Clustering Decision Making Units (DMUs) Using Full Dimensional Efficient Facets (FDEFs) of PPS with BCC Technology

M. R. Moazami Goudarzi\(^a\) and M. R. Jaber Ansari\(^a\)

\(^a\) Department of Mathematics, Borujerd Branch
Islamic Azad University, Borujerd, Iran
mrmoazamig@gmail.com (Mohammad Reza Moazami Goudarzi)

Abstract

In a recent paper by Po et al. [Po, R.W., Guh, Y.Y., Yang, M.S., 2009. A new clustering approach using data envelopment analysis], the authors have tried to show how the piecewise production functions derived from the DEA method can be used to cluster the decision making units. In particular, they use the CCR model to set up their clustering approach. Related with the afore-mentioned approach, there exist some problems which will be scrutinized in this paper. Meanwhile, a new method will be introduced for clustering DMUs using FDEFs of PPS with BCC Technology.

Keywords: Data envelopment analysis; Clustering; strong defining hyperplanes; BCC model

1 Introduction

The technique of clustering is as old as science itself. Clustering is the task of grouping similar objects into clusters. The process involves classification of existing data such that the variation in the data in the same group is as low as possible and that between groups is very high. This, in other words, means that the clusters should be tight within themselves and as far away from each other as possible. Clustering is loosely defined as a technique or a method used to find groups in data. More rigorously, clustering can be defined as "Partitioning the data set into subgroups called clusters such that data points in the same cluster are more 'similar' to each other compared to data points in other clusters.” Cluster analysis enables the analyst to find the underlying structure in data and many hypotheses that would answer many varied questions regarding the data. For example, how the data are related to each other within the population, can any inference be drawn on account of the before-mentioned relations, can an inductive argument be proposed to
allow for a result to be applicable to all the data within a subset and more so can the same argument be continued to include the whole set of data? Several techniques have been proposed to obtain meaningful classification of data. Hierarchical clustering (Hartigan, 1975; Kaufman and Rousseew, 1990), mixture-model clustering (McLachlan and Basford, 1988; McLachlan and Krishnan, 1997), learning network clustering (Grossberg, 1976; Lippmann, 1987; Tsao et al., 1994; Kohonen, 2001), objective-function-based clustering, and partition clustering (Bezdek, 1981; Yang, 1993) are the categories of cluster analysis. Algorithms k-means (Duda and Hart, 1973; Hartigan, 1975), fuzzy c-means (FCM) (Bezdek, 1981; Yang, 1993; Wu and Yang, 2002), and possibilistic c-means (PCM) (Krishnapuram and Keller, 1993) are clustering procedures that minimize total dissimilarity and can also be considered as feature analysis techniques.

In a recent paper by Po et al. [Po, R.W., Guh, Y.Y., Yang, M.S., 2009. A new clustering approach using data envelopment analysis], the authors have tried to show how the piecewise production functions derived from the DEA method can be used to cluster the data with input and output items. Related with the approach, there exist some problems, including: (1) It is well known that the optimal solutions of the envelopment form are often highly degenerate and, therefore, the multiplier form of the CCR model may have alternative optimal solutions, it is possible that the hyperplanes derived by the approach mentioned and following the clusters produced are not unique. (2) In the process of the approach, if the optimal weights in the multiplier form of CCR model are strictly positive, the corresponding hyperplane is taken into consideration as a reference for clustering, but it is possible not to achieve any strictly positive \((u^*, v^*)\) in evaluating all DMUs using CCR model. (3) Some of the obtained clusters using the approach may have overlapping units. In this paper, we cluster DMUs using FDEFs of PPS with BCC Technology that overcomes these shortcomings.

The rest of this paper is organized as follows: Section 2 discusses the preliminaries from which the proposed clustering approach is developed. Section 3 includes the characteristics and structure of FDEFs of PPS. In section 4 we present the DEA-based clustering method introduced by Po et al. (2009). Section 5 gives the proposed method. The problems relation with the method introduced by Po et al. (2009) and comparison to proposed method are presented in Section 6. Finally, conclusions are stated in Section 7.

2 Preliminaries

Assume that we have \(n\) observed DMUs \((X_j, Y_j)\), where, \(j = 1, 2, \ldots, n\), and every DMU \(X_j\) produces the same \(s\) outputs in (possibly) different amounts, \(y_{rj}, r = 1, 2, \ldots, s\), using the same inputs, \(x_{ij}, i = 1, 2, \ldots, m\), also in (possibly) different amounts. All inputs and outputs are assumed to be nonnegative, but at least one input and one output are positive, i.e., \(X_j = (x_1j, \ldots, x_mj) \geq 0, X_j \neq 0\) and \(Y_j = (y_1j, \ldots, y_sj) \geq 0, Y_j \neq 0\).

The DEA methodology comprises a wide variety of mathematical programming models for performance measurement and performance benchmark. DEA models differ with respect to the assumptions imposed on production relationships, and the measure used
for evaluating efficiency relative to the frontier (Dekker and Post, 2000). One of the general DEA models is the CCR model (Charnes et al. 1978). This model is constructed relating to the empirically defined production possibility set \( T \) with constant return to scale assumption that is specified by the following postulates:

**postulate 1.** *(Inclusion of observation).* The observed activities \((X_j, Y_j), (j = 1, ..., n)\) belong to \( T \).

**postulate 2.** *(Convexity).* If the activities \((X_i, Y_i)\) and \((X_j, Y_j)\) belong to \( T \) then the activity \((\lambda X_i + (1 - \lambda)X_j, \lambda Y_i + (1 - \lambda)Y_j)\) belong to \( T \), for any scaler \( \lambda \in [0, 1] \).

**postulate 3.** *(Constant return to scale).* If the activities \((X_i, Y_i)\) belong to \( T \) then the activity \((\lambda X_i, \lambda Y_i)\) belong to \( T \), for any scaler \( \lambda \in R \).

**postulate 4.** *(Monotonicity).* If the activity \((X, Y)\) is in \( T \), any semipositive activity \((X', Y')\) with \( X \geq X', Y \leq Y' \) is included in \( T \). That is, any activity with input no less than \( X \) and with output no greater than \( Y \) in any component also belongs to \( T \).

**postulate 5.** *(Minimum extrapolation).* If the production possibility set \( T' \) satisfied Postulate 1,2,3 and 4, then \( T \subseteq T' \).

The unique production possibility set with constant return to scale assumption determined by the above mentioned postulate is given by:

\[
T_c = \{(X, Y) \mid X \geq \sum_{j=1}^{n} \lambda_j X_j, Y \leq \sum_{j=1}^{n} \lambda_j Y_j, \lambda_j \geq 0, j = 1, ..., n\}
\]

Consider DMU_{jo} \((jo \in \{1, ..., n\})\) as the DMU under evaluation. For evaluation the efficiency of the DMU relative to \( T_c \), we consider the following mathematical program:

\[
\begin{align*}
\min & \quad \theta_{jo} \\
\text{s.t.} & \quad (\theta_{jo}, X_{jo}, Y_{jo}) \in T_c.
\end{align*}
\]

Considering the structure of \( T_c \), the model can be written as follows:

\[
\begin{align*}
\min & \quad \theta_{jo} \\
\text{s.t.} & \quad \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_{jo} x_{ij}, \quad i = 1, ..., m, \\
& \quad \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{rjo}, \quad r = 1, ..., s, \\
& \quad \lambda_j \geq 0.
\end{align*}
\]

**Definition 2.1.** *(Efficiency).* DMU_{jo} is CCR-efficient when in the optimal solution(s)

(i) \( \theta_{jo}^* = 1 \).

(ii) All slack variables are zero in alternative optimal solutions.

To account for slack variables in all alternative solutions, we may use the following model,
which maximizes the sum of the input excesses and output shortfalls, keeping $\theta_{jo} = \theta^*_{jo}$ (the optimal objective value of Model (1)).

$$\max \omega = es^- + es^+$$

$$s.t. \begin{align*}
\sum_{j=1}^{n} \lambda_j x_{ij} + s^-_i &= \theta^*_{jo} x_{ijo}, \quad i = 1, ..., m, \\
\sum_{j=1}^{n} \lambda_j y_{rj} - s^+_i &= y_{rjo}, \quad r = 1, ..., s, \\
\lambda_j &\geq 0, \quad s^-_i \geq 0, \quad s^+_i \geq 0,
\end{align*}$$

The dual form of (2), which evaluates DMU$_{jo}$, is as follows:

$$\max e_{jo} = u^t y_{jo}$$

$$s.t. \begin{align*}
v^t x_{jo} &= 1, \\
u^t y_j - v^t x_j &\leq 0, \quad j = 1, ..., n, \\
u &\geq 0, v \geq 0.
\end{align*}$$

If the postulate 3, that is one of the postulate that used to construction of $T_v$, will be omitted. So, new production possibility set will be resulted as $T_v$ by:

$$T_v = \{(X,Y) \mid X \geq \sum_{j=1}^{n} \lambda_j X_j, Y \leq \sum_{j=1}^{n} \lambda_j Y_j, \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1, ..., n\},$$

in which $v$ represent the variant return to scales. For evaluating the efficiency score of DMU$_{jo}$ relative to $T_v$, we consider the below program:

$$\min \theta_{jo}$$

$$s.t. \begin{align*}
\sum_{j=1}^{n} \lambda_j x_{ij} &\leq \theta_{jo} x_{ijo}, \quad i = 1, ..., m, \\
\sum_{j=1}^{n} \lambda_j y_{rj} &\geq y_{rjo}, \quad r = 1, ..., s, \\
\sum_{j=1}^{n} \lambda_j &= 1, \\
\lambda_j &\geq 0.
\end{align*}$$

DMU$_{jo}$ is BCC-efficient if and only if $\theta^*_{jo} = 1$ in (4) and the optimal value of (5) is equal
Clustering decision making units

\[ \max \omega = cs^- + cs^+ \]

\[ s.t. \sum_{j=1}^{n} \lambda_j x_{ij} + s^-_i = \theta^*_j x_{ijo}, \quad i = 1, \ldots, m, \quad (5) \]

\[ \sum_{j=1}^{n} \lambda_j y_{rj} - s^+_i = y_{rjo}, \quad r = 1, \ldots, s, \]

\[ \sum_{j=1}^{n} \lambda_j = 1, \]

\[ \lambda_j \geq 0, \quad s^-_i \geq 0, \quad s^+_i \geq 0. \]

**Definition 2.2.** (BCC-Projection). For a BCC-inefficient DMU \( j_o \), we define its BCC-projection, based on an optimal solution for Model (5), as follows:

\[(\hat{X}_{j_o} = \theta^*_{j_o} X_{j_o} - s^-_{j_o}, \hat{Y}_{j_o} = Y_{j_o} + s^{++}_o)\]

The improved activity \((\hat{X}_{j_o}, \hat{Y}_{j_o})\) is BCC-efficient.

The dual multiplier form of model (4) is expressed as:

\[
\begin{align*}
\max & \quad u^t y_{j_o} + u_o \\
\text{s.t.} & \quad v^t x_{j_o} = 1, \\
& \quad u^t y_j - v^t x_j + u_o e \leq 0, \quad j = 1, \ldots n, \\
& \quad u \geq 0, v \geq 0. \\
\end{align*}
\]

(6)

DMU \( j_o \) is BCC-efficient if there exists at least one optimal solution \((u^*, v^*)\) of model (6), with \((u^*, v^*) > 0\) and \(u^t y_{j_o} + u_o = 1\); otherwise, DMU \( j_o \) is BCC-inefficient.

The set of all DMUs corresponding to positive \( \lambda^*_j \)'s is called the reference set to DMU \( j_o \), i.e., \( E_{j_o} = \{DMU_j | \lambda^*_j > 0 \text { in some optimal solution of (5)}\} \). DMU \( j_o \) is extreme efficient if and only if \( E_{j_o} = \{o\} \).

In the next sections, we will need to determine all extreme BCC-efficient DMUs; for this
purpose, we solve the following model for each DMU:

\[
\begin{align*}
\text{max} & \quad \gamma_{jo} = \sum_{j=1, j \neq jo}^{n} \lambda_j \\
\text{s.t.} & \quad \sum_{j=1}^{n} \lambda_j x_j \leq x_{jo}, \\
& \quad \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{jo}, \\
& \quad \sum_{j=1}^{n} \lambda_j = 1, \\
& \quad \lambda_j \geq 0, j = 1, ..., n
\end{align*}
\] (7)

Theorem 1. A DMU \((X_{jo}, Y_{jo})\) is extrem BCC-efficient if and only if the optimal value of the objective function \(\gamma_{jo}^*\) is zero.

Proof. See Appendix.

3 Characteristics and structure of FDEFs of PPS

Let \(P \subseteq S^d\) be a convex set. A linear inequality \(CX \leq c_o\) is valid for \(P\) if it is satisfied for all \(X \in P\). A face of \(P\) is any set of the form \(F = P \cap \{X \in S^d : CX = c_o\}\) where \(CX \leq c_o\) is a valid inequality for \(P\).

The dimension of a face is the dimension of its affine hull. The faces of dimension 0, 1, \(\dim(P)-2\) and \(\dim(P)-1\) are called vertices, edges, ridges and facets respectively.

For DMUs with \(m\) inputs and \(s\) outputs, \(T_v\) is a convex set of dimension \(m + s\). So the dimension of each facet of \(T_v\) is \(m + s - 1\). Therefore, each facet of \(T_v\) contains at least \(m + s\) DMUs (virtual or actual) that are affine independent.

A hyperplane \(H\) in \(E^n\) is a set of the form \(\{X \mid P^dX = k\}\) where \(P\) is a nonzero vector in \(E^n\) and \(k\) is a scaler. Here \(P\) is usually called the normal or the gradient to the hyperplane.

A hyperplane \(H\) in an \((m + s)\) dimensional input-output space passing through the point represented by the vectors \((x_o, y_o)\) can be expressed by the equation:

\[
H : u(y - y_o) - v(x - x_o) = 0
\] (8)

where \(u \in R^s\) and \(v \in R^m\) are coefficient vectors. Now we define \(u_o\), a scaler, by \(u_o = uy_o - vx_o\). Thus, the hyperplane in (8) can be expressed as: \(uy - vx - u_o = 0\).

Generally, a hyperplane divides the space into two halfspaces. If the hyperplane \(H\) contains the production possibility set \(P\) in only one of the halfspace, then it said to be a
Supporting hyperplane of $P$ at the point $(x_o, y_o)$, that is a supporting hyperplane touches the PPS at this point. More concretely, for every $(x, y) \in P$, associated with any DMU, $uy - vx - u_o \leq 0$ holds.

In the evaluation of DMU $o \in \{1, \ldots, n\}$, if $(u^*, v^*, u_o^*)$ is an optimal solution of Model (6), then $H : u^*y - v^*x - u_o^* = 0$ is supporting hyperplane of the $T_v$. So the set $F = T_v \cap H$ is a face of $T_v$.

Hyperplane $H = \{X \mid P^tX = k, P \geq 0\}$ is strong if none of components of $P$ are zero, and is weak if some components of $P$ are zero, and the corresponding face, $F = T_v \cap H$ is called strong face.

Consider DMU $o$ in Fig. 1. Using Model (6), it can be seen that there are alternative optimal solutions which define an infinite number of hyperplanes passing through DMU $o$, of which only two hyperplanes ($H_1, H_2$) are defining hyperplanes.

![Figure 2.1. $H_1$ and $H_2$ are defining and $H$ is supporting but not defining.](image)

To completely characterize the FDEFs of $T_v$, we need the following definitions and preliminaries:

**Definition 3.1.** Suppose that $H : u^*t y - v^*t x + u_o^* = 0$ is a supporting hyperplane of $T_v$, then $F = H \cap T_v$ is called a full dimensional efficient facet (FDEF) of $T_v$ if:

(i) there exists at least one affine independent set with $m + s$ elements of BCC-efficient DMUs lying on $F = H \cap T_v$, and

(ii) all multipliers are strictly positive i.e. $(u^*, v^*, u_o^*) > 0$.

The hyperplane satisfying in the above definition is called strong defining hyperplane of $T_v$. 

**Figure 2.1. $H_1$ and $H_2$ are defining and $H$ is supporting but not defining.**
Given a convex set, a nonzero vector \( \mathbf{d} \) is called \((\text{recession})\) direction of the set, if for each \( \mathbf{x}_o \) in the set, the ray \{\( \mathbf{x}_o + \lambda \mathbf{d} : \lambda \geq 0 \)\} also belongs to the set. Hence starting at any point \( \mathbf{x}_o \) in the set, one can recede along \( \mathbf{d} \) for any step length \( \lambda \geq 0 \) and remain within the set. Clearly if the set is bounded (a set is bounded if there is a number \( k \) such that \( \| \mathbf{x} \| < k \) for each point \( \mathbf{x} \) in the set), then it has no directions.

A polyhedral set is the intersection of a finite number of halfspaces. A bounded polyhedral set is called a \( \text{polytope} \). In general any point in a bounded polyhedral set can be represented as a convex combination of its extreme points.

**Lemma 1.** Suppose that \( H : \mathbf{v}_t^\prime \mathbf{y} - \mathbf{v}_t^\prime \mathbf{x} + \pi_o = 0 \) is a strong supporting hyperplane of \( T_v \). An \((m + s)\)-point \( v \) is an extreme point of the set \( H \cap T_v \) if and only if it is an extreme BCC-efficient DMU lying on \( H \).

*Proof.* See Appendix.

**Theorem 2.** Suppose that \( H : \mathbf{v}_t^\prime \mathbf{y} - \mathbf{v}_t^\prime \mathbf{x} + \pi_o = 0 \) is a strong supporting hyperplane of \( T_v \) then \( H \cap T_v \) is a polytope.

*Proof.* See Appendix.

We can go further and prove the following theorem.

**Theorem 3.** Suppose that \( H \) is a strong defining hyperplane of \( T_v \). Then there exists at least one affine independent set with \( m + s \) elements of the extremes BCC-efficient DMUs in \( H \cap T_v \).

*Proof.* See Appendix.

Theorem (4) mentioned below plays a key role in the proposed method

**Theorem 4.** Suppose that DMU \( o \) is an extreme BCC-efficient DMU. If there is at least one strong hyperplane passing through DMU \( o \) and \((u^*, v^*, u^*_o)\) is an optimal solution of the following model in which \((u^*, v^*) > 0\), then \( H^* : u^* t^\prime \mathbf{y} - v^* t^\prime \mathbf{x} + u^*_o = 0 \) is a strong defining hyperplane of \( T_v \).

\[
\min_I I_o = \sum_{j \in E} I_j
\]
\[s.t.\]
\[
v^t \mathbf{x}_o = 1,
\]
\[
u^t \mathbf{y}_o + u_o = 1,
\]
\[
u^t \mathbf{y}_j - v^t \mathbf{x}_j + u_o + t_j = 0, \quad j \in E,
\]
\[
t_j - \mathbf{M} I_j \leq 0, \quad j \in E,
\]
\[
I_j - \mathbf{M} t_j \leq 0, \quad j \in E,
\]
\[
I_j \in \{0, 1\}, \quad j \in E,
\]
\[
t_j \geq 0, \quad j \in E,
\]
\[
u \geq 0, \quad v \geq 0,
\]
where set $E$ is the set of all extreme BCC-efficient observed units and $M$ is a sufficiently large positive quantity and can be chosen as $M = \sum_{i=1}^{m} \frac{\max_{j \in E} \{x_{ij}\}}{\min_{j \in E} \{x_{ij}\}}$.

**Proof.** See Appendix. \qed

### 3.1 A method for finding all strong defining hyperplanes of $T_v$

Consider extreme BCC-efficient observed unit, DMU$_o$, and evaluate it by model (9). By theorem (4) if there exists at least one FDEF containing DMU$_o$ (equivalently if there exists at least one strong defining hyperplane passing through DMU$_o$) then the optimal solution of model (9) is the gradient of a strong defining hyperplane passing through DMU$_o$ and it is positive for variables $u$ and $v$. Suppose that $I^*_o = \{E\} - k$ and $(u^*, v^*, u^*_o)$, in which $(u^*, v^*) > 0$, are the optimal objective and optimal solution of model (9), respectively. Let $H^*_o : u^*y - v^*x + u^*_o = 0$, and save $F_o = H^*_o \cap T_v$ as a strong defining hyperplane of $T_v$ and let $J_o = \{j : I^*_j = 0\}$. In fact $J_o$ is the set of all extreme BCC-efficient DMUs lying on $H^*_o$. Next, add the following constraint to the constraints of model (9):

$$\sum_{j \in J_o} |I_j| - \sum_{j \in J_o} |I_j| \leq I^*_o - 1 \quad (10)$$

Again evaluate DMU$_o$ by model (9), if there exists another strong defining hyperplane except $H^*_o$ passing through DMU$_o$, then theorem (4) and adding constraint (10) to the constraints of model (9), guarantee that the model (9) by new added constraint in evaluation DMU$_o$ will give the gradient of alternative strong defining hyperplane passing through DMU$_o$ as an alternative optimal solution (note that in model (9) $t^*_j = 0$ does not imply $I^*_j = 0$ so to prevent model (9) from giving the solution in the previous step the constraints $I_j - Mt_j \leq 0, j \in E$ should be added to the constraints of model (9). Save this strong defining hyperplane and construct the corresponding set $J_o$ to it. Add the corresponding constraint to constraints of model (9), and evaluate DMU$_o$ by model (9). Now it has two new constraints. (note that the pervious constraint added to model (9) corresponding to the first saved strong defining hyperplane must also be inserted in evaluation of DMU$_o$.

If there does not exists another strong defining hyperplane except $H^*_o$ passing through DMU$_o$, then the algorithm will be terminated for DMU$_o$.

Suppose that the implementation of algorithm is repeated $t$-times for DMU$_o$. Therefore $t$ strong defining hyperplanes are determined. Note that in final step, model (9) will have exactly $t$ new constraints corresponding to $t$ strong defining hyperplanes determined in previous steps. So after the implementation of algorithm for DMU$_o$, all the strong defining hyperplanes of $T_v$ passing through it and all of extreme. BCC-efficient DMUs lying on these hyperplanes will be determined. For computational purposes it is better to have the algorithm in the next stage implemented for DMU$_p$ which the number of determined strong defining hyperplanes passing through it, is less than the number of strong defining hyperplanes passing through other OMUs. Note that in order to prevent model (9) from giving the gradients of previous determined strong defining hyperplanes...
that are simultaneously passing through DMU\_o and DMU\_p, in the implementation of algorithm for DMU\_p, the constraint I\_o = 1 always must be added to the constraints of model (9) in all stages.

In general, in the implementation of the algorithm for the rth extreme BCC-efficient DMU, the constraints I\_p = 1 corresponding to DMU\_p which algorithm is implemented for them up to now must be added to constraints of the model (9). Note that since there exists at least one linear independent set with m + s elements of extreme BCC-efficient DMUs lying on each FDEF of T\_v, (see theorem 3), if | E\_r | > | E | − (m + s) then the algorithm will be automatically terminated.

By considering the structure of the algorithm, the following lemma and theorem guarantee that the algorithm will give the gradients of all strong defining hyperplanes of T\_v before termination.

**Lemma 2.** Suppose that DMU\_p and DMU\_q are two extreme BCC-efficient DMUs lying on two distinct strong defining hyperplanes namely H\_p and H\_q (exclude their intersection). Then each strict convex combination of DMU\_p and DMU\_q is strong efficient, if it is not radial inefficient.

**Proof.** See Appendix.

**Theorem 5.** In the implementation of the mentioned algorithm for DMU\_o in set E, while there exists a strong defining hyperplane passing through DMU\_o, the optimal solution of model (9) for variables u and v will be positive.

**Proof.** See Appendix.

4 DEA-based clustering method introduced by Po et al.

The DEA method uses a piecewise linear approximation to the efficient frontier, which is determined by the efficient DMUs and envelopes (production functions) with different virtual multipliers u\^* and v\^* (u\^*y − v\^*x = 0). Therefore, there are different ways of combining inputs to yield outputs. The basic idea of the DEA-based clustering approach uses the different production functions derived from Model (3) to conduct a cluster analysis for a group of DMUs, which can be summarized as follows:

**Step1.** Evaluate the efficiency ratio for each DMU, find all production functions, and then identify the DMU whose efficiency ratio needs to be re-evaluated according to the following procedure:

- Let p = 0; Let PF(p) = φ and C(p) = φ.
- Let q = 0; Let R(q) = φ.

**LOOP** for k = 1 to n
Obtain the efficiency ratio, $\text{Eff}_k$, of the kth DMU and its solution of virtual multipliers $v_i^*$, $i = 1, \ldots, m$, and $u_r^*$, $r = 1, \ldots, s$, using the Model (1).

These obtained $v_i^*$ and $u_r^*$ will be in one of the following cases.

**Case 1.** $v_i^*$ and $u_r^*$ are both nonzero.

Derive the frontier of the production function with

$$ f(x_1; x_2; \ldots; x_m; y_1; y_2; \ldots; y_s) = \sum_{r=1}^s u_r^* y_r - \sum_{i=1}^m v_i^* x_i = 0. $$

**IF** the derived production function exists in one of $PF(1), \ldots, PF(p)$, say $PF(h)$, **THEN** the kth DMU is classified into the Cluster $C(h)$.

**ELSE** let $p = p + 1$. Assign $f(x_1; x_2; \ldots; x_m; y_1; y_2; \ldots; y_s)$ as a production function in $PF(p)$ and classify the kth DMU into the Cluster $C(p)$.

**Case 2.** $v_i^*$ and $u_r^*$ are both not nonzero.

This means the kth DMU $(x_{k1}, x_{k2}, \ldots, x_{km}, y_{k1}, y_{k2}, \ldots, y_{ks})$ is surrounded by an edge frontier. Thus, its efficiency ratio should be re-evaluate. Let $q = q + 1$. Assign the kth DMU to $R(q)$.

**ENDLOOP** (Now, there exist production functions in $PF(1), PF(2), \ldots, PF(p)$ with Clusters $C(1), C(2), \ldots, C(p)$ and there are $q$ DMUs in $R(1), R(2), \ldots, R(q)$ (surrounded by edge frontiers).

**Step 2.** Re-evaluate the efficiency ratio and reclassify the DMU surrounded by edge frontiers according to the following loop:

**LOOP** for $j = 1$ to $q$ Multiply the input items of the $R(j)$th DMU by $t$ and then substitute the $R(j)$th DMU data by $(tx_{j1}, tx_{j2}, \ldots, tx_{jm}, y_{j1}, y_{j2}, \ldots, y_{js})$.

**LOOP** for $w = 1$ to $p$

Take the $R(j)$th DMU data $(tx_{R(j)1}, tx_{R(j)2}, \ldots, tx_{R(j)m}, y_{R(j)1}, y_{R(j)2}, \ldots, y_{R(j)s})$ into $PF(w)$.

Obtain the value of $t$ such that the production function

$$ f(tx_{R(j)1}, tx_{R(j)2}, \ldots, tx_{R(j)m}, y_{R(j)1}, y_{R(j)2}, \ldots, y_{R(j)s}) = 0. $$

Let $t(w) = t$.

**ENDLOOP**

Take the index $k^*$ where $t(k^*) = \max\{t(1), t(2), \ldots, t(p)\}$.

Re-evaluate the efficiency ratio of the $R(j)$th DMU to be $t(k^*)$.

Assign the $R(j)$th DMU to be in the Cluster $C(k^*)$.

**ENDLOOP**

**Step 3.** Obtain the final clusters $C(1), C(2), \ldots, C(p)$. Moreover, the efficiency ratios $Eff_1, Eff_2, \ldots$,
$Eff_n$ for all DMUs are also obtained.

Table 1

<table>
<thead>
<tr>
<th>DMU</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>D5</th>
<th>D6</th>
<th>D7</th>
<th>D8</th>
<th>D9</th>
<th>D10</th>
<th>D11</th>
<th>D12</th>
<th>D13</th>
<th>D14</th>
<th>D15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>1</td>
<td>3/2</td>
<td>2</td>
<td>4</td>
<td>3/2</td>
<td>2</td>
<td>5/2</td>
<td>7/2</td>
<td>5</td>
<td>5/2</td>
<td>7/2</td>
<td>7/2</td>
<td>5</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Output</td>
<td>1</td>
<td>3/2</td>
<td>2</td>
<td>5/2</td>
<td>1/2</td>
<td>4/5</td>
<td>3/2</td>
<td>11/5</td>
<td>12/5</td>
<td>2/5</td>
<td>1</td>
<td>9/5</td>
<td>2</td>
<td>3/5</td>
<td>11/5</td>
</tr>
</tbody>
</table>

5 Proposed method

In the proposed method, we use the different strong defining hyperplanes derived from the multiplier form of the BCC model to conduct a cluster analysis for a group of DMUs. The method can be expressed in a series of steps as follows.

Step 1. We find all of the Strong Defining Hyperplanes of $T_v$ by using the introduced method in (3.1).

Consider $H_1, H_2, ..., H_l$ are all of the hyperplanes.
Let $C(j) = \phi$ for $(j = 1, ..., n)$.

LOOP for $k = 1$ to $n$

step 2. We will examine the coordinate of DMU$_i$ in all of $H_1, H_2, ..., H_l$ equations. Among all strong defining hyperplanes satisfied by DMU$_i$, we select the one that has the highest $u_o^*$, and consider that, it is $H_t$. If so,

classify the $i$th DMU into the Cluster $C(t)$.

If DMU$_i$ is satisfied in none of the $H_j$s, then we will go to step 3.

step 3.

We will obtain BCC-projection of DMU$_i$ by using definition (2.2), and then we will apply the step 2 on it.

Consider that, the result of the second step is the projection belongs to $C(p)$. In this manner, classify the

$i$th DMU into the Cluster $C(p)$.

ENDLOOP

Step 4. Obtain the final clusters $C(1), C(2), \ldots, C(p)$.

To illustrate the approach, we render an example below:

Example 1. Table 1 displays data for 15 DMUs with one input and one output. Figure 4.1 show them graphically. Clearly the units $D1, D3, \text{ and } D4$ are extreme BCC-efficient. Therefore, $E = \{D1, D2, D3\}$. For more explanations, the procedure is presented stop by stop.

Stage 1:

Let $E_T = \phi$ and put $D1 \in E$
Step 1-1. Evaluate $D_1$ by Model (9). $I_{D_1}^* = 1$ and $(u^*, v^*, u^*_{D_1}) = (1, 1, 0)$ is the optimal solution of Model (9). Since $(u^*, v^*) > 0)$, then $H_{D_1} : y - x = 0$ is the strong defining hyperplane of $T_v$; So $S = \{H_{D_1}\}$.

Furthermore $t_{D_3}^* = 0$, that is, $H_{D_1}$ passes through $D_3$, so set $J_{D_1} = \{D_1, D_3\}$, $J_{D_1}^c = \{D_4\}$ and construct the following inequality

$$\sum_{j \in J_{D_1}^c} |I_j| - \sum_{j \in J_{D_1}} |I_j| \leq I_{D_1}^* - 1 = 0 \quad (11)$$

Step 1-2. Add constraint (11) to the constraint of Model (9) and again evaluate $D_1$ by Model (9). $I_{D_1}^* = 1$ so the cardinal of the new set $J_{D_1}$ for $D_1$ is less than $m + s = 2$; therefore, the algorithm terminates for $D_1$.

Hence, $E_T = \{D_1\}$ and set $S$ does not change.

Add the constraint $I_{D_1} = 1$ to the constraint of Model (9) and perform the following steps:

Stage 2:

Step 2-1. Put $D_3 \in E - E_T$ and evaluate it by Model (9). $I_{D_3}^* = 1$ and $(u^*, v^*, u^*_{D_3}) = (4, 1, -6)$ is the optimal solution of Model (9). Since $(u^*, v^*) > 0)$, then $H_{D_3} : 4y - x - 6 = 0$ is the strong defining hyperplane of $T_v$; So $S = \{H_{D_1}, H_{D_3}\}$.

Table 2

<table>
<thead>
<tr>
<th>DMU</th>
<th>$\theta_{BCC}^*$</th>
<th>$s^{-*}$</th>
<th>$s^{**}$</th>
<th>BCC-Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>(1,1)</td>
</tr>
<tr>
<td>D2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>(3/2,3/2)</td>
</tr>
<tr>
<td>D3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>(2,2)</td>
</tr>
<tr>
<td>D4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>(4,5/2)</td>
</tr>
<tr>
<td>D5</td>
<td>2/3</td>
<td>0</td>
<td>1/2</td>
<td>(1,1)</td>
</tr>
<tr>
<td>D6</td>
<td>1/2</td>
<td>0</td>
<td>4/5</td>
<td>(1,1)</td>
</tr>
<tr>
<td>D7</td>
<td>3/5</td>
<td>0</td>
<td>0</td>
<td>(3/2,3/2)</td>
</tr>
<tr>
<td>D8</td>
<td>4/5</td>
<td>0</td>
<td>0</td>
<td>(14/5,11/5)</td>
</tr>
<tr>
<td>D9</td>
<td>18/25</td>
<td>0</td>
<td>0</td>
<td>(18/5,12/5)</td>
</tr>
<tr>
<td>D10</td>
<td>2/5</td>
<td>0</td>
<td>3/5</td>
<td>(1,1)</td>
</tr>
<tr>
<td>D11</td>
<td>2/7</td>
<td>0</td>
<td>0</td>
<td>(1,1)</td>
</tr>
<tr>
<td>D12</td>
<td>18/35</td>
<td>0</td>
<td>0</td>
<td>(9/5,9/5)</td>
</tr>
<tr>
<td>D13</td>
<td>2/5</td>
<td>0</td>
<td>0</td>
<td>(2,2)</td>
</tr>
<tr>
<td>D14</td>
<td>1/4</td>
<td>0</td>
<td>2/5</td>
<td>(1,1)</td>
</tr>
<tr>
<td>D15</td>
<td>7/15</td>
<td>0</td>
<td>0</td>
<td>(14/5,11/5)</td>
</tr>
</tbody>
</table>
Furthermore $t^*_D = 0$, that is, $H_D$ passes through $D$, so set $J_D = \{D, D4\}$, $J^*_D = \{D1\}$ and construct the following inequality

$$\sum_{j \in J^*_D} |I_j| - \sum_{j \in J_D} |I_j| \leq I^*_D - 1 = 0 \quad (12)$$

**Step 2-2.** Add constraint (12) to the constraint of Model (9) and again evaluate $D$ by Model (9). $I^*_D = 1$

so the cardinal of the new set $J_D$ for $D$ is less than $m + s = 2$ therefore the algorithm terminates for $D$.

Hence, $E = \{D1, D3\}$ and set $S$ does not change.

**Stage 3:** Since $|E| - |E_T| < m + s$ therefore the algorithm totally terminated.

Indeed, two strong defining hyperplanes are found $H_D1 : y - x = 0$ and $H_D3 : 4y - x - 6 = 0$. After that, we applied the BCC model with phase I and phase II processes and obtained the results displayed in Table 2.

By running the rest of the algorithm, we see that the BCC-Projection corresponding to DMUs D4, D8, D9 and D15 satisfy $H_D$ and the others satisfy $H_D1$. Therefore, the 15 DMUs are classified into the following two clusters (see Fig.4.1):

- **Cluster 1:** DMU1, DMU2, DMU3, DMU5, DMU6, DMU7, DMU10, DMU11, DMU12, DMU13, DMU14.

- **Cluster 2:** DMU4, DMU8, DMU9, DMU15.

![Figure 4.1. An illustration of Proposed approach in Example 1](image-url)
6 Comparison to Po et al.'s method.

Here, we carry out a comparison between the two approaches to clarify deviations and non-interfaces as follows:

(1) Po et al. obtain the solution of multipliers \( u^* \) and \( v^* \) for a DMU using the CCR model. If \((u^*, v^*) > 0\) then \( u^*y - v^*x = 0 \) is derived as the frontier of the production function. Since the structure of the envelopment form imposes strong degeneracy then the multiplier form produces alternative optimal solutions. However, this algorithm is designed regardless of this fact. We clarify the above using an example problem. For this purpose, we use five DMUs, A, B, C, D and E, each with two inputs and one output, as shown in Table 3. The CCR Model for these DMUs is illustrated by Figure 5.1, which depicts Input 1 and Input 2 values of all DMUs. Since the output value is 1 for all the DMUs, we can compare their efficiencies via the input values. The frontier of the production function consists of the bold line ABC. Using model 3, it can be seen that there are alternative optimal solutions which define an infinite number of hyperplanes passing through B. One of the optimal solution to evaluate DMU B is \((v^*_1 = 0.5, v^*_2 = 0.5, u^* = 1)\), while its related hyperplane, i.e., \( x_1 + x_2 = 2 \), the dotted line in Figure 5.1, is not a member of the piecewise production functions, of which only two hyperplanes \((H_1 \text{ and } H_2)\) are strong defining hyperplanes. Note that Po et al. (2009) argue that the obtained hyperplane must be portion of piecewise production functions. Now we intend to abolish this controversial problem. Throughout the method proposed in this paper, on finding the extreme BCC-efficient DMUs by Model (7), Model (9) is run on the extreme BCC-efficient DMUs, the result of which would be a thoroughly and inevitably strong defining hyperplane. To tell the truth, Model (9) yields the equation of the hyperplane that, among all strong defining hyperplanes, has the greatest number of efficient DMUs on it.

Table 3
The data set for the computational example.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input1</td>
<td>0.5</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Input2</td>
<td>2.5</td>
<td>1</td>
<td>0.5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Output</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 5.1. Hyperplane H is not defining. $H_1$ and $H_2$ are defining.

(2) In Po et al.’s approach, if the optimal weights in the multiplier form of CCR model are strictly positive, the corresponding hyperplane is taken into consideration as a reference for clustering. But, it is possible not to achieve any strictly positive ($u^*, v^*$) in evaluating all DMUs using CCR model. our method by means of the BCC-projection of an inefficient DMU conveniently.

(3) Some of the obtained clusters using the preceding approach may have the overlapping units. For more explanation, we consider Example 1 in Po et al. [Po, R.W., Guh, Y.Y., Yang, M.S., 2009. A new clustering approach using data envelopment analysis]. The example concludes with three clusters as follows:

Cluster 1: DMU 1, DMU 2, DMU 5, DMU 7, DMU 8, DMU 9, DMU 10.

Cluster 2: DMU 2, DMU 3, DMU 6, DMU 11, DMU 12.


As can be seen, DMU 2 belongs to both clusters 1 and 2. DMU 3 is also a member of clusters 2 and 3. In other words, the approach does not necessarily partition the set of DMUs. As in the method administered here, the result of executing Model (9) over
the extreme BCC-efficient DMUs is a strong defining hyperplane on which the greatest possible number of the extreme BCC-efficient DMUs would lie, in addition to the DMU being under evaluation, even if this result is not unique and separable. It is safe to say it excludes and highlights one that will have most $u_o^*$ (i.e., having highly upgraded return to scale).

7 Conclusions

In this paper, we suggested modifying the DEA-based clustering approach, recently introduced by Po et al.[22]. We showed that three problems exist in relation with the approach mentioned above. We suggested a method for the clustering of decision making units based on strong defining hyperplanes in DEA. The presented method can overcome the problems pointed out above. To illustrate the proposed approach, we introduced an examples. Finally, developing the proposed method to other DEA models suggested for further research.

Appendix

Proof of Theorem 1. We first note that DMU $j_0$ is extreme BCC-efficient if and only if the solution $s^- = 0, s^+ = 0, \lambda_{j_0} = 0; j = 1, ..., n, j \neq j_0, \theta_{j_0} = 1$ is the only feasible solution of model (4). Now suppose that DMU $j_0$ is extreme BCC-efficient. By contradiction if $\gamma^*_j > 0$, there exists an optimal solution of model (7) such that for at least some index $t, t \neq j_0, \lambda^*_t > 0$. This solution is also feasible for model (4) with $\theta_{j_0} = 1$. This is a contradiction. On the other hand, suppose that DMU $j_0$ is not extreme BCC-efficient. Then there exists a feasible solution $(\theta_{j_0}, \lambda, s^-, s^+)$ of model (2) such that for at least some index $t, t \neq j_0, \lambda_t > 0$. Either $\theta_{j_0} = 1$ or $\theta_{j_0} < 1$ the solution $(\lambda, s^-, s^+)$ is a feasible solution of model (7) and $\gamma^*_j > 0$.

Proof of Lemma 1. Let $(x_o, y_o)$ be an BCC-efficient DMU which lies on $H$. If $(x_o, y_o)$ is not an extreme point of the set $H \cap T_v$ then

$$(x_o, y_o) = \alpha(x_\alpha, y_\alpha) + \beta(x_\beta, y_\beta), \quad (13)$$

where $(x_\alpha, y_\alpha)$ and $(x_\beta, y_\beta)$ lies on $H \cap T_v$ and $\alpha + \beta = 1, \alpha, \beta > 0$. Since $(x_\alpha, y_\alpha)$ and $(x_\beta, y_\beta)$ are in $H \cap T_v$, so we have

$$(x_\alpha, y_\alpha) = \sum_{j=1}^{n} \lambda^\alpha_j (x_j, y_j), \quad \sum_{j=1}^{n} \lambda^\alpha_j = 1, \quad \lambda^\alpha_j \geq 0, \quad j = 1, ..., n,$$

$$(x_\beta, y_\beta) = \sum_{j=1}^{n} \lambda^\beta_j (x_j, y_j), \quad \sum_{j=1}^{n} \lambda^\beta_j = 1, \quad \lambda^\beta_j \geq 0, \quad j = 1, ..., n,$$
for some $\lambda^x$ and $\lambda^y$. By substituting in (13) we have $(x_o, y_o) = \sum_{j=1}^{n} \lambda_j (x_j, y_j)$ where $\lambda_j = \alpha \lambda^x_j + \beta \lambda^y_j$, $\sum_{j=1}^{n} \lambda_j = 1$, $\lambda_j \geq 0$, $j = 1, ..., n$. It is clear that $\lambda$ is an optimal solution of model (4) in which there exist some index $t \neq o$ with $\lambda_t > 0$. This is a contradiction. On the other hand, suppose that $v$ is an extreme point of the set $H \cap T_v$. Let $v = (v^x, v^y) \in R^{m+s} \geq 0$ where $v^x \in R^m \geq 0$ and $v^y \in R^s \geq 0$, we will show that $v$ is an extreme BCC-efficient DMU lying on $H$. In other words, there exists some index $k \in E$ where $v = DMU_k = (x_k, y_k)$. Otherwise, (if is not the case) there exists some $\hat{\lambda} = (\hat{\lambda}_1, ..., \hat{\lambda}_n)$ that:

$$ (v^x, v^y) = \sum_{j=1}^{n} \hat{\lambda}_j (x_j, y_j) \quad \sum_{j=1}^{n} \hat{\lambda}_j = 1, \quad \hat{\lambda}_j \geq 0, \quad j = 1, ..., n, \quad (14) $$

By using the above relation it is clear that for each positive $\hat{\lambda}_k \geq 0$, $DMU_k \in H$. Since $v$ is an extreme point of the set $H \cap T_v$. Hence there is not more than once positive $\hat{\lambda}_j$ in (14). Therefore, there is exactly one positive namely $\hat{\lambda}_l = 1$ and $v = DMU_l = (x_l, y_l)$. Now it is clear $DMU_l$ is an extreme BCC-efficient DMU.

**Proof of Theorem 2.** It is clear that $H \cap T_v$ is a convex polyhedral set. It is sufficient to show that $H \cap T_v$ has no recession direction. By contradiction, suppose that $d = (d^x, d^y)$ is a recession direction of it, where $d^x \in R^m$ and $d^y \in R^s$. By definition, we have $(x_o, y_o) + \alpha (d^x, d^y) \in H \cap T_v$, for each $\alpha \geq 0$ and $(x_o, y_o) \in H \cap T_v$. Since for each positive $\alpha$, $(x_o + \alpha d^x, y_o + \alpha d^y)$ is a BCC-efficient DMU lying on $H$, then for each $\alpha \geq 0$, there exists $\lambda^x$ such that:

$$ \sum_{j=1}^{n} \lambda^x_j (x_j, y_j) = (x_o + \alpha d^x, y_o + \alpha d^y) \quad \sum_{j=1}^{n} \lambda^x_j = 1, \quad \lambda^x_j \geq 0, \quad j = 1, ..., n. $$

Therefore, for each $\alpha \geq 0$, $(x_o + \alpha d^x, y_o + \alpha d^y)$ can be represented as the convex combination of $DMU_1, ..., DMU_n$. In other words for each $\alpha \geq 0$, $(x_o + \alpha d^x, y_o + \alpha d^y)$ belongs to the set $V=\text{conv}\{(DMU_1, ..., DMU_n)\}$. Since $\{DMU_1, ..., DMU_n\}$ is finite so $V$ is bounded. This is a contradiction.

**Proof of Theorem 3.** By previous Theorem, $H \cap T_v$ is a polytope, since $H$ is a strong defining, there exists at least one affine independent set with $m+s$ elements of the BCC-efficient DMUs in $H \cap T_v$. Therefore, $H \cap T_v$ is an $(m+s-1)$-polytope. Suppose that $DMU_1, ..., DMU_n = \text{ext}(H \cap T_v)$, then $H \cap T_v = \text{conv}\{(DMU_1, ..., DMU_n)\}$ By lemma 2 each $DMU_i$ is an extreme BCC-efficient DMU. It is clear that $k \geq m+s$. Since $H \cap T_v$ is an $(m+s-1)$-polytope, hence there is some affinely independent $(m+s)$-subfamily of $\{DMU_1, ..., DMU_n\}$, and the result is in hand.

**Proof of Theorem 4.** Suppose that $(u^x, v^*, u^o)$ is an optimal solution of model (9). Since there exists at least one strong defining hyperplane passing through $DMU_o$, then
by theorem (2) we have $I_o^* = |E| - k \leq |E| - (m + s)$. Consider the following model:

$$
\begin{align*}
\max & \quad u' y_o + u_o \\
\text{s.t.} & \quad v' x_o = 1, \\
& \quad u' y_j - v' x_j + u_o \leq 0, \quad j \in E, \\
& \quad u \geq 0, v \geq 0.
\end{align*}
$$

In fact model (15) is the BCC multiplier model with constraint restricted $j \in E$. It is clear that $(u^*, v^*, u_o^*)$ is an optimal solution of model (15). There are two cases:

Cases(i). $(u^*, v^*, u_o^*)$ is an extreme (basic feasible) optimal solution of model (15). Since $(u^*, v^*) > 0$, there exist $m + s$ linearly independent constraints of $u' y_j - v' x_j + u_o^* \leq 0, j \in E$ binding at $(u^*, v^*)$. Suppose that $u'^i y_j - v'^i x_j + u_o^* = 0, i = 1, ..., m + s$ is these constraints, then the following matrix is row full rank:

$$
\begin{pmatrix}
-x_{j1} & y_{j1} \\
-x_{j2} & y_{j2} \\
\vdots & \vdots \\
-x_{jm+s} & y_{jm+s}
\end{pmatrix}
$$

And it is row equivalent with the following matrix:

$$
\begin{pmatrix}
-x_{j1} & y_{j1} \\
x_{j1} - x_{j2} & y_{j1} - y_{j2} \\
\vdots & \vdots \\
x_{j1} - x_{jm+s} & y_{j1} - y_{jm+s}
\end{pmatrix}
$$

So the set $\{(x_{j1} - x_{j2}, y_{j1} - y_{j2})\}_{i=2}^{m+s}$ is liner independent. Hence there exist $m + s$ affinely independent of extreme BCC-efficient DMUs lying on $H_o^* \cap T_v$. Therefore, by definition (3), $H_o^* \cap T_v$ is a FDEF of $T_v$.

Cases(ii). $(u^*, v^*, u_o^*)$ is not an extreme (basic feasible) optimal solution of model (15). We will show that this won’t occur. Suppose that $(\bar{u}^i, \bar{v}^i), ..., (\bar{u}^h, \bar{v}^h)$ are the gradient of all the strong hyperplanes passing through $\text{DMU}_o$, where $(\bar{u}^i, \bar{v}^i) > 0, i = 1, ..., h$ and also $(\tilde{u}^1, \tilde{v}^1), ..., (\tilde{u}^l, \tilde{v}^l)$. It is clear that $(\bar{u}^i, \bar{v}^i, \bar{y}_o^i), i = 1, ..., h$ and $(\tilde{u}^i, \tilde{v}^i, \tilde{y}_o^i), i = 1, ..., l$ are all the extreme optimal solutions (Basic optimal feasible) of model (15). Since $(u^*, v^*, u_o^*)$ is not an extreme optimal solution of model (12), $(u^*, v^*)$ can be represented as a convex
combination of vectors \((\bar{w}^i, \bar{v}^i) > 0, i = 1, ..., h\) and \((\tilde{w}^i, \tilde{v}^i), i = 1, ..., l\). In other words:

\[
(u^*, v^*) = \sum_{i=1}^{h} \lambda_i (\bar{w}^i, \bar{v}^i) + \sum_{i=1}^{l} \tilde{\lambda}_i (\tilde{w}^i, \tilde{v}^i)
\]

\[
\sum_{i=1}^{h} \lambda_i + \sum_{i=1}^{l} \tilde{\lambda}_i = 1,
\]

\[
\lambda_i \geq 0, i = 1, ..., h \quad \tilde{\lambda}_i \geq 0, i = 1, ..., l.
\]

\[\tag{16}\]

There are two cases:

\textbf{Case(1).} There exists such a combination as (16) in which for some index \(r \in \{1, ..., h\}, \bar{\lambda}_r \neq 0\). So all of the extreme BCC-efficient DMUs lying on \(H^*_o\), are also lying on \(H^r : \bar{w} y - \bar{v} x + \bar{w}_o = 0\). And also lies on \(H^r\) only these extreme efficient DMUs. This is because if there exists another extreme BCC-efficient DMU lying on \(H^*_o\) in addition to those extreme BCC-efficient DMUs, lying on \(H^*_o\), then \((\bar{w}^r, \bar{v}^r, \bar{w}_o)\) is a feasible solution to model (9) where its objective value is less than \(I^*_o\), and this is a contradiction.

\textbf{Case(2).} There is no combination as (16) in which \(\lambda_i \neq 0\) for some index \(i \in \{1, ..., h\}\). In other word, in any combination such as (13), \(\bar{\lambda}_i = 0\), for all indices \(i\). Therefore, the strong face, \(H^*_o \cap T_v\), is not contained in any FDEF of \(T_v\). Again this is a contradiction. So \((u^*, v^*, u^*_o)\) is an extreme optimal solution of model (15), and this is an extreme optimal solution of model (9).

\textbf{Conclusion 1.} Suppose that DMU\(_o\) is an extreme BCC-efficient DMU and the vector \((u^*, v^*, u^*_o)\) is an optimal solution of model (9) in which \((u^*, v^*) > 0\), then it is an extreme (basic feasible) optimal solution of model (9) via the simplex method.

\textbf{Proof of Lemma 2.} Let DMU\(_l = \lambda\)DMU\(_p + (1 - \lambda)\)DMU\(_q\) where \(0 < \lambda < 1\). Suppose that DMU\(_l\) is not radial inefficient, then it is not an interior point of \(T_v\), so it lies on frontier of \(T_v\). If this frontier is strong, we are done. Otherwise it lies on weak frontier. Since DMU\(_p\) and DMU\(_q\) also lie on this frontier, and they belong to the reference set of DMU\(_l\). By theorem (3.4) in Cooper et al. (1999), each nonnegative combination of the elements of reference set is strong efficient, DMU\(_l\) is strong efficient. This is a contradiction.

\textbf{Proof of Theorem 5.} By virtue of the type of added constraints through the implementation of algorithm, it is sufficient to prove that the optimal solution of model (9) for variables \(u\) and \(v\) will be positive. By contradiction suppose that in the optimal solution of model (9), \((u^*, v^*, u^*_o)\), \((u^*, v^*)\) is not positive and \(I^*_o = |E| - k\). Since there exists at least one strong defining hyperplane passing through DMU\(_o\) then \(k \geq m + s\). Let \(R = \{\text{DMU}_{1}, ..., \text{DMU}_{k}\}\) be the set of all extreme BCC-efficient DMUs lying on \(H^*_o : u^*_o y - v^*_o x + u^*_o = 0\). Since \(H^*_o\) is a weak defining hyperplane, so each DMU\(_j\) \((j = 1, ..., k)\) lies on the intersection \(H^*_o\) with at least one strong defining hyperplane of \(T_v\). Suppose that \(H_1, ..., H_l\) are all these strong defining hyperplanes. There are two cases:
Case(I). There exists some index \( t, t \in \{1, ..., l\} \), such that all DMU\(_j\) \((j = 1, ..., k)\) lie on \( H_t \). Then, by considering the optimal value of objective function, there does not exist another extreme BCC-efficient DMU lying on \( H_t \). Since \( H_t \) is a strong defining hyperplane passing through DMU\(_o\), then the set \( R \) is affinely independent. This shows that there are two distinct \((m+s)-\)dimensional hyperplanes \( H^*_o \) and \( H_t \) passing through the set \( R \). This is a contradiction.

Case (II). There exists at least two DMUs in set \( R \), namely DMU\(_p\) and DMU\(_q\) lying on the intersection of \( H^*_o \) with two distinct strong defining hyperplanes of \( T_v \) namely \( H_p \) and \( H_q \) respectively. Since each convex combination of DMU\(_p\) and DMU\(_q\) lies on \( H^*_o \), then it is weak efficient. On the other hand by lemma (2) it is strong efficient. This is a contradiction.

References


Received: October, 2011