双曲錐への射影計算

	東京大学	*永野隆之	NAGANO Takayuki
05000745	統計数理研究所	ロウレンソ 武流野 フィゲラ	LOURENÇO Bruno Figueira
01308490	東京大学/理化学研究所	武田朗子	TAKEDA Akiko

1. はじめに

本研究では双曲錐（hyperbolicity cone）への射影を計算する一次法を提案する．双曲錐は半正定値行列錐や対称錐等を特殊ケースとして含む，高い表現能力を持った重要なクラスである．双曲錐制約を含む最適化問題は他分野でも応用が期待されている [5] が，効率的な解法の研究は未だ発展途上である．そこで，本研究では双曲錐制約を含む最適化問題の効率的な解法を構築するための足掛かりとして，最適化アルゴリズムにおいて重要な役割を果たす射影を計算する手法を提案する．

射影を計算するための既存手法として，内点法 [2] や Renegar[4] が提案した平滑化と加速法を利用した一次法などが挙げられるが，これらの手法には一反復当たりの計算量が大きいなどの問題点がある．本研究ではこれらの既存手法が抱える問題点を克服する手法を提案する．

2. 双曲錐

$\mathcal{E} \subseteq \mathbb{R}^n$ を線形部分空間とし，$p:\mathcal{E}\to\mathbb{R}$ を斉次多項式とする．ある $e\in\mathcal{E}$ に対して，$p(e)\neq 0$ かつ任意の $x\in\mathcal{E}$ に対して t の一変数多項式 $t\mapsto p(x-te)$ が実根のみを持つとき，p を e を方向ベクトルとする双曲多項式と呼ぶ．そして，$t\mapsto p(x-te)$ の実根を x の（p,e に関する）固有値と呼び，x の最小固有値を $\lambda_{\min}(x)$，x の零固有値の個数を $\mathrm{mult}(x)$ と表記する．また，$K=\{x\in\mathcal{E}\mid\lambda_{\min}(x)\geq 0\}$ を p,e に関する双曲錐と呼ぶ．

3. 提案手法

双曲多項式 $p:\mathcal{E}\to\mathbb{R}$ と p の方向ベクトル $e\in\mathcal{E}$，$u\in\mathcal{E}$ が与えられたときに，u の p,e に対応する双曲錐 K への射影を求めることを考える．

双曲錐は閉凸錐であるので，Moreau の分解定理より

$$u=\mathcal{P}_K(u)+\mathcal{P}_{-K^*}(u)$$

が成立する．ここで，$\mathcal{P}_C(x)$ は x の C への射影，K^* は K の双対錐である．この関係式より，$\mathcal{P}_{-K^*}(u)$ が求まれば $\mathcal{P}_K(u)$ は計算できるので，以後は $\mathcal{P}_{-K^*}(u)$ を求めることを考えていく．

$\mathcal{P}_{-K^*}(u)$ の計算は次のような最適化問題として定式化できる．

$$\begin{array}{ll} \min_x & \frac{1}{2}\|x-u\|^2 \\ \text{s.t.} & x\in -K^* \end{array} \tag{1}$$

問題 (1) は問題を変形した上で Frank-Wolfe 法 (FW 法)[1] を適用することで効率的に解くことができる．ここからはその手順を紹介する．以後は簡便のため $f(x)=\frac{1}{2}\|x-u\|^2$ と書くことにする．

まず，FW 法は実行可能領域がコンパクトでないと適用できないため，問題 (1) には FW 法を適用することができない．そこで，問題 (1) を実行可能領域がコンパクトで等価な問題に変形する．問題 (1) の最適解 $\mathcal{P}_{-K^*}(u)$ は $\langle\mathcal{P}_{-K^*}(u),-e\rangle\leq\|u\|\|e\|$ を満たすので，問題 (1) は次の問題と等価である．

$$\begin{array}{ll} \min_x & f(x) \\ \text{s.t.} & \langle x,-e\rangle\leq\|u\|\|e\| \\ & x\in -K^* \end{array} \tag{2}$$

問題 (2) は実行可能領域がコンパクトなので，FW 法を適用することができる．問題 (2) に FW 法を適用すると毎反復で次の子問題を解くことになる．

$$\begin{array}{ll} \min_x & \langle\nabla f(x_k),x\rangle \\ \text{s.t.} & \langle x,-e\rangle\leq\|u\|\|e\| \\ & x\in -K^* \end{array} \tag{3}$$

この子問題 (3) の最適解は解析的に求めることができる．ここからはその最適解を導出していく．

(i)$-\nabla f(x_k)\in K$ のとき

双対錐の定義より $\mathbf{0}$ が問題 (3) の最適解となる．

(ii)$-\nabla f(x_k)\notin K$ のとき

この場合には，子問題 (3) は不等式制約を等式制約で置き換えた次の問題 (4) と等価になる．

$$\begin{array}{ll} \min_x & \langle\nabla f(x_k),x\rangle \\ \text{s.t.} & \langle x,-e\rangle=\|u\|\|e\| \\ & x\in -K^* \end{array} \tag{4}$$

問題 (4) の双対問題は次のようになる．

$$
\begin{array}{lll}
\max_{y,z} & \|u\|\|e\| \cdot y & \\
\text{s.t.} & z \in -K & \qquad (5) \\
 & -ye + z = \nabla f(x_k) &
\end{array}
$$

双対問題 (5) の最適解は

$$
\begin{aligned}
y_{\mathrm{opt}} &= \lambda_{\min}(-\nabla f(x_k)) \\
z_{\mathrm{opt}} &= \nabla f(x_k) + \lambda_{\min}(-\nabla f(x_k))e
\end{aligned}
$$

となる．この双対問題の最適解 z_{opt} と K の幾何学的な性質を組み合わせることで，

$$
x_{\mathrm{opt}} = -\frac{\|u\|\|e\|}{\langle \nabla p^{(r-1)}(-z_{\mathrm{opt}}), e\rangle} \nabla p^{(r-1)}(-z_{\mathrm{opt}}) \quad (6)
$$

が問題 (4) の最適解の一つとなることを示せる．ここで，$r = \mathrm{mult}(-z_{\mathrm{opt}})$，$p^{(i)} = D_e^{(i)} p$（$p$ を e 方向に i 回方向微分して得られる多項式）である．

以上より問題 (2) の子問題 (3) の最適解が得られたので，最後にステップサイズ α_k の定め方について言及しておく．定数ステップサイズや単調減少ステップサイズ，Armijo 条件で定まるステップサイズはもちろん利用できるが，今回扱っている問題については目的関数 $f(x) = \frac{1}{2}\|x - u\|^2$ の勾配が 1-Lipschitz 連続であることが分かるため，それを用いたステップサイズも利用できる．
また，直線探索による最適ステップサイズは，降下方向を d_k として

$$
\alpha_{\mathrm{opt}} = -\frac{\langle d_k, \nabla f(x_k)\rangle}{\|d_k\|^2} \ (> 0) \qquad (7)
$$

であり，これも簡単に計算できるものになっている．いずれのステップサイズを利用する場合でも $\alpha \in (0,1]$ に限定する必要がある点には注意する．以上を踏まえた提案手法全体の流れを Algorithm 1 に示す．

4.　収束解析

FW 法は，目的関数が凸関数かつ実行可能領域が凸集合であれば，反復回数を k として $O(1/k)$ のレートで収束することが知られている [3]．問題 (2) は目的関数が凸なので，この問題は $O(1/k)$ のレートで最適解に収束することが保証される．

5.　数値実験

数値実験の結果は当日紹介する．

Algorithm 1: 提案手法

$x_0 \in -K^*$ を選ぶ．
for $k = 0, 1, \ldots$ **do**
　もし x_k が停止条件を満たせば $u - x_k$ を出力して停止する．
　if $-\nabla f(x_k) \in K$ **then**
　　$s_k = \mathbf{0}$
　else
　　$z_k = \nabla f(x_k) + \lambda_{\min}(-\nabla f(x_k))e$
　　$r_k = \mathrm{mult}(-z_k)$
　　$s_k = -\frac{\|u\|\|e\|}{\langle \nabla p^{(r_k-1)}(-z_k), e\rangle} \nabla p^{(r_k-1)}(-z_k)$
　end if
　ステップサイズ $\alpha_k \in (0,1]$ を選ぶ．
　$x_{k+1} = x_k + \alpha_k(s_k - x_k)$
end for

6.　おわりに

本研究では，FW 法を利用した双曲錐への射影を計算する一次法を提案した．今回は射影計算，つまり目的関数を $\frac{1}{2}\|x - u\|^2$ に限定してアルゴリズムを提案したが，目的関数をより一般の関数に拡張しても同様のアルゴリズムを構築できる可能性がある．この拡張については当日触れたいと思う．

参考文献

[1] M. Frank and P. Wolfe. An algorithm for quadratic programming. *Naval research logistics quarterly*, 3(1-2):95–110, 1956.

[2] O. Güler. Hyperbolic polynomials and interior point methods for convex programming. *Mathematics of Operations Research*, 22(2):350–377, 1997.

[3] E. S. Levitin and B. T. Polyak. Constrained minimization methods. *USSR Computational mathematics and mathematical physics*, 6(5):1–50, 1966.

[4] J. Renegar. Accelerated first-order methods for hyperbolic programming. *Mathematical Programming*, 173(1):1–35, 2019.

[5] J. Saunderson. Certifying polynomial nonnegativity via hyperbolic optimization. *SIAM Journal on Applied Algebra and Geometry*, 3(4):661–690, 2019.