

多项式优化入门

王杰

中国科学院数学与系统科学研究院

北京师范大学, 2024 年秋季



课程内容

1. 半定规划
2. 平方和理论
3. 测度和矩
4. 矩-平方和松弛分层
5. 变量稀疏 (CS)
6. 项稀疏 (TS)
7. 扩展与应用
8. 软件与实验

多项式优化问题

- 问题形式：

$$f_{\min} := \begin{cases} \inf_{\mathbf{x} \in \mathbb{R}^n} & f(\mathbf{x}) \\ \text{s.t.} & g_j(\mathbf{x}) \geq 0, \quad j = 1, \dots, m \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \dots, \ell \end{cases}$$

- 应用：最优电力流，计算机视觉，组合优化，神经网络，信号处理，图像处理，量子信息……

多项式优化问题

- 问题形式：

$$f_{\min} := \begin{cases} \inf_{\mathbf{x} \in \mathbb{R}^n} & f(\mathbf{x}) \\ \text{s.t.} & g_j(\mathbf{x}) \geq 0, \quad j = 1, \dots, m \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \dots, \ell \end{cases}$$

- 应用： 最优电力流， 计算机视觉， 组合优化， 神经网络， 信号处理， 图像处理， 量子信息……

- 二次约束二次规划 (QCQP):

$$\begin{cases} \inf_{\mathbf{x} \in \mathbb{R}^n} & \mathbf{x}^\top A_0 \mathbf{x} + b_0^\top \mathbf{x} \\ \text{s.t.} & \mathbf{x}^\top A_i \mathbf{x} + b_i^\top \mathbf{x} - c_i \geq (=) 0, \quad i = 1, \dots, m \end{cases}$$

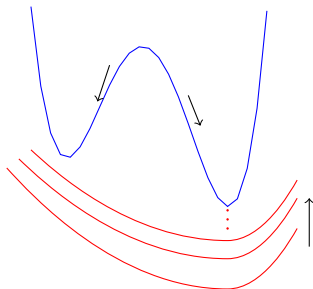
- 逼近: 任何连续函数可用多项式函数进行逼近

离散优化

- **± 1 二元变量:** $x \in \{-1, +1\} \iff x^2 - 1 = 0$
- **$0/1$ 二元变量:** $x \in \{0, 1\} \iff x(x - 1) = 0$
- **整数变量:** $x \in \{1, 2, \dots, t\} \iff (x - 1)(x - 2) \cdots (x - t) = 0$

多项式优化的非凸性

- **非凸**: 鞍点、局部最优解



- **NP-难**: 全局最优解

为什么研究多项式优化？

- **强大的建模能力**：二次约束二次规划、组合优化、混合整数（非）线性规划等
- 与基础数学紧密的内在联系：实（凸）代数几何、测度论
- 与理论计算机科学的联系：近似算法、计算复杂度理论
- 可求得全局最优值（解）：矩-平方和半定松弛分层

为什么研究多项式优化？

- **强大的建模能力**：二次约束二次规划、组合优化、混合整数（非）线性规划等
- **与基础数学紧密的内在联系**：实（凸）代数几何、测度论
- **与理论计算机科学的联系**：近似算法、计算复杂度理论
- **可求得全局最优值（解）**：矩-平方和半定松弛分层

为什么研究多项式优化？

- 强大的建模能力：二次约束二次规划、组合优化、混合整数（非）线性规划等
- 与基础数学紧密的内在联系：实（凸）代数几何、测度论
- 与理论计算机科学的联系：近似算法、计算复杂度理论
- 可求得全局最优值（解）：矩-平方和半定松弛分层

为什么研究多项式优化？

- 强大的建模能力：二次约束二次规划、组合优化、混合整数（非）线性规划等
- 与基础数学紧密的内在联系：实（凸）代数几何、测度论
- 与理论计算机科学的联系：近似算法、计算复杂度理论
- 可求得全局最优值（解）：矩-平方和半定松弛分层

- 2010, P. A. Parrilo (MIT): Semidefinite programming and complex algebraic geometry
- 2014, Monique Laurent (CWI): Optimization over polynomials: selected topics
- 2014, Boaz Barak (Harvard): Sum-of-squares proofs and the quest toward optimal algorithms
- 2018, J. B. Lasserre (CNRS): The moment-SOS hierarchy
- 2018, R. R. Thomas (UW): Spectrahedral lifts of convex sets
- 2018, P. Raghavendra (UC Berkeley) 和 D. Steurer (ETH Zürich): High dimensional estimation via sum-of-squares proofs

多项式优化的松弛分层 (hierarchy of relaxations)

矩-平方和分层

Moment-SOS 分层

Lasserre 分层

例子 (矩松弛)

$$\left\{ \begin{array}{l} \inf_{\mathbf{x}} \quad x_1^2 + x_1 x_2 + x_2^2 \\ \text{s.t.} \quad 1 - x_1^2 \geq 0, 1 - x_2^2 \geq 0 \end{array} \right. \iff \left\{ \begin{array}{l} \inf_{\mathbf{x}} \quad x_1^2 + x_1 x_2 + x_2^2 \\ \text{s.t.} \quad \begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & x_1^2 & x_1 x_2 \\ x_2 & x_1 x_2 & x_2^2 \end{bmatrix} = [1, x_1, x_2] \cdot [1, x_1, x_2]^T \succeq 0, \\ 1 - x_1^2 \geq 0, 1 - x_2^2 \geq 0 \end{array} \right.$$

$$\iff \left\{ \begin{array}{l} \inf_{\mathbf{y}} \quad y_{2,0} + y_{1,1} + y_{0,2} \\ \text{s.t.} \quad \begin{bmatrix} 1 & y_{1,0} & y_{0,1} \\ y_{1,0} & y_{2,0} & y_{1,1} \\ y_{0,1} & y_{1,1} & y_{0,2} \end{bmatrix} \succeq 0, \\ 1 - y_{2,0} \geq 0, 1 - y_{0,2} \geq 0, \\ \exists \mathbf{x} \in \mathbb{R}^2 \text{ s.t. } \mathbf{y} = (x_1, x_2, x_1^2, x_1 x_2, x_2^2) \end{array} \right. \xrightarrow{\text{relax}} \left\{ \begin{array}{l} \inf_{\mathbf{y}} \quad y_{2,0} + y_{1,1} + y_{0,2} \\ \text{s.t.} \quad \begin{bmatrix} 1 & y_{1,0} & y_{0,1} \\ y_{1,0} & y_{2,0} & y_{1,1} \\ y_{0,1} & y_{1,1} & y_{0,2} \end{bmatrix} \succeq 0, \\ 1 - y_{2,0} \geq 0, 1 - y_{0,2} \geq 0 \end{array} \right.$$

测度与优化

- $S = \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0, h_1(\mathbf{x}) = 0, \dots, h_\ell(\mathbf{x}) = 0\}$
- 原多项式优化问题等价于

$$\inf_{\mu \in \mathcal{M}(S)_+} \left\{ \int_S f(\mathbf{x}) d\mu : \mu(S) = 1 \right\}$$



$$\inf_{\mathbf{y}} \left\{ L_{\mathbf{y}}(f) = \sum_{\alpha \in \text{supp}(f)} f_{\alpha} y_{\alpha} : \exists \mu \in \mathcal{M}(S)_+ \text{ s.t. } \mathbf{y} \sim \mu \text{ 和 } y_0 = 1 \right\}$$

- **问题:** 什么样的序列 $\mathbf{y} = (y_{\alpha})_{\alpha \in \mathbb{N}^n}$ 存在支撑在 S 上的 Borel 表示测度?

测度与优化

- $S = \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0, h_1(\mathbf{x}) = 0, \dots, h_\ell(\mathbf{x}) = 0\}$
- 原多项式优化问题等价于

$$\inf_{\mu \in \mathcal{M}(S)_+} \left\{ \int_S f(\mathbf{x}) \, d\mu : \mu(S) = 1 \right\}$$

\Leftrightarrow

$$\inf_{\mathbf{y}} \left\{ L_{\mathbf{y}}(f) = \sum_{\alpha \in \text{supp}(f)} f_{\alpha} y_{\alpha} : \exists \mu \in \mathcal{M}(S)_+ \text{ s.t. } \mathbf{y} \sim \mu \text{ 和 } y_0 = 1 \right\}$$

- **问题：**什么样的序列 $\mathbf{y} = (y_{\alpha})_{\alpha \in \mathbb{N}^n}$ 存在支撑在 S 上的 Borel 表示测度？

测度与优化

- $S = \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0, h_1(\mathbf{x}) = 0, \dots, h_\ell(\mathbf{x}) = 0\}$
- 原多项式优化问题等价于

$$\inf_{\mu \in \mathcal{M}(S)_+} \left\{ \int_S f(\mathbf{x}) \, d\mu : \mu(S) = 1 \right\}$$

\Leftrightarrow

$$\inf_{\mathbf{y}} \left\{ L_{\mathbf{y}}(f) = \sum_{\alpha \in \text{supp}(f)} f_{\alpha} y_{\alpha} : \exists \mu \in \mathcal{M}(S)_+ \text{ s.t. } \mathbf{y} \sim \mu \text{ 和 } y_0 = 1 \right\}$$

- **问题：**什么样的序列 $\mathbf{y} = (y_{\alpha})_{\alpha \in \mathbb{N}^n}$ 存在支撑在 S 上的 Borel 表示测度？

定理

假设 $\mathcal{Q}(\mathbf{g}) + \mathcal{I}(\mathbf{h})$ 满足 *Archimedean 条件*. 序列 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ 存在支撑在 S 上的 *Borel 表示测度* 当且仅当对所有的 i, j, r 和 $\alpha \in \mathbb{N}^n$, $M_r(\mathbf{y}) \succeq 0$ 和 $M_{r-d_j^g}(g_j \mathbf{y}) \succeq 0, L_{\mathbf{y}}(\mathbf{x}^\alpha h_i) = 0$.

- $d_j^g := \lceil \deg(g_j)/2 \rceil, j = 1, \dots, m$

矩松弛分层

- r 阶矩松弛: (Lasserre 2001)

$$\lambda_r := \begin{cases} \inf_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_r(\mathbf{y}) \succeq 0 \\ & \mathbf{M}_{r-d_j^g}(g_j \mathbf{y}) \succeq 0, \quad j = 1, \dots, m \\ & L_{\mathbf{y}}(\mathbf{x}^{\alpha} h_i) = 0, \quad \forall |\alpha| \leq 2r - \deg(h_i), i = 1, \dots, \ell \\ & y_0 = 1 \end{cases}$$

➤ 等价于一个 SDP

➤ $\lambda_r \nearrow f_{\min}, r \rightarrow \infty$

例子 (对偶 SOS 松弛)

$$\begin{cases} \inf_{\mathbf{x}} & x_1^2 + x_1 x_2 + x_2^2 \\ \text{s.t.} & 1 - x_1^2 \geq 0, 1 - x_2^2 \geq 0 \end{cases} \iff \begin{cases} \sup_{\lambda} & \lambda \\ \text{s.t.} & x_1^2 + x_1 x_2 + x_2^2 - \lambda \geq 0, \forall \mathbf{x} \in \mathbb{R}^2 \text{ s.t. } (1 - x_1^2 \geq 0, 1 - x_2^2 \geq 0) \end{cases}$$

$$\xrightarrow{\text{strengthen}} \begin{cases} \sup_{\lambda, \sigma_i} & \lambda \\ \text{s.t.} & x_1^2 + x_1 x_2 + x_2^2 - \lambda = \sigma_0 + \sigma_1(1 - x_1^2) + \sigma_2(1 - x_2^2), \\ & \sigma_0, \sigma_1, \sigma_2 \in \text{SOS} \end{cases}$$

对偶等价问题

- 原多项式优化问题的对偶:

$$f_{\min} = \sup_{\lambda} \{ \lambda : f(\mathbf{x}) - \lambda \geq 0, \forall \mathbf{x} \in S \}$$

- $P_S(\mathbf{x}) := \{g(\mathbf{x}) \in \mathbb{R}[\mathbf{x}] \mid g(\mathbf{x}) \geq 0, \forall \mathbf{x} \in S\}$

↪ SOS 多项式 (Parrilo 2000 & Lasserre 2001)

↪ SONC/SAGE 多项式 (Illiman & de Wolf 2016, Murray et al. 2021)

对偶等价问题

- 原多项式优化问题的对偶:

$$f_{\min} = \sup_{\lambda} \{ \lambda : f(\mathbf{x}) - \lambda \geq 0, \forall \mathbf{x} \in S \}$$

- $P_S(\mathbf{x}) := \{g(\mathbf{x}) \in \mathbb{R}[\mathbf{x}] \mid g(\mathbf{x}) \geq 0, \forall \mathbf{x} \in S\}$

↪ SOS 多项式 (Parrilo 2000 & Lasserre 2001)

↪ SONC/SAGE 多项式 (Illiman & de Wolf 2016, Murray et al. 2021)

对偶等价问题

- 原多项式优化问题的对偶:

$$f_{\min} = \sup_{\lambda} \{ \lambda : f(\mathbf{x}) - \lambda \geq 0, \forall \mathbf{x} \in S \}$$

- $P_S(\mathbf{x}) := \{g(\mathbf{x}) \in \mathbb{R}[\mathbf{x}] \mid g(\mathbf{x}) \geq 0, \forall \mathbf{x} \in S\}$
 - ↪ **SOS 多项式** (Parrilo 2000 & Lasserre 2001)
 - ↪ **SONC/SAGE 多项式** (Illiman & de Wolf 2016, Murray et al. 2021)

Putinar's Positivstellensatz

定理 (Putinar's Positivstellensatz, 1993)

假设 $Q(\mathbf{g}) + \mathcal{I}(\mathbf{h})$ 满足 *Archimedean* 条件. 如果 f 在 S 上是严格正的, 那么

$$f \in Q(\mathbf{g}) + \mathcal{I}(\mathbf{h}).$$

对偶 SOS 松弛

- r 阶对偶 SOS 松弛: (Parrilo 2000 & Lasserre 2001)

$$\lambda_r^* := \begin{cases} \sup_{\lambda} & \lambda \\ \text{s.t.} & f - \lambda \in \mathcal{Q}(\mathbf{g})_{2r} + \mathcal{I}(\mathbf{h})_{2r} \end{cases}$$



$$\lambda_r^* = \begin{cases} \sup_{\lambda, \sigma_j, \tau_i} & \lambda \\ \text{s.t.} & f - \lambda = \sigma_0 + \sum_{j=1}^m \sigma_j g_j + \sum_{i=1}^{\ell} \tau_i h_i \\ & \sigma_0, \sigma_1, \dots, \sigma_m \in \Sigma[\mathbf{x}], \tau_1, \dots, \tau_{\ell} \in \mathbb{R}[\mathbf{x}] \\ & \deg(\sigma_0) \leq 2r, \deg(\sigma_j g_j) \leq 2r, \deg(\tau_i h_i) \leq 2r \end{cases}$$

矩-平方和松弛分层

$$\begin{array}{ccc} & f_{\min} & \\ & \swarrow & \searrow \\ & \vdots & \vdots \\ & \forall l & \forall l \\ \text{(矩松弛)} & \lambda_r & \text{"="} & \lambda_r^* & \text{(对偶平方和松弛)} \\ & \forall l & & \forall l \\ & \vdots & & \vdots \\ & \forall l & & \forall l \\ & \lambda_{r_{\min}} & \text{"="} & \lambda_{r_{\min}}^* \end{array}$$

$$r_{\min} := \max\{\lceil \deg(f)/2 \rceil, \lceil \deg(g_1)/2 \rceil, \dots, \lceil \deg(g_m)/2 \rceil, \lceil \deg(h_1)/2 \rceil, \dots, \lceil \deg(h_\ell)/2 \rceil\}$$

二元优化问题

$$\begin{cases} \inf_{\mathbf{x}} & \mathbf{x}Q\mathbf{x}^\top \\ \text{s.t.} & x_i^2 = 1, \quad i = 1, \dots, n \end{cases} \iff \begin{cases} \inf_{\mathbf{X}} & \langle Q, \mathbf{X} \rangle \\ \text{s.t.} & \mathbf{X} = \mathbf{x}^\top \mathbf{x} \succeq 0 \\ & X_{ii} = 1, \quad i = 1, \dots, n \end{cases}$$
$$\iff \begin{cases} \inf_{\mathbf{X}} & \langle Q, \mathbf{X} \rangle \\ \text{s.t.} & \mathbf{X} \succeq 0, \quad \text{rank}(\mathbf{X}) = 1 \\ & X_{ii} = 1, \quad i = 1, \dots, n \end{cases} \xrightarrow{\text{relax}} \begin{cases} \inf_{\mathbf{X}} & \langle Q, \mathbf{X} \rangle \\ \text{s.t.} & \mathbf{X} \succeq 0 \\ & X_{ii} = 1, \quad i = 1, \dots, n \end{cases}$$

强对偶性

定理

如果下述条件之一成立：

- 可行域 S 有内点；
- 多项式优化问题存在球（面）约束，

则多项式优化问题的**矩松弛**和**对偶平方和松弛**满足**强对偶性**，即有

$$\lambda_r = \lambda_r^*.$$

渐进收敛性与有限收敛性

- 假设 Archimedean 条件成立:

- ▶ $\lambda_r \nearrow f_{\min}$ 和 $\lambda_r^* \nearrow f_{\min}$ 当 $r \rightarrow \infty$ (Putinar's Positivstellensatz, 1993)

- ▶ 如果多项式优化问题有唯一最优解 x^* , 则 $\lim_{r \rightarrow \infty} L_{y^r}(x) = x^*$

- ▶ 一般地, 有限收敛性成立 (Nie 2014)

渐进收敛性与有限收敛性

- 假设 Archimedean 条件成立:

- ▶ $\lambda_r \nearrow f_{\min}$ 和 $\lambda_r^* \nearrow f_{\min}$ 当 $r \rightarrow \infty$ (Putinar's Positivstellensatz, 1993)

- ▶ 如果多项式优化问题有唯一最优解 \mathbf{x}^* , 则 $\lim_{r \rightarrow \infty} L_{y^r}(\mathbf{x}) = \mathbf{x}^*$

- ▶ 一般地, 有限收敛性成立 (Nie 2014)

渐进收敛性与有限收敛性

- 假设 Archimedean 条件成立:

- ▶ $\lambda_r \nearrow f_{\min}$ 和 $\lambda_r^* \nearrow f_{\min}$ 当 $r \rightarrow \infty$ (Putinar's Positivstellensatz, 1993)

- ▶ 如果多项式优化问题有唯一最优解 \mathbf{x}^* , 则 $\lim_{r \rightarrow \infty} L_{y^r}(\mathbf{x}) = \mathbf{x}^*$

- ▶ 一般地, 有限收敛性成立 (Nie 2014)

检测全局最优性和提取全局最优解

定理

设 y 是 r 阶矩松弛的一个最优解.

• 如果 $\text{rank } \mathbf{M}_{r_{\min}}(\mathbf{y}) = 1$, 则

(i) $\lambda_r = f_{\min}$;

(ii) 可提取 1 个全局最优解.

• 如果存在 $r_{\min} \leq r' \leq r$ 使得 $\text{rank } \mathbf{M}_{r'}(\mathbf{y}) = \text{rank } \mathbf{M}_{r'-d_S}(\mathbf{y})$, 则

(i) $\lambda_r = f_{\min}$;

(ii) 可提取 $\text{rank } \mathbf{M}_{r'}(\mathbf{y})$ 个全局最优解.

► $d_S := \max\{d_1^g, \dots, d_m^g, d_1^h, \dots, d_\ell^h\}$ ($d_i^h := \lceil \deg(h_i)/2 \rceil$)

有限收敛性

定理

假设 $\mathcal{V}_{\mathbb{R}}(\mathcal{I}(\mathbf{h}))$ 是有限个点. 则存在 $r \in \mathbb{N}$ 使得 $\lambda_r = f_{\min}$.

定理

假设 $S = \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0\}$, 且 f 和 $-g_j$ 是凸的. 假定 Archimedean 条件和 Slater 条件成立, 且 $\nabla^2 f(\mathbf{x}^*) \succ 0$, 对每个全局最优解 $\mathbf{x}^* \in S$. 则存在 $r \in \mathbb{N}$ 使得 $\lambda_r = f_{\min}$.

- SOS 矩阵: $F(\mathbf{x}) = R(\mathbf{x})^\top R(\mathbf{x})$
- SOS-凸: $\nabla^2 f$ 是 SOS 矩阵

定理

假设 $S = \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0\}$, 且 f 和 $-g_j$ 是 SOS-凸的. 假定 Slater 条件成立. 则 $\lambda_{r_{\min}} = \lambda_{r_{\min}}^* = f_{\min}$.

矩-平方和分层的理论研究

- **收敛速率**: 紧集、球面、单形、超立方体 (Klerk & Laurent)
- **近似比**: 一阶 \rightsquigarrow 最大割问题 ≈ 0.878 (Goemans & Williamson, 1995)
- **准确性**: 一阶、二阶 (秩 1 矩阵补全)、高阶 (Hua & Qu, 2021)
- **紧化**: Lagrange 乘子的多项式表达 (Nie, 2019)

矩-平方和分层的计算瓶颈

- r 阶 SOS 松弛对应 SDP 问题的规模：
 - PSD 矩阵阶数: $\binom{n+r}{r}$
 - 等式约束个数: $\binom{n+2r}{2r}$
- $r = 2, n < 30$ (MOSEK)
- 利用结构:
 - 商环
 - 对称性
 - 稀疏性

矩-平方和分层的计算瓶颈

- r 阶 SOS 松弛对应 SDP 问题的规模：
 - PSD 矩阵阶数: $\binom{n+r}{r}$
 - 等式约束个数: $\binom{n+2r}{2r}$
- $r = 2, n < 30$ (MOSEK)
- 利用结构：
 - 商环
 - 对称性
 - 稀疏性

矩-平方和分层的计算瓶颈

- r 阶 SOS 松弛对应 SDP 问题的规模：
 - PSD 矩阵阶数: $\binom{n+r}{r}$
 - 等式约束个数: $\binom{n+2r}{2r}$
- $r = 2, n < 30$ (MOSEK)
- 利用结构：
 - 商环
 - 对称性
 - 稀疏性

下次课

- 项稀疏 (TS)

参考文献

- Jean B. Lasserre, **Moments, Positive Polynomials and Their Applications**, Imperial College Press, 2010.
- Jean B. Lasserre, **An Introduction to Polynomial and Semi-Algebraic Optimization**, Cambridge University Press, 2015.
- Jie Wang and Victor Magron, **Sparse Polynomial Optimization: Theory and Practice**, World Scientific Publishing, 2023.

更多信息见个人主页

<https://wangjie212.github.io/jiewang>