

科学研究費助成事業 研究成果報告書

平成 27 年 5 月 26 日現在

機関番号：12608

研究種目：若手研究(B)

研究期間：2012～2014

課題番号：24710161

研究課題名(和文)非線形半正定値計画問題に対する錐への接近を考慮した信頼領域法の構築と実装

研究課題名(英文)An iterative method for nonlinear semidefinite programming with the distance information to a cone boundary

研究代表者

山下 真(Yamashita, Makoto)

東京工業大学・情報理工学(系)研究科・准教授

研究者番号：20386824

交付決定額(研究期間全体):(直接経費) 2,300,000円

研究成果の概要(和文):非線形半正定値計画問題は、変数行列の固有値に関する制約を課した数理最適化問題であり、金融工学や不確実なデータなどを扱う最適化などに利用できる。

本研究では、上下限制約付き非線形半正定値計画問題に対して、現在の反復点と半正定値行列のなす錐の境界までの間の距離情報を探索方向に取り込むことにより、効率的に求解する反復計算手法を提案した。

数値実験を通して、本研究の提案手法は3次以上の項を含むような非線形関数に対して実行可能方向法より短時間で求解を達成しており、錐の境界までの情報の有効性が示された。

研究成果の概要(英文):Nonlinear semidefinite programs are optimization problems with a constraint on the eigenvalues of a variable matrix, and the range of applications includes optimization with uncertainty and financial engineering.

We propose an iterative method for box-constrained nonlinear semidefinite programs using the distance information from the current point to the boundary of a positive semidefinite cone.

We verified through numerical tests that the proposed method is more effective for box-constrained problems that have strongly nonlinear functions than a feasible direction method.

研究分野：応用数学

キーワード：数理最適化 非線形最適化 半正定値計画問題 数値計算

1. 研究開始当初の背景

非線形半正定値計画問題 (Nonlinear semidefinite programs, 以下 Nonlinear SDP) とは、制約に行列の半正定値条件を含み、目的関数が非線形関数となる数理最適化問題である。一般には、以下の形式で表現される。

$$\min f(x) \text{ s.t. } X = C - \sum_{i=1}^m A_i x_i \succeq O \quad (1-1)$$

ここで変数はベクトル $x \in R^m$ と n 次元対称行列 X であり、入力関数 $f: R^m \rightarrow R$ と n 次元対称行列 C, A_1, A_2, \dots, A_m である。また、 $X \succ O$ は行列 X に関する半正定値条件 (X のすべての固有値が非負という条件) であり、この制約が Nonlinear SDP を特徴づけている。

Nonlinear SDP は統計や金融などにも利用されており、近年では不確実性を伴う最適化などへの活用も議論されている。Nonlinear SDP を求解するための計算手法としては、内点法をベースとした手法やペナルティ関数・バリア関数を用いる手法などが提案されてきた。

Nonlinear SDP の実行可能領域のうち半正定値制約 $X \succ O$ の形状を図示すると、図1のように錐の形状をなす。

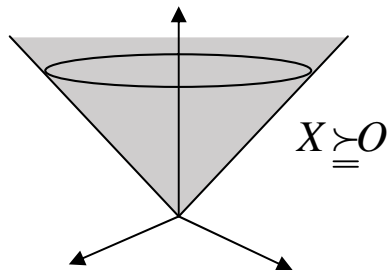


図1：行列半正定値制約のなす形状

Nonlinear SDP の数学的性質としては、錐の境界に最適解が存在する 경우가多いことが知られている。しかしながら、既存の内点法をそのまま手法を適用すると、この錐の境界条件を十分に捉えることが困難であり、錐の境界を考慮に入れた Nonlinear SDP への数学的解法が必要とされている。

2. 研究の目的

本研究では、錐の境界を考慮した数学的解法の構築を考察する。このため、上下限制約付き非線形最適化問題

$$\min f(x) \text{ s.t. } l \leq x \leq u \quad (2-1)$$

に対して Coleman and Li が 1996 年に提案した信頼領域法 (以下 CL1996) の手法をベースとすることとした。なお、この上下限制約付き制約問題(2-1)において、変数ベクトルは $x \in R^n$ であり、 $l \leq x \leq u$ は変数ごとの上下限制約 $l_i \leq x_i \leq u_i$ を示している。変数同

士は独立しており、行列の半正定値条件よりも簡単な問題となっている。

信頼領域法は反復計算の一種であり、 k 回目の反復点 x^k から探索方向 d^k を計算し、 $x^{k+1} = x^k + d^k$ として次の反復点への移動を、最適解に収束するまで繰り返す計算である。CL1996 の特徴としては、探索方向 d^k を計算する過程で現在の反復点 x^k における境界までの距離情報を取り込んでいる点が挙げられ、制約付き最適化問題に対する基本的なアプローチとして利用されている。

本研究の主な目的は、錐の境界への距離情報を取り込むことで、以下の上下限付き非線形半正定値計画問題 (box-constrained nonlinear semidefinite programs, 以下 box-constrained SDPs)

$$\min f(X) \text{ s.t. } O \preceq X \preceq I \quad (2-2)$$

への効率的な計算手法を構築することである。ここで変数は対称行列 X であり、「固有値が 0 から 1 の範囲」という上下限制約が付加されている。

この box-constrained SDP は、一般の非線形半正定値計画問題の特殊系とみなせるが、入力データの不確実性を数学的に記述できるなどの性質を保有しており、また、他の手法の子問題などにも表れるため、box-constrained SDP に対する計算手法は様々に活用できる。

また、本研究で構築した計算手法について、ソフトウェアによる実装を行う。実装したソフトウェアはフリーソフトウェアとして公開し、広く活用できるようにする。

3. 研究の方法

本研究では、box-constrained SDP に対する計算手法の構築を信頼領域法に近い枠組みを利用して行うこととした。通常の信頼領域法では、3つの要素、(1)探索方向、(2)2次近似関数、(3)信頼半径の更新、が重要な構成要素である。本研究のアプローチにおいても、この3つの点を考慮した。

(1) 探索方向

変数がベクトルである問題(2-1)に対して通常の信頼領域法では、反復点 x^k における探索方向 d^k を計算するために

$$\min g^k(d^k) \text{ s.t. } \|d^k\| \leq \Delta^k \quad (3-1)$$

となるような補助問題を解く。ここで、 g^k は(2-1)の目的関数 f の x^k における2次近似関数であり、

$$g^k(d) = f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T \nabla^2 f(x^k) d$$

のように1階の微分情報である勾配 ∇f と2階の微分情報であるヘッセ行列 $\nabla^2 f$ を用い

て構成される。また、 Δ^k は信頼半径と呼ばれ、近似関数 $g^k(d)$ が $f(x^k + d)$ を精度高く近似できる範囲を表している。

しかしながら、今回の box-constrained SDP (2-2) に対して、同様の補助問題を解くことは計算コストの観点から困難である。実際に、行列 X の次元を n とすると、対称行列であることから独立した変数は $\frac{n(n+1)}{2}$ 個

存在し、ヘッセ行列は $\left(\frac{n(n+1)}{2}\right)^2$ 個の要素を保持する必要がある。例えば、 $n = 500$ のようなサイズであっても、14GB のメモリ空間を必要とし、 $n \geq 2000$ などのサイズを通常の PC で解くことは不可能である。

そこで、本研究では、探索方向の計算に補助問題を解くのではなく、探索方向をまず求めてから、信頼半径に入る長さに調整するアプローチをとった。これにより、 $n \geq 5000$ なども扱うことを可能とした。

本研究での探索方向としては、現在の反復点 X^k について、

$$D^k = -P \begin{pmatrix} V_+^{1/2} \Gamma_+ V_+^{1/2} & \gamma P_+^T X^k P_- \\ \gamma P_-^T X^k P_+ & V_-^{1/2} \Gamma_- V_-^{1/2} \end{pmatrix} P^T \quad (3-2)$$

と与えることとした。ここで表記としては、まず P, Γ は $\nabla f(X^k)$ の固有値分解 $\nabla f(X^k) = P \Gamma P^T$ を満たす行列とする。また、固有値の行列 Γ を、固有値の符号により Γ_+ と Γ_- にわけ、それぞれに対応する固有ベクトルを列ベクトルとして構成される行列を P_+, P_- とする。さらに行列 V_+, V_- を $V_+ = P_+^T X^k P_+$, $V_- = P_-^T (I - X^k) P_-$ により定義する。また、 γ は $\gamma = \|\nabla f(X^k)\|_2$ として定義される。

これらの定義のうち、 V_+, V_- はそれぞれ X^k から錐形状の制約 $X \succeq O, I - X \succeq O$ までどれだけ距離があるかを示している。この2つの行列を介して、(3-2)の探索方向 D^k は錐の境界に関する距離情報を取り込んでおり、CL1996 の性質に則した探索方向として設計できている。

また、一般に固有値に重複があるとき、固有ベクトルは連続関数とは限らない。この性質を考慮せずに設計した探索方向 \hat{D} を用いてしまうと、ステップ長 α^k により次の反復点を $X^{k+1} = X^k + \alpha^k \hat{D}$ としようとしても、 $O \preceq X^{k+1} \preceq I$ とするためには、 $\alpha^k = 0$ となり反復点が更新されず最適解に到達できないという問題が起こりうる。

今回提案した探索方向 D^k では、(3-2)式に表れる γ の性質を利用すると、ある $\alpha_{\max} > 0$ が存在し、 $0 \leq \alpha \leq \alpha_{\max}$ となる α について $O \preceq X^k + \alpha D^k \preceq I$ を満足することが保証され、上記のような問題点を回避できる。

(2) 2次近似関数

通常の信頼領域法の補助問題(3-1)に対して、本研究の計算手法では以下のような補助問題を用いる。

$$\min g_2^k(\alpha) \quad s.t. \quad 0 \leq \alpha \leq \min\{\alpha_{\max}, \Delta^k\} \quad (3-3)$$

ここで2次近似関数は

$$g_2^k(\alpha) = f(X^k) + \alpha \langle \nabla f(X^k) | D^k \rangle + \frac{\alpha^2}{2} \langle D^k | \nabla^2 f(X^k) | D^k \rangle$$

と設計した。この2次近似関数を図示したのが図2である。

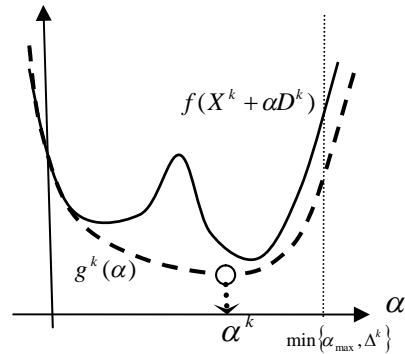


図2：2次近似関数の概念図

本研究のアプローチでは探索方向を計算してから補助問題を計算しているため、ヘッセ行列 $\nabla^2 f(X)$ の全要素のメモリ上への保持は不要となっている。例えば、 $f(X) = \log \langle X | X \rangle$ という目的関数であれば2階の微分に関する情報は

$$\langle D | \nabla^2 f(X) | D \rangle = \frac{\langle D | D \rangle}{\langle X | X \rangle} - 2 \frac{\langle X | D \rangle^2}{\langle X | X \rangle^2}$$

と与えられ、1反復あたりの計算コストの面でもヘッセ行列の逆行列計算などを回避できる点が有利となる。

(3) 信頼半径の更新

信頼半径 Δ^k については、元の目的関数 $f(X)$ と2次近似関数 $g^k(\alpha)$ の乖離が小さければ近似の精度が高いと判断して拡大をする。より具体的には、(3-3)の補助問題で求めたステップ長を α^k としたとき、近似の精度を

$$r_k = \frac{f(X^k) - f(X^k + \alpha^k D^k)}{f(X^k) - g_2^k(\alpha^k)}$$

として求め、

$$\Delta^{k+1} = \begin{cases} \eta_1 \Delta^k & \text{if } r_k < \mu_1 \\ \Delta^k & \text{if } \mu_1 \leq r_k \leq \mu_2 \\ \eta_2 \Delta^k & \text{if } r_k > \mu_2 \end{cases}$$

により信頼半径を Δ^{k+1} に更新する。ただし、 $\mu_1, \mu_2, \eta_1, \eta_2$ は $0 < \mu_1 < \mu_2 < 1$, $0 < \eta_1 < 1 < \eta_2$ を満たすパラメータである。

この信頼領域の調整の枠組みにより、2次近似関数の近似精度が高い反復では、より大きなステップ長 α^k を利用可能なが保証され、全体の反復回数を押さえる効果がある。

以上に記したように、本研究で構築した計算手法をまとめると、現在の反復点 X^k から、まず探索方向 D^k を計算し、補助問題によりステップ長 α^k を求める。これらの情報から信頼半径 Δ^k を更新する。この計算を一次の最適性条件に十分に近づくまで反復を行う。

4. 研究成果

(1) 提案した計算手法の理論的性質の解明

理論的性質の解明として、本研究では一次の最適性条件への大域的収束性を証明した。

対称行列 X^* の box-constrained SDP に対する一次の最適性条件を満たすことの定義は、

$$\langle \nabla f(X^*) | X - X^* \rangle \geq 0 \quad \text{for } O \preceq X \preceq I$$

である。この定義は、関数

$$\underline{f}(\hat{X}) = \min \{ \langle \nabla f(\hat{X}) | X - \hat{X} \rangle : O \preceq X \preceq I \}$$

を定義したときに、 X^* が

$$\underline{f}(X^*) = 0$$

を満たすこととしても表現できる。

本研究では、提案した計算手法で生成される点列 $\{X^k\}$ により

$$\lim_{k \rightarrow \infty} \underline{f}(X^k) = 0$$

が成立することを証明し、これにより一次の最適性条件に収束することを示した。

この証明の過程では、固有ベクトルから構成される行列 P_+, P_- が行列 X の連続関数と限らないために、通常信頼領域法の証明方法をそのまま利用することは困難であった。本研究では、まず、上記の条件との同値条件を導出し、 $\langle \Gamma_+ | V_+ \rangle$ と $\langle \Gamma_- | V_- \rangle$ で表現できることを示した。さらに、 $\underline{f}(X)$ の下限も $\langle \Gamma_+ | V_+ \rangle$ および $\langle \Gamma_- | V_- \rangle$ で評価できることを利用し、 $\langle \Gamma_+ | V_+ \rangle$ と $\langle \Gamma_- | V_- \rangle$ の取り得る範囲を示すことで P_+, P_- が連続関数とならないという難点を回避している。

これら証明の過程で得た数学的性質は、錐の境界を考慮するアルゴリズムを更に発展させる場合にも、適用できると考えられる。

(2) 数値実験結果

本研究で設計した提案手法の有効性を確認するために、計算時間により他の手法と比較したのが以下の表1である。比較対象としては同じく box-constrained SDP (2-2) に対して Xu らによって提案された実行可能方向法 (FD)、および PENLAB に実装された Penalty/Barrier Method (PBM) とした。また、数値実験で使用した関数は以下の3タイプである。

タイプ1 : $f(X) = -2\langle C | X \rangle + \langle X | X \rangle$

タイプ2 : $f(X) = \frac{\langle X | X \rangle^3}{n^3}$

タイプ3 :

$$f(X) = \langle C | X \rangle - \log \det(X + \bar{\varepsilon}I) - \log \det((1 + \bar{\varepsilon})I - X)$$

ここで、パラメータとして C は乱数で発生した対称行列、 $\bar{\varepsilon} = 0.02$ とした。数値実験に用いた環境は、CPU が AMD Opteron 4386 (3.1GHz)、メモリが 128GB である。なお、表1における OOM は Out of Memory (メモリ不足) を示している。

表1 : 計算時間による性能評価

タイプ	提案手法	FD	PBM
タイプ1			
n = 100	0.30 秒	0.04 秒	323.70 秒
n = 5000	3497 秒	1112 秒	OOM
タイプ2			
n = 100	0.11 秒	0.05 秒	10210 秒
n = 5000	367.7 秒	625.7 秒	OOM
タイプ3			
n = 100	0.11 秒	0.30 秒	3247 秒
n = 5000	1708.11 秒	4931.70 秒	OOM

表1の結果より、提案手法と実行可能方向法は Penalty/Barrier Method よりも短い計算時間で計算しており、メモリ使用量も抑えられていることが見て取れる。Penalty/Barrier Method では、2階の微分情報であるヘッセ行列 $\nabla^2 f(X)$ をすべて保持しているだけでなく、計算過程で $\nabla^2 f(X)$ の逆行列も利用しており、大規模な問題に適していない。提案手法では、 $\nabla^2 f(X)$ の情報を $\langle D | \nabla^2 f(X) | D \rangle$ という2次形式でのみ利用するため、より少ない計算コストで計算可能である。

また、表1ではタイプ1の関数については実行可能方向法が提案手法よりも有効であるが、より高次の項を含むタイプ2, 3については提案手法が短い計算時間で計算している。提案手法では、錐の境界への距離を探索方向 D^k に V_+, V_- を介して取り込んでいる。この結果、提案手法では少ない反復回数で一次の最適性条件を満たすことが可能であり、錐の境界を考慮した計算手法の有効性が示唆されている。

本研究では1次の最適性条件について証明を与えたが、理論的な発展としては2次の最適性条件についての議論も考えられる。ただし、Nonlinear SDP に対する2次の最適性条件はヘッセ行列 $\nabla^2 f(X)$ に追加項を加える必要があり、これまでの本研究の仮定よりも強い仮定が必要と推測される。

(3) ソフトウェアの公開および学術雑誌への投稿

本研究で設計した計算手法はソフトウェアとして実装し、これをフリーソフトとしてホ

ホームページ上で公開を行った。これにより、本研究の計算手法を他の研究などにも活用可能である。

また、一次の最適性条件への収束などの理論的解析については、数値実験結果と合わせて論文としてまとめて、学術雑誌への投稿を行った。

(4) SDP に関する研究の発展

Nonlinear SDP には、行列多項式や L_1 ノルムなどを目的関数に用いる場合も含まれている。このような問題に対するアルゴリズムについて発表を行った。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計4件)

(1) Makoto Yamashita and Kazuhide Nakata, "Fast implementation for semidefinite programs with positive matrix completion," *Optimization Methods and Software*, 掲載決定, 査読有
DOI: 10.1080/10556788.2015.1014554

(2) Jon Ahlinder, Tim J Mullin, and Makoto Yamashita, "Using semi-definite programming to optimize unequal deployment of genotypes to a clonal seed orchard," *Tree Genetics & Genomes*, Vol. 10, No. 1, 2014, pp 27 - 34, 査読有
DOI: 10.1007/s11295-013-0659-z

(3) Mizutani Tomohiko, and Makoto Yamashita, "Correlative Sparsity Structures and Semidefinite Relaxations for Concave Cost Transportation Problems with Change of Variables," *Journal of Global Optimization*, Vol. 56, No. 3, 2013, pp. 1073 - 1100, 査読有
DOI: 10.1007/s10898-012-9924-1

(4) James S. M. Anderson, Maho Nakata, Ryo Igarashi, Katsuki Fujisawa, and Makoto Yamashita, "The second-order reduced density matrix method and the two-dimensional Hubbard model," *Computational and Theoretical Chemistry*, Vol 1003, 2013, pp. 22 - 27, 査読有
DOI: 10.1016/j.comptc.2012.08.018

[学会発表](計2件)

(1) Makoto Yamashita, Mituhiro Fukuda, Takashi Nakagaki, "Dual Approach Based on Spectral Projection Gradient Method for log-det SDP with ell_1 Norm," SIAM Optimization 2014, Town and Country Resort and Convention Center, (San Diego,

California, USA), 2014年5月19日.

(2) Makoto Yamashita, I-Lin Wang "An approach based on shortest path and connectivity consistency for sensor network localization problems," ISMP (International Symposium on Mathematical Programming) 2012, TU Berlin, (Berlin, Germany), 2012年8月22日.

[その他]
ホームページ等
実装したソフトウェアの公開用ページ
「Download page for boxSDP-matlab」
<http://www.is.titech.ac.jp/~makoto/software/boxSDP/>

6. 研究組織

(1)研究代表者
山下 真 (YAMASHITA MAKOTO)
東京工業大学・大学院情報理工学研究科・
准教授
研究者番号：20386824