

球面上の交互射影法を用いた 2 錐間の角度の計算

田中未来 数理・推論研究系 数理最適化グループ, 統計的機械学習研究センター 准教授

1 問題設定

- \mathbb{R}^n 上の 2 つの非自明な閉凸錐 \mathcal{X}, \mathcal{Y} に対し, 次のように定義される \mathcal{X}, \mathcal{Y} 間の最小角 $\angle_{\min}(\mathcal{X}, \mathcal{Y})$ を計算する問題を考える (図 1):

$$\angle_{\min}(\mathcal{X}, \mathcal{Y}) := \min_{\substack{x \in \mathcal{X} \setminus \{0\} \\ y \in \mathcal{Y} \setminus \{0\}}} \arccos \frac{x^\top y}{\|x\| \|y\|} = \min_{\substack{x \in \mathcal{X} \cap \mathbb{S}^{n-1} \\ y \in \mathcal{Y} \cap \mathbb{S}^{n-1}}} \arccos x^\top y.$$

(ここで $\mathbb{S}^{n-1} = \{u \in \mathbb{R}^n : \|u\| = 1\}$ は \mathbb{R}^n 上の単位球面.)

- 応用: 線形常微分方程式の理論 (Obert, 1991), 順序データの正準相関分析 (Tenenhaus, 1988), 画像分類 (Sogi et al., 2018), 凸錐の幾何 (Goldberg and Shaked-Monderer, 2014; Seeger and Sossa, 2016b), etc.
- 1 次の最適性条件を満たす解の列挙は指数個の一般化固有値問題を解く必要あり (Seeger and Sossa, 2016a,b; Orlitzky, 2020).

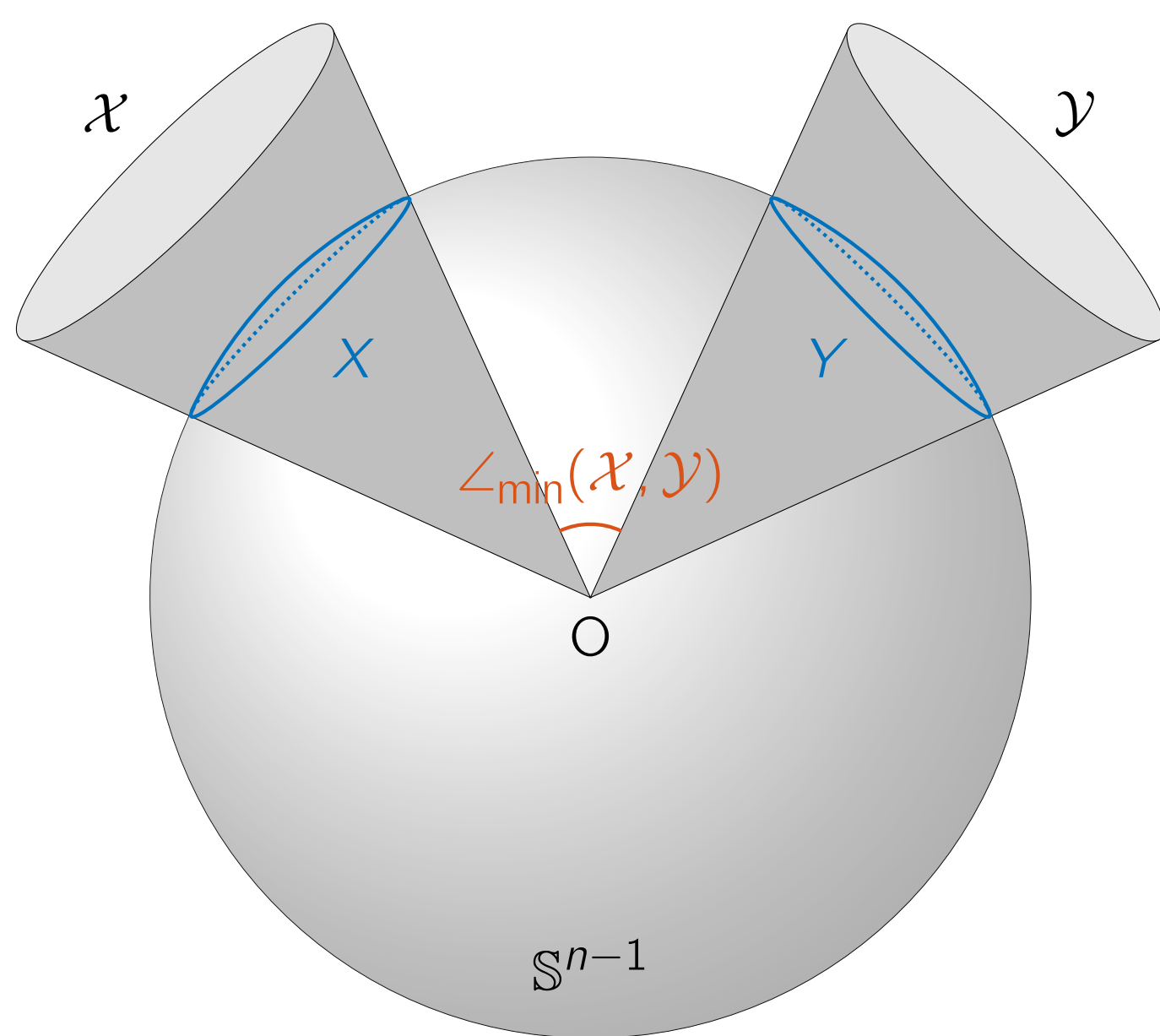


図 1: \mathcal{X}, \mathcal{Y} 間の最小角

2 交互射影法の適用: 射影の計算法と収束性

- \mathcal{X}, \mathcal{Y} 間の最小角をなす x, y は次の問題の最適解:

$$\min_{\substack{x \in \mathcal{X} \cap \mathbb{S}^{n-1} \\ y \in \mathcal{Y} \cap \mathbb{S}^{n-1}}} \|x - y\|.$$

- 一般に $X, Y \subseteq \mathbb{R}^n$ が与えられたときの問題 $\min_{x \in X, y \in Y} \|x - y\|$ は次のような交互射影法で解ける:

$$\begin{aligned} y^{(k+1)} &\in \text{proj}_Y(x^{(k)}), \\ x^{(k+1)} &\in \text{proj}_X(y^{(k+1)}) \quad (k = 0, 1, \dots). \end{aligned} \quad (1)$$

- 交互射影法の収束性に関する既存研究:

- X, Y が凸な場合のサーベイ: Bauschke and Borwein (1996); Escalante and Raydan (2011)
- X, Y が非凸で $X \cap Y \neq \emptyset$ だが点列が収束しない例: Bauschke and Noll (2013).
- X, Y が非凸で X と Y が真に交わるときの局所的収束性: Lewis et al. (2009); Bauschke et al. (2013b,a); Drusvyatskiy et al. (2015); Noll and Rondepierre (2016).
- X, Y が非凸で $X \cap Y = \emptyset$ を許容したときの大域的収束性と局所的収束性のための十分条件: Zhu and Li (2019).

- 研究上の疑問点:

- 非凸集合 $\mathcal{X} \cap \mathbb{S}^{n-1}, \mathcal{Y} \cap \mathbb{S}^{n-1}$ への射影は効率よく計算できるか?
- $X = \mathcal{X} \cap \mathbb{S}^{n-1}, Y = \mathcal{Y} \cap \mathbb{S}^{n-1}$ の場合, $X \cap Y = \emptyset$ でも交互射影法 (1) によって生成される点列は収束するか?

2.1 射影の計算

- 次の定理が成り立つ (Bauschke et al., 2018, Theorem 8.1):

定理 1. \mathbb{R}^n 上の非自明な閉凸錐 \mathcal{X} および点 $y \in \mathbb{S}^{n-1}$ が $y \notin -\mathcal{X}^*$ を満たすとき, $\text{proj}_{\mathcal{X} \cap \mathbb{S}^{n-1}}(y) = \text{proj}_{\mathcal{X}}(y) / \|\text{proj}_{\mathcal{X}}(y)\|$ が成り立つ.

(ここで $\mathcal{X}^* := \{s \in \mathbb{R}^n : x^\top s \geq 0 \ (x \in \mathcal{X})\}$ は \mathcal{X} の双対錐.)

- 定理 1 は, 初期点 $x^{(0)} \in \mathbb{S}^{n-1}$ が $\mathcal{Y} \cap \mathbb{S}^{n-1}$ に十分近ければ, 交互射影法 (1) を次のように書き換えることができることを意味する:

$$\begin{aligned} y^{(k+1)} &:= \text{proj}_{\mathcal{Y} \cap \mathbb{S}^{n-1}}(x^{(k)}) = \frac{\text{proj}_{\mathcal{Y}}(x^{(k)})}{\|\text{proj}_{\mathcal{Y}}(x^{(k)})\|}, \\ x^{(k+1)} &:= \text{proj}_{\mathcal{X} \cap \mathbb{S}^{n-1}}(y^{(k+1)}) = \frac{\text{proj}_{\mathcal{X}}(y^{(k+1)})}{\|\text{proj}_{\mathcal{X}}(y^{(k+1)})\|} \quad (k = 0, 1, \dots). \end{aligned} \quad (2)$$

- 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.

2.2 点列の収束性

- 一般に, 次の定理が成り立つ (Zhu and Li, 2019, Theorem 1)

定理 2. 半代数的集合 X, Y に対する交互射影法 (1) により生成される点列 $\{(x^{(k)}, y^{(k)})\}$ が有界で, 任意の $k \geq k_0$ に対して $x^{(k)} \in \bar{X}, y^{(k)} \in \bar{Y}$ となる $\bar{X} \subseteq X, \bar{Y} \subseteq Y, k_0 \in \mathbb{Z}_{++}$ が存在して, 次を満たす:

- 3 点条件: 定数 $\alpha > 0$ と関数 $\phi_\alpha : Y \times Y \rightarrow \mathbb{R}_+$ が存在して, 次が成り立つ: 任意の $y, y' \in Y$ に対して, $\phi_\alpha(y, y') \geq \alpha \|y - y'\|^2$. 任意の $y \in \bar{Y}, \tilde{x} \in \bar{X}, \tilde{y} \in \text{proj}_Y(\tilde{x})$ に対して, $\phi_\alpha(y, \tilde{y}) + \|\tilde{x} - \tilde{y}\|^2 \leq \|\tilde{x} - y\|^2$.
- X の \bar{Y} に関する局所縮小条件: 定数 $\epsilon > 0, \beta > 0$ が存在して, $\|\tilde{y} - y\| \leq \epsilon$ であるような任意の $y, \tilde{y} \in \bar{Y}$ に対して, $\|\text{proj}_X(\tilde{y}) - \text{proj}_X(y)\| \leq \beta \|\tilde{y} - y\|$ が成り立つ.

このとき, $\{(x^{(k)}, y^{(k)})\}$ は停留点に収束する.

- 定理 1 を用いると, $X = \mathcal{X} \cap \mathbb{S}^{n-1}, Y = \mathcal{Y} \cap \mathbb{S}^{n-1}$ について, 定理 2 の 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, 結果として次が成り立つ:

定理 3. \mathbb{R}^n 上の非自明な閉凸錐 \mathcal{X}, \mathcal{Y} が $\mathcal{X} \setminus (-\mathcal{Y}^*) \neq \emptyset$ を満たすものとする. このとき, 任意の $x^{(0)} \in \mathcal{X} \setminus (-\mathcal{Y}^*) \cap \mathbb{S}^{n-1}$ を初期点とした交互射影法 (2) により生成される点列は停留点に収束する.

3 カーネル化

- \mathcal{H} を特徴空間, $\phi : \mathbb{R}^n \rightarrow \mathcal{H}$ を特徴写像とし, \mathcal{H} 上の 2 つの凸多面錐 $\mathcal{X}^\phi := \text{cone}\{\phi(x_1), \dots, \phi(x_p)\}, \mathcal{Y}^\phi := \text{cone}\{\phi(y_1), \dots, \phi(y_q)\}$ の最小角の計算は次の問題に帰着できる:

$$\min_{\substack{x \in \mathcal{X}^\phi \cap \mathbb{S}_{\mathcal{H}} \\ y \in \mathcal{Y}^\phi \cap \mathbb{S}_{\mathcal{H}}}} \|x - y\|_{\mathcal{H}}.$$

- $x, y \in \mathcal{H}$ を保持することや既存の収束解析を適用することが困難.
- 次の問題に帰着し, $\xi \in \mathbb{R}_+^p, \eta \in \mathbb{R}_+^q$ を保持することになると交互射影法 (2) をカーネル化できる:

$$\min_{\substack{\xi \in \mathbb{R}_+^p, X^\phi \xi \in \mathbb{S}_{\mathcal{H}} \\ \eta \in \mathbb{R}_+^q, Y^\phi \eta \in \mathbb{S}_{\mathcal{H}}}} \|X^\phi \xi - Y^\phi \eta\|_{\mathcal{H}}^2.$$

- カーネル化した交互射影法をブロック座標降下法とみなすと, 生成される点列の任意の集積点が停留点であることがいえる (Tseng, 2001).