

# L1 正則化項を持つ無制約最適化問題に対するニュートン型近接勾配法の上界近似を用いた加速と切り替え規則について

成島 康史<sup>1</sup>, 辺 浩<sup>1</sup><sup>1</sup> 慶應義塾大学,

e-mail : narushima(at)ae.keio.ac.jp

## 1 はじめに

本講演では,  $\ell_1$  正則化項を含む無制約最適化問題

$$\min_{x \in \mathbb{R}^n} f(x) := g(x) + h(x) \quad (1)$$

に対する反復法を考える. ただし,  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  は  $L$ -平滑な関数,  $h(x) = \lambda \|x\|_1$  ( $\lambda > 0$ ) とし,  $\|\cdot\|_1$  を  $\ell_1$  ノルムとする. このような問題に対する反復法として, 近接勾配法がよく知られており, その加速法として, ニュートン型近接勾配法が提案されている [1]. ニュートン型近接勾配法では, 点列  $\{x_k\}$  を

$$x_{k+1} = \operatorname{argmin}_{x \in \mathbb{R}^n} \left\{ g(x_k) + \nabla g(x_k)^\top (x - x_k) + \frac{1}{2\alpha_k} \|x - x_k\|_{B_k}^2 + h(x) \right\} \quad (2)$$

によって更新する. ここで,  $x_k \in \mathbb{R}^n$  は  $k$  回目の近似解であり,  $\alpha_k > 0$  はステップ幅と呼ばれる. また,  $B_k$  は  $\nabla^2 f(x_k)$ , またはその近似行列であり,  $\|x\|_B = \sqrt{x^\top B x}$  は正定値対称行列  $B$  を重み行列とした重み付きノルムである. ニュートン型近接勾配法は  $g$  の 2 次近似を用いているため, 局所的に速い収束が期待できるが, その一方で毎回の反復で部分問題 (2) を解く手間がネックとなる. そこで, 本講演では,  $\ell_1$  ノルムを構成する絶対値関数を 2 次関数で近似した部分問題を考え, その解を (2) の近似解として用いる方法を提案する. さらに, 近似行列  $B_k$  の生成にメモリーレス準ニュートン法を用いることで, 近似解を容易に計算できる方法を与える. 一方, 近似解を用いた場合には, 大域的な収束性の保証がないため, (大域的収束性が保証された方法との) 切り替え規則を組み込んだアルゴリズムを提案する.

## 2 部分問題の上界近似について

今回, 我々は  $\ell_1$  ノルムを構成する絶対値関数  $\phi(z) = |z|$  を  $\hat{\phi}(z) = \frac{c^\dagger}{2} z^2 + \frac{c}{2}$  によって近似することを考える. ただし,  $c^\dagger$  は  $c = 0$  のときは  $c^\dagger = 0$ ,  $c \neq 0$  のときは  $c^\dagger = 1/c$  であるような一般化逆数であるとする.

図 2 は  $\phi(z) = |z|$  と  $c = 4$  とした場合の  $\hat{\phi}(z)$  のグラフを表している.  $\phi(z)$  と  $\hat{\phi}(z)$  は  $z = \pm c$  で接しており,  $c$  が十分小さいときには  $\hat{\phi}$  は区間  $[-c, c]$  では  $\phi(z)$  の上界近似となっていることがわかる. したがって,  $c = (x_k)_i$  として,  $\|x\|_1 \approx \sum_{i=1}^n \hat{\phi}((x_k)_i)$  という上界近似を用いることとする. ただし,  $(x_k)_i$  を  $x_k$  の第  $i$  成分とする. このとき, (2) は

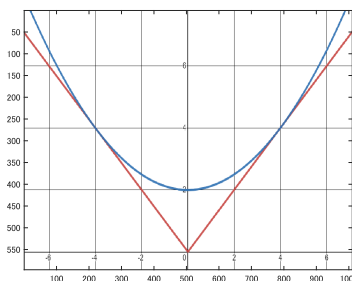


図 1.  $\phi(z)$  と  $\hat{\phi}(z)$  のグラフ ( $c = 4$ )

$$x_{k+1} = \operatorname{argmin}_{x \in \mathbb{R}^n} \left\{ g(x_k) + \nabla g(x_k)^\top (x - x_k) + \frac{1}{2\alpha_k} \|x - x_k\|_{B_k}^2 + \lambda \sum_{i=1}^n \frac{|(x_k)_i|^\dagger}{2} x^2 + \frac{|(x_k)_i|}{2} \right\}$$

によって計算されることになる。近似行列  $B_k$  が正定値対称のとき、上記の部分問題の目的関数は強凸関数となるため、部分問題は容易に解くことができ、

$$x_{k+1} = (B_k + \alpha_k \lambda D_k)^{-1} (B_k x_k - \alpha_k \nabla g(x_k)) \quad (3)$$

が得られる。ただし、 $D_k$  は  $(i, i)$  成分が  $|(x_k)_i|^\dagger$  であるような対角行列である。さらに、近似行列  $B_k$  として、メモリーレス BFGS 公式

$$B_k = I - \frac{s_{k-1} s_{k-1}^\top}{s_{k-1}^\top s_{k-1}} + \frac{y_{k-1} y_{k-1}^\top}{s_{k-1}^\top y_{k-1}}, \quad s_{k-1} = x_k - x_{k-1}, \quad y_{k-1} = \nabla g(x_k) - \nabla g(x_{k-1})$$

を用いることで、 $(B_k + \alpha_k \lambda D_k)^{-1}$  を陽に表すことができる（紙面に関係上、詳細は割愛する）。なお、その表現を用いた場合には、(3) を数回の内積のみで計算可能であるため、部分問題の計算の手間を大幅に削減することができる。

### 3 切り替え規則を用いたアルゴリズムの提案

まず、一般論として、(2) の近似解を  $x_{k+1}$  として用いた場合の大域的収束性を考える。今回は、以下を満たすステップ幅  $\alpha_k$  を用いるものとする：

$$\epsilon_2 \leq \alpha_k \leq \frac{1}{\theta_k + \epsilon_3}, \quad \frac{\frac{L}{2} + \epsilon_1}{m_k} \leq \theta_k \leq \bar{\theta}. \quad (4)$$

ただし、 $\epsilon_1, \epsilon_2, \epsilon_3$  は小さな正の定数、 $\bar{\theta}$  は十分大きな正の定数、 $L$  を  $\nabla g$  のリプシッツ定数とする。また、 $m_k$  は正定値対称な近似行列  $B_k$  の最小固有値であるとする。 $x_{k+1}$  は (2) の近似解なので、最適性条件には残差  $r_k$  が存在することとなる、つまり、 $r_k \in \nabla g(x_k) + \frac{1}{\alpha_k} B_k (x_{k+1} - x_k) + \partial h(x_{k+1})$  となる。もし、 $r_k = 0$  の場合は、 $x_k$  は (2) の厳密解と一致する。一方、大域的収束性を保証するために、 $x_{k+1}$ （とその時の残差  $r_k$ ）は以下の条件を満たすような (2) の近似解であるとする：

$$r_k^\top (x_{k+1} - x_k) \leq \left( \frac{1}{\alpha_k} - \theta_k \right) \|x_{k+1} - x_k\|_{B_k}^2, \quad \|r_k\| \leq R \|x_{k+1} - x_k\|. \quad (5)$$

このとき、以下の定理が得られる。

**定理 1.** 問題 (1) を考える。ただし、 $h$  を凸関数とする。さらに、正の定数  $m$  と  $M$  が存在して、すべての  $k$  に対して  $m\|v\|^2 \leq v^\top B_k v \leq M\|v\|^2$  (for  $\forall v \in \mathbb{R}^n$ ) が成り立つとする。 $\{x_k\}$  を (4)–(5) を満たすニュートン型近接勾配法によって生成される点列とする。このとき、もし  $\{x_k\}$  が有界ならば、その任意の集積点は (1) の停留点である。さらに、もし  $g$  が強凸関数ならば、 $\{x_k\}$  は (1) の唯一解に  $R-1$  次収束する。

ステップ幅として (4) を満たす  $\alpha_k$  を選択し、点列の更新に (3) を用いた場合、必ずしも (5) を満たすとは限らない。よって、ここでは、(3) を既存のニュートン型近接勾配法の加速ととらえて、条件 (5) を満たす場合のみ (3) を用い、条件を満たさない場合には、大域的収束性が保証された通常のニュートン型近接勾配法への切り替えを行うことで、大域的な収束性が保証されたアルゴリズムが構築される。

### 参考文献

- [1] J.D. Lee, Y. Sun, and M.A. Saunders, Proximal Newton-type methods for minimizing composite functions, SIAM Journal on Optimization, Vol. 24 (2014), 1420–1443.

# Riemannian generalized conditional gradient methods

CHEN Kangming<sup>1</sup>, FUKUDA Ellen Hidemi<sup>1</sup>

<sup>1</sup>Kyoto University

e-mail : kangming@amp.i.kyoto-u.ac.jp

## 1 Introduction

The Conditional Gradient (CG) Method, also known as the Frank-Wolfe algorithm, is a fundamental technique in constrained optimization, developed by Marguerite Frank and Philip Wolfe in 1956. It is particularly useful for minimizing convex functions over complex convex constraint sets where projection is computationally challenging. Recent explorations of the CG method on Riemannian manifolds [1] and the Generalized Conditional Gradient (GCG) method have been detailed in various studies [2]. Recent explorations of the CG method on Riemannian manifolds [1] and the GCG method have been detailed in various studies [2]. This work investigates the GCG method on Riemannian manifolds, focusing on three types of step size and the implementation of subproblem solutions.

## 2 The generalized conditional gradient method on Riemannian manifolds

We consider the following unconstrained vector optimization problem:

$$\begin{aligned} \min \quad & F(x) := f(x) + g(x) \\ \text{s.t.} \quad & x \in \mathcal{X}, \end{aligned} \tag{1}$$

where  $\mathcal{X} \subseteq \mathcal{M}$  is a compact and geodesically convex set and  $F: \mathcal{M} \rightarrow \mathbb{R}$  is a composite function where  $f: \mathcal{M} \rightarrow \mathbb{R}$  is continuously differentiable and  $g: \mathcal{M} \rightarrow \mathbb{R}$  is a proper, closed, geodesically convex and lower semicontinuous function with compact domain.

$\mathcal{M}$  is a manifold endowed with a Riemannian metric  $(\eta_x, \sigma_x) \mapsto \langle \eta_x, \sigma_x \rangle_x \in \mathbb{R}$ , where  $\eta_x$  and  $\sigma_x$  are tangent vectors in the tangent space of  $\mathcal{M}$  at  $x$ . The tangent space of a manifold  $\mathcal{M}$  at  $x \in \mathcal{M}$  is denoted as  $T_x\mathcal{M}$ , and the tangent bundle of  $\mathcal{M}$  is denoted as  $T\mathcal{M} := \{(x, d) \mid d \in T_x\mathcal{M}, x \in \mathcal{M}\}$ . The norm of  $\eta \in T_x\mathcal{M}$  is defined as  $\|\eta\|_x := \sqrt{\langle \eta, \eta \rangle_x}$ . For a map  $F: \mathcal{M} \rightarrow \mathcal{N}$  between two manifolds  $\mathcal{M}$  and  $\mathcal{N}$ ,  $DF(x): T_x\mathcal{M} \rightarrow T_{F(x)}\mathcal{N}$  denotes the derivative of  $F$  at  $x \in \mathcal{M}$ . The Riemannian gradient  $\text{grad}f(x)$  of a function  $f: \mathcal{M} \rightarrow \mathbb{R}$  at  $x \in \mathcal{M}$  is defined as a unique tangent vector at  $x$  satisfying  $\langle \text{grad}f(x), \eta \rangle_x = Df(x)[\eta]$  for any  $\eta \in T_x\mathcal{M}$ .

Compared to the update iterations  $x^{k+1} = x^k + t_k d_k$  in Euclidean space, the process in the Riemannian setting is generalized as  $x^{k+1} = R_{x^k}(t_k d^k)$ , where  $d^k$  is a descent direction,  $t_k$  is a step size, and  $R$  is a retraction that projects points from the tangent space of the manifold onto the manifold itself. This adaptation accounts for the curved geometry of the manifold, ensuring that the iterates remain on the manifold throughout the optimization process.

In Euclidean space, the GCG method aims to minimize composite functions  $f(x) = h(x) + g(x)$  over a convex set  $\mathcal{X}$ . At each iteration  $k$ , the method solves the subproblem:

$x^{k+1} = \arg \min_{s \in \mathcal{X}} \langle \nabla h(x^k), s \rangle + g(s)$ . Following this idea, we can formulate our algorithm.

---

**Algorithm 1** Riemannian GCG method

---

**Step 0. Initialization:**

Choose  $x^0 \in \mathcal{X}$  and initialize  $k = 0$ .

**Step 1. Compute the search direction:**

Compute an optimal solution  $p(x^k)$  and the optimal value  $\theta(x^k)$  as

$$p(x^k) = \arg \min_{u \in \mathcal{X}} \langle \text{grad} f(x^k), R_{x^k}^{-1}(u) \rangle_{x^k} + g(u) - g(x^k),$$

$$\theta(x^k) = \langle \text{grad} f(x^k), R_{x^k}^{-1}(u) \rangle_{x^k} + g(u) - g(x^k).$$

Define the search direction by  $d(x^k) = R_{x^k}^{-1}(p(x^k))$ .

**Step 2. Compute the Step size:**

Compute the step size  $\lambda_k$ .

**Step 3. Update:**

Update the current iterate  $x^{k+1} = R_{x^k}(\lambda_k d(x^k))$ .

**Step 4. Convergence check:**

If a convergence criteria is met, stop; otherwise, set  $k = k + 1$  and return to Step 1.

---

In this work, we consider three types of step sizes.

**Armijo step size:** Let  $\zeta \in (0, 1)$ , choose the step size  $\lambda_k$  satisfies  $F(R_{x^k}(\lambda_{k_\ell} d(x^k))) \leq F(x^k) + \zeta \lambda_{k_\ell} \theta(x^k)$ .

**Adaptive step size:**  $\lambda_k := \min \{1, -\theta(x^k)/(LD^2(p(x^k), x^k))\}$ , where  $D(x, y)$  is the geodesic distance between  $x$  and  $y$  and  $L$  is the smoothness constant of  $f$ .

**Diminishing step size:**  $\lambda_k := 2/(k + 2)$ .

### 3 Convergence results

The convergence results are established as follows:

**Theorem 1.** Let  $x^*$  be an optimal point of the problem and  $\{x^k\}$  be generated by Algorithm 1 with adaptive or diminishing step size. Assume that  $f$  is  $L$ -smooth and  $g$  is geodesically convex. Then  $\{x^k\}$  satisfies  $F(x^k) - F(x^*) = O(1/k)$ .

**Theorem 2.** Let  $\{x^k\}$  be generated by Algorithm 1 with the Armijo step size. Assume that  $f$  is  $L$ -smooth and  $g$  is geodesically convex and lower semicontinuous. Then  $\lim_{k \rightarrow \infty} \theta(x^k) = 0$  and every limit point  $x^* \in \mathcal{X}$  of the sequence  $\{x^k\}$  is a stationary point.

### 参考文献

- [1] M. Weber, S. Sra. Riemannian optimization via Frank-Wolfe methods. *Mathematical Programming*, 2023, 199(1): 525-556.
- [2] P. B. Assunção, O. P. Ferreira, L. F. Prudente. A generalized conditional gradient method for multiobjective composite optimization problems. *Optimization*, 2023: 1-31.

# 対称錐上の共正值錐および完全正值錐の面的構造

西島 光洋<sup>1,2</sup>, ロウレンソ 武流野 フィゲラ<sup>2</sup>

<sup>1</sup> 東京工業大学, <sup>2</sup> 統計数理研究所

e-mail: nishijima.m.ae@m.titech.ac.jp

## 1 概要

非凸な構造および組合せ構造を持つ様々な最適化問題は、対称錐から定まる一般化された意味での共正值錐および完全正值錐——対称錐上の共正值錐および完全正值錐を制約に持つ錐線型計画問題として等価に表現できる。本発表では、対称錐上の共正值錐および完全正值錐の面的構造に関する結果を報告する。まず、対称錐の面を用いて、対称錐上の共正值錐および完全正值錐の面を構成する。特に、対称錐の次元が2以上のとき、対応する共正值錐が面露出という性質を決して満たさないことを示す。また、このようにして構成された、共正值錐および完全正值錐のある面のすべての部分面を列挙し、その次元と露出性を決定する。

## 2 準備

$K$  を閉凸錐とする。 $K$  の非空な凸部分錐  $F$  が次の条件

$$\forall a, b \in K \ (a + b \in F \implies a, b \in F)$$

を満たすとき、 $F$  を  $K$  の面と呼ぶ。面のうち、 $K$  とその支持超平面との共通部分として書ける面を露出面と呼ぶ。そして、 $K$  のすべての面が露出しているとき、 $K$  を面露出という。

$(\mathbb{V}, \bullet)$  を有限次元実内積空間とし、 $\mathcal{S}(\mathbb{V})$  を  $\mathbb{V}$  上の自己随伴な線型変換全体からなる集合とする。 $\mathbb{V}$  に含まれる閉錐  $\mathbb{K}$  に対して、 $\mathcal{S}(\mathbb{V})$  の部分集合

$$\text{COP}(\mathbb{K}) := \{\mathcal{A} \in \mathcal{S}(\mathbb{V}) \mid x \bullet \mathcal{A}(x) \geq 0 \ \forall x \in \mathbb{K}\}$$

$$\text{CP}(\mathbb{K}) := \left\{ \sum_{i=1}^k a_i \otimes a_i \mid a_i \in \mathbb{K} \ \forall i = 1, \dots, k \right\}$$

をそれぞれ、 $\mathbb{K}$  上の共正值錐および完全正值錐と呼ぶ。（ここで内積空間  $(\mathbb{V}, \bullet)$  に対して、テンソル空間  $\mathbb{V} \otimes \mathbb{V}$  と  $\mathbb{V}$  上の線型変換全体の空間とは、元  $a \otimes b$  を写像  $x \mapsto (b \bullet x)a$  に移す写像によって、同一視できることに注意せよ。）

特に本発表では、 $\mathbb{K}$  が対称錐の場合に焦点を当てる。対称錐とは自己双対な等質錐のことである。非負象限錐  $\mathbb{R}_+^n := \{x \in \mathbb{R}^n \mid x_i \geq 0 \ \forall i = 1, \dots, n\}$  および2次錐（Lorentz 錐） $\{x \in \mathbb{R}^n \mid x_1 \geq \sqrt{x_2^2 + \dots + x_n^2}\}$  が対称錐の例として知られている。対称錐はEuclid 的 Jordan 代数と呼ばれる代数系と一対一の関係にある [1]。本研究でも、Euclid 的 Jordan 代数を介して、対称錐上の共正值錐および完全正值錐の面的構造を調査する。

## 3 対称錐の面から誘導される共正值錐および完全正值錐の面

我々の研究の主要な貢献として、対称錐の面を用いて、対称錐上の共正值錐の露出していない面を構成した。

**定理 1** ([2, Corollary 3.2]).  $\mathbb{K}$  を実内積空間  $\mathbb{V}$  における対称錐として、 $\mathbb{F}$  を  $\mathbb{K}$  の面とする。そして  $\mathcal{P}_{\text{span } \mathbb{F}}: \mathbb{V} \rightarrow \text{span } \mathbb{F}$  を  $\text{span } \mathbb{F}$  への直交射影とし、 $\mathcal{P}_{\text{span } \mathbb{F}}^*$  をその随伴写像とする。このとき、次の

集合

$$\{\mathcal{P}_{\text{span } \mathbb{F}}^* \circ \mathcal{A} \circ \mathcal{P}_{\text{span } \mathbb{F}} \mid \mathcal{A} \in \mathcal{COP}(\mathbb{F})\} \quad (1)$$

は  $\mathcal{COP}(\mathbb{K})$  の面である．特に， $\mathbb{K}$  の面  $\mathbb{F}$  が  $\{0\}$  でも  $\mathbb{K}$  自身でもないとき， $\mathcal{COP}(\mathbb{K})$  の面 (1) は露出していない．

特に，対称錐の 1 次元の面  $\mathbb{R}_+c$  に対応する，共正值錐の面 (1) は  $\mathbb{R}_+c \otimes c$  と書いて，対称錐  $\mathbb{K}$  の次元が 2 以上のとき，この面は露出していない [2, Theorem 3.3]. この結果は，任意の  $i = 1, \dots, n$  に対して， $e_i \in \mathbb{R}^n$  を第  $i$  成分が 1 でそれ以外の成分が 0 であるベクトルとしたとき， $\mathbb{R}_+e_ie_i^\top$  が  $\mathcal{COP}(\mathbb{R}_+^n)$  の露出していない面であるという事実 [3, Theorem 4.4] の拡張になっている．ただし， $\mathbb{K}$  が非負象限錐の場合と一般の対称錐の場合とで，定理 1 の証明の難しさは大きく異なる． $\mathbb{K}$  が一般の対称錐の場合，集合 (1) が  $\mathcal{COP}(\mathbb{K})$  の面であることを示すために，11 通りもの場合分けを必要とした一方で， $\mathbb{K}$  が非負象限錐の場合は 2 通りの場合分けだけで証明できる．

以下の定理で示すように，完全正值錐に対しても，定理 1 と類似の結果を示すことができる．

**定理 2.**  $\mathbb{K}$  を実内積空間における閉凸錐として， $\mathbb{F}$  を  $\mathbb{K}$  の面とする．このとき， $\mathcal{CP}(\mathbb{F})$  は  $\mathcal{CP}(\mathbb{K})$  の露出面である．

定理 2 は，錐  $\mathbb{K}$  の面が共正值錐ないし完全正值錐の面を誘導するという点で，定理 1 と似ている．しかし，定理 1 が成り立つ（ことが現時点で証明されている） $\mathbb{K}$  は対称錐に限定されているのに対して，定理 2 は一般の閉凸錐  $\mathbb{K}$  に対して成り立つ．また，定理 1 では共正值錐の非露出面が一般に得られるのに対して，定理 2 では完全正值錐の露出面が得られる．

## 4 2 次錐上の共正值錐および完全正值錐の面的構造

定理 1 および定理 2 で得られた面以外に，対称錐上の共正值錐および完全正值錐が他にどのような面を持つのか，そしてその面の露出性はどうかといった疑問が自然に生じる．その疑問に部分的に答えるために我々は，2 次錐上の共正值錐および完全正值錐に着目する．Euclid 的 Jordan 代数の理論ならびに定理 1 および定理 2 を用いると，次元 2 以上の対称錐  $\mathbb{K}$  上の共正值錐および完全正值錐は，（同型を無視すると）2 次錐上の共正值錐および完全正值錐をそれぞれ面に持つことが分かる．したがって，2 次錐上の共正值錐および完全正值錐の面的構造を調べることで，一般の対称錐上の共正值錐および完全正值錐の面的構造に関する知見を得ることができる．この事実を動機として，2 次錐上の共正值錐および完全正值錐の部分面を列挙し，その部分面の次元と露出性を決定する．より詳しい結果については当日報告する予定である．

**謝辞** 本研究は，日本学術振興会科学研究費助成事業（JP22KJ1327 および JP21H03398）の助成を受けた．

## 参考文献

- [1] J. Faraut and A. Korányi, Analysis on Symmetric Cones, Clarendon Press, 1994.
- [2] M. Nishijima and B.F. Lourenço, Non-facial exposedness of copositive cones over symmetric cones, arXiv.2402.12964 (2024).
- [3] P.J.C. Dickinson, Geometry of the copositive and completely positive cones, J. Math. Anal. Appl., 380 (1) (2011), 377–395.

# Bottleneck type DEA model reformulation under production trade-offs

堀 篤史<sup>1</sup>, 関谷和之<sup>1</sup><sup>1</sup> 成蹊大学

e-mail : kazuyuki-sekitani@st.seikei.ac.jp

## 1 Introduction

An envelopment Data Envelopment Analysis (DEA) model under production trade-offs is reduced into an equivalent dual form, a multiplier DEA model with weight restriction of input and/or output. Incorporating production trade-offs causes a practical issue of target setting by envelopment DEA models. This study shows that practical issue can be solved by DEA model reformulation with desirable properties of an efficiency measure.

## 2 Existing DEA models

In DEA a production possibility set is  $T := \{(\mathbf{x}, \mathbf{y}) \mid \text{inputs } \mathbf{x} \in \mathbb{R}_+^m \text{ can produce outputs } \mathbf{y} \in \mathbb{R}_+^s\}$ . For the production possibility set  $T$ , the strongly efficient frontier  $\partial^s(T)$  of  $T$  is defined by

$$\{(\mathbf{x}, \mathbf{y}) \in T \mid (\mathbf{x}, -\mathbf{y}) \geq (\mathbf{x}', -\mathbf{y}') \text{ and } (\mathbf{x}, -\mathbf{y}) \neq (\mathbf{x}', -\mathbf{y}') \implies (\mathbf{x}', \mathbf{y}') \notin T\}.$$

An efficiency measure is a mapping  $F : T \cap \mathbb{R}_{++}^{m+s} \rightarrow [0, 1]$ . Consider three axioms as desirable efficiency measures defined on the full space of inputs and outputs as follows:

Indication of Efficiency (I): For all  $(\mathbf{x}, \mathbf{y}) \in T \cap \mathbb{R}_{++}^{m+s}$ ,  $F(\mathbf{x}, \mathbf{y}) = 1$  if and only if  $(\mathbf{x}, \mathbf{y}) \in \partial^s(T)$ .

Strong monotonicity (SM): For all pairs  $(\mathbf{x}, \mathbf{y}) \in T \cap \mathbb{R}_{++}^{m+s}$  and  $(\mathbf{x}', \mathbf{y}') \in T \cap \mathbb{R}_{++}^{m+s}$  satisfying  $(\mathbf{x}, -\mathbf{y}) \leq (\mathbf{x}', -\mathbf{y}')$  and  $(\mathbf{x}, -\mathbf{y}) \neq (\mathbf{x}', -\mathbf{y}')$ ,  $F(\mathbf{x}, \mathbf{y}) > F(\mathbf{x}', \mathbf{y}')$ .

For any  $(\mathbf{x}, \mathbf{y}) \in T$  and  $(\mathbf{x} - \mathbf{d}^-, \mathbf{y} + \mathbf{d}^+) \in \partial^s(T)$ , we define

$$I^-(\mathbf{x}, \mathbf{d}^-) := \left\{ \frac{x_1 - d_1^-}{x_1}, \dots, \frac{x_m - d_m^-}{x_m} \right\}, \quad I^+(\mathbf{y}, \mathbf{d}^+) := \left\{ \frac{y_1}{y_1 + d_1^+}, \dots, \frac{y_s}{y_s + d_s^+} \right\}$$

and  $I(\mathbf{x}, \mathbf{y}, \mathbf{d}^-, \mathbf{d}^+) := I^-(\mathbf{x}, \mathbf{d}^-) \cup I^+(\mathbf{y}, \mathbf{d}^+)$ .

For any  $\mathbf{z} \in \mathbb{R}_+^n$ , the generalized mean with exponent  $t$  of  $\mathbf{z}$  is

$$M_t(\mathbf{z}) := \begin{cases} \left( \frac{1}{n} \sum_{i=1}^n z_i^t \right)^{1/t} & \text{if } t > 0 \text{ or } \mathbf{z} \in \mathbb{R}_{++}^n \\ 0 & \text{otherwise} \end{cases}$$

The SBM DEA model [1] and Russell graph measure DEA model [2] can be reduced to

$$\min \left\{ \frac{M_1(I^-(\mathbf{x}, \mathbf{d}^-))}{M_{-1}(I^+(\mathbf{y}, \mathbf{d}^+))} \mid (\mathbf{x} - \mathbf{d}^-, \mathbf{y} + \mathbf{d}^+) \in T, (\mathbf{d}^-, \mathbf{d}^+) \geq (\mathbf{0}, \mathbf{0}) \right\} \quad (1)$$

$$\min \left\{ M_1(I(\mathbf{x}, \mathbf{y}, \mathbf{d}^-, \mathbf{d}^+)) \mid (\mathbf{x} - \mathbf{d}^-, \mathbf{y} + \mathbf{d}^+) \in T, (\mathbf{d}^-, \mathbf{d}^+) \geq (\mathbf{0}, \mathbf{0}) \right\} \quad (2)$$

respectively.

Let  $(\mathbf{x}_j, \mathbf{y}_j) \in \mathbb{R}_{++}^{m+s}$  for  $j = 1, \dots, n$  be an observed input–output data, and let  $P$  ( $Q$ ) be a matrix representing production trade-offs between inputs (output). Podinovski [3] provided the following production possibility set:

$$T := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_{++}^{m+s} \left| \sum_{j=1}^n \lambda_j \mathbf{x}_j - P\boldsymbol{\alpha} \leq \mathbf{x}, \sum_{j=1}^n \lambda_j \mathbf{y}_j + Q\boldsymbol{\beta} \geq \mathbf{y}, \boldsymbol{\lambda} \geq \mathbf{0}, \boldsymbol{\alpha} \geq \mathbf{0}, \boldsymbol{\beta} \geq \mathbf{0} \right. \right\}. \quad (3)$$

**Theorem 1** Consider the production possibility set  $T$  of (3). Suppose  $(\mathbf{x}, \mathbf{y}) \in T \cap \mathbb{R}_{++}^{m+s}$  and let  $(\boldsymbol{\lambda}^*, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \mathbf{d}^{-*}, \mathbf{d}^{+*})$  be an optimal solution of (1) or (2). Then,  $(\mathbf{x} - \mathbf{d}^{-*}, \mathbf{y} + \mathbf{d}^{+*}) \in \partial^s(T)$ .

In DEA,  $(\mathbf{x} - \mathbf{d}^{-*}, \mathbf{y} + \mathbf{d}^{+*})$  is called the target of  $(\mathbf{x}, \mathbf{y})$ . If  $(\mathbf{0}, \mathbf{y}') \in T$  and  $\mathbf{y}' \neq \mathbf{0}$ , then there exists  $(\mathbf{x}, \mathbf{y}) \in T \cap \mathbb{R}_{++}^{m+s}$  with the target  $(\mathbf{0}, \mathbf{y} + \mathbf{d}^{+*})$  by an optimal solution of (1) or (2). Moreover, both efficiency measure of (1) and that of (2) violate (SM).

### 3 Bottleneck type reformulation

Let  $M_{-\infty}(\mathbf{z}) = \min\{z_1, \dots, z_n\}$  and  $M_{\infty}(\mathbf{z}) = \max\{z_1, \dots, z_n\}$  and replace  $M_1$ ,  $M_{-1}$  and  $T$  of (1) and (2) with  $M_{-\infty}$ ,  $M_{\infty}$  and  $\partial^s(T)$ , respectively, then DEA models (1) and (2) are reformulated to

$$\max \left\{ \frac{M_{-\infty}(I^-(\mathbf{x}, \mathbf{d}^-))}{M_{\infty}(I^+(\mathbf{y}, \mathbf{d}^+))} \mid (\mathbf{x} - \mathbf{d}^-, \mathbf{y} + \mathbf{d}^+) \in \partial^s(T), (\mathbf{d}^-, \mathbf{d}^+) \geq (\mathbf{0}, \mathbf{0}) \right\} \quad (4)$$

$$\max \{ M_{-\infty}(I(\mathbf{x}, \mathbf{y}, \mathbf{d}^-, \mathbf{d}^+)) \mid (\mathbf{x} - \mathbf{d}^-, \mathbf{y} + \mathbf{d}^+) \in \partial^s(T), (\mathbf{d}^-, \mathbf{d}^+) \geq (\mathbf{0}, \mathbf{0}) \}, \quad (5)$$

respectively. The optimal value of (4) is denoted by  $F_S(\mathbf{x}, \mathbf{y})$  and that of (5) is denoted by  $F_R(\mathbf{x}, \mathbf{y})$ . Both  $F_S$  and  $F_R$  have desirable properties for some matrices  $P$  and  $Q$  as follows:

**Theorem 2** Assume

$$\{\mathbf{v}^\top \in \mathbb{R}_+^m \mid \mathbf{v}P \leq \mathbf{0}\} \subseteq \mathbb{R}_{++}^m \cup \{\mathbf{0}\} \text{ and} \quad (6)$$

$$\emptyset \neq \{\mathbf{u}^\top \in \mathbb{R}_+^s \mid \mathbf{u}Q \leq \mathbf{0}, \mathbf{u} \neq \mathbf{0}\} \subseteq \mathbb{R}_{++}^s, \quad (7)$$

then  $\partial^s(T) \cap \mathbb{R}_{++}^{m+s} = \partial^w(T) \cap \mathbb{R}_{++}^{m+s}$ . Moreover for any  $(\mathbf{x}, \mathbf{y}) \in T \cap \mathbb{R}_{++}^{m+s}$ , any target of (4) and (5) is a positive vector,  $F_S(\mathbf{x}, \mathbf{y}) = F_R(\mathbf{x}, \mathbf{y}) \in (0, 1]$  and  $F_S$  ( $F_R$ ) satisfies (I) and (SM).

Check of (6) and (7) and optimal computing (4) and (5) for all observed input–output data can be solved by at most  $(2n + 1)(m + s)$  linear programming problems.

### 4 Conclusion

The bottleneck type reformulation DEA model (4), which is equivalent to (5), has the desirable properties of target and efficiency measure for  $P$  and  $Q$  satisfying (6) and (7).

### 参考文献

- [1] K. Tone, “A slacks-based measure of efficiency in data envelopment analysis,” *European journal of operational research*, vol. 130, no. 3, pp. 498–509, 2001.
- [2] R. Färe, S. Grosskopf, and C. K. Lovell, *The measurement of efficiency of production*, vol. 6. Springer Science & Business Media, 1985.
- [3] V. V. Podinovski, “Production trade-offs and weight restrictions in data envelopment analysis,” *Journal of the Operational Research Society*, vol. 55, no. 12, pp. 1311–1322, 2004.