

交互射影法の理論と応用

田中 未来

二つの閉集合 X, Y が与えられたとき、それぞれの集合上の点の組のうち互いに最も近いものを求める問題を考える。この問題に交互最小化を適用すると、 X への射影と Y への射影を交互に繰り返すアルゴリズムが得られる。このアルゴリズムは交互射影法と呼ばれる。問題が凸な場合の交互射影法の収束性は古くから知られていたが、問題が非凸な場合の収束性は近年になって明らかにされつつある。本稿では問題が凸な場合と非凸な場合のそれぞれにおける交互射影法の収束性について概説する。また、非凸な問題の応用として二つの閉凸錐の間の最小角を求める問題に対する効率のよいアルゴリズムを紹介する。

キーワード：射影, 交互射影法, 凸最適化, 非凸最適化, 収束解析, 閉凸錐, 球面上の最適化

1. 導入

n 次元 Euclid 空間 \mathbb{R}^n 上の m 個の閉集合 X_i ($i = 1, \dots, m$) に対する

$$\text{find } \mathbf{x}^* \text{ such that } \mathbf{x}^* \in \bigcap_{i=1}^m X_i \quad (1)$$

という問題は科学技術計算における基本的な問題であり、多くの分野での応用が知られている [1, Chapter 1]. 特に、 $m = 2$ のとき、 $X := X_1, Y := X_2$ とおき、Euclid ノルムを $\|\cdot\|$ で表すことにすると、問題 (1) が解をもつためには、最適化問題

$$\min_{\mathbf{x} \in X, \mathbf{y} \in Y} \|\mathbf{x} - \mathbf{y}\| \quad (2)$$

の最適解が存在し、最適値が 0 であることが必要かつ十分である。また、 $X \cap Y = \emptyset$ の場合であっても問題 (2) の最適値は X と Y の間の距離とみなせる基本的な量である。本稿では問題 (2) を解くためのアルゴリズムを考える。

問題 (2) に対し、 \mathbf{x} と \mathbf{y} を交互に最小化するアルゴリズムは次のように書ける：

$$\begin{cases} \mathbf{y}^{(k+1)} \in \text{proj}_Y(\mathbf{x}^{(k)}), \\ \mathbf{x}^{(k+1)} \in \text{proj}_X(\mathbf{y}^{(k+1)}) \end{cases} \quad (k = 1, 2, \dots). \quad (3)$$

ここで $\text{proj}_X, \text{proj}_Y$ はそれぞれ X, Y への射影を表し、次のように定義される：

$$\begin{aligned} \text{proj}_X(\mathbf{y}) &:= \operatorname{argmin}_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\|, \\ \text{proj}_Y(\mathbf{x}) &:= \operatorname{argmin}_{\mathbf{y} \in Y} \|\mathbf{x} - \mathbf{y}\|. \end{aligned} \quad (4)$$

ここで argmin は最適解集合を返すものとする。本稿では式 (3) で表されるアルゴリズムを交互射影法と呼ぶ。特に、問題 (2) が凸 — すなわち X, Y がいずれも凸集合であるとき、 $\text{proj}_X, \text{proj}_Y$ は単集合を返すので、 $\text{proj}_X, \text{proj}_Y$ を問題 (4) の最適解を返す関数とみなして、交互射影法を次のように書くことも多い：

$$\begin{cases} \mathbf{y}^{(k+1)} := \text{proj}_Y(\mathbf{x}^{(k)}), \\ \mathbf{x}^{(k+1)} := \text{proj}_X(\mathbf{y}^{(k+1)}) \end{cases} \quad (k = 1, 2, \dots).$$

問題 (4) の最適解はしばしば陽に書き下すことができる。そのため、交互射影法は実装が容易なことが多く、多くの応用で用いられている。

交互射影法は $m \geq 3$ であるような問題 (1) に対するアルゴリズムに拡張することができる。最も自然な拡張は、巡回射影法

$$\begin{cases} \mathbf{x}_1^{(k)} \in \text{proj}_{X_1}(\mathbf{x}_0^{(k)}), \\ \vdots \\ \mathbf{x}_m^{(k)} \in \text{proj}_{X_m}(\mathbf{x}_{m-1}^{(k)}), \\ \mathbf{x}_0^{(k+1)} := \mathbf{x}_m^{(k)}, \end{cases} \quad (k = 1, 2, \dots)$$

だが、射影平均法

$$\mathbf{x}^{(k+1)} \in \frac{1}{m} \sum_{i=1}^m \text{proj}_{X_i}(\mathbf{x}^{(k)}) \quad (k = 1, 2, \dots)$$

も交互射影法の変種とみなすことができる。実際、 $X := \prod_{i=1}^m X_i, Y := \{(\mathbf{y}^\top, \dots, \mathbf{y}^\top)^\top \in \mathbb{R}^{mn} : \mathbf{y} \in \mathbb{R}^n\}$ とおくと、 $\text{proj}_X((\mathbf{y}^\top, \dots, \mathbf{y}^\top)^\top) = \prod_{i=1}^m \text{proj}_{X_i}(\mathbf{y})$, $\text{proj}_Y((\mathbf{x}_1^\top, \dots, \mathbf{x}_m^\top)^\top) = \{(1/m) \sum_{i=1}^m \mathbf{x}_i\}$ なので、

たなか みらい

統計数理研究所数理・推論研究系

〒 190-8562 東京都立川市緑町 10-3

mirai@ism.ac.jp

X, Y に対する交互射影法は X_1, \dots, X_m に対する射影平均法に一致する。そのため、射影平均法の収束性は交互射影法の収束性から従う。また、射影平均法は巡回射影法と比べて並列分散処理が容易であるという利点もある。

問題 (2) が凸な場合の交互射影法の収束性は古くから知られていたが、非凸な場合の収束性は近年になって明らかにされつつある。本稿では交互射影法の収束性を問題が凸な場合 (2 節) と非凸な場合 (3 節) に分けて概説する。さらに 4 節では、問題 (2) が非凸な場合の応用として、著者らが最近進めている二つの錐の間の角度の計算を紹介する。

ここで、本稿で用いる記号の定義を行う。 \mathbb{R}^n 上の単位球面 $\{\mathbf{u} \in \mathbb{R}^n : \|\mathbf{u}\| = 1\}$ を S^{n-1} と表す。また、集合 X, Y に対してこれらの Minkowski 和および Minkowski 差をそれぞれ $X + Y, X - Y$ と表し、 X の内部および相対的内部をそれぞれ $\text{int}(X), \text{relint}(X)$ と表す。また、錐 $\mathcal{X} \subseteq \mathbb{R}^n$ の双対錐を $\mathcal{X}^* := \{\mathbf{s} \in \mathbb{R}^n : \mathbf{x}^\top \mathbf{s} \geq 0 \ (\mathbf{x} \in \mathcal{X})\}$ と定める。なお、本稿においてベクトルは列ベクトルとする。複数のベクトル $\mathbf{x}_1, \dots, \mathbf{x}_m$ をつなぎ合わせたベクトルは本来 $(\mathbf{x}_1^\top, \dots, \mathbf{x}_m^\top)^\top$ と書くべきものだが、本稿では記号の煩雑さを避けるために $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ と表記する。

2. 凸問題に対する交互射影法の収束性

ここでは問題が凸な場合の交互射影法の収束性について述べる。なお、問題が凸な場合の収束解析は Hilbert 空間上で議論されることが多いが、本稿では一貫性のために記述を Euclid 空間上でのものに限定する。詳細についてはサーベイ [1, 2] やその参考文献を参照されたい。

交互射影法の起源は古く、von Neumann [3] によって X, Y が部分空間である場合の大域的収束性が示されている。この結果はその後 Halperin [4] によって X_1, \dots, X_m が部分空間である場合の巡回射影法の大域的収束性に拡張されている。一般の凸集合で $X \cap Y = \emptyset$ の場合でも成り立つ大域的収束性は Cheney and Goldstein [5] によって示されている。以下ではこれを整理したものを紹介する。

Cheney and Goldstein [5] による交互射影法の大域的収束性は次に紹介する凸解析における基本的な性質を用いて初等的に示すことができる。この補題の証明は凸解析の基本的な教科書 [6] ((iv) は [5, Theorem 3]) などを参照されたい。

補題 2.1. \mathbb{R}^n 上の任意の空でない凸集合 X について次が成り立つ：

- (i) 任意の $\mathbf{y} \in \mathbb{R}^n$ に対して $\text{proj}_X(\mathbf{y})$ は単集合である。
- (ii) 任意の $\mathbf{y} \in \mathbb{R}^n$ に対し、 $\mathbf{x}^* = \text{proj}_X(\mathbf{y})$ であることの必要十分条件は、任意の $\mathbf{x} \in X$ に対して $(\mathbf{y} - \mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \leq 0$ が成り立つことである。
- (iii) 任意の $\mathbf{y}, \mathbf{y}' \in \mathbb{R}^n$ に対して $\|\text{proj}_X(\mathbf{y}) - \text{proj}_X(\mathbf{y}')\| \leq \|\mathbf{y} - \mathbf{y}'\|$ が成り立つ。したがって関数 proj_X は連続である。
- (iv) (iii) で等号が成立するとき $\|\mathbf{y} - \text{proj}_X(\mathbf{y})\| = \|\mathbf{y}' - \text{proj}_X(\mathbf{y}')\|$ が成り立つ。

この補題を用いると次の定理を示すことができる (補題 2.1 の (i), (ii) から定理 2.2 の (v) が、(iii), (iv), (v) から (vi) が示せる)。

定理 2.2 (Cheney and Goldstein [5]). \mathbb{R}^n 上の任意の空でない凸集合 X, Y に対して次が成り立つ：

- (v) \mathbf{x}^* が $\text{proj}_X \circ \text{proj}_Y$ の不動点であることの必要十分条件は、 $(\mathbf{x}^*, \text{proj}_Y(\mathbf{x}^*))$ が問題 (2) の最適解となることである。
- (vi) 問題 (2) に最適解 $(\mathbf{x}^*, \mathbf{y}^*)$ が存在することを仮定する (この仮定は X, Y の少なくとも一方が有界であるときや、 X, Y のいずれれもが多面体であるときなどに満たされる)。このとき、任意の $\mathbf{x}^{(0)} \in \mathbb{R}^n$ を初期点とした交互射影法によって生成される点列 $\{\mathbf{x}^{(k)}\}$ は \mathbf{x}^* に収束する。

定理 2.2 の (vi) では問題 (2) に最適解が存在することを仮定した。最適解が存在しない場合、次に示すように交互射影法は非有界な点列を生成する。

例 2.3 (点列が非有界となる例 (図 1 左)). \mathbb{R}^2 上の二つの集合 $X := \mathbb{R} \times \{0\}, Y := \{(y_1, y_2) : y_1 y_2 \geq 1\}$ に対し交互射影法を適用すると、生成される点列 $\{\mathbf{x}^{(k)}\}, \{\mathbf{y}^{(k)}\}$ は初期点によらず $\mathbf{x}^{(k)} - \mathbf{y}^{(k)} \rightarrow \mathbf{0}, \mathbf{y}^{(k+1)} - \mathbf{x}^{(k)} \rightarrow \mathbf{0}$ を満たすが、それぞれの点列は非有界である。

次に、点列が収束する場合の収束の速さについて述べる。交互射影法の収束の速さは速くても線形収束であり、状況によっては劣線形収束となる。これらの定義は次のとおり。

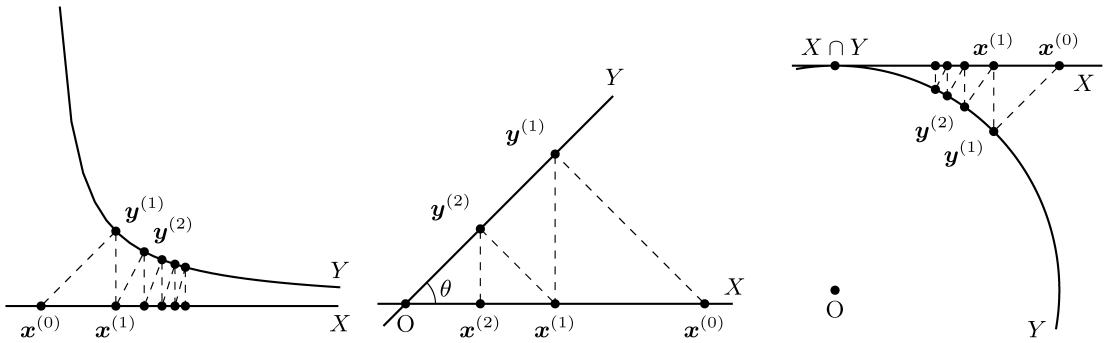


図 1 左：点列が非有界となる例 (例 2.3). 中：点列が線形収束する例 (例 2.5). 右：点列が劣線形収束する例 (例 2.6).

定義 2.4 (線形収束, 劣線形収束). 点 \mathbf{x}^* に収束する点列 $\{\mathbf{x}^{(k)}\}$ が

$$\limsup_{k \rightarrow \infty} \frac{\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\|}{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|} < 1$$

を満たすとき, $\{\mathbf{x}^{(k)}\}$ は \mathbf{x}^* に線形収束するという. また, 点 \mathbf{x}^* に収束する点列 $\{\mathbf{x}^{(k)}\}$ が

$$\limsup_{k \rightarrow \infty} \frac{\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\|}{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|} = 1$$

を満たすとき, $\{\mathbf{x}^{(k)}\}$ は \mathbf{x}^* に劣線形収束するという.

交互射影法の収束の速さの解析も歴史が長く, 古くは Aronszajn [7] が X, Y が部分空間である場合について交互射影法が生成する点列が $X \cap Y$ の元に線形収束することを示している. 単純化すれば次の例のようになる.

例 2.5 (点列が線形収束する例 (図 1 中)). \mathbb{R}^2 上の二つの集合 $X = \mathbb{R} \times \{0\}, Y = \{(r \cos \theta, r \sin \theta) : r \in \mathbb{R}\}$ ($0 < \theta < \pi/2$ は定数) に対し $(1, 0)$ を初期点 $\mathbf{x}^{(0)}$ とした交互射影法は点列 $\{\mathbf{x}^{(k)} := (\cos^{2k} \theta, 0)\}$ を生成する. この点列 $\{\mathbf{x}^{(k)}\}$ は $\mathbf{0}$ に線形収束する.

その後, Gubin et al. [8, Theorem 1 (a)] は $X \cap \text{int}(Y) \neq \emptyset$ の場合について線形収束性を示している. このような仮定は本質的で, 次に示すように X と Y が接するように交わる場合は線形収束性が失われることがある.

例 2.6 (点列が劣線形収束する例 (図 1 右)). \mathbb{R}^2 上の二つの集合 $X := \mathbb{R} \times \{1\}, Y := \{\mathbf{y} : \|\mathbf{y}\| \leq 1\}$ に対し $(1, 1)$ を初期点 $\mathbf{x}^{(0)}$ とした交互射影法は点列 $\{\mathbf{x}^{(k)} := (1/\sqrt{k+1}, 1)\}$ を生成する. この点列 $\{\mathbf{x}^{(k)}\}$ は $(0, 1)$

に劣線形収束する.

さらに, Bauschke and Borwein [9] は次に示すように交互射影法が線形収束性をもつための [8, Theorem 1 (a)] より緩い十分条件を示している. また, この定理を用いて [9, Theorem 6.3] では射影平均法が線形収束性をもつための十分条件を示している.

定理 2.7 (Bauschke and Borwein [9, Corollary 4.5]). \mathbb{R}^n 上の任意の空でない凸集合 X, Y に対して $\mathbf{0} \in \text{relint}(X - Y)$ が成り立つことを仮定する (この仮定は $\mathbf{0} \in \text{int}(X - Y)$ が成り立つときや, $X - Y$ が部分空間であるときなどに満たされる). このとき, 交互射影法によって生成される点列 $\{\mathbf{x}^{(k)}\}$ は線形収束する.

3. 非凸問題に対する交互射影法の収束性

問題 (2) において X, Y の少なくとも一方が非凸な場合, 交互射影法の収束性を示すためには議論により一層の配慮を要する. まずは, $X \cap Y \neq \emptyset$ かつ生成される点列 $\{\mathbf{x}^{(k)}\}$ が有界であっても, $\{\mathbf{x}^{(k)}\}$ の収束先が $X \cap Y$ の元とならない例や, $\{\mathbf{x}^{(k)}\}$ が収束列にならない例をみてみよう.

例 3.1 (収束先が共通部分の元とならない例 [10, Example 4.1 を改変] (図 2)). 正の整数 n に対し

$$x_{\pm n} := \pm \frac{1}{n}, \quad y_{\pm n} := \pm \frac{3n+2}{3n(n+1)}$$

とし, $x_0, y_0 := 0$ とする. これらを用いて \mathbb{R} 上の二つの集合を $X := \{x_n : n \in \mathbb{Z}\}, Y := \{y_n : n \in \mathbb{Z}\}$ と定める. この X, Y に対し交互射影法を適用したときに生成される点列 $\{\mathbf{x}^{(k)}\}$ を考える. 初期点 $\mathbf{x}^{(0)}$ が $\mathbf{0}$ のとき, 明らかに $x^{(k)} \rightarrow 0$ である. 一方, $x^{(0)} \neq 0$ のとき, ある $N \in \mathbb{Z} \setminus \{0\}$ が存在して $\text{proj}_Y(x^{(0)}) = y_N \neq 0$ が成り立つ. 任意の $n \in \mathbb{Z}$ に対して $\text{proj}_X(y_n) = x_n$,

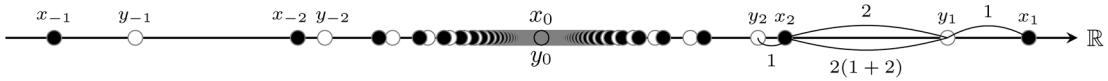


図2 収束先が共通部分の元とならない例 (例 3.1). 黒い点と白い点はそれぞれ X, Y の元を表す. また, $X \cap Y = \{0\}$ である.

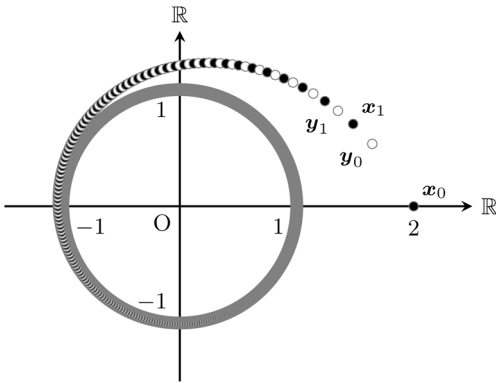


図3 点列は有界だが収束しない例 (例 3.2). 黒い点と白い点はそれぞれ X, Y の元を表す. また, $X \cap Y = \mathbb{S}^1$ である. ただし, 各点の位置は Bauschke and Noll [11] によるものとは少し異なる.

$\text{proj}_Y(x_n) = y_n$ であることに注意すると, $k \rightarrow \infty$ のとき $x^{(k)} \rightarrow x_N$ がいえるが, $x_N \notin X \cap Y = \{0\}$ である. つまりこの例では, 任意の初期点 $x^{(0)}$ について交互射影法が生成する点列 $\{x^{(k)}\}$ は収束するが, ほとんどすべての初期点 $x^{(0)} \neq 0$ についてはその収束先 x^* について期待される $x^* \in X \cap Y$ は成り立たない.

例 3.2 (点列は有界だが収束しない例 [11] (図 3)). 非有界な単調増加列 $\{\theta_n\}$ を用い, 非負の整数 $n \in \mathbb{Z}_+$ に対し $s_n := (1 + e^{-\theta_n})(\cos \theta_n, \sin \theta_n)$ とし, \mathbb{R}^2 上の二つの集合を

$$X := \{s_{2n} : n \in \mathbb{Z}_+\} \cup \mathbb{S}^1,$$

$$Y := \{s_{2n+1} : n \in \mathbb{Z}_+\} \cup \mathbb{S}^1$$

と定める. ここで \mathbb{S}^1 は \mathbb{R}^2 上の単位円周を表す. このとき, $\{\theta_n\}$ を適当にとることにより, $y_n = \text{proj}_Y(x_n)$, $x_{n+1} = \text{proj}_X(y_n)$ となるようにできる. この X, Y に対し, $x_0 = (2, 0)$ を初期点 $x^{(0)}$ として交互射影法を適用すると, 生成される点列 $\{x^{(k)}\}, \{y^{(k)}\}$ は有界で $x^{(k)} - y^{(k)} \rightarrow 0, y^{(k+1)} - x^{(k)} \rightarrow 0$ を満たすが, それぞれの点列は収束列ではない.

このような現象は X, Y が有限個の凸集合の和集合として書くことができないことに起因する. 実際, X, Y が有限個の凸集合の和集合として書くことができる

場合, 例 3.2 のような現象は起こらないことを次の定理は主張している.

定理 3.3 (Bauschke and Noll [11, Theorem 3.4]). \mathbb{R}^n 上の集合 X, Y がそれぞれ有限個の空でない閉凸集合の和集合であるものとする. さらに, これらに対して交互射影法を適用して得られる点列 $\{x^{(k)}\}, \{y^{(k)}\}$ が有界で $x^{(k)} - y^{(k)} \rightarrow 0, y^{(k+1)} - x^{(k)} \rightarrow 0$ であることを仮定する. このとき, $\{x^{(k)}\}, \{y^{(k)}\}$ は収束し, その収束先は $X \cap Y$ の元である.

次に, 問題が非凸な場合に交互射影法が局所的収束性をもつための十分条件について述べる. 問題が非凸な場合の収束解析は $X \cap Y \neq \emptyset$ の場合に対するものがほとんどである. $X \cap Y \neq \emptyset$ であっても例 2.6 のように X, Y が接するように交わる場合は収束解析は困難なため, X, Y が真に交わることを仮定した収束解析が多い¹. Lewis and Malick [12] は多様体 X, Y が横断的に交わる場合の局所的線形収束性を示している. この結果を拡張することで Lewis et al. [13] は一般の X, Y が真に交わる場合の局所線形収束性を示している. その後, Bauschke et al. [14, 15] や Drusvyatskiy et al. [16] は, X, Y が真に交わることを表す条件を緩めたうえで線形収束性を示している. Noll and Rondepierre [17] は X, Y が接するように交わることを許容する代わりに別の幾何的な仮定をおき, Hölder 正則性と組み合わせることで局所的収束性を示している.

一方で, $X \cap Y = \emptyset$ のときにも適用できる収束解析はあまり知られておらず, 著者の知る限り, Zhu and Li [18] によるもの以外に現時点では存在しない. Zhu and Li [18] は 3 点条件と局所縮小条件を仮定することで大域的収束性を示している. ここで課される 3 点条件は, 上述の Noll and Rondepierre [17] の収束解析で課される幾何的な条件と同種の条件である. Zhu and Li [18] は, $X \cap Y \neq \emptyset$ でなければ成立しない Hölder 正則条件の代わりに, $X \cap Y = \emptyset$ でも成立

¹ ここで, “接するように交わる” は “tangentially intersect” を, “真に交わる” は “nontangentially intersect” をそれぞれ意味する.

する局所縮小条件を課している. この局所縮小条件は補題 2.1 の (iii) を緩めた条件とみなすことができる. また, Zhu and Li [18] は Kurdyka–Lojasiewicz 条件 [19, Section 3] の下で局所的収束性を議論している. Kurdyka–Lojasiewicz 条件は非平滑非凸最適化問題に対するアルゴリズムの解析において近年よく用いられる条件であり, いわゆる Hölder エラーバウンドと本質的に等価である [20, Theorem 5].

このような X, Y が非凸な場合に対するより緩い条件の下での収束解析は現在進行中の研究課題といえる.

4. 応用: 2 錐間の角度の計算

問題 (2) において X, Y が非凸集合である場合の応用としては圧縮センシングにおけるタイトフレームの設計 [18, 21] などが知られている. ここでは著者らが進めている 2 錐間の角度の計算を応用例の一つとして紹介する. なお, 以下では \mathbb{R}^n 上の閉凸錐で $\{0\}$ でも \mathbb{R}^n 全体でもないものを非自明な閉凸錐と呼ぶ.

4.1 問題設定と応用

\mathbb{R}^n 上の二つの非自明な閉凸錐 \mathcal{X}, \mathcal{Y} が与えられたとき, これらの間のなんらかの意味での角度を計算することを考える. いま, $\mathbf{x} \in \mathcal{X} \setminus \{0\}, \mathbf{y} \in \mathcal{Y} \setminus \{0\}$ が与えられたとき, \mathbf{x} と \mathbf{y} のなす角は

$$\angle(\mathbf{x}, \mathbf{y}) := \arccos \frac{\mathbf{x}^\top \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

と定められる. $\angle(\mathbf{x}, \mathbf{y})$ は \mathbf{x} が $\mathcal{X} \setminus \{0\}$ を, \mathbf{y} が $\mathcal{Y} \setminus \{0\}$ を動くとき最小値と最大値をもつ. これらをそれぞれ \mathcal{X} と \mathcal{Y} のなす最小角 $\angle_{\min}(\mathcal{X}, \mathcal{Y})$ および最大角 $\angle_{\max}(\mathcal{X}, \mathcal{Y})$ とする. いま, \mathcal{X}, \mathcal{Y} が錐であることから, \mathbf{x}, \mathbf{y} の動く範囲を $\mathbf{x} \in \mathcal{X} \cap \mathbb{S}^{n-1}, \mathbf{y} \in \mathcal{Y} \cap \mathbb{S}^{n-1}$ に制限してもその最小値, 最大値は変化しない. すなわち, 次が成り立つ:

$$\begin{aligned} \angle_{\min}(\mathcal{X}, \mathcal{Y}) &= \min_{\substack{\mathbf{x} \in \mathcal{X} \cap \mathbb{S}^{n-1}, \\ \mathbf{y} \in \mathcal{Y} \cap \mathbb{S}^{n-1}}} \arccos \mathbf{x}^\top \mathbf{y}, \\ \angle_{\max}(\mathcal{X}, \mathcal{Y}) &= \max_{\substack{\mathbf{x} \in \mathcal{X} \cap \mathbb{S}^{n-1}, \\ \mathbf{y} \in \mathcal{Y} \cap \mathbb{S}^{n-1}}} \arccos \mathbf{x}^\top \mathbf{y}. \end{aligned}$$

$\angle_{\min}(\mathcal{X}, \mathcal{Y})$ を図で表すと図 4 のようになる. なお, $\angle_{\min}(\mathcal{X}, \mathcal{Y}) = \pi - \angle_{\max}(\mathcal{X}, -\mathcal{Y})$ であるため, 最小角と最大角の計算は本質的に同じ問題である.

2 錐間の角度はさまざまな分野で用いられている. 2 錐間の最小角は線形常微分方程式の理論 [22] や順序データの正準相関分析 [23] で用いられている. また, Sogi et al. [24] は 2 錐間の角度を用いて画像の分類を

行っている. 具体的には, 二つの画像集合が与えられたとき, それぞれ画像集合の特徴量を用いて凸多面錐を定め, 対応する凸多面錐の間の角度を用いて二つの画像集合の間の類似度を定めている.

最適化に関連する応用としては, 共正値行列錐や関連する行列錐のなす角度が錐最適化に関する基礎研究の文脈で調べられている. ここで, n 次対称行列 $\mathbf{M} \in \mathbb{R}^{n \times n}$ が共正値であるとは, 任意の $\mathbf{v} \in \mathbb{R}_+^n$ に対して $\mathbf{v}^\top \mathbf{M} \mathbf{v} \geq 0$ となることをいい, n 次共正値行列全体からなる集合は閉凸錐をなすことが知られている. これを n 次共正値行列錐と呼び, \mathcal{C}^n と表記する. \mathcal{C}^n (の双対錐) の上での錐線形最適化問題は, ある種の非凸 2 次最適化問題や組合せ最適化問題を表現できることが知られている [25]. そのため, \mathcal{C}^n の性質を調べることは錐最適化における重要な研究課題となっている. いま, \mathcal{S}_+^n を n 次半正定値行列錐, \mathcal{N}^n を n 次対称非負行列錐としたとき, \mathcal{C}^n はしばしば $\mathcal{S}_+^n + \mathcal{N}^n$ で近似される. 実際, $n \geq 5$ では $\mathcal{C}^n \supsetneq \mathcal{S}_+^n + \mathcal{N}^n$ だが, $n \leq 4$ では $\mathcal{C}^n = \mathcal{S}_+^n + \mathcal{N}^n$ となることが知られている. このような背景の下, Goldberg and Shaked-Moderer [26] は,

- すべての自然数 n に対して $\angle_{\max}(\mathcal{S}_+^n, \mathcal{N}^n) = \angle_{\max}(\mathcal{S}_+^n + \mathcal{N}^n, \mathcal{S}_+^n + \mathcal{N}^n)$ は成り立つか?
- すべての自然数 n に対して $\angle_{\max}(\mathcal{S}_+^n, \mathcal{N}^n) = \angle_{\max}(\mathcal{C}^n, \mathcal{C}^n)$ は成り立つか?

などの問題を提示しており, これらは著者の知る限り現時点でも未解決である. また, このような研究と関連して $\angle_{\max}(\mathcal{S}_+^n, \mathcal{N}^n)$ の計算も進められており, Seeger and Sossa [27] はいくつかの n について $\angle_{\max}(\mathcal{S}_+^n, \mathcal{N}^n)$ の下界を与えている. $\angle_{\max}(\mathcal{S}_+^n, \mathcal{N}^n)$

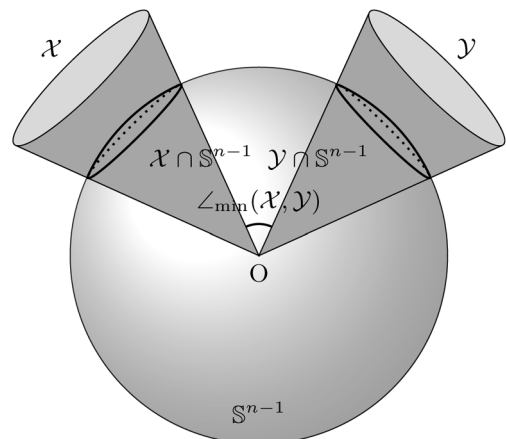


図 4 錐 \mathcal{X}, \mathcal{Y} のなす最小角 $\angle_{\min}(\mathcal{X}, \mathcal{Y})$.

の正確な値を計算する方法は著者の知る限り現時点では存在しない。

\mathcal{X}, \mathcal{Y} が凸多面錐であっても、一般には $\angle_{\max}(\mathcal{X}, \mathcal{Y})$ の計算は容易ではないが、 $\angle_{\max}(\mathcal{X}, \mathcal{Y}) \leq \pi/2$ すなわち $\mathcal{X} \subseteq \mathcal{Y}^*$ であれば、その計算が容易であることが知られている。

定理 4.1 (Orlitzky [28, Corollary 2]). $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_p) \in \mathbb{R}^{n \times p}$, $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_q) \in \mathbb{R}^{n \times q}$ とし、 \mathcal{X}, \mathcal{Y} を $\mathcal{X} = \{\mathbf{X}\boldsymbol{\xi} : \boldsymbol{\xi} \in \mathbb{R}_+^p\}$, $\mathcal{Y} = \{\mathbf{Y}\boldsymbol{\eta} : \boldsymbol{\eta} \in \mathbb{R}_+^q\}$ で定められる凸多面錐とする。さらに、 $\mathcal{X} \subseteq \mathcal{Y}^*$ を仮定する。このとき、 $\angle(\mathbf{x}_i, \mathbf{y}_j) = \angle_{\max}(\mathcal{X}, \mathcal{Y})$ とする $i \in \{1, \dots, p\}$, $j \in \{1, \dots, q\}$ が存在する。

4.2 交互射影法の適用

ここでは、閉凸錐 \mathcal{X} と \mathcal{Y} の間の最小角を求める問題を問題 (2) に帰着して交互射影法を適用する方法について述べる。いま、 \arccos が単調減少関数であることと、 $\mathbf{x}, \mathbf{y} \in \mathbb{S}^{n-1}$ のとき $\|\mathbf{x}\|^2 = \|\mathbf{y}\|^2 = 1$ であることを利用すると、 $\angle_{\min}(\mathcal{X}, \mathcal{Y})$ の計算は $X := \mathcal{X} \cap \mathbb{S}^{n-1}$, $Y := \mathcal{Y} \cap \mathbb{S}^{n-1}$ とした問題 (2) を解くことに帰着できる。しかしながら、前述のとおり X, Y が非凸な場合、問題 (2) を大域的に解くことは一般に困難なので、以下では停留点を求めるために交互射影法を用いることを考える。この問題に交互射影法を適用する場合、次の二つの疑問が生じる：

- $X := \mathcal{X} \cap \mathbb{S}^{n-1}, Y := \mathcal{Y} \cap \mathbb{S}^{n-1}$ への射影は効率よく計算できるか?
- 生成される点列 $\{\mathbf{x}^{(k)}, \mathbf{y}^{(k)}\}$ は停留点に収束するか?

一般に、非凸集合への射影を計算するためには非凸最適化問題を解く必要があり、その大域的最適解を効率よく計算できるかどうかは自明ではない。また、生成される点列の収束性も自明ではない。実際、非凸集合 $\mathcal{X} \cap \mathbb{S}^{n-1}$, $\mathcal{Y} \cap \mathbb{S}^{n-1}$ は有限個の凸集合の和集合として書くことができないため、たとえ $\mathcal{X} \cap \mathcal{Y} \neq \emptyset$ であったとしても定理 3.3 などを用いることはできない。

まずは、 $\mathcal{X} \cap \mathbb{S}^{n-1}, \mathcal{Y} \cap \mathbb{S}^{n-1}$ への射影の効率のよい計算方法について述べる。いま、 $\mathcal{X} \subseteq \mathbb{R}^n$ を直線を含まない閉凸錐とする。このとき、 $\mathcal{X} \cap \mathbb{S}^{n-1}$ は \mathbb{S}^{n-1} 上で測地的凸集合となる [29, Proposition 2]. Ferreira et al. [29] はこのような球面上の測地的凸集合への射影の性質を明らかにしている。具体的には次が成り立つ。

命題 4.2 (Ferreira et al. [29, Proposition 5] より抜

粋). \mathbb{R}^n 上の直線を含まない非自明な閉凸錐 \mathcal{X} および点 $\mathbf{y} \in \mathbb{S}^{n-1}$ に対し、ある $\mathbf{x} \in \mathcal{X} \cap \mathbb{S}^{n-1}$ が存在して $\mathbf{x}^\top \mathbf{y} > 0$ であるとき、 $\text{proj}_{\mathcal{X} \cap \mathbb{S}^{n-1}}(\mathbf{y})$ は単集合である。

命題 4.3 (Ferreira et al. [29, Proposition 6] より抜粋). \mathbb{R}^n 上の直線を含まない非自明な閉凸錐 \mathcal{X} および点 $\mathbf{y} \in \mathbb{S}^{n-1}$ が $\text{proj}_{\mathcal{X}}(\mathbf{y}) \neq \mathbf{0}$ を満たすとき、 $\text{proj}_{\mathcal{X} \cap \mathbb{S}^{n-1}}(\mathbf{y}) = \text{proj}_{\mathcal{X}}(\mathbf{y}) / \|\text{proj}_{\mathcal{X}}(\mathbf{y})\|$ が成り立つ。

命題 4.2 は、球面上の測地的凸集合 $\mathcal{X} \cap \mathbb{S}^{n-1}$ の場合でも \mathcal{X} とのなす角が鋭角となる近傍 $\mathbb{S}^{n-1} \setminus (-\mathcal{X}^*)$ においては補題 2.1 の (i) が成立することを意味する。球面上の測地的凸集合についてはその鋭角近傍においても (ii) 以降は成立しないことがあるが、鋭角近傍条件は重要である。実際、Ferreira et al. [29] では陽に述べられていないが、次が成り立つ。

命題 4.4. \mathbb{R}^n 上の直線を含まない非自明な閉凸錐 \mathcal{X} および点 $\mathbf{y} \in \mathbb{S}^{n-1}$ について、 $\text{proj}_{\mathcal{X}}(\mathbf{y}) \neq \mathbf{0}$ となるための必要十分条件は $\mathbf{y} \notin -\mathcal{X}^*$ である。

命題 4.3, 4.4 より次を得る。このことは \mathcal{X} が凸多面錐のとき Tenenhaus [23, Corollary] によって示されている。ここで錐に課される直線を含まないという条件は外すことができることも知られている [30, Theorem 8.1].

系 4.5. \mathbb{R}^n 上の直線を含まない非自明な閉凸錐 \mathcal{X} および点 $\mathbf{y} \in \mathbb{S}^{n-1}$ が $\mathbf{y} \notin -\mathcal{X}^*$ を満たすとき、 $\text{proj}_{\mathcal{X} \cap \mathbb{S}^{n-1}}(\mathbf{y}) = \text{proj}_{\mathcal{X}}(\mathbf{y}) / \|\text{proj}_{\mathcal{X}}(\mathbf{y})\|$ が成り立つ。

このことは初期点 $\mathbf{x}^{(0)} \in \mathbb{S}^{n-1}$ が $\mathcal{Y} \cap \mathbb{S}^{n-1}$ に十分近ければ、次に示すように交互射影法の中で現れる非凸集合への射影が閉凸錐への射影を正規化することで計算できることを示唆している。いま、 $\mathbf{x}^{(0)} \in \mathcal{X} \cap \mathbb{S}^{n-1}$ が $\mathbf{x}^{(0)} \notin -\mathcal{Y}^*$ を満たすものとする。このとき系 4.5 から、 $\mathbf{y}^{(1)} := \text{proj}_{\mathcal{Y} \cap \mathbb{S}^{n-1}}(\mathbf{x}^{(0)}) = \text{proj}_{\mathcal{Y}}(\mathbf{x}^{(0)}) / \|\text{proj}_{\mathcal{Y}}(\mathbf{x}^{(0)})\|$ とできる。さらに、 $(\mathbf{x}^{(0)})^\top \mathbf{y}^{(1)} > 0$ すなわち $\mathbf{y}^{(1)} \notin -\mathcal{X}^*$ となることから、 $\mathbf{x}^{(1)} := \text{proj}_{\mathcal{X} \cap \mathbb{S}^{n-1}}(\mathbf{y}^{(1)}) = \text{proj}_{\mathcal{X}}(\mathbf{y}^{(1)}) / \|\text{proj}_{\mathcal{X}}(\mathbf{y}^{(1)})\|$ とできる。以下、帰納的に $\mathcal{X} \cap \mathbb{S}^{n-1}, \mathcal{Y} \cap \mathbb{S}^{n-1}$ への射影を \mathcal{X}, \mathcal{Y} への射影の

正規化で計算できる。これらの計算は閉凸錐 \mathcal{X}, \mathcal{Y} への射影の計算が容易であれば容易である。たとえば、 \mathbb{R}_+^n への射影は各要素の 0 への切り上げで、 \mathcal{S}_+^n への射影は各固有値の 0 への切り上げで計算できる。また、凸多面錐への射影は凸 2 次最適化問題を解くことで計算できる。このとき、前の反復の最適解を初期解として用いるなどのホットスタートが有効である可能性がある。

次に、生成される点列の収束性について述べる。ここでも系 4.5 は重要な役割を果たす。具体的には、 $\text{proj}_{\mathcal{X} \cap \mathcal{S}^{n-1}}(\mathbf{y}) = \text{proj}_{\mathcal{X}}(\mathbf{y}) / \|\text{proj}_{\mathcal{X}}(\mathbf{y})\|$, $\text{proj}_{\mathcal{Y} \cap \mathcal{S}^{n-1}}(\mathbf{x}) = \text{proj}_{\mathcal{Y}}(\mathbf{x}) / \|\text{proj}_{\mathcal{Y}}(\mathbf{x})\|$ を用いることで、前述した Zhu and Li [18] の収束解析における 3 点条件および局所縮小条件が成立することを示すことができる。したがって、 \mathcal{X}, \mathcal{Y} が $\mathcal{X} \setminus (-\mathcal{Y}^*) \neq \emptyset$ を満たすとき、任意の $\mathbf{x}^{(0)} \in (\mathcal{X} \setminus (-\mathcal{Y}^*)) \cap \mathcal{S}^{n-1}$ を初期点とした交互射影法は停留点に収束することがいえる。

ここでの仮定 $\mathcal{X} \setminus (-\mathcal{Y}^*) \neq \emptyset$ が成り立たないとき、 $\mathcal{X} \subseteq (-\mathcal{Y}^*)$ が成り立つ。したがって、 \mathcal{X}, \mathcal{Y} が凸多面錐であれば、 $\mathcal{X} \cap (-\mathcal{Y}^*) = \emptyset$ のときは定理 4.1 が適用できて、交互射影法を使うまでもなく $\angle_{\min}(\mathcal{X}, \mathcal{Y}) = \pi - \angle_{\max}(\mathcal{X}, -\mathcal{Y})$ を厳密に計算できる。また、 \mathcal{X}, \mathcal{Y} の少なくとも一方が凸多面錐で $\mathcal{X} \setminus (-\mathcal{Y}^*) \neq \emptyset$ を満たすとき、 $\mathbf{x}^{(0)} \in (\mathcal{X} \setminus (-\mathcal{Y}^*)) \cap \mathcal{S}^{n-1}$ を満たす点 $\mathbf{x}^{(0)}$ を求める方法を示唆する次の命題が成り立つ。

命題 4.6. $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_p) \in \mathbb{R}^{n \times p}$ とし、 \mathcal{X} を $\mathcal{X} = \{\mathbf{X}\boldsymbol{\xi} : \boldsymbol{\xi} \in \mathbb{R}_+^p\}$ で定められる凸多面錐とし、 \mathcal{Y} を \mathbb{R}^n 上の直線を含まない非自明な閉凸錐とする。このとき、 $\mathcal{X} \setminus (-\mathcal{Y}^*) \neq \emptyset$ となるための必要十分条件は、 $\mathbf{x}_i^\top \text{proj}_{\mathcal{Y}}(\mathbf{x}_i) > 0$ となる $i \in \{1, \dots, p\}$ が存在することであり、このとき $\mathbf{x}_i \in \mathcal{X} \setminus (-\mathcal{Y}^*)$ が成り立つ。特に、 $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_q) \in \mathbb{R}^{n \times q}$, $\mathcal{Y} = \{\mathbf{Y}\boldsymbol{\eta} : \boldsymbol{\eta} \in \mathbb{R}_+^q\}$ のときの必要十分条件は、 $\mathbf{x}_i^\top \mathbf{y}_j > 0$ となる $i \in \{1, \dots, p\}$, $j \in \{1, \dots, q\}$ が存在することであり、このとき $\mathbf{x}_i \in \mathcal{X} \setminus (-\mathcal{Y}^*)$ が成り立つ。

謝辞 本稿を執筆する機会を与えてくださった高野祐一先生、RAMP 数理最適化シンポジウムでの発表の機会を与えてくださった福田エレン秀美先生、本研究を開始する契機を与えてくださった日野英逸先生に感謝する。なお、本研究は日本学術振興会科学研究費補助金若手研究 JP19K15247 の助成を受けている。

参考文献

- [1] R. Escalante and M. Raydan, *Alternating Projection Methods*, SIAM, 2011.
- [2] H. H. Bauschke and J. M. Borwein, “On projection algorithms for solving convex feasibility problems,” *SIAM Review*, **38**, pp. 367–426, 1996.
- [3] J. von Neumann, *Functional Operators, Volume II*, Princeton University Press, 1950.
- [4] I. Halperin, “The product of projection operators,” *Acta Scientiarum Mathematicarum (Szeged)*, **23**, pp. 96–99, 1962.
- [5] W. Cheney and A. A. Goldstein, “Proximity maps for convex sets,” In *Proceedings of the American Mathematical Society*, **10**, pp. 448–450, 1959.
- [6] D. P. Bertsekas, A. Nedić and A. E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, 2003.
- [7] N. Aronszajn, “Theory of reproducing kernels,” *Transactions of the American Mathematical Society*, **68**, pp. 337–404, 1950.
- [8] L. G. Gubin, B. T. Polyak and E. V. Raik, “The method of projections for finding the common point of convex sets,” *USSR Computational Mathematics and Mathematical Physics*, **27**, pp. 1–24, 1967.
- [9] H. H. Bauschke and J. M. Borwein, “On the convergence of von Neumann’s alternating projection algorithm for two sets,” *Set-Valued Analysis*, **1**, pp. 185–212, 1993.
- [10] P. L. Combettes and H. J. Trussell, “Method of successive projection for finding a common point of sets in metric spaces,” *Journal of Optimization Theory and Applications*, **67**, pp. 487–507, 1990.
- [11] H. H. Bauschke and D. Noll, “On cluster points of alternating projections,” *Serdica Mathematical Journal*, **39**, pp. 355–364, 2013.
- [12] A. S. Lewis and J. Malick, “Alternating projections on manifolds,” *Foundations of Computational Mathematics*, **33**, pp. 216–234, 2008.
- [13] A. S. Lewis, D. R. Luke and J. Malick, “Local linear convergence for alternating and averaged nonconvex projections,” *Foundations of Computational Mathematics*, **9**, pp. 485–513, 2009.
- [14] H. H. Bauschke, D. R. Luke, H. M. Phan and X. Wang, “Restricted normal cones and the method of alternating projections: Theory,” *Set-Valued and Variational Analysis*, **21**, pp. 431–473, 2013.
- [15] H. H. Bauschke, D. R. Luke, H. M. Phan and X. Wang, “Restricted normal cones and the method of alternating projections: Applications,” *Set-Valued and Variational Analysis*, **21**, pp. 475–501, 2013.
- [16] D. Drusvyatskiy, A. D. Ioffe and A. S. Lewis, “Transversality and alternating projections for nonconvex sets,” *Foundations of Computational Mathematics*, **15**, pp. 1637–1651, 2015.
- [17] D. Noll and A. Rondepierre, “On local convergence of the method of alternating projections,” *Foundations of Computational Mathematics*, **16**, pp. 425–455, 2016.
- [18] Z. Zhu and X. Li, “Convergence analysis of alternating projection method for nonconvex sets,” *arXiv:1802.03889v2*, 2019.
- [19] H. Attouch and J. Bolte, “On the convergence of the proximal algorithm for nonsmooth functions involving

- analytic features,” *Mathematical Programming*, **116**, pp. 5–16, 2009.
- [20] J. Bolte, T. P. Nguyen, J. Peypouquet and B. W. Suter, “From error bounds to the complexity of first-order descent methods for convex functions,” *Mathematical Programming*, **165**, pp. 471–507, 2017.
- [21] J. A. Tropp, I. S. Dhillon, R. W. Heath, Jr. and T. Strohmer, “Designing structured tight frames via an alternating projection method,” *IEEE Transactions on Information Theory*, **51**, pp. 425–455, 2005.
- [22] D. G. Obert, “The angle between two cones,” *Linear Algebra and its Applications*, **144**, pp. 63–70, 1991.
- [23] M. Tenenhaus, “Canonical analysis of two convex polyhedral cones and applications,” *Psychometrika*, **53**, pp. 503–524, 1988.
- [24] N. Sogi, T. Nakayama and K. Fukui, “A method based on convex cone model for image-set classification with CNN features,” In *Proceedings of 2018 International Joint Conference on Neural Networks*, pp. 1–8, 2018.
- [25] S. Burer, “On the copositive representation of binary and continuous nonconvex quadratic programs,” *Mathematical Programming*, **120**, pp. 479–495, 2009.
- [26] F. Goldberg and N. Shaked-Monderer, “On the maximal angle between copositive matrices,” *Electronic Journal of Linear Algebra*, **27**, pp. 837–850, 2014.
- [27] A. Seeger and D. Sossa, “Critical angles between two convex cones II. Special cases,” *TOP*, **24**, pp. 66–87, 2016.
- [28] M. Orliczky, “When a maximal angle among cones is nonobtuse,” *Computational and Applied Mathematics*, **39**, Article number 83, 2020.
- [29] O. P. Ferreira, A. N. Iusem and S. Z. Németh, “Projections onto convex sets on the sphere,” *Journal of Global Optimization*, **57**, pp. 663–676, 2013.
- [30] H. H. Bauschke, M. N. Bui and X. Wang, “Projecting onto the intersection of a cone and a sphere,” *SIAM Journal on Optimization*, **28**, pp. 2158–2188, 2018.