# Методы оптимизации
# Лекция 3: Примеры задач выпуклой оптимизации


Александр Катруца

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


27 сентября 2021 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* $f : X \subset \mathbb{R}^n \to \mathbb{R}$ *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:*

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

# Основные результаты по выпуклым функциям

## 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* $\Leftrightarrow \mathrm{epi}\ f$ *выпуклое множество.*

# Основные результаты по выпуклым функциям

### 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*

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

# Основные результаты по выпуклым функциям

## Определение

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

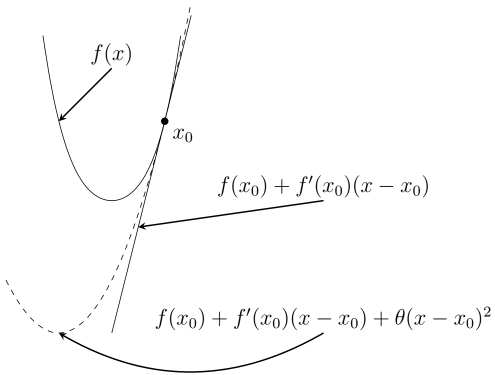
$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in X.$$

## 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

# Стандартная форма записи задачи выпуклой оптимизации

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

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

# Стандартная форма записи задачи выпуклой оптимизации

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

- $f_0$ – 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

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

# Линейное программирование (LP)

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

# Линейное программирование (LP)

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

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

# Линейное программирование (LP)

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

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

# Линейное программирование (LP)

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

- 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)

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

- 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)

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

- 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)

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

- 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 $b_1, \ldots, b_m$

# Линейное программирование (LP)

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

- 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 $b_1, \ldots, b_m$
- Известно, что в $j$-ом продукте содержится $a_{ij}$ $i$-го питательного вещества

# Линейное программирование (LP)

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

- 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 $b_1, \ldots, b_m$
- Известно, что в $j$-ом продукте содержится $a_{ij}$ $i$-го питательного вещества
- Необходимо определить количество каждого продукта

# Квадратичное программирование с квадратичными ограничениями (QCQP)

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

# Квадратичное программирование с квадратичными ограничениями (QCQP)

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

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

# Квадратичное программирование с квадратичными ограничениями (QCQP)

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

- 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)

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

- 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

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

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

# Линейные системы с неквадратными матрицами

- Линейная задача наименьших квадратов

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

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

- Поиск решения минимальной нормы

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

# Линейные системы с неквадратными матрицами

- Линейная задача наименьших квадратов

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

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

- Поиск решения минимальной нормы

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

  **Q**: Каков геометрический смысл у решения?

# Линейные системы с неквадратными матрицами

- Линейная задача наименьших квадратов

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

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

- Поиск решения минимальной нормы

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

  **Q**: Каков геометрический смысл у решения?

  **Q**: К какой задаче сводится похожая задача?

$$\begin{aligned} &\min_{\mathbf{x}} \|\mathbf{x}\|_\infty \\ \text{s.t. } &\mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

# 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 $\mathbf{\Sigma}$

# Задача о составлении оптимального портфеля

## Дано

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

## Классическая задача составления оптимального портфеля

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

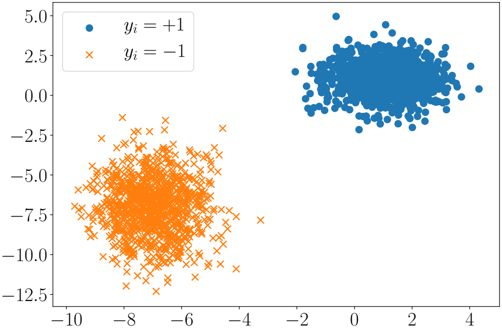
- Минимум риска при минимально допустимом доходе
- Существуют многочисленные вариации, которые не выводят задачу из класса QCQP или SOCP

# 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 $y_i < 0$

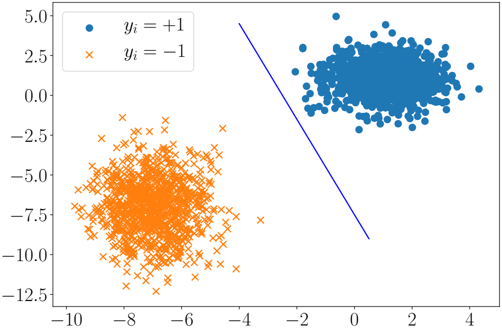
# Задача классификации

- 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 $y_i < 0$



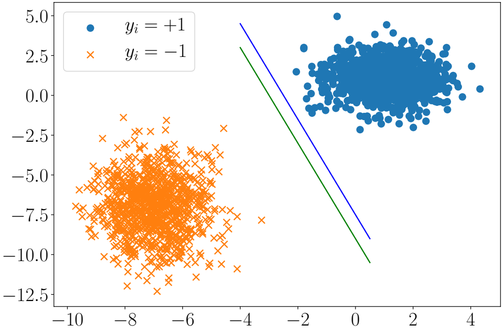
# Задача классификации

- 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 $y_i < 0$



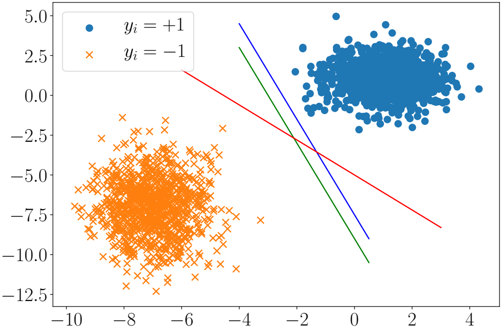
# Задача классификации

- 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 $y_i < 0$



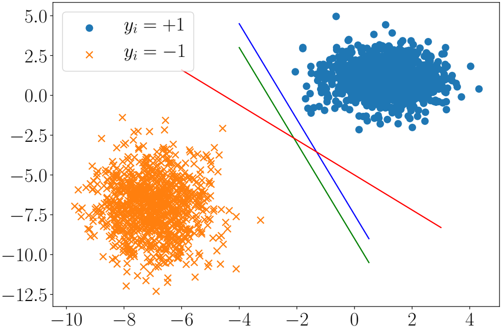
# Задача классификации

- 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

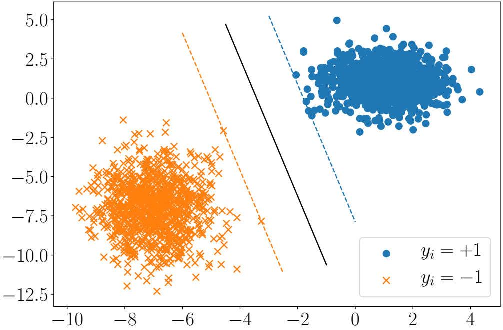
$$d = \frac{|c_1 - c_2|}{\|\mathbf{a}\|} = \frac{2}{\|\mathbf{a}\|_2}$$

## Финальная задача

$$\min_{\mathbf{a}, b} \frac{1}{2} \|\mathbf{a}\|_2^2$$

$$\text{s.t. } y_i(\mathbf{a}^\top \mathbf{x}_i + b) \geq 1, \; i = 1, \dots, m$$

# Оптимальная разделяющая гиперплоскость



# Оптимизация на конусе второго порядка (SOCP)

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

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 $\to$ SOCP

- Пусть $\mathbf{x}^\top \mathbf{P} \mathbf{x} + \mathbf{q}^\top \mathbf{x} + r \le 0$ 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 $\|\mathbf{L}^\top \mathbf{x} + \mathbf{L}^{-1}\tilde{\mathbf{q}}\|_2 \le \sqrt{\tilde{\mathbf{q}}^\top \mathbf{P}^{-1} \tilde{\mathbf{q}} - r}$

## Отношения между рассмотренными типами задач

LP $\subset$ QCQP $\subset$ SOCP

# Задача геометрического программирования (geometric programming)

## Определения

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

# Задача геометрического программирования (geometric programming)

## Определения

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

## Общий вид задачи

$$
\begin{aligned}
& \min f_0(\mathbf{x}) \\
\text{s.t. } & f_i(\mathbf{x}) \leq 1 \\
& h_j(\mathbf{x}) = 1,
\end{aligned}
$$

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 $\tilde{f}(\mathbf{x}) = e^{\mathbf{a}^\top \mathbf{y} + b}$

# Выпуклая форма задачи GP

- 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 $\tilde{f}(\mathbf{x}) = \sum_{k=1}^{K} e^{\mathbf{a}_k^\top \mathbf{y} + b_k}$

# Выпуклая форма задачи GP

- 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

# Выпуклая форма задачи GP

- 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

# Выпуклая форма задачи GP

- 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

$$
\begin{aligned}
&\min_{\mathbf{y}} \sum_{k=1}^{K} e^{\mathbf{a}_{k0}^\top \mathbf{y} + b_{k0}} \\
\text{s.t. } &\sum_{k=1}^{K} e^{\mathbf{a}_{ki}^\top \mathbf{y} + b_{ki}} \leq 1 \\
&e^{\mathbf{d}_j^\top \mathbf{y} + p_j} = 1
\end{aligned}
$$

# Выпуклая форма задачи GP

- 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

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

# Пример

- Пусть есть $n$ трансмиттеров и $n$ ресиверов

# Пример

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

# Пример

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

# Пример

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

# Пример

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

# Пример

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

## Задача минимизации мощности

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

# Пример

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

## Задача минимизации мощности

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

Больше примеров можно найти в этом туториале[1]

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

# Коническая форма задачи GP

- После равносильных преобразований задач достаточно описать неравенство

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

в конической форме

# Коническая форма задачи GP

- После равносильных преобразований задач достаточно описать неравенство

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

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

# Коническая форма задачи GP

- 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 $e^{z_k - t} \le u_k$

# Коническая форма задачи GP

- После равносильных преобразований задач достаточно описать неравенство

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

  в конической форме
- Эквивалентная запись $\sum_{k=1}^{m} e^{z_k - t} \leq 1$
- Введём переменную $e^{z_k - t} \leq u_k$
- Тогда $\sum_{k=1}^{m} u_k \leq 1$ — линейное неравенство

# Коническая форма задачи GP

- После равносильных преобразований задач достаточно описать неравенство

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

  в конической форме
- Эквивалентная запись $\sum_{k=1}^{m} e^{z_k - t} \leq 1$
- Введём переменную $e^{z_k - t} \leq u_k$
- Тогда $\sum_{k=1}^{m} u_k \leq 1$ — линейное неравенство

### Определение

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

# Коническая форма задачи GP

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

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

# 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

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

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

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

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

Стандартная форма

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

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

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

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

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

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

Стандартная форма

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

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

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

# LP и SOCP как задачи SDP

## LP

- $\mathbf{G} = 0$
- $\mathbf{F}_i$ 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 и SOCP как задачи SDP

## LP

- $\mathbf{G} = 0$
- $\mathbf{F}_i$ 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

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

# Немного про дополнение по Шуру

## Блочное исключение

- $\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix}$
- Выразим $\mathbf{y} = \mathbf{C}^{-1}(\mathbf{h} - \mathbf{B}^\top \mathbf{x})$
- Первое уравнение сводится к $\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{C}^{-1}(\mathbf{h} - \mathbf{B}^\top \mathbf{x}) = (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\top)\mathbf{x} + \mathbf{B}\mathbf{C}^{-1}\mathbf{h} = \mathbf{g}$
- Получили дополнение по Шуру для матрицы $\mathbf{C}$

# Немного про дополнение по Шуру

## Блочное исключение

- $\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix}$
- Выразим $\mathbf{y} = \mathbf{C}^{-1}(\mathbf{h} - \mathbf{B}^\top \mathbf{x})$
- Первое уравнение сводится к $\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{C}^{-1}(\mathbf{h} - \mathbf{B}^\top\mathbf{x}) = (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\top)\mathbf{x} + \mathbf{B}\mathbf{C}^{-1}\mathbf{h} = \mathbf{g}$
- Получили дополнение по Шуру для матрицы $\mathbf{C}$

## Доказательство факта с предыдущего слайда

- Пусть $\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix} \succeq 0$ тогда по определению $g(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} + 2\mathbf{y}^\top \mathbf{B}\mathbf{x} + \mathbf{y}^\top \mathbf{C}\mathbf{y} \geq 0$
- Значит функция $f(\mathbf{x}) = \min_{\mathbf{y}} g(\mathbf{x}, \mathbf{y}) \geq 0$
- Функция $g$ выпукла при каждом фиксированном $\mathbf{x}$, тогда $\mathbf{y}^* = -\mathbf{C}^{-1}\mathbf{B}\mathbf{x}$ и $f(\mathbf{x}) = \mathbf{x}^\top(\mathbf{A} - \mathbf{B}^\top\mathbf{C}^{-1}\mathbf{B})\mathbf{x}$
- Значит $\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\top \succeq 0$

# 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
$$
\begin{aligned}
&\max \frac{1}{2} \sum_{i<j} w_{ij}(1 - x_i x_j)\\
&\text{s.t. } x_i \in \{-1, +1\}
\end{aligned}
$$
- В матрично-векторном виде
$$
\begin{aligned}
&\min \frac{1}{2} \mathbf{x}^\top \mathbf{W} \mathbf{x}\\
&\text{s.t. } x_i \in \{-1, +1\}
\end{aligned}
$$
- Эквивалентный вид
$$
\begin{aligned}
&\min \frac{1}{2} \mathrm{trace}(\mathbf{W}\mathbf{X})\\
&\text{s.t. } \mathbf{X} \succeq 0, \; \mathrm{rank}(\mathbf{X}) = 1, \; \mathrm{diag}(\mathbf{X}) = \mathbf{1}
\end{aligned}
$$

# 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

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

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

- SDP релаксация

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

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

# 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

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

# 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
$$\{\mathbf{x} \mid \|\mathbf{x}\|_2^2 \leq 1\} \to \{\mathbf{u} \mid \|\mathbf{u}\|_2^2 \leq 1, \; \mathbf{u} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$$
- 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

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

- 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

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

- 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 $\mathbf{A} \succ 0$

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

# 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

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

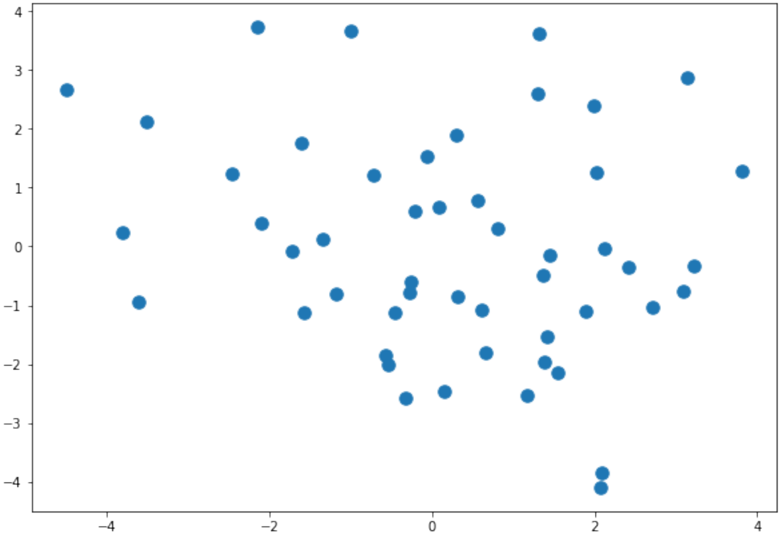
- Тогда площадь увеличивается в $\det(\mathbf{A}^{-1})$ раз.
- Детерминант не является выпуклой/вогнутой функцией
- $\log\det(\mathbf{A}^{-1}) = -\log\det(\mathbf{A})$ – выпуклая функция при $\mathbf{A} \succ 0$

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

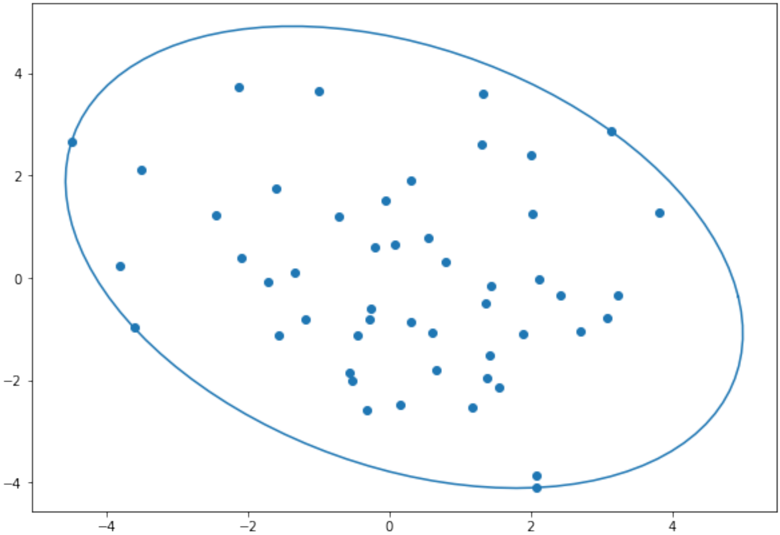
## Эллипсоид Лёвнера-Джона (Löwner-John)

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

# Пример построения экстремального эллипсоида



# Пример построения экстремального эллипсоида



# Равносильные преобразования задач

# Равносильные преобразования задач

## Запись через надграфик

$$
\begin{aligned}
&\min_{\mathbf{x}} f_0(\mathbf{x}) \\
\text{s.t. } &\mathbf{A}\mathbf{x} = \mathbf{b} \\
&f_i(\mathbf{x}) \le 0, \; i = 1, \dots, m
\end{aligned}
\quad \Longrightarrow \quad
\begin{aligned}
&\min_{\mathbf{x},\, t} t \\
\text{s.t. } &\mathbf{A}\mathbf{x} = \mathbf{b} \\
&f_i(\mathbf{x}) \le 0, \; i = 1, \dots, m \\
&f_0(\mathbf{x}) \le t
\end{aligned}
$$

# Равносильные преобразования задач

## Запись через надграфик

$$
\begin{aligned}
&\min_{\mathbf{x}} f_0(\mathbf{x}) \\
\text{s.t. } & \mathbf{A}\mathbf{x} = \mathbf{b} \\
& f_i(\mathbf{x}) \le 0, \; i = 1, \ldots, m
\end{aligned}
\quad \Longrightarrow \quad
\begin{aligned}
&\min_{\mathbf{x},\, t} t \\
\text{s.t. } & \mathbf{A}\mathbf{x} = \mathbf{b} \\
& f_i(\mathbf{x}) \le 0, \; i = 1, \ldots, m \\
& f_0(\mathbf{x}) \le t
\end{aligned}
$$

## Преобразования ограничений

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

# Равносильные преобразования задач

## Запись через надграфик

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

## Преобразования ограничений

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

## Перенос ограничений в целевую функцию

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

# Главное

- Стандартная форма задачи выпуклой оптимизации

# Главное

- Стандартная форма задачи выпуклой оптимизации
- Линейное программирование

# Главное

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

# Главное

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

# Главное

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

# Главное

- 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

# Главное

- 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