Невыпуклые гладкие релаксации задач квадратичной комбинаторной оптимизации

Роман Погодин, Михаил Кречетов

МФТИ, Skoltech

Постановка задачи

Задачи вида

$$\max \; x^\top A x, \\ \text{s.t. } x \in \{-1, 1\}^n.$$

SDP релаксация: $xx^\top = X$, но убираем $\operatorname{rank} X = 1$:

$$\max \; \operatorname{Tr}(AX), \\ \text{s.t. } \operatorname{diag} X = 1_n, \\ X^\top = X, \; X \succeq 0.$$

Ответ x ищется рандомизированно.Для $A \succeq 0$ в среднем дает разрез не хуже $2/\pi$ от максимального

Постановка задачи

Один из примеров – поиск максимального разреза

$$\max \frac{1}{2}\sum_{i<j} w_{ij}(1 - x_i x_j),$$
$$\text{s.t. } x_i \in \{-1, 1\},\ i = 1, \dots n.$$

Эквивалентно в терминах Лапласиана:

$$\max \frac{1}{4} x^\top L x,$$
$$\text{s.t. } x \in \{-1, 1\}^n.$$

Здесь гарантии лучше – средний разрез не хуже ≈ 0.878 от максимального

Ранг решений

Мы хотим решения маленького ранга. Почему?

- Меньше вариантов при рандомизации
- При $\operatorname{rank} X = 1$ точное решение

Что мы знаем про ранг:

Theorem (Pataki, 1998)

Разрешимая SDP-задача имеет решение X, такое что $\operatorname{rank} X = r,\ r(r+1) \leq 2m$, где m – число ограничений.

В нашем случае $m = n$

Поиск малоранговых решений

Что можно делать:

- Представить $X = VV^\top$, где V размера $n \times r$, решать такую задачу [Burer, Monteiro, 2003]
- Разрешить выполнение ограничений только приближенно [Ye, Zhang, 2007]
- Минимизировать ранг полученного решения

Последняя постановка для Max Cut:

$$
\begin{aligned}
\min\ & \mathrm{rank}\, X \\
\text{s.t. } & \mathrm{diag}\, X = 1_n, \\
& X^\top = X,\ X \succeq 0, \\
& \mathrm{Tr}\,(LX) = 4 \cdot SDP.
\end{aligned}
$$

Идея – релаксировать разрывную функцию ранга и ограничение на функционал

Релаксации ранга. Сингулярные значения

Приблизим отношением сингулярных значений:

$$\operatorname{rank} X \approx \sum \frac{\sigma_i^2(X)}{\sigma_i^2(X) + \varepsilon} = \operatorname{Tr} X^\top (XX^\top + \varepsilon I)^{-1} X.$$

Улучшение

$$\operatorname{rank} X \approx \Phi(X, \varepsilon) = (1 + \varepsilon^q) \operatorname{Tr} \left(X^\top (XX^\top + \varepsilon I)^{-1} X \right)$$

гарантирует

$$|\operatorname{rank} X - \Phi(X, \varepsilon)| \le \varepsilon^q \max \left\{ \operatorname{rank} X, \sum_{i=1}^{\operatorname{rank} X} \left| \frac{\varepsilon^{1-q}}{\sigma_i^2(X)} - 1 \right| \right\}.$$

Релаксации ранга. Шаттеновские нормы

При $p < 1$ такая квази-норма:

$$\|X\|_{S_p}^p = \sum_{i \geq 1} \sigma_i^p = \mathrm{Tr} \left(X^\top X \right)^{\frac{p}{2}}.$$

Предел по p даст функцию ранга. Сгладим:

$$\|X\|_{S_{p,\varepsilon}}^p = \sum_{i \geq 1} \left(\sigma_i^2 + \varepsilon \right)^{\frac{p}{2}} = \mathrm{Tr} \left(X^\top X + \varepsilon I \right)^{\frac{p}{2}} \approx \|X\|_{S_p}^p.$$

Релаксации ранга. Log-Det

Используем

$$\mathrm{rank}\, X \approx \log\det\left(X + \varepsilon I\right).$$

Из ряда Тейлора

$$\log\det\left(X + \varepsilon I\right) \approx \log\det\left(X_n + \varepsilon I\right) + \mathrm{Tr}\,\left((X_n + \varepsilon I)^{-1}(X - X_n)\right).$$

Это приведет к следующей итеративной процедуре:

$$X_{n+1} = \operatorname*{argmin}_{X \in K} \mathrm{Tr}\,\left((X_n + \varepsilon I)^{-1} X\right),$$

Методы решения. IRLS

Для Шаттеновских норм введем вес

$$A(X) = \left(X^\top X + \varepsilon I\right)^{\frac{p-2}{2}}.$$

Тогда такие итерации сойдутся к ККТ-точке:

$$\begin{aligned}
\min\ & \frac{p}{2}\mathrm{Tr}\left(AX^T X\right) \\
\text{s.t. } & \operatorname{diag} X = 1_n, \\
& X^\top = X,\ X \succeq 0, \\
& 4 \cdot W^* \leq \mathrm{Tr}\,(LX) \leq 4 \cdot SDP.
\end{aligned}$$

Методы решения. Градиентный спуск

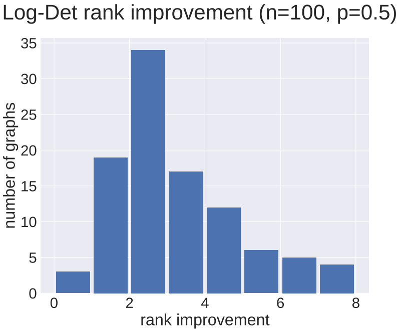
Для Шаттеновских норм и сингулярных значений репараметризуем задачу:

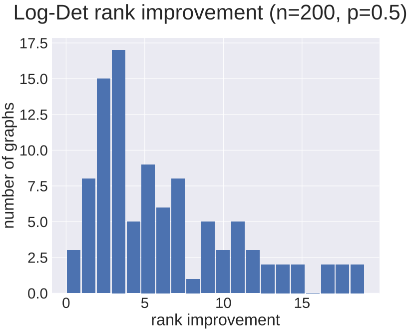
$$X = \begin{pmatrix} 1 & x_1 & \dots & x_{n-1} \\ x_1 & 1 & x_n & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

- Нужны проекции только на PSD-конус и нижнюю грань функционала
- Для Шаттена при шаге $\alpha \leq \frac{1}{2p}\varepsilon^{(2-p)/p}$ не выходит из PSD-конуса
- Для сингулярных значений при шаге $\alpha < \frac{\varepsilon}{4(1+\varepsilon^q)}$ не выходит из PSD-конуса
- Быстрее IRLS, применимо для больших задач

Результаты. Улучшения ранга

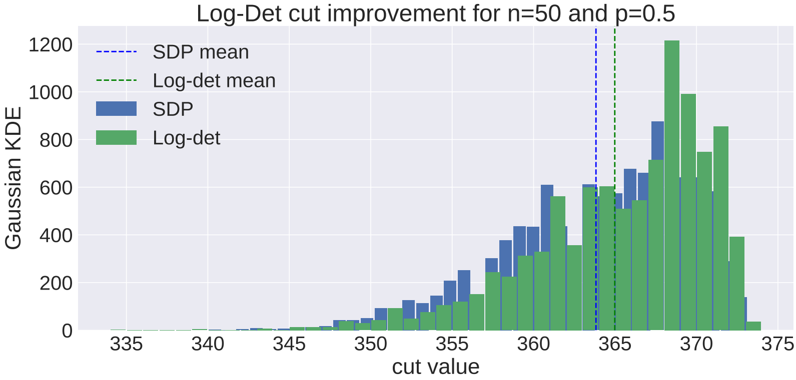
Шаттеновские нормы и сингулярные значения улучшают ранг, но значительно реже. Log-det:





Результаты. Улучшения среднего разреза

Все методы сохраняют среднее значение разреза ± 1,
но уменьшают дисперсию (хотя $\mathrm{Tr}\,(LX)$ не равен оптимальному)



Выводы

В работе проанализирована минимизация ранга для получения малоранговых решений SDP-релаксации.
Показано:

- сходимость IRLS
- гарантии на шаг градиентных спусков, что дает применимость для больших размерностей

Эксперименты показали:

- Снижение ранга улучшает распределение разрезов
- Log-Det лучше других понижает ранг задачи
- Релаксация значения функционала улучшает ранг решения, но не ухудшает средний разрез