

非線形半正定値計画問題に対する 主双対信頼領域内点法の大域的収束性

東京理科大学・理学部応用数学科 矢部博 (Hiroshi Yabe)
Department of Applied Mathematics, Tokyo University of Science
株式会社 NTT データ数理システム 山下浩 (Hiroshi Yamashita)
NTT DATA Mathematical Systems Inc.
株式会社 NTT データ数理システム 原田耕平 (Kouhei Harada)
NTT DATA Mathematical Systems Inc.

1 はじめに

本論文では次の非線形半正定値計画 (SDP) 問題を扱う。

$$\begin{aligned} & \text{minimize} && f(x), && x \in \mathbb{R}^n, \\ & \text{subject to} && g(x) = 0, && X(x) \succeq 0 \end{aligned} \quad (1)$$

ただし, $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $X: \mathbb{R}^n \rightarrow \mathbb{S}^p$ は十分滑らかであるとし, \mathbb{S}^p は p 次の実対称行列の集合とする. ここで, $X(x) \succeq 0$, $X(x) \succ 0$ はそれぞれ行列 $X(x)$ が半正定値, 正定値であることを意味する. SDP 問題は, 制御問題, 固有値問題, 金融問題などいろいろな分野で発生するもので, 応用範囲が広い. ([2, 3, 4, 10, 14, 15]). そうした理由から, 近年, 非線形 SDP 問題を解くための数値解法の研究が注目されている. 数値解法には拡張ラグランジュ法 [2, 8], 逐次線形 SDP 法 [3, 1, 4], 内点法 [18, 16, 7] などがあげられるが, 本論文では主双対内点法に焦点をあてる. 非線形 SDP 問題を解くための数値解法全般については, Yamashita and Yabe [17] のサーベイ論文を参照されたい.

主双対内点法の研究については, Yamashita, Yabe and Harada [18] が直線探索法の枠組みで数値解法を提案し大域的収束性を示している. また, Yamashita and Yabe [16] は局所的収束性について解析している. 一方, 原田・山下・矢部 [5] は直線探索とは別のアプローチとして, 信頼領域法を利用した非線形 SDP 問題に対する主双対内点法を提案し, 数値実験結果を報告している. 本論文では, 原田らの研究に基づいて, 主双対信頼領域内点法を扱いその大域的収束性を示す. さらにメリット関数の 2 次近似モデルについても議論し, いくつかの近似関数を提案する.

2 主双対信頼領域内点法の外部反復

本節では, 最適性の条件を述べるとともに記号の説明をする. \mathbb{S}^p の行列 X と Z の内積 $\langle X, Z \rangle$ を $\langle X, Z \rangle = \text{tr}(XZ)$ で定義する. ただし, $\text{tr}(M)$ は行列 M のトレースである. 問

題(1)のラグランジュ関数を

$$L(w) = f(x) - y^T g(x) - \langle X(x), Z \rangle$$

で定義する。ただし、 $w = (x, y, Z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{S}^p$ であり、 $y \in \mathbb{R}^m$ と $Z \in \mathbb{S}^p$ はそれぞれ等式制約、半正定値制約に対するラグランジュ乗数である。このとき、ラグランジュ関数の勾配ベクトルは

$$\nabla_x L(w) = \nabla f(x) - A_0(x)^T y - \mathcal{A}^*(x)Z$$

で与えられる。ここで、

$$A_0(x) = \begin{pmatrix} \nabla g_1(x)^T \\ \vdots \\ \nabla g_m(x)^T \end{pmatrix} \in \mathbb{R}^{m \times n}$$

であり、 $\mathcal{A}^*(x)$ は $\mathcal{A}(x) : \mathcal{A}(x)v = \sum_{i=1}^n v_i A_i(x)$ ($v \in \mathbb{R}^n$) の随伴作用素で

$$\mathcal{A}^*(x)Z = \begin{pmatrix} \langle A_1(x), Z \rangle \\ \vdots \\ \langle A_n(x), Z \rangle \end{pmatrix}$$

で定義される。また、 $A_i(x) (i = 1, \dots, n)$ は

$$A_i(x) = \frac{\partial X}{\partial x_i}$$

である。このとき、問題(1)のKarush-Kuhn-Tucker(KKT)条件は

$$r_0(w) \equiv \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ X(x)Z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$X(x) \succeq 0, \quad Z \succeq 0$$

で与えられる。以下では、 $X(x) \succ 0$ 、 $Z \succ 0$ を満たす点 $w = (x, y, Z)$ を内点と呼ぶ。主双対内点法では、バリアパラメータ $\mu > 0$ を導入して、相補性条件 $X(x)Z = 0$ を $X(x)Z = \mu I$ で置き換えて、次のバリアKKT(BKKT)条件を扱う。

$$r(w, \mu) \equiv \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ X(x)Z - \mu I \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (2)$$

$$X(x) \succ 0, \quad Z \succ 0$$

また、以下ではノルム $\|r(w, \mu)\|$ を

$$\|r(w, \mu)\| = \sqrt{\left\| \begin{pmatrix} \nabla_x L(w) \\ g(x) \end{pmatrix} \right\|^2 + \|X(x)Z - \mu I\|_F^2}$$

で定義し、 $\|\cdot\|$ をベクトルの l_2 ノルム、 $\|\cdot\|_F$ を行列のフロベニウスノルムとする。

Yamashita, Yabe and Harada [18] に倣って、次のような主双対内点法の外部反復を考える。

Algorithm SDPIP

Step 0. $\varepsilon > 0, M_c > 0$ を設定し, $k = 0$ とおく. $\mu_k \downarrow 0$ を満たす正の数数列 $\{\mu_k\}$ を設定する (反復の進行にともなって逐次決めてもよい).

Step 1. 次式を満たす内点 w_{k+1} を求める. これを近似 BKKT 点と呼ぶ.

$$\|r(w_{k+1}, \mu_k)\| \leq M_c \mu_k$$

Step 2. もし $\|r_0(w_{k+1})\| \leq \varepsilon$ ならば反復を終了する.

Step 3. $k := k + 1$ と更新して, Step 1 へ戻る. □

外部反復の収束性については, Yamashita, Yabe and Harada [18] で収束定理が示されている. 文献 [18] では, Step 1 で近似 BKKT 点 w_{k+1} を求めるための内部反復として直線探索法に基づいた主双対内点法が提案された. 一方, 原田・山下・矢部 [5] は Step 1 の内部反復として, 信頼領域法に基づいた主双対内点法を提案している. 本論文では原田らに倣って主双対信頼領域内点法を扱う.

以下では, 次節以降で用いる記号について簡単に説明する. 行列の積 $X(x)Z$ に対応して次の積を定義する.

$$X(x) \circ Z = \frac{X(x)Z + ZX(x)}{2}$$

$X(x) \succeq 0, Z \succeq 0, \mu \geq 0$ に対して, $X(x) \circ Z = \mu I$ と $X(x)Z = \mu I$ が同値であることが知られている. 行列 $U \in \mathbb{S}^p$ と正則行列 $P \in \mathbb{R}^{p \times p}, Q \in \mathbb{R}^{p \times p}$ に対して作用素

$$(P \circledast Q)U = \frac{1}{2}(PUQ^T + QUP^T)$$

と対称化クロネッカー積

$$(P \otimes_S Q)\text{svec}(U) = \text{svec}((P \circledast Q)U)$$

を定義する. ただし, svec は次式で定義される.

$$\text{svec}(U) = (U_{11}, \sqrt{2}U_{21}, \dots, \sqrt{2}U_{p1}, U_{22}, \sqrt{2}U_{32}, \dots, \sqrt{2}U_{p2}, U_{33}, \dots, U_{pp})^T \in \mathbb{R}^{p(p+1)/2}$$

また, svec の逆作用素として smat を定義すれば

$$(P \circledast Q)U = \text{smat}((P \otimes_S Q)\text{svec}(U))$$

となる. さらに,

$$(P \circledast Q)^{-1}U = \text{smat}((P \otimes_S Q)^{-1}\text{svec}(U))$$

を定義する. また, 行列 $M \in \mathbb{S}^p$ の最小固有値と最大固有値をそれぞれ $\lambda_{\min}(M), \lambda_{\max}(M)$ で表わす.

3 主双対信頼領域内点法

本節では、非線形最適化問題に対して Yamashita, Yabe and Tanabe [19] が提案した信頼領域法に基づいた主双対内点法の考えに倣って、非線形 SDP 問題に対する主双対信頼領域内点法の内部反復法を提案し、与えられたバリアパラメータ $\mu > 0$ に対して近似 BKKT 点を求めることを考える。以下では混乱が生じない場合には $X(x)$ を単に X と書くことにする。また、常に $X \succ 0$, $Z \succ 0$ が成り立っているものと仮定する。まず、主双対変数 (X, Z) に対する以下のようなスケーリングを考える。

$$\tilde{X} = TXT^T, \quad \tilde{Z} = T^{-T}ZT^{-1}$$

ただし、 $T \in \mathbb{R}^{p \times p}$ は正則なスケーリング行列である。ここで、 $XZ = \mu I$ を $\tilde{X} \circ \tilde{Z} = \mu I$ で置き換えれば、式 (2) の代わりに

$$\tilde{r}_S(w, \mu) \equiv \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ \tilde{X} \circ \tilde{Z} - \mu I \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (3)$$

を得る。

3.1 ニュートン方向と最急降下方向

方程式 (3) にニュートン法を適用することを考える。そのために、ニュートン方向を $\Delta w_N = (\Delta x_N, \Delta y_N, \Delta Z_N) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{S}^p$ で表わし、 $\Delta X_N = \sum_{i=1}^n (\Delta x_N)_i A_i(x) \in \mathbb{S}^p$ を定義する。また、 ΔX_N と ΔZ_N をスケーリングした行列を

$$\Delta \tilde{X}_N = T \Delta X_N T^T, \quad \Delta \tilde{Z}_N = T^{-T} \Delta Z_N T^{-1}$$

とおく。このとき、文献 [18] と同様にニュートン方程式は

$$\begin{aligned} G \Delta x_N - A_0(x)^T \Delta y_N - \mathcal{A}^*(x) \Delta Z_N &= -\nabla_x L(x, y, Z) \\ A_0(x) \Delta x_N &= -g(x) \\ \Delta \tilde{X}_N \tilde{Z} + \tilde{Z} \Delta \tilde{X}_N + \tilde{X} \Delta \tilde{Z}_N + \Delta \tilde{Z}_N \tilde{X} &= 2\mu I - \tilde{X} \tilde{Z} - \tilde{Z} \tilde{X} \end{aligned}$$

で与えられる。ただし、 G はラグランジュ関数 $L(w)$ のヘッセ行列、もしくはその近似行列であり、不定値行列でもかまわない。このとき、文献 [18] の定理 2 より次式が得られる。

$$\begin{pmatrix} G + H & -A_0(x)^T \\ -A_0(x) & 0 \end{pmatrix} \begin{pmatrix} \Delta x_N \\ \Delta y_N \end{pmatrix} = - \begin{pmatrix} \nabla f(x) - A_0(x)^T y - \mu \mathcal{A}^*(x) X^{-1} \\ -g(x) \end{pmatrix}$$

$$\Delta Z_N = \mu X^{-1} - Z - (T^T \circ T^T)(\tilde{X} \circ I)^{-1}(\tilde{Z} \circ I)(T \circ T) \Delta X_N \quad (4)$$

ただし、行列 $H \in \mathbb{R}^{n \times n}$ は

$$H_{ij} = \left\langle \tilde{A}_i(x), (\tilde{X} \circ I)^{-1}(\tilde{Z} \circ I) \tilde{A}_j(x) \right\rangle, \quad \tilde{A}_i(x) = T A_i(x) T^T \quad (5)$$

で定義される。もし \tilde{X} と \tilde{Z} が正定値対称で $\tilde{X} \tilde{Z} = \tilde{Z} \tilde{X}$ を満たし、かつ、行列 $A_i(x)$, $i = 1, \dots, n$ が線形独立ならば、行列 H は正定値対称行列になる。([18] の定理 3 を参照)

ニュートン方向に対して、最急降下方向 $\Delta w_{SD} = (\Delta x_{SD}, \Delta y_{SD}, \Delta Z_{SD})^T \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{S}^p$ を次式で定義する。

$$\begin{pmatrix} D + H & -A_0(x)^T \\ -A_0(x) & 0 \end{pmatrix} \begin{pmatrix} \Delta x_{SD} \\ \Delta y_{SD} \end{pmatrix} = - \begin{pmatrix} \nabla f(x) - A_0(x)^T y - \mu A^*(x) X^{-1} \\ -g(x) \end{pmatrix}$$

$$\Delta Z_{SD} = \mu X^{-1} - Z - (T^T \odot T^T)(\tilde{X} \odot I)^{-1}(\tilde{Z} \odot I)(T \odot T)\Delta X_{SD} \quad (6)$$

ただし、行列 D は正定値対称行列であり、通常は対角行列（例えば単位行列）が選ばれる。また、 $\Delta X_{SD} = \sum_{i=1}^n (\Delta x_{SD})_i A_i(x) \in \mathbb{S}^p$ と定義する。もし $D + H$ が正定値対称で行列 $A_0(x)$ がフルランクならば、最急降下方向 Δw_{SD} は一意に定まる。このとき、以下のことが成り立つ。これは有限回の反復で内部反復が終了する場合に対応している。

- (i) もし $\Delta w_N = 0$ ならば、 w は BKKT 点になる。
- (ii) もし $(\Delta x_N, \Delta Z_N) = 0$ ならば、 $(x, y + \Delta y_N, Z)$ は BKKT 点になる。
- (iii) もし $\Delta x_N = 0$ ならば、 $(x, y + \Delta y_N, Z + \Delta Z_N)$ は BKKT 点になる。

本節の最後にスケーリング行列について触れておく。スケーリング行列 T は $\tilde{X}\tilde{Z} = \tilde{Z}\tilde{X}$ が成り立つように選ばれるものであるが、次の2種類が知られている。

(i) HRVW/KSH/M 方向

$T = X^{-1/2}$ のとき $\tilde{X} = I$, $\tilde{Z} = X^{1/2} Z X^{1/2}$ となり、(5), (4), (6) は

$$\begin{aligned} H_{ij} &= \text{tr}(A_i(x) X^{-1} A_j(x) Z), \\ \Delta Z_N &= \mu X^{-1} - Z - \frac{1}{2}(X^{-1} \Delta X_N Z + Z \Delta X_N X^{-1}) \\ \Delta Z_{SD} &= \mu X^{-1} - Z - \frac{1}{2}(X^{-1} \Delta X_{SD} Z + Z \Delta X_{SD} X^{-1}) \end{aligned} \quad (7)$$

で与えられる。これは線形 SDP 問題における HRVM/KSH/M 方向に対応している [6, 9, 11].

(ii) NT 方向

$T = W^{-1/2}$ (ただし、 $W = X^{1/2}(X^{1/2} Z X^{1/2})^{-1/2} X^{1/2}$) のとき、 $\tilde{X} = W^{-1/2} X W^{-1/2} = W^{1/2} Z W^{1/2} = \tilde{Z}$ となり、(5), (4), (6) は

$$\begin{aligned} H_{ij} &= \text{tr}\{A_i(x) W^{-1} A_j(x) W^{-1}\} \\ \Delta Z_N &= \mu X^{-1} - Z - W^{-1} \Delta X_N W^{-1} \\ \Delta Z_{SD} &= \mu X^{-1} - Z - W^{-1} \Delta X_{SD} W^{-1} \end{aligned}$$

で与えられる。これは線形 SDP 問題における NT 方向に対応している [12, 13].

3.2 主双対メリット関数

本節では、信頼領域法で使われるメリット関数を提案する。基本的には、直線探索法の枠組みで Yamashita, Yabe and Harada [18] が扱った関数と同じであるが、後で述べるように2次近似モデルの導出が今回新しい結果である。以後、スケーリング行列 T は $\tilde{X}\tilde{Z} = \tilde{Z}\tilde{X}$ が成り立つように選ばれるものとし、簡単のために次の記号を導入する。

$$u = (x, Z) \in \mathbb{R}^n \times \mathbb{S}^p, \quad \Delta u = (\Delta x, \Delta Z) \in \mathbb{R}^n \times \mathbb{S}^p$$

$$U(u) = \begin{pmatrix} X(x) & 0 \\ 0 & Z \end{pmatrix}, \quad \Delta U = \begin{pmatrix} \Delta X & 0 \\ 0 & \Delta Z \end{pmatrix}$$

ただし, $\Delta X = \sum_{i=1}^n (\Delta x)_i A_i(x)$ である. ここで, Δu は $\Delta u_N = (\Delta x_N, \Delta Z_N)$ または $\Delta u_{SD} = (\Delta x_{SD}, \Delta Z_{SD})$ を意味し, そのノルム $\|\Delta u\|$ は

$$\|\Delta u\| = \sqrt{\|\Delta x\|^2 + \|\Delta Z\|_F^2}$$

で定義される.

3.4 節で述べる内部反復のアルゴリズムの大域的収束性を実現するために次の主双対メリット関数を用いる (ν は正の実数).

$$F(u) = F_{BP}(x) + \nu F_{PD}(u) \quad (8)$$

ただし, $F_{BP}(x)$ は主バリアペナルティ関数

$$F_{BP}(x) = f(x) - \mu \log(\det X(x)) + \rho \|g(x)\|_1$$

であり (ρ はペナルティパラメータ), $F_{PD}(u)$ は主双対バリア関数

$$F_{PD}(u) = \langle X(x), Z \rangle - \mu \log(\det X(x) \det Z)$$

である.

これらの関数の 1 次近似は次のように与えられる.

$$\begin{aligned} F_l(u; \Delta u) &= F(u) + \Delta F_l(u; \Delta u) \\ F_{BP_l}(x; \Delta x) &= F_{BP}(x) + \Delta F_{BP_l}(x; \Delta x) \\ F_{PD_l}(u; \Delta u) &= F_{PD}(u) + \Delta F_{PD_l}(u; \Delta u) \end{aligned}$$

ここで, $\Delta F_l(u; \Delta u)$, $\Delta F_{BP_l}(x; \Delta x)$, $\Delta F_{PD_l}(u; \Delta u)$ は方向微分係数に相当し次式で与えられる.

$$\Delta F_l(u; \Delta u) = \Delta F_{BP_l}(x; \Delta x) + \nu \Delta F_{PD_l}(u; \Delta u) \quad (9)$$

$$\Delta F_{BP_l}(x; \Delta x) = \nabla f(x)^T \Delta x - \mu \operatorname{tr}(X^{-1} \Delta X) + \rho (\|g(x) + A_0(x) \Delta x\|_1 - \|g(x)\|_1) \quad (10)$$

$$\Delta F_{PD_l}(u; \Delta u) = \operatorname{tr}(\Delta X Z + X \Delta Z - \mu X^{-1} \Delta X - \mu Z^{-1} \Delta Z)$$

特に, $g(x) + A_0(x) \Delta x = 0$ が満たされる場合には, (10) は

$$\Delta F_{BP_l}(x; \Delta x) = \nabla f(x)^T \Delta x - \mu \operatorname{tr}(X^{-1} \Delta X) - \rho \|g(x)\|_1$$

となる. 文献 [18] の補題 2, 3 と同様に最急降下方向について次の評価式が得られる.

$$\begin{aligned} \Delta F_{BP_l}(x; \Delta x_{SD}) &\leq -\Delta x_{SD}^T (D + H) \Delta x_{SD} - (\rho - \|y + \Delta y_{SD}\|_\infty) \|g(x)\|_1 \\ \Delta F_{PD_l}(u; \Delta u_{SD}) &\leq 0 \\ \Delta F_l(u; \Delta u_{SD}) &= \Delta F_{BP_l}(x; \Delta x_{SD}) + \nu \Delta F_{PD_l}(u; \Delta u_{SD}) \\ &\leq -\Delta x_{SD}^T (D + H) \Delta x_{SD} - (\rho - \|y + \Delta y_{SD}\|_\infty) \|g(x)\|_1 \end{aligned} \quad (11)$$

ただし, (11) において等号が成り立つことと $XZ = \mu I$ が成り立つことは同値である.

上記の結果を用いれば, 次の定理が得られる. これらの一部は [18] の定理 5 に対応している.

定理 1 行列 $D + H$ は正定値対称とし、 $\rho > \|y + \Delta y_{SD}\|_\infty$ が成り立つと仮定する。このとき次のことが成り立つ。

- (i) 探索方向 Δx_{SD} と Δu_{SD} は主バリアペナルティ関数および主双対メリット関数に対して降下方向になる。また、もし $\Delta x_{SD} \neq 0$ ならば、 $\Delta F_{BPq}(x; \Delta x_{SD}) < 0$ および $\Delta F_l(u; \Delta u_{SD}) < 0$ が成り立つ。
- (ii) $\Delta F_{BPq}(x; \Delta x_{SD}) = 0$ が成り立つことと $(x, y + \Delta y_{SD}, Z + \Delta Z_{SD})$ が $BKKT$ 点になることは同値である。
- (iii) $\Delta F_l(u; \Delta u_{SD}) = 0$ が成り立つことと $(x, y + \Delta y_{SD}, Z)$ が $BKKT$ 点になることは同値である。

3.3 メリット関数の 2 次近似モデル

本節では、メリット関数 (8) の 2 次近似

$$F_q(u; \Delta u) = F(u) + \Delta F_l(u; \Delta u) + \frac{1}{2}Q(u; \Delta u)$$

について考える。ただし、 $\Delta F_l(u; \Delta u)$ は (9) で定義した量であり、 $Q(u; \Delta u) = Q_{BP}(x; \Delta x) + \nu Q_{PD}(u; \Delta u)$ は任意の実数 α に対して $Q(u; \alpha \Delta u) = \alpha^2 Q(u; \Delta u)$ が成り立つように選ばれた 2 次関数である。また、2 次近似の Δu 方向の変化量を

$$\begin{aligned} \Delta F_q(u; \Delta u) &\equiv F_q(u; \Delta u) - F(u) = \Delta F_l(u; \Delta u) + \frac{1}{2}Q(u; \Delta u) \\ \Delta F(u; \Delta u) &\equiv F(u + \Delta u) - F(u) \end{aligned}$$

で定義する。

後述する収束定理 (定理 2) では上記の条件を満たす 2 次近似モデルであれば何でもよいが、ここでは 2 次近似モデルの具体的な形を導出することにする。そのために F_{BPq} と F_{PDq} の 2 次近似を次のように表わす。

$$F_{BPq}(x; \Delta x) = F_{BP}(x) + \Delta F_{BPq}(x; \Delta x)$$

$$F_{PDq}(u; \Delta u) = F_{PD}(u) + \Delta F_{PDq}(u; \Delta u)$$

ここで $\Delta F_{BPq}(x; \Delta x)$ は

$$\begin{aligned} \Delta F_{BPq}(x; \Delta x) &= \nabla f(x)^T \Delta x + \frac{1}{2} \Delta x^T \nabla^2 f(x) \Delta x - \mu \text{tr}(X^{-1} \Delta X) \\ &\quad - \frac{\mu}{2} \langle X^{-1}, \Delta x^T \nabla^2 X(x) \Delta x \rangle + \frac{\mu}{2} \langle X^{-1} \Delta X X^{-1}, \Delta X \rangle \\ &\quad + \rho \left(\|g(x) + A_0(x) \Delta x + \frac{1}{2} \Delta x^T \nabla^2 g(x) \Delta x\|_1 - \|g(x)\|_1 \right) \end{aligned}$$

で定義される。ただし、 $\Delta x^T \nabla^2 X(x) \Delta x$ はその (i, j) 要素が $\Delta x^T \nabla^2 X_{ij}(x) \Delta x$ であるような行列であり、 $\Delta x^T \nabla^2 g(x) \Delta x$ はその第 i 成分 $\Delta x^T \nabla^2 g_i(x) \Delta x$ であるようなベクトルであ

る。もし Δx が $g(x) + A_0(x)\Delta x = 0$ を満たすならば、上記の式は

$$\begin{aligned}\Delta F_{BPq}(x; \Delta x) &= \Delta F_{BPl}(x; \Delta x) + \frac{1}{2}\Delta x^T \nabla^2 f(x)\Delta x \\ &\quad - \frac{\mu}{2}\langle X^{-1}, \Delta x^T \nabla^2 X(x)\Delta x \rangle + \frac{\mu}{2}\langle X^{-1}\Delta X X^{-1}, \Delta X \rangle \\ &\quad + \frac{\rho}{2}\|\Delta x^T \nabla^2 g(x)\Delta x\|_1\end{aligned}\quad (12)$$

となる。この変化量を

$$\Delta F_{BPq}(x; \Delta x) \approx \Delta F_{BPl}(x; \Delta x) + \frac{1}{2}\Delta x^T \nabla^2 L(w)\Delta x + \frac{1}{2}\langle \Delta X, (\mu X^{-1} \odot X^{-1})\Delta X \rangle \quad (13)$$

で近似することが考えられる。また、 $Z = \mu X^{-1}$ とおくことによって $(\mu X^{-1} \odot X^{-1})\Delta X = (Z \odot X^{-1})\Delta X$ となるので、

$$\langle \Delta X, (\mu X^{-1} \odot X^{-1})\Delta X \rangle = \Delta x^T H \Delta x$$

が得られる。ただし、行列 H の (i, j) 要素は (7) で与えられる。したがって、 $\Delta F_{BPq}(x; \Delta x)$ は次式で表わされる。

$$\Delta F_{BPq}(x; \Delta x) \approx \Delta F_{BPl}(x; \Delta x) + \frac{1}{2}\Delta x^T (\nabla^2 L(w) + H)\Delta x \quad (14)$$

次に $\Delta F_{PDq}(u; \Delta u)$ について考える。これは

$$\begin{aligned}\Delta F_{PDq}(u; \Delta u) &= \Delta F_{PDI}(u; \Delta u) \\ &\quad + \langle \Delta X, \Delta Z \rangle + \frac{1}{2}(\mu \langle X^{-1}\Delta X X^{-1}, \Delta X \rangle + \mu \langle Z^{-1}\Delta Z Z^{-1}, \Delta Z \rangle) \\ &\quad + \frac{1}{2}(\langle \Delta x^T \nabla^2 X(x)\Delta x, Z \rangle - \mu \langle X^{-1}, \Delta x^T \nabla^2 X(x)\Delta x \rangle) \\ &= \Delta F_{PDI}(u; \Delta u) + \langle \Delta X, \Delta Z \rangle \\ &\quad + \frac{1}{2}\langle \Delta X, H_X(X)\Delta X \rangle + \frac{1}{2}\langle \Delta Z, H_Z(Z)\Delta Z \rangle \\ &\quad + \frac{1}{2}(\langle \Delta x^T \nabla^2 X(x)\Delta x, Z \rangle - \mu \langle X^{-1}, \Delta x^T \nabla^2 X(x)\Delta x \rangle)\end{aligned}\quad (15)$$

で与えられる。ただし、 $H_X(X)$ と $H_Z(Z)$ はそれぞれ $\mu X^{-1} \odot X^{-1}$ 、 $\mu Z^{-1} \odot Z^{-1}$ を意味する。ここで、 $2\langle \Delta X, \Delta Z \rangle + \langle \Delta X, H_X(X)\Delta X \rangle + \langle \Delta Z, H_Z(Z)\Delta Z \rangle$ を

$$\left\langle \Delta U, \begin{pmatrix} H_X(X) & \mathcal{I} \\ \mathcal{I} & H_Z(Z) \end{pmatrix} \Delta U \right\rangle$$

と表わせば、

$$\begin{aligned}\Delta F_{PDq}(u; \Delta u) &= \Delta F_{PDI}(u; \Delta u) + \frac{1}{2}\left\langle \Delta U, \begin{pmatrix} H_X(X) & \mathcal{I} \\ \mathcal{I} & H_Z(Z) \end{pmatrix} \Delta U \right\rangle \\ &\quad + \frac{1}{2}(\langle \Delta x^T \nabla^2 X(x)\Delta x, Z \rangle - \mu \langle X^{-1}, \Delta x^T \nabla^2 X(x)\Delta x \rangle)\end{aligned}$$

となる。なお、次式が成り立つことに注意しておく。

$$\begin{aligned}\langle \Delta X, H_X(X)\Delta X \rangle &= \mu \langle X^{-1}\Delta X X^{-1}, \Delta X \rangle = \mu \|X^{-1/2}\Delta X X^{-1/2}\|_F^2 \\ \langle \Delta Z, H_Z(Z)\Delta Z \rangle &= \mu \text{tr}(Z^{-1}\Delta Z Z^{-1}\Delta Z) = \mu \|Z^{-1/2}\Delta Z Z^{-1/2}\|_F^2\end{aligned}$$

さらに $Z = \mu X^{-1}$ とおけば,

$$\begin{aligned} \Delta F_{PDq}(u; \Delta u) &\approx \Delta F_{PDI}(u; \Delta u) + \langle \Delta X, \Delta Z \rangle \\ &\quad + \frac{1}{2} \langle \Delta X, H_X(X) \Delta X \rangle + \frac{1}{2} \langle \Delta Z, H_Z(Z) \Delta Z \rangle \\ &= \Delta F_{PDI}(u; \Delta u) + \frac{1}{2} \left\langle \Delta U, \begin{pmatrix} H_X(X) & \mathcal{I} \\ \mathcal{I} & H_Z(Z) \end{pmatrix} \Delta U \right\rangle \end{aligned} \quad (16)$$

が得られる。

以上の準備のもと, 2次近似 $Q(u; \Delta u) = Q_{BP}(x; \Delta x) + \nu Q_{PD}(u; \Delta u)$ の具体的な形を導出する。まず, (12), (13), (14) に基づいて

$$\begin{aligned} Q_{BP1} &= \Delta x^T \nabla^2 f(x) \Delta x - \mu \langle X^{-1}, \Delta x^T \nabla^2 X(x) \Delta x \rangle \\ &\quad + \mu \langle X^{-1} \Delta X X^{-1}, \Delta X \rangle + \rho \|\Delta x^T \nabla^2 g(x) \Delta x\|_1 \\ Q_{BP2} &= \Delta x^T \nabla^2 L(w) \Delta x + \langle \Delta X, H_X(X) \Delta X \rangle \\ Q_{BP3} &= \Delta x^T (\nabla^2 L(w) + H) \Delta x \end{aligned}$$

を定義する。同様に, (15) と (16) に基づいて

$$\begin{aligned} Q_{PD1} &= 2 \langle \Delta X, \Delta Z \rangle + \mu \langle X^{-1} \Delta X X^{-1}, \Delta X \rangle + \mu \langle Z^{-1} \Delta Z Z^{-1}, \Delta Z \rangle \\ &\quad + \langle \Delta x^T \nabla^2 X(x) \Delta x, Z - \mu X^{-1} \rangle \\ Q_{PD2} &= 2 \langle \Delta X, \Delta Z \rangle + \langle \Delta X, H_X(X) \Delta X \rangle + \langle \Delta Z, H_Z(Z) \Delta Z \rangle \end{aligned}$$

を定義する。このとき, 具体的な2次近似として次の4種類が考えられる。

- (i) $Q(u; \Delta u) = Q_{BP1}(x; \Delta x) + \nu Q_{PD1}(u; \Delta u)$
- (ii) $Q(u; \Delta u) = Q_{BP2}(x; \Delta x) + \nu Q_{PD2}(u; \Delta u)$
- (iii) $Q(u; \Delta u) = Q_{BP3}(x; \Delta x) + \nu Q_{PD2}(u; \Delta u)$
- (iv) $Q(u; \Delta u) = 0$

3.4 Algorithm SDPTR

前節までで探索方向とメリット関数について説明した。本節では, 主双対信頼領域内点法の内部反復にあたる Algorithm SDPTR を述べる。そのために, 信頼領域法で扱うコーシー点に関連して, 次のようなステップ幅 $\alpha_\delta^*(u, \Delta u)$ を定義する。

$$\alpha_\delta^*(u, \Delta u) = \arg \min \{ F_q(u; \alpha \Delta u) \mid \alpha \leq 1, \|\alpha \Delta u\| \leq \delta, \alpha \in [0, \gamma \bar{\alpha}(u, \Delta u)] \} \quad (17)$$

ただし, δ は信頼領域半径であり, $\bar{\alpha}(u, \Delta u)$ は次のように定義される。

- (i) もし $X(x)$ が線形関数ならば,

$$\bar{\alpha}(u, \Delta u) = \frac{1}{\max\{-\lambda_{\min}(U^{-1} \Delta U), 0\}}$$

とおく。ここで, 分母の値がゼロのときには $\bar{\alpha}(u, \Delta u)$ は ∞ であると解釈する。

(ii) さもなければ, 与えられた定数 $\tau \in (0, 1)$ に対して $\lambda_{\min}(U(u + \tau^i \Delta u)) \geq 0$ を満たす最小の非負整数 i を見つけて $\bar{\alpha}(u, \Delta u) = \tau^i$ とおく.

ステップ幅 $\alpha_{\delta}^*(u, \Delta u)$ は, 与えられた制約条件のもとで Δu に沿って 2 次関数 $F_q(u; \alpha \Delta u)$ を最小化するステップ幅で, 特に $\Delta u = \Delta u_{SD}$ に対して $u + \alpha_{\delta}^*(u, \Delta u_{SD}) \Delta u_{SD}$ はコーシー点と呼ばれている.

以上の準備のもと, 信頼領域法に基づいて BKKT 点を求めるための主双対信頼領域内点法の内部反復は次のようなアルゴリズムで与えられる.

Algorithm SDPTR

Step 0. 初期点 $w_0 = (x_0, y_0, Z_0)$ ($X_0 > 0, Z_0 > 0$) と初期信頼領域半径 $\delta_0 > 0$ を与える.

パラメータ $\mu, \rho, M > 1, \varepsilon, \gamma \in (0, 1)$ を与え, $k = 0$ とおく.

Step 1. もし $\|r(w_k, \mu)\| \leq \varepsilon$ ならば, 内部反復を終了する.

Step 2. ニュートン方向 Δu_{Nk} と最急降下方向 Δu_{SDk} を求める. もし次の条件

$$\|\Delta u_{Nk}\| \leq M \|\Delta u_{SDk}\| \quad (18)$$

が成り立たないならば, 行列 G_k を調整して (18) が成り立つようにする.

Step 3. 探索方向 Δu_{Nk} と Δu_{SDk} を用いて,

$$\bar{s}_k = (1 - \theta) \Delta u_{Nk} + \theta \Delta u_{SDk}$$

とおく. ここで, パラメータ $\theta \in [0, 1]$ は, $s_k = \alpha_{\delta_k}^*(u_k, \bar{s}_k) \bar{s}_k$ が次の式を満たすように選ばれる.

$$\Delta F_q(u_k; s_k) \leq \frac{1}{2} \Delta F_q(u_k; \alpha_{\delta_k}^*(u_k, \Delta u_{SDk}) \Delta u_{SDk})$$

Step 4. 次式のように更新して, 新しい信頼領域半径 δ_{k+1} を求める.

$$\delta_{k+1} = \begin{cases} \frac{1}{2} \delta_k & \text{if } \Delta F(u_k; s_k) > \frac{1}{4} \Delta F_q(u_k; s_k) \\ 2\delta_k & \text{if } \Delta F(u_k; s_k) \leq \frac{3}{4} \Delta F_q(u_k; s_k) \\ \delta_k & \text{otherwise.} \end{cases}$$

Step 5. もし $\Delta F(u_k; s_k) \leq 0$ ならば, $u_{k+1} = u_k + s_k$ および $y_{k+1} = y_k + \Delta y_{Nk}$ とおく. さもなければ, $w_{k+1} = w_k$ とする.

Step 6. $k := k + 1$ として Step 1 へ戻る. □

4 Algorithm SDPTR の大域的収束性

Algorithm SDPTR の大域的収束性について議論する. そのために以下の仮定を設ける.

仮定

(PD1) 関数 $f, g_i, i = 1, \dots, m, X$ は 2 回連続微分可能である。

(PD2) Algorithm SDPTR によって生成される点列 $\{u_k\}$ はあるコンパクト集合 $\Omega \subset \mathbb{R}^n \times \mathbb{S}^p$ に含まれる。

(PD3) 集合 Ω において、行列 $A_0(x)$ はフルランクであり、行列 $A_1(x), \dots, A_n(x)$ は線形独立である。

(PD4) 行列 D_k は一様正定値かつ一様有界である。また、行列 G_k は一様有界である。

(PD5) スケーリング行列 T_k は、 \tilde{X}_k と \tilde{Z}_k が可換であり、かつ、行列の列 $\{T_k\}$ と $\{T_k^{-1}\}$ が一様有界であるように選ばれる。

(PD6) ペナルティパラメータ ρ は、すべての k に対して $\rho > \|y_k + \Delta y_{SDk}\|_\infty$ が成り立つように十分大きく選ばれる。□

前節で述べたように、もし $\Delta u_{SDk} = 0$ ならば $w_{k+1} = (x_k, y_k + \Delta y_{NK}, Z_k)$ は BKKT 点になるので有限回の反復で終了する。以下では無限点列が生成される場合を考えるので、すべての k に対して $\Delta u_{SDk} \neq 0$ が成り立つと仮定する。

コーシー点に関連して次の補題が得られる。

補題 1 $u \in \mathbb{R}^n \times \mathbb{S}^p$ ($X(x) \succ 0, Z \succ 0$), $\Delta u = (\Delta x, \Delta Z) (\neq 0) \in \mathbb{R}^n \times \mathbb{S}^p$, $\delta > 0$ は任意に与えられているとする。このとき、もし $\Delta F_l(u; \Delta u) < 0$ かつ

$$g(x) + A_0(x)\Delta x = 0$$

が成り立つならば、(17) で定義されるステップ幅は

$$\alpha_\delta^*(u, \Delta u) = \min \left\{ 1, \frac{\delta}{\|\Delta u\|}, \gamma \bar{\alpha}(u, \Delta u), -\frac{\Delta F_l(u; \Delta u)}{\max\{Q(u, \Delta u), 0\}} \right\}$$

と表わされる。ただし、括弧内の分数で分母がゼロになる場合はその値は無限大であると解釈する。さらに 2 次近似モデルの減少量について以下の評価式が成り立つ。

$$\Delta F_q(u; \alpha_\delta^*(u, \Delta u)\Delta u) \leq \frac{1}{2} \alpha_\delta^*(u, \Delta u) \Delta F_l(u; \Delta u)$$

次の補題は点列 $\{\Delta w_{SDk}\}$ の有界性を保証するもので、文献 [18] の補題 5(iv) に対応する。

補題 2 仮定 (PD1)–(PD6) が成り立つとし、点列 $\{w_k\}$ は Algorithm SDPTR によって生成される無限点列であるとする。このとき探索方向の列 $\{\Delta w_{SDk}\}$ は有界である。

以上の準備のもとで、Algorithm SDPTR の大域的収束性が次の定理によって示される。

定理 2 仮定 (PD1)–(PD6) が成り立つとし、点列 $\{w_k\}$ は Algorithm SDPTR によって生成される無限点列であるとする。このとき、BKKT 点となるような集積点が存在する。

5 おわりに

本論文では非線形半正定値計画問題を解くための主双対信頼領域内点法を提案し、その大域的収束性を示した。また主双対メリット関数に関連して、信頼領域法で使われる2次近似モデルについていくつかの近似モデルを導出した。今後は数値実験を行って、こうした2次近似モデルの違いによって計算効率がどのように影響を受けるのかを調べる必要がある。さらに、主双対信頼領域内点法の局所的収束性の解析は今後の課題である。

謝辞

本研究の一部はJSPS 科研費 JP17K00039 の助成を受けて行われている。

参考文献

- [1] Correa, R. and Ramirez, C.H.: A global algorithm for nonlinear semidefinite programming. *SIAM Journal on Optimization* **15**, 303-318 (2004).
- [2] Fares, B., Apkarian, P. and Noll, D.: An augmented Lagrangian method for a class of LMI-constrained problems in robust control theory. *International Journal of Control* **74**, 348-360 (2001).
- [3] Fares, B., Noll, D. and Apkarian, P.: Robust control via sequential semidefinite programming. *SIAM Journal on Control and Optimization* **40**, 1791-1820 (2002).
- [4] Freund, R.W., Jarre, F. and Vogelbusch, C.H.: Nonlinear semidefinite programming: sensitivity, convergence, and an application in passive reduced-order modeling. *Mathematical Programming* **109**, 581-611 (2007).
- [5] 原田耕平, 山下浩, 矢部博: 信頼領域法を利用した非線形半正定値計画問題に対する主双対内点法. *京都大学数理解析研究所講究録* **1879**, 125-133 (2014).
- [6] Helmberg, C., Rendl, F., Vanderbei, R.J. and Wolkowicz, H.: An interior-point method for semidefinite programming. *SIAM Journal on Optimization* **6**, 342-361 (1996).
- [7] Kato, A., Yabe, H. and Yamashita, H.: An interior point method with a primal-dual quadratic barrier penalty function for nonlinear semidefinite programming. *Journal of Computational and Applied Mathematics* **275**, 148-161 (2015).
- [8] Kóčvara, M. and Stingl, M.: PENNON : A code for convex nonlinear and semidefinite programming. *Optimization Methods and Software* **18**, 317-333 (2003).
- [9] Kojima, M., Shindoh, S. and Hara, S.: Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. *SIAM Journal on Optimization* **7**, 86-125 (1997).
- [10] Konno, H., Kawadai, N. and Wu, D.: Estimation of failure probability using semidefinite Logit model. *Computational Management Science* **1**, 59-73 (2003).

- [11] Monteiro, R.D.C.: Primal-dual path-following algorithms for semidefinite programming. *SIAM Journal on Optimization* **7**, 663-678 (1997).
- [12] Nesterov, Y.E. and Todd, M.J.: Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research* **22**, 1-42 (1997).
- [13] Nesterov, Y.E. and Todd, M.J.: Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization* **8**, 324-364 (1998).
- [14] Qi, H. and Sun, D.: A quadratically convergent Newton method for computing the nearest correlation matrix. *SIAM Journal on Matrix Analysis and Applications* **28**, 360-385 (2006).
- [15] Vandenberghe, L., Boyd, S. and Wu, S.-P.: Determinant maximization with linear matrix inequality constraints. *SIAM Journal on Matrix Analysis and Applications* **19**, 499-533 (1998).
- [16] Yamashita, H. and Yabe, H.: Local and superlinear convergence of a primal-dual interior point method for nonlinear semidefinite programming. *Mathematical Programming* **132**, 1-30 (2012).
- [17] Yamashita, H. and Yabe, H.: A survey of numerical methods for nonlinear semidefinite programming. *Journal of the Operations Research Society of Japan* **58**, 24-60 (2015).
- [18] Yamashita, H., Yabe, H. and Harada, K.: A primal-dual interior point method for nonlinear semidefinite programming. *Mathematical Programming*. **135**, 89-121 (2012).
- [19] Yamashita, H., Yabe, H. and Tanabe, T.: A globally and superlinearly convergent primal-dual interior point trust region method for large scale constrained optimization. *Mathematical Programming* **102**, 111-151 (2005).