Оценки неотрицательных и полуопределенных рангов матриц

Рязанов Андрей Владимирович

Московский физико-технический институт
Skolkovo Institute of Science and Technology

План

Сложность расширения

- Подъёмы многогранников
- Сложность расширения
- Неотрицательные и положительно полуопределённые (PSD-) сложности

Теорема о разложении

- Матричные разложения
- Неотрицательный и PSD- ранги
- Теорема о разложении

Оценки на PSD-ранг

- Существующие ограничивающие функционалы
- Исследование ограничивающих функционалов

Подъёмы многогранников

Линейная оптимизация над выпуклыми многогранниками играет центральную роль в оптимизации. Для этой задачи существуют полиномиальные алгоритмы решения, однако все они зависят от количества ограничений (количества граней многогранника).

"Сложный" многогранник: гипероктаэдр (2^n граней):

$$P = \{x \in \mathbb{R}^n | \pm x_1 \pm x_2 \pm \dots \pm x_n \le 1\}.$$

Подъёмы многогранников

Подход

Представить "сложный" многогранник в виде проекции "простого": гипероктаэдр (2^n граней):

$$P = \{x \ \in \mathbb{R}^n | \ \pm x_1 \pm x_2 \pm \cdots \pm x_n \leq 1\}.$$

Но $P = \pi_x(Q)$, где у Q всего $2n+1$ граней (Q – *подъём* P):

$$Q = \{(x,y) \in \mathbb{R}^{2n} | -y_i \leq x_i \leq y_i, \ \sum_{i=1}^{n} y_i = 1\}.$$

$P = \pi(Q)$, π – линейное отображение. Тогда:
$\min_{x \in P} c(x) = \min_{x \in \pi(Q)} c(x) = \min_{y \in Q} c(\pi(y)) = \min_{y \in Q} f(y)$.

Сложность расширения

Расширение P – представление $P = \pi(K \cap L)$, K – конус, на сечениях которого оптимизация эффективна, L – аффинное подпространство.

- $K = \mathbb{R}^k_+ = \{x \in \mathbb{R}^n | x_i \geq 0\}$ – неотрицательный ортант;
- $K = \mathbf{S}^r_+$ – конус положительно полуопределённых матриц размера $r \times r$ (SDP конус).

Сложности расширений:

- неотрицательная: минимальное k такое, что $P = \pi(\mathbb{R}^k_+ \cap L)$;
- PSD: минимальное r такое, что $P = \pi(\mathbf{S}^r_+ \cap L)$.

Подъёмы многогранников

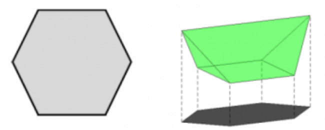
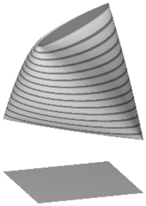


Рис.: Неотрицательный подъём



$$Q = \left\{ x \in \mathbb{R}^3 : \begin{pmatrix} 1 & x_1 & x_2 \\ x_1 & 1 & x_3 \\ x_2 & x_3 & 1 \end{pmatrix} \in \mathcal{S}^3_+ \right\}$$

Рис.: PSD-подъём

Оценки сложности расширений

В предположении $P \neq NP$ сложности расширений для многогранников NP-сложных комбинаторных задач (Max Cut, TSP, Max Clique, Knapsack, и т. д.) должны быть сверхполиномиальными по размерности. Однако на данный момент доказано это только для нескольких частных случаев:

Неотрицательная сложность расширения

- Symmetric TSP (Yannakakis, 1998)
- Factorization Theorem (Gouveia, Parrilo, Thomas, 2012)
- TSP, CUT, CORR (Fiorini et. al. 2012)

PSD сложность расширения

- TSP and CUT, Stable Set (Lee, Raghavendra, Steurer, 2014)

Теорема о разложении

Матрица невязок

$$P = \{x \in \mathbb{R}^n \mid \langle x, a_j \rangle \leq b_j,\ j \in \overline{1, m}\}$$

Матрица невязок $S_P \in \mathbb{R}_+^{n \times m}$:

$$S_P(i, j) = b_j - \langle v_i, a_j \rangle,$$

v_i – i-тая вершина P.

$$\operatorname{rank} S_P = \dim P + 1.$$

Теорема о разложении

Матричные разложения

Пусть $A \in \mathbb{R}^{n\times m}$ – неотрицательная матрица.

Неотрицательное матричное разложение: $A = BC$, где $B \in \mathbb{R}_+^{m\times k}$, $C \in \mathbb{R}_+^{k\times n}$. Эквивалентно: два набора векторов $\{b_i\}_{i=1}^m$, $\{c_j\}_{j=1}^n$, $b_i, c_j \in \mathbb{R}_+^k$ таких, что $A(i,j) = \langle b_i, c_j\rangle$.

PSD-разложение – два набора SDP-матриц $\{B_i\}_{i=1}^m$, $\{C_j\}_{j=1}^n$, $B_i, C_j \in \mathbf{S}_+^r$ таких, что $A(i,j) = \langle B_i, C_j\rangle$.

Неотрицательный и PSD- ранги

- *Неотрицательный ранг* $\operatorname{rank}_+ A$ – наименьшее $k \in \mathbb{N}$ такое, что существует неотрицательное разложение;
- *PSD-ранг* $\operatorname{rank}_{\mathrm{psd}} A$ – наименьшее $r \in \mathbb{N}$ такое, что существует PSD-разложение.

P – многогранник, S_P – матрица невязок.

Теорема о разложении

Неотрицательная сложность расширения P равняется $\mathrm{rank}_+ S_P$. (Yannakakis, 1998)

PSD-сложность расширения P равняется $\mathrm{rank}_{\mathrm{psd}} S_P$. (Gouveia, Parrilo, Thomas, 2012), (Fiorini et. al. 2012)

Используется несколько подходов для нахождения оценок на неотрицательный и PSD-ранги:

- Теория информации (взаимная и общая информация)
- Коммуникационная сложность (протоколы недетерминированной и квантовых коммуникаций)

Оценки

Известная оценка через взаимную информацию:
M – матрица совместного распределения A и B. Тогда

$$B_1(M) = 2^{I(A:B)},$$

$I(A:B) = H(A) - H(A|B)$.

Оценки на PSD-ранг

Оценки

(Lee, T., Wei, Z., de Wolf, R.,2016):
M – левая стохастическая матрица. Тогда

$$B_2(M) = \max_{\{q_i\}_{i=1}^m} \frac{1}{\sum_{i,j=1}^m q_i q_j F(M_i, M_j)^2}$$

$$B_3(M) = \sum_{i=1}^n \max_j M(i,j)$$

$$B_4(M) = \sum_{i=1}^n \max_{\{q_j^{(i)}\}_{j=1}^m} \frac{\sum_{k=1}^m q_k^{(i)} M(i,k)}{\sqrt{\sum_{s,t=1}^m q_s^{(i)} q_t^{(i)} F(M_s, M_t)^2}},$$

$F(p,q) = \sum_{i=1}^n \sqrt{p_i q_i}$ – согласованность (fidelity).

Оценки на PSD-ранг

$$\boxed{\operatorname{rank}_{\mathrm{psd}} M \geq B_i(M)}$$

Ключевой вопрос: *могут ли эти оценки на* $\operatorname{rank}_{\mathrm{psd}}$ *быть сверхполиномиальными по размерности?*

Оценки на PSD-ранг

$$\boxed{\operatorname{rank}_{\mathrm{psd}} M \geq B_i(M)}$$

Ключевой вопрос: *могут ли эти оценки на* $\operatorname{rank}_{\mathrm{psd}}$ *быть сверхполиномиальными по размерности?*

Ответ: *нет, не могут!*

Оценки на B_i

Полученные результаты

Пусть $M \in \mathbb{R}^{n\times m}$, $\operatorname{rank} M = r$. Тогда $B_i(M) \leq poly(r, \ln m)$:

$$B_1(M) \leq r$$
$$B_2(M) \leq (\ln m + 1)^2 r^2$$
$$B_3(M) \leq r$$
$$B_4(M) \leq (\ln m + 1) r^2$$

Подход

Преобразование зануления строк

$M \in \mathbb{R}_+^{n\times m}$, $\operatorname{rank} M = r < n$. Пусть первые $r+1$ строк $\overline{m_1}, \overline{m_2}\dots, \overline{m_{r+1}}$ ненулевые. Они линейно зависимы $\Rightarrow \exists \{\alpha_i\}_{i=1}^{r+1} : \sum_{i=1}^{r+1} \alpha_i \overline{m_i} = \overline{0}$.

Матрица M^ε:

- для $1 \le i \le (r+1)$: $\quad \overline{m_i^\varepsilon} = \overline{m_i}(1+\varepsilon\alpha_i)$
- для $i > (r+1)$: $\quad \overline{m_i^\varepsilon} = \overline{m_i}$

M^ε – *ε-преобразование* матрицы M. Без фиксации ε такое преобразование называется *преобразованием зануления строк*.

Подход

Преобразование зануления строк

$$\Delta_\alpha \triangleq \left[-\frac{1}{\max_i \alpha_i},\ -\frac{1}{\min_i \alpha_i}\right] = [\beta, \gamma], \quad \beta < 0, \gamma > 0$$

$$(1 + \varepsilon\alpha_i) \geq 0 \quad \forall i \in \overline{1, (r+1)} \qquad \Leftrightarrow \qquad \varepsilon \in \Delta_\alpha.$$

Свойства

1. Когда ε равняется одному из концов отрезка Δ_α, как минимум один из коэффициентов $(1 + \varepsilon\alpha_i)$ равняется нулю.
2. При преобразовании зануления строк сумма элементов по столбцам не меняется.
3. $B_i(M)$ линейно по ε:

$$B_i(M^\varepsilon) = B_i(M) + \varepsilon \cdot \Delta.$$

Заключение

- Изучена связь между сложностями расширений и неотрицательными и PSD- рангами
- Исследованы известные ограничивающие функционалы на PSD-ранги
- Доказано, что функционалы полиномиальны по рангу матрицы, а значит по размерности многогранника (для матриц невязок)
- В частности, получена оценка на взаимную информацию:

$$I(A:B) \leq \log \operatorname{rank} M$$