

非線形SDPに対する主双対内点法の超一次収束性

山川 雄也, 山下 信雄

京都大学 大学院情報学研究科 数理工学専攻

Yamakawa Yuya¹, Yamashita Nobuo

Department of Applied Mathematics and Physics, Kyoto University

1 はじめに

本稿では, 次に定義する非線形半正定値計画問題 (Nonlinear Semidefinite Programming, 以下, 非線形 SDP) を扱う.

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

ここで, 関数 $f: \mathbf{R}^n \rightarrow \mathbf{R}$, $g: \mathbf{R}^n \rightarrow \mathbf{R}^m$, $X: \mathbf{R}^n \rightarrow \mathbf{S}^d$ はそれぞれ二回連続的微分可能な関数である. また, $\mathbf{S}^d$ は d 次実対称行列の集合を表し, $X(x) \succeq 0$ ($X(x) \succ 0$) は行列 $X(x)$ が半正定値 (正定値) であることを表す. 非線形 SDP は様々な数理計画問題を包括する広いクラスの問題であり, 線形計画問題, 二次錐計画問題, 線形半正定値計画問題 (線形 SDP), 非線形計画問題などは非線形 SDP へ定式化することが可能である.

非線形 SDP(1) の解法として, 主双対内点法がいくつか提案されている. それらの内点法には, 中心化 KKT 条件に基づくもの [4] と, シフト付き中心化 KKT 条件に基づくもの [2] があり, それぞれ大域的収束性が示されている. 一方, 中心化 KKT 条件に基づく内点法に対しては超一次収束性は示されているが [4], シフト付き中心化 KKT 条件に基づく主双対内点法の超一次収束性はまだ示されていない. そこで本稿では, 中心化 KKT 条件とシフト付き中心化 KKT 条件を含む近似 KKT 条件に基づいたアルゴリズムを提案し, いくつかの仮定の下でアルゴリズムの超一次収束性を示す.

以下では, 本稿の構成を記す. 2 章では非線形 SDP(1) の最適性条件, 近似 KKT 条件に基づいたアルゴリズムを紹介する. 3 章では, いくつかの仮定を与え, それらの仮定の下で提案アルゴリズムの超一次収束性を示す. 最後に, まとめについて記す.

2 非線形SDPに対する最適性条件と主双対内点法

非線形 SDP(1) の KKT 条件は次式で与えられる.

$$\begin{aligned} \nabla_x L(w) &= 0, \quad g(x) = 0, \quad X(x) \circ Z = 0, \\ X(x) &\succeq 0, \quad Z \succeq 0 \end{aligned} \tag{2}$$

¹e-mail: yamakawa@amp.i.kyoto-u.ac.jp

ただし,

$$L(w) \equiv f(x) - g(x)^T y - \text{svec}(X(x))^T \text{svec}(Z), \quad X(x) \circ Z \equiv \frac{X(x)Z + ZX(x)}{2}$$

$$\text{svec}(U) = [U_{11}, \sqrt{2}U_{21}, \dots, \sqrt{2}U_{d1}, U_{22}, \sqrt{2}U_{32}, \dots, \sqrt{2}U_{d2}, U_{33}, \dots, U_{dd}]^T, (U \in \mathbf{S}^d)$$

で定義する. ラグランジュ関数 L における, $y \in \mathbf{R}^m$, $Z \in \mathbf{S}^d$ はそれぞれ $g(x) = 0$, $X(x) \succeq 0$ に対するラグランジュ乗数である. また, $w = [x, y, \text{svec}(Z)] \in \mathbf{R}^l$, $l \equiv n + m + \frac{d(d+1)}{2}$ である.

ここで, パラメータ $\mu > 0$, $\kappa \in [0, \infty)$ を導入して, 以下の近似 KKT 条件を考える.

$$r_\kappa(w, \mu) \equiv \begin{bmatrix} \nabla_x L(w) \\ g(x) + \kappa \mu y \\ X(x) \circ Z - \mu I \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad X(x) \succ 0, \quad Z \succ 0 \quad (3)$$

$\kappa = 0$ のとき (3) は中心化 KKT 条件, $\kappa = 1$ のときはシフト付き中心化 KKT 条件となる. 以下では, $X(x) \succ 0$, $Z \succ 0$ を満たす点を内点と呼ぶ.

次に, KKT 点を求めるためのアルゴリズムについて紹介する. 提案アルゴリズムでは, (3) に対するニュートン方程式

$$\begin{bmatrix} \nabla_{xx}^2 L(w) & -J_g(x) & -A(x) \\ J_g(x) & \kappa \mu I & 0 \\ (Z \otimes_S I)A(x) & 0 & (X \otimes_S I) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \text{svec}(\Delta Z) \end{bmatrix} = \begin{bmatrix} -\nabla_x L(w) \\ -g(x) - \kappa \mu y \\ \text{svec}(\mu I - X \circ Z) \end{bmatrix} \quad (4)$$

を解き, ニュートン方向を求め点列を更新する. ただし, J_g は関数 g のヤコビ行列であり,

$$A(x) \equiv [\text{svec}(A_1(x)), \dots, \text{svec}(A_n(x))],$$

$$A_i(x) \equiv \frac{\partial}{\partial x_i} X(x), \quad i = 1, \dots, n$$

である. 以上を用いて, 近似 KKT 条件 (3) に基づいたアルゴリズムを以下に記述する.

Algorithm 1

Step 0. $\varepsilon > 0, \kappa \geq 0$ と $0 < \tau < 1$ をとり, 内点 $w_0 = [x_0, y_0, Z_0]$ を与え, $k = 0$ とする.

Step 1. もし $\|r_\kappa(w_k, 0)\| \leq \varepsilon$ を満たせば終了する.

Step 2. $\mu_k = \|r_\kappa(w_k, 0)\|^{1+\tau}$ とする. ニュートン方程式 (4) を解き Δw_k を求め $w_{k+1} = w_k + \Delta w_k$ と更新する.

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

このアルゴリズムでは大域的収束が保証されていない. 大域的収束するためには, [2] などのように, 適当なメリット関数を用いてステップ幅 α_k を定め, $w_{k+1} = w_k + \alpha_k \Delta w_k$ とすれば良い. ここでは局所的な収束のみを議論するため, そのような直線探索は省略している.

3 Algorithm 1 の超一次収束性

本節では, KKT 条件 (2) を満たす点 (KKT 点) を $w^* = [x^*, y^*, \text{svec}(Z^*)]$ とし, ニュートン方程式 (4) のヤコビ行列を $M(w, \mu)$ とする. このヤコビ行列は $M(w, \mu) = M_0(w) + \kappa\mu M_I$ と分解できる. ただし,

$$M_0(w) \equiv \begin{bmatrix} \nabla_{xx}^2 L(w) & -J_g(x) & -A(x) \\ J_g(x) & 0 & 0 \\ (Z \otimes_S I)A(x) & 0 & (X \otimes_S I) \end{bmatrix}, \quad M_I \equiv \begin{bmatrix} 0 & 0 & 0 \\ 0 & I_m & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

であり, I_m は $m \times m$ の単位行列である. これらを用いて, まず Algorithm 1 の超一次収束性に必要な仮定を以下に与える.

(仮定 1) $\nu_L > 0$ が存在して $M_0(w)$ は $\mathcal{V}_L \equiv \{w \in \mathbf{R}^l \mid \|w - w^*\| \leq \nu_L\}$ 上でリプシッツ連続である.

(仮定 2) 問題 (1) に対する最適性の二次の十分条件が x^* において成り立つ.

(仮定 3) 狭義相補性条件が x^* において成り立つ.

(仮定 4) 非退化条件が x^* において成り立つ.

これらの仮定の詳細については [4] を参照して欲しい.

まず, (仮定 1) より任意の $w_1, w_2 \in \mathcal{V}_L$ に対して, 次の不等式を満たす正の定数 L_M が存在する.

$$\|M_0(w_1) - M_0(w_2)\| \leq L_M \|w_1 - w_2\|$$

これと [1, 3] より得られる不等式を以下に与える.

Lemma 1 [1, 3] (仮定 1) が成り立つとする. このとき, 任意の $w_1, w_2 \in \mathcal{V}_L, \mu \geq 0$ に対して

$$\|r_\kappa(w_1, \mu) - r_\kappa(w_2, \mu) - M(w_2, \mu)(w_1 - w_2)\| \leq L_M \|w_1 - w_2\|^2$$

が成り立つ. また, ある正の数 L_r が存在して,

$$\|r_\kappa(w_1, \mu) - r_\kappa(w_2, \mu)\| \leq L_r \|w_1 - w_2\|$$

が成り立つ. □

続いて, (仮定 2)–(仮定 4) より $M_0(w^*)$ の正則性が示される.

Theorem 1 [4] (仮定 2)–(仮定 4) が成り立つとする. このとき, $M_0(w^*)$ は正則となる.

□

Theorem 1 と陰関数定理を用いれば、次の Lemma が成り立つ。

Lemma 2 [4] (仮定 2)–(仮定 4) が成り立つとする。このとき、ある区間 $[0, \gamma]$ で定義される微分可能な連続関数 $\bar{w}(\mu) = [\bar{x}(\mu), \bar{y}(\mu), \text{svec}(\bar{Z}(\mu))] \in \mathbf{R}^l$ が存在し、

$$\begin{aligned} \bar{w}(0) &= w^*, \quad r_\kappa(\bar{w}(\mu), \mu) = 0 \quad \text{for any } \mu \in [0, \gamma], \\ X(\bar{x}(\mu)) &\succ 0, \quad \bar{Z}(\mu) \succ 0 \quad \text{for any } \mu \in (0, \gamma] \end{aligned}$$

を満たす。

□

以下では、 $\{\bar{w}(\mu) | \mu \in [0, \gamma]\}$ を中心パスと呼ぶことにする。また、Theorem 1 より (仮定 2)–(仮定 4) の下であれば行列 $M_0(w^*)$ が正則であるから、十分小さなある $\varepsilon \in (0, 1)$ が存在し、以下を満たす行列 G は正則である。

$$\|G - M_0(w^*)\|_F < \varepsilon$$

さらに、行列 $M_0(w)$ の連続性から、 $\|w - w^*\| \leq \nu_M$ ならば

$$\|M_0(w) - M_0(w^*)\|_F \leq \frac{1}{3}\varepsilon$$

であるとする。この ν_M と ν_L を用いて w^* の近傍 $\mathcal{V}$ を以下のように定義する。

$$\mathcal{V} \equiv \{w \in \mathbf{R}^l \mid \|w - w^*\| \leq \nu \equiv \min\{\nu_M, \nu_L\}\}$$

これらを用いて、ニュートン方程式 (4) のヤコビ行列 $M(w, \mu)$ の正則性を保証する補題を与える。

Lemma 3 (仮定 2)–(仮定 4) が成り立つとする。このとき、任意の $w \in \mathcal{V}$, $\mu \in [0, s]$ に対して $M(w, \mu)$ は正則となる。ただし、 $s \equiv \frac{\varepsilon}{6(\kappa+1)\sqrt{m}}$ とする。 □

Lemma 3 と集合 $\mathcal{V}$, $[0, s]$ の有界性から

$$U_M \equiv \sup_{w \in \mathcal{V}, \mu \in [0, s]} \|M(w, \mu)^{-1}\|_F, \quad U_y \equiv \sup_{w \in \mathcal{V}} \|y\|^2$$

を定義すれば、これらは有界である。

次に、 $\mathcal{V}$ の部分集合を以下に定義する。

$$\mathcal{N}(\mu) \equiv \{w \in \mathcal{V} \mid \|r_\kappa(w, \mu)\| \leq \alpha\mu, X(x) \succ 0, Z \succ 0\}$$

ただし、 $0 < \alpha < 1$ である。これと、 $0 < \theta \leq \min\{\gamma, s\}$ を満たす θ を用いて

$$\Theta(\theta) \equiv \cup_{\mu \in [0, \theta]} \mathcal{N}(\mu)$$

を定義する。Lemma 2 より、この集合は中心パスの近傍と考えることが出来る。さらに、これらを用いて以下を定義する。

$$U_w(\theta) \equiv \sup_{w \in \Theta_1(\theta), \mu \in [0, \theta]} \|w - \bar{w}(\mu)\|$$

定義から $t \rightarrow 0$ のとき $U_w(t) \rightarrow 0$ であることに注意する。このとき、ある θ_0 が存在して

$$L_M U_M U_w(\theta_0) = \frac{2}{3}$$

を満たす。これらを用いて、超一次収束性の証明で必要となる不等式を与える。

Lemma 4 (仮定 1)–(仮定 4) が成り立つとし、 θ は $0 \leq \theta \leq \theta_1$, $\theta_1 \equiv \min\{\gamma, s, \theta_0\}$ を満たすとする。このとき、任意の $w \in \Theta(\theta), \mu \in [0, \theta]$ に対して

$$\|w - \bar{w}(\mu)\| \leq U_r \|r_\kappa(w, \mu)\|$$

を満たす正の数 U_r が存在する。 □

続いて、 θ をうまく選ぶことで、Algorithm 1 により生成される点列が $\Theta(\theta)$ に含まれることが示される。

Lemma 5 (仮定 1)–(仮定 4) が成り立つとする。このとき、Algorithm 1 により生成される点列 $\{w_k\}$ と $\{\mu_k\}$ は

$$\begin{aligned} w_k &\in \mathcal{N}(\mu_{k-1}) \subset \Theta(\tilde{\theta}) & k = 1, 2, \dots \\ \mu_0 < \theta, \quad 0 \leq \mu_k &\leq c\mu_{k-1} & k = 1, 2, \dots \end{aligned}$$

を満たす。ただし、 $\tilde{\theta} \equiv \min\{\theta_1, \theta_2, \theta_3, \theta_4\}$ であり、

$$\begin{aligned} \theta_2 &\equiv \left[\frac{2}{3(\alpha + \sqrt{\kappa U_y + d})^{1+\tau}} \right]^{\frac{1}{\tau}}, \quad \theta_3 \equiv \left[\frac{\nu}{U_r + U_M(1 + \sqrt{\kappa U_y + d})} \right]^{1+\tau}, \\ \theta_4 &\equiv \left[\frac{\alpha}{L_M U_M^2 (1 + \sqrt{\kappa U_y + d})^2} \right]^{\frac{1+\tau}{1-\tau}}, \quad c \equiv (\alpha + \sqrt{\kappa U_y + d})^{1+\tau} \theta^\tau \end{aligned}$$

である。 □

Lemma 5 より、初期点を $w_0 \in \Theta(\tilde{\theta})$ とすれば、Algorithm 1 により生成される点列は $w_k \in \Theta(\tilde{\theta})$, $k = 1, 2, \dots$ を満たすことが示された。これを用いて、Algorithm 1 の超一次収束性が示される。

Theorem 2 (仮定 1)–(仮定 4) が成り立つとする。このとき、初期点を $w_0 \in \Theta_2(\tilde{\theta})$ として選べば、Algorithm 2 により生成される点列 $\{w_k\}$ は w^* へ超一次収束する。 □

4 まとめ

本稿では、主双対内点法に基づいた Algorithm 1 を提案し、その超一次収束性に関する議論を行った。Algorithm 1 は、中心化 KKT 条件 [4] とシフト付き中心化 KKT 条件 [2] を包括した、近似 KKT 条件に基づいたアルゴリズムである。収束証明では、いくつかの仮定を与え、それらの仮定の下で Algorithm 1 により生成される点列が KKT 点へ超一次収束することを示した。

参考文献

- [1] Ortega, J., Rheinboldt, W.: Iterative solution of nonlinear equations in several variables, Academic Press, 1970.
- [2] Kato, A., Yabe, H., Yamashita, H.: An interior point method with a primal-dual quadratic barrier penalty function for nonlinear semidefinite programming, Technical Report, Department of Mathematical Information Science, Tokyo University of Science, 2012.
- [3] Wenyu, Sun., Ya-Xiang, Yuan.: Optimization theory and methods, Springer Science+Business Media, 2006.
- [4] Yamashita, H., Yabe, H.: Local and superlinear convergence of a primal-dual interior point method for nonlinear semidefinite programming, Mathematical Programming **132**, 1-30 (2012).