

密级 _____



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

多项式优化和凸代数几何中的若干问题研究

作者姓名 _____ 王础

指导教师 _____ 支丽红 研究员

中国科学院数学与系统科学研究院

学位类别 _____ 理学博士

学科专业 _____ 应用数学

培养单位 _____ 中国科学院数学与系统科学研究院

2016 年 5 月

Some topics in polynomial optimization and convex algebraic geometry

By
Chu Wang

A Dissertation Submitted to
University of Chinese Academy of Sciences
In partial fulfillment of the requirement
For the degree of
Doctor of Science

Academy of Mathematics and Systems Science
Chinese Academy of Sciences

May, 2016

摘要

多项式优化在凸优化, 代数几何和图论等领域具有广泛的应用. 多项式优化问题可以松弛为基于矩量矩阵的半正定规划问题序列. 这个半正定规划问题序列称为矩量矩阵松弛序列. 矩量矩阵松弛问题的对偶问题是平方和松弛问题. 当可行集有界时, 矩量矩阵松弛和平方和松弛问题的最优值收敛到多项式优化问题的最优值. 在实代数几何中, 基于矩量矩阵的半正定规划方法可以用来计算理想的实根理想.

代数几何理论可以用来研究多项式优化问题的代数结构. 一些符号算法可以计算多项式优化问题的最优值. 然而, 对于非负多项式的结构, 半代数集表示的复杂性和高效性的相互关系, 经典的代数方法面临着挑战. 凸性提供了一个新的观点和框架来解决这些问题. 这个领域称为凸代数几何. 在凸代数几何中, 有很多重要的问题被深入研究. 例如通过基于矩量矩阵的谱多面体来近似半代数集的闭凸包问题, 凸半代数集和谱多面体投影的等价表示问题以及通过代数次数衡量多项式优化问题的代数复杂度问题等.

本文主要研究凸代数几何以及多项式优化中的若干问题. 第一个问题是给定实系数多项式环中的理想, 如何计算它的实根理想. 一个基于准素分解的符号方法可以用来计算它的实根理想. 当它的实根理想是零维时, 一个基于矩量矩阵松弛的数值算法可以求解它的实根理想. 平坦扩张条件是这个算法的终止条件. 当它的实根理想是正维时, 平坦扩张条件不满足. 基于偏微分方程中的对合理论, 我们提出一个新的终止条件. 当终止条件满足时, 算法返回一个介于它和它的实根理想之间的理想的Pommaret基. 在一些假设下, 算法返回它的实根理想的生成元.

第二个问题是给定一个基本闭半代数集, 如何计算它的闭凸包. 当半代数集有界时, Theta体序列和Lasserre半正定表示序列可以收敛到它的闭凸包. 当半代数集无界时, Theta体序列和Lasserre半正定表示序列的收敛性无法保证. 我们将半代数集映射到一个高维空间中的半代数集, 然后将它投影到这个高维空间的单位球上, 进而将无界半代数集的凸包求解问题转换成为有界的情形. 我们构造了推广Theta体序列和推广Lasserre半正定表示序列. 假设半代数集在无穷远处是闭的并且它在高维空间中的像所生成的闭凸锥是尖的闭凸锥, 我们

构造的推广Theta体序列和推广Lasserre半正定表示序列可以收敛到半代数集的闭凸包. 这两个假设条件对收敛性是必要的. 如果有一个假设条件不满足, 收敛性可能不成立.

第三个问题是参数化多项式优化问题的求解. 在参数化多项式优化问题中, 对任意的参数值, 相应的多项式优化问题存在一个最优值. 利用实代数几何中的Tariski定理可以说明最优值是参数值的半代数函数. 这个半代数函数记作最优值函数. 如何利用多项式方程来描述最优值函数? 当优化问题的可行集不可约, 光滑并且有界时, 对偶仿射簇可以用来求解最优值函数图像的Zariski闭包. 当可行集有界并且不光滑时, 我们构造了对偶仿射簇序列. 序列中对偶仿射簇的并可以用来包含最优值函数的图像. 当可行集光滑, 无界并且可行集的闭凸包不包含直线时, 通过广义关键值理论, 我们仍然可以利用对偶仿射簇来描述最优值函数图像的Zariski闭包. 我们设计一个基于极仿射簇的算法的参数化形式. 这个参数化算法返回一个多项式序列, 使得对任意的参数值, 在序列中存在一个多项式可以被约化为非零的单变元多项式并且它的根包含特定化优化问题的最优值.

最后一个问题刻画一个闭凸集的锥举起. 我们将凸体的锥举起定理扩展到一般的闭凸集情形. 给定一个闭凸集, 根据它是否满维, 是否是一个锥的平移, 是否包含直线, 我们定义相应的松弛算子. 我们推广了锥举起的定义: 闭凸集不仅是一个给定闭凸锥的仿射切片在线性映射下的像, 而且它的回收锥是闭凸锥的线性切片在相同线性映射下的像. 给定一个闭凸集, 我们利用推广松弛算子的锥分解来刻画它的推广锥举起.

关键词: 矩量矩阵; Lasserre松弛; Theta体; 谱多面体; 对偶仿射簇; 极仿射簇; 锥分解; 非负秩; 半正定秩; 锥举起.

Abstract

Polynomial optimization has a lot of applications in convex optimization, algebraic geometry, graph theory and so on. It can be relaxed to hierarchies of semidefinite optimization problems which are called moment relaxations. The dual of moment relaxations are sum of squares(SOS) relaxations. The optimal value of moment relaxations and SOS relaxations can converge to the optimal value of the polynomial optimization when the feasible set is bounded. In real algebraic geometry, moment relaxations can be used to compute the real radical of an ideal.

Algebraic geometry plays an important role in studying the algebraic structures of polynomial optimization problems. There are some symbolic algorithms for the computation of the optimal value in a polynomial optimization problem. However, classical algebraic methods will meet challenges when we deal with some topics, such as the structure of nonnegative polynomials and the interplay between efficiency and complexity of different representations of semialgebraic sets. Convexity provides a new viewpoint and framework for addressing these questions. This emerging area has become known as convex algebraic geometry. In convex algebraic geometry, there are a lot of important problems that have been studied deeply. For example, computing the closed convex hull of a semialgebraic set by constructing explicit semidefinite representations based on moment matrices, the equivalence between a convex semialgebraic set and the projected spectrahedron and measuring the algebraic complexity of an optimization problem by using algebraic degree.

This thesis focuses on studying some problems in polynomial optimization and convex algebraic geometry. The first problem is that given an ideal whose background field is real field, how to compute its real radical. A symbolic algorithm based on the primary decomposition can be used to compute its real radical. When its real radical is zero dimensional, a new algorithm based on moment relaxations can be used to compute the generators of its real radical. Flat

extension condition has been used as a termination criterion of this algorithm. When its real radical is positive dimensional, flat extension condition can not be satisfied. We propose a new termination criterion based on the involution theory in solving differential equations. When our termination criterion is satisfied, the algorithm returns a Pommaret basis of an ideal between the ideal and its real radical. Under some assumptions, it returns the generators of its real radical.

The second problem is that given a semialgebraic set, how to compute its closed convex hull. When the semialgebraic set is compact, hierarchies of Theta bodies and Lasserre's semidefinite representations converge to its closed convex hull. When the semialgebraic set is not compact, hierarchies of Theta bodies and Lasserre's semidefinite representations may not converge. We map the semialgebraic set to another semialgebraic set in a higher dimensional linear space and then project it to the unit sphere in the higher dimensional linear space so that the computation of the closed convex hull for a noncompact semialgebraic set can be transferred to that for a compact one. We construct hierarchies of modified Theta bodies and modified Lasserre's semidefinite representations. Assuming that the semialgebraic set is closed at infinity and the closed convex cone generated by its image in the higher dimensional linear space is closed and pointed, we prove that hierarchies of modified Theta bodies and modified Lasserre's semidefinite representations converge to its closed convex hull. Both assumptions are necessary for the convergence and the convergence will not be guaranteed if one of them is not satisfied.

The third problem is the computation of the parameterized optimization problem. In the parameterized optimization problem, for each parameter's value, there exists an optimal value in the corresponding optimization problem. Tariski's theorem in real algebraic geometry ensures that the optimal value is a semialgebraic function of the parameter's value so that it is called optimal value function. How to describe the optimal value function by polynomial equations? Dual variety can be used to compute the Zariski closure of the graph of the optimal value function when the feasible set is irreducible, compact and smooth. When the feasible set is compact but not smooth, we construct a sequence of dual varieties and the union of these dual varieties can be used to contain the

graph of the optimal value function. When the feasible set is smooth but not compact and the closed convex hull of the feasible set contains no lines, we show that the dual variety can be used to compute the Zariski closure of the graph of the optimal value function by the theory of generalized critical value. We propose a parameterized version of an algorithm which is based on the polar variety and the parameterized algorithm returns a sequence of polynomials such that for each parameter's value, there exists a polynomial in the sequence which can be reduced to a nonzero univariate polynomial whose roots contain the optimal value in the specialized optimization problem.

The last problem is characterizing the cone lift of a closed convex set. We generalize the factorization theorem of convex body to the case of general closed convex sets. Given a closed convex set, we define a slack operator associated to it according to whether the convex set is full dimensional, whether it is a translated cone and whether it contains lines. We generalize the definition of cone lift: not only the convex set is the image of a linear map of a given closed convex cone's affine slice, but also its recession cone is the image of the linear map of the closed convex cone's linear slice. Given a closed convex set, we use the cone factorization of the generalized slack operator to characterize its generalized cone lift.

Key words: Moment matrix; Lasserre relaxation; Theta body; spectrahedra; dual variety; polar variety; cone factorization; nonnegative rank; semidefinite rank; cone lift.

目 录

摘要	i
Abstract	iii
目录	vii
第一章 引言	1
1.1 问题概述	1
1.2 论文的结构和主要成果	4
第二章 预备知识	5
2.1 凸分析基础	5
2.1.1 凸集	5
2.1.2 凸锥	7
2.1.3 凸函数	8
2.2 矩量矩阵	8
2.3 代数几何基础	9
2.3.1 基本概念	9
2.3.2 平方和多项式	10
第三章 实根理想的计算	13
3.1 问题和背景	13
3.2 预备知识	14
3.2.1 矩量矩阵的性质	14
3.2.2 对合除法和对合基	15
3.3 终止条件和主要结果	18
3.3.1 终止条件	18

3.3.2 一个计算Pommaret基的算法	19
3.3.3 终止条件的证明	20
3.3.4 推广到 $V(I_{\mathbb{R}}(I) \cap \mathcal{A})$	25
3.4 数值试验	26
3.5 总结评论	34
第四章 凸代数几何	37
4.1 问题背景	37
4.2 无界凸集的半正定表示	41
4.2.1 近似 $\text{cl}(\text{co}(S))$ 的闭凸包	41
4.2.2 $\text{cl}(\text{co}(S))$ 的谱多面体近似	47
4.3 假设2的详细讨论	51
4.3.1 S 在无穷远处是闭的	51
4.3.2 $\text{co}(\text{cl}(\tilde{S}^{\circ}))$ 的尖性	53
第五章 参数化优化问题的求解	57
5.1 问题背景	57
5.2 预备知识	59
5.2.1 对偶仿射簇	59
5.2.2 广义关键值	59
5.3 无界情形的对偶性	60
5.3.1 光滑无界情形	60
5.3.2 奇异有界的情形	62
5.3.3 对偶仿射簇的坏参数	64
5.4 完整的算法	65
5.4.1 背景	65
5.4.2 极仿射簇	65
5.4.3 子函数	68
5.4.4 参数化优化问题算法的基本步骤	71
5.4.5 坏参数的算法	72

第六章 凸集的锥举起	75
6.1 介绍	75
6.2 预备知识	76
6.3 一个无界凸集的锥举起	77
6.3.1 C 是满维的情形	77
6.3.2 C 不是满维的	91
6.4 多面体的锥举起	92
6.4.1 K 是一个多面体锥	93
6.4.2 K 是一个半正定锥	95
6.4.3 判定多面体的松弛矩阵	96
第七章 结论与展望	99
参考文献	101
发表文章目录	113
致谢	115

插 图

4.1 例子4.1中的半代数集 S 和切线 l .	39
4.2 例子4.1中, 左:由 S 生成的锥 \tilde{S}^c 和由 l 生成的超平面 \tilde{l} ; 右:单位球上 的一个有界部分 \tilde{S} (被蓝色的线分割.)	42
4.3 例子4.1中, $\text{cl}(\text{co}(S))$ 的谱多面体近似 $\tilde{\Omega}_3(\tilde{G})$ (用阴影部分表示) ...	48
4.4 例子4.8中的半代数集 S (红色曲线)和谱多面体 $\tilde{\Omega}_3(\tilde{G})$ 的投影(用阴 影部分表示).	49
4.5 例子4.8中 $\text{cl}(\text{co}(S)) \cap \mathbf{B}_r$ 的Lasserre半正定表示.	50
4.6 例子4.9中的曲线 S (红色部分)和谱多面体 $\tilde{\Omega}_3(\tilde{G})$ (用阴影部分表 示).	53
4.7 例子4.10中的 $S'_{(0,0)}$ 的谱多面体近似 $\tilde{\Omega}_3\left(\tilde{G}'_{(0,0)}\right)$ (用阴影部分表示) ..	56
5.1 例子5.2中的星形线.....	63
6.1 例子6.1	78
6.2 例子6.1	80
6.3 例子6.3	86
6.4 例子6.4	94

第一章 引言

1.1 问题概述

多项式优化问题是经典的优化问题. 它的数学形式如下:

$$\inf_{x \in \mathbb{R}^n} f(x) \quad \text{使得 } g_1(x) \geq 0, \dots, g_m(x) \geq 0,$$

其中, $f(x)$ 和 $g_1(x), \dots, g_m(x)$ 是多项式. 线性规划, 稳定集问题, 极大割问题都是多项式优化问题. 由于多项式函数是光滑函数, 可以利用经典的数值方法, 例如牛顿法, 共轭梯度法等方法来进行局部最优值的数值求解. 由于多项式优化问题不是凸优化问题, 所以无法确保这些数值方法可以求解全局最优值. 一个想法是将这个问题转化或松弛成凸优化问题. 基于这个想法, 已有大量的工作. Lasserre等学者利用测度理论, 矩量矩阵和多项式平方和理论构造一个半正定松弛序列, 参见 [48]. 这个半正定松弛序列的最优值收敛到原多项式优化问题的最优值.

多项式方程组的零点求解问题也是非常重要的问题. 它和多项式优化问题密切相关. 当多项式方程组根的个数有限时, 可以利用同伦方法 [1] 来进行数值求解. 符号方法主要包括Gröbner基 [15], Ritt-Wu特征列方法 [75, 99]. 通过求解多项式方程组中多项式的Gröbner基, 我们可以将问题转化为单变元多项式的零点求解问题. 当多项式方程组根的个数无限时, 求解所有的根是不可能的. 在这个情形下, 给定一个实系数多项式方程组, 有两个重要的问题值得考虑: 第一个问题是如何在对应实代数集的每一个连通分支上取得有限多个样本点. 第二个问题是如何求解实系数多项式环中的理想的实根理想. 对于第一个问题, 一些基于关键点的方法可以用来求解这些样本点, 见 [3, 4, 8, 81]. 对于第二个问题, 已有一些符号方法来求解理想的实根理想. 如果实根理想是零维的, 文献 [26] 的结果可以求解理想的实根理想. 如果实根理想是正维的, 普遍使用的方法将正维情形转化成零维情形. 参见Gianni, Trager 和Zacharias在文献 [26] 中的结果以及Krick和Logar在文献 [45] 中的结果. Becker和Neuhaus在文献 [10] 中提出了一个基于准素分解的符号方法来计算一个理想的实根理想(参见 [61, 93, 100, 102]).

一个基于矩量矩阵的数值方法在文献 [49–51, 54] 中被提出. 当一个理想 I 对应的实代数集是零维时, 这个方法可以用来求解 I 的实根理想 $\sqrt{\mathbb{R}I}$ 以及实代数集

的所有点. 基于矩量矩阵的方法还可以用来求解零维半代数集的 \mathcal{A} -根理想. 当理想 I 对应的实代数集是正维时, 由于基于矩量矩阵的数值方法缺少一个收敛的判定准则, 所以不能直接用来求解 I 的实根理想.

给定一个基本闭半代数集 S . 如何刻画 S 的闭凸包是组合优化问题中的一类重要问题并且已经被广泛研究(见 [11, 40, 41, 66]). 当 S 是一个有界的实代数集时, Ranestad和Sturmfels在文献 [72]中利用代数几何的方法来计算包含 S 的闭凸包边界的仿射簇. Rostalski和Sturmfels在文献 [79]中利用射影簇来刻画 S 的闭凸包的极集的边界. 在相同情形下, 文献 [30]和 [47]分别利用Theta体序列和Lasserre半正定表示序列来逼近 S 的闭凸包.

考虑下面的参数化优化问题:

$$\begin{aligned} c_0^* := \sup_{x \in \mathbb{R}^n} \quad & \mathbf{c}^T x = \mathbf{c}_1 x_1 + \cdots + \mathbf{c}_n x_n \\ & h_1(x) = \cdots = h_p(x) = 0, \end{aligned} \tag{1.1}$$

这里, 多项式 $h_1, \dots, h_p \in \mathbb{R}[X_1, \dots, X_n]$, $\mathbf{c} := (\mathbf{c}_1, \dots, \mathbf{c}_n)$ 是参数.

在问题(1.1)中, 最优值 c_0^* 可以看作是参数 \mathbf{c} 的半代数函数, 即最优值函数. 我们的目的是利用多项式系统来描述最优值函数的图像. 柱代数分解(CAD, 参见 [17])能够用来计算一个半代数集的投影(等价于消除量词), 因此能够用来求解问题(1.1). 更准确地说, 通过引入布尔算子 \wedge , 我们将问题(1.1)关联到布尔表达式

$$(h_1(X) = 0) \wedge \cdots \wedge (h_p(X) = 0) \wedge (\mathbf{c}_0 - \mathbf{c}^T X \geq 0) \tag{1.2}$$

这里 $X = (X_1, \dots, X_n)$.

通过选取满足 X 变量大于 \mathbf{c} 变量的序可以计算 \mathbb{R}^{2n+1} 中由(1.2)定义的半代数集的柱代数分解. 在投影阶段, 我们得到半代数集的投影的定义多项式的集合. 这个集合可以用来描述参数空间在 \mathbb{R}^n 中的胞元边界. 然而, CAD算法的复杂度是变元个数和定义多项式次数的双指指数函数. 这限制了它的实际应用. 如果多项式的个数超过4, 它可能会输出很多不相关的多项式. 一些复杂度更低的方法可以用来求解多项式优化问题, 见 [6, 35–37, 80, 83]. 在无参数情形, 这些方法求解多项式优化问题的最优值所得到的单变元多项式的次数是变元个数的单指指数函数.

我们将求解无参数优化问题的算法推广到参数化优化问题中, 从而描述最优值函数并且求解参数化优化问题. 令 $\Phi \in \mathbb{Q}[\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_n]$ 为一个多项式系统,

其对应的代数集包含最优化函数的图像. Φ 关于变量 \mathbf{c}_0 最小可能的次数称为问题(1.1)的代数次数. 代数次数可以衡量问题(1.1)的代数复杂度. 代数次数在很多文献中被提及, 参见 [12, 16, 34, 43, 64, 65, 71].

令 $\mathbf{h} = \{h_1, \dots, h_p\}$ 为问题(1.1)中的多项式集合并且令

$$V = \{v \in \mathbb{C}^n \mid h_1(v) = \dots = h_p(v) = 0\}. \quad (1.3)$$

假设 $\langle \mathbf{h} \rangle$ 是等维根理想. 在文献 [12, 定理 5.23] 中, Rostalski 和 Sturmfels 指出当 V 不可约, 光滑并且 $V \cap \mathbb{R}^n$ 有界时, 最优化函数 Φ 的图像的 Zariski 闭包可以由 V 对应的对偶仿射簇的定义多项式刻画. $V \cap \mathbb{R}^n$ 的有界性可以确保对应的优化问题的最优化是有限并且可达的. 它对 [12, 定理 5.23] 的证明是必要的.

然而, 当 $V \cap \mathbb{R}^n$ 无界时, 对某一个参数值, 其对应的优化问题的最优化可能是无限的或者是有限不可达的, 即最优化是渐近关键值(参见 [44, 46, 70])或者是目标函数在奇异点上的值. 因此 [12, 定理 5.23] 的证明在这个情形下是无效的并且 V 对应的对偶仿射簇可能不是最优化函数的 Zariski 闭包. 此外, 对求解问题(1.1)的最优化, V 对应的对偶仿射簇的定义多项式存在一个不足之处, 即对某一个参数值, 定义多项式可能被约化成零多项式. 我们无法从对偶仿射簇的定义多项式中得到关于最优化的任何信息.

最后, 我们考虑给定一个优化问题, 如何减少描述该优化问题的可行集的变量和约束的个数. 当可行集是多面体时, 显然地, 它可以表示成一个非负象限的仿射切片在一个线性映射下的像. Yannakakis 在文献 [101] 中证明一个有界多面体的松弛矩阵的非负秩是使得上述表示成立的非负象限的最小维数, 即松弛矩阵的非负秩可以表示刻画有界多面体所需要的最少的变量和约束个数. 文献 [24, 31] 中, 通过松弛算子的锥分解, Yannakakis 的结果被推广到凸体的锥举起问题. 尽管文献 [31] 中的结果对凸体成立, 我们注意到一个不包含直线的无界闭凸集由极点和极方向生成. 因此判别一个闭凸集是否存在锥举起的问题不能由已有结论直接推广. 此外, 如果一个闭凸集包含直线, 它不存在极点和极方向. 如果原点不是闭凸集的内点, 它的极集所对应的线性函数不能完全刻画它(见例子 6.1). 我们研究如何将凸体的锥举起和松弛算子的定义推广到一般的闭凸集情形并且说明当一个闭凸集不是凸体时, 它的锥举起和松弛算子的锥分解之间的关系.

1.2 论文的结构和主要成果

在第三章中, 当理想 I 对应的实代数集 $V_{\mathbb{R}}(I)$ 是正维时, 我们提出一个终止条件来验证等式 $\langle \ker M_t(y) \rangle = I(V_{\mathbb{R}}(I))$ 是否成立. 遗憾的是, 我们在 [54, §2.4.3] 中没有完全解决这个问题. 尽管如此, 我们提出一个基于几何对合理论 [87–89] 的终止条件(3.6)来计算一个满足关系 $I \subseteq J \subseteq I(V_{\mathbb{R}}(I))$ 的理想 $J = \langle \ker M_{t-2}(y) \rangle$ 在分次逆字典序下的Pommaret基. Pommaret基是特殊的Gröbner基. Pommaret基可以用来计算一个理想的深度, 射影维数和一个模的Castelnuovo-Mumford正则性. 当实代数集 $V_{\mathbb{R}}(I)$ 是正维时, 在第3.4节的例子中, 我们的终止条件可以成功地计算出实根理想 $I(V_{\mathbb{R}}(I))$ 的一组Pommaret 基. 在一般情况下, 我们不能确保当终止条件(3.6)满足时, 算法返回的多项式集合是 I 的实根理想的生成元.

在第四章中, 我们将文献 [30, 47] 中的结果推广到无界半代数集的情形. 我们利用齐次化技巧构造了推广Theta体序列和推广Lasserre半正定表示序列. 我们指出当尖性和半代数集在无穷远处是闭的条件满足时, 这些序列可以收敛到无界半代数集的闭凸包. 此外, 如果PP-BDR性质成立并且当阶数充分大时, 推广Theta体和推广Lasserre半正定表示是准确的. 通过一些例子, 我们说明尖性和半代数集在无穷远处是闭的条件对于推广Theta体序列和推广Lasserre半正定表示序列的收敛性是必要的. 最后, 我们提供一些策略来解决尖性和半代数集在无穷远处是闭的条件不满足的情形.

在第五章中, 我们考虑一个参数化多项式优化问题的可行集是无界或者非光滑的代数集情形. 当可行集光滑, 无界并且可行集的闭凸包不包含直线时, 我们通过对偶仿射簇计算一个多项式. 这个多项式的定义超曲面是包含最优值函数图像的Zariski闭包. 利用对偶仿射簇, 我们设计一个算法, 使得对几乎所有的参数值, 对应的优化问题的最优值是一个非零单变元多项式的根. 为了求解参数化优化问题的最优值, 我们设计一个基于极仿射簇的算法的参数化算法. 算法返回一个多项式系统, 使得对任意的参数值, 多项式系统可以被约化为一个非零单变元多项式, 单变元多项式的根包含对应的特定化优化问题的最优值.

在第六章中, 我们考虑无界闭凸集的锥举起问题. 任给一个不包含直线的闭凸集, 我们可以利用它的极集的极点和极方向来刻画这个闭凸集. 利用这个结果和凸集的表示定理, 我们推广了文献 [27, 29, 32] 中的松弛算子和锥举起的定义. 我们将文献 [27, 29, 32] 中关于凸体的锥举起与松弛算子的锥分解的等价关系推广到一般闭凸集的情形.

第二章 预备知识

符号 $\mathbb{R}(\mathbb{C})$ 表示实(复)数域. 子空间 $L \subset \mathbb{R}^n$ 的正交补记作 L^\perp . 给定一个域 $\mathbb{K} = \mathbb{R}(\mathbb{C})$, 域 \mathbb{K} 中的 n 元多项式环记为 $\mathbb{K}[x] := \mathbb{K}[x_1, \dots, x_n]$. \mathbb{N} 表示非负整数的集合. $\mathbb{K}[x]$ 中的多项式定义为 $p := \sum_{\alpha \in \mathbb{N}^n} p_\alpha x^\alpha$, 其中只有有限多个系数 $p_\alpha \in \mathbb{K}$ 是非零的并且它的首项 $\text{lt}_\prec(p)$ 定义为对应系数 p_α 不为0的单项式中关于项序 \prec 最大的单项式 x^α . $\langle \text{lt}_\prec(I) \rangle$ 记作由理想 I 中多项式的首项生成的理想. 给定 $\mathbb{R}[x]$ 中的一个多项式 h , 定义 $\text{vec}(h) = (h_\alpha)_{\alpha \in \mathbb{N}^n}$. 对 $t \in \mathbb{N}$, 定义 $\mathbb{N}_t^n := \{\alpha \in \mathbb{N}^n \mid |\alpha| := \sum_{i=1}^n \alpha_i \leq t\}$. 对 $\alpha \in \mathbb{N}^n$, 单项式 $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ 的次数是 $|\alpha| := \sum_{i=1}^n \alpha_i$. 所有单项式均包含在集合 $\mathbb{T}^n := \{x^\alpha \mid \alpha \in \mathbb{N}^n\}$ 中. 集合 $\mathbb{T}_t^n := \{x^\alpha \mid \alpha \in \mathbb{N}_t^n\}$ 由次数至多为 $t \in \mathbb{N}$ 的单项式组成. 给定一个非负整数 t , $\mathbb{K}[x]_t$ 表示次数至多为 t 的多项式集合. 符号 $[x]_t$ 表示次数至多为 t 的单项式所生成的序列:

$$[x]_t := [1, x_1, \dots, x_n, x_1^2, x_1 x_2, \dots, x_1^t, x_1^{t-1} x_2, \dots, x_n^t].$$

2.1 凸分析基础

2.1.1 凸集

给定 \mathbb{R}^n 中的一个非空子集 C , 如果对任意的 $x \in C$, $y \in C$ 和 $\lambda \in (0, 1)$, 有 $(1 - \lambda)x + \lambda y \in C$, 则称 C 是凸集. 对 \mathbb{R}^n 中的任意子集 W , W 的凸包 $\mathbf{co}(W)$ 定义为包含 W 的最小凸集. $\mathbf{cl}(W)$ 记作 W 的闭包. W 的闭凸包(即凸包的闭包)记作 $\mathbf{cl}(\mathbf{co}(W))$. 一个凸集 C 的仿射闭包 $\mathbf{aff}(C)$ 定义为包含 C 的最小仿射集. 凸集 $C \subset \mathbb{R}^n$ 的相对内点 $\mathbf{ri}(C)$ 定义为 C 作为 $\mathbf{aff}(C)$ 中的子集的内点.

一个凸集 C 的面是 C 的一个子集 C' , 使得对 C 中的任意一个闭线段, 如果存在相对内点在 C' 中, 则闭线段的端点也在 C' 中. C 的零维面定义为 C 的极点. 如果 C' 是 C 的半线面, 则 C' 的方向称为 C 的极方向.

令 S_0 和 S_1 分别为 \mathbb{R}^n 中的点和方向的集合. $S = S_0 \cup S_1$ 的凸包 $\mathbf{co}(S)$ 定义为 \mathbb{R}^n 中使得条件 $S_0 \subseteq C$ 和 $S_1 \subseteq 0^+ C$ 成立的最小凸集 C . 代数上, 一个向量 x 属于 $\mathbf{co}(S)$ 当且仅当存在非负整数 m, k , 非负实数 $\lambda_1, \dots, \lambda_m$, 使得下面的表达式成

立:

$$x = \lambda_1 x_1 + \cdots + \lambda_k x_k + \lambda_{k+1} x_{k+1} + \cdots + \lambda_m x_m, \quad \sum_{i=1}^k \lambda_i = 1,$$

其中对 $1 \leq i \leq k$, $x_i \in S_0$, 对 $k+1 \leq i \leq m$, $x_i \in S_1$.

定理 2.1. [76, 定理 18.5] 给定一个不包含直线的闭凸集 C . 令 S 为 C 中极点和极方向的集合, 则 $C = \mathbf{co}(S)$.

一个非空凸集 $C \subset \mathbb{R}^n$ 的极集 (*polar*) 定义如下:

$$C^\circ = \{x \in \mathbb{R}^n \mid \text{对任意的 } y \in C, \langle x, y \rangle \leq 1\}.$$

显然地, $C^{\circ\circ} = \mathbf{cl}(\mathbf{co}(C \cup \{0\}))$.

众所周知, \mathbb{R}^n 中的有界集的凸包是闭的. 然而, 对一般的无界集, 结论不一定成立. 例如, 令

$$V := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 x_2 = 1, x_2 \geq 0\} \cup \{(0, 0)\},$$

则

$$\mathbf{co}(V) = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 > 0, x_2 > 0\} \cup \{(0, 0)\}.$$

因此, $\mathbf{co}(V)$ 不是闭集.

如果一个对称矩阵 A 是半正定的, 则记作 $A \succeq 0$. \mathbb{R}^n 中的谱多面体 (*spectrahe dron*) 定义为如下集合:

$$\{x \in \mathbb{R}^n \mid A_0 + \sum_{i=1}^k x_i A_i \succeq 0\},$$

这里, A_0, A_1, \dots, A_k 为实对称矩阵. \mathbb{R}^n 中谱多面体的投影 (*projected spectrahe dron*) 定义为如下集合:

$$\{x \in \mathbb{R}^n \mid \text{存在 } y \in \mathbb{R}^k, A_0 + \sum_{i=1}^k x_i A_i + \sum_{i=1}^k y_i A_i \succeq 0\},$$

这里, A_0, A_1, \dots, A_k 是实对称矩阵.

2.1.2 凸锥

给定一个集合 $K \subseteq \mathbb{R}^n$, 如果 K 关于正标量乘法运算是封闭的, 则称 K 是一个锥. 给定一个集合 $C \subseteq \mathbb{R}^n$, 称 $\text{cone}(C)$ 为由 C 生成的锥, 即包含 C 的最小锥. 给定一个凸锥 K , 它的极射线定义为以原点为端点的半线面. K 的对偶锥定义为

$$K^* = \{c \in \mathbb{R}^n \mid \text{对任意的 } x \in K, \langle c, x \rangle \geq 0\}.$$

对 \mathbb{R}^n 中的任意子空间 L , 有 $L^* = L^\perp$.

一个锥 K 不一定是凸集, 但是它的对偶锥 K^* 一定是闭凸锥. K 的对偶锥的对偶锥 K^{**} 是 K 的闭凸包. 因此, 如果 K 是一个闭凸锥, 则 $K^{**} = K$.

命题 2.2. [76, 推论 16.4.2] 给定 \mathbb{R}^n 中的闭凸锥 K_1 和 K_2 , 则

$$(K_1 + K_2)^* = K_1^* \cap K_2^* \quad \text{并且} \quad (K_1 \cap K_2)^* = \text{cl}(K_1^* + K_2^*). \quad (2.1)$$

特别地, 对 \mathbb{R}^n 中的任意子空间 L , 有

$$(K_1 \cap L)^* = \text{cl}(K_1^* + L^\perp).$$

定义 2.1. 给定一个闭凸锥 K , 如果 $K \cap -K = \{0\}$, 则称 K 是尖的(pointed).

定理 2.3. 给定 \mathbb{R}^n 中的闭锥 K . 下面的说法是等价的:

1. $\text{co}(K)$ 是尖的闭凸集.
2. $\text{co}(K)$ 不包含过原点的直线.
3. 存在向量 $c = (c_1, \dots, c_n) \in \mathbb{R}^n$, 使得对任意的 $x \in \text{co}(K) \setminus \{0\}$, 有 $\langle c, x \rangle > 0$.

证明. 证明和 [20, 定理 11, 引理 2] 相同. □

注 2.1. 尽管尖性定义在闭凸锥上, 对于一个闭锥 K , 如果 $\text{co}(K) \cap -\text{co}(K) = \{0\}$, 则称 $\text{co}(K)$ 是尖的.

在 \mathbb{R}^n 中的非空凸锥 K 的极集定义为

$$K^\circ = \{x \in \mathbb{R}^n \mid \text{对任意的 } y \in K, \langle x, y \rangle \leq 0\}.$$

在 \mathbb{R}^n 中的非空凸集 C 的回收(*recession*)锥 0^+C 定义为:

$$\{y \mid \text{对任意的 } \lambda > 0, x \in C, x + \lambda y \in C\}.$$

凸集 C 的线性空间定义为 $0^+C \cap (-0^+C)$. 闭凸集 C 不包含直线当且仅当 0^+C 是尖的. 闭凸集 $C \subseteq \mathbb{R}^n$ 有界当且仅当 0^+C 只包含零向量. 闭凸集 C 的极方向是 0^+C 的极射线.

定理 2.4. [76, 定理 8.3] 给定一个非空闭凸集 C 和一个非零向量 y . 如果存在向量 $x \in C$, 使得半线 $\{x + \lambda y \mid \lambda \geq 0\} \subseteq C$, 则 $y \in 0^+C$.

2.1.3 凸函数

f 是定义在 $S \subseteq \mathbb{R}^n$ 上的实值(包括 $\pm\infty$)函数. f 的上方图形定义为

$$\text{epi}(f) = \{(x, \mu) \in \mathbb{R}^{n+1} \mid x \in S, \mu \in \mathbb{R}, \mu \geq f(x)\}.$$

如果 $\text{epi}(f)$ 是 \mathbb{R}^{n+1} 中的凸集, 则称 f 是定义在凸集 S 上的凸函数. 函数 f 的有效域 $\text{dom}(f)$ 定义为 $\text{epi}(f)$ 在 \mathbb{R}^n 上的投影:

$$\begin{aligned} \text{dom}(f) &= \{x \in \mathbb{R}^n \mid \text{存在 } \mu \in \mathbb{R}, \text{ 使得 } (x, \mu) \in \text{epi}(f)\} \\ &= \{x \in \mathbb{R}^n \mid f(x) < +\infty\}. \end{aligned}$$

如果 f 的值域不包含 $-\infty$, 则称 f 是正常的(*proper*)凸函数.

定理 2.5. [21, 定理 2.2] 假设 $C \subseteq \mathbb{R}^n$ 是无界闭凸集. 如果 0^+C 是尖的闭凸锥, 则

1. $(0^+C)^\circ$ 是 n 维闭凸集.
2. $\text{int}((0^+C)^\circ) \subseteq \text{dom}(c_0^*(\mathbf{c} \mid C)) \subseteq (0^+C)^\circ$.

2.2 矩量矩阵

给定 \mathbb{R}^n 中的测度 μ , $y_\alpha := \int x^\alpha \mu(dx)$ 称为第 α 阶矩量. 序列 $(y_\alpha)_{\alpha \in \mathbb{N}^n}$ 称为测度 μ 的矩量序列, 测度 μ 称为序列 $(y_\alpha)_{\alpha \in \mathbb{N}^n}$ 的表示测度. 给定 $t \in \mathbb{N}^n$, 截断序列 $(y_\alpha)_{\alpha \in \mathbb{N}_t^n}$ 称为测度 μ 的第 t 阶矩量序列.

给定一个序列 $y = (y_\alpha)_{\alpha \in \mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$, 矩量(*moment*)矩阵

$$M(y) := (y_{\alpha+\beta})_{\alpha, \beta \in \mathbb{N}^n}$$

是一个行和列以单项式集合 $\mathbb{T}^n := \{x^\alpha \mid \alpha \in \mathbb{N}^n\}$ 为索引的实对称矩阵。定义序列 $hy := M(y)\text{vec}(h) \in \mathbb{R}^{\mathbb{N}^n}$ 。如果 $M(y)p := M(y)\text{vec}(p) = 0$, 则称多项式 p 在矩阵量矩阵 $M(y)$ 的核中。给定一个截断矩量序列 $y = (y_\alpha)_{\alpha \in \mathbb{N}_{2t}^n} \in \mathbb{R}^{\mathbb{N}_{2t}^n}$, 定义截断矩量矩阵

$$M_t(y) := (y_{\alpha+\beta})_{\alpha, \beta \in \mathbb{N}_t^n}.$$

截断矩量矩阵以单项式集合 $\mathbb{T}_t^n := \{x^\alpha \mid \alpha \in \mathbb{N}_t^n, |\alpha| := \sum_{i=1}^n \alpha_i \leq t\}$ 为索引。我们考虑次数小于等于 t 的多项式空间 $\mathbb{R}[x]_t$ 。给定 $\mathbb{R}[x]_t$ 中的多项式 p , 如果 $M_t(y)\text{vec}(p) = 0$, 则称 p 在 $M_t(y)$ 的核中, 即

$$\ker M_t(y) := \{p \in \mathbb{R}[x]_t \mid M_t(y)\text{vec}(p) = 0\}. \quad (2.2)$$

2.3 代数几何基础

2.3.1 基本概念

给定 n 元多项式环 $\mathbb{R}[x] := \mathbb{R}[x_1, \dots, x_n]$ 中的多项式 h_1, \dots, h_m 和理想 $I = \langle h_1, \dots, h_m \rangle$. I 的复(实)仿射簇分别定义为

$$V_{\mathbb{C}}(I) := \{x \in \mathbb{C}^n \mid \text{对任意的 } f \in I, f(x) = 0\}, \quad V_{\mathbb{R}}(I) := V_{\mathbb{C}}(I) \cap \mathbb{R}^n.$$

\mathbb{C}^n 中的仿射簇 V 在 $\mathbb{K}[x]$ 中的零化(vanishing)理想定义为

$$I(V) := \{f \in \mathbb{K}[x] \mid \text{对任意的 } v \in V, f(v) = 0\}.$$

$\mathbb{K}[x]$ 中的理想 I 的根理想定义为

$$\sqrt{I} := \{f \in \mathbb{K}[x] \mid \text{存在 } k \in \mathbb{N}, \text{使得 } f^k \in I\}.$$

$\mathbb{R}[x]$ 中的理想 I 的实根理想定义为

$$\sqrt{\mathbb{R}I} := \left\{ f \in \mathbb{R}[x] \mid \text{存在 } k \in \mathbb{N}, q_1, \dots, q_r \in \mathbb{R}[x], \text{使得 } f^{2k} + \sum_{i=1}^r q_i^2 \in I \right\}.$$

显然地, 它们之间的包含关系为 $I \subseteq \sqrt{I} \subseteq \sqrt{\mathbb{R}I}$. 如果 $I = \sqrt{I}$ ($I = \sqrt{\mathbb{R}I}$), 则称 I 为根(实根)理想。

定理 2.6. [13, 定理 4.14] 给定 $\mathbb{R}[x_1, \dots, x_n]$ 中的理想 I . 则 I 是实根理想当且仅当 $I = I(V_{\mathbb{R}}(I))$.

多项式 h_1, \dots, h_m 关于 x_1, \dots, x_n 的雅克比矩阵定义如下:

$$\begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial h_m}{\partial x_1} & \frac{\partial h_m}{\partial x_2} & \cdots & \frac{\partial h_m}{\partial x_n} \end{pmatrix}.$$

假设理想 $I := \langle h_1, \dots, h_m \rangle$ 是 d 维的等维根理想, 则仿射簇 $V := V(I)$ 的奇异点集 $\text{sing}(V)$ 由多项式 h_1, \dots, h_m 和它们的雅克比矩阵中所有的 $n - d$ 阶子式定义. V 的光滑(正则)点集定义为 $V \setminus \text{sing}(V)$. 如果 $\text{sing}(V) = \emptyset$, 则称仿射簇 V 是光滑的.

2.3.2 平方和多项式

给定 $\mathbb{R}[X]$ 中的多项式 $p(X)$. 如果存在多项式 $u_1(X), \dots, u_s(X) \in \mathbb{R}[X]$, 使得 $p(X) = \sum_{i=1}^s u_i(X)^2$, 则称多项式 $p(X)$ 是平方和(SOS)多项式. 符号 Σ^2 表示平方和多项式的集合.

给定多项式集合 $G := \{g_1, \dots, g_m\}$ 和基本闭半代数集 $S := \{x \mid g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$. G 的二次模定义为

$$\mathcal{Q}(G) := \left\{ \sum_{j=0}^m \sigma_j g_j \mid g_0 = 1, \sigma_j \in \Sigma^2 \right\}.$$

G 的第 k 阶二次模 (quadratic module) 定义为

$$\mathcal{Q}_k(G) := \left\{ \sum_{j=0}^m \sigma_j g_j \mid g_0 = 1, \sigma_j \in \Sigma^2, \deg(\sigma_j g_j) \leq 2k \right\}.$$

对任意的 $p \in \mathcal{Q}(G)$ 和 $x \in S$, 有 $p(x) \geq 0$.

定义 2.2. 如果存在 $\psi \in \mathcal{Q}(G)$, 使得集合 $\{x \mid \psi(x) \geq 0\}$ 在 \mathbb{R}^n 中有界, 则称 $\mathcal{Q}(G)$ 满足阿基米德条件.

注意到阿基米德条件可以推导出 S 是有界的, 但是反过来却不一定是对的, 对任意的有界集 S , 取充分大的 M , 将约束条件 $M - \|x\|_2^2$ 添加到 S 的定义多项式中, 新的定义多项式集合对应的二次模满足阿基米德条件.

定理 2.7. [69, PUTINAR正零点定理] 假设 $\mathcal{Q}(G)$ 满足阿基米德条件. 如果多项式 $p \in \mathbb{R}[X]$ 在 S 上恒大于 0, 则存在 $k \in \mathbb{N}$, 使得 $p \in \mathcal{Q}_k(G)$.

定义 2.3. [47, 定义 3] 如果存在 $k \in \mathbb{N}$ 和 $\mathbb{R}[x]_1$ 中的一个零测集 V , 使得在 S 上恒为正的仿射函数 p , 如果 $\text{vec}(p) \notin V$, 可以推出 $p \in \mathcal{Q}_k(G)$, 则称 G 满足第 k 阶 Putinar-Prestel 有界表示(PP-BDR) 条件.

第三章 实根理想的计算

3.1 问题和背景

给定多项式环 $\mathbb{R}[x_1, \dots, x_n]$ 中的 m 个多项式 h_1, \dots, h_m . 定义

$$d_j := \lceil \deg(h_j)/2 \rceil, \quad d := \max_{1 \leq j \leq m} d_j. \quad (3.1)$$

对 $t \geq d$, 定义集合

$$\mathcal{K}_t := \{y \in \mathbb{R}^{\mathbb{N}_{2t}^n} \mid y_0 = 1, M_t(y) \succeq 0, M_{t-d_j}(h_j y) = 0, j = 1, \dots, m\}. \quad (3.2)$$

如果 $y \in \mathcal{K}_t$ 的截断矩量矩阵 $M_t(y)$ 在 \mathcal{K}_t 中具有最大秩, 则称 y 是 \mathcal{K}_t 的母(generic)元素. 母元素的集合记为

$$\mathcal{K}_t^{gen} := \{y \in \mathcal{K}_t \mid M_t(y) \text{ 的秩在 } \mathcal{K}_t \text{ 中是最大的}\}. \quad (3.3)$$

当实代数集 $V_{\mathbb{R}}(I)$ 是零维时, Lasserre等学者在文献 [18, 48] 中使用平坦扩张条件来检验对 \mathcal{K}_t^{gen} 中的向量 y , $\ker M_s(y)$ ($1 \leq s \leq t$)中的多项式是否可以生成实根理想 $I(V_{\mathbb{R}}(I))$. 当 $V_{\mathbb{R}}(I)$ 是正维时, 这个条件是无效的. 下面的例子由Fialkow在文献 [22, 例子 3.2]中提出. 这个例子说明 $V_{\mathbb{R}}(I)$ 是正维时, 即使矩量矩阵的核包含实根理想的生成元, 平坦扩张条件仍然不成立.

例 3.1. 考虑如下定义的3阶截断矩量矩阵 $M_3(y)$:

$$M_3(y) = \begin{pmatrix} 1 & 0 & 0 & 1 & 2 & 5 & 0 & 0 & 0 & x \\ 0 & 1 & 2 & 0 & 0 & 0 & 2 & 5 & 14 & 42 \\ 0 & 2 & 5 & 0 & 0 & x & 5 & 14 & 42 & 132 \\ 1 & 0 & 0 & 2 & 5 & 14 & 0 & 0 & x & 0 \\ 2 & 0 & 0 & 5 & 14 & 42 & 0 & x & 0 & 0 \\ 5 & 0 & x & 14 & 42 & 132 & x & 0 & 0 & 0 \\ 0 & 2 & 5 & 0 & 0 & x & 5 & 14 & 42 & 132 \\ 0 & 5 & 14 & 0 & x & 0 & 14 & 42 & 132 & r \\ 0 & 14 & 42 & x & 0 & 0 & 42 & 132 & r & s \\ x & 42 & 132 & 0 & 0 & 0 & 132 & r & s & t \end{pmatrix}.$$

定义变量 x_1, x_2 的序 $x_1 \prec x_2$, 我们使用分次逆字典序来指定单项式 $x_1^{\alpha_1}x_2^{\alpha_2}$
($0 \leq \alpha_1 + \alpha_2 \leq 3$)的序并且排列矩阵 $M_3(y)$ 的行和列.

当 $x = 0, r = 429, s = 1422, t = 4798$ 时, 可得 $M_3(y) \succeq 0$, $\text{rank } M_3(y) = 9$ 以及 $\ker M_3(y) = \{x_2 - x_1^3\}$. 与零维情形不同, 尽管矩量矩阵 $M_3(y)$ 的核只包含一个多项式 $x_2 - x_1^3$ 并且这个多项式是实根理想 $I = I(V_{\mathbb{R}}(I)) = \langle x_2 - x_1^3 \rangle$ 的Gröbner基, Fialkow在文献 [22]中指出截断矩量序列 $y \in \mathcal{K}_3$ 不能扩展到下一阶, 并且 y 没有表示测度.

在这一章, 我们考虑下面的问题:

问题1. 给定 n 元多项式环 $\mathbb{R}[x]$ 中具有正维实根理想的理想 I . 如何计算实根理想 $I(V_{\mathbb{R}}(I))$ 的生成元?

利用偏微分方程中的对合理论, 我们在算法3.1中提出终止条件(3.6). 如果终止条件(3.6)成立, 算法返回一个介于理想 I 和它的实根 $\sqrt[{\mathbb{R}}]{I}$ 之间的理想 J 的Pommaret基. 如果终止条件(3.6)成立并且一些假设满足, 算法返回 $\sqrt[{\mathbb{R}}]{I}$ 的生成元. 在零维情形下, 我们研究终止条件(3.6)和平坦扩张条件的关系.

3.2 预备知识

3.2.1 矩量矩阵的性质

在文献 [18, 48, 52, 53, 60] 中, 矩量矩阵核的性质被广泛研究并且具有十分重要的应用.

引理 3.1. [48, 命题 3.6] 给定矩量矩阵 $M(y)$ 的核 $\ker M(y) := \{p \in \mathbb{R}[x] \mid M(y)\text{vec}(p) = 0\}$. 则 $\ker M(y)$ 是 $\mathbb{R}[x]$ 中的理想. 如果 $M(y) \succeq 0$, 则 $\ker M(y)$ 是一个实根理想.

尽管截断矩量矩阵 $M_t(y)$ 核中的多项式的集合不是一个理想, 它仍然具有重要的性质.

命题 3.2. [48, 引理 3.5, 3.9] 给定向量 $y \in \mathbb{R}^{\mathbb{N}_{2t}^n}$ 并且假设截断矩量矩阵 $M_t(y)$ 是半正定的.

(i) 如果 $f, g \in \mathbb{R}[x]$ 满足条件 $\deg(fg) \leq t - 1$, 则 $f \in \ker M_t(y) \implies fg \in \ker M_t(y)$.

- (ii) 给定多项式 $p \in \mathbb{R}[x]$, 如果存在 $k \in \mathbb{N}$ 和 $\sigma \in \sum \mathbb{R}[x]^2$, 使得 $p^{2k} + \sigma \in \ker M_t(y)$, 则 $p \in \ker M_t(y)$.
- (iii) 对 $1 \leq s \leq t$, 可得 $\ker M_t(y) \cap \mathbb{R}[x]_s = \ker M_s(y)$.

母元素 $y \in \mathcal{K}_t^{gen}$ 具有重要的性质. 下面的结果引自 [48, 引理 3.1] 和 [78, 引理 7.28, 7.39].

命题 3.3. 假设 $y \in \mathcal{K}_t^{gen}$,

- (i) 对 $1 \leq s \leq t$, 可得 $\ker M_s(y) \subseteq \sqrt[\mathbb{R}]{I}$. 对任意的 $z \in \mathcal{K}_t$, 有 $\ker M_s(y) \subseteq \ker M_s(z)$.
- (ii) 如果 $t \leq t'$ 并且 $y' \in \mathcal{K}_{t'}^{gen}$, 则 $\ker M_t(y) \subseteq \ker M_{t'}(y')$.
- (iii) 假设 $\{g_1, \dots, g_k\}$ 是 $\sqrt[\mathbb{R}]{I}$ 的有限生成元. 则存在 $t_0 \in \mathbb{N}$, 使得对任意的 $z \in \mathcal{K}_t$ 和 $t \geq t_0$, 有 $g_1, \dots, g_k \in \ker M_t(z)$.
- (iv) 当 t 充分大时, 可得 $\langle \ker M_t(y) \rangle = \sqrt[\mathbb{R}]{I}$.

3.2.2 对合除法和对合基

定义 3.1. 给定单项式 x^ν 的指数(exponent) $\nu = [\nu_1, \dots, \nu_n] \in \mathbb{N}^n$. 如果 k 是使得 $\nu_k \neq 0$ 成立的最小值, 则称 $\nu(x^\nu)$ 的类是 k , 记作 $\text{cls}(\nu) = k$ ($\text{cls}(x^\nu) = k$). 一个多项式 f 的首项指数 $\text{le}_\prec(f)$ 定义为它的首项的指数. 一个多项式 f 的类 $\text{cls}(f)$ 定义为它的首项 $\text{lt}_\prec(f)$ 的类.

给定一个项序 \prec , 如果对相同次数的单项式 x^μ 和 x^ν , $\text{cls}(\mu) \prec \text{cls}(\nu)$ 能推出 $x^\mu \prec x^\nu$, 则称项序是与类相关的.

定义 3.2. 给定变量序 $x_1 \prec \dots \prec x_n$, 分次逆字典序 \prec_{tdeg} 按照下面的方式定义: 如果 $|\alpha| < |\beta|$ 或者 $|\alpha| = |\beta|$ 并且指数 $\alpha - \beta$ 的第一个非零分量是正的, 则 $x^\alpha \prec_{\text{tdeg}} x^\beta$.

分次逆字典序 \prec_{tdeg} 是与类相关的序. 在这一章, 我们用 \prec_{tdeg} 来定义单项式的次序以及排列矩阵 $M_t(y)$ 的行与列. 集合 \mathbb{N}^n 通过定义加法来构成交换幺半群 $(\mathbb{N}^n, +)$. 给定指数 $\nu \in \mathbb{N}^n$, 它的锥定义为 $\mathcal{C}(\nu) = \nu + \mathbb{N}^n$. 如果 $\mu \in \mathcal{C}(\nu)$, 则称 ν 整除 μ , 记作 $\nu | \mu$. 对合除法 L 可以在幺半群 $(\mathbb{N}^n, +)$ 上定义.

定义 3.3. [89, 定义 3.1.1] 对有限集合 $\mathcal{B} \subseteq \mathbb{N}^n$ 和任意的 $\nu \in \mathbb{N}^n$, 定义乘性指标(multiplicative indices) $N_{L,\mathcal{B}}(\nu) \subseteq \{1, \dots, n\}$ 和 \mathbb{N}^n 的子幺半群 $L(\nu, \mathcal{B}) = \{\mu \in \mathbb{N}^n \mid \text{对任意的 } j \notin N_{L,\mathcal{B}}(\nu), \mu_j = 0\}$. 此外, 对合锥(involutory cones) $\mathcal{C}_{L,\mathcal{B}}(\nu) = \nu + L(\nu, \mathcal{B}) \subseteq \mathbb{N}^n$ 必须满足两个条件:

- (i) 如果 $\mu, \nu \in \mathcal{B}$ 满足条件 $\mathcal{C}_{L,\mathcal{B}}(\mu) \cap \mathcal{C}_{L,\mathcal{B}}(\nu) \neq \emptyset$, 则包含关系 $\mathcal{C}_{L,\mathcal{B}}(\mu) \subseteq \mathcal{C}_{L,\mathcal{B}}(\nu)$, $\mathcal{C}_{L,\mathcal{B}}(\nu) \subseteq \mathcal{C}_{L,\mathcal{B}}(\mu)$ 至少有一个成立.
- (ii) 如果 $\mathcal{B}' \subset \mathcal{B}$, 则对任意的 $\nu \in \mathcal{B}'$, 有 $N_{L,\mathcal{B}}(\nu) \subseteq N_{L,\mathcal{B}'}(\nu)$.

给定任意的指数 $\mu \in \mathbb{N}^n$ 和 $\nu \in \mathcal{B}$, 如果 $\mu \in \mathcal{C}_{L,\mathcal{B}}(\nu)$, 则称 ν 对合整除 μ , 记作 $\nu \mid_{L,\mathcal{B}} \mu$. 这时, ν 称为 μ 的对合除子(involutory divisor).

定义 3.4. [89, 例子 3.1.7] 给定 $\nu \in \mathbb{N}^n$, 乘性指标 $N_{L,\mathcal{B}}(\nu)$ 定义为 $\{1, \dots, \text{cls}(\nu)\}$, 则乘性指标 $N_{L,\mathcal{B}}(\nu)$ 对应的除法 L 是 Pommaret 除法.

注 3.1. Pommaret 除法是全局定义的除法. 指数 $\nu \in \mathcal{B}$ 的乘性指标与集合 \mathcal{B} 无关. Pommaret 除法的定义在 [89, 引理 3.1.8] 中提及.

定义 3.5. [89, 定义 3.1.9] 一个有限集合 $\mathcal{B} \subset \mathbb{N}^n$ 的对合域定义为

$$\langle \mathcal{B} \rangle_L = \bigcup_{\nu \in \mathcal{B}} \mathcal{C}_{L,\mathcal{B}}(\nu). \quad (3.4)$$

给定对合除法 L , 如果集合 \mathcal{B} 满足 $\langle \mathcal{B} \rangle_L = \langle \mathcal{B} \rangle$, 则称 \mathcal{B} 为弱对合的或者为幺半群理想 $\langle \mathcal{B} \rangle$ 的弱对合基. 如果 (3.4) 的并是互不相交的, 即对合锥的相交是空集, 则集合 \mathcal{B} 是一个强对合基, 简称对合基.

定义 3.6. [89, 定义 3.4.1] 令 $I \subseteq \mathbb{K}[x]$ 为一个理想. 对 \mathbb{N}^n 中的对合除法 L 和有限的集合 $\mathcal{H} \subset I$, 如果 $\text{le}_\prec(\mathcal{H})$ 是幺半群理想 $\text{le}_\prec(I)$ 的弱对合基, 则称 \mathcal{H} 是 I 的弱对合基. 如果 $\text{le}_\prec(\mathcal{H})$ 是 $\text{le}_\prec(I)$ 的强对合基并且 \mathcal{H} 中多项式的首项指数互不相同, 则称集合 \mathcal{H} 是理想 I 的强对合基.

$\mathbb{K}[x]$ 中的理想不一定存在有限的 Pommaret 基(参见 [89]).

定义 3.7. [89, 定义 4.3.1] 给定理想 I , 如果对项序 \prec , I 存在有限的 Pommaret 基, 则称坐标系关于理想 $I \subseteq \mathbb{K}[x]$ 和项序 \prec 是 δ -正则的.

定理 3.4. [89, 定理 4.3.15] 给定项序 \prec 和理想 $I \subseteq \mathbb{K}[x]$, 存在一个坐标系, 使得在这个坐标系下, I 存在有限的 Pommaret 基.

定义 3.8. [89, 定义 3.4.2] 给定一个有限多项式集合 $\mathcal{F} \subset \mathbb{K}[x] \setminus \{0\}$ 和 \mathbb{N}^n 中的对合除法 L . 任意的多项式 $f \in \mathcal{F}$ 对应一个乘法变量的集合

$$X_{L,\mathcal{F},\prec}(f) = \{x_i \mid i \in N_{L,\text{le-}\prec,\mathcal{F}}(\text{le-}\prec f)\}.$$

\mathcal{F} 的对合域定义为集合

$$\langle \mathcal{F} \rangle_{L,\prec} = \sum_{f \in \mathcal{F}} \mathbb{K}[X_{L,\mathcal{F},\prec}(f)] \cdot f \subseteq \langle \mathcal{F} \rangle.$$

定理 3.5. [89, 定理 3.4.4] 给定 \mathbb{N}^n 中的对合除法 L . 对一个非零理想 $I \subseteq \mathbb{K}[x]$, 一个有限集合 $\mathcal{H} \subset I \setminus \{0\}$, 在项序 \prec 下, 下面的结论是等价的:

(i) 集合 $\mathcal{H} \subset I$ 是 I 的弱对合基.

(ii) 理想 I 中的任意多项式 f 具有下面的表达式:

$$f = \sum_{h \in \mathcal{H}} P_h \cdot h, \quad (3.5)$$

这里, 对与非零多项式 $P_h \in \mathbb{K}[X_{L,\mathcal{H},\prec}(h)]$ 对应的多项式 $h \in \mathcal{H}$, 有 $\text{lt}_{\prec}(P_h \cdot h) \preceq \text{lt}_{\prec}(f)$.

\mathcal{H} 是一组强对合基当且仅当表达式 (3.5) 是唯一的.

注 3.2. 由定义 3.5, 定义 3.6 和表达式 (3.5) 可知弱对合基是 Gröbner 基.

推论 3.6. [89, 推论 3.4.5] 给定理想 $I \subseteq \mathbb{K}[x]$ 和 I 的弱对合基 \mathcal{H} . 则 $\langle \mathcal{H} \rangle_{L,\prec} = I$. 如果 \mathcal{H} 是 I 的强对合基, 则当 I 看作是一个 \mathbb{K} -线性空间时, I 存在直和分解

$$I = \bigoplus_{h \in \mathcal{H}} \mathbb{K}[X_{L,\mathcal{H},\prec}(h)] \cdot h.$$

命题 3.7. [89, 命题 3.4.7] 给定一个理想 $I \subseteq \mathbb{K}[x]$ 和 I 关于对合除法 L 的对合基 \mathcal{H} , 存在 \mathcal{H} 的子集 \mathcal{H}' , 使得 \mathcal{H}' 是 I 的一组强对合基.

定义 3.9. 如果将 $\mathbb{K}[x]$ 看作是一个线性空间, 则理想 I 和截断理想 $I_t = I \cap \mathbb{K}[x]_t$ 均为 $\mathbb{K}[x]$ 中的子空间. 给定一个项序和一个多项式集合 $G = \{g_1, \dots, g_s\}$. 如果 G 是 I_t 中线性无关的一组基并且 G 中多项式的首项单项式互不相同, 则称 G 是 I_t 的一组约化基.

3.3 终止条件和主要结果

在这一节, 当 $V_{\mathbb{R}}(I)$ 是正维时, 我们通过对合理论提出一个终止条件. 当终止条件成立时, 算法返回一个介于 I 和 $I(V_{\mathbb{R}}(I))$ 的理想 J 的 Pommaret 基.

3.3.1 终止条件

给定理想 $I := \langle h_1, \dots, h_m \rangle \subseteq \mathbb{R}[x]$, 由(3.1), (3.2)和(3.3), 有

$$d := \max_{1 \leq j \leq m} d_j, \quad d_j := \lceil \deg(h_j)/2 \rceil.$$

对每一个 $t \geq d$, 定义集合

$$\mathcal{K}_t := \{y \in \mathbb{R}^{\mathbb{N}_{2t}^n} \mid y_0 = 1, M_t(y) \succeq 0, M_{t-d_j}(h_j y) = 0, j = 1, \dots, m\},$$

并且

$$\mathcal{K}_t^{gen} := \{y \in \mathcal{K}_t \mid M_t(y) \text{ 的秩在 } \mathcal{K}_t \text{ 中是最大的}\}.$$

给定第 t 阶截断矩量矩阵 $M_t(y)$ 和 $\ell < t$, 截断矩量矩阵 $M_{t-\ell}(y)$ 表示 $M_t(y)$ 的第 $t - \ell$ 阶主子矩阵, 并且它的行和列以 $\mathbb{N}_{t-\ell}^n$ 为索引.

令 α_j 为 $\ker M_{t-2}(y)$ 中多项式的一组约化基中类为 j 的 $t - 2$ 次多项式的个数. 尽管 $\ker M_{t-2}(y)$ 中多项式的约化基并不唯一, 由于它们可以互相线性表出, 所以具有相同的首项集合. 因此 $\sum_{j=1}^n j\alpha_j$ 的值不依赖 $\ker M_{t-2}(y)$ 中约化基的选取. 此外, 根据命题3.2, 对 $1 \leq s \leq t$, $\ker M_t(y) \cap \mathbb{R}[x]_s = \ker M_s(y)$, 在算法3.1的步骤2中, $\ker M_{t-2}(y)$ 中多项式的约化基可以通过在 $\ker M_{t-1}(y)$ 中多项式的约化基中选取次数至多为 $t - 2$ 的多项式得到.

定理 3.8. 给定 $\mathbb{R}[x]$ 中的理想 $I = \langle h_1, \dots, h_m \rangle$, 假设对整数 $t \geq 2d$ 和 $y \in \mathcal{K}_t^{gen}$, 等式

$$\sum_{j=1}^n j\alpha_j = \operatorname{corank} M_{t-1}(y) - \operatorname{corank} M_{t-2}(y) \tag{3.6}$$

成立. 则 $M_{t-2}(y)$ 的零空间的一组约化基是理想 $J = \langle \ker M_{t-2}(y) \rangle$ 在单项式序 $\prec_{\operatorname{tdeg}}$ 下的弱 Pommaret 基并且

$$I \subseteq J \subseteq I(V_{\mathbb{R}}(I)), \quad V_{\mathbb{R}}(I) = V_{\mathbb{C}}(J) \cap \mathbb{R}^n. \tag{3.7}$$

定理3.8的证明利用命题3.3和定理3.11, 其证明在第3.3.3节中给出. 我们的算法需要在 \mathcal{K}_t 中找到母元素 y . 与文献 [48]相同, 截断矩量序列 y 可以通过利用自对偶嵌入的内点算法求解下面的半正定规划得到:

$$\min \quad 0 \quad \text{使得 } y \in \mathcal{K}_t \quad (3.8)$$

参见 [96, 98].

3.3.2 一个计算Pommaret基的算法

我们提出一个基于求解(3.8)的算法. 算法返回一个介于 I 和 $\mathrm{I}(\mathrm{V}_{\mathbb{R}}(I))$ 之间的理想 $J = \langle \ker M_{t-2}(y) \rangle$ 的Pommaret基.

算法 3.1. 计算一个理想 J 的Pommaret基, 使得 $I \subseteq J \subseteq \mathrm{I}(\mathrm{V}_{\mathbb{R}}(I))$.

输入: 理想 I 的生成元 $\{h_1, \dots, h_m\}$ 和定义在变量 x_1, \dots, x_n 上的单项式序 \prec_{tdeg} .

输出: 理想 $J = \langle \ker M_{t-2}(y) \rangle$ 的Pommaret基.

步骤 1 对 $t \geq 2d$, 求解半正定规划问题(3.8)中 \mathcal{K}_t 的母元素.

步骤 2 计算 $\ker M_{t-1}(y)$ 中多项式的一组约化基.

- 选取在 $\ker M_{t-1}(y)$ 中次数至多为 $t-2$ 的多项式集合 $\{g_1, \dots, g_{s+r}\}$, 其中对 $1 \leq i \leq s$, $\deg(g_i) = t-2$, 对 $s+1 \leq i \leq s+r$, $\deg(g_i) < t-2$. 集合 $\{g_1, \dots, g_{s+r}\}$ 为 $\ker M_{t-2}(y)$ 中多项式的一组约化基.
- 计算 $\sum_{j=1}^n j\alpha_j$, 其中 α_j 表示 $\{g_1, \dots, g_s\}$ 中类为 j 的多项式的个数.

步骤 3 通过计算 $\ker M_{t-1}(y)$ 中多项式的一组约化基中 $t-1$ 次多项式的个数来得到表达式 $\mathrm{corank} M_{t-1}(y) - \mathrm{corank} M_{t-2}(y)$ 的值.

步骤 4 检验条件(3.6)是否满足.

- 如果满足, 则 $\{g_1, \dots, g_{s+r}\}$ 是 $\langle \ker M_{t-2}(y) \rangle$ 的一组弱Pommaret基并且被进一步约化成Pommaret基.
- 否则, 令 $t := t + 1$ 并且返回步骤1.

在第3.3.3节中, 我们证明算法3.1是正确的并且在关于 $\sqrt[3]{I}$ 的 δ -正则坐标系下有限步终止. 算法利用Matlab软件中的GloptiPoly工具包 [42]并且通过第3.4节的例子展示算法的实际效果.

3.3.3 终止条件的证明

在这一节, 我们证明算法3.1的正确性并且在 \sqrt{I} 的 δ -正则坐标系下有限终止.

假设1. 给定理想 $I = \langle h_1, \dots, h_m \rangle \subseteq \mathbb{R}[x]$. 假设存在整数 $t \geq 2d$, 使得对 $y \in \mathcal{K}_t^{gen}$, 条件(3.6)满足. 给定 $\ker M_{t-2}(y)$ 中多项式的一组约化基 $\{g_1, \dots, g_{s+r}\}$. 不妨设

$$\deg(g_i) = t - 2, \quad 1 \leq i \leq s \text{ 并且 } \deg(g_i) < t - 2, \quad s + 1 \leq i \leq s + r.$$

引理 3.9. 在假设1下, 多项式集合

$$\{x_1g_1, \dots, x_{j_1}g_1, \dots, x_1g_s, \dots, x_{j_s}g_s, g_1, \dots, g_{s+r}\}$$

是 $\ker M_{t-1}(y)$ 中多项式的一组约化基, 这里 $i = 1, \dots, s$, $j_i = \text{cls}(g_i)$.

证明. 对 $k = 1, \dots, n$, $i = 1, \dots, s + r$, 由于 $\deg(x_k g_i) \leq t - 1$, 通过命题3.2(i), 可得 $x_k g_i \in \ker M_{t-1}(y)$. 实际上, 集合 $\{g_1, \dots, g_{s+r}\}$ 中多项式的首项互不相同. 根据定义3.9, 多项式

$$x_1g_1, \dots, x_{j_1}g_1, \dots, x_1g_s, \dots, x_{j_s}g_s \tag{3.9}$$

的首项互不相同, 并且次数均为 $t - 1$. 因此, 它们是线性无关的. 假设在 $\{g_1, \dots, g_s\}$ 中有 α_j 个类为 j 的多项式, 则由(3.9), $\ker M_{t-1}(y)$ 中有 $\sum_{j=1}^n j\alpha_j$ 个线性无关的 $t - 1$ 次多项式. 另一方面, $\ker M_{t-1}(y)$ 中 $t - 1$ 次线性无关的多项式的个数等于 $\text{corank } M_{t-1}(y) - \text{corank } M_{t-2}(y)$. 因此, 由条件(3.6)和命题3.2(iii)可得结论的正确性. \square

注 3.3. 如果假设1成立, $\ker M_{t-1}(y)$ 中的任意多项式 f 可以表示成下面的线性组合:

$$f = \sum_{k=1}^s \sum_{i=1}^{\text{cls}(g_k)} c_{ik} x_i g_k + \sum_{k=1}^{s+r} \lambda_k g_k, \tag{3.10}$$

这里, $c_{ik} \in \mathbb{R}$ 并且对 $1 \leq i \leq \text{cls}(g_k)$ 和 $1 \leq k \leq s$, 有 $\text{lt}_{\prec}(c_{ik} x_i g_k) \preceq_{\text{tdeg}} \text{lt}_{\prec}(f)$. 对 $1 \leq k \leq s + r$, 有 $\lambda_k \in \mathbb{R}$ 并且 $\text{lt}_{\prec}(\lambda_k g_k) \preceq_{\text{tdeg}} \text{lt}_{\prec}(f)$. 注意集合

$$\{x_1g_1, \dots, x_{j_1}g_1, \dots, x_1g_s, \dots, x_{j_s}g_s, g_1, \dots, g_{s+r}\}$$

中多项式的首项互不相同. 在单项式序 \prec_{tdeg} 下, 如果 c_{ik} 不都为零, 则存在唯一的非零系数 $c_{i_0 k_0}$, 使得 $\text{lt}_{\prec}(x_{i_0} g_{k_0}) = \text{lt}_{\prec}(f)$. 如果所有的 c_{ik} 均为零, 则存在唯一的整数 $k \in [1, s+r]$, 使得 $\lambda_k \neq 0$ 并且 $\text{lt}_{\prec}(g_k) = \text{lt}_{\prec}(f)$. 这个性质在接下来的定理证明中被使用.

引理 3.10. 如果假设 1 成立, 对所有的单项式 x^{μ} 和集合 $\{g_{s+1}, \dots, g_{s+r}\}$ 中次数小于 $t-2$ 的多项式 g_j , 多项式 $x^{\mu} g_j$ 存在下面的表达式:

$$x^{\mu} g_j = \sum_{k=1}^s p_k g_k + \sum_{k=s+1}^{s+r} \lambda_k g_k, \quad (3.11)$$

其中 $p_k \in \mathbb{R}[x]$, $\lambda_k \in \mathbb{R}$ 并且 $\text{lt}_{\prec}(p_k g_k) \preceq_{\text{tdeg}} \text{lt}_{\prec}(x^{\mu} g_j)$, $k = 1, \dots, s$, $\text{lt}_{\prec}(\lambda_k g_k) \preceq_{\text{tdeg}} \text{lt}_{\prec}(x^{\mu} g_j)$, $k = s+1, \dots, s+r$.

证明. 如果 $\deg(x^{\mu} g_j) \leq t-1$, 通过命题 3.2(i), 可得 $x^{\mu} g_j \in \ker M_{t-1}(y)$. 根据注 3.3, 可得表达式(3.11). 否则, 假设 $x^{\mu} = x^{\mu_1} x^{\mu_2}$ 满足条件 $\deg(x^{\mu_2} g_j) = t-1$, 可得

$$\begin{aligned} x^{\mu} g_j &= x^{\mu_1} x^{\mu_2} g_j = x^{\mu_1} \left(\sum_{k=1}^s p_k g_k + \sum_{k=s+1}^{s+r} \lambda_k g_k \right) \\ &= \sum_{k=1}^s x^{\mu_1} p_k g_k + \sum_{k=s+1}^{s+r} \lambda_k x^{\mu_1} g_k. \end{aligned}$$

对整数 $k \in [s+1, s+r]$, 令 $x^{\mu_1} g_k$ 重复上面的步骤. 由于 $\deg(x^{\mu_1}) < \deg(x^{\mu})$, 经过有限步, 可得形如(3.11)的表达式. \square

定理 3.11. 如果假设 1 成立, $\ker M_{t-2}(y)$ 中多项式的一组约化基 $\{g_1, \dots, g_{s+r}\}$ 是理想 $\langle \ker M_{t-2}(y) \rangle$ 的弱 Pommaret 基.

证明. 只需证明 $\langle \ker M_{t-2}(y) \rangle$ 中的任意一个多项式 f 可以表示成

$$f = \sum_{k=1}^s p_k g_k + \sum_{k=s+1}^{s+r} \lambda_k g_k, \quad (3.12)$$

这里 $\lambda_k \in \mathbb{R}$, $p_k \in \mathbb{R}[x_1, \dots, x_{\text{cls}(g_k)}]$. 由于对整数 $k \in [1, s+r]$, $\text{lt}_{\prec}(p_k g_k)$ 和 $\text{lt}_{\prec}(g_k)$ 互不相同, 如果 f 满足(3.12), 则对整数 $k \in [1, s]$, $\text{lt}_{\prec}(p_k g_k) \preceq_{\text{tdeg}} \text{lt}_{\prec}(f)$ 并且对整

数 $k \in [s+1, s+r]$, 可得 $\text{lt}_\prec(\lambda_k g_k) \preceq_{\text{tdeg}} \text{lt}_\prec(f)$. 根据定理3.5, 集合 $\{g_1, \dots, g_{s+r}\}$ 是理想 $\langle \ker M_{t-2}(y) \rangle$ 的弱Pommaret基.

由于 $\{g_1, \dots, g_{s+r}\}$ 是 $\ker M_{t-2}(y)$ 中多项式的一组约化基, 理想 $\langle \ker M_{t-2}(y) \rangle$ 中任意的多项式 f 可以表示成

$$f = \sum_{j=1}^{s+r} p_j g_j,$$

其中 $p_j \in \mathbb{R}[x]$, $j = 1, \dots, s+r$. 因此, 只需证明对于任意的 $\mu \in \mathbb{N}^n$ 和整数 $k \in [1, s+r]$, $x^\mu g_j$ 存在形如(3.12)的表达式.

令 $f = x^\mu g_j$. 如果 $\deg(f) \leq t-1$, 通过引理3.9, 可得 f 存在形如(3.12)的表达式. 否则, 对首项 $\text{lt}_\prec(f) = t_0$ 归纳来证明, 即假设对 $\mu \in \mathbb{N}^n$ 和整数 $k \in [1, s+r]$, 只要 $\text{lt}_\prec(f) \prec_{\text{tdeg}} t_0$, 则 f 存在形如(3.12)的表达式. 当 $\text{lt}_\prec(f) = t_0$ 时, f 存在形如(3.12)的表达式.

如果 $x^\mu \in \mathbb{R}[x_1, \dots, x_{\text{cls}(g_j)}]$, 直接可证. 否则, 不失一般性, 令 x_{i_1} 为 x^μ 中关于 g_j 的一个非乘法变量, 即 $i_1 \notin \{1, \dots, \text{cls}(g_j)\}$. 由于对整数 $j \in [1, s+r]$, $\deg(g_j) \leq t-2$, 通过命题3.2(i), 可得 $x_{i_1} g_j \in \ker M_{t-1}(y)$. 通过引理3.9和注3.3, 可得

$$\begin{aligned} f &= x^\mu g_j = (x^\mu / x_{i_1}) x_{i_1} g_j \\ &= (x^\mu / x_{i_1}) \left(\sum_{k=1}^s \sum_{i=1}^{\text{cls}(g_k)} c_{ik} x_i g_k + \sum_{k=1}^{s+r} \lambda_k g_k \right) \\ &= \sum_{k=1}^s \sum_{i=1}^{\text{cls}(g_k)} c_{ik} (x^\mu / x_{i_1}) x_i g_k + \sum_{k=1}^{s+r} \lambda_k (x^\mu / x_{i_1}) g_k. \end{aligned} \quad (3.13)$$

根据注3.3, 考虑下面两种情形:

- (i) 如果所有的 c_{ik} 均为0, 则存在唯一的整数 $j_1 \in [1, s+r]$, 使得 $\lambda_{j_1} \neq 0$ 并且 $\text{lt}_\prec(\lambda_{j_1} (x^\mu / x_{i_1}) g_{j_1}) = t_0$;
- (ii) 否则, 存在整数 $j_1 \in [1, s]$ 和 $i_2 \in [1, \text{cls}(g_{j_1})]$, 使得

$$c_{i_2 j_1} \neq 0, \quad \text{lt}_\prec(c_{i_2 j_1} (x^\mu / x_{i_1}) x_{i_2} g_{j_1}) = t_0.$$

在这两种情况下, (3.13)中所有其它项的次数均小于 t_0 , 通过归纳, 它们存在形如(3.12)的表达式. 此外, 这两种情况不能同时存在. 因此只需验证在第一种情况下, 多项式 $\lambda_{j_1}(x^\mu/x_{i_1})g_{j_1}$ 或者在第二种情况下, 多项式 $c_{i_2j_1}(x^\mu/x_{i_1})x_{i_2}g_{j_1}$ 都存在形如(3.12)的表达式.

情形(i): 如果 $x^\mu/x_{i_1} \in \mathbb{R}[x_1, \dots, x_{\text{cls}(g_{j_1})}]$, 则 $\lambda_{j_1}(x^\mu/x_{i_1})g_{j_1}$ 存在形如(3.12)的表达式. 否则, 对 $(x^\mu/x_{i_1})g_{j_1}$ 进行归纳. 由于 $\text{lt}_\prec(\lambda_{j_1}(x^\mu/x_{i_1})g_{j_1}) = \text{lt}_\prec(x^\mu g_j) = t_0$, 因此 $\deg(g_j) < \deg(g_{j_1})$, 即

$$\text{lt}_\prec(g_j) \prec_{\text{tdeg}} \text{lt}_\prec(g_{j_1}).$$

情形(ii): 如果 $x^\mu/x_{i_1} \in \mathbb{R}[x_1, \dots, x_{\text{cls}(g_{j_1})}]$, 由于 x_{i_2} 是 $\text{lt}_\prec(g_{j_1})$ 的乘法变量, 因此 $(x^\mu/x_{i_1})x_{i_2} \in \mathbb{R}[x_1, \dots, x_{\text{cls}(g_{j_1})}]$ 并且 $(x^\mu/x_{i_1})x_{i_2}g_{j_1}$ 存在形如(3.12)的表达式. 否则, 由于 x_{i_1} 不是 $\text{lt}_\prec(g_j)$ 的乘法变量并且 x_{i_2} 是 $\text{lt}_\prec(g_{j_1})$ 的乘法变量, 可得

$$\text{cls}(g_j) < \text{cls}(x_{i_1}), \quad \text{cls}(x_{i_2}) \leq \text{cls}(g_{j_1}).$$

由于 $\text{lt}_\prec(c_{i_2j_1}(x^\mu/x_{i_1})x_{i_2}g_{j_1}) = t_0$, 因此 $\text{lt}_\prec(x_{i_2}g_{j_1}) = \text{lt}_\prec(x_{i_1}g_j)$ 并且

$$\text{cls}(x_{i_2}) = \text{cls}(x_{i_2}g_{j_1}) = \text{cls}(x_{i_1}g_j) < \text{cls}(x_{i_1}). \quad (3.14)$$

因此 $x_{i_2} \prec_{\text{tdeg}} x_{i_1}$. 如果 $\text{lt}_\prec(g_{j_1}) \preceq_{\text{tdeg}} \text{lt}_\prec(g_j)$, 可得 $\text{lt}_\prec(x_{i_2}g_{j_1}) \prec_{\text{tdeg}} \text{lt}_\prec(x_{i_1}g_j)$, 导出矛盾. 因此,

$$\text{lt}_\prec(g_j) \prec_{\text{tdeg}} \text{lt}_\prec(g_{j_1}).$$

在两种情形下, 如果归纳不终止, 可得满足

$$\text{lt}_\prec(g_j) \prec_{\text{tdeg}} \text{lt}_\prec(g_{j_1}) \prec_{\text{tdeg}} \cdots \prec_{\text{tdeg}} \text{lt}_\prec(g_{j_i}) \prec_{\text{tdeg}} \text{lt}_\prec(g_{j_{i+1}}) \prec_{\text{tdeg}} \cdots \prec_{\text{tdeg}} t_0$$

的单项式序列. 由于这个多项式序列严格递增并且每一个单项式的首项均小于 $\text{lt}_\prec(f) = t_0$, 上面的归纳有限步终止. 因此 f 存在形如(3.12)的表达式. \square

定理 3.12. 在 $\sqrt{\mathbb{R}I}$ 的 δ -正则坐标系下, 算法 3.1 有限步终止并且返回正整数 $t \geq 2d$, 使得对 $y \in \mathcal{K}_t^{gen}$, 条件(3.6)满足.

证明. 在 δ -正则坐标系下, $I(V_{\mathbb{R}}(I))$ 存在有限的对合基 $\mathcal{H} = \{h_1, \dots, h_s\}$. 根据命题3.3(iii), 存在正整数 $t_1 \in \mathbb{N}$, 使得对所有的 $y \in \mathcal{K}_t$ 和 $t \geq t_1$, Pommaret 基 $\{h_1, \dots, h_s\}$ 包含在 $\ker M_t(y)$ 中.

由于 \mathcal{H} 是 $I(V_{\mathbb{R}}(I))$ 的 Pommaret 基, 根据推论 3.6, 对 $t \geq t_1 + 2$, $I(V_{\mathbb{R}}(I))_{t-2}$ 存在下面的分解:

$$I(V_{\mathbb{R}}(I))_{t-2} = \bigoplus_{h_k \in \mathcal{H}} \mathbb{R}[x_1, \dots, x_{\text{cls}(h_k)}]_{t-2-\deg(h_k)} \cdot h_k. \quad (3.15)$$

令

$$T = \{x^u h_k \mid x^u \in \mathbb{R}[x_1, \dots, x_{\text{cls}(h_k)}], \deg(x^u) \leq t - 2 - \deg(h_k), 1 \leq k \leq s\}. \quad (3.16)$$

根据命题 3.2(i), 可得 $T \subseteq \ker M_{t-2}(y)$. 因此, 通过(3.15)和(3.16), 可得

$$I(V_{\mathbb{R}}(I))_{t-2} \subseteq \ker M_{t-2}(y).$$

另一方面, 由于 y 是母元素, 通过命题 3.3(i), 可得

$$\ker M_{t-2}(y) \subseteq I(V_{\mathbb{R}}(I))_{t-2}.$$

因此, 可得 $\ker M_{t-2}(y) = I(V_{\mathbb{R}}(I))_{t-2}$ 和下面的分解:

$$\ker M_{t-2}(y) = \bigoplus_{h_k \in \mathcal{H}} \mathbb{R}[x_1, \dots, x_{\text{cls}(h_k)}]_{t-2-\deg(h_k)} \cdot h_k. \quad (3.17)$$

由于 \mathcal{H} 是 $I(V_{\mathbb{R}}(I))$ 的 Pommaret 基, 根据定义 3.6, T 中多项式的首项互不相同. 因此, T 是 $\ker M_{t-2}(y)$ 中的一组约化基. 通过定理 3.8, 只需证明条件(3.6)对 T 中的多项式成立.

与分解(3.17)相似, $\ker M_{t-1}(y)$ 存在直和分解:

$$\ker M_{t-1}(y) = \bigoplus_{h_k \in \mathcal{H}} \mathbb{R}[x_1, \dots, x_{\text{cls}(h_k)}]_{t-1-\deg(h_k)} \cdot h_k. \quad (3.18)$$

根据(3.18), $\ker M_{t-1}(y)$ 中任意的 $t-1$ 次多项式 f 存在下面的表达式:

$$\begin{aligned} f &= \sum_{k=1}^s \sum_{0 \leq |\mu| \leq t-1-\deg(h_k)} c_{\mu k} x^\mu h_k \quad (x^\mu \in \mathbb{R}[x_1, \dots, x_{\text{cls}(h_k)}]) \\ &= \sum_{k=1}^s \sum_{|\mu|=t-1-\deg(h_k)} c_{\mu k} x^\mu h_k + \sum_{k=1}^s \sum_{0 \leq |\mu| \leq t-2-\deg(h_k)} c_{\mu k} x^\mu h_k \\ &= \sum_{k=1}^s \sum_{|\mu|=t-1-\deg(h_k)} c_{\mu k} x^{\text{cls}(x^\mu)} (x^\mu / x_{\text{cls}(x^\mu)}) h_k + \sum_{k=1}^s \sum_{0 \leq |\mu| \leq t-2-\deg(h_k)} c_{\mu k} x^\mu h_k. \end{aligned}$$

由于 $x_{\text{cls}(x^\mu)}$ 是多项式 $(x^\mu/x_{\text{cls}(x^\mu)})h_k \in T$ 的乘法变量, $\ker M_{t-1}(y)$ 中任意的多项式可以由 T 和 T_1 中的多项式表示, 这里

$$T_1 = \{x_i g \mid 1 \leq i \leq \text{cls}(g), g \in T, \deg(g) = t-2\}.$$

T_1 和 T 中多项式的首项互不相同. 因此 $T \cup T_1$ 是 $\ker M_{t-1}(y)$ 中多项式的一组基. 此外, T 是 $\ker M_{t-2}(y)$ 中多项式的一组约化基并且 T_1 中多项式首项的集合与 $\ker M_{t-1}(y)$ 中所有 $t-1$ 次多项式首项的集合相同. 因此, T_1 中多项式的个数等于 $\text{corank } M_{t-1}(y) - \text{corank } M_{t-2}(y)$. 另一方面, 令 α_j 为 T 中类为 j 的 $t-2$ 次多项式的个数. 由于集合 T_1 由 T 中次数为 $t-2$ 的多项式与它们的乘法变量的积组成. T_1 中多项式的个数等于 $\sum_{j=1}^n j \alpha_j$. 因此, 条件(3.6)成立. \square

3.3.4 推广到 $V(I_{\mathbb{R}}(I) \cap \mathcal{A})$

考虑半代数集

$$\mathcal{A} := \{x \in \mathbb{R}^n \mid f_1(x) \geq 0, \dots, f_s(x) \geq 0\}, \quad (3.19)$$

这里 $f_1, \dots, f_s \in \mathbb{R}[x]$. \mathcal{A} -仿射簇 $V_{\mathcal{A}}(I)$ 定义为

$$V_{\mathcal{A}}(I) = V_{\mathbb{R}}(I) \cap \mathcal{A}.$$

对于任意的 $\nu \in \{0, 1\}^s$, $f^\nu := f_1^{\nu_1} f_2^{\nu_2} \cdots f_s^{\nu_s}$.

定义 3.10. [59] 理想 I 的 \mathcal{A} -根理想定义为

$$\sqrt[\mathcal{A}]{I} := \left\{ p \in \mathbb{R}[x] \mid \text{存在 } k \in \mathbb{N}, \sigma_\nu \in \sum \mathbb{R}[x]^2, \text{ 使得 } p^{2k} + \sum_{\nu \in \{0,1\}^s} \sigma_\nu f^\nu \in I \right\}.$$

如果 $I = \sqrt[\mathcal{A}]{I}$, 则称 I 为 \mathcal{A} -根理想.

定理 3.13. [94, 半代数集零点定理] 令 I 为 $\mathbb{R}[x]$ 中的理想并且 \mathcal{A} 由 (3.19) 定义. 则 $\sqrt[\mathcal{A}]{I}$ 是 \mathcal{A} -根理想并且 $\sqrt[\mathcal{A}]{I} = I(V_{\mathbb{R}}(I) \cap \mathcal{A})$.

为了计算 \mathcal{A} -根理想 $\sqrt[\mathcal{A}]{I}$, 考虑集合

$$\mathcal{K}_{t,\mathcal{A}} := \mathcal{K}_t \cap \left\{ y \in \mathbb{R}^{\mathbb{N}_{2t}^n} \mid \text{对任意的 } \nu \in \{0, 1\}^s, M_{t-d_{f^\nu}}(f^\nu y) \succeq 0 \right\}, \quad (3.20)$$

这里 $d_{f^\nu} = \lceil \deg(f^\nu)/2 \rceil$. 显然地, 集合 $\mathcal{K}_{t,\mathcal{A}}$ 是 \mathcal{K}_t 的子集. 集合 $\mathcal{K}_{t,\mathcal{A}}$ 的定义基于 $\sqrt[\mathcal{A}]{I}$ 的定义和半代数集的零点定理. $\mathcal{K}_{t,\mathcal{A}}$ 中母元素的集合定义为

$$\mathcal{K}_{t,\mathcal{A}}^{gen} := \{y \in \mathcal{K}_{t,\mathcal{A}} \mid \text{rank } M_t(y) \text{ 在 } \mathcal{K}_{t,\mathcal{A}} \text{ 中是最大的}\}.$$

引理 3.14. [48, 注 4.9] 令 $\{g_1, \dots, g_k\}$ 为理想 $\sqrt[4]{I}$ 的生成元. 则存在 $t_0 \in \mathbb{N}$, 使得对任意的 $y \in \mathcal{K}_{t,\mathcal{A}}$ 和 $t \geq t_0$, 可得 $g_1, \dots, g_k \in \ker M_t(y)$.

根据引理 3.14, 存在 $t_0 \in \mathbb{N}$, 使得对任意的 $y \in \mathcal{K}_{t,\mathcal{A}}^{gen}$, $t \geq t_0$, 有 $\langle \ker M_t(y) \rangle = \sqrt[4]{I}$. 因此, 对充分大的 t , $\sqrt[4]{I}$ 的信息包含在母元素 $y \in \mathcal{K}_{t,\mathcal{A}}$ 中. 以上讨论的命题和定理对于 $\mathcal{K}_{t,\mathcal{A}}$ 中的母元素 y 均成立.

如果将 \mathcal{K}_t 和 $\sqrt[4]{I}$ 分别替换为 $\mathcal{K}_{t,\mathcal{A}}$ 和 $\sqrt[4]{I}$, 下面的定理是定理 3.8 对半代数集情形的推广. 其证明与定理 3.11 和定理 3.12 相同.

定理 3.15. 令 $t \geq 2d$. 假设条件 (3.6) 对 $y \in \mathcal{K}_{t,\mathcal{A}}$ 成立. $\ker M_{t-2}(y)$ 中多项式的约化基是 $\langle \ker M_{t-2}(y) \rangle$ 关于单项式序 \prec_{tdeg} 的弱 Pommaret 基并且

$$I \subseteq \langle \ker M_{t-2}(y) \rangle \subseteq I(V_{\mathbb{R}}(I) \cap \mathcal{A}).$$

注 3.4. 为了计算 $\langle \ker M_{t-2}(y) \rangle$ 的弱 Pommaret 基, 我们将半代数集 \mathcal{A} 的定义多项式 $\{f_1, \dots, f_s\}$ 添加到上面算法的输入中. 对 $\nu \in \{0, 1\}^s$, 将约束 $M_{t-d_f\nu}(f^\nu y) \succeq 0$ 添加到半正定规划问题 (3.8) 的约束条件中.

3.4 数值试验

在这一节, 我们列举一些由算法 3.1 所计算的例子的结果. 例子引自 [78, 88, 90, 95]. 给定一个矩阵和容忍度 τ , 如果矩阵的奇异值满足 $\sigma_1 \geq \dots \geq \sigma_k > \tau > \sigma_{k+1}$ 或者 $\sigma_k / \sigma_{k+1} > 10^3$, 则这个矩阵的数值秩定义为 k .

Spang 在 SINGULAR 中设计一个符号算法 realrad 来计算在有理数域的超越扩张上理想的实根理想. 由于算法 realrad 是基于 Wu-Ritt 的特征集方法和 Gröbner 基的计算, 它具有双指数复杂度. 我们的算法基于线性代数和半正定规划, 具有多项式的复杂度. 然而, 算法返回的结果有数值误差. 通过 realrad, 我们能够证明例子 3.2, 3.3, 3.5 是实根理想, 即 $I = \sqrt[4]{I}$ 以及例子 3.4, 3.6 不是实根理想, 即 $I \subset \sqrt[4]{I}$.

例 3.2. 考虑 [95, 397 页, 方程 (9.60)] 中的 2 维理想 $I = \langle h_1, h_2, h_3 \rangle$,

$$\begin{aligned} h_1 &= x_1^2 + x_1x_2 - x_1x_3 - x_1 - x_2 + x_3, \\ h_2 &= x_1x_2 + x_2^2 - x_2x_3 - x_1 - x_2 + x_3, \\ h_3 &= x_1x_3 + x_2x_3 - x_3^2 - x_1 - x_2 + x_3. \end{aligned}$$

截断矩量矩阵 $M_{t-\ell}(y)$ 的秩和余秩序列在表3.1和表3.2中表示。令 $\tau = 10^{-5}$ 并且 $x_1 \prec_{\text{tdeg}} x_2 \prec_{\text{tdeg}} x_3$ 。对 $t=4$, 可得

$$\sum_{j=1}^3 j\alpha_j = 6, \text{ 并且 corank } M_{4-1} - \text{corank } M_{4-2} = 6.$$

因此条件(3.6)满足。对 $t = 4$, 算法3.1返回的Pommaret基为

$$\begin{aligned} & \{x_1 + x_2 - x_3 - x_1^2 - x_1x_2 + x_1x_3, x_1 + x_2 - x_3 - x_1x_2 - x_2^2 + x_2x_3, \\ & \quad 3x_1 + 3x_2 - 3x_3 - x_1^2 - 2x_1x_2 - x_2^2 + x_3^2\}. \end{aligned}$$

由表3.3可知, 条件(3.6)对 $t = 5, 6, 7$ 成立。使用函数realrad计算这个例子, 可以验证 $I = \langle \ker M_{4-2}(y) \rangle = \sqrt[3]{I}$ 以及 $\ker M_{4-2}(y)$ 中的多项式包含 $\sqrt[3]{I}$ 的弱Pommaret基。因此, 当 $t \geq 4$ 时, 条件(3.6)成立。

表 3.1: $M_{t-\ell}(y)$ 的秩				表 3.2: $M_{t-\ell}(y)$ 的余秩			
阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$	阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$
$t=4$	16	11	7	$t=4$	19	9	3
$t=5$	22	16	11	$t=5$	34	19	9
$t=6$	29	22	16	$t=6$	55	34	19
$t=7$	37	29	22	$t=7$	83	55	34

表 3.3: $\ker M_{t-2}(y)$ 中多项式的约化基的类 α_j

阶数	α_1	α_2	α_3	$\sum_{j=1}^3 j\alpha_j$
$t=4$	1	1	1	$1 \times 1 + 2 \times 1 + 3 \times 1 = 6$
$t=5$	3	2	1	$1 \times 3 + 2 \times 2 + 3 \times 1 = 10$
$t=6$	6	3	1	$1 \times 6 + 2 \times 3 + 3 \times 1 = 15$
$t=7$	10	4	1	$1 \times 10 + 2 \times 4 + 3 \times 1 = 21$

例 3.3. 考虑文献 [88, 20页, 例子 1.4.6]中的多项式系统 $P = \{h_1, h_2\}$,

$$h_1 = x_1^2 - x_2,$$

$$h_2 = x_1x_2 - x_3.$$

对项序 $x_3 \prec_{\text{tdeg}} x_1 \prec_{\text{tdeg}} x_2$, 有 $\text{cls}(x_1) = 2$, $\text{cls}(x_2) = 3$, $\text{cls}(x_3) = 1$. 对 $\tau = 10^{-8}$, 截断矩量矩阵 $M_{t-\ell}(y)$ 的秩和余秩序列在表3.4和表3.5中表示.

表 3.4: $M_{t-\ell}(y)$ 的秩			
阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$
t=3	12	7	4
t=4	16	10	7
t=5	20	13	10

表 3.5: $M_{t-\ell}(y)$ 的余秩			
阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$
t=3	8	3	0
t=4	19	10	3
t=5	36	22	10

表 3.6: $\ker M_{t-2}(y)$ 中多项式的约化基的类 α_j				
阶数	α_1	α_2	α_3	$\sum_{j=1}^3 j\alpha_j$
t=3	0	0	0	$1 \times 0 + 2 \times 0 + 3 \times 0 = 0$
t=4	0	2	1	$1 \times 0 + 2 \times 2 + 3 \times 1 = 7$
t=5	3	3	1	$1 \times 3 + 2 \times 3 + 3 \times 1 = 12$

对 $t=4$, 可得

$$\sum_{j=1}^3 j\alpha_j = 7, \text{ 并且 corank } M_{4-1} - \text{corank } M_{4-2} = 7.$$

因此, 条件(3.6)成立. 当 $t = 4$ 时, 算法3.1返回的Pommaret基是

$$\{x_1^2 - x_2, x_1x_2 - x_3, x_2^2 - x_1x_3\}.$$

这与函数realrad所返回的结果相同. 此外, 由于 $x_2^2 - x_1x_3 = x_1(x_1x_2 - x_3) - x_2(x_1^2 - x_2)$, 可得 $I = \langle \ker M_{4-2}(y) \rangle = \sqrt[3]{I}$.

例 3.4. 考虑文献 [78, 123页, 例子 7.41]中的理想 $I = \langle h_1, h_2 \rangle$,

$$\begin{aligned} h_1 &= x_1^2 + x_2^2 + x_3^2 - 2, \\ h_2 &= x_1^2 + x_2^2 - x_3. \end{aligned}$$

使用函数realrad, 返回实根理想 $\sqrt{\mathbb{R}I}$ 的生成元

$$\{x_3 - 1, x_1^2 + x_2^2 - 1\}. \quad (3.21)$$

由于 $x_3 - 1 \notin I$, I 严格包含在 $\sqrt{\mathbb{R}I}$ 中, 即 $I \subset \sqrt{\mathbb{R}I}$. 令 $\tau = 10^{-8}$ 和 $x_1 \prec_{\text{tdeg}} x_2 \prec_{\text{tdeg}} x_3$. 截断矩量矩阵 $M_{t-\ell}(y)$ 的秩和余秩在表3.7和表3.8中表示.

表 3.7: $M_{t-\ell}(y)$ 的秩				表 3.8: $M_{t-\ell}(y)$ 的余秩			
阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$	阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$
t=3	7	5	3	t=3	13	5	1
t=4	9	7	5	t=4	26	13	5
t=5	11	9	7	t=5	45	26	13
t=6	13	11	9	t=6	71	45	26

表 3.9: $\ker M_{t-2}(y)$ 中多项式的约化基的类 α_j

阶数	α_1	α_2	α_3	$\sum_{j=1}^3 j\alpha_j$
t=3	0	0	1	$1 \times 0 + 2 \times 0 + 3 \times 1 = 3$
t=4	1	2	1	$1 \times 1 + 2 \times 2 + 3 \times 1 = 8$
t=5	4	3	1	$1 \times 4 + 2 \times 3 + 3 \times 1 = 13$
t=6	8	4	1	$1 \times 8 + 2 \times 4 + 3 \times 1 = 19$

对 $t = 4$, 可得

$$\sum_{j=1}^3 j\alpha_j = 8, \text{ 并且 corank } M_{4-1} - \text{corank } M_{4-2} = 8.$$

因此, 条件(3.6)成立. 当 $t = 4$ 时, 算法3.1返回的Pommaret基是

$$\{-1 + x_3, -1 + x_1^2 + x_2^2\}. \quad (3.22)$$

这与使用realrad返回的结果(3.21)相同. 因此 $\langle \ker M_{4-2}(y) \rangle = \sqrt[3]{I}$ 并且 $\ker M_{4-2}(y)$ 的约化基(3.22)是 $\sqrt[3]{I}$ 的Pommaret基. 因此, 当 $t \geq 4$ 时, 条件(3.6)成立.

例 3.5. 考虑文献 [90, 61页, 例子 2.4.12]中的理想 $I = \langle h_1, h_2, h_3 \rangle$, 这里

$$\begin{aligned} h_1 &= x_3^2 + x_2x_3 - x_1^2, \\ h_2 &= x_1x_3 + x_1x_2 - x_3, \\ h_3 &= x_2x_3 + x_2^2 + x_1^2 - x_1. \end{aligned}$$

令 $\tau = 10^{-7}$ 并且定义项序为 $x_1 \prec_{\text{tdeg}} x_2 \prec_{\text{tdeg}} x_3$. 截断矩量矩阵 $M_{t-\ell}(y)$ 的秩和余秩在表3.10和表3.11中表示.

表 3.10: $M_{t-\ell}(y)$ 的秩				表 3.11: $M_{t-\ell}(y)$ 的余秩			
阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$	阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$
t=4	13	10	7	t=4	22	10	3
t=5	16	13	10	t=5	40	22	10
t=6	19	16	13	t=6	65	40	22
t=7	22	19	16	t=7	98	65	40

表 3.12: $\ker M_{t-2}(y)$ 中多项式的约化基的类 α_j

阶数	α_1	α_2	α_3	$\sum_{j=1}^3 j\alpha_j$
t=4	1	1	1	$1 \times 1 + 2 \times 1 + 3 \times 1 = 6$
t=5	4	2	1	$1 \times 4 + 2 \times 2 + 3 \times 1 = 11$
t=6	8	3	1	$1 \times 8 + 2 \times 3 + 3 \times 1 = 17$
t=7	13	4	1	$1 \times 13 + 2 \times 4 + 3 \times 1 = 24$

当 $t = 4, \dots, 7$ 时, 条件(3.6)不成立. 事实上, Seiler在文献 [90]中指出坐标系 (x_1, x_2, x_3) 不是理想 I 的 δ -正则坐标系. 如果对坐标系做和 [90]相同的线性变换 $\tilde{x}_1 = x_3, \tilde{x}_2 = x_2 + x_3, \tilde{x}_3 = x_1$, 可得多项式系统 $\tilde{P} = \{\tilde{x}_1\tilde{x}_2 - \tilde{x}_3^2, \tilde{x}_2\tilde{x}_3 - \tilde{x}_1, \tilde{x}_2^2 - \tilde{x}_3\}$.

令 \tilde{I} 为由 \tilde{P} 生成的理想. 利用函数realrad返回的 $\sqrt{\tilde{I}}$ 的生成元为

$$\{\tilde{x}_1\tilde{x}_2 - \tilde{x}_3^2, \tilde{x}_2\tilde{x}_3 - \tilde{x}_1, \tilde{x}_2^2 - \tilde{x}_3, \tilde{x}_3^3 - \tilde{x}_1^2\}. \quad (3.23)$$

由于 $\tilde{x}_3^3 - \tilde{x}_1^2 = -\tilde{x}_3(\tilde{x}_1\tilde{x}_2 - \tilde{x}_3^2) + \tilde{x}_1(\tilde{x}_2\tilde{x}_3 - \tilde{x}_1)$, 可得 $\tilde{I} = \sqrt{\tilde{I}}$.

选择项序 $\tilde{x}_1 \prec_{\text{tdeg}} \tilde{x}_2 \prec_{\text{tdeg}} \tilde{x}_3$ 和 $\tau = 10^{-8}$, $M_{t-\ell}(y)$ 的秩与余秩序列在表3.13和表3.14中表示.

表 3.13: $M_{t-\ell}(y)$ 的秩				表 3.14: $M_{t-\ell}(y)$ 的余秩			
阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$	阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$
t=4	13	10	7	t=4	22	10	3
t=5	16	13	10	t=5	40	22	10
t=6	19	16	13	t=6	65	40	22

表 3.15: $\ker M_{t-2}(y)$ 中多项式的约化基的类 α_j

阶数	α_1	α_2	α_3	$\sum_{j=1}^3 j\alpha_j$
t=4	0	2	1	$1 \times 0 + 2 \times 2 + 3 \times 1 = 7$
t=5	3	3	1	$1 \times 3 + 2 \times 3 + 3 \times 1 = 12$
t=6	7	4	1	$1 \times 7 + 2 \times 4 + 3 \times 1 = 18$

对 $t = 4$, 可得

$$\sum_{j=1}^3 j\alpha_j = 7, \text{ 并且 corank } M_{4-1} - \text{corank } M_{4-2} = 7.$$

因此, 条件(3.6)成立. 当 $t = 4$ 时, 由算法3.1计算的Pommaret基为

$$\{\tilde{x}_1\tilde{x}_2 - \tilde{x}_3^2, \tilde{x}_2\tilde{x}_3 - \tilde{x}_1, \tilde{x}_2^2 - \tilde{x}_3\}. \quad (3.24)$$

由(3.23)可知(3.24)是 $\sqrt{\tilde{I}}$ 的Pommaret基.

例 3.6. 考虑理想 $I = \langle h_1, h_2 \rangle$,

$$\begin{aligned} h_1 &= (x_1 - x_2)(x_1 + x_2)^2(x_1 + x_2^2 + x_2), \\ h_2 &= (x_1 - x_2)(x_1 + x_2)^2(x_1^2 + x_2^2). \end{aligned}$$

使用函数realrad返回 $\sqrt[3]{I}$ 的生成元为

$$\{x_1^2 - x_2^2\}. \quad (3.25)$$

由于 $x_1^2 - x_2^2 \notin I$, I 严格包含在 $\sqrt[3]{I}$ 中, 即 $I \subset \sqrt[3]{I}$.

令 $\tau = 10^{-4}$ 和 $x_1 \prec_{\text{tdeg}} x_2$, 截断矩量矩阵 $M_{t-\ell}(y)$ 的秩与余秩序列在表3.16和表3.17中表示.

表 3.16: $M_{t-\ell}(y)$ 的秩				表 3.17: $M_{t-\ell}(y)$ 的余秩			
阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$	阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$
t=7	15	13	11	t=7	21	15	10
t=8	17	15	13	t=8	28	21	15
t=9	19	17	15	t=9	36	28	21

表 3.18: $\ker M_{t-2}(y)$ 中多项式的约化基的类 α_j

阶数	α_1	α_2	$\sum_{j=1}^2 j\alpha_j$
t=7	3	1	$1 \times 3 + 2 \times 1 = 5$
t=8	4	1	$1 \times 4 + 2 \times 1 = 6$
t=9	5	1	$1 \times 5 + 2 \times 1 = 7$

对 $t = 7$, 可得

$$\sum_{j=1}^2 j\alpha_j = 5, \text{ 并且 corank } M_{7-1} - \text{corank } M_{7-2} = 5.$$

因此, 条件(3.6)成立. 当 $t = 7$ 时, 算法3.1返回的Pommaret基是

$$\{-x_1^2 + x_2^2\}. \quad (3.26)$$

对比(3.25)和(3.26), 有 $\langle \ker M_{7-2}(y) \rangle = \sqrt[3]{I}$ 并且约化基(3.26)是 $\sqrt[3]{I}$ 的Pommaret基.

在这个例子中, 如果容忍度 $\tau < 10^{-4}$, 截断矩量矩阵 $M_{t-\ell}(y)$ 的秩和余秩序列与表3.16和表3.17不同. 我们无法得到 $\sqrt[3]{I}$ 的Pommaret基 $\{-x_1^2 + x_2^2\}$.

例 3.7. 我们计算 $I = \langle h_1, h_2 \rangle$ 的 \mathcal{A} -根理想 $\mathrm{I}(\mathrm{V}_{\mathbb{R}}(I) \cap \mathcal{A})$,

$$h_1 = (x_1 - x_2)(x_1 + x_2)(x_1 + x_2^2 + x_2),$$

$$h_2 = (x_1 - x_2)(x_1 + x_2)(x_1^2 + x_2^2),$$

并且

$$\mathcal{A} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq 1, x_2 \geq 1\}.$$

令 $\tau = 10^{-8}$ 并且 $x_1 \prec_{\mathrm{tdeg}} x_2$. 截断矩量矩阵 $M_{t-\ell}(y)$ 的秩与余秩序列在表3.19和表3.20中表示.

表 3.19: $M_{t-\ell}(y)$ 的秩			
阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$
t=6	8	6	5
t=7	9	7	6
t=8	10	8	7

表 3.20: $M_{t-\ell}(y)$ 的余秩			
阶数	$\ell = 0$	$\ell = 1$	$\ell = 2$
t=6	20	15	10
t=7	27	21	15
t=8	35	28	21

表 3.21: $\ker M_{t-2}(y)$ 中多项式的约化基的类 α_j

阶数	α_1	α_2	$\sum_{j=1}^2 j\alpha_j$
t=6	3	1	$1 \times 3 + 2 \times 1 = 5$
t=7	4	1	$1 \times 4 + 2 \times 1 = 6$
t=8	5	1	$1 \times 5 + 2 \times 1 = 7$

对 $t = 6$, 可得

$$\sum_{j=1}^2 j\alpha_j = 5, \text{ 并且 } \mathrm{corank} M_{6-1} - \mathrm{corank} M_{6-2} = 5.$$

因此, 条件(3.6)成立. 当 $t = 6$ 时, 算法3.1返回的Pommaret基为

$$\{-x_1 + x_2\}.$$

3.5 总结评论

在这一章, 我们介绍一个返回介于 I 和 $\text{I}(\text{V}_{\mathbb{R}}(I))$ 之间的理想 J 的 Pommaret 基的半正定规划方法. 在这个方法中, J 可以由截断矩量矩阵核中的多项式生成. 我们的方法基于文献 [48, 49, 54, 74, 78, 87, 88, 90] 的工作. 通过对合理论和半正定矩量矩阵的性质, 我们在半正定规划方法(3.8)中设计一个新的终止条件(3.6). 我们还证明该算法在 δ -正则坐标系下的有限终止性.

注 3.5. 在零维情形下, 如果平坦扩张对阶数 $s \leq t$ 成立(对 $y \in \mathcal{K}_t^{gen}$, $\text{rank } M_s(y) = \text{rank } M_{s-1}(y)$), 则 $\langle \ker M_s(y) \rangle = \text{I}(\text{V}_{\mathbb{R}}(I))$ 并且 $\text{corank } M_s(y) - \text{corank } M_{s-1}(y) = \binom{n+s-1}{s}$. 因此, $\ker M_s(y)$ 中多项式的首项集合是所有 s 次单项式的集合. 因此可得 $\sum_{j=1}^n j \alpha_j = \binom{n+s}{s+1}$.

如果 $s < t - 1$, 通过命题3.2(i), 可得 $\text{rank } M_s(y) = \text{rank } M_{s+1}(y)$, 即

$$\text{corank } M_{s+1}(y) - \text{corank } M_s(y) = \binom{n+s}{s+1}.$$

因此, 在 δ -正则坐标系下, $\ker M_s(y)$ 满足条件(3.6)并且 $\ker M_s(y)$ 中多项式的约化基是理想 $\langle \ker M_s(y) \rangle$ 关于单项式序 \prec_{tddeg} 的一组弱 Pommaret 基.

如果 $s = t - 1$, 通过命题3.2(i)和命题3.3(ii), 可以直接说明对 $y' \in \mathcal{K}_{t+1}^{gen}$, $M_s(y')$ 满足平坦扩张条件, 即 $\text{rank } M_s(y') = \text{rank } M_{s-1}(y')$. 因此, $\text{rank } M_s(y') = \text{rank } M_{s+1}(y')$, $M_s(y')$ 满足条件(3.6)并且 $J = \langle \ker M_s(y') \rangle = \text{I}(\text{V}_{\mathbb{R}}(I))$. 如果 $s = t$, 类似地, 可以说明对 $y'' \in \mathcal{K}_{t+2}^{gen}$, $M_s(y'')$ 满足条件(3.6)并且 $J = \langle \ker M_s(y'') \rangle = \text{I}(\text{V}_{\mathbb{R}}(I))$.

注 3.6. 从第3.4节的表中, 可以验证只要第 t 阶截断矩量矩阵满足条件(3.6), 更高阶的截断矩量矩阵也满足条件(3.6). 此外, 可以验证 $y_1 \in \mathcal{K}_{t'}^{gen}$ 和 $y_2 \in \mathcal{K}_{t'+1}^{gen}$ 满足条件

$$\text{rank } M_{t'-\ell}(y_1) = \text{rank } M_{(t'+1)-(\ell+1)}(y_2), \quad t' \geq t, \ell = 1, 2. \quad (3.27)$$

然而对一般的多项式系统, 我们不能确保这个性质的正确性. 显然地, 如果(3.6)和(3.27)对更高阶的矩量矩阵均成立, 则对任意的 $k \geq 0$ 和任意的母向量 $y \in \mathcal{K}_{t+k}^{gen}$, 可得

$$\text{rank } M_{t+k-\ell}(y) = HP_{\sqrt[|I|]{I}}^{\text{aff}}(t+k-\ell), \quad (3.28)$$

这里 $HP_{\sqrt[|I|]{I}}^{\text{aff}}(t+k-\ell)$ 是 $\sqrt[|I|]{I}$ 的仿射希尔伯特多项式. 仿射希尔伯特多项式可以确定不属于 $\sqrt[|I|]{I}_{\leq t+k-\ell}$ 的多项式 $\mathbb{K}[x]_{\leq t+k-\ell}$ 的维数, 参见 [57, 定理 5.20]. 此

外, $\ker M_{t-\ell}(y)$ 中多项式的约化基是 $J = \langle \ker M_{t-\ell}(y) \rangle = I(V_{\mathbb{R}}(I))$ 在单项式序 \prec_{tdeg} 下的弱Pommaret基, 参见 [57, 定理 5.21].

注 3.7. 令 $I = \langle h_1, \dots, h_m \rangle$ 并且 $\dim(I) = s > 0$. 对 $\sqrt[{\mathbb{R}}]{I}$ 中任意的多项式 f , 存在 $t, r \in \mathbb{N}$ 和多项式 $q_1, \dots, q_r, p_1, \dots, p_m \in \mathbb{R}[x]$, 使得

$$f^{2r} + \sum_{i=1}^r q_i^2 = \sum_{i=1}^m p_i h_i. \quad (3.29)$$

对整数 $i \in [1, m]$, $\deg(p_i h_i)$ 的界 $D(n, \deg(f), s)$ 在 [56, 84, 85] 中提出. 由于文献 [62] 提出 $\sqrt[{\mathbb{R}}]{I}$ 的次数上界 k_1 , 可得 $\deg(p_i h_i) \leq k_2 = D(n, k_1, s)$. 通过 (3.29) 和命题 3.2(ii), $\sqrt[{\mathbb{R}}]{I}$ 所有的生成元包含在 $\ker M_{k_2}(y)$ 中. 因此, 如果条件 (3.6) 对 $t - 2 > k_2$ 满足, 则 $\ker M_{t-2}(y)$ 中多项式的一组约化基是 $J = \langle \ker M_{t-2}(y) \rangle = I(V_{\mathbb{R}}(I))$ 关于单项式序 \prec_{tdeg} 的一组弱Pommaret基.

一般情形下, 估计 $\sqrt[{\mathbb{R}}]{I}$ 的次数上界 k_1 是很困难的. 由于齐次理想满足 q 阶 Castelnuovo-Mumford 正则性当且仅当选择适当的坐标系时, 齐次理想存在关于单项式序 \prec_{tdeg} 的 q 次 Pommaret 基, 参见 [89, 定理 5.5.15]. 因此, 理想 $J = \langle \ker M_{t-2}(y) \rangle$ 的 Pommaret 基在某些条件(例如, [73]) 下可以估计 $I(V_{\mathbb{R}}(I))$ 的次数上界.

条件 (3.6) 只有在关于某个容忍度下可以被验证. 如果不能选取合适的容忍度, 我们无法确保给出一个有意义的答案.

第四章 凸代数几何

4.1 问题背景

在上一章, 我们利用基于矩量矩阵的半正定规划方法来求解一个理想的实根理想. 在这一章, 我们研究如何利用谱多面体的投影来近似一个半代数集的闭凸包. 计算半代数集的闭凸包或者闭凸包的近似具有广泛的应用. 例如, 它将一个非凸优化问题转换成一个凸优化问题, 从而降低计算复杂度. 凸优化的另一个好处是它的任何一个局部最优值都是全局最优值, 在求解时不存在陷入局部极小值的问题.

给定多项式集合 $G := \{g_1(x), \dots, g_m(x)\}$ 和基本闭半代数集 S :

$$S := \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \dots, g_m(x) \geq 0\}, \quad (4.1)$$

这里, $g_i(X) \in \mathbb{R}[X_1, \dots, X_n]$, $i = 1, \dots, m$. 我们考虑下面的问题:

问题2. 如何利用谱多面体的投影来近似 S 的闭凸包?

令 \mathcal{M}_{2k} 为第 $2k$ 阶矩量序列的集合. 对任意的 $y \in \mathcal{M}_{2k}$, y 在 $\mathbb{R}[X]_{2k}$ 中的里斯函数 \mathcal{L}_y 定义为

$$\text{对任意的 } q(X) \in \mathbb{R}[X]_{2k}, \quad \mathcal{L}_y \left(\sum_{\alpha} q_{\alpha} X_1^{\alpha_1} \cdots X_n^{\alpha_n} \right) := \sum_{\alpha} q_{\alpha} y_{\alpha}.$$

利用局部矩量矩阵 $M_{k-d_p}(py)$ 的结构, 可得

$$\text{对任意的 } q(X) \in \mathbb{R}[X]_{k-d_p}, \quad \text{vec}(q)^T M_{k-d_p}(py) \text{vec}(q) = \mathcal{L}_y(p(X)q(X)^2). \quad (4.2)$$

给定 \mathbb{R}^n 中的有界基本闭半代数集 S , Lasserre 在文献 [47] 中提出了 $\mathbf{co}(S)$ 的半正定近似表示序列. 令 $s(k) := \binom{n+k}{n}$ 并且对 $j = 1, \dots, m$, $k_j := \lceil \deg g_j / 2 \rceil$. 对任意的 $k \in \mathbb{N}$, 定义

$$\Omega_k(G) := \left\{ x \in \mathbb{R}^n \left| \begin{array}{l} \text{存在 } y \in \mathbb{R}^{s(2k)}, \text{ 使得 } \mathcal{L}_y(1) = 1, \\ \mathcal{L}_y(X_i) = x_i, \quad i = 1, \dots, n, \quad M_k(y) \succeq 0, \\ M_{k-k_j}(g_j y) \succeq 0, \quad j = 1, \dots, m, \end{array} \right. \right\}. \quad (4.3)$$

在文献 [47] 中, 有下面的结果:

定理 4.1. [47, 定理 6] 给定形如(4.1)的半代数集 S . 如果 G 的第 k 阶二次模 $\mathcal{Q}_k(g_1, \dots, g_m)$ 满足阿基米德条件, 则对任意的 $\epsilon > 0$, 存在正整数 N , 使得当 $k \geq N$ 时, 有 $\text{co}(S) \subseteq \Omega_k(G) \subseteq \text{co}(S) + \epsilon B$, 其中 B 是单位球.

与 $\{\Omega_k(G)\}$ 联系紧密的另一个表示序列是定义在实代数集上的Theta体, 参见 [30, 33]. Theta体的定义可以推广到半代数集上. 令 S 为一个基本闭半代数集并且 $\mathbb{R}[X]_1$ 为 $\mathbb{R}[X]$ 中所有的线性多项式的集合, 则

$$\text{cl}(\text{co}(S)) = \bigcap_{p \in \mathbb{R}[X]_1, p|_S \geq 0} \{x \in \mathbb{R}^n \mid p(x) \geq 0\}. \quad (4.4)$$

G 的第 k 阶Theta体将表达式(4.4)松弛为

$$\text{TH}_k(G) := \{x \in \mathbb{R}^n \mid \text{对任意的 } p \in \mathcal{Q}_k(G) \cap \mathbb{R}[X]_1, p(x) \geq 0\}. \quad (4.5)$$

显然可得

$$\text{TH}_1(G) \supseteq \text{TH}_2(G) \supseteq \cdots \supseteq \text{TH}_k(G) \supseteq \text{TH}_{k+1}(G) \supseteq \cdots \supseteq \text{cl}(\text{co}(S)).$$

当 $\mathcal{Q}(G)$ 满足阿基米德条件时, 通过Putinar正零点定理, 有

$$\text{cl}(\text{co}(S)) = \bigcap_{k=1}^{\infty} \text{TH}_k(G). \quad (4.6)$$

定理 4.2. 如果二次模 $\mathcal{Q}_k(G)$ 是闭的, 则 $\text{TH}_k(G) = \text{cl}(\Omega_k(G))$.

证明. 证明和 [30, 定理 2.8] 中实代数集的情形相同. \square

注 4.1. 当 S 存在非空内点时, 对任意的 $k \in \mathbb{N}$, 二次模 $\mathcal{Q}_k(G)$ 是闭的, 参见 [68, 86]. 因此对任意的 $k \in \mathbb{N}$, $\text{TH}_k(G) = \text{cl}(\Omega_k(G))$.

阿基米德条件对半正定表示序列(4.3)和Theta体序列(4.5)的收敛性具有至关重要的作用. 然而, 对无界的半代数集 S , 阿基米德条件不满足. 我们不能确保(4.3)或者(4.5)收敛到 $\text{cl}(\text{co}(S))$. 在下面的例子可以证实这一点:

例 4.1. 考虑基本半代数集

$$S := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq 0, x_1^2 - x_2^3 \geq 0\}. \quad (4.7)$$

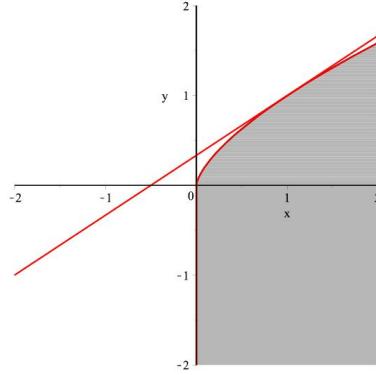
图 4.1: 例子4.1中的半代数集 S 和切线 l .

图4.1中, S 由曲线右半部分的灰色阴影区域表示. 令 $G := \{X_1, X_1^2 - X_2^3\}$. 显然有 $\text{cl}(\text{co}(S)) = S$. 红色的线表示 S 在 $(1, 1)$ 处的切线 $l(X_1, X_2) := 1 + 2X_1 - 3X_2 = 0$. 对任意的 $c_1X_1 + c_2X_2 + c_0 \in \mathcal{Q}_k(G) \cap \mathbb{R}[X]_1, c_0, c_1, c_2 \in \mathbb{R}$, 有

$$c_1X_1 + c_2X_2 + c_0 = \sigma_0(X_1, X_2) + \sigma_1(X_1, X_2)X_1 + \sigma_2(X_1, X_2)(X_1^2 - X_2^3),$$

这里, $\sigma_0, \sigma_1, \sigma_2 \in \Sigma^2$. 令 $X_1 = 0$, 可得

$$c_2X_2 + c_0 = \sigma_0(0, X_2) - X_2^3\sigma_2(0, X_2).$$

$\sigma_0(0, X_2)$ 和 $-X_2^3\sigma_2(0, X_2)$ 中的最高次项不能相互消去. 因此, $\sigma_2(0, X_2) = 0$ 并且 $\sigma_0(0, X_2)$ 是一个常数. 因此 $c_2 = 0$ 并且对任意的 $k \in \mathbb{N}$, 有

$$\text{TH}_k(G) = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq 0\}.$$

因此, (4.5)中的 $\text{TH}_k(G)$ 序列不收敛到 $\text{cl}(\text{co}(S))$. 此外, 由于 S 存在非空内点, 根据注4.1, 对任意 $k \in \mathbb{N}$, $\mathcal{Q}_k(G)$ 是闭的. 通过定理4.2, 对任意的 $k \in \mathbb{N}$, $\text{TH}_k(G) = \text{cl}(\Omega_k(G))$. 因此, (4.3)中的Lasserre松弛序列 $\Omega_k(G)$ 不收敛到 $\text{cl}(\text{co}(S))$, 并且 $\text{TH}_k(G)$ 不收敛到 $\text{co}(S)$. 这是因为除了 $X_1 = 0$, 对任意的 $k \in \mathbb{N}$, S 的任何切线不能被 $\mathcal{Q}_k(G) \cap \mathbb{R}[X]_1$ 对应的直线近似. 特别地, 对任意的 $\epsilon > 0$, $k \in \mathbb{N}$, 有 $l_\epsilon := l + \epsilon \notin \mathcal{Q}_k(G)$.

注 4.2. 由于半代数集和多面体的投影都是无界的, 它们只能在图中截断表示.

令 $\tilde{g}_1, \dots, \tilde{g}_m$ 分别为 g_1, \dots, g_m 经过齐次化后的多项式. 我们将 S 映射成高维空间的锥 $\tilde{S}^\circ \subseteq \mathbb{R}^{n+1}$:

$$\tilde{S}^\circ := \{\tilde{x} \in \mathbb{R}^{n+1} \mid \tilde{g}_1(\tilde{x}) \geq 0, \dots, \tilde{g}_m(\tilde{x}) \geq 0, x_0 > 0\}.$$

令 $\tilde{X} := (X_0, X_1, \dots, X_n)$. $\mathcal{Q}_k(\tilde{G})$ 表示由下面集合中的多项式生成的第 k 阶二次模:

$$\tilde{G} := \left\{ \tilde{g}_1, \dots, \tilde{g}_m, X_0, \|\tilde{X}\|_2^2 - 1, 1 - \|\tilde{X}\|_2^2 \right\}.$$

此外,

$$\tilde{S} := \left\{ \tilde{x} \in \mathbb{R}^{n+1} \mid \tilde{g}_1(\tilde{x}) \geq 0, \dots, \tilde{g}_m(\tilde{x}) \geq 0, x_0 \geq 0, \|\tilde{x}\|_2^2 = 1 \right\}.$$

将 $\mathbb{P}[\tilde{X}]_1$ 记作 $(\mathbb{R}[\tilde{X}]_1 \setminus \mathbb{R}) \cup \{0\}$, 其中 $\mathbb{R}[\tilde{X}]_1$ 是 $\mathbb{R}[\tilde{X}]$ 中线性多项式的集合. Theta 体序列 $\widetilde{\text{TH}}_k(\tilde{G})$ 定义如下:

$$\widetilde{\text{TH}}_k(\tilde{G}) := \left\{ x \in \mathbb{R}^n \mid \text{对任意的 } \tilde{l} \in \mathcal{Q}_k(\tilde{G}) \cap \mathbb{P}[\tilde{X}]_1, \tilde{l}(1, x) \geq 0 \right\},$$

和 Lasserre 松弛 $\tilde{\Omega}_k(\tilde{G})$:

$$\tilde{\Omega}_k(\tilde{G}) := \left\{ x \in \mathbb{R}^n \mid \begin{array}{l} \text{存在 } y \in \mathbb{R}^{\tilde{s}(2k)}, \text{ 使得 } \mathcal{L}_y(X_0) = 1, \\ \mathcal{L}_y(X_i) = x_i, i = 1, \dots, n, \\ M_{k-1}(X_0 y) \succeq 0, M_{k-1}((\|\tilde{X}\|_2^2 - 1)y) = 0, \\ M_k(y) \succeq 0, M_{k-k_j}(\tilde{g}_j y) \succeq 0, j = 1, \dots, m, \end{array} \right\},$$

这里, 对任意的 $k \in \mathbb{N}$, 有 $\tilde{s}(k) = \binom{n+k+1}{n+1}$ 并且 $k_j = \lceil \deg g_j / 2 \rceil$.

给定一个无界基本闭半代数集 S .

- 假设 S 在无穷远处是闭的(参见 [63])并且凸锥 $\mathbf{co}(\mathbf{cl}(\tilde{S}^\circ))$ 是尖的闭凸锥.
 - 首先, 我们证明 $\widetilde{\text{TH}}_k(\tilde{G})$ 序列和 $\mathbf{cl}(\tilde{\Omega}_k(\tilde{G}))$ 序列收敛到 $\mathbf{cl}(\mathbf{co}(S))$. 如果 $\mathcal{Q}_k(\tilde{G})$ 是闭的, 则对 $k \in \mathbb{N}$, 有 $\widetilde{\text{TH}}_k(\tilde{G}) = \mathbf{cl}(\tilde{\Omega}_k(\tilde{G}))$.
 - 如果 \tilde{G} 满足第 k 阶 PP-BDR 条件, $\mathbf{cl}(\mathbf{co}(S)) = \widetilde{\text{TH}}_{k'}(\tilde{G}) = \mathbf{cl}(\tilde{\Omega}_{k'}(\tilde{G}))$. 如果 PP-BDR 条件不成立, 可以证明对 $\epsilon > 0$, $\mathbf{cl}(\tilde{\Omega}_k(\tilde{G}))$ 在以原点为圆心的任意的球上一致收敛到 $\mathbf{cl}(\mathbf{co}(S))$.
- 我们指出闭性和尖性对于谱多面体序列 $\widetilde{\text{TH}}_k(\tilde{G})$ 的收敛性是必要的.
 - 我们指出 S 在无穷远处是闭的条件与 S 的生成元有关. 在很多情况下, 我们可以在 S 的定义多项式中添加一些线性多项式使 S 在无穷远处是闭的条件成立.

- 如果 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 不是尖的闭凸锥, 我们将 S 通过坐标轴分成 2^n 个部分. 如果 S 在无穷远处是闭的, 并且每一个部分均满足 PP-BDR 条件. 我们可以对每一部分计算推广 Lasserre 半正定表示序列并且将它们粘在一起.

4.2 无界凸集的半正定表示

当基本闭半代数集 S 无界时, 我们构造了推广 Theta 体序列和 Lasserre 半正定表示序列来近似 $\text{cl}(\text{co}(S))$. 主要的思想是通过齐次化, 将 S 映射成 \mathbb{R}^{n+1} 中的锥 \tilde{S}° . 我们引用文献 [23, 55] 中的技术来处理无界半代数集的半正定近似. 当 S 在无穷远处是闭的并且 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 是尖的闭凸锥时. 推广 Theta 体序列和推广 Lasserre 半正定表示序列收敛到 $\text{cl}(\text{co}(S))$. 一些例子可以说明闭性和尖性对于收敛性必要的.

4.2.1 近似 $\text{cl}(\text{co}(S))$ 的闭凸包

考虑多项式 $f(X) \in \mathbb{R}[X]$ 和它的齐次化多项式 $\tilde{f}(\tilde{X}) \in \mathbb{R}[\tilde{X}]$, 这里 $\tilde{X} = (X_0, X_1, \dots, X_n)$ 并且 $\tilde{f}(\tilde{X}) = X_0^d f(X/X_0)$, $d = \deg(f)$. 对给定的半代数集

$$S := \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \dots, g_m(x) \geq 0\}, \quad (4.8)$$

定义

$$\begin{aligned} \tilde{S}^\circ &:= \{\tilde{x} \in \mathbb{R}^{n+1} \mid \tilde{g}_1(\tilde{x}) \geq 0, \dots, \tilde{g}_m(\tilde{x}) \geq 0, x_0 > 0\}, \\ \tilde{S}^c &:= \{\tilde{x} \in \mathbb{R}^{n+1} \mid \tilde{g}_1(\tilde{x}) \geq 0, \dots, \tilde{g}_m(\tilde{x}) \geq 0, x_0 \geq 0\}, \\ \tilde{S} &:= \{\tilde{x} \in \mathbb{R}^{n+1} \mid \tilde{g}_1(\tilde{x}) \geq 0, \dots, \tilde{g}_m(\tilde{x}) \geq 0, x_0 \geq 0, \|\tilde{x}\|_2^2 = 1\}. \end{aligned} \quad (4.9)$$

注意到 \tilde{S}° 只依赖于 S , 然而 \tilde{S}^c 和 \tilde{S} 不仅依赖于 S , 而且还依赖于半代数集 S 的表达式的选择.

命题 4.3. [38, 命题 2.1] $f(x)$ 在 S 上非负当且仅当在 $\text{cl}(\tilde{S}^\circ)$ 上, 有 $\tilde{f}(\tilde{x}) \geq 0$.

推论 4.4. 对任意的线性函数 $f \in \mathbb{R}[X]_1$, $f(x)$ 在 $\text{cl}(\text{co}(S))$ 上非负当且仅当在 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 上, 有 $\tilde{f}(\tilde{x}) \geq 0$.

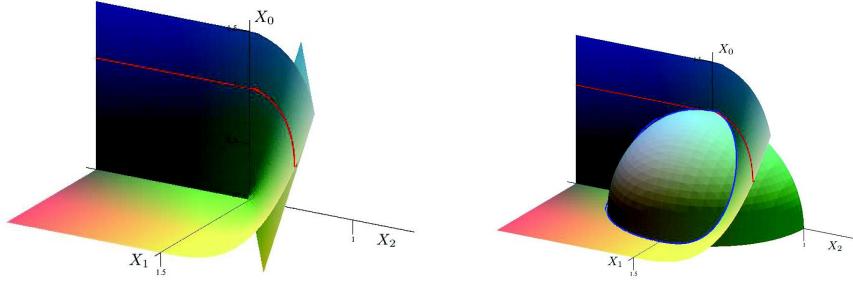


图 4.2: 例子4.1中, 左:由 S 生成的锥 \tilde{S}^c 和由 l 生成的超平面 \tilde{l} ; 右:单位球上的一个有界部分 \tilde{S} (被蓝色的线分割.)

证明. 由于 $f(X)$ 是线性函数, $f(x)$ 在 $\text{cl}(\text{co}(S))$ 上非负当且仅当 $f(x)$ 在 S 上非负. $\tilde{f}(\tilde{x})$ 在 $\text{co}(\text{cl}(\tilde{S}^o))$ 上非负当且仅当 $\tilde{f}(\tilde{x})$ 在 $\text{cl}(\tilde{S}^o)$ 上非负. 根据命题4.3, 结论成立. \square

例 4.2. 在例子4.1中, 通过(4.7)和(4.9), 可得

$$\begin{aligned}\tilde{S}^o &= \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_1 \geq 0, x_0 x_1^2 - x_2^3 \geq 0, x_0 > 0\}, \\ \tilde{S}^c &= \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_1 \geq 0, x_0 x_1^2 - x_2^3 \geq 0, x_0 \geq 0\}, \\ \tilde{S} &= \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_1 \geq 0, x_0 x_1^2 - x_2^3 \geq 0, x_0 \geq 0, \|\tilde{x}\|_2^2 = 1\}.\end{aligned}$$

在图4.2的左边, \mathbb{R}^3 中的锥 \tilde{S}^c 和由 l 对应的超平面 $\tilde{l}(X_0, X_1, X_2) := X_0 + 2X_1 - 3X_2 = 0$ 相切. \tilde{l} 相切于 $\text{co}(\text{cl}(\tilde{S}^o))$ 验证了推论4.4. 相交在单位球上的部分 \tilde{S} 由图4.2的右边表示.

定义 4.1. [63] 如果 $\text{cl}(\tilde{S}^o) = \tilde{S}^c$, S 在无穷远处是闭的.

注 4.3. 文献 [38]指出在无穷远处是闭的条件是一般性的条件. 考虑所有形如(4.8)的基本闭半代数集 S , 其中 n, m 和 g_i 的最高次数 d 是固定的. 则使得 S 在无穷远处是闭的条件不成立的所有 g_i 的系数 $\text{vec}(g_i)$ 在系数空间 $\mathbb{R}[X]_d$ 的一个代数集中.

由于 \tilde{S}^c 和 \tilde{S} 依然依赖于不等式 $g_1(x) \geq 0, \dots, g_m(x) \geq 0$, S 在无穷远处闭的条件不仅依赖于 S , 而且依赖于 S 的定义多项式. 见下面的例子:

例 4.3. 考虑集合 $S := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 - x_1^2 \geq 0\}$. 则

$$\tilde{S}^\circ = \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_0 x_2 - x_1^2 \geq 0, x_0 > 0\},$$

$$\tilde{S}^c = \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_0 x_2 - x_1^2 \geq 0, x_0 \geq 0\},$$

$$\tilde{S} = \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_0 x_2 - x_1^2 \geq 0, x_0 \geq 0, \|\tilde{x}\|_2 = 1\}.$$

由于 $\tilde{S}^c \setminus \text{cl}(\tilde{S}^\circ) = \{(0, 0, x_2) \in \mathbb{R}^3 \mid x_2 < 0\} \neq \emptyset$, S 在无穷远处不是闭的. 然而, 如果令 $S' := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 - x_1^2 \geq 0, 1 + x_2 \geq 0\}$, 很容易验证 $S = S'$. 但是 S' 在无穷远处是闭的. 在一般情形下, 对一些在无穷远处不是闭的半代数集 S , 可以添加一些线性多项式到 S 的定义多项式中, 使得 S 在无穷远处是闭的.

假设2. (i) S 在无穷远处是闭的. (ii) 凸锥 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 是尖的闭凸锥.

令 $\mathbb{P}[\tilde{X}]_1$ 为 $\mathbb{R}[\tilde{X}]$ 中常数项为 0 的线性多项式和零多项式的并, 并且

$$\tilde{G} := \{\tilde{g}_1, \dots, \tilde{g}_m, X_0, \|\tilde{X}\|_2^2 - 1, 1 - \|\tilde{X}\|_2^2\}. \quad (4.10)$$

定义 4.2. 形如(4.8)的无界半代数集 S 的第 k 阶推广 Theta 体定义为

$$\widetilde{\text{TH}}_k(\tilde{G}) := \{x \in \mathbb{R}^n \mid \text{对任意的 } \tilde{l} \in \mathcal{Q}_k(\tilde{G}) \cap \mathbb{P}[\tilde{X}]_1, \tilde{l}(1, x) \geq 0\}. \quad (4.11)$$

定理 4.5. 令 $S \subseteq \mathbb{R}^n$ 为形如(4.8)的半代数集. 如果假设2成立, 则对每一个 $k \in \mathbb{N}$, $\text{cl}(\text{co}(S)) \subseteq \widetilde{\text{TH}}_k(\tilde{G})$ 并且

$$\text{cl}(\text{co}(S)) = \bigcap_{k=1}^{\infty} \widetilde{\text{TH}}_k(\tilde{G}). \quad (4.12)$$

证明. 首先需要说明对任意的 $k \in \mathbb{N}$, $\text{cl}(\text{co}(S)) \subseteq \widetilde{\text{TH}}_k(\tilde{G})$. 固定 $\tilde{l} \in \mathcal{Q}_k(\tilde{G}) \cap \mathbb{P}[\tilde{X}]_1$, 则在 \tilde{S} 上, 有 $\tilde{l}(\tilde{x}) \geq 0$. 由于 \tilde{l} 是齐次线性函数, 在 \tilde{S}^c 上, 可得 $\tilde{l}(\tilde{x}) \geq 0$. 由于 $\tilde{S}^\circ \subseteq \tilde{S}^c$, 在 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 上, 可得 $\tilde{l}(\tilde{x}) \geq 0$. 通过推论4.4, 在 $\text{cl}(\text{co}(S))$ 上, 可得 $\tilde{l}(1, x) = l(x) \geq 0$. 因此对每一个 $k \in \mathbb{N}$, $\text{cl}(\text{co}(S)) \subseteq \widetilde{\text{TH}}_k(\tilde{G})$. 因此, (4.11) 中的推广 Theta 体序列构成了一个 $\text{co}(S)$ 的闭凸近似序列:

$$\widetilde{\text{TH}}_1(\tilde{G}) \supseteq \widetilde{\text{TH}}_2(\tilde{G}) \supseteq \cdots \supseteq \widetilde{\text{TH}}_k(\tilde{G}) \supseteq \widetilde{\text{TH}}_{k+1}(\tilde{G}) \supseteq \cdots \supseteq \text{cl}(\text{co}(S)). \quad (4.13)$$

现在证明这个序列收敛到 $\text{cl}(\text{co}(S))$. 假设 $u \notin \text{cl}(\text{co}(S))$, 我们说明对某一个 k , $u \notin \widetilde{\text{TH}}_k(\tilde{G})$. 由于 $\text{cl}(\text{co}(S))$ 是闭凸集. 根据超平面分离定理, 存在向量 $(f_0, \mathbf{f}) \in \mathbb{R}^{n+1}$, 使得

$$\langle \mathbf{f}, u \rangle < f_0, \text{ 并且对任意的 } x \in \text{cl}(\text{co}(S)), \langle \mathbf{f}, x \rangle > f_0.$$

令 $\tilde{f}(\tilde{X}) := \sum_{i=1}^n f_i X_i - f_0 X_0 \in \mathbb{R}[\tilde{X}]$, 则

$$\tilde{f}(1, u) < 0, \text{ 并且对任意的 } x \in \mathbf{cl}(\mathbf{co}(S)), \tilde{f}(1, x) = f(x) > 0.$$

通过推论4.4, 可得

$$\text{对任意的 } x \in \mathbf{co}(\mathbf{cl}(\tilde{S}^\circ)), \tilde{f}(\tilde{x}) \geq 0. \quad (4.14)$$

由于 $\mathbf{co}(\mathbf{cl}(\tilde{S}^\circ))$ 是尖的闭凸锥, 由定理2.3可知存在一个多项式 $\tilde{h}(\tilde{X}) = \sum_{i=0}^n h_i X_i \in \mathbb{P}[\tilde{X}]_1$, 使得在 $\mathbf{co}(\mathbf{cl}(\tilde{S}^\circ)) \setminus \{0\}$ 上, 有 $\tilde{h}(\tilde{x}) > 0$. 存在充分小的 $\epsilon > 0$, 使得 $(\tilde{f} + \epsilon \tilde{h})(1, u) < 0$ 并且将 $\tilde{f} + \epsilon \tilde{h}$ 重新命名为 \tilde{f} , 则

$$\tilde{f}(1, u) < 0, \text{ 并且对任意的 } x \in \mathbf{cl}(\tilde{S}^\circ) \setminus \{0\}, \tilde{f}(\tilde{x}) > 0. \quad (4.15)$$

如果 S 在无穷远处是闭的, 则 $\mathbf{cl}(\tilde{S}^\circ) \cap \{\tilde{x} \in \mathbb{R}^{n+1} \mid \|\tilde{x}\|_2 = 1\} = \tilde{S}$, 因此

$$\tilde{f}(1, u) < 0, \text{ 并且对任意的 } x \in \tilde{S}, \tilde{f}(\tilde{x}) > 0. \quad (4.16)$$

由于 $\mathcal{Q}(\tilde{G})$ 满足阿基米德条件, 通过Putinar正零点定理, 存在 $k' \in \mathbb{N}$, 使得 $\tilde{f} \in \mathcal{Q}_{k'}(\tilde{G}) \cap \mathbb{P}[\tilde{X}]_1$. 由于 $\tilde{f}(1, u) < 0$, 有 $u \notin \widetilde{\text{TH}}_{k'}(\tilde{G})$. 因此

$$\bigcap_{k=1}^{\infty} \widetilde{\text{TH}}_k(\tilde{G}) \subseteq \mathbf{cl}(\mathbf{co}(S)). \quad (4.17)$$

最后, 通过(4.13)和(4.17), 可得 $\mathbf{cl}(\mathbf{co}(S)) = \bigcap_{k=1}^{\infty} \widetilde{\text{TH}}_k(\tilde{G})$. \square

例 4.4. 在例子4.1中, 对每一个 $(0, u_1, u_2) \in \tilde{S}^c \setminus \tilde{S}^\circ$, 令

$$u^{(\epsilon)} := \left(\epsilon, u_1, \sqrt[3]{\epsilon u_1^2 + u_2^3} \right).$$

则 $\{u^{(\epsilon)}\}_{\epsilon>0} \subseteq \tilde{S}^\circ$ 并且 $\lim_{\epsilon \rightarrow 0} u^{(\epsilon)} = (0, u_1, u_2)$. 因此, 有 $\tilde{S}^c \setminus \tilde{S}^\circ \subseteq \mathbf{cl}(\tilde{S}^\circ)$ 并且 S 在无穷远处是闭的. 此外, 可以验证

$$\tilde{g}(X_0, X_1, X_2) := 2X_0 + 2X_1 - 3X_2$$

在 $\mathbf{co}(\mathbf{cl}(\tilde{S}^\circ)) \setminus \{0\}$ 上恒为正. 通过定理2.3, $\mathbf{co}(\mathbf{cl}(\tilde{S}^\circ))$ 是尖的闭凸锥. 因此 S 满足假设2.

令 ϵ 趋近于0, $\tilde{l} + \epsilon\tilde{g}$ 在 \tilde{S} 上恒大于0并且趋近于 \tilde{l} . 此外, 由于 $\mathcal{Q}(\tilde{G})$ 满足阿基米德条件. 通过Putinar正零点定理, 对每一个 $\epsilon > 0$, $\tilde{l} + \epsilon\tilde{g}$ 属于下面多项式所生成的二次模:

$$\tilde{G} := \{X_0, X_1, X_0X_1^2 - X_2^3, X_0^2 + X_1^2 + X_2^2 - 1, 1 - X_0^2 - X_1^2 - X_2^2\}.$$

定义

$$\widetilde{\text{TH}}_k(\tilde{G}) := \{(x_1, x_2) \in \mathbb{R}^2 \mid \text{对任意的 } \tilde{l} \in \mathcal{Q}_k(\tilde{G}) \cap \mathbb{P}[X_0, X_1, X_2]_1, \tilde{l}(1, x_1, x_2) \geq 0\},$$

$$\text{可得 } \text{cl}(\text{co}(S)) = \bigcap_{k=1}^{\infty} \widetilde{\text{TH}}_k(\tilde{G}).$$

推论 4.6. $S \subset \mathbb{R}^n$ 是形如(4.8)的半代数集. 如果假设2满足, 并且 \tilde{G} 的第 k' 阶PP-BDR条件成立, 则 $\text{cl}(\text{co}(S)) = \widetilde{\text{TH}}_{k'}(\tilde{G})$.

证明. 通过(4.13), 需要验证 $\text{cl}(\text{co}(S)) \supseteq \widetilde{\text{TH}}_{k'}(\tilde{G})$. 假设存在 $u \in \widetilde{\text{TH}}_{k'}(\tilde{G})$, 使得 $u \notin \text{cl}(\text{co}(S))$. 根据(4.16), 存在 $\mathbb{R}[\tilde{X}]$ 中的线性多项式 \tilde{f} , 使得 $\tilde{f}(0) = 0$, $\tilde{f}(1, u) < 0$ 并且 $\tilde{f}(\tilde{x})$ 在 \tilde{S} 上恒为正. 由于 \tilde{G} 满足第 k' 阶PP-BDR条件, 可得 $\tilde{f} \in \mathcal{Q}_{k'}(\tilde{G}) \cap \mathbb{P}[\tilde{X}]_1$. 由于 $\tilde{f}(1, u) < 0$, 可得 $u \notin \widetilde{\text{TH}}_{k'}(\tilde{G})$. 矛盾. 因此, $\text{cl}(\text{co}(S)) = \widetilde{\text{TH}}_{k'}(\tilde{G})$. \square

注 4.4. 定义透视投影 φ 为

$$\begin{aligned} \varphi : \{(x_0, x) \in \mathbb{R}^{n+1} \mid x_0 > 0\} &\rightarrow \mathbb{R}^n \\ (x_0, x) &\mapsto \frac{x}{x_0}. \end{aligned} \tag{4.18}$$

令

$$H := \{(0, x) \mid x \in \mathbb{R}^n\}. \tag{4.19}$$

为了近似无界集 S 的闭凸包 $\text{cl}(\text{co}(S))$, 考虑有界集 \tilde{S} 的经典Theta体的透视射影:

$$\begin{aligned} \varphi \left(\text{TH}_k(\tilde{G}) \setminus H \right) &= \{x \in \mathbb{R}^n \mid \text{存在 } t > 0, (t, tx) \in \text{TH}_k(\tilde{G})\} \\ &= \{x \in \mathbb{R}^n \mid \text{存在 } t > 0, \text{ 对任意的 } \tilde{l} \in \mathcal{Q}_k(\tilde{G}) \cap \mathbb{P}[\tilde{X}]_1, \tilde{l}(t, tx) \geq 0\}. \end{aligned} \tag{4.20}$$

显然地,

$$\text{对每一个 } k \in \mathbb{N}, \varphi \left(\text{TH}_k(\tilde{G}) \setminus H \right) \subseteq \widetilde{\text{TH}}_k(\tilde{G}). \tag{4.21}$$

我们在推论4.13中指出当 $\mathcal{Q}_k(\tilde{G})$ 是闭的时, 这两个集合相等.

假设²对定理4.5的结论是必要的.

例 4.5. 我们说明在例子4.3中, S 在无穷远处闭的条件不成立. 显然有 $\text{cl}(\text{co}(S)) = S$ 并且 $\text{cl}(\tilde{S}^\circ)$ 是凸集. 定义 $\tilde{f}(\tilde{X}) := X_0 + X_2$. 由于在 $\text{co}(\text{cl}(\tilde{S}^\circ)) \setminus \{0\}$ 上, $\tilde{f}(\tilde{x}) > 0$, 可以推出 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 是尖的闭凸锥. 因此可得 $(0, 0, -1), (1, 0, 0) \in \tilde{S}$ 并且对任意的 $\lambda \in [0, 1]$, 有 $(1 - \lambda, 0, -\lambda) \in \text{co}(\tilde{S}) \subseteq \text{TH}_k(\tilde{G})$. 通过(4.21), 可得

$$\text{对任意的 } \lambda \in [0, 1], \left(0, -\frac{\lambda}{1-\lambda}\right) \in \varphi(\text{TH}_k(\tilde{G}) \setminus H) \subseteq \widetilde{\text{TH}}_k(\tilde{G}).$$

因此, 对任意的 $k \in \mathbb{N}$, 可得 $\{(0, x_2) \in \mathbb{R}^2 \mid x_2 \leq 0\} \subseteq \widetilde{\text{TH}}_k(\tilde{G})$. 由于 $\text{cl}(\text{co}(S)) = S$, 有 $\bigcap_{k=1}^{\infty} \widetilde{\text{TH}}_k(\tilde{G}) \neq \text{cl}(\text{co}(S))$.

例 4.6. 考虑集合 $S := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2^3 - x_1^2 \geq 0\}$. 显然可得 $\text{cl}(\text{co}(S)) = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \geq 0\}$ 并且

$$\begin{aligned} \tilde{S}^\circ &= \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_2^3 - x_0 x_1^2 \geq 0, x_0 > 0\}, \\ \tilde{S}^c &= \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_2^3 - x_0 x_1^2 \geq 0, x_0 \geq 0\}. \end{aligned}$$

可以证明 $\tilde{S}^c \setminus \tilde{S}^\circ = \{(0, x_1, x_2) \in \mathbb{R}^3 \mid x_2 \geq 0\}$. 在例子4.1中使用相似的证明方法, 可得 S 在无穷远处是闭的. 因此, 由于 $\lim_{\epsilon \rightarrow 0} (\epsilon, \pm 1, \sqrt[3]{\epsilon}) = (0, \pm 1, 0)$ 并且 $(0, \pm 1, 0) \in \text{cl}(\tilde{S}^\circ)$, 可得 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 不是尖的闭凸锥. 令

$$\tilde{G} = \{X_0, X_2^3 - X_0 X_1^2, X_0^2 + X_1^2 + X_2^2 - 1, 1 - X_0^2 - X_1^2 - X_2^2\}.$$

对于每一个 $k \in \mathbb{N}$, $\widetilde{\text{TH}}_k(\tilde{G}) = \mathbb{R}^2$. 假设 $c_0 X_0 + c_1 X_1 + c_2 X_2 \in \mathcal{Q}_k(\tilde{G})$, 则

$$c_0 X_0 + c_1 X_1 + c_2 X_2 = \tilde{\sigma}_0 + \tilde{\sigma}_1 X_0 + \tilde{\sigma}_2 (X_2^3 - X_0 X_1^2) + \tilde{h}(X_0^2 + X_1^2 + X_2^2 - 1), \quad (4.22)$$

这里对 $i = 0, 1, 2$, $\tilde{\sigma}_i \in \Sigma^2$ 并且 $\tilde{h} \in \mathbb{R}[\tilde{X}]$. 将(4.22)中的变量 (X_0, X_1, X_2) 赋值为 $(0, \pm 1, 0)$, 可得 $c_1 = 0$. 将(4.22)中的变量 (X_0, X_1, X_2) 赋值为 $(0, \pm 1, X_2)$, 可得

$$c_2 X_2 = \tilde{\sigma}_0(0, \pm 1, X_2) + \tilde{\sigma}_2(0, \pm 1, X_2) X_2^3 + \tilde{h}(0, \pm 1, X_2) X_2^2. \quad (4.23)$$

因此 $\tilde{\sigma}_0(0, \pm 1, X_2)$ 的常数项为零, 从而等式(4.23)的右边可以被 X_2^2 整除. 因此 $c_2 = 0$. 通过(4.11), 对每一个 $k \in \mathbb{N}$, 有 $\widetilde{\text{TH}}_k(\tilde{G}) = \mathbb{R}^2$. 这说明 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 是尖的闭凸锥的条件不能在定理4.5中被去掉.

4.2.2 $\text{cl}(\text{co}(S))$ 的谱多面体近似

现在, 我们从对偶的角度来研究推广Theta体 $\widetilde{\text{TH}}_k(\tilde{G})$. 考虑以 $(n+1)$ 元数组 $\alpha := (\alpha_0, \alpha_1, \dots, \alpha_n) \in \mathbb{N}^{n+1}$ 为索引的矩量序列 y . 每一个矩量序列 y , 对应的 $\mathbb{R}[\tilde{X}]$ 中的里斯函数记作 \mathcal{L}_y . 令 $\tilde{s}(k) = \binom{n+k+1}{n+1}$ 和 $k_j = \lceil \deg g_j / 2 \rceil$.

定义 4.3. 形如(4.8)的无界半代数集 S 的第 k 阶推广 Lasserre 半正定表示定义为

$$\widetilde{\Omega}_k(\tilde{G}) := \left\{ x \in \mathbb{R}^n \middle| \begin{array}{l} \text{存在 } y \in \mathbb{R}^{\tilde{s}(2k)}, \text{ 使得 } \mathcal{L}_y(X_0) = 1, \\ \mathcal{L}_y(X_i) = x_i, i = 1, \dots, n, \\ M_{k-1}(X_0 y) \succeq 0, M_{k-1}((\|\tilde{X}\|_2^2 - 1)y) = 0, \\ M_k(y) \succeq 0, M_{k-k_j}(\tilde{g}_j y) \succeq 0, j = 1, \dots, m \end{array} \right\}. \quad (4.24)$$

命题 4.7. 对 $k \in \mathbb{N}$, 推广 Lasserre 半正定表示 $\widetilde{\Omega}_k(\tilde{G})$ 可以通过经典 Lasserre 半正定表示 $\Omega_k(\tilde{G})$ 的透视投影来得到, 即 $\widetilde{\Omega}_k(\tilde{G}) = \varphi(\Omega_k(\tilde{G}) \setminus H)$.

证明. 通过定义 4.3, 可得

$$\Omega_k(\tilde{G}) = \left\{ \tilde{x} \in \mathbb{R}^{n+1} \middle| \begin{array}{l} \text{存在 } y \in \mathbb{R}^{\tilde{s}(2k)}, \text{ 使得 } \mathcal{L}_y(1) = 1, \\ \mathcal{L}_y(X_i) = x_i, i = 0, \dots, n, \\ M_{k-1}(X_0 y) \succeq 0, M_{k-1}((\|\tilde{X}\|_2^2 - 1)y) = 0, \\ M_k(y) \succeq 0, M_{k-k_j}(\tilde{g}_j y) \succeq 0, j = 1, \dots, m \end{array} \right\}. \quad (4.25)$$

对任意的 $w \in \widetilde{\Omega}_k(\tilde{G})$, (4.24) 中存在形如 $(y_0, 1, w, \dots)$ 的向量 $y \in \mathbb{R}^{\tilde{s}(2k)}$. 由于 $M_k(y) \succeq 0$ 并且 $\mathcal{L}_y(X_0) = 1$, 可得 $y_0 > 0$. 因此 y/y_0 满足 (4.25) 中的条件, 即 $w = (w/y_0)/(1/y_0) \in \varphi(\Omega_k(\tilde{G}) \setminus H)$. 对任意的 $w \in \varphi(\Omega_k(\tilde{G}) \setminus H)$, 存在 $t > 0$ 和 $y = (1, t, tw, \dots) \in \mathbb{R}^{\tilde{s}(2k)}$, 使得 y 满足 (4.25) 中的条件. 则 y/t 满足 (4.24) 中的条件. 因此 $w \in \widetilde{\Omega}_k(\tilde{G})$. \square

定理 4.8. 对任意的 $k \in \mathbb{N}$, 可得 $\text{cl}(\text{co}(S)) \subseteq \text{cl}(\widetilde{\Omega}_k(\tilde{G})) \subseteq \widetilde{\text{TH}}_k(\tilde{G})$.

证明. 显然可得对任意的 $x \in S$, 存在 $t > 0$, 使得 $(t, tx) \in \widetilde{S}$. 则对每一个 $k \in \mathbb{N}$, $(t, tx) \in \Omega_k(\tilde{G})$. 通过命题 4.7, 可得 $x \in \varphi(\Omega_k(\tilde{G}) \setminus H) = \widetilde{\Omega}_k(\tilde{G})$. 由于 $\text{cl}(\widetilde{\Omega}_k(\tilde{G}))$ 是闭凸集, 可得 $\text{cl}(\text{co}(S)) \subseteq \text{cl}(\widetilde{\Omega}_k(\tilde{G}))$.

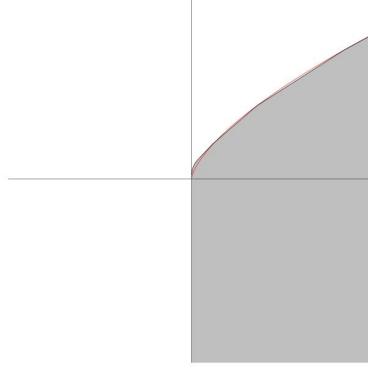


图 4.3: 例子4.1中, $\text{cl}(\text{co}(S))$ 的谱多面体近似 $\tilde{\Omega}_3(\tilde{G})$ (用阴影部分表示)

由于 $\Omega_k(\tilde{G}) \subset \text{TH}_k(\tilde{G})$, 通过(4.21)和命题4.7, 可得

$$\tilde{\Omega}_k(\tilde{G}) = \varphi\left(\Omega_k(\tilde{G}) \setminus H\right) \subseteq \varphi\left(\text{TH}_k(\tilde{G}) \setminus H\right) \subseteq \widetilde{\text{TH}}_k(\tilde{G}).$$

由于 $\widetilde{\text{TH}}_k(\tilde{G})$ 是闭的, 因此结论成立. \square

例 4.7. 对例子4.1使用软件包Bermeja [77], 可以绘制出谱多面体 $\tilde{\Omega}_3(\tilde{G})$. 如图4.3所示, 推广Lasserre半正定表示 $\tilde{\Omega}_3(\tilde{G})$ 近似 $\text{cl}(\text{co}(S))$ 的程度非常好.

推论 4.9. 令 $S \subseteq \mathbb{R}^n$ 为形如(4.8)的半代数集. 如果假设2满足, 则

1. $\text{cl}(\text{co}(S)) = \bigcap_{k=1}^{\infty} \text{cl}\left(\tilde{\Omega}_k(\tilde{G})\right).$
2. 如果 \tilde{G} 的第 k' 阶 PP-BDR 条件成立, 则 $\text{cl}(\text{co}(S)) = \text{cl}\left(\tilde{\Omega}_{k'}(\tilde{G})\right).$

证明. 通过定理4.5, 推论4.6和定理4.8, 很容易证明结论. \square

由于 PP-BDR 条件不经常成立, 与文献 [47, 第2.5节, 定理6] 相似, 我们构造 $\text{cl}(\text{co}(S))$ 的近似表示序列. 对 $r \in \mathbb{R}$, 令 $\mathbf{B}_r := \{x \in \mathbb{R}^n \mid \|x\|_2 \leq r\}$.

定理 4.10. 假设 $S \subseteq \mathbb{R}^n$ 为形如(4.8)的半代数集. 如果假设2成立, 则对任意的 $\epsilon > 0$ 和 $\text{cl}(\text{co}(S)) \cap \mathbf{B}_r \neq \emptyset$, 其中 $r > 0$, 存在 $k_{r,\epsilon} \in \mathbb{N}$, 使得

$$\text{cl}(\text{co}(S)) \cap \mathbf{B}_r \subseteq \text{cl}\left(\tilde{\Omega}_{k_{r,\epsilon}}(\tilde{G})\right) \cap \mathbf{B}_r \subseteq (\text{cl}(\text{co}(S)) + \epsilon \mathbf{B}_1) \cap \mathbf{B}_r.$$

证明. 通过命题4.7, 证明与 [47, 第2.5节, 定理6] 的证明相同. \square

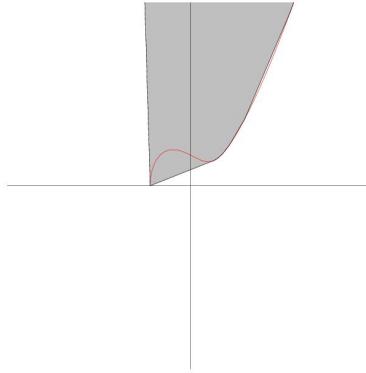


图 4.4: 例子4.8中的半代数集 S (红色曲线)和谱多面体 $\tilde{\Omega}_3(\tilde{G})$ 的投影(用阴影部分表示).

例 4.8. 考虑半代数集 S (图4.4中的红色曲线):

$$S := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^3 - x_2^2 - x_1 + 1 = 0, x_2 \geq 0\}.$$

显然可得

$$\begin{aligned}\tilde{S}^\circ &= \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_1^3 - x_0 x_2^2 - x_0^2 x_1 + x_0^3 = 0, x_2 \geq 0, x_0 > 0\}, \\ \tilde{S}^c &= \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_1^3 - x_0 x_2^2 - x_0^2 x_1 + x_0^3 = 0, x_2 \geq 0, x_0 \geq 0\}.\end{aligned}$$

显然地, $\tilde{S}^c \setminus \tilde{S}^\circ = \{(0, 0, x_2) \in \mathbb{R}^3 \mid x_2 \geq 0\}$. 可以证明 $\tilde{S}^c = \text{cl}(\tilde{S}^\circ)$, 即 S 在无穷远处是闭的. 由于在 $\text{cl}(\tilde{S}^\circ) \setminus \{0\}$ 上, $X_0 + X_2 > 0$, 因此在 $\text{co}(\text{cl}(\tilde{S}^\circ)) \setminus \{0\}$ 上, $X_0 + X_2 > 0$. 通过定理2.3, $\text{co}(\text{cl}(\tilde{S}^\circ))$ 是尖的闭凸锥. 因此 S 满足假设2. 谱多面体 $\tilde{\Omega}_3(\tilde{G})$ 的透视是图4.4中的阴影部分.

注 4.5. 如定理4.10和推论4.9所示, 对 $k \in \mathbb{N}$, 谱多面体 $\tilde{\Omega}_k(\tilde{G})$ 的投影是 $\text{cl}(\text{co}(S))$ 的外近似. 如果固定一个球 \mathbf{B}_r , 则限制在 \mathbf{B}_r 上, $\tilde{\Omega}_k(\tilde{G})$ 一致收敛到 $\text{cl}(\text{co}(S))$. 如果用球 \mathbf{B}_r 截断 S , 然后利用形如(4.3)的半正定表示来近似相交而成的有界集的闭凸包. 在通常情况下, 不能得到 $\text{cl}(\text{co}(S)) \cap \mathbf{B}_r$ 的外近似. 例如在例4.8中, 与图4.4相比, 图4.5所示的 $\text{cl}(\text{co}(S \cap \mathbf{B}_r))$ 的Lasserre半正定表示不是 $\text{cl}(\text{co}(S)) \cap \mathbf{B}_r$ 的外近似.

通过推广文献 [30, 33] 中相应结论的证明, 我们说明对每一个 $k \in \mathbb{N}$, 如果 $\mathcal{Q}_k(\tilde{G})$ 是闭的, 则 $\widetilde{\text{TH}}_k(\tilde{G}) = \text{cl}(\tilde{\Omega}_k(\tilde{G}))$.

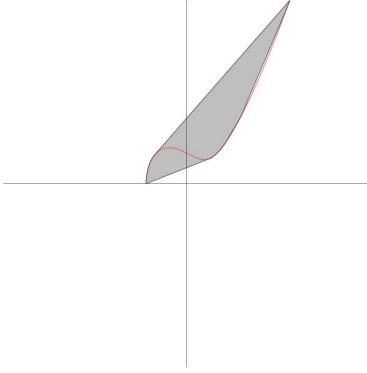


图 4.5: 例子4.8中 $\text{cl}(\text{co}(S)) \cap \mathbf{B}_r$ 的Lasserre半正定表示.

令 $M := \{(1, x) \mid x \in \mathbb{R}^n\}$ 并且 $\mathcal{Q}_k^1(\tilde{G}) = \mathcal{Q}_k(\tilde{G}) \cap \mathbb{P}[\tilde{X}]_1$. 通过对偶锥的定义, 可得

$$\mathcal{Q}_k^1(\tilde{G})^* \cap M = \{1\} \times \widetilde{\text{TH}}_k(\tilde{G}). \quad (4.26)$$

令 $\text{proj}(\mathcal{Q}_k(\tilde{G})^*)$ 是 $\mathcal{Q}_k(\tilde{G})^*$ 到 $(\mathbb{P}[\tilde{X}]_1)^*$ 的投影. 显然可得

$$\text{proj}(\mathcal{Q}_k(\tilde{G})^*) \cap M = \{1\} \times \widetilde{\Omega}_k(\tilde{G}). \quad (4.27)$$

如果 $\mathcal{Q}_k(\tilde{G})$ 是闭的, 通过命题2.2, 可得

$$\mathcal{Q}_k^1(\tilde{G})^* = \text{cl}(\text{proj}(\mathcal{Q}_k(\tilde{G})^*)). \quad (4.28)$$

引理 4.11. 如果 $\mathcal{Q}_k(\tilde{G})$ 是闭的, 超平面 M 与 $\text{ri}(\text{proj}(\mathcal{Q}_k(\tilde{G})^*))$ 相交.

证明. 通过 [76, 定理 6.3] 和(4.28), 这等价于证明 M 与 $\text{ri}(\mathcal{Q}_k^1(\tilde{G})^*)$ 相交. 固定向量 $u \in S$, 可得 $\tilde{l} := (1, u)/\|(1, u)\|_2 \in \widetilde{S}$ 并且 $\tilde{l} \in \mathcal{Q}_k^1(\tilde{G})^*$. 令

$$D := \{t_0 \in \mathbb{R} \mid \text{存在 } t \in \mathbb{R}^n, \text{ 使得 } (t_0, t) \in \mathcal{Q}_k^1(\tilde{G})^*\}.$$

由于 $X_0 \in \mathcal{Q}_k^1(\tilde{G})$ 并且对任意的 $c \geq 0$, 有 $c \cdot \tilde{l} \in \mathcal{Q}_k^1(\tilde{G})^*$, 因此 $D = [0, \infty)$ 从而 $1 \in \text{ri}(D)$. 通过 [76, 定理 6.8], 可得 M 与 $\text{ri}(\mathcal{Q}_k^1(\tilde{G})^*)$ 相交. \square

定理 4.12. 如果 $\mathcal{Q}_k(\tilde{G})$ 是闭的, 则 $\widetilde{\text{TH}}_k(\tilde{G}) = \text{cl}(\widetilde{\Omega}_k(\tilde{G}))$.

证明. 通过(4.26), (4.27), (4.28), [76, 推论 6.5.1]和引理4.11, 可得

$$\begin{aligned} \{1\} \times \text{cl}\left(\tilde{\Omega}_k(\tilde{G})\right) &= \text{cl}\left(\text{proj}\left(\mathcal{Q}_k(\tilde{G})^*\right) \cap M\right) \\ &= \text{cl}\left(\text{proj}\left(\mathcal{Q}_k(\tilde{G})^*\right)\right) \cap M \\ &= \mathcal{Q}_k^1(\tilde{G})^* \cap M \\ &= \{1\} \times \widetilde{\text{TH}}_k(\tilde{G}). \end{aligned}$$

因此 $\widetilde{\text{TH}}_k(\tilde{G}) = \text{cl}\left(\tilde{\Omega}_k(\tilde{G})\right)$. □

推论 4.13. 如果 $\mathcal{Q}_k(\tilde{G})$ 是闭的, 则推广 Theta体(4.11)等于 \tilde{S} 的 Theta体(4.5)的透视投影, 即

$$\widetilde{\text{TH}}_k(\tilde{G}) = \varphi\left(\text{TH}_k(\tilde{G}) \setminus H\right).$$

证明. 通过(4.21), 命题4.7以及定理4.12可以直接证明. □

注 4.6. 由于 \tilde{S} 不存在内点, 不能通过与注4.1相同的方式来验证 $\mathcal{Q}_k(\tilde{G})$ 的闭性. 然而, 如果原始的集合 S 存在内点, 则 \tilde{S} 的零化理想 $I = \langle \|\tilde{X}\|_2^2 - 1 \rangle$ 和 $\mathcal{Q}_k(\tilde{G})$ 在商环 $\mathbb{R}[\tilde{X}]/I$ 中的像在 $\mathbb{R}[\tilde{X}]_{2k}/I$ 中是闭的, 参见([58, 68]). 因此, 在定理4.12和推论4.13中的等式对定义在 $\mathbb{R}[\tilde{X}]/I$ 中的推广Theta体和Lasserre半正定表示成立. 二次模 $\mathcal{Q}_k(\tilde{G})$ 是闭的条件有必要作进一步的研究.

4.3 假设2的详细讨论

如果假设2满足, 我们构造了收敛到 $\text{cl}(\text{co}(S))$ 的半正定表示序列. 在这一节, 我们将进一步讨论假设2不成立的情形.

4.3.1 S 在无穷远处是闭的

[38]说明在一般情况下, 半代数集在无穷远处是闭的. 但是, 在无穷远处是闭的条件对某些半代数集不成立.

半代数集 U 定义如下:

$$U := \left\{ x \in \mathbb{R}^n \mid \begin{array}{l} h_i(x) = 0, i = 1, \dots, m_1, \\ g_j(x) \geq 0, \deg(g_j) \text{ 是偶数}, j = 1, \dots, m_2 \end{array} \right\}. \quad (4.29)$$

令

$$\begin{aligned}\tilde{U}^o &= \left\{ \tilde{x} \in \mathbb{R}^{n+1} \mid \tilde{h}_i(\tilde{x}) = 0, \tilde{g}_j(\tilde{x}) \geq 0, x_0 > 0, i = 1, \dots, m_1, j = 1, \dots, m_2 \right\}, \\ \tilde{U}^c &= \left\{ \tilde{x} \in \mathbb{R}^{n+1} \mid \tilde{h}_i(\tilde{x}) = 0, \tilde{g}_j(\tilde{x}) \geq 0, x_0 \geq 0, i = 1, \dots, m_1, j = 1, \dots, m_2 \right\}.\end{aligned}$$

命题 4.14. 假设 U 无界. 如果 $\text{co}(\text{cl}(\tilde{U}^o))$ 是尖的闭凸锥, 则 U 在无穷远处不是闭的.

证明. 由于 U 无界, 存在序列 $\{u^{(k)}\}_{k=1}^\infty \subseteq U$, 使得 $\lim_{k \rightarrow \infty} \|u^{(k)}\|_2 = \infty$. 由于 $\{(1, u^{(k)})/\|(1, u^{(k)})\|_2\} \subseteq \tilde{U}^o$ 是有界的, 存在非零向量 $\tilde{u} = (0, u_1, \dots, u_n) \in \text{cl}(\tilde{U}^o)$.

如果 $\text{co}(\text{cl}(\tilde{U}^o))$ 是尖的闭凸锥, 通过定理 2.3, 有 $-\tilde{u} \notin \text{cl}(\tilde{U}^o)$. 由于对 $j = 1, \dots, m_2$, $\deg(g_j)$ 是奇数, 显然可得 \tilde{u} 和 $-\tilde{u}$ 属于 \tilde{U}^c . 因此 $\text{cl}(\tilde{U}^o) \neq \tilde{U}^c$ 并且 U 在无穷远处不是闭的. \square

注 4.7. 考虑形如(4.8)的半代数集 S . 对 $i = 1, \dots, m$, 令 \hat{g}_i 为 g_i 的最高次齐次部分并且 $D_S = \tilde{S}^c \setminus \text{cl}(\tilde{S}^o)$. 如果 S 在无穷远处不是闭的, 则

$$\emptyset \neq D_S \subseteq \{(0, x) \in \mathbb{R}^{n+1} \mid \hat{g}_1(x) \geq 0, \dots, \hat{g}_m(x) \geq 0\}.$$

将 D_S 分解成 $D_S = D_S^1 \cup D_S^2$, 这里

$$D_S^1 = \left\{ (0, x) \in \mathbb{R}^{n+1} \mid (0, x) \in \tilde{S}^c \setminus \text{cl}(\tilde{S}^o), (0, -x) \in \text{cl}(\tilde{S}^o) \right\}$$

并且

$$D_S^2 = \left\{ (0, x) \in \mathbb{R}^{n+1} \mid (0, x) \in \tilde{S}^c \setminus \text{cl}(\tilde{S}^o), (0, -x) \notin \text{cl}(\tilde{S}^o) \right\}.$$

对形如(4.29)的半代数集 S , 根据命题 4.14 的证明, 有 $D_S^1 \neq \emptyset$. 如果 $\text{co}(\text{cl}(\tilde{S}^o))$ 是尖的闭凸锥, 则存在一个线性函数 $\tilde{l} \in \mathbb{P}[\tilde{X}]_1$, 使得在 $\text{co}(\text{cl}(\tilde{S}^o)) \setminus \{0\}$ 上, 有 $\tilde{l}(\tilde{x}) > 0$. 将不等式 $\tilde{l}(\tilde{x}) \geq 0$ 添加到 \tilde{S}^o 的定义表达式中, 或者将 $\tilde{l}(1, x) \geq 0$ 添加到 S 的定义表达式中. 显然可得 S 等于 \tilde{S}^o , 但是 $D_S^1 = \emptyset$. 因此, 将更多的定义多项式添加到半代数集 S 的表达式后, S 更有可能在无穷远处是闭的.

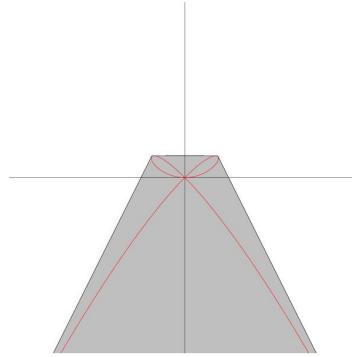


图 4.6: 例子4.9中的曲线 S (红色部分)和谱多面体 $\tilde{\Omega}_3(\tilde{G})$ (用阴影部分表示).

例 4.9. 考虑二次曲线(图4.6中的红色部分):

$$S := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^4 - x_1^2 x_2 + x_2^3 = 0\}.$$

显然可得

$$\tilde{S}^\circ = \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_1^4 - x_0 x_1^2 x_2 + x_0 x_2^3 = 0, x_0 > 0\},$$

$$\tilde{S}^c = \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_1^4 - x_0 x_1^2 x_2 + x_0 x_2^3 = 0, x_0 \geq 0\}.$$

我们通过证明线性函数 $X_0 - X_2$ 在 $\text{co}(\text{cl}(\tilde{S}^\circ)) \setminus \{0\}$ 上恒大于0来证明闭凸锥 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 是尖的.

对非零向量 $\tilde{u} = (u_0, u_1, u_2) \in \text{cl}(\tilde{S}^\circ)$, 有 $u_1^4 - u_0 u_1^2 u_2 + u_0 u_2^3 = 0$. 如果 $u_2 = 0$, 则 $u_1 = 0$ 并且 $u_0 - u_2 > 0$. 假设 $u_2 \neq 0$, 则 $u_1^2 u_2 - u_2^3 \neq 0$. 否则可得 $u_1^2 = u_2^2$ 并且 $u_2^4 - u_0 u_2^3 + u_0 u_2^3 = 0$, 因此 $u_2 = 0$, 矛盾. 因此由

$$u_0 - u_2 = \frac{u_1^4}{u_1^2 u_2 - u_2^3} - u_2 = \frac{(u_1^2 - u_2^2)^2 + u_1^2 u_2^2}{u_1^2 u_2 - u_2^3} > 0$$

可以推出 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 是尖的闭凸集. 由于 S 具有形如(4.29)的表达式, 通过命题4.14, S 在无穷远处不是闭的. 通过注4.7, 我们将 $1 - x_2 \geq 0$ 添加到 S 的表达式中, 使得 S 在无穷远处是闭的. 图4.6为新表达式的谱多面体 $\tilde{\Omega}_3(\tilde{G})$.

4.3.2 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 的尖性

如果 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 不是尖的, 我们将 S 沿着坐标轴分成 2^n 个部分. 令

$$\mathcal{E} := \{e = (e_1, \dots, e_n) \in \mathbb{R}^n \mid e_i \in \{0, 1\}, i = 1, \dots, n\}, \quad (4.30)$$

并且对于每一个 $e \in \mathcal{E}$,

$$S_e := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, (-1)^{e_j} x_j \geq 0, i = 1, \dots, m, j = 1, \dots, n\}. \quad (4.31)$$

则 $S = \bigcup_{e \in \mathcal{E}} S_e$ 并且 $|\mathcal{E}| = 2^n$. 对每一个 $e \in \mathcal{E}$, 按照(4.9)和(4.10)定义 \tilde{S}_e° , \tilde{S}_e^c , \tilde{S}_e 和 \tilde{G}_e . 通过定理2.3, 对每一个 $e \in \mathcal{E}$, $\text{co}(\text{cl}(\tilde{S}_e^\circ))$ 和 $\text{co}(\tilde{S}_e^c)$ 是尖的闭凸锥.

定理 4.15. 令 $S \subseteq \mathbb{R}^n$ 为形如(4.8)的半代数集. 假设

1. S 在无穷远处是闭的.

2. 对任意的 $e \in \mathcal{E}$, \tilde{G}_e 的PP-BDR条件成立.

则 $\text{cl}(\text{co}(S))$ 是谱多面体的投影的闭包.

证明. 固定一个整数 k' , 使得每一个 \tilde{G}_e 满足第 k' 阶PP-BDR条件. 注意到对某一个 $e \in \mathcal{E}$, S_e 可能在无穷远处不是闭的. 然而,

$$\text{对任意的 } e \in \mathcal{E}, \text{ cl}(\text{co}(S_e)) \subseteq \text{cl}\left(\tilde{\Omega}_{k'}(\tilde{G}_e)\right) \subseteq \text{cl}(\text{co}(S)). \quad (4.32)$$

通过定理4.8, 可得

$$\text{对任意的 } e \in \mathcal{E}, \text{ cl}(\text{co}(S_e)) \subseteq \text{cl}\left(\tilde{\Omega}_{k'}(\tilde{G}_e)\right) \subseteq \widetilde{\text{TH}}_{k'}(\tilde{G}_e).$$

固定向量 $u \notin \text{cl}(\text{co}(S))$. 根据(4.14), 存在线性函数 $\tilde{f} \in \mathbb{P}[\tilde{X}]_1$, 使得 $\tilde{f}(1, u) < 0$ 并且在 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 上, 有 $\tilde{f}(\tilde{x}) \geq 0$. 由于 $\text{cl}(\tilde{S}^\circ) = \tilde{S}^c$ 并且对每一个 $e \in \mathcal{E}$, $\tilde{S}_e^c \subseteq \tilde{S}^c$, 在 $\text{co}(\tilde{S}_e^c)$ 上, 有 $\tilde{f}(\tilde{x}) \geq 0$. 由于 $\text{co}(\tilde{S}_e^c)$ 是尖的闭凸锥. 通过定理2.3, 存在线性函数 $\tilde{g} \in \mathbb{P}[\tilde{X}]_1$, 使得在 $\text{co}(\tilde{S}_e^c)$ 上, 有 $\tilde{g}(\tilde{x}) > 0$. 选取一个充分小的 $\epsilon > 0$, 使得 $(\tilde{f} + \epsilon \tilde{g})(1, u) < 0$ 并且将 $\tilde{f} + \epsilon \tilde{g}$ 重新命名为 \tilde{f} . 则在 \tilde{S}_e^c 上, 有 $\tilde{f}(\tilde{x}) > 0$. 特别地, $\tilde{f}(\tilde{x})$ 在 \tilde{S}_e 上恒为正. 由于 \tilde{G}_e 满足第 k' 阶PP-BDR条件, 可得 $\tilde{f} \in \mathcal{Q}_{k'}(\tilde{G}_e)$. 由于 $\tilde{f}(1, u) < 0$, 可得 $u \notin \widetilde{\text{TH}}_{k'}(\tilde{G}_e)$. 因此 $\widetilde{\text{TH}}_{k'}(\tilde{G}_e) \subseteq \text{cl}(\text{co}(S))$ 并且(4.32)成立.

因此可得

$$\begin{aligned}
\text{cl}(\text{co}(S)) &= \text{cl}\left(\text{co}\left(\bigcup_{e \in \mathcal{E}} S_e\right)\right) \\
&= \text{cl}\left(\text{co}\left(\bigcup_{e \in \mathcal{E}} \text{cl}(\text{co}(S_e))\right)\right) \\
&\subseteq \text{cl}\left(\text{co}\left(\bigcup_{e \in \mathcal{E}} \text{cl}\left(\tilde{\Omega}_{k'}(\tilde{G}_e)\right)\right)\right) \\
&\subseteq \text{cl}(\text{co}(\text{cl}(\text{co}(S)))) \\
&= \text{cl}(\text{co}(S)).
\end{aligned}$$

因此

$$\text{cl}(\text{co}(S)) = \text{cl}\left(\text{co}\left(\bigcup_{e \in \mathcal{E}} \text{cl}\left(\tilde{\Omega}_{k'}(\tilde{G}_e)\right)\right)\right) = \text{cl}\left(\text{co}\left(\bigcup_{e \in \mathcal{E}} \tilde{\Omega}_{k'}(\tilde{G}_e)\right)\right).$$

由于对任意的 $k' \in \mathbb{N}$, $\tilde{\Omega}_{k'}(\tilde{G}_e)$ 是谱多面体的投影, 通过 [39, 定理 2.2], 可得

$$\text{cl}\left(\text{co}\left(\bigcup_{e \in \mathcal{E}} \tilde{\Omega}_{k'}(\tilde{G}_e)\right)\right) = \text{cl}\left(\left\{\sum_{e \in \mathcal{E}} \lambda_e x^{(e)} \mid \sum_{e \in \mathcal{E}} \lambda_e = 1, \lambda_e \geq 0, x^{(e)} \in \tilde{\Omega}_{k'}(\tilde{G}_e)\right\}\right)$$

是谱多面体投影的闭包. \square

注 4.8. 如果集合 $\mathcal{E}' \subseteq \mathcal{E}$ 满足条件 $S = \bigcup_{e \in \mathcal{E}'} S_e$, 将定理4.15证明中的 \mathcal{E} 替换成 \mathcal{E}' , 定理4.15仍然成立.

例子4.6(续) 对半代数集 $S = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2^3 - x_1^2 \geq 0\}$, 可得 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 不是尖的闭凸锥并且推广Theta体不收敛到 $\text{cl}(\text{co}(S))$. 通过注4.8, 我们将 S 分成两个部分:

$$\begin{aligned}
S_{(0,0)} &:= \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2^3 - x_1^2 \geq 0, x_1 \geq 0, x_2 \geq 0\}, \\
S_{(1,0)} &:= \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2^3 - x_1^2 \geq 0, -x_1 \geq 0, x_2 \geq 0\}.
\end{aligned}$$

显然可得 $\tilde{G}_{(0,0)}$ 和 $\tilde{G}_{(1,0)}$ 满足第1阶PP-BDR条件. 因此, 对任意的整数 $k' \geq 1$, 可得

$$\begin{aligned}
\tilde{\Omega}_{k'}(\tilde{G}_{(0,0)}) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0\}, \\
\tilde{\Omega}_{k'}(\tilde{G}_{(1,0)}) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \leq 0, x_2 \geq 0\}.
\end{aligned}$$

因此对任意的整数 $k' \geq 1$, $\text{cl}(\text{co}(\tilde{\Omega}_{k'}(\tilde{G}_{(0,0)}) \cup \tilde{\Omega}_{k'}(\tilde{G}_{(1,0)}))) = \text{cl}(\text{co}(S))$.

然而, 如果某一个 \tilde{G}_e 不满足第 k' 阶 PP-BDR 条件, 根据定理4.15 的证明, 对某一个 $e \in \mathcal{E}$, $\text{cl}(\tilde{\Omega}_{k'}(\tilde{G}_e))$ 可能不是 $\text{cl}(\text{co}(S))$ 的子集. 在这个情形下, $\text{cl}(\text{co}(\bigcup_{e \in \mathcal{E}} \tilde{\Omega}_{k'}(\tilde{G}_e)))$ 可能严格包含 $\text{cl}(\text{co}(S))$.

例 4.10. 将例子4.6中的半代数集 S 关于原点逆时针旋转 45° , 可得

$$S' := \{(x_1, x_2) \in \mathbb{R}^2 \mid -\sqrt{2}(x_1 - x_2)^3 - 2(x_1 + x_2)^2 \geq 0\},$$

(图4.7中红色曲线的左半部分). 则 $\text{cl}(\text{co}(S))$ 是直线 $X_2 = X_1$ 的左半平面.

通过注4.8, $\mathcal{E}' = \{(0,0), (1,0), (1,1)\}$ 将形如(4.31)的半代数集 S' 分解为 $S' = \bigcup_{e \in \mathcal{E}'} S'_e$. 对任意的正整数 $k' \geq 1$, 可得

$$\tilde{\Omega}_{k'}(\tilde{G}'_{(1,0)}) = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \leq 0, x_2 \geq 0\}.$$

$S'_{(0,0)}$ 的第3阶Lasserre半正定表示 $\tilde{\Omega}_3(\tilde{G}'_{(0,0)})$ 是图4.7的阴影部分.

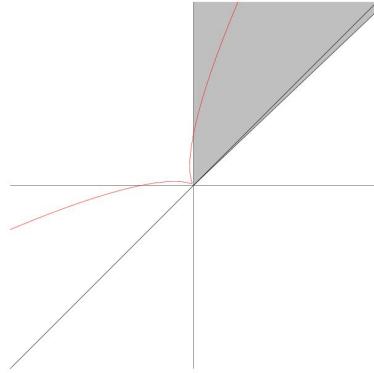


图 4.7: 例子4.10中的 $S'_{(0,0)}$ 的谱多面体近似 $\tilde{\Omega}_3(\tilde{G}'_{(0,0)})$ (用阴影部分表示)

如图所示, 切线 $X_2 = X_1$ 被参数为 $a < 1$ 的直线 $X_2 = aX_1$ 近似. 相同的情形发生在第三象限. 从图4.7中可以得到 $\text{cl}(\text{co}(\bigcup_{e \in \mathcal{E}'} \tilde{\Omega}_3(\tilde{G}'_e))) = \mathbb{R}^2$ 严格包含 $\text{cl}(\text{co}(S))$. 因此, 如果 $\text{co}(\text{cl}(\tilde{S}^\circ))$ 不是尖的闭凸锥并且对于某一个 $e \in \mathcal{E}$, \tilde{G}_e 不满足 PP-BDR 条件. 近似 $\text{cl}(\text{co}(S))$ 会非常复杂.

第五章 参数化优化问题的求解

5.1 问题背景

在这一章, 我们考虑参数化优化问题:

$$c_0^* := \sup_{x \in \mathbb{R}^n} \mathbf{c}^T x = \mathbf{c}_1 x_1 + \cdots + \mathbf{c}_n x_n \quad (5.1)$$

使得 $h_1(x) = \cdots = h_p(x) = 0,$

其中, $h_1, \dots, h_p \in \mathbb{R}[X_1, \dots, X_n]$ 并且 $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$ 是参数. 令 $\mathbf{h} = (h_1, \dots, h_p)$. 假设 $\langle \mathbf{h} \rangle$ 是等维根理想. 由于 c_0^* 可以看作是参数 \mathbf{c} 的半代数函数, 我们考虑下面的两个问题:

问题3.

- 计算一个多项式 $\Phi \in \mathbb{R}[c_0, \mathbf{c}]$, 使得对一般的 (generic) 参数值 $\gamma \in \mathbb{R}^n$, 可得 $\Phi(c_0, \gamma) \neq 0$ 并且 $\Phi(c_0^*, \gamma) = 0$.
- 计算一个多项式向量序列 $\{\Phi_i\} \subset \mathbb{R}[c_0, \mathbf{c}]$, 使得对任意的 $\gamma \in \mathbb{R}^n$, 存在 $k \in \mathbb{N}$, 下面的等式成立:

$$\Phi_k(c_0, \gamma) \neq 0, \quad \Phi_k(c_0^*, \gamma) = 0.$$

令

$$V := \{x \in \mathbb{C}^n \mid h_1(x) = 0, \dots, h_p(x) = 0\}. \quad (5.2)$$

当 V 不可约, 光滑并且 $V \cap \mathbb{R}^n$ 有界时, Rostalski 和 Sturmels 在文献 [79] 中利用对偶仿射簇来求解参数化优化问题的代数次数. 他们的结果可以用来解决问题3的第一个子问题.

定理 5.1. [79] 令 $V^* \subset (\mathbb{P}^n)^*$ 为 V 的射影闭包的对偶仿射簇. 如果 V 不可约, 光滑并且 $V \cap \mathbb{R}^n$ 有界, 则 V^* 是不可约的超曲面并且它的定义多项式是 $\Phi(-c_0, c_1, \dots, c_n)$, 其中 Φ 表示最优化函数.

定理 5.1 的证明基于当 V 不可约, 光滑并且 $V \cap \mathbb{R}^n$ 有界时, 对任意的参数值, 对应优化问题的最优化有限可达并且最优化点是关键点的事实. 当 $V \cap \mathbb{R}^n$ 无界或者 V 不光滑时, 对某一个参数值 γ , 其最优化可能是无限的或者有限可达但最优化点不是关键点. 在这两种情形下, 定理的结论可能不成立.

我们考虑下面两个情形:

- V 光滑, $V \cap \mathbb{R}^n$ 无界.
- $V \cap \mathbb{R}^n$ 不光滑, $V \cap \mathbb{R}^n$ 有界.

当 V 光滑时, 我们说明最优化函数的图像可以通过对偶仿射簇 V^* 来表示, 即对某一个参数值 γ , 只要对应优化问题的最优化 c_0^* 有限, 则 $(-c_0^* : \gamma_1 : \cdots : \gamma_n) \in V^*$. 我们设计一个算法来求解一般的参数值的优化问题(5.1). 算法返回多项式向量 (Φ, Z) , 使得

- $\Phi \in \mathbb{Q}[\mathbf{c}_0, \mathbf{c}]$ 和 $Z \in \mathbb{Q}[\mathbf{c}]$;
- 对任意的 $\gamma \notin V(Z)$, 如果问题(5.1)的最优化 c_0^* 有限, 则多项式 $\Phi(\mathbf{c}_0, \gamma) \neq 0$ 并且它的根包含问题(5.1)的最优化 c_0^* .

当 V 不可约, 光滑并且 $V \cap \mathbb{R}^n$ 的闭凸包不包含直线时, 可以推出和 [12, 定理5.23] 相同的结论. 当 V 不光滑, $V \cap \mathbb{R}^n$ 有界时, 我们递归地构造有限多个对偶仿射簇来包含最优化函数的图像.

对一些参数值, 对偶仿射簇的定义多项式可能被约化成零多项式. 对这些参数值, 对偶仿射簇无法提供对应优化问题的最优化的任何信息. 我们推广了 [37] 中的算法. 当 V 光滑时, [37] 中的算法返回一个次数是变元个数单指指数函数的单变元多项式, 其根包含对应优化问题的最优化.

我们利用 [37] 中的算法设计一个参数化的算法. 算法返回多项式向量序列

$$(\Phi_1, Z_1, \mathbf{P}_1), \dots, (\Phi_k, Z_k, \mathbf{P}_k),$$

使得

- $\Phi_i \in \mathbb{Q}[\mathbf{c}_0, \mathbf{c}]$, $Z_i \in \mathbb{Q}[\mathbf{c}]$ 并且对 $1 \leq i \leq k$, $\mathbf{P}_i \subset \mathbb{Q}[\mathbf{c}]$ 生成一个素理想.
- 对 $1 \leq i \leq k$, $V(\mathbf{P}_i) - V(Z_i) \neq \emptyset$.
- 对任意的 $\gamma = (\gamma_1, \dots, \gamma_n) \in \mathbb{R}^n$, 使得 $\gamma \in V(\mathbf{P}_i) - V(Z_i)$, 当参数值为 γ 时, $\Phi_i(\mathbf{c}_0, \gamma_1, \dots, \gamma_n) \neq 0$ 并且它的根包含优化问题(5.1)的全局最优化.

5.2 预备知识

5.2.1 对偶仿射簇

定义 $X := (X_1, \dots, X_n)$. 对 $\mathbb{R}[X]$ ($\mathbb{R}[X_0, X]$) 中的理想(齐次理想) I , $V(I)$ 定义为 I 在 $\mathbb{C}^n(\mathbb{P}^n(\mathbb{C}))$ 中的仿射簇(射影簇). 令 V_{reg} 为 $V(I)$ 的光滑点的集合. 为了简便, 我们将 $\mathbb{P}^n(\mathbb{C})$ 缩写为 \mathbb{P}^n . 假设齐次理想 $I = \langle f_1, \dots, f_p \rangle$ 是多项式环 $\mathbb{R}[X_0, X]$ 中的等维根理想并且 $V(I) \subset \mathbb{P}^n$.

射影空间 $(\mathbb{P}^n)^*$ 中的点 $u = (u_0 : u_1 : \dots : u_n)$ 表示超平面 $\{x \in \mathbb{P}^n \mid \sum_{i=0}^n u_i x_i = 0\}$. 如果光滑点 x 在超平面中, 并且向量 (u_0, u_1, \dots, u_n) 在点 x 处的雅克比矩阵 $\text{Jac}(I)$ 的行空间中, 则称 u 与 $V(I)$ 相切于光滑点 x . 余法仿射簇 $\text{CN}(V)$ 定义为下面集合的 Zariski 闭包:

$$\{(x, u) \in \mathbb{P}^n \times (\mathbb{P}^n)^* \mid x \in V_{\text{reg}} \text{ 并且 } u \text{ 与 } V(I) \text{ 在点 } x \text{ 处相切}\}.$$

对偶仿射簇 $V(I)^*$ 是 $\text{CN}(V)$ 到第二个因子的投影. 更准确地说, 对偶仿射簇是下面集合的 Zariski 闭包:

$$\{u \in (\mathbb{P}^n)^* \mid u \text{ 与 } V(I) \text{ 在某个光滑点处相切}\}.$$

5.2.2 广义关键值

给定 \mathbb{R}^n 中的向量 v , $\|v\|$ 表示 v 的标准欧几里得范数. 假设 V 是光滑的仿射簇并且定义多项式映射 $f : V \rightarrow \mathbb{R}$. f 在 V 上的非正则点(*non-properness*)的集合定义为

$$\{z \in \mathbb{R} \mid \text{存在 } y_k \in V, \text{ 使得 } \|y_k\| \rightarrow \infty, f(y_k) \rightarrow z\}.$$

将 $K_0(f, V)$ 记作 f 在 V 上的关键值(critical value). f 在 V 上的渐近关键值定义为

$$K_\infty(f, V) = \left\{ y \in \mathbb{R} \mid \begin{array}{l} \text{存在 } x^{(k)} \in V, \text{ 使得 } x^{(k)} \rightarrow \infty, \\ f(x^{(k)}) \rightarrow y, \|x^{(k)}\| \nu(d_{x^{(k)}} f) \rightarrow 0 \end{array} \right\},$$

其中 $d_{x^{(k)}} f$ 表示 f 在 $x^{(k)}$ 处的微分并且 $\nu(d_{x^{(k)}} f)$ 表示 $d_{x^{(k)}} f$ 限制在 V 中点 $x^{(k)}$ 的切空间 $T_{x^{(k)}} V$ 上与退化线性映射空间的距离. 如果将 B 记作以切空间 $T_{x^{(k)}} V$ 的一组正交基为列向量的矩阵并且令 J 为 f 在点 $x^{(k)}$ 处的雅克比矩阵的转置, 则

$$\nu(d_{x^{(k)}} f) = \inf_{\|y\|=1} \|B^T J y\|.$$

f 的广义关键值的集合定义为

$$K(f, V) = K_0(f, V) \cup K_\infty(f, V).$$

文献 [46, 定理 3.1] 和 [44, 定理 3.3, 推论 4.1] 已经证明 $K(f, V)$ 是一个有限集合.

定理 5.2. [44, 70] 如果 f 有上界, 即 $f^* = \sup_{x \in V} f(x)$ 是有限的, 则 $f^* \in K(f, V)$.

5.3 无界情形的对偶性

对形如(5.2)的仿射簇 V , 令

$$C_h = \text{cl}(\text{co}(V \cap \mathbb{R}^n)),$$

即 $V \cap \mathbb{R}^n$ 的闭凸包. 问题(5.1)等价于

$$c_0^* = \sup \mathbf{c}^T x \quad \text{使得 } x \in C_h. \quad (5.3)$$

$\text{dom}(c_0^*(\mathbf{c} \mid C_h))$ 是使得 $\gamma^T x$ 在 C_h 上的最优值有限的参数值 $\gamma \in \mathbb{R}^n$ 的集合. 定义 V^* 为 V 的射影闭包的对偶仿射簇. 当 C_h 不包含直线 ($0^+ C_h$ 是尖的闭凸锥) 时, 问题(5.1) 的最优值函数与 V^* 具有重要的联系, 见 [21]. 现在, 我们证明一个更强的结论并且设计一个解决问题3的第一个子问题的算法.

5.3.1 光滑无界情形

在这一节, 我们假设形如(5.2)的仿射簇 V 是光滑的. 我们将文献 [12, 定理 5.23] 中的结果推广到无界的情形.

定理 5.3. 假设形如(5.2)中的仿射簇 V 是光滑的. 对每一个 $\gamma \in \text{dom}(c_0^*(\mathbf{c} \mid C_h))$,

$$(-c_0^* : \gamma_1 : \cdots : \gamma_n) \in V^*. \quad (5.4)$$

证明. 给定 $\gamma \in \text{dom}(c_0^*(\mathbf{c} \mid C_h))$. $f(X) = \gamma^T X$ 在 $V \cap \mathbb{R}^n$ 上的最优值 c_0^* 是有限的. 当最优值 c_0^* 可达时, 见 [12, 定理 5.23] 的证明. 假设 c_0^* 不可达, 则通过定理 5.2 可知, c_0^* 是 f 在 $V \cap \mathbb{R}^n$ 上的渐近关键值. 因此, 存在序列 $\{x^{(k)}\} \subseteq V \cap \mathbb{R}^n$, 使得 $\|x^{(k)}\| \rightarrow \infty$, $f(x^{(k)}) \rightarrow c_0^*$ 并且 $\|x^{(k)}\| \nu(d_{x^{(k)}} f) \rightarrow 0$. 通过 [97, 引理 2.1], 对每一个 $x^{(k)}$, V 在点 $\{x^{(k)}\}$ 处的法空间中存在一个向量 $\gamma^{(k)}$, 使得 $\|\gamma^{(k)} - \gamma\| = \nu(d_{x^{(k)}} f)$. 因此 $\|x^{(k)}\| \|\gamma^{(k)} - \gamma\| \rightarrow 0$, 从而 $\|\gamma^{(k)} - \gamma\| \rightarrow 0$ 并且 $(\gamma^{(k)})^T x^{(k)} \rightarrow c_0^*$. 可以验证

$$(-(\gamma^{(k)})^T x^{(k)} : \gamma_1^{(k)} : \cdots : \gamma_n^{(k)}) \in V^*.$$

由 V^* 是闭集可知, $(-c_0^* : \gamma_1 : \cdots : \gamma_n) \in V^*$. \square

推论 5.4. 如果 V 不可约, 光滑并且 C_h 不包含直线, 则 V^* 是不可约的超曲面并且它的定义多项式是 $\Phi(-c_0, c_1, \dots, c_n)$, 其中 Φ 表示最优化函数.

证明. 由于 C_h 不包含直线, $0^+(C_h)$ 是尖的闭凸集. 根据定理 2.5, $(0^+C_h)^\circ$ 是 n 维闭凸集并且 $\text{int}((0^+C_h)^\circ) \subseteq \text{dom}(c_0^*(\mathbf{c} | C_h))$. 在仿射空间中, 集合

$$\left\{ (-c_0^* : \gamma_1 : \dots : \gamma_n) \in (\mathbb{P}^n)^* \mid \gamma \in \text{int}((0^+C_h)^\circ) \right\} \quad (5.5)$$

的 Zariski 闭包的维数至少为 n . 在射影空间中, 通过 [2, 定理 12(i), §3, 第 9 章], 集合(5.5)的 Zariski 闭包的维数至少是 $n - 1$. 通过定理 5.3 和 [12, 命题 5.10], 可得 $\dim(V^*) = n - 1$.

由于 V 不可约, 由文献 [25, 命题 1.3] 可知, V 是不可约的超曲面并且通过文献 [2, 命题 10(ii), §4, 第 9 章], V^* 是集合(5.5)的 Zariski 闭包, 因此结论成立. \square

给定一个性质. 如果存在关于未定元的非空 Zariski 开集, 使得当未定元在这个开集中取值时, 性质成立, 则称这个性质对于未定元的一般的(generic)值均成立.

可以利用定理 5.3 设计一个解决问题 3 的第一个子问题的算法. 理想 $J \subseteq \mathbb{Q}[\mathbf{c}_0, \mathbf{c}, \mu, X]$ 定义如下:

$$\langle \mathbf{c}^T X - \mathbf{c}_0, h_1, \dots, h_p, \cup_{i=1}^n \mathbf{c}_i - \sum_{j=1}^p \mu_j \frac{\partial h_j}{\partial X_i} \rangle,$$

其中 $\mu = (\mu_1, \dots, \mu_p)$ 是未定元向量. 由于 $\langle \mathbf{h} \rangle$ 是等维根理想, 对任意的 $(c_0^*, \gamma, \bar{\mu}, \bar{x}) \in V(J)$, c_0^* 是函数 $\gamma^T X$ 在仿射簇 V 上的关键点 \bar{x} 上的值.

算法 5.1. GenericParametricOptimization(\mathbf{h})

输入: 等维根理想的生成元 $h_1, \dots, h_p \in \mathbb{Q}[X]$

输出: 多项式向量 (Φ, Z) , 使得

- $\Phi \in \mathbb{Q}[\mathbf{c}_0, \mathbf{c}]$ 并且 $Z \in \mathbb{Q}[\mathbf{c}]$.
- 假设参数值 $\gamma \in \text{dom}(c_0^*(\mathbf{c} | C_h))$ 满足 $Z(\gamma) \neq 0$, 则 $\Phi(\mathbf{c}_0, \gamma) \neq 0$ 并且它的根包含了问题(5.1)关于参数值 γ 的最优化 c_0^* .

步骤1 在分块字典序 $X \succ \mu \succ \mathbf{c} \succ \mathbf{c}_0$ 下, 计算理想 $J \cap \mathbb{Q}[\mathbf{c}_0, \mathbf{c}]$ 的约化 Gröbner 基 G .

步骤2 令 Γ 为 G 中包含变量 \mathbf{c}_0 的多项式集合.

步骤3 令 Φ 是 Γ 中关于 \mathbf{c}_0 次数最小的多项式.

步骤4 在多项式环 $\mathbb{Q}[\mathbf{c}_1, \dots, \mathbf{c}_n][\mathbf{c}_0]$ 中, 令 Z 为 Φ 中所有系数的平方和.

定理 5.5. 在算法 5.1 中, 可得 $\Gamma \neq \emptyset$, $V^* = V(\Gamma)$ 并且算法是正确的.

证明. 通过定理5.3, 只需要说明 $\Gamma \neq \emptyset$. 定义投影映射 $\pi_{n+1} : (\mathbf{c}_0, \mathbf{c}, \mu, X) \rightarrow (\mathbf{c}_0, \mathbf{c})$. 则 G 是对应的消元理想的Gröbner基. 通过闭包定理 [2], $\pi_{n+1}(V(J)) \subseteq V(G)$ 并且存在子仿射簇 $W \subsetneqq V(G)$, 使得 $V(G) \setminus W \subseteq \pi_{n+1}(V(J))$.

给定向量 $(c_0^*, \gamma) \in V(G)$. 假设 $\Gamma = \emptyset$, 则 $\mathbb{C} \times \gamma \subseteq V(G)$. 通过Sard引理, $\mathbb{C} \times \gamma \cap \pi_{n+1}(V(J))$ 是空集或者有限集. 因此, 存在 $\mathbb{C} \times \gamma$ 中至多有限的点集 T , 使得 $\mathbb{C} \times \gamma \setminus T \subseteq V(G) \setminus \pi_{n+1}(V(J)) \subseteq W$.

由于 W 是闭集, $\mathbb{C} \times \gamma \subseteq W$. 特别地, $(c_0^*, \gamma) \in W$ 可推出 $V(G) = W$, 矛盾. \square

注 5.1. 推论5.4说明当 V 不可约, 光滑并且 C_h 不包含直线时, 算法5.1返回的集合 Γ 中只有一个多项式. 如果 C_h 包含直线, Γ 可能会包含多个多项式并且 V^* 可能不是集合(5.5)的Zariski闭包, 见例子5.1.

与 [80, 定理 6]相似, 通过算法5.1和判定实代数集非空的方法, 我们可以判定 γ ($\gamma \notin V(Z)$)是否属于 $\text{dom}(c_0^*(\mathbf{c} \mid C_h))$. 如果属于, 我们可以计算 c_0^* .

例 5.1. 仿射簇

$$V := h(X_1, X_2) = X_1^2 X_2 - 1$$

不可约, 光滑, $V \cap \mathbb{R}^2$ 无界并且 C_h 包含直线. 令 $\gamma = (0, -1)$, 显然可得 $c_0^* = 0$. 运行算法5.1, 返回 $\Gamma = \{4\mathbf{c}_0^3 + 27\mathbf{c}_1^2\mathbf{c}_2\}$ 并且

$$\Phi = 4\mathbf{c}_0^3 + 27\mathbf{c}_1^2\mathbf{c}_2, \quad V(Z) = \emptyset.$$

因此 $V^* = V(\Gamma) \subseteq \mathbb{P}^2$ 并且 $(0 : 0 : -1) \in V^*$.

由于 $\text{dom}(c_0^*(\mathbf{c} \mid C_h)) = \{\gamma \in \mathbb{R}^2 \mid \gamma_1 = 0, \gamma_2 < 0\}$ 并且对任意的 $\gamma \in \text{dom}(c_0^*(\mathbf{c} \mid C_h))$, $c_0^* = 0$, 可得集合

$$\{(-c_0^* : \gamma_1 : \gamma_2) \in (\mathbb{P}^2)^* \mid \gamma \in \text{dom}(c_0^*(\mathbf{c} \mid C_h))\} \quad (5.6)$$

的Zariski闭包是零维仿射簇 $\{(0 : 0 : \gamma_2) \in (\mathbb{P}^2)^* \mid \gamma_2 \in \mathbb{C}\}$. 由于 $\dim V^* = 1$, V^* 不是集合(5.6)的Zariski闭包.

5.3.2 奇异有界的情形

现在我们假设 V 不可约, 不光滑并且 $V \cap \mathbb{R}^n$ 有界. 在这个情形下, 我们指出包含关系(5.4)可能不成立.

例 5.2. 考虑平面代数曲线星形线 V , 其定义多项式如下

$$h(X_1, X_2) = (X_1^2 + X_2^2 - 1)^3 + 27X_1^2X_2^2.$$

显然地, 对定义在 $V \cap \mathbb{R}^2$ 上的线性函数, 对应优化问题的最优值点是四个奇异点 $\{(\pm 1, 0), (0, \pm 1)\}$ 中的一个.

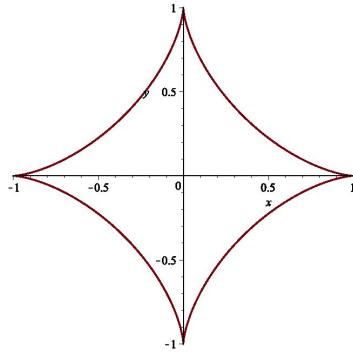


图 5.1: 例子5.2中的星形线

我们有 $V^* = V(\Gamma)$, 其中

$$\Gamma = \{-\mathbf{c}_1^2 \mathbf{c}_2^2 + \mathbf{c}_0^2 \mathbf{c}_1^2 + \mathbf{c}_2^2 \mathbf{c}_0^2\}.$$

给定一个参数值 $\gamma \in \mathbb{R}^2$, 则 $c_0^* = \max\{|\gamma_1|, |\gamma_2|\} > 0$. 可以验证当 $\gamma_1 \neq 0$ 或者 $\gamma_2 \neq 0$ 时, 可得 $(-c_0^* : \gamma_1 : \gamma_2) \notin V^*$, 即(5.4)不成立.

接下来, 我们递归地构造有限多个对偶仿射簇, 使得(5.4)对这些仿射簇的并成立. 类似的结果参见文献 [92].

算法5.2与算法5.1的输入相同并且返回有限个多项式向量 (Φ_k, Z_k) , 使得对任意的 $\gamma \in \text{dom}(c_0^*(\mathbf{c} | C_{\mathbf{h}}))$, 存在 $k \in \mathbb{N}$, 条件 $Z_k(\gamma) \neq 0$ 和 $\Phi_k(c_0^*, \gamma) = 0$ 成立.

算法 5.2. SingularParametricOptimization(\mathbf{h})

步骤 1 令 $k = 1$ 并且 $V_k = V$.

步骤 2 计算等维分解 $V_k = \cup_i V_{k,i}$, 其中 $V_{k,i} = V(I_{k,i})$ 并且 $I_{k,i}$ 是等维根理想.

步骤 3 对每一个 i , 运行函数 GenericParametricOptimization($I_{k,i}$), 并且令 $(V^{(k)})^* = \cup_i V_{k,i}^*$.

步骤 4 计算集合 $\Gamma_k \subseteq \mathbb{Q}[\mathbf{c}_0, \mathbf{c}]$, 使得 $V(\Gamma_k) = (V^{(k)})^*$.

步骤 5 令 Φ_k 为多项式集合 Γ_k 中关于 \mathbf{c}_0 次数最小的多项式.

步骤 6 令 Z_k 是多项式环 $\mathbb{Q}[\mathbf{c}_1, \dots, \mathbf{c}_n][\mathbf{c}_0]$ 中 Φ_k 的系数的平方和.

步骤 7 计算每一个集合 $V_{k,i}$ 的奇异点集 $\tilde{V}_{k,i}$ 和集合 $V_{k+1} = \cup_i \tilde{V}_{k,i}$. 如果 $V_{k+1} \neq \emptyset$, 则 $k = k + 1$ 并且回到步骤2.

下一个定理说明算法5.2的正确性和有限终止性.

定理 5.6. 算法 5.2 有限步终止. 当算法终止时, 对任意的 $\gamma \in \mathbb{R}^n$, 可得

$$(-c_0^* : \gamma_1 : \dots : \gamma_n) \subseteq \cup_{k=1}^l (V^{(k)})^*.$$

证明. 由于 $\tilde{V}_{k,i}$ 是 $V_{k,i}$ 的奇异点的集合, $\dim(\tilde{V}_{k,i}) < \dim(V_{k,i})$ 并且算法在有限步终止. 由于 $V \cap \mathbb{R}^n$ 有界, 对每一个参数 γ , 最优值 c_0^* 是有限可达的. 如果最优值点 x^* 是一个光滑点, 通过定理 5.3, $(-c_0^* : \gamma_1 : \dots : \gamma_n) \in V^*$. 如果 x^* 是 V 的一个奇异点, 则存在 $k, i \in \mathbb{N}$, 使得 x^* 是 $V_{k,i}$ 的光滑点并且 $(-c_0^* : \gamma_1 : \dots : \gamma_n) \in V_{k,i}^*$. \square

例子 5.2(续) V 的奇异点的集合 $\text{sing}(V)$ 由

$$\{h, X_1^5 - X_1, X_1^3 X_2 + X_1 X_2, 3X_1^4 - X_1^2 + 2X_2^2 - 2\}$$

定义并且 $\text{sing}(V) = \{(\pm 1, 0), (0, \pm 1)\}$.

运行算法 5.2, 可知集合 Γ_2 只包含下面的多项式:

$$\begin{aligned} & (\mathbf{c}_0 - \mathbf{c}_2)(\mathbf{c}_0 + \mathbf{c}_2)(\mathbf{c}_0 - \mathbf{c}_1)(\mathbf{c}_0 + \mathbf{c}_1)(\mathbf{c}_0^2 + \mathbf{c}_1^2 - 2\mathbf{c}_1\mathbf{c}_2 + \mathbf{c}_2^2) \\ & (\mathbf{c}_0^2 + \mathbf{c}_1^2 + 2\mathbf{c}_1\mathbf{c}_2 + \mathbf{c}_2^2). \end{aligned}$$

可以验证对每一个 $\gamma \in \mathbb{R}^2$, 可得 $(-c_0^* : \gamma_1 : \gamma_2) \in V(\Gamma_2) = (V^{(2)})^*$.

5.3.3 对偶仿射簇的坏参数

令 (Φ, Z) 为算法 5.1 返回的多项式向量. 对 $V(Z)$ 中的任意参数值, Φ 被约化成零多项式. 多项式 Φ 不能提供关于 $V(Z)$ 中参数值的优化问题(5.1)的最优值的任何信息. $V(Z)$ 称为对偶仿射簇的坏参数. 接下来, 我们考虑关于坏参数值的参数化优化问题(5.1) 的求解.

考虑多项式优化问题

$$f^* := \max_{x \in \mathbb{R}^n} f(x) \quad \text{使得 } h_1(x) = \dots = h_p(x) = 0,$$

这里, $f \in \mathbb{R}[X]$ 并且 $V = V(h_1(x), \dots, h_p(x))$. 如果 f 在 $V \cap \mathbb{R}^n$ 上有上界, 则 $f^* \in K(f, V \cap \mathbb{R}^n)$. 令

$$V_{\mathbf{h}, f} = \{(x, x_{n+1}) \in \mathbb{C}^{n+1} \mid x \in V, x_{n+1} - f(x) = 0\}.$$

通过定理 5.3, 可得 $(-f^* : 0 : \dots : 0 : 1) \in V_{\mathbf{h}, f}^*$.

例 5.3. [46, 例子 2.1] 令 $f = (X_1 + X_1^2 X_2 + X_1^4 X_2 X_3)^2$. 对 $V_{\mathbf{h}, f}$ 运行算法 5.1, 返回 $\Gamma = \{\Phi\}$, 这里

$$\begin{aligned} \Phi = & 1073741824\mathbf{c}_0^{12}\mathbf{c}_2^4\mathbf{c}_4^2 + 268435456\mathbf{c}_0^{11}\mathbf{c}_1^2\mathbf{c}_2^4\mathbf{c}_4 - \\ & + 9865003008\mathbf{c}_0^{10}\mathbf{c}_2^4\mathbf{c}_3\mathbf{c}_4^3 + \dots + 520093696\mathbf{c}_0^9\mathbf{c}_1\mathbf{c}_2^3\mathbf{c}_3^2\mathbf{c}_4^3. \end{aligned}$$

$\Phi(\mathbf{c}_0, 0, 0, 0, 1) \equiv 0$, 我们不能得到 f^* 的任何信息. 实际上, $V(Z) = V(\mathbf{c}_2\mathbf{c}_4, \mathbf{c}_3\mathbf{c}_4, \mathbf{c}_3\mathbf{c}_2\mathbf{c}_1)$. 因此, 对 $\mathbf{c}_2 = 0, \mathbf{c}_3 = 0, \Phi(\mathbf{c}_0, \mathbf{c}_1, 0, 0, \mathbf{c}_4) \equiv 0$.

5.4 完整的算法

5.4.1 背景

在这一节, 我们设计一个算法来解决问题3的第二个子问题. 与上一节相同, $\mathbf{h} := (h_1, \dots, h_p) \subset \mathbb{Q}[X_1, \dots, X_n]$ 生成一个等维根理想, 并且 V 为由(5.2)定义的代数集. 我们假设 V 是光滑的并且令 r 为它的余维数.

一些坏参数 $\gamma = (\gamma_1, \dots, \gamma_n)$ 很可能发生. 对坏参数, 对偶仿射簇 V^* 的定义多项式被约化成为零多项式. 对偶仿射簇对求解线性函数 $x \rightarrow \gamma^T x$ 在 $V \cap \mathbb{R}^n$ 上的最优值 c_0^* 没有任何帮助. 例如, 在例子5.3中, 多项式 Φ 被参数值 $\gamma = (\gamma_1, 0, 0, \gamma_4)$ 约化成零多项式. 在这一节, 我们设计一个算法来解决对偶仿射簇的坏参数所带来的困难. 这个算法是文献 [37]中的算法的参数化形式.

我们的算法返回一个多项式向量 (Φ, Z, \mathbf{P}) , 其中 $\Phi \in \mathbb{Q}[\mathbf{c}_0, \mathbf{c}]$, $\mathbf{P} \subset \mathbb{Q}[\mathbf{c}]$ 并且 $Z \in \mathbb{Q}[\mathbf{c}] - \langle \mathbf{P} \rangle$. 对任意的 $(\gamma_1, \dots, \gamma_n) \in V(\mathbf{P}) - V(Z)$, $\Phi(\mathbf{c}_0, \gamma_1, \dots, \gamma_n)$ 是非零多项式并且它的根包含函数 $x \rightarrow \gamma^T x$ 在 $V \cap \mathbb{R}^n$ 上的最优值.

接下来, 算法在 $\sqrt{\langle \mathbf{P} \rangle + \langle Z \rangle}$ 的准素分支上递归地求解参数化优化问题. 我们在整环 $\mathbb{Q}[\mathbf{c}_1, \dots, \mathbf{c}_n]/\mathcal{P}$ 上来运行算法, 其中 $\mathcal{P} \subset \mathbb{Q}[\mathbf{c}_1, \dots, \mathbf{c}_n]$ 是一个素理想. 算法在整环上进行计算, 计算过程中所有的操作在这个环上有效. 唯一的差别在于我们需要计算多项式模 \mathcal{P} 的伪逆, 因此可以在 $\mathbb{Q}[\mathbf{c}_1, \dots, \mathbf{c}_n]/\mathcal{P}$ 的分式域上进行计算.

在整环上运行函数BasicParametricOptimization. 它是文献 [37, 第3节]中算法SetContainingLocalExtrema的参数化推广. 它的主要优势在于返回的多项式系统中多项式的最高次数是变元个数的单指数函数(参见 [37, 第6节, 引理 6.8]). 柱代数分解返回多项式系统中多项式的最高次数是变元个数的双指数函数. 在我们介绍算法的递归步骤之前, 我们先描述在文献 [37]中算法包含的函数和细节.

5.4.2 极仿射簇

在这一节, 我们介绍极(polar)仿射簇(参见 [5, 7]及相关文献)和诺特位置(参见 [81])的性质. 令 $\mathbf{P} \subset \mathbb{Q}[\mathbf{c}]$ 为素理想 \mathcal{P} 的生成元集, 并且令 $\mathbb{A} = \mathbb{Q}[\mathbf{c}]/\mathcal{P}$. 因此 \mathbb{A} 是一个整环.

对 $\gamma \in V(\mathbf{P})$ 和 $1 \leq i \leq n$, 考虑第*i*个标准投影 π_i :

$$\pi_i : (x_1, \dots, x_n) \rightarrow (x_1, \dots, x_i)$$

和投影 π_γ :

$$\pi_\gamma : x = (x_1, \dots, x_n) \rightarrow \gamma^T x = \gamma_1 x_1 + \dots + \gamma_n x_n.$$

对 $1 \leq i \leq \dim(V) = n - r$, 定义映射 $\pi_{\gamma,i}$:

$$\pi_{\gamma,i} : x = (x_1, \dots, x_n) \rightarrow (\gamma^T x, x_1, x_2, \dots, x_i).$$

对 $\vartheta \in \mathbb{C}$, 将 $V_{\gamma,\vartheta}$ 记作由 $V \cap \pi_\gamma^{-1}(\vartheta)$ 定义的代数集. 考虑下面的定义:

- 关于指标 $\{i_1, \dots, i_r\} \subset \{1, \dots, p\}$ 以及指标 $1 \leq i \leq n - r - 1$ 的截断雅克比矩阵 $\text{Jac}([h_{i_1}, \dots, h_{i_r}, \gamma^T X], \mathbf{X}_{>i})$

$$\begin{pmatrix} \frac{\partial h_{i_1}}{\partial x_{i+1}} & \frac{\partial h_{i_1}}{\partial x_{i+2}} & \cdots & \frac{\partial h_{i_1}}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial h_{i_r}}{\partial x_{i+1}} & \frac{\partial h_{i_r}}{\partial x_{i+2}} & \cdots & \frac{\partial h_{i_r}}{\partial x_n} \\ \gamma_{i+1} & \gamma_{i+2} & \cdots & \gamma_n \end{pmatrix}$$

的所有 $(r + 1)$ 阶子式的集合定义为 $\mathbf{M}(\mathbf{h}, \gamma, i)$. 为方便起见, 令 $\mathbf{M}(\mathbf{h}, \gamma, n - r) = \emptyset$. 当 γ 是参数 $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$ 时, 子式的集合记作 $\mathbf{M}(\mathbf{h}, \mathbf{c}, i)$.

- $\{1, \dots, p\}$ 中所有子集 $\{i_1, \dots, i_r\}$ 对应的雅克比矩阵

$$\begin{pmatrix} \frac{\partial h_{i_1}}{\partial x_1} & \frac{\partial h_{i_1}}{\partial x_2} & \cdots & \frac{\partial h_{i_1}}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial h_{i_r}}{\partial x_1} & \frac{\partial h_{i_r}}{\partial x_2} & \cdots & \frac{\partial h_{i_r}}{\partial x_n} \\ \gamma_1 & \gamma_2 & \cdots & \gamma_n \end{pmatrix}$$

的所有 $r + 1$ 阶子式的集合记作 $S(\mathbf{h}, \gamma)$. 当 γ 是参数 $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$ 时, 子式的集合记作 $S(\mathbf{h}, \mathbf{c})$.

令 $\gamma \in \mathbb{C}^n - \{\mathbf{0}\}$, $\vartheta \in \mathbb{C}$. 假设 $V_{\gamma,\vartheta}$ 光滑并且 $\langle \mathbf{h}, \gamma^T X - \vartheta \rangle$ 是等维根理想. 关于 $\pi_{\gamma,i}$ 和 $V_{\gamma,\vartheta}$ 的极仿射簇 $W(\mathbf{h}, \gamma, \vartheta, i)$ 定义为 π_i 在 $V_{\gamma,\vartheta}$ 上的关键点的集合, 即

$$W(\mathbf{h}, \gamma, \vartheta, i) = V(\mathbf{h}, \mathbf{M}(\mathbf{h}, \gamma, i), \gamma^T X - \vartheta), \quad 1 \leq i \leq n - r.$$

定义

$$W(\mathbf{h}, \gamma, i) := V(\mathbf{h}, \mathbf{M}(\mathbf{h}, \gamma, i)), \quad 1 \leq i \leq n - r.$$

关于 π_γ 和 V 的极仿射簇 $C(\mathbf{h}, \gamma)$ 定义为 π_γ 在 V 上的关键点的集合. 即

$$C(\mathbf{h}, \gamma) := V(\mathbf{h}, S(\mathbf{h}, \gamma)).$$

现在, 我们讨论极仿射簇在一般坐标系下的性质. 给定 $\mathbf{A} \in \mathrm{GL}_n(\mathbb{C})$, $S \subset \mathbb{C}^n$. $S^\mathbf{A}$ 记作映射 $x \rightarrow \mathbf{A}^{-1}x$ 在 S 上的像. $\mathbf{h}^\mathbf{A}$ 记作 $\{\mathbf{h}_1(\mathbf{A}x), \dots, \mathbf{h}_p(\mathbf{A}x)\}$.

令 $\mathbf{A} \in \mathrm{GL}_n(\mathbb{Q})$. 我们考虑 $V(\mathbf{P})$ 中满足下面条件的参数值 γ : 存在一个非空的Zariski开子集 $\mathcal{O} \subset \mathbb{C}^n$, 使得对任意的 $\vartheta \in \mathcal{O}$, 下面的条件成立:

\mathfrak{P}_1 : $\langle \mathbf{h}, \gamma^T X - \vartheta \rangle$ 是等维根理想并且 $V_{\gamma, \vartheta}$ 光滑. 这等价于 $(\mathbf{h}, \gamma^T X - \vartheta)$ 的雅克比矩阵在 $V_{\gamma, \vartheta}$ 的任意一点上的秩为 $r + 1$.

$\mathfrak{P}_2(\mathbf{A})$: 对整数 $i \in [1, n - r]$, 极仿射簇 $W(\mathbf{h}^\mathbf{A}, \gamma, \vartheta, i)$ 在 π_{i-1} 的诺特位置上.

如果参数值 γ 满足假设 \mathfrak{P}_1 和 $\mathfrak{P}_2(\mathbf{A})$, 则称参数值 γ 满足假设 $\mathfrak{P}(\mathbf{A})$.

命题 5.7. [37, 命题 4.2] 假设 $\mathbf{A} \in \mathrm{GL}_n(\mathbb{C})$. 如果参数值 $\gamma \in \mathbb{C}^n - \{\mathbf{0}\}$ 满足假设 $\mathfrak{P}(\mathbf{A})$, 则下面的命题成立:

- 集合

$$\cup_{i=1}^{n-r} ((W(\mathbf{h}^\mathbf{A}, \gamma, i) - C(\mathbf{h}^\mathbf{A}, \gamma)) \cap \pi_{i-1}^{-1}(\mathbf{0}))$$

的Zariski闭包 $\mathcal{C}_\gamma^\mathbf{A}$ 的维数至多为1.

- $\pi_\gamma(C(\mathbf{h}^\mathbf{A}, \gamma))$ 与 π_γ 在 $\mathcal{C}_\gamma^\mathbf{A}$ 上的非正则点的集合的并是有限的并且包含映射 π_γ 在 $\mathcal{C}_\gamma^\mathbf{A}$ 上的极大值.

由 [37, 命题 4.3]可知, 对任意的 $\gamma \in V(\mathbf{P})$, 存在一个非空的Zariski开集 $\mathcal{A} \subset \mathrm{GL}_n(\mathbb{C})$, 使得对任意的 $\mathbf{A} \in \mathcal{A} \cap \mathrm{GL}_n(\mathbb{Q})$, 假设 $\mathfrak{P}(\mathbf{A})$ 成立.

在求解问题3的第二个子问题时, 需要证明一个更强的结论: 存在非空的Zariski开子集 $\mathcal{A} \subset \mathrm{GL}_n(\mathbb{C})$, 使得对任意的 $\mathbf{A} \in \mathcal{A} \cap \mathrm{GL}_n(\mathbb{Q})$, 存在 $V(\mathbf{P})$ 的一个Zariski稠密子集 U , 使得对任意的 $\gamma \in U$, 假设 $\mathfrak{P}(\mathbf{A})$ 成立.

函数BasicParametricOptimization返回多项式向量 $(Z, \Phi(\mathbf{c}_0, \mathbf{c}))$, 使得 $V(\mathbf{P}) - V(Z) \neq \emptyset$ 并且将 γ 限制在 $V(\mathbf{P}) - V(Z)$ 上时, $\Phi(\mathbf{c}_0, \gamma)$ 定义了在命题5.7的第二个结论中, $\pi_\gamma(C(\mathbf{h}^\mathbf{A}, \gamma))$ 与 π_γ 在 $\mathcal{C}_\gamma^\mathbf{A}$ 上的非正则点的集合的并.

5.4.3 子函数

在这一节, 假设 $\mathbf{h} = (h_1, \dots, h_p) \subset \mathbb{Q}[X_1, \dots, X_n]$ 生成一个等维根理想, 并且 V 为形如(5.2)的代数集. 多项式向量 $\mathbf{P} \subset \mathbb{Q}[\mathbf{c}]$ 表示一个素理想 \mathcal{P} 的生成元. 我们将 \mathbb{A} 记作整环 $\mathbb{Q}[\mathbf{c}]/\mathcal{P}$, 将 \mathbb{K} 记作 \mathbb{A} 的分式域并且将 $\bar{\mathbb{K}}$ 记作 \mathbb{K} 的代数闭包. 我们考虑环 $\mathbb{A}[X_1, \dots, X_n]$ 中的多项式系统.

基本函数 现在, 考虑下面的函数:

函数 `SingularMinors` 将 \mathbf{h}, \mathbf{P} 作为输入并且返回 $\tilde{\mathbf{G}} = (\mathbf{h}, \mathbf{S}(\mathbf{h}, \mathbf{c}))$.

函数 `SpecialCurve` 将 \mathbf{h}, \mathbf{P} 作为输入并且返回 $\tilde{\mathbf{F}} = (\tilde{\mathbf{F}}_1, \dots, \tilde{\mathbf{F}}_{n-r})$. $\tilde{\mathbf{F}}_1 := \{\mathbf{h}, \mathbf{M}(\mathbf{h}, \mathbf{c}, 1)\}$. $\tilde{\mathbf{F}}_i := \{\mathbf{h}, \mathbf{M}(\mathbf{h}, \mathbf{c}, i), X_1, \dots, X_{i-1}\}$, $2 \leq i \leq n-r$.

函数 `PointsPerComponents` 将 \mathbf{h} 作为输入并且返回 $V = V(\mathbf{h})$ 的一个零维子仿射簇. 零维子仿射簇与 $V \cap \mathbb{R}^n$ 的所有连通分支都相交.

函数 `ValuesTakenByPoly` 将 \mathbb{C}^n 中的零维仿射簇 V 的定义多项式 \mathbf{Q} 和 \mathbf{P} 作为输入并且返回多项式向量 $\Phi \subset \mathbb{Q}[\mathbf{c}_0, \mathbf{c}]$ 和多项式 $Z \in \mathbb{Q}[\mathbf{c}] - \langle \mathbf{P} \rangle$, 使得对 $\gamma \in V(\mathbf{P}) - V(Z)$, $\Phi(\mathbf{c}_0, \gamma)$ 定义集合 $\{\gamma^T x \mid x \in V\}$.

函数 `ParametricValuesTakenByPoly` 将域 $\bar{\mathbb{K}}$ 中的零维仿射簇 V 的定义多项式向量 \mathbf{Q} 和多项式向量 $\mathbf{P} \subset \mathbb{Q}[\mathbf{c}]$ 作为输入. 它返回多项式向量 $\Phi \subset \mathbb{Q}[\mathbf{c}_0, \mathbf{c}]$ 和多项式 $Z \in \mathbb{Q}[\mathbf{c}] - \langle \mathbf{P} \rangle$, 使得对 $\gamma \in V(\mathbf{P}) - V(Z)$, $\Phi(\mathbf{c}_0, \gamma)$ 定义集合 $\{\gamma^T x \mid x \in V\}$.

根据函数 `ParametricValuesTakenByPoly` 的运行原理, 下面的引理是显然的.

引理 5.8. 令 (Φ, Z) 为 `ParametricValuesTakenByPoly`(\mathbf{Q}, \mathbf{P}) 的输出, 则 $Z \notin \langle \mathbf{P} \rangle$.

函数 `CheckProperness` 将 $\mathbf{h}, \mathbf{A} \in \mathrm{GL}_n(\mathbb{Q}), \mathbf{P} \subset \mathbb{Q}[\mathbf{c}]$ 作为输入. 当 $V(\mathbf{P})$ 中不存在使得 $\mathfrak{P}(\mathbf{A})$ 成立的参数值时, 函数 `CheckProperness` 返回 (0) . 否则它返回 $Z \in \mathbb{Q}[\mathbf{c}] - \langle \mathbf{P} \rangle$, 使得对任意的 $\gamma \in V(\mathbf{P}) - V(Z)$, 假设 $\mathfrak{P}(\mathbf{A})$ 成立. 这个函数能够刻画使得假设 $\mathfrak{P}(\mathbf{A})$ 成立的参数值 γ .

引理 5.9. 使用上面的记号和假设. 存在一个非空的 Zariski 开集 $\mathcal{A} \subset \mathrm{GL}_n(\mathbb{C})$, 使得对任意的 $\mathbf{A} \in \mathcal{A} \cap \mathrm{GL}_n(\mathbb{Q})$, 下面的结论成立:

令 Z 为 `CheckProperness`($\mathbf{h}, \mathbf{A}, \mathbf{P}$) 的输出. 则 $V(\mathbf{P}) - V(Z)$ 是 $V(\mathbf{P})$ 的 Zariski 稠密子集, 并且对 $\gamma \in V(\mathbf{P}) - V(Z)$, 性质 $\mathfrak{P}(\mathbf{A})$ 成立.

证明. 由于 $Z \notin \langle \mathbf{P} \rangle$, 并且 $\langle \mathbf{P} \rangle$ 是素理想, $V(\mathbf{P}) - V(Z)$ 在 $V(\mathbf{P})$ 中是 Zariski 稠密的.

只需证明存在一个 Zariski 开集 $\mathcal{A} \subset \mathrm{GL}_n(\mathbb{C})$, 使得对任意的 $\mathbf{A} \in \mathcal{A} \cap \mathrm{GL}_n(\mathbb{Q})$, 存在多项式 Z , 只要 $\gamma \in V(\mathbf{P}) - V(Z)$, 性质 $\mathfrak{P}(\mathbf{A})$ 成立.

通过 [37, 命题 4.3], 对任意的 $\gamma \in V(\mathbf{P})$, 存在一个非空的Zariski开集 $\mathcal{A} \in GL_n(\mathbb{C})$, 使得假设 $\mathfrak{P}(\mathbf{A})$ 对 γ 成立. 我们现在证明存在 $Z \notin \langle \mathbf{P} \rangle$, 使得对任意的 $\mathbf{A} \in \mathcal{A}$ 和 $\gamma' \in V(\mathbf{P}) - V(Z)$, γ' 满足假设 $\mathfrak{P}(\mathbf{A})$.

以 $\mathbb{K}(\mathbf{c}_0)$ 为定义(background)域. 由理想 $\langle \mathbf{h}, \mathbf{c}^T X - \mathbf{c}_0 \rangle$ 的定义多项式和 $Jac(\mathbf{h}, \mathbf{c}^T X - \mathbf{c}_0)$ 的 $r + 1$ 阶子式所生成的理想 Gröbner 基 G 为 $\{1\}$. 实际上, 如果结论不成立, 令 Z' 为出现在 G 的计算中所有分母的乘积, 则对 $\gamma \in V(\mathbf{P}) - V(Z')$, 假设 \mathfrak{P}_1 不成立, 矛盾. 因此 $G = \{1\}$.

给定一个 $n \times n$ 矩阵 \mathfrak{A} , 其未定元为 $\mathfrak{A}_{i,j}$. 文献 [82] 指出可以通过判定参数值使下面的多项式

$$\mathbf{h}^{\mathfrak{A}}, M(\mathbf{h}^{\mathfrak{A}}, \mathbf{c}, i), \mathbf{c}^T \mathfrak{A} X - \mathbf{c}_0.$$

在定义域 $\mathbb{K}(\mathfrak{A}_{i,j})$ 中生成理想的极小 Gröbner 基的分母非零来验证参数在诺特位置. 我们通过使得计算过程中出现在分母位置的多项式关于 \mathbf{c}, \mathbf{c}_0 的系数不为零来定义 Zariski 开集 $\mathcal{A} \subset GL_n(\mathbb{C})$. 这个集合是非空的. 否则, 可得任意的 $\gamma \in V(\mathbf{P})$ 均满足假设 $\mathfrak{P}(\mathbf{A})$, 导出矛盾.

对 $\mathbf{A} \in \mathcal{A}$, 将 Z'' 定义为由下面的多项式

$$\mathbf{h}^{\mathbf{A}}, M(\mathbf{h}^{\mathbf{A}}, \mathbf{c}, i), \mathbf{c}^T \mathbf{A} X - \mathbf{c}_0$$

在定义域 \mathbb{K} 中生成理想的极小 Gröbner 基的计算过程中出现的所有分母. 取 $Z = Z' Z''$, 证毕. \square

考虑 $\mathbb{A}[X_1, \dots, X_n]$ 中的多项式向量 $\mathbf{F} = (f_1, \dots, f_s)$ 和 $\mathbf{G} = (g_1, \dots, g_k)$. $I \subset \mathbb{K}[X_1, \dots, X_n]$ 为饱和理想 $\langle \mathbf{F} \rangle : \langle \mathbf{G} \rangle^\infty$. 假设 I 是 1 维的等维理想. 令 $\mathbf{F}_\gamma \subset \mathbb{Q}[X]$ 为 \mathbf{F} 中的多项式对未定元 \mathbf{c} 赋值为 γ 之后构成的多项式向量.

函数 `ParametricCurveRepresentation` 将 \mathbf{F} , \mathbf{G} 和 \mathbf{P} 作为输入, 将理想 $\langle \mathbf{F} \rangle : \langle \mathbf{G} \rangle^\infty$ 和 $\mathbb{Q}[\mathbf{c}] - \mathcal{P}$ 中的多项式 Z 作为输出. 对任意的 $\gamma = (\gamma_1, \dots, \gamma_n) \in V(\mathbf{P}) - V(Z)$, $\langle \mathbf{F} \rangle : \langle \mathbf{G} \rangle^\infty$ 的生成元对未定元 \mathbf{c} 赋值为 $\gamma = (\gamma_1, \dots, \gamma_n)$ 之后是 $\langle \mathbf{F}_\gamma \rangle : \langle \mathbf{G} \rangle^\infty$ 的生成元.

引理 5.10. 令 (\mathbf{Q}, Z) 为 `ParametricCurveRepresentation`($\mathbf{F}, \mathbf{G}, \mathbf{P}$) 的输出. $\langle \mathbf{P} \rangle + \langle Z \rangle$ 的维数小于 $\langle \mathbf{P} \rangle$ 的维数.

证明. 不失一般性, 假设在一般的(generic)坐标系下. 通过计算定义域 \mathbb{K} 中的 Gröbner 基和利用 $\mathbb{K}(X_1)[X_2, \dots, X_n]$ 中的线性代数知识, 可以计算 I 的有理参数化(参见 [9, 19]). 在计算过程中, 一些模 \mathcal{P} 不为 0 多项式可以作多项式除法中的除因子. 这些多项式的乘积 Z 是一个有效的输出并且模 \mathcal{P} 不为 0, 因此 Z 模 \mathcal{P} 不为 0. 由于 \mathcal{P} 是素理想, $\mathcal{P} + \langle Z \rangle$ 的维数小于 \mathcal{P} 的维数. \square

函数 `UnionParametricCurve` 以多项式向量 $\mathbf{F} = (\mathbf{F}_1, \dots, \mathbf{F}_l)$, \mathbf{G} 以及 \mathbf{P} 为输入, 其中理想 $\langle \mathbf{F}_k \rangle : \langle \mathbf{G} \rangle^\infty \subset \mathbb{K}[X_1, \dots, X_n]$ 的维数是 1 并且对 $1 \leq k \leq l$, $\langle \mathbf{F}_k \rangle :$

$\langle \mathbf{G} \rangle^\infty$ 是等维的. 它返回理想 $\bigcap_{k=1}^l \langle \mathbf{F}_k \rangle : \langle \mathbf{G} \rangle^\infty$ 的生成元和多项式 $Z \in \mathbb{Q}[\mathbf{c}] - \mathcal{P}$, 使得对任意的 $\gamma = (\gamma_1, \dots, \gamma_n) \in V(\mathbf{P}) - V(Z)$, $\bigcap_{k=1}^l \langle \mathbf{F}_k \rangle : \langle \mathbf{G} \rangle^\infty$ 的生成元将未定元 \mathbf{c} 赋值为 $\gamma = (\gamma_1, \dots, \gamma_n)$ 之后是 $\bigcap_{k=1}^l \langle \mathbf{F}_{k\gamma} \rangle : \langle \mathbf{G} \rangle^\infty$ 的生成元, 其中 $\mathbf{F}_{k\gamma} \subset \mathbb{Q}[X]$ 为 \mathbf{F}_k 中的多项式对未定元 \mathbf{c} 赋值为 γ 之后构成的多项式向量. 下面的引理是引理 5.10 的直接推论.

引理 5.11. 令 (\mathbf{Q}, Z) 为 $\text{UnionCurveParametric}(\mathbf{F}, \mathbf{G}, \mathbf{P})$ 的输出. 则 $\langle \mathbf{P} \rangle + \langle Z \rangle$ 的维数小于 $\langle \mathbf{P} \rangle$ 的维数.

函数 $\text{ParametricIntersection}$ 以曲线 $C \subset \bar{\mathbb{K}}^n$ 的定义多项式向量 \mathbf{Q} , 多项式向量 $\mathbf{G} \in \mathbb{A}[X_1, \dots, X_n]$ 和 \mathbf{P} 为输入, 其中向量 \mathbf{G} 定义了 $\bar{\mathbb{K}}^n$ 中的代数集 H . 假设 $\dim C \cap H = 0$, 根据文献 [28], 函数返回 $C \cap H$ 的生成元和一个多项式 $Z \notin \langle \mathbf{P} \rangle$, 使得对任意的 $\gamma \in V(\mathbf{P}) - V(Z)$, $C \cap H$ 的生成元可以特定化为 $C_\gamma \cap H_\gamma$ 的生成元, 其中 C_γ, H_γ 分别定义为 C, H 对未定元 \mathbf{c} 赋值为 γ 之后构成的仿射簇.

引理 5.12. 令 (Φ, Z) 为 $\text{ParametricIntersection}(\mathbf{Q}, \mathbf{G}, \mathbf{P})$ 的输出. 则 $\dim(\langle \mathbf{P} \rangle + \langle Z \rangle) < \dim(\langle \mathbf{P} \rangle)$.

计算非正则点的集合 假设 \mathbf{Q} 定义了代数曲线 $C_1 \subset \bar{\mathbb{K}}^n$. 存在 $V(\mathbf{P})$ 的 Zariski 稠密子集 U , 使得对 $\gamma \in U$, $V(\mathbf{Q}_\gamma)$ 定义代数曲线 C_γ .

函数 $\text{ParametricSetOfNonProperness}$ 以 $\mathbf{c}^T x, \mathbf{Q}$ 和 \mathbf{P} 为输入, 并且返回多项式向量 (Φ, Z) , 使得 $V(\mathbf{P}) - V(Z) \neq \emptyset$, $V(\mathbf{P}) - V(Z) \subset U$ 并且对 $\gamma \in V(\mathbf{P}) - V(Z)$, $\Phi(\mathbf{c}_0, \gamma) \neq 0$ 并且它的根包含映射 $x \rightarrow \gamma^T x$ 在 C_γ 上的非正则点.

将 C_2 记作代数曲线 $\{(\gamma_0, x) \mid x \in C_1, \gamma_0 = \mathbf{c}^T x\}$, 将 \mathfrak{C}_1 记作 C_1 在 $\mathbb{P}^n(\bar{\mathbb{K}})$ 中的射影闭包. 对 $x = (x_0 : x_1 : \dots : x_n) \in \mathfrak{C}_1$ 和 $x_0 \neq 0$, 将 \tilde{x} 记作点 $\left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right)$ 并且将 \mathfrak{C}_2 记作集合 $\{(\gamma_0, x) \mid x \in \mathfrak{C}_1, \mathbf{c}^T \tilde{x} = \gamma_0\}$.

根据文献 [2] 的算法, 函数 $\text{ParametricSetOfNonProperness}$ 包含下面的步骤:

- 计算射影闭包 \mathfrak{C}_2 .
- 计算 \mathfrak{C}_2 与无穷远超平面 $X_0 = 0$ 的交集.

上述步骤通过计算定义域 \mathbb{K} 中的 Gröbner 基来实现. 计算过程中产生的所有分母记作 Z . 显然可得 $Z \notin \langle \mathbf{P} \rangle$.

引理 5.13. 令 (Φ, Z) 为 $\text{ParametricSetOfNonProperness}(\mathbf{Q}, \mathbf{P})$ 的输出. 则 $\dim(\langle \mathbf{P} \rangle + \langle Z \rangle) < \dim(\langle \mathbf{P} \rangle)$.

5.4.4 参数化优化问题算法的基本步骤

现在, 考虑函数BasicParametricOptimization. 它是文献 [37] 中算法SetContainingLocalExtrema的参数化形式.

算法将求解多项式函数在 $V \cap \mathbb{R}^n$ 上的最优化问题转化成求解相同的多项式函数在一条曲线上的最优化问题. 这两个优化问题的最优化相同. 我们将原算法BasicParametricOptimization中的函数推广到参数化形式.

输入: $\mathbf{h} = (h_1, \dots, h_p) \subset \mathbb{Q}[X_1, \dots, X_n]$ 和 $\mathbf{P} \subset \mathbb{Q}[\mathbf{c}]$

性质: \mathbf{P} 生成一个素理想并且 $V(\mathbf{h})$ 是一个光滑的等维仿射簇.

输出: 多项式向量 (Φ, Z) , 使得

- $\Phi \in \mathbb{Q}[\mathbf{c}_0, \mathbf{c}]$ 并且 $Z \in \mathbb{Q}[\mathbf{c}]$;
- Z 模 $\langle \mathbf{P} \rangle$ 不为 0;
- 对任意的 $\gamma \in V(\mathbf{P}) - V(Z)$, 有 $\Phi(\mathbf{c}_0, \gamma) \neq 0$ 并且它的根包含函数 π_γ 在 $V \cap \mathbb{R}^n$ 上的最优化.

BasicParametricOptimization(\mathbf{h}, \mathbf{P})

1. $R = \text{PointsPerComponents}(\mathbf{h})$
2. $(\Phi_0, Z_0) = \text{ValuesTakenByPoly}(R, \mathbf{c}^T X, \mathbf{P})$
3. 随机选取 $\mathbf{A} \in \text{GL}_n(\mathbb{C})$
4. $Z'_0 = \text{CheckProperness}(\mathbf{h}, \mathbf{A}, \mathbf{P})$
5. $\tilde{\mathbf{F}} = \text{SpecialCurve}(\mathbf{h}^\mathbf{A}, \mathbf{P})$, $\tilde{\mathbf{G}} = \text{SingularMinors}(\mathbf{h}^\mathbf{A}, \mathbf{P})$
6. $(R_1, Z_1) = \text{UnionParametricCurve}(\tilde{\mathbf{F}}, \tilde{\mathbf{G}}, \mathbf{P})$
7. $(\Phi_1, Z'_1) = \text{ParametricSetofNonProperness}(R_1, \mathbf{P}, \mathbf{c}^T X)$
8. $R_2 = \text{ParametricIntersection}(R_1, \tilde{\mathbf{G}}, \mathbf{P})$
9. $(\Phi_2, Z_2) = \text{ParametricValuesTakesByPoly}(R_2, \mathbf{c}^T X, \mathbf{P})$
10. 定义 $Z = Z_0 Z'_0 Z_1 Z'_1 Z_2$ 并且 $\Phi = \Phi_0 \Phi_1 \Phi_2$

11. 返回 (Φ, Z)

定理 5.14. 令 (Φ, Z) 为BasicParametricOptimization(\mathbf{h}, \mathbf{P})的输出. 则 $Z \notin \langle \mathbf{P} \rangle$ 并且它的输出是正确的.

证明. 由于 $\langle \mathbf{P} \rangle$ 是素理想并且通过引理5.8, 5.9, 5.11, 5.12和5.13, Z 的因式 Z_0, Z'_0, Z_1, Z'_1 和 Z_2 不属于 $\langle \mathbf{P} \rangle$. 因此, $Z \notin \langle \mathbf{P} \rangle$.

只需证明算法输出的正确性. 由于 \mathbf{A} 随机选择, 我们假设 \mathbf{A} 属于引理5.9中的非空Zariski开集.

对任意的 $\gamma \in V(\mathbf{P}) - V(Z)$, 通过引理5.9, 假设 $\mathfrak{P}(\mathbf{A})$ 成立. 因此不失一般性, 假设文献 [37] 中的算法SetContainingLocalExtrema利用函数BasicParametric Optimization中选取的矩阵 \mathbf{A} 运行. 函数SetContainingLocalExtrema对输入 $\gamma^T X$ 和 \mathbf{h} 返回一个多项式 $\varphi \in \mathbb{Q}[\mathbf{c}_0]$. φ 的根包含映射 $x \rightarrow \gamma^T x$ 限制在 $V \cap \mathbb{R}^n$ 中的最优值. 由引理5.8, 5.11, 5.12和5.13可知多项式 φ 等于 $\Phi(\mathbf{c}_0, \gamma)$. 因此, 算法SetContainingLocalExtrema [37, 命题 4.2] 的正确性可以证明函数BasicParametricOptimization的正确性. \square

5.4.5 坏参数的算法

现在我们列出用于求解问题3的算法ParametricOptimizationRec. 它包含函数BasicParametricOptimization以及将多项式向量 $\mathbf{P} \subset \mathbb{Q}[\mathbf{c}]$ 和多项式 $Z \in \mathbb{Q}[\mathbf{c}]$ 作为输入的函数PrimeDecomposition. PrimeDecomposition返回多项式向量序列 $\mathbf{P}_1, \dots, \mathbf{P}_k$, 使得

$$\sqrt{\langle \mathbf{P} \rangle + \langle Z \rangle} = \cap_{i=1}^k \langle \mathbf{P}_i \rangle$$

并且对整数 $i \in [1, k]$, $\langle \mathbf{P}_i \rangle$ 是素理想.

假设对偶仿射簇的坏参数由多项式向量 \mathbf{P}_0 定义. 使用准素分解, 不妨设 $\langle \mathbf{P}_0 \rangle$ 是素理想. 下面是算法ParametricOptimizationRec的结构.

算法 5.3. 求解关于对偶仿射簇的坏参数 $\mathbf{P}_0 \subset \mathbb{Q}[\mathbf{c}]$ 的参数化最优化问题.

输入: \mathbf{h}, \mathbf{P}_0 .

输出: 多项式序列 $(\Phi_i, Z_i, \mathbf{P}_i)$, 其中 $V(\mathbf{P}_0) = \cup_{i=1}^k V(\mathbf{P}_i)$. 给定 $\gamma \in \mathbb{C}^n$, 如果 $\gamma \in V(\mathbf{P}_i) - V(Z_i)$, 映射 $x \rightarrow \gamma^T x$ 的最优值是非零多项式 $\Phi_i(\mathbf{c}_0, \gamma)$ 的根.

步骤 1 如果 $\langle \mathbf{P}_0 \rangle = \langle 1 \rangle$, 则返回 $[]$.

步骤 2 调用函数 $(\Phi, Z) = \text{BasicParametricOptimization}(\mathbf{h}, \mathbf{P}_0)$.

步骤 3 通过函数 $(\mathbf{P}_1, \dots, \mathbf{P}_k) = \text{PrimeDecomposition}(\mathbf{P}_0, Z)$ 对理想 $\langle \mathbf{P}_0, Z \rangle$ 作准素分解.

步骤 4 对整数 $i \in [1, k]$, 递归运行算法 $\mathbf{L}_i = \text{ParametricOptimizationRec}(\mathbf{h}, \mathbf{P}_i)$.

步骤 5 返回 (Φ, Z, \mathbf{P}_0) 与 $\mathbf{L}_1, \dots, \mathbf{L}_i$ 的并.

定理 5.15. 算法ParametricOptimizationRec有限步终止并且是正确的.

证明. 正确性直接根据算法的递归方式以及BasicParametricOptimization的正确性归纳(见定理5.14).

我们现在证明终止性. 再次使用定理5.14, 由于 \mathbf{P}_0 是素理想并且 $Z \notin \langle \mathbf{P}_0 \rangle$, 算法5.3的步骤2返回满足 $\dim(\langle \mathbf{P}_0 \rangle + \langle Z \rangle) < \dim(\mathbf{P}_0)$ 的多项式 Z , 当每次递归步骤完成时, 维数严格下降. \square

例子5.3(续) 由于 $\langle \mathbf{c}_2\mathbf{c}_4, \mathbf{c}_3\mathbf{c}_4, \mathbf{c}_3\mathbf{c}_2\mathbf{c}_1 \rangle$ 表示算法5.1中的坏参数, 在每个递归步骤中, 考虑它的素分支 $\langle \mathbf{c}_1, \mathbf{c}_4 \rangle, \langle \mathbf{c}_2, \mathbf{c}_4 \rangle, \langle \mathbf{c}_3, \mathbf{c}_4 \rangle$ 和 $\langle \mathbf{c}_2, \mathbf{c}_3 \rangle$. 我们只列出对素分支 $\mathbf{P} = \langle \mathbf{c}_2, \mathbf{c}_3 \rangle$ 运行递归步骤所得到的结果, 其它的类似讨论. 特别对子函数BasicParametricOptimization. 通过非正则点集的计算, Φ_1 的无平方部分和 $Z_1Z'_1$ 分别为 $\mathbf{c}_0\mathbf{c}_1\mathbf{c}_4$ 和 $\mathbf{c}_1\mathbf{c}_4$. 我们需要对 $\mathbf{P}_1 = \langle \mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3 \rangle$ 和 $\mathbf{P}_2 = \langle \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4 \rangle$ 分别进行递归计算. 在 \mathbf{P}_1 中, 目标函数是 X_4 , 在 \mathbf{P}_2 中目标函数是 X_1 . 在 \mathbf{P}_1 情形下, 问题被约化成目标函数为 X_4 的非参数优化问题. 运行 [37]中的算法, 返回表示最优值的多项式 $\Phi_1 = \mathbf{c}_0$. CAD方法不能处理这个例子.

第六章 凸集的锥举起

6.1 介绍

线性规划问题在优化问题中占有重要的地位. 在线性规划问题中, 通过减少可行集的定义表达式中约束和变量的个数可以降低计算复杂度. 例如, 考虑下面的多面体:

$$C_n := \{x \in \mathbb{R}^n \mid \pm x_1 \pm x_2 \pm \cdots \pm x_n \leq 1\}.$$

在 C_n 的表达式中有 n 个变量和 2^n 个线性约束. 当 n 增加时, 表达式的个数会以指数级的速度增加. 如果在优化问题中利用这个表达式来表示 C_n , 那么计算复杂度关于变量的个数是指数级的. 因此只能处理小规模的优化问题.

现在考虑 C_n 的另一个表达式:

$$C_n := \{x \in \mathbb{R}^n \mid \text{存在 } y_1, \dots, y_n, \text{ 使得 } \sum_{i=1}^n y_i = 1, -y_i \leq x_i \leq y_i, i = 1, \dots, n\}.$$

C_n 的第二个表达式增加了 n 个变量, 但是线性约束的个数是 $2n$. 对同一个优化问题, 当 n 充分大时, 在优化问题中使用第二个表达式可以节省计算量. 在文献 [101] 中, Yannakakis 将有界多面体的最简表示问题转换成松弛矩阵的非负分解问题. 文献 [31] 推广了 [101] 的结果并且考虑将凸体的最简表示问题转换成锥举起问题. 在这一章, 我们考虑一般闭凸集的锥举起问题.

问题4. 给定一个闭凸集 $C \subset \mathbb{R}^n$.

- 给定一个满维的闭凸锥 $K \subset \mathbb{R}^m$, 是否存在一个仿射空间 $L \subset \mathbb{R}^m$ 和一个线性映射 $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^n$, 使得 $C = \pi(K \cap L)$?
- 给定一个锥序列 (K^k) (例如 (\mathbb{R}_+^k) , (S_+^k)), 如何求解最小的 k , 使得存在一个仿射空间 $L \subset \mathbb{R}^m$ 和一个线性映射 $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^n$, 满足 $C = \pi(K^k \cap L)$?

对一般的闭凸集, 由于锥举起和松弛算子的定义已经不再适用, 文献 [31] 和 [101] 的结果不成立.

假设 C 是 \mathbb{R}^n 中的闭凸集并且 K 是 \mathbb{R}^m 中的闭凸锥. 我们将文献 [31] 中凸体的锥分解定理推广到一般凸集的情形. 我们的主要结果如下:

- 当 C 包含原点时, 由于 $C = C^{oo}$, C 可以由 C^o 中的向量刻画. 当原点不是 C 的内点时, 我们指出凸集 C 可以被 $C^o = \{l \mid \text{对任意的 } x \in C, \langle l, x \rangle \leq 1\}$, $0^+C^o = \{l \mid \text{对任意的 } x \in C, \langle l, x \rangle \leq 0\}$ 和 $C_3 = \{l \mid \text{对任意的 } x \in C, \langle l, x \rangle \leq -1\}$ 对应的线性函数刻画.
- 我们将判定一个凸体 C 是否为一个凸锥的仿射切片在线性映射下的像推广到 C 无界并且原点不是 C 的内点以及 C 包含直线的情形. 我们介绍两个方法来刻画闭凸集 C 的锥举起. 第一个方法是基于 C 的极集中的所有点, 第二个方法基于 C 的极集的极点和极方向. 如果 C 包含直线, 第二个方法还要利用 C 的线性空间的正交基. 尽管第一个方法可以验证任意一个凸集的锥举起的存在性, 它很难被使用, 见注6.3. 因此, 我们主要研究第二个方法. 我们将文献 [31] 中的定义1, 2 和定理1 推广到一般凸集的情形并且说明一个凸集的推广锥举起可以由相应的松弛算子所刻画. 松弛算子根据凸集是否满维, 是否是一个锥的平移, 是否包含直线分别定义.
- 我们考虑多面体的锥举起并且指出当 C 和 K 是多面体时, 结论能够被加强. 当 K 是一个半正定锥时, 我们给出了多面体的半正定秩的下界, 从而推广了文献 [32] 中的结果. 我们将文献 [27, 29, 32] 中关于判别一个有界多面体的松弛矩阵的非负秩问题推广到无界的情形并且利用无界多面体的维数来刻画它的松弛矩阵的秩.

6.2 预备知识

令 \mathcal{S}^m 为 $m \times m$ 实对称矩阵空间. 令 $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ 为 \mathbb{R}^n 中的两个向量, 则 x, y 在 \mathbb{R}^n 中的内积表示成 $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$. 如果一个闭凸集 C 有界并且原点是 C 的内点, 则称 C 是一个凸体. 令 \mathbb{R}_+^m 为 m 维非负象限锥并且 \mathcal{S}_+^m 为实对称 $m \times m$ 半正定(psd)矩阵锥.

给定一个集合 C , 如果存在一个锥 C_0 和一个向量 $x \in \mathbb{R}^n$, 使得 $C = x + C_0$, 则称 C 是平移锥(锥的平移).

定义 6.1. [14] 如果闭凸锥 K 的所有面 F 满足 $K^* + F^\perp$ 是闭的条件, 则称闭凸锥 K 是好锥.

在文献 [67] 中, Pataki指出 K 是好锥当且仅当对 K 中的所有面 F , $F^* = K^* + F^\perp$. 很多凸锥都是好的, 例如多面体锥, 二阶锥和实对称半正定矩阵锥.

$\delta(\cdot, C)$ 的示性函数定义为

$$\delta(x, C) = \begin{cases} 0 & x \in C, \\ +\infty & x \notin C. \end{cases}$$

C 的支撑函数 $\delta^*(x, C)$ 定义为

$$\delta^*(x, C) = \sup\{\langle x, y \rangle \mid y \in C\}.$$

$\text{dom}\delta^*(x, C) = \{x \mid \delta^*(x, C) < +\infty\}$ 称为 C 的障碍锥.

定理 6.1. [76, 定理 8.7] 令 f 为一个正常的闭凸函数. 则对 $\alpha \in \mathbb{R}$, 集合 $\{x \mid f(x) \leq \alpha\}$ 的所有非空水平集具有相同的回收锥和线性空间.

推论 6.2. [76, 推论 14.2.1] 一个非空闭凸集 C 的障碍锥的极集等于 C 的回收锥.

定理 6.3. [76, 定理 13.1] 假设 C 是一个凸集. 则 $x \in \text{cl}(C)$ 当且仅当对任意的向量 x^* , 可得 $\langle x, x^* \rangle \leq \delta^*(x, C)$.

6.3 一个无界凸集的锥举起

闭凸集 C 的极点的集合记作 $\text{ext}_1(C)$. 如果 C 是闭凸锥, C 的极射线的集合记作 $\text{ext}_2(C)$. 在这一章, 我们用向量来表示极射线并且将 $\text{ext}_2(C)$ 记作这些向量的集合.

如果 $C \subset \mathbb{R}^n$ 是一个凸体, 根据 [31, 定义 1], C 的一个 K -举起定义为集合 $Q = K \cap L$, 这里 $L \subset \mathbb{R}^m$ 是一个仿射子空间, $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ 是一个线性映射并且下面的等式成立:

$$C = \pi(K \cap L). \quad (6.1)$$

如果 L 与 K 的内点相交, 则称 Q 是 C 的一个正则 K -举起. 如果存在映射

$$A : \text{ext}_1(C) \rightarrow K, \quad B : \text{ext}_1(C^\circ) \rightarrow K^*$$

使得对所有的 $(x, y) \in \text{ext}_1(C) \times \text{ext}_1(C^\circ)$, 有 $S_C(x, y) = \langle A(x), B(y) \rangle$, 则称松弛算子 S_C 是 K -可分的(参见 [31, 定理 2]).

6.3.1 C 是满维的情形

假设 $C \subset \mathbb{R}^n$ 是一个满维闭凸集. 定义

$$C^\circ = \{x \mid \delta^*(x, C) \leq 1\}, \quad 0^+ C^\circ = \{x \mid \delta^*(x, C) \leq 0\}, \quad C_3 = \{x \mid \delta^*(x, C) \leq -1\}.$$

闭凸集 C^o , C_3 不包含直线并且 0^+C^o 是包含 C_3 的尖的闭凸锥. 令

$$D_1 = \text{ext}_1(C^o) \setminus 0, D_2 = \text{ext}_2(0^+C^o) \cap \{x \mid \delta^*(x, C) = 0\}, D_3 = \text{ext}_1(C_3).$$

通过定理6.1, 可得 $0^+C^o = 0^+C_3$. 令

$$D_{32} = \text{ext}_2(0^+C^o) \cap \{x \mid \delta^*(x, C) = -1\}.$$

显然可得 $D_{32} \subseteq D_3$ 并且 D_{32} 不一定等于 D_3 .

例 6.1. 给定有界闭凸集

$$C = \{(x, y) \mid x + y \geq 1, x + y \leq 3, y - x \geq -1, y - x \leq 1\}.$$

则

$$\begin{aligned} C^o &= \{(x, y) \mid 2x + y \leq 1, x + 2y \leq 1, x \leq 1, y \leq 1\}, \\ 0^+C^o &= \{(x, y) \mid x \leq 0, y \leq 0\}, \\ C_3 &= \{(x, y) \mid x \leq -1, y \leq -1\}. \end{aligned}$$

如图6.1所示. 此外, $D_{32} = \emptyset$ 并且

$$D_1 = \{(-1, 1), (\frac{1}{3}, \frac{1}{3}), (1, -1)\}, D_2 = \{(-1, 0), (0, -1)\}, D_3 = \{(-1, -1)\}.$$

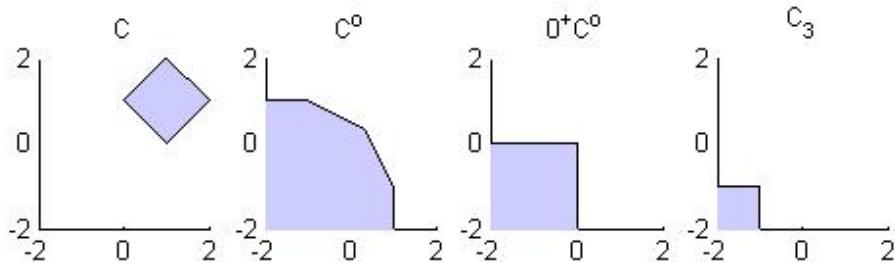


图 6.1: 例子6.1

注 6.1. 根据定理2.1和定理6.1, 可得

1. 凸集 C^o 可以表示成 D_1 中的向量和 D_2 , D_{32} 中向量方向的凸组合.
2. 凸锥 0^+C^o 可以表示成 D_2 , D_{32} 中向量方向的凸组合.
3. 凸集 C_3 可以表示成 D_3 中的向量与 D_2 , D_{32} 中向量方向的凸组合.

定理 6.4. 给定一个满维闭凸集 $C \subset \mathbb{R}^n$, 下面的陈述是正确的:

1. 集合 $D_1 = \emptyset$ 当且仅当 C° 是一个闭凸锥. 如果 $D_1 \neq \emptyset$, 则对任意的 $x \in D_1$, 可得 $\delta^*(x, C) = 1$.
2. 如果原点是 C 的内点, 集合 $D_2 = \emptyset$. 当 C 无界并且原点是 C 的边界点时, $D_2 \neq \emptyset$ 并且 0^+C° 中的每个极射线都是 D_2 中向量的方向.
3. 集合 $D_3 = \emptyset$ 当且仅当 C 包含原点. 如果 $D_3 \neq \emptyset$, 则对任意的 $x \in D_3$, 可得 $\delta^*(x, C) = -1$.

此外, C° 生成的凸锥是 $\text{dom}\delta^*(x, C)$.

证明. 由于 C° 不包含直线, $D_1 = \emptyset$ 当且仅当原点是 C° 唯一的极点, 即 C° 是一个闭凸锥. 如果存在 $x \in D_1$, 使得 $\delta^*(x, C) < 1$, 则存在 $\lambda > 0$, 使得 $\delta^*((1 + \lambda)x, C) \leq 1$ 并且 $\delta^*((1 - \lambda)x, C) \leq 1$. 因此 $(1 - \lambda)x$ 和 $(1 + \lambda)x$ 都在 C° 中. 这与 x 是 C° 的极点相矛盾.

当 C 包含原点作为内点时, C° 有界并且 $0^+C^\circ = \{0\}$. 因此 $D_2 = \emptyset$. 如果 C 包含原点, 对每一个 $x \in \text{dom}\delta^*(x, C)$, 可得 $\delta^*(x, C) \geq 0$. 如果原点是它的边界点, 则存在 C 在原点处的支撑超平面. 因此 C° 无界并且 0^+C° 包含一个非零向量. 由于对任意的 $x \in 0^+C^\circ$, 可得 $\delta^*(x, C) = 0$, D_2 可以表示 0^+C° 中的所有极射线.

显然可得 $D_3 = \emptyset$ 当且仅当对所有的 $y \in \mathbb{R}^n$, 可得 $\delta^*(y, C) \geq 0$. 通过定理 6.3, 这等价于 C 包含原点. 因此, $D_3 = \emptyset$ 当且仅当 C 包含原点. 相似的方法可以说明对任意的 $x \in D_3$, 可得 $\delta^*(x, C) = -1$.

对每一个 $x \in \text{cone}(C^\circ)$, 存在 $\lambda \geq 0$ 和 $y \in C^\circ$, 使得 $x = \lambda y$. 因此可得

$$\delta^*(x, C) = \lambda \delta^*(y, C) < \infty \quad \text{并且 } x \in \text{dom}\delta^*(x, C).$$

另一方面, 对每一个 $x \in \text{dom}\delta^*(x, C)$, 如果 $\delta^*(x, C) = M > 0$, 则 $x/M \in C^\circ$ 并且 $x \in \text{cone}(C^\circ)$. 因此, $\text{cone}(C^\circ) = \text{dom}\delta^*(x, C)$. \square

注 6.2. 当 C 不包含原点时, 很难验证 D_2 是空集. 例子 6.1 中的凸集 C 不包含原点, $D_2 = \{(-1, 0), (0, -1)\}$. 因此, 对凸集 $C = \{(x, y) \mid y \geq x + 1, y \geq -x + 1\}$, 可得

$$C^\circ = 0^+C^\circ = \{(x, y) \mid x + y \leq 0, y - x \leq 0\}.$$

0^+C° 的极射线为 $l_1 = (-1, -1)$ 和 $l_2 = (1, -1)$, 并且可得

$$\delta^*(l_1, C) = \delta^*(l_2, C) = -1 < 0.$$

因此, $D_2 = \emptyset$.

当原点是 C 的内点时, 通过定理 6.4, D_2, D_3 是空集并且 C 可以由 D_1 中的向量刻画. 然而, 当原点不是 C 的内点时, 对应于 D_1 或者 $D_1 \cup D_2$ 的线性函数不能完全刻画 C .

例子6.1(续)

对 $l_1 \in D_1$, 线性函数 $f(x) = \langle l_1, x \rangle$ 在 C 上的最大值是 1, 因此,

$$E_1 = \{(x, y) \mid c_1x + c_2y \leq 1, (c_1, c_2) \in D_1\} = \{(x, y) \mid -x + y \leq 1, x - y \leq 1, x + y \leq 3\}.$$

对 $l_2 \in D_2$, 线性函数 $f(x) = \langle l_2, x \rangle$ 在 C 上的最大值是 0, 因此,

$$E_2 = \{(x, y) \mid c_1x + c_2y \leq 0, (c_1, c_2) \in D_2\} = \{(x, y) \mid x \geq 0, y \geq 0\}.$$

对 $l_3 \in D_3$, 线性函数 $f(x) = \langle l_3, x \rangle$ 在 C 上的最大值是 -1, 因此,

$$E_3 = \{(x, y) \mid c_1x + c_2y \leq -1, (c_1, c_2) \in D_3\} = \{(x, y) \mid x + y \geq 1\}.$$

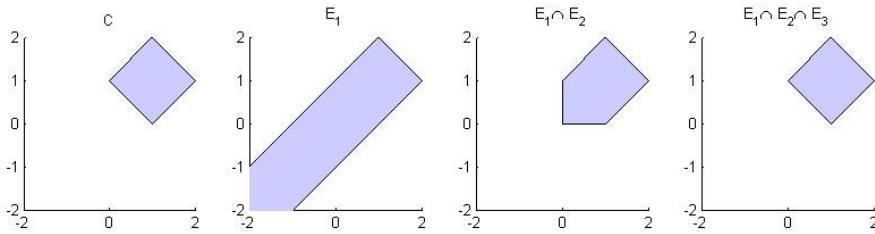


图 6.2: 例子6.1

下面的定理说明一个满维闭凸集 C 可以由 D_1, D_2 和 D_3 中的向量刻画.

定理 6.5. 假设 $C \subset \mathbb{R}^n$ 为一个满维闭凸集. 则

$$x \in C \iff \begin{cases} \langle l_1, x \rangle \leq 1 & l_1 \in D_1, \\ \langle l_2, x \rangle \leq 0 & l_2 \in D_2, \\ \langle l_3, x \rangle \leq -1 & l_3 \in D_3. \end{cases} \quad (6.2)$$

证明. 由于 C 是满维闭凸集, C° 不包含直线并且 D_1, D_2 和 D_3 都是良定义(well defined). 必要性是显然的. 另一方面, 假设 x 满足(6.2)右端的条件, 需要证明 $x \in C$. 通过定理6.3, 只需要证明对每一个 $l \in \text{dom}\delta^*(x, C)$, $\langle l, x \rangle \leq \delta^*(x, C)$. 令 $l_0 = \delta^*(l, C)$, 我们证明对下面的情形, 有 $\langle l, x \rangle \leq l_0$:

- 如果 $l_0 > 0$, 则 $l/l_0 \in C^\circ$. 通过注6.1, 存在 $\lambda_i^1 \geq 0, \lambda_j^2 \geq 0, \lambda_k^3 \geq 0$ 和 $x_i \in D_1, y_j \in D_2, z_k \in D_3$ 满足下面的等式:

$$l/l_0 = \sum_i \lambda_i^1 x_i + \sum_j \lambda_j^2 y_j + \sum_k \lambda_k^3 z_k, \quad \sum_i \lambda_i^1 = 1.$$

通过 D_1, D_2, D_3 的定义, 可得

$$\langle l/l_0, x \rangle = \sum_i \lambda_i^1 \langle x_i, x \rangle + \sum_j \lambda_j^2 \langle y_j, x \rangle + \sum_k \lambda_k^3 \langle z_k, x \rangle \leq \sum_i \lambda_i^1 = 1.$$

- 如果 $l_0 = 0$, 则 $l \in 0^+C^\circ$. 通过注6.1, 存在 $\lambda_j^2 \geq 0$, $\lambda_k^3 \geq 0$ 和 $y_j \in D_2$, $z_k \in D_{32}$ 满足下面的等式:

$$l = \sum_j \lambda_j^2 y_j + \sum_k \lambda_k^3 z_k.$$

因此, $\langle l, x \rangle = \sum_j \lambda_j^2 \langle y_j, x \rangle + \sum_k \lambda_k^3 \langle z_k, x \rangle \leq 0$.

- 如果 $l_0 < 0$, 则 $l/|l_0| \in C_3$. 通过注6.1, 存在 $\lambda_i^1 \geq 0$, $\lambda_j^2 \geq 0$, $\lambda_k^3 \geq 0$ 和 $x_i \in D_3$, $y_j \in D_2$, $z_k \in D_{32}$ 满足下面的等式:

$$l/|l_0| = \sum_i \lambda_i^1 x_i + \sum_j \lambda_j^2 y_j + \sum_k \lambda_k^3 z_k, \quad \sum_i \lambda_i^1 = 1.$$

因此, $\langle l/|l_0|, x \rangle = \sum_i \lambda_i^1 \langle x_i, x \rangle + \sum_j \lambda_j^2 \langle y_j, x \rangle + \sum_k \lambda_k^3 \langle z_k, x \rangle \leq \sum_i -\lambda_i^1 = -1$.

□

定理 6.6. 假设 $C \subset \mathbb{R}^n$ 是一个满维闭凸集. 如果存在 $x \in \mathbb{R}^n$, 使得对任意的 $l_1 \in D_1$, $1 - \langle l_1, x \rangle = 0$, 对任意的 $l_2 \in D_2$, $-\langle l_2, x \rangle = 0$, 对任意的 $l_3 \in D_3$, $-1 - \langle l_3, x \rangle = 0$. 则 x 是 C 唯一的极点并且 C 是一个平移锥.

证明. 通过定理6.5, $x \in C$. 通过定理6.4和注6.1, 对任意的 $l \in \text{dom}\delta^*(x, C)$, 存在 $\lambda_i^1 \geq 0$, $i = 1, \dots, i_1$, $\lambda_j^2 \geq 0$, $j = 1, \dots, j_2$, $\lambda_k^3 \geq 0$, $k = 1, \dots, k_3$ 和 $x_i \in D_1$, $y_j \in D_2$, $z_k \in D_{32}$, 使得 $l = \sum_{i=1}^{i_1} \lambda_i^1 x_i + \sum_{j=1}^{j_2} \lambda_j^2 y_j + \sum_{k=1}^{k_3} \lambda_k^3 z_k$. 由于对任意的 $y_j \in D_2$, 有 $\delta^*(y_j, C) = 0$, 下面的不等式成立:

$$\begin{aligned} \delta^*(l, C) &\leq \sum_{i=1}^{i_1} \lambda_i^1 \delta^*(x_i, C) + \sum_{k=1}^{k_3} \lambda_k^3 \delta^*(z_k, C) = \sum_{i=1}^{i_1} \lambda_i^1 - \sum_{k=1}^{k_3} \lambda_k^3 \\ &= \sum_{i=1}^{i_1} \lambda_i^1 \langle x_i, x \rangle + \sum_{k=1}^{k_3} \lambda_k^3 \langle z_k, x \rangle = \langle l, x \rangle. \end{aligned}$$

因此, 对每一个 $l \in \mathbb{R}^n$, 有 $\delta^*(l, C) \leq \langle l, x \rangle + \delta(l, \text{dom}\delta^*(x, C))$. 不等式两边取闭包, 有

$$\begin{aligned} \delta^*(l, C) &\leq \langle l, x \rangle + \delta(l, \text{cl}(\text{dom}\delta^*(x, C))) \\ &= \langle l, x \rangle + \delta(l, (0^+C)^\circ) \quad (\text{通过推论6.2}) \\ &= \delta^*(l, x + 0^+C). \end{aligned}$$

因此, $C \subseteq x + 0^+C$. 另一方面, 由于 $x \in C$, 有 $x + 0^+C \subseteq C$. 因此, $C = x + 0^+C$, 即 C 是一个平移锥并且 x 是它唯一的极点. □

6.3.1.1 C 不包含直线

当 $C \subset \mathbb{R}^n$ 是一个不包含直线的满维闭凸集时, 松弛算子 S_C 定义为

$$S_C = \begin{cases} S_C^1(x, y) = 1 - \langle x, y \rangle & (x, y) \in C \times D_1, \\ S_C^2(x, y) = -\langle x, y \rangle & (x, y) \in C \times D_2, \\ S_C^3(x, y) = -1 - \langle x, y \rangle & (x, y) \in C \times D_3. \end{cases}, \quad (6.3)$$

其中 D_1, D_2 和 D_3 互不相交并且对某一个凸集 C 可能是空集. 如果它们中的一个是空集, 我们只需要从定义中删除对应的松弛算子.

定义 6.2. 假设 $K \subset \mathbb{R}^m$ 是一个闭凸锥并且 $C \subset \mathbb{R}^n$ 是不包含直线的满维闭凸集. 如果存在映射

$$A : C \rightarrow K, B_1 : D_1 \rightarrow K^*, B_2 : D_2 \rightarrow K^*, B_3 : D_3 \rightarrow K^*$$

使得对 $(x, y) \in C \times D_i$, $S_C^i(x, y) = \langle A(x), B_i(y) \rangle$, $i = 1, 2, 3$, 则称由(6.3)定义的松弛算子 S_C 存在 K -分解.

定理 6.7. 令 $K \subset \mathbb{R}^m$ 为满维闭凸锥并且 $C \subset \mathbb{R}^n$ 为不包含直线的满维闭凸集. 假设 C 不是一个平移锥. 如果 C 存在由(6.1)定义的正则 K -举起, 则由(6.3)定义的松弛算子 S_C 存在 K -分解. 相反的, 如果由(6.3)定义的松弛算子 S_C 存在 K -分解, 则 C 存在由(6.1)定义的 K -举起.

定理6.8的证明可以经过适当改动来说明定理6.7的正确性.

例 6.2. 考虑 $C = \{x \mid x \geq -1\}$. 令 K 为 \mathcal{S}_+^3 并且

$$L = \left\{ \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{S}^3 \mid a_{11} = 1, a_{13} = 0, a_{23} = 0, a_{33} = a_{12} + 1 \right\}.$$

定义由 \mathcal{S}^3 到 \mathbb{R}^1 的线性映射: $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \rightarrow a_{33}$. 显然可得 C 存在 K -举起, 即 $C = \pi(K \cap L)$.

现在验证由(6.3)定义的松弛算子 S_C 是 K -可分的. 由于原点是 C 的内点, 根据定理6.4, D_2 和 D_3 是空集. 由于 $C^\circ = \{x \mid -1 \leq x \leq 0\}$, 可得 $D_1 = \text{ext}_1(C^\circ) \setminus \{0\} = \{(-1)\}$. 定义映射 $A : C \rightarrow K, B_1 : D_1 \rightarrow K^*$:

$$A(x) = \begin{pmatrix} 1 & x & 0 \\ x & x^2 & 0 \\ 0 & 0 & x+1 \end{pmatrix}, B_1(y) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -y \end{pmatrix}.$$

对所有的 $(x, y) \in C \times D_1$, 可得 $1 - \langle x, y \rangle = \langle A(x), B_1(y) \rangle$. 因此松弛算子 S_C 是 K -可分的.

注 6.3. 尽管定义6.2和定理6.7将文献 [31] 中的定理推广到一般凸集的情形, 推广的过程不是很显然. 理由是我们必须定义映射 A 并且验证 S_C 是否对 C 中所有的点都是可分的. 这个验证过程是困难的. 即使 C 是一个多面体, 它也包含了无穷多个点.

通过定理2.1, 如果 C 不包含直线, 它的每一个点都可以表示成极点和极方向的凸组合. 通过回收锥, 松弛算子和锥分解的定义, 我们将 [31] 中的结果推广到无界闭凸集的情形.

定义 6.3. 令 $K \subset \mathbb{R}^m$ 为一个闭凸锥. 一个无界闭凸集 $C \subset \mathbb{R}^n$ 的 K -举起定义为使得下面条件满足的集合 $Q = K \cap L$:

$$C = \pi(K \cap L), 0^+C = \pi(K \cap 0^+L) \quad (6.4)$$

这里 $L \subset \mathbb{R}^m$ 是一个仿射子空间并且 $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ 是一个线性映射. 如果 $L \cap \text{int}(K) \neq \emptyset$, 则称 Q 是 C 的一个正则 K -举起.

条件 $0^+C = \pi(K \cap 0^+L)$ 不是多余的并且一般不能由条件 $C = \pi(K \cap L)$ 得到. 见下面的例子:

例子6.2(续) 尽管等式 $C = \pi(K \cap L)$ 成立, 由于

$$0^+L = \left\{ \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathcal{S}^3 \mid a_{11} = 0, a_{13} = 0, a_{23} = 0, a_{33} = a_{12} \right\}$$

并且 $\pi(K \cap 0^+L) = \{0\}$, 可得 $\mathbb{R}_+^1 = 0^+C \neq \pi(K \cap 0^+L)$.

定义满维闭凸集 C 的松弛算子 S_C 为

$$S_C = \begin{cases} S_C^1(x, y) = 1 - \langle x, y \rangle & (x, y) \in \text{ext}_1(C) \times D_1, \\ S_C^2(x, y) = -\langle x, y \rangle & (x, y) \in \text{ext}_1(C) \times D_2, \\ S_C^3(x, y) = -1 - \langle x, y \rangle & (x, y) \in \text{ext}_1(C) \times D_3, \\ S_{0^+C}^i(x, y) = -\langle x, y \rangle & (x, y) \in \text{ext}_2(0^+C) \times D_i, i = 1, 2, 3. \end{cases} \quad (6.5)$$

在这个定义中, D_1 , D_2 和 D_3 互不相交并且可能对某一个闭凸集 C 是空集. 如果它们中的一个空集, 则从定义中去掉相应的松弛算子.

定义 6.4. 假设 $K \subset \mathbb{R}^m$ 是一个闭凸锥, $C \subset \mathbb{R}^n$ 是一个不包含直线的满维闭凸集. 如果存在映射

$$\begin{aligned} A_1 : \text{ext}_1(C) &\rightarrow K, A_2 : \text{ext}_2(0^+C) \rightarrow K, \\ B_1 : D_1 &\rightarrow K^*, B_2 : D_2 \rightarrow K^*, B_3 : D_3 \rightarrow K^* \end{aligned}$$

使得

- $S_c^i(x, y) = \langle A_1(x), B_i(y) \rangle \quad (x, y) \in \text{ext}_1(C) \times D_i, i = 1, 2, 3,$
- $S_{0^+C}^i(x, y) = \langle A_2(x), B_i(y) \rangle \quad (x, y) \in \text{ext}_2(0^+C) \times D_i, i = 1, 2, 3,$

则称由(6.5)定义的松弛算子 S_C 是 K -可分的.

定理 6.8. 假设 $K \subset \mathbb{R}^m$ 是一个满维闭凸锥, $C \subset \mathbb{R}^n$ 是不包含直线的满维闭凸集并且 C 不是一个平移锥. 如果 C 存在一个由(6.4)定义的正则 K -举起, 则由(6.5)定义的松弛算子 S_C 是 K -可分的. 相反的, 如果由(6.5)定义的 S_C 是 K -可分的, 则 C 存在一个由(6.4)定义的 K -举起.

证明. 假设 C 存在一个正则 K -举起, 则存在 \mathbb{R}^m 中的仿射子空间 $L = w_0 + L_0$, 这里 L_0 是一个线性子空间, $w_0 \in \text{int}(K)$ 和线性映射 $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^n$, 使得 $C = \pi(K \cap L)$, $0^+C = \pi(K \cap 0^+L)$. 由于 $0^+L = L_0$, 有 $0^+C = \pi(K \cap L_0)$. 现在, 需要通过 K -举起来定义映射 A_1, A_2 和 B_1, B_2, B_3 来分解松弛算子 S_C .

对任意的 $x_1 \in \text{ext}_1(C)$, 存在凸集 $w_1 \in K \cap L$, 使得 $\pi(w_1) = x_1$. 定义 $A_1(x_1) := w_1$. 此外, 对任意的 $x_2 \in \text{ext}_2(0^+C)$, 存在凸集 $w_2 \in K \cap L_0$, 使得 $\pi(w_2) = x_2$. 定义 $A_2(x_2) := w_2$.

B_1, B_2 和 B_3 的定义和文献 [31, 定理 1]中的定义相似, 都利用强对偶条件. 唯一的区别是对 $l_1 \in D_1$, $\max\{\langle l_1, x \rangle \mid x \in C\} = 1$, 对 $l_2 \in D_2$, $\max\{\langle l_2, x \rangle \mid x \in C\} = 0$ 并且对 $l_3 \in D_3$, $\max\{\langle l_3, x \rangle \mid x \in C\} = -1$. 这里省略了证明. 对任意的 $y_1 \in D_1$, 定义 $B_1(y_1) := z - \pi^*(y_1)$, 这里 z 是 $L_0^\perp \cap (K^* + \pi^*(y_1))$ 中满足 $\langle w_0, z \rangle = 1$ 的任意向量. 对任意的 $y_2 \in D_2$, 定义 $B_2(y_2) := z - \pi^*(y_2)$, 这里 z 是 $L_0^\perp \cap (K^* + \pi^*(y_2))$ 中满足 $\langle w_0, z \rangle = 0$ 的任意向量. 对任意的 $y_3 \in D_3$, 定义 $B_3(y_3) := z - \pi^*(y_3)$, 这里 z 是 $L_0^\perp \cap (K^* + \pi^*(y_3))$ 中满足 $\langle w_0, z \rangle = -1$ 的任意向量. 只需要验证 S_c^i 和 $S_{0^+C}^i$ 存在定义6.4中的 K -分解. S_c^i 的 K -分解可以使用与 [31, 定理 1] 相同的方法来验证. 对于任意的 $x_2 \in \text{ext}_2(0^+C)$ 和 $y_i \in D_i$, $i = 1, 2, 3$, 有

$$\begin{aligned} \langle x_2, y_i \rangle &= \langle \pi(w_2), y_i \rangle = \langle w_2, \pi^*(y_i) \rangle = \langle w_2, z - B_i(y_i) \rangle \\ &= -\langle w_2, B_i(y_i) \rangle = -\langle A_2(x_2), B_i(y_i) \rangle. \end{aligned}$$

因此, 根据定义6.4, $S_{0^+C}^i$ 是 K -可分的.

另一方面, 假设 S_C 是 K -可分的, 即存在映射 A_1, A_2 和 B_1, B_2, B_3 , 使得对任意的 $(x, y) \in \text{ext}_1(C) \times D_i$, $i = 1, 2, 3$, 有 $S_c^i(x, y) = \langle A_1(x), B_i(y) \rangle$ 并且对任意的 $(x, y) \in \text{ext}_2(0^+C) \times D_i$, $i = 1, 2, 3$, 有 $S_{0^+C}^i(x, y) = \langle A_2(x), B_i(y) \rangle$. 定义仿射空间

$$\begin{aligned} L = &\{(x, z) \in \mathbb{R}^n \times \mathbb{R}^m \mid \text{对任意的 } y_1 \in D_1, 1 - \langle x, y_1 \rangle = \langle z, B_1(y_1) \rangle, \\ &\text{对任意的 } y_2 \in D_2, -\langle x, y_2 \rangle = \langle z, B_2(y_2) \rangle, \\ &\text{对任意的 } y_3 \in D_3, -1 - \langle x, y_3 \rangle = \langle z, B_3(y_3) \rangle\}. \end{aligned}$$

令 L_K 为 L 到第二个分量 z 的投影.

首先, 需要证明 $0 \notin L_K$. 如果 $0 \in L_K$, 则存在 $x \in \mathbb{R}^n$, 使得对任意的 $y_1 \in D_1$, 有 $1 - \langle x, y_1 \rangle = 0$, 对任意的 $y_2 \in D_2$, 有 $-\langle x, y_2 \rangle = 0$, 对任意的 $y_3 \in D_3$,

有 $-1 - \langle x, y_3 \rangle = 0$. 通过定理6.6, C 是一个平移锥, 与假设矛盾. 由于对任意的 $x \in \text{ext}_1(C)$, 有 $A_1(x) \in K \cap L_K$, 因此 $K \cap L_K \neq \emptyset$.

对任意的 $x \in \mathbb{R}^n$, 如果存在 $z \in K$, 使得 $(x, z) \in L$, 则对任意的 $y_1 \in D_1$, 有 $\langle x, y_1 \rangle = 1 - \langle z, B_1(y_1) \rangle \leq 1$, 对任意的 $y_2 \in D_2$, $\langle x, y_2 \rangle = -\langle z, B_2(y_2) \rangle \leq 0$ 并且对任意的 $y_3 \in D_3$, 有 $\langle x, y_3 \rangle = -1 - \langle z, B_3(y_3) \rangle \leq -1$. 通过定理6.5, 有 $x \in C$. 因此, $\pi(K \cap L_K) \subseteq C$.

由于 C 不包含直线, 可以证明对任意的 $z \in K \cap L_K$, 存在唯一的 $x_z \in \mathbb{R}^n$, 使得 $(x_z, z) \in L$. 因此, 由 z 到 x_z 的仿射映射是一个良定义. 因此 L_K 不包含原点. 我们将定义在 $K \cap L_K$ 上的仿射映射扩展到 \mathbb{R}^m 上的线性映射. 为了证明 $C = \pi(K \cap L_K)$, 需要证明 $C \subseteq \pi(K \cap L_K)$.

对任意的 $x \in C$, 存在 $\lambda_i^1 \geq 0$, $i = 1, \dots, i_1$, $\lambda_j^2 \geq 0$, $j = 1, \dots, j_2$, 使得

$$x = \sum_{i=1}^{i_1} \lambda_i^1 x_i + \sum_{j=1}^{j_2} \lambda_j^2 y_j, \quad \sum_{i=1}^{i_1} \lambda_i^1 = 1,$$

这里 $x_i \in \text{ext}_1(C)$ 并且 $y_j \in \text{ext}_2(0^+C)$. 令 $z = \sum_{i=1}^{i_1} \lambda_i^1 A_1(x_i) + \sum_{j=1}^{j_2} \lambda_j^2 A_2(y_j)$. 由于 S_C 是 K -可分的, 很容易验证 $z \in K \cap L_K$ 并且 $x = \pi(z) \in \pi(K \cap L_K)$. 因此 $C = \pi(K \cap L_K)$.

此外, 还需要证明 $0^+C = \pi(K \cap 0^+L_K)$. 由于 $C = \pi(K \cap L_K)$, $0^+C \supseteq \pi(K \cap 0^+L_K)$. 另一方面, 对任意的 $x \in \text{ext}_2(0^+C)$, 根据 L 的定义, 有 $A_2(x) \in K \cap 0^+L_K$. 因此 $0^+C = \pi(K \cap 0^+L_K)$. \square

一般情况下, 给定一个闭凸集 C , 由(6.4)定义的 K -举起与由(6.4)定义的松弛算子的 K -分解不等价. 文献 [31, 推论 1] 证明当 K 是好锥并且 C 是凸体时二者的等价性. 在下面我们将结论推广到 C 是不包含直线的满维闭凸集情形.

推论 6.9. 假设 K 是一个好锥并且满维闭凸集 C 不包含直线. 如果 C 存在一个 K -举起(不一定是正则的), 则 S_C 存在一个 K -分解.

证明. 由于 C 存在一个由(6.4)定义的 K -举起 $Q = K \cap L$. 令 F 为 K 中包含 Q 的极小面. 因此 $Q \cap \mathbf{ri}(F) \neq \emptyset$ 并且 $Q = F \cap L$. 此外, 我们有 $0^+C = \pi(K \cap 0^+L) = \pi(0^+Q) = \pi(0^+(F \cap L)) = \pi(F \cap 0^+L)$. 因此, Q 存在一个由(6.4)定义的正则 F -举起. 定义(6.4)中的松弛算子 S_C 的 F -分解可以通过与定理6.8相同的方式来定义. 映射 A_1 和 A_2 可以映射到 K . 由于 K 是一个好锥并且 F 是 K 的一个面, 我们有 $F^* = K^* + L^\perp$. 通过与 [31, 推论 1] 相同的方法, 对 $1 \leq i \leq 3$, 定义相同的映射 B_i , 使得松弛算子 S_C 是 K -可分的. \square

下面的例子说明对 $i = 1, 2, 3$ 和 $(x, y) \in \text{ext}_1(C) \times D_i$, $S_c^i(x, y)$ 的 K -分解不能确保闭凸集 C 存在由(6.4)定义的 K -举起. 对 $i = 1, 2, 3$ 和 $(x, y) \in \text{ext}_2(0^+C) \times D_i$, 考虑 $S_{0^+C}^i(x, y)$ 的 K -分解是必要的.

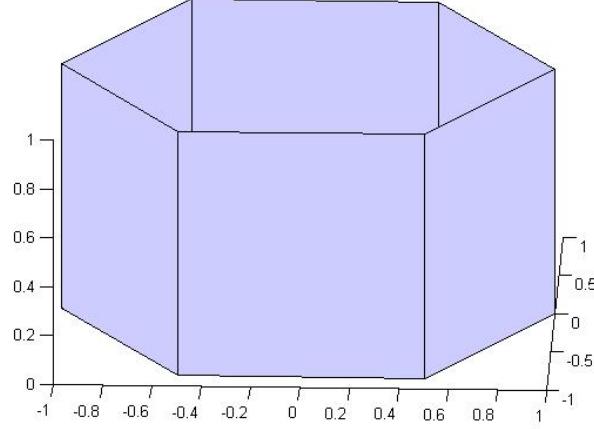


图 6.3: 例子6.3

例 6.3. \mathbb{R}^3 中的多面体 C 定义如下:

$$C = \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 : \begin{pmatrix} 1 & \frac{\sqrt{3}}{3} & 0 \\ 0 & \frac{2\sqrt{3}}{3} & 0 \\ -1 & \frac{\sqrt{3}}{3} & 0 \\ -1 & -\frac{\sqrt{3}}{3} & 0 \\ 0 & -\frac{2\sqrt{3}}{3} & 0 \\ 1 & -\frac{\sqrt{3}}{3} & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leq \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

回收锥 $0^+C = \text{cone}(\{(0,0,1)\})$ 并且 C 的顶点为 $\{(cos(i\pi/3), sin(i\pi/3), 0), i = 0, \dots, 5\}$. 根据定理6.7, C 的松弛矩阵是

$$S := \begin{pmatrix} 0 & 0 & 1 & 2 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 & 2 & 2 & 0 \\ 2 & 1 & 0 & 0 & 1 & 2 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

[31, 例子 2] 已经说明 S 的 6×6 子矩阵 S_H 存在 \mathbb{R}_+^5 -分解. 现在我们说明矩阵 S 不存在 \mathbb{R}_+^5 -分解. 假设矩阵 S 存在 \mathbb{R}_+^5 -分解:

$$\begin{pmatrix} S_H & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$

由于 $A_{21}B_{11} + A_{22}B_{21} = 0$, 有 $A_{22}B_{21} = 0$. 现在证明 $A_{22} = 0$. 否则, 有 $B_{21} = 0$ 并且 $S_H = A_{11}B_{11}$. 这与 S_H 不存在 \mathbb{R}_+^4 -分解矛盾. 通过 $A_{21}B_{12} + A_{22}B_{22} = 1$ 和 $A_{22} = 0$, 有 $B_{12} \neq 0$. 此外, 由于 $A_{11}B_{12} + A_{12}B_{22} = 0$, A_{11} 存在由零元素组成的列. 因此, S_H 在 \mathbb{R}_+^4 中存在非负分解. 矛盾. 因此根据定理 6.8, C 不存在 \mathbb{R}_+^5 -举起.

在定理 6.8 中的假设中, C 不是一个平移锥. 如果 C 是一个平移锥, C 的 K -举起定义为

$$C = b + \pi(K \cap L) \quad (6.6)$$

这里 $b \in \mathbb{R}^n$ 是一个常数向量, L 是一个线性空间并且 $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ 是一个线性映射. 在这个情况下, 只需要刻画 $\text{ext}_2(0^+C)$. 不失一般性, 假设 $b = 0$ 并且 C 是一个锥. 松弛算子 S_C 定义为

$$S_C(x, y) = -\langle x, y \rangle \quad (x, y) \in \text{ext}_2(C) \times \text{ext}_2(C^o). \quad (6.7)$$

如果存在映射

$$A : \text{ext}_2(C) \rightarrow K, \quad B : \text{ext}_2(C^o) \rightarrow K^*$$

使得

- $S_C(x, y) = \langle A(x), B(y) \rangle \quad (x, y) \in \text{ext}_2(C) \times \text{ext}_2(C^o),$

则称松弛算子 S_C 是 K -可分的.

定理 6.10. 假设 $K \subset \mathbb{R}^m$ 是满维闭凸锥并且 $C \subset \mathbb{R}^n$ 是尖的满维闭凸锥. 如果 C 存在由 (6.6) 定义的正则 K -举起, 则由 (6.7) 定义的松弛算子 S_C 是 K -可分的. 相反的, 如果由 (6.7) 定义的松弛算子 S_C 是 K -可分的, 则 C 存在由 (6.6) 定义的 K -举起.

如果 C 存在一个正则 K -举起, 则 S_C 是 K -可分的证明与定理 6.8 中的类似. 接下来, 一个简短的证明可以说明如果 S_C 是 K -可分的, C 存在一个 K -举起.

证明. 假设 S_C 是 K -可分的, 定义线性空间

$$L = \{(x, z) \mid \text{对任意的 } y \in \text{ext}_2(C^o), -\langle x, y \rangle = \langle z, B(y) \rangle\}.$$

令 L_K 为 L 到第二个分量 z 的投影. 由于 C 不包含直线, 对任意的 $z \in L_K$, 存在唯一的 $x_z \in \mathbb{R}^n$, 使得 $(x_z, z) \in L$. 因此, 线性映射 $\pi : L_K \rightarrow x_z$ 是一个良定义. 由于 L 是一个线性空间, 可以将 π 扩展成定义在 \mathbb{R}^m 上的线性映射.

对任意的 $x \in \mathbb{R}^n$, 如果存在 $z \in K$, 使得 $(x, z) \in L$, 则对所有的 $y \in \text{ext}_2(C^o)$, 有 $\langle x, y \rangle \leq 0$. 由于 C 是一个闭凸集, $x \in C^{oo} = \text{cl}(C) = C$. 因此 $\pi(K \cap L) \subseteq C$. 在另一方面, 由于 S_C 是 K -可分的. 则对任意的 $x \in \text{ext}_2(C)$, 有 $(x, A(x)) \in L$. 因此 $C \subseteq \pi(K \cap L)$. \square

6.3.1.2 C 包含直线的情形

当 C 是包含直线的满维闭凸集时, 定义6.2和定理6.7可以直接推广. 然而, 当 C 包含直线时, 它不存在极点并且 0^+C 不存在极射线. 定义6.4和定理6.8需要适当的调整.

令 L_1 为 C 的线性空间并且 $\{l_1, \dots, l_s\}$ 是 L_1 的一组正交基. C 可以分解为

$$C = C_0 + L_1, \quad (6.8)$$

这里, $C_0 = C \cap L_1^\perp$ 是一个不包含直线的闭凸集并且 L_1^\perp 是 L_1 的正交补.

引理 6.11. L_1^\perp 是 C° 的仿射闭包.

证明. 通过 $C = C_0 + L_1$, 可得 $C^\circ = C_0^\circ \cap L_1^\perp$. 由于闭凸集 C_0 不包含直线, 0^+C_0 不包含直线. 通过推论6.2和定理6.4, 有 $\text{cl}(\text{cone}(C_0^\circ)) = (0^+C_0)^\circ$. 由于 $(0^+C_0)^\circ$ 包含内点, $\text{cone}(C_0^\circ)$ 包含内点. 由于 $C_0 \subset L_1^\perp$ 中, $L_1 \subset \text{cone}(C_0^\circ)$. 因此 $\text{cone}(C^\circ) = \text{cone}(C_0^\circ) \cap L_1^\perp$ 与 L_1^\perp 的内点相交. 此外, $C^\circ = C_0^\circ \cap L_1^\perp$ 与 L_1^\perp 的内点相交. 因此, L_1^\perp 是 C° 的仿射闭包. \square

当 C 是包含直线的满维闭凸集时, 定义松弛算子 S_C :

$$S_C = \begin{cases} S_{C_0}^1(x, y) = 1 - \langle x, y \rangle & (x, y) \in \text{ext}_1(C_0) \times D_1, \\ S_{C_0}^2(x, y) = -\langle x, y \rangle & (x, y) \in \text{ext}_1(C_0) \times D_2, \\ S_{C_0}^3(x, y) = -1 - \langle x, y \rangle & (x, y) \in \text{ext}_1(C_0) \times D_3, \\ S_{0^+C_0}^i(x, y) = -\langle x, y \rangle & (x, y) \in \text{ext}_2(0^+C_0) \times D_i, i = 1, 2, 3, \\ S_{L_1}(x, y) = \langle x, y \rangle & (x, y) \in \{l_1, \dots, l_s\} \times \{l_1, \dots, l_s\}. \end{cases} \quad (6.9)$$

定义 6.5. 假设 $K \subset \mathbb{R}^m$ 是满维闭凸锥并且 $C \subset \mathbb{R}^n$ 是包含直线的无界满维闭凸集. 如果存在映射

$$\begin{aligned} A_1 : \text{ext}_1(C_0) &\rightarrow K, \quad A_2 : \text{ext}_2(0^+C_0) \rightarrow K, \quad A_3 : \{l_1, \dots, l_s\} \rightarrow K, \\ B_1 : D_1 &\rightarrow K^*, \quad B_2 : D_2 \rightarrow K^*, \quad B_3 : D_3 \rightarrow K^*, \quad F : \{l_1, \dots, l_s\} \rightarrow \mathbb{R}^m \end{aligned}$$

使得

- $S_{C_0}^i(x, y) = \langle A_1(x), B_i(y) \rangle \quad (x, y) \in \text{ext}_1(C_0) \times D_i, i = 1, 2, 3,$
- $S_{0^+C_0}^i(x, y) = \langle A_2(x), B_i(y) \rangle \quad (x, y) \in \text{ext}_2(0^+C_0) \times D_i, i = 1, 2, 3,$
- $S_{L_1}(x, y) = \langle A_3(x), F(y) \rangle \quad (x, y) \in \{l_1, \dots, l_s\} \times \{l_1, \dots, l_s\},$
- $\langle A_3(x), B_i(y) \rangle = 0 \quad (x, y) \in \{l_1, \dots, l_s\} \times D_i, i = 1, 2, 3,$

- $\langle A_1(x), F(y) \rangle = 0 \quad (x, y) \in \text{ext}_1(C_0) \times \{l_1, \dots, l_s\},$
- $\langle A_2(x), F(y) \rangle = 0 \quad (x, y) \in \text{ext}_2(0^+C_0) \times \{l_1, \dots, l_s\},$

则称由(6.9)定义的松弛算子 S_C 是 K -可分的.

定理 6.12. 令 $K \subset \mathbb{R}^m$ 是满维闭凸锥. 假设 $C \subset \mathbb{R}^n$ 是形如(6.8)的满维闭凸集并且 C_0 不是平移锥. 如果 C 存在一个由(6.4)定义的正则 K -举起, 则由(6.9)定义的松弛算子 S_C 是 K -可分的. 相反的, 如果由(6.9)定义的松弛算子 S_C 是 K -可分的, 则 C 存在一个由(6.4)定义的 K -举起.

证明. 由于 C 存在正则 K -举起, 存在仿射空间 $L = w_0 + L_0 \subset \mathbb{R}^m$, 这里 L_0 是线性子空间, $w_0 \in \text{int}(K)$ 和线性映射 $\pi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ 满足条件 $C = \pi(K \cap L)$, $0^+C = \pi(K \cap 0^+L)$. 由于 $0^+L = L_0$, 有 $0^+C = \pi(K \cap L_0)$. 现在, 定义映射 $A_1, A_2, A_3, B_1, B_2, B_3$ 和 F 来构造 K -举起. 可以利用和定理6.8相同的方式来定义 A_1, A_2, B_1, B_2, B_3 . 对任意的 $l_i, i = 1, \dots, s$, 存在 $w_i \in K \cap L_0$, 使得 $\pi(w_i) = l_i$. 因此, 对 $i = 1, \dots, s$, 定义 $A_3(l_i) := w_i$. 此外, 对 $i = 1, \dots, s$, 定义 $F(l_i) := \pi^*(l_i)$.

对 $i = 1, 2, 3$, 定理6.5 中的 $S_{C_0}^i, S_{0^+C_0}^i$ 可以利用和定理6.8相同的方式来验证. 对任意的 $x, y \in \{l_1, \dots, l_s\}$, 可得

$$\langle x, y \rangle = \langle \pi(A_3(x)), y \rangle = \langle A_3(x), F(y) \rangle.$$

对任意的 $x \in \{l_1, \dots, l_s\}$, $y \in D_i$, 可得

$$\langle A_3(x), B_i(y) \rangle = \langle A_3(x), z - \pi^*(y) \rangle = -\langle \pi(A_3(x)), y \rangle = -\langle x, y \rangle = 0.$$

对任意的 $x \in \text{ext}_1(C_0)$, $y \in \{l_1, \dots, l_s\}$, 可得

$$\langle A_1(x), F(y) \rangle = \langle \pi(A_1(x)), y \rangle = \langle x, y \rangle = 0.$$

对任意的 $x \in \text{ext}_2(0^+C_0)$, $y \in \{l_1, \dots, l_s\}$, 可得

$$\langle A_2(x), F(y) \rangle = \langle \pi(A_2(x)), y \rangle = \langle x, y \rangle = 0.$$

因此, S_C 是 K -可分的.

假设 S_C 是 K -可分的. 定义仿射空间 L :

$$\begin{aligned} L = & \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^m \mid x = x_1 + x_2, \quad x_1 \in L_1^\perp, \quad x_2 \in L_1, \\ & \text{对任意的 } y_1 \in D_1, \quad 1 - \langle x_1, y_1 \rangle = \langle z, B_1(y_1) \rangle, \\ & \text{对任意的 } y_2 \in D_2, \quad -\langle x_1, y_2 \rangle = \langle z, B_2(y_2) \rangle, \\ & \text{对任意的 } y_3 \in D_3, \quad -1 - \langle x_1, y_3 \rangle = \langle z, B_3(y_3) \rangle, \\ & \text{对 } i = 1, \dots, s, \quad \langle x_2, l_i \rangle = \langle z, F(l_i) \rangle\}. \end{aligned}$$

令 L_K 为 L 到分量 z 的投影.

现在, 证明 $0 \notin L_K$. 如果 $0 \in L_K$, 则存在 $x = x_1 + x_2$, 使得对任意的 $y_1 \in D_1$, 有 $1 - \langle x_1, y_1 \rangle = 0$, 对任意的 $y_2 \in D_2$, 有 $-\langle x_1, y_2 \rangle = 0$, 对任意的 $y_3 \in D_3$, 有 $-1 - \langle x_1, y_3 \rangle = 0$. 通过定理6.6, C_0 是一个平移锥, 矛盾. 此外, 对每一个 $x \in \text{ext}_1(C_0)$, 有 $A_1(x) \in K \cap L_K$. 因此 $K \cap L_K \neq \emptyset$.

现在证明对每一个 $z \in K \cap L_K$, 存在唯一的 $x \in \mathbb{R}^n$, 使得 $(x, z) \in L$. 假设对 $z \in K \cap L_K$, 存在两个不同的点 $x^1 = x_1^1 + x_2^1$ 和 $x^2 = x_1^2 + x_2^2$, 使得 $(x^1, z), (x^2, z)$ 在 L 中. 由于对 $y \in D_1 \cup D_2 \cup D_3$, $\langle x_1^1 - x_1^2, y \rangle = 0$ 并且 D_1, D_2 和 D_3 生成 C° , 对 $y \in C^\circ$, 有 $\langle x_1^1 - x_1^2, y \rangle = 0$. 根据引理6.11, L_1^\perp 是 C° 的仿射闭包, 因此 $x_1^1 - x_1^2 \in L_1$. 由于 $x_1^1 - x_1^2 \in L_1^\perp$, 有 $x_1^1 - x_1^2 = 0$. 此外, 对整数 $i \in [1, s]$, 有 $\langle x_2^1 - x_2^2, l_i \rangle = 0$ 并且 $x_2^1 - x_2^2 \in L_1$, 因此 $x_2^1 - x_2^2 = 0$. 因此, 由 z 到 x_z 的仿射映射 π 是一个良定义. 由于 $0 \notin L_K$, π 可以扩展成定义在 \mathbb{R}^m 上的线性映射. 通过使用和定理6.8 相同的证明方法, 可得 $C = \pi(K \cap L_K)$ 和 $0^+C = \pi(K \cap 0^+L_K)$. \square

注 6.4. 如果 K 是一个好锥, 定理6.12 中的正则条件是不必要的. 这时, C 的 K -举起和松弛算子 S_C 的 K -分解是等价的.

当 C 是包含直线的平移闭锥时, C 存在形如(6.8)的分解. 这时可以得到与定理6.10 相同的结果. 凸集 C 的松弛算子 S_C 定义为

$$S_C = \begin{cases} S_{C_0}(x, y) = -\langle x, y \rangle & (x, y) \in \text{ext}_2(C_0) \times \text{ext}_2(C^\circ), \\ S_{L_1}(x, y) = -\langle x, y \rangle & (x, y) \in \{l_1, \dots, l_s\} \times \{l_1, \dots, l_s\}. \end{cases} \quad (6.10)$$

定义 6.6. 如果存在映射

$$\begin{aligned} A_2 : \text{ext}_2(C_0) &\rightarrow K, \quad A_3 : \{l_1, \dots, l_s\} \rightarrow K, \\ B : \text{ext}_2(C^\circ) &\rightarrow K^*, \quad F : \{l_1, \dots, l_s\} \rightarrow \mathbb{R}^m \end{aligned}$$

使得

- $S_{C_0}(x, y) = \langle A_2(x), B(y) \rangle \quad (x, y) \in \text{ext}_2(C_0) \times \text{ext}_2(C^\circ),$
- $S_{L_1}(x, y) = \langle A_3(x), F(y) \rangle \quad (x, y) \in \{l_1, \dots, l_s\} \times \{l_1, \dots, l_s\},$
- $\langle A_2(x), F(y) \rangle = 0 \quad (x, y) \in \text{ext}_2(C_0) \times \{l_1, \dots, l_s\},$
- $\langle A_3(x), B(y) \rangle = 0 \quad \{l_1, \dots, l_s\} \times \text{ext}_2(C^\circ).$

则称由(6.10) 定义的松弛算子 S_C 是 K -可分的.

定理 6.13. 令 $K \subset \mathbb{R}^m$ 是满维闭凸锥并且 $C \subset \mathbb{R}^n$ 是包含直线的满维平移闭锥. 则 C 存在形如(6.8)的分解. 如果 C 存在由(6.6) 定义的正则 K -举起, 则由(6.10) 定义的松弛算子 S_C 是 K -可分的. 相反的, 如果由(6.10) 定义的松弛算子 S_C 是 K -可分的, 则 C 存在由(6.6) 定义的 K -举起.

定理6.13 的证明方法和定理6.10 以及定理6.12 的证明方法相同.

6.3.2 C 不是满维的

当闭凸集 C 不满维时, 它存在一个非平凡的仿射闭包.

定理 6.14. 假设 $C \subset \mathbb{R}^n$ 是一个闭凸集. 极集 C° 包含直线当且仅当 C 包含在一个非平凡的线性空间. 当 C 包含原点, C 不是满维的当且仅当 C° 包含直线.

证明. C° 包含直线当且仅当存在 $a \in \mathbb{R}^n$, 使得 $\delta^*(a, C) \leq 0$ 并且 $\delta^*(-a, C) \leq 0$, 即 $C \subseteq \{x \mid a^T x = 0\}$.

当 C 包含原点时, C 不满维当且仅当存在 $a \in \mathbb{R}^n$, 使得 $C \subseteq \{x \mid a^T x = 0\}$. \square

假设凸集 C 不满维并且不包含直线. 这时不能确定 C° 是否包含直线. 如果 C° 不包含直线, 与满维情形的结论相同. 如果 C° 包含直线, C° 中不存在极点和极方向并且集合 D_1, D_2 和 D_3 是空集. 令 L_2 为 C° 的线性空间. 假设 $C' = C + L_2$ 满足 $C' = C^\circ \cap L_2^\perp$. 因此, 闭凸集 C' 不包含直线. 显然可得 $0^+C' = 0^+C^\circ \cap L_2^\perp$. 由于 $C_3 = \{x \mid \delta^*(x, C) \leq -1\}$, 因此 $C'_3 = C_3 \cap L_2^\perp$ 不包含直线. 令

$$D'_1 = \text{ext}_1(C') \setminus 0, \quad D'_2 = \text{ext}_2(0^+C') \cap \{x \mid \delta^*(x, C') = 0\}, \quad D'_3 = \text{ext}_1(C'_3).$$

令

$$D'_{32} = \text{ext}_2(0^+C') \cap \{x \mid \delta^*(x, C') = -1\}.$$

则 $D'_{32} \subseteq D'_3$.

定理 6.15. 假设闭凸集 $C \subset \mathbb{R}^n$ 不满维并且不包含直线. $x \in C$ 当且仅当对每一个 $l \in D'_1$, $\langle l, x \rangle \leq 1$, 对每一个 $l \in D'_2$, 有 $\langle l, x \rangle \leq 0$, 对每一个 $l \in D'_3$, 有 $\langle l, x \rangle \leq -1$ 并且 $x \in L_2^\perp$, 其中 L_2 是 C° 的线性空间.

证明. 与定理6.5的证明相同. \square

假设闭凸集 C 不满维并且不包含直线. 对 $i = 1, 2, 3$, 通过将 D_i 替换成定理6.5和定义6.4中的 D'_i , 可以定义松弛算子 S_C 和它的 K -分解, 则第6.3.1.1节中的所有结果可以被推广到 C 不满维的情形. 尽管第6.3.1.2节中的结果可以被推广到闭凸集 C 不满维并且不包含直线的情形. 这会更加复杂, 我们在这里忽略了细节.

6.4 多面体的锥举起

与 [31, 第3节]相似, 我们推广了有界多面体的锥举起的结论. 假设多面体 $C \subset \mathbb{R}^n$ 定义如下:

$$\begin{aligned} C = \{x \in \mathbb{R}^n : f_1(x) \leq \alpha_1, \dots, f_{k_1}(x) \leq \alpha_{k_1}, g_1(x) \leq 0, \dots, g_{k_2}(x) \leq 0, \\ h_1(x) \leq -\beta_1, \dots, h_{k_3}(x) \leq -\beta_{k_3}\}, \end{aligned} \quad (6.11)$$

这里对 $i = 1, \dots, k_1$, $\alpha_i > 0$, 对 $j = 1, \dots, k_3$, $\beta_j > 0$. 0^+C 具有下面的形式:

$$\begin{aligned} 0^+C = \{x \in \mathbb{R}^n : f_1(x) \leq 0, \dots, f_{k_1}(x) \leq 0, g_1(x) \leq 0, \dots, g_{k_2}(x) \leq 0, \\ h_1(x) \leq 0, \dots, h_{k_3}(x) \leq 0\}. \end{aligned} \quad (6.12)$$

假设凸集 C 由点 c_1, \dots, c_t 和方向 r_1, \dots, r_s 生成. 我们推广了文献 [31, 101] 中的松弛矩阵的定义.

定义 6.7. 定义 C 的松弛矩阵为 $[S_1^T, S_2^T, S_3^T]^T$, 这里

1. $S_1 \in \mathbb{R}^{k_1 \times (t+s)}$, 其中对 $i = 1, \dots, k_1$, $j = 1, \dots, t$, 矩阵第 (i, j) -位置是 $\alpha_i - f_i(c_j)$ 并且对 $i = 1, \dots, k_1$, $j = 1, \dots, s$, 第 $(i, t+j)$ -位置是 $-f_i(r_j)$.
2. $S_2 \in \mathbb{R}^{k_2 \times (t+s)}$, 其中对 $i = 1, \dots, k_2$, $j = 1, \dots, t$, 矩阵第 (i, j) -位置是 $-g_i(c_j)$ 并且对 $i = 1, \dots, k_2$, $j = 1, \dots, s$, 第 $(i, t+j)$ -位置是 $-g_i(r_j)$.
3. $S_3 \in \mathbb{R}^{k_3 \times (t+s)}$, 其中对 $i = 1, \dots, k_3$, $j = 1, \dots, t$, 矩阵第 (i, j) -位置是 $-\beta_i - h_i(c_j)$ 并且对 $i = 1, \dots, k_3$, $j = 1, \dots, s$, 第 $(i, t+j)$ -位置是 $-h_i(r_j)$.

假设 C 是一个不包含直线的满维多面体. 如果 f_i, g_i, h_i 表示 C 的面; $\alpha_i = 1, \beta_j = 1$, $i = 1, \dots, k_1$, $j = 1, \dots, k_3$; c_1, \dots, c_t 和 r_1, \dots, r_s 分别为 C 的极点和极方向, 则 S 称为 C 的标准松弛矩阵.

定义 6.8. [31, 定理 7] 令 $M = (M_{ij}) \in \mathbb{R}_+^{p \times q}$ 为非负矩阵, K 为闭凸锥. 则 M 的 K -分解形如 $M_{ij} = \langle a_i, b_j \rangle$, 其中 $a_1, \dots, a_p \in K$ 并且 $b_1, \dots, b_q \in K^*$.

定义 6.8 将非负矩阵 [101] 的非负分解推广到任意闭凸锥的情形. 我们推广文献 [24, 定理 13], [31, 定理 3] 和 [101, 定理 3] 的结果并且证明了一般多面体的 K -举起和松弛矩阵的 K -分解的等价性.

当 $C \subset \mathbb{R}^n$ 是不包含直线的满维多面体并且 $K \subset \mathbb{R}^m$ 是满维的多面体锥时, 松弛算子的 K -分解等价于 C 的松弛矩阵的 K -分解. 当 C 满维时, 定理 6.16 可以由定理 6.8 和定理 6.10 直接推导.

定理 6.16. 假设 $K \subset \mathbb{R}^m$ 是满维闭凸锥. 如果不包含直线的满维多面体 $C \subset \mathbb{R}^n$ 存在一个正则 K -举起, 则 C 的每一个松弛矩阵存在一个 K -分解. 相反的, 如果 C 的一个松弛矩阵存在一个 K -分解, 则 C 存在一个 K -举起.

6.4.1 K 是一个多面体锥

在一般情形下, $0^+C = \pi(K \cap 0^+L)$ 不是多余的并且不能由条件 $C = \pi(K \cap L)$ 推出。当 C 和 K 都是多面体时, 由(6.1)定义的 K -举起和由(6.4)定义的 K -举起是等价的。

引理 6.17. 令 $C \subset \mathbb{R}^n$ 为不包含直线的满维多面体并且 $K \subset \mathbb{R}^m$ 为满维多面体锥, 则 C 存在一个由(6.1)定义的 K -举起当且仅当它存在一个由(6.4)定义的 K -举起。

证明. 只需证明如果存在一个仿射空间 L 和一个线性映射 $\pi: \mathbb{R}^m \rightarrow \mathbb{R}^n$, 使得 $C = \pi(K \cap L)$, 则 $0^+C = \pi(K \cap 0^+L)$ 。如果定义 Q 为 $K \cap L$, 则 Q 是一个多面体。对任意的 $x \in Q$, 存在 Q 的极点 $\alpha_1, \dots, \alpha_t$ 和 0^+Q 的非零极方向 $\alpha_{t+1}, \dots, \alpha_{t+s}$, 使得 $x = \lambda_1\alpha_1 + \dots + \lambda_t\alpha_t + \lambda_{t+1}\alpha_{t+1} + \dots + \lambda_{t+s}\alpha_{t+s}$, 这里对 $i = 1, \dots, t+s$, $\lambda_1 + \dots + \lambda_t = 1$ 并且 $\lambda_i \geq 0$ 。因此, $\pi(x) = \lambda_1\pi(\alpha_1) + \dots + \lambda_t\pi(\alpha_t) + \lambda_{t+1}\pi(\alpha_{t+1}) + \dots + \lambda_{t+s}\pi(\alpha_{t+s})$ 。因此, 0^+C 可以由 $\pi(\alpha_{t+1}), \dots, \pi(\alpha_{t+s})$ 生成。另一方面, 由于 $\alpha_{t+1}, \dots, \alpha_{t+s}$ 可以生成 $0^+Q = K \cap 0^+L$, 因此 $\pi(\alpha_{t+1}), \dots, \pi(\alpha_{t+s})$ 可以生成 $\pi(K \cap 0^+L)$ 。因此 $0^+C = K \cap 0^+L$, 证毕。□

定理 6.18. 假设 $C \subset \mathbb{R}^n$ 是一个不包含直线的满维多面体并且 $K \subset \mathbb{R}^m$ 是一个满维多面体锥。如果 C 不是一个平移锥, 则 C 存在一个由(6.1)定义的 K -举起当且仅当由(6.7)定义的 C 松弛矩阵存在一个由(6.8)定义的 K -分解。如果 C 是一个平移锥, 如果将由(6.1)定义的 K -举起替换成由(6.6)定义的 K -举起, 可以得到相同的结果。

证明. 根据引理 6.17, 多面体 C 存在定义(6.1)中的 K -举起当且仅当它存在定义(6.4)中的 K -举起。 L 与 $\text{int}(K)$ 相交的正则性条件确保定理 6.16 的证明中优化问题的强对偶条件成立(见文献 [31] 中的定理 1 的证明)。当 K 是一个多面体锥时, 在证明中的优化问题是一个线性规划问题并且只要 $K \cap L \neq \emptyset$, 强对偶条件自然成立。□

例 6.4. \mathbb{R}^2 中的多面体 C 定义如下:

$$C = \left\{ (x_1, x_2) \in \mathbb{R}^2 : \begin{pmatrix} 0 & 1 \\ -2 + \sqrt{3} & 1 \\ 1 - \sqrt{3} & \sqrt{3} - 1 \\ -1 & 2 - \sqrt{3} \\ -1 & -2 + \sqrt{3} \\ 1 - \sqrt{3} & 1 - \sqrt{3} \\ -2 + \sqrt{3} & -1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 2 \\ \sqrt{3} \\ 1 \\ 2 - \sqrt{3} \\ \sqrt{3} - 2 \\ -2\sqrt{3} + 3 \\ \sqrt{3} - 2 \\ 0 \end{pmatrix} \right\}.$$

通过定理 6.18, C 存在 \mathbb{R}_+^6 -举起当且仅当矩阵 S 存在 \mathbb{R}_+^6 -分解。我们将系数矩阵记作 H , 将右边的向量记作 d 。

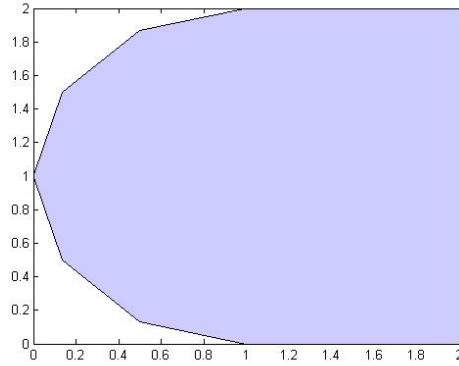


图 6.4: 例子6.4

C 的松弛矩阵为

$$S = \begin{bmatrix} 0 & 1 - \frac{\sqrt{3}}{2} & \frac{1}{2} & 1 & \frac{3}{2} & 1 + \frac{\sqrt{3}}{2} & 2 & 0 \\ 0 & 0 & 2 - \sqrt{3} & \sqrt{3} - 1 & 3 - \sqrt{3} & \sqrt{3} & 2 & 2 - \sqrt{3} \\ 2 - \sqrt{3} & 0 & 0 & 2 - \sqrt{3} & \sqrt{3} - 1 & 3 - \sqrt{3} & \sqrt{3} & \sqrt{3} - 1 \\ \sqrt{3} - 1 & 2 - \sqrt{3} & 0 & 0 & 2 - \sqrt{3} & \sqrt{3} - 1 & 3 - \sqrt{3} & 1 \\ 3 - \sqrt{3} & \sqrt{3} - 1 & 2 - \sqrt{3} & 0 & 0 & 2 - \sqrt{3} & \sqrt{3} - 1 & 1 \\ \sqrt{3} & 3 - \sqrt{3} & \sqrt{3} - 1 & 2 - \sqrt{3} & 0 & 0 & 2 - \sqrt{3} & \sqrt{3} - 1 \\ 2 & \sqrt{3} & 3 - \sqrt{3} & \sqrt{3} - 1 & 2 - \sqrt{3} & 0 & 0 & 2 - \sqrt{3} \\ 2 & 1 + \frac{\sqrt{3}}{2} & \frac{3}{2} & 1 & \frac{1}{2} & 1 - \frac{\sqrt{3}}{2} & 0 & 0 \end{bmatrix}.$$

S 的一个 \mathbb{R}_+^6 -分解为 $S = U \cdot V$, 这里

$$U = \begin{bmatrix} 1 & 1 & 0 & 1 - \frac{\sqrt{3}}{2} & 0 & 0 \\ 1 & -2\sqrt{3} + 4 & 2 - \sqrt{3} & 0 & 0 & 0 \\ \sqrt{3} - 1 & 0 & \frac{\sqrt{3}}{2} - \frac{1}{2} & 0 & 0 & \frac{\sqrt{3}}{2} - \frac{1}{2} \\ 2 - \sqrt{3} & 0 & 0 & 2 - \sqrt{3} & 0 & 1 \\ 0 & 0 & 0 & 2 - \sqrt{3} & 2 - \sqrt{3} & 1 \\ 0 & 0 & \frac{\sqrt{3}}{2} - \frac{1}{2} & 0 & \sqrt{3} - 1 & \frac{\sqrt{3}}{2} - \frac{1}{2} \\ 0 & -2\sqrt{3} + 4 & 2 - \sqrt{3} & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 - \frac{\sqrt{3}}{2} & 1 & 0 \end{bmatrix},$$

$$V = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & \sqrt{3} & 2 & 0 \\ 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{3}-1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & \sqrt{3} & 1 & 0 & 0 & 0 & 0 & 0 \\ \sqrt{3}-1 & 0 & 0 & 0 & 0 & 0 & \sqrt{3}-1 & 1 \end{bmatrix}.$$

C 的一个 \mathbb{R}_+^6 -举起是

$$C = \{(x_1, x_2) \mid \text{存在 } y \in \mathbb{R}_+^6, \text{使得 } Hx + Uy = d\}.$$

如果在方程 $Hx + Uy = d$ 中对变量 x_1 和 x_2 作消元, 可得

$$\begin{aligned} \{y \in \mathbb{R}_+^6 \mid y_1 &= 1 + (\sqrt{3}-1)y_4 + \frac{\sqrt{3}+1}{2}y_6 - \frac{\sqrt{3}+1}{2}y_3 - y_5, \\ y_2 &= \frac{1}{2} - \frac{1}{2}y_4 - \frac{\sqrt{3}+1}{4}y_6 + \frac{\sqrt{3}+1}{4}y_3\}. \end{aligned}$$

6.4.2 K 是一个半正定锥

有界多面体 C 的半正定秩是使得 C 存在 \mathcal{S}_+^k -举起最小的 k (参见文献 [31, 32]). 定义可以推广到无界多面体的情形. 由于对称半正定锥是一个好锥, C 的半正定秩等于 C 的松弛矩阵的半正定秩. 下面的引理推广了 [32, 命题 3.8]的结果.

引理 6.19. 假设 C 是不包含直线的满维多面体. 多面体 $C \subset \mathbb{R}^n$ 存在一个半正定秩是 k 的面, 则 C 的半正定秩至少为 $k+1$.

证明. 令 F 为 C 的一个面. 假设 F 的松弛矩阵 S_F 的半正定秩为 k . 令 $\alpha_1, \dots, \alpha_s$ 为 F 的顶点并且 $\alpha_{s+1}, \dots, \alpha_{s+t}$ 为 F 的极方向. C 中除了 F 的其他面记作 F_1, \dots, F_r . 由于 $F \neq C$, 存在一个不属于 F 的顶点或者极方向 α , 使得 $F(\alpha) > 0$. C 的松弛矩阵 S_C 包含一个 $(r+1) \times (s+t+1)$ 子矩阵, 其行索引为 F_1, \dots, F_r, F , 列索引为 $\alpha_1, \dots, \alpha_s, \alpha_{s+1}, \dots, \alpha_{s+t}, \alpha$ 并且有下面的形式:

$$S' = \begin{pmatrix} S_F & w \\ 0 & F(\alpha) \end{pmatrix} \text{ 这里 } w \in \mathbb{R}_+^r, F(\alpha) > 0.$$

根据 [32, 命题 2.6], S' 的半正定秩为 $k+1$. 因此 S_C 的半正定秩至少为 $k+1$. \square

有界多面体的半正定秩的下界在文献 [32, 命题 3.2]中给出并且可以推广到一般多面体的情形.

定理 6.20. 如果 $C \subset \mathbb{R}^n$ 是不包含直线的满维多面体, 则 C 的半正定秩至少为 n .

证明. 证明与 [32, 命题 3.2] 中的证明相似. 唯一的差别是当 $n = 1$ 时, C 是半线. 因此存在一个 1×2 的松弛矩阵. 显然可得对称矩阵的半正定秩是 1. 假设当维数小于 $n - 1$ 时是正确的. 在 C 中选取一个 $n - 1$ 维的面 F 并且 F 的半正定秩至少 $n - 1$. 通过引理 6.19, C 的半正定秩至少为 n . \square

注 6.5. 存在一个不包含直线的满维多面体 $C \subset \mathbb{R}^n$, 使得 C 的半正定秩为 n . 例如, 考虑 n 维非负象限 $\mathbb{R}_+^n = \{x \mid x_i \geq 0, i = 1, \dots, n\}$. \mathbb{R}_+^n 的松弛矩阵 $(0, I_n)$ 的半正定秩为 n , 这里 I_n 是一个单位阵并且 0 是一个零向量.

6.4.3 判定多面体的松弛矩阵

Gouveia 等在文献 [29] 中利用算法来判定给定一个非负矩阵是某一个有界多面体的松弛矩阵. 这个结果可以推广到刻画一般多面体的松弛矩阵. 下面的引理与 [29, 引理 10] 类似.

引理 6.21. 非负矩阵 S 是多面体 C 的松弛矩阵当且仅当它是一个不包含直线的满维多面体的松弛矩阵.

证明. 由于多面体 C 的松弛矩阵也是它的平移的松弛矩阵. 可以假设 C 包含原点. 如果 C 包含直线, 则 $C = C_0 + L_1$, 这里 $C_0 = C \cap L_1^\perp$ 是不包含直线的凸集并且 L_1^\perp 是 L_1 的正交补.

令 C 为由线性不等式 $f_i(x) \leq \alpha_i, g_j(x) \leq 0$ 定义的多面体, 这里 $\alpha_i > 0, 1 \leq i \leq k_1$ 并且 $1 \leq j \leq k_2$. C 中的每一个点是点 c_1, \dots, c_t 和方向 r_1, \dots, r_s 的凸组合. 根据定义 6.7, 多面体 C 的松弛矩阵 S 可以分解成

$$S = U \cdot V = \begin{pmatrix} \alpha_1 & -f_1 \\ \vdots & \vdots \\ \alpha_{k_1} & -f_{k_1} \\ 0 & -g_1 \\ \vdots & \vdots \\ 0 & -g_{k_2} \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ c_1 & \cdots & c_t & r_1 & \cdots & r_s \end{pmatrix}. \quad (6.13)$$

由于与 f_i 和 g_j 对应的线性函数在 C 上有上界, f_i 和 g_j 与 L_1 正交. 令 Q 为 L_1^\perp 的正交基, 则 $f_i \cdot (I - QQ^T) = 0$ 并且 $g_j \cdot (I - QQ^T) = 0$, 这里 I 表示单位矩阵, 0 表示零向量. 我们有下面的等式:

$$\begin{aligned} S &= U \cdot \begin{pmatrix} 1 & 0 \\ 0 & I - QQ^T + QQ^T \end{pmatrix} \cdot V \\ &= U \cdot \begin{pmatrix} 0 & 0 \\ 0 & I - QQ^T \end{pmatrix} \cdot V + U \cdot \begin{pmatrix} 1 & 0 \\ 0 & QQ^T \end{pmatrix} \cdot V \\ &= U \cdot \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & Q^T \end{pmatrix} \cdot V. \end{aligned}$$

令 $U' = U \cdot \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$ 并且 $V' = \begin{pmatrix} 1 & 0 \\ 0 & Q^T \end{pmatrix} \cdot V$. 则 $S = U' \cdot V'$ 是多面体 $Q^T C_0$ 的松弛矩阵.

如果 C_0 不满维, 则 $\text{aff}(C_0)$ 是一个非平凡的线性空间. 通过类似的变换, 可得 S 是 C_0 在 $\text{aff}(C_0)$ 中的松弛矩阵. \square

下面的定理和证明与 [29, 定理 6] 类似.

定理 6.22. 一个秩不小于 2 的非负矩阵 $S \in \mathbb{R}_+^{p \times q}$ 是一个多面体的松弛矩阵当且仅当 S 是一个多面体锥的松弛矩阵并且松弛矩阵的行空间中存在一个分量只有 0 和 1 的向量.

文献 [29, 定理 14], [27, 推论 5] 和 [32, 引理 3.1] 利用有界多面体的秩来刻画它的松弛矩阵的秩. 当 C 是一个尖的多面体锥时, 它的维数等于文献 [29, 引理 13]) 中松弛矩阵的秩. 结果可以扩展到一般多面体的情形.

定理 6.23. 假设 $C \subset \mathbb{R}^n$ 是不包含直线的 n 维多面体. 如果 C 不是平移锥, 则松弛矩阵 S 的秩是 $n + 1$.

证明. 假设 C 不是平移锥. 由于 C 经过平移之后其松弛矩阵的秩不改变并且 C 的所有松弛矩阵都有相同的秩, 假设 C 包含原点并且它的标准松弛矩阵形如 (6.13). 现在证明矩阵 U 是列满秩. 否则, 存在向量 $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, 使得 $U \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$. 如果 $x_1 \neq 0$, 令 $x_1 = 1$. 则对所有的 $f \in D_1$, $1 - f(x_2) = 0$ 并且对所有的 $g \in D_2$, $g(x_2) = 0$. 通过引理 6.6, 可得 C 是一个平移锥. 矛盾. 如果 $x_1 = 0$, 由于对每一个向量 $y \in C^\circ$, 存在 $\lambda_i^1 \geq 0$, $1 \leq i \leq k_1$, $\lambda_j^2 \geq 0$, $1 \leq j \leq k_2$ 和 $f_i \in D_1$, $g_j \in D_2$, 使得 $y = \sum_{i=1}^{k_1} \lambda_i^1 f_i + \sum_{j=1}^{k_2} \lambda_j^2 g_j$, 因此 $\langle x_2, y \rangle = 0$. 由于 $\dim(C^\circ) = n$, C° 包含内点. 由于对每一个 $y \in C^\circ$, $\langle y, x_2 \rangle = 0$, 因此 $x_2 = 0$. 因此 U 是一个列满秩矩阵. 此外, 由于 C 是 n 维多面体, \mathbb{R}^{n+1} 中由向量 $\begin{pmatrix} 1 \\ c_1 \end{pmatrix}, \dots, \begin{pmatrix} 1 \\ c_t \end{pmatrix}, \begin{pmatrix} 0 \\ r_1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ r_s \end{pmatrix}$ 生成的锥的维数是 $n + 1$. 因此, 矩阵 V 的秩为 $n + 1$, 从而 S 的秩为 $n + 1$. \square

推论 6.24. 假设多面体 C 存在分解 $C = C_0 + L_1$, 这里 L_1 是 C 的线性空间并且 $C_0 = C \cap L_1^\perp$. 如果 C 不是一个平移凸锥, 则松弛矩阵 S 的秩为 $\dim(C_0) + 1$.

文献 [91, 定理 3.2] 说明矩阵空间 $\mathbb{R}^{m \times n}$ 中秩为 3 的非负矩阵的非负秩存在一个上界 $[6 \min\{m, n\}/7]$. 通过定理 6.23 和推论 6.24, 对 \mathbb{R}^2 中任意的多面体(平移锥除外), 如果它的松弛矩阵 $S \in \mathbb{R}^{m \times n}$ 的秩为 3, 当 $\min(m, n) \geq 7$ 时, S 具有非平凡的非负秩分解. 这个结论为例子 6.4 中的松弛矩阵 $S \in \mathbb{R}^{8 \times 8}$ 存在 \mathbb{R}_+^7 -分解提供了理论保证.

第七章 结论与展望

本文主要研究多项式优化和凸代数几何领域的以下四个问题:

- 给定一个理想, 如何计算它的实根理想?
- 给定一个无界闭半代数集, 如何计算它的闭凸包?
- 给定一个参数化优化问题, 如何刻画它的最优值函数?
- 给定一个闭凸集, 如何利用最少的变量和约束条件来刻画它?

在第三章中, 我们讨论利用基于矩量矩阵的Lasserre半正定松弛序列来求解一个理想的实根理想. 当理想的实根理想是正维时, 已有算法由于缺乏一个有效的终止条件而无法计算它的实根理想. 我们设计一个基于偏微分方程中的对合理论的终止条件(3.6)和算法3.1. 在一般的坐标系下运行算法3.1, 则终止条件(3.6)在有限步成立. 当终止条件成立时, 算法返回一个介于理想和它的实根理想之间的理想 J 的Pommaret基. 给定一个半代数集 \mathcal{A} , 终止条件(3.6)可以计算一个理想的 \mathcal{A} -根理想. 在这个情形下, 可以得到和计算理想的实根理想相同的结论.

在第四章中, 我们讨论计算无界闭半代数集的闭凸包问题. 我们构造了推广Theta体序列和推广Lasserre半正定表示序列. 当假设2成立时, 推广Theta体序列和推广Lasserre半正定表示序列可以收敛到半代数集的闭凸包. 此外, 如果PP-BDR性质成立并且当阶数充分大时, 推广Theta体和推广Lasserre半正定表示是准确的. 假设2对推广Theta体序列和推广Lasserre半正定表示序列的收敛性是必要的. 最后我们讨论当假设2不满足时可采用的一些策略.

在第五章中, 我们讨论如何求解参数化优化问题. 当可行集无界并且光滑时, 我们利用对偶仿射簇来描述最优值函数. 当可行集的闭凸包不包含直线时, 对一般的参数值, 对偶仿射簇的定义多项式可以约化成一个非零的单变元多项式, 并且它的根可以表示对应优化问题的最优值. 当可行集有界并且不光滑时, 我们构造一个对偶仿射簇序列来描述最优值函数. 我们设计一个基于极仿射簇算法的参数化形式. 这个参数化算法返回一个多项式序列, 使得对任意的参数

值, 序列中的一个多项式被约化成非零单变元多项式并且它的根包含对应优化问题的最优值.

在第六章中, 我们讨论无界闭凸集的锥举起问题. 任给一个闭凸集, 我们利用它的极集的极点和极方向来刻画这个闭凸集. 利用这个结果和凸集的表示定理, 我们推广了凸体的松弛算子和锥举起的定义. 给定一个满维闭凸锥和一个不包含直线的满维闭凸集. 如果闭凸集不是平移锥, 我们证明闭凸集存在一个正则锥举起可以推出它的松弛算子是具有锥分解. 松弛算子具有锥分解可以推出闭凸集存在一个锥举起. 我们将这个结果扩展到闭凸集包含直线, 不满维或者闭凸集是平移锥的情形.

今后的工作主要包括下面几个方面:

- 在计算理想的实根理想的生成元问题中, 如何利用矩量矩阵的结构来刻画实根理想生成元的次数上界?
- 在无界实代数集的闭凸包求解问题中, 实代数集在无穷远处是闭的条件对一般的实代数集不一定成立. 如何构造谱多面体近似序列来逼近实代数集的闭凸包?
- 在可行集是半代数集的参数化优化问题中, 如何利用代数集或者半代数集来刻画最优值函数?

参考文献

- [1] Algorithm 795: Phcpack: A general-purpose solver for polynomial systems by homotopy continuation. *ACM Transactions on Mathematical Software*, 25(2):251–276, 1999.
- [2] *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Undergraduate Texts in Mathematics, 3rd edition. Springer, 233 Spring Street, New York, NY 10013, USA, 2007.
- [3] P. Aubry, F. Rouillier, and M. Safey El Din. Real solving for positive dimensional systems. *J. Symbolic Comput.*, 34(6):543–560, 2002.
- [4] B. Bank, M. Giusti, J. Heintz, and G. M. Mbakop. Polar varieties and efficient real elimination. *Math. Z.*, 238(1):115–144, 2001.
- [5] B. Bank, M. Giusti, J. Heintz, and L. M. Pardo. Generalized polar varieties: geometry and algorithms. *J. Complexity*, 21(4):377–412, 2005.
- [6] B. Bank, M. Giusti, J. Heintz, and M. Safey El Din. Intrinsic complexity estimates in polynomial optimization. *J. Complexity*, 30(4):430–443, 2014.
- [7] B. Bank, M. Giusti, J. Heintz, M. Safey El Din, and E. Schost. On the geometry of polar varieties. *Appl. Algebra Engrg. Comm. Comput.*, 21(1):33–83, 2010.
- [8] S. Basu, R. Pollack, and M.-F. Roy. On computing a set of points meeting every cell defined by a family of polynomials on a variety. *J. Complexity*, 13(1):28–37, 1997.
- [9] E. Becker, M. G. Marinari, T. Mora, and C. Traverso. The shape of the shape lemma. In *In Proc. Of The International Symposium On Symbolic And Algebraic Computation*, ISSAC ’93. ACM, 1993.

- [10] E. Becker and R. Neuhaus. Computation of real radicals of polynomial ideals. In *Computational algebraic geometry (Nice, 1992)*, volume 109 of *Progr. Math.*, pages 1–20. Birkhäuser Boston, Boston, MA, 1993.
- [11] A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization*. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2001. Analysis, algorithms, and engineering applications.
- [12] G. Blekherman, P. A. Parrilo, and R. R. Thomas, editors. *Semidefinite optimization and convex algebraic geometry*, volume 13 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2013.
- [13] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3)*. Springer-Verlag, Berlin, 1998.
- [14] J. Borwein and H. Wolkowicz. Regularizing the abstract convex program. *J. Math. Anal. Appl.*, 83(2):495–530, 1981.
- [15] B. Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynom*. Thesis, Mathematical Institute, University of Innsbruck, Austria, 1965.
- [16] F. Catanese, S. Hosten, A. Khetan, and B. Sturmfels. The maximum likelihood degree. *Amer. J. Math.*, 128(3):671–697, 2006.
- [17] G. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decomposition. *Lecture Notes in Computer Science*, 33(3):134–183, 1975.
- [18] R. E. Curto and L. A. Fialkow. Solution of the truncated complex moment problem for flat data. *Mem. Amer. Math. Soc.*, 119(568):x+52, 1996.

- [19] J. C. Faugère, P. Gianni, D. Lazard, and T. Mora. Efficient computation of zero-dimensional Gröbner bases by change of ordering. *J. Symbolic Comput.*, 16(4):329–344, 1993.
- [20] W. W. Fenchel. *Convex cones, sets, and functions*. Princeton University, 1953.
- [21] G. Feng, W. Chu, and Z. Lihong. Optimizing a linear function over a noncompact real algebraic variety. SNC ’14, pages 39–40, New York, NY, USA, 2014. ACM.
- [22] L. A. Fialkow. Solution of the truncated moment problem with variety $y = x^3$. *Trans. Amer. Math. Soc.*, 363(6):3133–3165, 2011.
- [23] L. A. Fialkow and J. Nie. The truncated moment problem via homogenization and flat extensions. *J. Funct. Anal.*, 263(6):1682–1700, 2012.
- [24] S. Fiorini, S. Massar, S. Pokutta, H. R. Tiwary, and R. de Wolf. Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds. In *STOC’12—Proceedings of the 2012 ACM Symposium on Theory of Computing*, pages 95–106. ACM, New York, 2012.
- [25] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, resultants and multidimensional determinants*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2008. Reprint of the 1994 edition.
- [26] P. Gianni, B. Trager, and G. Zacharias. Gröbner bases and primary decomposition of polynomial ideals. *J. Symbolic Comput.*, 6(2-3):149–167, 1988. Computational aspects of commutative algebra.
- [27] N. Gillis and F. Glineur. On the geometric interpretation of the nonnegative rank. *Linear Algebra Appl.*, 437(11):2685–2712, 2012.
- [28] M. Giusti, G. Lecerf, and B. Salvy. A Gröbner free alternative for polynomial system solving. *J. Complexity*, 17(1):154–211, 2001.

- [29] J. Gouveia, R. Grappe, V. Kaibel, K. Pashkovich, R. Z. Robinson, and R. R. Thomas. Which nonnegative matrices are slack matrices? *Linear Algebra Appl.*, 439(10):2921–2933, 2013.
- [30] J. Gouveia, P. A. Parrilo, and R. R. Thomas. Theta bodies for polynomial ideals. *SIAM J. Optim.*, 20(4):2097–2118, 2010.
- [31] J. Gouveia, P. A. Parrilo, and R. R. Thomas. Lifts of convex sets and cone factorizations. *Math. Oper. Res.*, 38(2):248–264, 2013.
- [32] J. Gouveia, R. Z. Robinson, and R. R. Thomas. Polytopes of minimum positive semidefinite rank. *Discrete Comput. Geom.*, 50(3):679–699, 2013.
- [33] J. Gouveia and R. R. Thomas. Spectrahedral approximations of convex hulls of algebraic sets. In *Semidefinite optimization and convex algebraic geometry*, volume 13 of *MOS-SIAM Ser. Optim.*, pages 293–340. SIAM, Philadelphia, PA, 2013.
- [34] H.-C. Graf von Bothmer and K. Ranestad. A general formula for the algebraic degree in semidefinite programming. *Bull. Lond. Math. Soc.*, 41(2):193–197, 2009.
- [35] A. Greuet, F. Guo, M. Safey El Din, and L. Zhi. Global optimization of polynomials restricted to a smooth variety using sums of squares. *J. Symbolic Comput.*, 47(5):503–518, 2012.
- [36] A. Greuet and M. Safey El Din. Deciding reachability of the infimum of a multivariate polynomial. In *ISSAC 2011—Proceedings of the 36th International Symposium on Symbolic and Algebraic Computation*, pages 131–138. ACM, New York, 2011.
- [37] A. Greuet and M. Safey El Din. Probabilistic algorithm for polynomial optimization over a real algebraic set. *SIAM J. Optim.*, 24(3):1313–1343, 2014.

- [38] F. Guo, L. Wang, and G. Zhou. Minimizing rational functions by exact Jacobian SDP relaxation applicable to finite singularities. *J. Global Optim.*, 58(2):261–284, 2014.
- [39] J. W. Helton and J. Nie. Sufficient and necessary conditions for semidefinite representability of convex hulls and sets. *SIAM J. Optim.*, 20(2):759–791, 2009.
- [40] J. W. Helton and J. Nie. Semidefinite representation of convex sets. *Math. Program.*, 122(1, Ser. A):21–64, 2010.
- [41] J. W. Helton and J. Nie. Semidefinite representation of convex sets. *Math. Program.*, 122(1, Ser. A):21–64, 2010.
- [42] D. Henrion and J.-B. Lasserre. GloptiPoly: global optimization over polynomials with Matlab and SeDuMi. *ACM Trans. Math. Software*, 29(2):165–194, 2003.
- [43] S. Hosten, A. Khetan, and B. Sturmfels. Solving the likelihood equations. *Found. Comput. Math.*, 5(4):389–407, 2005.
- [44] Z. Jelonek and K. Kurdyka. Quantitative generalized Bertini-Sard theorem for smooth affine varieties. *Discrete Comput. Geom.*, 34(4):659–678, 2005.
- [45] T. Krick and A. Logar. An algorithm for the computation of the radical of an ideal in the ring of polynomials. In *Applied algebra, algebraic algorithms and error-correcting codes (New Orleans, LA, 1991)*, volume 539 of *Lecture Notes in Comput. Sci.*, pages 195–205. Springer, Berlin, 1991.
- [46] K. Kurdyka, P. Orro, and S. Simon. Semialgebraic Sard theorem for generalized critical values. *J. Differential Geom.*, 56(1):67–92, 2000.
- [47] J. B. Lasserre. Convex sets with semidefinite representation. *Math. Program.*, 120(2, Ser. A):457–477, 2009.

- [48] J. B. Lasserre, M. Laurent, and P. Rostalski. Semidefinite characterization and computation of zero-dimensional real radical ideals. *Found. Comput. Math.*, 8(5):607–647, 2008.
- [49] J. B. Lasserre, M. Laurent, and P. Rostalski. A prolongation-projection algorithm for computing the finite real variety of an ideal. *Theoret. Comput. Sci.*, 410(27-29):2685–2700, 2009.
- [50] J. B. Lasserre, M. Laurent, and P. Rostalski. A unified approach to computing real and complex zeros of zero-dimensional ideals. In *Emerging applications of algebraic geometry*, volume 149 of *IMA Vol. Math. Appl.*, pages 125–155. Springer, New York, 2009.
- [51] J. B. Lasserre, M. Laurent, and P. Rostalski. A unified approach to computing real and complex zeros of zero-dimensional ideals. In *Emerging applications of algebraic geometry*, volume 149 of *IMA Vol. Math. Appl.*, pages 125–155. Springer, New York, 2009.
- [52] M. Laurent. Revisiting two theorems of Curto and Fialkow on moment matrices. *Proceedings of the American Mathematical Society*, 133(10):2965–2976, 2005.
- [53] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging applications of algebraic geometry*, volume 149 of *IMA Vol. Math. Appl.*, pages 157–270. Springer, New York, 2009.
- [54] M. Laurent and P. Rostalski. The approach of moments for polynomial equations. In *Handbook on semidefinite, conic and polynomial optimization*, volume 166 of *Internat. Ser. Oper. Res. Management Sci.*, pages 25–60. Springer, New York, 2012.
- [55] S. Lin. Unbounded convex semialgebraic sets as spectrahedral shadows, 2010.

- [56] H. Lombardi, D. Perrucci, and M.-F. Roy. An elementary recursive bound for effective positivstellensatz and hilbert 17-th problem. *ArXiv e-prints*, Apr. 2014.
- [57] Y. Ma. *Polynomial Optimization via Low-rank Matrix Completion and Semidefinite Programming*. Thesis, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, 2012.
- [58] M. Marshall. Optimization of polynomial functions. *Canad. Math. Bull.*, 46(4):575–587, 2003.
- [59] M. Marshall. *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008.
- [60] H. Möller. An inverse problem for cubature formulae. *Computational Technologies*, 9(13-20), 2004.
- [61] R. Neuhaus. Computation of real radicals of polynomial ideals. II. *J. Pure Appl. Algebra*, 124(1-3):261–280, 1998.
- [62] R. Neuhaus. Computation of real radicals of polynomial ideals. II. *J. Pure Appl. Algebra*, 124(1-3):261–280, 1998.
- [63] J. Nie. An exact Jacobian SDP relaxation for polynomial optimization. *Math. Program.*, 137(1-2, Ser. A):225–255, 2013.
- [64] J. Nie and K. Ranestad. Algebraic degree of polynomial optimization. *SIAM J. Optim.*, 20(1):485–502, 2009.
- [65] J. Nie, K. Ranestad, and B. Sturmfels. The algebraic degree of semidefinite programming. *Math. Program.*, 122(2, Ser. A):379–405, 2010.
- [66] P. A. Parrilo. Exact semidefinite representations for genus zero curves. Talk at the Banff workshop Positive Polynomials and Optimization, Banff, Canada. October 8-12, 2006.

- [67] G. Pataki. On the connection of facially exposed and nice cones. *J. Math. Anal. Appl.*, 400(1):211–221, 2013.
- [68] V. Powers and C. Scheiderer. The moment problem for non-compact semi-algebraic sets. *Adv. Geom.*, 1(1):71–88, 2001.
- [69] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.*, 42(3):969–984, 1993.
- [70] P. J. Rabier. Ehresmann fibrations and Palais-Smale conditions for morphisms of Finsler manifolds. *Ann. of Math. (2)*, 146(3):647–691, 1997.
- [71] K. Ranestad. Algebraic degree in semidefinite and polynomial optimization. In *Handbook on semidefinite, conic and polynomial optimization*, volume 166 of *Internat. Ser. Oper. Res. Management Sci.*, pages 61–75. Springer, New York, 2012.
- [72] K. Ranestad and B. Sturmfels. The convex hull of a variety. In *Notions of positivity and the geometry of polynomials*, Trends Math., pages 331–344. Birkhäuser/Springer Basel AG, Basel, 2011.
- [73] M. S. Ravi. Regularity of ideals and their radicals. *Manuscripta Math.*, 68(1):77–87, 1990.
- [74] G. Reid and L. Zhi. Solving polynomial systems via symbolic-numeric reduction to geometric involutive form. *J. Symbolic Comput.*, 44(3):280–291, 2009.
- [75] J. F. Ritt. *Differential Algebra*. American Mathematical Society Colloquium Publications, Vol. XXXIII. American Mathematical Society, New York, N. Y., 1950.
- [76] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
- [77] P. Rostalski. Bermeja-software for convex algebraic geometry. <http://math.berkeley.edu/~philipp/Software/Software>.

- [78] P. Rostalski. *Algebraic moments. real root finding and related topics.* PhD thesis, ETH Zurich, 2009.
- [79] P. Rostalski and B. Sturmfels. Dualities in convex algebraic geometry. *Rend. Mat. Appl.* (7), 30(3-4):285–327, 2010.
- [80] M. Safey El Din. Computing the global optimum of a multivariate polynomial over the reals. In *ISSAC 2008*, pages 71–78. ACM, New York, 2008.
- [81] M. Safey El Din and É. Schost. Polar varieties and computation of one point in each connected component of a smooth algebraic set. In *Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation*, pages 224–231 (electronic). ACM, New York, 2003.
- [82] M. Safey El Din and É. Schost. Properness defects of projections and computation of at least one point in each connected component of a real algebraic set. *Discrete Comput. Geom.*, 32(3):417–430, 2004.
- [83] M. Safey El Din and L. Zhi. Computing rational points in convex semialgebraic sets and sum of squares decompositions. *SIAM J. Optim.*, 20(6):2876–2889, 2010.
- [84] J. Schmid. On the degree complexity of hilbert’s 17th problem and the real nullstellensatz. *Habilitationsschrift, Universität Dortmund*, 1998.
- [85] J. Schmid. On the complexity of the real nullstellensatz in the 0-dimensional case. *Journal of Pure and Applied Algebra*, 151(3):301–308, 2000.
- [86] M. Schweighofer. Optimization of polynomials on compact semialgebraic sets. *SIAM J. Optim.*, 15(3):805–825 (electronic), 2005.
- [87] R. Scott. *Approximate Gröbner Bases.* Master thesis, University of Western Ontario, Canada, 2006.
- [88] R. Scott, G. Reid, W. Wu, and L. Zhi. Geometric involutive bases and applications to approximate commutative algebra. In *Approximate commutative*

- algebra*, Texts Monogr. Symbol. Comput., pages 99–124. SpringerWien-NewYork, Vienna, 2009.
- [89] W. M. Seiler. *Involution*, volume 24 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, 2010. The formal theory of differential equations and its applications in computer algebra.
- [90] W. M. Seiler. *Involution—the formal theory of differential equations and its applications in computer algebra*. PhD thesis, 2011.
- [91] Y. Shitov. An upper bound for nonnegative rank. *J. Combin. Theory Ser. A*, 122:126–132, 2014.
- [92] R. Sinn. Algebraic boundaries of convex semi-algebraic sets. *Res. Math. Sci.*, 2:Art. 3, 18, 2015.
- [93] S. Spang. *On the computation of the real radical*. Thesis, Technische Universität Kaiserslautern, 2007.
- [94] G. Stengle. A nullstellensatz and a positivstellensatz in semialgebraic geometry. *Math. Ann.*, 207:87–97, 1974.
- [95] H. Stetter. *Numerical polynomial algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2004.
- [96] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Rev.*, 38(1):49–95, 1996.
- [97] H. H. Vui and P. T. So’n. Representations of positive polynomials and optimization on noncompact semialgebraic sets. *SIAM J. Optim.*, 20(6):3082–3103, 2010.
- [98] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of Semidefinite Programming*. International Series in Operations Research & Management Science, 27. Kluwer Academic Publishers, Boston, MA, 2000. Theory, algorithms, and applications.

- [99] W.-T. Wu. *Mathematics mechanization*, volume 489 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht; Science Press, Beijing, 2000. Mechanical geometry theorem-proving, mechanical geometry problem-solving and polynomial equations-solving.
- [100] B. Xia and L. Yang. An algorithm for isolating the real solutions of semi-algebraic systems. *J. Symbolic Comput.*, 34(5):461–477, 2002.
- [101] M. Yannakakis. Expressing combinatorial optimization problems by linear programs. *J. Comput. System Sci.*, 43(3):441–466, 1991.
- [102] G. Zeng. Computation of generalized real radicals of polynomial ideals. *Sci. China Ser. A*, 42(3):272–280, 1999.

发表文章目录

- [1] Yue Ma, Chu Wang, Lihong Zhi. *A certificate for semidefinite relaxations in computing positive dimensional real varieties.* Journal of Symbolic Computation. 2016, Vol. 72, pp. 1-20.
- [2] Feng Guo, Chu Wang, Lihong Zhi. *Semidefinite Representations of Noncompact Convex Sets.* SIAM Journal on Optimization, 2015, Vol. 25, No. 1, pp. 377-395.
- [3] Feng Guo, Chu Wang, Lihong Zhi. *Optimizing a linear function over a non-compact real algebraic variety.* SNC'14 Proceedings of the 2014 Symposium on Symbolic-Numeric Computation, pp 39-40.
- [4] Feng Guo, Mohab Safey El Din, Chu Wang and Lihong Zhi. *Optimizing a parameteric linear function over a non-compact real algebraic variety.* ISSAC'15 Proceedings of the 2015 ACM on International Symposium on Symbolic and Algebraic Computation, pp 205-212.
- [5] Chu Wang, Lihong Zhi. Lifts of Non-compact Convex Sets and Cone Factorizations. eprint arXiv:1501. 00115.

简 历

王础, 男, 辽宁省, 1989年出生. E-mail: cwang@mmrc.iss.ac.cn

教育状况

- 2007.9-2011.7 大连理工大学, 数学学院, 信息与计算科学, 学士.
- 2011.9-2016.7 中国科学院数学与系统科学研究院, 系统所, 应用数学, 博士生, 导师: 支丽红研究员

研究领域

多项式优化, 凸代数几何, 计算代数几何, 凸分析

学术报告

- NIMS应用代数几何会议 韩国, 大田2014.7.15-28
- 符号数值计算会议 华东师范大学, 上海, 2014.7.28-31
- 国际符号与代数计算会议 巴斯大学, 英国, 巴斯, 2015.7.6-9
- SIAM应用代数几何会议 韩国, 大田, 2015.8.3-7

获奖情况

- 2012.9-2013.7 中科院系统所 三好学生
- 2013.9-2014.7 中科院系统所 三好学生
- 2014.9-2015.7 中科院系统所 三好学生
- 2015.9-2016.7 国家奖学金

致 谢

转眼间，我在中国科学院数学与系统科学研究院数学机械化中心度过了五年的时光。这五年来，有辛苦的工作，有成功的喜悦，有失败的沮丧，有太多的人和事值得记忆。在论文完成之际，我感觉非常充实，感慨良多。首先，我感谢我的恩师支丽红研究员。支老师严谨的科研态度，敏锐的眼光和勇于探索的精神让我在学术上找到了正确的研究方向。她的教诲和鞭策深深地激励着我，让我勇于克服在科研中遇到的困难，让我不断地取得进步。支老师积极乐观的生活态度和宽广豁达的胸襟让我学会了做人的道理。支老师为我树立了一生的典范，让我终身受益。

衷心的感谢Mohab Safey El Din教授在学术上对我的帮助。在和Mohab Safey El Din教授合作的过程中，我不仅学习了许多代数理论，同时也感受到了Mohab Safey El Din教授广阔的视野和对科学的热爱与钻研精神。

衷心的感谢郭峰师兄在学术合作中对我的帮助。郭峰师兄一丝不苟的学术态度让我体会到细节的重要性。通过和郭峰师兄的合作与交流，我的代数功底有了显著的提高。

衷心的感谢数学机械化中心的各位老师！特别感谢吴文俊院士、高小山研究员、李子明研究员、李洪波研究员、刘卓君研究员和王定康研究员。没有各位老师昔日的付出，就没有数学机械化中心今日的成就！也感谢周代珍老师和李佳老师的热心帮助！

衷心的感谢陈绍示师兄、梁野师兄、马玥师姐、李子佳师兄、李楠师兄、郭庆东师兄、刘琦师姐以及郝志伟师弟、杨志红师妹和姜文蝶师妹。通过在讨论班上的交流，我掌握了很多相关领域的知识，开阔了研究思路和视野。

最后，谨以此文献感谢我的父母。感谢你们对我的养育之恩和在生活中对我的鼓舞与支持！在此祝愿你们身体健康、万事如意！