European Journal of Operational Research 199 (2009) 276–284

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

A new clustering approach using data envelopment analysis

Rung-Wei Po [a], Yuh-Yuan Guh [b,*], Miin-Shen Yang [c]

[a] Institute of Technology Management, National Tsing Hua University, 101, Section 2, Kuang-Fu Road, Hsinchu 30013, Taiwan, ROC
[b] Graduate School of Business Administration, Chung Yuan Christian University, ChungLi 32023, Taiwan, ROC
[c] Department of Applied Mathematics, Chung Yuan Christian University, ChungLi 32023, Taiwan, ROC

ARTICLE INFO

Article history:
Received 16 July 2007
Accepted 22 October 2008
Available online 5 November 2008

Keywords:
Data envelopment analysis
Production
Cluster analysis
CCR model
DEA-based clustering
Piecewise production function

ABSTRACT

In this paper, we present a new clustering method that involves data envelopment analysis (DEA). The proposed DEA-based clustering approach employs the piecewise production functions derived from the DEA method to cluster the data with input and output items. Thus, each evaluated decision-making unit (DMU) not only knows the cluster that it belongs to, but also checks the production function type that it confronts. It is important for managerial decision-making where decision-makers are interested in knowing the changes required in combining input resources so it can be classified into a desired cluster/class. In particular, we examine the fundamental CCR model to set up the DEA clustering approach. While this approach has been carried for the CCR model, the proposed approach can be easily extended to other DEA models without loss of generality. Two examples are given to explain the use and effectiveness of the proposed DEA-based clustering method.

1. Introduction

Cluster analysis is a branch in statistical multivariate analysis and an unsupervised learning in pattern recognition (see Duda and Hart, 1973; Kaufman and Rousseeuw, 1990; Jain et al., 2000). It is a method for classifying like groups of a data set into the same cluster and unlike groups into different clusters. Clustering is a powerful data exploratory approach to forming data groups and to revealing the feature structure information of a given data set. It is a data-driven procedure for classifying a datum in one of a few classes by looking at proximity and homogeneity in feature space. Generally, we may roughly divide clustering methods into the following categories: hierarchical clustering (Hartigan, 1975; Kaufman and Rousseeuw, 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).

Conventionally, most clustering algorithms are procedures that minimize total dissimilarity; examples of such algorithms are 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). Let $A_1, A_2, \ldots, A_s$ be the features of data, and the units to be clustered be $\text{DMU}_1, \text{DMU}_2, \ldots, \text{DMU}_n$ where $x_i = (x_{i1}, x_{i2}, \ldots, x_{is})$ is a feature vector for DMU_j in the s-dimensional Euclidean space R^s. Consider $d(x_j, z_i)$ as the dissimilarity measure between x_j and the cluster center z_i. A general clustering method is to find c cluster centers $z_1, z_2, \ldots, z_c$ so that the total dissimilarity measure $J_s(z)$ with $J_s(z) = \sum_{i=1}^{c}\sum_{j=1}^{n} a_{ij} f(d(x_j, z_i))$ is minimized. $J_s(z)$ is usually defined as a distance-based function, and the problem here is to select a useful and reasonable distance measure $d(x_j, z_i)$.

On the other hand, the stated clustering approaches can be seen as a feature analysis technique. An assumption of underlying feature analysis is to regard the feature items $A_1, A_2, \ldots, A_s$ as multiple features so that the minimization of $J_s(z)$ presents the closer of data among their features and makes it more possible for these DMUs to be classified into the same cluster. However, the clustering results derived from the minimization of the total feature dissimilarity $J_s(z)$ may not be helpful in some cases of clustering DMUs, especially in production units. In these cases, we use their production data to cluster them. Suppose that the production data have feature items $A_1, A_2, \ldots, A_k, A_{k+1}, \ldots, A_s$ with A_1 to A_k being input items and A_{k+1} to A_s being output items. The clustering information obtained from the conventional clustering approaches can only reveal DMU is more similar to another one. However, the more important information we want to know is the production feature (functions) implied from the production data of all DMUs. i.e., $f_j(A_1, A_2, \ldots, A_k; A_{k+1}, A_{k+2}, \ldots, A_s) = 0$. From these

* Corresponding author. Tel.: +886 3 2655126; fax: +886 3 2651399.
E-mail address: yuhyuan@cycu.edu.tw (Y.-Y. Guh).

0377-2217/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2008.10.022

derived production functions, $f_1, f_2, \ldots$, all DMUs are classified into different clusters (production functions). Therefore, each DMU not only knows the cluster that it belongs to, but also knows the production function type that it confronts. Each DMU can compare its production feature with the other production functions so that the combination of its input resources or the combination of inputs and outputs can be readjusted. That is, for the case of data feature with input and output items, the cluster derived from production functions is more valuable than that derived from feature dissimilarity measures.

The idea of this study is to employ the production functions to cluster production data. The method supporting this idea is data envelopment analysis (DEA), as initiated and developed by Charnes et al. (1978). The DEA is a data-oriented method for evaluating the relative efficiency of DMUs where each DMU is an entity responsible for converting multiple inputs into multiple outputs. Since the fundamental of DEA uses the nonparametric mathematical programming approach to estimate piecewise frontiers and envelop the DMU data sets. In this study, each piecewise frontier is regarded as one cluster of production functions. Therefore, we use all piecewise frontiers as a base to cluster production data. That is, we give up traditional clustering approaches of feature dissimilarity and propose a new approach by adopting the production functions revealed by the observation data to cluster all DMUs.

The rest of this paper is organized as follows: Section 2 discusses the CCR model from which the proposed clustering approach is developed. Furthermore, the piecewise linear convex hull is described to establish the fundamental DEA clustering approach. Section 3 looks into the proposed DEA-based clustering approach. We focus on why and how piecewise production functions drawn from DEA models are employed to cluster data. The algorithm of the DEA-based clustering approach is then established. Section 4 gives two numerical examples to illustrate the proposed DEA clustering approach. Discussion is made using this empirical example with a comparison of the resultant clusters derived from distance-defined clustering approaches. Moreover, a two-level clustering approach for production data is proposed by combining the cluster obtained from production functions and efficiency ratio. Finally, conclusions are stated in Section 5.

2. Data envelopment analysis

The DEA method is a useful tool for evaluating the relative efficiency for a group of DMUs. Up to now, DEA has been widely studied and applied in various areas for 30 years since Charnes et al. (1978) first proposed the DEA method with the CCR model. Among them, the main forms of DEA models and their extensions include those of BCC model (Banker et al., 1984), the additive model (Charnes et al., 1985) and the imprecise DEA models (Cooper et al., 1999; Zhu, 2003). Modifications and extensions are the assurance region models (Thompson et al., 1986; Zanakis et al., 2007), super-efficiency models (Andersen and Petersen, 1993; Li et al., 2007), cone ratio models (Charnes et al., 1989, 1990). Stochastic and chance-constrained extensions are considered by Land et al. (1994); Olesen and Petersen (1995); Cooper et al. (1996); Lahdelma and Salminen (2006) and Cooper et al. (2002). A taxonomy and general model frameworks for DEA can be found in Gattoufi et al. (2004) and Kleine (2004). The CCR is the original model of DEA (see the M1 model), and is used in this study to explain the DEA-based clustering approach. Without loss of generality, the proposed approach is also suitable for other models of the DEA family.

The DEA model generalizes the usual input/output ratio measure of efficiency for a given unit in terms of a fractional linear program formulation. According to the economic notion of Pareto optimality, the DEA method states that a DMU is considered inefficient if some other DMUs or some combinations of other DMUs produce at least the same amount of output with less of the same resources input and not more of any other resources. Conversely, a DMU is considered Pareto efficient if the above is not possible. Suppose there are n DMUs to be evaluated, x_{ij} is the noted amount of the ith input for the jth DMU and$y_{rj}$is the noted amount of the rth output for the jth DMU. With decision variables outputs weights $u_1, u_2, \ldots, u_s$ (one for each item of output) and input weights $v_1, v_2, \ldots, v_m$ (one for each item of input) being selected, the mathematical formulation of the method is summarized below, where the relative efficiency of the DMU$_k$ is to be determined (see the M1 model).

M1 model

$$\begin{aligned} &\mathit{Max}_{u,v} && \mathit{Eff}_k = \frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}} \\ &\mathit{Subject\ to} && \frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}} \leqslant 1, \\ &&& u_r \geqslant \varepsilon, v_i \geqslant \varepsilon\ \forall r, i. \end{aligned}$$

The DEA model is essentially a fractional programming problem with a ratio of a weighted sum of outputs to a weighted sum of inputs where the weights for both inputs and outputs are to be selected in a manner that calculates the efficiency of the evaluated unit. Therefore, the original form of the DEA model is both nonlinear and nonconvex problem. Charnes et al. (1981) proved that fractional programming problem can be transformed into two equivalent linear programming formulations. The first formulation is "input-based", constraining the weighted sum of outputs to be unity and minimizes the inputs that can then be obtained (see the M2 model). The second formulation is "output-based", constraining the weighted sum of inputs to be unity and maximizes the outputs that can then be obtained (see the M2′ model). The choice of using an input-based or output-based model depends on the production process characterizing the firm (that is, minimize the use of inputs to produce a given output or maximize the output with given levels of inputs):

M2 model

$$\begin{aligned} &\mathit{Min}_{u,v} && \mathit{Eff}_k = \sum_{i=1}^{m} v_i x_{ik} \\ &\mathit{Subject\ to} && \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leqslant 0 \quad j = 1, \ldots, n, \\ &&& \sum_{r=1}^{s} u_r y_{rj} = 1, \\ &&& u_r \geqslant \varepsilon,\ \ v_i \geqslant \varepsilon\ \forall r, i. \end{aligned}$$

M2′ model

$$\begin{aligned} &\mathit{Max}_{u,v} && \mathit{Eff}_k = \sum_{r=1}^{s} u_r y_{rk} \\ &\mathit{Subject\ to} && \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leqslant 0 \quad j = 1, \ldots, n, \\ &&& \sum_{i=1}^{m} v_i x_{ik} = 1, \\ &&& u_r \geqslant \varepsilon,\ \ v_i \geqslant \varepsilon\ \forall r, i. \end{aligned}$$

Both M2 and M2′ models are linear programming forms of the DEA method. It is implicit that the methodology employed by the DEA method is the production function theory. In economics, the production function is a function that summarizes the process of converting multiple inputs into a single output. Thus, a general mathematical form for the production function in economics can be expressed as $y = f(x_1, x_2, x_3, \ldots, x_n)$, where y is a quantity of output and $x_1, x_2, x_3, \ldots, x_n$ are quantities of inputs. However, the DEA is a process of converting multiple inputs into multiple outputs, i.e., $g(y_1, y_2, y_3, \ldots, y_n) = f(x_1, x_2, x_3, \ldots, x_n)$. In fact, we can see that both M2 and M2′ models with the constraints $\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leqslant 0, j = 1, \ldots, n$, use the production function that converts multiple inputs into multiple outputs. Most previous studies had mentioned and discussed the properties of production function that are hidden in DEA methods (see Charnes et al., 1983; Banker et al., 1984; Seiford and Thrall, 1990; Chang and Guh, 1991; Andersen and Petersen, 1993; Olesen and Petersen, 1995; Cooper et al., 1996, 2002, 2007; Huang et al., 1997; Pitaktong et al., 1998; Zanakis et al., 2007 and Li et al., 2007).

Since the number of DMUs is usually much larger than the number of inputs, we prefer to express the linear programming in its duality form. Further, the duality form can interpret the geometric meaning of DEA and provide information about conservation of resources or expansion of outputs to have DMUs from inefficiency to efficiency. Therefore, we prefer to have its duality form as follows (see the M3 and M3′ model):

M3 model

$$\begin{aligned} &\textit{Min} && \textit{Eff}_k = (\Phi_k - \varepsilon \sum_{i=1}^{m} s_i^- - \varepsilon \sum_{i=1}^{s} s_r^+) \\ &\textit{Subject to} && x_{ik}\Phi_k - \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^- = 0 \quad i = 1, \ldots, m, \\ &&& \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = y_{rk} \; j = 1, \ldots, n, \\ &&& \lambda_j, s_i^-, s_r^+ \geqslant 0 \text{ for all } j, i, r. \end{aligned}$$

M3′ model

$$\begin{aligned} &\textit{Max} && \textit{Eff}_k = (\Phi_k + \varepsilon \sum_{i=1}^{m} s_i^- + \varepsilon \sum_{i=1}^{s} s_r^+) \\ &\textit{Subject to} && y_{rk}\Phi_k - \sum_{j=1}^{n} y_{rj}\lambda_j + s_r^+ = 0 \quad j = 1, \ldots, n, \\ &&& \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- = x_{ik} \; i = 1, \ldots, m, \\ &&& \lambda_j, s_i^-, s_r^+ \geqslant 0 \text{ for all } j, i, r. \end{aligned}$$

The above shows three types of CCR model. If $\textit{Eff}_k^*$ is the optimal value of $\textit{Eff}_k$, the DMU_k is said to be efficient if and only if $\textit{Eff}_k^* = 1$. If $\textit{Eff}_k^*$ is less than 1, DMU_k is inefficient. According to the efficiency ratio, DMUs may be grouped as good ($\textit{Eff}_k^* = 1$) and poor ($\textit{Eff}_k^* < 1$) performers, or clustered by assigning different efficiency ratio grades (see Yu et al., 1996; Thompson et al., 1997; Jahanshahloo et al., 2005; Bick et al., 2007; Cook and Bala, 2007). Although clustering by efficiency ratio gives some information about the rationality of output/input, it does not reveal the intrinsic relationship between the input and output production features. Therefore, this study adopts piecewise production functions derived from the DEA method to cluster data.

In M2 and M2′ models, it is obvious that the constraint $\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leqslant 0$ is an inequality formula of production functions. Solving M2 and M2′ models yields the virtual multipliers u_r^* and v_i^*. Thus, $\sum_{r=1}^{s} u_r^* y_{rj} - \sum_{i=1}^{m} v_i^* x_{ij} = 0$ is derived. Running either the M2 or M2′ model for $k = 1$ to n gives all production functions. Then, all DMUs are classified into different clusters by these piecewise production functions. Thus, a clustering method using production functions via the DEA method is implemented.

The piecewise linear convex hull approach to frontier estimation proposed by Farrel (1957) provides a non-parametric method for determining the relative efficiency of a DMU. Further works on identification of the empirically defined production possibility includes Charnes et al. (1982, 1983, 1987), Banker et al. (1984), Seiford and Thrall (1990), Jahanshahloo et al. (2007). However, they basically use Pareto-efficiency to generate these reference sets and describes DEA by floating a piecewise linear surface to rest on top of the observations (i.e., envelop the data). Suppose a simple model is erected with the input–output observations $(X_1, Y_1), \ldots, (X_n, Y_n)$ for each DMU. These DMUs for efficiency comparisons are assumed to use the same inputs and also to produce the same outputs even though it may be in varying amounts. Our objective is to characterize a production possibility set and, in particular to determine an efficient frontier according to these observed data. T is a production possibility set which has following properties:

$$T = \{(X, Y) | Y \geqslant 0 \text{ can be produced from } X \geqslant 0\}.$$

Postulate 1. *Convexity*. If $(X_j, Y_j) \in T$, $j = 1, \ldots, n$, and $\lambda_j \geqslant 0$ are nonnegative scalars such that $\sum_{j=1}^{n} \lambda_j = 1$, then $(\sum_{j=1}^{n} \lambda_j X_j, \sum_{j=1}^{n} \lambda_j Y_j) \in T$.
Postulate 2. *Inefficiency*. (a) If $(X_j, Y_j) \in T$ and $\overline{X} \geqslant X$ then $(\overline{X}, Y) \in T$. (b) If $(X_j, Y_j) \in T$ and $\overline{Y} \leqslant Y$ then $(X, \overline{Y}) \in T$.
Postulate 3. *Ray unboundness.* If $(X,Y) \in T$ then $(kX,kY) \in T$ for any $k > 0$.
Postulate 4. *Minimum extrapolation.* T is the intersection set of all $\widehat{T}$ satisfying Postulates 1, 2 and 3 and subject to the condition that each of the observed vectors $(X_j, Y_j) \in \widehat{T}$, $j = 1, \ldots, n$.

The slope along the piecewise efficient frontier of the production possibility set denotes different rates of change in outputs with respect to changes in inputs. Chang and Guh (1991), Huang et al. (1997), Pitaktong et al. (1998) and Cooper et al. (2007) had developed methods for identifying facet members of the Pareto-optimal frontier. The piecewise efficient facet stated by these authors has important implications for effective management of the resources employed to obtain desired feasible outputs. In particular, Huang et al. (1997) developed a series of linear programming for determining rates of change in facets. Up to date, we find that there is less consideration in using these facets (production function) as a reference to classify evaluated DMUs. In this study, we shall propose a clustering approach according to the properties of DEA and its production possibility set such that we can use these facets (production function) as a reference to classify evaluated DMUs. This DEA-based clustering method will be derived in the next section.

3. DEA-based clustering method

As stated above, the basic idea of DEA-based clustering approach uses the piecewise production functions derived from the M2 or M2′ model to conduct a cluster analysis for a group of DMUs. In this section, we will further explain the DEA-based clustering approach using the following demonstration.

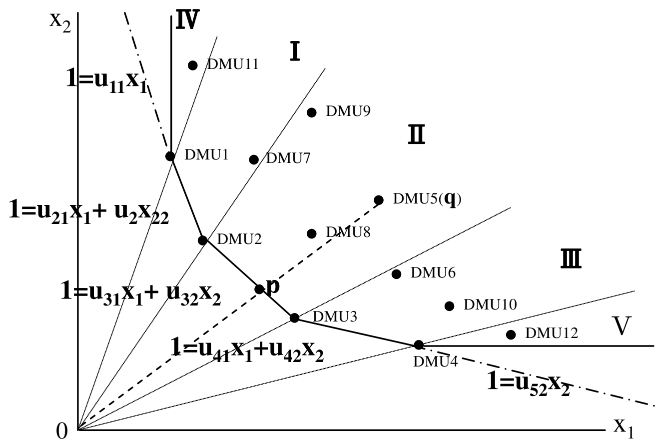

Fig. 1. An illustration of DEA-based clustering approach (DEA isoquant: Combination of x_1 and x_2 for producing one unit of output).

As illustrated in Fig. 1, the DEA method uses a piecewise linear approximation to the efficient frontier, which is determined by the efficient DMUs (DMU_1, DMU_2, DMU_3, DMU_4) and five envelopes (production functions) with different virtual multipliers u_i^* ($1 = u_{11}^* x_1$, $1 = u_{21}^* x_1 + u_{22}^* x_2$, $1 = u_{31}^* x_1 + u_{32}^* x_2$, $1 = u_{41}^* x_1 + u_{42}^* x_2$ and $1 = u_{52}^* x_2$). The slope of each segment of envelope determines the substitution possibility for a DMU on this local frontier to produce one unit of output. For example, the input substitution rate of the envelope $1 = u_{21}^* x_1 + u_{22}^* x_2$ is $-u_{21}^*/u_{22}^*$. Therefore, there are five different ways of combining inputs to yield outputs. For example, a product is fabricated by the machine (x_1) and manpower (x_2). If DMU_i and DMU_j use the production functions $1 = u_{21}x_1 + u_{22}x_2$ and $1 = u_{41}x_1 + u_{42}x_2$, respectively, to yield one unit of product, DMU_i is a labor-oriented industry and the DMU_j is a capital-oriented industry because $u_{11} \geqslant u_{21} \geqslant u_{31} \geqslant u_{41} \geqslant u_{51} = 0$ and $u_{52} \geqslant u_{42} \geqslant u_{32} \geqslant u_{22} \geqslant u_{12} = 0$. Thus, five clusters are established in this example and each piecewise envelope represents one type of production. Each DMU is clustered according to its corresponding production function:

Cluster I: DMU_1, DMU_2, DMU_7.
Cluster II: DMU_2, DMU_3, DMU_5, DMU_8, DMU_9.
Cluster III: DMU_3, DMU_4, DMU_6, DMU_{10}.
Cluster IV: DMU_1, DMU_{11}.
Cluster V: DMU_4, DMU_{12}.

Among these clusters, the virtual multipliers of production functions corresponding to Clusters I, II and III are all nonzero. This information is important for managerial decision-making while a DMU is interested in knowing its production feature relative to other DMUs, and refers to other production function features, which provide a direction to readjust the combination of its input resources, and/or the combination of inputs and outputs, so as to be reassigned into a desired cluster/class. However, the virtual multipliers of production functions corresponding to Clusters IV and V are not all nonzero, for example $1 = u_{11}x_1$ and $1 = u_{52}x_2$, these frontiers cannot be considered as effective clusters because they are degenerative, and there exists no substitution rate between input and output items. Thus, the inefficient DMUs (DMU_{11} and DMU_{12}) belonging to degenerative clusters will be reclassified into another effective cluster whose production functions are with all nonzero virtual multipliers. According to the minimum extrapolation postulate of the possibility production set, the nearest frontier of production function that the inefficient DMU confronts toward the original point should be considered as its cluster. In this illustration, both DMU_{11} and DMU_{12} are re-clustered to the frontier $1 = u_{21}^* x_1 + u_{22}^* x_2$ (Cluster I) and $1 = u_{41}^* x_1 + u_{42}^* x_2$ (Cluster III), respectively. Hence, the clusters with their DMU classification shown in Fig. 1 reduce to the following three types:

Cluster I: DMU_1, DMU_2, DMU_7, DMU_{11}.
Cluster II: DMU_2, DMU_3, DMU_5, DMU_8, DMU_9.
Cluster III: DMU_3, DMU_4, DMU_6, DMU_{10}, DMU_{12}.

Fig. 1 shows the geometric meaning of efficiency ratio determined by the DEA method. For example, DMU_5 is inefficient relative to the reference set of DMU_2 and DMU_3. The dashed line from q to the origin represents the contraction path for DMU_5. By connecting the piecewise envelope segment from DMU_2 to DMU_3, the efficiency ratio (Eff_5^*) of DMU_5 is evaluated as $\bar{o}\bar{p}/\bar{o}\bar{q}$ and obtained from the implementation of any one model of M1, M1′, M2 and M2′. However, if a DMU belongs to a degenerative cluster, we had discussed that it will be reclassified into a new cluster (the nearest frontier of production function), and thus its efficiency ratio will be re-evaluated by this frontier. For example, DMU_{11} and DMU_{12} are inefficient and belong to degenerative clusters initially. After being re-clustered to their nearest effective clusters $1 = u_{21}x_1 + u_{22}x_2$ and $1 = u_{41}x_1 + u_{42}x_2$, the efficiency ratios of DMU_{11} and DMU_{12} are re-evaluated by these piecewise envelope segments, respectively. The re-clustering and reevaluating algorithm is shown in the next section.

It is noted, according to DEA cluster analysis, the cluster for each DMU is identified. However, if the evaluated DMU falls over the intersection point of frontiers, it will be attributed to these clusters of frontiers simultaneously. For example, DMU_2 is classified into Clusters I and II for its location is at the intersection point of frontiers $1 = u_{21}^* x_1 + u_{22}^* x_2$ and $1 = u_{31}^* x_1 + u_{32}^* x_2$. Thus, the algorithm of the DEA-based clustering approach can be summarized as follows:

3.1. DEA-based clustering algorithm

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 $PF(p) = \phi$ and $C(p) = \phi$.

Let $q = 0$; Let $R(q) = \phi$.

LOOP for $k = 1$ to n

Obtain the efficiency ratio 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 M2 or M2′ model. 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 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.

It 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, it should re-evaluate its efficiency ratio. 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 $R(j)$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),…,C(p). Moreover, the efficiency ratios Eff_1, $Eff_2,\ldots,Eff_n$ for all DMUs are also obtained.

We mention that the proposed DEA-based clustering algorithm can systematically choose the effective clusters (the production functions whose virtual multipliers are all nonzero) and cancel the degenerative clusters simultaneously. The algorithm re-classifies the DMU which confronts the degenerative cluster (frontier) to the nearest effective cluster (frontier) toward the original point, and re-evaluate its efficiency ratio using this effective frontier.

It is also noted, as the numbers of inputs (m) and outputs (s) of DEA problem grow, the number of piecewise production functions (clusters) may increase drastically but most are degenerative (there exists $C_1^{m+s} + C_2^{m+s} + \cdots + C_{m+s-1}^{m+s}$ possible degenerative production functions). Consequently, it will possibly cause most DMUs to be clustered into separated and degenerative clusters, and thus the cluster classification is no longer meaningful. It is exactly the purpose of the proposed clustering algorithm, which takes advantage of the piecewise production function of DEA to cluster evaluated units while avoiding the disturbance of degenerative clusters.

4. Numerical examples

We now examine two numerical examples to demonstrate the DEA-based clustering approach and then to illustrate its applications in the real world.

Example 1. Consider an efficiency evaluation problem with 20 DMUs, each DMU with two inputs and one output. The simplified production data of DMU (input1, input2, output1) are shown as follows:

$$\begin{array}{lllll} DMU_1(1,5,1) & DMU_2(2,3,1) & DMU_3(3,2,1) & DMU_4(5,1,1) & DMU_5(2,5,1), \\ DMU_6(3,4,1) & DMU_7(3,8,1) & DMU_8(4,8,1) & DMU_9(5,9,1) & DMU_{10}(4,10,1), \\ DMU_{11}(6,5,1) & DMU_{12}(7,5,1) & DMU_{13}(7,4,1) & DMU_{14}(7,3,1) & DMU_{15}(8,4,1), \\ DMU_{16}(9,2,1) & DMU_{17}(10,3,1) & DMU_{18}(11,3,1) & DMU_{19}(10,1.5,1) & DMU_{20}(11,2,1). \end{array}$$

By using the M2 or M2′ model, for each DMU$_k$, its efficiency ratio Eff_k and the solution of virtual multipliers v_1^*, v_2^* and u_1^* are obtained. The analytical results are shown in Table 1.

By selecting the set of virtual multipliers v_1^*, v_2^* and u_1^* to be all nonzero, three frontiers of production functions are found, PF(1) = $y - 2/7x_1 - 1/7x_2 = 0$, PF(2) = $y - 1/7x_1 - 2/7x_2 = 0$ and PF(3) = $y - 1/5x_1 - 1/5x_2 = 0$. Therefore, the 20 DMUs are classified into the following three clusters (see Fig. 2):

Table 1
Analytical results derived from M2 or M2′ model in Example 1.

	Virtual multipliers			Efficiency ratio (Eff_k)	Evaluated by the frontier of
	v_1^*	v_2^*	u_1^*		
$DMU_1(1,5,1)$	2/7	1/7	1	1.0000000	$y = 2/7x_1 + 1/7x_2$
$DMU_2(2,3,1)$	2/7	1/7	1	1.0000000	$y = 2/7x_1 + 1/7x_2$
	1/5	1/5	1		$y = 1/5x_1 + 1/5x_2$
$DMU_3(3,2,1)$	1/7	2/7	1	1.0000000	$y = 1/7x_1 + 2/7x_2$
	1/5	1/5	1		$y = 1/5x_1 + 1/5x_2$
$DMU_4(5,1,1)$	1/7	2/7	1	1.0000000	$y = 1/7x_1 + 2/7x_2$
$DMU_5(2,5,1)$	2/9	1/9	7/9	0.7777778	$y = 2/7x_1 + 1/7x_2$
$DMU_6(3,4,1)$	1/7	1/7	5/7	0.7142857	$y = 1/5x_1 + 1/5x_2$
$DMU_7(3,8,1)$	1/7	5/70	1/2	0.5000000	$y = 2/7x_1 + 1/7x_2$
$DMU_8(4,8,1)$	1/8	5/80	7/40	0.4375000	$y = 2/7x_1 + 1/7x_2$
$DMU_9(5,9,1)$	2/19	1/19	7/19	0.3684211	$y = 2/7x_1 + 1/7x_2$
$DMU_{10}(4,10,1)$	1/9	5/90	7/18	0.3888889	$y = 2/7x_1 + 1/7x_2$
$DMU_{11}(6,5,1)$	1/11	1/11	5/11	0.4545455	$y = 1/5x_1 + 1/5x_2$
$DMU_{12}(7,5,1)$	5/60	5/60	5/12	0.4166667	$y = 1/5x_1 + 1/5x_2$
$DMU_{13}(7,4,1)$	2/30	2/15	7/15	0.4666667	$y = 1/7x_1 + 2/7x_2$
$DMU_{14}(7,3,1)$	1/13	2/13	7/13	0.5384615	$y = 1/7x_1 + 2/7x_2$
$DMU_{15}(8,4,1)$	5/80	1/8	7/16	0.4375000	$y = 1/7x_1 + 2/7x_2$
$DMU_{16}(9,2,1)$	1/13	2/13	7/13	0.5384615	$y = 1/7x_1 + 2/7x_2$
$DMU_{17}(10,3,1)$	5/80	1/8	7/16	0.4375000	$y = 1/7x_1 + 2/7x_2$
$DMU_{18}(11,3,1)$	1/17	2/17	1/17	0.4117647	$y = 1/7x_1 + 2/7x_2$
$DMU_{19}(10,1.5,1)$	0	2/3	2/3	0.6666667	$y = x_2$
				(0.53846)	Re-evaluated by $y = 1/7x_1 + 2/7x_2$
$DMU_{20}(11,2,1)$	0	1/2	1/2	0.5000000	$y = x_2$
				(0.4666667)	Re-evaluated by $y = 1/7x_1 + 2/7x_2$

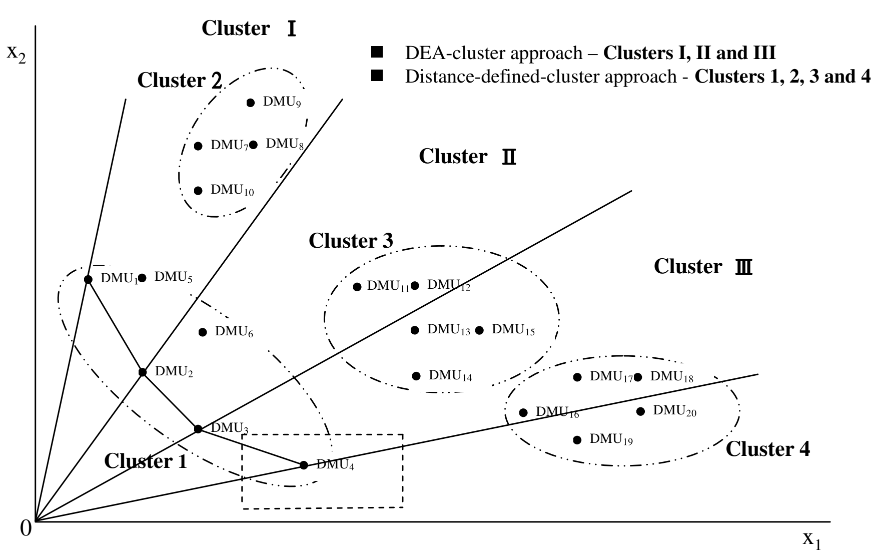

Fig. 2. Resultant clusters derived from DEA-cluster and distance-defined clustering approaches for Example 1. (DEA isoquant: combining x_1 and x_2 for producing one unit of output).

Cluster I: DMU_1, DMU_2, DMU_5, DMU_7, DMU_8, DMU_9, DMU_{10},
Cluster II: DMU_2, DMU_3, DMU_6, DMU_{11}, DMU_{12},
Cluster III: DMU_3, DMU_4, DMU_{13}, DMU_{14}, DMU_{15}, DMU_{16}, DMU_{17}, DMU_{18}, DMU_{19}, DMU_{20}.

However, if we use the general clustering approaches such as the distance-defined *c*-mean, the 20 DMUs will be classified into the following four clusters (see Fig. 2):

Cluster 1: DMU_1, DMU_2, DMU_3, DMU_4, DMU_5, DMU_6,
Cluster 2: DMU_7, DMU_8, DMU_9, DMU_{10},
Cluster 3: DMU_{11}, DMU_{12}, DMU_{13}, DMU_{14}, DMU_{15},
Cluster 4: DMU_{16}, DMU_{17}, DMU_{18}, DMU_{19}, DMU_{20}.

Fig. 2 points out a significant difference between the clusters derived from the DEA-cluster approach and distance-defined clustering approach. For example, according to the distance-defined clustering approach, DMU_1, DMU_2, DMU_3, DMU_4, DMU_5, and DMU_6 are classified into Cluster 1. On the contrary, according to the DEA-cluster approach, these six DMUs belong to three clusters of production functions.

That is, DMU_1, DMU_2 and DMU_5 are classified into Cluster I; DMU_2, DMU_3 and DMU_6 are classified into Cluster II; and DMU_4 and DMU_5 are classified into Cluster III.

Why is there discrepancy between results derived from these two approaches? It is because the distance-defined clustering approach ignores the input and output relationship between the features, and regards all items as multiple attributes. Thus, the DMUs are classified into the same cluster if their data attributes are closer. This result will give us an incorrect message that these DUMs have the same or similar production features. Indeed, they may belong to different production types. Hence, the clustering results derived from the DEA-based clustering approach using the production functions reveal the input–output relationships hidden in the feature items of inputs and outputs, so it is more meaningful and helpful for production units.

DMU_{19} and DMU_{20} confront the degenerative frontier ($y = x_2$). This study suggests that they should be reclassified into the nearest effective frontier (the frontier with nonzero virtual multipliers). In this example, it is observed DMU_{19} and DMU_{20} confront the nearest effective frontier $y = 1/7x_1 + 2/7x_2$, thus their efficiency ratio will be re-evaluated by this frontier. However, in complicated problems (with more data items of input and output), it is impossible to judge the nearest effective frontier by observation. Hence, for DMU_{19}, we follow the procedure of Step 2 stated in Section 3, taking $(x_1, x_2, y) = (10t(i), 1.5t(i), 1)$, $i = 1, 2, 3$ into $\text{PF}(1) = y - 2/7x_1 - 1/7x_2 = 0$, $\text{PF}(2) = y - 1/7x_1 - 2/7x_2 = 0$ and $\text{PF}(3) = y - 1/5x_1 - 1/5x_2 = 0$, respectively. The $t(i)$ value is calculated, giving $t(1) = 0.32558$, $t(2) = 0.53846$ and $t(3) = 0.43483$. By taking the maximal value, the efficiency ratio for DMU_{19} is re-evaluated as 0.53846. In addition, DMU_{19} is classified into the cluster determined by the corresponding envelope $\text{PF}(2) = y - 1/7x_1 - 2/7x_2 = 0$.

Finally, to provide more information by clustering, we combine production function and efficiency ratio to propose a two-level cluster for production data. The first level is according to the production function that the evaluated DUMs confront. The second level is according to the efficiency ratio, which is divided into good ($Eff_k^* = 1$) and poor ($Eff_k^* < 1$) performance. The resultant clusters for these 20 DMUs are shown in Fig. 3.

As to the stability of resultant clusters derived from the DEA-based clustering method, it is determined by the proposed method whether or not it is robust to a slight change in the input–output data, and thus retains the existing reference set of frontiers (production functions). As seen in Fig. 1, the production function $y = 1/7x_1 + 2/7x_2$ is created by connecting the efficient DMU_4 and DMU_3. If the imprecise input data of DMU_4 are within the scope of the dashed-line rectangle, then the clustering result shown in Fig. 3 still holds. That is, the proposed DEA-based clustering algorithm is robust to a slight change in the input and output data sets. However, if the imprecise input data are beyond the scope of the dashed-line rectangle, or the input data are an outlier, they will change the reference set of frontiers. Hence, it is important to check data correction or diminish the bias of data priori to implement the DEA-based clustering method. We mention that this kind of sensitivity to outliers is always a problem for most clustering algorithms (see Bezdek, 1981; Yang, 1993; Wu and Yang, 2002). In clustering literature, several authors had discussed the robustness for clustering (see Jolion et al., 1991; Dave and Krishnapuram, 1997). The robustness of the DEA-based clustering algorithm will be another interesting research issue. In fact, a robust algorithm to outliers using the DEA approach merits further study.

Example 2. This example has 15 DMUs. Each DMU also has two inputs (x_1 and x_2) and one output (y) as shown in Fig. 4, in which the data are listed as follows:

$$\begin{array}{lllll}
\text{DMU}_{D1}(6,5,1) & \text{DMU}_{D2}(8,5,1) & \text{DMU}_{D3}(5,2,1) & \text{DMU}_{D4}(6,3,1) & \text{DMU}_{D5}(8,2,1), \\
\text{DMU}_{E1}(18,15,3) & \text{DMU}_{E2}(24,15,3) & \text{DMU}_{E3}(15,6,3) & \text{DMU}_{E4}(18,9,3) & \text{DMU}_{E5}(24,6,3), \\
\text{DMU}_{F1}(30,25,5) & \text{DMU}_{F2}(40,25,5) & \text{DMU}_{F3}(25,10,5) & \text{DMU}_{F4}(30,15,5) & \text{DMU}_{F5}(40,10,5).
\end{array}$$

According to the distance-based clustering method, these data can be divided into the three groups: **Cluster 1** (DMU_{D1}–DMU_{D5}), **Cluster 2** (DMU_{E1}–DMU_{E5}) and **Cluster 3** (DMU_{F1}–DMU_{F5}), Suppose the enveloped frontier of DMU_{D1}–DMU_{D5} is $u_1^*y = v_1^*x_1 + v_2^*x_2$. Since the data of DMU_{E1}–DMU_{E5} are two times those of DMU_{D1}–DMU_{D5}, respectively, the enveloped frontier of DMU_{E1}–DMU_{E5} is $2u_1^*y = 2v_1^*x_1 + 2v_2^*x_2$, the same as the enveloped frontier of DMU_{D1}–DMU_{D5}. Thus, according to the DEA-cluster approach, DMU_{E1}–DMU_{E5} and DMU_{D1}–DMU_{D5} are classified into the same cluster. Similarly, DMU_{F1}–DMU_{F5} and DMU_{D1}–DMU_{D5} are also classified into the same cluster. That is, these three groups of DMUs form one cluster (**Cluster I**). However, if we use the distance-function clustering approaches, they will be classified into three different clusters.

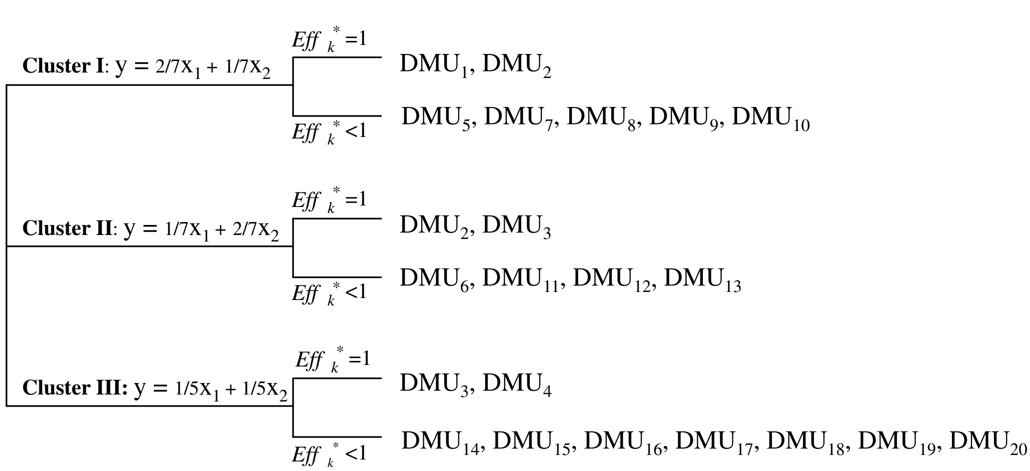

Fig. 3. Two-level cluster for production data.

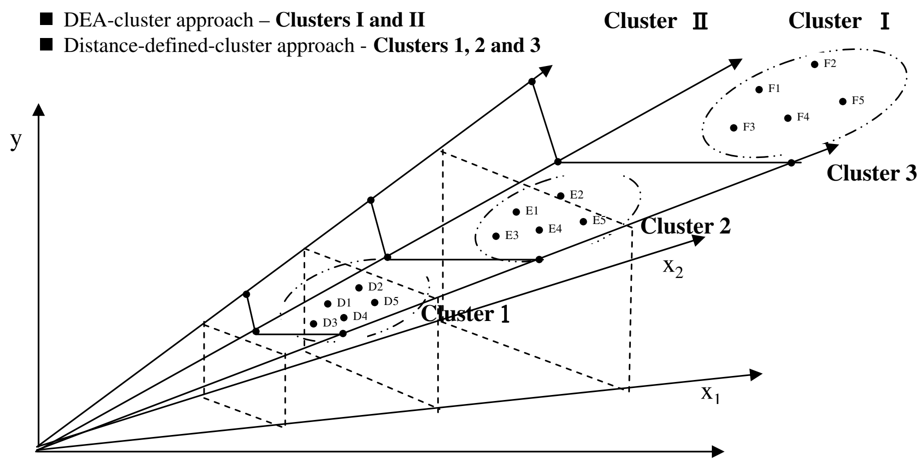

Fig. 4. Resultant clusters derived from DEA-cluster and distance-defined clustering approaches for Example 2.

This is an interesting result for cluster analysis. It means that the DEA-cluster result is not affected by the scale of data, thus the three groups of DMUs are classified into the same cluster. The reason is that the CCR model of the DEA method has constant return to scale (the production function has no constant term), so the model can automatically diminish the "multiple" effect among data. It is so-called "unit invariant" (invariant to the units of measurement chosen); that is, changing the units of measurement (for example, measuring the quantity of labors in person-hours instead of person-years) will not change the cluster it belongs to or its efficiency ratio. Nevertheless, the unit-invariant property may not exist in other DEA models (for example, the BCC model).

To summarize the analytical results from the above examples, the DEA-cluster approach not only clusters data, but also provides production functions to describe the relationship between the feature items. The proposed approach is suitable for clustering data which contain input and output items.

5. Conclusion

This study develops a DEA-based clustering approach. The proposed approach employs the piecewise production functions derived from the DEA method to cluster the data with input and output items. Compared with distance-defined clustering approaches that only provide the information of similarity features among DMUs, our proposed approach reveals the input–output relationships hidden in the data items of input and output. Thus, for each evaluated DMU, we know not only the cluster that it belongs to, but also the production function type that it confronts. It is very important for managerial decision-making where decision-makers are interested in knowing the changes required in combining input resources so that it can be re-classified into a different and desired cluster/class.

The focus of this paper is to examine the CCR model of DEA and then establish the DEA-based clustering. Without loss of generality, while this approach has been carried out for the CCR model, the proposed approach can be easily extended to other DEA models. We also showed that the clustering results drawn from the DEA-based clustering are unit-invariant, meaning that they are not affected by the scale of data.

It is important to point out, however, the DEA-based clustering approach is suitable for most clustering problems where there are inputs-and-outputs or cause-and-effect relationships between the features. For example, we can use the proposed approach in the analysis of industry classification, sorting of organizations by input–output data, classification for developed, developing and underdeveloped nations, and productivity sorting for staff, subunits, departments, institutes, and so on. In summary, in view of the advantages of the DEA-based clustering approach, it is uniquely poised for clustering problems. We believe that future researches are necessary to unleash the full potential of this DEA-based clustering approach. It thus has tremendous potential to be used for various clustering problems. Finally, we need to point out that the proposed DEA-based clustering algorithm is robust to a slight change in the input and output data sets, but not to outliers. Our future research will consider developing a robust-type DEA-based clustering algorithm.

Acknowledgements

The authors are grateful to the anonymous referees for their critical and constructive comments and suggestions. This work was supported in part by the National Science Council, Taiwan under Miin-Shen Yang's Grant NSC-95-2118-M-033-001-MY2.

References

Andersen, P., Petersen, N.C., 1993. A procedure for ranking efficient units in data envelopment analysis. Management Science 39, 1261–1264.

Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science 30, 1078–1092.

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithm. Plenum Press, New York.

Bick, A., Yang, F., Shandalov, S., Oron, G., 2007. Data envelopment analysis for assessing optimal operation of an immersed membrane bioreactor equipped with a draft tube for domestic wastewater reclamation. Desalination 204, 17–23.

Charnes, A., Cooper, W.W., Rousseau, J.J., Semple, J., 1987. Data envelopment analysis and axiomatic notions of efficiency and reference sets. Research Report CCS S58, Center for Cybernetic Studies, The University of Texas at Austin, Austin, Texas.

Charnes, A., Cooper, W.W., Seiford, L., Stutz, J., 1982. A multiplicative model for efficiency analysis. Socio-Economic Planning Science 16, 223–224.

Charnes, A., Cooper, W.W., Seiford, L., Stutz, J., 1983. Invariant multiplicative efficiency and piece-wise Cobb-Douglas envelopments. Operations Research Letters 2, 101–103.

Chang, K.P., Guh, Y.Y., 1991. Linear production functions and the data envelopment analysis. European Journal of Operational Research 52, 215–223.
Charnes, A., Cooper, W.W., Golany, B., Seiford, L., Stutz, J., 1985. Foundations of data envelopment analysis for Pareto–Koopmans efficient empirical production functions. Journal of Econometrics 30, 91–107.
Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision-making units. European Journal of Operational Research 2, 429–444.
Charnes, A., Cooper, W.W., Rhodes, E., 1981. Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through. Management Science 27, 668–697.
Charnes, A., Cooper, W.W., Huang, Z., Sun, B., 1990. Polyhedral cone-ratio models with an application to large commercial banks. Journal of Econometrics 46, 73–91.
Charnes, A., Cooper, W.W., Wei, Q.L., Huang, Z.M., 1989. Cone ratio data envelopment analysis and multi-objective programming. International Journal of Systems Science 20, 1099–1118.
Cook, W.D., Bala, K., 2007. Performance measurement and classification data in DEA: Input-oriented model. Omega 35, 39–52.
Cooper, W.W., Deng, H., Huang, Z.M., Li, S.X., 2002. Chance constrained programming approaches to technical efficiencies and inefficiencies in stochastic data envelopment analysis. Journal of the Operational Research Society 53, 1347–1356.
Cooper, W.W., Huang, Z.M., Li, S.X., 1996. Satisfying DEA models under chance constraints. The Annals of Operations Research 66, 279–295.
Cooper, W.W., Park, K.S., Yu, G., 1999. IDEA and AR-IDEA: Models for dealing with imprecise data in DEA. Management Science 45, 597–607.
Cooper, W.W., Ruiz, J.L., Sirvent, I., 2007. Choosing weights from alternative optimal solutions of dual multiplier models in DEA. European Journal of Operational Research 180, 443–458.
Dave, N., Krishnapuram, R., 1997. Robust clustering methods: A unified view. IEEE Transactions of Fuzzy Systems 5, 270–293.
Duda, R.O., Hart, P.E., 1973. Pattern Classification and Scene Analysis. Wiley, New York.
Farrel, M.J., 1957. The measurement of productivity efficiency. Journal of Royal Statistical Society Series A 120, 253–290.
Gattoufi, S., Oral, M., Reisman, A., 2004. A taxonomy for data envelopment analysis. Socio-Economic Planning Sciences 38, 141–158.
Grossberg, S., 1976. Adaptive pattern classification and universal recoding I: Parallel development and coding of neural feature detectors. Biological Cybernetics 23, 121–134.
Hartigan, J.A., 1975. Clustering Algorithms. Wiley, New York.
Huang, Z., Li, S.X., Rousseau, J.J., 1997. Determining rates of change in data envelopment analysis. Journal of the Operational Research Society 48, 591–599.
Jahanshahloo, G.R., Lotfi, F.H., Shoja, N., Tohidi, G., Razavyan, S., 2005. A one-model approach to classification and sensitivity analysis in DEA. Applied Mathematics and Computation 169, 887–896.
Jahanshahloo, G.R., Lotfi, F.H., Rezai, H.Z., Balf, F.R., 2007. Finding strong defining hyperplanes of production possibility set. European Journal of Operational Research 177, 42–54.
Jain, A.K., Duin, R.P.W., Mao, J., 2000. Statistical pattern recognition: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 4–37.
Jolion, J.M., Meer, P., Bataouche, S., 1991. Robust clustering with applications in computer vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 13, 791–802.
Kaufman, L., Rousseeuw, P.J., 1990. Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, New York.
Kleine, A., 2004. A general model framework for DEA. Omega 32, 17–23.
Kohonen, T., 2001. Self-Organizing Maps, third ed. Springer-Verlag, Berlin.
Krishnapuram, R., Keller, J.M., 1993. A possibilistic approach to clustering. IEEE Transactions on Fuzzy Systems 1, 98–110.
Lahdelma, R., Salminen, P., 2006. Stochastic multicriteria acceptability analysis using the data envelopment model. European Journal of Operational Research 170, 241–252.
Land, K.C., Lovell, C.A.K., Thore, S., 1994. Productivity and efficiency under capitalism and state socialism: An empirical inquiry using chance-constrained data envelopment analysis. Technological Forecasting and Social Change 46, 139–152.
Li, S., Jahanshahloo, G.R., Khodabakhshi, M., 2007. A super-efficiency model for ranking efficient units in data envelopment analysis. Applied Mathematics and Computation 184, 638–648.
Lippmann, R.P., 1987. An introduction to computing with neural nets. IEEE Transactions on Acoustics, Speech, Signal Processing, 4–22.
McLachlan, G.J., Basford, K.E., 1988. Mixture Models: Inference and Applications to Clustering. Marcel Dekker, New York.
McLachlan, G.J., Krishnan, T., 1997. The EM Algorithm and Extensions. Wiley, New York.
Olesen, O.B., Petersen, N.C., 1995. Chance constrained efficiency evaluation. Management Science 41, 442–457.
Pitaktong, U., Brockett, P.L., Mote, J.R., Rousseau, J.J., 1998. Identification of Pareto-efficient facets in data envelopment analysis. European Journal of Operational Research 109, 559–570.
Seiford, L.M., Thrall, R.M., 1990. Recent developments in DEA, the mathematical programming approach to frontier analysis. Journal of Econometrics 46, 7–38.
Thompson, R.G., Brinkmann, E.J., Dharmapala, P.S., Gonzalez-Lima, M.D., Thrall, R.M., 1997. DEA/AR profit ratios and sensitivity of 100 large US banks. European Journal of Operational Research 98, 213–229.
Thompson, R.G., Singleton, R.D., Thrall, R.M., Smith, B.A., 1986. Comparative site evaluations for locating a high energy physics laboratory in Texas. Interfaces 16, 35–49.
Tsao, E.C.K., Bezdek, J.C., Pal, N.R., 1994. Fuzzy Kohonen clustering networks. Pattern Recognition 27, 757–764.
Wu, K.L., Yang, M.S., 2002. Alternative *c*-means clustering algorithms. Pattern Recognition 35, 2267–2278.
Yang, M.S., 1993. A survey of fuzzy clustering. Mathematical and Computer Modelling 18, 1–16.
Yu, G., Wei, Q., Brockett, P., Zhou, L., 1996. Construction of all DEA efficient surfaces of the production possibility set under the generalized data envelopment analysis model. European Journal of Operational Research 95, 491–510.
Zanakis, S.H., Alvarez, C., Li, V., 2007. Socio-economic determinants of HIV/AIDS pandemic and nations efficiencies. European Journal of Operational Research 176, 1811–1838.
Zhu, J., 2003. Imprecise data envelopment analysis: A review and improvement with an application. European Journal of Operational Research 144, 513–529.