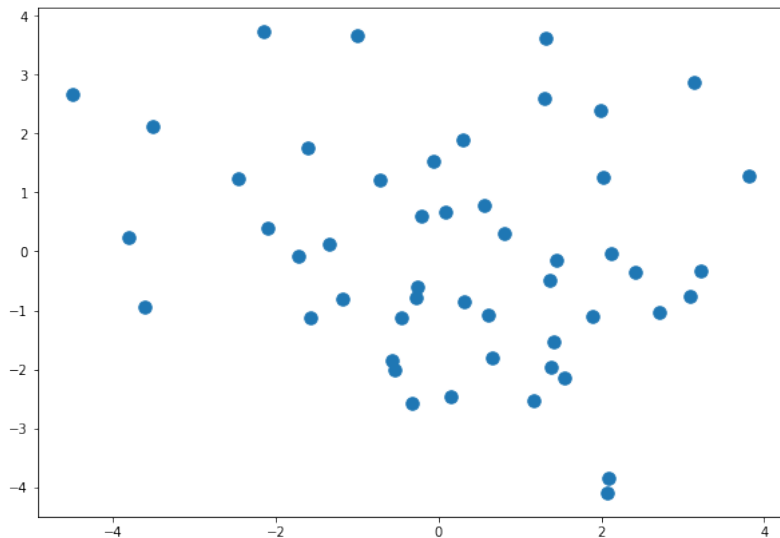
- ▶ Тогда площадь увеличивается в $\det(\mathbf{A}^{-1})$ раз.
- ▶ Детерминант не является выпуклой/вогнутой функцией
- ▶ $\log \det(\mathbf{A}^{-1}) = -\log \det(\mathbf{A})$ – выпуклая функция при $\mathbf{A} \succ 0$

$$\begin{aligned} & \min_{\mathbf{A}, \mathbf{b}} \log \det \mathbf{A}^{-1} \\ & \text{s.t. } \mathbf{A} \succ 0 \\ & \quad \|\mathbf{A}\mathbf{x}_i + \mathbf{b}\|_2 \leq 1 \end{aligned}$$

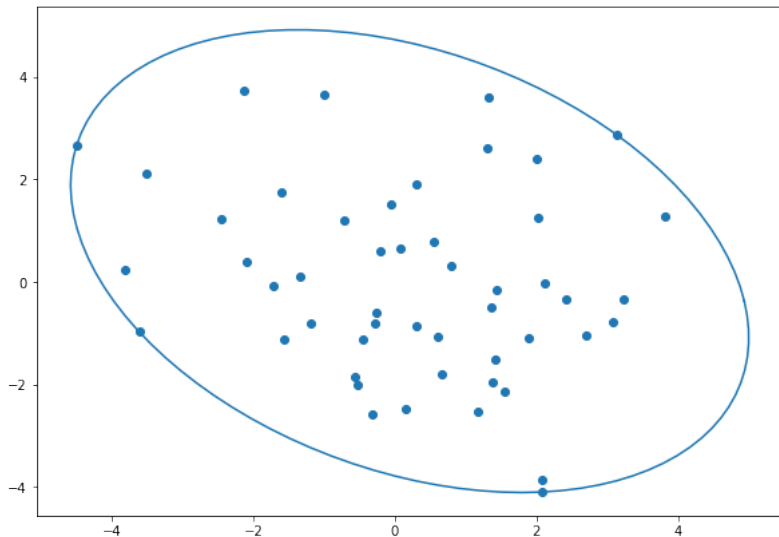
Эллипсоид Löwner-John

Постановка аналогична только не для точек, а для некоторого выпуклого множества

Пример построения экстремального эллипсоида



Пример построения экстремального эллипса



Равносильные преобразования задач

Равносильные преобразования задач

Запись через надграфик

$$\begin{array}{ll} \min_{\mathbf{x}} f_0(\mathbf{x}) & \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} & \\ f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, m & \end{array} \quad \Longrightarrow \quad \begin{array}{ll} \min_{\mathbf{x}, t} t & \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} & \\ f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, m & \\ f_0(\mathbf{x}) \leq t & \end{array}$$

Равносильные преобразования задач

Запись через надграфик

$$\begin{array}{ll} \min_{\mathbf{x}} f_0(\mathbf{x}) & \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} & \\ f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, m & \end{array} \quad \Longrightarrow \quad \begin{array}{ll} \min_{\mathbf{x}, t} t & \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} & \\ f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, m & \\ f_0(\mathbf{x}) \leq t & \end{array}$$

Преобразования ограничений

- ▶ $\mathbf{Ax} \leq \mathbf{b} \rightarrow \mathbf{Ax} + \mathbf{y} = \mathbf{b}, \mathbf{y} \geq 0$
- ▶ $\mathbf{x} \rightarrow \mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_1 \geq 0, ; \mathbf{x}_2 \geq 0$
- ▶ Формирование блочных матриц

Равносильные преобразования задач

Запись через надграфик

$$\begin{array}{ll} \min_{\mathbf{x}} f_0(\mathbf{x}) & \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} & \\ f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, m & \end{array} \quad \Longrightarrow \quad \begin{array}{l} \min_{\mathbf{x}, t} t \\ \text{s.t. } \mathbf{Ax} = \mathbf{b} \\ f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, m \\ f_0(\mathbf{x}) \leq t \end{array}$$

Преобразования ограничений

- ▶ $\mathbf{Ax} \leq \mathbf{b} \rightarrow \mathbf{Ax} + \mathbf{y} = \mathbf{b}, \mathbf{y} \geq 0$
- ▶ $\mathbf{x} \rightarrow \mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_1 \geq 0, ; \mathbf{x}_2 \geq 0$
- ▶ Формирование блочных матриц

Перенос ограничений в целевую функцию

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \Rightarrow \min_{\mathbf{x} \in \mathbb{R}^n} f_0(\mathbf{x}) + \mathbb{I}_X(\mathbf{x}), \quad \mathbb{I}_X(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in X, \\ +\infty, & \mathbf{x} \notin X. \end{cases}$$

- ▶ Стандартная форма задачи выпуклой оптимизации

Главное

- ▶ Стандартная форма задачи выпуклой оптимизации
- ▶ Линейное программирование

- ▶ Стандартная форма задачи выпуклой оптимизации
- ▶ Линейное программирование
- ▶ Коническая оптимизация: LP, SOCP и SDP

- ▶ Стандартная форма задачи выпуклой оптимизации
- ▶ Линейное программирование
- ▶ Коническая оптимизация: LP, SOCP и SDP
- ▶ Решение линейных систем с неквадратными матрицами

- ▶ Стандартная форма задачи выпуклой оптимизации
- ▶ Линейное программирование
- ▶ Коническая оптимизация: LP, SOCP и SDP
- ▶ Решение линейных систем с неквадратными матрицами
- ▶ Задача классификации и SVM

- ▶ Стандартная форма задачи выпуклой оптимизации
- ▶ Линейное программирование
- ▶ Коническая оптимизация: LP, SOCP и SDP
- ▶ Решение линейных систем с неквадратными матрицами
- ▶ Задача классификации и SVM
- ▶ Выпуклая релаксация задачи MAXCUT

- ▶ Стандартная форма задачи выпуклой оптимизации
- ▶ Линейное программирование
- ▶ Коническая оптимизация: LP, SOCP и SDP
- ▶ Решение линейных систем с неквадратными матрицами
- ▶ Задача классификации и SVM
- ▶ Выпуклая релаксация задачи MAXCUT
- ▶ Задача построения оптимального эллипсоида