Strongly convex optimization over hyperbolicity cones (双曲錐上の強凸関数最適化)

数理情報学専攻 48216224 永野 隆之
指導教員 武田 朗子 教授

1 はじめに

双曲錐 (hyperbolicity cone) は対称錐や半正定値錐を特別な場合として含む，高い一般性を持つ錐のクラスである．双曲錐はその幾何学的な性質が盛んに研究されているが，最適化の分野でも制約の表現能力の高さ等から関心を集めている．しかし双曲錐への射影を効率的に計算する手法が未だ見つかっていないために，双曲錐制約問題に適用できる最適化手法は限られている．内点法 [1] や Renegar [2] が提案した一次法を用いれば射影を計算することはできるが，これらの手法はある程度の精度の解を求めるのにも時間がかかるという問題点を抱えており，射影勾配法などのサブルーチンとして射影を計算するには不向きである．本研究では，まず特別な構造を持つ双曲錐に対する距離関数と射影作用素の表式を導出する．この結果は対称行列と半正定値錐との距離関数・射影の一般化になっている．この結果を適用できる双曲錐に対しては導出した式を用いることで射影を高速に計算できるが，実は適用可能な場面はあまり多くない．そこで，本研究では一般の双曲錐への射影計算に利用可能な，双対理論とフランクウルフ法 [3] に基づく一次法を併せて提案する．また，提案手法の収束解析と数値実験の結果も紹介する．

2 特別な構造を持つ双曲錐への射影

まず，双曲錐の定義を紹介する．

定義 1 (双曲多項式，固有値，双曲錐)**.** d 次の斉次多項式 $p:\mathbb{R}^n\to\mathbb{R}$ が $e\in\mathbb{R}^n$ に対して $p(e)\neq 0$ であり，さらに任意の $x\in\mathbb{R}^n$ に対して $t\mapsto p(x-te)$ 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**という．また，$t\mapsto p(x-te)$ の実根を (p,e) に対する x の**固有値**と呼び，$\lambda(x)=(\lambda_1(x),\ldots,\lambda_d(x))^\top$ と表す．ここで，$\lambda_1(x)\geq\cdots\geq\lambda_d(x)$ である．そして，$\Lambda(p,e)=\{x\in\mathbb{R}^n \mid \lambda_d(x)\geq 0\}$ を (p,e) に対する**双曲錐**という．

仮定の詳細は本稿では省略するが，特別な構造を持つ双曲錐に対する距離関数と射影は閉形式で表せる．

定理 1. p が isometric な双曲多項式であり，$\lambda(\mathbb{R}^n)=\mathbb{R}^d_\downarrow$ ならば，(p,e) が誘導するノルムに関して

$$\mathrm{dist}(x,\Lambda(p,e))^2=\sum_{i=1}^d \min\{\lambda_i(x),0\}^2.$$

さらに，$\lambda_i(x)\neq\lambda_j(x)$ $(i\neq j)$ ならば，

$$P_{\Lambda(p,e)}(x)=\sum_{i=1}^d \max\{\lambda_i(x),0\}\frac{\nabla p(x-\lambda_i(x)e)}{D_e p(x-\lambda_i(x)e)}.$$

ここで $\mathbb{R}^d_\downarrow=\{u\in\mathbb{R}^d \mid u_1\geq\cdots\geq u_d\}$ である．

半正定値錐はこの定理の仮定を満たすため，この結果は対称行列のスペクトル分解の拡張になっている．しかし，この定理における仮定は強く，性質の良い錐でも成り立たないことが多々ある．例えば，ℓ_∞ 錐は isometric だが $\lambda(\mathbb{R}^n)\neq\mathbb{R}^d_\downarrow$ である．

3 双曲錐上の強凸関数最適化アルゴリズム

この節では，定理 1 の仮定を満たさない一般の双曲錐に対する射影を計算するアルゴリズムを紹介する．提案手法は (P) に適用可能な一次法である．

$$\begin{cases}\min\ f(x)\\ \text{s.t.}\ \ Tx+u\in\Lambda(p,e)\end{cases}\tag{P}$$

ここで，$f:\mathbb{R}^n\to\mathbb{R}$ は強凸関数，$T:\mathbb{R}^n\to\mathbb{R}^m$ は線形写像，$u\in\mathbb{R}^m$ は所与の点である．$f(x)=\|x-b\|^2, T=I$，$u=0$ とすれば，(P) は b の双曲錐への射影問題となる．提案手法では，(P) の双対問題と等価な問題

$$\begin{cases}\min & (f^*\circ T^*)(y)+\langle u,y\rangle\\ \text{s.t.} & \langle Te,y\rangle\leq c_D\|e\|\\ & y\in\Lambda(p,e)^*\end{cases}\tag{D'}$$

をフランクウルフ法 (FW 法) [3] で解く．(P) と (D') の最適解 x_{opt}，y_{opt} の間には，KKT 条件から

$$x_{\mathrm{opt}}=\nabla f^*(T^*y_{\mathrm{opt}})\tag{1}$$

が成り立つので，(D') の最適解が見つかれば (P) の最適解が求められる．提案手法であえて双対問題 (D') に FW 法を適用するのは，(D') に FW 法を適用した際に現れる子問題

$$\begin{cases}\min & \langle\nabla(f^*\circ T^*)(y_k)+u,y\rangle\\ \text{s.t.} & \langle Te,y\rangle\leq c_D\|e\|\\ & y\in\Lambda(p,e)^*\end{cases}\tag{Sub}$$

の解が解析的に導出できるからである．

Algorithm 1 提案手法

1: 初期点 $y_0 \in \Lambda(p,e)^*$ を選ぶ
2: **for** $k = 0, 1, \ldots$ **do**
3: $\quad x_k = \nabla f^*(T^* y_k)$
4: $\quad x_k, y_k$ が停止条件を満たしていたら停止
5: $\quad$ **if** $Tx_k + u \in K$ **then**
6: $\quad\quad s_k = \mathbf{0}$
7: $\quad$ **else**
8: $\quad\quad z_k = (Tx_k + u) - \lambda_{\min}(Tx_k + u)Te$
9: $\quad\quad s_k = \frac{c_D \|e\|}{\langle Te, \nabla p^{(\mathrm{mult}(z_k)-1)}(z_k)\rangle} \nabla p^{(\mathrm{mult}(z_k)-1)}(z_k)$
10: $\quad$ ステップサイズ $\alpha_k \in (0,1]$ を選択
11: $\quad y_{k+1} = y_k + \alpha_k(s_k - y_k)$

定理 2. $\nabla(f^* \circ T^*)(y_k) + u \in \Lambda(p,e)$ のとき，0 が (Sub) の最適解の一つである．またそうでないときは

$$\frac{c_D \|e\|}{\langle Te, \nabla(D_e^{(\mathrm{mult}(z_k)-1)}p)(z_k)\rangle} \nabla(D_e^{(\mathrm{mult}(z_k)-1)}p)(z_k)$$

が (Sub) の最適解の一つである．ここで $z_k = \nabla(f^* \circ T^*)(y_k) + u - \lambda_{\min}(\nabla(f^* \circ T^*)(y_k) + u)Te$ であり，$\mathrm{mult}(z_k)$ は z_k の零固有値の個数である．

以上の流れを Algorithm 1 にまとめる．Algorithm 1 では (P) に対する出力 $\{x_k\}$ を第 3 行のように定めている．こうすることで，(1) より y_k が (D') の最適解に近づけば x_k も (P) の最適解に近づくことが期待できる．また x_k の定め方から $\nabla(f^* \circ T^*)(y_k) = Tx_k$ となる．

4 収束解析

Algorithm 1 の収束性に関する理論保証を紹介する．

定理 3. Algorithm 1 の出力 $\{x_k\}$ は以下を満たす．

(i) 適切な $\{\alpha_k\}$ のもとで $\lim_{k\to\infty} x_k = x_{\mathrm{opt}}$.

(ii) k 反復目の出力 x_k が実行可能解ならば
$$f(x_k) - f_{\mathrm{opt}} \le \langle Tx_k + u, y_k\rangle = G(y_k).$$

(iii) $T^*T \succ 0$ ならば，適切な $\{\alpha_k\}$ のもとで
$$\|x_k - x_{\mathrm{opt}}\| = O(1/\sqrt{k}).$$

(i) と (iii) は単調減少ステップサイズやアルミホルール等の下で成立する．また (ii) 右辺の $G(y_k)$ は，フランクウルフギャップと呼ばれる FW 法で最適性誤差の指標としてよく用いられる量である．提案手法では (D') に FW 法を適用しているが，(P) における最適性誤差もフランクウルフギャップで評価できる．

5 数値実験

$E_i^n(x)$ を n 変数 i 次基本対称式，$\mathbf{1}_n \in \mathbb{R}^n$ を全要素が 1 の n 次元ベクトルとする．$\Lambda(E_{n-1}^n, \mathbf{1}_n) \subset \mathbb{R}^n$ ($n = 10, 20$) への射影問題を提案手法と Renegar の手法 [2]，内点法ソルバ [4] で解き，各相対誤差を達成するまでに要した平均時間を表 1 にまとめる．また，錐を $\Lambda(E_3^{50}, \mathbf{1}_{50}) \subset \mathbb{R}^{50}$ に変えて同様の実験を行った結果を表 2 にまとめる．提案手法と Renegar の手法では双曲錐に対する固有値を計算しなくてはならないが，$\Lambda(E_{n-1}^n, \mathbf{1}_n)$ に対する固有値を計算するには $n-1$ 次方程式を解く必要があり，n が大きくなるとこの固有値計算は数値的に不安定になりやすい．一方で，$\Lambda(E_3^{50}, \mathbf{1}_{50})$ に対する固有値は 3 次方程式を解けば求まるので固有値計算は数値的に不安定になりにくい． 表 1,2 の結果から，提案手法は Renegar の手法よりも良い解を得られることと，提案手法は高精度な解を求めようとすると内点法よりも時間がかかってしまうが，ある程度の精度の解であれば内点法よりも高速で求められることが確認できる．また，$\Lambda(E_{n-1}^n, \mathbf{1}_n)$ での実験で n が大きくなると提案手法と Renegar の手法は数値的に不安定になるが，$\Lambda(E_3^{50}, \mathbf{1}_{50})$ での実験では不安定にならない．これらの結果から，提案手法と Renegar の手法の精度は次元の大きさよりも固有値の精度に依存していることが予想される．

表 1. $\Lambda(E_{n-1}^n, \mathbf{1}_n)$ への射影問題で各相対誤差を達成するために要した平均計算時間 (秒)

n	Method	10%	3%	1%	0.1%	0.01%
10	Ours	**0.00322**	**0.00370**	**0.00409**	**0.0569**	1.847
	Renegar's	0.405	0.835	7.86	-	-
	DDS	0.293	0.369	0.435	0.593	**0.687**
20	Ours	0.0191	0.0199	0.0287	0.137	-
	Renegar's	-	-	-	-	-
	DDS	1.266	1.604	1.906	2.440	**2.774**

表 2. $\Lambda(E_3^{50}, \mathbf{1}_{50})$ への射影問題で各相対誤差を達成するために要した平均計算時間 (秒)

Method	10%	5%	3%	1%	0.1%	0.01%
Ours	**0.941**	**0.941**	**0.941**	**0.941**	**0.941**	**0.941**
Renegar's	40.169	43.450	60.163	72.752	126.789	558.107
DDS	45.174	48.333	50.734	55.794	62.679	75.312

参考文献

[1] O.Güler. "Hyperbolic polynomials and interior point methods for convex programming." Mathematics of Operations Research 22(2),1997,pp.350-377

[2] J.Renegar. "Accelerated first-order methods for hyperbolic programming." Mathematical Programming 173(1),2019,pp.1-35

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

[4] M.Karimi and L. Tunçel. "Domain-Driven Solver (DDS) Version 2.0: a MATLAB-based Software Package for Convex Optimization Problems in Domain-Driven Form." arXiv preprint arXiv:1908.03075,2019