arXiv:2303.05081v1 [math.AG] 9 Mar 2023

奇异点集上的平方和表示

Ngoc Hoang Anh Mai* Victor Magron*†

2023年3月10日

摘要

刻画实代数簇 V 上的实多项式 f 何时可表示为多项式的平方和这一问题，可追溯至希尔伯特在 [20]年的开创性工作。本文聚焦于 f 在 V 上的实零点为 V的奇异点的情形，研究这一问题。通过运用最优性条件和不可约分解，我们对多项式优化中的以下关键问题给出了肯定答案：

Are there always exact semidefinite programs to compute the minimum value attained by a given polynomial over a given real algebraic variety? 我们的答案意味着，作为具有代数结构的凸规划与非凸规划之间桥梁的 Lasserre层级，不仅在一般位置情形下，而且在一般情形下都具有有限收敛性。因此，我们以构造性方式证明了 *each hyperbolic program is equivalent to a semidefinite program*。

关键词：平方和；非负定理；梯度理想；奇异轨迹；多项式优化；Karush–Kuhn–Tucker 条件；半定规划

目录

1 引言 2 2 预备知识 5

2.1 实代数簇.5 2.2 正则点集与奇异点集.6 2.3 一阶最优性条件.7 2.4 半代数集.9 2.5 平方和.10 2.6 正则点集上的非负定理.10 2.7 一般位置情形下的精确多项式优化 13

3 奇异点集的分解 15 4 奇异点集上的非负表示定理 17 5 高阶最优性条件 18

*康斯坦茨大学；德国康斯坦茨 D-78464。
†图卢兹大学；LAAS；法国图卢兹 F-31400。

1 引言

半定规划、正性表示定理与多项式优化。半定规划是凸优化的一个分支，由 Bellman 和 Fan 于 20 世纪 60 年代初发展起来 [6]。其目标是在正半定矩阵锥与一个仿射空间的交集上最小化线性目标函数。半定规划的一项重要应用是对一类具有代数结构的非凸规划进行松弛，这类规划称为多项式优化问题。这些问题的目标函数和约束函数均为多项式。

2001 年，Lasserre 在 [28] 中引入了一种恰当的半定规划层级，它能返回一列数值，并近似收敛到给定多项式优化问题的最优值。为此，他利用了正定性定理，即在基本半代数集上为正的多项式的表示；基本半代数集是由有限个多项式不等式构成的系统所定义的集合。例如，保证 Lasserre 层级收敛的 Putinar 正性证书定理 [48], 指出，每个在满足所谓阿基米德条件的紧基本半代数集 S上为正的多项式，都可以分解为定义 S的多项式的线性组合，其系数是多项式的平方和。

精确半定规划与非负性定理。我们的目标是开发精确半定规划，以确定给定多项式优化问题的最优值。在此语境下，“精确半定规划”一词是指：给定一个多项式优化问题的目标多项式和约束多项式，我们可以通过算法构造一个半定规划，使其最优值与原问题的最优值完全相等。获得此类规划意义重大，因为它使我们能够利用凸优化求解非凸问题。此外，通常很难界定凸松弛与原问题之间的差距，Goemans-Williamson 的最大割算法 [15]就是如此。

为了构造此类精确半定规划，我们需要建立非负定理，用以表示在基本半代数集上非负的多项式。Marshall 在 [37, 38] 中证明了一个与 Putinar 的定理形式相同的非负定理，其成立条件是所谓的边界 Hessian 条件。基于此，Nie 在 [45] 中证明，在与二阶最优性条件相关的一般性假设下，Lasserre层级有限收敛。Marshall 和 Nie 的研究依赖于 Scheiderer 在 [49–51] 中提出的局部-整体原理。Scheiderer 在那里证明了 Schm¨udgen 正定性定理的非严格扩展 [52]。该定理指出，在紧致基本半代数集 S 上为正的每个多项式f ，都可以表示为定义 S 的多项式之乘积的线性组合，其中权重为多项式的平方和。在证明中，Scheiderer 要求 S 上 f 的实零点满足有限性假设。关于 Burgdorf、Scheiderer 和 Schweighofer 提出的一些扩展，我们请读者参阅 [8] ；这些扩展允许去除实零点的有限性假设

同时仍保持 S 上 f 的紧致性。在强正则性假设下，Nie 在其研究 [43] 中利用输入多项式的雅可比矩阵，得到了精确半定规划。

基于最优性条件的非负定理。我们将半代数集 S 视为实流形。若点 a 属于 S，且 S 在 a 处的切空间维数小于（分别为等于） S的维数，则称 S 中的 a 为奇异点（分别为正则点）。 S 的奇异轨迹（分别为正则轨迹）是其所有奇异点（分别为正则点）构成的集合。

给定一个在基本半代数集 S上非负的多项式 f，Demmel、Nie 和 Powers 在 [11] 中给出了在交集 $S \cap V^{\mathrm{KKT}}$上得到非负定理的一种方法。这里， V_{KKT} 表示由在 S上最小化 f 的 Karush–Kuhn–Tucker 条件所定义的簇。第一作者利用 Fritz-John 条件，在 [33] 中将这一方法推广到如下情形：奇异轨迹 S^{sing} 在 f下的像（记为 $f(S^{\mathrm{sing}})$）是有限集，其中 S^{sing} 是 S 的奇异轨迹。他还在 [33] 中指出了若干例子，在这些例子中，先前的非负定理并不适用，因为 f 在 S 上的实零点属于 S^{sing}。在 [32]中，第一作者给出了这些非负定理的次数界，并分析了它们在多项式优化中的应用收敛速率。我们强调，Karush–Kuhn–Tucker 条件以及关于 $f(S^{\mathrm{sing}})$ 有限性的假设均属于一般位置情形。所谓一般位置，是指这些性质在具有给定次数的输入多项式的系数空间中的一个扎里斯基开集上成立。迄今为止，尚未处理剩余情形，即集合 $f(S^{\mathrm{sing}})$ 具有无限基数的情形。

关于上述两类一阶最优性条件，Karush– Kuhn–Tucker 条件和 Fritz-John 条件在刻画半代数集 S 上多项式 f 的实零点时发挥作用。此外，不难构造性地证明，在 S 与由这些最优性条件定义的代数簇的交集上， f 存在基于平方和的表示。为此，我们利用以下事实：在与 S 上 f 的最小化相关的 Karush–Kuhn–Tucker 条件所定义的代数簇 V^{KKT} 的每个连通分支上， f 都是常数。在假设 f 有限的条件下，对于由 Fritz-John 条件定义的代数簇 V^{FJ}，这一性质同样成立。这里的区别在于， V^{KKT} 包含 S 的正则点处的 f 的实零点，而 V^{FJ} 则包含 S 上 f 的所有实零点，即使它们是 S 的奇异点。

贡献。本文旨在为一般情形下在基本半代数集 S上非负的多项式 f 提供一些表示（及其次数界）。

一方面，我们将 S 转换为高维实空间中的实代数簇 V（即多项式组的公共实零点），并考虑 f在 V上的表示。设 $U := V$。然后，我们将 U 分解为有限个不可约分支 U_j。如果 U_j 是零维的，那么 f 在 U_j上等同于多项式平方和。否则， f 在 U_j^{KKT}上等同于多项式平方和； U_j^{KKT}是由 f 在 U_j上的最小化所对应的 Karush– Kuhn–Tucker 条件定义的 U_j 的子簇。在这种情况下，我们更新 $U := U_{\text{奇异轨迹}\, j}$，其中 U_j^{sing}是 U_j的奇异轨迹，并重复上述过程。该过程将在有限步后终止，因为

奇异轨迹 U_j^{sing} 的维数低于 U_j。

另一方面，我们给出了 f 上 U_j^{KKT}（或零维 U_j）的平方和表示的次数界。因此，我们分析了 Lasserre层级在各个 U_j^{KKT}（或零维 U_j）上最小化多项式 p 时的收敛速率。我们证明，在某些 U_j^{KKT}（或零维 U_j）上，Lasserre层级有限收敛至最小值 $p^\star$，该最小值由 p 在 V上取得。基于这一结果以及双曲锥的代数结构，我们证明每个双曲规划都可以写成一个半定规划。

双曲多项式与双曲规划。对于感兴趣的读者，双曲规划由 G¨uler 于 [16], 提出，研究的是在仿射空间与所谓的、由双曲多项式约束的双曲锥的交集上优化线性目标函数。这里，若实多项式 p 满足：对于所有向量 a，单变量多项式 $t \mapsto p(te - a)$ 只有实根，则称其关于给定向量 e 是双曲的。关于双曲多项式与双曲规划的近期研究，可参阅 [18, 24–26, 39–41, 53] 。这些研究主要致力于解决广义拉克斯猜想，其直接推论是每个双曲规划都等价于一个半定规划（例如，参见 [2, 注 1.7]）。在本文中，我们证明了后者。读者可能会想知道，这两个规划的等价性是否意味着它们的可行集相同，从而得以解决广义拉克斯猜想。在我们看来，广义拉克斯猜想讨论的是在不同优化问题中使用的固定锥。这与如下说法大不相同：给定目标函数 f 的双曲优化问题，可以通过一个特定的半定规划求解，而该半定规划取决于与 f相关联的双曲规划。

前人工作。在 [34], 中，第一作者针对在实代数簇 V 上非负的多项式 f，在 V的正则轨迹 V^{reg} 在 V中稠密的情况下，给出了基于平方和的表示。其核心思路是借助 Hironaka奇异性消解 [21,22]，将 V 视为正则簇 W 在由多项式向量定义的态射 φ下的像。然后，我们利用 Demmel–Nie–Powers非负性定理，得到 $f \circ \varphi$ 在 W_{KKT}上的表示； W_{KKT}是由在 W上最小化 $f \circ \varphi$ 的 Karush–Kuhn–Tucker 条件定义的簇。然而，当 V^{reg} 在 V 中不稠密时（例如， V 是 Whitney 伞或 Cartan伞），这种方法并不总是适用。相比之下，我们的方法在这些棘手情形下没有问题。

Bucero 和 Mourrain 在 [7] 中提出了一种方法，可在一般位置情形下，获得用于在半代数集 S 上最小化多项式 p 的精确半定规划。首先，他们考虑在一个新的半代数集 S_{FJ}上对 p 进行等价的最小化；该集合是 S 与由在 S 上最小化 p 的 Fritz-John 条件所定义的簇的投影的交集。其次，他们将 S_{FJ} 分解为两个半代数集 S_{KKT} 和 S_{sing}，然后分别考虑在这两个集合上最小化 p 。这里， S_{KKT} 是 S 与由在 S上最小化 p的 Karush–Kuhn–Tucker 条件所定义的簇的投影的交集。集合 S_{sing} 包含 S的奇异轨迹。他们获得了用于在 S_{KKT}上最小化 p 的精确半定松弛

这得益于 Demmel、Nie 和 Powers [11]的研究。对 S_{sing} 上的 p进行最小化是递归求解的。如果 S_{sing} 是零维的，他们就能获得用于在有限集 S_{sing}上最小化 p 的精确半定规划。他们的论文尚未完全处理正维 S_{sing} 的情形。

关于精确性这一性质，我们强调，本文中的表示具有平方和精确性。在他与 Baldi 的合作研究 [3] 中，第三作者考察了矩精确性。在实践中，它比平方和精确性更强，因为它能够给出极小化点，并检验多项式优化是否精确。此外，[3, 定理 4.14] 证明，如果由子式乘积定义的极性簇（且与 Karush–Kuhn–Tucker 簇的投影相关）是有限的，那么矩精确性也成立。

动机。本文所述方法在带互补约束的数学规划中具有潜在应用（例如，参见 [1, 31]）。这类优化问题颇具挑战性，因为互补约束通常违反所有标准约束资格条件，使得求解困难。众所周知，约束资格条件是非线性规划中 Karush–Kuhn–Tucker 条件成立的充分条件。因此，依赖 Karush–Kuhn–Tucker 条件来求解带互补约束的数学规划的经典方法有所局限。相比之下，我们的方法不需要任何约束资格条件

组织结构。本文安排如下：第 2 节介绍证明主要结果所需的实代数几何预备知识。第 3 节给出一种算法，用于求出包含实代数簇奇异点的零维和正维子簇。第 4 节陈述这些子簇上的一些 Nichtnegativstellens¨atze。第 5 节构造高阶最优性条件，以刻画多项式优化问题的全局极小点。第 6 节介绍主要算法，该算法能够为具有全局极小点的多项式优化问题求得精确的半定规划。第 7 节说明如何将双曲规划转化为半定规划。

我们给出一些有趣的例子来说明本文的结果，并使用软件 Oscar [46] 在 Julia 1.7.1 中对这些例子进行了一些计算。相关代码可通过以下链接获取：https://github.com/maihoanganh/SingularSOS。

2 预备知识

2.1 实代数簇

令 $\mathbb{R}[x]$ 表示变量向量 $x = (x_1, \ldots, x_n)$ 中实系数多项式构成的多项式环。令 $\mathbb{R}_r[x]$ 表示 $\mathbb{R}[x]$ 中次数至多为 r 的多项式所构成的线性空间。

给定 $h_1, \ldots, h_l \in \mathbb{R}[x]$，我们用 $V(h)$ 表示在 $\mathbb{R}^n$ 中由向量 $h = (h_1, \ldots, h_l)$ 定义的（实）代数簇，即，

$$V(h) := \{x \in \mathbb{R}^n \,:\, h_j(x) = 0\,,\, j = 1, \ldots, l\}\,. \tag{1}$$

在这种情况下，$h_1, \ldots, h_l$ 称为定义 $V(h)$ 的多项式。

给定 $h_1, \ldots, h_l \in \mathbb{R}[x]$，令 $I(h)[x]$ 为由 $h = (h_1, \ldots, h_l)$ 生成的理想，即，

$$I(h)[x] := \sum_{j=1}^{l} h_j \mathbb{R}[x] \,. \tag{2}$$

理想 $I(h)[x]$ 的实根基记为 $\sqrt[\mathbb{R}]{I(h)[x]}$，定义为

$$\sqrt[\mathbb{R}]{I(h)[x]} := \{f \in \mathbb{R}[x] \,:\, \exists m \in \mathbb{N} \,:\, -f^{2m} \in \Sigma^2[x] + I(h)[x]\} \,. \tag{3}$$

Krivine–Stengle 非负性定理 [23] 蕴含

$$\sqrt[\mathbb{R}]{I(h)[x]} := \{p \in \mathbb{R}[x] \,:\, p = 0 \text{ on } V(h)\} \,. \tag{4}$$

若 $I(h)[x]$ 满足 $I(h)[x] = \sqrt[\mathbb{R}]{I(h)[x]}$，则称其为实根基。

给定代数簇 $V \subset \mathbb{R}^n$，其消失理想记为 $I(V)$，即

$$I(V) := \{p \in \mathbb{R}[x] \,:\, p(x) = 0 \,, \forall x \in V\} \,. \tag{5}$$

注意， $I(V)$ 是一个实根基理想。若 V 由 $h = (h_1, \ldots, h_l)$ 定义，且 $h_j \in \mathbb{R}[x]$，则 $I(V) = \sqrt[\mathbb{R}]{I(h)[x]}$。要计算由给定多项式 $h_1, \ldots, h_l$ 定义 V 的 $I(V)$ 的生成元，可以使用例如 $[5, 42]$ 中的 Becker–Neuhaus 方法。因此，若 h_j 的次数至多为 d，则 $I(V)$ 的生成元次数有上界 $d^{2^{\mathcal{O}(n^2)}}$。此外，也可以通过矩矩阵的核（或矩序列的消灭子）来计算 $I(V)$ 的生成元，例如使用 Baldi–Mourrain 方法 [4]。进而，研究该方法所得到的生成元次数的界可能会很有意义。

给定代数簇 V in $\mathbb{R}^n$，如果不存在两个真子簇 V_1, V_2 in $\mathbb{R}^n$ 使得 $V = V_1 \cup V_2$，我们就称 V is 不可约。

我们回顾以下关于代数簇结构的结果，该结果见于 [10, 第 4.6 节、定理 4] 和 [54, 引理 4.9]：

引理 1. 设 V 是 $\mathbb{R}^n$中的一个代数簇。那么 V 可以写成有限个集合的并

$$V = V_1 \cup \cdots \cup V_r \,, \tag{6}$$

其中每个 V_j 都是 $\mathbb{R}^n$ 中的不可约代数簇，并且当$j \neq t$时满足 $V_j \not\subset V_t$ 。此外，如果 V 由 $\mathbb{R}_d[x]$中的多项式定义，那么 r 有一个仅依赖于 n 和 d的常数界。

我们称引理 1 中的 $V_1, \ldots, V_r$ 为 V的不可约分支。已有若干算法可以根据定义 V的多项式，求出定义 V_j, $j = 1, \ldots, r$的多项式。例如，Gianni、Trager 和 Zacharias 在 [14]中提出了一种不可约分解（也称为主分解）。

2.2 正则点集与奇异点集

给定满足 $h_j \in \mathbb{R}[x]$的 $h = (h_1, ..., h_l)$，我们用 $J(h)$ 表示与 h 相关联的雅可比矩阵，即:

$$J(h)(x) := \left(\frac{\partial h_j}{\partial x_t}(x) \right)_{1 \leq t \leq n, 1 \leq j \leq l} \,. \tag{7}$$

我们将代数簇 V in $\mathbb{R}^n$ 的维数定义为在 V 是实子流形的点处所能达到的最高维数，记作 $\dim(V)$。为方便起见，本文假设 $\dim(\emptyset) = 0$ 。

设 V 是 $\mathbb{R}^n$ 中一个维数为 d 的代数簇。设 $h_1, \ldots, h_l \in \mathbb{R}[x]$ 是 $I(V)$ 的生成元。令 $h = (h_1, \ldots, h_l)$。若雅可比矩阵 $J(h)(a)$ 的秩为 $n - d$，则称 $a \in V$ 为 V 的正则点。这里，实系数矩阵 M 的秩是满足所有 $(r+1)\times(r+1)$ 子式都为零的最大整数 r。V 的所有正则点构成的集合 V^{reg} 称为 V 的正则轨迹。令 V sing $= V \backslash V^{\mathrm{reg}}$。若 $a \in V^{\mathrm{sing}}$，即雅可比矩阵 $J(h)(a)$ 的秩小于 $n - d$，则称 $a \in V$ 为 V 的奇异点。集合 V^{sing} 称为 V 的奇异轨迹。V 在 $a \in V$ 处的切空间记为 $T_a(V)$，它是由下式给出的 $\mathbb{R}^n$ 的线性子空间

$$T_a(V) := \{u \in \mathbb{R}^n \,:\, J(h)(a)^\top u = 0\}\,. \tag{8}$$

不难证明，$a \in V$ 是 V 的正则点，当且仅当 $\dim(T_a(V)) = d = \dim(V)$。

给定满足 $h_j \in \mathbb{R}[x]$ 的 $h = (h_1, \ldots, h_l)$，用 $m_t(h)(x)$ 表示雅可比矩阵 $J(h)(x)$ 的 $t \times t$ 阶子式构成的向量。那么，$m_t(h)$ 的长度为 $\binom{n}{t} \times \binom{l}{t}$。$m_t(h)$ 的每个分量都属于 $\mathbb{R}[x]$，且次数至多为 $t \times \max_j \deg(h_j)$。

下面的引理陈述了奇异点集的一些基本性质：

引理 2. 设 V 是 $\mathbb{R}^n$ 中的一个维数为 d 的代数簇。设 $h_1, \ldots, h_l \in \mathbb{R}[x]$ 为 $I(V)$ 的生成元。令 $h = (h_1, \ldots, h_l)$。则满足以下条件：

1. V 的奇异点集是 $\mathbb{R}^n$ 中的一个代数簇，由 $(h, m_{n-d}(h))$ 定义。
2. V 的奇异轨迹的维数低于 V。

证明。第一个陈述的证明与 [56, 第 6 节2]类似。第二个陈述的证明见 [54, 定理 48]（另见 [55, 第 12.4 节2]）。

□

2.3 一阶最优性条件

给定 $h_0, h_1, \ldots, h_l$ 属于 $\mathbb{R}[x]$，考虑以下多项式优化问题：

$$h^\star := \inf_{x \in V(h)} h_0(x)\,, \tag{9}$$

其中 $V(h)$ 是由 $h = (h_1, \ldots, h_l)$ 在 $\mathbb{R}^n$ 中定义的代数簇。

注 1. 更一般的形式

$$\begin{array}{ll} \inf\limits_{y \in \mathbb{R}^r} & f(y) \\ s.t. & g_j(y) \geq 0\,,\ j = 1, \ldots, l\,, \end{array} \tag{10}$$

通过设置 $x = (y, z)$、$h_0(x) = f(y)$、$h_j(x) = g_j(y) - z_j^2$，可将其写成式 (9) 的一个实例。

给定 $p \in \mathbb{R}[x]$，我们用 ∇p 表示 p 的梯度，即 $\nabla p = (\frac{\partial p}{\partial x_1}, \ldots, \frac{\partial p}{\partial x_n})$。我们在以下引理中回顾 Karush–Kuhn–Tucker 条件：

引理 3. 设 h_0 属于 $\mathbb{R}[x]$。设 V 是 $\mathbb{R}^n$ 中维数为$d > 0$的代数簇。设 $h_1, \ldots, h_l \in \mathbb{R}[x]$ 是 $I(V)$ 的生成元。设 $x^\star$ 是问题（9）的一个局部极小点，且满足 $h = (h_1, \ldots, h_l)$。假设 $J(h)(x^\star)$ 的秩为 $n - d$。那么，问题（9）在$x^\star$处满足 Karush–Kuhn–Tucker 条件，即

$$\begin{cases} \exists(\lambda_1^\star, \ldots, \lambda_l^\star) \in \mathbb{R}^l : \\ h_j(x^\star) = 0\,,\ j = 1, \ldots, l\,, \\ \nabla h_0(x^\star) = \sum_{j=1}^{l} \lambda_j^\star \nabla h_j(x^\star)\,. \end{cases} \tag{11}$$

证明。根据假设，$x^\star$ 是流形 $V \subset \mathbb{R}^n$上的正则点。因此，存在一个微分同胚 $\Phi : U \to V$，其中 $U \subset \mathbb{R}^d$ 是某个开集，使得$x^\star = \Phi(t^\star)$，其中 $t^\star \in U$为某个值。Φ 在 $t \in U$ 处的微分由线性映射 $D\Phi_t : \mathbb{R}^d \to T_{\Phi(t)}V$, $u \mapsto J(\Phi)(t)^\top u$定义，其中 T_aV 是 V在 $a \in V$处的切空间（定义同式 (8)）。由于 $x^\star$ 是 V上的正则点，$D\Phi_{t^\star}$是双射。由此可得秩$(J(\Phi)(t^\star)) = d$，因此，根据秩-零化度定理，$J(\Phi)(t^\star)$的零空间（记为 $J(\Phi)(t^\star)^\perp$）的维数为 $n-d$。根据假设，$t^\star$ 是 $h_0 \circ \Phi$ 在 U上的局部极小点。这意味着$0 = \nabla(h_0 \circ \Phi)(t^\star) = J(\Phi)(t^\star) \times \nabla h_0(x^\star)$，从而 $\nabla h_0(x^\star)$属于 $J(\Phi)(t^\star)^\perp$。此外，对于 $j = 1, \ldots, l$，$(h_j \circ \Phi)(t) = 0$ 对所有 $t \in U$均成立。对t取梯度，得到 $J(\Phi)(t) \times \nabla h_j(\Phi(t)) = 0$，对所有 $t \in U$以及 $j = 1, \ldots, l$均成立。这意味着 $\nabla h_j(\Phi(t))$，$j = 1, \ldots, l$对所有 $t \in U$均属于 $J(\Phi)(t)$的零空间。根据假设，$\nabla h_j(x^\star)$，$j = 1, \ldots, l$ 的线性张成维数为 $n - d$。由于 $J(\Phi)(t^\star)^\perp$ 的维数为 $n - d$，$J(\Phi)(t^\star)^\perp$ 是 $\nabla h_j(x^\star)$,$j = 1, \ldots, l$的线性张成。因此，由于 $\nabla h_0(x^\star)$属于 $J(\Phi)(t^\star)^\perp$，结论得证。

□

注 2. 如 Freund 的讲义所示 *[12, 定理 11]*, 当满足线性无关约束资格条件时，Karush– Kuhn–Tucker 条件（11）在 $x^\star$ 处对问题（9）成立；也就是说，梯度 $\nabla h_j(x^\star)$、$j = 1$、…、l在 $\mathbb{R}^n$中线性无关，这等价于 $J(h)(x^\star)$的秩为 l。为便于比较，我们在引理 3 中采用一个较弱的假设，即 $J(h)(x^\star)$的秩为 $n - \dim(V)$。与 *[10, 练习 17 b、第495]*, 页类似，我们一般可得 $l \geq n - \dim(V)$。注意，扭曲三次曲线 $V = \{(t, t^2, t^3) \in \mathbb{R}^3 : t \in \mathbb{R}\}$ 是一个由 $x_2 - x_1^2$和 $x_3 - x_1^3$定义的一维簇，但理想 $I(V)$由包含三个多项式的向量$h = (x_1x_3 - x_2^2, x_2 - x_1^2, x_3 - x_1x_2)$生成。因此，在此例中有 $l = 3 > 2 = n - \dim(V)$。

为证明引理 3 中关于 $J(h)(x^\star)$ 的最大秩假设不能去掉，考虑以下例子：

例 1. 设 $n = 2$、$h_0 = x_1$和 $h = x_1^3 - x_2^2$。那么 $x^\star = (0, 0)$ 是问题（9）的唯一全局极小点。此外，在 $x^\star$ 处，问题（9）不满足 Karush–Kuhn–Tucker 条件。事实上，对于任意 $\lambda \in \mathbb{R}$，我们得到

$$\nabla h_0(x^\star) - \lambda \nabla h(x^\star) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \lambda \begin{bmatrix} 3x_1^{\star 2} \\ -2x_2^\star \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq 0\,. \tag{12}$$

注意，$V(h)$的维数为 $d = 1$，且 $J(h) = \nabla h = \begin{bmatrix} 3x_1^2 \\ -2x_2 \end{bmatrix}$。不难验证，$J(h)(x^\star)$ 的秩为 $0 < 1 = n - d$。

给定 $h := (h_0, h)$，其中 $h := (h_1, \ldots, h_l)$且 $h_j \in \mathbb{R}[x]$，我们将由以下定义的 Karush–Kuhn–Tucker 条件所对应的 $\mathbb{R}[x, \lambda]$ 中的多项式向量记为 h_{KKT}

$$\bar{h}_{\mathrm{KKT}} := (h, \nabla h_0 - \sum_{j=1}^{l} \lambda_j \nabla h_j)\,, \tag{13}$$

其中 $\lambda = (\lambda_1, \ldots, \lambda_l)$。条件 (11) 可写为：对于某个 $\lambda^\star \in \mathbb{R}^l$，有 $(x^\star, \lambda^\star) \in V(\bar{h}_{\mathrm{KKT}})$。

2.4 半代数集

给定满足 $g_j \in \mathbb{R}[x]$ 的 $g = (g_1, \ldots, g_m)$，我们用 $S(g)$ 表示与 g 相关联的基本半代数集，即

$$S(g) := \{x \in \mathbb{R}^n : g_j(x) \geq 0\,,\, j = 1, \ldots, m\}\,. \tag{14}$$

$\mathbb{R}^n$ 的一个半代数子集具有以下形式：

$$\bigcup_{i=1}^{t} \bigcap_{j=1}^{r_i} \{x \in \mathbb{R}^n : f_{ij}(x) *_{ij} 0\}\,, \tag{15}$$

其中 $f_{ij} \in \mathbb{R}[x]$ 和 $*_{ij} \in \{>, =\}$。注意，(15) 是有限个基本半代数集的并。

给定两个半代数集 $A \subset \mathbb{R}^n$ 和 $B \subset \mathbb{R}^m$，若映射 f：$A \to B$ 的图像 $\{(x, f(x)) : x \in A\}$ 是 $\mathbb{R}^{n+m}$ 中的一个半代数集，则称该映射为半代数映射。若半代数子集 $A \subset \mathbb{R}^n$ 满足：对于任意 A 中的 x, y，都存在连续的半代数映射 $\phi : [0, 1] \to A$，使得 $\phi(0) = x$ 且 $\phi(1) = y$，则称其为半代数路径连通的。

以下引理可见于 [47, 命题 1.6.2 (ii)]：

引理4. 半代数映射的复合仍为半代数映射。

以下引理见于 [47, 定理 1.8.1]：

引理5. 设 $f : (a, b) \to \mathbb{R}$ 是一个半代数函数。那么存在 $a = a_0 < a_1 < \cdots < a_s < a_{s+1} = b$，使得对于每个 $i = 0, \ldots, s$，限制$f|_{(a_i, a_{i+1})}$是解析的。

以下引理由中值定理可得：

引理6. 设 f：$[0, 1] \to \mathbb{R}$ 是一个连续的分段可微函数，即存在 $0 = a_1 < \cdots < a_r = 1$，使得 f 连续，且 f 在每个开区间(a_i, a_{i+1})上可微。假设 f 的次梯度为零。那么 $f(0) = f(1)$。

证明。对每个开区间 (a_i, a_{i+1}) 应用中值定理，可得$f(a_i) = f(a_{i+1})$。因此，由 $f(0) = f(a_1) = \cdots = f(a_r) = f(1)$ 可得所需结果。 □

给定 $n, d, s \in \mathbb{N}$，我们定义

$$c(n, d, s) := d(2d - 1)^{n+s-1}\,. \tag{16}$$

我们在下面的引理中重述 Coste 在 [9, 命题 4.13] 中给出的基本半代数集的连通分支数上界：

引理 7. 设 $g_1,\dots,g_m,h_1,\dots,h_l\in\mathbb{R}_d[x]$ 满足 $d\geq 2$。 $S(g)\cap V(h)$ 的（半代数路径）连通分支数不超过$c(n,d,m+l)$。

2.5 平方和

用 $\Sigma^2[x]$ （分别用 $\Sigma_r^2[x]$）表示 $\mathbb{R}[x]$（分别为 $\mathbb{R}_r[x]$）中多项式的平方和锥。给定 $g_1,\dots,g_m\in\mathbb{R}[x]$, 令 $P_r(g)[x]$ 为与 $g=(g_1,\dots,g_m)$ 关联的 $r\in\mathbb{N}$ 阶截断预序，即

$$P_r(g)[x]:=\{\sum_{\alpha\in\{0,1\}^m}\sigma_\alpha g^\alpha\,:\,\sigma_\alpha\in\Sigma^2[x]\,,\,\deg(\sigma_\alpha g^\alpha)\leq 2r\}\,,\tag{17}$$

其中 $\alpha=(\alpha_1,\dots,\alpha_m)$, 且 $g^\alpha:=g_1^{\alpha_1}\dots g_m^{\alpha_m}$。若 $m=0$, 则有 $P_r(g)[x]=\Sigma_r^2[x]$。

给定 $h_1,\dots,h_l\in\mathbb{R}[x]$, 令 $I_r(h)[x]$ 为由 h 定义的 r 阶截断理想，即

$$I_r(h)[x]:=\{\sum_{j=1}^{l}h_j\psi_j\,:\,\psi_j\in\mathbb{R}[x]\,,\,\deg(h_j\psi_j)\leq 2r\}\,.\tag{18}$$

我们用 bit(d) 表示 $d\in\mathbb{N}$ 的比特数，即

$$\mathrm{bit}(d):=\begin{cases}1 & \text{if } d=0\,,\\ k & \text{if } d\neq 0 \text{ and } 2^{k-1}\leq d<2^k.\end{cases}\tag{19}$$

给定 $n,d,s\in\mathbb{N}$，我们定义

$$b(n,d,s):=2^{2^{\left(2^{\max\{2,d\}^{4^n}+s^{2^n}\max\{2,d\}^{16^n\,\mathrm{bit}(d)}}\right)}}.\tag{20}$$

我们回顾 Lombardi、Perrucci 和 Roy 关于 Krivine–Stengle 非负表示的次数界，见以下两个引理:

引理 8. 设 $g_1,\dots,g_m,h_1,\dots,h_l$ in $\mathbb{R}_d[x]$。假设 $S(g)\cap V(h)=\emptyset$ with $g:=(g_1,\dots,g_m)$ and $h:=(h_1,\dots,h_l)$。令 $r=\frac{1}{2}\times b(n,d,m+l+1)$。则有 $-1\in P_r(g)[x]+I_r(h)[x]$。

引理 9. 设 $p,g_1,\dots,g_m,h_1,\dots,h_l$ in $\mathbb{R}_d[x]$。假设 p 在$S(g)\cap V(h)$上消失，其中 $g:=(g_1,\dots,g_m)$ and $h:=(h_1,\dots,h_l)$。令 $r:=\frac{1}{2}\times b(n,d,m+l+1)$ and $s:=2\lfloor r/d\rfloor$。则有 $-p^s\in P_r(g)[x]+I_r(h)[x]$。

2.6 正则点集上的非负表示定理

用 $|\cdot|$ 表示一个集合的基数，用 δ_{ij} 表示在 (i,j) 处的克罗内克 δ 函数 $\in\mathbb{N}^2$。

我们在下面的引理中陈述一种基于平方和的表示及其次数界，适用于在实代数簇上仅取有限个非负值的多项式:

引理 10. 设 $h_0,h_1,\dots,h_l$ 属于 $\mathbb{R}_d[x]$。假设 h_0 在$V(h)$上非负，且 $h_0(V(h))$ 有限，其中 $h=(h_1,\dots,h_l)$。令 $r:=|h_0(V(h))|$ 以及$u:=\frac{1}{2}\times b(n,d,l+1)$。那么存在 $\sigma\in\Sigma_w^2[x]$，使得 $w=\max\{d(r-1)\ u+d\}$，且 $h_0-\sigma$ 在 $V(h)$上消失。

证明。考虑以下两种情况：

- 情形 1： $r = 0$。显然， $V(h) = \emptyset$。引理 8 表明 $-1 \in \Sigma_u^2[x] + I_u(h)[x]$。这意味着存在 $q \in P_u(g)[x]$，使得 $-1 = q$ 在 $V(h)$ 上。我们写作 $h_0 = s_1 - s_2$，其中 $s_1 = (h_0 + \frac{1}{2})^2$ 和 $s_2 = h_0^2 + \frac{1}{4}$ 属于 $\Sigma^2[x]$。由此可得 $h_0 = s_1 + qs_2$ 在 $V(h)$ 上。令 $\sigma = s_1 + qs_2$，则有 $\sigma \in \Sigma_{u+d}^2[x] \subset \Sigma_w^2[x]$，因为 $w \geq u + d$。因此 $h_0 - \sigma$ 在 $V(h)$上消失。

- 情形 2： $r > 0$。根据假设，我们可以假设 $h_0(V(h)) = \{t_1, \ldots, t_r\} \subset [0, \infty)$，其中 $t_i \neq t_j$ 若 $i \neq j$。对于 $j = 1, \ldots, r$，令 $W_j := V(h, h_0 - t_j)$。那么 W_j 是由 $\mathbb{R}_d[x]$ 中的 $l+1$ 个多项式定义的实簇。显然， $h_0(W_j) = \{t_j\}$。定义以下多项式：

$$p_j(x) := \prod_{i \neq j} \frac{h_0(x) - t_i}{t_j - t_i}, \; j = 1, \ldots, r. \tag{21}$$

容易验证 $p_j(W_i) = \{\delta_{ji}\}$，且 $\deg(p_j) \leq d(r-1)$。注意，对于 $i = 1, \ldots, r$， $h_0 = t_i \geq 0$ 在 W_i 上。现在令 $\sigma = \sum_{i=1}^r t_i p_i^2$，得到$\sigma \in \Sigma_w^2[x]$，因为 $\deg(t_i p_i^2) \leq 2 \deg(p_i) \leq 2d(r-1) \leq 2w$。因此 $h_0 - \sigma$在 $V(h) = W_1 \cup \cdots \cup W_r$ 上消失，从而得到所需结果。

□

下面的引理与 [11, 引理 3.3] 类似，但其证明使用了实代数几何中的工具（而非复代数几何中的工具）：

引理 11. 设 $h_1, \ldots, h_l \in \mathbb{R}[x]$。设 h_0 是 $\mathbb{R}[x]$ 中的一个多项式。设 W 是 $V(h_{\mathrm{KKT}})$ 的一个半代数路径连通分支，其中 $h := (h_0, \ldots, h_l)$。那么 h_0 在 W 上为常数。

证明. 回忆 $\lambda := (\lambda_1, \ldots, \lambda_m)$。任取 W 中的两个点 $(x^{(0)}, \lambda^{(0)})$、$(x^{(1)}, \lambda^{(1)})$。我们断言 $h_0(x^{(0)}) = h_0(x^{(1)})$。根据假设，存在一个连续的半代数映射 $\phi\colon [0,1] \to W$，由 $\phi(\tau) = (x(\tau)\ \lambda(\tau))$ 定义，使得 $\phi(0) = (x^{(0)}, \lambda^{(0)})$ 且 $\phi(1) = (x^{(1)}, \lambda^{(1)})$。我们断言 $\tau \mapsto h_0(x(\tau))$ 在 $[0,1]$ 上为常数。拉格朗日函数

$$L(x, \lambda) := h_0(x) - \sum_{j=1}^m \lambda_j h_j(x) \tag{22}$$

在 $V(h_{\mathrm{KKT}})$ 上等于 $h_0(x)$，而后者包含 $\phi([0,1])$。根据引理4，函数 $L \circ \phi$ 是半代数的。此外，由于 L 和 ϕ 连续，函数 $L \circ \phi$ 也连续。由此可知，得益于引理5， $L \circ \phi$ 是连续的分段可微函数。注意，函数 $L \circ \phi$ 在 $[0,1]$ 上的次梯度为零。由引理6可得， $h_0(x(0)) = (L \circ \phi)(0) = (L \circ \phi)(1) = h_0(x(1))$。现在得到 $h_0(x^{(0)}) = h_0(x^{(1)})$，因此 h_0 在 W上为常数。

□

基于 Karush–Kuhn–Tucker 条件，我们在下述定理中陈述关于代数簇上非负多项式的基于平方和的表示及其次数界：

定理 1. 设 $h_0, h_1, \ldots, h_l$ 是 $\mathbb{R}_d[x]$ 中次数为 $d \geq 2$ 的多项式。假设 $\bar{}$ 满足： h_0 在 $V(h)$ 上非负，其中 $h = (h_1, \ldots, h_l)$。令 $h := (h_0, h)$，$\lambda := (\lambda_1, \ldots, \lambda_l)$，并且

$$w := \max\{d \times (c(n+l, d, n+l) - 1), \frac{1}{2} \times b(n+l, d, l+n+1) + d\}, \quad (23)$$

其中 $c(\cdot)$ 和 $b(\cdot)$ 分别按 (16) 和 (20) 定义。于是有以下结论：

$\bar{}$1. $h_0(V(h_{\mathrm{KKT}}))$ 的基数至多为 $c(n+l, d+1, n+l)$。2. 存在 $\sigma \in \Sigma_w^2[x, \lambda]$，使得 $h_0 - \sigma$ 在 $V(\bar{h}_{\mathrm{KKT}})$ 上为零。

$\bar{}$证明。利用引理 7，我们将 $V(h_{\mathrm{KKT}})$ 分解为半代数路径连通分支： $Z_1, \ldots, Z_s$，其中

$$s \leq c(n+l, d, n+l), \quad (24)$$

$\bar{}$因为 h_{KKT} 的每个分量的次数至多为 $d \geq 2$。因此，引理 11$\bar{}$表明 h_0 在每个 Z_i 上均为常数。由此， $h_0(V(h_{\mathrm{KKT}}))$ 是有限的。令 $r = |h_0(V(h_{\mathrm{KKT}}))|$。由 (24)，可得

$$r \leq s \leq c(n+l, d, n+l). \quad (25)$$

Set

$$u := \frac{1}{2} \times b(n+l, d, l+n+1). \quad (26)$$

利用引理 10，存在 $\sigma \in \Sigma_\xi^2[x, \lambda]$，其满足 $\xi = \max\{d(r-1)\ u + d\}\bar{}$，使得 $h_0 - \sigma$ 在 $V(h_{\mathrm{KKT}})$ 上为零。根据 (25) 和 (23)，可得 $\xi \leq w$，因此 $\sigma \in \Sigma_w^2[x, \lambda]$。 □

注 3. 关于定理 1 中的表示，留在 x-空间而不是提升到 (x, λ)-空间也可能很有意思。为此，我们需要将 Karush–Kuhn–Tucker 簇投影到 x-空间，并处理类似于 Bucero 和 Murrain 在 *[7]* 中给出的表示。另一个尚未解决／较难的问题是，为多项式表示获得更好的（例如，双指数）界。

我们在下述定理中给出 Nie 的 Nichtnegativstellensatz [44]，的次数界；该定理给出了在零维代数簇上非负的多项式的基于平方和的表示：

定理 2. 设 $h_0, h_1, \ldots, h_l$ 是 $\mathbb{R}_d[x]$ 中满足 $d \geq 2$ 的多项式。假设 $V(h)$ 与 $h = (h_1, \ldots, h_l)$ 的维数为零，且 h_0 在 $V(h)$ 上非负。令

$$w := \max\{d \times (c(n, d, l) - 1), \frac{1}{2} \times b(n, d, l+1) + d\}, \quad (27)$$

其中 $c(\cdot)$ 和 $b(\cdot)$ 分别如 (16) 和 (20) 中所定义。则以下结论成立：

1. *The cardinality of* $h_0(V(h))$ *is at most* $c(n, d, l)$.
2. 存在 $\sigma \in \Sigma_w^2[x]$，使得 $h_0 - \sigma$ 在 $V(h)$ 上消失。

证明。由于 $V(h)$ 是零维的， $V(h)$ 有有限个不同的点。由引理 7 可得 $s \le c(n,d,l)$, 因为 h 的每个项的次数至多为 $d \ge 2$。于是 h_0 在每个 Z_i 上都是常数，这意味着 $h_0(V(h_{\mathrm{KKT}}))$ 是有限的。令 $r = |h_0(V(h))|$。这意味着 $r \le s \le c(n,d,l)$。令 $u := \frac{1}{2} \times b(n,d,l+1)$。利用引理 10，存在 $\sigma \in \Sigma^2_\xi[x]$，满足 $\xi = \max\{d(r-1)\ u+d\}$, 使得 $h_0 - \sigma$ 在 $V(h)$ 上消失。不难证明 $\xi \le w$，因此 $\sigma \in \Sigma^2_w[x]$。

□

2.7 一般位置情形下的精确多项式优化

我们回顾一下 Lasserre 在 [28] 中最初提出的矩-SOS 松弛的一些预备知识。给定 $d \in \mathbb{N}$，令 $\mathbb{N}^n_d := \{\alpha \in \mathbb{N}^n : \sum_{j=1}^n \alpha_j \le d\}$。给定 $d \in \mathbb{N}$，我们用 v_d 表示 x 中次数不超过d的单项式向量，即 $v_d = (x^\alpha)^{\alpha \in \mathbb{N}^n_d}$，其中 $x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$。对于每个 $p \in \mathbb{R}_d[x]$, 我们写作$p = c(p)^\top v_d = \sum_{\alpha \in \mathbb{N}^n_d} p_\alpha x^\alpha$，其中 $c(p)$ 表示 p的系数向量，即 $c(p) = (p_\alpha)^{\alpha \in \mathbb{N}^n_d}$，其中 $p_\alpha \in \mathbb{R}$。给定对称矩阵 $A \in \mathbb{R}^{r \times r}$，若 A的每个特征值均为非负，我们称 A 为半正定矩阵，记作 $A \succeq 0$。

给定 $y = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subset \mathbb{R}$，令 $L_y : \mathbb{R}[x] \to \mathbb{R}$ 为由 $L_y(p) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha y_\alpha$ 定义的 Riesz 线性泛函，对每个 $p \in \mathbb{R}[x]$均成立。给定 $d \in \mathbb{N}$、 $p \in \mathbb{R}[x]$和 $y = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subset \mathbb{R}$，令 $M_d(y)$为由 $(y_{\alpha+\beta})_{\alpha,\ \beta \in \mathbb{N}^n_d}$定义的 d 阶矩阵。

下面的引理说明了平方和与半定规划之间的联系（例如，参见 [29, 命题 2.1]）：

引理 12. 设 $\sigma \in \mathbb{R}[x]$ 和 $d \in \mathbb{N}$ 满足 $2d \ge \deg(\sigma)$。那么， $\sigma \in \Sigma^2[x]$当且仅当存在 $G \succeq 0$ 使得 $\sigma = v_d^\top G v_d$。

给定 $k \in \mathbb{N}$和 h_0、h_1、.....、$h_l \in \mathbb{R}[x]$，考虑与 $h = (h_0$、h_1、.....、$h_l)$相关联的以下原始—对偶半定规划：

$$\begin{aligned} \tau_k(\bar{h}) := \inf_{y} \quad & L_y(h_0) \\ \text{s.t} \quad & M_k(y) \succeq 0\,, \\ & M_{k-r_t}(h_t y) = 0\,,\ t = 1,\dots,l\,,\ y_0 = 1\,, \end{aligned} \tag{28}$$

$$\begin{aligned} \rho_k(\bar{h}) := \sup_{\xi, G, u_t} \quad & \xi \\ \text{s.t} \quad & G \succeq 0\,, \\ & h_0 - \xi = v_k^\top G v_k + \textstyle\sum_{t=1}^l h_t u_t^\top v_{2(k-r_t)}\,, \end{aligned} \tag{29}$$

其中 $r_t = \lceil \deg(h_t)/2 \rceil$。利用引理 12，我们得到

$$\rho_k(\bar{h}) = \sup_{\xi \in \mathbb{R}} \{\xi : h_0 - \xi \in \Sigma_k[x] + I_k(h)[x]\}\,. \tag{30}$$

原始-对偶半正定规划（28）-（29）称为问题（9）的 k 阶矩-SOS 松弛。

我们在下面的引理中陈述了关于矩-SOS 松弛的一些最新结果：

引理 13. 设 $h_0, h_1, \dots, h_l \in \mathbb{R}_d[x]$。令 $h^\star$ 如 (9) 中所述，其中 $h = (h_1, \dots, h_l)$。令 $h = (h_0, h)$。假设 $h^\star > -\infty$。则下列结论成立：

1. 对每个 $k \in \mathbb{N}$, $\tau_k(h) \le \tau_{k+1}(h)$ 且 $\rho_k(h) \le \rho_{k+1}(h)$。2. 对每个 $k \in \mathbb{N}$, $\rho_k(h) \le \tau_k(h) \le h^\star$。3. 若存在 $q \in \Sigma_w[x]$，满足 $2w \ge d$，使得 $h_0 - h^\star - q$ 在$V(h)$ 上消失，则 $\rho_r(\bar{h}) = h^\star$，其中 $r = \frac{1}{2} \times b(n, 2w, l+1) + d$，而 $b(\cdot)$ 的定义如 (20) 所示。

证明。前两个陈述的证明是平凡的。我们在 [45, 定理 1 的证明1]中使用 Nie 的技巧来证明第三个陈述。（借助 Krivine–Stengle 的非负表示定理，它使我们无需使用实根基，同时仍能得到精确半定规划。）考虑以下两种情况：

- 情况 1：$V(h) = \emptyset$。于是我们得到 $h^\star = \infty$。令$s = \frac{1}{2} \times b(n, d, l+1)$。那么引理 9 表明 $-1 \in \Sigma_s^2[x] + I_s(h)[x]$。对于所有 $\xi \ge 0$，都有

$$h_0 - \xi = (1 + \frac{h_0}{4})^2 - (\xi + (1 - \frac{h_0}{4})^2) \in \Sigma_{s+d}[x] + I_{s+d}(h)[x]\,. \tag{31}$$

由于 $r \ge s + d$，这意味着对于所有 $\xi \ge 0$，$h_0 - \xi \in \Sigma_r[x] + I_r(h)[x]$，从而得到 ξ 是 (30) 的一个可行解，其值为 $\rho_r(h)$。因此我们得到 $\rho_r(h) = \infty = h^\star$。

- 情形 2：$V(h) \ne \emptyset$。于是得到 $h^\star < \infty$。令 $u = h_0 - h^\star - q$。根据假设，我们在 $V(h)$ 上得到 $u \in \mathbb{R}_{2w}[x]$ 和 $u = 0$。令 $s = 2\lfloor (r-d)/(2w) \rfloor$。由此，引理 9 表明，存在 $\sigma \in \Sigma_{r-d}^2[x]$ 使得 $u^{2s} + \sigma \in I_{r-d}(h)[x]$。令 $c = \frac{1}{2s}$。那么有 $1 + t + ct^{2s} \in \Sigma_s^2[t]$。因此，对所有 $\varepsilon > 0$，我们有

$$\begin{aligned} h_0 - h^\star + \varepsilon &= q + \varepsilon(1 + \tfrac{u}{\varepsilon} + c\left(\tfrac{u}{\varepsilon}\right)^{2s}) - c\varepsilon^{1-2s}(u^{2s} + \sigma) + c\varepsilon^{1-2s}\sigma \\ &\in \Sigma_{r-d}^2[x] + I_{r-d}(h)[x] \subset \Sigma_r^2[x] + I_r(h)[x]\,. \end{aligned} \tag{32}$$

那么，对所有 $\varepsilon > 0$，$h^\star - \varepsilon$ 都是 (30) 的一个可行解，其取值为$\rho_r(h)$。由此得到 $\rho_r(h) \ge h^\star - \varepsilon$，对所有 $\varepsilon > 0$均成立，进而我们得到 $\rho_r(h) \ge h^\star$。利用第二个陈述，我们可得 $\rho_r(h) = h^\star$，从而得到第三个陈述。

□

我们将定理 1 应用于多项式优化如下：

定理 3. 设 $h_0, h_1, \ldots, h_l \in \mathbb{R}_d[x]$。设 $h^\star$ 如问题（9）中所示，其中$h = (h_1, \ldots, h_l)$。假设问题（9）存在一个全局极小点，且该问题在此点满足 Karush–Kuhn–Tucker 条件。设 w 如（23）中所示。令

$$r := \frac{1}{2} \times b(n + l, 2w, l + n + 1) + d\,. \tag{33}$$

则 $\rho_r(h_0, h_{\mathrm{KKT}}) = h^\star$，其中 $h := (h_0, h)$，且 $b(\cdot)$ 的定义与（20）中相同。

证明。根据假设，存在 $(x^\star, \lambda^\star) \in V(h_{\mathrm{KKT}})$，使得 $x^\star$ 是（9）的全局极小点。这意味着

$$\begin{aligned} h^\star := \min_{x,\lambda} \quad & h_0(x) \\ \text{s.t.} \quad & (x, \lambda) \in V(\bar{h}_{\mathrm{KKT}})\,, \end{aligned} \tag{34}$$

根据假设，由定理 1 可知，存在 $q \in \Sigma_w^2(g)[x,\lambda]$ ，使得$h_0 - h^\star - q$在 $V(h_{\mathrm{KKT}})$ 上消失。应用引理 13 的第三个陈述，我们得到结论。

□

注 4. 设 $h = (h_1, \ldots, h_l)$，其中 $h_j \in \mathbb{R}[x]$，并令 $h^\star$ 如式 (9) 所示。假设 $h^\star$ 有限但未达到，即对于所有 $x \in V(h)$，均有 $h_0(x) - h^\star > 0$。（例如，可以取 (i) $h_0 = x_1$ 且 $h = (x_1 x_2^2 - 1)$，或 (ii) $h_0 = (x_1 x_2 - 1)^2 + x_1^2$ 且 $h = (0)$。）由定理 1 可知，集合 $h_0(V(h_{\mathrm{KKT}})) - h^\star$（其中 $h := (h_0, h)$）的取值有限个，但不包含零值。这是因为

$$0 \notin h_0(V(h)) - h^\star \supset h_0(V(\bar{h}_{\mathrm{KKT}})) - h^\star . \tag{35}$$

这意味着 $\inf(h_0(V(h_{\mathrm{KKT}})) - h^\star) = \delta > 0$，因此 $\inf h_0(V(h_{\mathrm{KKT}})) = h^\star + \delta$。注意，当且仅当 $h_0(V(h_{\mathrm{KKT}})) = \emptyset$ 时，$\delta = \infty$。因此我们得到 $\rho_r(h_0, h_{\mathrm{KKT}}) = h^\star + \delta > h^\star$，其中 w 如式 (23) 所示，r 如式 (33) 所示。关于这一可达到性问题，请参见注 12。

我们展示定理 2 在多项式优化中的应用：

定理 4. 设 $h_0, h_1, \ldots, h_l \in \mathbb{R}_d[x]$。设 $h^\star$ 如问题（9）中所述，其中$h = (h_1, \ldots, h_l)$。假设 $V(h)$ 是零维的。设 w 如（ 27 ）中所述。令

$$r := \frac{1}{2} \times b(n, 2w, l+1) + d . \tag{36}$$

则 $\rho_r(h_0, h) = h^\star$，其中 $b(\cdot)$ 的定义如（20）中所述。

证明. 根据假设，由定理 2 可知，存在 $q \in \Sigma_w^2[x]$ ，使得$h_0 - h^\star - q$在 $V(h)$ 上为零。应用引理 13 的第三个结论，即得所需结论。

□

3 奇异点集的分解

以下算法可用于求得包含实代数簇奇异点的零维子簇和正维子簇：

算法 1. 奇异点集的分解。

- 输入：代数簇 V 属于 $\mathbb{R}^n$ ，由 $\mathbb{R}[x]$ 中的多项式定义。
- 输出： V的子簇集合 A 和 B 。1. 设 $U := V$, $A := \emptyset$ 和 $B := \emptyset$。2. 令 $U_1, \ldots, U_r$为 U的不可约分支。3. 设 $A := A \cup \{U_1, \ldots, U_r\}$。4. 对于 $j = 1, \ldots, r$，执行：(a) 如果 U_j 是零维的，则设 $A := A \backslash \{U_j\}$ 和 $B := B \cup \{U_j\}$。(b) 否则，令 $U_{\text{奇异轨迹 } j}$为 U_j的奇异轨迹，设 $U := U_{\text{奇异轨迹 } j}$，并再次执行步骤 2、3、4。

备注 5. 为了验证 U_j 是否如算法 1 的步骤 4 (a) 所述为零维，我们请读者参阅 Lairez 和 Safey El Din 在 *[27]*中的研究。此外，他们的方法还可以计算一般实代数簇的维数。

注 6。在算法 1 的步骤 4 (b) 中，我们隐式地计算实根基 $I(U_j)$ 的生成元 h，以得到由 $(h, m_{n-d}(h))$ 定义的奇异轨迹 U_j^{sing}（见引理 2 的第一条陈述）。特别地，如果 U_j^{sing} 由 $(h, m_{n-d}(h))$ 定义，其中某个定义 U_j 且具有重数的多项式向量 h（例如，对于某个 $q \in \mathbb{R}[x]$，$h = (q^2)$）满足此条件，那么我们有 $U_j^{\mathrm{sing}} = U_j$，而算法 1 将陷入无限循环。基于降幂技术和所谓 Thom–Boardman 分层的其他方法可以避免这一问题。它们允许我们直接处理方程和导数（例如，参见 Hauenstein 和 Wampler 的工作 *[17]*）。这些方法可能效率不高，但与依赖于 Becker–Neuhaus 方法 *[5, 42]* 中实根基计算的方法相比，或许能为定义奇异点集的多项式给出更好的界。

下面的引理给出了算法 1 的输出的一些性质。

引理 14. 设 A 和 B 是算法 1 的输出。以下陈述成立：

1. A 和 B 是由 $\mathbb{R}[x]$中的多项式定义的 $\mathbb{R}^n$ 中的有限个代数簇的集合。2. A 中的每个簇都是不可约且正维的。3. B 中的每个簇都是不可约且零维的。
4. $V = \bigcup_{W \in A \cup B} W$.

证明. 证明第一条。根据引理 1 和引理 2 的第一条，算法 1 的步骤 4 (b) 中的 U_j 和 U_j^{sing} 是由 $\mathbb{R}[x]$中的多项式定义的 $\mathbb{R}^n$ 中的代数簇，因此 A 和 B中的元素也如此。要证明 A 和 B 是有限的，只需证明算法 1 会在有限步后终止。反设算法 1 永不终止。引理 1 指出，算法 1 的步骤 2 中的数量 r 是有限的。因此，步骤 4 (b) 会产生一个无限的不可约簇序列 $(W_t)_{t \in \mathbb{N}}$，其中 W_0 是 V 的一个不可约分支，且 W_{t+1} 是 W_t的奇异轨迹的一个不可约分支。这是不可能的，因为引理 2 的第二条给出，对所有 $t \in \mathbb{N}$，$\dim(V) \geq \dim(W_t) > \dim(W_{t+1}) \geq 0$。最后三条陈述由算法 1 的步骤1和步骤 4 (a) 得出。

□

现在给出以下示例，该示例基于 [33, 示例 18]:

例 2. 令 $n = 5$ 和 $V = V(h)$（其中 $h = (x_1^3 - x_4^2, (x_2 - x_3)^3 - x_5^2, x_1(x_2 - x_3))$）作为算法 1 的输入。在算法 1 的步骤 2 中，我们将V分解为不可约分支： $U_j = V(h^{(j)})$、$j = 1$、2，其中 $h^{(1)} = (x_5, x_2 - x_3, x_1^3 - x_4^2)$ 且 $h^{(2)} = (x_4, (x_2 - x_3)^3 - x_5^2, x_1)$。（$U_1$ 和 U_2的维数均为二。）令 U_j^{sing}为 U_j的奇异轨迹，$j = 1$，2。利用第一条陈述

引理 2，我们得到定义 $U_{\text{奇异}1}$ 的多项式向量 $h^{(1)}{}_{\text{奇异}} = (x_1, x_5, x_4, x_2 - x_3)$。显然， U_1^{sing}是一个一维不可约簇，且 U_1^{sing} 的奇异轨迹为空。类似地，我们得到相同的多项式向量 $h^{(2)}{}_{\text{奇异}} = (x_1, x_5, x_4, x_2 - x_3)$，它定义了 U_2^{sing}，因此得到 $U_{\text{奇异}2} = U_{\text{奇异}1}$。算法 1 的步骤 4 (a) 得到 $A = \{U_1, U_2, U_1^{\text{sing}}\}$ 和 $B = \{\emptyset\}$。

4 奇异点集上的非负性定理

下述定理给出了算法 1 返回的、基于平方和的子簇表示：

定理 5. 设 $q_0, h_1, \ldots, h_l$ 属于 $\mathbb{R}[x]$。假设 q_0 在 $V(h)$上非负，且满足 $h := (h_1, \ldots, h_l)$。设 A 和 B 是以 $V = V(h)$ 为输入运行算法 1 得到的输出。设 $c(\cdot)$ 和 $b(\cdot)$ 分别如 (16) 和 (20) 中所述。则以下结论成立：

1. 对于由 $q = (q_1, \ldots, q_r)$ 定义且满足 $q_j \in \mathbb{R}[x]$ 的 A 中的每个簇 U，有

   (a) 满足 $\bar{q} = (q_0, q)$ 的 $q_0(V(\bar{q}_{\text{KKT}}))$ 的基数至多为

   $$c(n + r, d + 1, n + r), \tag{37}$$

   其中 $d := \max\{\deg(q_j) : j = 0, \ldots, r\}$;

   (b) 存在 $\sigma \in \Sigma_w[x, \lambda]$，使得 $\lambda = (\lambda_1, \ldots, \lambda_l)$，并且

   $$w := \max\{d \times (c(n+r, d, n+r) - 1), \ \frac{1}{2} \times b(n+r, d, r+n+1) + d\} \tag{38}$$

   使得 $q_0 - \sigma$ 在 $V(\bar{q}_{\text{KKT}})$ 上消失。

2. 对于每个由 $q = (q_1, \ldots, q_r)$ 定义的簇 $U \in B$，其中 $q_j \in \mathbb{R}[x]$，则

   (a) $q_0(V(q))$ 的基数至多为 $c(n, d, r)$，其中 $d := \max\{\deg(q_j) : j = 0, \ldots, r\}$;

   (b) 存在 $\sigma \in \Sigma_w[x]$ 使得

   $$w := \max\{d \times (c(n, d, r) - 1), \frac{1}{2} \times b(n, d, r + 1) + d\} \tag{39}$$

   $q_0 - \sigma$ 在 $V(q)$ 上消失。

证明。为证明第一和第二个陈述，我们分别应用定理 1 和定理 2。注意，对于每个 $U \in A$， p_0 在 $U \cap \mathbb{R}^n$ 上非负，因为 q_0 在 $V \cap \mathbb{R}^n \supset U \cap \mathbb{R}^n$ 上非负。此外，引理 14 的第二个陈述表明，每个 $U \in B$ 的维数为零。 □

例 3。我们使用与例 2 相同的记号。设 $q_0 = x_1 + x_2 - x_3$。则 q_0 在 $V(h)$ 上非负。反设 Schm ¨udgen 正定性定理的非严格扩展适用于 q_0 在 $V(h)$ 上的情形，即$q_0 \in \Sigma_d^2[x] + I_d(h)[x]$ 对某个 $d \in \mathbb{N}$成立。令 $x_2 = x_3 = x_4 = x_5 = 0$，可得$x_1 = \sigma + x_1^3\psi$，其中 $\sigma \in \Sigma^2[x_1]$ 和 $\psi \in \mathbb{R}[x_1]$。这意味着 x_1 整除 σ，因此对某个 $\sigma_1 \in \Sigma^2[x_1]$有 $\sigma = x_1^2\sigma_1$。由此，$x_1 = x_1^2\sigma_1 + x_1^3\psi$ 推出

$1 = x_1(\sigma_1 + x_1\psi)$，这在令 $x_1 = 0$时不可能成立。不难验证定理 5.2 (b) 适用于此例。事实上，令$\lambda := (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$，并令 $\sigma := 0$，可推出 $\sigma \in \Sigma[x, \lambda]$，且 $q_0 - \sigma$在$U_1^{\mathrm{sing}} = V(h^{(1)}_{\text{奇异}}) \in A$上消失。因此， $q_0 - \sigma$在 $V(\bar{q}_{\mathrm{KKT}}) \subset V(h^{(1)}_{\text{奇异}}) \times \mathbb{R}^4$ 上消失，其中$\bar{q} = (q_0, h^{(1)}_{\text{奇异}})$。

5 高阶最优性条件

下面的引理给出了问题（9）满足一阶最优性条件时局部极小点的刻画：

引理 15. 设 V 是 $\mathbb{R}^n$ 中的一个代数簇。设 $h_1, \ldots, h_l \in \mathbb{R}[x]$ 是 $I(V)$ 的生成元。设 h_0 是 $\mathbb{R}[x]$ 中的一个多项式，使得问题（9）在$h = (h_1, \ldots, h_l)$ 下存在局部极小点 $x^\star$。那么以下两种情况之一成立：

1. 问题（9）在 $x^\star$ 处满足 Karush–Kuhn–Tucker 条件。
2. 点 $x^\star$ 位于 V 的奇异轨迹中。

证明. 若雅可比矩阵 $J(h)(x^\star)$ 的秩为 $n - \dim(V)$，则由引理 3 可知，问题（9）在 $x^\star$ 处满足 Karush–Kuhn–Tucker 条件。否则，我们有 $m_{n-d}(h)(x^\star) = 0$，因此 $x^\star$ 位于 V 的奇异轨迹中。这里，$m_r(h)(x)$ 是与 h 相关的雅可比矩阵 $J(h)(x)$ 的 $r \times r$ 阶子式构成的向量。 □

下面的定理说明，问题（9）的局部极小点属于算法 1 返回的哪些簇，以及问题（9）在这些点处满足哪些类型的最优性条件：

定理 6. 设 $h_0, h_1, ..., h_l$ 是 $\mathbb{R}[x]$ 中的多项式，且问题（9）在 $h = (h_1, \ldots, h_l)$ 下存在一个全局极小点 $x^\star$。设 A 和 B 是算法 1 以 $V = V(h)$ 为输入时的输出。则以下两种情况之一成立：

1. 存在 A 中的一个簇 U，其中 $q_1, \ldots, q_r \in \mathbb{R}[x]$ 是 $I(U)$ 的生成元，使得对于 $q = (q_1, \ldots, q_r)$， $x^\star$ 是问题的全局极小点

$$\min_{x \in V(q)} h_0(x) \tag{40}$$

并且问题（40）的 Karush–Kuhn–Tucker 条件在 $x^\star$ 处成立。

2. 存在由 $q = (q_1, \ldots, q_r)$ 定义的 B 中的一个簇 U，且满足 $q_j \in \mathbb{R}[x]$，使得 $x^\star$ 是问题（40）的全局极小点。

证明。根据引理 14 的最后一条陈述， $x^\star$ 属于$A \cup B$中的某个簇 U。根据引理 14 的第一条陈述， U 是一个不可约簇。设q_1，……， $q_r \in \mathbb{R}[x]$ 是消失理想 $I(U)$的生成元。那么，对于带有 $q = (q_1$,……， $q_r)$的优化问题 (40)， $x^\star$ 也是全局极小点。假设在 $x^\star$处，优化问题 (40) 的 Karush–Kuhn–Tucker 条件不成立。根据引理 15， $x^\star$位于 U的奇异轨迹 U^{sing}中。如果 U是零维的，那么 U 属于 B。否则，我们将 U^{sing} 分解为不可约分支（算法 1 的步骤 4 (b)

）。在这种情况下， $x^\star$ 属于U^{sing}的某个不可约分支。我们重复上述过程，直到只得到零维分支。因此结论得证。

□

注7. 根据算法 1 的步骤 4 (b) 和引理 2 的第一个陈述， U 中的每个 $A \cup B$ 都是由 $h_1, \ldots, h_l$ 的高阶偏导数构成的多项式所定义的不可约簇。

6 一般情形下的精确多项式优化

下面的算法使我们能够获得精确半定规划，以计算问题（9）的最小值 $h^\star$ 。

算法 2. 计算问题（9）的最小值 $h^\star$ 。

- 输入： $h_0, h_1, \ldots, h_l$ 属于 $\mathbb{R}[x]$。
- 输出： $\bar{h}^\star$

1. 将 A 和 B 设为以 $V = V(h)$ 为输入时算法 1 的输出，其中$h = (h_1, \ldots, h_l)$。

2. 设定 $T := \emptyset$。

3. 对于 A 中每个由 $q_1, \ldots, q_r \in \mathbb{R}[x]$ 生成的簇 U，其中它们是$I(U)$的生成元，执行:

   (a) 计算 $t := \rho_\xi(h_0, \bar{q}_{\mathrm{KKT}})$，其中 $q := (q_1, \ldots, q_r)$， $\bar{q} = (h_0, q)$， $d := \max\{\deg(h_0), \deg(q_j) : j = 1, \ldots, r\}$, w如 (37) 中所述，并且

   $$\xi := \frac{1}{2} \times b(n + r, 2w, r + n + 1) + d\,. \tag{41}$$

   (b) 设定 $T := T \cup \{t\}$。

4. 对于 B 中每个由 $q := (q_1, \ldots, q_r)$ 定义且满足 $q_j \in \mathbb{R}[x]$ 的簇 U，执行:

   (a) 计算 $t := \rho_\xi(h_0, q)$，其中 $d := \max\{\deg(h_0), \deg(q_j) : j = 1, \ldots, r\}$, w 如 (38) 中所示，并且

   $$\xi := \frac{1}{2} \times b(n, 2w, r + 1) + d\,. \tag{42}$$

   (b) 令 $T := T \cup \{t\}$。

5. 令 $h^\star := \min T$。

注 8. 算法 2 的步骤 3 (a) 和步骤 4 (a) 等价于精确求解半定规划（形式为 (29)）。关于求解半定规划，我们请读者参阅 Henrion、Naldi 和 Safey El Din 的精确算法，见 *[19]* 。此外，如果步骤 3 (a) 和步骤 4 (a) 中某个半定规划所关联的多项式优化问题的可行集为空，那么算法 2 中的集合 T 可能包含 ∞ 。

我们在下面的定理中陈述主要结果：

定理 7. 设 $h_0, h_1, \ldots, h_l \in \mathbb{R}_d[x]$。设 $h^\star$ 如问题（9）中所述，其中 $_h = (h_1, \ldots, h_l)$。假设问题（9）存在全局极小点。设 $h^\star$ 为算法 2 以 $h_0, h_1, \ldots, h_l$ 为输入时的输出，记作 $\bar{}$。则 $h^\star = h^\star$。

证明。设 $x^\star$ 是问题（9）的一个全局极小点。根据定理 6，以下两种情形之一成立：

- 情形 1：存在簇 U，它位于 A 中，其中 $q_1, \ldots, q_r \in \mathbb{R}[x]$ 是 $I(U)$ 的生成元，使得 $x^\star$ 是问题（40）的一个全局极小点，且满足$q = (q_1, \ldots, q_r)$，并且问题（40）的 Karush–Kuhn–Tucker 条件在 $x^\star$处成立。根据定理 3，我们得到 $\rho_\xi(h_0, \bar{q}_{\text{KKT条件}}) = h^\star$，其中 $\bar{q} = (h_0, q), d := \max\{\deg(h_0), \deg(q_j) : j = 0, \ldots, r\}$, w 如（37）中所述，ξ 如（41）中所述。
- 情形 2：存在一个簇 U，位于 B 中，由 $q = (q_1, \ldots, q_r)$ 定义，且满足$q_j \in \mathbb{R}[x]$，使得 $x^\star$ 是问题（40）的一个全局极小点。根据定理 4，我们得到 $\rho_\xi(h_0, q) = h^\star$，其中 $d := \max\{\deg(h_0), \deg(q_j) : j = 0, \ldots, r\}$,$w$ 如（38）中所述，ξ 如（42）中所述。

因此，根据算法 2 的步骤 3、4 和 5（以及注 4），即可得出该结果。

□

令 $\|\cdot\|_2$ 表示实向量的 l_2-范数。则 $\|x\|_2^2 = x_1^2 + \cdots + x_n^2$ 是关于 x的多项式。

注 9. 为了求出问题（9）的一个全局极小点，我们在 *[36,* 第 4 节*3]* 中应用算法 2 和添加球面约束方法，具体如下：使用算法 2 计算问题（9）的最小值 $h^\star$。令 $_h = (h, h_0 - h^\star)$，则 $V(h)$ 是问题（9）的全局极小点集。我们假设 $V(\bar{h})$ 非空，并在 $V(\bar{h})$ 中取一点 $x^\star$。令 $(a^{(t)})_{t=0}^n$为 $\mathbb{R}^n$ 中的有限点序列，使得 $a^{(t)} - a^{(0)}$、$t = 1$、.....、n 在 $\mathbb{R}^n$中线性无关。我们使用算法 2 计算以下数值：

$$\begin{aligned} \xi_0 &:= \min\{\|x - a^{(0)}\|_2^2 : x \in V(\bar{h})\}, \\ \xi_t &:= \min\{\|x - a^{(t)}\|_2^2 : x \in V(\bar{h}, b_0, \ldots, b_{t-1})\}, t = 1, \ldots, n. \end{aligned} \tag{43}$$

其中 $b_r = \xi_r - \|x - a^{(r)}\|_2^2$。那么存在唯一的实点 $x^\star$ in $V(h, b_0, \ldots, b_n)$，它满足非奇异线性方程组

$$(a^{(t)} - a^{(0)})^\top x^\star = -\frac{1}{2}(\xi_t - \xi_0 + \|a^{(0)}\|_2^2 - \|a^{(t)}\|_2^2),\ t = 1, \ldots, n. \tag{44}$$

我们通过以下示例来说明算法 2，该示例基于 [33,示例 18]:

例4. 设 $n = 5$、$h_0 = x_1 + x_2$、$h_1 = x_1^5 - x_3^2$、$h_2 = x_2^5 - x_4^2$ 和 $h_3 = -x_1x_2 - x_5^2$ 为算法 2 的输入。设 $h^\star$ 为问题（9）中的情形，其中 $h = (h_1, \ldots, h_l)$。那么不难证明 $h^\star = 0$。在算法 2 的步骤1中，我们得到

$$A = \{V(q^{(1)}), V(q^{(2)})\} \quad and \quad B = \{V(q^{(3)})\}, \tag{45}$$

其中 $q^{(1)} := (x_5, x_4, x_2, x_1^5 - x_3^2)$， $q^{(2)} := (x_5, x_3, x_2^5 - x_4^2, x_1)$，且 $q^{(3)} := (x_1, x_2, x_3, x_4, x_5)$。对于 $j = 1$，令 2，$\bar{q}^{(j)} := (h_0, q^{(j)})$，可得 $V(\bar{q}^{(j)}\ \text{KKT条件}) = \emptyset$。事实上，对于任意 $\lambda \in \mathbb{R}^4$，我们有

$$\begin{aligned} \nabla h_0(x) - \textstyle\sum_{i=1}^4 \lambda_j \nabla q_i^{(1)}(x) &= \begin{bmatrix} 1 - 5\lambda_4 x_1^4 \\ 1 - \lambda_3 \\ 2\lambda_4 x_3 \\ -\lambda_2 \\ -\lambda_1 \end{bmatrix}, \\ \nabla h_0(x) - \textstyle\sum_{i=1}^4 \lambda_j \nabla q_i^{(2)}(x) &= \begin{bmatrix} 1 - \lambda_4 \\ 1 - 5\lambda_3 x_2^4 \\ -\lambda_2 \\ 2\lambda_3 x_4 \\ -\lambda_1 \end{bmatrix}. \end{aligned} \tag{46}$$

那么，$\bar{q}^{(1)}\ \text{KKT条件}(x) = 0$ 意味着 $x_1^5 = x_3^2$，$1 - 5\lambda_4 x_1^4 = 2\lambda_4 x_3 = 0$，这不可能。此外，$\bar{q}^{(2)}\ \text{KKT条件}(x) = 0$ 意味着 $x_2^5 = x_4^2$，$1 - 5\lambda_3 x_2^4 = 2\lambda_3 x_4 = 0$，这不可能。对于 $j = 1, 2$，由于 $V(\bar{q}^{(j)}\ \text{KKT条件}) = \emptyset$，当 $\xi \in \mathbb{N}$ 足够大时，成立 $\rho_\xi(h_0, \bar{q}^{(j)}\ \text{KKT条件}) = \infty$。由于 $V(q^{(3)}) = \{0\}$，当 $\xi \in \mathbb{N}$ 足够大时，我们得到 $\rho_\xi(h_0, q^{(3)}) = 0$。因此最终得到 $T = \{\infty, 0\}$，从而有 $h^\star = \min T = 0 = h^\star$。

注 10. 我们可以完全依靠符号计算 *[31]* 来求出问题（9） $h^\star$ 的最小值，前提是该问题存在全局极小点 $x^\star$，而不必使用半定规划。（第一作者提出的方法 *[31]* 依赖于实根基生成元和 Groeb- ner 基的计算。该方法使我们能够求出单变量多项式，用以定义基本半代数集 S 在多项式 f下的像的 Zariski闭包。）为此，设 A 和 B 为以 $V = V(h)$ 为输入运行算法 1 的输出。根据定理 5，对于每个 U属于 A 且 q_1^U、……、$q_{r_U}^U \in \mathbb{R}[x]$ 是 $I(U)$ 的生成元时， $h_0(V(\bar{q}_{\text{KKT}}^U))$ 只有有限个取值，其中 $\bar{q}^U = (h_0, q_1^U, \ldots, q_{r_U}^U)$。此外，对于每个 U 属于 B， $h_0(U)$ 只有有限个取值。利用 *[31,* 算法 *1],*，我们可以计算出每个 U 属于 A（分别为 B）时， $h_0(V(\bar{q}_{\text{KKT}}^U))$（分别为 $h_0(U)$）的所有取值。对每个 U 属于 A（分别为 B），设 $h_U^\star$ 为 $h_0(V(\bar{q}_{\text{KKT}}^U))$（分别为 $h_0(U)$）中的最小值。注意，对于每个 U 属于 A（分别为 B），都有$V(\bar{q}_{\text{KKT}}^U) \subset V(h) \times \mathbb{R}^{r_U}$（分别为 $U \subset V(h)$，因此 $h_U^\star \geq h^\star$。定理 6 表明， $x^\star$ 属于某个 $V(\bar{q}_{\text{KKT}}^U)$ 的投影，其中 U 属于 A，或者 $x^\star$ 属于某个 B中的 U，这意味着 $h^\star = \min\{h_U^\star : U \in A \cup B\}$。

注 11. 如果我们将算法 1 的步骤 2 替换为“设置 $r := 1$ and $U_1 := U$。”，那么目前为止的所有结果仍然成立。不过，在这种情况下，定义 U_1 的多项式次数可能非常高（如例 2、3 和 4 所示）。

注 12. 定理 7 的结果要求下确界值 $h^\star$ 可达到。在 *[35],* 中，第一作者证明，任何具有有限下确界值 $h^\star$ 的形式为 (9) 的多项式优化问题，都可以通过符号变换转化为一维空间中的等价问题，且其最优值 $h^\star$ 可达到。为此，他使用了量词消去，以及依赖代数学基本定理和最大公因子的代数算法。设 d 为 h_j 的次数上界。他的符号算法

具有复杂度 $O(d^{O(n)})$，可生成等价问题中次数至多为 $d^{O(n)}$ 的目标多项式和约束多项式。

下面的推论说明了多项式优化问题与半定规划之间的等价性：

推论 1. 设 $h_0, h_1, \ldots, h_l \in \mathbb{R}_d[x]$。设 $h^\star$ 如问题（9）中所示，其中$h = (h_1, \ldots, h_l)$。那么，存在一个符号算法，以 $h_0, \ldots, h_l$ 为输入，并生成一个有限的半定规划序列，其最小值为 $(\zeta_r)_{r=1}^s$，使得（9）等价于最小值最小的半定规划 $\zeta_{\bar{r}} = \min_{r=1,\ldots,s} \zeta_r$。

证明. 若 $h^\star \in \{-\infty, \infty\}$，则不难构造一个下确界值为 $h^\star$的半定规划。若 $h^\star$ 有限但无法取到，根据注 12，我们可以将问题（9）转化为一个等价的多项式优化问题，使其最小值 $h^\star$ 可以取到。下面考虑剩余情形。假设（9）存在全局极小点。由定理 7 可知，算法 2 会生成有限个半定规划，其中最小的最小值恰为$h^\star$。因此结论成立。

□

注 13. 本文（尤其是推论 1）从理论上概述了为什么每个一般的多项式优化问题都可以通过递归求得的一系列半定规划来求解。然而，由于算法 2 中的界 (41) 和 (42) 非常宽泛，我们的方法在实践中效率应该很低。这是不可避免的，因为我们使用符号计算来获得结果。因此，我们不能指望本方法像近似方法那样高效。

7 双曲规划与半定规划的等价性

若多项式 $f \text{in } \mathbb{R}[x]$ 关于 $e = (e_1, \ldots, e_n) \in \mathbb{R}^n \text{if} f(e) \neq 0$ 是双曲的，且对于所有 $a \in \mathbb{R}^n$，单变量多项式 $t \mapsto f(te - a)$ 都只有实根，则称其为双曲多项式。

给定一个关于 $e \in \mathbb{R}^n$ 双曲的多项式 f，我们定义集合

$$\Lambda_+(f, e) := \{a \in \mathbb{R}^n : f(te - a) = 0 \Rightarrow t \geq 0\}, \tag{47}$$

这称为 f 沿方向 e 的双曲锥。

正如 G˚arding 在 [13] 中所示，$\Lambda_+(f, e)$ 是一个基本半代数集。下面的引理陈述了一个弱于该结论的结果：

引理 16. 设 f 是关于 $e \in \mathbb{R}^n$ 的 $\mathbb{R}[x]$ 双曲多项式。那么 $\Lambda_+(f, e)$ 是一个基本半代数集的投影，该集合由基于 e 和 f 的系数构造的多项式定义。

证明。设 t 是一个变量。令 $d := \deg(f)$，且 $\varphi_a(t) = f(te - a)$。那么 φ_a 是一个次数至多为 d 的 t 多项式。分别用 $\varphi_a^{(j)}$、$j = 0$、……、d 表示与单项式 t^j、$j = 0$、……、d 相关联的 φ_a 的系数。那么，每个映射 $a \mapsto \varphi_a^{(j)}$ 都是 a 的多项式。令 $a \in \Lambda_+(f, e)$。那么，恒等多项式 t 在零维代数簇上非负

$V(\varphi_a)$。假设对于某个 $r \in \mathbb{N}$， $V(\varphi_a) = \{t_1, \ldots, t_r\}$。那么有 $t_j \geq 0$、$j = 1$、.....、r。由于 φ_a 的次数至多为 d，我们得到 $r \leq d$。令

$$p_j(t) := \prod_{\substack{i=1 \\ i \neq j}}^{r} \frac{t - t_i}{t_j - t_i}\,, \; j = 1, \ldots, r\,, \tag{48}$$

我们得到 $p_j(t_i) = \delta_{ij}$. 令 $\sigma = \sum_{j=1}^{r} t_j p_j^2$，得到 $\sigma \in \Sigma_{r-1}^2[t]$. 此外，$t - \sigma$ 在 $V(\varphi_a)$ 上为零。由于 $r \leq d$， σ 的次数至多为 $2(d-1)$。这意味着 φ_a 和 σ 的次数都至多为 $2d$。令 $\xi := \frac{1}{2} \times b(1, 2d, 2)$ 且$s := 2\lfloor \xi/(2d) \rfloor$，其中 $b(\cdot)$ 的定义见 (20)。根据引理 9，有

$$\exists \sigma \in \Sigma_{d-1}^2[t] \,:\, -(t-\sigma)^s \in \Sigma_{\xi}^2[t] + I_{\xi}(\varphi_a)[t] \tag{49}$$

注意， s 和 ξ 仅取决于 d。由此，(49) 等价于“$t \geq 0$ 在$V(\varphi_a)$上”。此外，(49) 可以写成

$$\begin{cases} \exists G_j \succeq 0\,, \exists u \in \mathbb{R}^{2(\xi-d)+1} \,: \\ -(t - v_{d-1}^{\top} G_0 v_{d-1})^s = v_{\xi}^{\top} G_1 v_{\xi} + v_{2(\xi-d)}^{\top} u \varphi_a\,, \end{cases} \tag{50}$$

其中 v_η 是 t 中次数不超过 η的单项式构成的向量。不难看出，(50) 等价于

$$\begin{cases} \exists G_0 \in \mathbb{R}^{d \times d}\,, \exists G_1 \in \mathbb{R}^{(\xi+1)\times(\xi+1)}\,, \exists u \in \mathbb{R}^{2(\xi-d)+1} \,: \\ G_j^{\top} = G_j\,,\, F(G_j) \geq 0\,,\, H(G_0, G_1, u, \varphi_a^{(0)}, \ldots, \varphi_a^{(d)}) = 0\,, \end{cases} \tag{51}$$

其中 F 表示矩阵的主子式向量， H 是一个多项式向量。由于 $\varphi_a^{(j)}$是关于 a的多项式，(51) 等价于 a 属于一个基本半代数集在关于 a的坐标上的投影，这便得到了所需结果。

□

最后，我们陈述广义拉克斯猜想的一个推论（参见例如 [2,注 1.7]）：

推论 2. 设 f 是关于 $e \in \mathbb{R}^n$的 $\mathbb{R}[x]$ 双曲多项式。考虑以下双曲规划:

$$\begin{array}{ll} \inf\limits_{x \in \mathbb{R}^n} & c^{\top} x \\ s.t. & Ax = b\,,\, x \in \Lambda_+(f, e)\,, \end{array} \tag{52}$$

其中 $A \in \mathbb{R}^{m \times n}$、 $b \in \mathbb{R}^m$ 和 $c \in \mathbb{R}^n$ 是给定的。因此，存在一个符号算法，以 A、b、c、f、e 为输入，生成一列最小值为 $(\zeta_r)_{r=1}^{s}$ 的半定规划，使得 (52) 等价于最小值最小的半定规划 $\zeta_{\bar{r}} = \min_{r=1,\ \ldots,\ s} \zeta_r$。

证明。设 $h^\star$ 为问题（52）的下确界值。根据引理 16，问题（52）可写成最小值为 $h^\star$的多项式优化问题（9）（见注 1）。注意，目标多项式和约束多项式 $h_0, \ldots, h_l$由 A, b, c, f, e构成。根据推论 1，存在一个符号算法，以 $h_0, \ldots, h_l$ 为输入，并生成一个有限的半定规划序列，使得（9）等价于其中最小值最小的半定规划。因此结论得证。

□

注 14。推论 2 的框架提供了一种构造性方法，可为给定 A, b, c, f, e 的双曲规划（52）得到一个等价的半定规划。然而，不容易看出推论 2 的结果能够解决广义拉克斯猜想。这是因为双曲规划与半定规划等价，并不一定直接意味着它们的可行集相同。

致谢。作者谨向 Bernard Mourrain、MarkusSchweighofer 和 Claus Scheiderer 致谢，感谢他们就本文展开的宝贵讨论。

第一作者获得了 ANITI 的资助。第二作者获得了 Tremplin ERC Stg Grant ANR-18-ERC2-0004-01（T-COPS 项目）、FMJH Program PGMO（EPICS 项目）以及 EDF、Thales、Orange 和 Criteo 的资助。本研究还受益于欧盟地平线 2020 研究与创新计划下玛丽·斯克沃多夫斯卡-居里行动的资助，协议编号 813211（POEMA）；同时也获得了 AI 跨学科研究院 ANITI 的资助，该资助由法国“面向未来投资 PIA3”计划提供，资助协议编号为 n°ANR-19-PI3A-0004。

参考文献

[1] S. Albrecht 和 M. Ulbrich。运动逆最优控制中的互补约束数学规划。《优化方法与软件》，32(4):670–698，2017年。[2] N. Amini。多元匹配多项式的双曲性锥的谱半代数性。《代数组合学杂志》，50(2):165–190，2019年。[3] L. Baldi 和 B. Mourrain。多项式优化中的精确矩表示。arXiv 预印本 arXiv:2012.14652，2020年。[4] L. Baldi 和 B. Mourrain。通过矩优化计算实根基。载于《2021 年国际符号与代数计算研讨会论文集》，第 43–50 页，2021年。[5] E. Becker 和 R. Neuhaus。多项式理想实根基的计算。载于《计算代数几何》，第 1–20 页。施普林格，1993年。[6] R. Bellman 和 K. Fan。关于厄米矩阵变量线性不等式组的问题。《凸性》，7:1–11，1963年。[7] M. A. Bucero 和 B. Mourrain。半代数集上多项式优化的精确松弛。arXiv 预印本 arXiv:1307.6426，2013。[8] S. Burgdorf、C. Scheiderer 和 M. Schweighofer。纯态、非负多项式与平方和。《瑞士数学评论》，87(1):113–140，2012年。[9] M. Coste。半代数几何导论，2000年。

[10] D. Cox、J. Little 和 D. OShea。《理想、簇与算法：计算代数几何与交换代数导论》。Springer Science & Business Media，2013。

[11] J. Demmel、J. Nie 和 V. Powers。通过 KKT 理想表示非紧半代数集上的正多项式。《纯粹与应用代数学杂志》，209(1):189–200，2007。[12] R. M. Freund。约束优化问题的最优性条件。麻省理工学院（链接见此），2004年。[13] L. G˚arding。关于双曲多项式的一个不等式。《数学与力学杂志》，第957–965页，1959年。[14] P. Gianni、B. Trager 和 G. Zacharias。多项式理想的 Gr¨obner 基与主分解。《符号计算杂志》，6(2- 3):149–167，1988年。[15] M. X. Goemans 和 D. P. Williamson。利用半定规划改进最大割与可满足性问题的近似算法。《ACM 会刊》（JACM），42(6):1115–1145，1995年。[16] O. G¨uler。用于凸规划的双曲多项式与内点法。《运筹学数学》，22(2):350–377，1997年。[17] J. D. Hauenstein 和 C. W. Wampler。等奇异集与去膨胀。《计算数学基础》，13(3):371–403，2013。[18] J. W. Helton 和 V. Vinnikov。集合的线性矩阵不等式表示。《纯粹与应用数学通讯：库朗数学科学研究所发行的期刊》，60(5):654–674，2007年。[19] D. Henrion、S. Naldi 和 M. Safey El Din。退化可行集上半定规划的精确算法。载于《2018 ACM 符号与代数计算国际研讨会论文集》，第191–198页，2018年。.[20] D. Hilbert。关于定定形式表示为平方形式之和。《数学年刊》，32(3):342–350，1888年。[21] H. Hironaka。零特征域上代数簇的奇点消解：I。《数学年刊》，第109–203页，1964年。[22] H. Hironaka。零特征域上代数簇的奇点消解：II。《数学年刊》，第205–326页，1964年。[23] J.-L. Krivine。预序环。《数学分析杂志》，12(1):307–326，1964年。[24] M. Kummer。行列式表示与贝祖矩阵。《数学杂志》，285(1):445–459，2017年。[25] M. Kummer、S. Naldi 和 D. Plaumann。平面双曲曲线的谱半代数表示。《太平洋数学杂志》，303(1):243–263，2019年。[26] M. Kummer、D. Plaumann 和 C. Vinzant。双曲多项式、交错多项式与平方和。《数学规划》，153(1):223–245，2015年。

[27] P. Lairez 和 M. Safey El Din. 实代数集维数的计算。载于《2021年国际符号与代数计算研讨会论文集》，第257–264页，2021年。[28] J. B. Lasserre. 多项式全局优化与矩问题。《SIAM优化杂志》，11(3):796–817，2001年。[29] J. B. Lasserre. 多项式与半代数优化导论，第52卷。剑桥大学出版社，2015年。[30] H. Lombardi、D. Perrucci 和 M.-F. Roy. 有效正定性定理与希尔伯特第十七问题的初等递归界，第263卷。美国数学会，2020年。[31] N. H. A. Mai. 一种结合 Fritz John 条件强化的精确多项式优化符号算法。arXiv 预印本 arXiv:2206.02643，2022年。[32] N. H. A. Mai. 结合 Fritz John 条件强化的精确多项式优化复杂度。arXiv 电子预印本，第 arXiv–2205 页，2022年。[33] N. H. A. Mai. 结合 Fritz John 条件强化的精确多项式优化。arXiv 预印本 arXiv:2205.04254，2022年。[34] N. H. A. Mai. 正则点集稠密条件下奇异簇上的非负定理。arXiv 预印本 arXiv:2211.12440，2022年。[35] N. H. A. Mai. 多项式作用下半代数集像的闭包的半代数描述。arXiv 预印本 arXiv:2210.13933，2022年。[36] N. H. A. Mai、J.-B. Lasserre 和 V. Magron. 非紧半代数集上的正性证书与多项式优化。《数学规划》，第1–43页，2021年。[37] M. Marshall. 具有有限个零点的非负多项式的表示。载于《图卢兹科学学院年鉴：数学》，第 ´15卷，第599–609页，2006年。 ˆ[38] M. Marshall. 非负多项式的表示、次数界及其在优化中的应用。《加拿大数学杂志》，61(1):205–221，2009年。[39] S. Naldi 和 D. Plaumann. 双曲规划中的符号计算。《代数及其应用杂志》，17(10):1850192，2018年。[40] T. Netzer、D. Plaumann 和 A. Thom. 行列式表示与埃尔米特矩阵。《密歇根数学杂志》，62(2):407–420，2013。[41] T. Netzer 和 A. Thom. 有行列式表示与无行列式表示的多项式。《线性代数及其应用》，437(7):1579–1595，2012年。[42] R. Neuhaus. 多项式理想实根基的计算——II。《纯粹与应用代数学杂志》，124(1-3):261–280，1998年。

[43] J. Nie. 多项式优化的精确雅可比半正定规划松弛。《数学规划》，137(1):225–255，2013年。[44] J. Nie. 实代数簇上的多项式优化。《SIAM优化杂志》，23(3):1634–1646，2013年。[45] J. Nie. 最优性条件与 Lasserre层级的有限收敛。《数学规划》，146(1-2):97–121，2014年。[46] OSCAR – 开源计算机代数研究系统，版本0.9.0，2022年。[47] T. S. Pham 和 H. H. Vui. 多项式优化中的一般性，第3卷。世界科学出版社，2016年。[48] M. Putinar. 紧半代数集上的正多项式。《印第安纳大学数学杂志》，42(3):969–984，1993年。[49] C. Scheiderer. 实代数簇上正则函数的平方和。《美国数学会会刊》，352(3):1039–1069，2000年。[50] C. Scheiderer. 实代数曲线上的平方和。《数学杂志》，245(4):725–760，2003年。[51] C. Scheiderer. 实代数曲面上的平方和。《数学手稿》，119(4):395–410，2006年。[52] K. Schm¨udgen. 紧半代数集的 K-矩问题。《数学年刊》，289(1):203–206，1991年。[53] M. Schweighofer. 双曲性锥的谱半定松弛。arXiv 预印本 arXiv:1907.13611，2019年。[54] A. Sheffer. 多项式方法与关联理论，第197卷。剑桥大学出版社，2022年。[55] M. Sitharam、A. S. John 和 J. Sidman. 几何约束系统原理手册。Chapman and Hall/CRC，2018年。[56] K. Smith、L. Kahanp¨a¨a、P. Kek¨al¨ainen 和 W. Traves. 代数几何入门。Springer Science & Business Media，2004年。