arXiv:2303.03826v1 [math.OC] 7 Mar 2023

单项式曲线的凸包及一个稀疏正定性定理

GENNADIY AVERKOV 和 CLAUS SCHEIDER ER

Abstract. 考虑由参数式（$t^{m_1},\dots,t^{m_n}$）给出的单项式曲线的 K ，其中参数 t在区间 I内变化。我们通过构造性论证表明， K 可以由 $\mathcal{O}(d)$ 个线性矩阵不等式（LMIs）给出提升后的半定描述，每个线性矩阵不等式的大小为$\lfloor\frac{n}{2}\rfloor+1$，其中 $d=\max\{m_1,\dots,m_n\}$ 是该曲线的度。在对偶方面，我们证明：如果一个次数为 d 、至多含有 $2k+1$ 个单项式的单变量多项式 $p(t)$在 $\mathbb{R}_+$上非负，那么 p 可以表示为 $p=t^0\sigma_0+\cdots+t^{d-k}\sigma_{d-k}$，其中多项式 $\sigma_0,\dots,\sigma_{d-k}$是平方和多项式，且 $\deg(\sigma_i)\le 2k$。后者是针对稀疏多项式的单变量正性定理，其中 p的非负性可由平方和多项式予以证明，而其度仅取决于 p的稀疏性。我们的结果属于将多项式优化问题表述为具有小规模线性矩阵不等式的半定问题这一总体尝试。此类小规模描述从计算角度来看更易于处理。

1. 引言

多项式优化研究的是在半代数集 $S\subseteq\mathbb{R}^n$ 上最小化一个 n 元多项式 $p\in\mathbb{R}[x_1,\dots,x_n]$ 的问题。该问题的凸化方法包括选取有限集 $A\supseteq\operatorname{supp}(p)$，其中元素为 $\mathbb{Z}_+^n$，考察满足 $\alpha\in A$ 的单项式 $x^\alpha=x_1^{\alpha_1}\cdots x_n^{\alpha_n}$，并求出闭凸包的精确或近似描述

$$K=\overline{\operatorname{conv}\{(x^\alpha)_{\alpha\in A}\,:\,x\in S\}}$$

在某种优化框架下。基本思想是，优化多项式 $p=c_0+\sum_{\alpha\in A}c_\alpha x^\alpha\in\mathbb{R}[x_1,\dots,x_n]$ 意味着优化仿射线性函数λ：$(y_\alpha)_{\alpha\in A}\mapsto c_0+\sum_{\alpha\in A}c_\alpha y_\alpha$，其定义域为 K。通常，集合 K 的描述源自实代数中的正性定理，因为 λ 在 K 上的非负性对应于 p 在 S上的非负性。实代数中的大多数正性定理都采用平方和（sos）证书。这一事实建立了实代数与半定优化之间的联系，因为由 n个变量构成、度为 $\le 2d$ 的平方和多项式所成的锥$\Sigma_{n,2d}$ 是大小为 $\binom{n+d}{n}$的对称半正定矩阵锥的线性像。正如 Lasserre 在其开创性工作 [12] 中所提出的，并沿着这一思路，Putinar 正性定理 [16] 通过半定约束给出了 K 的一系列外逼近。尽管拉瑟尔方法为将多项式问题转化为半定问题提供了通用模板，但通常还需要进一步调整，才能使其在计算上可行。尽管半定规划可以在多项式时间内求解，但.....

Date：2023年3月8日。

在给定误差容限 $[3, 15]$, 的条件下求解半定规划，在实践中很快就会变得极具挑战性，因为半定约束的规模是一个关键参数。我们将半定问题视为在有限个线性矩阵不等式（LMIs）的约束下优化一个线性函数的问题。规模为 d 的线性矩阵不等式是条件 $M(y) \succeq 0$，它要求对称矩阵 $d \times d$ 矩阵 $M(y)$具有半正定性，其中各项是变量 $y = (y_1, \ldots, y_N)$的仿射线性函数。为了初步了解线性矩阵不等式规模较大如何增加工作量，考虑用凸问题$L(y) - \epsilon \sum_{i=1}^{\ell} \log \det M_i(y)$来近似半定问题 $\inf\{L(y) : M_1(y) \succeq 0, \ldots, M_\ell(y) \succeq 0\}$；该凸问题使用对数障碍函数和一个较小的参数$\epsilon > 0$。用梯度下降法求解该凸问题，需要计算障碍函数的梯度，而这要求对矩阵$M_i(y)$ $[6, 4.3.1]$求逆。然而，已知对规模较大的 $M_i(y)$求逆代价高昂。因此，尽管理论研究中通常会使用块对角矩阵，将 ℓ 个线性矩阵不等式建模为一个规模为 ℓd的线性矩阵不等式，但这种化简掩盖了效率问题，因为一个规模为 ℓd 的一般线性矩阵不等式，其计算成本远高于 ℓ 个规模为 d的线性矩阵不等式。

为解决这一问题，文献 $[4]$ 引入了以下术语:

定义 1.1. 设 $C \subseteq \mathbb{R}^n$ 为一个集合。将 C 表示为以下形式

$$C = \{x \in \mathbb{R}^n : M_1(x, y) \succeq 0, \ldots, M_\ell(x, y) \succeq 0 \text{ for some } y \in \mathbb{R}^m\} \qquad (1)$$

其中 M_i 是关于 (x, y) 的线性矩阵不等式，称为 C 的提升（或扩展）半正定表示。如果每个 M_i 的大小均为 $\leq d$，则称 (1) 是大小为 $\leq d$ 的提升表示。使得 C 具有大小为 $\leq d$ 的提升表示的最小 d，称为 C 的半正定扩展次数，记为 sxdeg(C)。如果不存在这样的 d，则令 sxdeg(C) $= \infty$。

研究 sxdeg(C)，可以跟踪优化 C 上线性函数所需线性矩阵不等式的大小，而不考虑其数量。对于计算而言，线性矩阵不等式的大小和数量都发挥作用，但大小是更为关键的参数。显然，当且仅当 C 是多面体时，sxdeg(C) $= 1$；同样容易看出，当且仅当 C 是二阶锥可表示的，sxdeg(C) ≤ 2。文献 $[4]$ 表明，n 元、$\leq 2d$ 度的平方和多项式构成的锥 $\Sigma_{n,\ 2d}$ 具有 sxdeg $= \binom{n+d}{n}$，这与拉瑟尔方法中线性矩阵不等式的大小相符。$\binom{n+d}{n}$ 在 n 和 d 中的快速增长，解释了为何求解 Lasserre 松弛所需的计算时间对 n 和 d 的选择极为敏感。许多研究人员都发现了线性矩阵不等式大小方面的这一问题。文献 $[1, 5, 21, 22]$ 提出了一些在实践中应对该问题的建议。文献 $[1, 5, 21, 22]$ 的共同总体思路是，有意选择一个锥 C，其 sxdeg(C) 由构造可知较小，将其作为 K 的易处理外逼近。或者，从对偶观点出发，选择一个由非负多项式构成的锥 P，这些多项式的结构比 Putinar 正性定理中的更为特殊，使得 P 具有较小的 sxdeg(P)。尽管这些方法在实践中似乎有所帮助，但它们纯粹是启发式方法，因为目前仍缺少正性定理来保证：正性始终能够以预期的、计算成本更低的方式得到认证。

本文通过研究单变量稀疏多项式优化问题中半正定表示的规模，迈出了弥补这一空白的第一步。考虑一组正整数 $A = \{m_1, \ldots, m_n\}$ 和一个非退化区间 $I \subseteq \mathbb{R}$。对于优化单变量多项式

$p = \sum_{a\in A} c_a t^a \in \mathbb{R}[t]$ 在区间 I上，我们希望找到 $K = \mathrm{conv}\{(t^{m_1}, \ldots, t^{m_n}) : t \in I\} \subseteq \mathbb{R}^n$ 的提升半正定表示。由于曲线 $(t^{m_1}, \ldots, t^{m_n})$ 是度为 m_1 的有理正规曲线 $(t^i)_{1\le i\le m_1}$ 的投影，拉瑟尔方法给出了规模为 $\lfloor \frac{m_1}{2} \rfloor + 1$ 的描述。当单项式的数量 $n = |A|$ 相对于度 m_1 较少时，人们希望找到规模更小的替代描述。我们的主要结果（定理 3.19）表明，此类 K 的半正定扩展次数至多为 $\lfloor \frac{n}{2} \rfloor + 1$，且这是可能的最佳界。因此，描述的规模仅取决于单项式的数量，而与度无关。该描述完全明确，并由一个稀疏正性定理（定理 3.9）得出。后者利用锥 $t^0\Sigma_{1,\ 2k} + \cdots + t^{d-2k}\Sigma_{1,\ 2k}$，刻画区间$I \subseteq \mathbb{R}_+$上至多含有 $2k+1$ 项、度为 d 的多项式 p 的非负性；该锥的半正定扩展次数为 $k+1$。我们证明中的关键技术要素，是对称多项式理论中的舒尔多项式所满足的 Jacobi 双交错式公式。

我们关于一元稀疏多项式非负性的结果，应有助于理解基于稀疏性的方法对任意维度优化问题的影响。一元问题一方面可能展示出在每个维度中都会出现的现象；另一方面，每当我们能在一维情形中发现某种局限性时，它也将以某种形式出现在 n维情形中。此外，我们的主要结果还可以作为沿多个方向开展进一步研究的起点。显然，如果能在多变量情形中获得类似结果，将会很有意义。虽然我们的 K 给定为一条单项式曲线的闭凸包，但人们可以进一步探讨：对于任意半代数曲线 K 的闭凸包，半正定扩张次数能否得到界。对于 $n = 2$， [17] 证明了平面中的每个闭凸半代数集K 都是二阶锥可表示的，即其 $\mathrm{sxdeg}(K) \le 2$。可以证明，对于任意半代数曲线 $S \subseteq \mathbb{R}^n$的闭凸包，界 $\mathrm{sxdeg}(K) \le \lfloor \frac{n}{2} \rfloor + 1$ 成立，但证明要复杂得多 [18]。

本文组织如下。第2节介绍一些预备知识后，第3节给出主要结果。第4节中，我们引入余正少项式之和锥，它们是非负多项式锥的稀疏对应物。我们证明，在一维情形下，这些锥具有由小规模线性矩阵不等式构成的半正定描述，并说明此类描述可导出基于小规模线性矩阵不等式的 Lasserre 松弛变体。

2. Preliminaries

2.1. 令 $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ 和 $\mathbb{N} = \{1, 2, 3, \ldots\}$ 和 $\mathbb{R}_+ = [0, \infty[$。集合 A 的基数记为 $|A|$。对于元组 $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n$，我们使用记号 $|\alpha| := \alpha_1 + \cdots + \alpha_n$。对于单变量多项式，我们通常用 t 表示变量，并写作 $\mathbb{R}[t]_d = \{f \in \mathbb{R}[t] : \deg(f) \le d\}$。多项式 $p = \sum_\alpha c_\alpha x^\alpha \in \mathbb{R}[x_1, \ldots, x_n]$ 的支撑是集合 $\mathrm{supp}(p) := \{\alpha \in \mathbb{Z}^n_+ : \ c_\alpha \ne 0\}$。我们用 $\mathbb{S}^m$ 表示实对称 $m \times m$ 矩阵空间，而 $\mathbb{S}^m_+$ 是 $\mathbb{S}^m$中的半正定矩阵锥。向量空间子集 M 的线性张成记为 $\mathrm{lin}(M)$。

2.2. 设 $C \subseteq \mathbb{R}^n$ 是一个凸锥。C的对偶锥为 $C^* = \{y \in \mathbb{R}^n : \ \forall x \in C \langle x, y\rangle \ge 0\}$，其中 $\langle x, y\rangle$ 表示 $\mathbb{R}^n$上的标准内积。C 的双对偶$C^{**} := (C^*)^*$ 等于 C的闭包，即 $C^{**} = \overline{C}$。如果 $C \cap (-C) = \{0\}$，则锥 C 是尖锥。当 $C \subseteq V$是任意实向量空间 V中的一个凸锥时，

对偶锥为 $C^* := \{y \in V^\vee : \forall x \in C\ y(x) \geq 0\}$，其中 $V^\vee$ 是 V的对偶向量空间。

C 的一个面是一个凸锥 F，满足 $F \subseteq C$，并且只要 x、 $y \in C$ 和 $x + y \in F$ 成立，就有 x、 $y \in F$。对于任意 $x \in C$，都存在唯一一个包含关系意义下极小的面 F，它属于 C且满足 $x \in F$；该面称为 x 在 C中的支撑面。凸锥的一维面称为极射线。众所周知，有限维闭且尖的凸锥是其极射线的闵可夫斯基和。

2.3. 我们通过给出 [13] 第10章中讨论的一般形式，简要说明多项式优化方法背后的锥对偶。给定 $\mathbb{R}^n$ 的一个子集 S，以及有限维向量子空间V中的多项式 p、 q，其中V属于 $\mathbb{R}[x_1, \ldots, x_n]$，且在 S上满足 $q > 0$。对于 q，一种常见的选择是常数$q = 1$。利用一个闭凸锥 $C \subseteq V$，可以对在 S 上最小化商 p/q 的问题进行松弛（即给出下界）；该闭凸锥满足：对于每个 $g \in C$，在 S 上有 $g \geq 0$。于是有

$$\inf_S \frac{p}{q} \geq \sup\{\lambda \in \mathbb{R} : p - \lambda q \in C\},$$

其中上确界是与该问题对偶的锥优化问题

$$\inf\{v(p) : v \in C^*,\ v(q) = 1\}.$$

在下面的命题中，通过将 V及其对偶空间 $V^\vee$ 分别与 $\mathbb{R}^N$相识别，其中 $N := \dim(V)$，上述对偶性表述为不涉及多项式的形式。

命题 2.4. 设 N 为正整数， $C \subseteq \mathbb{R}^N$ 为闭且尖的凸锥。则对于任意 $p \in C$ 和 $q \in C \setminus \{0\}$，有

$$\sup\{\lambda \in \mathbb{R} : p - \lambda q \in C\} = \inf\{\langle p, v\rangle : v \in C^*,\ \langle v, q\rangle = 1\}. \tag{2}$$

这是锥优化问题对偶性的一个特例，参见 [7] 2.1.4 中的讨论。证明相当直接，故略去。

3. Convex hulls of monomial curves

设 $I \subseteq \mathbb{R}$ 为非退化闭区间， $V \subseteq \mathbb{R}[t]$ 为有限维线性子空间，并设 $P = \{f \in V : f \geq 0 \text{ 在 } I\}$上的一个闭且尖的凸锥，其中 V。我们首先证明，锥 P 的每个面都可由在 I中各点处的适当消失条件描述；当 I 无界时，也可由无穷远处的消失条件描述。令 $\operatorname{ord}_s(f)$ 表示 $f \in \mathbb{R}[t]$ 在 $s \in \mathbb{R}$处的消失阶。

命题 3.1. 对于 $0 \neq f \in P$，令 $W_f = \{g \in V : \forall s \in I \operatorname{ord}_s(g) \geq \operatorname{ord}_s(f)\}$。令 $U_f = \{g \in W_f : \deg(g) \leq \deg(f)\}$，如果 I 无界；如果 I 紧致，则令 $U_f = W_f$。那么 U_f 是 f的支撑面 F_f 的线性张成。特别地，$\dim(F_f) = \dim(U_f)$。

对于 P的极射线，这意味着：

推论 3.2. 假设 I 紧致，令 $f \in V$ 张成 P的一条极射线。那么 f 在 I中至少有 $\dim(V) - 1$ 个根，按重数计。

证明。根据假设，我们有 $F_f = \mathbb{R}_+ f$，因此由命题 3.1 可得 $W_f = \mathbb{R} f$。由于 W_f 由 V 中在 I 内至少具有与 f 相同的根的元素组成，因此 f 在 I 中至少有 $\operatorname{codim}(W_f) = \dim(V) - 1$ 个根，按重数计。□

在非紧情形下，结果如下：

推论 3.3。设 I 无界，设 $f \in P$ 张成 P的一条极射线。记$d = \deg(f)$，且 $V_d := \{g \in V: \deg(g) \le d\}$。那么 f在 I中至少有 $\dim(V_d) - 1$ 根，按重数计。

证明。论证与推论 3.2 相同，因为 $\mathbb{R}f$ 由 V_d中在 I 里至少有与 f相同的根的元素组成。□

命题 3.1 的证明。 f的支撑面为 $F_f = \{p \in P: \exists \gamma > 0$，其中$f - \gamma p \in P\}$。显然，若 p， $q \in P$满足 $f = p + q$，则 p， $q \in W_f$；而且当 I 无界时，$\deg(p)$、$\deg(q) \le \deg(f)$。因此 $F_f \subseteq U_f$ 成立。为证明$U_f \subseteq \operatorname{lin}(F_f)$，注意对每个 $g \in U_f$，都存在常数 $c > 0$，使得$c|g| \le f$ 在 I上成立。事实上，对每个 $s \in I$，都有常数 $c_s > 0$，使得在 s 的某个邻域 J_s 上， $c_s|g(t)| \le f(t)$。若 I 包含一条右半直线，则存在 a、 $c > 0$，使得对所有 $t \in J_\infty = [a, \infty[$，$c|g(t)| \le f(t)$；左半直线的情形同理。通过对 I 取有限子覆盖，可找到所需的常数 c。

由前述注释可知，我们可以找到 U_f 的一个线性基 $g_1, \dots, g_r$，使得 $|g_i| \le f$ 在 I 上满足 $i = 1, \dots, r$。由此可得 $f - g_i \in F_f (i = 1, \dots, r)$，因此$U_f = \operatorname{lin}(f - g_1, \dots, f - g_r, f)$ 包含于 $\operatorname{lin}(F_f)$。□

我们继续假设 $V \subseteq \mathbb{R}[t]$ 是一个有限维线性子空间，且 $\dim(V) = n + 1$。假设给定了 V 中元素在实直线上各点处的 n 个消失条件，其中允许高阶消失条件。如果这些条件在 V 上是独立的，那么以下引理给出了这些条件在 V 中本质上唯一解的显式公式：

引理 3.4。设 $p_0, \dots, p_n$ 是 V 的一个线性基。设 $\xi = (\xi_1, \dots, \xi_r)$ 是一个由两两不同的实数组成的元组，且 $b_1, \dots, b_r \ge 1$ 是满足 $\sum_{i=1}^r b_i = n$ 的整数。以下条件等价：

(i) 矩阵

$$A(t;\xi) = \begin{pmatrix} p_0(t) & \cdots & p_n(t) \\ p_0(\xi_1) & \cdots & p_n(\xi_1) \\ \vdots & & \vdots \\ p_0^{(b_1-1)}(\xi_1) & \cdots & p_n^{(b_1-1)}(\xi_1) \\ \vdots & & \vdots \\ p_0(\xi_r) & \cdots & p_n(\xi_r) \\ \vdots & & \vdots \\ p_0^{(b_r-1)}(\xi_r) & \cdots & p_n^{(b_r-1)}(\xi_r) \end{pmatrix} \tag{3}$$

大小为 $(n+1) \times (n+1)$ 且元素属于 $\mathbb{R}[t]$ 的矩阵，其行列式非零；

(ii) V的子空间 $\{f \in V: \operatorname{ord}_{\xi_i}(f) \ge b_i \text{ for } i = 1, \dots, r\}$ 维数为一。

当 (i) 和 (ii) 成立时，多项式 $\det A(t; \xi)$ 是 V中唯一（相差一个比例因子）的元素，它在 ξ_i 处的消失阶数至少为 b_i，其中 $i = 1, \dots, r$。

证明。注意，(ii) 中的维数始终为 ≥ 1。由 (3) 的下 n 行构成的矩阵 B 是线性映射 ϕ: $V \to \mathbb{R}^n$ 的矩阵，

$$p \mapsto \left(p(\xi_1), \dots, p^{(b_1-1)}(\xi_1), \dots, p(\xi_r), \dots, p^{(b_r-1)}(\xi_r)\right)$$

相对于 V 的基 $p_0, \dots, p_n$。子空间 (ii) 正是 ϕ 的核。(3) 的行列式是 V 的一个元素，并且当且仅当 B 有一个非零的 $n \times n$ 阶子式时，它才非零。这等价于 ϕ 是满射，因此也等价于 (ii)。最后一个断言显然成立，因为 $\det A(t; \xi)$ 是 V 的一个元素，它在 ξ_i 处有一个重数为 $\geq b_i$ 的零，其中 $i = 1, \dots, r$。□

3.5. 接下来我们回顾一些关于舒尔多项式的背景知识。设 $n \in \mathbb{N}$ 为一个固定的整数，设 $x = (x_0, \dots, x_n)$ 为一组不定元。我们考虑将 $\mu = (m_0, \dots, m_n)$ 分拆为 $n+1$ 个部分的分拆，其中 $m_0 \geq \cdots \geq m_n \geq 0$。一种特定的分拆为 $\delta = (n, n-1, \dots, 1, 0)$。给定如上所述的分拆 μ，行列式

$$F_\mu(x) \ := \ \det \begin{pmatrix} x_0^{m_0} & \cdots & x_0^{m_n} \\ \vdots & & \vdots \\ x_n^{m_0} & \cdots & x_n^{m_n} \end{pmatrix}$$

除非 m_i 两两不同，否则它恒等于零。在后一种情况下，$\lambda = \mu - \delta = \big(m_0 - n, \dots, m_{n-1} - 1, m_n\big)$ 是另一个分拆。显然，范德蒙德乘积

$$v(x) \ := \ \prod_{0 \leq i < j \leq n} (x_i - x_j) \ = \ F_\delta(x)$$

整除 $F_\mu(x)$。余因子是分拆 λ 的舒尔多项式 $s_\lambda(x)$。换言之，双交错公式

$$F_\mu(x) \ = \ v(x) \cdot s_\lambda(x) \tag{4}$$

成立。根据舒尔多项式的引入方式，恒等式 (4) 或者是 $s_\lambda(x)$ 的定义，或者是一个定理，参见 [19] 定理 7.15.1。

舒尔多项式 $s_\lambda(x)$ 作为变量 x_i 的多项式是对称的，具有 $|\lambda|$ 的齐次度，并且关于每个变量 x_i 的度为 λ_0。Schur 多项式具有一个显著性质：其所有系数都是非负整数。事实上，系数存在一种组合描述，参见 [19] 中的第 7.10 节。在我们的语境中，系数的整性不起作用，但非负性是一个关键性质。

3.6. 我们使用舒尔多项式，在 p_i 为单项式的情况下，推导出引理 3.4 中行列式的乘积公式。设 $\mu = (m_0, \dots, m_n)$ 是一个由不同部分组成的分拆，即满足 $m_0 > \cdots > m_n \geq 0$。对于 $i = 0, \dots, n$，将 $p_i(t) = t^{m_i}$ 写作，并将 $p_i(t)$ 的 j 阶导数记为 $p_i^{(j)}(t) := \frac{d^j}{dt^j} p(t)$ $(j \geq 0)$。此外，对于 $j \geq 0$，我们引入元组 $p(t) = \big(p_0(t) \dots, p_n(t)\big)$ 和 $p^{(j)}(t) = \big(p_0^{(j)}(t) \dots, p_n^{(j)}(t)\big)$。

令 $b = (b_0, \dots, b_r)$ 为满足 $\sum_i b_i = n + 1$ 的整数元组 $b_i \geq 1$，令 $y = (y_0, \dots, y_r)$ 为一个 $r+1$ 变量元组。我们考虑大小为 $n+1$ 的行列式 $F_{\mu,\, b}(y)$，其中对于每个 $i = 0, \dots, r$，都包含以下各行

$$p^{(j)}(y_i) \ = \ \big(p_0^{(j)}(y_i), \dots, p_0^{(j)}(y_i)\big)$$

对于 $j = 0$、1、. . .、$b_i - 1$。换句话说，令

$$F_{\mu,b}(y) \;:=\; \det \begin{pmatrix} p_0(y_0) & \cdots & p_n(y_0) \\ p_0'(y_0) & \cdots & p_n'(y_0) \\ \vdots & & \vdots \\ p_0^{(b_0-1)}(y_0) & \cdots & p_n^{(b_0-1)}(y_0) \\ \vdots & & \vdots \\ p_0(y_r) & \cdots & p_n(y_r) \\ \vdots & & \vdots \\ p_0^{(b_r-1)}(y_r) & \cdots & p_n^{(b_r-1)}(y_r) \end{pmatrix} \tag{5}$$

并再次令 $\lambda = \mu - \delta$。

命题 3.7。行列式$F_{\mu,\ b}(y)$具有乘积分解

$$F_{\mu,b}(y) \;=\; c \cdot v_b(y) \cdot s_{\lambda,b}(y)$$

其中

$$v_b(y) \;:=\; \prod_{0 \le i < j \le r} (y_i - y_j)^{b_i b_j}$$

and

$$s_{\lambda,b}(y) \;:=\; s_\lambda\big(\underbrace{y_0, \ldots, y_0}_{b_0},\ \ldots,\ \underbrace{y_r, \ldots, y_r}_{b_r}\big)$$

且 c 是一个常数，等于

$$c \;=\; \prod_{i=0}^{r} (-1)^{b_i(b_i-1)/2} \cdot (b_i - 1)!$$

证明。我们通过对双交错公式进行归纳推导，得出该断言。令 A_0 为行分别是 $p(x_0)$ . . . , $p(x_n)$ 的矩阵，因此

$$\det(A_0) \;=\; s_\lambda(x) \cdot \prod_{0 \le i < j \le n} (x_i - x_j)$$

根据 (4)。将第二行 $p(x_1)$ 替换为 $(p(x_1) - p(x_0))/(x_1 - x_0)$ 后，行列式变为原来的 $x_1 - x_0$ 分之一。现在若令 $x_1 := x_0$，则所得矩阵 A_1 的第二行为 $p'(x_0)$，其行列式为

$$\det(A_1) \;=\; -s_\lambda(x_0, x_0, x_2, \ldots, x_n) \cdot \prod_{j=2}^{n} (x_0 - x_j)^2 \cdot \prod_{2 \le i < j \le n} (x_i - x_j)$$

提取出 $x_1 - x_0$ 后，我们便从 $v(x)$ 中用掉了因子 $x_0 - x_1$，并得到一个负号。现在重复这一步。接下来，我们将 A_1 的第三行（即 $p(x_2)$）替换为

$$\frac{p(x_2) - p(x_0) - (x_2 - x_0)p'(x_0)}{(x_2 - x_0)^2}$$

将 $x_2 := x_0$ 特化后，所得矩阵 A_2 的行是

$$p(x_0),\ p'(x_0),\ \frac{1}{2}p''(x_0),\ p(x_3), \ldots$$

其行列式为

$$\det(A_2) \;=\; -s_\lambda(x_0,x_0,x_0,x_3,\dots,x_n)\cdot\prod_{j=3}^{n}(x_0-x_j)^3\cdot\prod_{3\le i<j\le n}(x_i-x_j)$$

依此类推。经过 b_0 多步后，各行变为

$$p(x_0),\ p'(x_0),\ \dots,\ \frac{1}{(b_0-1)!}p^{(b_0-1)}(x_0),\ p(x_{b_0}),\dots,p(x_n),$$

此时，我们已经加入了 $0+1+\cdots+(b_0-1)=\binom{b_0}{2}$ 个负号。

现在我们可以重复这一过程。下一步是对变量 $x_{b_0},\dots,x_{b_0+b_1-1}$进行 b_1 步操作。依此类推。最后，将保留下来的变量重新标记为 $y_0,\dots,y_r$. □

3.8. 做好这些准备后，我们进入本文的主要结果。设 $\mu=(m_0,\dots,m_n)$，其中 $m_0>\cdots>m_n\ge 0$，记作 $k=1+\lfloor\frac{n}{2}\rfloor$。设 $\mathbb{R}[t]$ 为变量 t上的多项式环。我们考虑 $\mathbb{R}[t]$ 的子空间 $V=V_\mu$，它由单项式 $t^{m_0},\dots,t^{m_n}$张成。设 $I\subseteq\mathbb{R}$ 为一个区间，并令

$$P\;:=\;\{f\in V\colon f\ge 0\text{ on }I\}.$$

注意， P 是 V中的闭凸锥。关键结果如下，即非负实零点定理:

定理 3.9。设 (a) $I=[0,1]$ 或 (b) $I=\mathbb{R}_+=[0,\infty[$。则每个 $f\in P$ 都可以写成有限和

$$\begin{aligned}&\text{(a)}\quad f=\sum_i g_i(t)^2p_i(t)+(1-t)\sum_j h_j(t)^2q_j(t),\\&\text{(b)}\quad f=\sum_i g_i(t)^2p_i(t),\end{aligned}\tag{6}$$

其中， g_i、 p_i、 h_j、 q_j是多项式，且 $\deg(g_i)$、$\deg(h_j)\le\lfloor\frac{n}{2}\rfloor=k-1$，使得每个被加项的度为 $\le m_0$，并且 p_i、 q_j 的系数均为非负。

反之，当 f 具有定理中的表示形式时，显然$f\ge 0$ 在 I上。

3.10. 为开始证明定理 3.9（无论情形 (a) 还是 (b)），令 Q 为所有具有表示式 (6) 的 $f\in V$的集合。则 Q 是一个凸锥，并且 $Q\subseteq P$。令 $E\subseteq P$为所有生成 P的极射线的多项式 f 的集合。由于锥 P 是闭且尖的， P中的每个元素都是 E中有限多个元素之和。为证明 $Q=P$，因此只需对任意给定的 $f\in E$证明 $f\in Q$。

3.11. 让我们再进行两步归约。假设对于所有单项式子空间，只要 $f\in E$ 满足 $f(0)>0$，就已证明 $f\in Q$ 成立。设 $f\in E$ 满足~$f(0)=0$。写作 $f=t^wf$，其中 $w=\mathrm{ord}_0(f)$，则存在一个指标 $0\le s\le n$~满足 $m_s=w$。子空间 $V=\mathrm{lin}(t^{m_0-w},\dots,t^{m_{s-1}-w},1)$ 属于 $\mathbb{R}[t]$，其维数为~ ~ ~$s+1\le n+1$，且包含 f。此外， $f\ge 0$ 在 I上成立，且 $f(0)>0$。为明确起见，~假设我们处于情形 (a)，因此 $I=[0,1]$。根据假设，定理对 f~和子空间 V成立。这意味着存在一个恒等式

$$\tilde f\;=\;\sum_i g_i(t)^2p_i(t)+(1-t)\sum_j h_j(t)^2q_j(t)\tag{7}$$

其中 $\deg(g_i)$、$\deg(h_j) \le \lfloor \frac{s}{2} \rfloor$，且 p_i、q_j 具有非负系数，使得 (7) 中的每一项的度均为 $\le m_0 - w$。将恒等式乘以 t^w，我们便得到 f 的一个恒等式，正如定理 3.9(b) 所述。在情形 (b) 中，论证完全相同。

3.12. 因此，我们需要为每个 $f \in E$（其中 $f(0) > 0$）建立恒等式 (6)。为此，显然可以舍弃所有次数大于 $d = \deg(f)$ 的单项式。换言之，我们可以假设 $\deg(f) = m_0$。这一步约化仅在情形 (b) 中发挥作用。（事实上，如果 $I = [0,1]$，则 $\deg(f) = m_0$ 是自动成立的，这一点将在下面的证明中看到。）

3.13. 根据笛卡尔符号法则（例如 [11] 推论 1.10.3），每个非零 $f \in V$ 至多有 n 个严格正根，按重数计。另一方面，推论 3.2 和 3.3 表明，每个 $f \in E$（其 $\deg(f) = m_0$）在 I 中至少有 n 个根。如果进一步有 $f(0) > 0$，则可知 f 恰好有 n 个严格正根，并且它们全部位于 I 中。

3.14. 设 $f \in E$ 满足 $f(0) > 0$，且 $\deg(f) = m_0$。注意，现在 $m_n = 0$，因为 $f(0) \neq 0$。由 3.13 可知，f 恰有 n 个正根，且它们都位于 I 中。只需证明每个这样的 f 都满足恒等式 (6)。

设 $\xi_1, \dots, \xi_r$ 是 f 的不同正根，并令 $b_i = \operatorname{ord}_{\xi_i}(f)$ $(i = 1, \dots, r)$。除情况 (a) 中 $\xi_i = 1$ 外，每个 b_i 都是偶整数。由于 f 在 I 中有 n 个根，因此 $\sum_{i=1}^{r} b_i = n$ 成立。

考虑 3.6 中的行列式 (5)，其中 $b_0 := 1$，且 $p_i = t^{m_i}$ $(i = 0, \dots, n)$。代入 $y_0 = t$ 和 $y_i = \xi_i$ $(i = 1, \dots, r)$ 后，命题 3.7 表明该行列式可分解为

$$F_{\mu,b}(t, \xi_1, \dots, \xi_r) \;=\; \gamma \prod_{i=1}^{r} (t - \xi_i)^{b_i} \cdot s_\lambda\big(t, \underbrace{\xi_1, \dots, \xi_1}_{b_1}, \;\dots,\; \underbrace{\xi_r, \dots, \xi_r}_{b_r}\big)$$

其中 $\gamma \neq 0$ 是一个常数。最后一个因子 $s_\lambda(t, \xi_1, \dots, \xi_r)$ 是关于 t 的多项式。由于舒尔多项式 s_λ 的所有系数都是 ≥ 0，且对所有 i 都有 $\xi_i > 0$，因此该多项式不恒为零（且其度为 $\lambda_0 = m_0 - n$）。因此，行列式非零，而引理 3.4 表明它与 f 相差一个比例因子。这意味着 f 可以分解为

$$f \;=\; \gamma' \prod_{i=1}^{r} (t - \xi_i)^{b_i} \cdot s_\lambda(t, \underbrace{\xi_1, \dots, \xi_1}_{b_1}, \dots, \underbrace{\xi_r, \dots, \xi_r}_{b_r}) \tag{8}$$

其中有一个常数 $\gamma' \neq 0$。第一个因子的形式为

$$\prod_{i=1}^{r} (t - \xi_i)^{b_i} \;=\; g(t)^2 u(t)$$

其中 $\deg(g) = k - 1 = \lfloor \frac{n}{2} \rfloor$ and $u(t) = 1$ 或 $t - 1$。仅当情形 (a)（$I = [0,1]$）中 n 为奇数时，才会出现情形 $u(t) = t - 1$。(8) 中的最后一个因子 $s_\lambda(t, \xi_1, \dots)$ 是关于 t 的一个具有非负系数的多项式。因此，除去一个非零常数因子后，右端具有定理 3.9 所述的形式。由于 $f|_I \ge 0$，我们看到，在情形 (a) 中 n 为偶数时，或在情形 (b) 中，$\gamma' > 0$。当情形 (a) 中 n 为奇数时，对于某个 i，有 $\xi_i = 1$，且 b_i 为奇数，因此在这种情形下 $\gamma' < 0$。定理得证。□

令 $\Sigma = \Sigma_{2k} := \{g \in \mathbb{R}[t]\colon \deg(g) \leq 2k,\ g \geq 0 \text{ 在 } \mathbb{R}\}$ 上。每个 $g \in \Sigma_{2k}$ 都可以写成 $g = g_1^2 + g_2^2$，其中 g_1、g_2 是度为 $\leq k$ 的多项式。定理 3.9 还可以表述为以下等价形式：

推论 3.15。再次令 (a) $I = [0,1]$ 或 (b) $I = \mathbb{R}_+$。在定理 3.8 的假设下，包含关系

$$\text{(a)}\ P \subseteq \Big(\Sigma + t\Sigma + \cdots + t^{m_0-2k}\Sigma\Big) + (1-t)\Big(\Sigma + t\Sigma + \cdots + t^{m_0-2k-1}\Sigma\Big)$$

or

$$\text{(b)}\ P \subseteq \Big(\Sigma + t\Sigma + \cdots + t^{m_0-2k}\Sigma\Big)$$

分别成立。

注 3.16。推论 3.15 的 (b) 可视为一元多项式的稀疏（或少项式）正性定理（更确切地说，非负定理）：每个在 $\mathbb{R}_+$上非负且具有 m 个单项式的多项式 $f \in \mathbb{R}[t]$ 都可以表示为一个（有限）和

$$f = \sum_{i \geq 0} t^i \big(p_i(t)^2 + q_i(t)^2\big)$$

其中 p_i、q_i 是度为 $\leq \lfloor \frac{m-1}{2} \rfloor$的多项式。下一节将更详细地拓展这一观点。

注 3.17。如前所述，令 $V \subseteq \mathbb{R}[t]$ 为由 $n+1$个单项式生成的线性子空间，并令 $P = \{f \in V\colon\ f \geq 0 \text{ 在 } I\}$ 上，其中 (a) $I = [0,1]$ 或 (b) $I = \mathbb{R}_+$。无论哪种情况，推论 3.15 都允许我们直接得到 P的分块半定表示，其分块大小至多为 $k+1$。令 $\varphi\colon \mathsf{S}^{k+1} \to \mathbb{R}[t]_{2k}$为线性映射，将对称矩阵 $M = (a_{ij})_{0 \leq i,\ j \leq k}$ 映射为

$$\varphi(M) := (1, t, \ldots, t^k) \cdot M \cdot (1, t, \ldots, t^k)^\top = \sum_{i,j=0}^{k} a_{ij} t^{i+j}$$

那么 $\varphi(\mathsf{S}_+^{k+1}) = \Sigma_{2k}$。在情形 (a) 中，考虑如下线性映射

$$\phi\colon \big(\mathsf{S}^{k+1}\big)^{2(m_0-2k)+1} \to \mathbb{R}[t]$$

其定义为

$$\big(M_0, \ldots, M_{m_0-2k};\, N_0, \ldots, N_{m_0-2k-1}\big) \mapsto \sum_{i=0}^{m_0-2k} t^i \varphi(M_i) + (1-t) \sum_{j=0}^{m_0-2k-1} t^j S\varphi(N_j)$$

对于 (b)，考虑线性映射 $\phi\colon \big(\mathsf{S}^{k+1}\big)^{m_0-2k+1} \to \mathbb{R}[t]$，其定义为

$$\big(M_0, \ldots, M_{m_0-2k}\big) \mapsto \sum_{i=0}^{m_0-2k} t^i \varphi(M_i)$$

无论哪种情形，根据推论 3.15，我们都有 $P = V \cap \mathrm{im}(\phi)$。这是 P 的一个显式分块半定表示，其块大小为 $k+1$。

注 3.18. 对于 f中在整个实轴上非负的稀疏多项式 $\mathbb{R}[t]$，通常不存在类似于稀疏分解

$$f = \sum_{i \in 2\mathbb{Z}_+} t^i \left(p_i(t)^2 + q_i(t)^2\right)$$

注 3.16 中所述的那种分解。例如， $f = 3t^4 - 4t^3 + 1 = (t-1)^2(3t^2+2t+1)$在 $\mathbb{R}$ 上非负，但无法写成 $f = g_1(t) + t^2 g_2(t)$，其中 g_1、g_2 均为线性多项式的平方和。

不过，根据 3.17 中的情形，我们很容易为 $\mathbb{R}$}上的 $f \geq 0$ 构造一个块大小为 $k+1$的分块半定表示： $P = \{f \in V$。只需注意到， $f(t)$ 在 $\mathbb{R}$上成立，等价于 $f(t)$ 和 $f(-t)$ 均在 $\mathbb{R}_+$上成立。

由锥对偶性，我们得到关于单项式空间曲线凸包的所需定理：

定理 3.19。设 $m_1,\ldots,m_n \geq 1$ 为整数，且 $I \subseteq \mathbb{R}$ 是一个半代数集。集合 K 的闭凸包

$$S = \left\{(t^{m_1}, \ldots, t^{m_n}) \colon t \in I\right\}$$

在 $\mathbb{R}^n$ 中具有至多为 $\lfloor \frac{n}{2} \rfloor + 1$的半正定扩展次数。

证明。集合 K 被视为 P 的对偶锥 P^* 的一个仿射线性截面。由于 $\operatorname{sxdeg}(P^*) = \operatorname{sxdeg}(P)$（[17] 命题 1.7），只需证明 $\operatorname{sxdeg}(P) \leq \lfloor \frac{n}{2} \rfloor + 1$。我们不妨假设 m_i 两两不同。当 $I = [0,1]$ 或 $I = \mathbb{R}_+$ 时，命题 $\operatorname{sxdeg}(P) \leq \lfloor \frac{n}{2} \rfloor + 1$ 已在注 3.17 中证明。我们简述如何将其他集合 I 的情形实质上归约到这两种情形，而不详述全部细节。显然，可以假设 I 是闭集。若 $I = I_1 \cup I_2$ 且 I_1、I_2 为半代数集，根据 [17]命题 1.6，只需分别证明 I_1 和 I_2 的情形。这样，我们便归约为考察 $I = [a,b]$ 或 $I = [a,\infty[$，其中 $0 \leq a < b < \infty$。情形 $a = 0$ 已经处理完毕（在紧致情形下假设 $b = 1$ 并非关键）。设 $I = [a,b]$ 且 $0 < a < b < \infty$。那么，每个生成 P 的极射线的 $f \in P$ 在 I 中恰有 n 个根，按重数计。因此，定理 3.9 的证明中的归约步骤 3.11 并不需要。否则，我们只需沿用该定理的证明。这样，我们便得到 P 中每个元素的一种表示，其形式类似于 3.9(a)，但权重为 1、$t-a$、$b-t$ 和 $(t-a)(b-t)$，而不只是 1 和 $1-t$。当 $I = [a,\infty[$ 且 $a > 0$ 时，我们可以用类似的方法处理。□

注记 3.20。对于 $I = [0,1]$ 或 $I = \mathbb{R}_+$，可以通过对偶化，利用注记 3.17、3.18 得到 K 的一个块大小为 $k = \lfloor \frac{n}{2} \rfloor + 1$ 的显式分块半定表示。

注记 3.21。自然会问，半正定扩张次数 $\operatorname{sxdeg}(K) \leq \lfloor \frac{n}{2} \rfloor + 1$ 的界是否可以推广到比单项式曲线的凸包更一般的情形。事实上，只要 $K \subseteq \mathbb{R}^n$ 是 $\mathbb{R}^n$ [18]中一维半代数集的闭凸包，该界一般都成立。然而，与单项式情形相比，证明要困难得多。即使曲线由多项式而非单项式参数化，也必须在曲线上足够小的区间上进行局部论证，而且似乎不存在分块半定表示的显式形式。当曲线不是有理曲线时，证明在技术上会变得更加复杂。

4. Sparse non-negativity certificates in polynomial optimization

4.1. 余正少项式之和与稀疏半正定松弛。我们将重新审视推论 3.15，重点关注其中的 (b) 部分。若单变量多项式 $f \in \mathbb{R}[t]$ 是一个 k项式，则称其为少项式；条件是 f 为线性组合

至多包含 k 个单项式 t^i。若 f 在 $\mathbb{R}_+$上满足 $f \geq 0$，则称其为共正。给定有限集 $J \subseteq \mathbb{Z}_+ := \{0, 1, 2, \dots\}$，令

$$\mathrm{CP}(J) := \{p \in \mathbb{R}[t] : p \geq 0 \text{ on } \mathbb{R}_+ \text{ and } \mathrm{supp}(p) \subseteq J\}.$$

这是第 3 节中考虑的锥 P，对应于 J的单项式曲线。令

$$\mathrm{SOCF}_{k,d} := \sum_{\substack{J \subseteq \{0,\dots,d\}: \\ |J|=k}} \mathrm{CP}(J)$$

缩写 SOCF 代表余正少项式之和。因此，$f \in \mathbb{R}[t]_d$ 属于 $\mathrm{SOCF}_{k,\,d}$ 当且仅当 f 可以表示为有限个余正 k项式之和，且其度至多为 d。显然，可以将这一设置从单变量多项式推广到多变量情形。未来探索这一设置可能会很有趣。

如前所述，令 $\Sigma_{2k} \subseteq \mathbb{R}[t]$ 表示次数为$\leq 2k$的平方和多项式集合。第 3 节的主要结果蕴含了锥 SOCF 的如下正性定理：

推论 4.1。对于所有满足 $d > 2k$ 的整数 k、$d \geq 1$，有

$$\mathrm{SOCF}_{2k+1,d} = t^0 \Sigma_{2k} + \cdots + t^{d-2k} \Sigma_{2k}. \tag{9}$$

此外，$\mathrm{sxdeg}(\mathrm{SOCF}_{2k+1,\,d}) = k + 1$。

对于 $d = 2k$，注意到 $\mathrm{SOCF}_{2k+1,\,2k}$ 恰好是所有次数不超过 $2k$ 的共正单变量多项式所构成的锥，即它是该锥

$$\{f \in \mathbb{R}[t]_{2k} : f \geq 0 \text{ on } \mathbb{R}_+\} = \Sigma_{2k} + t\,\Sigma_{2k-2},$$

其中，单变量情形下共正性的后一种描述是众所周知的。

证明。包含关系“$\subseteq$”成立，这是因为对于每个满足 $|J| \leq 2k+1$ 的集合 $J \subseteq \{0, \dots, d\}$，锥 $\mathrm{CP}(J)$ 都包含于 (9) 式右端（推论 3.15(b)）。反向包含关系显然成立。最后一个结论由恒等式 (9) 结合引理 1.4(d) 得出。□

为描述 SOCF 锥的对偶锥，我们通过 $v_i = v(t^i)$ 将 $v \in (\mathbb{R}[t]_d)^\vee$ 和 $(v_0, \dots, v_d) \in \mathbb{R}^{d+1}$ 进行对应。

引理 4.2. 设 $d, k \geq 1$ 满足 $d > 2k$。$\mathrm{SOCF}_{2k+1,\,d}$ 的对偶是锥

$$\big\{v = (v_0, \dots, v_d) \in \mathbb{R}^{d+1} \colon M_{0,k}(v) \succeq 0, \dots, M_{d-2k,k}(v) \succeq 0\big\},$$

其中

$$M_{s,k}(v) := \big(v_{s+i+j}\big)_{i,j=0,\dots,k} \in \mathbb{R}^{(k+1)\times(k+1)}.$$

证明. 由推论 4.1 中的 (9) 可得，因为 $(C_1 + C_2)^* = C_1^* \cap C_2^*$，且因为 $(\Sigma_{2k})^* = \{u = (u_0, \dots, u_k) \colon M_{0,\,k}(u) \succeq 0\}$；例如参见 [7] 第 4.6 节。□

引理 4.2 中的记号 $M_{s,\,k}(v)$ 用于下面的命题，该命题为问题 $\min_{\mathbb{R}_+} f$ 提供原始和对偶稀疏松弛，其中 $f \in \mathbb{R}[t]_d$，并可灵活选择稀疏度阈值 $2k + 1$。

命题 4.3. 设 k、d 为正整数，且 $2k < d$ 和 $f \in \mathbb{R}[t]_d$。那么 $\mathbb{R} \cup \{\pm\infty\}$ 中的以下元素相同：

(1) $\sup\{\lambda \in \mathbb{R} \colon f - \lambda \in \mathrm{SOCF}_{2k+1,d}\}$,

(2) $\sup\{\lambda \in \mathbb{R} \colon f - \lambda \in t^0\,\Sigma_{2k} + \cdots + t^{d-2k}\,\Sigma_{2k}\}$,
(3) $\inf\{\langle f, v\rangle \colon v = (v_0, \ldots, v_d) \in \mathbb{R}^{d+1},\ v_0 = 1,\ M_{0,k}(v), \ldots, M_{d-2k,k}(v) \succeq 0\}$.

证明. 由推论 4.1 中的式 (9)，结合命题 2.4，即可得出此结论。 □

注 4.4. 对于情形 $k = 0$，命题 4.3 中的松弛是平凡的，其最优值为 $f(0)$ 或 $-\infty$，具体取决于 $f - f(0)$ 的所有系数是否均为非负。但即使是最简单的非平凡情形 $k = 1$，也展现出与当前关于多项式优化中稀疏松弛的研究之间的联系，具体而言，是与依赖所谓 SAGE 多项式和 SONC 多项式的松弛之间的联系；分别参见 [8] 和 [14, 10], 。这两类多项式都源于将同一思想分别应用于在 $\mathbb{R}^n_+$ 和 $\mathbb{R}^n$ 上的非负性。SAGE 多项式是由若干多项式相加得到的，这些多项式在 $\mathbb{R}^n_+$ 上非负，且其支撑包含于某个单纯形回路中。这里，单纯形回路是指一个包含极小的仿射相关集合，其凸包是一个单纯形。由于 $\mathbb{R}$ 中的单纯形回路只是含有三个元素的集合，我们可知，锥 $\mathrm{SOCF}_{3,\ d}$ 是所有次数不超过 d 的单变量 SAGE 多项式构成的锥。因此，情形 $k = 1$ 给出了 $\min_{\mathbb{R}_+} f$ 的原始和对偶 SAGE 松弛。例如，参见 [10]，其中讨论了 SONC 锥和 SAGE 锥的对偶性及其在优化中的应用。我们还注意到，推论 4.1 的特殊情形 $k = 1$ 相当于用所谓的约化回路来描述非负 SAGE 多项式；参见 [10]。也就是说，在单变量情形下，共正（$2k + 1$）项式情境中的约化回路的对应概念，是一组 $2k + 1$ 个连续整数值，因为 $t^i\Sigma_{2k}$ 中多项式的支撑是 $\{i, \ldots, i + 2k\}$ 的子集。

我们强调，SAGE 和 SONC 多项式的稀疏度阈值受维数约束，因为在 n 维中，一个电路的元素数不可能超过 $n+2$。与此相反，在我们的设定中，稀疏度阈值 $2k+1$ 可以任意变化，从而将 SAGE 松弛与 Lasserre 的稠密标准松弛连接起来。尽管我们意识到我们的结果仅限于单变量情形，但我们希望这些结果能为任意维数n中的类似研究提供启发。也就是说，理解具有可变稀疏度阈值的 SOS 和矩松弛的性质将会很有意义。

注 4.5。上述关于 $\mathbb{R}_+$上稀疏多项式优化的讨论，也可用于处理任意区间$I \subseteq \mathbb{R}_+$上的稀疏多项式优化。例如，当 $2k < d-1$ 且 $I = [0, 1]$时，鉴于推论 3.15(a) 和 4.1，次数至多为 d、并且在 $[0, 1]$ 上非负的 $(2k + 1)$-项式之和所构成的锥可描述为 $\mathrm{SOCF}_{2k+1,\ d} + (t - 1)\,\mathrm{SOCF}_{2k,\ d-1}$。对于所涉及的两个 SOCF 锥，我们都有大小为 $k + 1$的半正定描述。这导出了优化问题 $\inf_{t\in[0,1]} p(t)$ 的一个半正定规划形式，其中 p为具有 $|\mathrm{supp}(p)| \leq 2k + 1$ 的多项式，使用了 $\mathcal{O}(d)$ 个大小为 $k + 1$的线性矩阵不等式，类似于命题 4.3 中给出的规划形式 (3)。

4.2. 具有弦图稀疏性的稀疏半正定松弛。我们考虑图 $G = (V, E)$，其中 $|V| < \infty$ 以及边 $e \in E$是 V的二元子集。G 中长度为 k 的环，是 G 的一个连通子图，具有 k 个节点，且每个节点的度均为二。G 中环 C 的一条弦，是 G 的一条边，它连接 C 中的两个节点，但不是 C中的一条边。如果图 G 中的每个长度至少为四的环 G 都在 G 中具有一条弦，则称图为弦图。图 $G = (V, E)$ 的一个团 $W \subseteq V$ 是这样一组节点：其中任意两个不同节点 W 之间都由一条

边相连。以下结果给出了具有弦稀疏模式的半正定矩阵锥的一种便捷表示：

定理 4.6（参见 [2] 和 [20] 第 9.2 节）。设 $G = (V, E)$ 是一个具有 $n = |V| < \infty$ 的图。我们将锥

$$\mathsf{S}^V_{+,E} := \left\{(a_{ij})_{i,j\in V} \in \mathsf{S}^n_+ \,:\, a_{ij} = 0 \text{ when } i \neq j \text{ and } \{i,j\} \notin E\right\}.$$

若 G 是弦图，则此锥可分解为

$$\mathsf{S}^V_{+,E} = \mathsf{S}^{V_1}_+ + \cdots + \mathsf{S}^{V_N}_+,$$

其中 $V_1, \ldots, V_N \subseteq V$ 是 G 的所有包含极大的团，且对于 $W \subseteq V$，

$$\mathsf{S}^W_+ := \left\{(a_{ij})_{i,j\in V} \in \mathsf{S}^n_+ \,:\, a_{ij} = 0 \text{ for } (i,j) \notin W \times W\right\}.$$

注 4.7. $\mathsf{S}^V_{+,E}$ 是所谓的具有 $G = (V, E)$ 稀疏模式的半正定矩阵锥。若 G 是弦图，则已知 $\mathsf{S}^V_{+,E}$ 具有良好的理论和计算性质，因此可以比针对 S^n_+ 的优化问题更高效地求解针对锥 $\mathsf{S}^V_{+,E}$ 的锥优化问题。例如，在 $\mathsf{S}^V_{+,E}$ 中的矩阵所对应的图 $G = (V, E)$ 为弦图时，可以更高效地执行 Cholesky 分解，因此在 $\mathsf{S}^V_{+,E}$ 上计算障碍函数也更高效。更多详情请参见 [23] 附录 A。

注 4.8. 由定理 4.6 可知，对于弦图$G = (V, E)$，$\mathsf{S}^V_{+,E}$ 的半正定扩展次数等于 G的团数，即 G的最大团大小。

对于满足 $k < n$的正整数 n 和 k，锥

$$\mathsf{S}^n_{+,k} = \left\{(a_{ij})_{i,j=0,\ldots,n-1} \in \mathsf{S}^n_+ \,:\, a_{ij} = 0 \text{ for } |i-j| > k\right\}$$

是图 $G = (V, E)$对应的锥 $\mathsf{S}^V_{+,E}$，其顶点集为 $V = \{0, \ldots, n-1\}$，边集为 E，由满足 $0 < |i-j| \leq k$的 $\{i, j\}$ 构成。在 [20] 第 8.2 节中，$\mathsf{S}^n_{+,k}$称为带宽为$2k+1$的带状稀疏模式半正定矩阵锥。显然，定义 $\mathsf{S}^n_{+,k}$ 的图 G 是弦图，而 G的极大团是满足 $0 \leq i < n-k$的 $\{i, \ldots, i+k\}$。我们将展示，与 S^n_+ 用于表示平方和锥和截断二次模的方式类似，锥 $\mathsf{S}^n_{+,k}$ 可用于表示锥 $\mathsf{SOCF}_{2k+1,n}$。也就是说，命题 4.3 中的优化问题 (1) 可表述为关于锥 $\mathsf{S}^n_{+,k}$的锥问题。若求解器能够利用弦稀疏性，这类表述可能具有计算优势。

命题 4.9. 设 d、k 为满足 $2k < d$ 的正整数。那么，锥 $\mathsf{SOCF}_{2k+1,\ d}$可表示为 $\mathsf{S}^{e+1}_{+\ k}$，$\times \mathsf{S}^{e}_{+,\ k}$ 的线性像（若 $d = 2e$ 为偶数），以及 $\mathsf{S}^{e+1}_{+\ k} \times \mathsf{S}^{e+1}_{+\ k}$，的线性像（若 $d = 2e+1$ 为奇数）。

证明。我们只考虑偶数 d 且满足 $d = 2e$的情形，因为奇数 d 的情形完全类似。我们使用注 3.17 中的线性映射 $\varphi : \mathsf{S}^{e+1} \to \mathbb{R}[t]_d$且满足 $\varphi(A) := \sum_{i,\ j=0}^{e} a_{ij} t^{i+j}$ 对于 $A = (a_{ij})_{i,\ j=0,\ \ldots\ldots e} \in \mathsf{S}^{e+1}$。如注 3.17 所述，$\varphi(\mathsf{S}^W_+) = \Sigma_{2k}$ 对于 $W = \{0, \ldots, k\}$，其中符号 S^W_+ 借自定理 4.6。后者推出 $\varphi(\mathsf{S}^{V_i}_+) = t^{2i}\Sigma_{2k}$ 对于 $V_i := \{i, \ldots, i+k\}$ 且满足 $0 \leq i \leq d-2k$。因此，将表示

$\mathrm{SOCF}_{2k+1,\ d}$在 (9) 中的和按照 $t^0,\dots,t^{d-2k}$中指数的奇偶性分成两部分，我们得到

$$\begin{aligned}\mathrm{SOCF}_{2k+1,d} &= \sum_{i=0}^{e-k} t^{2i}\Sigma_{2k} + t\sum_{i=0}^{e-k-1} t^{2i}\Sigma_{2k}\\ &= \phi\Big(\underbrace{\sum_{i=0}^{e-k}\mathsf{S}_+^{V_i}}_{=:K}\Big) + t\,\phi\Big(\underbrace{\sum_{i=0}^{e-k-1}\mathsf{S}_+^{V_i}}_{=:L}\Big).\end{aligned}$$

因此，$\mathrm{SOCF}_{2k+1,\ d}$是 $K\times L$ 在映射 $(A,B)\ \mapsto\varphi(A)\ +t\,\varphi(B)$下的线性像。根据定理 4.6，$K$ 和 L 分别是 $\mathsf{S}_{+\,k}^{e+1}$，和 $\mathsf{S}_{+,\ k}^{e}$的副本。这便得到了所需的断言。□

备注 4.10。Wang、Magron 和 Lasserre [21, 22] 建议使用弦图 $G=(V,E)$对应的锥 $\mathsf{S}_{+,\ E}^{V}$来近似稀疏平方和多项式。对于给定的有限集 $A\subseteq\mathbb{Z}_+^n$，他们考虑锥

$$\Sigma(A) := \{f\in\mathbb{R}[x_1,\dots,x_n]\ :\ f \text{ sos and } \operatorname{supp}(f)\subseteq A\}$$

他们提出了两种迭代算法（一种在 [22] 中，另一种在 [22]中），以 A 为输入，生成一个弦图 $G=(V,E)$，其中包含 $V\subseteq\mathbb{Z}_+^n$，从而可以使用相应锥 $\mathsf{S}_{+,\ E}^{V}$ 在映射下的像

$$\varphi:(a_{\alpha,\beta})_{\alpha,\beta\in V}\mapsto\sum_{\alpha,\beta\in V}a_{\alpha,\beta}x^{\alpha+\beta}$$

作为 $\Sigma(A)$ 的近似。 [21] 中的算法生成一个 $G=(V,E)$，它是若干团的不相交并。 [21] 中的图 $G=(V,E)$ 保证满足等式

$$\Sigma(A)=\big\{f\in\phi(\mathsf{S}_{+,E}^{V})\ :\ \operatorname{supp}(f)\subseteq A\big\} \tag{10}$$

（见 [21]中的定理 3.3），但无法保证 G 的团数较小，或者更确切地说，G 的团数与 A 的性质之间的依赖关系仍有待探究。由于由 [21] 中的算法生成的图 $G=(V,E)$ 可能包含较大的团，因此，[22] 中提出了另一种方法，该方法生成边数较少的图$G=(V,E)$。这种方法之所以是启发式方法，是因为对于等式 $\Sigma(A)=\phi(\mathsf{S}_{+,E}^{V})$ 没有理论保证（见 [22]中的例 3.5）。研究锥 $\Sigma(A)$ 的半正定扩展次数，并尝试以非启发式的方式将这些锥与锥 $\mathsf{S}_{+,E}^{V}$ 联系起来，将会很有意义。

References

[1] A. A. Ahmadi, A. Majumdar：DSOS 和 SDSOS 优化：平方和与半定优化的更易处理替代方案。SIAM J. Appl. Algebra Geom. 3, 193–230 (2019). [2] J. Agler, W. Helton, S. McCullough, L. Rodman：具有给定稀疏模式的半正定矩阵。Linear Algebra Appl. 107, 101–149 (1988). [3] F. Alizadeh：半定规划中的内点法及其在组合优化中的应用。SIAM J. Optim. 5, 13–51 (1995). [4] G. Averkov：多项式优化半定方法中线性矩阵不等式的最优规模。SIAM J. Appl. Algebra Geo m. 3, 128–151 (2019). [5] G. Averkov, B. Peters, S. Sager：通过单项式模式凸化多项式优化问题。预印本，arXiv:1901.05675。

[6] A. Ben-Tal, A. Nemirovski：现代凸优化讲义：分析、算法与工程应用。SIAM, MPS, 费城 PA, 2001. [7] G. Blekherman, P. A. Parrilo, R. R. Thomas（编）：半定优化与凸代数几何。SIA M, MOS, 费城 PA, 2013.[8] M. Dressler, S. Iliman, T. de Wolff：非负回路多项式之和的正性定理。SIAM J. Appl. Algebra Geom. 1, 536–555 (2017). [9] H. Fawzi：用二阶锥表示正半定锥。Math. Program. 175 109–118 (2019). [10] L. Katth¨an, H. Naumann, T. Theobald：SAGE 与 SONC 多项式及其对偶理论的统一框架。Math. Comp. 90, 1297–1322 (2021). [11] M. Knebusch, C. Scheiderer： *Real Algebra—a First Course*。Universitext, Springer, 2022. [12] J. B. Lasserre：利用多项式进行全局优化与矩问题。SIAM J. Optim. 11 796–817 (2000/01). [13] J. B. Lasserre： *An Introduction to Polynomial and Semi-algebraic Optimization*. 剑桥大学出版社，剑桥, 2015. [14] R. Murray, V. Chandrasekaran, A. Wierman：牛顿多面体与相对熵优化。Found. Comput. Math. 21, 1703–1737 (2021). [15] Y. Nesterov, A. Nemirovskii: *Interior-point Polynomial Algorithms in Convex Programming*. **13**, 1994

SIAM Studies Appl. Math. SIAM, 费城, PA, .[16] M. Putinar：紧半代数集上的正多项式。Indiana Univ. Math. J. 42, 969–984 (1993). [17] C. Scheiderer：平面凸集的二阶锥表示。SIAM J. Appl. Algebra Geom. 5, 114–139 (2021). [18] C. Scheiderer：进行中的工作，2023. [19] R. Stanley： *Enumerative Combinatorics*，第 2 卷。Cambridge Studies in Advanced Mathematics 62, Cambridge University Press, 剑桥, 1999.[20] L. Vandenberghe, M. S. Andersen：弦图与半定优化。Found. Trends Optim. 1, 241–433 (2015). [21] J. Wang, V. Magron, J. B. Lasserre：TSSOS：利用项稀疏性的矩-SOS 层级。SIAM J. Optim. 31, 30–58 (2021). [22] J. Wang, V. Magron, J. B. Lasserre：Chordal-TSSOS：利用弦扩展的项稀疏性矩-SOS 层级。SIAM J. Optim. 31, 114–141 (2021). [23] Y. Zheng, G. Fantuzzi, A. Papachristodoulou, A. Wynn：用于具有弦稀疏性的半定规划的快速 ADMM。2017 年美国控制会议（ACC），IEEE，第 3335-3340 页。

G. Averkov, Brandenburgische Technische Universität Cottbus-Senftenberg, Platz der Deutschen Einheit 1, D-03046, Germany*Email address*: averkov@b-tu.de

C. Scheiderer, Fachbereich Mathematik und Statistik, Universität Konstanz, D-78457 Konstanz, Germany
Email address: claus.scheiderer@uni-konstanz.de