非線形半正定値最適化問題に対する
2次の最適性保証付き主双対内点法

05001374 東京大学 *新幡駿 ARAHATA Shun
02006704 理化学研究所 奥野貴之 OKUNO Takayuki
01308490 東京大学/理化学研究所 武田朗子 TAKEDA Akiko

1. はじめに

本研究では，以下の非線形半正定値最適化問題（nonlinear semidefinite optimization problem; NSDP）を考える：

$$\begin{aligned} &\underset{x \in \mathbb{R}^n}{\text{minimize}} && f(x) \\ &\text{subject to} && X(x) \in \mathbb{S}^m_+. \end{aligned} \tag{1}$$

ここで，$f : \mathbb{R}^n \to \mathbb{R}$ と $X : \mathbb{R}^n \to \mathbb{S}^m$ は C^2 級とする．$\mathbb{S}^m, \mathbb{S}^m_+, \mathbb{S}^m_{++}$ はそれぞれ $m \times m$ の実対称行列，半正定値対称行列，正定値対称行列の集合を表す．NSDP の応用先として，制御，ファイナンス，構造最適化などがある．本稿では $X(x)$ を X と示すことにして，狭義実行可能領域 $\mathcal{X}$ を $\mathcal{X} := \left\{x \in \mathbb{R}^n : X \in \mathbb{S}^m_{++}\right\}$ と定める．$\mathcal{X}$ の凸包 $\mathrm{conv}(\mathcal{X})$ において，f の勾配，ヘシアンがそれぞれ L_1, L_2 リプシッツ連続であると仮定する．さらに制約については，任意の $x \in \mathrm{conv}(\mathcal{X})$ において，同じ L_1 と新しく L_0 を用いて

$$\sum_{i=1}^n \|\mathcal{A}_i(x)\|_{\mathrm{F}} \le L_0,\ \sum_{i=1}^n \sum_{j=1}^n \left\| \frac{\partial^2 X}{\partial x_i \partial x_j}(x) \right\|_{\mathrm{F}} \le L_1$$

が成り立つことと，同じ L_2 を用いて

$$\sum_{i=1}^n \sum_{j=1}^n \left\| \frac{\partial^2 X}{\partial x_i \partial x_j}(x) - \frac{\partial^2 X}{\partial x_i \partial x_j}(z) \right\|_{\mathrm{F}} \le L_2 \|x - z\|$$

が任意の $x, z \in \mathrm{conv}(\mathcal{X})$ について成り立つことを仮定する．ここで，$\mathcal{A}_i(x) := \frac{\partial X}{\partial x_i}(x)$ である．

本研究では，NSDP に対して，ヘシアンの最小固有値方向を用いることで 2 次の最適性の必要条件を保証する主双対内点法ベースのアルゴリズムを新たに提案した．さらに，提案したアルゴリズムに，2 次の停留点 (second-order stationary point; SOSP) への収束に対する最悪反復計算量の保証を与えた．非線形計画問題 (nonlinear optimization problem; NLP) に対しては，[1] などの 2 次の停留点への収束保証がついたアルゴリズムがいくつか存在する．一方，NSDP を NLP に帰着することなく 2 次の最適性を保証するアルゴリズムは本研究が初めてである．

2. 最適性条件と制約想定

ラグランジュ乗数行列を $\Lambda \in \mathbb{S}^m$ として，問題 (1) のラグランジュ関数を $L(x, \Lambda) := f(x) - \langle X, \Lambda \rangle$ と定める．ラグランジュ関数を用いて Karush–Kuhn–Tucker (KKT) 条件は次のように書かれる:

$$\nabla_x L(x, \Lambda) = 0,\ X\Lambda = O,\ X, \Lambda \in \mathbb{S}^m_+$$

Mangasarian–Fromovitz constraint qualification (MFCQ) のもと，KKT 条件は局所的最適解であるための必要条件であることが知られている．次に 2 次の必要条件を述べるが，まず点 x における $\mathcal{L}(x)$ を次のように定義する: $\mathcal{L}(x) := \{d \in \mathbb{R}^n \mid E^\top(\Delta X(x; d))E = O\}$. ここで，$\Delta X(x; d) := \sum_{i=1}^n \mathcal{A}_i(x) d_i$，$E$ は $X(x)$ の零空間の基底を列に並べたものである．KKT 点 $(\bar{x}, \bar{\Lambda})$ において，

$$d^\top(\nabla^2_{xx} L(\bar{x}, \bar{\Lambda}) + H(\bar{x}, \bar{\Lambda}))d \ge 0 \tag{2}$$

が任意の $d \in \mathcal{L}(\bar{x})$ について成立するとき，$(\bar{x}, \bar{\Lambda})$ は2次の必要条件を満たし，SOSP と呼ばれる．ここで，$\left(H(\bar{x}, \bar{\Lambda})\right)_{ij} := 2\,\mathrm{tr}\left(\mathcal{A}_i(\bar{x}) X(\bar{x})^\dagger \mathcal{A}_j(\bar{x}) \bar{\Lambda}\right)$ である．狭義相補性条件と非退化条件のもと上記の 2 次の必要条件は，KKT 点 $(\bar{x}, \bar{\Lambda})$ が局所的最適解であるための必要条件となることが知られている．NSDP の MFCQ，狭義相補性条件，非退化条件については，[2] に書かれている．

3. アルゴリズム

アルゴリズム 1は主双対内点法に基づいており，点列を狭義実行可能領域 $\mathcal{X} \times \mathbb{S}^m_{++}$ 内に生成する．

$\mu, \nu > 0$ として，次の主双対メリット関数 ψ_μ を導入する: $\psi_\mu(x, Z) := F_{\mathrm{PB}}(x) + \nu F_{\mathrm{BC}}(x, Z)$. ここで $F_{\mathrm{PB}}(x) := f(x) - \mu \log\det X, F_{\mathrm{BC}}(x, Z) := \langle X, Z\rangle - \mu\log\det X - \mu \log\det Z$ であり，[3] によって提案された．アルゴリズム 1は勾配方向と最小固有値方向を用いて ψ_μ を最小化することで次の条件を満たす点を求める：

$$
\begin{aligned}
&X, Z \in \mathbb{S}^m_{++}, \\
&\|\nabla_x \psi_\mu(x, Z)\| \le \varepsilon_g\big(1 + \mu\big\|X^{-1}\big\|_{\mathrm{F}} + \|Z\|_{\mathrm{F}}\big), \\
&\|\nabla_Z \psi_\mu(x, Z)\|_{\mathrm{F}} \le \varepsilon_\mu(1 + \mu\big\|Z^{-1}\big\|_{\mathrm{F}}), \\
&\lambda_{\min}\big(\nabla^2_{xx}\psi_\mu(x, Z)\big) \ge -\varepsilon_H\big(1 + \mu\big\|X^{-1}\big\|_{\mathrm{F}} + \|Z\|_{\mathrm{F}}\big)^2.
\end{aligned}
$$

この条件を満たす点 (x, Z) を $(\varepsilon_g, \varepsilon_\mu, \varepsilon_H)$-SOSP$(\mu, \nu)$ と呼ぶ.

アルゴリズム中のステップサイズの詳細は当日述べるが，リプシッツ定数 L_0, L_1, L_2 を用いて，$(x^{\ell+1}, Z_{\ell+1}) \in \mathcal{X} \times \mathbb{S}^m_{++}$ となるように定めている. 特に, この定め方によって ψ_μ の減少量を下から抑えられることができ, 最終的に最悪計算量に関する以下の結果を得た.

定理 1. メリット関数 ψ_μ が $\mathcal{X} \times \mathbb{S}^m_{++}$ で下に有界であると仮定する．点 (x_1, Z_1) を アルゴリズム 1の初期点として， アルゴリズム 1は $(\varepsilon_g, \varepsilon_\mu, \varepsilon_H)$-SOSP$(\mu, \nu)$ に次の反復回数のオーダーで収束する:

$$
\mathcal{O}\left(\frac{\psi_\mu(x_1, Z_1) - \psi^*_\mu}{\min\left\{\varepsilon_g^2, \varepsilon_\mu^2, \varepsilon_H^3\right\}}\right).
$$

ここで，ψ^*_μ は $\mathcal{X} \times \mathbb{S}^m_{++}$ での ψ_μ の下限である． ◇

実は，定理 1 により保証される $(\varepsilon_g, \varepsilon_\mu, \varepsilon_H)$-SOSP$(\mu, \nu)$ について $\varepsilon_g, \varepsilon_\mu, \varepsilon_H, \mu, \nu$ を小さくして生成した点列の集積点が問題 (1) の SOSP になっている (定理 2).

定理 2. $(\varepsilon_{g_k}, \varepsilon_{\mu_k}, \varepsilon_{Hk})$-近似 SOSP$(\mu_k, \nu_k)$ のなす点列 $\big\{\big(x_k, Z^k\big)\big\}$ を取り，以下を仮定する.

1. 点列 $\big\{\big(x_k, Z^k\big)\big\}$ は有界である.
2. 点列 $\{\varepsilon_{g_k}\}, \left\{\frac{\varepsilon_{\mu_k}}{\nu_k \mu_k}\right\}, \{\nu_k\}, \{\varepsilon_{Hk}\}, \{\mu_k\}$ が 0 に収束する.
3. 点列 $\big\{\mu_k X(x_k)^{-1}\big\}$ が有界であるか，点列 $\big\{\lambda_{\min}\big(\mu_k X(x_k)^{-1}\big)\big\}$ が発散する.
4. 点列 $\big\{\big(x_k, Z^k\big)\big\}$ の集積点において，狭義相補性条件と非退化条件が成り立つ.

このとき，集積点は 2 次の必要条件 (2) を満たす. ◇

アルゴリズム 1: 負固有方向を用いた主双対内点法（概略）

Input: $x^1 \in \mathcal{X}, Z_1 \in \mathbb{S}^m_{++}$, μ, ν, ε_g, ε_μ, ε_H, L_0, L_1, L_2
for $\ell = 1, \ldots$**:**
  if $\big\|\nabla_Z \psi_\mu\big(x^\ell, Z_\ell\big)\big\|_{\mathrm{F}} > \varepsilon_\mu(1 + \mu\big\|Z_\ell^{-1}\big\|_{\mathrm{F}})$**:**
    方向 d_{Z_ℓ} として，$\nabla_Z \psi_\mu\big(x^\ell, Z_\ell\big)$ を用い，ステップサイズルールにより α_ℓ を定める．$Z_{\ell+1} \leftarrow Z_\ell + \alpha_\ell d_{Z_\ell}$, $x^{\ell+1} \leftarrow x^\ell$
  else if $\big\|\nabla_x \psi_\mu\big(x^\ell, Z_\ell\big)\big\| > \varepsilon_g\big(1 + \mu\big\|X_\ell^{-1}\big\|_{\mathrm{F}} + \|Z_\ell\|_{\mathrm{F}}\big)$**:**
    方向 d_{x^ℓ} として，$\nabla_x \psi_\mu\big(x^\ell, Z_\ell\big)$ を用い，ステップサイズルールにより α_ℓ を定める．$x^{\ell+1} \leftarrow x^\ell + \alpha_\ell d_{x^\ell}$, $Z_{\ell+1} \leftarrow Z_\ell$
  else if $\lambda_{\min}\big(\nabla^2_{xx}\psi_\mu\big(x^\ell, Z_\ell\big)\big) < -\varepsilon_H\big(1 + \mu\big\|X_\ell^{-1}\big\|_{\mathrm{F}} + \|Z_\ell\|_{\mathrm{F}}\big)^2$**:**
    方向 d_{x^ℓ} として，$\nabla^2_{xx}\psi_\mu\big(x^\ell, Z_\ell\big)$ の正規化された最小固有値方向を用い，ステップサイズルールにより α_ℓ を求める．$x^{\ell+1} \leftarrow x^\ell + \alpha_\ell d_{x^\ell}$, $Z_{\ell+1} \leftarrow Z_\ell$
  else:
    Output(x^ℓ, Z_ℓ)

参考文献

[1] O. Hinder and Y. Ye: *Worst-case iteration bounds for log barrier methods for problems with nonconvex constraints.* 2020. arXiv: `1807.00404 [math.OC]`.

[2] J. Bonnans and A. Shapiro: *Perturbation Analysis of Optimization Problems.* Springer, 2000.

[3] H. Yamashita, H. Yabe, and K. Harada: A primal–dual interior point method for nonlinear semidefinite programming. *Mathematical Programming* 135.1 (2012), pp. 89–121.