arXiv:2601.02797v1 [math.OC] 6 Jan 2026

对数多项式优化

JIYOUNG CHOI、JIAWANG NIE、XINDONG TANG 和 SUHAN ZHONG

Abstract. 我们研究一类优化问题，其目标函数由对数多项式函数之和构成。这种形式的动机来自最大似然估计等统计估计原理，以及交叉熵和Kullback-Leibler散度等损失函数。我们提出一种基于截断 K-矩问题的矩松弛层级，用于求解对数多项式优化。我们给出该层级具有紧性的充分条件，并在紧性成立时提出一种提取全局优化解的数值方法。此外，我们将最优性条件融入松弛，以便通过便于处理的拉格朗日乘子表达式更好地适用于对数多项式优化。我们还展示了各种应用和数值实验，以说明该方法的有效性。

1. 引言

考虑以下对数多项式优化问题

$$\begin{cases} \max\limits_{x\in\mathbb{R}^n} & f(x) = \sum\limits_{i=1}^{m} a_i \log p_i(x) \\ s.t. & c_j(x) = 0 \ (j \in \mathcal{E}), \\ & c_j(x) \geq 0 \ (j \in \mathcal{I}), \end{cases} \tag{1.1}$$

其中， a_1， $\ldots$， a_m 为正标量， $\mathcal{E}$和 $\mathcal{I}$是两个有限且不相交的标签集，并且对所有 $j \in \mathcal{E} \cup J$， p_1， $\ldots$， p_m， c_j均为多项式。记 $c_\mathcal{E} = (c_j)_{j\in\mathcal{E}}$和$c_\mathcal{I} = (c_j)_{j\in\mathcal{I}}$。(1.1)的可行集可写为

$$K := \{x \in \mathbb{R}^n : c_\mathcal{E}(x) = 0, \, c_\mathcal{I}(x) \geq 0\}. \tag{1.2}$$

为确保 (1.1) 中的目标函数定义良好，我们假设 $p_i \geq 0$ over K，且存在一点 $\hat{x} \in K$，使得对于本文中的所有 $i \in \mathcal{E}\cup\mathcal{I}$，均有 $p_i(\hat{x}) > 0$。为方便起见，记

$$p^a(x) := p_1(x)^{a_1} \cdots p_m(x)^{a_m} = e^{f(x)}.$$

对数多项式优化 (1.1) 通常是非凸的。如果所有 a_i 都是正整数，那么 p^a 就是多项式。在这种情况下，问题 (1.1) 等价于在 K 上最大化 p^a。虽然可以通过矩-SOS松弛全局求解多项式优化的重构问题 [17, 24], ，但当指数 a_i 较大时，计算成本会显著增加。如果任意一个 a_i 不是整数，问题 (1.1) 通常无法表示为多项式优化问题。因此，当对于所有 i， a_i 都是一般的正标量时，很难计算出有用的上界，或证明 (1.1) 的全局最优性。

对数多项式有许多应用。在组合优化中，对数凹多项式是研究计数问题的重要工具 [2]。在统计学习中，许多常见的损失函数，例如对数似然函数和交叉熵，都包含对数项。这些函数广泛

应用于密度估计 [6, 9, 10, 16, 34]和分类 [5, 11, 12, 20, 21, 36]任务，其中待学习的参数通常受到结构约束 [13,14]，例如非负性、归一化或诱导稀疏正则化 [4, 15, 18, 33, 35]。假设所选模型由多项式函数参数化，那么学习过程就归结为求解对数多项式优化。下面我们通过以下应用示例来说明这一框架。

最大似然估计（*MLE*）是统计学中的一个核心问题。考虑由多项式模型 $p(\xi; x)$ 生成的样本序列 $\mathcal{D}$，其中参数 x 受限于可行集 K。令 $S = \mathrm{supp}(\mathcal{D}) = \{\xi^{(1)}, \ldots, \xi^{(m)}\}$ 表示 $\mathcal{D}$的支撑集。我们用 a_i 来统计 $\mathcal{D}$ 中等于 $\xi^{(i)}$的样本数量，其中 $i = 1, \ldots\ldots, m$。与 $\mathcal{D}$ 相关的似然函数以及相应的对数似然函数分别定义为

$$L(\mathcal{D}; x) := \prod_{i=1}^{m} p(x; \xi^{(i)})^{a_i}, \quad \log L(\mathcal{D}; x) = \sum_{i=1}^{m} a_i \log p(x; \xi^{(i)}).$$

最大似然估计旨在找到使似然函数最大化的参数值，这等价于使对数似然最大化。

交叉熵和*Kullback-Leibler*（*KL*）散度是分类任务中两种常见的损失函数。考虑一个支撑集为$S = \{\xi^{(1)}, \ldots, \xi^{(m)}\}$的离散随机变量 ξ，其真实分布 P以 $p(\xi)$作为其概率质量函数（PMF）。令 $Q(x)$为预测分布，其多项式概率质量函数为$q(\xi; x)$。P 与 $Q(x)$之间的交叉熵为

$$H(P, Q(x)) := -\sum_{i=1}^{m} p(\xi^{(i)}) \log q(\xi^{(i)}; x).$$

从 P 到 $Q(x)$ 的 KL 散度为

$$D_{\mathrm{KL}}(P \,\|\, Q(x)) := \sum_{i=1}^{m} p(\xi^{(i)}) \big(\log p(\xi^{(i)}) - \log q(\xi^{(i)}; x) \big).$$

在学习过程中，目标是确定一个最优参数 $x \in K$，使得 $H(P, Q(x))$ 或 $D_{\mathrm{KL}}(P \,\|\, Q(x))$ 最小化。

贡献。 本文提出了一种用于对数多项式优化的矩松弛层级。一般的多项式优化问题可以通过 Moment-SOS 层级全局求解 [17]。这促使我们将该方法扩展到对数多项式优化。我们为 (1.1) 构建了一个矩松弛层级，其中每个松弛问题都具有对数线性目标和半定约束。我们研究了这些松弛的性质，并给出了验证其有限收敛的充分条件（即在有限松弛阶数下松弛是紧的）。确认有限收敛后，我们给出了提取全局最优解的便捷方法。特别地，我们分析了由 SOS-凹多项式构成的特殊对数多项式优化类别。

本文的组织结构如下。第2节介绍符号约定和一些预备知识。第3节提出基于截断 K-矩问题的矩松弛方法。第4节给出松弛紧性的理论保证，并讨论可提取全局优化解的条件。第5节介绍拉格朗日乘子表达式（LMEs），以获得强化矩松弛。数值实验

与应用见第6节。结论及一些讨论见第7节。

2. Preliminaries

记号。符号 $\mathbb{N}$ （分别为 $\mathbb{R}$）表示非负整数（分别为实数）的集合。对于正整数 n, $[n] := \{1, \dots, n\}$, $\mathbb{R}^n$ 表示 n 维实欧几里得空间。对于向量 $u \in \mathbb{R}^n$, $\|u\|$ 表示标准欧几里得范数, δ_u 表示支撑于 u 的狄拉克测度。$\mathbf{1}_n \in \mathbb{R}^n$ 表示全为 $\mathbf{1}$ 的向量, $\mathbf{0}_{n_1 \times n_2}$ 表示维数为 $n_1 \times n_2$ 的零矩阵。对于记号 $\mathbf{1}_n$ 和 $\mathbf{0}_{n_1 \times n_2}$, 如果上下文中维数明确, 则可以省略下标。设 A 为实对称矩阵。不等式 $A \succeq 0$（分别为 $A \succ 0$）表示 A 为半正定（分别为正定）矩阵。令 $x := (x_1, \dots, x_n)$ 为变量向量。我们用 $\mathbb{R}[x]$ 表示关于 x 的实系数多项式环, 并用 $\mathbb{R}[x]_d$ 表示其中次数不超过 d 的多项式子集。对于函数 $f(x)$, $\nabla f(x)$ 表示其全梯度, $\nabla_{x_i} f(x)$ 表示其关于 x_i 的梯度。对于 $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, 记 $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, 其次数由 $|\alpha| := \alpha_1 + \cdots + \alpha_n$ 计数。给定次数 d, 我们记幂集

$$\mathbb{N}^n_d := \{\alpha \in \mathbb{N}^n : |\alpha| \le d\}.$$

x 中所有次数不超过 d 的单项式组成的列向量记为

$$[x]_d := \begin{bmatrix} 1 & x_1 & \dots & x_n & x_1^2 & x_1 x_2 & \dots & x_n^d \end{bmatrix}^T. \tag{2.1}$$

2.1. 非负多项式。 对于多项式 p、$q \in \mathbb{R}[x]$ 以及子集 I、$J \subseteq \mathbb{R}[x]$, 它们的乘积和分别定义为

$$p \cdot I := \{pq : q \in I\} \quad \text{and} \quad I + J := \{a + b : a \in I, b \in J\}.$$

若子集 $I \subseteq \mathbb{R}[x]$ 满足 $I + I \subseteq I$ 和 $\mathbb{R}[x] \cdot I \subseteq I$, 则称其为 $\mathbb{R}[x]$ 的理想。若多项式$\sigma \in \mathbb{R}[x]$ 满足 $\sigma = p_1^2 + \cdots + p_k^2$（其中$p_i \in \mathbb{R}[x]$为某些多项式）, 则称其为平方和（*SOS*）。所有平方和多项式的集合记为 $\Sigma[x]$。

设 $h = (h_1、\dots、h_m)$和 $g = (g_1、\dots、g_t)$为两个多项式元组。 h 的实零点集是集合 $Z(h) := \{x \in \mathbb{R}^n : h_1(x) = \cdots = h_m(x) = 0\}$。由 h 生成的理想为

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

显然, 每个 $p \in$ 理想(h) 都在 $Z(h)$ 上消失。考虑由 g 确定的半代数集:

$$S(g) := \{x \in \mathbb{R}^n : g_1(x) \ge 0, \dots, g_t(x) \ge 0\}.$$

若多项式 f 可以分解为如下形式, 则它在 $S(g)$ 上非负:

$$f = \sigma_0 + \sigma_1 g_1 + \cdots + \sigma_t g_t, \tag{2.2}$$

对某个 $\sigma_i \in \Sigma[x]$成立。由 g 生成的二次模由所有形如 (2.2) 的多项式组成, 即

$$\text{QM}[g] := \Sigma[x] + g_1 \cdot \Sigma[x] + \cdots + g_t \cdot \Sigma[x].$$

如果 $p \in$ 理想$[h] + \text{QM}[g]$, 那么对于所有 $x \in Z(h) \cap S(g)$, $p(x) \ge 0$。但反过来并不总是成立。如果理想$[h] + \text{QM}[g]$ 包含一个多项式 q, 满足 $\{x \in \mathbb{R}^n : q(x) \ge 0\}$ 是一个紧集, 则称理想$[h] + \text{QM}[g]$ 为 阿基米德的。在这一

条件下，如果一个多项式 f 在 $Z(h)\cap S(g)$ 上严格为正，那么它必属于理想$[h]+\mathrm{QM}[g]$。这一结果称为 *Putinar* 正定性定理 [32]。

2.2. 矩阵与局部化矩阵。 以$\alpha \in \mathbb{N}^n_{2k}$ 为索引的实向量空间记为 $\mathbb{R}^{\mathbb{N}^n_{2k}}$。该空间中的向量 $y = (y_\alpha)_{\alpha\in\mathbb{N}^n_{2k}}$ 称为次数为 $2k$的截断多重序列 (tms)。对于整数 $k \in [0, d]$, 由 y 生成的 k阶矩阵 定义为

$$M_k[y] := [y_{\alpha+\beta}]_{\alpha,\beta\in\mathbb{N}^n_k}.$$

$M_k[y]$ 的行和列由 $\mathbb{N}^n_k$中的多重指标索引，通常按分次字典序排列。例如，当 $n = 2$ 且 $k = 2$时，矩阵 $M_2[y]$为

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

每个多项式 $f \in \mathbb{R}[x]_{2k}$ 都可以用其系数向量 $\mathrm{vec}(f) := (f_\alpha)_{\alpha\in\mathbb{N}^n_{2k}}$表示，即 $f(x) = \mathrm{vec}(f)^T[x]_{2k}$。因此，tms $y \in \mathbb{R}^{\mathbb{N}^n_{2k}}$在 $\mathbb{R}[x]_{2k}$ 上定义了一个线性泛函:

$$\langle f, y\rangle := \sum_{\alpha\in\mathbb{N}^n_{2k}} f_\alpha y_\alpha = \mathrm{vec}(f)^T y. \tag{2.3}$$

对于多项式 $q \in \mathbb{R}[x]_{2k}$，令 $s := k - \lceil \deg(q)/2\rceil$。乘积 $q(x)[x]_s[x]_s^T$ 是一个 $\binom{n+s}{s}\times\binom{n+s}{s}$ 阶对称多项式矩阵。其展开式可写为

$$q(x)[x]_s[x]_s^T = \sum_{\alpha\in\mathbb{N}^n_{2k}} x^\alpha Q_\alpha$$

其中 Q_α 为某些常对称矩阵。与多项式 q 和 tms $y \in \mathbb{R}^{\mathbb{N}^n_{2k}}$ 相关联的 k 阶 局部化矩阵 定义为

$$L_q^{(k)}[y] := \sum_{\alpha\in\mathbb{N}^n_{2k}} y_\alpha Q_\alpha.$$

对于 $q = 1$ （常值单位多项式）的特殊情形，该定义还原为矩阵，即 $L_1^{(k)}[y] = M_k[y]$。

2.3. 截断 K-矩问题。 若 tms $y = (y_\alpha)_{\alpha\in\mathbb{N}^n_{2k}}$ 存在 μ 上的博雷尔测度 K，则称其

$$y_\alpha = \int_K x^\alpha d\mu \quad \text{for all} \quad \alpha \in \mathbb{N}^n_{2k}. \tag{2.4}$$

这样的测度 μ称为的表示测度。μ的支撑集记为 $\mathrm{supp}(\mu)$，它是满足 $\mu(\mathbb{R}^n \setminus T) = 0$ 的最小闭集 $T \subseteq \mathbb{R}^n$。如果一个测度的支撑集是有限集，则称其为有限原子测度。具体而言，如果 $\mathrm{supp}(\mu) = \{v_1, \ldots, v_r\}$，则称测度 μ为 r-原子测度，并可写成

$$\mu = \sum_{i=1}^r \lambda_i \delta_{v_i} \quad \text{for some positive weights } \lambda_1, \ldots, \lambda_r > 0.$$

要使 tms $y \in \mathbb{R}^{\mathbb{N}^n_{2k}}$ 在 K 上像（2.4）中那样存在表示测度 μ，必须满足以下矩阵条件（见 [24]）：

$$M_k[y] \succeq 0, \quad L^{(k)}_{c_j}[y] \succeq 0\,(j \in \mathcal{I}), \quad L^{(k)}_{c_j}[y] = 0\,(j \in \mathcal{E}). \tag{2.5}$$

3. A semidefinite relaxation hierarchy

本节提出一种矩松弛方法，用于求解对数多项式优化问题（1.1）。

3.1. 一种凸松弛。首先，我们引入对数多项式优化问题（1.1）的一种自然凸松弛。定义次数

$$d := \max\{\lceil \deg(p)/2 \rceil, \lceil \deg(c_{\mathcal{E}})/2 \rceil, \lceil \deg(c_{\mathcal{I}})/2 \rceil\}. \tag{3.1}$$

（1.1）中的每个多项式都属于多项式环 $\mathbb{R}[x]_{2k}$，其中 $k \geq d$。通过将每个涉及的单项式 x^α 替换为辅助矩变量 y_α，我们得到（1.1）的如下等价矩重构：

$$\begin{cases} f_{\max} := \max\limits_{z} & \sum\limits_{i=1}^{m} a_i \log\langle p_i, z\rangle \\ s.t. & z = [x]_{2d} \text{ for some } x \in K. \end{cases} \tag{3.2}$$

这里，$\langle \cdot,\ \cdot \rangle$ 是（2.3）中定义的双线性运算，$[x]_{2d}$ 是（2.1）中的 $2d$ 次单项式向量，而 K 是（1.2）给出的可行集。问题（3.2）的目标函数是凹函数，因为每个 $a_i > 0$ 和对数函数都是凹函数。然而，其可行集通常是非凸的，并且可能没有非空内部。因此，大多数通用非线性优化方法都不适用于求解（3.2）。为了解决这一问题，我们定义由（3.2）的可行集生成的锥包：

$$\mathscr{R}_d(K) := \Big\{ z = \sum_{j=1}^{l} \lambda_j [v_j]_{2d} : \lambda_j \geq 0,\ v_j \in K,\ l \in \mathbb{N} \Big\}. \tag{3.3}$$

由此可知，$\mathscr{R}_d(K) \cap \{z_0 = 1\}$ 是集合 $\{[x]_{2d} :\ x \in K\}$的凸包。鉴于对所有 $i \in [m]$ 和 $x \in K$，$p_i(x) \geq 0$，每个 $\log\langle p_i,\ z\rangle$ 在 $\mathscr{R}_d(K) \cap \{z_0 = 1\}$上均非负。这是因为任意 $z \in \mathscr{R}_d(K) \cap \{z_0 = 1\}$ 都可以分解为 $z = \sum_{j=1}^{l} \lambda_j [v_j]_{2d}$，其中某个 $l \in \mathbb{N}$ 满足所有 $v_j \in K$ 和 $\lambda_j \geq 0$，$\lambda_1 + \cdots + \lambda_l = 1$。因此，

$$\langle p_i, z\rangle = \Big\langle p_i, \sum_{i=1}^{l} \lambda_j [v_j]_{2d} \Big\rangle = \sum_{j=1}^{l} \lambda_j \langle p_i, [v_j]_{2d}\rangle = \sum_{j=1}^{m} \lambda_j p_i(v_j) \geq 0.$$

由此可得到（3.2）的一个自然凸松弛：

$$\begin{cases} f_{\mathrm{mom}} := \max\limits_{z} & \sum\limits_{i=1}^{m} a_i \log\langle p_i, z\rangle \\ s.t. & z_0 = 1,\ z \in \mathscr{R}_d(K). \end{cases} \tag{3.4}$$

显然，$f_{\max} \leq f_{\mathrm{mom}}$。特别地，当 $f_{\max} = f_{\mathrm{mom}}$时，优化问题（3.4）被称为（3.2）的紧 松弛。我们指出，松弛间隙 $f_{\mathrm{mom}} - f_{\max}$ 可以任意大。以下例子说明了这一现象。

例3.1。 考虑单变量对数多项式优化问题

$$\begin{cases} \max\limits_{x\in\mathbb{R}} & 0.5\log(1+\epsilon-x)+0.5\log(1+\epsilon+x) \\ s.t. & 1-x^2=0, \end{cases} \tag{3.5}$$

其中 $\epsilon>0$ 是一个较小的标量。由于可行集 $K=\{1,-1\}$ 只包含两个点，我们可以显式表示锥包 $\mathscr{R}_1(K)$，从而得到 (3.5) 的凸松弛:

$$\begin{cases} \max\limits_{z\in\mathbb{R}^3} & 0.5\log(1+\epsilon-z_1)+0.5\log(1+\epsilon+z_1) \\ s.t. & z_0=z_2=1,\ -1\le z_1\le 1. \end{cases}$$

(3.5) 及其凸松弛的最优值分别为

$$f_{\max}=0.5\log(\epsilon)+0.5\log(2+\epsilon),\quad f_{\mathrm{mom}}=\log(1+\epsilon).$$

显然，松弛间隙 $f_{\mathrm{mom}}-f_{\max}\ge 0-0.5\log(\epsilon)\to\infty$ as $\epsilon\to 0$。

现在给出使松弛 (3.4) 成为紧松弛的充分条件。

命题 3.2. 假设 K 非空。那么，若满足以下任一条件，凸松弛 (3.4)就是对数多项式优化 (1.1) 的紧松弛。

(i) 矩重构 (3.2) 的最优值等于

$$\begin{cases} \min & \gamma_1+\cdots+\gamma_m \\ s.t. & \gamma_i-a_i\log p_i(x)\ge 0 \ \textit{for all}\ x\in K\,(i\in[m]), \\ & \gamma_1,\ldots,\gamma_m\in\mathbb{R}. \end{cases} \tag{3.6}$$

(ii) 存在凸松弛 (3.4)的一个最大化解 $z^*=\sum_{j=1}^{l}\lambda_j[v_j^*]_{2d}$，其所有 $\lambda_j\ge 0$ 和 $v_j^*\in K$ 均满足 $p(v_1^*)=p(v_2^*)=\cdots=p(v_l^*)$。

证明. 原始的对数多项式优化 (1.1) 等价于其矩重构 (3.2)。令 $f_{\max}, f_{\mathrm{mom}}$ 分别表示 (3.2) 和 (3.4) 的最优值。只需证明 $f_{\max}\ge f_{\mathrm{mom}}$。

(i) 若 $f_{\max}=\infty$，则 $f_{\mathrm{mom}}\ge f_{\max}=\infty$。设 $f_{\max}<\infty$是 (3.6) 的最优值。令 $\hat{}\gamma=(\hat{}\gamma_1、\ldots\ldots、\hat{}\gamma_m)$ 为 (3.6) 的极小化解。则 $f_{\max}=\mathbf{1}^T\hat{}\gamma$。注意，不等式 $\gamma_i\ge a_i\log p_i(x)$ 等价于 $e^{\gamma_i/a_i}\ge p_i(x)$。(3.6) 中 $\hat{}\gamma$ 的可行性确保对所有 $i\in[m]$,

$$e^{\hat{\gamma}_i/a_i}\ge\max_{x\in K}p_i(x)=\max_{x\in K}\langle p_i,[x]_{2d}\rangle.$$

由于 $\mathscr{R}_d(K)\cap\{z_0=1\}$ 是 $\{[x]_{2d}$ 的凸包: $x\in K\}$，由此可得

$$e^{\hat{\gamma}_i/a_i}-\langle p_i,z\rangle\ge 0\ \Leftrightarrow\ \hat{\gamma}_i-a_i\log\langle p_i,z\rangle\ge 0$$

对于 (3.4) 的所有可行点。因此， $f_{\max}=\mathbf{1}^T\hat{}\gamma\ge f_{\mathrm{mom}}$。

(ii) 在给定条件下，若对于所有 $i\in[m]$, $p_i(v_1^*)=\cdots=p_i(v_l^*)$，则

$$\begin{aligned} f_{\mathrm{mom}}&=\sum_{i=1}^{m}a_i\log\big\langle p_i,\sum_{j=1}^{l}\lambda_j[v_j^*]_{2d}\big\rangle=\sum_{i=1}^{m}a_i\log\big(\sum_{j=1}^{l}\lambda_j\langle p_i,[v_j^*]_{2d}\rangle\big)\\ &=\sum_{i=1}^{m}a_i\log\big(\sum_{j=1}^{l}\lambda_j p_i(v_j^*)\big)=\sum_{i=1}^{m}a_i\log\langle p_i,[v_1^*]_{2d}\rangle\le f_{\max}. \end{aligned}$$

因此，当条件 (i) 或 (ii) 成立时，凸松弛 (3.4) 是紧的。 □

对于 $m=1$ 这一特殊情形，对数多项式优化 (1.1) 等价于在 K 上最小化 $p_1(x)$ 的多项式优化。下面的结果可直接由命题 3.2 推出。

推论 3.3。当 $m=1$ 时，松弛(3.4) 是紧的。

在实践中，问题 (3.4) 仍然难以求解。这是因为，对于多变量情形，矩锥 $\mathscr{R}_d(K)$ 通常没有方便的参数表达式。另一方面，矩锥 $\mathscr{R}_d(K)$ 常用半定约束来近似。这促使我们进一步构造 (3.4) 的半定松弛层级。

3.2. **矩松弛层级**。回想一下，K 由多项式约束决定

$$c_j(x)=0\,(j\in\mathcal{E}),\quad c_j(x)\geq 0\,(j\in\mathcal{I}).$$

设 $z\in\mathscr{R}_d(K)$ 为任意截断矩向量。对于所有 $k\geq d$，都存在一个 k 阶的截断矩序列扩展 z，使得

$$y=(y_\alpha)\in\mathbb{R}^{\mathbb{N}^n_{2k}}\text{ with }y_\alpha=z_\alpha\,\forall\alpha\in\mathbb{N}^d_n\quad\text{and}$$

$$M_k[y]\succeq 0,\quad L^{(k)}_{c_j}[y]=0\,(j\in\mathcal{E}),\quad L^{(k)}_{c_j}[y]\succeq 0\,(j\in\mathcal{I}).$$

在上述内容中，$M_k[y]$ 是 k阶矩矩阵，每个 $L^{(k)}_{c_j}[y]$ 都是与 c_j相关联的 k阶局部化矩阵，这些矩阵已在第 2.2 小节中介绍。对于递增的 $k=d,\ d+1,$ ……，我们可以构造 (3.4) 的半定约束矩松弛层级：

$$\left\{\begin{array}{rl} f_{\mathrm{mom},k}:=\max\limits_{y} & \sum\limits_{i=1}^{m}a_i\log\langle p_i,y\rangle\\ s.t. & y_0=1,\,y\in\mathbb{R}^{\mathbb{N}^n_{2k}},\,M_k[y]\succeq 0,\\ & L^{(k)}_{c_j}[y]=0\,(j\in\mathcal{E}),\\ & L^{(k)}_{c_j}[y]\succeq 0\,(j\in\mathcal{I}),\end{array}\right.\tag{3.7}$$

其中，整数 k 称为松弛阶数。为与前面的凸松弛（3.4）区分，我们将问题（3.7）称为 k 阶矩松弛，相应的层级称为矩层级。注意，每个矩松弛（3.7）都是凸的，但不同于标准半定规划，因为它具有对数线性目标函数。在计算实践中，可以使用内点法的锥规划求解器对其进行全局求解。令 $f_{\mathrm{mom},k}$ 表示（3.7）在第 k 松弛阶数下的最优值。根据前述分析，显然有

$$f_{\mathrm{mom}}\leq\cdots\leq f_{\mathrm{mom},k+1}\leq f_{\mathrm{mom},k}\leq\cdots\leq f_{\mathrm{mom},d}.$$

下面我们证明，该松弛层级具有渐近收敛性。

定理 3.4。假设理想$[c_\mathcal{E}]+QM[c_\mathcal{I}]$ 是阿基米德的，且（3.4） 的最优值等于（3.6）的最优值。那么 $f_{\mathrm{mom},k}\to f_{\mathrm{mom}}$ 当 $k\to\infty$ 时。

证明。令 $\hat{}\gamma=(\hat{}\gamma_1,\ \ldots\ldots,\ \hat{}\gamma_m)$ 表示最小化问题 (3.6) 的优化器。那么对于任意 $\epsilon>0$，$\hat{}\gamma_i+\epsilon-a_i\log p_i(x)>0$ 对于所有 $x\in K$ 和 $i\in[m]$，等价地，

$$e^{(\hat{\gamma}_i+\epsilon)/a_i}-p_i(x)>0\text{ on }K\text{ for all }i\in[m].$$

假设理想 $[c_{\mathcal{E}}] + \mathrm{QM}[c_{\mathcal{I}}]$ 是阿基米德的。根据 Putinar 正定性定理，存在 $N_0(\epsilon) > 0$，使得对于所有 $i \in [m]$ 和 $k \geq N_0(\epsilon)$，

$$e^{(\hat{\gamma}_i+\epsilon)/a_i} - p_i(x) \in \mathrm{Ideal}[c_{\mathcal{E}}]_{2k} + \mathrm{QM}[c_{\mathcal{I}}]_{2k},$$

其中，$\mathrm{Ideal}[c_{\mathcal{E}}]_{2k} + \mathrm{QM}[c_{\mathcal{I}}]_{2k}$ 是 $\mathrm{Ideal}[c_{\mathcal{E}}] + \mathrm{QM}[c_{\mathcal{I}}]$的 $2k$次截断。根据 [24, 定理 2.5.2]，，(3.7) 的每个可行点都属于 $\mathrm{Ideal}[c_{\mathcal{E}}]_{2k} + \mathrm{QM}[c_{\mathcal{I}}]_{2k}$的对偶锥。这意味着

$$\langle e^{(\hat{\gamma}_i+\epsilon)/a_i} \cdot 1 - p_i, y\rangle = e^{(\hat{\gamma}_i+\epsilon)/a_i} - \langle p_i, y\rangle \geq 0$$

(3.7) 的所有可行点均满足。由于 $\langle p_i,\ y\rangle \geq 0$ （对于 $\langle p_i,\ y\rangle = 0$这一特殊情况，以下不等式仍成立，其中 $\log(0) = -\infty$），取对数可得，

$$\hat{\gamma}_i + \epsilon - a_i \log\langle p_i, y\rangle \geq 0 \quad \forall i \in [m].$$

将上述不等式对 $i \in [m]$ 求和，得到

$$m\epsilon + \mathbf{1}^T\hat{\gamma}_i - \sum_{i=1}^{m} a_i \log\langle p_i, y\rangle \geq 0$$

(3.7) 的所有可行点均满足。注意，由于 (3.4) 和 (3.6) 具有相同的最优值，$f_{\mathrm{mom}} = \mathbf{1}^T\hat{}\gamma$。因此，对于所有 $k \geq N_0(\epsilon)$，都有 $f_{\mathrm{mom},\ k} \leq f_{\mathrm{mom}} + m\epsilon$。令$\epsilon \to 0$，可得 $N_0(\epsilon) \to \infty$，故 $f_{\mathrm{mom},\ k} \to f_{\mathrm{mom}}$，因为 $k \to \infty$。□

回顾一下，$f_{\max}$ 表示原始对数多项式优化的最优值。结合命题 3.2 和定理 3.4，我们可以得到以下结果。

推论 3.5。假设 *Ideal*$[c_{\mathcal{E}}] + QM[c_{\mathcal{I}}]$ 是阿基米德的，且命题 *3.2* 的条件 *(i)* 成立。那么 $f_{\mathrm{mom},\ k} \to f_{\max}$ 当 $k \to \infty$时。

4. Tightness analysis of moment relaxations

如果矩松弛 (3.7) 在某个有限松弛阶数下与 (1.1) 具有相同的最优值，则称矩松弛 (3.7) 对 (1.1) 是紧的 （或具有有限收敛性）。本节研究如何证明这种紧性，以及在确认紧性后如何获得极大化点。特别地，我们研究由 SOS-凹多项式给出的特殊类对数多项式优化问题，该类问题总能实现有限收敛。

4.1. 证明矩松弛的紧性。

设 k 为松弛阶数。令$f_{\mathrm{mom},\ k}$ 表示矩松弛 (3.7) 的最优值，f_{mom}表示凸松弛 (3.4) 的最优值，以及 $f_{\max}$ 表示原始对数多项式优化 (1.1) 的最优值。由于 $f_{\max} \leq f_{\mathrm{mom}} \leq f_{\mathrm{mom},\ \mathrm{k}}$，$f_{\max} = f_{\mathrm{mom},\ \mathrm{k}}$ 的必要条件是 $f_{\max} = f_{\mathrm{mom}}$。如果 (3.7) 的一个最优解 y^* 满足 $y^*|_{2d} \in \mathscr{R}_d(K)$，则该条件成立，其中 $y^*|_{2d}$ 表示 y^*的 $2d$阶截断。为验证这种从属关系，我们可以应用一个称为平坦截断的方便秩条件，其形式为

$$\exists t \in [d, k] \quad \text{s.t.} \quad \mathrm{rank}\, M_t[y^*] = \mathrm{rank}\, M_{t-d}[y^*]. \tag{4.1}$$

在平坦截断条件下，y^* 存在唯一的由 $r =$ 支撑、包含 $M_t[y^*]$ 秩 $M_t[y^*]$ 个互异点的 K 表示原子测度。换言之，存在正标量 $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ 以及互异点 $v_1^*, \ldots, v_r^* \in K$，使得

$$y^*|_{2t} = \lambda_1[v_1^*]_{2t} + \cdots + \lambda_r[v_r^*]_{2t}, \tag{4.2}$$

其中， $y^*|_{2t}$ 是 y^* 的 $2t$ 阶截断。特别地，根据 [24, 定理 2.7.7], ，这种分解（4.2）在重排意义下是唯一的。由于 $t \geq d$，分解（4.2）保证了 $y^*|_{2d} \in \mathscr{R}_d(K)$。随后，我们总结使（3.7）成为（1.1）的紧松弛的充分条件。

定理 4.1。设 y^* 是 k 阶矩松弛(3.7)的优化器。若 y^* 满足平坦截断条件(4.1)，其中 $t \in [d,k]$，则$f_{\mathrm{mom},\ k} = f_{\mathrm{mom}}$，且 $y^*|_{2t}$ 可表示为(4.2) 中的分解，其中包含 $r =$ 秩为 $M_t[y^*]$的互异点 v_1^*、……、$v_r^* \in K$。若进一步有 $p(v_1^*) = \cdots = p(v_r^*)$，则 $f_{\max} =$$f_{\mathrm{mom}} = f_{\mathrm{mom},\ k}$，并且所有 v_1^*、……、v_r^* 都是 *(1.1)* 的全局最大化点。

证明。在平坦截断条件下， $z^* = y^*|_{2d}$ 是 (3.4) 的可行点。由于每个 $\deg(p_i) \leq 2d$, 我们有

$$f_{\mathrm{mom,k}} = \sum_{i=1}^{m} a_i \log\langle p_i, y^*\rangle = \sum_{i=1}^{m} a_i \log\langle p_i, z^*\rangle \leq f_{\mathrm{mom}}. \tag{4.3}$$

将上述不等式与$f_{\mathrm{mom}} \leq f_{\mathrm{mom,k}}$结合，我们得出 $f_{\mathrm{mom}} =$$f_{\mathrm{mom,k}}$。分解式 (4.2) 由平坦截断条件 [24,定理 2.7。7]推出。若 $p(v_1^*) = \cdots = p(v_r^*)$，则根据命题 3.2 有 $f_{\max} = f_{\mathrm{mom}}$。因此，$f_{\max} = f_{\mathrm{mom}} = f_{\mathrm{mom,k}}$。要证明对于每个 $l \in [r]$， v_l^* 都是全局最大化点，只需证明 $f(v_l^*) = f_{\mathrm{mom,k}} = f_{\max}$。事实上，我们有

$$f_{\mathrm{mom,k}} = \sum_{i=1}^{m} a_i \log\Big(\sum_{j=1}^{r} \lambda_j p_i(v_j^*)\Big) = \sum_{i=1}^{m} a_i \log\big(p_i(v_l^*)\big) = f(v_l^*)$$

对每个 $l \in [r]$ 均成立。因此， $v_1^*, \ldots, v_r^*$都是 (1.1) 的全局最大化点。 □

注意，秩 $M_d[y^*] = 1$ 当且仅当 $y^*|_{2d} = [v^*]_{2d}$，其中 $v^* = (y_{e_1}^*, \ldots, y_{e_n}^*)$。对于这种特殊情形，定理 4.1 中的所有条件都自动满足。

推论 4.2。设 y^* 是 k 阶矩松弛(3.7)的优化器。若秩为 $M_d[y^*] = 1$，则 $f_{\max} = f_{\mathrm{mom},\ \mathrm{k}}$和 $v^* = (y_{e_1}^*,\ \ldots,\ y_{e_n}^*)$ 是(1.1)的全局最大化点。

我们注意到，平坦截断条件对于对数多项式优化的充分性，与其对于标准多项式优化的充分性大不相同。设 y^* 是 k 阶矩松弛 (3.7) 的优化器。若其满足平坦截断条件 (4.1)，则 (3.7) 是 (3.4) 的紧松弛，这意味着 $f_{\mathrm{mom}} = f_{\mathrm{mom},\ \mathrm{k}}$。然而，如例3.1所示，凸松弛 (3.4) 不一定是矩重构 (3.2) 的紧松弛。因此，即使 y^* 具有带有候选解v_1^*，……，$v_r^* \in K$的分解 (4.2)，仍有可能出现 $f_{\max} < f_{\mathrm{mom}} = f_{\mathrm{mom},\ \mathrm{k}}$。有趣的是，在此类情形下，点 v_i^* 可能是也可能不是 (1.1) 的全局最大化点。下面的例子将说明这两种可能性。

例4.3。考虑以下对数多项式优化问题:

$$\begin{cases} \max\limits_{x\in\mathbb{R}} & f(x) \coloneqq \log((x-1)^2 + a) + \log((x+1)^2 + a) \\ s.t. & x(x-1)(x+1) = 0, \end{cases} \tag{4.4}$$

其中 $a > 0$ 是一个参数。可行集 $K = \{1, -1, 0\}$ 以及次数界$d = 2$。很容易计算

$$f(1) = \log(a(4+a)), \quad f(-1) = \log(a(4+a)), \quad f(0) = 2\log(1+a).$$

考虑松弛阶数为 $k=3$的矩松弛 (3.7)。(4.4) 中的等式约束等价于 $x^3-x=0$，由此在矩松弛中产生线性等式约束 $y_5=y_3=y_1$ 和 $y_6=y_4=y_2$。通过消去变量，(4.4) 的三阶矩松弛可简化为

$$(4.5)\qquad \begin{cases} \max\limits_{(y_1,y_2)} & \log(y_2-2y_1+1+a)+\log(y_2+2y_1+1+a) \\ s.t. & y_1,y_2\in\mathbb{R},\ \begin{bmatrix} 1 & y_1 & y_2 & y_1 \\ y_1 & y_2 & y_1 & y_2 \\ y_2 & y_1 & y_2 & y_1 \\ y_1 & y_2 & y_1 & y_2 \end{bmatrix} \succeq 0. \end{cases}$$

以上，目标函数等于 $\log((y_2^2+1+a)^2-4y_1^2)$，而半正定约束意味着 $0\le y_1^2\le y_2^2\le y_2\le 1$. 因此，对于 (4.5) 的所有可行点，均满足

$$\log((y_2^2+1+a)^2-4y_1^2)\le\log((1+1+a)^2-4(0))=2\log(a+2).$$

特别地，上述等式在 $y_1^*=1$ 和 $y_2^*=0$处成立。因此，(4.4) 的三阶矩松弛的最优值和优化器为

$$f_{\mathrm{mom},3}=2\log(a+2),\quad y^*=(1,0,1,0,1,0,1),$$

其中 $y^*=0.5[u_1]_6+0.5[u_2]_6$ 的支撑集为可行点 $u_1=1, u_2=-1$。 (i) 当 $a=1$时，我们有

$$f(1)=f(-1)=\log(5)>f(0)=\log(4).$$

在这种情况下，$f_{\max}=\log(5)$，且 $u_1=1$ 和 $u_2=-1$ 都是 (4.4) 的极大化点。然而，$f_{\mathrm{mom},3}=\log(9)$，严格大于 $f_{\max}$。事实上，可以验证，对于所有 $k\ge 3$，均有 $f_{\mathrm{mom}}=f_{\mathrm{mom},k}=\log(9)$，这表明矩松弛层级与原始对数多项式优化问题之间存在非零间隙。

(ii) 当 $a=0.1$ 时，我们有

$$f(1)=f(-1)=\log(0.41)<f(0)=\log(1.21).$$

这意味着 $f_{\max}=\log(1.21)$，其在唯一优化器 $u_3=0$ 处取得。在这种情况下，$f_{\mathrm{mom},3}=\log(4.41)>f_{\max}$ 以及 u_1，u_2 不再是式 (4.4) 的全局最大化点。还可以进一步验证，对于所有 $k\ge 3$，松弛 (3.7) 始终不是紧的。

命题 3.2 给出了关于 $f_{\max}=f_{\mathrm{mom}}$ 的两个充分条件，而这两个条件都需要进行数值验证。回顾可知，凸多项式优化的矩松弛总是表现出有限收敛。这促使我们研究由凹/凸多项式定义的对数多项式优化。

4.2. **凹对数多项式优化**。如果 K 是凸集，且每个 p_i 在 K 上都是凹函数，即对于所有 u，$v\in K$、$i\in[m]$ 和 $\lambda\in[0,1]$，则问题 (1.1) 称为凹对数多项式优化问题，

$$\lambda u+(1-\lambda)v\in K,\quad p_i(\lambda u+(1-\lambda)v)\ge\lambda p_i(u)+(1-\lambda)p_i(v).$$

在满足平坦截断条件时，凹对数多项式优化具有紧致的矩松弛。

定理 4.4。假设(1.1) 是凹的，且 y^* 是其 k阶矩松弛 *(3.7)* 的优化器。若 y^* 满足平坦截断条件 *(4.1)*，则$f_{\max}=f_{\mathrm{mom}}=f_{\mathrm{mom},k}$且 $v^*=(y^*_{e_1},\dots,y^*_{e_n})$ 是 *(1.1)* 的全局最大化点。

证明。 假设对于某个 $t \in [d,k]$， y^* 满足平坦截断条件 (4.1)。令 $r =$ 的秩为 $M_r[y^*]$。存在正标量 λ_i 和互异点 $v_i^* \in K$ （$i \in [r]$ ），使得

$$y^*|_{2d} = \lambda_1[v_1^*]_{2d} + \cdots + \lambda_r[v_r^*]_{2d}.$$

由于 $y_0^* = 1$ 且 (1.1) 是一个凹对数多项式优化问题，我们有$\lambda_1 + \cdots + \lambda_r = 1$ 且 $v^* = (y_{e_1}^*, \ldots\ldots, y_{e_n}^*) = \lambda_1 v_1^* + \cdots + \lambda_r v_r^* \in K$. 这意味着

$$\langle p_i, y^* \rangle = \sum_{j=1}^{r} \lambda_j \langle p_i, [v_j^*]_{2d} \rangle = \sum_{j=1}^{r} \lambda_j p_i(v_j^*) \le p_i(v^*)$$

对所有 $i \in [m]$均成立。根据假设， y^* 是 (3.7) 的一个优化器。由于对数函数是单调递增的，我们进一步有

$$f_{\mathrm{mom,k}} = \sum_{i=1}^{m} a_i \log\langle p_i, y^* \rangle \le \sum_{i=1}^{m} a_i \log p_i(v^*) \le f_{\max}.$$

这意味着 $f_{\max} = f_{\mathrm{mom,\ k}}$ 且 v^* 是 (1.1) 的一个极大化点。 □

我们指出，定理 4.1 和定理 4.4 描述了两类不同的对数多项式优化问题，它们均能得到紧的矩松弛，同时都要求满足平坦截断条件。具体而言，定理 4.1 给出了相应条件，在这些条件下，由 $y^*|_{2d}$ 所表示的原子测度的支撑集中的所有点 v_1^*, ……， v_r^*都是 (1.1) 的极大化点。相比之下，定理 4.4 关注凹对数多项式优化，其中这些 v_1^*，……， v_r^*不一定是 (1.1) 的全局极大化点。

然而，判断次数为四次或更高的多项式的凸性或凹性是 NP难的 [1]。在计算实践中，SOS 凸/凹多项式是凸/凹多项式的子集，备受关注，因为可以通过求解半定规划来验证它们 [24, 引理 7.1。3]若存在一个矩阵多项式 $R(x)$，使其海森矩阵 $\nabla^2 q = R^T R$，则称多项式 $q(x)$ 为*SOS* 凸； 若 $-q$ 是 SOS 凸的，则称其为平方和凹 。接下来我们证明，对于每个 $k \ge d$ ，如果 (1.1) 由平方和凹多项式定义，则 $f_{\max} = f_{\mathrm{mom,\ k}}$ 。

定理 4.5。 假设对于所有 $j \in \mathcal{I}$， c_j 都是平方和凹的； $c_{\mathcal{E}}$ 是一个线性多项式组，并且对于所有 $x \in K$，每个 $p_i(x)$ （ $i \in [m]$ ）都是平方和凹的。那么对于所有 $k \ge d$，$f_{\mathrm{mom,\ k}} = f_{\max}$。此外，如果 *(3.7)* 存在优化器 y^*，那么 $v^* = (y_{e_1}^*, \ldots\ldots, y_{e_n}^*)$ 是(1.1)的全局最大化点。

证明。 给定 $k \ge d$，设 y为 (3.7) 的一个可行点，并记 $\pi(y) := (y_{e_1}, \ldots, y_{e_n})$。对于每个 $j \in \mathcal{I}$， $\langle c_j, y \rangle \ge 0$ 因为 $L_{c_j}^{(k)}[y] \succeq 0$。对于每个 $j \in \mathcal{E}$，由于 c_j 是线性的，我们有 $c_j(\pi(y)) = 0$ ，因为它对应于 $L_{c_j}^{(k)}[y]$ 的第 (1，1) 个元素。由于 $M_d[y] \succeq 0$和 $y_0 = 1$，平方和凹性与詹森不等式 [24, 定理 7.1.6] 蕴含

$$\langle p_i, y \rangle \le p_i(\pi(y)), \quad 0 \le \langle c_j, y \rangle \le c_j(\pi(y)),$$

对所有 $i \in [m]$ 和 $j \in \mathcal{I}$均成立。因此，我们有 $\pi(y) \in K$ ，并且

$$\sum_{i=1}^{m} a_i \log\langle p_i, y \rangle \le \sum_{i=1}^{m} a_i \log p_i(\pi(y)) \le f_{\max}. \tag{4.6}$$

注意，松弛阶数 k 和可行点 y 可以任意选取。通过在 (3.7) 的可行集上对 y 最大化 (4.6)，我们得到

$f_{\mathrm{mom,\ k}} \le f_{\max}$，这意味着 $f_{\mathrm{mom,\ k}} = f_{\max}$ 对所有 $k \ge d$ 都成立。假设 (3.7) 存在一个优化器 y^*。那么根据定理 4.4，$v^* = \pi(y^*)$ 是 (1.1) 的一个全局最大化点。□

定理 4.5 为在凸集上优化平方和凹函数乘积提供了一种有效的思路。考虑如下形式的多项式优化问题

$$\left\{\begin{array}{ll}\max\limits_{x\in\mathbb{R}^n} & \theta(x) \coloneqq p_1(x)p_2(x)\cdots p_m(x)\\ s.t. & c_j(x) = 0\ (j\in\mathcal{E}),\\ & c_j(x)\ge 0\ (j\in\mathcal{I}).\end{array}\right. \tag{4.7}$$

令 K 表示 (4.7) 的可行集。假设 p_1、……、p_m 在 K 上平方和凹（不一定互不相同），对于 $j\in\mathcal{E}$，每个 c_j 都是线性的，并且对于 $j\in\mathcal{I}$，每个 c_j 都是平方和凹。当 $m>1$ 时，其目标函数 $\theta(x)$ 通常不是凹函数。因此，(4.7) 的标准矩-SOS松弛在最低松弛阶数下可能并不紧。若对于 $i\in[m]$，每个 $p_i(x)$ 在 K 上均为正，则 (4.7) 可以等价地转化为如下对数多项式优化问题

$$\max_{x\in K} \log\theta(x) = \sum_{i=1}^{m}\log p_i(x).$$

根据定理 4.5，重构后的问题在最低松弛阶数下具有紧的矩松弛。

5. Tighter relaxations with LMEs

本节介绍一种称为拉格朗日乘子表达式的技术，用于构造比（3.7）更紧的（1.1）的矩松弛。

为方便起见，假设 $\mathcal{E}\cup\mathcal{I} = [\ell] \coloneqq \{1,\dots,\ell\}$。令 $f(x)$ 为目标函数，x^* 为（1.1）的极大化点。在满足某些约束资格条件时，存在一个拉格朗日乘子向量 $\lambda = (\lambda_j)$，对应于 $j=1,\dots,\ell$，使得 $(x^*,\ \lambda)\in\mathcal{K}$，其中

$$\mathcal{K} \coloneqq \left\{(x,\lambda)\in K\times\mathbb{R}^{\ell}\ \middle|\ \begin{array}{c}\nabla f(x) + \sum\limits_{j\in\mathcal{E}\cup\mathcal{I}}\lambda_j\nabla c_j(x) = 0,\\ \lambda_j\ge 0,\ \lambda_j c_j(x)\ge 0\,(j\in\mathcal{I})\end{array}\right\}. \tag{5.1}$$

（5.1）中的约束称为 Karush-Kuch-Tucker（KKT）最优性条件。每个 $(x,\ \lambda)\in\mathcal{K}$ 都称为一个临界对，其中仅 x 称为临界点。在满足某些约束资格条件时，对数多项式优化（1.1）的所有极大化点都是临界点。假设存在一个多项式或有理元组 $\tau = (\tau_1,\dots,\tau_\ell)$，使得

$$\lambda_j = \tau_j(x) \quad \text{for all} \quad (x,\lambda)\in\mathcal{K}. \tag{5.2}$$

因此，将（5.1）中的 λ 替换为 τ，即可得到临界点的参数化近似。这样的 τ 称为拉格朗日乘子表达式（LME）。对数矩不等式是构造多项式优化 [27]中紧松弛的有效工具。近年来，它们已被应用于由多项式给出的复杂优化问题，包括双层优化、广义纳什均衡问题和变分不等式 [7, 28, 29, 30, 31]。

例5.1。我们给出盒约束、单纯形约束和球约束的显式对数矩不等式。

(i) 若 $K=\{x\in\mathbb{R}^n : a_j\le x_j\le b_j,\ \ j\in[n]\}$，则 $\ell=2n$ 且对于 $\tau=(\tau_1,\ \ldots,\ \tau_{2n})$，(5.2) 成立，其中

$$\tau_{2j-1}=\frac{\nabla_{x_j}f(x)(x_j-b_j)}{b_j-a_j},\ \tau_{2j}=\frac{\nabla_{x_j}f(x)(x_j-a_j)}{b_j-a_j}\quad\forall j\in[n].$$

(ii) 如果 $K=\{x\in\mathbb{R}^n:\ 1-\mathbf{1}^Tx\ge 0,\ \ x\ge 0\}$，那么 $\ell=n+1$，且（5.2）对以下情况成立

$$\tau=x^T\nabla f(x)\mathbf{1}-(0,\nabla_{x_1}f(x),\ldots,\nabla_{x_n}f(x)),$$

其中，$\mathbf{1}\in\mathbb{R}^n$ 是全1向量。

(iii) 如果 $K=\{x\in\mathbb{R}^n:\ \|x-a\|\le b\}$，那么 $\ell=1$，且（5.2）对以下情况成立

$$\tau=\frac{(x-a)^T\nabla f(x)}{2b^2}.$$

接下来，我们介绍如何应用对数矩不等式为（1.1）构造更紧的矩松弛。考虑具有盒约束、单纯形约束或球约束的对数多项式优化问题（1.1）。然后，例5.1中明确给出了适用于 $j\in\mathcal{E}\cup\mathcal{U}$ 的多项式对数矩阵等式 $\tau_j(x)$。回顾

$$f(x)=\sum_{i=1}^m a_i\log p_i(x),\quad \nabla f(x)=\sum_{i=1}^m\frac{a_i}{p_i(x)}\nabla p_i(x).$$

为方便起见，记 $p^{\mathbf{1}}=p_1p_2\cdots p_m$ 和

$$f_i(x)\coloneqq\frac{p^{\mathbf{1}}(x)}{p_i(x)},\quad\forall i\in[m].$$

假设 $p>0$ 在 K上。那么，可行点 $x\in K$ 是 (1.1) 的临界点，当且仅当它满足

$$\begin{cases}\sum\limits_{i=1}^m a_if_i(x)\nabla p_i(x)+p^{\mathbf{1}}(x)\sum\limits_{j\in\mathcal{E}\cup\mathcal{I}}\tau_j(x)\nabla c_j(x)=0,\\ \tau_j(x)\ge 0\,(j\in\mathcal{I}),\quad \tau_j(x)c_j(x)=0\,(j\in\mathcal{I}).\end{cases}$$

该多项式系统为 (1.1) 提供了强化约束。为方便起见，记多项式元组为

$$\Phi\coloneqq\{c_{\mathcal{E}}\}\cup\Big\{\sum_{i=1}^m a_if_i(x)\nabla p_i(x)+p^{\mathbf{1}}(x)\sum_{j\in\mathcal{E}\cup\mathcal{I}}\tau_j(x)\nabla c_j(x)\Big\}\cup\{\tau_jc_j:j\in\mathcal{I}\},$$
$$\Psi\coloneqq\{c_{\mathcal{I}}\}\cup\{\tau_j:j\in\mathcal{I}\}.$$

对于由多项式组成的向量 ϕ，$\{\phi\}$ 表示由 ϕ 的各项构成的多项式集合。我们可以将 (1.1) 等价地重写为

$$\begin{cases}\max\limits_{x\in\mathbb{R}^n} & \sum\limits_{i=1}^m a_i\log p_i(x)\\ s.t. & \varphi(x)=0\,(\varphi\in\Phi),\\ & \psi(x)\ge 0\,(\psi\in\Psi).\end{cases}\tag{5.3}$$

值得注意的是，(5.3) 的可行集记为

$$K_1\coloneqq\{x\in\mathbb{R}^n:\varphi(x)=0\,(\varphi\in\Phi),\,\psi(x)\ge 0\,(\psi\in\Psi)\},$$

通常是 K 的真子集。(5.3) 也是一个对数多项式优化问题。因此，可以类似于 (3.4) 得到它的凸松弛：

$$\left\{\begin{array}{ll} f_{\mathrm{lme}} \coloneqq \max\limits_{z} & \sum\limits_{i=1}^{m} a_i \log\langle p_i, z\rangle \\ s.t. & z_0 = 1,\ z \in \mathscr{R}_d(K_1). \end{array}\right. \tag{5.4}$$

令 f_{lme} 表示 (5.5) 的最优值。由于 $K_1 \subseteq K$，必有 $f_{\mathrm{lme}} \leq f_{\mathrm{mom}}$，其中 f_{mom} 是标准凸松弛 (3.4) 的最优值。此外，其矩松弛层级可以类似于 (3.7) 构造。令

$$d_1 \coloneqq \max\{\lceil \deg(q) \rceil : q \in \Phi \cup \Psi\}.$$

(5.4) 的 k 阶矩松弛（对每个 $k \geq d_1$）为

$$\left\{\begin{array}{ll} f_{\mathrm{lme},k} \coloneqq \max\limits_{y} & \sum\limits_{i=1}^{m} a_i \log\langle p_i, y\rangle \\ s.t. & y_0 = 1,\ y \in \mathbb{R}^{\mathbb{N}^n_{2k}},\ M_k[y] \succeq 0, \\ & L_{\varphi}^{(k)}[y] = 0\,(\varphi \in \Phi), \\ & L_{\psi}^{(k)}[y] \succeq 0\,(\psi \in \Psi). \end{array}\right. \tag{5.5}$$

令$f_{\mathrm{lme},\ k}$ 表示 (5.5) 的最优值。由于 (3.7) 的所有约束都包含在 (5.5) 中，(3.7) 的可行集是 (5.5) 对应可行集的子集。因此，对所有 $k \geq d_1$，有 $f_{\mathrm{lme},\ k} \leq f_{\mathrm{mom},\ k}$。

更紧松弛也可以类似地构造并推广到更一般的情形。假设对所有 $x \in K$，线性无关约束资格条件（LICQC）均成立，且所有 p_i 在 K 上均为正。根据 [29] 中的命题 3.6，存在分母在 K上严格为正的有理数 LME。综上，我们得到以下结果。

命题 5.2。 假设存在有理数 *LME* $\tau = (\tau_1, \ldots, \tau_l)$ 满足(5.2)，且其分母在 K 上严格为正。那么凸问题 (5.4) 和(5.5) 是 (1.1)的松弛，其最优值满足

$$f_{\mathrm{mom},k} \quad \geq \quad f_{\mathrm{mom}} \quad \geq \quad f_{\max}$$

$$f_{\mathrm{lme},k} \quad \geq \quad f_{\mathrm{lme}} \quad \geq \quad f_{\max}.$$

我们指出，可能出现 $f_{\mathrm{lme}} < f_{\mathrm{mom}}$，因此对于所有$k$，$f_{\mathrm{mom},k} > f_{\max}$，但利用对数矩不等式的矩松弛层级 (5.5) 是紧的。此类情况可参见例 6.2。此外，我们还指出，定理 3.4、4.1、4.4、4.5 以及推论 3.5、4.2 中的类似结论，对更紧松弛 (5.5) 同样成立。为简洁起见，我们省略详细证明，并在下文给出一个例子，以更好地说明 (3.7) 与 (5.5) 之间的性能差异。

例 5.3。 考虑由 ABO 血型系统 [23] 推导出的对数多项式优化问题：

$$\left\{\begin{array}{ll} \max\limits_{x \in \mathbb{R}^3} & 182\log(x_1^2 + 2x_1x_3) + 60\log(x_2^2 + 2x_2x_3) \\ & \quad +17\log(2x_1x_2) + 176\log(x_3^2) \\ s.t. & 1 - \mathbf{1}^T x \geq 0,\ x \geq \mathbf{0}, \end{array}\right. \tag{5.6}$$

其中，$\mathbf{1} \in \mathbb{R}^3$ 是全 1 向量，$\mathbf{0} \in \mathbb{R}^3$ 是全 0 向量。该优化问题具有单纯形约束，其对数矩不等式明确为

见例5.1 (ii)。我们比较标准矩松弛 (3.7) 与更紧松弛 (5.5) 的数值表现。所有计算均使用 YALMIP [19]和半正定规划求解器 MOSEK [3]完成。对于更紧的矩松弛 (5.5)，当 $r=1$ 时，$t=k=3$ 满足平坦截断条件 (4.1)。根据推论 4.2，我们得到 (5.6) 的最优值和优化器为

$$f_{\max}=f_{\mathrm{lme},3}=-492.5353,\quad x^*=(0.2644,0.0932,0.6424).$$

详细的计算结果见表1。显然,

k	2	3	4	5	6
$f_{\mathrm{mom},k}$	-491.8158	-491.8158	-491.8158	-491.8158	-491.8158
$f_{\mathrm{lme},k}$	-492.2927	-492.5353	-492.5353	-492.5353	-492.5353

TABLE 1. 例 5.3 中$f_{\mathrm{mom},k}$ 与$f_{\mathrm{lme},k}$ 的比较

与标准矩松弛 (3.7) 相比，更紧松弛 (5.5) 求解对数多项式优化问题 (5.6) 的效率更高。

实际上，在求解对数多项式优化问题时，选择 (3.7) 还是 (5.5) 可能颇为棘手。请注意，次数 $d_1\geq d$ 和 d_1 取决于 m。通常，当 m 较小且 (1.1) 具有盒约束、单纯形约束和球约束等简单约束时，我们更倾向于采用更紧松弛 (5.5)。这是因为在这种情况下，最低松弛阶数 d，d_1 相近，而且 (1.1) 具有便于处理的对数矩不等式。因此，预期在较低松弛阶数下求解 (5.5)，相比 (3.7) 能得到更优的上界。然而，随着 m 增大，由于存在乘积项 $p^{\mathbf{1}}$ 和 f_i，间隙 d_1-d 可能会显著增大。在这种情况下，即使采用初始松弛阶数 $k=d_1$，求解 (5.5) 的计算开销也可能很大。因此，对于较大的 m，使用标准矩层级 (3.7) 进行计算通常更有效。

6. NUMERICAL EXPERIMENTS AND APPLICATIONS

本节给出对数多项式优化的具体数值示例及应用。矩松弛（3.7）使用软件 YALMIP [19] 和半正定规划求解器 MOSEK [3]求解。计算在 MATLAB R2024b 中完成，所用笔记本电脑配备第 12 代 Intel(R) Core(TM) i7-1270P 2.20GHz CPU 和 32GB 内存。为提高可读性，计算结果保留四位小数。

6.1. **具体数值示例**。

例6.1. 考虑如下对数多项式优化问题:

$$\begin{cases}\max\limits_{x\in\mathbb{R}^5} & \frac{30}{218}\log p_1(x)+\frac{97}{218}\log p_2(x)+\frac{91}{218}\log p_3(x)\\ s.t. & \|x\|\leq 1,\end{cases}$$

其中

$$\begin{aligned}p_1(x) &= (x_3+x_4)^2+\left(1+2(x_1+x_5)+3(x_2+x_4)\right)^2+0.01,\\ p_2(x) &= (x_2+x_4-x_5)^2+(2x_2+3x_5)^2+0.02,\\ p_3(x) &= (x_1+x_3+x_4)^2+(x_1-x_3+x_4)^2+0.03.\end{aligned}$$

显然，目标函数在可行集上定义良好。在 $k=2$ 处求解松弛问题 (3.7) 耗时 1.10 秒。相应的矩

优化器在 $t = 2$处满足平坦截断条件 (4.1)，其中 $r = 1$。根据推论 4.2，矩松弛是紧的。我们得到全局最优值$f_{\max} = f_{\mathrm{mom},\ 2} = 1.6216$，以及全局最大化点

$$x^* = (0.4717, 0.4518, 0.0036, 0.5276, 0.5432).$$

例 6.2。考虑以下对数多项式优化问题：

$$\begin{cases} \max\limits_{x\in\mathbb{R}^3} & \sum\limits_{i=1}^{6} a_i \log p_i(x) \\ s.t. & 1 - \mathbf{1}^T x \geq 0,\ x \geq \mathbf{0}, \end{cases} \tag{6.1}$$

其中，$a = (a_1、\ldots、a_6)$ 是给定的系数向量，且

$$\begin{array}{lll} p_1(x) = x_1^3 + 3x_1^2x_2 + 3x_1^2x_3, & p_2(x) = 3x_1x_2^2 + 6x_1x_2x_3, & p_3(x) = 3x_1x_3^2, \\ p_4(x) = x_2^3 + 3x_2^2x_3, & p_5(x) = 3x_2x_3^2, & p_6(x) = x_3^3, \end{array}$$

问题 (6.1) 具有单纯形约束，其对数矩不等式已在例 5.1(ii) 中明确给出。

(i) 考虑系数向量

$$a = (0.0968,\ 0.1419,\ 0.2194,\ 0.0839,\ 0.2839,\ 0.1742).$$

对于标准矩松弛 (3.7)，最低松弛阶数为 $d = 2$。当松弛阶数为 $k = 4$ 时，求解 $f_{\mathrm{mom},\ 4} =$−1.7181 大约耗时 1.77 秒；其中，相应的矩优化器满足平坦截断条件，且 $t = 3$ 和 $r = 2$，从而可得到分解 (4.2)：

$$\begin{array}{ll} v_1^* = (0.2195, 0.1929, 0.5877), & f(v_1^*) = -1.7308, \\ v_2^* = (0.0537, 0.3100, 0.6363), & f(v_2^*) = -2.0655. \end{array}$$

根据定理 4.1，我们有 $f_{\mathrm{mom}} = f_{\mathrm{mom},\ 4}$。但这并不能保证原问题的最优性。注意，$f(v_1^*)$ $f(v_2^*)$ 为真实最优值 $f_{\max}$ 提供下界。因此，我们得到

$$|f_{\max} - f_{\mathrm{mom}}| \leq |f(v_1^*) - f_{\mathrm{mom},4}| \leq 0.0127.$$

对于更紧的矩松弛 (5.5)，最低松弛阶数为 $d_1 = 3$。在这一最低松弛阶数 $k = 3$ 下，得到 $f_{\mathrm{lme},\ 3} =$−1.7194 大约耗时 1.24 秒；其中，相应的矩优化器满足平坦截断条件，且 $t = 3$ 和 $r = 1$。根据推论 4.2，该矩松弛是紧的。我们得到全局最优值和优化器

$$f_{\max} = f_{\mathrm{lme},3} = -1.7194, \quad x^* = (0.1873, 0.2161, 0.5967).$$

(ii) 令系数向量 a 由 MATLAB 函数 rand 随机生成。表 2 报告了五个实例的计算结果。我们分别用 k_{mom} 和k_{lme} 表示在满足平坦截断条件时，(3.7) 和 (5.5) 的松弛阶数。对于松弛 (3.7)，数值 r 是满足平坦截断条件的矩优化器所对应矩阵的秩。“间隙”一列列出了 f_{mom} 与 $f(v_i^*)$之间的最小数值差，其中v_i^* 是矩优化器 (3.7) 所允许的原子测度支撑集中的任意一点。对于更紧的矩松弛 (5.5)，在全部五个实例中，其矩优化器均满足平坦截断条件，且 $t = 3$ 和 $r = 1$ 。根据推论 4.2，矩松弛是紧的，即 $f_{\max} = f_{\mathrm{lme}}$。

实例	松弛 (3.7)				松弛 (5.5)	
	k_{mom}	r	f_{mom}	gap	k_{lme}	$f_{\max}=f_{\mathrm{lme}}$
#1	4	2	-4.6833	0.2033	3	-4.6968
#2	4	3	-3.9524	0.0225	3	-3.9658
#3	5	3	-2.7490	0.5442	3	-2.8979
#4	6	3	-6.6434	0.3871	3	-6.7018
#5	5	4	-7.0067	0.5815	3	-7.0672

TABLE 2. 例 6.2 (ii) 的计算结果

例6.3。 考虑以下多项式优化问题:

$$\begin{cases} \max\limits_{x\in\mathbb{R}^5} & \theta(x)=\big((x_3+x_4+x_5)^2-x_1x_2\big)^{20}\big(8-(x_5-x_2)^2\big)^{25} \\ s.t. & \|x\|\le 2,\ (x_3+x_4+x_5)^2-x_1x_2\ge 0, \\ & 8-(x_5-x_2)^2\ge 0. \end{cases}$$

由于 $\deg(\theta)=45$，直接应用矩-SOS松弛的计算开销很大。我们可以将问题重新表述为一个等价的对数多项式优化问题来求解。存在两种可能的重新表述方式:

$$\begin{cases} \max\limits_{x\in\mathbb{R}^5} & 20\log\big((x_3+x_4+x_5)^2-x_1x_2\big)+25\log\big(8-(x_5-x_2)^2\big) \\ s.t. & \|x\|\le 2,\ (x_3+x_4+x_5)^2-x_1x_2\ge 0, \\ & 8-(x_5-x_2)^2\ge 0; \end{cases} \tag{6.2}$$

$$\begin{cases} \max\limits_{x\in\mathbb{R}^5} & 20\log\big((x_3+x_4+x_5)^2-x_1x_2\big) \\ & \quad +25\log\big(2\sqrt{2}-(x_5-x_2)\big)+25\log\big(2\sqrt{2}+(x_5-x_2)\big) \\ s.t. & \|x\|\le 2,\ (x_3+x_4+x_5)^2-x_1x_2\ge 0, \\ & 8-(x_5-x_2)^2\ge 0. \end{cases} \tag{6.3}$$

由于这些对数多项式优化问题的约束相对复杂，我们考虑问题 (6.2) 的标准矩松弛 (3.7)。

(i) 对于问题 (6.2)，在松弛阶数 $k=2$ 下，约用 1 秒便得到了 $f_{\mathrm{mom},\,2}=99.7252$。相应的矩优化器在 $t=2$ 处满足平坦截断条件 (4.1)。它具有如下分解 (4.2):

$$\begin{aligned} v_1^* &= (0.0673,-0.3654,-1.2400,-1.2400,-0.8870),\\ v_2^* &= (-0.0673,0.3654,1.2400,1.2400,0.8870). \end{aligned}$$

记多项式 $p_1(x)=(x_3+x_4+x_5)^2-x_1x_2$ 和 $p_2(x)=8-(x_5-x_2)^2$。由于 p_1, p_2 是偶函数，因此显然有 $p_1(v_1^*)=p_1(v_2^*)$ 和 $p_2(v_1^*)=p_2(v_2^*)$。因此，根据定理 4.1，我们有

$$f_{\max}=f_{\mathrm{mom},2}=99.7252$$

且两者都是原优化问题的全局最大化点。

(ii) 对于问题 (6.3)，在松弛阶数 $k=2$ 下，求解耗时 1.21 秒$f_{\mathrm{mom},\,2}=101.6842$。相应的矩优化器满足平坦截断条件 (4.1)，其中 $t=2$。它具有分解式 (4.2)，其中

$$\begin{aligned} v_1^* &= (0.0001,-0.0002,-1.1546,-1.1546,-1.1546),\\ v_2^* &= (-0.0001,0.0002,1.1546,1.1546,1.1546). \end{aligned}$$

记多项式 $p_1(x) = (x_3 + x_4 + x_5)^2 - x_1x_2$、 $p_2(x) = 2\sqrt{2} - (x_5 - x_2)$ 和 $p_3(x) = 2\sqrt{2} + (x_5 - x_2)$。由于 $p_2(v_1^*) \neq p_2(v_2^*)$，定理 4.1 的条件不满足。根据 (i) 中的计算结果，(6.3) 的矩松弛并不紧。

例 6.4。考虑如下形式的对数多项式优化问题

$$\begin{cases} \max\limits_{x\in\mathbb{R}^{10}} & \sum\limits_{i=1}^{5} a_i \log p_i(x) \\ s.t. & p_i(x) - i \geq 0, \quad (i \in [5]), \\ & 6 - (x_6 - x_7 - x_8)^2 - x_9^2 \geq 0, \\ & 7 - (x_1 - x_{10})^2 - (x_2 - x_9)^2 \geq 0, \\ & (x_1 + x_2 + x_3)^2 + (x_4 + x_5)^2 - 8 \geq 0, \\ & (x_2 + x_4 - x_6 + x_8 - x_{10})^2 - 9 \geq 0, \end{cases}$$

其中每个 $p_i(x)$ 具有如下形式

$$p_i(x) := \alpha_i - \big(\beta_i(x_{2i-1} + b_i)^2 + \gamma_i(x_{2i} + c_i)^2 + \delta_i(x_{2i-1} + d_i)(x_{2i} + e_i)\big).$$

设权重 a_i 和 p_i 的参数如下。

i	a_i	α_i	β_i	γ_i	δ_i	b_i	c_i	d_i	e_i
1	3	10	2	3	-3	0.5	0.5	-0.5	0.5
2	2	12	3	2	-1	-0.6	0.1	-0.6	-0.1
3	1	11	4	1	0	0.7	-0.2	0	0
4	1	15	2	5	-2	-0.8	0.3	-0.8	0.3
5	5	13	-3	2	0	-0.9	0.4	0	0

通过检查它们的海森矩阵的半正定性，可以很容易地验证 $p_1, \ldots, p_4$ 是凹函数，而 p_5 不是。求解 $k = 2$的松弛问题 (3.7) 耗时 1.95 秒。 $t = 2$和 $r = 1$满足平坦截断条件 (4.1)。根据推论 4.2，矩松弛是紧的。我们得到全局最优值 $f_{\max} = f_{\text{mom},2} = 36.3994$, 以及优化器

$$x^* = (-1.4038, -1.5957, 0.4259, -0.4182, -0.7548, 0.6579, 0.9798, -0.3904, -2.4485, -0.0622).$$

例 6.5。考虑以下对数多项式优化问题:

$$\begin{cases} \max\limits_{x\in\mathbb{R}^{12}} & \sum\limits_{i=1}^{10} \log p_i(x) \\ s.t. & (x_1 + x_2 + x_3 + x_4 + x_5 + x_6)^2 \leq 15, \\ & (x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12})^2 \leq 15, \\ & (x_1 - x_3 + x_5 - x_7 + x_9 - x_{11})^2 \leq 8, \\ & (x_2 - x_4 + x_6 - x_8 + x_{10} - x_{12})^2 \leq 8, \\ & (x_1 + x_{12})^2 + (x_2 + x_{11})^2 \leq 9, \\ & (x_3 - x_{10})^2 + (x_4 - x_9)^2 + (x_5 - x_8)^2 \leq 9 \\ & \sum\limits_{i=1}^{6} x_{2i}^2 + 2(x_1^2 + x_3^2 + x_7^2 + x_9^2) + 3(x_5^2 + x_{11}^2) \leq 20, \\ & \sum\limits_{i=1}^{6} (x_{2i-1} - x_{2i}) \leq 5, \ (x_1 + x_6 + x_{12})^2 \leq 4, \\ & x_1 - x_{12} \leq 3, \ \mathbf{1}^T x \geq 0, \end{cases}$$

在上式中，对每个 $i = 1,\ldots,10$,

$$p_i(x) \coloneqq \alpha_i - \beta_i\big(h_i(x)\big)^4,$$

其中，对 $i = 1$、.....、5,

$$h_i(x) \coloneqq (x_{2i} + a_i) + (x_{2i+1} + b_i),$$

并且对 $i = 6$、.....、10（$j_1 \bmod j_2$ 表示 j_1 除以 j_2所得的余数）

$$\begin{aligned} h_i(x) \coloneqq \quad & (-1)^{i-4}(x_{i-5} + a_i) + (-1)^{i+1}(x_i + b_i) \\ & +(-1)^{(i+4) \bmod 12}(x_{(i+4) \bmod 12+1} + c_i). \end{aligned}$$

此外， p_i、 h_i 中的参数列于下表:

i	1	2	3	4	5	6	7	8	9	10
α_i	20	25	18	22	30	15	28	19	21	26
β_i	1.5	1.2	2.5	1.1	1.6	1.9	2.3	1.7	2.4	1.4
a_i	-0.2	-1	0.4	0.4	-1.1	-0.8	-0.3	-1.7	1.2	1.7
b_i	-1.6	-0.4	-1	0.8	-1.5	-0.7	0	-1	-1.9	0.9
c_i	-	-	-	-	-	-1.3	0.7	-0.7	0.8	-0.5

通过检验海森矩阵的半正定性，可以很容易地验证所有 p_i 都是平方和凹的，且可行集是凸集。因此，这是一个凹对数多项式优化问题。采用松弛阶数 $k = 2$求解标准矩松弛（3.7）耗时 3.19 秒。相应的矩优化器不满足平坦截断（4.1）。然而，根据定理 4.5，对于凹对数多项式优化问题，矩松弛（3.7）始终是紧的。因此，$f_{\max} = f_{\text{mom},\ 2} = 30.3426$，且得到的全局最大化点为

$$\begin{aligned} x^* = (&1.5452, 0.2145, -0.8337, 1.7120, -1.1664, 0.4258, \\ &-0.2089, 0.3101, 0.5313, -2.1910, 0.4054, -0.7443). \end{aligned}$$

6.2. **应用**。

例6.6（亲子关系分析问题 [22]）。 考虑一个包含 n 个等位基因的单一基因座，其中涉及一位母亲和 n 位可能的父亲。假设有 t 种可能的基因型。令 x_j 表示对于 $j = 1,\ldots,k$，随机选取的后代的生物学父亲是第 j 位男性的概率。我们的目标是利用一窝可能由多个父本繁衍的幼崽的基因型数据，估计概率 $x =$（$x_1,\ldots,$ x_n）。这可以通过求解

$$\left\{ \begin{array}{ll} \max\limits_{x\in\mathbb{R}^n} & \sum\limits_{i=1}^{t} N_i \log\Big(\sum\limits_{j=1}^{n} P_{ij}x_j\Big) \\ s.t. & \mathbf{1}^T x = 1,\ x \geq \mathbf{0}, \end{array} \right. \tag{6.4}$$

其中，每个 N_i 表示一窝幼崽中具有第 i 种可能基因型的后代数量，每个 P_{ij} 是后代从第 j 位潜在父亲处获得第 i 种基因型的给定概率。令 $\bar{x} =$(.....，x_{n-1})。通过代换 $x_n = 1 - \mathbf{1}_{n-1}^T\bar{x}$，我们可以将 (6.4) 重新表述为

$$\left\{ \begin{array}{ll} \max\limits_{\bar{x}\in\mathbb{R}^{n-1}} & \sum\limits_{i=1}^{t} N_i \log\Big(P_{in} + \sum\limits_{j=1}^{n-1}(P_{ij} - P_{in})x_j\Big) \\ s.t. & \bar{x} = (x_1,\ldots,x_{n-1}) \geq 0, \\ & 1 - \mathbf{1}_{n-1}^T\bar{x} \geq 0. \end{array} \right. \tag{6.5}$$

这是一个带有单纯形约束的对数多项式优化问题。因此，可以使用例5.1 (ii) 中的对数矩不等式，通过更紧的矩松弛 (5.5) 求解。令 N_i，P_{ij} 为从 [22] 中选取的模拟数据。我们通过更紧的矩松弛 (5.5) 求解 (6.5)。当松弛是紧的时，令 $x = (\bar{x}, 1-\mathbf{1}^T\bar{x})$，即可恢复原问题 (6.4) 的优化器。对于每个实例，当计算得到的矩优化器满足平坦截断条件 (4.1) 时，相应矩阵的秩为 1。根据推论 4.2，矩松弛对于原问题是紧的。详细的模拟数据和数值结果见表 3。在表中，我们将参数 $N = (N_1, \ldots\ldots, N_t)$ 和 $P = (P_{ij})$ 记为 $\in \mathbb{R}^{t\times n}$。$k$ 表示平坦截断成立时的松弛阶数，r 表示与矩优化器相对应的矩阵的秩。记号 x^* 表示计算得到的原对数多项式优化问题 (6.4) 的优化器。“time”列报告了求解第 k 阶更紧矩松弛所用的时间，以秒计。

n	t	N	P	k	r	x^*	time
2	2	$\begin{bmatrix}77\\23\end{bmatrix}$	$\begin{bmatrix}0.5 & 1\\ & 0\end{bmatrix}$	2	1	$\begin{bmatrix}0.4600\\0.5400\end{bmatrix}$	1.01
2	2	$\begin{bmatrix}63\\37\end{bmatrix}$	$\begin{bmatrix} & 1\\0.5 & 0\end{bmatrix}$	2	1	$\begin{bmatrix}0.7400\\0.2600\end{bmatrix}$	1.35
2	3	$\begin{bmatrix}49\\40\\11\end{bmatrix}$	$\begin{bmatrix}0.5 & 0.875\\0.25 & 0.125\\0.25 & 0\end{bmatrix}$	3	1	$\begin{bmatrix}0.8813\\0.1187\end{bmatrix}$	1.40
2	3	$\begin{bmatrix}83\\2\\15\end{bmatrix}$	$\begin{bmatrix}0.5 & 0.25\\0.5 & 0.25\\0 & 0\end{bmatrix}$	2	1	$\begin{bmatrix}0.6939\\0.3061\end{bmatrix}$	1.61
2	3	$\begin{bmatrix}63\\17\\20\end{bmatrix}$	$\begin{bmatrix}0.5 & 0.25\\0.5 & 0.25\\0 & 0\end{bmatrix}$	2	1	$\begin{bmatrix}0.5181\\0.4819\end{bmatrix}$	1.10
2	4	$\begin{bmatrix}59\\8\\16\\17\end{bmatrix}$	$\begin{bmatrix}0.25 & 0.5\\0.25 & 0.5\\0.25 & 0\\0 & 0\end{bmatrix}$	3	1	$\begin{bmatrix}0.6599\\0.3401\end{bmatrix}$	1.03
2	5	$\begin{bmatrix}7\\9\\4\\33\\47\end{bmatrix}$	$\begin{bmatrix}0.25 & 0\\0.25 & 0\\0.25 & 0\\0.25 & 0.5\\0 & 0\end{bmatrix}$	4	1	$\begin{bmatrix}0.2463\\0.7537\end{bmatrix}$	1.49
3	3	$\begin{bmatrix}39\\38\\23\end{bmatrix}$	$\begin{bmatrix}0.5 & 0 & 0.875\\0.25 & 0.75 & 0.125\\0.25 & 0.25 & 0\end{bmatrix}$	3	1	$\begin{bmatrix}0.6400\\0.2800\\0.0800\end{bmatrix}$	1.18
3	4	$\begin{bmatrix}29\\21\\88\\62\end{bmatrix}$	$\begin{bmatrix} & 0 & 0\\ & 0.75 & 0.25\\0.25 & 0.25 & \\0 & 0 & \end{bmatrix}$	3	1	$\begin{bmatrix}0.2240\\0.0000\\0.7760\end{bmatrix}$	1.14

TABLE 3. 例6.6的数值结果

例 6.7（潜在类别模型 [8]）。潜在类别模型（LCM） [22] 是一类有限混合模型，用于根据观测到的类别数据识别总体中隐藏的、不可观测的类别组。设 $y = (y_1, \ldots\ldots, y_d)$ 为类别随机变量向量，其中每个 y_j 都从 c_j 个类别中取值，例如 $\{1, \ldots, c_j\}$。假设每个 y_j 相互独立地分布，其分布取决于 T个子组中的一个，该子组由指示变量 $s \in \{1, \ldots, T\}$决定。令$I(\cdot)$表示艾弗森括号函数，使得当 P 为真时，$I(P) = 1$，否则为零。则 y 在第 t个子组中的概率密度函数为

$$\prod_{j=1}^{d} \prod_{l=1}^{c_j} \pi_{t,j,l}^{I(y_j=l)} \quad \text{with} \quad \pi_{t,j,l} = \Pr(y_j = l | s = t),$$

其中，Pr(·) 是概率表示法。潜在类别模型的总体混合密度随后可表示为加权和

$$p(y|\eta, \pi) = \sum_{t=1}^{T} \Big(\eta_t \prod_{j=1}^{d} \prod_{l=1}^{c_j} \pi_{t,j,l}^{I(y_j=l)} \Big),$$

其中，每个 $\eta_t \geq 0$ 表示第 t个子组的权重，所有权重之和为一。考虑从 LCM 分布中抽取的 N个样本，记为 $y^{(1)}, \ldots\ldots, y^{(N)}$。则 LCM 的最大似然估计可表述为

$$\left\{ \begin{array}{ll} \max\limits_{\eta, \pi} & \sum\limits_{i=1}^{N} \log \left(\sum\limits_{t=1}^{T} \eta_t \prod\limits_{j=1}^{d} \prod\limits_{l=1}^{c_j} \pi_{t,j,l}^{I(y_j^{(i)}=l)} \right) \\ s.t. & \eta = (\eta_1, \ldots, \eta_T), \\ & \pi = (\pi_{t,j,l})_{t \in [T], j \in [d], l \in [c_j]}, \\ & \eta \geq 0, \, \pi \geq 0, \, \sum\limits_{t=1}^{T} \eta_t = 1, \\ & \sum\limits_{l=1}^{c_j} \pi_{t,j,l} = 1 \, (t \in [T], \, j \in [d]). \end{array} \right. \tag{6.6}$$

我们考虑三个实例，并采用与 [8]相同的设置。对于每个实例，我们从相应的 LCM 中模拟生成 $N = 500$ 组样本，重复十次，并在初始阶数下用其标准矩松弛 (3.7) 近似 (6.6)。通过将所得近似值与真实参数的对数似然进行比较，评估近似质量；该对数似然见 [8]。因此，近似质量取决于所选的矩松弛阶数和采样数据。最佳数值结果如下。

(i) 对于 $d = 1$, $T = 2$, 假设每个 y_i 都是一个二元变量，即 $y_i \in \{0, 1\}$。令

$$\begin{array}{llll} \eta_1 = 0.5, & \eta_2 = 0.5, & \pi_{1,1,1} = 0.4, & \pi_{1,1,2} = 0.6, \\ \pi_{2,1,1} = 0.8, & \pi_{2,1,2} = 0.2. & & \end{array}$$

在我们的数值实验中，我们花了约 0.78 秒得到矩上界$f_{\text{mom, 2}} = -335.2519$。这一结果与真实的对数似然非常接近 d $f^* = -335.29$。

(ii) 对于 $d = 1$, $T = 3$, 假设每个 y_i 都是一个二元变量，即 $y_i \in \{0, 1\}$。令

$$\begin{array}{llllll} \eta_1 = 0.5, & \eta_2 = 0.3, & \eta_3 = 0.2, & \pi_{1,1,1} = 0.4, & \pi_{1,1,2} = 0.6, \\ \pi_{2,1,1} = 0.8, & \pi_{2,1,2} = 0.2 & \pi_{3,1,1} = 0.1, & \pi_{3,1,2} = 0.9. & & \end{array}$$

在我们的数值实验中，我们花了约 0.57 秒得到了矩上界 $f_{\text{mom, 2}} = -344.45$。这个结果非常接近真实的对数似然$f^* = -344.49$。

(iii) 对于 $d = 2$，$T = 2$，假设每个 y_i 都是二元变量，即 $y_i \in \{0, 1\}$。令

$$\eta_1 = 0.5, \quad \eta_2 = 0.5, \quad \pi_{1,1,1} = 0.4, \quad \pi_{1,1,2} = 0.6, \quad \pi_{2,1,1} = 0.8,$$
$$\pi_{2,1,2} = 0.2 \quad \pi_{1,2,1} = 0.1, \quad \pi_{1,2,2} = 0.9 \quad \pi_{2,2,1} = 0.6, \quad \pi_{2,2,2} = 0.4.$$

在我们的数值实验中，我们花了约 0.53 秒得到矩上界 $f_{\text{mom, 2}} = -657.79$。这一结果接近真实的对数似然$f^* = -661.30$。

7. Conclusion and Discussions

本文研究了对数多项式优化问题，其特征是目标函数为对数多项式，并带有多项式等式约束和不等式约束。利用截断的 K-矩问题，我们为原始非凸表述推导出一系列矩松弛。在特定条件下，我们证明松弛是紧的，并能给出精确的全局最优解。我们还建立了用于判定紧性以及在全局解存在时提取全局解的准则。此外，我们引入拉格朗日乘子表达式这一技术，为带有盒约束、单纯形约束和球约束的对数多项式优化构造更紧的松弛。命题 5.2 总结了这两类矩松弛与原始对数多项式优化之间的关系。

所提出的方法在最大似然估计等应用，以及涉及熵相关优化的更广泛领域中具有实际意义。如何提高这种矩松弛方法的可扩展性，是一个值得探索的未来研究方向。一种有前景的途径是利用矩阵的稀疏性和结构。另一项潜在工作是将该框架扩展到非多项式约束或动态约束。

致谢 JIAWANG NIE 获得美国国家科学基金会资助（DMS- 2513254）的部分支持，XINDONG TANG 获得 HKBU-15303423 和HKBU- 22304125 的部分支持。

References

[1] A. A. Ahmadi, A. Olshevsky, P. A. Parrilo, and J. N. Tsitsiklis, *NP-hardness of deciding convexity of quartic polynomials and related problems* 137, 453 ，《数学规划》，–476，2013年。[2] N. Anari, K. Liu, S.O. Gharan and C. Vinzant, *Log-concave polynomials II: High-dimensional walks and an FPRAS for counting bases of a matroid* 2 ，《数学年刊》（）199（1）259 - 299，2024年。[3] MOSEK ApS, *The MOSEK Optimization Toolbox for MATLAB Manual*，9.0 版，2019年。[4] O. Banerjee, L. El Ghaoui, and A. d'Aspremont, *Model selection through sparse maximum likelihood estimation for multivariate Gaussian or binary data*，《机器学习研究杂志》，9，第 485–516 页，2008 年。[5] C. M. Bishop, *Pattern recognition and machine learning. Information Science and Statistics* 2006 ，施普林格，纽约，。[6] T. Carpenter, I. Diakonikolas, A. Sidiropoulos, and A. Stewart, *Near-optimal sample complexity bounds for maximum likelihood estimation of multivariate log-concave densities*，载于第 31 届学习理论会议（COLT）论文集，75，第 1234–1262 页，2018 年。[7] J. Choi, J. Nie, X. Tang, and S. Zhong, *Generalized Nash equilibrium problems with quasi-linear constraints* arXiv:2405.03926 2024 ，预印本，，。

[8] H. Chen, L. Han, and A. Lim, *Beyond the EM algorithm: constrained optimization methods for latent class model*《统计学通讯：模拟与计算》，51，5222- 5244，2020年。[9] A. P. Dempster, N. M. Laird, and D. B. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*，《皇家统计学会会刊》。B辑（方法论），39(1)，第 1–38 页，1977年。[10] R. A. Fisher, *On the mathematical foundations of theoretical statistics*，《伦敦皇家学会哲学汇刊》。A辑，数学或物理学论文集，222，第 309–368 页，1922年。[11] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*，麻省理工学院出版社，马萨诸塞州剑桥，2016年。[12] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio *Generative adversarial nets* 27 , , ，载于第届神经信息处理系统国际会议（NeurIPS）论文集，2，第 2672– 2680 页，2014年。[13] G. S. Gusmão and A. J. Medford, *Maximum-likelihood estimators in physics-informed neural networks for high-dimensional inverse problems*，《计算机与化学工程》，181，108547，2024年[14] D. Hansen, D. C. Maddix, S. Alizadeh, G. Gupta, and M. W. Mahoney, *Learning physical models that can respect conservation laws*，《物理学D辑：非线性现象》，457，133952 2024 , , S. Mathur, S. Sidheekh, and S. Natarajan, *A Unified Framework for* [15] B. A. Karanam *Human-Allied Learning of Probabilistic Circuits*，载于 AAAI人工智能会议论文集，39(17)，第 17779–17787 页，2025年。[16] R. Korlakai Vinayak, W. Kong, G. Valiant, and S. M. Kakade, *Maximum likelihoodestimation for learning populations of parameters*，载于第 36 届国际机器学习会议（ICML）论文集，97，第6448–6457页，2019年。[17] J.B. Lasserre, *Global optimization with polynomials and the problem of moments*，SIAM 优化杂志，11(3)，第796–817 页，2001年。[18] S. Lauritzen, C. Uhler, and P. Zwiernik, *Maximum Likelihood Estimation in Gaussian models under total positivity* 4 4 1835 1863, 2019 ，《统计学年刊》， （），第 – 页。[19] J. Löfberg, *YALMIP: A Toolbox for Modeling and Optimization in MATLAB*，载于 CACSD会议论文集，台湾台北，2004年[20] S. Mannor, D. Peleg, and R. Rubinstein, *The cross entropy method for classification*，载于第 22 届国际机器学习会议（ICML）论文集，第561—568 页，2005年。[21] A. Mao, M. Mohri, and Y. Zhong, *Cross-entropy loss functions: theoretical analysisand applications*，载于第 40 届国际机器学习会议（ICML）论文集，202，第 23803–23828 页，2023年。[22] C. E. McCulloch, *Maximum likelihood estimation in a multinomial mixture model*，技术报告 BU-934-MA，生物统计学部，康奈尔大学，纽约州伊萨卡，1987年。[23] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*，威利出版社，纽约，1997年。[24] J. Nie, *Moment and Polynomial Optimization*，工业与应用数学学会，费城，2023年。[25] , *Certifying convergence of Lasserre's hierarchy via flat truncation*，《数学规划》，142，第 485–510 页，2013年。[26] , *Optimality conditions and finite convergence of Lasserre's hierarchy*，《数学规划》，146，第 97–121 页，2014年。[27] , *Tight relaxations for polynomial optimization and Lagrange multiplier expressions*，《数学规划》，178，第 1–37 页，2019年。[28] J. Nie, D. Sun, X. Tang, and M. Zhang, *Solving polynomial variational inequality problems via Lagrange multiplier expressions and Moment-SOS relaxations*，预印本，arXiv:2303.12036，2023年。[29] J. Nie and X. Tang, *Convex generalized Nash equilibrium problems and polynomial optimization* 1485 ，《数学规划》，198，第 –1518 页，2023年。[30] J. Nie, X. Tang, and S. Zhong, *Rational generalized Nash equilibrium problems*，SIAM 优化杂志，33，第 1587–1620 页，2023年。

[31] J. Nie, L. Wang, J.J. Ye, and S. Zhong, *Lagrange multiplier expression method for bilevel polynomial optimization* 31, 2368 2395, 2021 , SIAM 优化杂志，第 - 页。[32] M. Putinar, *Positive polynomials on compact semi-algebraic sets*,《印第安纳大学数学杂志》，42，第 969–984 页，1993年。[33] A. Rayas, R. Anguluri, and G. Dasarathy, *Learning the Structure of Large NetworkedSystems Obeying Conservation Laws*，载于第36届神经信息处理系统年会论文集（NeurIPS 2022），第35卷，第14637–14650页，2022年。[34] B. W. Silverman, *Density estimation for statistics and data analysis*, Chapman and Hall, 1986年。[35] A. Srivastava, A. Bayati and S. M. Salapaka, *Sparse Linear Regression with Constraints:A Flexible Entropy-Based Framework*，载于2024年欧洲控制会议（ECC）论文集，第2105–2110页，2024年。[36] Z. Zhang and M. R. Sabuncu, *Generalized cross entropy loss for training deep neuralnetworks with noisy labels*，载于第32届神经信息处理系统国际会议（NeurIPS）论文集，第8792–8802页，2018年。

Jiyoung Choi, Jiawang Nie, Department of Mathematics, University of California San Diego, 9500 Gilman Drive, La Jolla, CA, USA, 92093.*Email address*: jichoi@ucsd.edu, njw@math.ucsd.edu

Xindong Tang Department of Mathematics Hong Kong Baptist University Kowloon , Tong Kowloon Hong Kong.

, , ,,*Email address*: xdtang@hkbu.edu.hkSuhan Zhong, School of Mathematical Sciences, Shanghai Jiao Tong University,Shanghai, China, 200240.*Email address*: suzhong@sjtu.edu.cn