2021年6月18日　統計数理研究所　オープンハウス

2 錐間の大域的最小角を求めるための分枝限定法

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

- $\mathbb{R}^d$ 上の 2 つの非自明な閉凸錐 $\mathcal{X}$, $\mathcal{Y}$ に対し, $\mathcal{X}$, $\mathcal{Y}$ 間の最小角 $\angle_{\min}(\mathcal{X},\mathcal{Y})$ を次のように定義する (図 1):

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

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

- 等価な問題:

$$\omega^* := \min_{\substack{x\in\mathcal{X}\cap\mathbb{S}^{d-1}\\ y\in\mathcal{Y}\cap\mathbb{S}^{d-1}}} \|x-y\|^2. \qquad \text{(NCP)}$$

- 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).
- (NCP) に対する交互射影法の効率のよい計算と収束性 (Tanaka, 2020).

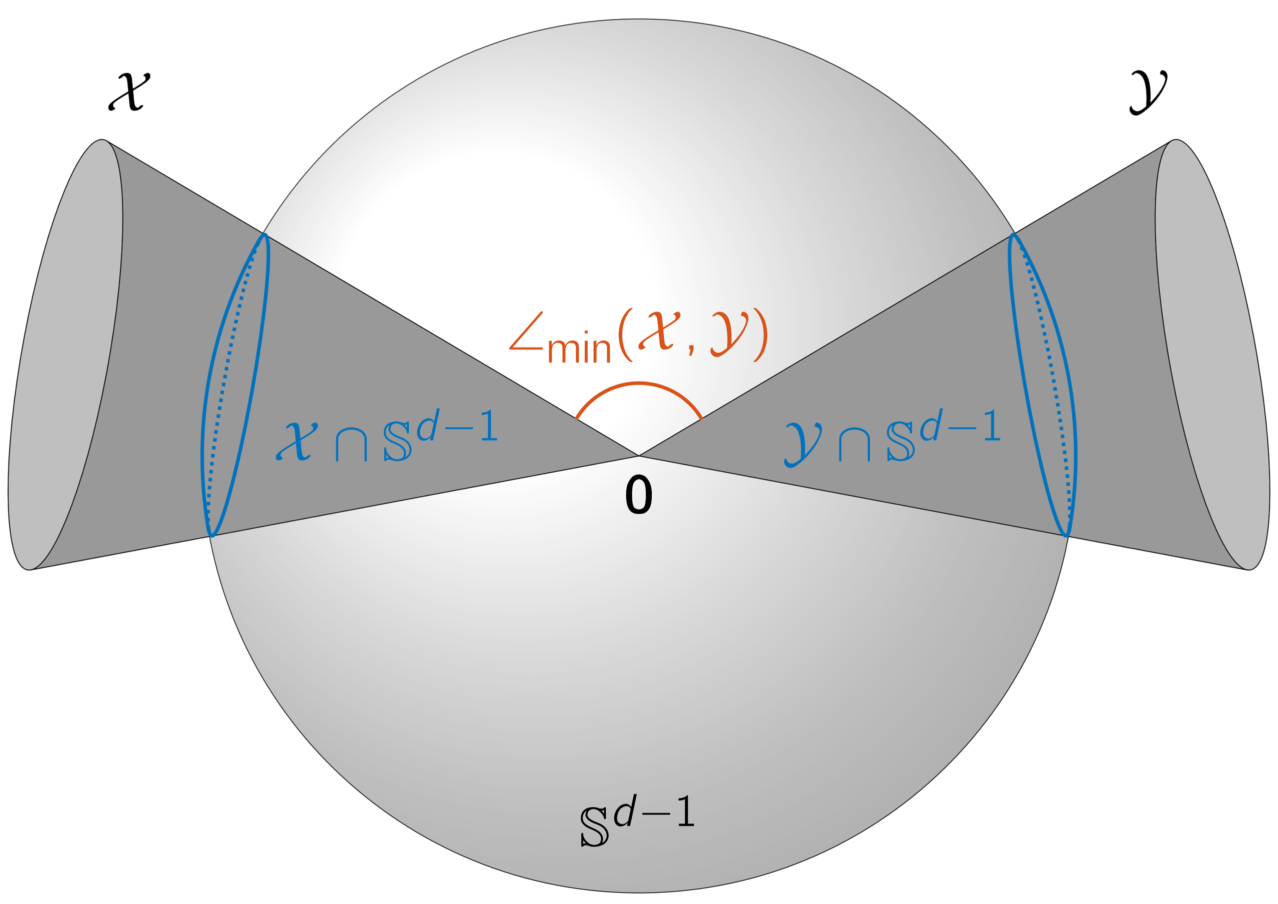

図 1: $\mathcal{X}$, $\mathcal{Y}$ 間の最小角

2　凸緩和: 多面錐の場合

- $\mathcal{X}$, $\mathcal{Y}$ が多面錐の場合を考える.
- 具体的には, $\|x_1\| = \cdots = \|x_m\| = \|y_1\| = \cdots = \|y_n\| = 1$ を満たす $X := (x_1,\dots,x_m) \in \mathbb{R}^{d\times m}$, $Y := (y_1,\dots,y_n) \in \mathbb{R}^{d\times n}$ を用いて $\mathcal{X} := \{X\xi : \xi \in \mathbb{R}^m_+\}$, $\mathcal{Y} := \{Y\eta : \eta \in \mathbb{R}^n_+\}$ と書けるものとする.

定理 1. (NCP) の実行可能領域をその凸包で置き換えた凸緩和問題は次の問題と等価:

$$\begin{aligned}&\text{minimize} && \|x-y\|^2\\ &\text{subject to} && x = X\xi,\ \xi\in\mathbb{R}^m_+,\ \mathbf{1}^\top\xi \ge 1,\\ &&& y = Y\eta,\ \eta\in\mathbb{R}^n_+,\ \mathbf{1}^\top\eta\ge 1.\end{aligned} \qquad \text{(QP)}$$

定理 2. $(\hat{x},\hat{y})$ を (QP) の最適解としたとき, $(\hat{x}/\|\hat{x}\|, \hat{y}/\|\hat{y}\|)$ は (NCP) の実行可能解で, $\|\hat{x}\|^{-1}\|\hat{y}\|^{-1}$ 近似解.

3　分枝限定法: 単体的な多面錐の場合

- 前節で考えた多面錐の場合で, 特に $m = n = d$ であるときを考える.
- (QP) を利用して (NCP) を大域的に解くための分枝限定法をアルゴリズム 1 に示す.
- 単体上での凹関数の大域的最小化のための分枝限定法 (Horst and Tuy, 1996) に基づく.

アルゴリズム 1 (NCP) を大域的に解くための分枝限定法

根ノードでの計算: $(\mathcal{X}^0,\mathcal{Y}^0) := (\mathcal{X},\mathcal{Y})$ に対応する (QP) を解き, ω^* の下界 $\underline{\omega}(\mathcal{X}^0,\mathcal{Y}^0)$, (NCP) の実行可能解 (x^0,y^0), ω^* の上界 $\bar{\omega}^0$ を求め, $\mathcal{N}^0 := \{(\mathcal{X}^0,\mathcal{Y}^0)\}$, $k := 0$ とする.
for $k := 0,1,\dots$ **do**
　if $\mathcal{N}^k = \emptyset$ **do**
　　最適値として $\bar{\omega}^k$ を, 最適解として (x^k,y^k) を出力して終了.
　end if
　分枝操作: $(\mathcal{X}^k,\mathcal{Y}^k) \in \mathcal{N}^k$ を選び, $\mathcal{X}^k$, $\mathcal{Y}^k$ の少なくとも一方を $\bigcup_i \mathcal{X}^k_i$, $\bigcup_j \mathcal{Y}^k_j$ と分割し, $\mathcal{N}^{k+1/2} := \mathcal{N}^k \setminus \{(\mathcal{X}^k,\mathcal{Y}^k)\} \cup \bigcup_{i,j}\{(\mathcal{X}^k_i,\mathcal{Y}^k_j)\}$ と更新.
　子ノードでの計算: 各 $(\mathcal{X}^k_i,\mathcal{Y}^k_j)$ に対応する (QP) を解き, ω^* の下界 $\underline{\omega}(\mathcal{X}^k_i,\mathcal{Y}^k_j)$, (NCP) の実行可能解 (x^k_{ij},y^k_{ij}), ω^* の上界 $\bar{\omega}(\mathcal{X}^k_i,\mathcal{Y}^k_j)$ を求め, 暫定最適解 (x^k,y^k), 暫定最適値 $\bar{\omega}^k$ を更新.
　限定操作: $\mathcal{N}^{k+1} := \{(\mathcal{X},\mathcal{Y}) \in \mathcal{N}^{k+1/2} : \underline{\omega}(\mathcal{X},\mathcal{Y}) \le \bar{\omega}^k\}$ と更新.
end for

- アルゴリズム 1 が有限列を生成するときの出力の最適性は明らか.
- アルゴリズム 1 が無限列を生成するときの収束解析を行なった.

定理 3. アルゴリズム 1 が無限列 $\{\mathcal{X}^k\}$, $\{\mathcal{Y}^k\}$ を生成するとき, 生成された任意のフィルターが悉皆的であるとする. このとき, $\{\bar{\omega}^k\}$ は ω^* に収束する.

- 無限列 $\{\mathcal{X}^k\}$ の無限部分列 $\{\mathcal{X}^{k_l}\}$ について, 任意の $l = 1,2,\dots$ に対して $\mathcal{X}^{k_{l+1}} \subseteq \mathcal{X}^{k_l}$ が成り立つとき, $\{\mathcal{X}^{k_l}\}$ をフィルターとよぶ.
- 半直線に収束するフィルターを悉皆的なフィルターとよぶ.
- 生成される任意のフィルターが悉皆的となる分割として最長辺分割などが知られている.
- 実用的には許容誤差 $\epsilon > 0$ を用いた限定操作 $\mathcal{N}^{k+1} := \{(\mathcal{X},\mathcal{Y}) \in \mathcal{N}^{k+1/2} : \underline{\omega}(\mathcal{X},\mathcal{Y}) \le \bar{\omega}^k - \epsilon\}$ を行なう. これにより, 有限回の計算で ϵ 近似解を出力するアルゴリズムとできる.

4　一般の閉凸錐の場合への拡張

- アルゴリズム 1 を一般の閉凸錐 $\mathcal{X}$, $\mathcal{Y}$ に適用できるように拡張する.
- 方針: $\mathcal{X}$, $\mathcal{Y}$ とは別に補助的な単体的錐 $\tilde{\mathcal{X}}$, $\tilde{\mathcal{Y}}$ を用意する.
- 具体的には, (x^*,y^*) を (NCP) の大域的最適解の 1 つとしたとき, $\|x_1\| = \cdots = \|x_d\| = \|y_1\| = \cdots = \|y_d\| = 1$ を満たす $X := (x_1,\dots,x_d) \in \mathbb{R}^{d\times d}$, $Y := (y_1,\dots,y_d) \in \mathbb{R}^{d\times d}$ を用いて書ける $\tilde{\mathcal{X}} := \{X\xi : \xi \in \mathbb{R}^d_+\} \ni x^*$, $\tilde{\mathcal{Y}} := \{Y\eta : \eta\in\mathbb{R}^d_+\} \ni y^*$ をとる.
- (QP) と同様の凸緩和として次の問題を考える:

$$\begin{aligned}&\text{minimize} && \|x-y\|^2\\ &\text{subject to} && x\in\mathcal{X},\ x = X\xi,\ \xi\in\mathbb{R}^d_+,\ \mathbf{1}^\top\xi \ge 1,\\ &&& y\in\mathcal{Y},\ y = Y\eta,\ \eta\in\mathbb{R}^d_+,\ \mathbf{1}^\top\eta\ge 1.\end{aligned} \qquad \text{(QCP)}$$

これは $\tilde{\mathcal{X}}\cap\mathbb{S}^{d-1}$, $\tilde{\mathcal{Y}}\cap\mathbb{S}^{d-1}$ をそれらの凸包で置き換えた凸緩和問題と等価.

- アルゴリズム 1 において, (QP) の代わりに (QCP) を解き, $\mathcal{X}$, $\mathcal{Y}$ の代わりに $\tilde{\mathcal{X}}$, $\tilde{\mathcal{Y}}$ を分割すれば, $\mathcal{X}$, $\mathcal{Y}$ が一般の閉凸錐であるような (NCP) に対するアルゴリズムとでき, 同様の収束解析が成り立つ.
- (QCP) は錐最適化問題なので頻繁には解きたくない. $\tilde{\mathcal{X}} \subseteq \mathcal{X}$, $\tilde{\mathcal{Y}} \subseteq \mathcal{Y}$ を検出できれば (QCP) における制約 $x\in\mathcal{X}$, $y\in\mathcal{Y}$ が冗長になるが, その効率のよい検出方法は自明でない.