Методы оптимизации Лекция 3: Примеры задач выпуклой оптимизации

Александр Катруца

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



21 сентября 2020 г.

На прошлой лекции

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

Уточнение про конусы и отношение порядка

Определение

Конус K называется правильным (proper), если

- К выпуклый
- К замкнутый
- ightharpoonup K не содежит прямых
- ightharpoonup внутренность K непуста

Уточнение про конусы и отношение порядка

Определение

Конус K называется правильным (proper), если

- К выпуклый
- К замкнутый
- ightharpoonup K не содежит прямых
- ightharpoonup внутренность K непуста

Дополнение к прошлой лекции

Упорядочивание элементов относительно конусов справедливо только для правильных конусов.

Определение

```
Функция f: X \subset \mathbb{R}^n \to \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)
```

Определение

```
Функция f: X \subset \mathbb{R}^n \to \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 $\Leftrightarrow \operatorname{epi} f$ выпуклое множество.

Определение

Функция $f: X \subset \mathbb{R}^n \to \mathbb{R}$ 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 $\Leftrightarrow \operatorname{epi} f$ 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

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

Определение

Функция $f: X \subset \mathbb{R}^n \to \mathbb{R}$ 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 $\Leftrightarrow \operatorname{epi} f$ 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

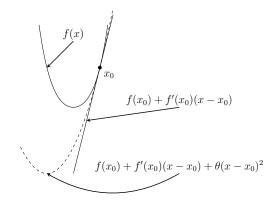
$$f(\mathbf{x}) - f(\mathbf{x}^*) \ge \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

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

Определение

Определение

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

 $ightharpoonup \mathcal{C}^n$ выпуклое множество

Определение

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

Определение

- $ightharpoonup \mathcal{C}^n$ выпуклое множество
- $ightharpoonup \mathbf{S}_+^n \subset \mathcal{C}^n$
- lacktriangle Задача проверки $\mathbf{X}
 ot\in \mathcal{C}^n$ является со-NP полной!

Определение

- $ightharpoonup \mathcal{C}^n$ выпуклое множество
- $ightharpoonup \mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки $\mathbf{X} \not\in \mathcal{C}^n$ является со-NP полной!
- ightharpoonup Задача конической оптимизации с конусом \mathcal{C}^n является NP-трудной

Определение

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

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

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

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

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

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

Допустимое множество невыпукло

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

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

- ▶ Допустимое множество невыпукло
- ▶ Целевая функция выпукла

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

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

- ▶ Допустимое множество невыпукло
- Целевая функция выпукла

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

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

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

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

• A – диагонализуема, поэтому есть базис из собственных векторов $[{\bf v}_1,\dots,{\bf v}_n]$ причём предположим $|\lambda_1|>|\lambda_2|>\dots>|\lambda_n|$

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

- A диагонализуема, поэтому есть базис из собственных векторов $[{\bf v}_1,\dots,{\bf v}_n]$ причём предположим $|\lambda_1|>|\lambda_2|>\dots>|\lambda_n|$
- $\|\mathbf{x}_k \mathbf{v}_1\|_2 \le \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k \|\mathbf{x}_0 \mathbf{v}_1\|_2$

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

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

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

- **A** диагонализуема, поэтому есть базис из собственных векторов $[{\bf v}_1,\dots,{\bf v}_n]$ причём предположим $|\lambda_1|>|\lambda_2|>\dots>|\lambda_n|$
- $\|\mathbf{x}_k \mathbf{v}_1\|_2 \le \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k \|\mathbf{x}_0 \mathbf{v}_1\|_2$
- 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

$$\min f_0(\mathbf{x})$$

s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$
 $f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, p$

- $ightharpoonup f_0$ выпуклая целевая функция
- $ightharpoonup f_i$ выпуклые функции для ограничений типа неравенств
- Ограничения типа равенств только линейные

Стандартная форма записи задачи выпуклой оптимизации

$$\min f_0(\mathbf{x})$$

s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$
 $f_i(\mathbf{x}) \leq 0, \ i = 1, \dots, p$

- ▶ f₀ выпуклая целевая функция
- $ightharpoonup f_i$ выпуклые функции для ограничений типа неравенств
- Ограничения типа равенств только линейные

Ограничения стандартной формы записи

Выпуклое множество может быть задано более общим образом

$$\min x^{2}$$
s.t. $(x-2)^{2}=0$

$$x^{3} \ge 0$$

$$\min \mathbf{c}^{\top} \mathbf{x}$$

s.t.
$$Ax = b$$

$$x_i \ge 0, \ i = 1, \dots, n$$

$$\min \mathbf{c}^{\top} \mathbf{x}$$

s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$
 $x_i \ge 0, \ i = 1, \dots, n$

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

$$\min \mathbf{c}^{\top} \mathbf{x}$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$x_i > 0, \ i = 1, \dots, n$$

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

 $\min \mathbf{c}^{\mathsf{T}} \mathbf{x}$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

s.t.
$$\mathbf{A}\mathbf{x} = \mathbf{b}$$
 $x_i \geq 0, \ i = 1, \dots, n$

- 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

 $\min \mathbf{c}^{\top} \mathbf{x}$ s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$ $x_i > 0, \ i = 1, \dots, n$

- 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

$$\min \mathbf{c}^{\top} \mathbf{x}$$

s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$
 $x_i \ge 0, \ i = 1, \dots, n$

- 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

lacktriangle Дано n продуктов, цена единицы каждого c_i

$$\min \mathbf{c}^{\top} \mathbf{x}$$

s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$
 $x_i \ge 0, \ i = 1, \dots, n$

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

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

$$\min \mathbf{c}^{\top} \mathbf{x}$$

s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$
 $x_i \ge 0, \ i = 1, \dots, n$

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

- lacktriangle 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

$$\min \mathbf{c}^{\top} \mathbf{x}$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$x_i > 0, \ i = 1, \dots, n$$

- 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

- lacktriangle 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

$$\min \frac{1}{2} \mathbf{x}^{\top} \mathbf{P}_0 \mathbf{x} + \mathbf{q}^{\top} \mathbf{x} + r_0$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\frac{1}{2} \mathbf{x}^{\top} \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^{\top} \mathbf{x} + r_i \le 0, \ i = 1, \dots, n$$

$$\min \frac{1}{2} \mathbf{x}^{\top} \mathbf{P}_0 \mathbf{x} + \mathbf{q}^{\top} \mathbf{x} + r_0$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\frac{1}{2} \mathbf{x}^{\top} \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^{\top} \mathbf{x} + r_i \leq 0, \ i = 1, \dots, n$$

lacktriangle Задача будет выпукла, если все $\mathbf{P}_i \in \mathbf{S}^n_+$

$$\min \frac{1}{2} \mathbf{x}^{\top} \mathbf{P}_0 \mathbf{x} + \mathbf{q}^{\top} \mathbf{x} + r_0$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\frac{1}{2} \mathbf{x}^{\top} \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^{\top} \mathbf{x} + r_i \leq 0, \ i = 1, \dots, n$$

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

$$\min \frac{1}{2} \mathbf{x}^{\top} \mathbf{P}_0 \mathbf{x} + \mathbf{q}^{\top} \mathbf{x} + r_0$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\frac{1}{2} \mathbf{x}^{\top} \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^{\top} \mathbf{x} + r_i \le 0, \ i = 1, \dots, n$$

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

Линейная задача наименьших квадратов

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

Возможно добавление линейных ограничений вида $l_i \leq x_i \leq u_i$

Линейная задача наименьших квадратов

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

Возможно добавление линейных ограничений вида $l_i \leq x_i \leq u_i$

▶ Поиск решения минимальной нормы

$$\min \frac{1}{2} \|\mathbf{x}\|_2^2$$
 s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$

Линейная задача наименьших квадратов

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

Возможно добавление линейных ограничений вида $l_i \leq x_i \leq u_i$

▶ Поиск решения минимальной нормы

$$\min \frac{1}{2}\|\mathbf{x}\|_2^2$$
 s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$

Q: Каков геометрический смысл у решения?

Линейная задача наименьших квадратов

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

Возможно добавление линейных ограничений вида $l_i \leq x_i \leq u_i$

▶ Поиск решения минимальной нормы

$$\min \frac{1}{2}\|\mathbf{x}\|_2^2$$
 s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$

Q: Каков геометрический смысл у решения?

Q: К какой задаче сводится похожая задача?

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

s.t.
$$Ax = b$$

ightharpoonup n активов

- \triangleright n активов
- lacktriangle Изменение относительной цены активов случайный вектор со средним $ar{\mathbf{p}}$ и ковариационной матрицей $oldsymbol{\Sigma}$

- \triangleright n активов
- lacktriangle Изменение относительной цены активов случайный вектор со средним $ar{\mathbf{p}}$ и ковариационной матрицей $oldsymbol{\Sigma}$
- lacktriangle Минимально допустимый средний доход ar r

Дано

- \triangleright n активов
- lacktriangle Изменение относительной цены активов случайный вектор со средним $ar{\mathbf{p}}$ и ковариационной матрицей $oldsymbol{\Sigma}$
- lacktriangle Минимально допустимый средний доход ar r

Классическая задача составления оптимального портфеля

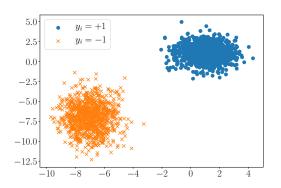
$$\min \mathbf{x}^{\top} \mathbf{\Sigma} \mathbf{x}$$
s.t. $\bar{\mathbf{p}}^{\top} \mathbf{x} \geq \bar{r}$

$$\sum_{i=1}^{n} x_i = 1, \ x_i \geq 0, \ i = 1, \dots, n$$

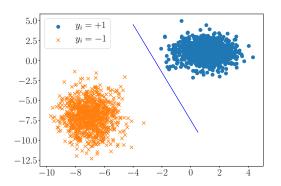
- Минимум риска при минимально допустимом доходе
- Существуют многочисленные вариации, которые не выводят задачу из класса QCQP или SOCP

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

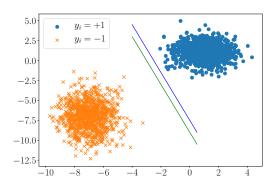
- lacktriangle Дана выборка (\mathbf{x}_i,y_i) , где $\mathbf{x}_i\in\mathbb{R}^n$, а $y_i\in\{-1,+1\}$
- $f extbf{ iny}$ Необходимо построить гиперплоскость так, чтобы ${f a}^{ op}{f x}_i+b>0$, если $y_i=+1$ и ${f a}^{ op}{f x}_i+b<0$, если $y_i<0$



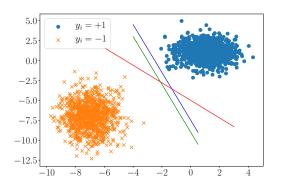
- lacktriangle Дана выборка (\mathbf{x}_i,y_i) , где $\mathbf{x}_i\in\mathbb{R}^n$, а $y_i\in\{-1,+1\}$
- $f extbf{ iny}$ Необходимо построить гиперплоскость так, чтобы ${f a}^{ op}{f x}_i+b>0$, если $y_i=+1$ и ${f a}^{ op}{f x}_i+b<0$, если $y_i<0$



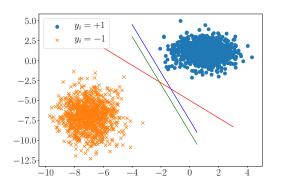
- lacktriangle Дана выборка (\mathbf{x}_i,y_i) , где $\mathbf{x}_i\in\mathbb{R}^n$, а $y_i\in\{-1,+1\}$
- $f extbf{ iny}$ Необходимо построить гиперплоскость так, чтобы ${f a}^{ op}{f x}_i+b>0$, если $y_i=+1$ и ${f a}^{ op}{f x}_i+b<0$, если $y_i<0$



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



- lacktriangle Дана выборка (\mathbf{x}_i,y_i) , где $\mathbf{x}_i\in\mathbb{R}^n$, а $y_i\in\{-1,+1\}$
- $f extbf{ iny Heofxoдимо}$ построить гиперплоскость так, чтобы ${f a}^{ op}{f x}_i+b>0$, если $y_i=+1$ и ${f a}^{ op}{f x}_i+b<0$, если $y_i<0$



Q: Как однозначно задать разделяющую гиперплоскость?

▶ Для опорных объектов каждого класса выполнено

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

▶ 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

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

▶ Для опорных объектов каждого класса выполнено

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

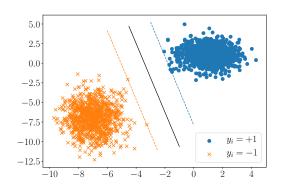
• Расстояние между гиперплоскостями

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

Финальная задача

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

Оптимальная гиперплоскость



Оптимизация на конусе второго порядка (SOCP)

$$\min \mathbf{f}^{\top} \mathbf{x}$$
s.t. $\|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\|_2 \le \mathbf{c}_i^{\top} \mathbf{x} + d_i$

$$\mathbf{F} \mathbf{x} = \mathbf{g}$$

Коническая форма записи

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

$QCQP \rightarrow SOCP$

- ▶ Пусть $\mathbf{x}^{\top}\mathbf{P}\mathbf{x} + \mathbf{q}^{\top}\mathbf{x} + r \leq 0$ и $0 \prec \mathbf{P} = \mathbf{L}\mathbf{L}^{\top}$
- $(\mathbf{L}^{\top}\mathbf{x})^{\top}(\mathbf{L}^{\top}\mathbf{x}) + 2\tilde{\mathbf{q}}^{\top}\mathbf{L}^{-\top}\mathbf{L}^{\top}\mathbf{x} + \|\mathbf{L}^{-1}\tilde{\mathbf{q}}\|_{2}^{2} \le \|\mathbf{L}^{-1}\tilde{\mathbf{q}}\|_{2}^{2} r$
- ▶ 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 ightarrow SOCP: преобразование целевой функции

$$\min_{\mathbf{x}} \mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0} \implies \min_{\mathbf{x}, t} t + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0}$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} \le t$$

QCQP ightarrow SOCP: преобразование целевой функции

$$\min_{\mathbf{x}} \mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0} \implies \min_{\mathbf{x}, t} t + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0}$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} \le t$$

lacktriangle Tak kak $\mathbf{P}_0 \succ 0$, to $\mathbf{P} = \mathbf{L}\mathbf{L}^{ op}$

QCQP o SOCP: преобразование целевой функции

$$\min_{\mathbf{x}} \mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0} \implies \min_{\mathbf{x}, t} t + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0}$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} \le t$$

- lacktriangle Tak kak $\mathbf{P}_0 \succ 0$, to $\mathbf{P} = \mathbf{L} \mathbf{L}^ op$
- ightharpoonup Тогда $\mathbf{x}^{ op}\mathbf{P}_0\mathbf{x} \leq t$ сводится к $\|\mathbf{L}^{ op}\mathbf{x}\|_2^2 \leq t$

QCQP o SOCP: преобразование целевой функции

$$\min_{\mathbf{x}} \mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0} \implies \min_{\mathbf{x}, t} t + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0}$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} \le t$$

- lacktriangle Tak kak $\mathbf{P}_0 \succ 0$, to $\mathbf{P} = \mathbf{L} \mathbf{L}^ op$
- ▶ 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: $\mathcal{Q}_{\mathsf{rot}}^n = \{ (\mathbf{x}, y, z) \in \mathbb{R}^{n+2} \mid \|\mathbf{x}\|_2 \leq \sqrt{yz}, \ y \geq 0, \ z \geq 0 \}$

QCQP ightarrow SOCP: преобразование целевой функции

$$\min_{\mathbf{x}} \mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0} \implies \min_{\mathbf{x}, t} t + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0}$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} \le t$$

- lacktriangle Tak kak $\mathbf{P}_0 \succ 0$, to $\mathbf{P} = \mathbf{L} \mathbf{L}^ op$
- ► 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: $Q_{\text{rot}}^n = \{(\mathbf{x}, y, z) \in \mathbb{R}^{n+2} \mid ||\mathbf{x}||_2 \leq \sqrt{yz}, \ y \geq 0, \ z \geq 0\}$
- ▶ Условие $\|\mathbf{x}\|_2 \leq \sqrt{yz}, \; y,z \geq 0 \Leftrightarrow \|(2\mathbf{x},y-z)\|_2 \leq y+z$

QCQP o SOCP: преобразование целевой функции

$$\min_{\mathbf{x}} \mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0} \implies \min_{\mathbf{x}, t} t + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0}$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} \le t$$

- lacktriangle Tak kak $\mathbf{P}_0 \succ 0$, to $\mathbf{P} = \mathbf{L} \mathbf{L}^ op$
- ▶ 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 $(\mathbf{L}^{\top}\mathbf{x},t,1) \in \mathcal{Q}_{\mathsf{rot}}^n$

QCQP o SOCP: преобразование целевой функции

$$\min_{\mathbf{x}} \mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0} \implies \min_{\mathbf{x}, t} t + \mathbf{q}_{0}^{\top} \mathbf{x} + r_{0}$$
s.t. $\mathbf{A} \mathbf{x} = \mathbf{b}$

$$\mathbf{x}^{\top} \mathbf{P}_{0} \mathbf{x} \le t$$

- lacktriangle Tak kak $\mathbf{P}_0 \succ 0$, to $\mathbf{P} = \mathbf{L} \mathbf{L}^ op$
- ▶ 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: $Q_{\text{rot}}^n = \{(\mathbf{x}, y, z) \in \mathbb{R}^{n+2} \mid ||\mathbf{x}||_2 \leq \sqrt{yz}, \ y \geq 0, \ z \geq 0\}$
- ▶ Условие $\|\mathbf{x}\|_2 \leq \sqrt{yz}, \; y,z \geq 0 \Leftrightarrow \|(2\mathbf{x},y-z)\|_2 \leq y+z$
- ▶ Тогда $(\mathbf{L}^{\top}\mathbf{x},t,1) \in \mathcal{Q}_{\mathsf{rot}}^n$

Отношения между рассмотренными типами задач

$$LP \subset QCQP \subset SOCP$$

Задача геометрического программирования (geometric programming)

Определения

- lacktriangle Функцию вида $f(\mathbf{x}) = cx_1^{a_1} \dots x_n^{a_n}$, где $c>0, a_i \in \mathbb{R}$, $\mathrm{dom} f = \mathbb{R}^n_{++}$ называют обобщённым мономом
- lacktriangle Функцию вида $f(\mathbf{x})=\sum_{k=1}^K c_k x_1^{a_{1k}}\dots x_n^{a_{nk}}$, где $c_k>0, a_{ik}\in\mathbb{R}$ и $\mathrm{dom} f=\mathbb{R}^n_{++}$ называют позиномом

Задача геометрического программирования (geometric programming)

Определения

- lacktriangle Функцию вида $f(\mathbf{x}) = cx_1^{a_1} \dots x_n^{a_n}$, где $c>0, a_i \in \mathbb{R}$, $\mathrm{dom} f = \mathbb{R}^n_{++}$ называют обобщённым мономом
- lacktriangle Функцию вида $f(\mathbf{x})=\sum_{k=1}^K c_k x_1^{a_{1k}}\dots x_n^{a_{nk}}$, где $c_k>0, a_{ik}\in\mathbb{R}$ и $\mathrm{dom} f=\mathbb{R}^n_{++}$ называют позиномом

Общий вид задачи

$$\min f_0(\mathbf{x})$$

s.t. $f_i(\mathbf{x}) \leq 1$
 $h_i(\mathbf{x}) = 1$,

где f_i — позиномы, а h_j — обобщённые мономы

Выпуклая форма задачи GP

▶ Преобразование переменных: $x_i = e^{y_i}$

Выпуклая форма задачи GP

- Преобразование переменных: $x_i = e^{y_i}$
- ▶ Обобщенный моном превращается в $ilde{f}(\mathbf{x}) = e^{\mathbf{a}^{\top}\mathbf{y} + b}$

- ▶ Преобразование переменных: $x_i = e^{y_i}$
- ▶ Обобщенный моном превращается в $\widetilde{f}(\mathbf{x}) = e^{\mathbf{a}^{\top}\mathbf{y} + b}$
- lacktriangle Позином превращается в $ilde{f}(\mathbf{x}) = \sum_{k=1}^K e^{\mathbf{a}_k^{ op} \mathbf{y} + b_k}$

- Преобразование переменных: $x_i = e^{y_i}$
- ▶ Обобщенный моном превращается в $\tilde{f}(\mathbf{x}) = e^{\mathbf{a}^{\top}\mathbf{y} + b}$
- ▶ Позином превращается в $\tilde{f}(\mathbf{x}) = \sum_{k=1}^K e^{\mathbf{a}_k^{\top}\mathbf{y} + b_k}$
- ▶ Общий вид переписывается в выпуклой форме

- Преобразование переменных: $x_i = e^{y_i}$
- ▶ Обобщенный моном превращается в $\tilde{f}(\mathbf{x}) = e^{\mathbf{a}^{\top}\mathbf{y} + b}$
- ▶ Позином превращается в $\tilde{f}(\mathbf{x}) = \sum_{k=1}^K e^{\mathbf{a}_k^{\top}\mathbf{y} + b_k}$
- ▶ Общий вид переписывается в выпуклой форме

- Преобразование переменных: $x_i = e^{y_i}$
- ▶ Обобщенный моном превращается в $\tilde{f}(\mathbf{x}) = e^{\mathbf{a}^{\top}\mathbf{y} + b}$
- lacktriangle Позином превращается в $ilde{f}(\mathbf{x}) = \sum_{k=1}^K e^{\mathbf{a}_k^{ op} \mathbf{y} + b_k}$
- Общий вид переписывается в выпуклой форме

$$\min_{\mathbf{y}} \sum_{k=1}^{K} e^{\mathbf{a}_{k0}^{\top} \mathbf{y} + b_{k0}}$$
s.t.
$$\sum_{k=1}^{K} e^{\mathbf{a}_{ki}^{\top} \mathbf{y} + b_{ki}} \leq 1$$

$$e^{\mathbf{d}_{j}^{\top} \mathbf{y} + p_{j}} = 1$$

- Преобразование переменных: $x_i = e^{y_i}$
- ▶ Обобщенный моном превращается в $\tilde{f}(\mathbf{x}) = e^{\mathbf{a}^{\top}\mathbf{y} + b}$
- ▶ Позином превращается в $\tilde{f}(\mathbf{x}) = \sum_{k=1}^K e^{\mathbf{a}_k^{\top}\mathbf{y} + b_k}$
- ▶ Общий вид переписывается в выпуклой форме

$$\begin{aligned} & \min_{\mathbf{y}} \sum_{k=1}^{K} e^{\mathbf{a}_{k0}^{\top} \mathbf{y} + b_{k0}} & \min_{\mathbf{y}} \log \left(\sum_{k=1}^{K} e^{\mathbf{a}_{k0}^{\top} \mathbf{y} + b_{k0}} \right) \\ \text{s.t. } & \sum_{k=1}^{K} e^{\mathbf{a}_{ki}^{\top} \mathbf{y} + b_{ki}} \leq 1 & \text{s.t. } \log \left(\sum_{k=1}^{K} e^{\mathbf{a}_{ki}^{\top} \mathbf{y} + b_{ki}} \right) \leq 0 \\ & e^{\mathbf{d}_{j}^{\top} \mathbf{y} + p_{j}} = 1 & \mathbf{d}_{j}^{\top} \mathbf{y} + p_{j} = 0 \end{aligned}$$

lacktriangle Пусть есть n трансмиттеров и n ресиверов

- lacktriangle Пусть есть n трансмиттеров и n ресиверов
- lacktriangle Мощность сигнала p_i на трансмиттере

- ightharpoonup Пусть есть n трансмиттеров и n ресиверов
- lacktriangle Мощность сигнала p_i на трансмиттере
- lacktriangle Мощность сигнала на i-ом ресивере $G_{ii}p_i$, $G_{ii}>0$

- ightharpoonup Пусть есть n трансмиттеров и n ресиверов
- ightharpoonup Мощность сигнала p_i на трансмиттере
- ▶ Мощность сигнала на i-ом ресивере $G_{ii}p_i$, $G_{ii}>0$
- lacktriangle Мощность интерференции на i-ом ресивере $\sum_{j
 eq i} G_{ij} p_j$

- ightharpoonup Пусть есть n трансмиттеров и n ресиверов
- lacktriangle Мощность сигнала p_i на трансмиттере
- ▶ Мощность сигнала на i-ом ресивере $G_{ii}p_i$, $G_{ii}>0$
- lacktriangle Мощность интерференции на i-ом ресивере $\sum_{j
 eq i} G_{ij} p_j$
- lacktriangle Отношение сингал-шум (SINR) $rac{G_{ii}p_i}{\sigma_i+\sum_{j
 eq i}G_{ij}p_j}$, где σ_i мощность шума на i-ом ресивере

- ightharpoonup Пусть есть n трансмиттеров и n ресиверов
- lacktriangle Мощность сигнала p_i на трансмиттере
- ▶ Мощность сигнала на i-ом ресивере $G_{ii}p_i$, $G_{ii}>0$
- lacktriangle Мощность интерференции на i-ом ресивере $\sum_{j
 eq i} G_{ij} p_j$
- lacktriangle Отношение сингал-шум (SINR) $\frac{G_{ii}p_i}{\sigma_i+\sum_{j
 eq i}G_{ij}p_j}$, где σ_i мощность шума на i-ом ресивере

Задача минимизации мощности

$$\begin{split} \min_{\mathbf{p}} \sum_{i=1} p_i \\ \text{s.t. } p^{\min} &\leq p_i \leq p^{\max} \\ \frac{G_{ii}p_i}{\sigma_i + \sum_{j \neq i} G_{ij}p_j} \geq S_i \end{split}$$

- ightharpoonup Пусть есть n трансмиттеров и n ресиверов
- lacktriangle Мощность сигнала p_i на трансмиттере
- lacktriangle Мощность сигнала на i-ом ресивере $G_{ii}p_i$, $G_{ii}>0$
- lacktriangle Мощность интерференции на i-ом ресивере $\sum_{j
 eq i} G_{ij} p_j$
- lacktriangle Отношение сингал-шум (SINR) $rac{G_{ii}p_i}{\sigma_i+\sum_{j
 eq i}G_{ij}p_j}$, где σ_i мощность шума на i-ом ресивере

Задача минимизации мощности

$$egin{aligned} \min & \sum_{i=1}^n p_i \ ext{s.t.} & p^{\min} & \leq p_i \leq p^{\max} \ & rac{G_{ii}p_i}{\sigma_i + \sum_{i
eq i} G_{ij}p_j} \geq S_i \end{aligned}$$

Больше примеров можно найти в этом туториале 1

https://web.stanford.edu/~boyd/papers/pdf/gp_tutorial.pdf

 После равносильных преобразований задач достаточно описать неравенство

$$\log\left(\sum_{k=1}^{m} e^{z_k}\right) \le t$$

в конической форме

 После равносильных преобразований задач достаточно описать неравенство

$$\log\left(\sum_{k=1}^{m} e^{z_k}\right) \le t$$

в конической форме

lacktriangle Эквивалентная запись $\sum_{k=1}^m e^{z_k-t} \leq 1$

 После равносильных преобразований задач достаточно описать неравенство

$$\log\left(\sum_{k=1}^{m} e^{z_k}\right) \le t$$

в конической форме

- ▶ Эквивалентная запись $\sum_{k=1}^{m} e^{z_k t} \le 1$
- ▶ Введём переменную $e^{z_k-t} \le u_k$

 После равносильных преобразований задач достаточно описать неравенство

$$\log\left(\sum_{k=1}^{m} e^{z_k}\right) \le t$$

в конической форме

- lacktriangle Эквивалентная запись $\sum_{k=1}^m e^{z_k-t} \leq 1$
- ▶ Введём переменную $e^{z_k-t} \le u_k$
- ▶ Тогда $\sum_{k=1}^m u_k \le 1$ линейное неравенство

 После равносильных преобразований задач достаточно описать неравенство

$$\log\left(\sum_{k=1}^{m} e^{z_k}\right) \le t$$

в конической форме

- ▶ Эквивалентная запись $\sum_{k=1}^{m} e^{z_k t} \le 1$
- ▶ Введём переменную $e^{z_k-t} \le u_k$
- ▶ Тогда $\sum_{k=1}^m u_k \le 1$ линейное неравенство

Определение

Экспоненциальным конусом называется такой конус $K_{\text{exp}} = \{(x,y,z) \mid x \geq ye^{z/y}, \ y > 0\} \cup \{(x,0,z) \mid x > 0, \ z < 0\}$

 После равносильных преобразований задач достаточно описать неравенство

$$\log\left(\sum_{k=1}^{m} e^{z_k}\right) \le t$$

в конической форме

- ▶ Эквивалентная запись $\sum_{k=1}^{m} e^{z_k t} \le 1$
- ▶ Введём переменную $e^{z_k-t} \le u_k$
- ▶ Тогда $\sum_{k=1}^m u_k \le 1$ линейное неравенство

Определение

Экспоненциальным конусом называется такой конус

$$K_{\text{exp}} = \{(x, y, z) \mid x \ge ye^{z/y}, \ y > 0\} \cup \{(x, 0, z) \mid x \ge 0, \ z \le 0\}$$

 $e^{z_k-t} \le u_k \Leftrightarrow (u_k, 1, z_k - t) \in K_{\text{exp}}$

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

Коническая форма

$$\min_{\mathbf{x}} \mathbf{c}^{\top} \mathbf{x}$$

s.t.
$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

$$\mathbf{G} + \sum_{i=1}^{n} x_i \mathbf{F}_i \leq 0,$$

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

Стандартная форма

$$\min_{\mathbf{X}} \operatorname{trace}(\mathbf{CX})$$

s.t.
$$\operatorname{trace}(\mathbf{A}_i \mathbf{X}) = b_i$$

 $\mathbf{X} \succeq 0$,

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 Стандартная форма
$$\min_{\mathbf{x}} \mathbf{c}^{\top} \mathbf{x}$$
 $\min_{\mathbf{X}} \operatorname{trace}(\mathbf{C}\mathbf{X})$ s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$ s.t. $\operatorname{trace}(\mathbf{A}_i\mathbf{X}) = b_i$ $\mathbf{X} \succeq 0,$ где $\mathbf{C} \in \mathbf{S}^n$ и все $\mathbf{A}_i \in \mathbf{S}^n$

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

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

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

Коническая форма Стандартная форма
$$\min_{\mathbf{x}} \mathbf{c}^{\top} \mathbf{x}$$
 $\min_{\mathbf{X}} \operatorname{trace}(\mathbf{C}\mathbf{X})$ s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$ s.t. $\operatorname{trace}(\mathbf{A}_i\mathbf{X}) = b_i$ $\mathbf{X} \succeq 0,$ где $\mathbf{C} \in \mathbf{S}^n$ и все $\mathbf{A}_i \in \mathbf{S}^n$

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

- ▶ 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 Стандартная форма
$$\min_{\mathbf{x}} \mathbf{c}^{\top} \mathbf{x}$$
 $\min_{\mathbf{X}} \operatorname{trace}(\mathbf{C}\mathbf{X})$ s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$ s.t. $\operatorname{trace}(\mathbf{A}_i\mathbf{X}) = b_i$ $\mathbf{X} \succeq 0,$ где $\mathbf{C} \in \mathbf{S}^n$ и все $\mathbf{A}_i \in \mathbf{S}^n$

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

- ▶ 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

- $\mathbf{G} = 0$
- $ightharpoonup \mathbf{F}_i$ такие, что $\sum\limits_{i=1}^n x_i \mathbf{F}_i = -\mathrm{diag}(\mathbf{x})$

LP

- G = 0
- $ightharpoonup \mathbf{F}_i$ такие, что $\sum\limits_{i=1}^n x_i \mathbf{F}_i = -\mathrm{diag}(\mathbf{x})$

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

$$\mathbf{G} = 0$$

$$ightharpoonup \mathbf{F}_i$$
 такие, что $\sum\limits_{i=1}^n x_i \mathbf{F}_i = -\mathrm{diag}(\mathbf{x})$

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

- $\|\mathbf{x}\|_2 \le t \Leftrightarrow \begin{bmatrix} t & \mathbf{x}^{\top} \\ \mathbf{x} & t\mathbf{I} \end{bmatrix} \succeq 0$
- ▶ Аналогично для Q^n :

$$\|\mathbf{A}^{\top}\mathbf{x} + \mathbf{b}_i\|_2 \le \mathbf{c}^{\top}\mathbf{x} + d_i \Leftrightarrow \begin{bmatrix} \mathbf{c}^{\top}\mathbf{x} + d_i & (\mathbf{A}^{\top}\mathbf{x} + \mathbf{b}_i)^{\top} \\ \mathbf{A}^{\top}\mathbf{x} + \mathbf{b}_i & (\mathbf{c}^{\top}\mathbf{x} + d_i)\mathbf{I} \end{bmatrix} \succeq 0$$

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

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

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

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

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

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

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

▶ В матрично-векторном виде

$$\min \frac{1}{2}\mathbf{x}^{\top}\mathbf{W}\mathbf{x}$$
 s.t. $x_i \in \{-1, +1\}$

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

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

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

▶ В матрично-векторном виде

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

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

Эквивалентный вид

$$\min \frac{1}{2} trace(\mathbf{WX})$$
 s.t. $\mathbf{X} \succeq 0$, $rank(\mathbf{X}) = 1$, $diag(\mathbf{X}) = \mathbf{1}$

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

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

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

▶ В матрично-векторном виде

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

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

Эквивалентный вид

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

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

SDP релаксация

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

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

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

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

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

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

- ightharpoonup 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

$$\{\mathbf{x} \mid \|\mathbf{x}\|_{2}^{2} \le 1\} \to \{\mathbf{u} \mid \|\mathbf{u}\|_{2}^{2} \le 1, \ \mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$$

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

- ightharpoonup 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

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

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

Экстремальный эллипсоид для данного множества точек

- ightharpoonup 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

$$\{\mathbf{x} \mid \|\mathbf{x}\|_{2}^{2} \le 1\} \to \{\mathbf{u} \mid \|\mathbf{u}\|_{2}^{2} \le 1, \ \mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$$

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

Экстремальный эллипсоид для данного множества точек

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

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

- lacktriangle 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 $A \succ 0$

$$\min_{\mathbf{A}, \mathbf{b}} \log \det \mathbf{A}^{-1}$$
s.t. $\mathbf{A} \succ 0$

$$\|\mathbf{A}\mathbf{x}_i + \mathbf{b}\|_2 \le 1$$

Экстремальный эллипсоид для данного множества точек

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

$$\{\mathbf{x} \mid \|\mathbf{x}\|_{2}^{2} \le 1\} \to \{\mathbf{u} \mid \|\mathbf{u}\|_{2}^{2} \le 1, \ \mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$$

- lacktriangle 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 $A \succ 0$

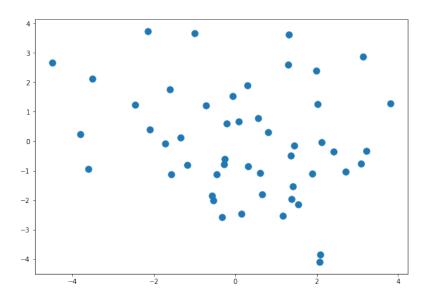
$$\min_{\mathbf{A}, \mathbf{b}} \log \det \mathbf{A}^{-1}$$
s.t. $\mathbf{A} \succ 0$

$$\|\mathbf{A}\mathbf{x}_i + \mathbf{b}\|_2 \le 1$$

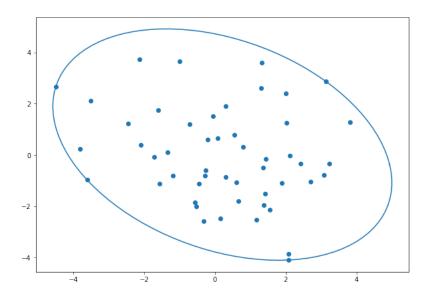
Эллипсоид Löwner-John

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

Пример построения экстремального эллипсоида



Пример построения экстремального эллипсоида



Запись через надграфик

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

$$\Rightarrow \quad \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\
f_0(\mathbf{x}) \le 0, \ i = 1, \dots, m$$

Запись через надграфик

$$\min_{\mathbf{x}} f_0(\mathbf{x}) \qquad \qquad \min_{\mathbf{x}, t} t$$
s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$

$$f_i(\mathbf{x}) \le 0, \ i = 1, \dots, m$$

$$f_0(\mathbf{x}) \le t$$

$$\min_{\mathbf{x}, t} t$$
s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}$

$$f_i(\mathbf{x}) \le 0, \ i = 1, \dots, m$$

Преобразования ограничений

- $Ax \le b \to Ax + y = b, y \ge 0$
- $\mathbf{x} \to \mathbf{x}_1 \mathbf{x}_2, \ \mathbf{x}_1 \ge 0, \ ; \mathbf{x}_2 \ge 0$
- Формирование блочных матриц

Запись через надграфик

$$\min_{\mathbf{x}} f_0(\mathbf{x}) \\ \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\ f_i(\mathbf{x}) \le 0, \ i = 1, \dots, m \end{cases} \implies \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b} \\ f_0(\mathbf{x}) \le t$$

Преобразования ограничений

- $Ax < b \rightarrow Ax + v = b, v > 0$
- $\mathbf{x} \to \mathbf{x}_1 \mathbf{x}_2, \ \mathbf{x}_1 > 0, \ ; \mathbf{x}_2 > 0$
- 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

- ▶ 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