基于 SOS 的超压缩不等式与 Frankl-Rödl 图

周晟煜

清华大学

2025 年 12 月 30 日

摘要

- 在平方和（SOS）证明系统中，给出了反向超压缩不等式的形式化表述及其完全证明。
- 作为推论，证明了常数阶的 SOS 算法能够解决 Frankl-Rödl 图的整数间隙问题。
- 对于任意常数 $0 < \gamma \leq 1/4$，度数为 $4\lceil \frac{1}{4\gamma} \rceil$ 的 SOS/Lasserre 层级能够证明：

  Frankl-Rödl 图 FR_γ^n 的最大独立集大小为 $o(1)$。

背景：SOS 证明系统与半定规划

什么是 SOS (Sum-of-Squares) 证明系统?

- SOS 是一个基于代数的证明系统，用于证明多项式不等式 $P(x) \geq 0$。
- 核心思想：如果 $P(x)$ 可以写成一组多项式的平方和（$P(x) = \sum Q_i(x)^2$），那么 $P(x)$ 必然非负。
- **算法意义**：寻找 d 阶 SOS 证明等价于求解一个大小为 $n^{O(d)}$ 的半定规划（SDP）。
- 它是目前已知最强大的 SDP 松弛层级之一（Lasserre 层级）。

为什么研究 SOS?

- 在理论计算机科学中，我们想知道高效算法（如 P, 准 P）的极限在哪里。
- SOS 层级为许多优化问题（如最大割、唯一博弈）提供了最佳近似算法。

背景：Frankl-Rödl 图与整数间隙

Frankl-Rödl 图 (FR_{γ}^{n})

- 顶点是 n 维超立方体 $\{-1, 1\}^n$。
- 边连接那些汉明距离恰好为 $(1-\gamma)n$ 的点对。
- 汉明距离：两个二进制向量中不同坐标的数量。
- 该图的最大独立集非常小（密度趋于 0），因此色数非常大。

整数间隙 (Integrality Gap)

- 这是一个著名的“困难实例”。
- 尽管真实色数很大，但许多标准的 SDP 松弛（如 Lovász Theta 函数）会被“欺骗”，认为该图包含很大的独立集（即认为色数很小）。
- 这被称为“整数间隙”：凸松弛的最优解与整数规划的真实解之间存在巨大差距。

研究动机

- 证明常数阶的 SOS 确实能证明 Frankl-Rödl 图的独立集很小。
- 这种图的性质与布尔函数的分析性质（特别是噪声算子）密切相关。
- 在分析学中，控制噪声算子范数的工具是**超压缩不等式**。
- 要在 SOS 中证明图的性质，我们必须先在 SOS 系统中建立对应的分析工具——即**反向超压缩不等式**。

基本定义：布尔函数分析

定义 (噪声算子 T_ρ)

设 $f\colon\{-1,1\}^n \to \mathbb{R}$，$-1 \le \rho \le 1$，定义：

$$T_\rho f(x) = \mathbb{E}_y[f(y)|x]$$

其中 y 是 x 的 ρ-相关副本，即 $\mathbb{E}[x_i y_i] = \rho, \mathbb{E}[x_i] = \mathbb{E}[y_i] = 0$ 且各坐标独立。等价地，在傅里叶展开下，

$$T_\rho f = \sum_{S\subseteq[n]} \rho^{|S|}\hat{f}(S)\chi_S$$

定义 (p-范数)

$$||f||_p = \mathbb{E}_{x\sim\{-1,1\}^n}[|f(x)|^p]^{1/p}$$

这里期望是针对均匀分布取的。

SOS 证明系统的形式化定义

定义 (SOS 证明)

设 X 为变量集，$A = \{q_1 \geq 0, ..., q_m \geq 0\} \cup \{r_1 = 0, ..., r_{m'} = 0\}$ 为约束条件。我们说 A 以度数 k **SOS-证明** $p \geq 0$，记作 $A \vdash_k p \geq 0$，当且仅当存在 $v_1, ..., v_{m'}$ 和 SOS 多项式 $u_0, u_1, ..., u_m$ 使得：

$$p = u_0 + \sum_{i=1}^{m} u_i q_i + \sum_{j=1}^{m'} v_j r_j$$

且表达式中每一项的度数都不超过 k。

- **SOS 反驳**：如果 $A \vdash_k -1 \geq 0$，则称 A 具有 k 阶 SOS 反驳（说明集合 A 无解）。
- 单变量非负多项式是 SOS。
- 齐次双变量非负多项式是 SOS。
- 当 $A = \emptyset$ 时，有时使用简写 $\vdash_k p \geq 0$，这简单地意味着 p 是 SOS 且 $\deg(p) \leq k$。

预备知识

引理 (凸性)

对于任何 $k \in N^+$，我们有

$$\vdash_{2k} (\frac{X+Y}{2})^{2k} \leq \frac{X^{2k}+Y^{2k}}{2}$$

证明.

由齐次双变量非负多项式是 SOS 以及

$$\frac{X^{2k}+Y^{2k}}{2} - (\frac{X+Y}{2})^{2k} \geq 0$$

$\square$

超压缩不等式

超压缩不等式 (Hypercontractive Inequality)

设 $f: \{-1,1\}^n \to \mathbb{R}$，$1 \le p \le q \le \infty$，$0 \le \rho \le \sqrt{(p-1)/(q-1)}$。那么

$$||T_\rho f||_q \le ||f||_p$$

反向超压缩不等式 (Reverse Hypercontractive Inequality)

设 $f: \{-1,1\}^n \to \mathbb{R}^{\ge 0}$，$-\infty \le q \le p \le 1$，$0 \le \rho \le \sqrt{(1-p)/(1-q)}$。那么

$$||T_\rho f||_q \ge ||f||_p$$

本文证明的 SOS 版本定理

我们主要证明如下特殊情况的 SOS 版本：

定理 (正向)

对于偶整数 $q = 2s$，在矩条件下，存在不等式 $||T_\rho f||_q \le ||f||_2$ 的 *SOS* 证明。

定理 (反向)

设 $k \in N^+$，$\rho \le 1 - 1/(2k)$。对于任意函数 f, g，存在如下不等式的 *SOS* 证明：

$$\mathbb{E}_{(x,y)}[f(x)^{2k}g(y)^{2k}] \ge \mathbb{E}[f]^{2k}\mathbb{E}[g]^{2k}$$

基础情况 ($n=1$)

定理 (正向)

对于偶整数 $q=2s$，在矩条件下，存在不等式 $||T_\rho f||_q \leq ||f||_2$ 的 SOS 证明。

证明.

由于齐次性，设 $f(x)=1+\epsilon x$。展开两边：

$$LHS = \mathbb{E}[(1+\rho\epsilon x)^{2s}] = \sum \binom{2s}{2j} \rho^{2j}\epsilon^{2j}\mathbb{E}[x^{2j}]$$

$$RHS = (1+\epsilon^2)^s = \sum \binom{s}{j} \epsilon^{2j}$$

矩条件：$\mathbb{E}[x_i^2]=1$，$\mathbb{E}[x_i^{2j-1}]=0$，$\mathbb{E}[x_i^{2j}] \leq (2s-1)^j \frac{\binom{s}{j}}{\binom{2s}{2j}}$。 □

归纳步骤

从 $n = 1$ 推广到一般 n 需要处理乘积项。在 SOS 中，我们不能直接使用 Hölder 不等式。

引理

对于偶整数 v，有 *SOS* 证明：

$$\prod_{i=1}^{v} G_i H_i \leq \frac{1}{\binom{v}{v/2}} \sum_{|T|=v/2} \prod_{i \in T} G_i^2 \prod_{i \in [v] \setminus T} H_i^2$$

- 利用归纳法将 n 维函数拆解为 $f(x) = x_n g + h$。
- 展开 $T_\rho f$ 的 $2s$ 次方。
- 利用上述引理将交叉项 $G_i H_i$ 分解为平方项。
- 最后使用 Zeilberger 算法证明产生的组合系数恒等式。

反向超压缩

定理 4.1

$$\vdash_{4k} \mathbb{E}[f(x)^{2k} g(y)^{2k}] \geq \mathbb{E}[f]^{2k} \mathbb{E}[g]^{2k}$$

在 $n = 1$ 时，涉及 $f(1), f(-1), g(1), g(-1)$ 四个变量。通过变量代换和齐次化，进一步将其归约为证明双变量多项式 $P_k(a, b)$ 的非负性。

两点不等式

设 $k \in \mathbb{N}^+$ 且 $\rho^* = 1 - \frac{1}{2k}$。那么

$$\vdash_{4k} \begin{aligned} P_k(a,b) :=& (\frac{1}{4} + \frac{1}{4}\rho^*)((1+a)^{2k}(1+b)^{2k} + (1-a)^{2k}(1-b)^{2k}) \\ &+ (\frac{1}{4} - \frac{1}{4}\rho^*)((1+a)^{2k}(1-b)^{2k} + (1-a)^{2k}(1+b)^{2k}) - 1 \end{aligned} \geq 0$$

两点不等式的计算机辅助证明

为了证明 $P_k(a,b)$ 是 SOS，我们使用变量代换：

$$r = a - b, \quad s = a + b, \quad t = ab$$

将 P_k 重写为关于 s^2 的多项式：

$$P_k(a,b) = Q_{k,0}(t) + Q_{k,1}(t)s^2 + \cdots + Q_{k,k}(t)s^{2k}$$

关键步骤：

- 如果能证明每个系数多项式 $Q_{k,i}(t)$ 在实数上非负，则由单变量 SOS 事实，整体即为 SOS。
- 我们利用 Zeilberger 算法找到了 $Q_{k,i}(t)$ 的闭式解。
- 剩下的就是 dirty work 了。

归纳步骤

设 $n>1$。我们将 f 和 g 根据第 n 个坐标 x_n 的值分解为 f_0, f_1 和 g_0, g_1（例如 $f_0(x')$ 表示 $f(x',1)$）。期望可以展开为：

$$\begin{aligned}\mathbb{E}_{(x,y)}[f(x)^{2k}g(y)^{2k}] = (\frac{1}{4}+\frac{1}{4}\rho)\mathbb{E}[f_0(x)^{2k}g_0(y)^{2k}] + (\frac{1}{4}+\frac{1}{4}\rho)\mathbb{E}[f_1(x)^{2k}g_1(y)^{2k}] \\ + (\frac{1}{4}-\frac{1}{4}\rho)\mathbb{E}[f_0(x)^{2k}g_1(y)^{2k}] + (\frac{1}{4}-\frac{1}{4}\rho)\mathbb{E}[f_1(x)^{2k}g_0(y)^{2k}]\end{aligned}$$

应用四次归纳假设（对于 $n-1$ 维的情况 $\vdash_{4k}$）：

$$\begin{aligned}\vdash_{4k} \text{RHS} \geq (\frac{1}{4}+\frac{1}{4}\rho)\mathbb{E}[f_0]^{2k}\mathbb{E}[g_0]^{2k} + (\frac{1}{4}+\frac{1}{4}\rho)\mathbb{E}[f_1]^{2k}\mathbb{E}[g_1]^{2k} \\ + (\frac{1}{4}-\frac{1}{4}\rho)\mathbb{E}[f_0]^{2k}\mathbb{E}[g_1]^{2k} + (\frac{1}{4}-\frac{1}{4}\rho)\mathbb{E}[f_1]^{2k}\mathbb{E}[g_0]^{2k}\end{aligned}$$

最后，将 $n=1$ 的基础情况应用于上述不等式的右侧，即得证：

$$\geq \left(\frac{\mathbb{E}[f_0]+\mathbb{E}[f_1]}{2}\right)^{2k}\left(\frac{\mathbb{E}[g_0]+\mathbb{E}[g_1]}{2}\right)^{2k} = \mathbb{E}[f]^{2k}\mathbb{E}[g]^{2k} \quad \square$$

SOS 中的 Frankl-Rödl 定理

定理

给定 *Frankl-Rödl* 图 FR_γ^n，对于度数 $4\lceil\frac{1}{4\gamma}\rceil$ 的 *SOS* 系统，可以反驳以下关于独立集大小的断言：

$$\frac{1}{|V|}\sum_{x\in V} f(x) \geq Cn^{-\gamma/10}$$

即 *SOS* 能够“看出”该图没有大的独立集。

我们将图的组合性质翻译为分析语言，并利用反向超压缩不等式导出矛盾。

分析语言

定理

设 $n \in \mathbb{N}^+$ 且 $\frac{1}{\log n} \le \gamma \le \frac{1}{4}$ 使得 $(1-\gamma)n$ 为偶整数。给定 *Frankl-Rödl* 图 $FR_\gamma^n = (V, E)$，对于每个 $x \in V = \{-1, 1\}^n$，设 $f(x)$ 为一个不定元。那么存在度数为 $4\lceil \frac{1}{4\gamma} \rceil$ 的 *SOS* 反驳对应表述 $Max - IS(G) \ge O(n^{-\gamma/10})$ 的系统，即：

$$\left\{\begin{array}{ll} f(x)^2 = f(x) & \forall x \in V, \\ f(x)f(y) = 0 & \forall (x,y) \in E, \\ \frac{1}{|V|}\sum_{x \in V} f(x) \ge Cn^{-\gamma/10} & \end{array}\right\} \vdash_{4\lceil \frac{1}{4\gamma} \rceil} -1 \ge 0$$

其中 C 是任意常数。

证明步骤 1：算子转化

- 定义图的邻接算子 $S_d f(x) = \mathbb{E}_{y \sim N(x)}[f(y)]$。
- 由于 f 是独立集的指示函数，如果在距离 d 处有边，则 $\langle f, S_d f \rangle = 0$。
- **技术难点**：图算子有奇偶性限制。我们需要引入混合算子

$$S'_d = \frac{1}{2} S_d + \frac{1}{2} S_{d+1}$$

- (Siavosh Benabbas, Hamed Hatami, and Avner Magen.)：设 $d = n - c$。对于 $\rho = 1 - 2d/n$，有：

$$\langle f, S'_d f \rangle - \langle f, T_\rho f \rangle = \sum \hat{f}(U)^2 \delta(U)$$

其中误差项 $\delta(U)$ 满足

$$\|\delta(U)\|_\infty = O\left(\max\left\{n^{-1/5}, \frac{n}{c^2} \log^2\left(\frac{c^2}{n}\right)\right\}\right)$$

证明步骤 2：导出矛盾

- 我们将 f 根据汉明权重的奇偶性分解为 f_0 和 f_1。由于 d 是偶数，图的边只连接同奇偶性的点，因此独立集约束蕴含：

$$\mathcal{A} \vdash_2 \langle f_0, S'_d f_0 \rangle + \langle f_1, S'_d f_1 \rangle = 0$$

- 由步骤 1 得到

$$\{f(x)f(y) = 0 \forall \Delta(x,y) = d\} \vdash_2 \langle f_0, T_{\rho'} f_0 \rangle + \langle f_1, T_{\rho'} f_1 \rangle \leq \delta \mathbb{E}[f^2]$$

- 转换为正相关形式 $\rho = 1 - 2\gamma$，并提升至 $2k$ 次幂：

$$\begin{aligned}\{f^2(x) = f(x), \forall x, f(x)f(y) = 0 \forall \Delta(x,y) = d\} &\vdash_{4k} \\ &\mathbb{E}[f_0(x)^{2k} g_0(y)^{2k}] + \mathbb{E}[f_1(x)^{2k} g_1(y)^{2k}] \leq \delta \mathbb{E}[f]\end{aligned}$$

- $\{f^2(x) = f(x), \forall x\} \vdash_2 \delta \mathbb{E}[f] = \delta \mathbb{E}[2f - f^2] = \delta \mathbb{E}[1 - (1-f)^2] \leq \delta$

证明步骤 2：导出矛盾

- 应用反向超压缩定理 4.1 及凸性不等式：

$$
\begin{aligned}
&\left\{f^2(x) = f(x), \forall x, f(x)f(y) = 0 \forall \Delta(x,y) = d\right\} \vdash_{4k} \delta \\
&\geq \mathbb{E}[f_0]^{2k}\mathbb{E}[g_0]^{2k} + \mathbb{E}[f_1]^{2k}\mathbb{E}[g_1]^{2k} \\
&= \mathbb{E}[f_0]^{4k} + \mathbb{E}[f_1]^{4k} \geq 2\left(\frac{\mathbb{E}[f]}{2}\right)^{4k} = 2^{1-4k}\mathbb{E}[f]^{4k}
\end{aligned}
$$

- 代入条件 $\mathbb{E}[f] \geq Cn^{-\gamma/10}$ 得到：

$$
C^{4k}n^{-1/10} \leq 2^{4k-1}\delta \sim 2^{4k-1}n^{-1/5}
$$

- 即我们证明了 $\mathcal{A} \vdash_{4k} -1 \geq 0$，从而完成反驳。

结论

- 建立了反向超压缩不等式的 SOS 证明。
- 利用该工具解决了 Frankl-Rödl 图的整数间隙问题。
- 展示了计算机代数在证明复杂性中的威力。
- **开放问题**：
  - 是否存在更优雅的、无需计算机辅助的证明?
  - 能否将结果推广到 γ 更小的情况（如 $\sqrt{\frac{\log n}{n}}$)?

参考文献

Manuel Kauers, Ryan O'Donnell, Li-Yang Tan, and Yuan Zhou. Hypercontractive inequalities via SOS, and the Frankl-Rödl graph. 2013.

谢谢！