

多项式优化入门

王杰

中国科学院数学与系统科学研究院

北京师范大学, 2024 年秋季



课程内容

1. 半定规划
2. 平方和理论
3. 测度和矩
4. 矩-平方和松弛分层
5. 项稀疏 (TS)
6. 变量稀疏 (CS)
7. 扩展与应用
8. 软件与实验

矩-平方和松弛分层

$$f_{\min} := \begin{cases} \inf_{\mathbf{x} \in \mathbb{R}^n} & f(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \geq 0, \quad i \in [m] \\ & h_j(\mathbf{x}) = 0, \quad j \in [\ell] \end{cases}$$

$$\begin{cases} \inf_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_r(\mathbf{y}) \succeq 0, y_0 = 1 \\ & \mathbf{M}_{r-d_i}(g_i \mathbf{y}) \succeq 0, i \in [m] \\ & L_{\mathbf{y}}(\mathbf{x}^{\alpha} h_j) = 0, \forall |\alpha| \leq 2r - \deg(h_j), j \in [\ell] \end{cases} \quad \begin{cases} \sup_{\lambda} & \lambda \\ \text{s.t.} & f - \lambda \in \mathcal{Q}(\mathbf{g})_{2r} + \mathcal{I}(\mathbf{h})_{2r} \end{cases}$$

矩-平方和松弛分层

$$f_{\min} := \begin{cases} \inf_{\mathbf{x} \in \mathbb{R}^n} & f(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \geq 0, \quad i \in [m] \\ & h_j(\mathbf{x}) = 0, \quad j \in [\ell] \end{cases}$$

$$\begin{cases} \inf_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_r(\mathbf{y}) \succeq 0, y_0 = 1 \\ & \mathbf{M}_{r-d_i}(g_i \mathbf{y}) \succeq 0, i \in [m] \\ & L_{\mathbf{y}}(\mathbf{x}^{\alpha} h_j) = 0, \forall |\alpha| \leq 2r - \deg(h_j), j \in [\ell] \end{cases} \quad \begin{cases} \sup_{\lambda} & \lambda \\ \text{s.t.} & f - \lambda \in \mathcal{Q}(\mathbf{g})_{2r} + \mathcal{I}(\mathbf{h})_{2r} \end{cases}$$

性质

- 假定 Archimedean 条件:

- ▶ $\lambda_r \nearrow f_{\min}, \lambda_r^* \nearrow f_{\min}, r \rightarrow \infty$

- ▶ 如果多项式优化问题有唯一最优解 \mathbf{x}^* , 那么 $\lim_{r \rightarrow \infty} L_{y^r}(\mathbf{x}) = \mathbf{x}^*$

- ▶ 一般地, 有限收敛性成立

- 平坦性条件或秩一: 达到全局最优, 可提取全局最优解

性质

- 假定 Archimedean 条件:

- ▶ $\lambda_r \nearrow f_{\min}, \lambda_r^* \nearrow f_{\min}, r \rightarrow \infty$

- ▶ 如果多项式优化问题有唯一最优解 \mathbf{x}^* , 那么 $\lim_{r \rightarrow \infty} L_{y^r}(\mathbf{x}) = \mathbf{x}^*$

- ▶ 一般地, 有限收敛性成立

- 平坦性条件或秩一: 达到全局最优, 可提取全局最优解

性质

- 假定 Archimedean 条件:

- ▶ $\lambda_r \nearrow f_{\min}, \lambda_r^* \nearrow f_{\min}, r \rightarrow \infty$

- ▶ 如果多项式优化问题有唯一最优解 \mathbf{x}^* , 那么 $\lim_{r \rightarrow \infty} L_{y^r}(\mathbf{x}) = \mathbf{x}^*$

- ▶ 一般地, 有限收敛性成立

- 平坦性条件或秩一: 达到全局最优, 可提取全局最优解

性质

- 假定 Archimedean 条件:

- ▶ $\lambda_r \nearrow f_{\min}, \lambda_r^* \nearrow f_{\min}, r \rightarrow \infty$

- ▶ 如果多项式优化问题有唯一最优解 \mathbf{x}^* , 那么 $\lim_{r \rightarrow \infty} L_{y^r}(\mathbf{x}) = \mathbf{x}^*$

- ▶ 一般地, 有限收敛性成立

- 平坦性条件或秩一: 达到全局最优, 可提取全局最优解

稀疏单项式基

- 无约束: Newton 多面体 \rightsquigarrow 稀疏单项式基

- f, g_i, h_j 稀疏:

- ▶ 构造一组稀疏单项式基 $\mathbf{v}(\mathbf{x}) \subsetneq [\mathbf{x}]_r$

- $\rightsquigarrow M_{\mathbf{v}}(\mathbf{y}) := L_{\mathbf{y}}(\mathbf{v}(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x})^{\top})$

稀疏单项式基

- 无约束: Newton 多面体 \rightsquigarrow 稀疏单项式基

- f, g_i, h_j 稀疏:

- ▶ 构造一组稀疏单项式基 $\mathbf{v}(\mathbf{x}) \subsetneq [\mathbf{x}]_r$

- $\rightsquigarrow \mathbf{M}_{\mathbf{v}}(\mathbf{y}) := L_{\mathbf{y}}(\mathbf{v}(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x})^T)$

商环

- 等式约束: $h_j(\mathbf{x}) = 0, \quad j = 1, \dots, \ell$

$$f - \lambda \in \mathcal{Q}(\mathbf{g}) + \mathcal{I}(\mathbf{h}) \iff f - \lambda \in \mathcal{Q}(\mathbf{g}) \pmod{\mathcal{I}(\mathbf{h})}$$

- 利用 Gröbner 基计算一纽约化单项式基
- 在商环 $\mathbb{R}[\mathbf{x}]/\mathcal{I}(\mathbf{h})$ 上建立等式:

$$f - \lambda = \sigma_0 + \sum_{i=1}^m \sigma_i g_i$$

- ↪ 用 Gröbner 基计算 normal form

商环

- 等式约束: $h_j(\mathbf{x}) = 0, \quad j = 1, \dots, \ell$

$$f - \lambda \in \mathcal{Q}(\mathbf{g}) + \mathcal{I}(\mathbf{h}) \iff f - \lambda \in \mathcal{Q}(\mathbf{g}) \pmod{\mathcal{I}(\mathbf{h})}$$

- ▶ 利用 Gröbner 基计算一纽约化单项式基
- ▶ 在商环 $\mathbb{R}[\mathbf{x}]/\mathcal{I}(\mathbf{h})$ 上建立等式:

$$f - \lambda = \sigma_0 + \sum_{i=1}^m \sigma_i g_i$$

- ↪ 用 Gröbner 基计算 normal form

例子（二元优化）

$$\begin{cases} \inf_{\mathbf{x}} & \mathbf{x}Q\mathbf{x}^\top \\ \text{s.t.} & x_i^2 = 1, \quad i = 1, \dots, n \end{cases}$$

- $\mathbf{v}(\mathbf{x}) = \{1, x_1, \dots, x_n, x_1x_2, \dots, x_{n-1}x_n\}$
- 等式约束：

$$\mathbf{x}Q\mathbf{x}^\top - \lambda = \mathbf{v}(\mathbf{x})G\mathbf{v}(\mathbf{x})^\top \quad \text{mod } \mathcal{I}(\mathbf{h})$$

例子（二元优化）

	基的大小	等式约束个数
原始	$\binom{n+2}{2}$	$\binom{n+4}{4}$
商环	$1 + n + \binom{n}{2}$	$1 + n + \binom{n}{2} + \binom{n}{3} + \binom{n}{4}$

$$n = 10$$

原始	66	1001
商环	56	386

$$n = 20$$

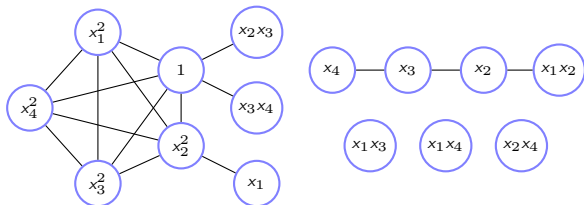
原始	231	10626
商环	211	6196

项稀疏型

- $f = \sum_{\alpha} f_{\alpha} \mathbf{x}^{\alpha} \in \mathbb{R}[\mathbf{x}]$, $\text{supp}(f) := \{\mathbf{x}^{\alpha} \mid f_{\alpha} \neq 0\}$
- 单项式基: $[\mathbf{x}]_r$
- 项稀疏型图 $G^{\text{tsp}}(V, E)$:
 - ▶ $V := [\mathbf{x}]_r$
 - ▶ $\{\mathbf{x}^{\alpha}, \mathbf{x}^{\beta}\} \in E \iff \mathbf{x}^{\alpha} \cdot \mathbf{x}^{\beta} = \mathbf{x}^{\alpha+\beta} \in \text{supp}(f) \cup \bigcup_{i=1}^m \text{supp}(g_i) \cup [\mathbf{x}]_r^2$

项稀疏型图

$$\left\{ \begin{array}{l} \inf_{\mathbf{x} \in \mathbb{R}^4} f(\mathbf{x}) = x_1^4 + x_1 x_2^2 + x_2 x_3 + x_3^2 x_4 \\ \text{s.t.} \quad g_1(\mathbf{x}) = 1 - x_1^2 - x_2^2 - x_3^2 \geq 0 \\ \quad \quad g_2(\mathbf{x}) = 1 - x_3 x_4 \geq 0 \end{array} \right.$$



TSSOS 迭代程序

- 对于图 $G(V, E)$, $V = [\mathbf{x}]_r$, 定义

$$\text{supp}(G) := \left\{ \mathbf{x}^{\alpha+\beta} \mid \alpha = \beta \text{ 或 } \{\mathbf{x}^\alpha, \mathbf{x}^\beta\} \in E \right\}$$

- 令 $G_0^{(0)} := G^{\text{tsp}}$. 定义图的升链 $(G_0^{(k)})_{k \geq 1}$:

- ① **支撑扩张**: 令图 $F_0^{(k)}$ 具有顶点集 $[\mathbf{x}]_r$ 和边集

$$E(F_0^{(k)}) := \left\{ \{\mathbf{x}^\alpha, \mathbf{x}^\beta\} \mid \mathbf{x}^{\alpha+\beta} \in \text{supp}(G_0^{(k-1)}) \right\}$$

- ② **弦扩张**: $G_0^{(k)} = (F_0^{(k)})'$

TSSOS 迭代程序

- 对于图 $G(V, E)$, $V = [\mathbf{x}]_r$, 定义

$$\text{supp}(G) := \left\{ \mathbf{x}^{\alpha+\beta} \mid \alpha = \beta \text{ 或 } \{\mathbf{x}^\alpha, \mathbf{x}^\beta\} \in E \right\}$$

- 令 $G_0^{(0)} := G^{\text{tsp}}$. 定义图的升链 $(G_0^{(k)})_{k \geq 1}$:

- ① **支撑扩张**: 令图 $F_0^{(k)}$ 具有顶点集 $[\mathbf{x}]_r$ 和边集

$$E(F_0^{(k)}) := \left\{ \{\mathbf{x}^\alpha, \mathbf{x}^\beta\} \mid \mathbf{x}^{\alpha+\beta} \in \text{supp}(G_0^{(k-1)}) \right\}$$

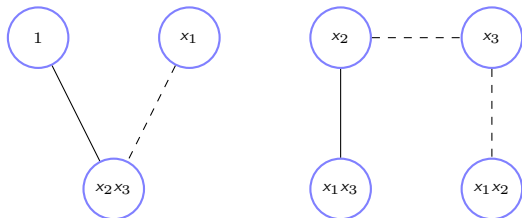
- ② **弦扩张**: $G_0^{(k)} = (F_0^{(k)})'$

支撑扩张

- 考虑图 $G(V, E)$:

$$V = \{1, x_1, x_2, x_3, x_2x_3, x_1x_3, x_1x_2\}$$

$$E = \{\{1, x_2x_3\}, \{x_2, x_1x_3\}\}$$

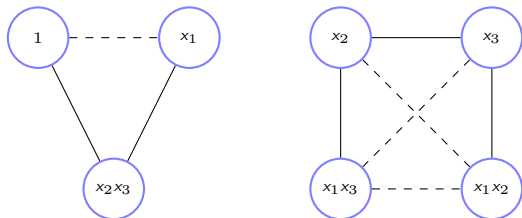


最大弦扩张

- 考虑图 $G(V, E)$:

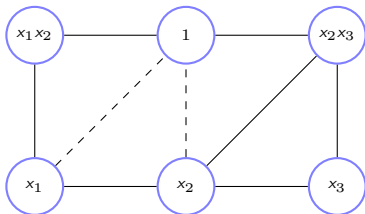
$$V = \{1, x_1, x_2, x_3, x_2x_3, x_1x_3, x_1x_2\}$$

$$E = \{\{1, x_2x_3\}, \{x_2, x_1x_3\}, \{x_1, x_2x_3\}, \{x_2, x_3\}, \{x_3, x_1x_2\}\}$$



(近似) 最小弦扩张

- $f = x_1^2 - 2x_1x_2 + x_2^2 - 2x_1^2x_2 + x_1^2x_2^2 - 2x_2x_3 + x_3^2 + x_2^2x_3 - x_2x_3^2 + x_2^2x_3^2$
- 单项式基: $\{1, x_1, x_2, x_3, x_1x_2, x_2x_3\}$



TSSOS 松弛

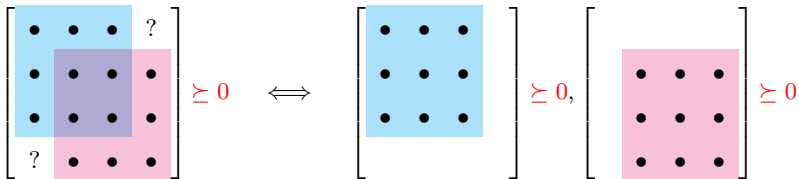
- 对每个不等式约束 $g_i \geq 0$, 用类似的方式定义图列 $(G_i^{(k)})_{k \geq 1}$
- 令 $B(G_i^{(k)})$ 是 $G_i^{(k)}$ 的邻接矩阵
- TSSOS 松弛:

$$\left\{ \begin{array}{l} \inf_{\mathbf{y}} \quad L_{\mathbf{y}}(f) \\ \text{s.t.} \quad B(G_0^{(k)}) \circ \mathbf{M}_r(\mathbf{y}) \in \mathbb{S}^+(G_0^{(k)}, ?) \\ \quad \quad B(G_i^{(k)}) \circ \mathbf{M}_{r-d_i}(g_i \mathbf{y}) \in \mathbb{S}^+(G_i^{(k)}, ?), \quad i \in [m] \\ \quad \quad y_0 = 1 \end{array} \right.$$

半正定补全矩阵刻画

定理

设 G 是弦图, C_1, \dots, C_t 是 G 的极大团集. 则 $Q \in \mathbb{S}^+(G, ?)$ 当且仅当 $Q_{C_i} \succeq 0, i = 1, \dots, t$.



TSSOS 松弛

- 1 取 $G_0^{(k)}$ 的所有极大团: $C_1^0, \dots, C_{t_0}^0$
- 2 对每个极大团 C_p^0 , 取 $M_r(\mathbf{y})$ 的主子矩阵 $\mathbf{M}_{C_p^0}(\mathbf{y})$
- 3 类似地, 取局部化矩阵 $M_{r-d_i}(g_i\mathbf{y})$ 的主子矩阵 $\mathbf{M}_{C_p^i}(\mathbf{y})$
- 4 构造分块矩-平方和松弛 \rightsquigarrow TSSOS 松弛

$$\lambda_r^{(k)} := \begin{cases} \inf & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_{C_p^0}(\mathbf{y}) \succeq 0, \quad p \in [t_0] \\ & \mathbf{M}_{C_p^i}(g_i\mathbf{y}) \succeq 0, \quad p \in [t_i], i \in [m] \\ & y_0 = 1 \end{cases}$$

TSSOS 松弛分层的性质

- 对于 QCQP, $\lambda_1^{(1)} = \lambda_1$
- 固定稀疏阶数 k , 序列 $(\lambda_r^{(k)})_{r \geq r_{\min}}$ 单调非减
- 固定松弛阶数 r , 序列 $(\lambda_r^{(k)})_{k \geq 1}$ 单调非减
- 使用最大弦扩张, 序列 $(\lambda_r^{(k)})_{k \geq 1}$ 有限步收敛到 λ_r

双层下界网络

$$\begin{array}{ccccccc} \lambda_{r_{\min}}^{(1)} & \leq & \lambda_{r_{\min}}^{(2)} & \leq & \cdots & \leq & \lambda_{r_{\min}} \\ \wedge & & \wedge & & & & \wedge \\ \lambda_{r_{\min}+1}^{(1)} & \leq & \lambda_{r_{\min}+1}^{(2)} & \leq & \cdots & \leq & \lambda_{r_{\min}+1} \\ \wedge & & \wedge & & & & \wedge \\ \vdots & & \vdots & & \vdots & & \vdots \\ \wedge & & \wedge & & & & \wedge \\ \lambda_r^{(1)} & \leq & \lambda_r^{(2)} & \leq & \cdots & \leq & \lambda_r \\ \wedge & & \wedge & & & & \wedge \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

SDSOS 优化

- DD 矩阵 $Q \in \mathbb{S}^n$: $Q_{ii} \geq \sum_{j \neq i} |Q_{ij}|$, $i \in [n]$
- SDD 矩阵 $Q \in \mathbb{S}^n$: 存在对角矩阵 D 使得 DQD 是对角占优矩阵
- SDSOS 多项式 $f(\mathbf{x})$: 存在一个 Gram 矩阵是 SDD 矩阵
- SDSOS 优化:

$$\lambda_{\text{sdsos}} := \sup_{\lambda \in \mathbb{R}} \{\lambda \mid f(\mathbf{x}) - \lambda \in \text{SDSOS}\}$$

▶ $\lambda_{\text{sdsos}} \leq \lambda^{(1)}$

SDSOS 优化

- DD 矩阵 $Q \in \mathbb{S}^n$: $Q_{ii} \geq \sum_{j \neq i} |Q_{ij}|$, $i \in [n]$
- SDD 矩阵 $Q \in \mathbb{S}^n$: 存在对角矩阵 D 使得 DQD 是对角占优矩阵
- SDSOS 多项式 $f(\mathbf{x})$: 存在一个 Gram 矩阵是 SDD 矩阵
- SDSOS 优化:

$$\lambda_{\text{sdsos}} := \sup_{\lambda \in \mathbb{R}} \{ \lambda \mid f(\mathbf{x}) - \lambda \in \text{SDSOS} \}$$

▶ $\lambda_{\text{sdsos}} \leq \lambda^{(1)}$

SDSOS 优化

- DD 矩阵 $Q \in \mathbb{S}^n$: $Q_{ii} \geq \sum_{j \neq i} |Q_{ij}|$, $i \in [n]$
- SDD 矩阵 $Q \in \mathbb{S}^n$: 存在对角矩阵 D 使得 DQD 是对角占优矩阵
- SDSOS 多项式 $f(\mathbf{x})$: 存在一个 Gram 矩阵是 SDD 矩阵
- SDSOS 优化:

$$\lambda_{\text{sdsos}} := \sup_{\lambda \in \mathbb{R}} \{ \lambda \mid f(\mathbf{x}) - \lambda \in \text{SDSOS} \}$$

► $\lambda_{\text{sdsos}} \leq \lambda^{(1)}$

符号对称性

- 符号对称性：二元向量 $\theta \in \{-1, 1\}^n$ 使得

$$\forall \mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = f(\theta \circ \mathbf{x})$$

- $f = x_1^2 + x_1x_2 + x_3^2$

- $\triangleright \theta_1 = (-1, -1, 1), \theta_2 = (1, 1, -1)$

符号对称性

- 符号对称性：二元向量 $\theta \in \{-1, 1\}^n$ 使得

$$\forall \mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = f(\theta \circ \mathbf{x})$$

- $f = x_1^2 + x_1x_2 + x_3^2$

- ▶ $\theta_1 = (-1, -1, 1), \theta_2 = (1, 1, -1)$

符号对称性诱导分块结构

- 二元矩阵 R : $f(\mathbf{x}), g_1(\mathbf{x}), \dots, g_m(\mathbf{x})$ 的所有符号对称性
- 在 $[\mathbf{x}]_r$ 上定义等价关系 \sim :

$$\mathbf{x}^\alpha \sim \mathbf{x}^\beta \iff R^\top(\boldsymbol{\alpha} + \boldsymbol{\beta}) \equiv 0 \pmod{2}$$

▶ $[\mathbf{x}]_r = \bigsqcup_{p=1}^{t_0} \mathcal{B}_p^0$

▶ $[\mathbf{x}]_{r-d_i} = \bigsqcup_{p=1}^{t_i} \mathcal{B}_p^i$

• $f = x_1^2 + x_1x_2 + x_3^2$: $\{1, x_1x_2, x_1^2, x_2^2, x_3^2\}, \{x_1, x_2\}, \{x_3\}, \{x_1x_3, x_2x_3\}$

符号对称性诱导分块结构

- 二元矩阵 R : $f(\mathbf{x}), g_1(\mathbf{x}), \dots, g_m(\mathbf{x})$ 的所有符号对称性
- 在 $[\mathbf{x}]_r$ 上定义等价关系 \sim :

$$\mathbf{x}^\alpha \sim \mathbf{x}^\beta \iff R^\top(\boldsymbol{\alpha} + \boldsymbol{\beta}) \equiv 0 \pmod{2}$$

▶ $[\mathbf{x}]_r = \bigsqcup_{p=1}^{t_0} \mathcal{B}_p^0$

▶ $[\mathbf{x}]_{r-d_i} = \bigsqcup_{p=1}^{t_i} \mathcal{B}_p^i$

• $f = x_1^2 + x_1x_2 + x_3^2$: $\{1, x_1x_2, x_1^2, x_2^2, x_3^2\}, \{x_1, x_2\}, \{x_3\}, \{x_1x_3, x_2x_3\}$

TSSOS 与符号对称性

- ▶ TSSOS 迭代给出的分块结构是符号对称性诱导的分块结构的加细

定理 (Wang, Magron, and Lasserre, 2021)

固定松弛阶数 r , 使用最大弦扩张. 则 TSSOS 迭代给出的分块结构在有限步内收敛到符号对称性诱导的分块结构.

TSSOS 与符号对称性

- ▶ TSSOS 迭代给出的分块结构是符号对称性诱导的分块结构的加细

定理 (Wang, Magron, and Lasserre, 2021)

固定松弛阶数 r , 使用最大弦扩张. 则 TSSOS 迭代给出的分块结构在有限步内收敛到符号对称性诱导的分块结构.

稀疏 Putinar's Positivstellensatz

定理 (Wang, Magron, and Lasserre, 2021)

假设 $\mathcal{Q}(g)$ 满足**阿基米德条件**. 令 R 是 $f(x), g_1(x), \dots, g_m(x)$ 的所有符号对称性给出的二元矩阵. 如果 f 在 S 上是严格正的, 则 f 有表示

$$f = \sigma_0 + \sum_{i=1}^m \sigma_i g_i,$$

其中 σ_i 是 SOS 满足 $R^T \alpha \equiv 0 \pmod{2}, \forall \mathbf{x}^\alpha \in \text{supp}(\sigma_i), i = 0, \dots, m$.

- TSSOS: 基于稀疏平方和松弛的多项式优化软件包
- GloptiPoly: 基于矩松弛的多项式优化软件包
- Yalmip: 通用的优化平台, 包含 SOS 优化模块
- MOSEK: 商业 SDP 求解器

随机生成的平方和多项式

- $f = \sum_{i=1}^t f_i^2 \in \mathbf{randpoly1}(n, 2d, t, \rho)$

n	$2d$	TSSOS (Min)	TSSOS (Max)	GloptiPoly	Yalmip
8	8	0.24	1.7	306	10
8	10	0.58	4.8	-	92
9	10	0.50	3.2	-	44
10	12	2.2	12	-	474
10	16	36	15	-	-
12	12	8.4	61	-	-

随机生成的具有单形牛顿多面体的多项式

- $f = c_0 + \sum_{i=1}^n c_i x_i^{2d} + \sum_{j=1}^{s-n-1} c'_j x^{\alpha_j} \in \mathbf{randpoly2}(n, 2d, s)$

n	$2d$	TSSOS (Min)	TSSOS (Max)	GloptiPoly	Yalmip
8	8	0.36	8.5	346	31
9	8	1.0	40	-	-
9	10	6.6	24	-	322
10	8	1.2	13	-	-
11	8	1.7	13	-	655
12	8	10	693	-	-

下次课

- 变量稀疏 (CS)

参考文献

- Jie Wang and Victor Magron, **Sparse Polynomial Optimization: Theory and Practice**, World Scientific Publishing, 2023.
- Jie Wang, Victor Magron, and Jean B. Lasserre, **TSSOS: A Moment-SOS hierarchy that exploits term sparsity**, SIAM Journal on Optimization, 2021.
- Jie Wang, Victor Magron, and Jean B. Lasserre, **Chordal-TSSOS: a moment-SOS hierarchy that exploits term sparsity with chordal extension**, SIAM Journal on Optimization, 2021.

更多信息见个人主页

<https://wangjie212.github.io/jiewang>