中国科学: 数学　2025 年　第 55 卷　第 2 期: 481～500
SCIENTIA SINICA Mathematica
《中国科学》杂志社
SCIENCE CHINA PRESS
CrossMark
click for updates

综　述

多项式优化简介

献给韩继业教授 90 寿辰

黄磊[1], 聂家旺[1]*, 袁亚湘[2]

1. Department of Mathematics, University of California San Diego, La Jolla, CA 92093, USA;
2. 中国科学院数学与系统科学研究院, 北京 100190
E-mail: leh010@ucsd.edu, njw@math.ucsd.edu, yyx@lsec.cc.ac.cn

收稿日期: 2024-04-27; 接受日期: 2024-09-22; 网络出版日期: 2024-12-18; * 通信作者

摘要　多项式优化是目标函数和约束函数均为多项式的一类特殊的非线性优化问题. 矩 - 平方和松弛算法是全局求解多项式优化问题的主要算法. 本文主要介绍多项式优化的矩 - 平方和松弛以及其有限收敛性理论、全局最优性的判别以及如何获取全局最优解. 当标准的矩 - 平方和松弛不具有收敛性时, 本文介绍如何利用 Lagrange 乘子表示来构造具有有限收敛性的矩 - 平方和松弛.

关键词　多项式优化　矩 - 平方和松弛　有限收敛　最优性条件　Lagrange 乘子表示

MSC (2020) 主题分类　90C23, 65K05, 90C22

1　引言

多项式优化 (polynomial optimization) 是一类重要的非线性优化问题, 其通常可以表达成如下形式:

$$\begin{cases} \min & f(x) \\ \text{s.t.} & c_i(x) = 0 \quad (i \in \mathcal{E}), \\ & c_j(x) \geqslant 0 \quad (j \in \mathcal{I}), \end{cases} \tag{1.1}$$

这里 $f(x)$, $c_i(x)$ 和 $c_j(x)$ 为关于变量 $x := (x_1, \ldots, x_n) \in \mathbb{R}^n$ 的多项式函数, $\mathcal{E}$ 和 $\mathcal{I}$ 分别是等式和不等式约束的不交的有限指标集. 记 K 为问题 (1.1) 的可行集, $f_{\min}$ 为 (1.1) 的最优值. 多项式优化包含了很多重要的优化问题 [27,31,32,34,35]. 与传统的非凸优化通常设计算法求解驻点或者局部最优解不同, 多项式优化研究通常关注于求解全局最优值和全局最优解 1).

矩 - 平方和松弛序列 (moment-SOS hierarchy) 是全局求解多项式优化的主要方法 [12]. 自它被提出以来, 该方法在优化领域引起了广泛的关注, 并被应用于求解各类实际问题. 矩 - 平方和松弛由

1) 本文若无特殊声明, 所有的最优解均指全局最优解.

英文引用格式: Huang L, Nie J W, Yuan Y-X. An introduction to polynomial optimization (in Chinese). Sci Sin Math, 2025, 55: 481–500, doi: 10.1360/SSM-2024-0129

一列半正定规划松弛构成. 在阿基米德性条件 (Archimedean condition) 下, Lasserre [12] 证明了矩 - 平方和松弛具有渐近收敛性. 但在实际计算中, 矩 - 平方和松弛常常具有有限收敛性. 这一现象在近期被 Nie 等 [31] 所揭示. Nie [20] 证明了, 若线性无关约束规范性条件、严格互补松弛性条件和二阶充分性条件在问题的每个最优解处都成立, 则矩 - 平方和松弛具有有限收敛性. 在一些特殊的情形下, 其有限收敛性也得到了证明, 如严格凸或者平方和 - 凸 [13]、等式约束的可行集是有限集等. 最为有效的验证矩 - 平方和松弛有限收敛性和提取 (1.1) 最优解的准则是平滑截断 (flat truncation). Nie [18] 证明了在一般性的假设下, 若矩 - 平方和松弛具有有限收敛性, 则当松弛阶充分大时, 矩松弛的每个最优解都具有平滑截断. 这一结论近期在文献 [11] 中被改进为在标准的最优性条件下, 当松弛阶充分大时, 矩松弛的每个最优解都具有平滑截断.

当最优性条件 (如严格互补条件和二阶充分性条件) 得不到满足时, 标准的矩 - 平方和松弛通常不具有有限收敛性. 文献 [23] 在约束多项式是非奇异的假设下, 利用一阶最优性条件构造了 Lagrange 乘子变量的多项式表示, 进而构造了具有紧性的矩 - 平方和松弛. 多项式优化也被成功地应用于全局求解广义 Nash 均衡问题 [27,28]、双层规划 [30,31]、鞍点问题 [35]、特征值优化 [1,9,21,22,29,33] 和矩阵优化 [5,8,17,39] 等.

本文余下内容的安排如下. 第 2 节介绍实代数几何和多项式优化的基本记号以及最优性条件. 第 3 节介绍无约束多项式优化的矩 - 平方和松弛以及高阶松弛. 第 4 节介绍约束多项式优化的矩 - 平方和松弛以及有限收敛性理论. 第 5 节介绍如何利用 Lagrange 乘子表示构造具有紧性的矩 - 平方和松弛. 第 6 节给出总结与展望.

2 预备知识

2.1 基本符号

符号 $\mathbb{N}$ ($\mathbb{R}$, $\mathbb{C}$) 记为所有非负整数 (实数, 复数) 的集合. $\lceil t \rceil$ 记为大于等于实数 $t \in \mathbb{R}$ 的最小整数. 对正整数 m, 记 $[m] := \{1, \dots, m\}$. 向量 v 的标准 Euclid 范数记为 $\|v\|$, A^{T} 为矩阵 A 的转置. $A \succeq 0$ 表示矩阵 A 是对称半正定的. 对变量 $x := (x_1, \dots, x_n)$ 和幂 $\alpha = (\alpha_1, \dots, \alpha_n)$, 记

$$x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad |\alpha| := \alpha_1 + \cdots + \alpha_n.$$

记 $\mathbb{R}[x] := \mathbb{R}[x_1, \dots, x_n]$ 为关于变量 x 的实系数多项式环. $\mathbb{R}[x]_d$ 为 $\mathbb{R}[x]$ 中所有次数小于或等于 d 的全体多项式构成的子集, $[x]_d$ 表示所有次数小于等于 d 的单项式构成的向量, 在加权字典序下可以表示为

$$[x]_d^{\mathrm{T}} = [1, x_1, x_2, \dots, x_1^2, x_1x_2, \dots, x_1^d, x_1^{d-1}x_2, \dots, x_n^d].$$

记单项式幂集 $\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n : |\alpha| \leqslant d\}$. 记 $\deg(p)$ 为多项式 $p \in \mathbb{R}[x]$ 的总次数. 设 p 为齐次多项式, 若 $p(x) > 0, \forall x \neq 0$, 则称 p 是正定的. 若 p 为关于变量 x 的函数, 则 ∇p 和 $\nabla^2 p$ 分别代表 p 关于 x 的梯度和 Hessian 矩阵.

2.2 理想、二次模和矩矩阵

本小节介绍多项式优化和实代数几何的基本概念, 详情可参见文献 [7,10,14,24]. 一个子集 $I \subseteq \mathbb{R}[x]$ 称为 $\mathbb{R}[x]$ 的一个理想, 如果 I 满足

$$I \cdot \mathbb{R}[x] \subseteq I, \quad I + I \subseteq I.$$

对一个多项式组 $h := (h_1, \ldots, h_m)$, Ideal$[h]$ 记为 h 生成的理想, 即

$$\mathrm{Ideal}[h] = h_1 \cdot \mathbb{R}[x] + \cdots + h_m \cdot \mathbb{R}[x].$$

对一个次数 k, 理想 Ideal$[h]$ 的第 k 阶截断为

$$\mathrm{Ideal}[h]_k = h_1 \cdot \mathbb{R}[x]_{k-\deg(h_1)} + \cdots + h_m \cdot \mathbb{R}[x]_{k-\deg(h_m)}.$$

理想 Ideal$[h]$ 所定义的实代数簇 (real variety) 为

$$V_{\mathbb{R}}(h) = \{x \in \mathbb{R}^n \mid h_1(x) = \cdots = h_m(x) = 0\}.$$

其所定义的复代数簇 (complex variety) 为

$$V_{\mathbb{C}}(h) = \{x \in \mathbb{C}^n \mid h_1(x) = \cdots = h_m(x) = 0\}.$$

理想 Ideal$[h]$ 被称为实根理想 (real radical), 若它满足

$$\mathrm{Ideal}[h] = \mathrm{Ideal}[V_{\mathbb{R}}(h)],$$

这里 Ideal$[V_{\mathbb{R}}(h)]$ 为在集合 $V_{\mathbb{R}}(h)$ 上都为零的多项式全体.

多项式 p 称为平方和多项式 (sum of squares, SOS), 如果存在 $p_1, \ldots, p_t \in \mathbb{R}[x]$ 使得 $p = p_1^2 + \cdots + p_t^2$. 所有平方和多项式构成的集合记为 $\Sigma[x]$. 对一个次数 k, 记截断

$$\Sigma[x]_k := \Sigma[x] \cap \mathbb{R}[x]_k.$$

一个多项式组 $g = (g_1, \ldots, g_\ell)$ 所生成的二次模 (quadratic module, QM) 为

$$\mathrm{QM}[g] := \Sigma[x] + g_1 \cdot \Sigma[x] + \cdots + g_\ell \cdot \Sigma[x]. \tag{2.1}$$

类似地, 对一个次数 k, 二次模 QM$[g]$ 的第 k 阶截断为

$$\mathrm{QM}[g]_k = \Sigma[x]_k + g_1 \cdot \Sigma[x]_{k-2\lceil \deg(g_1)/2 \rceil} + \cdots + g_\ell \cdot \Sigma[x]_{k-2\lceil \deg(g_\ell)/2 \rceil}. \tag{2.2}$$

集合 Ideal$[h]$ + QM$[g]$ 被称为阿基米德的, 如果存在 $R > 0$ 使得 $R - \|x\|^2 \in \mathrm{Ideal}[h] + \mathrm{QM}[g]$. 当 Ideal$[h]$ + QM$[g]$ 为阿基米德的时, 集合

$$K := \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \geqslant 0\}$$

一定是紧集. 这是因为对于每个 $x \in K$, 都有 $R - \|x\|^2 \geqslant 0$. 集合 K 称为实半代数集. 显然, 如果多项式 $p \in \mathrm{Ideal}[h] + \mathrm{QM}[g]$, 则 p 在 K 上非负. 反之, 则不一定是正确的, 例如, 当 $f = x_1^4x_2^2 + x_1^2x_2^4 + 1 - 3x_1^2x_2^2$ 为 Motzkin 多项式且 $K = \mathbb{R}^2$ 时 (参见文献 [38]). 但是, 如果 p 在 K 上是正的且集合 Ideal$[h]$ + QM$[g]$ 是阿基米德的, 则一定有 $p \in \mathrm{Ideal}[h] + \mathrm{QM}[g]$. 这个结论也被称为 Putinar 正多项式表示定理 (Putinar's Positivstellensatz) (参见文献 [37]).

对一个正整数 k, 记号 $\mathbb{R}^{\mathbb{N}^n_{2k}}$ 代表全体幂集为 $\alpha \in \mathbb{N}^n_{2k}$ 的实向量所构成的线性空间, 即每个 $y \in \mathbb{R}^{\mathbb{N}^n_{2k}}$ 都可以表示如下:

$$y = (y_\alpha)_{\alpha \in \mathbb{N}^n_{2k}}.$$

这样的 y 被称为 $2k$ 次截断多元序列 (truncated multi-sequences, tms). 对于 $t \leqslant 2k$, $y|_t$ 记为截断子序列:

$$y|_t = (y_\alpha)_{\alpha \in \mathbb{N}^n_t}.$$

对多项式 $p = \sum_{\alpha \in \mathbb{N}^n_{2k}} p_\alpha x^\alpha \in \mathbb{R}[x]_{2k}$, 定义双线性算子

$$\langle p, y \rangle = \sum_{\alpha \in \mathbb{N}^n_{2k}} p_\alpha y_\alpha. \tag{2.3}$$

当多项式 p 给定时, $\langle p, y \rangle$ 是关于 y 的一个线性泛函. 对次数 $t := k - \lceil \deg(p)/2 \rceil$, p 的 k 阶局部化矩阵 (k-th order localizing matrix of p) $L_p^{(k)}[y]$ 是一个关于 y 的对称线性矩阵泛函, 且使得对所有 $q \in \mathbb{R}[x]_t$, 下式成立:

$$\mathrm{vec}(q)^{\mathrm{T}}(L_p^{(k)}[y])\mathrm{vec}(q) = \langle pq^2, y \rangle, \tag{2.4}$$

其中 $\mathrm{vec}(q)$ 是多项式 q 的系数向量. 例如, 当 $n = k = 2$ 和 $p = x_1 - x_1 x_2$ 时, 矩阵 $L_p^{(2)}[y]$ 为

$$L_p^{(2)}[y] = \begin{bmatrix} y_{10} - y_{11} & y_{20} - y_{21} & y_{11} - y_{12} \\ y_{20} - y_{21} & y_{30} - y_{31} & y_{21} - y_{22} \\ y_{11} - y_{12} & y_{21} - y_{22} & y_{12} - y_{13} \end{bmatrix}.$$

特别地, 如果 $p = 1$, 则 $L_1^{(k)}[y]$ 也称为 y 的 k 阶矩量矩阵 (k-th order moment matrix of y), 相应地记为

$$M_k[y] := L_1^{(k)}[y].$$

例如, 当 $n = 2$, $k = 2$ 时, 矩阵 $M_2[y]$ 的构成如下:

$$M_2[y] = \begin{bmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix}.$$

对多项式 p, 令 $\mathscr{V}_p^{(2k)}[y]$ 为使得

$$\langle qp, y \rangle = (\mathscr{V}_p^{(2k)}[y])^{\mathrm{T}} \mathrm{vec} q, \quad \forall\, q \in \mathbb{R}[x]_{2k-\deg(p)} \tag{2.5}$$

的向量, 称 $\mathscr{V}_p^{(2k)}[y]$ 为 p 的局部化向量 (localizing vector of p). 例如, 当 $p = x_1^3 + x_2^3 - 1$, $k = 2$ 时, 向量

$$\mathscr{V}_p^{(4)}[y] = [y_{30} + y_{03} - y_{00} \;\; y_{40} + y_{13} - y_{10} \quad y_{31} + y_{04} - y_{01}]^{\mathrm{T}}.$$

集合 $\mathrm{Ideal}[h]_{2k} + \mathrm{QM}[g]_{2k}$ 的对偶锥为

$$\left\{ y \in \mathbb{R}^{\mathbb{N}^n_{2k}} \;\middle|\; \begin{array}{l} \mathscr{V}_{h_i}^{(2k)}[y] = 0 \ (i = 1, \ldots, m), \\ M_k[y] \succeq 0, L_{g_j}^{(k)}[y] \succeq 0 \ (j = 1, \ldots, \ell) \end{array} \right\}.$$

2.3 最优性条件

本小节回顾非线性优化中的最优性条件, 可参见文献 [40]. 假设 u 为 (1.1) 的局部最优解. 在 u 处, 积极不等式约束的指标集记为

$$J(u) := \{j \in \mathcal{I} \mid c_j(u) = 0\}. \tag{2.6}$$

称线性无关约束规范性条件 (linear independence constraint qualification condition, LICQC) 在 u 处满足, 若梯度集合 $\{\nabla c_i(u)\}_{i\in\mathcal{E}\cup J(u)}$ 是线性无关的. 在 LICQC 下, 存在 Lagrange 乘子向量

$$\lambda := (\lambda_i)_{i\in\mathcal{E}} \cup (\lambda_j)_{j\in\mathcal{I}}$$

满足 Karush-Kuhn-Tucker (KKT) 条件:

$$\nabla f(u) = \sum_{i\in\mathcal{E}} \lambda_i \nabla c_i(u) + \sum_{j\in\mathcal{I}} \lambda_j \nabla c_j(u), \tag{2.7}$$

$$0 \leqslant c_j(u) \perp \lambda_j \geqslant 0, \quad \forall\, j \in \mathcal{I}. \tag{2.8}$$

(对 $a, b \in \mathbb{R}$, 记 $a \perp b$ 表示 $a \cdot b = 0$.) 称方程 (2.7) 为一阶最优性条件 (first order optimality condition, FOOC), 称 (2.8) 为互补条件. 满足 (2.7) 和 (2.8) 的可行点称为 KKT 点. 进一步地, 若

$$0 \leqslant c_j(u) \perp \lambda_j \geqslant 0, \quad \lambda_j + c_j(u) > 0, \quad \forall\, j \in \mathcal{I}, \tag{2.9}$$

则称严格互补条件 (strict complementarity condition, SCC) 在 u 处成立. 对于 (2.7) 和 (2.8) 中的 Lagrange 乘子 λ_i, 相应的 Lagrange 函数为

$$L(x) := f(x) - \sum_{i\in\mathcal{E}} \lambda_i c_i(x) - \sum_{j\in\mathcal{I}} \lambda_j c_j(x). \tag{2.10}$$

记 Lagrange 函数的 Hessian 矩阵为

$$\nabla^2 L(x) := \nabla^2 f(x) - \sum_{i\in\mathcal{E}} \lambda_i \nabla^2 c_i(x) - \sum_{j\in\mathcal{I}} \lambda_j \nabla^2 c_j(x). \tag{2.11}$$

若 LICQC 在 u 处成立, 则二阶必要性条件 (second order necessary condition) 在 u 处成立:

$$v^{\mathrm{T}}(\nabla^2 L(u))v \geqslant 0, \quad \forall\, v \in \bigcap_{i\in\mathcal{E}\cup J(u)} \nabla c_i(u)^{\perp}, \tag{2.12}$$

其中 $\nabla c_i(u)^{\perp}$ 记为其正交补:

$$\nabla c_i(u)^{\perp} := \{v \in \mathbb{R}^n \mid c_i(u)^{\mathrm{T}} v = 0\}.$$

进一步地, 若

$$v^{\mathrm{T}}(\nabla^2 L(u))v > 0, \quad \forall\, 0 \neq v \in \bigcap_{i\in\mathcal{E}\cup J(u)} \nabla c_i(u)^{\perp}, \tag{2.13}$$

则称二阶充分性条件 (second order sufficient condition, SOSC) 在 u 处成立.

3　无约束多项式优化

无约束多项式优化是一类基本的优化问题. 本节介绍无约束多项式优化的矩 - 平方和松弛算法以及如何获取全局最优解. 无约束多项式优化可表示成如下形式:

$$\min_{x \in \mathbb{R}^n} \; f(x), \tag{3.1}$$

其中目标函数 $f(x)$ 为多项式. 注意到, 若问题 (3.1) 有下界, 即最优值 $f_{\min} > -\infty$, 则多项式 f 的次数一定是偶数. 否则, f 在 $\mathbb{R}^n$ 上一定是下无界的. 假设 f 的次数为 $2d$. 优化问题 (3.1) 有着非常广阔的应用, 如非线性最小二乘问题、最佳秩一张量逼近核传感器网络选址等. 当 f 的次数大于等于 4 时, 求解无约束多项式优化 (3.1) 是一个 NP (non-deterministic polynomial)- 难的问题 [16]. 以下假设最优值 $f_{\min}$ 有限且可以取到, 即 $f(x) \geqslant f_{\min}$ $(\forall\, x \in \mathbb{R}^n)$ 且存在 $u \in \mathbb{R}^n$ 使得 $f(u) = f_{\min}$.

3.1　标准的矩 - 平方和松弛

假设 $f \in \mathbb{R}[x]_{2d}$. 实数 γ 是 f 的一个下界当且仅当多项式 $f(x) - \gamma$ 是一个非负多项式, 即对任意的 $x \in \mathbb{R}^n$, 都有 $f(x) - \gamma \geqslant 0$. 将非负条件替换成平方和多项式, 则可以得到经典的 SOS 松弛 [12, 36]. 求解 (3.1) 的标准 SOS 松弛为

$$\begin{cases} \max \;\gamma \\ \text{s.t. } f - \gamma \in \Sigma_{n,2d}. \end{cases} \tag{3.2}$$

优化 (3.2) 的对偶问题为

$$\begin{cases} \min \;\langle f, y \rangle \\ \text{s.t. } M_d[y] \succeq 0, \\ \qquad y_0 = 1, \quad y \in \mathbb{R}^{\mathbb{N}^n_{2d}}. \end{cases} \tag{3.3}$$

问题 (3.3) 称为 (3.1) 的矩松弛. 称原始对偶对 (3.2) 和 (3.3) 为标准的矩 - 平方和松弛.

令 f_{sos} 和 f_{mom} 分别表示 (3.2) 和 (3.3) 的最优值. 容易验证原始对偶对 (3.2) 和 (3.3) 无对偶间隙, 即 $f_{\text{sos}} = f_{\text{mom}}$. 若 $f_{\text{sos}} = f_{\min}$, 则称矩 - 平方和松弛 (3.2) 和 (3.3) 为紧的. 注意到 (3.3) 存在内点, 因此强对偶性成立. 进一步地, 若 γ 是 (3.2) 的可行点, 则有 $f_{\min} \geqslant \gamma$. 因此, 下面的性质是直接的.

性质 3.1 [12, 14]　令 $f_{\min}, f_{\text{sos}}$ 和 f_{mom} 为如上不同优化问题的最优值, 则 $f_{\text{sos}} = f_{\text{mom}} \leqslant f_{\min}$. 进一步地, 若 $f_{\text{sos}} > -\infty$, 则 (3.2) 有最优解.

3.2　最优解的提取

通过求解矩 - 平方和松弛 (3.2) 和 (3.3), 可以得到下界 f_{sos} 和 f_{mom}. 最优值 $f_{\min}$ 通常是不知道的. 下面介绍如何判别矩 - 平方和松弛的紧性, 并讨论最优解的提取. 一个探测矩 - 平方和松弛的紧性和提取最优值的常用方法是平滑扩张条件 (flat extension condition) [2] 或者平滑截断条件 (flat truncation condition) [18]. 假设 y^* 是矩松弛 (3.3) 的一个最优解.

3.2.1　平滑扩张

为了提取最优解, 通常假设如下平滑扩张条件成立: 存在非负整数 $t < d$ 使得

$$\operatorname{rank} M_t[y^*] = \operatorname{rank} M_d[y^*]. \tag{3.4}$$

下面是一个经典结果.

定理 3.2 [2, 4, 18] 假设 y^* 是矩松弛 (3.3) 的最优解且秩条件 (3.4) 成立. 令 $r := \operatorname{rank} M_d[y^*]$, 则存在 r 个不同的点 $u_1, \ldots, u_r \in \mathbb{R}^n$ 和正整数 $\lambda_1, \ldots, \lambda_r > 0$ 使得

$$y^* = \lambda_1 [u_1]_{2d} + \cdots + \lambda_r [u_r]_{2d}, \tag{3.5}$$

并且每个 u_i 是 (3.1) 的最优解. 进一步地, 矩 - 平方和松弛是紧的, 即 $f_{\min} = f_{\text{sos}} = f_{\text{mom}}$.

3.2.2 平滑截断

秩条件 (3.4) 是矩 - 平方和 (3.2) 和 (3.3) 松弛紧性的充分条件, 但却不一定必要的. 例如, 当

$$f = (x_1x_2 - 1)^2 + (x_1x_3 - 1)^2 + (x_2x_3 - 1)^2$$

时, 有 $f_{\text{sos}} = f_{\min} = 0$, 但并不是 (3.3) 的所有最优解都满足 (3.4) (参见文献 [18]). 一般来讲, 若秩条件 (3.4) 成立且 $\operatorname{rank} M_d[y^*]$ 是最大时, 则 f 只有有限多个最优解 [14]. 但即使 f 只有有限多个最优解, 条件 (3.4) 也不一定成立. 然而, 当矩 - 平方和松弛是紧的时, 如果用更高阶的松弛, 秩条件 (3.4) 最终会成立.

对整数 $k \geqslant d$ ($2d$ 是 f 的次数), 考虑第 k 阶平方和松弛

$$\begin{cases} \max \ \gamma \\ \text{s.t. } f - \gamma \in \Sigma_{n,2k}, \end{cases} \tag{3.6}$$

其对偶优化问题为

$$\begin{cases} \min \ \langle f, y \rangle \\ \text{s.t. } M_k[y] \succeq 0, \\ \qquad y_0 = 1, \quad y \in \mathbb{R}^{\mathbb{N}^n_{2k}}. \end{cases} \tag{3.7}$$

由于 $f - \gamma \in \Sigma_{n,2k}$ 当且仅当 $f - \gamma \in \Sigma_{n,2d}$, 松弛 (3.6) 等价于 (3.2), 因此松弛 (3.6) 的最优值也等于 f_{sos}. 我们同样可以证明 (3.7) 的最优值也为 f_{mom}. 但是, 从提取最优解的角度来看, (3.7) 和 (3.3) 却是不等价的.

定义 3.3 [18] 称 (3.7) 的最优解 $y^* \in \mathbb{R}^{\mathbb{N}^n_{2k}}$ 满足平滑截断, 如果存在整数 $t_2 > t_1 \geqslant 0$ 使得

$$\operatorname{rank} M_{t_1}[y^*] = \operatorname{rank} M_{t_2}[y^*], \quad d \leqslant t_2 \leqslant k. \tag{3.8}$$

定理 3.4 [18] 假设 y^* 是 (3.7) 的最优解. 若 y^* 满足平滑截断 (3.8), 则存在 $r := \operatorname{rank} M_{t_2}[y^*]$ 个不同的点 $u_1, \ldots, u_r \in \mathbb{R}^n$ 和正实数 $\lambda_1, \ldots, \lambda_r > 0$ 使得

$$y^*|_{2t_2} \ = \ \lambda_1 [u_1]_{2t_2} + \cdots + \lambda_r [u_r]_{2t_2}, \tag{3.9}$$

且每个 u_i 都是 (3.1) 的最优解. 进一步地, 矩 - 平方和松弛是紧的, 即 $f_{\text{sos}} = f_{\text{mom}} = f_{\min}$.

当矩 - 平方和松弛是紧的时, 下面的定理证明如果 f 只有有限多个最优解, 则平滑截断 (3.8) 一定会被满足.

定理 3.5 [18] 若 $f_{\min} = f_{\text{sos}}$ 且 (3.1) 的最优解只有有限多个, 则松弛 (3.7) 可以取得最优值且当松弛阶 k 充分大时, (3.7) 的每个最优解 y^* 都满足平滑截断.

3.3 高阶松弛

由于并非所有的非负多项式都是 SOS (如 Motzkin 多项式 $x_1^4x_2^2+x_1^2x_2^4+1-3x_1^2x_2^2$、Robinson 多项式 $x_1^6+x_2^6+3x_1^2x_2^2-x_1^4(x_2^2+1)-x_2^4(x_1^2+1)-(x_1^2+x_2^2)$ 和 Choi-Lam 多项式 $x_1^4x_2^2+x_2^4+x_1^2-3x_1^2x_2^2$), 矩 - 平方和松弛 (3.2) 和 (3.3) 可能不能够求得 (3.1) 的全局最优值, 即 $f_{\text{sos}} < f_{\min}$. 然而, Artin 对 Hilbert 第 17 问题的证明告诉我们: 对每个非负多项式 p, 都存在非零平方和多项式 σ 使得 $\sigma \cdot p$ 是 SOS. 特别地, 若 p 是一个对称正定的齐次多项式, 则可以选 $\sigma = (x^{\mathrm{T}}x)^k$. 当 p 非齐次时, 可以选 $\sigma = (1+x^{\mathrm{T}}x)^k$. 这启发我们使用带分母形式的高阶矩 - 平方和松弛.

对整数 $k \geqslant 0$, γ 是 f 的下界当且仅当 $(1+x^{\mathrm{T}}x)^k(f-\gamma)$ 是非负多项式. 同样地, 将非负性用平方和多项式表示, 可以得到第 k 阶平方和松弛

$$\begin{cases} \max \ \gamma \\ \text{s.t. } (1+x^{\mathrm{T}}x)^k(f-\gamma) \in \Sigma_{n,2d+2k}. \end{cases} \tag{3.10}$$

记 f_k 为 (3.10) 的最优值. 显然, $f_k \leqslant f_{\min}$ 且序列 $\{f_k\}$ 是单调递增的, 即

$$f_0 \leqslant \cdots \leqslant f_k \leqslant \cdots \leqslant f_{\min}.$$

当 $k=0$ 时, 松弛 (3.10) 和 (3.2) 相同, 故 $f_0 = f_{\text{sos}}$. 优化 (3.10) 的对偶问题为 (参见文献 [24])

$$\begin{cases} \min \ \langle (1+x^{\mathrm{T}}x)^k f, y\rangle \\ \text{s.t. } \langle (1+x^{\mathrm{T}}x)^k, y\rangle = 1, \\ \qquad M_{k+d}[y] \succeq 0, \\ \qquad y \in \mathbb{R}^{\mathbb{N}^n_{2d+2k}}. \end{cases} \tag{3.11}$$

由强对偶定理可得 (3.11) 的最优值等于 f_k. 原始对偶对 (3.10) 和 (3.11) 被称为求解 (3.1) 的第 k 阶矩 - 平方和松弛. 称该松弛为紧的, 如果存在某个 k 使得 $f_k = f_{\min}$. 下面是渐近收敛性结果.

定理 3.6 假设 $f \in \mathbb{R}[x]_{2d}$, 记 f^{hom} 为 f 中的次数为 $2d$ 的部分. 若 f^{hom} 是对称正定的, 则 $f_{\min} > -\infty$ 且

$$\lim_{k\to\infty} f_k = f_{\min}.$$

当 f^{hom} 不是对称正定时, 收敛性 $f_k \to f_{\min}$ 是无法保证的. 假设 y^* 是矩松弛问题 (3.11) 的一个最优解. 为了提取 (3.1) 的最优解, 考虑如下秩条件: 存在整数 $t \in [0, d+k)$ 使得

$$\operatorname{rank} M_t[y^*] = \operatorname{rank} M_{d+k}[y^*]. \tag{3.12}$$

当 (3.12) 成立时, 一定有 $f_k = f_{\min}$ 且能够得到 (3.1) 的若干个最优解.

定理 3.7 [18] 假设 y^* 是矩松弛 (3.11) 的一个最优解. 如果秩条件 (3.12) 成立, 则存在 $r := \operatorname{rank} M_t[y^*]$ 个不同的点 $u_1, \ldots, u_r \in \mathbb{R}^n$ 和实数 $\lambda_1, \ldots, \lambda_r > 0$ 使得

$$y^*|_{2d+2k} = \lambda_1[u_1]_{2d+2k} + \cdots + \lambda_r[u_r]_{2d+2k}.$$

进一步地, 每个 u_i 都是 (3.1) 的最优解且 $f_k = f_{\min} = f(u_i)$.

4　约束多项式优化

本节讨论约束多项式优化问题. 矩 - 平方和松弛是求解约束多项式优化的基本方法. 本节介绍其渐近收敛性和有限收敛性以及最优解的提取等相关议题. 记约束多项式组

$$c_{\mathrm{eq}} := (c_i)_{i\in\mathcal{E}}, \quad c_{\mathrm{in}} := (c_j)_{j\in\mathcal{I}}. \tag{4.1}$$

一个实数 γ 是 $f_{\min}$ 的下界 (即 $\gamma \leqslant f_{\min}$) 当且仅当多项式 $f-\gamma$ 在 K 上非负. 确保非负性的一个充分性条件是

$$f-\gamma \in \mathrm{Ideal}[c_{\mathrm{eq}}] + \mathrm{QM}[c_{\mathrm{in}}],$$

其中 $\mathrm{Ideal}[c_{\mathrm{eq}}]$ 是 c_{eq} 生成的理想, $\mathrm{QM}[c_{\mathrm{in}}]$ 是 c_{in} 生成的二次模.

问题 (1.1) 的第 k 阶平方和松弛如下:

$$\begin{cases} \max\ \gamma \\ \text{s.t. } f(x)-\gamma \in \mathrm{Ideal}[c_{\mathrm{eq}}]_{2k} + \mathrm{QM}[c_{\mathrm{in}}]_{2k}. \end{cases} \tag{4.2}$$

由于集合 $\mathrm{Ideal}[c_{\mathrm{eq}}]_{2k} + \mathrm{QM}[c_{\mathrm{in}}]_{2k}$ 可以表示为半正定规划的可行集, 因此问题 (4.2) 可以转化成一个半正定规划问题. 这里不详细地写出其推导过程, 感兴趣的读者可参见文献 [14,24].

问题 (4.2) 的对偶问题为第 k 阶矩松弛

$$\begin{cases} \min\ \langle f, y\rangle \\ \text{s.t. } \mathscr{V}_{c_i}^{(2k)}[y] = 0\ (i\in\mathcal{E}), \\ \qquad L_{c_j}^{(k)}[y] \succeq 0\ (j\in\mathcal{I}), \\ \qquad M_k[y] \succeq 0, \\ \qquad y_0 = 1, \quad y\in\mathbb{R}^{\mathbb{N}_{2k}^n}. \end{cases} \tag{4.3}$$

松弛 (4.2) 和 (4.3) 构成了经典的矩 - 平方和松弛算法. 记 $f_{\mathrm{sos},k}$ 和 $f_{\mathrm{mom},k}$ 分别为 (4.2) 和 (4.3) 的最优值. 以下是矩 - 平方和松弛的基本性质.

性质 4.1　对每个松弛阶 k, 有

$$f_{\mathrm{sos},k} \leqslant f_{\mathrm{mom},k} \leqslant f_{\min}. \tag{4.4}$$

此外, $\{f_{\mathrm{sos},k}\}_{k=1}^{\infty}$ 和 $\{f_{\mathrm{mom},k}\}_{k=1}^{\infty}$ 都是单调增加的.

(i) 假设所有的等式约束优化是线性的或者 $\mathcal{E}=\emptyset$. 若存在 $v\in K$ 使得 $c_j(v)>0$ $(j\in\mathcal{I})$, 则 $f_{\mathrm{sos},k} = f_{\mathrm{mom},k}$ 且松弛 (4.2) 的最优值是可达的.

(ii) 假设存在 $R>0$ 使得 $R-\|x\|^{2k_0} \in \mathrm{Ideal}[c_{\mathrm{eq}}]_{2k_0} + \mathrm{QM}[c_{\mathrm{in}}]_{2k_0}$, 其中 $k_0\in\mathbb{N}$, 则对所有 $k\geqslant k_0$, 矩松弛 (4.3) 是可行的, 且最优值可达的.

矩 - 平方和松弛 (4.2) 和 (4.3) 称为紧的 (或者说具有有限收敛性), 若对某个松弛阶 k, 有 $f_{\mathrm{sos},k} = f_{\min}$. 若当 $k\to\infty$ 时, 有 $f_{\mathrm{sos},k}\to f_{\min}$, 则称矩 - 平方和松弛具有渐近收敛性. 下面是关于矩 - 平方和松弛的渐近收敛性的一个经典结果, 需要阿基米德假设.

定理 4.2 [12]　若集合 $\mathrm{Ideal}[c_{\mathrm{eq}}] + \mathrm{QM}[c_{\mathrm{in}}]$ 是阿基米德的, 则 $f_{\mathrm{sos},k}\to f_{\min}$.

将多项式优化问题的矩 - 平方和松弛建模成半正定规划问题的常用软件为 `GloptiPoly3`[6] 和 `YALMIP`[15]. 更加深入和系统的介绍, 可参见文献 [14,24].

定理 4.2 只是保证了矩 - 平方和松弛的渐近收敛性, 但是在实际计算中, 该方法往往都具有有限收敛性. 本节介绍矩 - 平方和松弛有限收敛性理论的主要结果, 其中包括最优值的有限收敛性以及能够在有限步内获得最优解.

4.1 最优解的提取

矩 - 平方和算法的另一个优势是, 我们常常可以从矩量矩阵 $M_k[w^*]$ 的信息中提取出 (1.1) 的最优解, 从而能够验证全局最优性. 这是非凸优化的绝大部分算法所不具备的. 在计算中, 检验有限收敛性的一个方便的判据是平滑截断[2,18]. 假设 $w^* \in \mathbb{R}^{\mathbb{N}^n_{2k}}$ 是松弛 (4.3) 的一个最优解. 记

$$\boxed{\begin{aligned} &d_c := \max_{i\in\mathcal{E}\cup\mathcal{I}}\left\{1, \left\lceil\frac{1}{2}\deg(c_i)\right\rceil\right\}, \\ &d_0 := \max\left\{d_c, \left\lceil\frac{1}{2}\deg(f)\right\rceil\right\}. \end{aligned}} \tag{4.5}$$

如果存在整数 $t \in [d_c, k]$ 使得

$$\operatorname{rank} M_{t-d_c}[w^*|_{2t}] = \operatorname{rank} M_t[w^*|_{2t}], \tag{4.6}$$

则称截断多元序列 w^* 具有平滑截断. 此时, 一定存在 $\lambda_1, \dots, \lambda_r > 0$, $u_1, \dots, u_r \in K$, 使得

$$\begin{aligned} &\lambda_1 + \cdots + \lambda_r = 1, \\ &w^*|_{2t} = \lambda_1[u_1]_t + \cdots + \lambda_r[u_r]_t. \end{aligned}$$

注意到 $\langle f, w^*\rangle = \lambda_1\langle f, [u_1]_t\rangle + \cdots + \lambda_r\langle f, [u_r]_t\rangle \geqslant f_{\min}$. 因此, 有 $f_{\mathrm{mom},k} = f_{\min}$ 且 $u_1, \dots, u_r$ 是 (1.1) 的最优解. 如何从矩量矩阵 $M_t(w^*)$ 中提取出 (1.1) 的最优解可参见文献 [14,24].

Nie[18] 首次给出了平滑截断会成立的一般性条件.

定理 4.3[18] 假设 (1.1) 只有有限多个最优解且存在 $\rho \in \mathrm{QM}[c_{\mathrm{in}}]$ 使得对每一个指标子集 $J \subseteq \mathcal{I}$ (其中 J 是 (1.1) 的某个最优解 u 处积极约束的指标集), 等式约束优化问题

$$\begin{cases} \min\ f(x) \\ \text{s.t. } c_i(x) = 0 \quad (i \in \mathcal{E} \cup J) \end{cases}$$

最多只有有限多个实的 KKT 点属于如下集合:

$$\Theta := \{x \in \mathbb{R}^n : f(x) = f_{\min}, c_{\mathrm{eq}}(x) = 0, \rho(x) \geqslant 0\}.$$

则当松弛阶 k 充分大时, 矩松弛 (4.3) 的每一个最优解都有平滑截断.

上述定理近期在文献 [11] 得到了进一步改进.

定理 4.4[11] 假设集合 $\mathrm{Ideal}[c_{\mathrm{eq}}] + \mathrm{QM}[c_{\mathrm{in}}]$ 是阿基米德的. 如果 LIQC, SCC 和 SOSC 在 (1.1) 的每一个最优解处都满足, 则当松弛阶 k 充分大时, 矩松弛 (4.3) 的每一个最优解都有平滑截断.

若 $\operatorname{rank} M_k[y^*]$ 是 (4.3) 的所有最优解中秩最大的, 则对 y^* 的每一个平滑截断 $y^*|_{2t}$, $\operatorname{rank} M_t[w^*|_{2t}]$ 等于 (1.1) 最优解的个数. (1.1) 的所有的最优解都可以从 $y^*|_{2t}$ 中得到. 但如果 (1.1) 有无穷多个最优解且 $\operatorname{rank} M_k[y^*]$ 是 (4.3) 最优解中秩最大的时, y^* 不可能有平滑截断. 我们总结如下.

性质 4.5 [14] 记 S 为 (1.1) 的最优解的集合, y^* 为矩松弛 (4.3) 的秩 $\operatorname{rank} M_k[y^*]$ 最大的最优解.

(i) 若 (4.6) ($t \in [d_0, k]$) 得到满足, 则 $\operatorname{rank} M_t[y^*] = |S|$, 这里 $|S|$ 是集合 S 的基数;

(ii) 若 S 是无限集, 则不存在 $t \in [d_0, k]$ 使得 (4.6) 成立.

当矩 - 平方和松弛没有有限收敛性时, 平滑截断 (4.6) 不可能得到满足. 然而, 平滑截断依然可以作为一个验证渐近收敛性和提取近似最优解的一个准则. 当 (1.1) 只有有限多个最优解时, 平滑截断 (4.6) 是渐近满足的. 由于矩松弛 (4.3) 可能不存在最优解, 我们需要考虑 (4.3) 的近似最优解. 称一个序列 $\{y^{(k)}\}_{k=1}^{\infty}$ 为渐近最优的, 若 $y^{(k)}$ 是 (4.3) 的可行点且

$$\lim_{k\to\infty} \langle f, y^{(k)} \rangle = \lim_{k\to\infty} f_{\mathrm{mom},k}.$$

定理 4.6 [18] 假设 $\mathrm{Ideal}[c_{\mathrm{eq}}] + \mathrm{QM}[c_{\mathrm{in}}]$ 是阿基米德的并且 (1.1) 的最优解集是非空有限集. 假设 $\{y^{(k)}\}_{k=1}^{\infty}$ 是 (4.3) 的渐近最优序列, 则当 $t \geqslant d_c + |S| - 1$ 时, 截断序列 $\{y^{(k)}|_{2t}\}_{k=1}^{\infty}$ 的每个极限点都是平滑的.

4.2 有限收敛性理论

矩 - 平方和松弛的有限收敛性和最优性条件存在着深刻的联系. 下面是 (1.1) 的矩 - 平方和松弛的有限收敛性的主要结果.

定理 4.7 [20] 假设集合 $\mathrm{Ideal}[c_{\mathrm{eq}}] + \mathrm{QM}[c_{\mathrm{in}}]$ 是阿基米德的. 若 LICQC, SCC 和 SOSC 在 (1.1) 的每一个最优解处都满足, 则矩 - 平方和松弛 (4.2) 和 (4.3) 具有有限收敛性, 即当松弛阶 k 充分大时, 有 $f_{\mathrm{sos},k} = f_{\mathrm{mom},k} = f_{\min}$.

例 4.8 [20] 考虑如下约束优化问题:

$$\begin{cases} \min\ x_1^6 + x_2^6 + x_3^6 + 3x_1^2x_2^2x_3^2 - x_1^4(x_2^2 + x_3^2) - x_2^4(x_3^2 + x_1^2) - x_3^4(x_1^2 + x_2^2) \\ \text{s.t. } x_1^2 + x_2^2 + x_3^2 = 1, \end{cases} \tag{4.7}$$

其目标函数 f 是 Robinson 多项式. 它是全局非负但不是平方和的多项式. 最优值 $f_{\min} = 0$ 且 (4.7) 的最优解为

$$\frac{1}{\sqrt{3}}(\pm1, \pm1, \pm1), \quad \frac{1}{\sqrt{2}}(\pm1, \pm1, 0), \quad \frac{1}{\sqrt{2}}(\pm1, 0, \pm1), \quad \frac{1}{\sqrt{2}}(0, \pm1, \pm1).$$

我们容易验证线性无关约束规范性条件和严格互补松弛条件在每个最优解处都成立. 实际上, 二阶充分性条件 (2.13) 也在所有最优解处都成立. 例如, 当 $u = \frac{1}{\sqrt{3}}(1,1,1)$ 时, Hessian 矩阵

$$\nabla^2 L(u) = \frac{4}{9}(3I_3 - uu^{\mathrm{T}})$$

在正交补空间 $u^{\perp}$ 上是正定的. 定理 4.7 告诉我们矩 - 平方和松弛对这个问题是有有限收敛的. 数值实验表明 $f_{\mathrm{sos},5} = f_{\min}$.

进一步地, Nie [20] 也证明了 LICQC, SCC 和 SOSC 实际上都是一般性的条件, 即这些条件在所出现的多项式空间的一个稠密开子集中都成立.

定理 4.9 [20] 设 d_0 和 d_i ($i \in \mathcal{E} \cup \mathcal{I}$) 为正整数, 则存在以 $f \in \mathbb{R}[x]_{d_0}$ 和 $c_i \in \mathbb{R}[x]_{d_i}$ ($i \in \mathcal{E} \cup \mathcal{I}$) 的系数为变量的有限多个非零多项式 $\phi_1, \ldots, \phi_L$ 使得: 若

$$\phi_1(f, c) \neq 0, \ldots, \phi_1(f, c) \neq 0,$$

则 LICQC, SCC 和 SOSC 在 (1.1) 的每个局部最优解处都满足.

定理 4.9 表明, 当 f 和 c_i $(i \in \mathcal{E} \cup \mathcal{I})$ 是一般性的多项式时, 矩 - 平方和松弛具有有限收敛性. 对有些例子, 即使有些最优性条件不成立, 矩 - 平方和松弛仍然具有有限收敛性. 然而, 条件 (2.7) 一般而言是有限收敛性所必须具有的.

性质 4.10 [20] 假设平方和松弛 (4.2) 的最优值可达. 若一阶最优性条件 (FOOC) (2.7) 在 (1.1) 的一个最优解处不成立, 则对所有的 k, 都有 $f_{\text{sos},k} < f_{\min}$.

4.3 有限集上的优化问题

4.3.1 有限实代数簇的情形

当等式约束的可行点是有限集时, 文献 [19] 证明了矩 - 平方和松弛具有有限收敛性. 记实代数簇

$$V_{\mathbb{R}}(c_{\text{eq}}) = \{x \in \mathbb{R}^n : c_i(x) = 0 \ (i \in \mathcal{E})\}.$$

定理 4.11 [19] 对多项式优化 (1.1), 如果实代数簇 $V_{\mathbb{R}}(c_{\text{eq}})$ 是有限的, 则矩 - 平方和松弛 (4.2) 和 (4.3) 具有有限收敛性, 即当松弛阶 k 充分大时, 有 $f_{\text{sos},k} = f_{\text{mom},k} = f_{\min}$. 进一步地, 若 $K \neq \emptyset$, 则矩松弛 (4.3) 有最优解且当 k 充分大时, 每个最优解 y^* 一定满足平滑截断 (4.6).

当 $V_{\mathbb{R}}(c_{\text{eq}})$ 是有限集时, 平方和松弛 (4.2) 可能在任何松弛阶 k 都达不到最优值 $f_{\text{sos},k}$. 例如, 考虑优化

$$\begin{cases} \min \ x_1 \\ \text{s.t.} \ x_1^2 + x_2^2 + \cdots + x_n^2 = 0. \end{cases}$$

我们知道对所有的 $k \geqslant 1$, 都有 $f_{\text{sos},k} = 0$. 然而, 对任意的 $\phi \in \mathbb{R}[x]$, 多项式 $\varphi = x_1 - (x_1^2 + x_2^2 + \cdots + x_n^2)\phi$ 都不是平方和的. 此时, 松弛 (4.2) 没有最优解. 然而, 当 $\text{Ideal}[c_{\text{eq}}]$ 是实根理想, 即 $\text{Ideal}[c_{\text{eq}}] = I(V_{\mathbb{R}}(c_{\text{eq}}))$, 且 k 充分大时, (4.2) 都取得到最优值. 当 $V_{\mathbb{R}}(c_{\text{eq}})$ 不是有限集时, 有如下性质.

性质 4.12 [19] 设 K 为 (1.1) 的可行集. 假设 $f_{\min}$ 有限且矩 - 平方和松弛 (4.2) 和 (4.3) 有有限收敛性. 若 $\text{Ideal}[c_{\text{eq}}] = I(K)$, 当 k 充分大时, (4.2) 都达到最优值.

例 4.13 [19] 考虑目标函数 $f = x_1 x_2$ 和约束条件

$$c_1(x) = (x_1^2 - 1)^2 + (x_2^2 - 1)^2 = 0, \quad c_2(x) = x_1 + x_2 - 1 \geqslant 0.$$

显然, $V_{\mathbb{R}}(c_1) = (\pm 1, \pm 1)$, $K = (1, 1)$, $f_{\min} = 1$. 令

$$a := \frac{1}{2}(x_1 + x_2 - 1)(x_1 - x_2)^2 \in \text{QM}[c_2]_4,$$
$$\hat{f} := f - 1 - a = \frac{1}{2}[(x_2^2 - 1)(x_1 - x_2 + 1) - (x_1^2 - 1)(x_1 - x_2 - 1)].$$

由于在 $V_{\mathbb{R}}(c_1)$ 上有 $\hat{f} \equiv 0$, 所以可得

$$\hat{f}^2 + q = \frac{1}{2}((x_1 - x_2)^2 + 1)c_1 \in \text{Ideal}[c_1]_6,$$

其中

$$q = \frac{1}{4}((x_1^2 - 1)(x_1 - x_2 + 1) + (x_2^2 - 1)(x_1 - x_2 - 1))^2.$$

对于任意 $\epsilon > 0$, 有 $\phi_\epsilon = -\frac{1}{4\epsilon}(\hat{f}^2 + q) \in \text{Ideal}[c_1]_6$ 且

$$\sigma_\epsilon = \epsilon\left(1 + \frac{\hat{f}}{2\epsilon}\right)^2 + \frac{1}{4\epsilon}q + a \in \text{QM}[c_2]_6.$$

对所有 $\epsilon > 0$, 有 $f - 1 + \epsilon = \phi_\epsilon + \sigma_\epsilon$. 因此, 当 $k \geqslant 3$ 时, $f_{\text{sos},k} = 1$.

4.3.2 有限实半代数集的情形

当实代数簇 $V_{\mathbb{R}}(c_{\mathrm{eq}})$ 不是有限集时, 可行解集合 K 仍然可能是有限的. 一个处理这种情形的自然技巧是引入新变量 $x_{n+1},\dots,x_{n+\ell}$, 其中 ℓ 是不等式约束的个数. 因此, K 可以等价地定义为

$$\begin{aligned} &c_i(x)=0 \quad (i\in\mathcal{E}),\\ &c_j(x)-x_{n+j}^2=0 \quad (j\in\mathcal{I}). \end{aligned}$$

显然, K 是有限集当且仅当上面的方程在空间 $\mathbb{R}^{n+\ell}$ 中只有有限多个实解. 当 K 有限时, 若使用上述新的定义方程, 由定理 4.11 可知, 矩 - 平方和松弛 (4.2) 和 (4.3) 具有有限收敛性. 但是新变量的引入会使得矩 - 平方和松弛的计算工作量更大.

当 K 是有限集时, 矩 - 平方和松弛 (4.2) 和 (4.3) 则在一个额外的假设下具有有限收敛性. 记理想

$$J:=(\mathrm{Ideal}[c_{\mathrm{eq}}]+\mathrm{QM}[c_{\mathrm{in}}])\cap-(\mathrm{Ideal}[c_{\mathrm{eq}}]+\mathrm{QM}[c_{\mathrm{in}}]). \tag{4.8}$$

理想 J 的维数定义为坐标环 (coordinate ring) $\frac{\mathbb{R}[x]}{J}$ 的维数, 即复代数簇 $V_{\mathbb{C}}(J)$ 的维数 (参见文献 [24]).

定理 4.14 [24] 假设 K 有限且 $\dim\frac{\mathbb{R}[x]}{J}\leqslant 1$, 则矩 - 平方和松弛 (4.2) 和 (4.3) 具有有限收敛性, 即当 k 充分大时, $f_{\mathrm{sos},k}=f_{\mathrm{mom},k}=f_{\min}$. 进一步地, 当 $K\neq\emptyset$ 时, 矩松弛 (4.3) 具有最优解且当 k 充分大时, 每个最优解 y^* 都满足平滑截断条件 (4.6).

5 Lagrange 乘子表示与紧松弛

经典的矩 - 平方和松弛对有些多项式优化问题可能不具有紧性. 本节引入紧松弛方法. 该方法基于最优性条件和 Lagrange 乘子表示. 在约束条件非奇异性的假设下, 本节给出新的具有有限收敛性的矩 - 平方和松弛序列, 并证明它对所有的目标函数都具有紧性.

假设 u 是 (1.1) 的一个局部最优解. 积极不等式约束的指标集为

$$J(u):=\{j\in\mathcal{I}\mid c_j(u)=0\}. \tag{5.1}$$

若线性无关约束规范性条件在 u 处成立, 则存在 Lagrange 乘子 λ_i $(i\in\mathcal{E}\cup\mathcal{I})$ 使得

$$\boxed{\begin{aligned} &\sum_{i\in\mathcal{E}}\lambda_i\nabla c_i(u)+\sum_{j\in\mathcal{I}}\lambda_j\nabla c_j(u)=\nabla f(u),\\ &c_j(u)\geqslant 0,\quad \lambda_j\geqslant 0,\quad \lambda_j c_j(u)=0 \quad (j\in\mathcal{I}). \end{aligned}} \tag{5.2}$$

记 Lagrange 乘子向量为

$$\lambda:=(\lambda_i)_{i\in\mathcal{E}\cup\mathcal{I}}.$$

考虑复数域上的驻点集合

$$\mathcal{K}:=\left\{(x,\lambda)\in\mathbb{C}^{n+|\mathcal{E}|+|\mathcal{I}|}\ \middle|\ \begin{array}{l}\sum\limits_{i\in\mathcal{E}\cup\mathcal{I}}\lambda_i\nabla c_i(x)=\nabla f(x),\\ c_i(x)=0\ (i\in\mathcal{E}),\\ \lambda_j c_j(x)=0\ (j\in\mathcal{I})\end{array}\right\}. \tag{5.3}$$

下面是一个很有用的结论.

定理 5.1 [3,25] 对所有的 $f, c_i, c_j \in \mathbb{C}[x]$ $(i \in \mathcal{E},\ j \in \mathcal{I})$, 目标函数值集合 $\{f(u) : (u, \lambda) \in \mathcal{K}\}$ 是一个有限集.

为了记号方便, 假设指标集 $\mathcal{E}$ 和 $\mathcal{I}$ 可写为 $\mathcal{E} \cup \mathcal{I} = \{1, \ldots, m\}$. 对变量 $(n_1, n_2, \ldots, n_k)$, 定义如下对称多项式:

$$\mathbf{S}_r(n_1, n_2, \ldots, n_k) = \sum_{i_1+i_2+\cdots+i_k=r} n_1^{i_1} \cdots n_k^{i_k}.$$

下面给出 (1.1) 的复数域上驻点个数的一个估计.

定理 5.2 [26] 假设 f 的次数是 d_0, c_i 的次数为 d_i. 若多项式 f 和 c_i 具有一般性的系数, 则 (1.1) 的复数域上的驻点数为

$$\sum_{\mathcal{E} \subseteq \{i_1, \ldots, i_k\} \subseteq \mathcal{E} \cup \mathcal{I}} d_{i_1} \cdot d_{i_2} \cdots d_{i_k} \cdot \mathbf{S}_{n-k}(d_0 - 1, d_{i_1} - 1, \ldots, d_{i_k} - 1).$$

由方程 (5.3) 可知, 若线性无关约束规范性条件成立, 则 Lagrange 乘子 λ_i 可以由梯度向量 $\nabla c_i(u)$ 唯一决定, 其中 $i \in \mathcal{E} \cup J(u)$. 记矩阵

$$G(x) := [\nabla c_1(x)\ \cdots\ \nabla c_m(x)].$$

若 $m \leqslant n$ 且 $\operatorname{rank} G(x) = m$, 则可以得到如下的有理表示:

$$\lambda = [G(x)^{\mathrm{T}} G(x)]^{-1} \cdot G(x)^{\mathrm{T}} \cdot \nabla f(x). \tag{5.4}$$

由于行列式 $\det[G(x)^{\mathrm{T}} G(x)]$ 通常次数很高, 所以矩阵逆 $[G(x)^{\mathrm{T}} G(x)]^{-1}$ 往往是难以计算的. 若 $m > n$, 则 $G(x)^{\mathrm{T}} G(x)$ 总是奇异的, 此时乘子表示 (5.4) 并不存在.

5.1 Lagrange 乘子表示

本小节展示如何获得 Lagrange 乘子表达. 假设总共有 m 个约束, 记

$$\mathcal{E} \cup \mathcal{I} := \{1, \ldots, m\}, \quad c := (c_1, \ldots, c_m), \quad \lambda := [\lambda_1\ \cdots\ \lambda_m]^{\mathrm{T}},$$

其中 λ 是 Lagrange 乘子向量. 若 (x, λ) 是驻点, 则 $\lambda_i c_i(x) = 0$, 其中 $i \in \mathcal{E} \cup \mathcal{I}$. 记

$$C(x) := \begin{bmatrix} \nabla c_1(x) & \nabla c_2(x) & \cdots & \nabla c_m(x) \\ c_1(x) & 0 & \cdots & 0 \\ 0 & c_2(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_m(x) \end{bmatrix}, \quad g(x) := \begin{bmatrix} \nabla f(x) \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{5.5}$$

因此, Lagrange 乘子向量 λ 满足如下线性方程:

$$C(x)\lambda = g(x). \tag{5.6}$$

注意到 $C(x)$ 是 $(m+n) \times m$ 的矩阵多项式. 若存在矩阵多项式 $L(x) \in \mathbb{R}[x]^{m \times (m+n)}$ 使得

$$L(x)C(x) = I_m, \tag{5.7}$$

则可以得到 $\lambda = L(x)C(x)\lambda = L(x)g(x)$. 因此, Lagrange 乘子向量 λ 可以表示成

$$\lambda = L(x)g(x) = L_1(x)\nabla f(x), \tag{5.8}$$

这里 $L_1(x)$ 是 $L(x)$ 的前 n 列所构成的子矩阵.

定义 5.3 若对所有的复数点 $x \in \mathbb{C}^n$, 有 $\operatorname{rank} C(x) = m$, 则称约束多项式组 $c = (c_1, \ldots, c_m)$ 是非奇异的.

实际上, 约束多项式组 c 是非奇异的当且仅当对所有的 $u \in \mathbb{C}^n$, 梯度 $\nabla c_i(u)$ 是线性无关的, 其中 $i \in \mathcal{E} \cup J(u)$, 这里 $J(u)$ 是 (5.1) 中积极不等式约束的指标集. 有趣的是, 矩阵多项式 $L(x)$ 的存在性等价于约束多项式的非奇异性.

性质 5.4 [23] 多项式组 c 非奇异当且仅当存在 $L(x) \in \mathbb{R}[x]^{m\times(m+n)}$ 满足 (5.7).

当多项式组 c 非奇异时, 以上性质告诉我们一定存在 $L(x) \in \mathbb{R}[x]^{m\times(m+n)}$ 满足 (5.7). 此时, 矩阵 $L(x)$ 可以通过求解线性方程组获得. 若矩阵 $L(x)$ 的次数小于等于 ℓ, 可以假设

$$L_{ij}(x) = \sum_{|\alpha|\leqslant \ell} l_{ij,\alpha} x^\alpha, \quad i \in [m], \quad j \in [m+n],$$

则 (5.7) 等价于

$$\sum_{k=1}^{m+n} \left(\sum_{|\alpha|\leqslant \ell} l_{ik,\alpha} x^\alpha \right) C_{kj} = \delta_{ij},$$

其中 δ_{ij} 为 Kronecker 记号. 以上等式意味着左右两边多项式有相同的系数, 因此求解 $L(x)$ 等价于求解线性方程组. 若方程组不可行, 可以增加次数到 $\ell+1$. 以此类推, 直到线性方程可行从而获得 $L(x)$.

这个性质实际上可以推广到一般的非奇异矩阵多项式.

性质 5.5 [23] 对矩阵多项式 $W(x) \in \mathbb{C}[x]^{s\times t}$, 其中 $s \geqslant t$, $\operatorname{rank} W(u) = t$ 对所有的复数点 $u \in \mathbb{C}^n$ 成立当且仅当存在 $P(x) \in \mathbb{C}[x]^{t\times s}$ 满足方程 $P(x)W(x) = I_t$.

实际上, 当多项式 c_i 具有一般性的系数时, 多项式组 $c = (c_1, \ldots, c_m)$ 是非奇异的. 这在文献 [23] 中得到证明.

定理 5.6 [23] 设 $d_1, \ldots, d_m$ 为正整数, 则存在空间 $\mathbb{R}[x]_{d_1} \times \cdots \times \mathbb{R}[x]_{d_m}$ 中的一个稠密开集 $\mathcal{U}$ 使得所有的 $c = (c_1, \ldots, c_m) \in \mathcal{U}$ 都是非奇异的.

以下是几个多项式乘子表示的例子.

- 考虑单纯形约束 $e^{\mathrm{T}}x - 1 = 0$, $x_1 \geqslant 0, \ldots, x_n \geqslant 0$,

$$\mathcal{E} = \{0\}, \quad \mathcal{I} = [n], \quad c_0(x) = e^{\mathrm{T}}x - 1, \quad c_j(x) = x_j \quad (j \in [n]).$$

其 Lagrange 乘子可以表示为

$$\lambda_0 = x^{\mathrm{T}}\nabla f(x), \quad \lambda_j = \partial_{x_j} f - x^{\mathrm{T}}\nabla f(x), \quad j = 1, \ldots, n. \tag{5.9}$$

- 考虑超立方体 $[-1,1]^n$ 约束

$$c_j := 1 - x_j^2 \geqslant 0, \quad j = 1, \ldots, n.$$

我们可以得到如下乘子表示:

$$\lambda_j = -\frac{1}{2} x_j \partial_{x_j} f, \quad j = 1, \ldots, n. \tag{5.10}$$

- 考虑球面约束 $1-x^{\mathrm{T}}x=0$ 或者球约束 $1-x^{\mathrm{T}}x\geqslant 0$, $\mathcal{E}\cup\mathcal{I}=\{1\}$, $c=(1-x^{\mathrm{T}}x)$ 且

$$x^{\mathrm{T}}\nabla f(x)=x^{\mathrm{T}}(-2\lambda_1 x)=2\lambda_1(1-x^{\mathrm{T}}x)-2\lambda_1=-2\lambda_1.$$

我们可以得到如下乘子表示:

$$\lambda_1=-\frac{1}{2}x^{\mathrm{T}}\nabla f(x). \tag{5.11}$$

5.2 紧的矩 - 平方和松弛序列

本小节介绍新的求解多项式优化 (1.1) 的矩 - 平方和松弛序列. 它对所有目标函数都具有紧性. 假设 (1.1) 的最优值是可达的. 令 $\lambda:=(\lambda_i)_{i\in\mathcal{E}\cup\mathcal{I}}$ 为 (5.2) 中的 Lagrange 乘子向量. 我们作如下假设.

假设 5.7 对每个 $i\in\mathcal{E}\cup\mathcal{I}$, 存在多项式 $p_i\in\mathbb{R}[x]$ 使得对所有 $(x,\lambda)\in\mathcal{K}$, 有 $\lambda_i=p_i(x)$.

假设 5.7 是一个一般性的假设. 例如, 若存在矩阵多项式 $L(x)$ 满足 (5.7), 则

$$\lambda=L(x)C(x)g(x)=L_1(x)\nabla f(x).$$

因此, 可以选择如下的多项式表示:

$$p_i=(L_1(x)\nabla f(x))_i, \tag{5.12}$$

其中下标 i 表示第 i 个分量.

记多项式集合如下:

$$\boxed{\begin{aligned}&\Phi:=\left\{\frac{\partial f}{\partial x_j}-\sum_{i\in\mathcal{E}\cup\mathcal{I}}p_i\frac{\partial c_i}{\partial x_j}\right\}_{j=1}^{n}\cup\{p_jc_j:j\in\mathcal{I}\},\\&\Psi:=\{p_j:j\in\mathcal{I}\}.\end{aligned}} \tag{5.13}$$

若 KKT 条件在最优解处得到满足, 则问题 (1.1) 等价于

$$\begin{cases}\min\ f(x)\\ \text{s.t. } c_i(x)=0\quad(i\in\mathcal{E}),\\ \qquad c_j(x)\geqslant 0\quad(j\in\mathcal{I}),\\ \qquad \phi(x)=0\quad(\phi\in\Phi),\\ \qquad \psi(x)\geqslant 0\quad(\psi\in\Psi).\end{cases} \tag{5.14}$$

我们可以直接用矩 - 平方和松弛求解约束优化 (5.14). 第 k 阶 SOS 松弛为

$$\begin{cases}\max\ \gamma\\ \text{s.t. } f-\gamma\in\mathrm{Ideal}[c_{\mathrm{eq}}]_{2k}+\mathrm{QM}[c_{\mathrm{in}}]_{2k}+\mathrm{Ideal}[\Phi]_{2k}+\mathrm{QM}[\Psi]_{2k}.\end{cases} \tag{5.15}$$

其对偶问题为第 k 阶矩松弛

$$\begin{cases} \min \ \langle f, y\rangle \\ \text{s.t.} \ \mathscr{V}_{c_i}^{(2k)}[y] = 0 \quad (i \in \mathcal{E}), \\ \qquad L_{c_j}^{(k)}[y] \succeq 0 \quad (j \in \mathcal{I}), \\ \qquad \mathscr{V}_{\phi}^{(2k)}[y] = 0 \quad (\phi \in \Phi), \\ \qquad L_{\psi}^{(k)}[y] \succeq 0 \quad (\psi \in \Psi), \\ \qquad y_0 = 1, \quad M_k[y] \succeq 0, \quad y \in \mathbb{R}^{\mathbb{N}^n_{2k}}. \end{cases} \tag{5.16}$$

因此, 得到一个新的矩 - 平方和松弛序列. 记 $f^c_{\mathrm{sos},k}$ 和 $f^c_{\mathrm{mom},k}$ 分别为 (5.15) 和 (5.16) 的最优值. 类似地, 我们有不等式

$$f^c_{\mathrm{sos},k} \leqslant f^c_{\mathrm{mom},k} \leqslant f_{\min}. \tag{5.17}$$

新松弛序列 (5.15) 和 (5.16) 称为紧的 (或者说是精确的, 或者是具有有限收敛性的), 若对所有充分大的 k, $f^c_{\mathrm{sos},k} = f_{\min}$ 都成立.

5.3　松弛的紧性

松弛序列 (5.15) 和 (5.16) 的一个重要性质是对所有目标函数都具有有限收敛性. 考虑如下假设.

假设 5.8　存在 $\rho \in \mathrm{QM}[c_{\mathrm{in}}, \Psi]$ 使得, 若 $u \in \mathcal{K}$, $f(u) < f_{\min}$, 则 $\rho(u) < 0$.

包含关系 $\rho \in \mathrm{QM}[c_{\mathrm{in}}, \Psi]$ 意味着, 若 u 是 (5.14) 的一个可行点, 则 $\rho(u) \geqslant 0$. 因此 $\rho(x) = 0$ 可以看成一个分隔可行驻点和不可行驻点的超曲面. 假设 5.8 是一个一般性的假设 (参见文献 [23]). 下面证明松弛序列 (5.15) 和 (5.16) 的紧性.

定理 5.9 [23]　假设 $\mathcal{K} \neq \emptyset$ 且假设 5.7 成立. 若下列条件之一成立:

(i) $\mathrm{Ideal}[c_{\mathrm{eq}}, \Phi] + \mathrm{QM}[c_{\mathrm{in}}, \Psi]$ 是阿基米德的;

(ii) $\mathrm{Ideal}[c_{\mathrm{eq}}] + \mathrm{QM}[c_{\mathrm{in}}]$ 是阿基米德的;

(iii) 假设 5.8 成立,

则当 k 充分大时, $f^c_{\mathrm{sos},k} = f^c_{\mathrm{mom},k} = f_{\min}$.

5.4　紧性的判定和最优解的提取

假设 y^* 是第 k 阶松弛 (5.16) 的一个最优解. 记多项式次数

$$d := \max\left\{\left\lceil \frac{1}{2}\deg(h)\right\rceil : h \in \{c_i\}_{i\in\mathcal{E}\cup\mathcal{I}} \cup \Phi \cup \Psi\right\}. \tag{5.18}$$

如果平滑截断成立: 存在整数 $t \in [d, k]$ 使得

$$\mathrm{rank} M_t[y^*] = \mathrm{rank} M_{t-d}[y^*], \tag{5.19}$$

则 $f^c_{\mathrm{sos},k} = f_{\min}$ 且可以得到 $r := \mathrm{rank} M_t(y^*)$ 个问题 (5.14) 的最优解. 一般而言, 条件 (5.19) 是提取最优解的一个充分和必要条件. 当 (5.14) 不存在 KKT 点时, 通过求解松弛 (5.15) 和 (5.16), 也可以验证其不可行性. 下面是探测松弛序列 (5.15) 和 (5.16) 的紧性和提取最优解的主要结论.

定理 5.10 [23] 如果假设 5.7 中的 Lagrange 乘子表示存在, 则松弛序列 (5.15) 和 (5.16) 具有如下性质.

(i) 如果对某个松弛阶 k, 松弛 (5.16) 是不可行的, 则优化 (5.14) 一定是不可行的, 即 (1.1) 没有 KKT 点.

(ii) 当假设 5.8 成立时, 如果 (5.14) 不可行, 则当松弛阶 k 充分大时, 松弛 (5.16) 是不可行的.

如果 (5.14) 是可行的, 则还有如下性质.

(iii) 若存在 $t \in [d,k]$ 使得 (5.19) 成立, 则 $f^c_{\mathrm{mom},k} = f_{\min}$.

(iv) 若假设 5.8 成立且 (5.14) 只有有限多个最优解, 则矩松弛 (5.16) 有最优解且当 k 充分大时, 每个最优解 y^* 都满足 (5.19).

若 (5.14) 有无穷多个最优解, 则平滑截断 (5.19) 通常不会成立. 这可以参考性质 4.5.

例 5.11 [23] 考虑优化问题

$$\begin{cases} \min \ x_1^2 + 50x_2^2 \\ \text{s.t.} \ \ x_1^2 - \dfrac{1}{2} \geqslant 0, \\ \qquad x_2^2 - 2x_1x_2 - \dfrac{1}{8} \geqslant 0, \\ \qquad x_2^2 + 2x_1x_2 - \dfrac{1}{8} \geqslant 0. \end{cases}$$

为了满足 (5.7), 矩阵 $L(x)$ 的第一列为

$$\begin{bmatrix} \frac{8\,x_1^3}{5} + \frac{x_1}{5} \\ \frac{288\,x_2\,x_1^4}{5} - \frac{16\,x_1^3}{5} - \frac{x_2\,x_1^2\,124}{5} + \frac{8\,x_1}{5} - 2\,x_2 \\ -\frac{288\,x_2\,x_1^4}{5} - \frac{16\,x_1^3}{5} + \frac{x_2\,x_1^2\,124}{5} + \frac{8\,x_1}{5} + 2\,x_2 \end{bmatrix},$$

第二列为

$$\begin{bmatrix} -\frac{8\,x_1^2\,x_2}{5} + \frac{4\,x_2^3}{5} - \frac{x_2}{10} \\ \frac{288\,x_1^3\,x_2^2}{5} + \frac{16\,x_1^2\,x_2}{5} - \frac{142\,x_1\,x_2^2}{5} - \frac{9\,x_1}{20} - \frac{8\,x_2^3}{5} + \frac{11\,x_2}{5} \\ -\frac{288\,x_1^3\,x_2^2}{5} + \frac{16\,x_1^2\,x_2}{5} + \frac{142\,x_1\,x_2^2}{5} + \frac{9\,x_1}{20} - \frac{8\,x_2^3}{5} + \frac{11\,x_2}{5} \end{bmatrix}.$$

目标函数是强制性的 (即集合 $\{x \in \mathbb{R}^n : f(x) \leqslant c\}$ 对任意的 $c \in \mathbb{R}$ 均是紧集), 因此 $f_{\min}$ 可以在驻点处达到. 该问题的最优值为 $f_{\min} = 56 + \frac{3}{4} + 25\sqrt{5} \approx 112.6517$, 其最优解为 $(\pm\sqrt{\frac{1}{2}}, \pm(\sqrt{\frac{5}{8}} + \sqrt{\frac{1}{2}}))$. 松弛序列 (4.2)–(4.3) 与 (5.15)–(5.16) 计算结果的比较见表 1 (为了简洁, 只显示四位有效数字). 我们可以观测到对所有 $k \geqslant 4$, 有 $f^c_{\mathrm{sos},k} = f_{\min}$.

表 1 例 5.11 的计算结果

阶数 k	(4.2)–(4.3)		(5.15)–(5.16)	
	$f_{\mathrm{mom},k}$	时间	$f^c_{\mathrm{mom},k}$	时间
3	6.7535	0.4611	56.7500	0.1309
4	6.9294	0.2428	112.6517	0.2405
5	8.8519	0.3376	112.6517	0.2167
6	16.5971	0.4703	112.6517	0.3788
7	35.4756	0.6536	112.6517	0.4537

6 总结与展望

本文介绍了多项式优化问题的矩 - 平方和松弛算法及其有限收敛性理论、全局最优性的判别以及全局最优解的获取. 当最优性条件不成立时, 经典的矩 - 平方和松弛通常并不具有有限收敛性. 本文进一步介绍了如何利用 Lagrange 乘子表示来构造具有有限收敛性的矩 - 平方和松弛.

多项式优化是一个正在蓬勃发展的研究领域, 仍然具有很多亟待解决的重要问题. 例如, 定理 4.7 证明了当松弛阶充分大时, 经典的矩 - 平方和松弛具有有限收敛性. 但是却很难得到如此收敛阶次数的准确估计. 在实际计算中, 我们发现对很多问题, 低阶的矩 - 平方和松弛常常已经是精确的. 如何给出低阶松弛精确性的一些充分性刻画, 是一个很重要的研究问题.

参考文献

1 Cui C F, Dai Y H, Nie J. All real eigenvalues of symmetric tensors. SIAM J Matrix Anal Appl, 2014, 35: 1582–1601
2 Curto R, Fialkow L. Truncated K-moment problems in several variables. J Oper Theory, 2005, 54: 189–226
3 Demmel J, Nie J, Powers V. Representations of positive polynomials on noncompact semialgebraic sets via KKT ideals. J Pure Appl Algebra, 2006, 209: 189–200
4 Henrion D, Lasserre J B. Detecting global optimality and extracting solutions in Gloptipoly. In: Positive Polynomials in Control. Lecture Notes in Control and Information Sciences, vol. 312. Berlin-Heidelberg: Springer, 2005, 293–310
5 Henrion D, Lasserre J B. Convergent relaxations of polynomial matrix inequalities and static output feedback. IEEE Trans Automat Control, 2006, 51: 192–202
6 Henrion D, Lasserre J B, Löfberg J. GloptiPoly 3: Moments, optimization and semidefinite programming. Optim Methods Softw, 2009, 24: 761–779
7 Huang L. Optimality conditions for homogeneous polynomial optimization on the unit sphere. Optim Lett, 2023, 17: 1263–1270
8 Huang L, Nie J. Tightness of the matrix moment-SOS hierarchy. arXiv:2403.17241, 2024
9 Huang L, Nie J, Yuan Y X. Generalized truncated moment problems with unbounded sets. J Sci Comput, 2023, 95: 15
10 Huang L, Nie J, Yuan Y X. Homogenization for polynomial optimization with unbounded sets. Math Program, 2023, 200: 105–145
11 Huang L, Nie J, Yuan Y X. Finite convergence of moment-SOS relaxations with non-real radical ideals. arXiv: 2309.15398, 2023
12 Lasserre J B. Global optimization with polynomials and the problem of moments. SIAM J Optim, 2001, 11: 796–817
13 Lasserre J B. Convexity in semialgebraic geometry and polynomial optimization. SIAM J Optim, 2009, 19: 1995–2014
14 Laurent M. Sums of squares, moment matrices and optimization over polynomials. In: Emerging Applications of Algebraic Geometry of IMA Volumes in Mathematics and Its Applications, vol. 149. New York: Springer, 2009, 157–270
15 Lofberg J. YALMIP: A toolbox for modeling and optimization in MATLAB. In: 2004 IEEE International Conference on Robotics and Automation. New York: IEEE, 2004, 284–289
16 Nesterov Y. Squared functional systems and optimization problems. In: High Performance Optimization. Dordrecht: Kluwer Academic, 2000, 405–440
17 Nie J. Polynomial matrix inequality and semidefinite representation. Math Oper Res, 2011, 36: 398–415
18 Nie J. Certifying convergence of Lasserre's hierarchy via flat truncation. Math Program, 2013, 142: 485–510
19 Nie J. Polynomial optimization with real varieties. SIAM J Optim, 2013, 23: 1634–1646
20 Nie J. Optimality conditions and finite convergence of Lasserre's hierarchy. Math Program, 2014, 146: 97–121
21 Nie J. Generating polynomials and symmetric tensor decompositions. Found Comput Math, 2017, 17: 423–465
22 Nie J. Symmetric tensor nuclear norms. SIAM J Appl Algebra Geom, 2017, 1: 599–625
23 Nie J. Tight relaxations for polynomial optimization and Lagrange multiplier expressions. Math Program, 2019, 178: 1–37
24 Nie J. Moment and Polynomial Optimization. Philadelphia: SIAM, 2023
25 Nie J, Demmel J, Sturmfels B. Minimizing polynomials via sum of squares over the gradient ideal. Math Program, 2006, 106: 587–606
26 Nie J, Ranestad K. Algebraic degree of polynomial optimization. SIAM J Optim, 2009, 20: 485–502
27 Nie J, Tang X. Convex generalized Nash equilibrium problems and polynomial optimization. Math Program, 2023,

198: 1485–1518
28 Nie J, Tang X, Zhong S. Rational generalized Nash equilibrium problems. SIAM J Optim, 2023, 33: 1587–1620
29 Nie J, Wang L. Semidefinite relaxations for best rank-1 tensor approximations. SIAM J Matrix Anal Appl, 2014, 35: 1155–1179
30 Nie J, Wang L, Ye J J. Bilevel polynomial programs and semidefinite relaxation methods. SIAM J Optim, 2017, 27: 1728–1757
31 Nie J, Wang L, Ye J J, et al. A Lagrange multiplier expression method for bilevel polynomial optimization. SIAM J Optim, 2021, 31: 2368–2395
32 Nie J, Yang L, Zhong S, et al. Distributionally robust optimization with moment ambiguity sets. J Sci Comput, 2023, 94: 12
33 Nie J, Yang Z. Hermitian tensor decompositions. SIAM J Matrix Anal Appl, 2020, 41: 1115–1144
34 Nie J, Yang Z, Zhang X. A complete semidefinite algorithm for detecting copositive matrices and tensors. SIAM J Optim, 2018, 28: 2902–2921
35 Nie J, Yang Z, Zhou G. The saddle point problem of polynomials. Found Comput Math, 2022, 22: 1133–1169
36 Parrilo P A. Semidefinite programming relaxations for semialgebraic problems. Math Program, 2003, 96: 293–320
37 Putinar M. Positive polynomials on compact semi-algebraic sets. Indiana Univ Math J, 1993, 42: 969–984
38 Reznick B. Some concrete aspects of Hilbert's 17th problem. Contemp Math, 2000, 253: 251–272
39 Scherer C W, Hol C W J. Matrix sum-of-squares relaxations for robust semi-definite programs. Math Program, 2006, 107: 189–211
40 Sun W, Yuan Y. Optimization Theory and Methods: Nonlinear Programming. New York: Springer, 2006

An introduction to polynomial optimization

Lei Huang, Jiawang Nie & Ya-Xiang Yuan

Abstract Polynomial optimization is a broad class of optimization problems whose objective functions and constraining functions are polynomials. The moment-SOS (sum of squares) hierarchy is the most powerful method to solve polynomial optimization globally. In this paper, we introduce moment-SOS methods for solving polynomial optimization. We also study their finite convergence theory, the certificate for global optimality, and how to extract minimizers. When the standard moment-SOS hierarchy does not have finite convergence, we introduce tight relaxations by using Lagrange multiplier expressions.

Keywords polynomial optimization, moment-SOS relaxations, finite convergence, optimality condition, Lagrange multiplier expressions

MSC(2020) 90C23, 65K05, 90C22

doi: 10.1360/SSM-2024-0129