数理解析研究所講究録 第2108巻 2019年 70-79

Optimization over the Efficient Set of a Linear Multiobjective Programming: Algorithm and Applications

Kuan Lu
Department of Industrial Engineering and Management
Tokyo Institute of Technology
Shinji Mizuno
Department of Industrial Engineering and Management
Tokyo Institute of Technology
Jianming Shi
Department of Business Economics
Tokyo University of Science

1 Introduction

Let $\boldsymbol{x} \in \mathbb{R}^n$, and $p \geqq 2$. Consider the multiobjective linear programming (MOLP) problem given below:

$$\text{MOLP} \left| \begin{array}{ll} \text{minimize} & \boldsymbol{C}\boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} \leqq \boldsymbol{b}, \boldsymbol{x} \geqq \boldsymbol{0}, \end{array} \right. \tag{1}$$

where $\boldsymbol{C} \in \mathbb{R}^{p\times n}$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{0}$ is a zero vector. Let $\boldsymbol{X} := \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{A}\boldsymbol{x} \leqq \boldsymbol{b}, \boldsymbol{x} \geqq \boldsymbol{0}\}$ be the set of feasible solutions. We assume that $\boldsymbol{X}$ is nonempty, and the set $\{\boldsymbol{C}\boldsymbol{x} | \boldsymbol{x} \in \boldsymbol{X}\}$ is bounded from below which ensures that the MOLP is solvable.

To define the optimal solution for the MOLP problem, we have the following definition.

Definition 1.1 (Efficient Solution)**.** *An $\boldsymbol{x}' \in \boldsymbol{X}$ is called an efficient solution of MOLP if there is no $\boldsymbol{x} \in \boldsymbol{X}$ such that $\boldsymbol{C}\boldsymbol{x} \leqq \boldsymbol{C}\boldsymbol{x}'$ and $\boldsymbol{C}\boldsymbol{x} \neq \boldsymbol{C}\boldsymbol{x}'$.*

Let $\boldsymbol{E}$ denote the set of efficient solutions of MOLP. Then, the optimization over the efficient set (OE) is defined as the following

$$\text{OE} \left| \begin{array}{ll} \text{minimize} & \phi(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \boldsymbol{E}, \end{array} \right. \tag{2}$$

where $\phi(\boldsymbol{x})$ is a real function.

In 1972, Philip [1] considered the problem of optimizing some real function over the efficient set (OE) of a multiobjective linear programming. There are many works on the OE problem and since generally the efficient set is nonconvex, the problem plays an important role in multiobjective programming and global optimization. Yamamoto's survey [2] classified the existing algorithms [1, 3, 4, 5, 6, 7, 8]. Since the efficient set can be described as the difference of two convex set, algorithms based on DC programming [9] have been raised. In fact, all these approaches are based in one way or another way on DC optimization or extensions [10]. Because that the dimension of the outcome space of problem is usually smaller than that of the decision space, algorithms that search an optimal solution in the outcome space has been proposed in Benson's works [11, 12]. Liu and Ehrgott [13] proposed primal and dual approaches for the problem which are also based this idea. Recently in Sun [14]'s work, the first mixed integer programming (MIP) approach has been proposed to solve the OE problem of linear cases under several conditions. In these previous researches, the running time usually grow rapidly with both the number of objective functions and the number of constraints of the original multiobjective programming.

In this paper, the approach of Lu et al. would be generalized to solve the OE problem. Based on their result, we propose a necessary sufficient condition of the efficient solution of a multiobjective linear programming which also can be seen as an application of Wendell and Lee [15]. Then with the condition, we can use a MIP approach to solve the OE problem. The proposed approach relax the condition and reduce the binary variables to solve the problem more efficiently.

In Section 2, we will show Sun' main result and ours, then we will compare our approach with the previous work's approach. A sketch proof will be given in Section 3. The necessary and sufficient condition for the efficient solution of a multiobjective linear problem will be established by using an auxiliary problem, then the formulation of the OE problem will be transformed into a MIP problems. In Section 4, we will introduce the minimum maximal flow problem which is known to be a typically OE problem [16, 17, 18], and with the MIP approach, the problem can be efficiently solved. In Section 5, we will introduce another problem called the least distance problem in data envelopment analysis, show that it is a OE problem, and provide a MIP approach. In Section 6, the conclusion will be given.

2 Mixed Integer Programming Approach

In this section, after introducing Sun's result [14] on the problem (2), we will show our main result and make a comparison of the results. Sun's result obtained under the following two assumptions on the problems (1) and (2):

Assumption 2.1. *For the problem (1), there is an $i \in \{1, 2, \ldots, p\}$ such that $\boldsymbol{c}_i^\top \boldsymbol{x}$ is one-to-one on $\boldsymbol{X}$, where $\boldsymbol{c}_i$ is the ith row of the matrix $\boldsymbol{C}$.*

Assumption 2.2. *The problem (2) has an optimal solution $\boldsymbol{x}^*$.*

We note that Assumption 2.1 is very strong, because it is not true if the dimension of the feasible region $\boldsymbol{X}$ is bigger than 1. Hence, Sun's result is valid only when the dimension of the feasible region $\boldsymbol{X}$ is at most 1. Assumption 2.2 is rather standard and used in the most analysis of the problem (2). Please note that the problem (1) is feasible and the efficient set $\boldsymbol{E}$ is nonempty under Assumption 2.2.

We state Sun's result [14] as the next proposition.

Proposition 2.1. *(Sun [14]) Under Assumptions 2.1 and 2.2, a solution of the following mixed integer programming (MIP) problem is a solution of (2)*

$$
\left|\begin{array}{ll}
\text{minimize} & \phi(\boldsymbol{x}) \\
\text{subject to} & \boldsymbol{s} = \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{\beta} \end{pmatrix} - \begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{C}_{(i)} \end{pmatrix} \boldsymbol{x}, \\
& \boldsymbol{r} = \boldsymbol{c}_i + \begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{C}_{(i)} \end{pmatrix}^\top \boldsymbol{u}, \\
& \boldsymbol{\beta} = \boldsymbol{C}_{(i)} \boldsymbol{y}, \quad \boldsymbol{A}\boldsymbol{y} \leqq \boldsymbol{b}, \\
& \boldsymbol{0} \leqq \boldsymbol{u} \leqq \theta \boldsymbol{\alpha}_1, \quad \boldsymbol{0} \leqq \boldsymbol{s} \leqq \theta(\boldsymbol{1} - \boldsymbol{\alpha}_1), \\
& \boldsymbol{0} \leqq \boldsymbol{x} \leqq \theta \boldsymbol{\alpha}_2, \quad \boldsymbol{0} \leqq \boldsymbol{r} \leqq \theta(\boldsymbol{1} - \boldsymbol{\alpha}_2), \\
& \boldsymbol{y} \geqq \boldsymbol{0}, \quad \boldsymbol{\alpha}_1 \in \{0,1\}^{m+p-1}, \quad \boldsymbol{\alpha}_2 \in \{0,1\}^n,
\end{array}\right. \tag{3}
$$

where $\boldsymbol{c}_i^\top \boldsymbol{x}$ *is one-to-one,* $\boldsymbol{C}_{(i)} = (\boldsymbol{c}_1^\top, ..., \boldsymbol{c}_{i-1}^\top, \boldsymbol{c}_{i+1}^\top, ..., \boldsymbol{c}_p^\top)^\top$, $\theta > 0$ *is a sufficiently large number, and and* $\mathbf{0}$ *is a vector whose each element is* 0.

The proof can be found in Sun's work [14].

We state our main result as the next theorem, whose proof is shown in the next section.

Theorem 2.1. *Under Assumption 2.2, a solution of the following MIP is a solution of (2):*

$$\left|\begin{array}{ll} \text{minimize} & \phi(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{z} = \boldsymbol{b}, \\ & \boldsymbol{A}^\top \boldsymbol{u} + \boldsymbol{C}^\top \boldsymbol{v} - \boldsymbol{w} = -\boldsymbol{C}^\top \mathbf{1}, \\ & \boldsymbol{z} \leqq \theta \boldsymbol{\alpha}, \quad \boldsymbol{u} \leqq \theta(\mathbf{1} - \boldsymbol{\alpha}), \\ & \boldsymbol{x} \leqq \theta \boldsymbol{\beta}, \quad \boldsymbol{w} \leqq \theta(\mathbf{1} - \boldsymbol{\beta}), \\ & \boldsymbol{\alpha} \in \{0,1\}^m, \quad \boldsymbol{\beta} \in \{0,1\}^n, \\ & \boldsymbol{x}, \boldsymbol{z}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \geqq \mathbf{0}, \end{array}\right. \tag{4}$$

where θ *is a sufficiently large number.*

Please note that we do not need the assumption 2.1 in this theorem. Moreover, compared with the problem (3), the size of the problem (4) is smaller, that is, the numbers of constraints and binary variables in (4) are less than those in (3). So we will be able to solve the problem (4) much faster than (3) in practical computation.

3 A Sketch Proof

In this section, we will give a sketch proof of Theorem 2.1.

At first, we refer a linear version of Wendell and Lee's Theorem 1[15], which gives a necessary and sufficient condition for an efficient solution $\boldsymbol{x} \in \boldsymbol{E}$.

Lemma 3.1. $\boldsymbol{x}' \in \boldsymbol{X}$ *is an efficient solution of the problem (1) if and only if* $\boldsymbol{y}$ *is an optimal solution of the following auxiliary problem, and* $\boldsymbol{y} = \boldsymbol{x}'$*:*

$$\left|\begin{array}{ll} \text{minimize} & (\boldsymbol{C}^\top \mathbf{1})^\top \boldsymbol{y} \\ \text{subject to} & -\boldsymbol{A}\boldsymbol{y} \geqq -\boldsymbol{b}, \\ & -\boldsymbol{C}\boldsymbol{y} \geqq -\boldsymbol{C}\boldsymbol{x}', \\ & \boldsymbol{y} \geqq \mathbf{0}, \end{array}\right. \tag{5}$$

where $\boldsymbol{y}$ *is a vector of variables, and* $\mathbf{1}$ *is a vector whose each element is* 1.

Since the problem in the lemma is linear programming (LP), we obtain the next result by using the optimal condition of LP.

Lemma 3.2. $\boldsymbol{x}' \in \mathbb{R}^n$ *is an efficient solution of the problem (1) if and only if there exists* $(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$ *such that*

$$\begin{array}{l} -\boldsymbol{A}\boldsymbol{x}' \geqq -\boldsymbol{b}, \\ \boldsymbol{A}^\top \boldsymbol{u} + \boldsymbol{C}^\top \boldsymbol{v} - \boldsymbol{w} = -\boldsymbol{C}^\top \mathbf{1}, \\ (-\boldsymbol{A}\boldsymbol{x}' + \boldsymbol{b})^\top \boldsymbol{u} = 0, \quad \boldsymbol{x}'^\top \boldsymbol{w} = 0, \\ \boldsymbol{x}', \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \geqq \mathbf{0}. \end{array} \tag{6}$$

From Lemma 3.1, the problem (2) is equivalent to the following problem

$$\left|\begin{array}{ll} \text{minimize} & \phi(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{z} = \boldsymbol{b}, \\ & \boldsymbol{A}^\top \boldsymbol{u} + \boldsymbol{C}^\top \boldsymbol{v} - \boldsymbol{w} = -\boldsymbol{C}^\top \boldsymbol{1}, \\ & \boldsymbol{z}^\top \boldsymbol{u} = 0, \quad \boldsymbol{x}^\top \boldsymbol{w} = 0, \\ & \boldsymbol{x}, \boldsymbol{z}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \geqq \boldsymbol{0}. \end{array}\right. \tag{7}$$

The problem (7) is a mathematical programming with linear complementarity constraints. And we remark that Lemma 3.2 is also equivalent to Yamamoto's (4) of THEOREM 2.7. [2]. We will show that the problem (7) can be solved as a mixed integer programming problem. Since the problem (2) has an optimal solution from Assumption 2.2, the problem (7) also has an optimal solution. Then we impose the following assumption on the problem (7).

Assumption 3.1. *Let $(\boldsymbol{x}^*, \boldsymbol{z}^*, \boldsymbol{u}^*, \boldsymbol{v}^*, \boldsymbol{w}^*)$ be an optimal solution of the problem (7). Assume that θ is sufficiently large number so that*

$$\|(\boldsymbol{x}^*, \boldsymbol{z}^*, \boldsymbol{u}^*, \boldsymbol{w}^*)\|_\infty \leqq \theta.$$

We can obtain the result of Theorem 2.1 from the next lemma.

Lemma 3.3. *Under Assumption 3.1, the problem (7) is equivalent to the following MIP problem:*

$$\left|\begin{array}{ll} \text{minimize} & \phi(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{z} = \boldsymbol{b}, \\ & \boldsymbol{A}^\top \boldsymbol{u} + \boldsymbol{C}^\top \boldsymbol{v} - \boldsymbol{w} = -\boldsymbol{C}^\top \boldsymbol{1}, \\ & \boldsymbol{z} \leqq \theta\boldsymbol{\alpha}, \quad \boldsymbol{u} \leqq \theta(\boldsymbol{1} - \boldsymbol{\alpha}), \\ & \boldsymbol{x} \leqq \theta\boldsymbol{\beta}, \quad \boldsymbol{w} \leqq \theta(\boldsymbol{1} - \boldsymbol{\beta}), \\ & \boldsymbol{\alpha} \in \{0,1\}^m, \quad \boldsymbol{\beta} \in \{0,1\}^n, \\ & \boldsymbol{x}, \boldsymbol{z}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \geqq \boldsymbol{0}. \end{array}\right. \tag{8}$$

4 Application I: the Minimum Maximal Flow Problem

The minimum maximal flow (MMF) problem is introduced by Shi and Yamamoto [16] in 1997 and it is known to be NP-hard. Let $\boldsymbol{N} = (\boldsymbol{V}, \boldsymbol{E}, s, t, \boldsymbol{c})$ be a network, where $\boldsymbol{V}$ is a set of $m+2$ nodes v_j ($j \in \{1, 2, \ldots, m+2\}$), $\boldsymbol{E}$ is a set of n edges e_k ($k \in \{1, 2, \ldots, n\}$), s is a source, t is a sink, and $\boldsymbol{c} \in \mathbb{R}^n$ is a positive vector whose k-th element c_k denotes the capacity of k-th edge e_k. In this paper, $\boldsymbol{V}$ includes both s and t ($s = v_{m+1}$ and $t = v_{m+2}$). Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ be the node-edge incidence matrix of the network $\boldsymbol{N}$, that is, each entry a_{jk} ($j \in \{1, 2, \ldots, m\}$ and $k \in \{1, 2, \ldots, n\}$) is defined as

$$a_{jk} = \begin{cases} 1 & \text{if edge } e_k \text{ leaves node } v_j, \\ -1 & \text{if edge } e_k \text{ enters node } v_j, \\ 0 & \text{otherwise.} \end{cases}$$

Then the set $\boldsymbol{X}$ of feasible flows is given by

$$\boldsymbol{X} = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0},\ \boldsymbol{0} \leqq \boldsymbol{x} \leqq \boldsymbol{c}\}, \tag{9}$$

where $\mathbf{0}$ is a vector whose each element is 0. The value of a flow $\boldsymbol{x}$ is computed by $\boldsymbol{d}^\top \boldsymbol{x}$, where the k-th element of $\boldsymbol{d} \in \mathbb{R}^n$ is

$$d_k = \begin{cases} 1 & \text{if edge } e_k \text{ leaves source } s, \\ -1 & \text{if edge } e_k \text{ enters source } s, \\ 0 & \text{otherwise.} \end{cases}$$

A feasible flow $\boldsymbol{x} \in \boldsymbol{X}$ is maximal if there are no feasible flows $\boldsymbol{y} \in \boldsymbol{X}$ such that $\boldsymbol{y} \geqq \boldsymbol{x}$ and $\boldsymbol{y} \neq \boldsymbol{x}$. The minimum maximal flow (MMF) problem is denoted as

$$\begin{array}{ll} \text{minimize} & \boldsymbol{d}^\top \boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \in \boldsymbol{X} \text{ is a maximal flow.} \end{array} \tag{10}$$

Easy to find that the maximal flow is the efficient solution of the following MOLP.

$$\begin{array}{ll} \text{maximize} & \boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \mathbf{0}, \\ & \mathbf{0} \leqq \boldsymbol{x} \leqq \boldsymbol{c}. \end{array} \tag{11}$$

So the maximal flow set $\boldsymbol{E} := \{\boldsymbol{x} \in \boldsymbol{X} |$ There exists no $\boldsymbol{y} \in \boldsymbol{X}$ such that $\boldsymbol{x} \leqq \boldsymbol{y}$ and $\boldsymbol{x} \neq \boldsymbol{y}\}$ is the efficient set, and the MMF problem is the following OE problem.

$$\begin{array}{ll} \text{minimize} & \boldsymbol{d}^\top \boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \in \boldsymbol{E} \end{array} \tag{12}$$

Then by applying the proposed MIP approach, the MMF problem can be reformulated into the following

$$\begin{array}{ll} \text{minimize} & \boldsymbol{d}^\top \boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}^\top \boldsymbol{z} + \boldsymbol{w} = \mathbf{1}, \\ & \boldsymbol{c} - \boldsymbol{x} \leqq c_{max} \boldsymbol{\alpha}, \\ & \boldsymbol{w} \leqq n(\mathbf{1} - \boldsymbol{\alpha}), \\ & \boldsymbol{A}\boldsymbol{x} = \mathbf{0}, \\ & \mathbf{0} \leqq \boldsymbol{x} \leqq \boldsymbol{c}, \\ & \boldsymbol{\alpha} \in \{0,1\}^n, \end{array} \tag{13}$$

where

$$c_{max} = \max\{c_k \mid k \in \{1, 2, \ldots, n\}\}.$$

Also, in this problem we use n instead of one of the θ by the following lemma.

Lemma 4.1. *Fix any $I \subset \{1, 2, \ldots, n\}$. If the MMF problem (10) has an optimal solution, then there exists an optimal solution $(\boldsymbol{x}^*, \boldsymbol{z}^*, \boldsymbol{u}^*, \boldsymbol{v}^*)$ which satisfies*

$$\|(\boldsymbol{z}^*, \boldsymbol{u}^*, \boldsymbol{v}^*)\|_\infty \leqq n.$$

The proof can be found in Lu et al.[19]. And in the experiment part, Lu et al. [19] showed computational advantages when compared with vertex search method and DC algorithm which are also popular algorithms to solve the OE problem. In fact, by performing computational experiments, the proposed approach is efficient to the MMF problem even for relatively large instances, where the number of edges is up to 15,000, and that the growth rate of running time of our approach is slower than the rates of previous works when the sizes of the instances grow.

5 Application II: the Least Distance Problem in Data Envelopment Analysis

In this section we will introduce some results and extension of Wang et al. [20] Data Envelopment Analysis (DEA) has been studied to measure the efficiency since it was introduced by Charnes et al. in 1978 [21]. In DEA, we assume that there are n decision making units (DMU), each DMU_j ($j = 1, 2, ..., n$) uses m inputs $\boldsymbol{x}_j \in \mathbb{R}^m_{++}$ to produce s outputs $\boldsymbol{y}_j \in \mathbb{R}^s_{++}$. A pair of inputs $\boldsymbol{x}$ and outputs $\boldsymbol{y}$ of a DMU is called an activity. And the variable returns to scale (VRS) [22] production possibility set $\boldsymbol{T}$ in DEA is defined as follows.

$$\boldsymbol{T} = \left\{ (\boldsymbol{x}, \boldsymbol{y}) \,\middle|\, \boldsymbol{X\lambda} \leqq \boldsymbol{x}, \boldsymbol{Y\lambda} \geqq \boldsymbol{y}, \boldsymbol{1}^\top \boldsymbol{\lambda} = 1, \boldsymbol{x}, \boldsymbol{y} > \boldsymbol{0}, \boldsymbol{\lambda} \geqq \boldsymbol{0} \right\}$$

where $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n)$ is the input data matrix, $\boldsymbol{Y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_n)$ is the output data matrix, $\boldsymbol{\lambda} \in \mathbb{R}^n$.

A DMU whose activity is $(\boldsymbol{x}, \boldsymbol{y})$ is called efficient if there is no pair $(\boldsymbol{x}', \boldsymbol{y}') \in \boldsymbol{T}$ such that $(\boldsymbol{x}, -\boldsymbol{y}) \geqq (\boldsymbol{x}', -\boldsymbol{y}')$ and $(\boldsymbol{x}, -\boldsymbol{y}) \neq (\boldsymbol{x}', -\boldsymbol{y}')$. Let efficient frontier $\boldsymbol{E}$ be the set of efficient DMUs. Obviously, the efficient DMU is the efficient solution of the following MOLP.

$$\begin{array}{ll} \text{minimize} & \begin{array}{c} \boldsymbol{x} \\ -\boldsymbol{y} \end{array} \\ \text{subject to} & (\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{T} \end{array} \tag{14}$$

The traditional DEA models try to find an efficient activity which maximizes the improvement from the assessed DMU. In recent years, Briec [23] proposed ℓ_p-norm distance model which is the first least distance problem (LDP) of DEA. Compared with the traditional model, the LDP try to provide a closer efficient activity which means it would cause less improvement to become efficient as well. Several researches [24, 25] tried to solve this problem, but failed to establish the whole efficient frontier. Recently, Wang et al. [20] try to establish the frontier by using an auxiliary problem, then use a branch and bound method to solve it. But since the efficient frontier is not convex in general, the computation is difficult. In this paper, we try to find out a efficient method to search the closest efficient activity in the whole efficient frontier with the idea of the OE problem. Let $\boldsymbol{E}$ be the efficient frontier, the LDP can be written as following in general.

$$\begin{array}{ll} \text{minimize} & d(\boldsymbol{x}, \boldsymbol{y}) \\ \text{subject to} & (\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{E} \end{array} \tag{15}$$

where $d(\boldsymbol{x}, \boldsymbol{y})$ can be any kind of distance function. So (15) is an OE problem.

Since the number of binary variables grows with the number of inequality constraints, we propose some techniques to reduce it. By exchange the position, let DMU_1 to DMU_k be efficient and all the other DMUs be inefficient, and let $\boldsymbol{X}_E = (\boldsymbol{x}_1, ..., \boldsymbol{x}_k)$ and $\boldsymbol{Y}_E = (\boldsymbol{y}_1, ..., \boldsymbol{y}_k)$. Then we can construct the following set $\boldsymbol{H}$.

$$\boldsymbol{H} = \left\{ (\boldsymbol{x}, \boldsymbol{y}) \,\middle|\, \boldsymbol{X}_E\boldsymbol{\lambda} = \boldsymbol{x}, \boldsymbol{Y}_E\boldsymbol{\lambda} = \boldsymbol{y}, \boldsymbol{1}^\top \boldsymbol{\lambda} = 1, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda} \geqq \boldsymbol{0} \right\}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^k$. When $\boldsymbol{T}' = \left\{ (\boldsymbol{x}, \boldsymbol{y}) \,\middle|\, \boldsymbol{X\lambda} = \boldsymbol{x}, \boldsymbol{Y\lambda} = \boldsymbol{y}, \boldsymbol{1}^\top \boldsymbol{\lambda} = 1, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda} \geqq \boldsymbol{0} \right\}$, we have the following lemma.

Lemma 5.1. $\boldsymbol{E} \subset \boldsymbol{H} \subset \boldsymbol{T}' \subset \boldsymbol{T}$

Because elements in $\boldsymbol{X}_E$, $\boldsymbol{Y}_E$, $\boldsymbol{X}$, and $\boldsymbol{Y}$ are all positive, the convex combination of $\boldsymbol{X}_E$ and $\boldsymbol{Y}_E$ and the convex combination of $\boldsymbol{X}$ and $\boldsymbol{Y}$ should be constantly positive which implies $\boldsymbol{H} \subset \boldsymbol{T}' \subset \boldsymbol{T}$. To show that $\boldsymbol{E} \subset \boldsymbol{H}$, Then we construct the following two MOLP.

$$\begin{array}{ll} \text{minimize} & \begin{array}{c} \boldsymbol{x} \\ -\boldsymbol{y} \end{array} \\ \text{subject to} & (\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{T}' \end{array} \tag{16}$$

$$\begin{array}{ll} \text{minimize} & \begin{array}{c} \boldsymbol{x} \\ -\boldsymbol{y} \end{array} \\ \text{subject to} & (\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{H} \end{array} \tag{17}$$

Obviously, all efficient solutions of (14) must also be efficient solutions of (14), and all efficient solutions of (16) must also be efficient solutions of (17). Otherwise, there must be some activities dominate it.

The ℓ_p-norm LDP problem of $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ is the OE of (14), then it is also the OE of (17). Then with Theorem2.1, the ℓ_p-norm LDP problem of $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ can be formulated into the following standard form.

$$\left|\begin{array}{ll}
\text{minimize} & ||(\boldsymbol{x}, \boldsymbol{y}) - (\boldsymbol{x}_0, \boldsymbol{y}_0)||_p \\
\text{subject to} & \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} & -\boldsymbol{X}_E \\ -\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{X}_E \\ \boldsymbol{0} & \boldsymbol{I} & -\boldsymbol{Y}_E \\ \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{Y}_E \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{1}^\top \\ \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{1}^\top \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \boldsymbol{\lambda} \end{pmatrix} + \begin{pmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ \boldsymbol{z}_3 \\ \boldsymbol{z}_4 \\ z_5 \\ z_6 \end{pmatrix} = \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ 1 \\ -1 \end{pmatrix}, \\
& \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} & -\boldsymbol{X}_E \\ -\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{X}_E \\ \boldsymbol{0} & \boldsymbol{I} & -\boldsymbol{Y}_E \\ \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{Y}_E \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{1}^\top \\ \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{1}^\top \end{pmatrix}^\top \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \boldsymbol{u}_3 \\ \boldsymbol{u}_4 \\ u_5 \\ u_6 \end{pmatrix} + \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{pmatrix} - \begin{pmatrix} \boldsymbol{w}_1 \\ \boldsymbol{w}_2 \\ \boldsymbol{w}_3 \end{pmatrix} = \begin{pmatrix} -\boldsymbol{1} \\ \boldsymbol{1} \\ \boldsymbol{0} \end{pmatrix}, \\
& \boldsymbol{z}_1 \leqq \theta\boldsymbol{\alpha}_1, \quad \boldsymbol{u}_1 \leqq \theta(\boldsymbol{1} - \boldsymbol{\alpha}_1), \quad \boldsymbol{z}_2 \leqq \theta\boldsymbol{\alpha}_2, \quad \boldsymbol{u}_2 \leqq \theta(\boldsymbol{1} - \boldsymbol{\alpha}_2), \\
& \boldsymbol{z}_3 \leqq \theta\boldsymbol{\alpha}_3, \quad \boldsymbol{u}_3 \leqq \theta(\boldsymbol{1} - \boldsymbol{\alpha}_3), \quad \boldsymbol{z}_4 \leqq \theta\boldsymbol{\alpha}_4, \quad \boldsymbol{u}_4 \leqq \theta(\boldsymbol{1} - \boldsymbol{\alpha}_4), \\
& z_5 \leqq \theta\alpha_5, \quad u_5 \leqq \theta(1 - \alpha_5), \quad z_6 \leqq \theta\alpha_6, \quad u_6 \leqq \theta(1 - \alpha_6), \\
& \boldsymbol{x} \leqq \theta\boldsymbol{\beta}_1, \quad \boldsymbol{w}_1 \leqq \theta(\boldsymbol{1} - \boldsymbol{\beta}_1), \\
& \boldsymbol{y} \leqq \theta\boldsymbol{\beta}_2, \quad \boldsymbol{w}_2 \leqq \theta(\boldsymbol{1} - \boldsymbol{\beta}_2), \\
& \boldsymbol{\lambda} \leqq \theta\boldsymbol{\beta}_3, \quad \boldsymbol{w}_3 \leqq \theta(\boldsymbol{1} - \boldsymbol{\beta}_3), \\
& \boldsymbol{\alpha}_1 \in \{0,1\}^m, \quad \boldsymbol{\alpha}_2 \in \{0,1\}^s, \quad \boldsymbol{\alpha}_3 \in \{0,1\}^m, \\
& \boldsymbol{\alpha}_4 \in \{0,1\}^s, \quad \alpha_5 \in \{0,1\}, \quad \alpha_6 \in \{0,1\}, \\
& \boldsymbol{\beta}_1 \in \{0,1\}^m, \quad \boldsymbol{\beta}_2 \in \{0,1\}^s, \quad \boldsymbol{\beta}_3 \in \{0,1\}^k, \\
& \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}, \boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3, \boldsymbol{z}_4, z_5, z_6, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4, u_5, u_6, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3 \geqq \boldsymbol{0}.
\end{array}\right. \tag{18}$$

Because the equality constraints of (17), $\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3, \boldsymbol{z}_4, z_5, z_6$, are all 0, the related binary variables can be deleted, and let $\boldsymbol{u}_1 = \boldsymbol{u}_1 - \boldsymbol{u}_2$, $\boldsymbol{u}_3 = \boldsymbol{u}_3 - \boldsymbol{u}_4$, and $u_5 = u_5 - u_6$. And because $\boldsymbol{x}$ and $\boldsymbol{y}$ are nonzero, $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are all 0. Also for $\boldsymbol{\lambda} \geqq \boldsymbol{0}$ and $\boldsymbol{1}^\top\boldsymbol{\lambda} = 1$, θ can be

1 in $\boldsymbol{\lambda}$-part. Then the LDP is equivalent to the following MIP.

$$\left|\begin{array}{ll} \text{minimize} & ||(\boldsymbol{x},\boldsymbol{y})-(\boldsymbol{x}_0,\boldsymbol{y}_0)||_p \\ \text{subject to} & \boldsymbol{X}_E\boldsymbol{\lambda}=\boldsymbol{x}, \\ & \boldsymbol{Y}_E\boldsymbol{\lambda}=\boldsymbol{y}, \\ & \boldsymbol{1}^\top\boldsymbol{\lambda}=1, \\ & \boldsymbol{u}_1 \geqq -\boldsymbol{1}, \\ & \boldsymbol{u}_3 \leqq \boldsymbol{1}, \\ & -\boldsymbol{X}_E^\top\boldsymbol{u}_1-\boldsymbol{Y}_E^\top\boldsymbol{u}_3+u_5-\boldsymbol{w}_3=\boldsymbol{0} \\ & \boldsymbol{\lambda}\leqq\boldsymbol{\beta}_3, \quad \boldsymbol{w}_3\leqq\theta(\boldsymbol{1}-\boldsymbol{\beta}_3), \\ & \boldsymbol{\beta}_3\in\{0,1\}^k, \boldsymbol{\lambda},\boldsymbol{w}_3\geqq\boldsymbol{0}. \end{array}\right. \tag{19}$$

Then we have a tool to find an efficient activity with the least improvements in the whole efficient frontier.

6 Conclusion

Based on optimal conditions of linear programming, a necessary and sufficient condition for efficiency has been established to describe the efficient set. Then we showed the MIP approach can be applied to solve the OE problem. We reduce $p-1$ binary variables and relax the assumption of previous research. The exponential growth of running time caused by the number of objective functions has been reduced because that the related variables are all continuous in proposed approach. Also, Lu et al[19] showed computational advantages when applied the proposed approach to solve the MMF problem. Then, we showed that the LDP is an OE problem and can use MIP approach to solve it. For the development of the state-of-art MIP solver, the proposed approach may be able to solve large-scale optimization problems over the efficient set efficiently.

Acknowledgment

This research is supported in part by Grant-in-Aid for Science Research (A) 26242027 of Japan Society for the Promotion of Science.

References

[1] Philip, J.: Algorithms for the vector maximization problem. Mathematical programming **2**(1), 207–229 (1972)

[2] Yamamoto, Y.: Optimization over the efficient set: overview. Journal of Global Optimization **22**(1-4), 285–317 (2002)

[3] Bolintineanu, S.: Minimization of a quasi-concave function over an efficient set. Mathematical Programming **61**(1-3), 89–110 (1993)

[4] Benson, H.P.: A finite, nonadjacent extreme-point search algorithm for optimization over the efficient set. Journal of Optimization Theory and Applications **73**(1), 47–64 (1992)

[5] Benson, H.P.: An algorithm for optimizing over the weakly-efficient set. European journal of operational research **25**(2), 192–199 (1986)

[6] Dauer, J.P., Fosnaugh, T.A.: Optimization over the efficient set. Journal of Global Optimization **7**(3), 261–277 (1995)

[7] Thach, P., Konno, H., Yokota, D.: Dual approach to minimization on the set of pareto-optimal solutions. Journal of Optimization Theory and Applications **88**(3), 689–707 (1996)

[8] Phong, T.Q., Tuyen, J.: Bisection search algorithm for optimizing over the efficient set. Vietnam J. Math **28**, 217–226 (2000)

[9] An, L.T.H., Tao, P.D., Muu, L.D.: Numerical solution for optimization over the efficient set by dc optimization algorithms. Operations Research Letters **19**(3), 117–128 (1996)

[10] Hoang, T.: Convex Analysis and Global Optimization, vol. 110. Springer (2016)

[11] Benson, H.P., Lee, D.: Outcome-based algorithm for optimizing over the efficient set of a bicriteria linear programming problem. Journal of Optimization Theory and Applications **88**(1), 77–105 (1996)

[12] Benson, H.P.: An outcome space algorithm for optimization over the weakly efficient set of a multiple objective nonlinear programming problem. Journal of Global Optimization **52**(3), 553–574 (2012)

[13] Liu, Z., Ehrgott, M.: Primal and dual algorithms for optimization over the efficient set. Optimization pp. 1–26 (2018)

[14] Sun, E.: On optimization over the efficient set of a multiple objective linear programming problem. Journal of Optimization Theory and Applications **172**(1), 236–246 (2017)

[15] Wendell, R.E., Lee, D.: Efficiency in multiple objective optimization problems. Mathematical programming **12**(1), 406–414 (1977)

[16] Shi, J., Yamamoto, Y.: A global optimization method for minimum maximal flow problem. Acta Mathematica Vietnamica **22**(1), 271–287 (1997)

[17] Muu, L.D., Thuy, L.Q.: On dc optimization algorithms for solving minmax flow problems. Mathematical Methods of Operations Research **80**(1), 83–97 (2014)

[18] Shigeno, M., Takahashi, I., Yamamoto, Y.: Minimum maximal flow problem: an optimization over the efficient set. Journal of Global Optimization **25**(4), 425–443 (2003)

[19] Lu, K., Mizuno, S., Shi, J.: A mixed integer programming approach for the minimum maximal flow. Journal of the Operations Research Society of Japan **64**(4), 261–271 (2018)

[20] Wang, X., Lu, K., Shi, J., Hasuike, T.: A branch and bound approach for the least distance problem in data envelopment analysis (working paper) (2018)

[21] Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. European journal of operational research **2**(6), 429–444 (1978). DOI 10.1016/0377-2217(78)90138-8

[22] Cooper, W.W., Seiford, L.M., Tone, K.: Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software. Springer Science & Business Media (2007). DOI 10.1007/978-0-387-45283-8

[23] Briec, W.: Hölder distance function and measurement of technical efficiency. Journal of Productivity Analysis **11**(2), 111–131 (1999). DOI 10.1023/A:1007764912174

[24] Aparicio, J., Ruiz, J.L., Sirvent, I.: Closest targets and minimum distance to the pareto-efficient frontier in DEA. Journal of Productivity Analysis **28**(3), 209–218 (2007). DOI 10.1007/s11123-007-0039-5

[25] Ando, K., Kai, A., Maeda, Y., Sekitani, K.: Least distance based inefficiency measures on the pareto-efficient frontier in DEA. Journal of the Operations Research Society of Japan **55**(1), 73–91 (2012). DOI 10.15807/jorsj.55.73

Kuan Lu
Department of Industrial Engineering and Management
Tokyo Institute of Technology
2-12-1 Ohokayama
Meguro-ku Tokyo 152-8552, Japan
E-mail: `lu.k.aa@m.titech.ac.jp`