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

测度

给定集合 $\mathbf{A} \subseteq \mathbb{R}^n$:

- $\mathscr{M}(\mathbf{A})$: 支撑在 $\mathbf{A}$ 上的有限符号 Borel 测度
- 测度 $\mu \in \mathscr{M}(\mathbf{A})$ 的支撑集:

$$\mathrm{supp}(\mu) := \mathrm{cl}(\{\mathbf{x} \in \mathbf{A} \mid \mu(\mathbf{B}) \neq 0, \forall \mathbf{B} \text{ 是 } \mathbf{x} \text{ 的开邻域}\})$$

- $\mathscr{M}_+(\mathbf{A})$: 支撑在 $\mathbf{A}$ 上的有限 Borel 测度

原子测度（atomic measure）

- Dirac 测度 $\delta_{\mathbf{x}}$：$\delta_{\mathbf{x}}(\mathbf{x}) = 1$，$\delta_{\mathbf{x}}(\mathbb{R}^n \setminus \{\mathbf{x}\}) = 0$

- s-原子测度：

$$\mu = \sum_{i=1}^{s} \lambda_i \delta_{\mathbf{x}_i}, \quad \mathbf{x}_1, \ldots, \mathbf{x}_s \in \mathbb{R}^n, \lambda_1, \ldots, \lambda_s > 0, \sum_{i=1}^{s} \lambda_i = 1$$

- 矩（moment）：$y_{\boldsymbol{\alpha}} = \int \mathbf{x}^{\boldsymbol{\alpha}} \mathrm{d}\mu$，$\boldsymbol{\alpha} \in \mathbb{N}^n$

矩问题

- 基础半代数集 $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0\}$

全矩问题

给定伪矩序列 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subseteq \mathbb{R}$，是否存在支撑在 $\mathbf{K}$ 上的有限 Borel 测度 μ(表示测度) 使得 $y_\alpha = \int \mathbf{x}^\alpha \mathrm{d}\mu, \forall \boldsymbol{\alpha} \in \mathbb{N}^n$?

截断矩问题

给定伪矩序列 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathscr{A}} \subseteq \mathbb{R}$，是否存在支撑在 $\mathbf{K}$ 上的有限 Borel 测度 μ(表示测度) 使得 $y_\alpha = \int \mathbf{x}^\alpha \mathrm{d}\mu, \forall \boldsymbol{\alpha} \in \mathscr{A}$?

Riesz 线性泛函

- 给定伪矩序列 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subseteq \mathbb{R}$，定义 Riesz 线性泛函 $L_{\mathbf{y}} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$:

$$f\left(= \sum_{\boldsymbol{\alpha}} f_\alpha \mathbf{x}^{\boldsymbol{\alpha}}\right) \mapsto L_{\mathbf{y}}(f) = \sum_{\boldsymbol{\alpha}} f_\alpha y_\alpha, \quad \forall f \in \mathbb{R}[\mathbf{x}]$$

定理 (Riesz–Haviland)

给定 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subseteq \mathbb{R}$，假设 $\mathbf{K} \subseteq \mathbb{R}^n$ 是闭的. 则 $\mathbf{y}$ 存在支撑在 $\mathbf{K}$ 上的 *Borel* 表示测度当且仅当 $L_{\mathbf{y}}(f) \geq 0, \forall f \in \mathcal{P}(\mathbf{K})$（$\mathbf{K}$ 上的非负多项式）.

矩方阵和局部化矩阵

- $\mathbf{y} = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n} \subseteq \mathbb{R}$
- $\mathbb{N}^n_r := \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid |\boldsymbol{\alpha}| \leq r\}$
- r 阶矩方阵（moment matrix）$\mathbf{M}_r(\mathbf{y})$:

$$[\mathbf{M}_r(\mathbf{y})]_{\beta\gamma} := L_{\mathbf{y}}(\mathbf{x}^{\beta+\gamma}) = y_{\beta+\gamma}, \quad \forall \beta, \gamma \in \mathbb{N}^n_r$$

- 给定 $g = \sum_\alpha g_\alpha \mathbf{x}^\alpha$, r 阶局部化矩阵（localizing matrix）$\mathbf{M}_r(g\mathbf{y})$:

$$[\mathbf{M}_r(g\mathbf{y})]_{\beta\gamma} := L_{\mathbf{y}}(\mathbf{x}^{\beta+\gamma} g) = \sum_\alpha g_\alpha y_{\alpha+\beta+\gamma}, \quad \forall \beta, \gamma \in \mathbb{N}^n_r$$

矩方阵和局部化矩阵

- $\mathbf{y} = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}^n} \subseteq \mathbb{R}$
- $\mathbb{N}_r^n := \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid |\boldsymbol{\alpha}| \le r\}$
- r 阶矩方阵（moment matrix）$\mathbf{M}_r(\mathbf{y})$:

$$[\mathbf{M}_r(\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := L_{\mathbf{y}}(\mathbf{x}^{\boldsymbol{\beta}+\boldsymbol{\gamma}}) = y_{\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}_r^n$$

- 给定 $g = \sum_{\alpha} g_{\alpha}\mathbf{x}^{\alpha}$, r 阶局部化矩阵（localizing matrix）$\mathbf{M}_r(g\mathbf{y})$:

$$[\mathbf{M}_r(g\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := L_{\mathbf{y}}(\mathbf{x}^{\boldsymbol{\beta}+\boldsymbol{\gamma}}g) = \sum_{\alpha} g_{\alpha}y_{\alpha+\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}_r^n$$

矩方阵和局部化矩阵

- $\mathbf{y} = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}^n} \subseteq \mathbb{R}$
- $\mathbb{N}^n_r := \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid |\boldsymbol{\alpha}| \le r\}$
- r 阶矩方阵（moment matrix）$\mathbf{M}_r(\mathbf{y})$:

$$[\mathbf{M}_r(\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := L_{\mathbf{y}}(\mathbf{x}^{\boldsymbol{\beta}+\boldsymbol{\gamma}}) = y_{\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_r$$

- 给定 $g = \sum_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$, r 阶局部化矩阵（localizing matrix）$\mathbf{M}_r(g\mathbf{y})$:

$$[\mathbf{M}_r(g\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := L_{\mathbf{y}}(\mathbf{x}^{\boldsymbol{\beta}+\boldsymbol{\gamma}} g) = \sum_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_r$$

矩方阵和局部化矩阵：例子

- $\mathbf{x} = x$, $g = 1 - x^2$:

$$\mathbf{M}_2(\mathbf{y}) = \begin{array}{c c} & \begin{array}{ccc} 1 & x & x^2 \end{array} \\ \begin{array}{c} 1 \\ x \\ x^2 \end{array} & \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \end{array}, \quad \mathbf{M}_1(g\mathbf{y}) = \begin{array}{c c} & \begin{array}{cc} 1 & x \end{array} \\ \begin{array}{c} 1 \\ x \end{array} & \begin{bmatrix} y_0 - y_2 & y_1 - y_3 \\ y_1 - y_3 & y_2 - y_4 \end{bmatrix} \end{array}$$

矩方阵和局部化矩阵：例子

- $\mathbf{x} = (x_1, x_2)$, $g = x_1 - x_1^2$:

$$
\mathbf{M}_2(\mathbf{y}) = \begin{array}{c} \\ 1 \\ \\ x_1 \\ x_2 \\ \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{array}
\begin{array}{c}
\begin{array}{ccccccccc} 1 & & x_1 & x_2 & & x_1^2 & x_1x_2 & x_2^2 \end{array} \\
\left[\begin{array}{c|cc|ccc}
y_{0,0} & y_{1,0} & y_{0,1} & y_{2,0} & y_{1,1} & y_{0,2} \\
\hline
y_{1,0} & y_{2,0} & y_{1,1} & y_{3,0} & y_{2,1} & y_{1,2} \\
y_{0,1} & y_{1,1} & y_{0,2} & y_{2,1} & y_{1,2} & y_{0,3} \\
\hline
y_{2,0} & y_{3,0} & y_{2,1} & y_{4,0} & y_{3,1} & y_{2,2} \\
y_{1,1} & y_{2,1} & y_{1,2} & y_{3,1} & y_{2,2} & y_{1,3} \\
y_{0,2} & y_{1,2} & y_{0,3} & y_{2,2} & y_{1,3} & y_{0,4}
\end{array}\right]
\end{array}
$$

$$
\mathbf{M}_1(g\mathbf{y}) = \begin{array}{c} \\ 1 \\ x_1 \\ x_2 \end{array}
\begin{array}{c}
\begin{array}{ccc} 1 & x_1 & x_2 \end{array} \\
\left[\begin{array}{ccc}
y_{1,0} - y_{2,0} & y_{2,0} - y_{3,0} & y_{1,1} - y_{2,1} \\
y_{2,0} - y_{3,0} & y_{3,0} - y_{4,0} & y_{2,1} - y_{3,1} \\
y_{1,1} - y_{2,1} & y_{2,1} - y_{3,1} & y_{1,2} - y_{2,2}
\end{array}\right]
\end{array}
$$

全矩问题的解：Archimedean 条件

- 基础半代数集 $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0\}$
- Archimedean 条件：存在 $N > 0$ 使得 $N - \|\mathbf{x}\|_2^2 \in \mathcal{Q}(\mathbf{g}) \rightsquigarrow \mathbf{K}$ 是紧集

定理 (Putinar's Positivstellensatz 对偶)

给定 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subseteq \mathbb{R}$，假设 *Archimedean* 条件成立. 则 $\mathbf{y}$ 存在支撑在 $\mathbf{K}$ 上的 *Borel* 表示测度当且仅当 $\mathbf{M}_r(\mathbf{y}) \succeq 0, \mathbf{M}_r(g_i\mathbf{y}) \succeq 0$，对所有的 $i \in [m]$ 和 $r \in \mathbb{N}$.

全矩问题的解：紧集

- 基础半代数集 $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0\}$

定理 (Schmüdgen's Positivstellensatz 对偶)

给定 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subseteq \mathbb{R}$，假设 $\mathbf{K}$ 是紧的. 则 $\mathbf{y}$ 存在支撑在 $\mathbf{K}$ 上的 *Borel* 表示测度当且仅当 $\mathbf{M}_r(\mathbf{y}) \succeq 0, \mathbf{M}_r(g_I\mathbf{y}) \succeq 0$，对所有的 $I \subseteq [m]$ 和 $r \in \mathbb{N}$. $(g_I := \prod_{i \in I} g_i)$

平坦扩张

- 给定 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n_{2r}} \subseteq \mathbb{R}$ 满足 $\mathbf{M}_r(\mathbf{y}) \succeq 0$
- 平坦扩张： $\mathbf{M}_{r+1}(\mathbf{y}) \succeq 0, \, \mathrm{rank}(\mathbf{M}_r(\mathbf{y})) = \mathrm{rank}(\mathbf{M}_{r+1}(\mathbf{y}))$

定理 (Curto & Fialkow, 1991)

给定 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n_{2r}} \subseteq \mathbb{R}$. 则 $\mathbf{y}$ 存在 $\mathrm{rank}(\mathbf{M}_r(\mathbf{y}))$-原子表示测度当且仅当 $\mathbf{M}_r(\mathbf{y}) \succeq 0$ 且 $\mathbf{M}_r(\mathbf{y})$ 存在平坦扩张 $\mathbf{M}_{r+1}(\mathbf{y})$.

截断 K-矩问题的解

- $d_i := \lceil \frac{\deg(g_i)}{2} \rceil$, $i \in [m]$
- $d_{\mathbf{K}} := \max\{d_1, \ldots, d_m\}$

定理 (Curto & Fialkow, 2000)

给定 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n_{2r}} \subseteq \mathbb{R}$. 则 $\mathbf{y}$ 存在支撑在 $\mathbf{K}$ 上的 $\mathrm{rank}(\mathbf{M}_r(\mathbf{y}))$-原子表示测度当且仅当

1. $\mathbf{M}_r(\mathbf{y}) \succeq 0$, $\mathbf{M}_{r-d_i}(g_i\mathbf{y}) \succeq 0$, $i \in [m]$;
2. $\mathrm{rank}(\mathbf{M}_r(\mathbf{y})) = \mathrm{rank}(\mathbf{M}_{r-d_{\mathbf{K}}}(\mathbf{y}))$.

原子测度表示

- $f \in \mathbb{R}[\mathbf{x}]_r$，令 $\mathbf{f}$ 是 f 的系数向量（$f = \mathbf{f}^\intercal[\mathbf{x}]_r$）
- $\ker(\mathbf{M}_r(\mathbf{y})) := \{f \in \mathbb{R}[\mathbf{x}]_r \mid \mathbf{M}_r(\mathbf{y})\mathbf{f} = \mathbf{0}\}$

定理

假设 $\mathbf{M}_r(\mathbf{y}) \succeq 0$ 且 $\mathrm{rank}(\mathbf{M}_r(\mathbf{y})) = \mathrm{rank}(\mathbf{M}_{r-1}(\mathbf{y})) =: s$. 则 $\mathbf{y}$ 存在原子测度表示：$\mu = \sum_{i=1}^{s} \lambda_i \delta_{\mathbf{x}_i}$，且 $(\ker(\mathbf{M}_r(\mathbf{y}))) = \mathcal{I}(\{\mathbf{x}_1, \ldots, \mathbf{x}_s\})$ 是一个零维实根理想.

乘法算子与乘法矩阵

- $\mathbf{M}_r(\mathbf{y}) \succeq 0$ 且 $\mathrm{rank}(\mathbf{M}_r(\mathbf{y})) = \mathrm{rank}(\mathbf{M}_{r-1}(\mathbf{y}))$
- 定义乘法算子 $\mathcal{M}_i, i \in [n]$：

$$
\begin{aligned}
\mathcal{M}_i : \mathbb{R}[\mathbf{x}]/(\ker(\mathbf{M}_r(\mathbf{y}))) &\longrightarrow \mathbb{R}[\mathbf{x}]/(\ker(\mathbf{M}_r(\mathbf{y}))) \\
f &\longmapsto x_i \cdot f
\end{aligned}
$$

- 乘法矩阵 $M_1, \ldots, M_n$：乘法算子 $\mathcal{M}_1, \ldots, \mathcal{M}_n$ 在某个基下的表示矩阵

Stickelberger's Theorem

定理 (Stickelberger's Theorem)

设 $I \subseteq \mathbb{R}[\mathbf{x}]$ 是一个零维理想. 则 I 的零点集是由乘法矩阵 $\{M_i\}_{i=1}^n$ 所有公共特征值形成的点集，即,

$$\mathcal{V}(I) = \{(\xi_1, \ldots, \xi_n) \in \mathbb{C}^n \mid \exists \mathbf{v} \in \mathbb{C}^n \,\mathrm{s.t.}\, \forall i \in [n], M_i\mathbf{v} = \xi_i\mathbf{v}\}\,.$$

提取支撑点集（Klep, Povh, & Volčič, 2018）

假设 $\mathbf{M}_r(\mathbf{y}) \succeq 0$ 且 $\mathrm{rank}(\mathbf{M}_r(\mathbf{y})) = \mathrm{rank}(\mathbf{M}_{r-1}(\mathbf{y})) = s$

- Step 1：计算奇异值分解 $M_{r-1}(\mathbf{y}) = UDU^\intercal$，其中 D 是正定对角阵，$U^\intercal U = I$
- Step 2：令 $M_i = \sqrt{D}^{-1} U^\intercal M_{r-1}(x_i\mathbf{y}) U \sqrt{D}^{-1}$， $i = 1, \ldots, n$
- Step 3：令 $M = \sum_{i=1}^n c_i M_i$, 其中 c_i 是随机生成的系数
- Step 4：计算谱分解 $M = QTQ^\intercal$
- Step 5：令 $\mathbf{x}_j := (\mathbf{q}_j^\intercal M_i \mathbf{q}_j)_{i=1}^n$, $j \in [s]$, 其中 $\{\mathbf{q}_j\}_{1\le j\le s}$ 是 Q 的列向量

例子

- $n = 2$:

$$\mathbf{M}_2(\mathbf{y}) = \begin{bmatrix} 1 & 1 & 1 & 2 & 0 & 3 \\ 1 & 2 & 0 & 4 & 0 & 0 \\ 1 & 0 & 3 & 0 & 0 & 9 \\ 2 & 4 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 9 & 0 & 0 & 27 \end{bmatrix}$$

$$\mathbf{M}_1(x_1\mathbf{y}) = \begin{bmatrix} 2 & 4 & 0 \\ 4 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{M}_1(x_2\mathbf{y}) = \begin{bmatrix} 3 & 0 & 9 \\ 0 & 0 & 0 \\ 9 & 0 & 27 \end{bmatrix}$$

例子

- 提取三个支撑点：

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

截断 K-矩问题的解

- $d_i := \lceil \frac{\deg(g_i)}{2} \rceil$, $i \in [m]$
- $d_{\mathbf{K}} := \max\{d_1, \ldots, d_m\}$

定理 (Curto & Fialkow, 2000)

给定 $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n_{2r}} \subseteq \mathbb{R}$. 若

1. $\mathbf{M}_r(\mathbf{y}) \succeq 0$, $\mathbf{M}_{r-d_i}(g_i\mathbf{y}) \succeq 0$, $i \in [m]$;
2. $\mathrm{rank}(\mathbf{M}_r(\mathbf{y})) = \mathrm{rank}(\mathbf{M}_{r-d_{\mathbf{K}}}(\mathbf{y}))$,

则 $\mathbf{y}$ 存在原子测度表示: $\mu = \sum_{i=1}^{s} \lambda_i \delta_{\mathbf{x}_i}$, 且 $\{\mathbf{x}_1, \ldots, \mathbf{x}_s\} \subseteq \mathbf{K}$.

下次课

- 矩-平方和松弛分层

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.
- Igor Klep, Janez Povh, and Volčič, **Minimizer Extraction in Polynomial Optimization Is Robust**, SIAM Journal on Optimization, 2018.

更多信息见个人主页

https://wangjie212.github.io/jiewang