日本オペレーションズ・リサーチ学会
2021年 秋季研究発表会

半正定値計画問題に対する下三角低ランク行列分解を利用した等価な非線形計画問題について

05000141　京都大学　*山川雄也　YAMAKAWA Yuya
ヤフー株式会社　池上哲哉　IKEGAMI Tetsuya
05000235　京都大学　福田エレン秀美　FUKUDA Ellen Hidemi
01704632　京都大学　山下信雄　YAMASHITA Nobuo

1. はじめに

本研究では，次の半正定値計画問題を扱う．

$$
\text{(SDP)}\qquad \begin{array}{ll} \underset{X\in\mathbb{S}^n}{\text{Minimize}} & \langle C, X\rangle \\ \text{subject to} & \mathcal{A}(X)=b,\quad X\succeq O \end{array}
$$

ここで，$\mathbb{S}^n$ は n 次実対称行列全体の集合を表すとし，$X \succeq O$ は行列 X が半正定値であることを意味する．また，関数 $\mathcal{A}\colon \mathbb{S}^n \to \mathbb{R}^m$ は $\mathcal{A}(X) := [\langle A_1, X\rangle, \dots, \langle A_m, X\rangle]^\top$ で定義されるとし，行列 $C, A_1, \dots, A_m \in \mathbb{S}^n$ およびベクトル $b \in \mathbb{R}^m$ は与えられているとする．

SDP は，線形計画問題や二次錐計画問題などを含む広いクラスの最適化問題であり，多くの分野で応用を持つことで知られる．また，SDP を解くため，これまで多くの解法が提案されてきた．中でも，主双対内点法は最も主要な解法の一つとして知られ，多くのソルバーで実装されている．しかし，小中規模の問題に対しては高精度な解を得られると言われているが，大規模な問題に対しては，解の近傍においてニュートン方程式の計算が不安定になるという欠点も指摘されている．

行列のサイズが大きく（n が大きく），制約の本数が少ない（m が小さい）大規模な SDP を解くために，Burer et al [1, 2] は，SDP と等価な非線形計画問題（NLP）を提案した．この既存モデルは，行列 $X \in \mathbb{S}^n$ を矩形行列 $R \in \mathbb{R}^{n\times r}$ の積 $RR^\top$ で置き換えることで SDP の半正定値制約を取り除き，一般的な NLP へと帰着される．さらに，$r \geq \max\{r \in \mathbb{N}\colon r(r+1)/2 \leq m\}$ を満たせば，既存モデルは SDP と等価となることも示しており，r の値によっては SDP よりも次元が小さい問題となるというメリットがある．しかし，このモデルは狭義局所的最小解が存在しないことから，最適性の二次の十分条件が成り立たないため，局所的に速い収束性を持つ手法が有効に機能しない可能性がある．また，m の値が十分に小さくなければ，提案モデルの次元は SDP の次元より大きくなる可能性もある．

本発表では，Burer らの提案モデルを改良し，SDP と等価な NLP を提案する．既存の Burer らのモデルでは $\mathbb{R}^{n\times r}$ を決定変数の空間としているが，提案モデルでは $\mathbb{R}^{n\times r}$ に含まれる下三角行列全体の集合を決定変数の空間とする．従って，既存モデルに比べて変数の次元は小さくなり，また，m の値が大きい場合でも，元の SDP の次元を超えることはない．さらに，決定変数の空間を制限したことで，適切な仮定の下での狭義局所的最適解の存在を理論的に保証することが可能となる．これにより，二次法などの手法を用いることで，解への速い収束性が期待される．ここでは，提案モデルに対して逐次二次計画法（SQP 法）も提案し，数値実験によりその有効性を示す．

2. SDP を解くための既存モデル

Burer et al [1] は，SDP を直接解く代わりに，以下のランク制約付き SDP を解くことを考えた．

$$
(\text{LRSDP}_r)\qquad \begin{array}{ll} \underset{X\in\mathbb{S}^n}{\text{Minimize}} & \langle C, X\rangle \\ \text{subject to} & \mathcal{A}(X)=b,\ X\succeq O, \\ & \operatorname{rank}(X)\leq r \end{array}
$$

この問題は，$r \geq r_m := \max\{r \in \mathbb{N}\colon r(r+1)/2 \leq m\}$ のとき，SDP と等価となることが知られている [1, Theorem 1]．ここで，$X \succeq O$ かつ $\operatorname{rank}(X) \leq r$ を満たす行列 $X \in \mathbb{S}^n$ は，ある矩形行列 $R \in \mathbb{R}^{n\times r}$ を用いて $X = RR^\top$ と分解されるので，LRSDP_r は以下のように書き換えることができる．

$$
(\text{NSDP}_r)\qquad \begin{array}{ll} \underset{R\in\mathbb{R}^{n\times r}}{\text{Minimize}} & \langle C, RR^\top\rangle \\ \text{subject to} & \mathcal{A}(RR^\top)=b \end{array}
$$

NSDP_r は，以下で述べるような SDP と関連した重要な性質を持つ．

命題 1 ([1]) R^* は NSDP_r の停留点，つまり，ある $v^* \in \mathbb{R}^m$ が存在して，$(C - \mathcal{A}^*(v^*))R^* = 0$ を満たすとする．ただし，$\mathcal{A}^*(v^*) := v_1^* A_1 + \cdots + v_m^* A_m$ である．もし，$C - \mathcal{A}^*(v^*) \succeq O$ であれば，$R^*(R^*)^\top$ は SDP の最適解である．□

命題 2 ([2]) $r \geq r_{m+1}$ であり，R^* は NSDP_r の局所最小解とする．もし，SDP は唯一解を持つならば，それは $R^*(R^*)^\top$ で与えられる．□

命題1および命題2で述べられているように，適切な条件を満たしていれば NSDP_r を解くことで SDP の解を得ることができる．しかし，NSDP_r の局所最小解 R に対して，任意の直交行列 Q を R の右から掛けた行列 RQ もまた NSDP_r の局所最小解となる．従って，Q をうまく選べば，RQ は R にいくらでも近づけることができるため，任意の局所最小解 R は狭義局所的最小解にはならない．つまり，二次の十分条件が成り立たないため，二次法などの速い収束性を持つ手法を適用しても，その性能を十分に発揮できない可能性がある．

3. 提案モデルとその性質

以下では，$\mathbb{R}^{n\times r}$ に含まれる下三角行列全体の集合を $\mathbb{L}^{n\times r}$ で表す．まず，半正定値行列の下三角行列を用いた分解に関する命題を与える．

命題 3 $X \succeq O$ かつ $\mathrm{rank}(X) = r < n$ を満たす行列 $X \in \mathbb{S}^n$ に対して，ある $S \in \mathbb{L}^{n\times r}$ が存在して，$X = SS^\top$ を満たす．□

命題3を利用すると，LRSDP_r は NSDP_r とは異なる次の最適化問題へと書き換えることができる．

$$(\text{T-NSDP}_r) \quad \begin{array}{ll} \underset{S\in\mathbb{L}^{n\times r}}{\text{Minimize}} & \langle C, SS^\top \rangle \\ \text{subject to} & \mathcal{A}(SS^\top) = b \end{array}$$

上記の T-NSDP_r が，本研究で提案するモデルであり，NSDP_r で見られるような類似の性質を持つ．

命題 4 S^* は T-NSDP_r の停留点，つまり，ある $w^* \in \mathbb{R}^m$ が存在して，$(C - \mathcal{A}^*(w^*))S^* = 0$ を満たすとする．もし，$C - \mathcal{A}^*(w^*) \succeq O$ であれば，$S^*(S^*)^\top$ は SDP の最適解である．□

命題 5 $r \geq r_{m+1}$ であり，S^* は T-NSDP_r の局所最小解とする．もし，SDP は唯一解を持つならば，それは $S^*(S^*)^\top$ で与えられる．□

T-NSDP_r は，NSDP_r にはない複数のメリットがある．一つは，NSDP_r に比べて，決定変数の空間の次元が，$r(r-1)/2$ 次元少なくなる．また，$r = n$ のとき，NSDP_r の決定変数の次元は元の SDP の次元を超えてしまうが，T-NSDP_r は最大でも $n(n+1)/2$ 次元であるため，SDP の次元を超えることはない．さらに，決定変数の空間を制限したことで，以下で述べるように適切な条件の下で狭義局所的最適解を持つことが示される．

定理 1 SDP は唯一解 X^* を持ち，以下を満たすと仮定する．

$$X^* = \begin{bmatrix} X_1^* & O \\ O & O \end{bmatrix}, \ X_1^* \in \mathbb{S}^\ell, \ \mathrm{rank}(X_1^*) = \ell$$

また，$r \geq r_{m+1}$ であり，$S^* \in \mathbb{L}^{n\times r}$ が T-NSDP_r の局所最適解であるとする．このとき，$S^* \in \mathbb{L}^{n\times r}$ は T-NSDP_r の 狭義局所最適解となる．□

ここで，定理1では X^* に対してブロック対角行列であることを要求しているが，これは一般性を失わず仮定することができる．定理1により，T-NSDP_r に対して二次法などの手法を適用すれば，最適解への速い収束性が期待される．本研究では，最適化手法として SQP 法を提案する．ただし，NSDP_r および T-NSDP_r は，ともに次の二次等式制約付き二次計画問題へと変形可能であるため，この問題に対して SQP 法を提案し，数値実験を行う．

$$\begin{array}{ll} \underset{x\in\mathbb{R}^d}{\text{Minimize}} & \langle Hx, x \rangle \\ \text{subject to} & \langle G_j x, x \rangle = b_j, \ j = 1, \ldots, m \end{array}$$

ただし，$d \in \mathbb{N}$，$H \in \mathbb{S}^d$ および $G_j \in \mathbb{S}^d$ $(j = 1, \ldots, m)$ は，NSDP_r や T-NSDP_r に応じて決定される．提案アルゴリズムの詳細および実験結果については当日発表する．

参考文献

[1] Burer, S., Monteiro, R.D.C.: A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. Math. Program. Ser. B **95**, 329–357 (2003)

[2] Burer, S., Monteiro, R.D.C.: Local minima and convergence in low-rank semidefinite programming. Math. Program. Ser A **103**, 427–444 (2005)