密级 _____

中国科学院研究生院

博士学位论文

基于低秩矩阵恢复和半正定规划的多项式优化方法

作者姓名: 马 玥

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

2012年5月

Polynomial Optimization via Low-rank Matrix Completion and Semidefinite Programming

By

Yue Ma

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

Academy of Mathematics and Systems Science
Chinese Academy of Sciences

September, 2012

摘　要

多项式优化问题是优化领域内最根本的问题之一. 在生产实践中, 有大量源于生物工程、控制、信号处理等领域的问题都可以归结为多项式优化问题. 本文着重研究多项式优化领域内的三个问题：精确验证多项式的全局非负性、求大规模多项式系统的实根, 给定具有正维实代数簇的多项式理想 I, 求实根理想 $I(V_{\mathbb{R}}(I))$ 的一组 Gröbner 基.

本文首先讨论如何通过求数目最少的多项式平方和分解来精确验证多项式的全局非负性. 该问题可以转化为低秩对称半正定矩阵的恢复问题. 针对此问题我们提出了一种新的一阶算法—改进的不动点迭代算法 (MFPC-BB), 并给出了算法的收敛性分析. 算法以不动点迭代算法中的算子分裂技术为基础, 通过改进阈值算子以适应对称半正定矩阵的恢复. 同时, 我们还引入 Barzilai-Borwein 技术进行参数的选取, 提高了算法的收敛速度. 在此方法基础上, 我们又提出一种加速的不动点迭代算法 (AFPC-BB), 将收敛速度提高为二次收敛. 数值实验显示 AFPC-BB 算法对于低秩对称半正定矩阵的近似或准确恢复较基于半定规划的方法提速明显, 更适合大规模问题的求解.

我们将多项式系统实根求解的问题转化为低秩矩量矩阵的恢复问题, 并利用 AFPC-BB 算法求解. 同时, 我们给出了算法的收敛性分析和在 Maple 和 Matlab 中的实现 (MMCRSolver). 对于较大规模的多项式系统, 如果只存在一个或少数几个实根, MMCRSolver 能够快速地将它们求解出来. 与此同时, 如果多项式系统有无穷多个实根, MMCRSolver 仍能求出其中部分孤立实根或是在代数流形上的实根.

给定实系数多项式环中具有正维实代数簇的理想 I, 我们提出了一种基于矩量矩阵半正定松弛方法的符号—数值混合算法求理想 I 的实根理想 $I(V_{\mathbb{R}}(I))$. 通过将几何对合理论与半正定矩量矩阵的性质相结合, 我们提出了正维情形下半正定松弛方法终止的判定定理. 基于猜想 5.1, 我们证明了在 δ-正则坐标系下, 判定定理中的条件一定在有限步的半正定松弛内满足, 并给出了实根理想 $I(V_{\mathbb{R}}(I))$ 的 Gröbner 基. 与此同时, 给定半代数集合 $\mathcal{S}$, 我们将算法推广到求 $\mathcal{S}$-根理想 $I(V_{\mathbb{R}}(I) \cap \mathcal{S})$ 的 Gröbner 基.

关键词： 多项式平方和分解, 低秩矩阵恢复, 核范数极小化, 半正定规划, 实根理想

Abstract

Polynomial optimization is one of the fundamental problems in the field of optimization with applications in a large range of areas, including biomedical engineering, control theory, signal processing, etc. This thesis presents a study of some important subclasses of polynomial optimization problems arising from various applications. We focus on the following three problems: exact certificate of global nonnegativity of polynomials, computing real roots of polynomial systems, computing a Gröbner basis of the real radical ideal $I(V_{\mathbb{R}}(I))$ of a positive-dimensional ideal I.

The problem of computing a representation for a real polynomial as a sum of minimum number of squares of polynomials can be casted as finding a symmetric positive semidefinite matrix of minimum rank subject to linear equality constraints. We propose algorithms for solving the minimum-rank Gram matrix completion problem, and show the convergence of these algorithms. Our methods are based on the fixed point continuation method. We also use the Barzilai-Borwein technique and a specific linear combination of two previous iterates to accelerate the convergence of modified fixed point continuation algorithms. Numerical experiments show the effectiveness of our algorithms for computing approximate and exact rational sum of squares decompositions of polynomials with rational coefficients.

Based on the above positive semidefinite matrix completion algorithms, we propose a new algorithm for computing real roots of polynomial equations or a subset of real roots in a given semi-algebraic set described by additional polynomial inequalities. The algorithm is based on using modified fixed point continuation method for solving Lasserre's hierarchy of moment relaxations. We establish convergence properties for our algorithm. For a large-scale polynomial system with only a few real solutions in a given area, we can extract them quickly. Moreover, for a polynomial system with infinitely many real solutions, our algorithm can also be used to find some isolated real solutions or real solutions on

the manifolds.

For an ideal $I \subseteq \mathbb{R}[x]$ with positive-dimensional real variety $V_{\mathbb{R}}(I)$, we propose a symbolic-numeric algorithm to compute a Gröbner base of the real radical ideal $I(V_{\mathbb{R}}(I))$ based on semidefinite relaxation. By using the geometric involutive theory, we prove a certificate for terminating the algorithm. The stopping criterion consists of conditions on ranks of moment matrices and Cartan characters. Based on the conjecture 5.1, we prove that, for a δ-regular coordinate system, these conditions are satisfiable by a finite number of semidefinite relaxations. Moreover, given a semialgebraic set $\mathcal{S}$, we extend our algorithm to compute a numeric Gröbner basis of the $\mathcal{S}$-radical ideal $I(V_{\mathbb{R}}(I) \cap \mathcal{S})$.

Keywords: Sums of squares decomposition of polynomials, low-rank matrix completion, nuclear norm minimization, semidefinite programming, real radical ideal

目　　录

表　　格

第一章　引言

1.1　问题和研究概述

多项式优化问题是优化领域内最根本的问题之一. 在生产实践中, 有大量源于生物医学工程、控制理论、信号处理、量子力学、计算机模拟、语音识别等领域的问题都可以归结为多项式优化问题. 此类问题通常可以写为以下形式：

$$\left.\begin{array}{lll} \min & f(x_1,\ldots,x_n) & \\ \text{s. t.} & g_j(x_1,\ldots,x_n)=0, & j=1,\ldots,s_1, \\ & g_j(x_1,\ldots,x_n)\geq 0, & j=s_1+1,\ldots,s_2, \end{array}\right\} \tag{1.1}$$

其中 $f,\ g_j\in\mathbb{R}[x_1,\ldots,x_n],\ j=1,\ldots,s_2$. 关于多项式优化问题的研究可以追溯到十九世纪初, Hilbert [49] 讨论了非负多项式函数与多项式平方和之间的关系. 在 1900 年法国巴黎的国际数学家大会上, Hilbert 提出了对以后的数学发展产生重大影响的 23 个数学问题, 其中第 17 个问题可叙述如下：对于任意非负多项式 $f\in\mathbb{R}[x_1,\ldots,x_n]$, 是否存在有理函数 $g_1,\ldots,g_s\in\mathbb{R}(x_1,\ldots,x_n)$ 使得 $f=\sum_{i=1}^{s}g_i^2$? 1927 年, 奥地利数学家 Artin [4] 对这一问题给出了肯定的证明并以此为实代数理论的发展奠定了基础. Dezell [35] 于 1984 年给出了解此问题的一个连续的构造性方法. 然而并非所有的非负多项式都具有多项式平方和分解, 例如 Motzkin 多项式 $f=x_1^4x_2^2+x_1^2x_2^4+1-3x_1^2x_2^2$, 其证明参见 [106]. 文献 [14] 中指出, 在所有次数大于或等于 4 的多元多项式集合中, 能够分解为平方和的多项式与非负多项式的比例随着变元数的增加而趋向于 0.

基于上述关于 Hilbert 第 17 问题的讨论, 我们考虑以下问题：如何给出有效的方法判定多项式 $f\in\mathbb{R}[x_1,\ldots,x_n]$ 的全局非负性, 即

$$f(x_1,\ldots,x_n)\geq 0,\ \forall(x_1,\ldots,x_n)\in\mathbb{R}^n. \tag{1.2}$$

这个问题非常重要, 前人对此做了大量的研究. 系统和控制科学中的许多问题最终都会转化为对多项式全局非负性的判定. 对于精确系数的单变元多项式, 我们可以利用代数几何中多项式相异根判别法则 [143] 简单地判断出多项式的全局非负性. 文献 [84] 中证明了当多元多项式 f 的次数大于等于 4 时, (1.2) 是 NP-难

问题. 对于高次多变元多项式非负性的判定, 多数方法在实际计算中并不可行, 例如, 量词消去法 [54].

如果一个多项式可以表示为多项式平方和的形式, 其全局非负性则不言而喻. 文章 [98] 中证明了实系数多项式 $f(x)$ 能分解为 $\mathbb{R}[x]$ 上多项式平方和的充分必要条件为

$$f(x) = [x]_d^T \cdot W \cdot [x]_d, \tag{1.3}$$

其中 $[x]_d$ 为所有次数小于等于 $d = \lceil \deg(f)/2 \rceil$ 的单项式构成的列向量, W 为实对称半正定矩阵, 也称之为 f 的 Gram 矩阵. 因此, 判定多项式的全局非负性问题可以转化为求解带限制条件的优化问题

$$\left.\begin{array}{ll} \min & 1 \\ \text{s. t.} & f(x) = [x]_d^T \cdot W \cdot [x]_d, \\ & W \succeq 0, \quad W^T = W. \end{array}\right\} \tag{1.4}$$

平方和松弛方法首先由 Shor [118] 引入到多项式全局最优化问题的求解中, 后来被扩展到有限制条件情形和有理函数最优化情形. Nesterov [84] 利用优化的理论和方法探索矩量 (Moment) 矩阵锥与非负多项式锥之间的对偶性质. 他证明了对于能够表示为平方和形式的非负多项式构成的锥, 其对偶矩量矩阵锥中的元素满足半正定性. Lasserre [60] 结合实代数几何将多项式优化问题松弛为一系列半正定规划 (Semidefinite Program) 问题, 并证明了此类半正定松弛问题的解收敛到原问题的最优解. Parrilo [90, 91] 也给出了基于半正定规划的构造多项式平方和分解的方法. 与此同时, Waki [135] 等人提出了一系列稀疏半正定松弛方法, 通过探索问题的稀疏性来提高算法的计算效率.

如上所述, 判定多项式的全局非负性可以转化为求解相应的半正定规划问题 (1.4). 而半正定规划问题可以通过 Matlab 中基于内点法 (Interior-point Method) 的软件包高效求解, 例如 SeDuMi [125], SDPT3 [128], SDPNAL [145]. 然而由于 Matlab 只能进行有限精度的计算, 所得结果往往带有较大的数值误差, 所求解只是近似地满足问题的等式或不等式限制. 文献 [55, 56, 93] 利用有理化正交投影, Gauss-Newton 迭代等工具, 将多项式近似平方和分解转化为准确有理系数平方和分解. 但是, 由于 Gauss-Newton 迭代精化过程的计算量与多项式平方和分解中平方数有关, 平方和数越多, Gauss-Newton 迭代的计算量越大. 而基于内点法的半正定规划软件包通常返回满足限制条件的最大秩的 Gram 矩

阵 [98, 定理 1], Gram 矩阵的秩等于平方和数. 这就促使我们考虑如何通过求平方和数最少的多项式平方和分解来精确验证给定多项式的全局非负性. 与此同时, 由于内点法在迭代过程中需要计算并储存相关函数的二阶信息, 当矩阵的维数大于 1000, 限制条件个数大于 6000 时, 内点法将不再适用. 这也促使我们探索更高效的算法来解决较大规模的问题. 事实上, 我们将 SDP 问题 (1.4) 转化为低秩矩阵恢复问题 (Low-rank Matrix Completion), 并在第三章中详细介绍问题的求解方法.

多项式优化领域内的另一个基本而重要的问题是多项式方程组求解. 该问题也是推动代数学发展的原始动力. 早在公元前, 我国古代数学家在研究初等几何问题时就开始使用多项式方程表示图形的边长与面积之间的关系. 在当今复杂科学研究中, 通常也是利用多项式系统来描述一些复杂物体的运动轨迹、压力、势场等量之间的关系. 给定多项式方程组

$$\begin{cases} g_1(x_1,\ldots,x_n) &= 0,\\ g_2(x_1,\ldots,x_n) &= 0,\\ \vdots & \\ g_{s_1}(x_1,\ldots,x_n) &= 0, \end{cases} \tag{1.5}$$

其中 $g_i \in \mathbb{R}[x_1,\ldots,x_n]$, $i=1,\ldots,s_1$. 由于在工程实践中, 变元通常表示某些物理量, 所以大量的实际问题最终转化为多项式方程组的实根求解问题, 或是求满足如下不等式限制条件的实根.

$$\begin{cases} g_{s_1+1}(x_1,\ldots,x_n) &\geq 0,\\ g_{s_1+2}(x_1,\ldots,x_n) &\geq 0,\\ \vdots & \\ g_{s_2}(x_1,\ldots,x_n) &\geq 0, \end{cases} \tag{1.6}$$

求解方法大体上分为两种: 符号方法和数值方法. 这两类方法各有优缺点. 符号解法能够精确地求出多项式系统的所有实根, 但速度较慢, 适合解决中小问题. 符号方法主要有: Gröbner 基方法, Ritt-Wu 特征列方法和多元结式的方法. Gröbner 基方法是 1965 年由 Buchberger 在其博士论文中提出的 [16]. Ritt-Wu 方法是美国数学家 Ritt [107] 在 20 世纪 50 年代引入的, 而后由吴文俊院士 [132] 对其进行改进和发展. 以上两种方法的基本思想都是在零点不变的前提

下对多项式方程组进行消元, 并转化为等价的三角形式便于求解. 多元结式方法的基本原理是从给定的方程组构造出包含多个方程的导出方程组, 将其看作线性方程组, 从而利用已有的“线性方法”来研究原非线性方程组的解. 除此之外, Collins [28] 提出了柱形代数分解的方法来求解多项式方程组. 该方法将任一代数多项式系统剖分成有限多个互不相交的半代数集, 剖分之后每一个半代数集胞腔上定义多项式的符号不变. 根据不同胞腔的符号变化情况进行实根隔离. 还有基于 Descartes 符号法则的实根隔离方法, 见 [110, 141, 142].

相对符号方法而言, 数值方法较为成熟, 速度较快, 可以解决的问题的规模较大, 但其往往只能求出有限精度的近似解. 数值方法主要有二分法、牛顿法和同伦方法 [134] 等. 近年来, Chesi [23–25] 及 Lasserre [46, 63–66] 等人致力于利用半正定规划等数值优化算法来求解多项式系统的实根. 正如文献 [64] 所述, 半正定规划技术及矩量松弛理论的优点在于它直接探索问题的实代数特性而求得方程组的实根, 从而避免计算与复根相关的信息. 文献 [64–66] 提出的基于矩量矩阵半正定松弛的方法是求解一系列 SDP 问题

$$\left.\begin{array}{ll} \min & 1 \\ \text{s. t.} & y_0 = 1, \\ & M_t(y) \succeq 0, \\ & M_{t-d_j}(g_j\, y) = 0, \quad j = 1, \ldots, s_1, \\ & M_{t-d_j}(g_j\, y) \succeq 0, \quad j = s_1+1, \ldots, s_2, \end{array}\right\} \tag{1.7}$$

其中 $d_j := \lceil \deg(g_j)/2 \rceil, j = 1, \ldots, s_2$. Lasserre [64] 等还给出了基于求解上述半正定松弛的问题 (1.7) 终止的判定定理:

定理 1.1. *[64] 对于零维多项式系统 (1.5), 设 $t \geq d$, $M_t(y)$ 为问题 (1.7) 的可行解, 且满足 $\operatorname{rank} M_t(y)$ 最大. 如果存在 $d \leq k \leq t$ 满足*

$$\operatorname{rank} M_k(y) = \operatorname{rank} M_{k-d}(y), \tag{1.8}$$

其中 $d = \max_{1\leq j\leq s_2} d_j$, $d_j = \lceil \deg(g_j)/2 \rceil, j = 1, \ldots, s_2$, 那么 $\langle \ker M_k(y) \rangle = I(V_{\mathbb{R}}(I))$. 多项式系统 (1.5),(1.6) 实根的个数等于 $\operatorname{rank} M_k(y)$.

半正定规划问题 (1.7) 可以由内点法求解. 虽然利用内点法返回的最大秩矩量矩阵 $M_t(y)$, 我们可以求得零维多项式系统的全部实根. 但是, 当秩条件 (1.8) 成立时, t 值通常较大. 由于 t 阶矩量矩阵的维数 $m = \binom{n+t}{t}$, 此

时 (1.7) 中等式限制条件个数 $p = \sum_{j=1}^{s_1} \frac{1}{2}\binom{n+t-d_j}{n}\left(\binom{n+t-d_j}{n}+1\right)$. 随着 t 值增大, 问题 (1.7) 的规模也会大幅度增加, 使得通常的半正定规划软件包, 如 SeDuMi [125], SDPT3 [128], 都无法计算出结果. 更重要的问题是如果给定多项式系统有无穷多个实根, 无论 t 值多大, 由内点法返回的最大秩矩量矩阵均不满足秩条件 (1.8). 针对上述难题, 我们将问题 (1.7) 转化为低秩矩量矩阵的恢复问题, 并结合矩阵恢复方法来求解较大规模多项式系统 (1.5),(1.6) 的部分或全部的实根 (详见第四章).

以上介绍的符号解法和数值解法大多适用于多项式方程组只有有限多个实根的情形. 如果给定的多项式系统有无穷多的实根, 基于多项式理想 $I = \langle g_1, \ldots, g_{s_1}\rangle$ 与代数簇 $V(I)$ (方程组的公共零点集) 的对应关系, 多项式方程组实根求解问题可以转化为对理想 I 的实根理想 $\sqrt[\mathbb{R}]{I}$ 的研究. 下面的定理给出了理想 $I(V_{\mathbb{R}}(I))$ 与实根理想 $\sqrt[\mathbb{R}]{I}$ 之间的关系.

定理 1.2 (实零点定理). *[15]* 理想 $I \subseteq \mathbb{R}[x]$, $\sqrt[\mathbb{R}]{I} = I(V_{\mathbb{R}}(I))$.

对于实根理想 $I(V_{\mathbb{R}}(I))$ 的计算要比根理想的计算更加困难. 对于零维多项式系统, Lasserre 等人提出了基于半正定规划的数值方法 [64, 66] 和符号–数值混合方法 [63, 65] 来计算实根理想 $I(V_{\mathbb{R}}(I))$ 的一组边界基 (Border Basis) 或 Gröbner 基. 对于正维多项式系统, Becker 和 Neuhaus [12] 提出了一种基于理想的准素分解的方法来求实根理想 $I(V_{\mathbb{R}}(I))$, 相关工作可参见 [87, 141, 144]. 除此之外, 还有一类方法基于实代数几何中的关键点方法, 能够在实代数簇的每一个连通分支上求出一点, 见 [5, 6, 8, 9, 113]. 但是, 随着问题规模的增大, 此类针对正维多项式系统的符号方法在计算过程中会出现表达式迅速膨胀, 内存需求增加, 计算速度降低等问题, 因而无法满足实际应用的需求. 我们提出了一种基于矩量矩阵半正定松弛 (1.7) 的符号–数值混合算法求理想 I 的实根理想 $I(V_{\mathbb{R}}(I))$ 的关于序 $\prec_{\text{tdeg}}$ 的一组 Gröbner 基 (详见第五章).

1.2　论文的结构和主要结果

本文的结构及主要贡献如下：

第二章中, 我们介绍了低秩矩阵恢复, 多项式全局非负性与平方和表示以及实代数几何基础知识.

第三章中, 我们讨论如何通过求数目最少的多项式平方和分解精确验证多项式的全局非负性. 此问题可以转化为对称半正定矩阵秩的极小化问题

$$\left.\begin{array}{ll}\min & \mathrm{rank}W \\ \text{s. t.} & f(x)=[x]_d^T\cdot W\cdot[x]_d, \\ & W\succeq 0,\, W^T=W,\end{array}\right\} \tag{1.9}$$

问题 (1.9) 可以凸松弛为仿射限制条件下矩阵核范数极小化问题

$$\left.\begin{array}{ll}\min & \|W\|_* \\ \text{s. t.} & \mathcal{A}(W)=b, \\ & W\succeq 0,\, W^T=W,\end{array}\right\} \tag{1.10}$$

其中 $\|W\|_*$ 为矩阵的核范数 (Nuclear Norm), 即矩阵奇异值的和. 针对半正定规划方法存在计算量大、速度慢、能够处理的矩阵规模小等问题, 我们提出了一种新的求解 Gram 矩阵核范数极小化问题 (1.10) 的一阶算法——改进的不动点迭代算法 (MFPC-BB), 并给出了算法的收敛性分析和在 Maple 和 Matlab 中的实现. 我们的算法以不动点迭代算法中的算子分裂技术为基础, 通过改进阈值算子 $\mathcal{T}$ 以适应对称半正定矩阵的恢复. 同时, 我们引入 Barzilai-Borwein 技术来进行步长参数的选取, 从而提高算法的收敛速度. 在此方法基础上, 我们又提出一种加速的不动点迭代算法 (AFPC-BB). 它既保持了 MFPC-BB 算法的简单易实现的特点, 又将收敛速度提高为二次收敛. 数值实验显示 AFPC-BB 算法对于低秩 Gram 矩阵的近似或准确恢复较基于半正定规划的方法提速明显, 更适合大规模问题的求解. 例如, 对于单项个数超过十二万的多元多项式, 利用 AFPC-BB 算法在一小时左右即可将与其相应的维数为 1500×1500 秩为 50 的 Gram 矩阵准确地恢复出来.

第四章中, 我们研究如何快速求解大规模多项式系统的实根. 通过将 Lasserre 提出的半正定规划模型 (1.7) 中的目标函数改为矩量矩阵的核范数 $\|M_t(y)\|_*$, 从而将多项式系统实根求解的问题转化为求矩量矩阵核范数极小化问题

$$\left\{\begin{array}{lll}\min & \|M_t(y)\|_* & \\ \text{s. t.} & y_0=1, & \\ & M_t(y)\succeq 0, & \\ & M_{t-d_j}(g_j\,y)=0, & j=1,\ldots,s_1, \\ & M_{t-d_j}(g_j\,y)\succeq 0, & j=s_1+1,\ldots,s_2.\end{array}\right. \tag{1.11}$$

对于半正定限制条件 $M_{t-d_j}(g_j y) \succeq 0,\ j = s_1+1, \ldots, s_2$, 通过引入松弛矩阵变元 Z_j, 将其转化为等式限制条件

$$M_{t-d_j}(g_j y) = Z_j, \quad Z_j = Z_j^T, \quad Z_j \succeq 0. \tag{1.12}$$

关于矩量矩阵 $M_t(y)$ 和矩阵变元 Z_j, 我们利用第三章给出的 AFPC-BB 算法交替地求上述带限制条件的低秩矩量矩阵的恢复问题 (1.11). 与此同时, 我们给出了算法的收敛性分析和在 Maple 和 Matlab 中的实现 (MMCRSolver). 如果对于 t 阶问题 (1.11), MMCRSolver 返回的低秩矩量矩阵的某个 k 阶子块 $M_k(y)$ 满足判定条件 (1.8), 那么可以通过求矩阵 $M_{k-1}(y)$ 的像空间的一组基和相应乘法矩阵 (Multiplication Matrix) 的公共特征向量得到多项式方程组的实根 [29, 46, 101].

目前我们的算法无法保证求出多项式系统的全部实根. 对于较大规模的多项式系统, 如果只存在一个或少数几个实根, MMCRSolver 能够快速地将它们求解出来. 如果多项式系统有无穷多个实根, 我们仍能求出部分孤立实根或是代数流形上的实根. 数值实验显示, 对某些利用半正定规划方法难以求解的例子, MMCRSolver 也能快速地求出其全部或部分的实根 (见表 4.1).

第五章中, 给定具有正维实代数簇的多项式理想 $I = \langle h_1, \ldots, h_m \rangle \subseteq \mathbb{R}[x]$ (无穷多个根), 我们提出了一种基于矩量矩阵半正定松弛的符号–数值混合算法求实根理想 $I(V_{\mathbb{R}}(I))$. 通过将几何对合理论与半正定矩量矩阵的性质相结合, 我们给出了正维情形下半正定松弛方法终止的判定定理. 记 $\mathrm{vec}(p) = (p_\alpha)_{\alpha \in \mathbb{N}^n}$ 为多项式 p 的系数向量. 令

$$\ker M_t(y) := \{p \in \mathbb{R}[x]_t \,|\, M_t(y)\mathrm{vec}(p) = 0\},$$
$$\mathcal{K}_t^{gen} := \{y \in \mathbb{R}^{\mathbb{N}^n_{2t}} \,|\, y_0 = 1, M_t(y) \succeq 0, M_{t-d_j}(h_j y) = 0, j = 1, \ldots, m, M_t(y) \text{ 秩最大}\}.$$

定理 1.3. *在 δ-正则坐标系下, 如果存在整数 (t, ℓ), $t \geq 2d, 1 \leq \ell \leq t - 2d$ 及 $y_1 \in \mathcal{K}_t^{gen}, y_2 \in \mathcal{K}_{t+1}^{gen}$ 满足下列条件*

$$\mathrm{rank}\, M_{t-\ell}(y_1) = \mathrm{rank}\, M_{(t+1)-(\ell+1)}(y_2), \tag{1.13}$$

$$\sum_{j=1}^{n} j\alpha_{t-\ell}^{(j)} \text{ 对于 } M_{t-\ell}(y_1) = \mathrm{corank}\, M_{(t+1)-\ell}(y_2) - \mathrm{corank}\, M_{(t+1)-(\ell+1)}(y_2), \tag{1.14}$$

那么 $\ker M_{t-\ell}(y_1)$ 是实根理想 $\sqrt[\mathbb{R}]{I}$ 关于序 $\prec_{\text{tdeg}}$ 的一组 *Gröbner* 基, 即

$$\langle \ker M_{t-\ell}(y_1) \rangle = \sqrt[\mathbb{R}]{I}. \tag{1.15}$$

基于猜想 5.1, 我们证明了在 δ-正则坐标系下, 定理 1.3 中条件 (1.13)-(1.14) 一定在有限步的半正定松弛内满足, 并给出了实根理想 $I(V_{\mathbb{R}}(I))$ 关于序 $\prec_{\text{tdeg}}$ 的一组 Gröbner 基. 条件 (1.13)-(1.14) 可以作为 Flat Extension 定理 5.1 中条件 (5.1) 在正维情形下的推广. 与此同时, 给定的半代数集合 $\mathcal{S} = \{x \in \mathbb{R}^n \mid f_1(x) \geq 0, \ldots, f_s(x) \geq 0\}$, 我们将算法推广到求理想 I 的 $\mathcal{S}$-根理想 $I(V_{\mathbb{R}}(I) \cap \mathcal{S})$ 的一组 Gröbner 基.

在最后一章, 我们总结了已有的工作成果并讨论了以后继续努力的方向.

第二章　预备知识

本章中, 我们简要地介绍低秩矩阵恢复, 多项式全局非负性与平方和表示, 实代数几何基础知识.

2.1　低秩矩阵恢复

低秩矩阵恢复问题是指, 对于矩阵 $M \in \mathbb{R}^{n_1 \times n_2}$, 已知其中 m 个元素 $\{M_{ij} : (i,j) \in \Omega\}$ $(m < n_1 n_2)$, 如何将未知元素合理准确地恢复出来. 矩阵恢复问题一个著名的应用是 Netflix 系统 [120]. Netflix 是世界上最大的在线影片租赁服务商, 从 2006 年 10 月份开始举办 Netflix 大奖赛. 它公开了大约一亿个 1～5 级的匿名电影评级, 来自约 48 万个客户对 1.8 万部电影的评价. 数据中所有个人信息都被删除, 仅包含了影片名称、评价星级和评价日期, 没有任何评价内容. 比赛要求参赛者预测 Netflix 客户分别喜欢什么影片, 要把预测的效率相对原推荐系统 Cinematch 提高百分之十以上. 这是一个典型的矩阵恢复问题, 即矩阵的每一行对应某个用户对电影的评级, 每一列表示某电影在所有用户中的评级, 但是每个用户只可能对部分电影进行评价. 因此, 可以通过矩阵恢复得出用户对每部电影的喜好程度. 矩阵恢复的问题还出现在许多工程及其应用科学领域中, 例如, 机器学习 [1–3], 控制 [77] 及计算机视觉中 [130]. 在很多的具体问题中, 信号或者数据经常面临缺失、损坏、受噪声污染等等问题. 如何在各种情况下得到干净、准确、结构性良好的数据, 就是矩阵恢复要解决的问题.

文献 [20] 中证明可以通过求解如下优化问题来实现低秩矩阵的恢复:

$$\left.\begin{array}{ll} \min & \operatorname{rank}(X) \\ \text{s. t.} & X_{ij} = M_{ij}, \quad (i,j) \in \Omega. \end{array}\right\} \tag{2.1}$$

在通常情况下, 非凸函数 $\operatorname{rank}(\cdot)$ 的组合性质导致问题 (2.1) 是 NP-难问题, 已有的精确算法都具有双指数形式的时间复杂度 [26]. 为了解决此问题, Fazel [38] 证明了对于 $W \in \{W \in \mathbb{R}^{n_1 \times n_2} : \|W\|_2 \leq 1\}$, 函数 $\operatorname{rank}(W)$ 的凸包络 (Convex Envelop), 即满足 $f(W) \leq \operatorname{rank}(W)$ 的最大凸函数 f 等于矩阵 W 的核范数, 即矩阵 W 的奇异值的和, 记为 $\|W\|_*$. 因此, 函数 $\operatorname{rank}(\cdot)$ 用其凸包络代

替, 问题 (2.1) 转化为凸优化问题

$$\left.\begin{array}{ll}\min & \|X\|_* \\ \text{s. t.} & X_{ij} = M_{ij}, \quad (i,j) \in \Omega.\end{array}\right\} \tag{2.2}$$

Candès 和 Recht [20] 证明当采样点集 $\{M_{ij} : (i,j) \in \Omega\}$ 满足给定条件, 如果采样数 m 满足

$$m \geq Cn^{6/5} r \log n,$$

其中 $n = \max(n_1, n_2)$, r 为矩阵的维数, C 为常数. 那么问题 (2.2) 的最优解唯一且等于原问题 (2.1) 的最优解, 而且得到的矩阵等于 M 的概率满足

$$p \geq 1 - cn^{-3}.$$

其中 c 为常数.

如果已知的不是采样点而是关于矩阵中元素的仿射条件, 那么低秩矩阵恢复问题 (2.1) 即可转化为仿射限制条件下矩阵秩的极小化问题

$$\left.\begin{array}{ll}\min & \operatorname{rank}(X) \\ \text{s. t.} & \mathcal{A}(X) = b,\end{array}\right\} \tag{2.3}$$

其中线性算子 $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^p$, $b \in \mathbb{R}^p$. 此问题同样可以凸松弛为

$$\left.\begin{array}{ll}\min & \|X\|_* \\ \text{s. t.} & \mathcal{A}(X) = b,\end{array}\right\} \tag{2.4}$$

上述问题 (2.3) 和 (2.4) 的等价性条件, Recht 等在文献 [100] 中做了详细的讨论. 他们提出了矩阵限制等距性质 (Restricted Isometry Property). 对任意的整数 r, $1 \leq r \leq \max(n_1, n_2)$, 定义 r-限制等距常数 $\delta_r(\mathcal{A})$ 为使得下式

$$(1 - \delta_r(\mathcal{A}))\|X\|_F \leq \|\mathcal{A}(X)\|_2 \leq (1 - \delta_r(\mathcal{A}))\|X\|_F, \tag{2.5}$$

对所有的秩小于等于 r 的矩阵 X 均成立的 $\delta_r(\mathcal{A})$ 的最小值. 他们还证明, 对于 $r \geq 1$, 当 $\delta_{5r}(\mathcal{A}) < 1/10$ 时, 问题 (2.3) 和 (2.4) 的解相等且唯一. 关于矩阵恢复理论的可行性研究可参见 [21, 76, 82, 99].

针对上述凸优化问题 (2.4), 已存在多种求解方法. 对于规模较小的问题, 可以将其转化成标准的半正定规划问题, 并利用内点法 [17, 18, 39, 70, 103, 122] 或

投影方法 [47, 48, 57, 74, 89, 145] 有效地求解, 且能够达到较高的精度. 但是, 由于通常基于内点法的半定规划软件包在运算过程中需要计算并储存相关函数的二阶信息, 当矩阵的维数大于 1000 或限制条件的个数大于 6000 时, 内点法将不再适用. 当矩阵的维数高达 5000, 限制条件的个数大于 10^5 时, 投影算法也难以计算出结果. 最近, 许多学者致力于发展一阶快速算法来求解矩阵恢复问题, 例如, SVT [19]、FPC [43, 71]、APG [129] 等. 与此同时, 一些具有较好的收敛速度 $\mathcal{O}(1/k^2)$ (k 为迭代数) 的加速的梯度算法也受到广泛关注, 例如 [11, 50, 72, 73, 83, 85, 86, 129, 131]. 虽然此类一阶算法求得的解在数值精度上不及内点法, 但是, 此类方法能解决维数高达 10^5 的较大规模矩阵的恢复问题.

2.2　多项式全局非负性与平方和表示

判断函数的全局非负性在系统科学和控制论的许多领域中有着重要的应用, 考虑多项式函数 $f(x) \in \mathbb{R}[x] = \mathbb{R}[x_1, \ldots, x_n]$. 如果多项式 $f(x)$ 满足

$$f(x) \geq 0, \forall x \in \mathbb{R}^n,$$

则称其为半正定 (Positive Semidefinite) 多项式. 通常情况下, 判定多项式的半正定性是非常困难的任务. 如果存在多项式 $u_i(x) \in \mathbb{R}[x]$ 使得

$$f(x) = \sum_j u_j(x)^2 \in \sum \mathbb{R}[x]^2$$

成立, 则称 $f(x)$ 具有平方和分解. 关于非负多项式与多项式平方和之间的关系的研究可以追溯到 Hilbert 十七问题. 如果多项式 $f(x)$ 具有平方和分解, 显然 $f(x)$ 是半正定的; 反之, 如果 $f(x)$ 是半正定的, 且满足如下三个条件之一:

1. $f(x)$ 为单变元多项式;
2. $\deg\ f(x) = 2$;
3. $\deg\ f(x) \leq 4$ 且只含有两个变元;

则 $f(x)$ 具有平方和分解. 然而一般情形下非负多项式不一定能分解为平方和的形式. Motzkin [80] 给出了一个简单的反例

$$M(x_1, x_2, x_3) = x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^6 - 3x_1^2 x_2^2 x_3^2 \geq 0.$$

1927 年, 奥地利数学家 Emil Artin [4] 证明了对于任意的半正定多项式 $f(x)$, 存在多项式 $v(x)$ 使得 $f(x)v(x)^2$ 可以分解为多项式平方和. Polya [95] 于 1928 年证明了如果 $f(x)$ 是偶数次正定多项式, 那么对充分大的正整数 r, $(\Sigma\, x_i^2)^r f$ 是 $\mathbb{R}[x_1,\ldots,x_n]$ 上单项式的平方和. Reznick [105] 于 1995 年证明了: 假设 $f(x)$ 是 $\mathbb{R}[x_1,\ldots,x_n]$ 中次数为 m 的齐次正定多项式, $\epsilon(f)$ 是 f 在单位球上的下确界和上确界的比, 如果 $r \geq \frac{nm(m-1)}{4\log 2\epsilon(f)} - \frac{n+m}{2}$, 那么 $(\Sigma_{i=1}^n x_i^2)^r f$ 是 $\mathbb{Q}[x_1,\ldots,x_n]$ 上线性形的 (m+2r) 次幂的非负 $\mathbb{R}$-线性组合. 更多相关介绍可参见文献 [106].

将 $\mathbb{R}[x]$ 中所有能表示成平方和形式的多项式的集合记为

$$\sum \mathbb{R}[x]^2 := \Big\{ f \in \mathbb{R}[x] \mid f = \sum_i g_i^2, g_i \in \mathbb{R}[x] \Big\}.$$

定理 2.1. *[98] 实系数多项式 $f(x)$ 能分解为 $\mathbb{R}[x]$ 上多项式平方和的充分必要条件为 $f(x)$ 可以表示为*

$$f(x) = [x]_d^T \cdot W \cdot [x]_d, \tag{2.6}$$

其中 $[x]_d$ 为所有次数小于等于 $d = \lceil \deg(f)/2 \rceil$ 的单项式构成的列向量, W 为实对称半正定矩阵, 也称之为 f 的 Gram 矩阵.

如果多项式 f 是稀疏的, 那么向量 $[x]_d$ 和矩阵 W 通常也是稀疏的. 我们可以通过分析牛顿多面体 (Newton Polytope) [27, 104] 来减小问题的规模. 满足等式条件 (2.6) 的矩阵 W 不唯一, 并构成全体对称矩阵集合 $\mathbb{S}$ 的一个仿射子空间

$$\mathcal{X} = \{W \mid W^T = W, f(x) = [x]_d^T \cdot W \cdot [x]_d\}. \tag{2.7}$$

如果仿射子空间 $\mathcal{X}$ 与对称半正定矩阵锥 $\mathbb{S}_+$ 的交集非空, $f(x)$ 即可分解为平方和形式. 如果矩阵 W 的元素是有理数, 则 f 能够分解为 $\mathbb{Q}[x]$ 中的多项式平方和.

多项式的平方和分解问题等价于求半正定规划

$$\left.\begin{array}{ll} \min & 1 \\ \text{s. t.} & f(x) = [x]_d^T \cdot W \cdot [x]_d, \\ & W \succeq 0, \quad W^T = W. \end{array}\right\} \tag{2.8}$$

的可行解. 关于如何利用半正定规划方法求多项式的平方和分解, 参见 [55, 56, 60, 61, 90, 91, 119].

例 2.1. 考虑如下例子 $f(x_1,x_2)=2x_1^4+2x_1^3x_2-x_1^2x_2^2+5x_2^4$,

$$\begin{aligned} f(x_1,x_2) &= 2x_1^4+2x_1^3x_2-x_1^2x_2^2+5x_2^4 \\ &= \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix}^T \begin{pmatrix} 2 & -\lambda & 1 \\ -\lambda & 5 & 0 \\ 1 & 0 & -\lambda+2\lambda \end{pmatrix} \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix}. \end{aligned}$$

令 $\lambda=3$, 对矩阵 W 做 Cholesky 分解,

$$W=\begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \end{pmatrix}^T \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \end{pmatrix},$$

因此, f 的平方和分解为

$$f(x_1,x_2)=\frac{1}{2}(2x_1^2-3x_2^2+x_1x_2)^2+\frac{1}{2}(x_2^2+3x_1x_2)^2.$$

由以上例子可以看出, 求半正定 Gram 矩阵中的参数 λ 实际上是矩阵恢复的过程. 我们将在第三章中详细讨论如何利用矩阵恢复的方法和技巧来求多项式精确的平方和分解.

2.3　实代数几何基础

实代数几何领域中最重要的研究对象之一为实代数簇 (Real Variety), 这也是本文研究的重点.

给定域 $\mathbb{K}=\mathbb{R}$ 或 $\mathbb{C}$, 其上的 n 元多项式环记为 $\mathbb{K}[x]:=\mathbb{K}[x_1,\ldots,x_n]$. 对于整数 $t\geq 0$, $\mathbb{K}[x]_t$ 表示所有次数小于等于 t 的多项式构成的集合. $\mathbb{N}$ 表示非负整数集. 对于 $t\in\mathbb{N}$, $\mathbb{N}^n_t:=\{\alpha\in\mathbb{N}^n \mid |\alpha|:=\sum_{i=1}^n\alpha_i\leq t\}$. 对于 $\alpha\in\mathbb{N}^n$, $x^\alpha=x_1^{\alpha_1}\cdots x_n^{\alpha_n}$, 此单项式的全次数为 $|\alpha|:=\sum_{i=1}^n\alpha_i$. 记 $\mathbb{T}^n:=\{x^\alpha\mid\alpha\in\mathbb{N}^n\}$ 为所有单项式的集合, $\mathbb{T}^n_t:=\{x^\alpha\mid\alpha\in\mathbb{N}^n_t\}$ 包含全次数小于等于 t 的单项式. 给定 $\mathbb{T}^n$ 上的一个单项序 $\prec$, 多项式 $p\in\mathbb{K}[x]$ 可以写为 $p=\sum_{\alpha\in\mathbb{N}^n}p_\alpha x^\alpha$, 其中只有有限多个系数 $p_\alpha\in\mathbb{K}$ 非零. 记 $\mathrm{vec}(p):=(p_\alpha)_{\alpha\in\mathbb{N}^n}$ 为多项式 p 的系数向量.

令 $P=\{p_1,\ldots,p_m\}\subseteq\mathbb{K}[x]=\mathbb{K}[x_1,\ldots,x_n]$. $I=\langle p_1,\ldots,p_m\rangle$ 是由 P 中多项式生成的理想. 理想 I 的代数簇为

$$V_{\mathbb{C}}(I):=\{x\in\mathbb{C}^n\mid f(x)=0,\forall\, f\in I\}.$$

实代数簇为

$$V_{\mathbb{R}}(I) = V_{\mathbb{C}}(I) \cap \mathbb{R}^n.$$

令 $V \subset \mathbb{C}^n$ 为代数簇, 定义

$$I(V) = \left\{ f \in \mathbb{C}[x] \mid \text{任给点} \, \xi \in V, \ f(\xi) = 0 \right\},$$

为 V 的消逝理想 (Vanishing Ideal).

定义 2.1. [75] 理想的实根 (Real Radical) 定义为

$$\sqrt[\mathbb{R}]{I} := \{ f \in \mathbb{R}[x] \mid f^{2k} + \sigma \in I \, \text{对某个} \, \sigma \in \sum \mathbb{R}[x]^2, k \in \mathbb{N} \backslash \{0\} \}.$$

如果 $I = \sqrt[\mathbb{R}]{I}$, 称 I 为实根理想 (Real Radical Ideal). 通常情况下, 有

$$I \subseteq \sqrt[\mathbb{R}]{I} \subseteq I(V_{\mathbb{R}}(I)).$$

下面的引理给出了实根理想的一个简单的描述.

引理 2.2. *理想 $I \in \mathbb{R}[x]$ 为实根理想的充分必要条件为*

$$\forall p_i \in \mathbb{R}[x], \sum_i p_i^2 \in I \Rightarrow p_i \in I.$$

与 Hilbert 零点定理, 即 $\sqrt{I} = I(V_{\mathbb{C}}(I))$ 相类似, 下面我们给出理想 $I(V_{\mathbb{R}}(I))$ 与理想 I 的实根 $\sqrt[\mathbb{R}]{I}$ 之间的关系.

定理 2.3 (实零点定理)**.** *[15] 对理想 $I \subseteq \mathbb{R}[x]$, $\sqrt[\mathbb{R}]{I} = I(V_{\mathbb{R}}(I))$.*

给定半代数集合

$$\mathcal{S} = \{ x \in \mathbb{R}^n \mid f_1(x) \geq 0, \ldots, f_s(x) \geq 0 \},$$

其中 $f_1, \ldots, f_s \in \mathbb{R}[x]$. $\mathcal{S}$-代数簇 $V_{\mathcal{S}}(I)$ 定义为实代数簇 $V_{\mathbb{R}}(I)$ 与给定的半代数集 $\mathcal{S}$ 的交集 $V_{\mathcal{S}}(I) = V_{\mathbb{R}}(I) \cap \mathcal{S}$. 我们将乘积 $f_1^{e_1} f_2^{e_2} \cdots f_s^{e_s}$, $e_i \in \{0, 1\}$ 记为 $\underline{f}^e$.

定义 2.2. 理想 I 的 $\mathcal{S}$-根定义为

$$\sqrt[\mathcal{S}]{I} := \left\{ p \in \mathbb{R}[x] \mid p^{2m} + \sum_{e \in \{0,1\}^k} \sigma_e \underline{f}^e \in I \ \text{对某个} \, \sigma_e \in \sum \mathbb{R}[x]^2, m \in \mathbb{N} \backslash \{0\} \right\}.$$

如果 $I = \sqrt[\mathcal{S}]{I}$, 那么理想 I 称为 $\mathcal{S}$-根理想.

下面两个定理来自Stengel [123], 说明 $\sqrt[\mathcal{S}]{I}$ 是理想, 并将理想 $I(V_{\mathbb{R}}(I) \cap \mathcal{S})$ 与 $\mathcal{S}$-根联系起来.

引理 2.4. *[123, 引理 1]* $\sqrt[\mathcal{S}]{I}$ *为* $\mathcal{S}$*-根理想.*

定理 2.5. *[123, 定理 1][半代数零点定理] 对任意理想* $I \in \mathbb{R}[x]$,

$$\sqrt[\mathcal{S}]{I} = I(V_{\mathcal{S}}(I)) = I(V_{\mathbb{R}}(I) \cap \mathcal{S}).$$

第三章　低秩 Gram 矩阵恢复

3.1　引言

问题 3.1. 给定次数等于 $2d$ 的多项式 $f(x) \in \mathbb{Q}[x_1, \ldots, x_n]$, 计算其精确的平方和分解, 即

$$f(x) = \sum_{i=1}^{k} f_i^2(x), \quad f_i(x) \in \mathbb{Q}[x_1, \ldots, x_n],$$

使得平方和数 k 最小.

满足等式条件

$$f(x) = [x]_d^T \cdot W \cdot [x]_d, \tag{3.1}$$

的所有矩阵 W 构成全体对称矩阵集合 $\mathbb{S}$ 的一个仿射子空间. 当此仿射子空间与对称半正定矩阵锥的交集非空时, $f(x)$ 即可分解为平方和形式. 由于向量 $[x]_d$ 中的元素不是代数无关的, 则 W 不唯一. 问题 3.1 可以转化为求满足等式限制条件 (3.1) 且秩最小的 Gram 矩阵

$$\left.\begin{array}{ll} \min & \operatorname{rank}(W) \\ \text{s. t.} & f(x) = [x]_d^T \cdot W \cdot [x]_d, \\ & W \succeq 0,\ W^T = W. \end{array}\right\} \tag{3.2}$$

对于单变元情形, Pourchet [97] 证明了每个非负的单变元有理系数多项式都能分解为 $\mathbb{Q}[x]$ 中五个多项式的平方和. 因此, 当 $n = 1$ 时, 满足条件 (3.2) 的 Gram 矩阵的最小秩的上界为 5. 对于多变元情形, Pfister [94] 证明了 $\mathbb{R}[x_1, \ldots, x_n]$ 中的半正定多项式都能分解为 2^n 个实系数有理函数的平方和. 尽管不是所有非负多项式都能分解成平方和的形式 (例如 Motzkin 多项式), 但是, 文献 [56] 中给出的非负多项式在乘以恰当的多项式后都可以分解为 10 项以内的有理多项式平方和. 这也是我们研究问题 (3.1) 的动机之一.

在通常情况下, 非凸函数 $\operatorname{rank}(\cdot)$ 的组合性质导致上述问题是 NP-难问题, 已有的精确算法都具有双指数形式的时间复杂度 [26]. 为了解决此问题, Fazel [38] 证明了对于 $W \in \{W \in \mathbb{R}^{n_1 \times n_2} : \|W\|_2 \leq 1\}$, 函数 $\operatorname{rank}(W)$ 的凸包络,

即满足 $f(W) \le \mathrm{rank}(W)$ 的最大凸函数 f 等于矩阵 W 的核范数, 即矩阵 W 的奇异值的和, 记为 $\|W\|_*$. 因此, 函数 $\mathrm{rank}(\cdot)$ 用其凸包络代替, 从而将问题 (3.2) 转化为凸优化问题.

将问题 (3.2) 中等式限制条件的右端展开并化简, 得到一组关于矩阵 W 中元素的线性方程组

$$\mathcal{A}(W) = b, \tag{3.3}$$

其中 $b = (b_1, \ldots, b_p) \in \mathbb{R}^p$, b_i 为 $f(x)$ 中单项式 $x^{\alpha_i} = x_1^{\alpha_{i,1}} \cdots x_n^{\alpha_{i,n}}$ 的系数. 线性算子 $\mathcal{A} : \mathbb{S}^m \to \mathbb{R}^p$ 作用到 W 上定义为矩阵内积的形式 $\langle A_i, W\rangle := \mathrm{Tr}(A_i^T W)$, 其中 $A_1, \ldots, A_p \in \mathbb{S}^m$. 利用 $[x]_d$ 中单项式的指数向量作为矩阵的指标, 则矩阵 A_i 中的元素定义如下: 当 $\beta_i + \gamma_i = \alpha_i$ 时, 矩阵中 β_i 行 γ_i 列的元素为 1, 其余的元素均为 0. 算子 $\mathcal{A}$ 的伴随算子记为 $\mathcal{A}^* : \mathbb{R}^p \to \mathbb{S}^m$. 因此, 矩阵秩的极小化问题 (3.2) 可以松弛为矩阵核范数极小化问题

$$\left.\begin{array}{ll} \min & \|W\|_* \\ \text{s. t.} & \mathcal{A}(W) = b, \\ & W \succeq 0,\ W^T = W. \end{array}\right\} \tag{3.4}$$

等式限制条件 $\mathcal{A}(W) = b$ 仍可做松弛处理, 问题相应地转化为

$$\left.\begin{array}{ll} \min & \|W\|_* \\ \text{s. t.} & \|\mathcal{A}(W) - b\|_2 \le \epsilon, \\ & W \succeq 0,\ W^T = W, \end{array}\right\} \tag{3.5}$$

或是 Lagrange 形式

$$\min_{W \in \mathbb{S}_+^m} \ \mu\|W\|_* + \frac{1}{2}\|\mathcal{A}(W) - b\|_2^2, \tag{3.6}$$

其中 $\mathbb{S}_+^m$ 表示 m 维对称半正定矩阵的集合, 参数 $\mu > 0$.

近年来, 许多学者致力于发展一阶快速算法来求解矩阵恢复问题, 例如, SVT [19]、FPC [43, 71]、APG [129] 等. 与此同时, 一些具有较好的收敛速度 $\mathcal{O}(1/k^2)$ (k 为迭代数) 的加速梯度算法也受到广泛关注, 例如 [11, 50, 72, 73, 83, 85, 86, 129, 131]. 虽然此类一阶算法求得的解在数值精度上不及内点法, 但是能解决较大规模的问题.

本章中, 我们提出求解问题 (3.6) 的一阶算法——改进的不动点迭代算法 (MFPC-BB), 并给出了算法的收敛性分析. 算法以不动点迭代算法中的算子分

裂技术为基础, 通过改进阈值算子 $\mathcal{T}_\nu$ 来求解低秩 Gram 矩阵的恢复问题. 同时, 引入 Barzilai-Borwein 技术动态地选取迭代步长, 从而提高了算法的收敛速度. 在此方法基础上, 我们又提出了一种加速的不动点迭代算法 (AFPC-BB). 它既保持了 MFPC-BB 简单易实现的特点, 又能够将原算法的线性收敛速度提高为二次收敛. 数值实验显示, AFPC-BB 算法在近似或准确计算有理系数多项式平方和分解的问题上的表现优于内点法或投影方法.

3.2　改进的不动点迭代算法

由于矩阵的核范数是不可微函数, 为了描述核范数极小化问题 (3.6) 的解所满足的最优性条件, 首先需讨论核范数的次微分(Subdifferential). 给定一个凸函数 $f:\mathbb{R}^{n_1\times n_2}\to\mathbb{R}$, f 在 $X^*\in\mathbb{R}^{n_1\times n_2}$ 处的次微分定义为如下紧致凸集

$$\partial f(X^*) := \{Z\in\mathbb{R}^{n_1\times n_2} : f(Y)\geq f(X^*)+\langle Z, Y-X^*\rangle, \forall\, Y\in\mathbb{R}^{n_1\times n_2}\}.$$

根据 [69, 定理 3.2] 和 [136] 中的讨论, 我们推导出核范数关于对称矩阵 $W\in\mathbb{S}^m$ 的次微分的表达式.

定理 3.1. 令 $W\in\mathbb{S}^m$, 则

$$\partial\|W\|_* = \{Q^{(1)}Q^{(1)T} - Q^{(2)}Q^{(2)T} + Z : Q^{(i)T}Z = 0, i=1,2, \text{ 且 } \|Z\|_2\leq 1\},$$

其中 $Q^{(1)}$ 和 $Q^{(2)}$ 分别为对应于正、负特征值的正交特征向量.

证明. 对于给定的实对称矩阵 $W\in\mathbb{S}^m$, 将其特征值排序为 $\lambda_1\geq\cdots\geq\lambda_t>0>\lambda_{t+1}\geq\cdots\geq\lambda_s$, $\lambda_{s+1}=\cdots=\lambda_m=0$. 设 $W=Q\Lambda Q^T$ 为矩阵 W 的Schur 分解, 其中 $\Lambda=\mathrm{diag}(\lambda_1,\ldots,\lambda_m)$, $Q\in\mathbb{R}^{m\times m}$ 为实正交矩阵, 并可分块记为

$$Q=\left(Q^{(1)},Q^{(2)},Q^{(3)}\right),\quad \Lambda=\begin{pmatrix}\Lambda^{(1)} & 0 & 0\\ 0 & \Lambda^{(2)} & 0\\ 0 & 0 & \Lambda^{(3)}\end{pmatrix},$$

其中 $Q^{(1)},Q^{(2)},Q^{(3)}$ 包含 $t, s-t, m-s$ 列且分别对应于特征值 $\Lambda^{(1)}=\mathrm{diag}(\lambda_1,\ldots,\lambda_t)$, $\Lambda^{(2)}=\mathrm{diag}(\lambda_{t+1},\ldots,\lambda_s)$, $\Lambda^{(3)}=\mathrm{diag}(\lambda_{s+1},\ldots,\lambda_m)$.

令 $\lambda=(\lambda_1,\ldots,\lambda_m)^T$, 由 [136] 中的分析可知

$$\partial\|\lambda\|_1 = \{y\in\mathbb{R}^m : y_i=1, i=1,\ldots,t;\ y_j=-1, j=t+1,\ldots,s;\ |y_k|\leq 1,\\ k=s+1,\ldots,m\}.$$

令 $Y \in \partial\|W\|_*$, 由 [69, 定理 3.1] 可知

$$Y = Q\,\mathrm{diag}(T)\,Q^T,$$

其中 $T \in \partial\|\lambda\|_1$. 从而有,

$$Y = Q^{(1)}Q^{(1)T} - Q^{(2)}Q^{(2)T} + Q^{(3)}DQ^{(3)T},$$

其中 D 为 $(m-s)\times(m-s)$ 对角矩阵, 其对角线上元素的绝对值均小于等于 1.

令 $Z = Q^{(3)}DQ^{(3)T}$, 则有 $Q^{(i)T}Z = 0, i = 1, 2$. 记 $\sigma_1(\cdot)$ 为矩阵最大奇异值, 则

$$\|Z\|_2 = Q^{(3)}DQ^{(3)T} \le \sigma_1(D) < 1,$$

□

下面将 [71, 定理 2] 中无限制条件的凸优化问题的最优性定理推广到求解带限制条件的问题 (3.6).

定理 3.2. *令 $f : \mathbb{S}^m \to \mathbb{R}$ 为真凸函数, 即 $f > -\infty$ 且在其定义域中至少存在一点使得 $f < +\infty$, 则 W^* 是*

$$\min_{W\in\mathbb{S}^m_+} f(W), \tag{3.7}$$

的最优解当且仅当 $W^* \in \mathbb{S}^m_+$, 且存在矩阵 $U \in \partial f(W^*)$, 使得

$$\langle U, V - W^*\rangle \ge 0, \quad \forall\, V \in \mathbb{S}^m_+. \tag{3.8}$$

证明. 首先, 假设 $U \in \partial f(W^*)$ 满足 (3.8). 由 f 在 $W^* \in \mathbb{S}^m_+$ 处次微分的定义可知

$$f(V) \ge f(W^*) + \langle U, V - W^*\rangle, \quad \forall\, V \in \mathbb{S}^m_+.$$

因此, 对所有 $V \in \mathbb{S}^m_+$, $f(V) \ge f(W^*)$, 从而 W^* 为 (3.7) 的最优解.

反之, 假设 W^* 为 (3.7) 的最优解但不满足条件 (3.8), 即对每个 $U \in \partial f(W^*)$, 都存在 $V \in \mathbb{S}^m_+$ 使得

$$\langle U, V - W^*\rangle < 0. \tag{3.9}$$

令 $Z(t) = tW^* + (1-t)V$, 其中 $t \in [0,1]$ 为参数. 由于 $Z(t)$ 在以 W^* 和 V 为端点的线段中, 且 $\mathbb{S}_+^m$ 为凸集, 则对所有 $t \in [0,1]$, $Z(t) \in \mathbb{S}_+^m$. 由 [108, 定理 23.4] 可知, f 在 $Z(1)$ 点关于向量 $W^* - V$ 的方向导数满足不等式

$$\begin{aligned} f'(Z(t); W^* - V)|_{t=1} &= f'(W^*; W^* - V) \\ &= \sup\{\langle W, W^* - V\rangle : W \in \partial f(W^*)\} \\ &\overset{(3.9)}{=} \langle U, W^* - V\rangle > 0. \end{aligned}$$

因此, 对一个足够小的 $\epsilon > 0$, 有 $f(Z(1-\epsilon)) < f(W^*)$, 与 W^* 是 (3.7) 的最优解矛盾. □

下面我们介绍阈值算子 $\mathcal{T}_\nu$ 并给出解问题 (3.6) 的改进的不动点迭代算法.

定义 3.1. 设矩阵 $W \in \mathbb{S}^m$ 的一个 Schur 分解为 $W = Q\Lambda Q^T$, 其中 $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$, $Q \in \mathbb{R}^{m\times m}$ 为实正交矩阵. 任给 $\nu \geq 0$, 阈值算子 $\mathcal{T}_\nu(\cdot)$ 定义为

$$\mathcal{T}_\nu(W) := Q\, \mathcal{T}_\nu(\Lambda)\, Q^T, \quad \mathcal{T}_\nu(\Lambda) = \mathrm{diag}(\{\lambda_i - \nu\}_+),$$

其中 $t_+ = \max(0, t)$.

给定 $\mu \geq 0$, 步长 $\tau > 0$ 以及初始点 X^0, 改进的不动点迭代算法为

$$\left\{ \begin{array}{rcl} Y^k & = & X^k - \tau\mathcal{A}^*(\mathcal{A}(X^k) - b), \\ X^{k+1} & = & \mathcal{T}_{\tau\mu}(Y^k), \end{array} \right. \tag{3.10}$$

其中 $k = 1, 2, \cdots$.

注 1. 基于对称矩阵 Y^k 的特征值分解的阈值算子 $\mathcal{T}_\nu$ 也在 [129, 注 3] 中出现. 而文献 [73] 中作者分析了基于特征值分解的阈值算子的收敛性.

定理 3.3. *设 $W^* \in \mathbb{S}_+^m$ 满足*

- *对于给定的 $\mu > 0$, $\|\mathcal{A}(W^*) - b\|_2 < \mu/m$,*
- *$W^* = \mathcal{T}_{\tau\mu}(h(W^*))$, 其中 $h(\cdot) = I(\cdot) - \tau\mathcal{A}^*(\mathcal{A}(\cdot) - b)$, $I(\cdot)$ 为恒等算子,*

则 W^* 是问题 (3.6) 的唯一的最优解.

证明. 令 $\nu = \tau\mu$, $Y^* = h(W^*) = W^* + E$, 其中 $E = -\tau\mathcal{A}^*(\mathcal{A}(W^*) - b)$, 则 $Y^* \in \mathbb{S}^m$.

先证 $\mathcal{T}_\nu(Y^*)$ 是问题

$$\min_{W\in\mathbb{S}^m_+} \nu\|W\|_* + \frac{1}{2}\|W - Y^*\|_F^2, \tag{3.11}$$

的唯一的最优解.

事实上, 由于问题 (3.11) 的目标函数为严格凸函数, 易知其存在唯一的极小值点, 因此, 我们只需证明其极小值点即为 $\mathcal{T}_\nu(Y^*)$. 将 Y^* 的特征值排列为

$$\lambda_1(Y^*) \geq \cdots \geq \lambda_t(Y^*) \geq \nu > \lambda_{t+1}(Y^*) \geq \cdots > 0 > \cdots \geq \lambda_s(Y^*),$$
$$\lambda_{s+1}(Y^*) = \cdots = \lambda_m(Y^*) = 0.$$

则 Y^* 的一个 Schur 分解可记为

$$Y^* = Q^{(1)}\Lambda^{(1)}Q^{(1)T} + Q^{(2)}\Lambda^{(2)}Q^{(2)T},$$

其中 $\Lambda^{(1)} = \mathrm{diag}(\lambda_1, \ldots, \lambda_t)$, $\Lambda^{(2)} = \mathrm{diag}(\lambda_{t+1}, \ldots, \lambda_s)$, $Q^{(1)}$、$Q^{(2)}$ 分别为相应于 $\Lambda^{(1)}$、$\Lambda^{(2)}$ 的分块矩阵.

记 $\widehat{X} = \mathcal{T}_\nu(Y^*)$, 则

$$\widehat{X} = Q^{(1)}(\Lambda^{(1)} - \nu I)Q^{(1)T},$$

因此,

$$Y^* - \widehat{X} = \nu(Q^{(1)}Q^{(1)T} + Z), \quad Z = \nu^{-1}Q^{(2)}\Lambda^{(2)}Q^{(2)T}.$$

显然 $Q^{(1)T}Z = 0$.

- 如果 $\lambda_{t+1}(Y^*) \geq |\lambda_s(Y^*)|$, 那么 $\|Z\|_2 = \lambda_{t+1}(Y^*)/\nu < 1$.
- 否则, 令 $y = \mathcal{A}(W^*) - b \in \mathbb{R}^p$, 则

$$\begin{aligned}\|E\|_F^2 &= \tau^2\|A_1y_1 + \cdots + A_py_p\|_F^2\\ &= \tau^2(\|A_1\|_F^2y_1^2 + \cdots + \|A_p\|_F^2y_p^2)\\ &\leq \tau^2m^2(y_1^2 + \cdots + y_p^2)\\ &< \tau^2\mu^2\end{aligned}$$

注意到 $E \in \mathbb{S}^m$ 且 $W^* \in \mathbb{S}^m_+$, 由 [44, 定理8.1.5] 可知

$$\|Z\|_2 = \frac{|\lambda_s(Y^*)|}{\nu} = \frac{\max\{|\lambda_1(E)|, |\lambda_n(E)|\}}{\nu} \le \frac{\|E\|_F}{\nu} < 1.$$

因此, $Y^* - \widehat{X} \in \nu\partial\|\widehat{X}\|_*$, 即 $0 \in \nu\partial\|\widehat{X}\|_* + \widehat{X} - Y^*$. 由定理3.2可知 $\mathcal{T}_\nu(Y^*)$ 为问题 (3.11) 的最优解.

注意到, 问题 (3.6) 的目标函数也是严格凸函数, 其最优解唯一. 如果 $W^* = \mathcal{T}_{\tau\mu}(Y^*)$, 即 W^* 为 (3.11) 的最优解, 由定理3.2可知存在 $U \in \nu\partial\|W^*\|_* + W^* - Y^*$, 使得

$$\langle U, V - W^*\rangle \ge 0, \quad \forall\, V \in \mathbb{S}^m_+.$$

令 $\widetilde{U} = U/\tau$, 将 $\nu = \tau\mu$ 和 $Y^* = W^* - \tau\mathcal{A}^*(\mathcal{A}(W^*) - b)$ 代入上述次微分表达式中, 得出 $\widetilde{U} \in \mu\partial\|W^*\|_* + \mathcal{A}^*(\mathcal{A}(W^*) - b)$, 因此,

$$\langle \widetilde{U}, V - W^*\rangle \ge 0, \quad \forall\, V \in \mathbb{S}^m_+.$$

再次应用定理 3.2, 可以推出 W^* 为问题 (3.6) 的最优解. □

3.3　收敛性分析

本节给出改进的不动点迭代算法 (3.10) 的收敛性分析. 在证明主要的收敛性定理之前, 我们首先给出两个引理说明阈值算子 $\mathcal{T}_\nu$ 和定理3.3中定义的函数 h 都具有非扩张性.

引理 3.4. *阈值算子 $\mathcal{T}_\nu$ 是非扩张的, 即对任意 $X_1, X_2 \in \mathbb{S}^m$,*

$$\|\mathcal{T}_\nu(X_1) - \mathcal{T}_\nu(X_2)\|_F \le \|X_1 - X_2\|_F, \tag{3.12}$$

而且

$$\|X_1 - X_2\|_F = \|\mathcal{T}_\nu(X_1) - \mathcal{T}_\nu(X_2)\|_F \Longleftrightarrow X_1 - X_2 = \mathcal{T}_\nu(X_1) - \mathcal{T}_\nu(X_2). \tag{3.13}$$

证明. 设 $X_1 = Q^{(1)}\Lambda^{(1)}Q^{(1)T}$ 和 $X_2 = Q^{(2)}\Lambda^{(2)}Q^{(2)T}$ 分别为 X_1 和 X_2 的 Schur 分解, 其中

$$\Lambda^{(1)} = \begin{pmatrix} \mathrm{diag}(\lambda_1) & 0 \\ 0 & 0 \end{pmatrix}, \quad \Lambda^{(2)} = \begin{pmatrix} \mathrm{diag}(\lambda_2) & 0 \\ 0 & 0 \end{pmatrix},$$

$\lambda_1 = (\alpha_1, \ldots, \alpha_s)^T$ 和 $\lambda_2 = (\beta_1, \ldots, \beta_t)^T$ 分别为 X_1、X_2 的特征值构成的向量, $Q^{(1)}$、$Q^{(2)}$ 为正交矩阵. 不妨设 $\alpha_1 \geq \cdots \geq \alpha_k \geq \nu > \alpha_{k+1} \geq \cdots \geq \alpha_s$, $\beta_1 \geq \cdots \geq \beta_l \geq \nu > \beta_{l+1} \geq \cdots \geq \beta_t$, 则

$$\widetilde{X}_1 := \mathcal{T}_\nu(X_1) = Q^{(1)}\widetilde{\Lambda}^{(1)}Q^{(1)T}, \quad \widetilde{X}_2 := \mathcal{T}_\nu(X_2) = Q^{(2)}\widetilde{\Lambda}^{(2)}Q^{(2)T},$$

其中

$$\bar{\Lambda}^{(1)} = \begin{pmatrix} \mathrm{diag}(\widetilde{\lambda}_1) & 0 \\ 0 & 0 \end{pmatrix}, \quad \bar{\Lambda}^{(2)} = \begin{pmatrix} \mathrm{diag}(\widetilde{\lambda}_2) & 0 \\ 0 & 0 \end{pmatrix},$$

$\widetilde{\lambda}_1 = (\alpha_1 - \nu, \ldots, \alpha_k - \nu)$, $\widetilde{\lambda}_2 = (\beta_1 - \nu, \ldots, \beta_l - \nu)$. 因此,

$$\begin{aligned}
&\|X_1 - X_2\|_F^2 - \|\widetilde{X}_1 - \widetilde{X}_2\|_F^2 \\
&= \mathrm{Tr}((X_1 - X_2)^T(X_1 - X_2)) - \mathrm{Tr}((\widetilde{X}_1 - \widetilde{X}_2)^T(\widetilde{X}_1 - \widetilde{X}_2)) \\
&= \mathrm{Tr}(X_1^T X_1 - \widetilde{X}_1^T\widetilde{X}_1 + X_2^T X_2 - \widetilde{X}_2^T\widetilde{X}_2) - 2\mathrm{Tr}(X_1^T X_2 - \widetilde{X}_1^T\widetilde{X}_2) \\
&= \sum_{i=1}^{s}\alpha_i^2 - \sum_{i=1}^{k}(\alpha_i - \nu)^2 + \sum_{i=1}^{t}\beta^2 - \sum_{i=1}^{l}(\beta_i - \nu)^2 - 2\mathrm{Tr}(X_1^T X_2 - \widetilde{X}_1^T\widetilde{X}_2).
\end{aligned}$$

注意到

$$\begin{aligned}
&\mathrm{Tr}(X_1^T X_2 - \widetilde{X}_1^T\widetilde{X}_2) \\
&= \mathrm{Tr}((X_1 - \widetilde{X}_1)^T(X_2 - \widetilde{X}_2) + (X_1 - \widetilde{X}_1)^T\widetilde{X}_2 + \widetilde{X}_1^T(X_2 - \widetilde{X}_2)) \\
&= \mathrm{Tr}((\Lambda^{(1)} - \widetilde{\Lambda}^{(1)})(\Lambda^{(2)} - \widetilde{\Lambda}^{(2)})) + \mathrm{Tr}((\Lambda^{(1)} - \widetilde{\Lambda}^{(1)})\widetilde{\Lambda}^{(2)}) + \mathrm{Tr}(\widetilde{\Lambda}^{(1)}(\Lambda^{(2)} - \widetilde{\Lambda}^{(2)})).
\end{aligned} \tag{3.14}$$

下面考虑 (3.14) 式的上界. 给定两个对称矩阵 $X, Y \in \mathbb{S}^m$, 则有

$$\mathrm{Tr}(XY) \leq \lambda(X)^T\lambda(Y),$$

等号成立当且仅当存在实正交矩阵 $Q \in \mathbb{R}^{m\times m}$ 满足

$$X = Q\mathrm{diag}(\lambda(X))Q^T, \quad Y = Q\mathrm{diag}(\lambda(Y))Q^T$$

其中 $\lambda(X) = (\lambda_1(X), \cdots, \lambda_n(X))^T$ 表示矩阵 X 的特征值构成的向量(见 [69, 定理 2.2]). 不妨设 $k \leq l \leq s \leq t$, 将此结果用到式 (3.14) 中可得

$$\mathrm{Tr}(X_1^T X_2 - \widetilde{X}_1^T\widetilde{X}_2) \leq \sum_{i=1}^{l}\alpha_i\nu + \sum_{i=l+1}^{s}\alpha_i\beta_i + \sum_{i=1}^{k}(\beta_i - \nu)\nu + \sum_{i=k+1}^{l}\alpha_i(\beta_i - \nu).$$

从而有,

$$\begin{aligned}
&\|X_1 - X_2\|_F^2 - \|\widetilde{X}_1 - \widetilde{X}_2\|_F^2 \\
&\geq \sum_{i=1}^{s} \alpha_i^2 - \sum_{i=1}^{k} (\alpha_i - \nu)^2 + \sum_{i=1}^{t} \beta^2 - \sum_{i=1}^{l} (\beta_i - \nu)^2 \\
&\quad - 2(\sum_{i=1}^{l} \alpha_i \nu + \sum_{i=l+1}^{s} \alpha_i \beta_i + \sum_{i=1}^{k} (\beta_i - \nu)\nu + \sum_{i=k+1}^{l} \alpha_i (\beta_i - \nu)) \\
&= (\sum_{i=l+1}^{s} \alpha_i^2 + \sum_{i=l+1}^{t} \beta_i^2 - 2 \sum_{i=l+1}^{s} \alpha_i \beta_i) + \sum_{i=k+1}^{l} (2\beta_i \nu - \nu^2 + \alpha_i^2 - 2\alpha_i \beta_i).
\end{aligned}$$

由 $t \geq s$ 且 $\alpha_i^2 + \beta_i^2 - 2\alpha_i \beta_i \geq 0$, 可知

$$\sum_{i=l+1}^{s} \alpha_i^2 + \sum_{i=l+1}^{t} \beta_i^2 - 2 \sum_{i=l+1}^{s} \alpha_i \beta_i \geq 0.$$

与此同时, 由于函数 $g(x) := 2\beta_i x - x^2 + \alpha_i^2 - 2\alpha_i \beta_i$ 在区间 $[-\infty, \beta_i]$ 上单调递增, 且 $\alpha_i \leq \nu \leq \beta_i$, $i = k+1, \ldots, l$, 则有,

$$2\nu\beta_i - \nu^2 + \alpha_i^2 - 2\alpha_i \beta_i > 0, \quad i = k+1, \ldots, l.$$

因此,

$$\|X_1 - X_2\|_F^2 - \|\widetilde{X}_1 - \widetilde{X}_2\|_F^2 \geq 0,$$

即 (3.12) 成立.

如果 $\|X_1 - X_2\|_F = \|\mathcal{T}_\nu(X_1) - \mathcal{T}_\nu(X_1)\|_F$, 则 $s = t, k = l$, 且 $\alpha_i = \beta_i, i = k+1, \ldots, s$, 说明 $\Lambda^{(1)} - \widetilde{\Lambda}^{(1)} = \Lambda^{(2)} - \widetilde{\Lambda}^{(2)}$ 且 $\mathrm{Tr}((X_1 - \widetilde{X}_1)^T (X_2 - \widetilde{X}_2))$ 达到其最大值. 由 [69, 定理 2.2] 可知存在正交矩阵 $Q \in \mathbb{R}^{m \times m}$ 使得

$$X_1 - \widetilde{X}_1 = Q(\Lambda^{(1)} - \widetilde{\Lambda}^{(1)})Q^T = Q(\Lambda^{(2)} - \widetilde{\Lambda}^{(2)})Q^T = X_2 - \widetilde{X}_2,$$

因此,

$$X_1 - X_2 = \mathcal{T}_\nu(X_1) - \mathcal{T}_\nu(X_2). \tag{3.15}$$

另一方面, 如果 (3.15) 成立, 显然有 $\|X_1 - X_2\|_F = \|\mathcal{T}_\nu(X_1) - \mathcal{T}_\nu(X_2)\|_F$. □

引理 3.5. *设 $\tau \in (0, 2/\|\mathcal{A}\|_2^2)$, 则算子 $h(\cdot) = I(\cdot) - \tau \mathcal{A}^*(\mathcal{A}(\cdot) - b)$ 是非扩张的, 即对任意的 $X_1, X_2 \in \mathbb{S}^m$,*

$$\|h(X_1) - h(X_2)\|_F \leq \|X_1 - X_2\|_F.$$

而且

$$\|h(X_1)-h(X_2)\|_F=\|X_1-X_2\|_F \Longleftrightarrow h(X_1)-h(X_2)=X_1-X_2.$$

此引理的证明参见 [71, 引理 2].

下面我们给出算法的收敛性定理, 说明改进的不动点迭代算法 (3.10) 收敛到问题 (3.6) 的最优解.

定理 3.6. *设 $\tau\in(0,2/\|\mathcal{A}\|_2^2)$, $W^*\in\mathbb{S}_+^m$ 满足*

- *对于给定的 $\mu>0$, $\|\mathcal{A}(W^*)-b\|_2<\mu/m$,*
- *$W^*=\mathcal{T}_{\tau\mu}(h(W^*))$, 其中 $h(\cdot)=I(\cdot)-\tau\mathcal{A}^*(\mathcal{A}(\cdot)-b)$, $I(\cdot)$ 为恒等算子,*

那么由改进的不动点迭代算法 (3.10) 所得的列 $\{X^k\}$ 收敛到问题 (3.6) 的最优解.

证明. 由定理 refeqivalence 可知, 满足上述两个条件的 W^* 为问题 (3.6) 的唯一的最优解, 下面只需证明 $\{X^k\}$ 收敛到 W^*.

令 $\nu=\tau\mu$, 由于算子 $\mathcal{T}_\nu(\cdot)$ 和 $h(\cdot)$ 都是非扩张的, 则 $\mathcal{T}_\nu(h(\cdot))$ 也是非扩张的, 则有 $\{X^k\}$ 在紧集中. 那么一定存在 $\{X^k\}$ 的收敛子列 $\{X^{k_j}\}$, 设 $\widetilde{X}=\lim_{j\longrightarrow\infty}X^{k_j}$ 为满足 $\|\mathcal{A}(\widetilde{X})-b\|_2<\mu/m$ 的极限点.

注意到 $W^*=\mathcal{T}_\nu(h(W^*))$, 则有

$$\|X^{k+1}-W^*\|_F=\|\mathcal{T}_\nu(h(X^k))-\mathcal{T}_\nu(h(W^*))\|_F\le\|h(X^k)-h(W^*)\|_F=\|X^k-W^*\|_F,$$

即 $\{\|X^k-W^*\|_F\}$ 是单调非增列, 从而有

$$\lim_{k\longrightarrow\infty}\|X^k-W^*\|_F=\|\widetilde{X}-W^*\|_F,$$

其中 $\widetilde{X}$ 是 $\{X^k\}$ 的任意一个极限点. 由算子 $\mathcal{T}_\nu(h(\cdot))$ 的连续性可知

$$\mathcal{T}_\nu(h(\widetilde{X}))=\lim_{j\longrightarrow\infty}\mathcal{T}_\nu(h(X^{k_j}))=\lim_{j\longrightarrow\infty}X^{k_j+1},$$

即 $\mathcal{T}_\nu(h(\widetilde{X}))$ 也是 $\{X^k\}$ 的一个极限点, 则有

$$\|\mathcal{T}_\nu(h(\widetilde{X}))-\mathcal{T}_\nu(h(W^*))\|_F=\|\mathcal{T}_\nu(h(\widetilde{X}))-W^*\|_F=\|\widetilde{X}-W^*\|_F.$$

利用引理 3.4 和引理 3.5, 得

$$\mathcal{T}_\nu(h(\widetilde{X})) - \mathcal{T}_\nu(h(W^*)) = h(\widetilde{X}) - h(W^*) = \widetilde{X} - W^*.$$

说明 $\mathcal{T}_\nu(h(\widetilde{X})) = \widetilde{X}$. 注意到, $\|\mathcal{A}(\widetilde{X}) - b\|_2 < \mu/m$, 由定理 3.3 可知 $\widetilde{X}$ 为问题 (3.6) 的最优解, 即 $\widetilde{X} = W^*$.

因此,

$$\lim_{k\longrightarrow\infty} \|X^k - W^*\|_F = 0,$$

即 $\{X^k\}$ 收敛到 W^*.

□

3.4　算法及实现

本节给出了求解低秩 Gram 矩阵恢复问题的改进的不动点迭代算法, 以及在算法的实现过程中相关参数的选取方法.

3.4.1　阈值算子的赋值

迭代公式 (3.10) 中主要的计算量在于每一步迭代过程中对矩阵做 Schur 分解 $Y^k = Q\Lambda Q^T$. 结合阈值算子 $\mathcal{T}_\nu$ 的定义, 我们只需计算矩阵 Y^k 中大于 ν 的特征值及相应的特征向量. 而在数值线性代数领域中, 计算矩阵的较大部分特征值及特征向量的数值算法已经较为成熟, 也有许多高质量的软件包供人们使用, 参见 [51–53]. 根据文献 [19, 129], 我们选取 Matlab 软件包 PROPACK [59] 来计算矩阵的 Schur 分解. PROPACK 能够计算出矩阵给定个数的较大特征值及相应的特征向量, 而不能自动地计算出大于或等于 ν 的特征值. 因此, 每一步迭代中, 在计算 Y^k 的 Schur 分解之前, 我们需要预先给定所需计算的特征值的个数 s_k. 我们采用以下方法：设 X^k 的 Schur 分解为 $X^k = Q^{k-1}\Lambda^{k-1}(Q^{k-1})^T$, 令 s_k 等于对角矩阵 Λ^{k-1} 的对角线上大于或等于 $\epsilon_k \parallel \Lambda^{k-1} \parallel_2$ 的元素的个数, 其中 $\epsilon_k > 0$ 为给定参数. 注意到, 按照上述方法计算得到的 s_k 逐渐减小. 如果在第 k 步迭代中 s_k 取值过小, 将会导致阈值算子 $\mathcal{T}_\nu$ 不满足非扩张性条件 (3.12). 因此, 在算法实现过程中, 如果超过 10 次违反条件 (3.12), 我们令 $s_k = s_k + 1$. 数值实验显示这样的处理会增加算法的鲁棒性.

3.4.2 Barzilai-Borwein 技术

在文献 [71] 中, 由于算子 $\mathcal{A}$ 来自于由独立同分布的高斯随机变量构成的矩阵中随机抽取的 p 个元素, 则问题 (3.6) 的目标函数的 Lipschitz 常数为 1. 因此作者始终取参数 $\tau = 1$. 本章中, 由定理 3.6 可知改进的不动点迭代算法 (3.10) 收敛的条件为 $\tau \in (0,\ 2/\|\mathcal{A}\|_2^2)$. 但是此选择可能过于保守, 导致算法的收敛速度较慢.

关于迭代步长的选取方式有很多. 受文献 [138, 140] 的启发, 我们提出一项基于 Barzilai-Borwein 方法 (BB) [7] 的选择步长 τ_k 的技术. 令 $g(\cdot) = \mathcal{A}^*(\mathcal{A}(\cdot)-b)$ 且 $g^k = \mathcal{A}^*(\mathcal{A}(X^k) - b)$. 改进的不动点迭代算法 (3.10) 首先以 τ 为步长沿着光滑函数 $\frac{1}{2}\|\mathcal{A}(X^k) - b\|_2^2$ 的负梯度方向 g^k下降, 而后利用阈值算子 $\mathcal{T}_\nu(\cdot)$ 来协调非光滑函数 $\|X\|_*$, 因此, τ 的选择只与 $\frac{1}{2}\|\mathcal{A}(X^k) - b\|_2^2$ 有关.

令

$$\Delta X = X^k - X^{k-1}, \quad \Delta g = g^k - g^{k-1}.$$

BB 技术是基于拟牛顿法给出的切线方程的两点近似,

$$\tau_k = \frac{\langle \Delta X, \Delta g\rangle}{\langle \Delta g, \Delta g\rangle}, \quad \text{或} \quad \tau_k = \frac{\langle \Delta X, \Delta X\rangle}{\langle \Delta X, \Delta g\rangle}.$$

为了避免参数 τ_k 取得过小或过大, 令

$$\tau_k = \max\{\tau_{min}, \min\{\tau_k, \tau_{max}\}\},$$

其中 $0 < \tau_{min} < \tau_{max} < \infty$ 为给定的参数.

利用 BB 方法求得步长从而加速梯度算法的收敛的技术在 [138] 中也出现过.

3.4.3 算法

给定一列单调下降趋于零的参数 $\mu_k \downarrow 0$, 问题 (3.6) 的极小值点趋于原问题 (3.4) 的最优解. 此类方法也称为外罚值方法 (Exterior Penalty Method), 是由 Fiacco 和 McCormick [40] 在上世纪六十年代提出的. 借鉴文献 [45, 71, 129], 我们的算法中采用如下的连续性技术: 设 $\bar{\mu}$ 为目标参数, 例如, 10^{-4}. 对于单调下降的参数 μ_k, 利用改进的不动点迭代算法来求解一系列的问题 (3.6). 在求解

参数为 μ_{k+1} 的问题 (3.6) 时, 起始点选为上一步迭代得到的参数为 μ_k 的问题 (3.6) 的近似最优解. 令

$$\mu_{k+1} = \max(\eta\mu_k, \bar{\mu}), \quad k = 1, \ldots, L-1,$$

其中 $0 < \eta < 1$ 表示相邻两个 μ 值减小的比例.

算法：*MFPC-BB*

输入：▸ 给定参数 $0 < \tau_{min} < \tau_0 < \tau_{max} < \infty$, $\mu_1 > \bar{\mu} > 0$, $\eta > 0$ 和误差界 $\epsilon > 0$.

输出：▸ 数值 Gram 矩阵.

- 令 $X^0 = 0$.
- 对于 $\mu = \mu_1, \ldots, \mu_L$, 进行循环
  1. 利用 BB 技术选择步长 τ_k 且满足 $\tau_{min} \le \tau_k \le \tau_{max}$.
  2. 计算 $Y^k = X^k - \tau_k \mathcal{A}^*(\mathcal{A}(X^k) - b)$ 及 Schur 分解 $Y^k = Q^k\, \Lambda^k\, (Q^k)^T$.
  3. 计算 $X^{k+1} = Q^k\, \mathcal{T}_{\tau_k\mu_k}(\Lambda^k)\, (Q^k)^T$.
- 如果满足停机条件, 那么返回 X_{opt}.

设 $F(X) = \mu\|W\|_* + \frac{1}{2}\|\mathcal{A}(W) - b\|_2^2$. 如文献 [11] 中所证

$$F(X^k) - F(X^*) \le \frac{C \parallel X^0 - X^* \parallel_F^2}{k}.$$

其中 C 为常数. 因此, 改进的不动点迭代算法的收敛速度为 $\mathcal{O}(1/k)$.

最近, Beck 和 Teboulle [11] 将 Nesterov [83] 的算法中的加速技巧应用到求解图像处理中产生的线性反问题上, 得到更好的收敛效果. 他们的算法在计算新的迭代步 X^{k+1} 时, 不只是利用 X^k, 而是要利用上两步迭代的结果 X^k 和 X^{k-1} 的线性组合. 我们将此技巧引入到 MFPC-BB 算法中, 给出加速的不动点迭代算法 (AFPC-BB). 新算法既保持了 MFPC-BB 简单易实现的特点, 又能够将收敛速度提高到 $\mathcal{O}(1/k^2)$.

算法：*AFPC-BB*

输入：▸ 给定参数 $0 < \tau_{min} < \tau_0 < \tau_{max} < \infty$, $\mu_1 > \bar{\mu} > 0$, $\eta > 0$ 及误差界 $\epsilon > 0$.

输出：▸ 数值 Gram 矩阵.

- 令$X^0 = 0, t_0 = 1$.
- 对 $\mu = \mu_1, \ldots, \mu_L$, 进行循环
  1. 利用 BB 技术选择步长 τ_k 且满足 $\tau_{min} \le \tau_k \le \tau_{max}$.
  2. 计算 $Z^k = X^k + \frac{t_{k-1}-1}{t_k}(X^k - X^{k-1})$.
  3. 计算 $Y^k = Z^k - \tau_k \mathcal{A}^*(\mathcal{A}(Z^k) - b)$ 及 Schur 分解 $Y^k = Q^k\, \Lambda^k\, (Q^k)^T$.
  4. 计算 $X^{k+1} = Q^k\, \mathcal{T}_{\tau_k\mu_k}(\Lambda^k)\, (Q^k)^T$.
  5. 计算 $t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$.
- 如果满足停机条件, 那么返回 $X_{\rm opt}$.

下面的定理说明将 MFPC-BB 算法中关于 X^k 的梯度步改为对 Z^k 实施, 算法的收敛速度将提高到 $\mathcal{O}(1/k^2)$.

定理 3.7. *[50, 129]* 设 $\{X^k\}$ 是由加速的不动点迭代算法得到的列. 则对任意 $k > 1$, 有

$$F(X^k) - F(X^*) \le \frac{C\|X^* - X^0\|_F^2}{(k+1)^2}.$$

其中 C 为常数.

3.4.4 停机准则

我们给出改进的不动点迭代算法的停机准则. 注意到, 在解问题 (3.6) 的第 k 次内部迭代的过程中, 参数 $\mu = \mu_k$ 是固定的. 有多种方式来确定此内部迭代的停机准则, 根据定理 3.2, 当 W 满足

$$\begin{cases} W = \mathcal{T}_\nu(h(W)), \\ \|\mathcal{A}(W) - b\|_2 < \epsilon. \end{cases} \tag{3.16}$$

时, W 是问题 (3.5) 的解. 为了保证 X^k 与最优解 W^* 足够接近, 只需检验 X^k 是否近似地满足条件 (3.16). 对于条件 (3.16) 中的第一个方程, 当 X^{k+1} 与 X^k 之间的相对误差满足

$$rel.err := \frac{\|X^{k+1} - X^k\|_F}{\max\{1, \|X^k\|_F\}} < \epsilon, \tag{3.17}$$

迭代过程即可停止.

对于条件 (3.16) 中的第二个方程, 我们使用如下停机准则

$$rel.err := \frac{\|\mathcal{A}(X^k) - b\|_2}{\|b\|_2} < \epsilon. \tag{3.18}$$

3.5　数值实验

在本节中, 我们给出 MFPC-BB 算法和 AFPC-BB 算法求半正定多项式数目最少的平方和分解的数值表现. 实验中, 我们首先按照如下方法构造半正定多项式 $f(x)$: 对于给定的整数 m, r 满足 $r < m$, 随机选取 -10 到 10 之间的整数构成 $m \times r$ 维矩阵 L. 令 $W = LL^T$, 从而得到对称半正定矩阵. 列向量 $[x]_d$ 中包含所有由 $x_1, \ldots, x_n$ 构成的次数小于等于 d 的单项式, 其中 $2 \le n \le 4$, $5 \le d \le 20$. 我们得到半正定多项式

$$f(x) = [x]_d^T \cdot W \cdot [x]_d \in \mathbb{Q}[x].$$

将矩阵 W 参数化, 并将上式展开, 令等式两端相同单项式的系数相等, 即可得到线性等式条件 $\mathcal{A}W = b$.

注意到, 秩为 r 的 $n \times n$ 阶对称矩阵的自由度为 $d_r = r(2n - r + 1)/2$. 令 FR 表示矩阵的自由度与已知限制条件的个数 p 的比值, 即 $FR = d_r/p$. 如果 FR 的值较大 (接近于 1), 说明限制条件的个数近似等于矩阵中未定元的个数, 此时矩阵 W 的恢复较为困难. 反之, 如果 FR 的值较小 (接近于 0), 那么矩阵恢复相对简单. 而当 $FR > 1$ 时, 满足给定限制条件的秩 r 矩阵有无穷多个, 矩阵恢复的难度大大增加.

本节中, 算法 MFPC、MFPC-BB 和 AFPC-BB 的终止性条件均为 (3.18), 其中 ϵ 为给定的误差界. 数值实验中, 初始点为 $X^0 = 0$. 在迭代之前, 我们需要预估计问题 (3.6) 的目标函数中第二项 $\frac{1}{2}\|\mathcal{A}X - b\|_2^2$ 的最小 Lipschitz 常数 $C = \|\mathcal{A}\|_2^2$, 而后令 Barzilai-Borwein 参数 $\tau_{max} = 10/C$, $\tau_{min} = 10^{-3}/C$, 其中 10 和 10^{-3} 是在实验过程中由经验所得.

我们在 MATLAB 中实现了算法 MFPC-BB 和 AFPC-BB. 下面所有实验数据都是在 HP xw8600 工作站 (Inter Xeon(R) 2.67GHz CPU, 3.00 GB of RAM)上运行所得. 程序代码可以在下面地址下载 http://www.mmrc.iss.ac.cn/~lzhi/Research/hybrid/FPCs/

3.5.1 随机 Gram 矩阵恢复的数值结果

在第一部分的实验中, 令 $\epsilon = 10^{-3}$, 比较算法 MFPC, MFPC-BB 和AFPC-BB 在求解规模较小的随机生成的低秩 Gram 矩阵恢复问题上的表现. 为了清楚地比较三种算法的收敛速度, 在这部分实验中, 我们不使用连续性技术和 PROPACK 软件包, 而是在每一步迭代中都计算矩阵完整的 Schur 分解.

问题				MFPC		MFPC-BB		AFPC-BB	
n	r	p	FR	迭代数	误差	迭代数	误差	迭代数	误差
100	10	579	1.6494	527	9.98e-4	434	9.97e-4	50	9.46e-4
200	10	1221	1.6011	797	9.99e-4	512	9.99e-4	59	9.84e-4
500	10	5124	0.9670	632	4.99e-3	499	4.99e-3	66	4.90e-3

表 3.1: MFPC, MFPC-BB 和AFPC-BB 在不使用连续性技术下的比较

表 3.1 中给出了矩阵的自由度与已知限制条件个数的比值 FR, 三个算法停止时分别达到的精度 (3.18) 及所需迭代数. 算法的计算效率由迭代数来衡量. 通过对表中结果的比较可以看出, 三种算法在达到几乎相同的误差界时, 与 MFPC 相比, MFPC-BB 所需的迭代步数更少. 这说明 Barzilai-Borwein 技术对于 MFPC 算法具有较好的加速效果. 三种算法相比, 显然 AFPC-BB 所需的迭代步数最少.

表 3.2 中给出使用了连续性技术的 AFPC-BB 算法对随机生成的 Gram 矩阵恢复问题的数值结果. 其中, 令目标参数 $\bar{\mu} = 10^{-4}\|\mathcal{A}^*b\|$, 而初始参数 $\mu_1 = 1/4\|\mathcal{A}^*b\|$. 迭代过程中, μ_k 的更新方法为 $\mu_k = \max(1/4\mu_{k-1}, \bar{\mu})$, 直到满足停机条件. 同时, 在迭代中我们使用 PROPACK 软件包来计算矩阵部分较大特征值及相应的特征向量. 对于数值矩阵秩的计算, 给定误差界 10^{-5}.

数值实验结果显示 AFPC-BB 算法只需 200 步以内的迭代及不到 10 分钟的时间即满足 $\frac{\|\mathcal{A}(X^k)-b\|_2}{\|b\|_2} < 10^{-3}$. 这组实验中所有例子均满足 $FR > 1$. 因此, 在

问题				结果		
n	r	p	FR	迭代数	秩	时间/秒
100	10	579	1.6494	76	10	1.48e+0
500	10	3309	1.4974	139	27	6.13e+1
1000	50	10621	4.5923	127	59	1.53e+2
1500	50	25573	2.8849	196	77	5.41e+2

表 3.2: 使用连续性技术的算法AFPC-BB 的数值结果

已知限制条件个数小于未定元个数的前提下, AFPC-BB 算法仍然能够将 Gram 矩阵较好地恢复出来.

3.5.2 准确的平方和分解

加速的不动点迭代算法 AFPC-BB 求得的数值 Gram 矩阵 W 近似满足

$$f(x) \approx [x]_d^T \cdot W \cdot [x]_d, \quad W \succeq 0. \tag{3.19}$$

为了得到半正定多项式 $f(x)$ 准确的平方和分解, 数值 Gram 矩阵需要达到较高的精度 [55, 56, 92, 93], 才能将其转化为有理数矩阵.

尽管一阶方法是求解大规模矩阵恢复问题唯一可行的方法, 然而利用 AFPC-BB 算法来得到高精度解 (例如：10^{-10}) 需要耗费很长时间. 按照文献 [55, 56] 所述, 我们利用 Gauss-Newton 迭代精化由 AFPC-BB 算法计算得到的近似满足 (3.19) 的矩阵 W. 首先, 我们选择合适的整数 r 作为矩阵的秩, 并计算矩阵 W 的 $\mathrm{L^TDL}$ 分解从而得到 $f(x)$ 近似的平方和展开

$$f(x) \approx \sum_{i=1}^{r}(\sum_{\alpha} c_{i,\alpha}x^{\alpha})^2.$$

然后, 利用标准的 Gauss-Newton 迭代来计算 $\Delta c_{i,\alpha}x^{\alpha}$ 使其满足

$$f(x) = \sum_{i=1}^{r}(\sum_{\alpha} c_{i,\alpha}x^{\alpha} + \Delta c_{i,\alpha}x^{\alpha})^2 + O(\sum_{i=1}^{r}(\sum_{\alpha} \Delta c_{i,\alpha}x^{\alpha})^2).$$

并将矩阵 W 更新为 $W + \Delta W$. 当

$$\theta = \|f(x) - [x]_d^T \cdot W \cdot [x]_d\|_2,$$

比一个给定的误差界小的时候停止迭代. 如果 θ 经过几步迭代后仍然不收敛, 我们可以增加 Gauss-Newton 迭代计算中的精度, 或改变 r 值重新尝试计算. 如果选取的 r 值很大, Gauss-Newton 迭代步非常耗时, 甚至无法计算出结果. 因此, 这也是我们计算平方和数最小的多项式平方和分解的重要原因之一. 当 (3.19) 的精度足够高时, 直接将数值 Gram 矩阵的每个元素舍入到与之最近的有理数, 即可得到 f 准确的平方和分解

$$f(x) = [x]_d^T \cdot \widetilde{W} \cdot [x]_d,\ \widetilde{W} \succeq 0. \tag{3.20}$$

注2 AFPC-BB 算法返回的数值 Gram 矩阵 W 通常不是满秩的, 此时满足 (3.20) 中等式条件的对称矩阵 $\widetilde{W}$ 所形成的超平面 $\mathcal{X}$ 与对称半正定矩阵锥相切, 即问题 (3.2) 的最优解不在对称半正定矩阵锥的内部, 而是在其边界上, 利用文献 [55, 56] 中的有理化正交投影的方法很难将 W 投影到对称半正定矩阵锥上.

加速的不动点迭代算法 AFPC-BB 能够返回低秩 Gram 矩阵, 而经典的半定规划软件包 SeDuMi [125] 返回极大秩 Gram 矩阵 (见 [32, 定理 2.1]). 近期出现的半定规划软件包 SDPNAL [145] 则能够返回相对较为低秩的矩阵.

在表 3.3 中, 我们使用稠密单项式构成的向量 $[x]_d$来构造例子, 并分别给出了 AFPC-BB 算法和基于半定规划的软件包 SeDuMi 和 SDPNAL 在准确地恢复低秩 Gram 矩阵问题上的表现. 同时, 还给出了以上述三种方法返回的矩阵作为初始值, 利用 Gauss-Newton 迭代在 Maple 13 中对于数值 Gram 矩阵精化所需的时间. 为了达到较好的恢复效果, 这组试验中, 三种算法的终止性条件均设为 $\mathcal{A}W - b$ 的相对误差小于 $\epsilon = 5 \times 10^{-4}$.

如表 3.3 所示, 对于前四个例子, 均可利用 Gauss-Newton 迭代 (3.5.2) 将三种方法返回的数值 Gram 矩阵精化到较高的精度. 通过将精化后所得矩阵的每一元素改写为离其最近的整数, 即可得到秩为 5 的有理 Gram 矩阵, 从而得到半正定多项式准确的平方和分解. 对于表中最后的例子, 当 $n = 500$ 时, SeDuMi 已经无法计算出结果.

在表 3.4 中, 我们利用由稀疏单项式构成的向量 $[x]_d$ 来构造较大规模的例子. 随着限制条件个数 p 的增加, 矩阵的自由度与 p 的比值 FR 减小. 对于表中所有例子, 虽然 SDPNAL 能够计算出相对低秩的数值 Gram 矩阵, 但是目前无法在较短的时间内利用它们求出半正定多项式准确的平方和分解. 另一方面, 即使不运行 Gauss-Newton 迭代的精化步骤, 我们也能够利用 AFPC-BB 算法

问题				结果				Gauss-Newton 迭代		
n	r	p	FR	算法	秩	θ	时间/秒	秩	θ	时间/秒
100	5	579	0.8463	AFPC-BB	9	8.415e-1	1.75e+0	5	1.935e-9	2.98e+1
				SDPNAL	16	2.600e-1	1.50e+0	5	8.852e-10	2.63e+1
				SeDuMi	100	5.373e-2	4.03e+0	5	1.102e-10	3.22e+1
200	5	1221	0.8108	AFPC-BB	14	3.629e+0	1.07e+1	5	6.950e-10	4.02e+2
				SDPNAL	21	2.828e+0	1.06e+1	5	6.912e-10	5.57e+2
				SeDuMi	200	2.579e-1	5.56e+1	5	7.176e-10	1.10e+3
300	5	1932	0.7712	AFPC-BB	14	2.232e+1	2.32e+1	5	1.379e-9	5.61e+2
				SDPNAL	25	2.505e+0	2.69e+1	5	1.075e-9	7.05e+2
				SeDuMi	300	4.748e-1	2.62e+2	5	1.131e-9	6.89e+2
400	5	2610	0.7624	AFPC-BB	15	1.252e+1	6.23e+1	5	5.825e-7	1.22e+3
				SDPNAL	27	2.086e+0	8.69e+1	5	2.341e-8	5.03e+3
				SeDuMi	399	3.384e-1	4.88e+2	5	4.390e-8	5.03e+3
500	5	5124	0.4859	AFPC-BB	17	2.483e+1	5.33e+1	5	1.479e-5	7.92e+3
				SDPNAL	38	6.333e+0	2.53e+2	5	4.913e-8	1.84e+4
				SeDuMi	–	–	–	–	–	–

表 3.3: 利用 AFPC-BB, SDPNAL, SeDuMi 及 Gauss-Newton 迭代求 f 准确的平方和分解

问题				AFPC-BB			SDPNAL		
n	r	p	FR	秩	θ	时间/秒	秩	θ	时间/秒
400	10	10078	0.3924	10	1.712e+1	2.46e+1	66	1.093e+1	1.43e+2
500	20	24240	0.4047	20	1.497e+1	4.48e+1	113	4.232e+1	6.72e+2
1000	10	27101	0.3673	10	2.207e+1	3.70e+2	99	8.801e+1	2.70e+3
1000	50	95367	0.5114	50	1.009e+1	6.56e+2	218	9.200e+1	9.92e+3
1500	10	45599	0.3280	10	3.310e+1	1.00e+3	121	3.408e+1	3.72e+4
1500	50	122742	0.6011	50	1.508e+1	3.84e+3	226	3.790e+1	1.36e+4

表 3.4: 利用 AFPC-BB 和 SDPNAL 求 f 准确的平方和分解.

返回的数值 Gram 矩阵直接恢复出半正定多项式准确的平方和分解. 例如, 最后一个例子中 Gram 矩阵的维数为 1500×1500, 秩为 50. 随机构造的半正定多项式为

$$f = 498w^{34}x^4z^2 - 160w^{31}x^3y^2z^3 + 58x^6z^2 + \underbrace{\cdots}_{122399\text{ 项}},$$

问题中等式限制条件的个数超过100000. AFPC-BB 算法能够在一小时内返回给定精度的数值 Gram 矩阵. 通过将此矩阵的每个元素舍入到与之最近的整数, 我们即可得到上述多项式准确的 50 项平方和分解.

第四章　低秩矩量矩阵恢复和多项式系统实根求解

4.1　引言

给定多项式方程组

$$\begin{cases} g_1(x_1,\ldots,x_n) &= 0,\\ g_2(x_1,\ldots,x_n) &= 0,\\ \vdots & \\ g_{s_1}(x_1,\ldots,x_n) &= 0, \end{cases} \tag{4.1}$$

其中 $g_i \in \mathbb{R}[x_1,\ldots,x_n]$, $i = 1,\ldots,s_1$. 上述非线性多项式方程组求解是一个基本而又重要的问题. 多年来, 许多专家学者致力于此项问题的研究, 参见 [9, 36, 68, 121, 124, 126] 以及被广泛使用的软件包, 例如：由 Jan Verschelde 给出的基于同伦算法的软件包 PHCpack [134]. 在工程实践、经济学、信息安全和动力学等方面有大量的实际问题最终转化为求多项式方程组的实根, 或是满足某些不等式限制条件

$$\begin{cases} g_{s_1+1}(x_1,\ldots,x_n) &\geq 0,\\ g_{s_1+2}(x_1,\ldots,x_n) &\geq 0,\\ \vdots & \\ g_{s_2}(x_1,\ldots,x_n) &\geq 0, \end{cases} \tag{4.2}$$

的实根, 其中 $g_i \in \mathbb{R}[x_1,\ldots,x_n]$, $i = s_1+1,\ldots,s_2$.

求多项式方程组实根的方法大体上分为两种: 符号方法和数值方法, 包括 Gröbner 基方法, 区间算法, 同伦方法等, 见 [10, 81, 88]. 近年来, Chesi [23–25] 及 Lasserre [46, 63–66] 等人致力于利用半正定规划等数值算法来求解多项式系统的实根. 正如文献 [64] 中所述, 使用半正定规划技术最大的优点在于它通过直接探索问题的实代数特性而求得方程组的实根, 从而避免了计算多余的与复根相关的信息. 文献 [64–66] 中提出一种矩量松弛方法, 即通过求解一系列半

正定规划问题

$$\begin{cases} \min & 1 \\ \text{s. t.} & y_0 = 1, \\ & M_t(y) \succeq 0, \\ & M_{t-d_j}(g_j\, y) = 0, \quad j = 1, \ldots, s_1, \\ & M_{t-d_j}(g_j\, y) \succeq 0, \quad j = s_1+1, \ldots, s_2, \end{cases} \tag{4.3}$$

得到一系列矩量矩阵 $M_t(y)$, 其中 $d_j := \lceil \deg(g_j)/2 \rceil, j = 1, \ldots, s_2$. 文献 [64] 中证明了如果得到的矩量矩阵 $M_t(y)$ 满足条件 (4.4), 那么可以通过求矩阵 $M_t(y)$ 的像空间的一组基和相应乘法矩阵的公共特征向量得到多项式方程组的实根. 下面的定理可以作为求解一系列半正定规划问题 (4.3) 的终止性判定条件 (见 [31]):

定理 4.1. *[64] 对于零维多项式系统 (4.1), 设 $t \geq d$, $M_t(y)$ 为问题 (4.3) 的可行解, 且满足 $\operatorname{rank} M_t(y)$ 最大. 如果存在 $d \leq k \leq t$ 满足*

$$\operatorname{rank}\, M_k(y) = \operatorname{rank}\, M_{k-d}(y), \tag{4.4}$$

其中 $d = \max_{1 \leq j \leq s_2} d_j$. 那么 $\langle \ker M_k(y) \rangle = I(V_{\mathbb{R}}(I))$. 多项式系统 (4.1), (4.2) 的实根个数等于 $\operatorname{rank}\, M_k(y)$.

对于规模较小的半正定规划问题 (4.3), 我们可以利用内点法求解. 然而由于内点法的计算复杂度为 $\mathcal{O}(pm^3 + p^2m^2 + p^3)$, 其中 m 为矩量矩阵 $M_t(y)$ 的维数, p 为限制条件的个数. 当矩阵的维数 $m > 1000$ 限制条件个数 $p > 6000$ 时, 内点法不再适用. 虽然利用内点法返回的最大秩矩量矩阵能够求得多项式系统的全部实根. 但是, 当秩条件 (4.4) 成立时, t 值通常较大. 由于 t 阶矩量矩阵的维数 $m = \binom{n+t}{t}$, 此时等式限制条件的个数 $p = \sum_{j=1}^{s_2} \frac{1}{2}\binom{n+t-d_j}{n}\left(\binom{n+t-d_j}{n} + 1\right)$. 随着 t 值增大, 半定规划问题 (4.3) 的规模也将大大增加. 更重要的问题在于如果给定多项式系统有无穷多个实根时, 无论 t 值多大, 由内点法返回的最大秩矩量矩阵均不满足秩条件 (4.4). 文献 [46, 62] 中, 作者提出了一个新的模型, 将问题 (4.3) 的目标函数改为矩量矩阵的迹, 并且通过数值实验说明他们给出的基于 SeDuMi 的软件包 GloptiPoly 在求多项式方程组的部分实根时非常有效, 见 [62, 表 6.3, 6.4]. 由于半正定矩量矩阵的迹等于矩阵的核范数, 于是我们将问题 (4.3)

转化为求解下列矩阵核范数极小化问题

$$\begin{cases} \min & ||M_t(y)||_* \\ \text{s. t.} & y_0 = 1, \\ & M_t(y) \succeq 0, \\ & M_{t-d_j}(g_j\, y) = 0, \quad j = 1, \ldots, s_1, \\ & M_{t-d_j}(g_j\, y) \succeq 0, \quad j = s_1+1, \ldots, s_2. \end{cases} \tag{4.5}$$

如第三章所述, 矩阵核范数极小化问题可以直接由内点法求解, 但由于此类二阶方法计算量以及对计算机内存的要求都较大, 因此, 不适于求解大规模矩阵恢复问题. 例如, [62, 表 6.3, 6.4] 中, 部分例子在矩量矩阵的阶 t 取值较小时 GloptiPoly 已经无法计算出结果.

本章中, 我们研究如何运用第三章中给出的 AFPC-BB 算法求解满足已知等式和不等式限制条件的低秩矩量矩阵的恢复问题, 进而求得

$$K := \{x \in \mathbb{R}^n \mid g_1(x) = 0, \ldots, g_{s_1}(x) = 0; g_{s_1+1}(x) \geq 0, \ldots, g_{s_2}(x) \geq 0\}. \tag{4.6}$$

中的部分实点. 目前我们的算法无法保证求出多项式系统的全部实根. 对于较大规模的多项式系统, 如果只存在一个或少数几个实根, 我们的方法能够快速地将它们求解出来. 当 K 中的实根个数无穷时, 我们仍能求出部分孤立实根或是代数流形上的实根. 数值实验显示, 对利用半正定规划方法难以求解的例子, 我们的算法也能快速地求出其全部或部分的实根 (见表 4.1).

4.2　计算多项式系统的实根

给定 $\mathbb{R}^n$ 上测度 μ, 称 $y_\alpha := \int x^\alpha \mu(dx)$ 为对应于 μ 的阶为 α 的矩量, 称序列 $(y_\alpha)_{\alpha\in\mathbb{N}^n}$ 为对应于 μ 的矩量序列, 定义相应的给定 $t \in \mathbb{N}$, 称截断序列 $(y_\alpha)_{\alpha\in\mathbb{N}^n_t}$ 为对应于 μ 的阶为 t 的矩量序列. 给定序列 $y = (y_\alpha)_{\alpha\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$, 其相应的 (无穷维的) 矩量矩阵定义为

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

给定整数 $t \geq 1$ 和截断的序列 $y = (y_\alpha)_{\alpha\in\mathbb{N}^n_{2t}} \in \mathbb{R}^{\mathbb{N}^n_{2t}}$, 对于 $\alpha, \beta \in \mathbb{N}^n_t$, 其相应的 t 阶矩量矩阵定义为

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

设 $M_t(y)$ 的维数为 $m = \binom{n+t}{t}$. 给定一组单项式基 $(x^\alpha)_{\alpha\in\mathbb{N}^n}$, 定义矩阵 B_α 如下 (见 [60]):

$$B_\alpha(\zeta,\eta) := \begin{cases} 1, & \zeta+\eta=\alpha, \\ 0, & \text{其他}. \end{cases}$$

$M_t(y)$ 可以表示为

$$M_t(y) = \sum_{\alpha\in\mathbb{N}^n_t} B_\alpha y_\alpha.$$

记 $g_j(x) = \sum_{\alpha\in\mathbb{N}^n} g_{j,\alpha}x^\alpha \in \mathbb{R}[x]$ 且只有有限多个非零系数 $g_{j,\alpha}\in\mathbb{R}$. 如果记 $M_t(y)$ 的 (k,l) 位置的元素为 y_β, 那么维数等于 $m_j = \binom{n+t-d_j}{t-d_j}$ 的 $(t-d_j)$ 阶局部化矩阵 (Localizing Matrix) $M_{t-d_j}(g_jy)$ 的 (k,l) 位置的元素为

$$M_{t-d_j}(g_jy)(k,l) := \sum_\alpha g_{j,\alpha}y_{\alpha+\beta}.$$

对 $j=1,\dots,s_1, 1\le k\le m_j, k\le l\le m_j$, 如下定义

$$A^{(j)}_{(k,l)} := \sum_\alpha g_{j,\alpha}B_{\alpha+\beta}, \quad b^{(j)}_{(k,l)} := g_{j,0},$$

将 $A^{(j)}_{(k,l)}, b^{(j)}_{(k,l)}$ 分别排序, 记为 $A_1,\dots,A_p$ 和 $b_1,\dots,b_p$, 其中 $p=\sum_{j=1}^{s_1}(m_j^2+m_j)/2$. 定义线性算子 $\mathcal{A}:\mathbb{S}^m\to\mathbb{R}^p$ 为

$$\mathcal{A}(M_t(y)) := (\langle A_1, M_t(y)\rangle,\dots,\langle A_p, M_t(y)\rangle)^T, \tag{4.7}$$

其中 $\mathbb{S}^m$ 为 m 维对称半正定矩阵集合, 内积 $\langle A_i, M_t(y)\rangle := \mathrm{Tr}(A_i^T M_t(y))$.

因此, 问题 (4.5) 中所有等式限制条件 $M_{t-d_j}(g_jy)=0, j=1,\dots,s_1$ 可以表示为如下形式

$$\mathcal{A}(M_t(y)) = b, \tag{4.8}$$

其中 $b=(b_1,\dots,b_p)^T$. 对于 $z\in\mathbb{R}^p$, 算子 $\mathcal{A}$ 的伴随算子 $\mathcal{A}^*:\mathbb{R}^p\longrightarrow\mathbb{S}^m$ 定义为

$$\mathcal{A}^*(z) := A_1z_1+\cdots+A_pz_p. \tag{4.9}$$

注 2. 注意到, 矩量矩阵 $M_t(y)$ 是对称的, 那么在实验过程中, 我们只需求其上三角部分. 将 $M_t(y)$ 的上三角部分的列首尾相接排成一个向量, 记为

$\mathbf{svec}(M_t(y)) \in \mathbb{R}^{(m^2+m)/2}$. 同时, 将矩阵 $A_i \in \mathbb{R}^{m\times m}$ 的上三角部分以相同的方式排成向量 $\mathbf{svec}(A_i) \in \mathbb{R}^{(m^2+m)/2}$, $i = 1, \ldots, p$, 则 (4.8) 等价于

$$A\,\mathbf{svec}(M_t(y)) = b,$$

其中

$$A = \begin{pmatrix} \mathbf{svec}(A_1)^T \\ \vdots \\ \mathbf{svec}(A_p)^T \end{pmatrix}.$$

如上定义的矩阵 A 非常稀疏, 在程序实现中, 我们只存储矩阵 A 的非零元素及下标 (i, j).

算子 $\mathcal{A}$ 的伴随算子 (4.9) 也可相应地转化为 $\mathcal{A}^*(z) = \mathbf{smat}(A^T z)$, 其中算子 $\mathbf{smat}$ 为 $\mathbf{svec}$ 的逆运算, 即 $\mathbf{smat}(A^T z) \in \mathbb{R}^{m\times m}$ 为对称矩阵, 其上三角部分的元素从左到右、从上到下排列构成向量 $A^T z$.

对于半正定限制条件 $M_{t-d_j}(g_j y) \succeq 0$, $j = s_1 + 1, \ldots, s_2$, 通过引入松弛矩阵变元 $Z_j \in \mathbb{S}^{m_j}$, 将其转化为如下等式限制条件.

$$M_{t-d_j}(g_j y) = Z_j, \quad Z_j \succeq 0. \tag{4.10}$$

按照算子 $\mathcal{A}$ 的定义, (4.10) 中的每一个等式都可以写成如下形式

$$\mathcal{C}_j(M_t(y)) = \mathbf{svec}(Z_j), \quad j = s_1 + 1, \ldots, s_2, \tag{4.11}$$

其中算子 $\mathcal{C}_j : \mathbb{S}^m \to \mathbb{R}^{(m_j^2+m_j)/2}$, $j = s_1 + 1, \ldots, s_2$.

基于上述算子的定义及对限制条件的转化, 矩量矩阵核范数极小化问题 (4.5) 可转化为

$$\begin{cases} \min & ||X||_* \\ \text{s. t.} & \mathcal{A}(X) = b, \\ & X = X^T, X \succeq 0, \\ & \mathcal{C}_j(X) = \mathbf{svec}(Z_j), \quad j = s_1 + 1, \ldots, s_2, \\ & Z_j = Z_j^T, Z_j \succeq 0, \quad j = s_1 + 1, \ldots, s_2. \end{cases} \tag{4.12}$$

线性等式限制条件可做松弛处理, 得到 Lagrange 形式的正则的极小化问题

$$\min_{X\in\mathbb{S}_+^m, Z_j\in\mathbb{S}_+^{m_j}, j=s_1+1,\ldots,s_2} \mu\|X\|_* + \frac{1}{2}\|\mathcal{A}(X)-b\|_2^2 + \frac{1}{2}\sum_{j=s_1+1}^{s_2}\|\mathcal{C}_j(X)-\mathbf{svec}(Z_j)\|_2^2, \tag{4.13}$$

其中 $\mathbb{S}_+^m, \mathbb{S}_+^{m_j}$ 分别表示 m、m_j 维对称半正定矩阵组成的集合, 参数 $\mu>0$.

注意到, 问题 (4.13) 的目标函数中变元 $Z_j, j=s_1+1,\ldots,s_2$, 相互独立, 并共同依赖于 X 的取值. 因此, 当 X 给定时, 可以分别求出 $Z_j, j=s_1+1,\ldots,s_2$. 给定 Z_j, 可以利用第三章给出的 AFPC-BB 算法求出低秩矩阵 X. 这种对于不同的变元分开求解的方法称为交替方向法 (Alternating Direction Method). 它能够有效地求解目标函数具有可分离结构的带线性限制条件的凸优化问题. 这种求解的思想最早由 Gabay 和 Mercier [41, 42] 提出, 关于此类方法近期的发展可参见 [22, 37, 74, 137].

对于给定的 $X=\hat{X}\in\mathbb{S}_+^m$, 问题 (4.13) 的目标函数的前两项为常数, 所有的变元 $Z_j, j=s_1+1,\ldots,s_2$ 相互独立. 问题的最优解可以通过求解下面的最小二乘问题而得到

$$\min_{Z_j\in\mathbb{S}_+^{m_j}} \|\mathcal{C}_j(\hat{X})-\mathbf{svec}(Z_j)\|_2^2, \tag{4.14}$$

其中 $j=s_1+1,\ldots,s_2$.

设对称矩阵 $Y\in\mathbb{S}^m$ 的谱分解为 $Y=\sum_i \lambda_i q_i q_i^T$, 其中 λ_i 为矩阵 Y 的特征值, q_i 为相应的正交特征向量. 则矩阵 $Y\in\mathbb{S}^m$ 在半正定锥 $\mathbb{S}_+^m$ 和极锥 $\mathbb{S}_-^m$ 上的投影分别为

$$Y_+ = \sum_{\lambda_i>0}\lambda_i q_i q_i^T, \quad Y_- = \sum_{\lambda_i<0}\lambda_i q_i q_i^T. \tag{4.15}$$

从而有

$$Y = Y_+ + Y_-. \tag{4.16}$$

因此, 问题 (4.14) 唯一的最优解为

$$\hat{Z}_j = \mathbf{smat}(\mathcal{C}_j(\hat{X}))_+. \tag{4.17}$$

另一方面, 对于给定的 $Z_j = \hat{Z}_j$, $j = s_1+1, \ldots, s_2$, 原问题 (4.13) 可以写为关于 X 的极小化问题

$$\min_{X\in\mathbb{S}_+^m} \mu\|X\|_* + \frac{1}{2}\|\mathcal{A}(X) - b\|_2^2 + \frac{1}{2}\sum_{j=s_1+1}^{s_2} \|\mathcal{C}_j(X) - \mathbf{svec}(\hat{Z}_j)\|_2^2. \tag{4.18}$$

利用上一章中介绍的加速的不动点迭代方法 (AFPC-BB) 来求解此问题. 给定 $\tau > 0$, 定义算子

$$h(X, Z_{s_1+1}, \ldots, Z_{s_2}) := X - \tau\mathcal{A}^*(\mathcal{A}(X) - b) - \sum_{j=s_1+1}^{s_2} \tau\mathcal{C}_j^*(\mathcal{C}_j(X) - \mathbf{svec}(Z_j)). \tag{4.19}$$

设 μ 为给定的正实数, X_0 为初始点. 求解问题 (4.18) 的迭代公式为

$$X^{k+1} = \mathcal{T}_{\tau\mu}(h(X^k, \hat{Z}_{s_1+1}, \ldots, \hat{Z}_{s_2})). \tag{4.20}$$

其中 $\mathcal{T}_{\tau\mu}$ 为第三章中定义 3.1 给出的阈值算子, $k = 0, 1, 2, \ldots$.

第三章中定理 3.3 证明了当算子 $\mathcal{A}$ 和 $\mathcal{C}_j$, $j = s_1+1, \ldots, s_2$, 满足一定的条件时, 由迭代公式 (4.20) 得到的序列 $\{X^k\}$ 收敛到问题 (4.18) 的最优解.

固定 (4.20) 中的迭代数为 1, 即可得到迭代公式

$$\begin{cases} Z_j^{k+1} = & \mathbf{smat}(\mathcal{C}_j(X^k))_+, \quad j = s_1+1, \ldots, s_2, \\ X^{k+1} = & \mathcal{T}_{\tau\mu}(h(X^k, Z_{s_1+1}^{k+1}, \ldots, Z_{s_2}^{k+1})). \end{cases} \tag{4.21}$$

下面给出问题 (4.13) 的最优解满足的充分条件.

定理 4.2. *给定 $\mu > 0$, 如果算子 $\mathcal{A}$ 和 $\mathcal{C}_j, j = s_1+1, \ldots, s_2$ 满足*

$$\|\mathcal{A}(X^*) - b\|_2^2 + \sum_{j=s_1+1}^{s_2} \|\mathcal{C}_j(X^*) - \mathbf{svec}(Z_j^*)\|_2^2 < \frac{\mu^2}{m \max_{1\le j\le s_2} \|g_j\|_2^2}. \tag{4.22}$$

而且, $X^* \in \mathbb{S}_+^m$ 和 $Z_j^* \in \mathbb{S}_+^{m_j}$, $j = s_1+1, \ldots, s_2$ 满足

$$\begin{cases} Z_j^* = \mathbf{smat}(\mathcal{C}_j(X^*))_+, \quad j = s_1+1, \ldots, s_2, \\ X^* = \mathcal{T}_{\tau\mu}(h(X^*, Z_{s_1+1}^*, \ldots, Z_{s_2}^*)), \end{cases} \tag{4.23}$$

那么 $(X^*, Z_{s_1+1}^*, \ldots, Z_{s_2}^*)$ 是问题 (4.13) 唯一的最优解.

证明. 给定 $X^* \in \mathbb{S}_+^m$, 由 (4.17) 式可知, 问题 (4.13) 关于变元 Z_j 的解为

$$\mathbf{smat}(\mathcal{C}_j(X^*))_+, \quad j = s_1+1, \ldots, s_2.$$

给定 $Z_j^* \in \mathbb{S}_+^{m_j}, j = s_1+1, \ldots, s_2$, 由于问题 (4.13) 的目标函数是严格凸函数, 则存在唯一最小值. 令 $\nu = \tau\mu$,

$$Y^* = h(X^*, Z_{s_1+1}^*, \ldots, Z_{s_2}^*) = X^* + E \in \mathbb{S}^m,$$

其中

$$E = -\tau\mathcal{A}^*(\mathcal{A}(X^*) - b) - \sum_{j=s_1+1}^{s_2} \tau\mathcal{C}_j^*(\mathcal{C}_j(X^*) - \mathbf{svec}(Z_j^*)).$$

不失一般性, 将 Y^* 的特征值按照如下方式排序

$$\lambda_1(Y^*) \geq \cdots \geq \lambda_{k_1}(Y^*) \geq \nu > \lambda_{k_1+1}(Y^*) \geq \cdots > 0 > \cdots \geq \lambda_k(Y^*),$$
$$\lambda_{k+1}(Y^*) = \cdots = \lambda_m(Y^*) = 0.$$

计算 Y^* 的 Schur 分解

$$Y^* = Q_1\Lambda_1 Q_1^T + Q_2\Lambda_2 Q_2^T,$$

其中 $\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_{k_1})$, $\Lambda_2 = \mathrm{diag}(\lambda_{k_1+1}, \ldots, \lambda_k)$. 从而有,

$$\mathcal{T}_\nu(Y^*) = Q_1(\Lambda_1 - \nu I)Q_1^T,$$

且

$$Y^* - \mathcal{T}_\nu(Y^*) = \nu(Q_1Q_1^T + Z), \quad Z = \nu^{-1}Q_2\Lambda_2 Q_2^T.$$

显然有 $Q_1^T Z = 0$.

- 如果 $\lambda_{k_1+1}(Y^*) \geq |\lambda_k(Y^*)|$, 那么 $\|Z\|_2 = \lambda_{k_1+1}(Y^*)/\nu < 1$.
- 否则, 有

$$\begin{aligned}\|E\|_F^2 \leq & \tau^2\|\mathcal{A}^*(\mathcal{A}(X^*) - b)\|_F + \tau^2 \sum_{j=s_1+1}^{s_2} \|\mathcal{C}_j^*(\mathcal{C}_j(X^*) - \mathbf{svec}(Z_j^*)\|_F^2 \\ \leq & \tau^2 m \max_{1\leq j\leq s_2} \|g_j\|_2^2\big(\|\mathcal{A}(X^*) - b\|_2^2 + \sum_{j=s_1+1}^{s_2} \|\mathcal{C}_j(X^*) - \mathbf{svec}(Z_j^*)\|_2^2\big) \\ < & \tau^2\mu^2.\end{aligned}$$

注意到 $E \in \mathbb{S}^m$ 且 $W^* \in \mathbb{S}^m_+$. 由 [44, 定理8.1.5], 可得

$$\|Z\|_2 = \frac{|\lambda_k(Y^*)|}{\nu} = \frac{\max\{|\lambda_1(E)|, |\lambda_m(E)|\}}{\nu} \leq \frac{\|E\|_F}{\nu} < 1.$$

由定理 3.1 知 $Y^* - \mathcal{T}_\nu(Y^*) \in \nu\partial\|\mathcal{T}_\nu(Y^*)\|_*$, 即 $0 \in \nu\partial\|\mathcal{T}_\nu(Y^*)\|_* + \mathcal{T}_\nu(Y^*) - Y^*$. 则

$$0 \in \mu\partial\|X^*\|_* - \mathcal{A}^*(\mathcal{A}(X^*) - b) - \sum_{j=s_1+1}^{s_2} \mathcal{C}_j^*(\mathcal{C}_j(X^*) - \mathbf{svec}(Z_j^*)).$$

因此, 由定理 3.2 知 $(X^*, \mathbf{smat}(\mathcal{C}_{s_1+1}(X^*))_+, \ldots, \mathbf{smat}(\mathcal{C}_{s_2}(X^*))_+)$ 为问题 (4.13) 唯一的最优解. □

4.3　收敛性分析

本节中, 分析迭代方法 (4.21) 的收敛性. 阈值算子 $\mathcal{T}_\nu$ 的非扩张性已经在引理 3.4 中证明, 即对任给的 $X_1, X_2 \in \mathbb{S}^m$,

$$\|\mathcal{T}_\nu(X_1) - \mathcal{T}_\nu(X_2)\|_F \leq \|X_1 - X_2\|_F.$$

而且

$$\|X_1 - X_2\|_F = \|\mathcal{T}_\nu(X_1) - \mathcal{T}_\nu(X_2)\|_F \Longleftrightarrow X_1 - X_2 = \mathcal{T}_\nu(X_1) - \mathcal{T}_\nu(X_2).$$

令 $Z_j = \mathbf{smat}(\mathcal{C}_j(X))_+, \quad j = s_1 + 1, \ldots, s_2$, 仍然记 h 为作用到矩阵 X 上的函数:

$$h(X) := h\left(X, \mathbf{smat}(\mathcal{C}_{s_1+1}(X))_+, \ldots, \mathbf{smat}(\mathcal{C}_{s_2}(X))_+\right). \tag{4.24}$$

下面的引理说明在某种条件下, $h(\cdot)$ 是非扩张的.

引理 4.3. *设步长 $\tau \in (a, b)$, 其中*

$$\begin{cases} a = 1/(\|\mathcal{A}\|_2^2 + \sum_{j=s_1+1}^{s_2} \|\mathcal{C}_j\|_2^2), \\ b = \min(3a, 1/2\sum_{j=s_1+1}^{s_2} \|\mathcal{C}_j\|_2^2). \end{cases} \tag{4.25}$$

那么 (4.24) 中定义的函数 $h(\cdot)$ 是非扩张的, 即对任意 $X_1, X_2 \in \mathbb{S}^m$,

$$\|h(X_1) - h(X_2)\|_F \leq \|X_1 - X_2\|_F.$$

而且

$$\|h(X_1) - h(X_2)\|_F = \|X_1 - X_2\|_F \Longleftrightarrow h(X_1) - h(X_2) = X_1 - X_2.$$

证明. 由 (4.16) 式知

$$\begin{aligned}
&\|X_1 - X_2\|_F^2 - \|X_{1+} - X_{2+}\|_F^2 \\
&= \langle X_1 - X_2, X_1 - X_2\rangle - \langle X_{1+} - X_{2+}, X_{1+} - X_{2+}\rangle \\
&= \|X_{1-} - X_{2-}\|_F^2 - 2\mathrm{Tr}(X_{1+}^T X_{2-} + X_{1-}^T X_{2+}) \\
&\geq 0.
\end{aligned}$$

因此, 对所有 $j = s_1 + 1, \ldots, s_2$, 有

$$\begin{aligned}
\|Z_{1,j} - Z_{2,j}\|_F &= \|\mathbf{smat}(\mathcal{C}_j(X_1))_+ - \mathbf{smat}(\mathcal{C}_j(X_2))_+\|_F \\
&\leq \|\mathbf{smat}(C_j(X_1)) - \mathbf{smat}(C_j(X_2))\|_F \\
&\leq \|\mathcal{C}_j\|_2 \, \|X_1 - X_2\|_F. \qquad (4.26)
\end{aligned}$$

可得

$$\begin{aligned}
&\|h(X_1) - h(X_2)\|_F \\
&\leq \|I - \tau\mathcal{A}^*\mathcal{A} - \sum_{j=s_1+1}^{s_2} \tau\mathcal{C}_j^*\mathcal{C}_j\|_2 \, \|X_1 - X_2\|_F + \sum_{j=s_1+1}^{s_2} \tau\|\mathcal{C}_j\|_2 \, \|Z_{1,j} - Z_{2,j}\|_F \\
&\leq \left(\|I - \tau\mathcal{A}^*\mathcal{A} - \sum_{j=s_1+1}^{s_2} \tau\mathcal{C}_j^*\mathcal{C}_j\|_2 + \sum_{j=s_1+1}^{s_2} \tau\|\mathcal{C}_j\|_2^2 \right) \|X_1 - X_2\|_F. \qquad (4.27)
\end{aligned}$$

当 $\tau \in (a, b)$ 时, 其中 a, b 如 (4.25) 式所示, 可以推出

$$\|I - \tau\mathcal{A}^*\mathcal{A} - \sum_{j=s_1+1}^{s_2} \tau\mathcal{C}_j^*\mathcal{C}_j\|_2 + \sum_{j=s_1+1}^{s_2} \tau\|\mathcal{C}_j\|_2^2 \leq 1. \qquad (4.28)$$

因此, $\|h(X_1) - h(X_2)\|_F \leq \|X_1 - X_2\|_F$. 与此同时, $\|h(X_1) - h(X_2)\|_F = \|X_1 - X_2\|_F$ 成立当且仅当上述不等式 (4.26)、(4.27) 和 (4.28) 中等号均成立. 从而有,

$$(I - \tau\mathcal{A}^*\mathcal{A} - \sum_{j=s_1+1}^{s_2} \tau\mathcal{C}_j^*\mathcal{C}_j)(X_1 - X_2) + \sum_{j=s_1+1}^{s_2} \tau\mathcal{C}_j^*(\mathbf{svec}(Z_{1,j}) - \mathbf{svec}(Z_{2,j})) = X_1 - X_2,$$

$$Z_{1,j} - Z_{2,j} = \mathbf{smat}(C_j(X_1)) - \mathbf{smat}(C_j(X_2)), \quad j = s_1 + 1, \ldots, s_2.$$

因此, $h(X_1) - h(X_2) = X_1 - X_2$. □

下面的定理说明迭代算法 (4.21) 收敛到问题 (4.13) 的最优解.

定理 4.4. 设步长 $\tau \in (a,b)$, 其中 a,b 如 *(4.25)* 式所定义. $X^* \in \mathbb{S}_+^m$ 和 $Z_j^* \in \mathbb{S}_+^{m_j}$, $j = s_1+1,\ldots,s_2$ 满足条件*(4.23)* 和 *(4.22)*. 那么由迭代公式 *(4.21)* 计算得到的序列

$$(X^k, Z_{s_1+1}^k, \ldots, Z_{s_2}^k)$$

收敛到问题 *(4.13)* 的最优解 $(X^*, Z_{s_1+1}^*, \ldots, Z_{s_2}^*)$.

证明. 此定理的证明与定理 3.6 类似. 令 $\nu = \tau\mu$. 由引理 3.4、4.3 可知算子 $\mathcal{T}_\nu(\cdot)$ 和 $h(\cdot)$ 均为非扩张的. 则 $\mathcal{T}_\nu(h(\cdot))$ 也是非扩张的. 因此, $\{X^k\}$ 为有界序列, 那么一定存在 $\{X^k\}$ 的收敛子列 $\{X^{k_j}\}$.

设 $\widetilde{X} = \lim_{j\longrightarrow\infty} X^{k_j}$ 且满足条件 (4.22). 因为 $X^* = \mathcal{T}_\nu(h(X^*))$, 有

$$\|X^{k+1} - X^*\|_F = \|\mathcal{T}_\nu(h(X^k)) - \mathcal{T}_\nu(h(X^*))\|_F \le \|h(X^k) - h(X^*)\|_F \le \|X^k - X^*\|_F.$$

即 $\{\|X^k - W^*\|_F\}$ 是单调非增列, 且收敛到 $\|\widetilde{X} - X^*\|_F$. 由函数 $\mathcal{T}_\nu(h(\cdot))$ 的连续性可知

$$\mathcal{T}_\nu(h(\widetilde{X})) = \lim_{j\longrightarrow\infty} \mathcal{T}_\nu(h(X^{k_j})) = \lim_{j\longrightarrow\infty} X^{k_j+1},$$

这说明 $\mathcal{T}_\nu(h(\widetilde{X}))$ 仍为 $\{X^k\}$ 的一个极限点. 从而有

$$\|\mathcal{T}_\nu(h(\widetilde{X})) - \mathcal{T}_\nu(h(X^*))\|_F = \|\mathcal{T}_\nu(h(\widetilde{X})) - X^*\|_F = \|\widetilde{X} - X^*\|_F.$$

由引理 3.4 和引理 4.3 可以推出

$$\mathcal{T}_\nu(h(\widetilde{X})) - \mathcal{T}_\nu(h(X^*)) = h(\widetilde{X}) - h(X^*) = \widetilde{X} - X^*,$$

即 $\mathcal{T}_\nu(h(\widetilde{X})) = \widetilde{X}$. 根据定理 4.2 可知 $\widetilde{X}$ 为问题 (4.13) 的最优解, 即 $\widetilde{X} = X^*$. 因此,

$$\lim_{k\longrightarrow\infty} \|X^k - X^*\|_F = 0,$$

即 $\{X^k\}$ 收敛到唯一的极限点 X^*.

同时, 利用 (4.26) 式可得

$$\|Z_j^k - Z_j^*\|_F \le \|\mathcal{C}_j\|_2 \|X^k - X^*\|_F.$$

因此,

$$\lim_{k\longrightarrow\infty} \|Z_j^k - Z_j^*\|_F = 0, \quad j = s_1+1,\ldots,s_2.$$

即 $(Z_{s_1+1}^k, \ldots, Z_{s_2}^k)$ 收敛到唯一的极限点 $(Z_{s_1+1}^*, \ldots, Z_{s_2}^*)$. □

4.4 算法及实现

本节中, 我们给出了求解矩量矩阵恢复问题的算法以及算法在实现中的细节. 同时我们也介绍了如何利用矩量矩阵求非线性多项式系统 (4.1) 和 (4.2) 的实根.

算法：*MMCRSolver*

输入：▸ 多项式 $g_1(x),\dots,g_{s_1}(x),g_{s_1+1}(x),\dots,g_{s_2}(x)\in\mathbb{R}[x]$, 参数 $0<\tau_{min}<\tau_0<\tau_{max}<\infty$, $\mu_1>\bar{\mu}>0$, $\eta>0$ 和误差界 $\epsilon>0$.

输出：▸ 实根 $v_1,\dots,v_r$.

1. 令 $t=\max_{1\le j\le s_2}\lceil \deg(g_j)/2\rceil$, $a_0=1$, 矩阵 X^0 中只有一个非零元 $X^0(1,1)=1$.

2. 计算 t 阶松弛下的算子 $\mathcal{A}$ 和 $\mathcal{C}_j$, $j=s_1+1,\dots,s_2$.

3. 对于 $\mu=\mu_1,\dots,\mu_L$, 进行循环

   (a) 利用 BB 技术选择步长 τ_k 且满足 $\tau_{min}\le\tau_k\le\tau_{max}$;

   (b) 计算 $Y^k=X^k+\frac{a_{k-1}-1}{a_k}(X^k-X^{k-1})$;

   (c) 计算 $Z_j^{k+1}=\mathbf{smat}(\mathcal{C}_j(Y^k))_+$, $j=s_1+1,\dots,s_2$;

   (d) 计算 $X^{k+1}=\mathcal{T}_{\tau_k\mu_k}(Y^k-\tau_k\mathcal{A}^*(\mathcal{A}Y^k-b)-\sum_{j=s_1+1}^{s_2}\tau_k\mathcal{C}_j^*(\mathcal{C}_j(Y^k)-\mathbf{svec}(Z_j^{k+1})))$;

   (e) 计算$a_{k+1}=\frac{1+\sqrt{1+4a_k^2}}{2}$;

   (f) 如果满足停机条件 (4.31) 或 (4.32), 那么返回 $X_{\rm opt}$.

4. 如果矩阵 $X_{\rm opt}$ 满足条件 (4.4), 那么

   (a) 利用 (4.29) 式计算乘法矩阵;

   (b) 利用 (4.30) 式计算实根.

5. 如果条件 (4.4) 不满足, 令 $t=t+1$ 并返回到第 2 步.

内部循环 3 的目的是返回低秩矩量矩阵. 我们采用连续性技术, 即对于单调下降的参数 μ_k, 利用迭代公式 (4.21) 来求解一系列的问题 (4.13). 在求解参数为 μ_{k+1} 的问题时, 起始点选为上一步迭代得到的参数为 μ_k 的问题 (4.13) 的近似最优解. 令

$$\mu_{k+1} = \max(\eta\mu_k, \bar{\mu}), \quad k = 1, \ldots, L-1,$$

其中 $0 < \eta < 1$ 表示相邻两个 μ 值减小的比例, $L = \lceil \log_\eta \mu_1/\bar{\mu} \rceil$. 另外, 定理 4.4 中的条件 $\tau \in (a, b)$ 保证了此内部循环的收敛性. 但是此选择可能过于保守, 导致算法的收敛速度较慢, 迭代步数较多. 因此, 在计算过程中, 我们使用上一章第 3.4.2 节中介绍的 Barzilai-Borwein 技术来选取参数 τ_k. 这部分计算中主要的计算量在于 3(d) 步中矩阵的 Schur 分解. 按照第 3.4.1 节中介绍的方法, 我们使用 Matlab 中的 PROPACK [59] 软件包计算对称矩阵的部分较大特征值和相应的特征向量. 每一步迭代中计算的特征值个数 s_k 会直接影响到 MMCRSover 最终返回的多项式系统实根的个数. 如果只需计算少量实根, 我们在迭代过程中利用 PROPACK 软件包求矩阵的部分较大特征值及相应的特征向量, 从而提高 MMCRSover 的计算效率. 表 4.1 中的例子“puma”清楚地证实了这一点.

4.4.1　计算矩量矩阵的秩

对于第 3 步返回的矩量矩阵 $M_t(y)$, 为了检验对于 $1 \le k \le t$ 条件 (4.4) 是否满足, 我们需要计算矩阵 $M_t(y)$ 的每一阶主子块 $M_k(y)$ $(k \le t)$ 的秩. 数值矩阵的秩在计算过程中敏感度较高. 因此, 我们采用数值稳定性较好的奇异值分解的方法来求矩阵的秩.

设矩阵 X 的奇异值分解为 $X = U\Sigma V^T$, 其中 U、V 为正交矩阵, Σ 为对角矩阵, 对角线元素为矩阵 X 的奇异值 (即矩阵 XX^T 的特征值). 在精确计算下, Σ 对角线上非零元素的个数 $r := \mathrm{rank}(X)$.

设 $\sigma_1(X) \ge \sigma_2(X) \ge \cdots \ge \sigma_m(X)$ 为矩阵 $X \in \mathbb{R}^{m\times m}$ 的奇异值. 给定容许误差 $\bar{\epsilon} > 0$, 矩阵的数值秩通常取满足 $\sigma_k(X) > \bar{\epsilon}$ 的最大的正整数 $k \in \mathbb{N}$. 如果相邻的两个奇异值之间的差距过大, 即 $\sigma_k(X) \ge \epsilon_{decay} \cdot \sigma_k(X)$, 那么取 $\mathrm{rank}(X) = k$. 这种判定方法也在文献 [33, 34, 64] 中使用过, 实验效果较好. 数值试验中, 通常取 $\bar{\epsilon} = 1e-8$, $\epsilon_{decay} = 10^3$.

在计算矩阵 $M_t(y)$ 的每一阶子块的数值秩的同时, 还能求出矩阵 $M_{k-1}(y)$ 的列向量组成空间的一组基. 设 $M_{k-1}(y)$ 的奇异值分解为 $M_{k-1}(y) = U\Sigma V^T$, 秩

为 r. 正交矩阵 U 中对应于非零奇异值 $\sigma_1 \geq \cdots \geq \sigma_r$ 的前 r 列 $\{u_1, \ldots, u_r\}$ 即构成矩阵 $M_{k-1}(y)$ 的列向量组成空间的一组基.

4.4.2 计算乘法矩阵和多项式系统的实根

此时矩量矩阵满足秩条件

$$\mathrm{rank} M_{k-1}(y) = \mathrm{rank} M_k(y).$$

参照文献 [101] 中的方法, 利用矩阵的奇异值分解 $M_{k-1}(y) = U\Sigma V^T$ 得到的 $M_{k-1}(y)$ 的像空间的一组基 $\{u_1, \ldots, u_r\}$ 来构造乘法矩阵.

对 $j = 1, \ldots, r$, 记 $[x]_{k-1} = (x^\alpha)_{\alpha \in \mathbb{N}^n_{k-1}}$ 为 n 变元次数小于等于 $k-1$ 的所有单项式构成的向量. 令

$$b_j = u_j^T [x]_{k-1},$$

得到一组多项式集合 $\mathcal{B} = \{b_1, \ldots, b_r\}$. 根据 [101] 中所述, x_j 关于 $\mathcal{B}$ 的乘法矩阵为

$$M_{x_j} = U_r^T \cdot N_{x_j} \cdot V^T \cdot S, \tag{4.29}$$

其中 $U_r = (u_1, \ldots, u_r)$, N_{x_j} 表示 $M_k(y)$ 中对应于单项式 $x_j \cdot [x]_{k-1}$ 的行构成的矩阵. S 为对角矩阵, 其对角线上的元素为 Σ 中前 r 个奇异值的倒数.

最后, 第 4(b) 步中, 多项式系统 (4.1), (4.2) 的实根可以通过求乘法矩阵 M_{x_j}, $j = 1, \ldots, n$ 的公共特征向量得到.

按照 [29, 46] 中的方法, 首先利用随机向量 $\omega = (\omega_1, \ldots, \omega_n)$ 构造乘法矩阵 M_{x_j} 的线性组合

$$M' = \sum_{j=1}^{n} \omega_j M_{x_j},$$

其中 $\omega_j \geq 0$, $\sum_{j=1}^n \omega_j = 1$. 而后计算矩阵 M' 的 Schur 分解

$$M' = QRQ^T,$$

其中 $Q = (q_1, \ldots, q_r)$ 为正交矩阵, R 为上三角矩阵, 其对角线元素为矩阵 M' 的特征值. 多项式系统的 r 个实根可以由下式计算得到:

$$v_j = (q_j^T M_{x_1} q_j, \ldots, q_j^T M_{x_n} q_j) \in \mathbb{R}^n, \quad j = 1, \ldots, r. \tag{4.30}$$

如果得到的根的精度太低, 那么以 $v_j, j = 1, \ldots, r$ 为初始点做牛顿迭代, 进而得到较高精度的解.

注 3. 对于零维多项式系统, 由定理 4.1 可知, 当 t 足够大时, 条件 (4.4) 总满足. 当 μ 趋于 0 时, 问题 (4.13) 的最优解渐进地收敛到问题 (4.5) 的全局最优解. 在数值实验中, 我们发现对于某些正维多项式系统, 由 MMCRSolver 返回的低秩矩量矩阵也满足条件 (4.4). 因此, 我们能够求出正维系统的部分孤立实根或代数流形上的实根.

4.5　数值实验

本节中, 给出了 MMCRSolver 在求多项式系统的实根问题上的表现. 实验中, 我们利用下面两个条件作为算法内部循环第 3 步的终止条件

$$\frac{\|\mathcal{A}(X_{\text{opt}}) - b\|_2}{\|b\|_2} < 0.005, \tag{4.31}$$

或

$$\frac{\|X^{k+1} - X^k\|_F}{\max(1, \|X^k\|_F)} < 10^{-4}. \tag{4.32}$$

在 MATLAB (Version 7.7.0.471)中运行 MMCRSolver. 下面所有实验数据都是在台式机 (Intel(R) Core(TM) i3-2100 CPU @ 3.10GHz and 2.00 GB of RAM)上运行所得. 程序代码可以在下面地址下载 http://www.mmrc.iss.ac.cn/~lzhi/Research/hybrid/FPCs/MMCRSolver

例 4.1. Camera Pose 这个例子来自于计算机视觉中相机的位置及角度的估算问题, 可以转化为求解下列近似系数的多项式方程组, 见 [102].

$$\begin{cases} g_1 &= x_1^2 + x_2^2 - 1.49071\, x_1x_2 - 4, \\ g_2 &= x_1^2 + x_3^2 - .400000\, x_1x_3 - 8, \\ g_3 &= x_1^2 + x_4^2 - .894427\, x_1x_4 - 4, \\ g_4 &= x_2^2 + x_3^2 - 1.49071\, x_2x_3 - 4, \\ g_5 &= x_2^2 + x_4^2 - .666667\, x_2x_4 - 8, \\ g_6 &= x_3^2 + x_4^2 - .894427\, x_3x_4 - 4. \end{cases}$$

当 $t=2$ 时, 矩量矩阵 $M_2(y)$ 满足

$$\operatorname{rank} M_2(y) = \operatorname{rank} M_1(y) = 2.$$

得到全部实根:

$$\begin{aligned} v_1 &= (-2.2361, -3.0000, -2.2361, -1.0023), \\ v_2 &= (2.2361, 3.0000, 2.2361, 1.0023). \end{aligned}$$

例 4.2. 下面的例子来自于文献 [13]. 此多项式系统有 20 个根, 其中 8 个为实根.

$$\begin{cases} g_1 &= 5{x_1}^9 - 6{x_1}^5 x_2 + x_1 {x_2}^4 + 2x_1 x_3, \\ g_2 &= -2{x_1}^6 x_2 + 2{x_1}^2 {x_2}^3 + 2x_2 x_3, \\ g_3 &= {x_1}^2 + {x_2}^2 - 0.265625. \end{cases}$$

这个多项式系统的特点是次数较高 $d=5$, 根的个数较多. 当 $t=6$ 时, 条件 (4.4) 成立

$$\operatorname{rank} M_6(y) = \operatorname{rank} M_1(y) = 3.$$

可得 3 个实根:

$$\begin{aligned} v_1 &= (0.5154, -0.0000, -0.0124), \\ v_2 &= (-0.5016, 0.1185, 0.0124), \\ v_3 &= (-0.0000, -0.5154, 0.0000). \end{aligned}$$

表 4.1 给出了MMCRSolver在一系列经典问题上的表现, 见(http://homepages.math.uic.edu/~jan/). 表中列出了变元数(var)、多项式系统的次数(deg)、矩量松弛阶 t 以及限制条件的个数 p. 我们还列出了分别由 MMCRSolver 和 GloptiPoly 两种算法计算满足条件 (4.4) 的矩量矩阵和求实根过程所用的 CPU 时间, 以及求得的实根的个数 (sol). 表中最后两列数据来自文献 [62] 中的表 6.3 和表 6.4.

如表所示, 对于前四个例子, 在两种方法求得的实根个数相同的前提下, MMCRSolver 用的时间明显较少. 例子“puma”一共有 16 个实根 (见 [79]). 尽管 MMCRSolver不能返回全部实根, 但是在矩量矩阵恢复过程中, 通过调

问题	var	deg	t	p	$CPU_{MMCRSolver}$/秒	sol	$CPU_{GloptiPoly}$/ 秒	sol
boon	6	4	4	21841	31.75	8	1220	8
eco8	8	3	3	11953	1.37	1	1310	1
heart	8	4	3	12853	53.09	2	1532	2
puma	8	2	3	14653	3.96	4	1136	4
puma	8	2	3	14653	6.61	13	1136	4
butcher	7	4	4	51877	214.38	1	-	-
d1	12	3	3	103559	76.55	4	-	-
kin1	12	2	3	103559	94.71	11	-	-
reimer5	5	6	6	107267	128.70	1	-	-

表 4.1: MMCRSolver 和 GloptiPoly 求得的实根个数及 CPU 时间的比较

节 PROPACK 所求的矩阵特征值的个数 s_k, MMCRSolver 能在 6.61 秒内返回此多项式系统的 13 个实根. 表中最后四个例子的规模相对较大, 限制条件的个数均超过 50000 或 100000, GloptiPoly 已无法求解如此规模的问题. 而 MMCRSolver 均能返回至少 1 个实根. 其中, 例子“butcher”来自于 POSSO 测试集, 是一个正维多项式系统. MMCRSolver 能够成功地求出此多项式系统在流形 $x_1 = x_3 = 0, x_5 = x_6 = -1$ 上的实根. 然而与基于同伦算法的软件包 PHCpack 相比, 仍有一些例子 (cassou, des18_3 和 rabmo) MMCRSolver 在短时间内不收敛或是不能返回正确的结果.

例 4.3. “puma”出自文献 [79].

$$
\begin{cases}
g_1 &= {x_1}^2 + {x_2}^2 - 1, \\
g_2 &= {x_3}^2 + {x_4}^2 - 1, \\
g_3 &= {x_5}^2 + {x_6}^2 - 1, \\
g_4 &= {x_7}^2 + {x_8}^2 - 1, \\
g_5 &= 0.0047x_1x_3 - 0.3578x_2x_3 - 0.2238x_1 - 0.0016x_2 - 0.9338x_4 + x_7 - 0.3571, \\
g_6 &= 0.2238x_1x_3 + 0.7623x_2x_3 + 0.2638x_1 - 0.07745x_2 - 0.6734x_4 \\
&\quad -0.6022x_6x_8 + 0.3578x_1 + 0.004731x_2 - 0.7623, \\
g_7 &= x_1 + 0.2238x_2 + 0.3461.
\end{cases}
$$

此多项式系统共有 16 个实根, 其中 4 个满足不等式限制条件：

$$\{x_5 \geq 0,\, x_6 \geq 0\}.$$

当 $t = 3$ 时, 矩量矩阵 $M_3(y)$ 满足

$$\mathrm{rank}M_1(y^*) = \mathrm{rank}M_3(y^*) = 4.$$

MMCRSolver 在 36.93 秒内即可求得半代数全部实根

$$v_1 = (0.6716, 0.7410, 0.9607, 0.2774, 0.6029, 0.7978, 0.9522, -0.3056),$$
$$v_2 = (0.1644, -0.9864, 0.2394, -0.9709, 0.9976, 0.0687, -0.6155, -0.7881),$$
$$v_3 = (0.1644, -0.9864, -0.9559, 0.2938, 0.9351, 0.3544, 0.9882, -0.1529),$$
$$v_4 = (0.6716, 0.7410, -0.2423, -0.9702, 0.9579, 0.2871, -0.5280, -0.8493).$$

第五章　正维多项式理想的实根的计算

5.1　引言

上一章中我们给出了一种求解大规模多项式系统部分实根的方法. 如果给定的多项式系统有无穷多的实根, 基于多项式理想 I 与代数簇 $V(I)$ 之间的对应关系, 多项式方程组的实根求解问题可以转化为对实根理想 $I(V_{\mathbb{R}}(I))$ 的研究.

问题 5.1. 设 $I = \langle h_1, \ldots, h_m \rangle \subseteq \mathbb{R}[x] = \mathbb{R}[x_1, \ldots, x_n]$ 是由多项式 $h_1, \ldots, h_m \in \mathbb{R}[x]$ 生成的理想, 且 $|V_{\mathbb{R}}(I)|$ 无穷, 求实根理想 $I(V_{\mathbb{R}}(I))$ 关于序 $\prec_{\text{tdeg}}$ 的 Gröbner 基.

实根理想 $I(V_{\mathbb{R}}(I))$ 的计算要比根理想的计算更加困难. 当 $|V_{\mathbb{R}}(I)|$ 有限时, Lasserre 等人提出了基于半正定规划的数值方法 [64, 66] 和符号–数值混合方法 [63, 65] 来计算实根理想 $I(V_{\mathbb{R}}(I))$ 的一组边界基 (Border Basis) 或 Gröbner 基. 他们基于矩量松弛的理论和方法都建立在 Flat Extension 定理的基础上.

定理 5.1 (Flat Extension 定理). *[31] 给定有限序列 $y \in \mathbb{R}^{\mathbb{N}^n_{2t}}$, 如果满足*

$$\text{rank} M_t(y) = \text{rank} M_{t-1}(y) \tag{5.1}$$

那么 y 能够延拓到 $\tilde{y} \in \mathbb{R}^{\mathbb{N}^n_{2t+2}}$, 使得 $\text{rank}\, M_t(\tilde{y}) = \text{rank}\, M_t(y)$.

对于具有正维实代数簇的多项式理想 I, Becker 和 Neuhaus [12] 提出了一种基于理想的准素分解的方法求正维实根理想 $I(V_{\mathbb{R}}(I))$, 相关工作可参见 [87, 141, 144]. 除此之外, 还有一类基于实代数几何中关键点的方法, 它们能够在实代数簇的每一个连通分支上求出一点, 见 [5, 6, 8, 9, 113]. 但是, 随着问题规模的增大, 此类针对正维多项式系统的符号方法在计算过程中会出现表达式迅速膨胀, 内存需求增加, 计算速度降低等问题, 因而无法满足实际应用的需求.

在正维情形下, 等式 (5.1) 不成立. 与定理 5.1 类似, 我们给出如下猜想.

猜想 5.1. 如果存在整数 (t, ℓ), $t \geq 2d, 1 \leq \ell \leq t - 2d$ 及 $y_1 \in \mathcal{K}_t^{gen}, y_2 \in \mathcal{K}_{t+1}^{gen}$ 满足条件 (5.4)-(5.5), 那么存在 $\tilde{y} \in \mathcal{K}_{t+2}$ 满足

$$\text{rank} M_{(t+1)-(\ell+1)}(y_2) = \text{rank} M_{(t+2)-(\ell+2)}(\tilde{y}). \tag{5.2}$$

本章中, 通过将几何对合理论与半正定矩量矩阵的性质相结合, 我们给出了半正定松弛

$$\left.\begin{array}{ll} \min & 1 \\ \text{s. t.} & y_0 = 1, \\ & M_t(y) \succeq 0, \\ & M_{t-d_j}(h_j\, y) = 0, \quad j = 1, \ldots, m. \end{array}\right\} \tag{5.3}$$

在正维情形下终止的判定定理 5.2. 对于 $t > \ell$, $M_{t-\ell}(y)$ 表示 $M_t(y)$ 的 $t-\ell$ 阶主子块, 相应的指标 $\alpha, \beta \in \mathbb{N}^n_t$ 满足 $|\alpha| \leq t-\ell$ 且 $|\beta| \leq t-\ell$. 对于 $t \geq d$, 令

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

其中的母元素 (generic) 构成的集合为

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

定理 5.2. *在 δ-正则坐标系下, 如果存在整数(t, ℓ), $t \geq 2d, 1 \leq \ell \leq t-2d$ 及 $y_1 \in \mathcal{K}_t^{gen}, y_2 \in \mathcal{K}_{t+1}^{gen}$ 满足下列条件*

$$\operatorname{rank} M_{t-\ell}(y_1) = \operatorname{rank} M_{(t+1)-(\ell+1)}(y_2), \tag{5.4}$$

$$\sum_{j=1}^{n} j\alpha_{t-\ell}^{(j)} \text{ 对于 } M_{t-\ell}(y_1) = \operatorname{corank} M_{(t+1)-\ell}(y_2) - \operatorname{corank} M_{(t+1)-(\ell+1)}(y_2), \tag{5.5}$$

那么 $\ker M_{t-\ell}(y_1)$ 是实根理想 $\sqrt[\mathbb{R}]{I}$ 关于序 $\prec_{\text{tdeg}}$ 的一组 Gröbner 基, 即

$$\langle \ker M_{t-\ell}(y_1) \rangle = \sqrt[\mathbb{R}]{I}. \tag{5.6}$$

定理中 $\alpha_{t-\ell}^{(j)}$, $j = 1, \ldots, n$ 为矩阵 $M_{t-\ell}(y_1)$ 的 Cartan 特征 (见定义 5.7). 基于定理 5.2, 我们给出如下算法计算实根理想 $\sqrt[\mathbb{R}]{I}$ 的 Gröbner 基.

算法 5.1. *计算实根理想 $\sqrt[\mathbb{R}]{I}$ 的 Gröbner 基*

输入: *理想 I 的生成集 $\{h_1, \ldots, h_m\}$ 和变元序.*

输出: *$\sqrt[\mathbb{R}]{I}$ 关于序 $\prec_{\text{tdeg}}$ 的 Gröbner 基.*

1. *求解问题 (5.3) 计算 $y \in \mathcal{K}_t^{gen}$. 对 $t \geq 2d$, $1 \leq \ell \leq t-2d$, 计算截断的矩量矩阵 $M_{t-\ell}(y)$ 的秩.*

2. 寻找最小的整数 t, 使得存在 ℓ 满足条件 (5.4) 和 (5.5). 对于固定的 t, 取满足条件 (5.4) 和 (5.5) 的最大的 ℓ.

3. 计算 $M_{t-\ell}(y)$ 的零空间的一组基 $\{v_1, \ldots, v_s\}$, 并返回多项式集合

$$\{v_1^T[x]_{t-\ell}, \ldots, v_s^T[x]_{t-\ell}\},$$

其中 $[x]_{t-\ell}$ 是由 n 个变元次数小于等于 $t-\ell$ 的所有单项式构成的向量.

注 4. 算法 5.1 的首要任务是求集合 $\mathcal{K}_t^{gen}$ 中的元素. 如文章 [64] 中所述, 利用内点法求解半定规划问题 (5.3), 即可得到满足限制条件的最大秩的矩量矩阵 (参见 [133, 139]). 为了检验条件 (5.4)-(5.5), 需要计算矩量矩阵 $M_t(y)$ 的秩和约化的行阶梯形式, 在此计算过程中存在数值稳定性的问题. 关于数值矩阵秩的计算, 我们使用矩阵的奇异值分解, 见 4.4.1 节. 如果 $M_t(y)$ 的奇异值满足 $\sigma_1 \geq \cdots \geq \sigma_r > 10^{-8} > \sigma_{r+1}$ 或 $\sigma_r/\sigma_{r+1} > 10^3$, 记数值矩阵 $M_t(y)$ 的秩为 r. 对于第 5.5 节数值实验中的例 5.6, 我们将误差界设为 10^{-4}. 关于矩阵 $M_{t-\ell}(y)$ 约化的行阶梯形式的计算, 我们也需要选择合适的误差界, 从而保证 $\ker M_{t-\ell}(y)$ 中没有信息丢失. 在数值实验中, 我们在每个例子中都给出了计算矩量矩阵的行阶梯形式所使用的误差界.

基于猜想 5.1, 在 δ-正则坐标系下, 我们证明了算法 5.1 的正确性和有限终止性, 即条件 (5.4)-(5.5) 一定在有限步半正定松弛内满足. 同时, 我们给出了实根理想 $I(V_{\mathbb{R}}(I))$ 关于序 $\prec_{\text{tdeg}}$ 的一组 Gröbner 基. 条件 (5.4)-(5.5) 可以作为 Flat Extension 定理 5.1 中条件 (5.1) 在正维情形下的推广. 与此同时, 给定半代数集 $\mathcal{S} = \{x \in \mathbb{R}^n \mid f_1(x) \geq 0, \ldots, f_s(x) \geq 0\}$, 我们将算法推广到求理想 I 的 $\mathcal{S}$-根理想 $I(V_{\mathbb{R}}(I) \cap \mathcal{S})$ 关于序 $\prec_{\text{tdeg}}$ 的一组 Gröbner 基.

5.2　预备知识

5.2.1　Hilbert 函数与代数簇的维数

将 $\mathbb{K}[x]_t$ 看作是域 $\mathbb{K}$ 上的一个向量空间, 其维数等于 $\binom{n+t}{t}$. 对任给的理想 $I \in \mathbb{K}[x]$, 令 $I_t = I \cap \mathbb{K}[x]_t$ 为 I 中次数小于等于 t 的多项式构成的集合.

定义 5.1. [30] 理想 I 的仿射 Hilbert 函数为以下关于非负整数 q 的多项式

$$HF_I^{\text{aff}}(q) = \dim \mathbb{K}[x_1, \ldots, x_n]_q - \dim I_q,$$

同时, 理想 I 的 Hilbert 函数定义为

$$HF_I(q) = HF_I^{\text{aff}}(q) - HF_I^{\text{aff}}(q-1).$$

当 q 足够大时, $HF_I^{\text{aff}}(q)\,(HF_I(q))$ 的值最终会满足某个多项式, 称为仿射 Hilbert 多项式 (Hilbert 多项式). 使得等式 $HP_I^{\text{aff}}(q) = HF_I^{\text{aff}}(q)$ 对所有 $q \geq q_0$ 均成立的最小的整数 q_0 称为理想 I 的正则指标 (Index of Regularity).

定义 5.2. 代数簇 $V \subseteq \mathbb{K}^n$ 的维数定义为理想 $I(V) \subseteq \mathbb{K}[x_1, \ldots, x_n]$ 的 Hilbert 多项式的次数, 记为 $\dim V$.

定义 5.3. 给定理想 I, 如果其相应的代数簇 $V(I)$ 是零维的, 即 $|V(I)| < \infty$, 称 I 为零维理想. 否则, 称理想 I 是正维理想.

对于零维理想 I, 其仿射 Hilbert 多项式 HP_I^{aff} 等于常数. 此时, 向量空间 $\mathbb{K}[x]/I$ 的维数 $\dim \mathbb{K}[x]/I = HP_I^{\text{aff}}$, 它与代数簇 $V(I)$ 的基数 $|V(I)|$ 的关系如下定理所述.

定理 5.3. *[30] 设 I 为 $\mathbb{K}[x_1, \ldots, x_n]$ 中的理想, 那么 $|V(I)| < \infty \Longleftrightarrow \dim \mathbb{K}[x]/I < \infty$. 而且, $|V(I)| \leq \dim \mathbb{K}[x]/I$, 其中等号成立当且仅当 I 为根理想.*

定理 5.4. *[30] 设 I 为 $\mathbb{K}[x_1, \ldots, x_n]$ 中的理想, 理想 I 与 $\sqrt{I}$ 的仿射 Hilbert 多项式具有相同的次数.*

给定理想 $I \in \mathbb{K}[x]$, 如果存在子集 $A \subseteq \mathbb{N}^n$ (可能无穷) 使得 I 包含所有形如 $\sum_{\alpha \in A} h_\alpha x^\alpha$ 的有限和组成的多项式, 其中 $h_\alpha \in \mathbb{K}[x]$, 那么称 I 为单项理想. 此时, 记 $I = \langle x^\alpha, \alpha \in A \rangle$. 事实上, Dickson 引理表明每个单项理想都是有限生成的, 见 [30].

定理 5.5. *[30] 设 I 为 $\mathbb{K}[x_1, \ldots, x_n]$ 中的理想, 在给定的分次序下, 单项理想 $\langle \mathrm{LT}(I) \rangle$ 与理想 I 的仿射 Hilbert 函数相同.*

理想的 Gröbner 基是代数几何中常用的概念之一, 它的定义依赖于单项序的选取.

定义 5.4. 给定 $\mathbb{T}^n$ 上的单项序 $\prec$, 如果有限多项式集合 $G = \{g_1, \ldots, g_s\}$ 满足

$$\langle \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_s) \rangle = \langle \mathrm{LT}(I) \rangle.$$

则称 G 为理想 $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ 的 Gröbner 基.

Hilbert 基定理确保任意理想均有 Gröbner 基, 并可通过算法构造出来, 例如 Buchberger 算法 [16]. 给定非零多项式 $f \in \mathbb{K}[x_1, \ldots, x_n]$, 利用多项式除法法则, f 除以理想 I 的 Gröbner 基中的元素 $g_1, \ldots, g_s$ 可得 $f = \sum_{j=1}^{s} u_j g_j + r$, 其中 $u_j, g_j, r \in \mathbb{K}[x_1, \ldots, x_n]$. 余式 r 是唯一确定的, 且 r 中的任何单项式都不能被 $\mathrm{LT}(g_j), j = 1, \ldots, s$ 整除. 在给定的分次序 (Graded Ordering) 下, $\deg(f) \geq \deg(u_j g_j)$, $j = 1, \ldots, s$.

5.2.2　矩量矩阵的相关性质

给定序列 $y = (y_\alpha)_{\alpha \in \mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$, 其相应的 (无穷维的) 矩量矩阵 $M(y) := (y_{\alpha+\beta})_{\alpha,\beta \in \mathbb{N}^n}$ 的核空间定义为如下多项式集合:

$$\ker M(y) := \{p \in \mathbb{R}[x] \mid M(y)\mathrm{vec}(p) = 0\}.$$

它是 $\mathbb{R}[x]$ 中的理想. 而且, 如果 $M(y) \succeq 0$, $\ker M(y)$ 为实根理想 (见 [31, 67, 78]). 类似的, 截断的矩量矩阵 $M_t(y)$ 的核空间定义为

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

这个集合只是 $\mathbb{R}[x]_t$ 的子集, 而不是理想. 但是在某些条件下, $\ker M_t(y)$ 也具有类似理想或实根理想的性质 (见 [31, 67, 78]):

引理 5.6. *[64]* 设 $M_t(y) \succeq 0$, 则

(i) 如果 $f, g \in \mathbb{R}[x]$, 且 $\deg(fg) \leq t - 1$, 那么 $f \in \ker M_t(y) \Longrightarrow fg \in \ker M_t(y)$.

(ii) 设 $p, q_j \in \mathbb{R}[x]$, $f := p^{2m} + \sum_j q_j^2 \in \mathbb{R}[x]_t$ 其中 $m \in \mathbb{N}$, $m \geq 1$. 那么 $f \in \ker M_t(y) \Longrightarrow p \in \ker M_t(y)$.

由半正定矩阵的基本性质可推出如下结果

$$M_t(y) \succeq 0 \Longrightarrow \ker M_t(y) \cap \mathbb{R}[x]_s = \ker M_s(y) \quad \text{for } 1 \leq s \leq t, \tag{5.7}$$

$$M_t(y), M_t(y') \succeq 0 \Longrightarrow \ker M_t(y + y') = \ker M_t(y) \cap \ker M_t(y'). \tag{5.8}$$

设理想 $I := \langle h_1, \ldots, h_m \rangle \subseteq \mathbb{R}[x]$, 定义

$$d := \max_{1 \leq j \leq m} d_j, \quad d_j := \lceil \deg(h_j)/2 \rceil, \quad j = 1, \ldots, m. \tag{5.9}$$

对于 $t \geq d$, 令

$$\mathcal{K}_t := \{y \in \mathbb{R}^{\mathbb{N}^n_{2t}} \mid y_0 = 1, M_t(y) \succeq 0, M_{t-d_j}(h_j y) = 0, j = 1, \ldots, m\}. \tag{5.10}$$

定义

$$\mathcal{K}_t^{gen} := \{y \in \mathcal{K}_t \mid M_t(y)\text{的秩最大}\}, \tag{5.11}$$

称集合 $\mathcal{K}_t^{gen}$ 中的元素为母元素. 下面介绍母元素的性质 (见 [64, 65, 109]).

引理 5.7. *[109]* 给定 $y \in \mathcal{K}_t$, 则下列叙述等价:

(i) $y \in \mathcal{K}_t^{gen}$.

(ii) $\mathrm{rank} M_t(y) = \max_{z \in \mathcal{K}_t} \mathrm{rank} M_t(z)$.

(iii) 对所有 $1 \leq s \leq t$, $\mathrm{rank} M_s(y) = \max_{z \in \mathcal{K}_t} \mathrm{rank} M_s(z)$.

(iv) 对所有 $z \in \mathcal{K}_t$, $\ker M_s(y) \subseteq \ker M_s(z)$ 对所有 $1 \leq s \leq t$ 均成立, 且有 $\ker M_s(y) \subseteq \sqrt[\mathbb{R}]{I}$.

由此引理得出, $\mathcal{K}_t$ 中所有母元素都具有相同的核空间.

引理 5.8. *[65]* 设 $t \leq t'$ 且 $y \in \mathcal{K}_t^{gen}$ 和 $y' \in \mathcal{K}_{t'}^{gen}$, 则

$$\ker M_t(y) \subseteq \ker M_{t'}(y'). \tag{5.12}$$

引理 5.9. *[64]* 设 $\{g_1, \ldots, g_k\}$ 为实根理想 $\sqrt[\mathbb{R}]{I}$ 的一组基, 存在 $t_0 \in \mathbb{N}$, 使得对所有 $t \geq t_0$, 都有 $g_1, \ldots, g_k \in \ker M_t(y)$, $y \in \mathcal{K}_t$.

定理 5.10. *[64]* 存在 $t_0 \in \mathbb{N}$ 使得对所有 $t \geq t_0$, $y \in \mathcal{K}_t^{gen}$, 都有 $\langle \ker M_t(y) \rangle = \sqrt[\mathbb{R}]{I}$.

给定半代数集合

$$\mathcal{S} = \{x \in \mathbb{R}^n \mid f_1(x) \geq 0, \ldots, f_s(x) \geq 0\},$$

其中 $f_1, \ldots, f_s \in \mathbb{R}[x]$. 为了计算 $\mathcal{S}$-根理想 $\sqrt[\mathcal{S}]{I}$, 我们考虑如下集合

$$\mathcal{K}_{t,\mathcal{S}} := \mathcal{K}_t \cap \left\{y \in \mathbb{R}^{\mathbb{N}^n_{2t}} \mid M_{t-d_{\underline{f}^e}}(\underline{f}^e y) \succeq 0 \text{ 对所有 } e \in \{0,1\}^s\right\}, \tag{5.13}$$

其中 $d_{\underline{f}^e} = \lceil \deg(\underline{f}^e)/2 \rceil$. 因为 $\sqrt[\mathcal{S}]{I}$ 是 $\mathcal{S}$-根理想, 与引理 5.9 和引理 5.10 相似, 我们给出如下结论.

引理 5.11. *设 $\{g_1,\dots,g_k\}$ 为理想 $\sqrt[s]{I}$ 的一组有限基, 存在 $t_0\in\mathbb{N}$ 使得对所有 $t\geq t_0$, 都有 $g_1,\dots,g_k\in\ker M_t(y)$, 其中 $y\in\mathcal{K}_{t,\mathcal{S}}$.*

证明. 对每个 $l\in\{1,\dots,k\}$, 由半代数零点定理 2.5, 存在 $m_l\in\mathbb{N}$ 和多项式 $\sigma_e\in\sum\mathbb{R}[x]^2$, $u_j(j\leq m)$ 满足

$$g_l^{2m_l}+\sum_{e\in\{0,1\}^k}\sigma_e\underline{f}^e=\sum_{j=1}^m u_jh_j.$$

令

$$t_0=1+\max(d,\deg(g_l^{2m_l}),\deg(\sigma_e\underline{f}^e),\deg(u_jh_j)).$$

对 $t\geq t_0$, 由于 $\deg(u_jh_j)\leq t-1$ 且 $h_j\in\ker M_t(y)$, 由引理 5.6 可知 $u_jh_j\in\ker M_t(y)$. 因此, $g_l^{2m_l}+\sum_{e\in\{0,1\}^k}\sigma_e\underline{f}^e\in\ker M_t(y)$ 且有

$$\operatorname{vec}(g_l^{m_l})^TM_t(y)\operatorname{vec}(g_l^{m_l})+\sum_{e\in\{0,1\}^k}\operatorname{vec}(\sigma_e)^TM_t(y)\operatorname{vec}(\underline{f}^e)=0,$$

由 $M_t(y)\succeq 0$, $f_i\geq 0, i=1,\dots,s$ 可以推出 $g_l^{m_l}\in\ker M_t(y)$. 如果 $m_l\in\mathbb{N}$ 为偶数, 那么 $g_l^{m_l}\in\ker M_t(y)\Longrightarrow g_l^{m_l/2}\in\ker M_t(y)$. 如果 m_l 为奇数, 由 $\deg(g_l^{m_l+1})\leq t-1$ 可知

$$g_l^{m_l}\in\ker M_t(y)\Longrightarrow g_l^{m_l+1}\in\ker M_t(y)\Longrightarrow g_l^{(m_l+1)/2}\in\ker M_t(y).$$

依次下去对 $m_l\geq 1$ 运用归纳法可得 $g_l\in\ker M_t(y)$. □

定理 5.12. *存在 $t_0\in\mathbb{N}$ 使得对所有 $t\geq t_0$ 和母元素 $y\in\mathcal{K}_{t,\mathcal{S}}^{gen}$, 都有 $\langle\ker M_t(y)\rangle=\sqrt[s]{I}$.*

证明. 设 $y\in\mathcal{K}_t^{gen}$, 任取点 $v\in V_{\mathbb{R}}(I)\cap\mathcal{S}$, 有 $[v]_{2t}:=(v^\alpha)_{\alpha\in\mathbb{N}^n_{2t}}\in\mathcal{K}_{t,\mathcal{S}}$ 且 $z=(y+[v]_{2t})/2\in\mathcal{K}_{t,\mathcal{S}}$. 显然有

$$\ker M_t((y+[v]_{2t})/2)=\ker M_t(y)\cap\ker M_t([v]_{2t}).$$

矩阵 $M_t(y)$ 的秩最大说明 $\ker M_t((y+[v]_{2t})/2)=\ker M_t(y)$ 且 $\ker M_t(y)\subseteq M_t([v]_{2t})$. 对所有 $p\in\ker M_t(y)$, 都有 $p\in M_t([v]_{2t})$ 且 $p(v)=0$. 由半代数零点定理 2.5 可得 $p\in\sqrt[s]{I}$, 从而有 $\ker M_t(y)\subseteq\sqrt[s]{I}$. 又由引理 5.11 可知 $\ker M_t(y)\supseteq\sqrt[s]{I}$. □

5.2.3 Cartan 指标与 Cartan 特征

本节介绍有关几何对合理论的一些基本概念.

定义 5.5. 给定关于变元 $x_1,\dots,x_n$ 的排序, 如果单项式 x^γ 的多元指标 $\gamma=(\gamma_1,\dots,\gamma_n)$ 左数第一个非零元为 γ_j, 则称单项式 x^γ 的类为 j.

例 5.1. 给定两个变元 $x_1\prec x_2$ 的二次单项式 x_1^2,x_1x_2,x_2^2, 那么 x_1^2,x_1x_2 的类为 1, x_2^2 的类为 2. 如果将变元的排序换为 $x_2\prec x_1$, 那么 x_1^2 的类为 2 而 x_1x_2,x_2^2 的类为 1.

定义 5.6. 给定单项序 $x_1\prec x_2\prec\cdots\prec x_n$, 多项式 $p\in\mathbb{R}[x]$ 的类定义为其首项 x^γ 的类. 如果 x^γ 的类为 j, 那么称 $x_1,\dots,x_j$ 为多项式 p 的乘子变元.

通常情况下, 我们选择分次反字典序 $(\prec_{\mathrm{tdeg}})$ 对单项式进行排序. 对于给定的变元序 $x_1\prec\cdots\prec x_n$, 分次反字典序的的定义如下: $x^\alpha\prec_{\mathrm{tdeg}}x^\beta\iff|\alpha|<|\beta|$ 或 $|\alpha|=|\beta|$ 且 $\alpha-\beta$ 左边第一个非零元大于零. 下面文献中矩量矩阵 $M_t(y)$ 的行和列所对应的指标 x^α,x^β 都使用分次反字典序来排列. 通过计算矩量矩阵 $M_t(y)$ 的约化的行阶梯形式, 可以将矩阵的列分为两部分: 主元部分和非主元部分.

为了检验条件 (5.5), 我们需要 Cartan 指标和 Cartan 特征的定义.

定义 5.7. 给定矩量矩阵 $M_t(y)$, 对 $j\in\{1,\dots,n\}$, Cartan 指标 $\beta_t^{(j)}$ 定义为矩阵 $M_t(y)$ 约化的行阶梯形式中类为 j 且次数等于 t 的主元列的个数. 相似的, 定义 Cartan 特征 $\alpha_t^{(j)}$ 为矩阵 $M_t(y)$ 约化的行阶梯形式中类为 j 且次数等于 t 的非主元列的个数.

引理 5.13. *n 元 t 次单项式中类为 j 的单项式的个数为*

$$N_t^{(j)}=\binom{n-j+t-1}{t-1}.$$

证明. 多元向量 $\gamma=(\gamma_1,\dots,\gamma_n)\in\mathbb{N}^n$ 中满足 $|\gamma|=t$ 的个数等于 $N_t=\binom{t+n-1}{n-1}$. 因为 x^γ 的类为 j, 那么有 $\gamma_1=\cdots=\gamma_{j-1}=0$ 且 $\gamma_j\neq0$. 因此 γ_j 可以取 1 到 t 中的任意整数, 而且剩余的 $\eta=(\gamma_{j+1},\dots,\gamma_n)\in\mathbb{N}^{n-j}$ 需满足 $|\eta|=t-\gamma_j$. 从而有

$$\sum_{\gamma_j=1}^{t}\begin{pmatrix}t-\gamma_j+n-j-1\\ n-j-1\end{pmatrix}=\begin{pmatrix}t-1+n-j-1\\ n-j-1\end{pmatrix}+\cdots+\begin{pmatrix}n-j-1\\ n-j-1\end{pmatrix}=\begin{pmatrix}t-1+n-j\\ n-j\end{pmatrix},$$

其中最后一个等号是根据 Fermat 组合恒等式. □

由引理 5.13 可知 Cartan 特征 $\alpha_t^{(j)}$ 与 Cartan 指标 $\beta_t^{(j)}$ 之间的关系满足：

$$\alpha_t^{(j)} + \beta_t^{(j)} = N_t^{(j)}. \tag{5.14}$$

由于 $\alpha_t^{(j)}$ 和 $\beta_t^{(j)}$ 依赖于变元的顺序, 也就是与坐标系相关. 坐标变换会导致上述两个值发生变化. 以后的讨论都需要在 δ-正则坐标系下进行.

定义 5.8. 在坐标系中, 如果矩量矩阵 $M_t(y)$ 中 $\sum_{j=1}^{n} j\alpha_t^{(j)}$ 取得最大值, 即不能通过坐标变化而达到更大的值, 那么称当前的坐标系为 δ-正则坐标系.

如 [117] 中所述, 每个坐标系都能通过坐标变换概率为 1 地转化为 δ-正则坐标系. 事实上, 我们只需将当前坐标向量乘以一个随机生成的对角线上元素为 1 的上三角矩阵, 即可得到δ-正则坐标系.

注意到, 矩量矩阵 $M_t(y)$ 的零空间为 $\mathbb{R}$ 上的向量空间. 相应的, 由多项式构成的向量空间 $\ker M_t(y)$ 有无穷多组基底. 然而, $M_t(y)$ 约化的行阶梯形式是唯一的, 它对应于 $M_t(y)$ 零空间唯一的一组约化基. 将这组约化基乘以单项式向量 $[x]_t$, 即可得到矩量矩阵核空间 $\ker M_t(y)$ 的一组基, 记为 B. B 中多项式的首项对应 $M_t(y)$ 约化的行阶梯形式中的非主元列. 如果 $p \in B$ 满足

(i) 首项 $\mathrm{LT}(p)$ 的系数等于 1;

(ii) p 中没有单项式属于集合 $\mathrm{LT}(B - \{p\})$.

那么称 B 为 $\ker M_t(y)$ 的一组约化基.

5.3　判定准则的验证

定理 5.2 中条件 (5.4) 需要检测相邻的 t 阶和 $t+1$ 阶矩量矩阵的秩. 而条件 (5.5) 涉及矩量矩阵的亏秩和 Cartan 特征的计算. 下面的定理及命题刻画了这两个条件的基本性质 (参见 Seiler [115–117]).

命题 5.14. *设 $t \geq 2d$, $1 \leq \ell \leq t - 2d$, $y_1 \in \mathcal{K}_t^{gen}$ 和 $y_2 \in \mathcal{K}_{t+1}^{gen}$ 满足条件 (5.5).*

(i) 令 $\ker M_{t-\ell}(y_1)$ 的约化基中次数等于 $t-\ell$ 的多项式为 $\{p_1, \ldots, p_s\}$, 类分别为 $j_1, \ldots, j_s$. 那么

$$\{x_1 p_1, \ldots, x_{j_1} p_1, \ldots, x_1 p_s, \ldots, x_{j_s} p_s\} \cup \ker M_{(t+1)-(\ell+1)}(y_2)$$

构成 $\ker M_{t+1-\ell}(y_2)$ 的一组基.

(ii) *若次数等于 $t-\ell$ 类为 i 的单项 x^γ 对应于矩阵 $M_{t-\ell}(y_1)$ 的非主元列, 则对所有 $j>i$, $x^{\gamma-e_i+e_j}$ 也对应于矩阵 $M_{t-\ell}(y_1)$ 的非主元列.*

证明. (i) 由于 $y_1\in\mathcal{K}_t^{gen}$, $y_2\in\mathcal{K}_{t+1}^{gen}$, 由引理 5.8 和 (5.7) 式可知

$$p_i\in\ker M_{t-\ell}(y_1)\subseteq\ker M_t(y_1)\subseteq\ker M_{t+1}(y_2),\quad i=1,\dots,s.$$

对 $\ell\geq 1$, $k=1,\dots,n$, 有 $\deg(x_kp_i)=t+1-\ell\leq t$. 根据引理 5.6(i) 和 (5.7) 式, 有

$$x_kp_i\in\ker M_{t+1}(y_2)\cap\mathbb{R}[x]_{t+1-\ell}=\ker M_{t+1-\ell}(y_2). \tag{5.15}$$

事实上, 由于多项式 $x_1p_1,\dots,x_{j_1}p_1,\dots,x_1p_s,\dots,x_{j_s}p_s$ 的首项次数等于 $t+1-\ell$ 且互不相同, 则多项式之间线性无关. 由于集合 $\{p_1,\dots,p_s\}$ 中有 $\alpha_{t-\ell}^{(j)}$ 个多项式的类为 j, 将所有类 j 多项式乘以其乘子变元, 即可得到 $\ker M_{t+1-\ell}(y_2)$ 中 $\sum_{j=1}^n j\alpha_{t-\ell}^{(j)}$ 个次数等于 $t+1-\ell$ 的多项式, 且线性无关. 另一方面, $\ker M_{t+1-\ell}(y_2)$ 中次数等于 $t+1-\ell$ 的线性无关多项式的个数等于 $\operatorname{corank}M_{(t+1)-\ell}(y_2)-\operatorname{corank}M_{(t+1)-(\ell+1)}(y_2)$. 因此, 由条件 (5.5) 成立推出 (i) 的结论正确.

(ii) 显然存在多项式 $p_i\in\{p_1,\dots,p_s\}$ 使得 $\mathrm{LT}(p_i)=x^\gamma$. 对所有 $j>i$, 因为 x_j 不是 p_i 的乘子变元, 所以 x_jp_i 的类仍为 i. 由于 $x_jp_i\in\ker M_{t+1-\ell}(y_2)$ 且 $\deg(x_jp_i)=t+1-\ell$, 由 (i) 可知, x_jp_i 可以表示成 $x_1p_1,\dots,x_{j_1}p_1,\dots,x_1p_s,\dots,x_{j_s}p_s$ 和 $\ker M_{(t+1)-(\ell+1)}(y_2)$ 中多项式的线性组合. 由于 $x_1p_1,\dots,x_{j_s}p_s$ 的首项互不相同且次数为 $t+1-\ell$, 则存在类大于或等于 i 的多项式 $p_k\in\{p_1,\dots,p_s\}$ 使得

$$\mathrm{LT}(x_ip_k)=\mathrm{LT}(x_jp_i)=x^{\gamma+e_j}.$$

因此, p_k 的首项可以表示为 $x^{\gamma-e_i+e_j}$, 对应于 $M_{t-\ell}(y_1)$ 约化的行阶梯形式中的一个非主元列. □

定理 5.15. *设 $y_1\in\mathcal{K}_t^{gen}$, $y_2\in\mathcal{K}_{t+1}^{gen}$, $y_3\in\mathcal{K}_{t+2}^{gen}$.*

(i) *若 (5.5) 式对于 (y_1,y_2,t,ℓ) 成立, 则 (5.5) 对于 $(y_2,y_3,t+1,\ell)$ 也成立, 即*

$$\sum_{j=1}^n j\alpha_{t+1-\ell}^{(j)}\text{ 对于 }M_{t+1-\ell}(y_2)=\operatorname{corank}M_{t+2-\ell}(y_3)-\operatorname{corank}M_{(t+2)-(\ell+1)}(y_3). \tag{5.16}$$

(ii) 若 *(5.4)*-*(5.5)* 对于 (y_1, y_2, t, ℓ) 成立, 则 *(5.4)* 对于 $(y_2, y_3, t+1, \ell)$ 也成立, 即

$$\operatorname{rank} M_{t+1-\ell}(y_2) = \operatorname{rank} M_{(t+2)-(\ell+1)}(y_3). \tag{5.17}$$

证明. 假设 $\ker M_{t-\ell}(y_1)$ 的约化的基中次数为 $t-\ell$ 的多项式为 $\{p_1, \ldots, p_s\}$, 且类分别为 $j_1, \ldots, j_s$. 由于条件 (5.5) 对 (y_1, y_2, t, ℓ) 成立, 由命题 5.14(i) 可知 $\{x_1p_1, \ldots, x_{j_1}p_1, \ldots, x_1p_s, \ldots, x_{j_s}p_s\} \cup \ker M_{(t+1)-(\ell+1)}(y_2)$ 为 $\ker M_{t+1-\ell}(y_2)$ 的一组基. 要证 (i), 需证 $\{x_1p_1, \ldots, x_{j_1}p_1, \ldots, x_1p_s, \ldots, x_{j_s}p_s\}$ 乘以其乘子变元所得的次数为 $t+2-\ell$ 的多项式和 $\ker M_{(t+2)-(\ell+1)}(y_3)$ 中多项式 (次数小于等于 $t+1-\ell$) 构成 $\ker M_{t+2-\ell}(y_3)$ 的一组基.

不妨设 $p \in \{p_1, \ldots, p_s\}$ 的类为 k, 首项 $\mathrm{LT}(p) = x^\gamma$. 对于 $i \le k$, 由 (5.15) 式可知 $x_ip \in \ker M_{t+1-\ell}(y_2)$, 且类为 i. 用 x_ip 乘以所有变元 $x_1, \ldots, x_n$ 所得多项式的首项为 $x^{\gamma+e_i+e_j}, j = 1, \ldots, n$. 由 (5.15) 式可知 $x_jx_ip \in \ker M_{t+2-\ell}(y_3)$. 下面对 x_jx_ip 分三种情况讨论:

$j \le i$: x_j 为 x_ip 的乘子变元.

$i < j \le k$: 由于 $j \le k$, 那么 x_j 为 $p \in \ker M_{t-\ell}(y_1)$ 的乘子变元. 将乘积 x_jx_ip 看做是由 p 先乘以 x_j 再乘以 x_i 而得. 因此, 每一步乘法都乘以乘子变元. 因此, 我们可以将其归为第一种情况.

$k < j$: 如第二种情况所述, 改变 x_i 和 x_j 的乘法顺序, 得 $t+2-\ell$ 次多项式 $x_i(x_jp) \in \ker M_{t+2-\ell}(y_3)$. 由于 $j > k$, 由命题 5.14 (ii) 可知单项 $x^{\gamma-e_k+e_j}$ 也对应于 $M_{t-\ell}(y_1)$ 的非主元列, 即存在 $q \in \ker M_{t-\ell}(y_1)$ 使得 $\mathrm{LT}(q) = x^{\gamma-e_k+e_j}$, 且 $x^{\gamma-e_k+e_j}$ 的类大于或等于 k. 则 x_k 为 q 的乘子变元. 因此, $\mathrm{LT}(x_kq) = x^{\gamma+e_j}$ 的类为 k 且对应于 $M_{t+1-\ell}(y_2)$ 的非主元列. 由于 x_i 为 x_kq 的乘子变元, 则多项式 x_ix_kq 的次数为 $t+2-\ell$, 首项为 $x^{\gamma+e_j+e_i}$, 包含在第一种情况中.

设 $\mathrm{LT}(x_ix_jp) = \mathrm{LT}(x_ix_kq) = x^{\gamma+e_j+e_i}$, 且二者均由首项等于 $x^{\gamma+e_j}$ 的多项式 x_jp 和 x_kq 乘以 x_i 而得. 不妨设 x_jp 是由 p 乘以非乘子变元 x_j 而得. 由命题 5.14 (i) 可知, $x_jp \notin \{x_1p_1, \ldots, x_{j_1}p_1, \ldots, x_1p_s, \ldots, x_{j_s}p_s\}$. 由于 $x_jp \in \ker M_{t+1-\ell}(y_2)$, 则可以表示为 $\{x_1p_1, \ldots, x_{j_1}p_1, \ldots, x_1p_s, \ldots, x_{j_s}p_s\}$ 以及 $\mathbb{R}[x]_{t-\ell}$ 中多项式的线性组合. 则 x_ix_jp 可以表示为上述线性组合与 x_i 的乘积. 如前所述, x_ix_jp 的首项可以由 x_ix_kq 约化掉, 且 x_i 为

$x_k q$ 的乘子变元. 按照此种方法依次考虑线性组合中的每一个单项, 最终可以将 $t+2-\ell$ 次多项式 $x_i x_j p$ 表示为所有满足第一种情形的多项式和 $\mathbb{R}[x]_{t+1-\ell}$ 中低次多项式的线性组合.

因此, $\ker M_{t+2-\ell}(y_3)$ 中所有次数等于 $t+2-\ell$ 的多项式都可以表示为 $\{x_1 p_1, \ldots, x_{j_1} p_1, \ldots, x_1 p_s, \ldots, x_{j_s} p_s\}$ 乘以其乘子变元所得的次数为 $t+2-\ell$ 的多项式和 $\ker M_{(t+2)-(\ell+1)}(y_3)$ 中多项式 (次数小于等于 $t+1-\ell$) 的线性组合. 也就是条件 (5.16) 成立.

(ii) 由于条件 (5.4)-(5.5) 对于 (y_1, y_2, t, ℓ) 成立, 由猜想 5.1 可知, 存在 $\tilde{y} \in \mathcal{K}_{t+2}$ 满足

$$\ker M_{t-\ell}(y_1) = \ker M_{(t+1)-(\ell+1)}(y_2) = \ker M_{(t+2)-(\ell+2)}(\tilde{y}). \tag{5.18}$$

由于 $y_3 \in \mathcal{K}_{t+2}^{gen}$, 由引理 5.7 (iii) 可知

$$\ker M_{(t+2)-(\ell+2)}(y_3) \subseteq \ker M_{(t+2)-(\ell+2)}(\tilde{y}) \overset{(5.18)}{=} \ker M_{(t+1)-(\ell+1)}(y_2). \tag{5.19}$$

又由于 $y_3|_{2t+2} \in \mathcal{K}_{t+1}$, 且 $y_2 \in \mathcal{K}_{t+1}^{gen}$, 由引理 5.7 (iii) 可知

$$\ker M_{(t+1)-(\ell+1)}(y_2) \subseteq \ker M_{(t+2)-(\ell+2)}(y_3).$$

从而有 $\ker M_{(t+1)-(\ell+1)}(y_2) = \ker M_{(t+2)-(\ell+2)}(y_3)$.

因此, 要证 (5.17) 只需考虑 $\ker M_{t+1-\ell}(y_2)$ 和 $\ker M_{(t+2)-(\ell+1)}(y_3)$ 中次数等于 $t+1-\ell$ 的部分. 由于条件 (5.5) 对于 (y_1, y_2, t, ℓ) 成立, 根据命题 5.14 (i) 和本定理 (i) 的结果可知 $\ker M_{t+2-\ell}(y_3)$ 中所有次数等于 $t+2-\ell$ 的多项式都可以表示为 $\{x_1 p_1, \ldots, x_{j_1} p_1, \ldots, x_1 p_s, \ldots, x_{j_s} p_s\}$ 分别乘以乘子变元得到的次数为 $t+2-\ell$ 的多项式和 $\ker M_{(t+2)-(\ell+1)}(y_3)$ 中多项式的线性组合. 对于 $y_2 \in \mathcal{K}_{t+1}^{gen}$, $y_3 \in \mathcal{K}_{t+2}^{gen}$, 由引理 5.7 (iii) 可知

$$\ker M_{(t+1)-\ell}(y_2) \subseteq \ker M_{(t+2)-\ell}(y_3) \cap \mathbb{R}[x]_{(t+1)-\ell} = \ker M_{(t+2)-(\ell+1)}(y_3).$$

假设存在 $t+1-\ell$ 次多项式 $q \in \ker M_{(t+2)-(\ell+1)}(y_3)$, 但是

$$q \notin \ker M_{(t+1)-\ell}(y_2).$$

那么 $\mathrm{LT}(q)$ 一定不同于 $\{\mathrm{LT}(x_1p_1),\ldots,\mathrm{LT}(x_{j_1}p_1),\ldots,\mathrm{LT}(x_1p_s),\ldots,\mathrm{LT}(x_{j_s}p_s)\}$. 设 x_i 是 q 的乘子变元, 那么 $\mathrm{LT}(x_iq)$ 一定与 $\{x_1p_1,\ldots,x_{j_1}p_1,\ldots,x_1p_s,\ldots,x_{j_s}p_s\}$ 分别乘以乘子变元所得的多项式的首项互不相同. 另一方面, 由 (5.15) 式知

$$x_iq \in \ker M_{t+2}(y_3) \cap \mathbb{R}_{t+2-\ell} \in \ker M_{t+2-\ell}(y_3).$$

导出矛盾. 因此, $\ker M_{t+1-\ell}(y_2) = \ker M_{(t+2)-(\ell+1)}(y_3)$, 即 (5.17) 成立. □

根据 [115, 定理 2.18] 和 [117, 定理 6.1.21], 我们给出如下定理保证条件 (5.5) 成立.

定理 5.16. *设 $I = \langle h_1,\ldots,h_m\rangle$ 为 $\mathbb{R}[x]$ 中的理想, $\mathcal{K}_t^{gen}$ 如 (5.11) 所定义. 任给 $\ell \geq 1$, 都存在整数 $t \geq 2d+\ell$ 使得对所有 $y_1 \in \mathcal{K}_t^{gen}$, $y_2 \in \mathcal{K}_{t+1}^{gen}$, 下式均成立*

$$\sum_{j=1}^{n} j\alpha_{t-\ell}^{(j)} \text{ 对于 } M_{t-\ell}(y_1) = \mathrm{corank}\, M_{t+1-\ell}(y_2) - \mathrm{corank}\, M_{(t+1)-(\ell+1)}(y_2).$$

该定理的证明参见 [117, 127].

定理 5.17. *设 $I = \langle h_1,\ldots,h_m\rangle$ 为 $\mathbb{R}[x]$ 中的理想. 存在整数 $t \geq 2d$, $1 \leq \ell \leq t-2d$, 使得 (5.4)-(5.5) 对所有 $y_1 \in \mathcal{K}_t^{gen}$, $y_2 \in \mathcal{K}_{t+1}^{gen}$ 均成立.*

此定理来自于 Cartan-Kähler 理论 [58] 中相似的定理. 其证明可参见 [96]. 我们给出算法 5.1 中第 2 步的具体计算方法, 并证明其正确性和有限终止性.

证明. 对于 $\ell = 1$, 由定理 5.16 可知存在 $t_1 \geq 2d$ 使得矩阵 $M_{t_1}(y_1)$ 满足条件 (5.5). 如果条件 (5.4) 不成立, 则有 $\mathrm{rank}M_{t_1-1}(y_1) > \mathrm{rank}M_{(t_1+1)-2}(y)$, 其中 $y_1 \in \mathcal{K}_{t_1}^{gen}, y \in \mathcal{K}_{t_1+1}^{gen}$. 令 $\ell = 2$, 寻找 $t_2 \geq t_1+1$ 使得矩阵 $M_{t_2-2}(y_2)$ 满足条件 (5.5). 我们得到如下的包含关系

$$\ker M_{t_1-1}(y_1) \subsetneq \ker M_{(t_1+1)-2}(y) \subseteq \ker M_{t_2-2}(y_2).$$

若条件 (5.4) 仍然不成立, 按照上述方法进行下去, 得到一列矩量矩阵 $M_{t_1-1}(y_1)$, $M_{t_2-2}(y_2),\ldots,M_{t_i-i}(y_i)$, 其中 $2d \leq t_1 < t_2 < \cdots < t_i$. 从而得到理想的升链

$$\langle \ker M_{t_1-1}(y_1)\rangle \subsetneq \langle \ker M_{t_2-2}(y_2)\rangle \subsetneq \cdots \subsetneq \langle \ker M_{t_i-i}(y_i)\rangle \subsetneq \cdots,$$

因为 $\mathbb{R}[x_1,\ldots,x_n]$ 是 Noether 环, 以上理想的升链最终会达到稳定, 即存在 $k>1$ 使得

$$\langle \ker M_{t_{k-1}-(k-1)}(y_k)\rangle \subsetneq \langle \ker M_{t_k-k}(y_k)\rangle = \langle \ker M_{t_{k+1}-(k+1)}(y_{k+1})\rangle = \cdots .$$

因此, 条件 (5.4) 对于 $i \geq k$ 总成立. 与此同时, 条件 (5.5) 也成立. □

下面我们讨论矩量矩阵的秩与实根理想 $\sqrt[\mathbb{R}]{I}$ 的仿射 Hilbert 函数之间的关系.

命题 5.18. *设 $I=\langle h_1,\ldots,h_m\rangle$ 为 $\mathbb{R}[x]$ 中的理想. 令 $t \geq 2d$. 如果存在 $1\leq \ell \leq t-2d$ 使得对所有 $k\geq 0$, $y\in \mathcal{K}_{t+k}^{gen}$, $y'\in \mathcal{K}_{t+k+1}^{gen}$ 都有*

$$\mathrm{rank} M_{t+k-\ell}(y) = \mathrm{rank} M_{(t+k+1)-(\ell+1)}(y'). \tag{5.20}$$

那么

$$\mathrm{rank} M_{t+k-\ell}(y) = HF^{aff}_{\sqrt[\mathbb{R}]{I}}(t+k-\ell). \tag{5.21}$$

证明. 由于 $y\in \mathcal{K}_{t+k}^{gen}$, 由引理 5.7 得 $\ker M_{t+k-\ell}(y) \subseteq \sqrt[\mathbb{R}]{I}\cap \mathbb{R}[x]_{t+k-\ell}$. 由定义 5.1 知

$$\mathrm{rank} M_{t+k-\ell}(y) \geq HF^{\mathrm{aff}}_{\sqrt[\mathbb{R}]{I}}(t+k-\ell).$$

如果上式中的等号不成立, 则有 $\ker M_{t+k-\ell}(y) \subsetneq \sqrt[\mathbb{R}]{I}\cap \mathbb{R}[x]_{t+k-\ell}$. 那么一定存在整数 $k'\geq k$, $y_1\in \mathcal{K}_{t+k'}^{gen}$, $y_2\in \mathcal{K}_{t+k'+1}^{gen}$, 使得

$$\ker M_{t+k'-\ell}(y_1)\cap \mathbb{R}[x]_{t+k-\ell} \subsetneq \ker M_{t+k'+1-\ell}(y_2)\cap \mathbb{R}[x]_{t+k-\ell}. \tag{5.22}$$

从而有

$$\ker M_{t+k'-\ell}(y_1) \subsetneq \ker M_{t+k'+1-\ell}(y_2)\cap \mathbb{R}[x]_{t+k'-\ell} = \ker M_{(t+k'+1)-(\ell+1)}(y_2).$$

因此, $\mathrm{rank} M_{t+k'-\ell}(y_1) > \mathrm{rank} M_{(t+k'+1)-(\ell+1)}(y_2)$. 与条件 (5.20) 矛盾. □

下面的引理说明, 如果条件 (5.5) 满足, 那么我们不需要通过计算次数等于 $t+k-\ell$ 的非主元列的个数来求高阶矩量矩阵 $M_{t+k-\ell}(y)$ 的 Cartan 特征 $\alpha_{t+k-\ell}$, 而是由如下递归公式得到.

引理 5.19. *[115]* 设 *(5.5)* 对 $t \geq 2d$, $1 \leq \ell \leq t-2d$ 成立. 那么, 对所有 $k \geq 0$, $y \in \mathcal{K}_{t+k}^{gen}$, 矩量矩阵 $M_{t+k-\ell}(y)$ 的 *Cartan* 指标和 *Cartan* 特征为

$$\alpha_{t+k-\ell}^{(j)} = \sum_{i=j}^{n} \binom{k+i-j-1}{k-1} \alpha_{t-\ell}^{(i)}, \quad 1 \leq j \leq n, \tag{5.23}$$

$$\beta_{t+k-\ell}^{(j)} = \sum_{i=j}^{n} \binom{k+i-j-1}{k-1} \beta_{t-\ell}^{(i)}, \quad 1 \leq j \leq n. \tag{5.24}$$

证明. 此引理的证明与 [115, 引理 3.6] 相似. 如果 (5.23) 成立, 由 (5.14) 式可知, (5.24) 式也成立. 下面利用数学归纳法来证明 (5.23) 式正确. 设 p 为矩阵 $\ker M_{t-\ell}(y_1)$ 的约化基中类大于或等于 j 次数等于 $t-\ell$ 的多项式, 那么 $x_j p$ 的类等于 j. 由命题 5.14 (i) 知所有上述 $x_j p$ 属于 $\ker M_{t+1-\ell}(y_2)$ 的一组基. 因此, $\alpha_{t+1-\ell}^{(j)} = \alpha_{t-\ell}^{(j)} + \cdots + \alpha_{t-\ell}^{(n)}$, 即 (5.23) 式对于 $j=1$ 成立.

假设 (5.23) 对 $j-1$ 成立. 由定理 5.15 (i) 可知, (5.5) 对所有 $k \geq 1$ 均成立. 利用 $j=1$ 时的分析可知 $\alpha_{(t+k)-\ell}^{(j)} = \alpha_{(t+k-1)-\ell}^{(j)} + \cdots + \alpha_{(t+k-1)-\ell}^{(n)}$. 依次进行, 可得

$$\begin{aligned} \alpha_{(t+k)-\ell}^{(j)} &= \sum_{\ell=j}^{n} \sum_{i=\ell}^{n} \begin{pmatrix} k+i-\ell-2 \\ k-2 \end{pmatrix} \alpha_{t-\ell}^{(\ell)} \\ &= \sum_{i=j}^{n} \begin{pmatrix} k+i-j-1 \\ k-1 \end{pmatrix} \alpha_{t-\ell}^{(i)}. \end{aligned}$$

□

定理 5.20. 如果条件 *(5.4)-(5.5)* 对 (y_1, y_2, t, ℓ) 成立, 那么对所有 $k \geq 0$

$$HP_{\mathbb{R}[\sqrt{I}]}(t+k-\ell) = \sum_{j=1}^{n} \binom{k+j-1}{k} \beta_{t-\ell}^{(j)}. \tag{5.25}$$

与此同时, 对所有 $k \geq 0$, $y \in \mathcal{K}_{t+k}^{gen}$ 均满足

$$\operatorname{rank} M_{t+k-\ell}(y) = HP_{\mathbb{R}[\sqrt{I}]}^{aff}(t+k-\ell). \tag{5.26}$$

证明. 对 $t, t+1, t+2, \ldots$ 依次利用定理 5.15 和命题 5.18 可知, 对 $k \geq 0$, $y \in \mathcal{K}_{t+k}^{gen}$ 均满足

$$\operatorname{rank} M_{t+k-\ell}(y) = HF_{\mathbb{R}[\sqrt{I}]}^{\mathrm{aff}}(t+k-\ell).$$

对于 $y' \in \mathcal{K}_{t+k-1}^{gen}$, 总有 $\ker M_{(t+k-1)-\ell}(y') \subseteq \ker M_{(t+k)-\ell}(y)$. 由于矩量矩阵的秩满足仿射 Hilbert 函数, 那么 $\ker M_{t+k-\ell}(y)$ 的约化基中不属于集合 $\ker M_{(t+k-1)-\ell}(y')$ 的多项式的次数一定等于 $t+1-\ell$.

对于 $1 \le j \le n$, 由引理 5.19 知, 矩阵 $M_{t+k-\ell}(y)$ 的 Cartan 指标 $\beta_{t+k-\ell}^{(j)}$ 满足

$$\beta_{t+k-\ell}^{(j)} = \sum_{i=j}^{n} \binom{k+i-j-1}{k-1} \beta_{t-\ell}^{(i)}.$$

因此, 矩阵 $M_{t+k-\ell}(y)$ 的类 $1 \le j \le n$ 的 Cartan 指标 $\beta_{t+k-\ell}^{j}$ 的和为

$$\begin{aligned}\beta_{t+k-\ell} = \sum_{j=1}^{n} \beta_{t+k-\ell}^{(j)} &= \sum_{j=1}^{n}\sum_{i=j}^{n} \binom{k+i-j-1}{k-1} \beta_{t-\ell}^{(i)} \\ &= \sum_{j=1}^{n} \binom{k+j-1}{k} \beta_{t-\ell}^{(j)}.\end{aligned}$$

因为 $\beta_{t+k-\ell} = \mathrm{rank} M_{t+k-\ell}(y) - \mathrm{rank} M_{(t+k-1)-\ell}(y')$ 给出了仿射 Hilbert 函数的变化. 对于固定的 t, ℓ 和 n, $\beta_{t+k-\ell}$ 可以看做是关于变元 k 的多项式. 因此,

$$\beta_{t+k-\ell} = HP_{\mathbb{R}\sqrt{I}}(t+k-\ell) = \sum_{j=1}^{n} \binom{k+j-1}{k} \beta_{t-\ell}^{(j)}.$$

与此同时, (5.26) 也成立. □

下面我们给出本节的主要结果, 可以作为通过求解一系列半定规划问题 (5.3) 来计算实根理想 $\mathbb{R}\sqrt{I}$ 的 Gröbner 基算法终止的判定定理.

定理 5.21. *设 $I = \langle h_1, \ldots, h_m \rangle$ 为 $\mathbb{R}[x]$ 中的理想. $t \ge 2d$, $1 \le \ell \le t-2d$, $y_1 \in \mathcal{K}_t^{gen}$, $y_2 \in \mathcal{K}_{t+1}^{gen}$, 如果条件 (5.4)-(5.5) 对于 (y_1, y_2, t, ℓ) 成立. 那么 $\ker M_{t-\ell}(y_1)$ 是实根理想 $\mathbb{R}\sqrt{I}$ 关于序 $\prec_{\mathrm{tdeg}}$ 的一组 Gröbner 基, 即*

$$\langle \ker M_{t-\ell}(y_1) \rangle = \mathbb{R}\sqrt{I}. \tag{5.27}$$

证明. 对所有 $k \ge 0$ 和 $y \in \mathcal{K}_{t+k}^{gen}$, 由定理 5.7 (iv) 可知,

$$\ker M_{t+k-\ell}(y) \subseteq \mathbb{R}\sqrt{I} \cap \mathbb{R}[x]_{t+k-\ell}.$$

由于条件 (5.4)-(5.5) 对 (y_1, y_2, t, ℓ) 成立, 根据定理 5.20 和仿射 Hilbert 多项式的定义可知, 对所有 $k \ge 0$,

$$\mathrm{rank}\, M_{t+k-\ell}(y) = HP_{\mathbb{R}\sqrt{I}}^{\mathrm{aff}}(t+k-\ell) = \dim \mathbb{R}[x]_{t+k-\ell} - \dim\, \mathbb{R}\sqrt{I}_{\,t+k-\ell}.$$

从而, 向量空间 $\ker M_{t+k-\ell}(y) = \sqrt[\mathbb{R}]{I}_{t+k-\ell}$. 设 B_k 为 $\ker M_{t+k-\ell}(y)$ 的约化基. 那么 B_k 为向量空间 $\sqrt[\mathbb{R}]{I}_{t+k-\ell}$ 的一组基. 由定理 5.5 可知, 向量空间 $\langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle_{t+k-\ell}$ 和 $\sqrt[\mathbb{R}]{I}_{t+k-\ell}$ 具有相同的维数, 则 $\mathrm{LT}(B_k)$ 是 $\langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle_{t+k-\ell}$ 的一组基, 从而有

$$\langle \mathrm{LT}(B_k)\rangle = \big\langle \langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle_{t+k-\ell}\big\rangle. \tag{5.28}$$

由于条件 (5.4)-(5.5) 对 (y_1, y_2, t, ℓ) 成立, 由命题 5.14 (i) 可知 $\ker M_{t+1-\ell}(y_2)$ 的约化基 B_1 中次数等于 $t+1-\ell$ 的多项式是由 $\ker M_{t-\ell}(y_1)$ 的约化基 B_0 中次数等于 $t-\ell$ 的多项式乘以其乘子变元得到的. 因此, $\langle \mathrm{LT}(B_0)\rangle = \langle \mathrm{LT}(B_1)\rangle$. 又由定理 5.15 知, 条件 (5.4)-(5.5) 对 $(t+k, \ell)$ 均成立, 因此, 对所有 $k \geq 0$,

$$\langle \mathrm{LT}(B_0)\rangle = \langle \mathrm{LT}(B_1)\rangle = \cdots = \langle \mathrm{LT}(B_k)\rangle = \cdots. \tag{5.29}$$

由 (5.28) 和 (5.29) 式可以推出

$$\big\langle \langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle_{t-\ell}\big\rangle = \big\langle \langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle_{t+1-\ell}\big\rangle = \cdots = \big\langle \langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle_{t+k-\ell}\big\rangle = \cdots.$$

注意到, 对足够大的 k, 总有 $\big\langle \langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle_{t+k-\ell}\big\rangle = \langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle$. 由此可知,

$$\langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle = \langle \mathrm{LT}(B_0)\rangle \subseteq \langle \mathrm{LT}(\ker M_{t-\ell}(y_1))\rangle \subseteq \langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle.$$

因此, $\langle \mathrm{LT}(\ker M_{t-\ell}(y_1))\rangle = \langle \mathrm{LT}(\sqrt[\mathbb{R}]{I})\rangle$. 由定义 5.4 知, $\ker M_{t-\ell}(y_1)$ 是实根理想 $\sqrt[\mathbb{R}]{I}$ 关于序 $\prec_{\mathrm{tdeg}}$ 的一组 Gröbner 基. □

5.4　计算 $\mathcal{S}$-根理想 $I(V_{\mathbb{R}}(I)\cap\mathcal{S})$

上节中所有命题和定理只与矩量矩阵的秩和核空间有关. 这些结果均可以用于计算 $\mathcal{S}$-根理想 $I(V_{\mathbb{R}}(I)\cap\mathcal{S})$, 其中 $\mathcal{S} = \{x \in \mathbb{R}^n \mid f_1(x) \geq 0, \ldots, f_s(x) \geq 0\}$ 为半代数集.

回顾 (5.13), 考虑子集 $\mathcal{K}_{t,\mathcal{S}} \subseteq \mathcal{K}_t$,

$$\mathcal{K}_{t,\mathcal{S}} := \mathcal{K}_t \cap \Big\{ y \in \mathbb{R}^{\mathbb{N}^n_{2t}} \mid M_{t-d_{\underline{f}^e}}(\underline{f}^e y) \succeq 0 \text{ 对所有 } e \in \{0,1\}^s \Big\},$$

其中 $d_{\underline{f}^e} = \lceil \deg(\underline{f}^e)/2 \rceil$ 且将 d 的定义改为

$$d := \max_{1\leq j\leq m, e\in\{0,1\}^s} \{d_j,\ d_{\underline{f}^e}\}, \tag{5.30}$$

当 t 足够大时, 由引理 5.11 和定理 5.12 知, $\mathcal{S}$-根理想 $\sqrt[\mathcal{S}]{I}$ 中的信息都包含于母元素集合

$$\mathcal{K}_{t,\mathcal{S}}^{gen} := \{y \in \mathcal{K}_{t,\mathcal{S}} \mid \mathrm{rank} M_t(y) \text{ 的秩最大}\}.$$

因此, 上一节中的定理和命题对 $y \in \mathcal{K}_{t,\mathcal{S}}^{gen}$ 均成立.

下面的定理可以看做是定理 5.21 关于半代数集 $\mathcal{S}$ 的推广.

定理 5.22. *设 $I = \langle h_1, \ldots, h_m\rangle$ 为 $\mathbb{R}[x]$ 中的理想. 对于 $t \geq 2d$, $1 \leq \ell \leq t-2d$, $y_1 \in \mathcal{K}_{t,\mathcal{S}}^{gen}$, $y_2 \in \mathcal{K}_{t+1,\mathcal{S}}^{gen}$, 如果条件 (5.4)-(5.5) 对于 (y_1, y_2, t, ℓ) 成立. 那么 $\ker M_{t-\ell}(y_1)$ 是 $\mathcal{S}$-根理想 $\sqrt[\mathcal{S}]{I}$ 关于序 $\prec_{\mathrm{tdeg}}$ 的一组 Gröbner 基. 与此同时*

$$\langle \ker M_{t-\ell}(y_1)\rangle = \sqrt[\mathcal{S}]{I}. \tag{5.31}$$

注 5. 对于 $\mathcal{S}$-根理想 $\sqrt[\mathcal{S}]{I}$ 的 Gröbner 基的计算, 只需将定义半代数集 $\mathcal{S}$ 的多项式 $\{f_1, \ldots, f_s\}$ 加入到算法 5.1 的输入中. 同时在半定规划问题 (5.3) 中加入限制条件 $M_{t-d_{\underline{f}^e}}(\underline{f}^e y) \succeq 0$, 对所有 $e \in \{0,1\}^s$.

5.5 数值实验

本节中, 给出算法 5.1 在计算实根理想 $I(V_{\mathbb{R}}(I))$ 的 Gröbner 基的问题上的数值表现. 下面的例子中, 均使用分次反字典序 $\prec_{\mathrm{tdeg}}$.

例 5.2. 考虑多项式系统 $P = \{p_1, p_2\}$ [114, p.20, Ex 1.4.6], 其中

$$p_1 = x_1^2 - x_2,$$
$$p_2 = x_1 x_2 - x_3,$$

$d = 1$. 对于 $t \geq 2$, $0 \leq \ell \leq t-2$, 矩量矩阵 $M_{t-\ell}(y)$ 的秩如表 5.1 所示.

由上表可以看出 $\mathrm{rank} M_{3-1} = \mathrm{rank} M_{4-2} = 7$. 因此, 对于 $t = 3$, $\ell = 1$ 条件 (5.4) 满足.

下面检测条件 (5.5) 是否对 $t = 3$, $\ell = 1$ 也满足. 3 个变元 3 阶矩量矩阵的维数为 20×20, 其中 5 到 10 列对应次数等于 2 的单项. 选择变元序 $x_3 \prec_{\mathrm{tdeg}} x_1 \prec_{\mathrm{tdeg}} x_2$ 和误差界 10^{-8}, M_3 的主元出现在第 $1, 2, 3, 4, 5, 6, 7, 11, 12, 13, 16, 20$ 列中, 其中,

$$5, 6, 7,$$

阶	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$
t=2	8	4	1			
t=3	12	7	4	1		
t=4	16	10	7	4	1	
t=5	20	13	10	7	4	1

表 5.1: 矩阵 $M_{t-\ell}(y)$ 的秩

列对应次数为 2 的单项. 那么次数等于 2 的非主元列对应的单项为

$$x_1^2,\ x_1x_2, x_2^2.$$

类分别为 $2,2,3$. 因此, 对于 M_{3-1}, 有

$$\sum_{j=1}^{3} j\alpha_{3-1}^{(j)} = 3\times 1 + 2\times 2 = 7.$$

与此同时,

$$\mathrm{corank}M_{4-1} - \mathrm{corank}M_{4-2} = (20-10)-(10-7) = 7.$$

因此, 条件 (5.4)-(5.5) 对于 $t=3$, $\ell=1$ 均满足.

因此, $\ker M_{3-1}$ 的约化基

$$\{-x_2+x_1^2,\ -x_3+x_1x_2,\ -x_3x_1+x_2^2\}$$

是实根理想 $\sqrt[\mathbb{R}]{I}$ 关于序 $\prec_{\mathrm{tdeg}}$ 的一组 Gröbner 基.

例 5.3. 考虑 2-维理想 $I=\langle p_1,p_2,p_3\rangle$ [124, p.397, Eq. (9.60)], 其中

$$\begin{aligned} p_1 &= x_1^2+x_1x_2-x_1x_3-x_1-x_2+x_3,\\ p_2 &= x_1x_2+x_2^2-x_2x_3-x_1-x_2+x_3,\\ p_3 &= x_1x_3+x_2x_3-x_3^2-x_1-x_2+x_3. \end{aligned}$$

矩量矩阵 $M_{t-\ell}$ 的秩如下表所示:

阶	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$
t=2	7	4	1			
t=3	11	7	4	1		
t=4	16	11	7	4	1	
t=5	22	16	11	7	4	1

表 5.2: 矩阵 $M_{t-\ell}(y)$ 的秩

显然, $\mathrm{rank}M_{3-1} = \mathrm{rank}M_{4-2} = 7$. 下面检验条件 (5.5). 选择变元序 $x_1 \prec_{\mathrm{tdeg}} x_2 \prec_{\mathrm{tdeg}} x_3$, 矩阵 M_3 的第 5 到 10 行对应次数等于 2 的单项式为

$$x_1^2 \prec x_1x_3 \prec x_1x_3 \prec x_2^2 \prec x_2x_3 \prec x_3^2.$$

误差界选为 10^{-8} 时, M_3 的非主元出现在如下列中

$$1, 2, 3, 4, 5, 6, 8, 11, 12, 14, 17.$$

其中次数等于 2 的非主元列对应的单项为

$$x_1x_3,\ x_2x_3,\ x_3^2,$$

类分别为 $1, 2, 3$. 对 M_{3-1}, 有 $\sum_{j=1}^{3} j\alpha_{3-1}^{(j)} = 6$. 与此同时,

$$\mathrm{corank}M_{4-1} - \mathrm{corank}M_{4-2} = 6.$$

因此, 条件 (5.4)-(5.5) 对于 $t=3$, $\ell=1$ 均满足.

因此, $\ker M_{3-1}$ 的约化基

$$\{x_1 + x_2 - x_3 - x_1^2 - x_1x_2 + x_1x_3,\ x_1 + x_2 - x_3 - x_1x_2 - x_2^2 + x_2x_3,\\ 3x_1 + 3x_2 - 3x_3 - x_1^2 - 2x_1x_2 - x_2^2 + x_3^2\}$$

是 $\sqrt[\mathbb{R}]{I}$ 关于序 $\prec_{\mathrm{tdeg}}$ 的一组 Gröbner 基.

例 5.4. 给定理想 $I = \langle p_1, p_2\rangle$ (见 [109, p.123, 例 7.41]), 其中

$$p_1 = x_1^2 + x_2^2 + x_3^2 - 2,$$
$$p_2 = x_1^2 + x_2^2 - x_3.$$

理想的实代数簇 $V_{\mathbb{R}}(I)$ 严格包含于 $V_{\mathbb{C}}(I)$. 利用文章 [114] 中的延拓-投影方法得到的多项式系统的维数表如 5.3 所示. 其中包含复根信息, t 表示延拓的阶数, ℓ 表示投影的次数.

阶	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
t=0	8	4	1		
t=1	12	8	4	1	
t=2	16	12	8	4	1

表 5.3: 多项式系统维数表

阶	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
t=2	5	3	1		
t=3	7	5	3	1	
t=4	9	7	5	3	1

表 5.4: 矩阵 $M_{t-\ell}(y)$ 的秩

表 5.4 中给出了通过求解一系列半定规划问题 (5.3) 得到的矩量矩阵的秩. 显然, $\mathrm{rank}M_{3-1}=\mathrm{rank}M_{4-2}=5$, $\mathrm{corank}M_{4-1}-\mathrm{corank}M_{4-2}=8$.

下面计算 $\sum_{j=1}^{3} j\alpha_{3-1}^{(j)}$. 选择变元序 $x_1 \prec_{\mathrm{tdeg}} x_2 \prec_{\mathrm{tdeg}} x_3$ 和误差界 10^{-8}, M_3 次数等于 2 的主元在第 5, 6 列, 对应于非主元的 2 次单项为

$$x_1x_3,\ x_2^2,\ x_2x_3,\ x_3^2,$$

类分别为 1, 2, 2 和 3. 从而有

$$\sum_{j=1}^{3} j\alpha_{3-1}^{(j)}=8.$$

因此, $\ker M_{3-1}$ 的约化基 $\{-1+x_3, -1+x_1^2+x_2^2\}$ 是 $\sqrt[\mathbb{R}]{I}$ 关于序 $\prec_{\mathrm{tdeg}}$ 的一组 Gröbner 基.

例 5.5. 给定理想 $I = \langle p_1, p_2, p_3 \rangle$ (见 [116, p.61, 例2.4.12]), 其中

$$\begin{aligned} p_1 &= x_3^2 + x_2x_3 - x_1^2, \\ p_2 &= x_1x_3 + x_1x_2 - x_3, \\ p_3 &= x_2x_3 + x_2^2 + x_1^2 - x_1. \end{aligned}$$

选择变元序 $x_3 \prec_{\text{tdeg}} x_1 \prec_{\text{tdeg}} x_2$. 矩量矩阵 $M_{t-\ell}$ 的秩如表 5.5 所示.

阶	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$
t=2	7	4	1			
t=3	10	7	4	1		
t=4	13	10	7	4	1	
t=5	16	13	10	7	4	1

表 5.5: 矩阵 $M_{t-\ell}(y)$ 的秩

对所有 $M_{t-\ell}$, 都有 $\sum_{j=1}^{3} j\alpha_{t-\ell}^{(j)} < \text{corank}M_{t+1-\ell} - \text{corank}M_{(t+1)-\ell+1}$. 这说明坐标系 (x_1, x_2, x_3) 不是 δ-正则坐标系. 然而, 通过坐标变换 $\tilde{x}_1 = x_2 + x_3$, $\tilde{x}_2 = x_1$ 和 $\tilde{x}_3 = x_3$ 及自约化, 可将多项式系统转化为 $\tilde{P} = \{\tilde{x}_1\tilde{x}_3 - \tilde{x}_2^2, \tilde{x}_1\tilde{x}_2 - \tilde{x}_3, \tilde{x}_1^2 - \tilde{x}_2\}$. 利用此多项式系统求得的矩量矩阵的秩与表 5.1 相同. 因此, 条件 (5.4)-(5.5) 对于 $t = 3$, $\ell = 1$ 成立.

例 5.6. 给定理想 $I = \langle p_1, p_2 \rangle$, 其中

$$\begin{aligned} p_1 &= x_1^4 + 2x_1^3x_2 - 2x_2^3x_1 - x_2^4 + x_2^2x_1^3 + x_2^3x_1^2 - x_2^4x_1 - x_2^5, \\ p_2 &= x_1^5 + x_1^4x_2 - x_2^4x_1 - x_2^5. \end{aligned}$$

理想 I 不是实根理想. 矩阵 $M_{t-l}(y)$ 的秩如表 5.6 所示.

给定变元序 $x_1 \prec_{\text{tdeg}} x_2$. 截断的矩量矩阵 M_{6-4} 的奇异值为

$$1.98859,\ 0.30359,\ 0.29970,\ 0.14185,\ 0.08482,\ 0.00001.$$

我们选定误差界 10^{-4}, 并在表 5.6 中列出矩量矩阵的秩.

如表 5.6 所示, $\text{rank}M_{6-1} = \text{rank}M_{7-2} = 11$, 并通过计算得 $\sum_{j=1}^{2} j\alpha_{6-1}^{(j)} = \text{corank}M_{7-1} - \text{corank}M_{7-2} = 5$.

阶	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$	$\ell=6$	$\ell=7$	$\ell=8$
t=6	13	11	9	7	5	3	1		
t=7	15	13	11	9	7	5	3	1	
t=8	17	15	13	11	9	7	5	3	1

表 5.6: 矩阵 $M_{t-\ell}(y)$ 的秩

因此, $\ker M_{6-4}$ 的约化基 $\{-x_1^2+x_2^2\}$ 是实根理想 $\sqrt[\mathbb{R}]{I}$ 关于序 $\prec_{\text{tdeg}}$ 的一组 Gröbner 基.

下面给出求 $\mathcal{S}$-根理想 $I(V_{\mathbb{R}}(I)\cap\mathcal{S})$ 的例子.

例 5.7. 给定理想 $I=\langle p_1,p_2\rangle$, 其中

$$p_1=x_1^3+x_1^2x_2-x_2^2x_1-x_2^3+x_2^2x_1^2-x_2^4,$$
$$p_2=x_1^4-x_2^4.$$

半代数集

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

矩阵 $M_{t-l}(y)$ 的秩如表 5.7 所示, 其中 $y\in\mathcal{K}_{t,\mathcal{S}}^{gen}$.

阶	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$	$\ell=6$
t=4	6	4	3	2	1		
t=5	7	5	4	3	2	1	
t=6	8	6	5	4	3	2	1

表 5.7: 矩阵 $M_{t-\ell}(y)$ 的秩, $y\in\mathcal{K}_{t,\mathcal{S}}^{gen}$

给定变元序 $x_1\prec_{\text{tdeg}}x_2$ 和误差界 10^{-8}, 如上表所示, $\text{rank}M_{5-1}=\text{rank}M_{6-2}=5$, 并通过计算得 $\sum_{j=1}^{2}j\alpha_{5-1}^{(j)}=\text{corank}M_{6-1}-\text{corank}M_{6-2}=5$.

因此, $\ker M_{5-1}$ 的约化基 $\{-x_1+x_2\}$ 是 $\mathcal{S}$-根理想 $I(V_{\mathbb{R}}(I)\cap\mathcal{S})$ 关于序 $\prec_{\text{tdeg}}$ 的一组 Gröbner 基.

阶	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$	$\ell=6$
t=4	9	7	5	3	1		
t=5	11	9	7	5	3	1	
t=6	13	11	9	7	5	3	1

表 5.8: 矩阵 $M_{t-\ell}(y)$ 的秩, $y \in \mathcal{K}_t^{gen}$

表 5.8 中列出了去掉不等式限制条件 $\mathcal{S}$ 后解半定规划问题 (5.3) 所得矩量矩阵的秩.

注意到 $\mathrm{rank}M_{5-1} = \mathrm{rank}M_{6-2} = 9$. 给定变元序 $x_1 \prec_{\mathrm{tdeg}} x_2$ 和误差界 10^{-5}, 通过计算可知 $\sum_{j=1}^{2} j\alpha_{5-1}^{(j)} = \mathrm{corank}M_{6-1} - \mathrm{corank}M_{6-2} = 4$. 因此, $\ker M_{5-1}$ 的约化基 $\{-x_1^2 + x_2^2\}$ 是实根理想 $\sqrt[\mathbb{R}]{I}$ 关于序 $\prec_{\mathrm{tdeg}}$ 的一组 Gröbner 基.

第六章　结论与展望

本文主要研究多项式优化领域内以下三个问题：精确验证多项式的全局非负性、求多项式系统的实根, 给定具有正维实代数簇的多项式理想 I, 求实根理想 $I(V_{\mathbb{R}}(I))$ 的一组 Gröbner 基.

第三章中, 我们讨论了如何通过求平方和数目最少的多项式平方和分解来精确验证给定多项式的全局非负性. 我们提出了一种新的求解 Gram 矩阵核范数极小化问题 (3.4) 的一阶算法——改进的不动点迭代算法 (MFPC-BB), 并给出了算法的收敛性分析以及在 Maple 和 Matlab 中的实现. 我们的算法以不动点迭代算法中的算子分裂技术为基础, 通过改进阈值算子 $\mathcal{T}$ 以适应于求解对称半正定矩阵的恢复问题. 同时, 我们还引入Barzilai-Borwein 技术来进行步长参数的选取, 从而提高算法的收敛速度. 在此方法基础上, 我们又提出一种加速的不动点迭代算法 (AFPC-BB). 它既保持了 MFPC-BB 算法的简单易实现的特点, 又能够将原算法线性的收敛速度提高为二次收敛. 数值实验显示 AFPC-BB 算法对于低秩 Gram 矩阵的近似或精确恢复较基于半定规划的方法提速明显, 更适合大规模问题的求解.

第四章中, 我们研究了如何快速求解大规模多项式系统的实根. 通过将 Lasserre 提出的半定规划模型 (4.3) 中的目标函数改为矩量矩阵的核范数 $\|M_t(y)\|_*$, 我们将多项式系统实根求解的问题转化为矩量矩阵核范数极小化问题 (4.5), 并利用第三章给出的 AFPC-BB 算法求解. 如果返回的矩量矩阵满足条件 (4.4), 那么可以通过求矩量矩阵的像空间的一组基和相应乘法矩阵的公共特征向量得到多项式方程组的实根. 同时, 我们给出了算法的收敛性分析和在 Maple 和 Matlab 中的实现 (MMCRSolver). 我们的算法不能保证求出多项式系统的全部实根. 对于较大规模的多项式系统, 如果只存在一个或少数几个实根, 我们的方法能够快速地将它们求解出来. 与此同时, 如果多项式系统有无穷多个实根, 我们仍能求出其中部分孤立实根或是在代数流形上的实根. 数值实验显示, 对于半正定规划的方法难以求解的例子, MMCRSolver 也能快速地求出其全部或部分的实根.

第五章中, 给定具有正维实代数簇的多项式理想 I, 我们提出了一种基于矩量矩阵半正定松弛 (5.3) 的符号—数值混合算法求理想 I 的实根理想 $I(V_{\mathbb{R}}(I))$ 的

一组 Gröbner 基. 通过将几何对合理论与半正定矩量矩阵的性质相结合, 我们给出了半正定松弛 (5.3) 在正维情形下终止性的判定定理 5.2. 基于猜想 5.1, 我们证明了在 δ-正则坐标系下, 定理 5.2 中的条件 (5.4)-(5.5) 一定在有限步的半正定松弛内满足, 并给出了实根理想 $I(V_{\mathbb{R}}(I))$ 关于序 $\prec_{\text{tdeg}}$ 的一组 Gröbner 基. 条件 (5.4)-(5.5) 可以作为 Flat Extension 定理 5.1 中条件 (5.1) 在正维情形下的推广. 与此同时, 给定的半代数集 $\mathcal{S} = \{x \in \mathbb{R}^n \mid f_1(x) \geq 0, \ldots, f_s(x) \geq 0\}$, 我们将算法推广到求理想 I 的 $\mathcal{S}$-根理想 $I(V_{\mathbb{R}}(I) \cap \mathcal{S})$ 的 Gröbner 基.

今后的工作主要包含以下几个方面:

1. 基于低秩矩量矩阵恢复的多项式系统实根求解算法得到的是有限精度的数值解, 我们希望在未来的工作中结合 Rump [111, 112] 的浮点数区间算法准确地刻画实根所在的范围.

2. 关于正维实根理想 $\sqrt[\mathbb{R}]{I}$ 的计算, 我们提出的判定准则 (5.4)-(5.5) 都建立在猜想 5.1 的基础上. 尽管数值实验中所有的例子都说明此猜想是正确的, 但是目前仍然没有理论证明. 在今后的工作中, 我们希望能证明此猜想.

3. 算法 5.1 是建立在求解一系列目标函数为常数的半正定规划问题上的. 我们希望将求正维实根理想的判定准则 (5.4)-(5.5) 推广到更一般的多项式优化问题上.

参考文献

[1] J. Abernethy, F. Bach, T. Evgeniou, and J.-P. Vert. Low-rank matrix factorization with attributes. Technical report, N24/06/MM, Ecole des Mines de Paris, 2006.

[2] Y. Amit, M. Fink, N. Srebro, and S. Ullman. Uncovering shared structures in multiclass classification. In *Proceedings of the Twenty-fourth International Conference on Machine Learning*, 2007.

[3] A. Argyriou, T. Evgeniou, and M. Pontil. Multi-task feature learning. In *Neural Information Processing Systems*, 2007.

[4] E. Artin. Über die zerlegung definiter funktionen in quadrate. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 5:100–115, 1927.

[5] P. Aubry, F. Rouillier, and M. Safey El Din. Real solving for positive dimensional systems. *J. Symbolic Comput.*, 34(6):543–560, 2002.

[6] B. Bank, M. Giusti, J. Heintz, and G.M. Mbakop. Polar varieties and efficient real elimination. *Math. Z.*, 238(1):115–144, 2001.

[7] J. Barzilai and J.M. Borwein. Two-point step size gradient methods. *IMA J. Numer. Anal.*, 8:141–148, 1988.

[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] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in real algebraic geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, 2003.

[10] D. Bates and F. Sottile. Khovanskii-rolle continuation for real solutions. A version of this article will appear in Foundations of Computational Mathematics, 2011.

[11] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2(1):183–202, 2009.

[12] 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.

[13] D. Bini and B. Mourrain. Polynomial test suite. 1996. Available at http://www.sop.inria.fr/saga/{POL}.

[14] Grigoriy Blekherman. There are significantly more nonegative polynomials than sums of squares. *Israel Journal of Mathematics*, 153(1):355–380, December 2006.

[15] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1998. Translated from the 1987 French original, Revised by the authors.

[16] B. Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*. PhD thesis, Mathematical Institute, University of Innsbruck, Austria, 1965.

[17] S. Burer and R.D.C. Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Math. Program.*, 95(2):329–357, 2003.

[18] S. Burer and R.D.C. Monteiro. Local minima and convergence in low-rank semidefinite programming. *Math. Program.*, 103(3):427–444, 2005.

[19] Jian-Feng Cai, Emmanuel J. Candès, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.*, 20(4):1956–1982, 2010.

[20] Emmanuel J. Candès and Benjamin Recht. Exact matrix completion via convex optimization. *Found. Comput. Math.*, 9(6):717–772, 2009.

[21] Emmanuel J. Candès and Terence Tao. The power of convex relaxation: near-optimal matrix completion. *IEEE Trans. Inform. Theory*, 56(5):2053–2080, 2010.

[22] Gong Chen and Marc Teboulle. A proximal-based decomposition method for convex minimization problems. *Math. Programming*, 64(1, Ser. A):81–101, 1994.

[23] G. Chesi, A. Garulli, A. Tesi, and A. Vicino. An LMI-based approach for characterizing the solution set of polynomial systems. In *Decision and Control, 2000. Proceedings of the 39th IEEE Conference on*, volume 2, pages 1501–1506, 2000.

[24] G. Chesi, A. Garulli, A. Tesi, and A. Vicino. Characterizing the solution set of polynomial systems in terms of homogeneous forms: an LMI approach. *International Journal of Robust and Nonlinear Control*, 13(13):1239–1257, 2003.

[25] G. Chesi and Y.S. Hung. Solving polynomial systems: an LMI-based approach. In *Decision and Control, 2006 45th IEEE Conference on*, pages 5132 –5137, dec. 2006.

[26] A.L. Chistov and D. Grigoriev. Complexity of quantifier elimination in the theory of algebraically closed fields. In *Proceedings of the Mathematical Foundations of Computer Science 1984*, pages 17–31, London, UK, 1984. Springer-Verlag.

[27] M.D. Choi, T.Y. Lam, and B. Reznick. Sums of squares of real polynomials. *Symp. in Pure Math.*, 58(2):103–126, 1995.

[28] G. E. Collins. Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition. volume 33, pages 134–183. Springer-Verlag, Berlin, 1975.

[29] Robert M. Corless, Patrizia M. Gianni, and Barry M. Trager. A reordered Schur factorization method for zero-dimensional polynomial systems with multiple roots. In *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation (Kihei, HI)*, pages 133–140 (electronic), New York, 1997. ACM.

[30] D. Cox, J. Little, and D. O'Shea. *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1997.

[31] R. Curto and L. Fialkow. Solution of the truncated complex moment problem for flat data. *Memoirs of the American Mathematical Society*, 119(568):1–62, 1996.

[32] E. de Klerk, C. Roos, and T. Terlaky. Initialization in semidefinite programming via a self-dual skew-symmetric embedding. *Oper. Res. Lett.*, 20(5):213–221, 1997.

[33] James Demmel and Bo Kågström. The generalized schur decomposition of an arbitrary pencil a-λb - robust software with error bounds and applications. part i: theory and algorithms. *ACM Trans. Math. Softw.*, 19(2):160–174, 1993.

[34] James Demmel and Bo Kågström. The generalized schur decomposition of an arbitrary pencil a-λb - robust software with error bounds and applications. part ii: software and applications. *ACM Trans. Math. Softw.*, 19(2):175–201, 1993.

[35] C.D. Dezell. A continuous, constructive solution to hilbert's 17th problem,. *Invent. math.*, 76(3):365–384, 1984.

[36] Alicia Dickenstein and Ioannis Z. Emiris. *Solving Polynomial Equations*, volume 14 of *Algorithms and Computation in Mathematics*. Springer, 2005.

[37] Jonathan Eckstein and Dimitri P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Programming*, 55(3, Ser. A):293–318, 1992.

[38] M. Fazel. *Matrix rank minimization with applications*. PhD thesis, Stanford University, 2002.

[39] M. Fazel, H. Hindi, and S.P. Boyd. A rank minimization heuristic with application to minimum order system approximation. In *Proceedings of the 2001 American Control Conference*, pages 4734–4739, 2001.

[40] A.V. Fiacco and G.P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Wiley, 1968. Reprinted by SIAM Publications, 1990.

[41] D. Gabay. Applications of the method of multipliers to variational inequalities. In M. Fortin and R. Glowinski, editors, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, volume 15, pages 299–331. Elsevier, 1983.

[42] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers Mathematics with Applications*, 2:1416–1438, 1976.

[43] D. Goldfarb and S. Ma. Convergence of fixed point continuation algorithms for matrix rank minimization. *Foundations of Computational Mathematics*, 11(2):183–210, 2011.

[44] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[45] Elaine T. Hale, Wotao Yin, and Yin Zhang. Fixed-point continuation for l_1-minimization: methodology and convergence. *SIAM J. Optim.*, 19(3):1107–1130, 2008.

[46] D. Henrion and J.B. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. In *Positive polynomials in control*, volume 312 of *Lecture Notes in Control and Inform. Sci.*, pages 293–310. Springer, Berlin, 2005.

[47] D. Henrion and J. Malick. Projection methods in convex optimization. *LAAS-CNRS Research Report 10730*, 2010.

[48] D. Henrion and J. Malick. Projection methods for conic feasibility problems; application to sum-of-squares decompositions. *Optimization Methods and Software*, 26(1):23–46, 2011.

[49] D. Hilbert. Über ternäre definite Formen. *Acta Math*, 17:169–197, 1893.

[50] S. Ji and J. Ye. An accelerated gradient method for trace norm minimization. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pages 457–464, New York, NY, USA, 2009. ACM.

[51] Zhongxiao Jia. Composite orthogonal projection methods for large matrix eigenproblems. *Sci. China Ser. A*, 42:578 - 585, 1999.

[52] Zhongxiao Jia. Polynomial characterizations of the approximate eigenvectors by the refined arnoldi method and an implicitly restarted refined arnoldi algorithm. *Linear Algebra Appl.*, 287:191 - 214, 1999.

[53] Zhongxiao Jia and G.W. Stewart. An analysis of the rayleigh ╟ritz method for approximating eigenspaces. *Math. Comput.*, 70:637 - 647, 2000.

[54] M. Jirstrand. Nonlinear control system design by quantifier elimination. *J. Symbolic Computation*, 24:137–152, 1997.

[55] E. Kaltofen, B. Li, Z. Yang, and L. Zhi. Exact certification of global optimality of approximate factorizations via rationalizing sums-of-squares with floating point scalars. In *ISSAC '08: Proceedings of the twenty-first international symposium on Symbolic and algebraic computation*, pages 155–164, New York, NY, USA, 2008. ACM.

[56] E. Kaltofen, B. Li, Z. Yang, and L. Zhi. Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients, 2009. Accepted for publication in J. Symbolic Comput.

[57] Michal Kočvara and Michael Stingl. On the solution of large-scale SDP problems by the modified barrier method using iterative solvers. *Math. Program.*, 109(2-3, Ser. B):413–444, 2007.

[58] M. Kuranishi. On E. Cartan's prolongation theorem of exterior differential systems. *Amer. J. Math.*, 79:1–47, 1957.

[59] R.M. Larsen. PROPACK - software for large and sparse SVD calculations. Available from: http://soi.stanford.edu/~rmunk/PROPACK/.

[60] J.B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817 (electronic), 2001.

[61] J.B. Lasserre. Polynomials nonnegative on a grid and discrete representations. *Transactions of the American Mathematical Society*, 354, 2001.

[62] J.B. Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press, 2009.

[63] J.B. Lasserre, M. Laurent, B. Mourrain, P. Trébuchet, and P. Rostalski. Moment matrices, border bases and radical computation. Preprint, 2011.

[64] J.B. Lasserre, M. Laurent, and P. Rostalski. Semidefinite characterization and computation of zero-dimensional real radical ideals. *Foundations of Computational Mathematics*, 8:607–647, 2008.

[65] J.B. Lasserre, M. Laurent, and P. Rostalski. A prolongation-projection algorithm for computing the finite real variety of an ideal. *Theoretical Computer Science*, 410(27-29):2685–2700, 2009.

[66] 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.

[67] M. Laurent. Revisiting two theorems of Curto and Fialkow on moment matrices. *Proceedings of the American Mathematical Society*, 133(10):2965–2976, 2005.

[68] D. Lazard. Thirty years of polynomial system solving, and now? *J. Symbolic Comput.*, 44(3):222–231, 2009.

[69] A. S. Lewis. Convex analysis on the Hermitian matrices. *SIAM J. Optim.*, 6(1):164–177, 1996.

[70] Z. Liu and L. Vandenberghe. Interior-point method for nuclear norm approximation with application to system identification. *SIAM J. Matrix Anal. Appl.*, 31:1235–1256, 2009.

[71] S. Ma, D. Goldfarb, and L. Chen. Fixed point and bregman iterative methods for matrix rank minimization. *Math. Program.*, pages 1–33, 2009.

[72] Y. Ma. The minimum-rank Gram matrix completion via fixed point continuation method (in Chinese). *Journal of Systems Science and Mathematical Sciences*, 30(11):1501–1511, 2010.

[73] Yue Ma and Lihong Zhi. The minimum-rank Gram matrix completion via modified fixed point continuation method. In *ISSAC 2011: Proceedings of the 36th international symposium on Symbolic and algebraic computation*, pages 241–248, New York, NY, USA, 2011. ACM.

[74] Jérôme Malick, Janez Povh, Franz Rendl, and Angelika Wiegele. Regularization methods for semidefinite programming. *SIAM J. Optim.*, 20(1):336–356, 2009.

[75] M. Marshall. *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008.

[76] Raghu Meka, Prateek Jain, and Inderjit S. Dhillon. Matrix completion from power-law distributed samples. In *NIPS*, 2009.

[77] M. Mesbahi and G. P. Papavassilopoulos. On the rank minimization problem over a positive semidefinite linear matrix inequality. *IEEE Trans. Automat. Control*, 42(2):239–243, 1997.

[78] H.M. Möller. An inverse problem for cubature formulae. *Computational Technologies*, 9(13-20), 2004.

[79] Alexander Morgan and Vadim Shapiro. Box-bisection for solving second-degree systems and the problem of clustering. *ACM Trans. Math. Software*, 13(2):152–167, 1987.

[80] T.S. Motzkin. *The arithmetic-geometric inequality*, pages 205–224. Inequalities. Academic Press, 1967.

[81] B. Mourrain and J. P. Pavone. Subdivision methods for solving polynomial equations. *J. Symb. Comput.*, 44:292–306, March 2009.

[82] Sahand Negahban, Pradeep D. Ravikumar, Martin J. Wainwright, and Bin Yu. A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers. In *Proceedings of the Conference on Neural Information Processing Systems (NIPS)*, pages 1348–1356, 2009.

[83] Y. Nesterov. A method of solving a convex programming problem with convergence rate $O(1/k^2)$. *Soviet Mathematics Doklady*, 27:372–376, 1983.

[84] Y. Nesterov. Squared functional systems and optimization problems. In H. Frenk, K. Roos, T. Terlaky, and S. Zhang, editors, *High Performance Optimization*, pages 405–440. Kluwer Academic Publishers, 2000.

[85] Y. Nesterov. Smooth minimization of non-smooth functions. *Math. Program.*, 103(1):127–152, 2005.

[86] Y. Nesterov. Gradient methods for minimizing composite objective function. Technical report, 2007.

[87] R. Neuhaus. Computation of real radicals of polynomial ideals. II. *J. Pure Appl. Algebra*, 124(1-3):261–280, 1998.

[88] A. Neumaier. *Interval Methods for Systems of Equations*, volume 37 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1991.

[89] J. Nie. Regularization methods for sum of squares relaxations in large scale polynomial optimization. Technical report, 2009. Available: http://arxiv.org/abs/0909.3551.

[90] Pablo A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. Dissertation (Ph.D.), California Institute of Technology, 2000.

[91] Pablo A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96(2):293–320, 2003.

[92] H. Peyrl and P.A. Parrilo. A Macaulay 2 package for computing sum of squares decompositions of polynomials with rational coefficients. In *ProcSNC07*, pages 207–208, 2007.

[93] H. Peyrl and P.A. Parrilo. Computing sum of squares decompositions with rational coefficients. *Theoretical Computer Science*, 409:269–281, 2008.

[94] A. Pfister. Zur Darstellung definiter Funktionen als Summe von Quadraten. *Inventiones Math.*, 4(4):229–236, 1967.

[95] G. Polya. Über positive Darstellung von Polynomen Viereljschr. *Ges. Zürich*, 73:141–145, 1928.

[96] J.F. Pommaret. *Systems of Partial Differential Equations and Lie Pseudogroups*. Gordon & Breach, 1978.

[97] Y. Pourchet. Sur la représentation en somme de carrés des polynômes à une indéterminée sur un corps de nombres algébriques. *Acta Arith.*, 19:89–104, 1971.

[98] Victoria Powers and Thorsten Wörmann. An algorithm for sums of squares of real polynomials. *J. Pure Appl. Algebra*, 127(1):99–104, 1998.

[99] B. Recht, W. Xu, and B. Hassibi. Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization. In *CDC*, pages 3065–3070, 2008.

[100] Benjamin Recht, Maryam Fazel, and Pablo A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, 2010.

[101] G. Reid and L. Zhi. Solving polynomial systems via symbolic-numeric reduction to geometric involutive form. *J. Symbolic Comput.*, 44(3):280–291, 2009.

[102] Gregory J. Reid, Jianliang Tang, and Lihong Zhi. A complete symbolic-numeric linear method for camera pose determination. In *ISSAC 2003*, pages 215–223, New York, NY, USA, 2003. ACM.

[103] J.D.M. Rennie and N. Srebro. Fast maximum margin matrix factorization for collaborative prediction. In *Proceedings of the 22nd international conference on Machine learning*, ICML '05, pages 713–719, 2005.

[104] B. Reznick. Extremal PSD forms with few terms. *Duke Mathematical Journal*, 45(2):363–374, 1978.

[105] B. Reznick. Uniform denominators in Hilbert's Seventeen Problem. *Math. Z*, 220:75–97, 1995.

[106] B. Reznick. Some concrete aspects of Hilbert's 17th problem. In Charles N. Delzell and James J. Madden, editors, *Real Algebraic Geometry and Ordered Structures*, volume 253 of *Contemporary Mathematics*, pages 251–272. AMS, Providence, RI, USA, 2000.

[107] J.F. Ritt. *Differential Algebra*. New York: American Mathematical Society, 1950.

[108] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, 1972.

[109] P. Rostalski. *Algebraic moments. real root finding and related topics*. PhD thesis, ETH Zurich, 2009.

[110] F. Rouillier. Solving zero-dimensional systems through the rational univariate representation. *Appl. Algebra Engrg. Comm. Comput.*, 9(5):433–461, 1999.

[111] Siegfried M. Rump. INTLAB - INTerval LABoratory. In Tibor Csendes, editor, *Developments in Reliable Computing*, pages 77–104. Kluwer Academic Publishers, Dordrecht, 1999.

[112] Siegfried M. Rump. Verification methods: Rigorous results using floating-point arithmetic. *Acta Numerica*, 19:287–499, 2010.

[113] M. Safey El Din and E. Schost. Polar varieties and computation of one point in each connected component of a smooth real algebraic set. In *Proceedings of the 16th international symposium on Symbolic and algebraic computation*, ISSAC '03, pages 224–231, New York, NY, USA, 2003. ACM.

[114] 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. SpringerWienNewYork, Vienna, 2009.

[115] W.M. Seiler. *Analysis and Application of the Formal Theory of Partial Differential Equations*. Dissertation, Lancaster University, Germany, 1994.

[116] W.M. Seiler. Involution - the formal theory of differential equations and its applications in computer algebra and numerical analysis. *Habilitation Thesis, Univ. of Mannheim*, 2002.

[117] W.M. Seiler. *Involution: The Formal Theory of Differential Equations and its Applications in Computer Algebra*, volume 25 of *Algorithms and Computation in Mathematics*. Springer-Verlag, 2010.

[118] N.Z. Shor. Quadratic optimization problems. *Soviet Journal of Computer and Systems Sciences*, 25:1–11, 1987.

[119] N.Z. Shor. An approach to obtaining global extremums in polynomial mathematical programming problems. *Cybernetics*, 23(5):695–700, 1988.

[120] ACM SIGKDD and Netflix. *Proceedings of KDD Cup and Workshop*, 2007.

[121] A. Sommese and C. Wampler. *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*. World Scientific Press, Singapore, 2005.

[122] N. Srebro, J.D.M. Rennie, and T.S. Jaakkola. Maximum-margin matrix factorization. In *Advances in Neural Information Processing Systems*, 2005.

[123] G. Stengle. A nullstellensatz and positivstellensatz in semialgebraic geometry. *Math. Ann.*, 207:87–97, 1994.

[124] H.J. Stetter. *Numerical Polynomial Algebra*. SIAM, 2004.

[125] J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11/12:625–653, 1999.

[126] Bernd Sturmfels. *Solving systems of polynomial equations*, volume 97 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2002.

[127] W.J. Sweeney. The D-Neumann problem. *Acta. Math.*, 120:223–251, 1968.

[128] K.-C. Toh, M.J. Todd, and R.H. Tütüncü. SDPT3 - a matlab software package for semidefinite programming. *Optimization Methods and Software*, 11:545–581, 1998.

[129] K.-C. Toh and S. Yun. An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems. Technical report, 2009. Available: http://www.optimization-online.org/DBHTML/2009/03/2268.html.

[130] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: a factorization method. *International Journal of Computer Vision*, 9(2).

[131] P. Tseng. On accelerated proximal gradient methods for convex-concave optimization. *Submitted to SIAM J. Optim*, 2008.

[132] Wen tsun Wu. *Mathematics Mechanization*. Science Press and Kluwer Academic Publishers, 2000.

[133] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Rev.*, 38(1):49–95, 1996.

[134] J. Verschelde. Algorithm 795: PHCpack: A general-purpose solver for polynomial systems by homotopy continuation. *ACM Transactions on Mathematical Software*, 25(2):251–276, 1999.

[135] Hayato Waki, Sunyoung Kim, Masakazu Kojima, and Masakazu Muramatsu. Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization*, 17:218–242, 2006.

[136] G. A. Watson. Characterization of the subdifferential of some matrix norms. *Linear Algebra Appl.*, 170:33–45, 1992.

[137] Zaiwen Wen, Donald Goldfarb, and Wotao Yin. Alternating direction augmented Lagrangian methods for semidefinite programming. *Math. Program. Comput.*, 2(3-4):203–230, 2010.

[138] Zaiwen Wen, Wotao Yin, Donald Goldfarb, and Yin Zhang. A fast algorithm for sparse reconstruction based on shrinkage, subspace optimization, and continuation. *SIAM J. Sci. Comput.*, 32(4):1832–1857, 2010.

[139] 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.

[140] Stephen J. Wright, Robert D. Nowak, and Mário A.T. Figueiredo. Sparse reconstruction by separable approximation. *IEEE Trans. Signal Process.*, 57(7):2479–2493, 2009.

[141] Bican Xia and Lu Yang. An algorithm for isolating the real solutions of semi-algebraic systems. *J. Symb. Comput.*, 34(5):461–477, 2002.

[142] Bican Xia and Ting Zhang. Real solution isolation using interval arithmetic. *Computers & Mathematics with Applications*, 52(6-7):853–860, 2006.

[143] L. Yang, X.R. Hou, and Z.B. Zeng. A complete discrimination system for polynomials. *Sci. China, E*, 39:628 – 646, 1996.

[144] G. Zeng. Computation of generalized real radicals of polynomial ideals. *Sci. China Ser. A*, 42(3):272–280, 1999.

[145] Xin-Yuan Zhao, Defeng Sun, and Kim-Chuan Toh. A Newton-CG augmented Lagrangian method for semidefinite programming. *SIAM J. Optim.*, 20(4):1737–1765, 2010.

发表文章目录

[1] 马玥. 基于改进的不动点迭代算法的低秩Gram矩阵的恢复. 系统科学与数学, 2010, 30(11), 1501-1511.

[2] Yue Ma and Lihong Zhi. The minimum-rank Gram matrix completion via modified fixed point continuation method. In ISSAC 2011: Proceedings of the 36th international symposium on symbolic and algebraic computation (San Jose, CA, USA, 2011), ACM, pp. 241-248.

[3] Yue Ma and Lihong Zhi. Computing Real Solutions of Polynomial Systems via Low-rank Matrix Completion. ISSAC 2012, Grenoble, France, 2012.

[4] Yue Ma and Lihong Zhi. Semidefinite Characterization and Computation of Positive-Dimensional Real Radical Ideals. Preprint.2012.

简　　历

马玥, 女, 黑龙江省, 1984 年出生.　E-mail: yma@mmrc.iss.ac.cn

教育状况

2009.9–2012.7　中国科学院数学与系统科学研究院, 系统所, 应用数学, 博士, 导师：支丽红研究员.

2007.9–2009.7　吉林大学, 数学研究所, 计算数学, 硕士, 导师：雷娜教授.

2003.9–2007.7　吉林大学, 数学学院, 信息与计算科学, 学士.

会议活动

1. 第四届全国计算机数学会议, 并作报告, 广州, 2011.11.25-28
2. 可信计算国际会议, 并作报告（英文）, 广西, 2011.7.17-20
3. 第36届国际符号与代数计算年会（ISSAC 2011）, 并作报告（英文）, San Jose, 美国, 2011.6.8-11
4. 清华大学第一届“清华软件日”, 并作报告（英文）, 北京, 2011.4.11-12
5. 第三届全国计算机数学会议, 并作报告, 上海, 2010.10.19-22

获奖经历

2012.4　　中科院系统所“三好学生”

2011.5　　中国科学院博时奖学金

2011.4　　清华大学第一届”清华软件日”获得唯一最佳学生报告奖

致　　谢

三年的博士生生活一晃而过, 回首走过的岁月, 心中倍感充实, 论文即将完成之日, 感慨良多. 首先诚挚的感谢我的恩师支丽红研究员, 这篇论文的内容从选题到研究再到写作都是在支老师的悉心指导下完成的, 恩师对我的指导和影响之大, 无以言表, 自己取得的点滴成绩无不凝聚着恩师的心血. 恩师国际化的视野, 前沿而精髓的学术造诣, 严谨的治学之道, 宽厚仁慈的胸怀, 积极乐观的生活态度, 为我树立了一生学习的典范, 她的教诲与鞭策将激励我在今后人生的道路上励精图治, 开拓创新.

衷心感谢聂家旺教授、梁野师兄和李楠在论文研究讨论中的贡献和帮助! 聂家旺教授对学术科研充满了激情, 在同他讨论的过程中, 我学习到了很多知识. 在同梁野师兄和李楠一起学习期间, 从他们那里我获得了很多启发和热情的帮助.

郑重感谢我的硕士导师, 吉林大学数学学院雷娜教授. 雷老师是我一生的良师益友, 在我艰难奋战的日子里, 她给予我许多学习和生活上的各种帮助, 为我排解各种困难.

衷心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、王定康研究员、刘卓军研究员. 从他们那里我学习到了很多的知识. 同时感谢周代珍老师和丁健敏老师的热心帮助.

衷心感谢现在或曾经在数学机械化中心学习的杨争峰师兄、李冰玉师姐、吴晓丽师姐、赵尚威师姐、张梅师姐、李斌师兄、梁野师兄、郭峰、李楠、李子佳、郭庆东、刘琦、李喆以及实验室其他同学. 也感谢一起上讨论班的老师和同学. 通过讨论班上的讨论交流让我学习到了很多知识. 在此, 我还要感谢我的“饭团”同志们, 靳庆芳、李伟、柳刚、戴兆鹏、金凯, 他们给我的科研生活增添了绚丽的色彩, 在我需要帮助的时候, 他们总是第一时间放下自己的工作无私地为我做一切, 在我心情低落时为我分忧, 在我信心不足时给我打气, 从生活的点点滴滴中给我帮助和鼓励.

最后, 谨以此文献给我挚爱的双亲和男友, 你们在背后的默默支持是我前进的动力. 在此祝愿你们身体健康, 工作顺利, 心情愉快!