半正定値行列錐の凸多面錐近似
—その評価と応用—

汪　玉柱

半正定値計画問題は組合せ最適化，制御理論など，幅広い分野で応用されている．近年，大規模な半正定値計画問題を解く有効な手法として，半正定値計画問題を緩和して解く切除平面法が研究されている．半正定値計画問題の緩和で使われる半正定値行列錐の近似集合は，切除平面法の効率に大きな影響を与えるため，注目を集めている．本稿では，著者の博士論文を基に，半正定値行列錐の近似集合をいくつか解説し，各近似の近似精度に対する評価を述べ，半正定値計画問題に対する切除平面法についても紹介する．

キーワード：半正定値計画問題，半正定値行列錐の凸多面錐近似，切除平面法

1. はじめに

半正定値計画問題 (Semidefinite optimization problem: SDP) は凸最適化，組合せ最適化，制御理論など，幅広い分野に応用でき，1990 年代から注目されている．SDP の標準形は下のように定義される：

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

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，$C \in \mathbb{S}^n$, $A_j \in \mathbb{S}^n$, $b_j \in \mathbb{R}$ $(j = 1,\dots,m)$, $\langle A, B\rangle := \sum_{i,j=1}^n A_{i,j}B_{i,j}$ 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．

SDP は線形計画問題 (Linear optimization problem: LP) と二次錐計画問題 (Second-order cone optimization problem: SOCP) を含む問題クラスであり，凸二次計画問題，固有値最適化問題などの凸最適化問題も記述できる．また，SDP は凸最適化問題の特殊なケースとして，max-cut 問題，2 次ナップザック問題などの組合せ最適化や非凸最適化に対して凸緩和を与えることもできる [1].

SDP は LP の拡張であり，内点法により任意の精度まで多項式時間で解くことができ [2]，内点法を実装した最先端の SDP ソルバーも開発されている（e.g., Mosek [3]）．しかし，内点法は毎回の反復でサイズが大きい密な行列を計算し，保存する必要があるため，その計算は SDP のサイズが大きくなると困難になる．

近年，大規模な SDP を効率よく解くための手法が数多く開発されており [4], その中で，切除平面法は SDP の精度の良い近似解を高速に与えることができるため，注目を集めている (e.g., 文献 [5–7])．切除平面法は，SDP を比較的に求解が容易な問題 (e.g., LP または 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，現在の緩和解を次の反復の実行可能領域から切除し，SDP の最適解に近づける．

上記の切除平面法で，SDP の半正定値制約 $X \in \mathcal{S}^n_+$ はまず $X \in \mathcal{K}_{\text{out}}$ に緩和される．ただし, $\mathcal{S}^n_+ \subseteq \mathcal{K}_{\text{out}} \subseteq \mathbb{S}^n$．このような集合 $\mathcal{K}_{\text{out}}$ を $\mathcal{S}^n_+$ の外部近似と呼ぶ．制約 $X \in \mathcal{K}_{\text{out}}$ が線形で与えられる場合, 緩和された問題は LP となり，強力な LP ソルバー (e.g., Gurobi [8]) で解くことができる．さらに，$X \in \mathcal{K}_{\text{out}}$ が疎な制約（e.g., $\langle X, A\rangle \geq 0$，ただし，$A$ は疎な行列）で与えられると，効率的な計算が期待できる．このように計算しやすく，解の精度の高い緩和問題を構築するためには，疎で近似精度の高い $\mathcal{S}^n_+$ の外部近似集合が必要である．

本稿では，著者の博士論文を基に，$\mathcal{S}^n_+$ の代表的な近似集合をいくつか取り上げ，凸多面錐近似を中心に説明し，各近似の近似精度に関する評価も解説する．また，$\mathcal{S}^n_+$ の凸多面錐近似の応用として，半正定値計画問題に対する切除平面法についても紹介する．

2. 準備

本節では準備として，凸錐と凸多面錐の定義を述べ，$\mathcal{S}^n_+$ の性質および $\mathcal{S}^n_+$ の近似の定義について説明する．

実対称行列空間 $\mathbb{S}^n$ 上の空集合でない集合 $\mathcal{K}$ に対し，任意のスカラー $\alpha, \beta \geq 0$ と行列 $X, Y \in \mathcal{K}$ で

おう　ぎょくちゅう
日鉄ソリューションズ株式会社
yuzhu.wang.b85@jp.nssol.nipponsteel.com

$\alpha X + \beta Y \in \mathcal{K}$ が成り立つとき，$\mathcal{K}$ を凸錐と呼ぶ．

集合 $\mathcal{A}$ からなる凸錐は

$$\mathrm{cone}(\mathcal{A}) := \{\sum_{i=1}^{k} \alpha_i X_i \mid k \in \mathbb{N}, \alpha_i \geq 0, X_i \in \mathcal{A}\}$$

と定義される．ただし，$\mathbb{N}$ は非負整数の集合である．$\mathcal{A} \subseteq \mathcal{B}$ であれば，$\mathrm{cone}(\mathcal{A}) \subseteq \mathrm{cone}(\mathcal{B})$ が成り立つ．有限個要素の集合からなる凸錐を有限生成錐と呼ぶ．

凸多面錐は凸錐の特殊なケースである．凸錐 $\mathcal{K} \subseteq \mathbb{S}^n$ が正の整数 m と行列 $A_1, \ldots, A_m \in \mathbb{S}^n$ で

$$\mathcal{K} = \{X \in \mathbb{S}^n \mid \langle A_j, X \rangle \leq 0, j = 1, \ldots, m\}$$

を満たせば，$\mathcal{K}$ を凸多面錐と呼ぶ．凸多面錐は有限生成錐であり，逆も成り立つ [9]．

凸錐 $\mathcal{K} \subseteq \mathbb{S}^n$ の双対錐は

$$\mathcal{K}^* := \{X \in \mathbb{S}^n \mid \forall Y \in \mathcal{K}, \langle X, Y \rangle \geq 0\}$$

と定義される．凸錐 $\mathcal{K}$ の双対錐 $\mathcal{K}^*$ は閉凸錐である．また，$\mathcal{K}_2 \subseteq \mathcal{K}_1$ であれば，$\mathcal{K}_1^* \subseteq \mathcal{K}_2^*$ が成り立つ．$\mathcal{K} = \mathcal{K}^*$ を満たす集合を自己双対と呼ぶ．たとえば，$\mathcal{S}_+^n$ は自己双対（i.e., $(\mathcal{S}_+^n)^* = \mathcal{S}_+^n$）な閉凸錐である．半正定値行列の同値条件として，以下のことが知られている：

定理 1（文献 [10] (Theorem 1.10)）**.** $X \in \mathbb{S}^n$ に対し，下記の三つの条件は同値である．

- X は半正定値行列，i.e., $\forall d \in \mathbb{R}^n, d^T X d \geq 0$.
- X の固有値がすべて非負．
- 正の整数 m，ベクトル $v_1, \ldots, v_m \in \mathbb{R}^n$ が存在し，$X = \sum_{i=1}^m v_i v_i^T$.

$\mathcal{S}_+^n \subseteq \mathcal{K}_{\mathrm{out}} \subseteq \mathbb{S}^n$ を満たす集合 $\mathcal{K}_{\mathrm{out}}$ を $\mathcal{S}_+^n$ の外部近似と呼び，$\mathcal{K}_{\mathrm{in}} \subseteq \mathcal{S}_+^n$ を満たす集合 $\mathcal{K}_{\mathrm{in}}$ を $\mathcal{S}_+^n$ の内部近似と呼ぶ．双対の関係を用いると，$\mathcal{S}_+^n$ の内部近似 $\mathcal{K}_{\mathrm{in}}$ から外部近似 $(\mathcal{K}_{\mathrm{in}})^*$ を作ることができ，$\mathcal{S}_+^n$ の外部近似 $\mathcal{K}_{\mathrm{out}}$ から内部近似 $(\mathcal{K}_{\mathrm{out}})^*$ も作れる．

3. 半正定値行列錐の近似集合

本節では，$\mathcal{S}_+^n$ の代表的な近似集合である $\mathcal{FW}_n(k)$, $\mathcal{S}^{n,k}$, $\mathcal{DD}_n$, $\mathcal{SDD}_n$, $\mathrm{cone}(\mathcal{B}_+^n \cup \mathcal{B}_-^n)$ と $\mathcal{SDB}_n(\mathcal{H})$ について紹介する．

3.1 Factor width に基づく凸錐近似

Boman et al. [11] は，半正定値行列に対して Factor width という概念を定義し，$\mathcal{S}_+^n$ の内部近似 $\mathcal{FW}_n(k)$ 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 $A = \sum_i v_i v_i^T$，ただし，$\max_i z(v_i) = k$，の中での k の最小値である．Factor width が最大 k の行列の集合を $\mathcal{FW}_n(k)$ と表記し，下記のように定義される．

定義 1. $\mathcal{FW}_n(k) := \{\sum_{i=1}^m v_i v_i^T \mid v_i$ の非ゼロ要素が最大 k 個 $(i = 1, \ldots, m)$, $m \geq 1\}$.

上記の定義から，$\mathcal{FW}_n(k)$ が任意の $k \in \{1, \ldots, n\}$ に対して凸錐であることは検証できる．また，以下の包含関係が自然に導かれる：

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

よって，任意の $k \in \{1, \ldots, n\}$ に対して，$\mathcal{FW}_n(k)$ は $\mathcal{S}_+^n$ の内部近似である．

Blekherman et al. [12] は Factor width に基づいて $\mathcal{S}_+^n$ の外部近似 $\mathcal{S}^{n,k}$ を提案した．

定義 2. $\mathcal{S}^{n,k} := \{X \in \mathbb{S}^n \mid X$ の $k \times k$ 主小行列はすべて半正定値 $\}$.

任意の $k \in \{1, \ldots, n\}$ に対し，$\mathcal{S}^{n,k}$ は $\mathcal{FW}_n(k)$ の双対錐であることが知られている [12]．よって，$\mathcal{S}^{n,k}$ は凸錐で，$\mathcal{S}_+^n$ の外部近似である：

$$\mathcal{S}_+^n = (\mathcal{S}_+^n)^* \subseteq (\mathcal{FW}_n(k))^* = \mathcal{S}^{n,k}.$$

ある行列が $\mathcal{FW}_n(k)$（または $\mathcal{S}^{n,k}$）に属するかどうかを判断するには，$k \times k$ 行列に対する半正定値制約を $\frac{n!}{k!(n-k)!}$ 個もつ SDP を解く必要がある．その計算は $k \geq 3$ となると非効率的になる．

3.2 優対角性に基づく凸錐近似

Ahamadi et al. [5] は優対角行列の集合 $\mathcal{DD}_n$ とその拡張である $\mathcal{SDD}_n$ を $\mathcal{S}_+^n$ の内部近似として分析し，切除平面法に応用した．

定義 3. $\mathcal{DD}_n := \{A \in \mathbb{S}^n \mid A_{i,i} \geq \sum_{j \neq i} |A_{i,j}| \ (i = 1, \ldots, n)\}$.

$\mathcal{DD}_n$ は n^2 個の extreme ray（端線）をもつ有限生成錐であり [13]，凸多面錐でもある．$\mathcal{DD}_n$ の双対錐である $\mathcal{DD}_n^*$ も凸多面錐で，以下のように表せる [5]:

$$\mathcal{DD}_n^* = \{X \in \mathbb{S}^n \mid X_{i,i} + X_{j,j} \pm 2X_{i,j} \geq 0 \ (1 \leq i \leq j \leq n)\}.$$

$\mathcal{SDD}_n$ は以下のように定義される：

定義 4. $\mathcal{SDD}_n := \{A \in \mathbb{S}^n \mid$ 対角成分が正の対角行列 $D \in \mathbb{S}^n$ が存在し，$DAD \in \mathcal{DD}_n\}$.

$\mathcal{SDD}_n$ と $\mathcal{FW}_n(2)$ 等価関係は証明されている：

定理 2（文献 [11] (Theorem 8 and 9)）**.** $n \geq 2$ に対して，$\mathcal{SDD}_n = \mathcal{FW}_n(2)$ が成り立つ.

よって，$\mathcal{SDD}_n$ が凸錐であることがわかる．定理 2 と $\mathcal{FW}_n(2)$ の双対錐を用いると，$\mathcal{SDD}_n^* = (\mathcal{FW}_n(2))^* = \mathcal{S}^{n,2}$ を証明できる.

$\mathcal{DD}_n$ と $\mathcal{SDD}_n$ が $\mathcal{S}_+^n$ の内部近似であること：

$$\mathcal{DD}_n \subseteq \mathcal{SDD}_n \subseteq \mathcal{S}_+^n$$

は良く知られている [5]．よって，$\mathcal{DD}_n^*$ と $\mathcal{SDD}_n^*$ は $\mathcal{S}_+^n$ の外部近似である：$\mathcal{S}_+^n \subseteq \mathcal{SDD}_n^* \subseteq \mathcal{DD}_n^*$.

ある行列が $\mathcal{DD}_n$（または $\mathcal{DD}_n^*$）に属するかどうかの問題は LP であり，強力な LP ソルバーを用いて解くことができる．また，ある行列が $\mathcal{SDD}_n$（または $\mathcal{SDD}_n^*$）に属するかどうかの問題は SOCP であり [5]，内点法を用いて効率的に解くことも期待できる.

3.3 半正定値基に基づく凸多面錐近似

定理 1 を用いると，$\mathcal{S}_+^n$ は以下のように書き直せる：

$$\mathcal{S}_+^n = \mathrm{cone}(\{xx^T \mid x \in \mathbb{R}^n\}).$$

半正定値行列の部分集合 $\mathcal{B} \subseteq \{xx^T \mid x \in \mathbb{R}^n\} \subseteq \mathcal{S}_+^n$ から凸錐を作ることで，$\mathcal{S}_+^n$ の内部近似を構築できる：

$$\mathrm{cone}(\mathcal{B}) \subseteq \mathrm{cone}(\{xx^T \mid x \in \mathbb{R}^n\}) = \mathcal{S}_+^n.$$

限られた数の半正定値行列から精度の良い $\mathcal{S}_+^n$ の凸多面錐近似を生成するために，Tanaka and Yoshise [14] は半正定値基（SD 基）を定義した.

以下では，i 番目の要素が 1，それ以外が 0 の n 次元ベクトルを e_i とする.

定義 5. 半正定値基（SD 基）は下記のように定義される：

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

$\mathcal{B}_+^n$ を SD 基 Type I, $\mathcal{B}_-^n$ を SD 基 Type II と呼ぶ.

$\mathcal{B}_+^n$ と $\mathcal{B}_-^n$ はともに $\frac{n(n+1)}{2}$ 個の半正定値行列からなる $\mathbb{S}^n$ 上の基底集合である．よって，$\mathcal{B}_+^n \cup \mathcal{B}_-^n$ の凸錐を作ることで，$\mathcal{S}_+^n$ の凸多面錐近似を作れる：

$$\mathrm{cone}(\mathcal{B}_+^n \cup \mathcal{B}_-^n) \subseteq \mathcal{S}_+^n.$$

その近似と $\mathcal{DD}_n$ の等価関係も証明されている [13]:

$$\mathrm{cone}(\mathcal{B}_+^n \cup \mathcal{B}_-^n) = \mathcal{DD}_n. \tag{3}$$

3.4 拡張半正定値基に基づく凸多面錐近似

定義 5 からわかるように，集合 $\mathcal{B}_+^n$ と $\mathcal{B}_-^n$ は最大四つの非ゼロ要素しかもたない $n \times n$ 行列（i.e.，疎な行列）で構成されている．その基底行列の疎性は，SD 基を用いた $\mathcal{S}_+^n$ の近似の計算効率を考えるうえで重要な性質である．著者らは，基底行列の疎性を失わないように SD 基の拡張を提案し，$\mathcal{S}_+^n$ の疎な凸多面錐近似を構築した [7].

定義 6. パラメーター $\alpha \in \mathbb{R}^n$ を用いて拡張した SD 基は以下のように定義される：

$$\bar{\mathcal{B}}_n(\alpha) := \{(e_i + \alpha e_j)(e_i + \alpha e_j)^T \mid 1 \leq i \leq j \leq n\}.$$

拡張 SD 基 $\bar{\mathcal{B}}_n(\alpha)$ は，SD 基と同様，最大四つの非ゼロ要素しかもたない $n \times n$ 行列（i.e.，疎な行列）から構成される．パラメーターが $\alpha \in \mathbb{R} \setminus \{0, -1\}$ を満たす場合，$\bar{\mathcal{B}}_n(\alpha)$ は $\frac{n(n+1)}{2}$ 個の半正定値行列からなる $\mathbb{S}^n$ 上の基底集合になる．定義 5 と定義 6 により，

$$\bar{\mathcal{B}}_n(1) \cup \bar{\mathcal{B}}_n(-1) = \mathcal{B}_+^n \cup \mathcal{B}_-^n \tag{4}$$

が成り立つ．パラメーター α を調整することで，既存の SD 基行列と異なる新たな行列を生成できる.

命題 1（文献 [7] (Proposition 3)）**.** パラメーター $\alpha_1, \alpha_2 \in \mathbb{R}$ が $\alpha_1 \neq 0, -1$, $\alpha_2 \neq 0, \alpha_1$ を満たせば，任意の $1 \leq i < j \leq n$ に対して，以下の関係が成り立つ：

$$(e_i + \alpha_2 e_j)(e_i + \alpha_2 e_j)^T \notin \mathrm{cone}(\bar{\mathcal{B}}_n(\alpha_1)).$$

$\{1, -1\} \subseteq \mathcal{H} \subseteq \mathbb{R}$ を満たすパラメーター集合 $\mathcal{H}$ で生成される拡張 SD 基の凸錐を取ることで，$\mathcal{S}_+^n$ の凸多面錐近似を作ることができる.

定義 7. $\mathcal{H}: \{1, -1\} \subseteq \mathcal{H} \subseteq \mathbb{R}$ をパラメーター集合とし，拡張 SD 基を用いた $\mathcal{S}_+^n$ の凸多面錐内部近似は以下のように定義される：

$$\mathcal{SDB}_n(\mathcal{H}) := \operatorname{cone}\left(\bigcup_{\alpha \in \mathcal{H}} \bar{\mathcal{B}}_n(\alpha)\right).$$

式 (3)，式 (4) と定義 7 を用いると，

$$\mathcal{SDB}_n(\{1,-1\}) = \mathcal{DD}_n \tag{5}$$

がわかる．さらに，$\{1,-1\} \subseteq \mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \mathbb{R}$ を満たす集合 $\mathcal{H}_1, \mathcal{H}_2$ に対し，以下の包含関係を証明できる：

$$\mathcal{DD}_n \subseteq \mathcal{SDB}_n(\mathcal{H}_1) \subseteq \mathcal{SDB}_n(\mathcal{H}_2) \subseteq \mathcal{S}^n_+. \tag{6}$$

また，$\mathbb{R}$ をパラメーター集合として使う場合，拡張 SD 基の近似と $\mathcal{SDD}_n$ の等価関係も証明できる．

定理 3（文献 [7] (Corollary 2)）**.** $\mathcal{SDB}_n(\mathbb{R}) = \mathcal{SDD}_n$.

式 (2)，式 (6)，定理 2 と定理 3 を用いると，ここまで紹介した $\mathcal{S}^n_+$ の内部近似と外部近似の包含関係を以下のようにまとめられる．

系 1. $\mathcal{H} : \{1,-1\} \subseteq \mathcal{H} \subseteq \mathbb{R}$ をパラメーター集合とする．$\mathcal{S}^n_+$ の内部近似について，以下の関係が成り立つ：

$$\begin{array}{c} \mathcal{DD}_n \subseteq \mathcal{SDB}_n(\mathcal{H}) \subseteq \mathcal{SDD}_n \\ \qquad\qquad\qquad\qquad\quad \| \\ \mathcal{S}^n_+ = \mathcal{FW}_n(n) \supseteq \ldots \supseteq \mathcal{FW}_n(2). \end{array}$$

$\mathcal{S}^n_+$ の外部近似について，以下の関係が成り立つ：

$$\begin{array}{c} \mathcal{S}^n_+ = \mathcal{S}^{n,n} \subseteq \ldots \subseteq \mathcal{S}^{n,2} \\ \qquad\qquad\qquad\qquad \| \\ \mathcal{DD}^*_n \supseteq \mathcal{SDB}^*_n(\mathcal{H}) \supseteq \mathcal{SDD}^*_n. \end{array}$$

4. 半正定値行列錐の近似に対する評価

本節では，$\mathcal{S}^n_+$ の近似集合に対する評価を紹介する．$\mathcal{S}^n_+$ の近似集合は内部近似と外部近似に分けられるが，双対関係を用いることで，内部近似から外部近似，外部近似から内部近似を作ることができる．$\mathcal{S}^n_+$ の外部近似の応用は数多く知られている．たとえば，$\mathcal{DD}^*_n$ と $\mathcal{SDD}^*_n$ は SDP の前処理手法である Partial facial reduction 手法で使われており [15]，$\mathcal{SDB}^*_n(\mathcal{H})$ は SDP に対する切除平面法で初期近似として使われている [7]．以下では，$\mathcal{S}^n_+$ の外部近似，特に $\mathcal{S}^{n,k}$，$\mathcal{DD}^*_n$ と $\mathcal{SDD}^*_n$ に対する評価を解説する．

Blekherman et al. [12] は $\mathcal{S}^n_+$ の外部近似を評価する方法として，与えられた近似の行列から $\mathcal{S}^n_+$ への距離に基づく Frobenius ノルム正規化距離を提案した．外部近似集合 $\mathcal{S}$ から $\mathcal{S}^n_+$ への Frobenius ノルム正規化距離は次のように定義される：

表 1 $\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+)$ の上界と下界 [12]

k の範囲	上界	下界
$2 \le k \le n$	$\frac{n-k}{n+k-2}$	$\frac{n-k}{\sqrt{(k-1)^2 n + n(n-1)}}$

$$\overline{\mathrm{dist}}_F(\mathcal{S}, \mathcal{S}^n_+) := \sup_{X \in \mathcal{S}, \|X\|_F = 1} \|X - \mathrm{P}_{\mathcal{S}^n_+}(X)\|_F. \tag{7}$$

ただし，$\|X\|_F = \langle X, X \rangle = \sum_{i=1}^n \sum_{j=1}^n X^2_{i,j}$ は行列 X の Frobenius ノルムで，$\mathrm{P}_{\mathcal{S}^n_+}(X) := \mathrm{argmin}_{Y \in \mathcal{S}^n_+} \|X - Y\|_F$ は行列 X から $\mathcal{S}^n_+$ への射影である．

Blekherman et al. [12] の手法の特徴は，Frobenius ノルムの値が 1 であるという制約の下で距離を評価することである．外部近似集合 $\mathcal{S}$ に対し，この制約の下で，$\{X \in \mathcal{S} \mid \|X\|_F = 1\}$ は有界な集合となり，$\overline{\mathrm{dist}}_F(\mathcal{S}, \mathcal{S}^n_+)$ も有界な値を取る．

4.1 Frobenius ノルム正規化距離による $\mathcal{S}^{n,k}$, $\mathcal{DD}^*_n$ と $\mathcal{SDD}^*_n$ の評価

Blekherman et al. [12] は $\mathcal{S}^{n,k}$ を対象に，$\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+)$ の上界と下界を与えた（表 1）．

著者ら [16] は Frobenius ノルム正規化距離を用いて $\mathcal{DD}^*_n$ と $\mathcal{SDD}^*_n$ の評価を行った．

定理 4（文献 [16] (Theorem 1)）**.** $n \ge 4$ とすると，

$$\overline{\mathrm{dist}}_F(\mathcal{DD}^*_n, \mathcal{S}^n_+) = \overline{\mathrm{dist}}_F(\mathcal{SDD}^*_n, \mathcal{S}^n_+) = \frac{n-2}{n}.$$

式 (5) と定理 3 が示したように，$\mathcal{DD}^*_n = \mathcal{SDB}^*_n(\{1,-1\})$ と $\mathcal{SDD}^*_n = \mathcal{SDB}^*_n(\mathbb{R})$ は異なる集合である．しかし，定理 4 により，$\mathcal{SDD}^*_n \subseteq \mathcal{S} \subseteq \mathcal{DD}^*_n$ を満たす集合 $\mathcal{S}$ から $\mathcal{S}^n_+$ への Frobenius ノルム正規化距離はすべて同じ値をもつことがわかる．Frobenius 正規化距離が $\mathcal{DD}^*_n$ と $\mathcal{SDD}^*_n$ を区別できない理由の一つとして，正規化手法 $\|X\|_F = 1$ が厳しく，集合 $\{X \in \mathbb{S}^n \mid \|X\|_F = 1\}$ を有界にしていることが考えられる．ここで，正規化手法 (e.g., $f : \mathbb{S}^n \to \mathbb{R}$) に必要な性質は，ただ $\{X \in \mathcal{DD}^*_n \mid f(X) = 1\}$ と $\{X \in \mathcal{SDD}^*_n \mid f(X) = 1\}$ だけを有界にすることである．

著者ら [16] は (7) の正規化手法を調整し，$\mathcal{DD}^*_n$ と $\mathcal{SDD}^*_n$ の近似の評価を区別できる新たな指標：Trace

正規化距離を提案した.

4.2 Trace 正規化距離による $\mathcal{DD}_n^*$ と $\mathcal{SDD}_n^*$ の評価

外部近似集合 $\mathcal{S}$ から $\mathcal{S}_+^n$ への Trace 正規化距離は次のように定義される：

$$\overline{\mathrm{dist}}_T(\mathcal{S},\mathcal{S}_+^n) := \sup_{X\in\mathcal{S},\mathrm{Tr}(X)=1} \|X-\mathrm{P}_{\mathcal{S}_+^n}(X)\|_F.$$

ただし，$\mathrm{Tr}(X) = \sum_{i=1}^n X_{i,i}$ は行列 X のトレース (Trace) である．正規化手法 $\mathrm{Tr}(X)=1$ の下で，$\{X\in\mathcal{DD}_n^* \mid \mathrm{Tr}(X)=1\}$ と $\{X\in\mathcal{SDD}_n^* \mid \mathrm{Tr}(X)=1\}$ が有界な集合であり，$\{X\in\mathbb{S}^n \mid \mathrm{Tr}(X)=1\}$ が有界でないことは証明できる．

Blekherman et al. [12] の $\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k},\mathcal{S}_+^n)$ の上界と下界の証明方法を調整することで，著者ら [16] は $\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*,\mathcal{S}_+^n)$ の上界と下界が一致することを証明し，$\mathcal{SDD}_n^*$ から $\mathcal{S}_+^n$ への Trace 正規化距離の値を与えた．

定理 5（文献 [16] (Theorem 2)）**.** $n\geq 2$ とすると，

$$\overline{\mathrm{dist}}_T(\mathcal{SDD}_n^*,\mathcal{S}_+^n) = \frac{n-2}{n}.$$

$\{X\in\mathcal{DD}_n^* \mid \mathrm{Tr}(X)=1\}$ を $\mathcal{DDT}_n^*$ と表記すると，

$$\overline{\mathrm{dist}}_T(\mathcal{DD}_n^*,\mathcal{S}_+^n) = \max_{X\in\mathcal{DDT}_n^*} \|X-\mathrm{P}_{\mathcal{S}_+^n}(X)\|_F$$

が自明である．著者ら [16] は $\mathcal{DDT}_n^*$ の凸多面体構造を用いて $\overline{\mathrm{dist}}_T(\mathcal{DD}_n^*,\mathcal{S}_+^n)$ の値を与えた．

$\mathbb{S}^n$ 上のコンパクト（i.e., 有界かつ閉）な凸集合上の連続凸関数は，その集合の端点で最大値を取ることが知られている．また，距離関数 $\|X-\mathrm{P}_{\mathcal{S}_+^n}(X)\|_F$ が連続凸関数で，集合 $\mathcal{DDT}_n^*$ が有界な閉凸集合であることも知られている．よって，$\|X-\mathrm{P}_{\mathcal{S}_+^n}(X)\|_F$ は $\mathcal{DDT}_n^*$ の端点で最大値を取る．実は，$\mathcal{DD}_n^*$ から $\mathcal{S}_+^n$ への Trace 正規化距離は $\mathcal{DDT}_n^*$ の任意の端点から $\mathcal{S}_+^n$ までの距離を計算することで得られる．このことは，$\mathcal{DDT}_n^*$ のすべての端点が同じような特殊な構造をもつことと，その構造により，$\mathcal{DDT}_n^*$ のすべての端点から $\mathcal{S}_+^n$ までの距離が同じであると証明可能なことから導かれる．定理 6 の証明の詳細については文献 [16] を参照されたい．

定理 6（文献 [16] (Theorem 3)）**.** $n\geq 2$ とすると，

$$\overline{\mathrm{dist}}_T(\mathcal{DD}_n^*,\mathcal{S}_+^n) = \frac{\sqrt{n}-1}{2}.$$

定理 5 と定理 6 により，任意の $\mathcal{SDD}_n^* \subseteq \mathcal{S} \subseteq \mathcal{DD}_n^*$ を満たす集合 $\mathcal{S}$ の Trace 正規化距離は $\frac{n-2}{n} \leq \overline{\mathrm{dist}}_T(\mathcal{S},\mathcal{S}_+^n) \leq \frac{\sqrt{n}-1}{2}$ を満たす．よって，$\{1,-1\}\subseteq\mathcal{H}\subseteq\mathbb{R}$ を満たすパラメーター集合 $\mathcal{H}$ を用いて作られた $\mathcal{SDB}_n^*(\mathcal{H})$ も下記を満たす：

$$\frac{n-2}{n} \leq \overline{\mathrm{dist}}_T(\mathcal{SDB}_n^*(\mathcal{H}),\mathcal{S}_+^n) \leq \frac{\sqrt{n}-1}{2}.$$

集合 $\{X\in\mathcal{SDB}(\mathcal{H})_n^* \mid \mathrm{Tr}(X)=1\}$ はコンパクトな凸多面体であるため，端点の構造を用いることで，$\mathcal{DDT}_n^*$ と同様に $\overline{\mathrm{dist}}_T(\mathcal{SDB}_n^*(\mathcal{H}),\mathcal{S}_+^n)$ を求めることができる．だが，$\{X\in\mathcal{SDB}(\mathcal{H})_n^* \mid \mathrm{Tr}(X)=1\}$ の端点の構造は $\mathcal{DDT}_n^*$ のときより複雑なため，定理 6 と同じ方法で $\overline{\mathrm{dist}}_T(\mathcal{SDB}_n^*(\mathcal{H}),\mathcal{S}_+^n)$ の値を求めることは困難である．著者の博士論文では，$\mathcal{DDT}_n^*$ の端点から $\mathcal{SDB}_n^*(\mathcal{H})$ までの距離を用いて凸多面錐近似 $\mathcal{SDB}_n^*(\mathcal{H})$ の評価も行った．興味のある読者は著者の博士論文を参照されたい．

5. 半正定値計画問題に対する切除平面法

本節では，$\mathcal{S}_+^n$ の外部近似の応用として，Ahmadi et al. [5] に基づいて，問題 (1) に対する切除平面法を紹介する．ただし，以下の仮定を置く：

仮定 1. 正のスカラー $T>0$ が存在し，問題 (1) の任意の実行可能解 X は $\mathrm{Tr}(X)\leq T$ を満たす．

$\mathcal{P}^n$ を $\mathcal{S}_+^n$ の外部近似集合とする：$\mathcal{S}_+^n\subseteq\mathcal{P}^n\subseteq\mathbb{S}^n$．問題 (1) の半正定値制約 $X\in\mathcal{S}_+^n$ を $X\in\mathcal{P}^n$ に置き換えることで，問題 (1) の初期緩和問題を得ることができる：

$$\begin{aligned} &\min\ \langle C,X\rangle \\ &\ \text{s.t.}\ \langle A_j,X\rangle = b_j,\ j=1,\ldots,m, \\ &\qquad X\in\mathcal{P}^n, \end{aligned} \tag{8}$$

初期緩和問題 (8) に関して，以下を仮定する：

仮定 2. 問題 (8) の実行可能領域は有界な集合である．

初期緩和問題 (8) を解いて得られた解を X^* とする．X^* が半正定値行列であれば，X^* が元問題 (1) の最適解になり，切除平面法は停止する．X^* が半正定値でないとき，$\langle C,X^*\rangle$ は元問題 (1) の最適値の下界を与え，X^* の負の固有値をもつ．その場合，切除平面法は X^* は負の固有値に対応する固有ベクトル $d_1,\ldots,d_k$ を用いて，以下の切除平面

$$\langle d_i d_i^T, X\rangle \geq 0 \quad (i=1,\ldots,k)$$

を緩和問題 (8) の制約に加えることで，新たな緩和問題を作る：

$$\begin{aligned} &\min\ \langle C, X\rangle \\ &\text{s.t. } \langle A_j, X\rangle = b_j,\ j=1,\ldots,m, \\ &\qquad \langle d_i d_i^T, X\rangle \geq 0, i=1,\ldots,k, \\ &\qquad X \in \mathcal{P}^n. \end{aligned} \tag{9}$$

$d_1,\ldots,d_k$ が X^* の負の固有値の固有ベクトルであるため，$\langle d_i d_i^T, X^*\rangle < 0$ $(i=1,\ldots,k)$ が成り立つ．よって，初期緩和問題 (8) の最適解 X^* が新たな緩和問題 (9) の実行可能領域からカットされ，緩和問題 (9) は元問題 (1) のよりタイトな緩和問題になる．以上の手順を繰り返すことで，元問題 (1) に対する一連の非減少な下界を得ることができる．

以上で紹介した SDP に対する切除平面法はアルゴリズム 1 にまとめられる．

アルゴリズム 1. 問題 (1) に対する切除平面法.

Step 0 許容誤差を $\epsilon > 0$ とし，仮定 2 を満たす $\mathcal{S}_+^n$ の初期外部近似を $\mathcal{P}_0^n$ とする．$k \leftarrow 0$ を設定する．

Step 1 下記の問題を解き，得られた解を X_k^* とする．

$$\begin{aligned} &\min\ \langle C, X\rangle \\ &\text{s.t. } \langle A_j, X\rangle = b_j,\ j=1,\ldots,m, \\ &\qquad X \in \mathcal{P}_k^n, \end{aligned} \tag{10}$$

Step 2 問題 (10) が実行不可能なら，元問題 (1) も実行不可能なので，終了する．X_k^* の最小固有値 $\lambda_{\min}(X_k^*)$ が $\lambda_{\min}(X_k^*) > -\epsilon$ を満たせば，X_k^* は元問題 (1) の ϵ 近似解になるので，終了する．

Step 3 X_k^* の最小固有値に対応する固有ベクトルを d_k とし，切除平面を制約に加える．

$$\mathcal{P}_{k+1}^n \leftarrow \{X \in \mathcal{P}_k^n \mid \langle d_k d_k^T, X\rangle \geq 0\}.$$

Step 4 $k \leftarrow k+1$ とし，Step 1 に戻る．

$\mathcal{SDD}_n^*$ を初期近似とした場合のアルゴリズム 1 の収束性は証明されている [6]．また，$\mathcal{DD}_n^*$ または $\mathcal{SDB}_n^*(\mathcal{H})$ を初期近似とするアルゴリズム 1 の収束性についても同様に証明できる．

定理 7（文献 [6] (Theorem 1)）**.** 問題 (1) が仮定 1 を満たし，アルゴリズム 1 の初期近似集合 $\mathcal{P}_0^n$ が仮定 2 を満たすとすると，任意の許容誤差 $\epsilon > 0$ に対し，アルゴリズム 1 は有限回の反復で終了する．

仮定 1 の下で，$\mathcal{DD}_n^*, \mathcal{SDD}_n^*$ と $\mathcal{SDB}_n^*(\mathcal{H})$ を (8) の初期近似 $\mathcal{P}^n$ とすると，仮定 2 が成り立つことが証明できる．$\mathcal{SDD}_n^*$ を初期近似とした場合，緩和問題 (8) は SOCP になり，$\mathcal{DD}_n^*$ または $\mathcal{SDB}_n^*(\mathcal{H})$ を初期近似にする場合，緩和問題 (8) は LP になる．著者らは最大独立集合問題の SDP 緩和問題に対して，$\mathcal{DD}_n^*, \mathcal{SDD}_n^*$ と $\mathcal{SDB}_n^*(\mathcal{H})$ を初期近似とした切除平面法の計算効率を比較し，$\mathcal{SDB}_n^*(\mathcal{H})$ の切除平面法が一番効率的であることを示した [7].

6. おわりに

本稿では，著者の博士論文を基に，半正定値行列錐のいくつかの近似集合，特に SD 基に基づく凸多面錐近似について述べ，それぞれの近似精度の評価を説明し，半正定値計画問題に対する切除平面法についても紹介した．本稿が半正定値行列錐の近似集合に関する研究に少しでも興味をもたらすことができれば幸いである．

謝辞 本稿を執筆する機会を下さりました編集委員の先生方，著者の博士論文に有益なコメントをくださった吉瀬章子先生に，この場を借りて深く御礼申し上げます．

参考文献

[1] H. Wolkowicz, R. Saigal and L. Vandenberghe, *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, Springer Science & Business Media, 2000.

[2] F. Alizadeh, “Combinatorial optimization with interior point methods and semi-definite matrices,” Ph. D. thesis, University of Minnesota, 1991.

[3] MOSEK ApS, The MOSEK optimization toolbox for MATLAB manual. Version 9.0., http://www.mosek.com（2021 年 3 月 12 日閲覧）

[4] A. Majumdar, G. Hall and A. A. Ahmadi, “Recent scalability improvements for semidefinite programming with applications in machine learning, control, and robotics,” *Annual Review of Control, Robotics, and Autonomous Systems* **3**, pp. 331–360, 2020.

[5] A. A. Ahmadi, S. Dash and G. Hall, “Optimization over structured subsets of positive semidefinite matrices via column generation,” *Discrete Optimization*, **24**, pp. 129–151, 2017.

[6] D. Bertsimas and R. Cory-Wright, “On polyhedral and second-order cone decompositions of semidefinite optimization problems,” *Operations Research Letters*,

48, pp. 78–85, 2020.
[7] Y. Wang, A. Tanaka and A. Yoshise, “Polyhedral approximations of the semidefinite cone and their application,” *Computational Optimization and Applications*, **78**, pp. 893–913, 2021.
[8] Gurobi Optimization, Gurobi optimizer reference manual, http://www.gurobi.com（2021 年 8 月 15 日閲覧）
[9] A. Schrijver, *Theory of Linear and Integer Programming*, John Wiley & Sons, 1998.
[10] A. Berman and N. Shaked-Monderer, *Completely Positive Matrices*, World Scientific, 2003.
[11] E. G. Boman, D. Chen, O. Parekh and S. Toledo, “On factor width and symmetric H-matrices,” *Linear Algebra and Its Applications*, **405**, pp. 239–248, 2005.
[12] G. Blekherman, S. S. Dey, M. Molinaro and S. Sun, “Sparse PSD approximation ofthe PSD cone,” *Mathematical Programming*, **191**, pp. 981–1004, 2022.
[13] G. Barker and D. Carlson, “Cones of diagonally dominant matrices”, *Pacific Journal of Mathematics*, **57**, pp. 15–32, 1975.
[14] A. Tanaka and A. Yoshise, “LP-based tractable subcones of the semidefinite plus nonnegative cone,” *Annals of Operations Research*, **265**, pp. 155–182, 2018.
[15] F. Permenter and P. Parrilo, “Partial facial reduction: simplified, equivalent SDPs via approximations of the PSD cone,” *Mathematical Programming*, **171**, pp. 1–54, 2018.
[16] Y. Wang and A. Yoshise, “Evaluating approximations of the semidefinite cone with trace normalized distance,” *arXiv preprint*, arXiv:2105.13579, 2021.