日本オペレーションズ・リサーチ学会
2019年 秋季研究発表会

SD 基に基づく半正定値行列錐の凸多面錐近似

05001183　筑波大学システム情報工学研究科　*汪 玉柱　Wang Yuzhu
01703540　筑波大学システム情報系　吉瀬 章子　Yoshise Akiko

1. はじめに

錐最適化問題の一種である半正定値計画問題 (SDP) は組合せ最適化問題の凸緩和を与えることができるため，注目を集めている．SDP のほかに，二重非負正値錐 (DNN) 計画問題や完全正値錐計画問題 (CP) などの錐最適化問題も非凸二次計画問題などの求解が難しい問題に対してよい近似解を与えられる．しかし，現在の SDP ソルバーで大規模な SDP を解くのは非効率的で，DNN 問題と CP は SDP 以上に求解が難しいと知られている．本研究の目的は SDP，DNN 問題と CP のような求解が難しい錐最適化問題に効率の良い近似解法を提案することである．

SDP の標準形は下のように定義される：

$$
\begin{aligned}
\min \quad & \langle C, X\rangle \\
\text{s.t.} \quad & \langle A_j, X\rangle = b_j \ (j = 1, .., m), \\
& X \in \mathcal{S}^n_+,
\end{aligned} \tag{1}
$$

ただし，$\mathbb{S}^n := \{X \in \mathbb{R}^{n\times n} \mid \forall i < j,\ X_{i,j} = X_{j,i}\}$ が実対称行列空間で，$C \in \mathbb{S}^n$，$A_j \in \mathbb{S}^n$，$b_j \in \mathbb{R}$ $(j = 1, .., m)$，$\langle A, B\rangle := \sum_{i,j=1}^n A_{i,j}B_{i,j}$ が $\mathbb{S}^n$ 上の内積で，$\mathcal{S}^n_+ := \{X \in \mathbb{S}^n \mid \forall d \in \mathbb{R}^n,\ d^TXd \geq 0\}$ が半正定値行列錐である．

大規模 SDP と DNN 問題を効率よく解く手法として，切除平面法が知られている [2],[4]．切除平面法は SDP を比較的に求解が容易な線形計画問題 (LP) 或いは二次錐計画問題 (SOCP) に緩和し，毎回の反復で切除平面を緩和問題の制約に加えることで，現在の緩和解を次回反復の実行可能領域から切除し，SDP の最適解に近づく．具体的に，切除平面法は最初に $\mathcal{S}^n_+$ の外部近似 $\mathcal{K}_{\text{out}} : \mathcal{S}^n_+ \subseteq \mathcal{K}_{\text{out}} \subseteq \mathbb{S}^n$ を構築し，(1) の半正定値制約 $X \in \mathcal{S}^n_+$ を $X \in \mathcal{K}_{\text{out}}$ に緩和し，初期緩和問題を得る．もし $\mathcal{K}_{\text{out}}$ は凸多面錐 (resp. 二次錐制約) の表現があれば，初期緩和問題が LP(resp. SOCP) になる．

本研究は大規模な SDP や DNN 問題など，求解が難しい錐最適化問題に効率の良い近似解法を提案することを目的として，切除平面法に着目し，精度が良くて計算が容易な初期緩和問題を構築するために，半正定値行列錐 $\mathcal{S}^n_+$ の近似に注目する．

2. 半正定値行列錐 $\mathcal{S}^n_+$ の近似

[1] は Factor width という半正定値行列の概念を定義し，$\mathcal{S}^n_+$ の部分集合を構築した．$z(v)$ を $v \in \mathbb{R}^n$ の非ゼロ要素数を表す記号とし，ある行列 $A \in \mathcal{S}^n_+$ の factor width とは，あらゆる分解 $A = \sum_i v_iv_i^T$, $\max_i z(v_i) = k$ の中に k の最小値である．さらに集合を定義する:$\mathcal{FW}(k) := \{A \in \mathbb{S}^n \mid A \text{ has factor width at most } k\}$．定義により，$A \in \mathbb{S}^n$ に対し，もし分解 $A = \sum_i v_iv_i^T$, $\max_i z(v_i) = k$ が存在すれば，$A \in \mathcal{FW}(k)$．以下の包含関係が自然に導かれる：

$$
\mathcal{FW}(1) \subseteq \mathcal{FW}(2) \subseteq \ldots \subseteq \mathcal{FW}(n) = \mathcal{S}^n_+.
$$

[2] は優対角行列の集合 $\mathcal{DD}_n$ とスケーリングされた優対角行列の集合 $\mathcal{SDD}_n$ を用いて $\mathcal{S}^n_+$ を近似した．$\mathcal{DD}_n := \{A \in \mathbb{S}^n \mid A_{i,i} \geq \sum_{j\neq i}|A_{i,j}|\ (i = 1, .., n)\}$，もしある対角成分が正の対角行列 D が存在し，$DAD \in \mathcal{DD}_n$ であれば，$A \in \mathcal{SDD}_n$．[1] と [2] により，包含関係

$$
\mathcal{DD}_n \subseteq \mathcal{FW}(2) = \mathcal{SDD}_n \subseteq \mathcal{S}^n_+
$$

が知られている．[2] によると，$\mathcal{DD}_n$(resp. $\mathcal{SDD}_n$) 上の最適化問題は LP(resp. SOCP) で解ける．

[3] は半正定値基 (SD 基) という概念を定義し，$\mathcal{S}^n_+$ を近似した．与えられた正規直交行列 $P = (p_1, .., p_n) \in \mathbb{R}^{n\times n}$ から生成される SD 基は以下のように定義される：

$$
\begin{aligned}
\mathcal{B}_+(P) :=\ & \{(p_i + p_j)(p_i + p_j)^T \mid 1 \leq i \leq j \leq n\}, \\
\mathcal{B}_-(P) :=\ & \{(p_i + p_i)(p_i + p_i)^T \mid i = 1, .., n\} \\
& \cup \{(p_i - p_j)(p_i - p_j)^T \mid 1 \leq i < j \leq n\}.
\end{aligned}
$$

$\mathcal{B}_+(P)$ が SD 基 Type I と呼び，$\mathcal{B}_-(P)$ が SD 基 Type II と呼ぶ．$\mathcal{B}_+(P)$ と $\mathcal{B}_-(P)$ は共に $\frac{n(n+1)}{2}$ 個

の半正定値行列からなる $\mathbb{S}^n$ 上の基底集合である．SD 基からなる凸錐とその双対錐を取ることで，$\mathcal{S}^n_+$ の凸多面錐の内部近似と外部近似を構築できる：

$$\mathcal{S}_{\text{in}} := \text{cone}(\mathcal{B}_+(P) \cup \mathcal{B}_-(P)) \subseteq \mathcal{S}^n_+ \subseteq (\mathcal{S}_{\text{in}})^*,$$

ただし，集合 $\mathcal{K}$ からなる凸錐が $\text{cone}(\mathcal{K}) := \{\sum_{i=1}^k \alpha_i X_i \mid X_i \in \mathcal{K}, \alpha_i \geq 0, k \in \mathbb{Z}_{\geq 0}\}$ で，$\mathbb{Z}_{\geq 0}$ が非負整数の集合である．集合 $\mathcal{K}$ の双対錐は $(\mathcal{K})^* := \{A \in \mathbb{S}^n \mid \forall B \in \mathcal{K}, \langle A, B\rangle \geq 0\}$ で定義される．

3.　SD 基の拡張

単位行列 I で [3] の SD 基を生成すると，$\mathcal{S}^n_+$ の疎な凸多面錐近似を作ることができる．疎性を持つ $\mathcal{S}^n_+$ の近似は計算上望ましいので，本研究は SD 基に注目し，まずは SD 基の性質を分析した．

$\mathcal{O}^n$ を正規直交行列の集合とし，$I = (e_1, .., e_n)$ を単位行列とする．本研究はまず SD 基を用いた $\mathcal{S}^n_+$ の表現を与えた：

$$\mathcal{S}^n_+ = \text{cone}\left(\bigcup_{P \in \mathcal{O}^n} \mathcal{B}_+(P)\right) = \text{cone}\left(\bigcup_{P \in \mathcal{O}^n} \mathcal{B}_-(P)\right).$$

そして，本研究は SD 基を用いた $\mathcal{S}^n_+$ の凸多面錐近似と $\mathcal{DD}_n$ の等価関係を証明した：

$$\mathcal{DD}_n = \text{cone}(\mathcal{B}_+(I) \cup \mathcal{B}_-(I)).$$

本研究は SD 基を用いた凸多面錐近似の精度を高める可能性を発見し，疎性を失わない SD 基の拡張を提案した．

$P = (p_1, .., p_n) \in \mathcal{O}^n$ を正規直交行列とし，パラメーター $\alpha \in \mathbb{R}$ を用いた SD 基の拡張を以下のように定義する：

$$\bar{\mathcal{B}}(\alpha, P) := \{(p_i + \alpha p_j)(p_i + \alpha p_j)^T \mid 1 \leq i \leq j \leq n\}.$$

本研究はさらに拡張 SD 基の性質を証明した．$P \in \mathcal{O}^n$ が与えられた時，拡張 SD 基の以下の性質が成り立つ：

- $\alpha \neq 0, -1$ ならば，$\bar{\mathcal{B}}(\alpha, P)$ は線形独立で，$\mathbb{S}^n$ の基底集合である．
- $\alpha_1 \in \mathbb{R} \setminus \{0, -1\}$ がある時，$\alpha_2 \in \mathbb{R} \setminus \{0, \alpha_1\}$ ならば，$(p_i + \alpha_2 p_j)(p_i + \alpha_2 p_j)^T \notin \text{cone}(\bar{\mathcal{B}}(\alpha_1, P))\ \ (1 \leq i < j \leq n)$.

さらに，本研究は拡張 SD 基による $\mathcal{SDD}_n$ の表現とその証明を与えた．$I = (e_1, .., e_n)$ を単位行列とすると，

$$\text{cone}\left(\bigcup_{\alpha \in \mathbb{R}} \bar{\mathcal{B}}(\alpha, I)\right) = \mathcal{FW}(2) = \mathcal{SDD}_n.$$

パラメーターの集合を $\{1, -1\}$ にする時，$\cup_{\alpha \in \{1,-1\}} \bar{\mathcal{B}}(\alpha, I) = \mathcal{B}_+(I) \cup \mathcal{B}_-(I)$ が成り立つ．よって，あるパラメーターの集合 $\mathcal{H} : \{1, -1\} \subseteq \mathcal{H} \subseteq \mathbb{R}$ を使うと，疎な凸多面錐近似 $\text{cone}\left(\bigcup_{\alpha \in \mathcal{H}} \bar{\mathcal{B}}(\alpha, I)\right)$：

$$\mathcal{DD}_n \subseteq \text{cone}(\cup_{\alpha \in \mathcal{H}} \bar{\mathcal{B}}(\alpha, I)) \subseteq \mathcal{SDD}_n$$

を生成することができる．

4.　拡張 SD 基の応用

本研究は提案した拡張 SD 基に基づき，半正定値行列錐の凸多面錐近似を錐最適化問題に応用し、計算機実験を行って提案手法の有効性を検証した．実験結果については当日説明する．

参考文献

[1] Boman, Erik G., et al. "On factor width and symmetric H-matrices." Linear algebra and its applications 405 (2005): 239-248.

[2] A. A. Ahmadi, S. Dash, and G. Hall, Optimization over structured subsets of positive semidefinite matrices via column generation, Discrete Optimization, 24 (2017), pp. 129-151.

[3] A. Tanaka and A. Yoshise, Lp-based tractable subcones of the semidefinite plus nonnegative cone, Annals of Operations Research, 265 (2018), pp. 155-182.

[4] Krishnan, Kartik, and John E. Mitchell. "A unifying framework for several cutting plane methods for semidefinite programming." Optimization methods and software 21.1 (2006): 57-74.