Методы оптимизации
Лекция 13: Полуопределённое программирование (SDP)

Александр Катруца

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

13 декабря 2019 г.

На прошлой лекции

- Линейное программирование

На прошлой лекции

- Линейное программирование
- Структура допустимого множества

На прошлой лекции

- Линейное программирование
- Структура допустимого множества
- Симплекс-метод

От LP к SDP

LP в стандартной форме

$$
\begin{aligned}
&\min_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle \\
\text{s.t. } &\mathbf{A}\mathbf{x} = \mathbf{b} \\
&\mathbf{x} \geq 0
\end{aligned}
$$

- Замена вектора на матрицу
- $\mathbb{R}^n_+ \to \mathbf{S}^n_+$
- Скалярные произведения между векторами $\to$ скалярные произведения между матрицами

Постановка задачи SDP: напоминание

- Стандартная форма

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

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

Постановка задачи SDP: напоминание

- Стандартная форма

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

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

- Двойственная 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 $\mathbf{F}_i \in \mathbf{S}^n$.

Геометрия допустимого множества

- В LP допустимое множество – многогранник
- В SDP ситуация иная:

  $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{G} + \sum\limits_{i=1}^{n} x_i \mathbf{F}_i \preceq 0\}$

Геометрия допустимого множества

- В LP допустимое множество – многогранник
- В SDP ситуация иная:

  $$\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{G} + \sum_{i=1}^{n} x_i \mathbf{F}_i \preceq 0\}$$

Q: как описать множество $\mathcal{X}$?

Отличия от LP

Двойственность

- Условие Слейтера существенно для сильной двойственности!
- Возможны **разные** конечные значения у прямой и двойственной задачи
- Минимум также может не достигаться

Отличия от LP

Двойственность

- Условие Слейтера существенно для сильной двойственности!
- Возможны **разные** конечные значения у прямой и двойственной задачи
- Минимум также может не достигаться

Методы решения

- Нет алгоритма, сходящегося за конечное число шагов
- Аналог симплекс-метода будет работать не за конечное число шагов
- Нет аналога базисного допустимого решения

Отличия от LP

Двойственность

- Условие Слейтера существенно для сильной двойственности!
- Возможны **разные** конечные значения у прямой и двойственной задачи
- Минимум также может не достигаться

Методы решения

- Нет алгоритма, сходящегося за конечное число шагов
- Аналог симплекс-метода будет работать не за конечное число шагов
- Нет аналога базисного допустимого решения

Упражнение

Покажите, что LP $\subset$ SOCP $\subset$ SDP

Приложения

- Минимизация спектрального радиуса

Приложения

- Минимизация спектрального радиуса
  - Пусть $\mathbf{A}(\mathbf{x}) = x_1\mathbf{A}_1 + \ldots + x_n\mathbf{A}_n$, $\mathbf{A}_i \in \mathbf{S}^n$

Приложения

- Минимизация спектрального радиуса
  - Пусть $\mathbf{A}(\mathbf{x}) = x_1\mathbf{A}_1 + \ldots + x_n\mathbf{A}_n$, $\mathbf{A}_i \in \mathbf{S}^n$
  - Задача
$$\min_{\mathbf{x}} \lambda_{\max}(\mathbf{A}(\mathbf{x}))$$

Приложения

- Минимизация спектрального радиуса
  - Пусть $\mathbf{A}(\mathbf{x}) = x_1\mathbf{A}_1 + \ldots + x_n\mathbf{A}_n$, $\mathbf{A}_i \in \mathbf{S}^n$
  - Задача
$$\min_{\mathbf{x}} \lambda_{\max}(\mathbf{A}(\mathbf{x}))$$
  - SDP форма
$$\begin{aligned} &\min_{\mathbf{x},t} t \\ \text{s.t. } &\mathbf{A}(\mathbf{x}) - t\mathbf{I} \preceq 0 \end{aligned}$$

Приложения

- Минимизация спектрального радиуса
  - Пусть $\mathbf{A}(\mathbf{x}) = x_1\mathbf{A}_1 + \ldots + x_n\mathbf{A}_n$, $\mathbf{A}_i \in \mathbf{S}^n$
  - Задача
$$\min_{\mathbf{x}} \lambda_{\max}(\mathbf{A}(\mathbf{x}))$$
  - SDP форма
$$\begin{aligned} &\min_{\mathbf{x},t} t \\ \text{s.t. } &\mathbf{A}(\mathbf{x}) - t\mathbf{I} \preceq 0 \end{aligned}$$

- Задача о минимальном по площади эллипсоиде, покрывающем набор точек $\mathbf{x}_i$
$$\begin{aligned} &\min_{\mathbf{A},\mathbf{b}} -\log\det(\mathbf{A}) \\ \text{s.t. } &\|\mathbf{A}\mathbf{x}_i + \mathbf{b}\|_2 \leq 1 \\ &\mathbf{A} \succeq 0 \end{aligned}$$

Выпуклые релаксации невыпуклых задач QP

- Исходная невыпуклая задача QP

$$
\begin{aligned}
\min \mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x}\\
\text{s.t. } \mathbf{x}^\top \mathbf{A}_i \mathbf{x} + \mathbf{b}_i^\top \mathbf{x} \leq 0
\end{aligned}
$$

Выпуклые релаксации невыпуклых задач QP

- Исходная невыпуклая задача QP

$$
\begin{aligned}
\min\ & \mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x} \\
\text{s.t. }\ & \mathbf{x}^\top \mathbf{A}_i \mathbf{x} + \mathbf{b}_i^\top \mathbf{x} \leq 0
\end{aligned}
$$

- Равносильное преобразование

$$
\begin{aligned}
\min\ & \operatorname{trace}(\mathbf{A}\mathbf{X}) + \mathbf{b}^\top \mathbf{x} \\
\text{s.t. }\ & \operatorname{trace}(\mathbf{A}_i \mathbf{X}) + \mathbf{b}_i^\top \mathbf{x} \leq 0 \\
& \mathbf{X} = \mathbf{x}\mathbf{x}^\top
\end{aligned}
$$

Выпуклые релаксации невыпуклых задач QP

- Исходная невыпуклая задача QP

$$\begin{aligned} &\min \mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x} \\ \text{s.t. } &\mathbf{x}^\top \mathbf{A}_i \mathbf{x} + \mathbf{b}_i^\top \mathbf{x} \leq 0 \end{aligned}$$

- Равносильное преобразование

$$\begin{aligned} &\min \operatorname{trace}(\mathbf{A}\mathbf{X}) + \mathbf{b}^\top \mathbf{x} \\ \text{s.t. } &\operatorname{trace}(\mathbf{A}_i \mathbf{X}) + \mathbf{b}_i^\top \mathbf{x} \leq 0 \\ &\mathbf{X} = \mathbf{x}\mathbf{x}^\top \end{aligned}$$

- Релаксация невыпуклого ограничения на ранг $\mathbf{X}$

$$\begin{aligned} &\min_{\mathbf{X},\mathbf{x}} \operatorname{trace}(\mathbf{A}\mathbf{X}) + \mathbf{b}^\top \mathbf{x} \\ \text{s.t. } &\operatorname{trace}(\mathbf{A}_i \mathbf{X}) + \mathbf{b}_i^\top \mathbf{x} \leq 0 \\ &\mathbf{X} - \mathbf{x}\mathbf{x}^\top \succeq 0 \end{aligned}$$

Выпуклые релаксации невыпуклых задач QP

- Исходная невыпуклая задача QP

$$
\begin{aligned}
&\min \mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x} \\
\text{s.t. } &\mathbf{x}^\top \mathbf{A}_i \mathbf{x} + \mathbf{b}_i^\top \mathbf{x} \leq 0
\end{aligned}
$$

- Равносильное преобразование

$$
\begin{aligned}
&\min \operatorname{trace}(\mathbf{A}\mathbf{X}) + \mathbf{b}^\top \mathbf{x} \\
\text{s.t. } &\operatorname{trace}(\mathbf{A}_i \mathbf{X}) + \mathbf{b}_i^\top \mathbf{x} \leq 0 \\
&\mathbf{X} = \mathbf{x}\mathbf{x}^\top
\end{aligned}
$$

- Релаксация невыпуклого ограничения на ранг $\mathbf{X}$

$$
\begin{aligned}
&\min_{\mathbf{X}, \mathbf{x}} \operatorname{trace}(\mathbf{A}\mathbf{X}) + \mathbf{b}^\top \mathbf{x} \\
\text{s.t. } &\operatorname{trace}(\mathbf{A}_i \mathbf{X}) + \mathbf{b}_i^\top \mathbf{x} \leq 0 \\
&\mathbf{X} - \mathbf{x}\mathbf{x}^\top \succeq 0
\end{aligned}
$$

- Последнее ограничение равносильно $\begin{bmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{x}^\top & 1 \end{bmatrix} \succeq 0$

Релаксация Лагранжа

- К невыпуклой задаче можно построить **выпуклую** двойственную
- Двойственная даст оценку снизу на оптимальное значение прямой задачи
- Оказывается, что этот подход даст точно такую же оценку

MAXCUT

- Дан неориентированный граф $G = (V, E)$ с матрицей весов $\mathbf{W} \in \mathbf{S}^n$ и $\mathbf{W} \geq 0$

MAXCUT

- Дан неориентированный граф $G = (V, E)$ с матрицей весов $\mathbf{W} \in \mathbf{S}^n$ и $\mathbf{W} \geq 0$
- Найти такое подмножество вершин $S \subset V$, что сумма весов рёбер из вершин, лежащих в S, в вершины из $V \setminus S$ **максимальна**

MAXCUT

- Дан неориентированный граф $G = (V, E)$ с матрицей весов $\mathbf{W} \in \mathbf{S}^n$ и $\mathbf{W} \geq 0$
- Найти такое подмножество вершин $S \subset V$, что сумма весов рёбер из вершин, лежащих в S, в вершины из $V \setminus S$ **максимальна**

MAXCUT

- Дан неориентированный граф $G = (V, E)$ с матрицей весов $\mathbf{W} \in \mathbf{S}^n$ и $\mathbf{W} \geq 0$
- Найти такое подмножество вершин $S \subset V$, что сумма весов рёбер из вершин, лежащих в S, в вершины из $V \setminus S$ **максимальна**

Q: а как решается задача минимизации этой же величины?

MAXCUT

- Дан неориентированный граф $G = (V, E)$ с матрицей весов $\mathbf{W} \in \mathbf{S}^n$ и $\mathbf{W} \geq 0$
- Найти такое подмножество вершин $S \subset V$, что сумма весов рёбер из вершин, лежащих в S, в вершины из $V \setminus S$ **максимальна**

Q: а как решается задача минимизации этой же величины?

Формализация задачи

$$\max_{\mathbf{x}} \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(1 - x_i x_j)$$

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

Обозначим оптимальное значение целевой функции c^*

SDP релаксация

- Введём матрицу $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$

SDP релаксация

- Введём матрицу $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$
- Тогда

$$
\begin{aligned}
&\max_{\mathbf{x}} \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(1 - x_{ij}) \\
\text{s.t. } & x_i \in \{+1, \ -1\} \\
& \mathbf{X} = \mathbf{x}\mathbf{x}^\top
\end{aligned}
$$

SDP релаксация

- Введём матрицу $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$
- Тогда

$$
\begin{aligned}
\max_{\mathbf{x}} \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(1 - x_{ij}) \\
\text{s.t. } x_i \in \{+1,\ -1\} \\
\mathbf{X} = \mathbf{x}\mathbf{x}^\top
\end{aligned}
$$

- Первое ограничение переписывается в виде $\mathrm{diag}(\mathbf{X}) = \mathbf{1}$

SDP релаксация

- Введём матрицу $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$
- Тогда

$$
\begin{aligned}
&\max_{\mathbf{x}} \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(1 - x_{ij}) \\
\text{s.t. } & x_i \in \{+1, \ -1\} \\
& \mathbf{X} = \mathbf{x}\mathbf{x}^\top
\end{aligned}
$$

- Первое ограничение переписывается в виде $\mathrm{diag}(\mathbf{X}) = \mathbf{1}$
- Ограничение на ранг $\mathbf{X}$ заменяют на $\mathbf{X} \in \mathbf{S}_+^n$

SDP релаксация

- Введём матрицу $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$
- Тогда

$$\begin{aligned} & \max_{\mathbf{x}} \frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n w_{ij}(1 - x_{ij}) \\ \text{s.t. } & x_i \in \{+1, \; -1\} \\ & \mathbf{X} = \mathbf{x}\mathbf{x}^\top \end{aligned}$$

- Первое ограничение переписывается в виде $\mathrm{diag}(\mathbf{X}) = \mathbf{1}$
- Ограничение на ранг $\mathbf{X}$ заменяют на $\mathbf{X} \in \mathbf{S}^n_+$
- Оптимальное значение целевой функции обозначим p^*

SDP релаксация

- Введём матрицу $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$
- Тогда

$$
\begin{aligned}
&\max_{\mathbf{x}} \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(1 - x_{ij}) \\
\text{s.t. } & x_i \in \{+1, \ -1\} \\
& \mathbf{X} = \mathbf{x}\mathbf{x}^\top
\end{aligned}
$$

- Первое ограничение переписывается в виде $\mathrm{diag}(\mathbf{X}) = \mathbf{1}$
- Ограничение на ранг $\mathbf{X}$ заменяют на $\mathbf{X} \in \mathbf{S}_+^n$
- Оптимальное значение целевой функции обозначим p^*
- Поскольку допустимое множество существенно расширилось, то $p^* \geq c^*$

Как восстановить решение?

- В результате решения релаксированной задачи получена матрица $\mathbf{X}$
  Q: как восстановить разрез?

Как восстановить решение?

- В результате решения релаксированной задачи получена матрица $\mathbf{X}$
  Q: как восстановить разрез?

Алгоритм Goemans-Williamson'a

1. Сгенерировать случайный вектор $\mathbf{v}$ на единичной сфере
2. $S = \{i \mid \langle \mathbf{v}, \mathbf{u}_i \rangle \geq 0\}$, где $\mathbf{u}_i$ формируют матрицу $\mathbf{U} : \mathbf{X} = \mathbf{U}^\top \mathbf{U}$

Как восстановить решение?

- В результате решения релаксированной задачи получена матрица $\mathbf{X}$
  Q: как восстановить разрез?

Алгоритм Goemans-Williamson'a

1. Сгенерировать случайный вектор $\mathbf{v}$ на единичной сфере
2. $S = \{i \mid \langle \mathbf{v}, \mathbf{u}_i \rangle \geq 0\}$, где $\mathbf{u}_i$ формируют матрицу $\mathbf{U} : \mathbf{X} = \mathbf{U}^\top \mathbf{U}$

Оценка точности

Алгоритм Goemans-Williamson'a в среднем даёт решение r^* исходной задачи с точностью не хуже чем

$$c^* \geq r^* \geq 0.878p^*$$

Авторы получили премию Фалкерсона в 2000 г. за этот алгоритм

Доказательство

1. Поскольку алгоритм предлагает каждый вектор $\mathbf{u}_i$ отобразить в ± 1 случайно, то можем вычислить матожидание величины разреза C

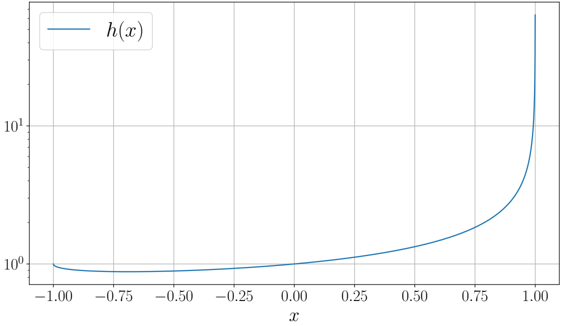
$$
\begin{aligned}
\mathbb{E}_{\mathbf{v}}(C) &= \mathbb{E}_{\mathbf{v}} \frac{1}{2} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} [\mathrm{sign}(\mathbf{v}^\top \mathbf{u}_i) \neq \mathrm{sign}(\mathbf{v}^\top \mathbf{u}_j)] \right) \\
&= \frac{1}{2} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \mathbb{P}(\mathrm{sign}(\mathbf{v}^\top \mathbf{u}_i) \neq \mathrm{sign}(\mathbf{v}^\top \mathbf{u}_j)) \right)
\end{aligned}
$$

2. Вероятность того, что знак $\mathbf{v}^\top \mathbf{u}_i$ отличается от знака $\mathbf{v}^\top \mathbf{u}_j$, равна $\frac{\angle(\mathbf{u}_i, \mathbf{u}_j)}{\pi} = \frac{\arccos(\mathbf{u}_i^\top \mathbf{u}_j)}{\pi}$
3. В итоге $\mathbb{E}_{\mathbf{v}}(C) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \frac{\arccos(\mathbf{u}_i^\top \mathbf{u}_j)}{\pi}$

4. Далее сведём это выражение к известному

$$\mathbb{E}_{\mathbf{v}}(C) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\frac{2\arccos(\mathbf{u}_i^\top \mathbf{u}_j)}{\pi(1-\mathbf{u}_i^\top \mathbf{u}_j)}\frac{1-\mathbf{u}_i^\top \mathbf{u}_j}{2}$$

5. Найдём минимум функции $h(x) = \frac{2\arccos(x)}{\pi(1-x)}$ для $x \in [-1, 1)$



6. Оказывается, что $h(x^*) \approx 0.8785 = \alpha_{GW}$

7. Тогда

$$\mathbb{E}_{\mathbf{v}}(C) \geq \alpha_{GW} \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(1 - \mathbf{u}_i^\top \mathbf{u}_j) = \alpha_{GW} p^* \geq \alpha_{GW} c^*$$

8. Так как метод Goemans-Williamson'a выдаёт некоторый разрез, то для матожидания выполнено

$$c^* \geq \mathbb{E}_{\mathbf{v}}(C) \geq \alpha_{GW} p^* \geq \alpha_{GW} c^*$$

9. В итоге

$$c^* \leq p^* \leq \frac{1}{\alpha_{GW}} c^* \approx 1.1382 c^*$$

Другая интерпретация задачи MAXCUT

$$
\begin{aligned}
\max_{x_i=\pm 1} \frac{1}{4} \sum_{i,j=1}^{n} w_{ij}(1 - x_i x_j) &= \max_{x_i=\pm 1} \frac{1}{4} \sum_{i,j=1}^{n} w_{ij} \left(\frac{x_i^2 + x_j^2}{2} - x_i x_j \right) \\
&= \max_{x_i=\pm 1} \frac{1}{4} \left(-\sum_{i,j=1}^{n} w_{ij} x_i x_j + \frac{1}{2} \sum_{i=1}^{n} \left[\sum_{j=1}^{n} w_{ij} \right] x_i^2 + \frac{1}{2} \sum_{j=1}^{n} \left[\sum_{i=1}^{n} w_{ij} \right] x_j^2 \right) \\
&= \max_{x_i=\pm 1} \frac{1}{4} \left(-\sum_{i,j=1}^{n} w_{ij} x_i x_j + \frac{1}{2} \sum_{i=1}^{n} \deg(i) x_i^2 + \frac{1}{2} \sum_{j=1}^{n} \deg(j) x_j^2 \right) \\
&= \max_{x_i=\pm 1} \frac{1}{4} \left(\sum_{i=1}^{n} \deg(i) x_i^2 - \sum_{i,j=1}^{n} w_{ij} x_i x_j \right) = \max_{x_i=\pm 1} \frac{1}{4} \mathbf{x}^\top \mathbf{L} \mathbf{x},
\end{aligned}
$$

где $\mathbf{L}$ — лапласиан графа, то есть $\mathbf{L} = \mathbf{D} - \mathbf{W}$, где $\mathbf{D}$ — диагональная матрица, на диагонали стоят степени вершин.

Можно ли улучшить оценку?

- Известно, что получение оценки в $\frac{16}{17}$ от оптимального значения уже является NP-сложной задачей!
- Поиск алгоритма субэкспоненциальной сложности, который бы находил оценку лучше, является открытой проблемой!

Резюме

- Постановки задач полуопределённой оптимизации
- Приложения
- Методика релаксации невыпуклых задач
- Алгоритм Goemans-Williamson'a