International Journal of Computational Intelligence Systems, Vol. 4, No. 4 (June, 2011), 606-618

A MODIFIED SUPER-EFFICIENCY DEA APPROACH FOR SOLVING MULTI-GROUPS CLASSIFICATION PROBLEMS

Jie Wu[1], **Qingxian An**[1, *], **Liang Liang**[1]

[1]*School of Management*
University of Science and Technology of China
He Fei, An Hui Province, P.R. China 230026
Email: jacky012@mail.ustc.edu.cn;aqxian03@mail.ustc.edu.cn; lliang@ustc.edu.cn

Abstract

Among the various discriminant analysis (DA) methods, researchers have investigated several directions in this area: statistics, econometrics, computer data mining technologies and mathematical programming. Recently, as a nonparametric mathematical programming approach, Data envelopment analysis has been applied in DA area and received great attention. In this paper, we propose a new discriminant approach based upon the relative distance measured by super-efficiency data envelopment analysis (DEA). This approach may generally avoid the drawbacks that usually occur in statistics discriminations of constructing function to determine a DMU's category. On the other hand, this approach may maintain discriminant capabilities by incorporating the non-parametric feature of DEA into DA. At the same time, it can also inherit the advantages of avoiding the process of dealing with different dimensional data in DEA. Our approach can be used to classify a sample's category by the discrimination results, even in the multiple-groups situation. Therefore, it can be applied to the discriminant analysis in various real-life cases.

Keywords: Data Envelopment Analysis (DEA); Super-efficiency; Discriminant analysis; Relative distance

Received 19 0ctober 2010
Accepted 20 June 2011

[*] *School of Management, University of Science and Technology of China, He Fei, An Hui Province 230026, China*

1. Introduction

Discriminant analysis (DA) model can be described as: there are *k* categories, $G_1,...,G_k$. Given a new sample, we should determine which category it belongs to. It is a classification method that can predict group membership of a newly sampled observation. This problem exists in many situations and plays a key role in decision making.

Most researches on DA have focused on the following problem, proposing new mathematical programming models and evaluating the classificatory performance of proposed models against that of the standard parametric classification procedures (Bal and Örkcü, 2010). Sueyoshi (2006) summarized previous research on DA and classified them into four groups as follows. One of the four groups is statistics, such as Fisher's linear discriminant function and Smith's quadratic discriminant function. Econometrics is another group, such as logit and probit models. The third group is computer data mining technologies, such as Neural Network and Decision Tree. The final group is mathematical programming. Recently, fuzzy logic is also used in this field (see, e.g., Zio et al., 2008; Chen and Chen, 2008).

Fisher's linear discriminant function (Fisher, 1936) and Smith's quadratic discriminant function (Smith, 1947) are popular statistical approaches to solve the problem under the assumption of multivariate normality and variance-covariance homogeneity. Without the assumption, Chang and Kuo (2008) indicated that the DA performance of linear programming had been proven superior to the former methods for classification purposes in many studies. Computer data mining technologies also can deal with the DA problem, however, so for these methods have the similar methodological shortcomings of no theoretical support on optimality (Sueyoshi, 2006). Because of these, nonlinear DA was proposed and popularly applied in terms of its classification and prediction capabilities. But most of the nonlinear DA methods, especially in econometrics, also have deficiency. They need to pre-specify the nonlinear function form to make a separation hyperplane. This procedure is not impossible but very difficult for us to prescribe such nonlinear discriminant function to fit a real data set. Relatively, the nonparametric methods are less restrictive and so receive more attention for classification problems.

In previous studies on nonparametric DA, Sueyoshi (2006) and Cooper *et al.* (1999) had proven that the piecewise linear discriminant function was more flexible than conventional linear discriminant function in terms of its discriminant capability. Data envelopment analysis (DEA) is a non-parametric programming technique for evaluating the relative efficiency of a set of homogenous decision making units (DMUs) with multiple inputs and multiple outputs. It has been applied in many areas, such as schools, hospitals, shops, bank branches and so on (Wu et al., 2010; Ozgen et al., 2011; Lopez et al., 2010; Doreswamy, 2010; Xu et al., 2009 and Kaya, 2010). DEA can also form a piecewise linear function for discriminant. Retzlaff-Roberts (1996) and Sueyoshi (1999) identified differences and similarities between DEA and DA, respectively. Now, DEA has been applied in the discriminant area and received a great deal of attention because it can maintain discriminant capabilities by incorporating the non-parametric feature of DEA into DA. Sueyoshi (1999) applied DEA in DA area and proposed Data Envelopment Analysis-Discriminant Analysis (DEA-DA) approach firstly. DEA-DA approach is a type of non-parametric DA approach that provides a set of weights of a piecewise linear discriminant function(s), and consequently yields an evaluation score(s) for the determination of group membership (Sueyoshi, 2006). This approach has been well developed by Sueyoshi (2001, 2004, 2005, 2006), Sueyoshi and Kirihara (1998), and Sueyoshi and Hwang (2004). Different from the above DEA and DA combination approach, we propose a new discriminant approach based on DEA method that just applied DEA in DA to measure the relative distance of new sample to each category. The basic idea was inspired by the relative distance, which can be measured through DEA models. Firstly we built some super-efficiency FG models, based on super-efficiency model by Andersen and Petersen (1993) and FG model by Färe and Grosskopf (1985). The super-efficiency FG model based on the best practice frontier was used to measure relative distances of a new sample to the best frontier. And super-efficiency FG model that based on the worst frontier was used to measure the relative distance of a new

sample to the worst frontier. According to the range of relative distance value, we divided them into four groups and calculated the relative distance of the new sample to each category by the relative distances of it to the best and worst frontier. Finally, we determined the new sample's category by its relative distances. This distance is different from the prior distance, such as Euclidean distance, Chebyshev distance, Mahalanobis distance, Minkowski distance and so on. Our distance is relative, while the others are absolute. The relative distance may more convenient than others because it can deal with the raw data directly without the need for standardization. Additionally, it also inherits the advantage of other kind distance for DA that should not form the discriminant function.

The remainder of this paper was organized as follows. Section 2 discussed basic theories through geometrical illustration and transformational thoughts of classifying the new sample. In section 3, modified super-efficiency models were built and applied in the proposed procedure. Section 4 introduced the solving steps. An illustration based upon an example with multiple inputs and outputs was given in section 5. Finally, conclusion and discussion were presented in section 6.

2. Geometrical illustration

2.1. The best-practice and worst-practice frontier

Most of DEA models always establish an efficient frontier (best-practice) among the units based on a comparison process in which the ratio scales of the weighted sum of the outputs to that of the inputs are evaluated. The DMUs on the frontier are efficient, otherwise are inefficient. The best-practice frontier can be defined as

$$F_b(\vartheta) = \{(x,y) \in \vartheta \mid \forall (x',y') \in \mathbb{R}^{m+s} ((x',-y') \leq\neq (x,-y) \Rightarrow (x',y') \notin \vartheta)\} \subseteq \vartheta.$$

At the same time, we also should gain the worst-practice frontier. According to the definition of best-practice frontier, we can obtain the worst-practice frontier similarly (Liu and Chen, 2009; Jahanshahloo and Afzalinejad, 2006).

$$F_w(\vartheta) = \{(x,y) \in \vartheta \mid \forall (x',y') \in \mathbb{R}^{m+s} ((-x',y') \leq\neq (-x,y) \Rightarrow (x',y') \notin \vartheta)\} \subseteq \vartheta.$$

To illustrate the difference between the best-practice frontier and the worst-practice frontier, we use an example of two inputs and one output data as shown in Table 1. All outputs normalized to 1 for simplicity. The best-practice and worst-practice frontiers of the example are presented in Fig. 1 and Fig. 2.

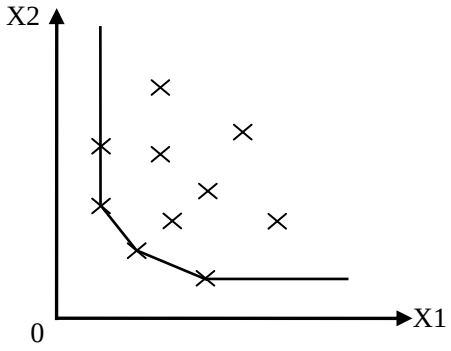

Fig.1. Best-practice frontier

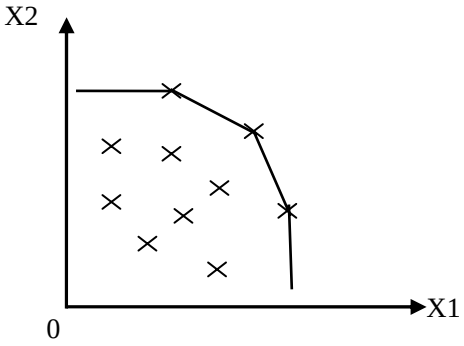

Fig.2. Worst-practice frontiers

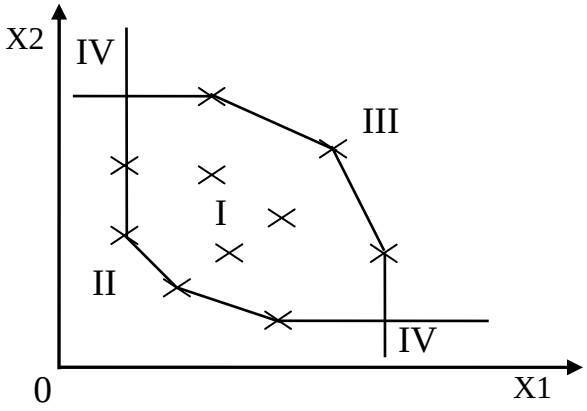

Fig.3. Four different locations of new sample

2.2. Locations of the new sample (DMU) to a certain category

According to the best-practice and worst-practice frontier of the certain category, the locations can be divided into three parts. To illustrate them, we use an example of two inputs and one output data as shown in Table 1. All outputs normalized to 1 for simplicity. The locations of the new sample are presented in Fig. 3. This will be used in section 3.

2.3. Transformational thoughts

For the discriminant analysis of multi-output samples (the number of output variables is *S*. Each variable has l_j styles, of which $j=1,2,...,s$), we can classify the samples as groups of $l_1 \times l_2 \times ... \times l_s$. After quantitative analysis of raw output information, we regard them as the output value of DEA. It should be noted that the outputs value reflect the category of the sample. For example, 1 refers to "ill" and 0 refers to "healthy" in medicine. If a DMU's output is 1, it will be classified into the "ill" category. We also treat the value of various factors which affect the sample type as the input value. For each type of the output, the new sample forms a new decision making unit (DMU). Therefore, a new sample will form $l_1 \times l_2 \times ... \times l_s$ new DMU totally. But, we don't put all DMUs together in comparing. We evaluate the new DMU just by putting it into raw samples which own the same output type. Based on the super-efficiency DEA method, we measure relative distance of the new DMU to the best-practice and the worst-practice frontier, and finally determine its category according some rules.

3. Modified DEA models used in the proposed procedure

3.1. A new super-efficiency DEA

Data envelopment analysis (DEA), as a non-parametric programming technique, provides a relative efficiency measure for peer decision making units (DMUs) with multiple inputs and multiple outputs, on the basis of evaluation of the private sectors by the economists Farell (1957). DEA was first proposed by Charnes, Cooper and Rhodes in 1978. DEA has been extensively applied in performance evaluation and benchmarking of schools, hospitals, bank branches, production plants, and so on (Cooper *et al*., 2004). As an evaluation method, it is suitable for evaluations with multi-input multi-output complex systems. Since the CCR model, there has been an impressive growth both in theoretical developments and applications of DEA. DEA researchers have developed a number of updated models, such as variable returns to scale (VRS) model, additive model, multilevel models, super-efficiency models and so on (Cook and Seiford, 2008). In this paper, we build super-efficiency FG model by integrating super-efficiency models (Andersen and Petersen, 1993) and FG model (R. Fä re and Grosskopf, 1985).

We assume that there are a set of *n* evaluation objects. Each object is named as $DMU_j, j=1,2,...,n.$ Each DMU produces *s* different outputs using *m* different inputs. The inputs of DMU_j is $x_j=(x_{1j},x_{2j},...,x_{mj})^T$, and the outputs is $y_j=(y_{1j},y_{2j},...,y_{sj})^T$, $x_j \geq 0, y_j \geq 0, \quad j=1,2,...,n$. That is to say, its components are non-negative and at least one of them is positive.

3.1.1. Super-efficiency model

The input-oriented super-efficiency measure θ' for an observation $(X_k,Y_k), k \in \{1,2,...,N\}$ is obtained by solving the following linear program:

$$\begin{aligned} &\theta'^{*} = min\ \theta' \\ &\text{s.t.} \quad \sum_{\substack{j=1 \\ j\neq k}}^{n} \lambda_j x_{ij} \leq \theta' x_{ik}, i=1,2,...,n \\ &\qquad \sum_{\substack{j=1 \\ j\neq k}}^{n} \lambda_j y_{rj} \geq y_{rk}, r=1,2,...,n \\ &\qquad \lambda_j \geq 0,\ j=1,2,...,n. \end{aligned} \tag{1}$$

From the model (1), we can see that when the super-efficiency model is applied, the observation "*k*" under evaluation is not included in the reference set. The super-efficiency model provides a means of evaluating the extent to which such changes could occur without violating that DMU's status as an efficient unit. Super-efficiency has been applied in many situations, such as detection of influential observations, DEA sensitivity analysis, acceptance decision rules, two-person ratio efficiency games and so forth. However, the super-efficiency DEA model may not have

feasible solutions for efficient DMUs (see, e.g., Zhu, 1996; Dulá and Hickman, 1997; Seiford and Zhu 1998a,b,c).

3.1.2. FG model

The model, firstly proposed by Färe and Grosskopf in 1985, was applied as a non-parametric cost approach to scale efficiency. The model is as follows:

$$\begin{aligned} \min \quad & \theta' \\ \text{s.t.} \quad & \sum_{j=1, j\neq j_0}^{n} \lambda_j X_j \le \theta' X_0 \\ & \sum_{j=1, j\neq j_0}^{n} \lambda_j Y_j \ge Y_0 \\ & \sum_{j=1,, j\neq j_0}^{n} \lambda_j \le 1 \\ & \lambda_j \ge 0,\ j = 1,2,...,n \end{aligned} \tag{2}$$

Comparing model (2) with BCC model proposed by Banker *et al.* (1984), we can find that the FG model has different constraint on variable λ. The reason of choosing this input-oriented model is that all DMUs in the same category have the same outputs.

3.1.3. Super-efficiency FG model based on the best practice frontier

The presentation is as formula (3)

$$\begin{aligned} \max \quad & V_P^{I'} = u^T Y_0 - \mu_0 \\ \text{s.t.} \quad & u^T Y_j - v^T X_j - \mu_0 \le 0 \\ & v^T X_0 = 1 \\ & j = 1,2,...,n,\ j \neq j_0 \\ & u \ge 0, v \ge 0, u \neq 0, v \neq 0, \mu_0 \le 0 \end{aligned} \tag{3}$$

Its dual model, is as formula (4)

$$\begin{aligned} \min \quad & \theta' \\ \text{s.t.} \quad & \sum_{j=1, j\neq j_0}^{n} \lambda_j X_j \le \theta' X_0 \\ & \sum_{j=1, j\neq j_0}^{n} \lambda_j Y_j \ge Y_0 \\ & \sum_{j=1,, j\neq j_0}^{n} \lambda_j \le 1 \\ & \lambda_j \ge 0,\ j = 1,2,...,n,\ j \neq j_0 \end{aligned} \tag{4}$$

The value θ' which is calculated by the above model reflects the relative efficiency of the evaluated DMU based on the best frontier formatted by all the decision-making units except the evaluating DMU.

Definition 1: When $u^T Y - v^T X - \mu_0 > 0$ or $\theta' > 1$, we consider (X, Y) is at the negative side of the best-practice frontier. When $u^T Y - v^T X - \mu_0 < 0$ or $\theta' < 1$, we consider (X, Y) is at the positive side of the best-practice frontier.

3.1.4. Super-efficiency FG model based on the worst frontier

The presentation is as formula (5)

$$\begin{aligned} \min \quad & V_P^{I'} = u^T Y_0 - \mu_0 \\ \text{s.t.} \quad & u^T Y_j - v^T X_j - \mu_0 \ge 0 \\ & v^T X_0 = 1 \\ & j = 1,2,...,n,\ j \neq j_0 \\ & u \ge 0, v \ge 0, u \neq 0, v \neq 0, \mu_0 \le 0 \end{aligned} \tag{5}$$

Its dual model is as formula (6)

$$\begin{aligned} \max \quad & \theta'' \\ \text{s.t.} \quad & \sum_{j=1, j\neq j_0}^{n} \lambda'_j X_j \ge \theta'' X_0 \\ & \sum_{j=1, j\neq j_0}^{n} \lambda'_j Y_j \le Y_0 \\ & \sum_{j=1,, j\neq j_0}^{n} \lambda'_j \le 1 \\ & \lambda'_j \ge 0,\ j = 1,2,...,n,\ j \neq j_0 \end{aligned} \tag{6}$$

The value θ' which was calculated by the above models reflects the relative efficiency of the evaluated DMU based on the worst frontier formatted by all DMUs except the evaluating DMU.

Definition 2 : When $u^T Y - v^T X - \mu_0 < 0$ or $\theta^{"} < 1$, we consider (X, Y) is at the negative side of the worst-practice frontier. When $u^T Y - v^T X - \mu_0 > 0$ or $\theta^{"} > 1$, we consider (X, Y) is at the positive side of the worst-practice frontier.

As the outputs of all decision-making units of the same type are unanimous in classification, model (4) and model (6) can be transformed into a corresponding model (7)

$$\begin{aligned} &min \quad \theta^{'} \\ &\text{s.t.} \quad \sum_{j=1, j\neq j_0}^{n} \lambda_j X_j \le \theta^{'} X_0 \\ &\qquad \sum_{j=1,, j\neq j_0}^{n} \lambda_j \le 1 \\ &\qquad \lambda_j \ge 0, j = 1, 2, ..., n, j \ne j_0 \end{aligned} \tag{7}$$

and model (8)

$$\begin{aligned} &max \quad \theta^{"} \\ &\text{s.t.} \quad \sum_{j=1, j\neq j_0}^{n} \lambda_j' X_j \ge \theta^{"} X_0 \\ &\qquad \sum_{j=1,, j\neq j_0}^{n} \lambda_j' \le 1 \\ &\qquad \lambda_j' \ge 0, j = 1, 2, ..., n, j \ne j_0 \end{aligned} \tag{8}$$

Taking into account the possibility of the existence of non-feasible solution in super-efficiency，firstly our paper need to prove the existence of optimum solution of the above mentioned super efficient models. The conclusion is that they both have the optimum solutions. We give the proof as follows.

Theorem 1: Super-efficiency DEA model (7) and model (8) both have the optimal solution.

Proof: For the model (7), assume the vector $\zeta' = (\lambda_j = 0, j = 1, ..., n, j \ne j_0; \theta' = 0)$, so ζ' is a group of feasible solutions of model (7).

Because $\sum_{j=1,, j\neq j_0}^{n} \lambda_j = 0 \le 1$ and $\sum_{j=1, j\neq j_0}^{n} \lambda_j' X_j = 0, \theta^{'} X_0 = 0$, $\sum_{j=1, j\neq j_0}^{n} \lambda_j' X_j \le \theta^{'} X_0$. Therefore, we prove that model (7) has feasible solution. Additionally，$\theta^{'} \ge \min_{k\in\{1,...,m\}} (\sum_{j=1, j\neq 0}^{n} \lambda_j X_{kj} / X_{k0})$, so the objectives θ' have lower bound. In summary, model (7) has optimal solution. Similarly, model (8) also has optimal solution.

To facilitate our study, model (7) and model (8) can be further transformed into model (9)

$$\begin{aligned} &min \quad \theta^{'} = \max_{k\in\{1,...,m\}} (\sum_{j=1, j\neq 0}^{n} \lambda_j X_{kj} / X_{k0}) \\ &\text{s.t.} \quad \sum_{j=1,, j\neq j_0}^{n} \lambda_j \le 1 \\ &\qquad \lambda_j \ge 0, j = 1, 2, ..., n, j \ne j_0 \end{aligned} \tag{9}$$

and model (10).

$$\begin{aligned} &max \quad \theta^{"} = \min_{k\in\{1,...,m\}} (\sum_{j=1, j\neq 0}^{n} \lambda_j X_{kj} / X_{k0}) \\ &\text{s.t.} \quad \sum_{j=1,, j\neq j_0}^{n} \lambda_j \le 1 \\ &\qquad \lambda_j \ge 0, j = 1, 2, ..., n, j \ne j_0 \end{aligned} \tag{10}$$

3.2. Model results analysis and discriminant rules

3.2.1. Model results Analysis

Because the model has an optimal solution, we can express $\theta_i^{'}$, $\theta_i^{"}$ as $\theta_i^{'} = \max_{k\in\{1,...,m\}} (\sum_{j=1, j\neq i}^{n} \lambda_{jk}^{*} X_{jk} / X_{ik})$, $\theta_i^{"} = \min_{k\in\{1,...,m\}} (\sum_{j=1, j\neq i}^{n} \lambda_{jk}^{'*} X_{jk} / X_{ik})$, of which *i* represents the new DMU *i*, λ_{jk}^{*} is the optimal solution of the model (9), $\lambda_{jk}^{'*}$ is the optimal solution of model (10). In reality, the optimum

solution values $\theta_i^{'}, \theta_i^{''}$ may have the following four cases: $\theta_i^{'} \le 1$ and $\theta_i^{''} \ge 1$; $\theta_i^{'} \le 1$ and $\theta_i^{''} \le 1$; $\theta_i^{'} \ge 1$ and $\theta_i^{''} \ge 1$; $\theta_i^{'} \ge 1$ and $\theta_i^{''} \le 1$.

Definition 3: θ_i is the relative distance of the new DMU to a certain category, which meets the following three properties:

(i) When the new DMU is at the positive side of both the best-practice frontier and the worst-practice frontier, the relative distance is 0;

(ii) When a new DMU of the samples is at the positive side of one of the best-practice frontier or the worst-practice frontier, then the distance to the frontier is 0; and their relative distance θ_i is the distance of a new DMU to the another frontier.

(iii)When the new DMU is not at negative side of the certain samples production frontier, the farther the distance from the samples frontier is, the greater the relative distance is.

Proposition 1: When the new DMU in I district, then $\theta_i^{'} \le 1$, $\theta_i^{''} \ge 1$, so $\theta_i = 0$.

Proof: In this situation, the new DMU is at the positive side of both the best-practice frontier and the worst-practice frontier. By definition 3, the relative distance of the new DMU is 0, so it can be classified into this style directly.

Proposition 2: When the new DMU in II district, then $\theta_i^{'} \ge 1$, $\theta_i^{''} \ge 1$, so $\theta_i = \theta_i^{'} - 1$.

Proof: In this situation, the new DMU is at the positive side of both the worst-practice frontier. By definition 3, its distance is 0. At the same time, it is at the negative side of both the best-practice frontier, so this distance should be taken into account. The relative distance of the new DMU to the best-practice frontier is $\max_{k\in\{1,\ldots,m\}} (\sum_{j=1, j\neq i}^{n} \lambda_{jk}^{*} X_{jk} - X_{ik}) / X_{ik}$. Overall, the relative distance θ when DMU is classified as such category is

$$\theta = \max_{k\in\{1,\ldots,m\}} (\sum_{j=1, j\neq i}^{n} \lambda_{jk}^{*} X_{jk} - X_{ik}) / X_{ik} .$$

$$\max_{k\in\{1,\ldots,m\}} (\sum_{j=1, j\neq i}^{n} \lambda_{jk}^{*} X_{jk} - X_{ik}) / X_{ik}$$

Since $= \max_{k:1\to m} (\sum_{j=1, j\neq i}^{n} \lambda_{jk}^{*} X_{jk} / X_{ik} - 1)$,

$$= \min_{k:1\to m} (\sum_{j=1, j\neq i}^{n} \lambda_{jk}^{*} X_{jk} / X_{ik}) - 1 = \theta_i^{'} - 1$$

therefore $\theta = \theta_i^{'} - 1$.

Proposition 3: When the new DMU in III district, then $\theta_i^{'} \le 1$, $\theta_i^{''} \le 1$, so $\theta = 1 - \theta_i^{''}$.

Proof: In this situation, the new DMU is at the positive side of both the best-practice frontier. By definition 3, its distance is 0. At the same time, it is at the negative side of the worst-practice frontier, so this distance should be taken into account. The relative distance of the new DMU to the worst-practice frontier is $\max_{k:1\to m} (X_{ik} - \sum_{j=1, j\neq i}^{n} \lambda_{jk}' X_{jk}) / X_{ik}$. Overall, the relative distance θ when DMU is classified as such category is $\max_{k:1\to m} (X_{ik} - \sum_{j=1, j\neq i}^{n} \lambda_{jk}' X_{jk}) / X_{ik}$.

Since

$$\max_{k:1\to m} (X_{ik} - \sum_{j=1, j\neq i}^{n} \lambda_{jk}' X_{jk}) / X_{ik}$$

$$= \max_{k:1\to m} (1 - \sum_{j=1, j\neq i}^{n} \lambda_{jk}' X_{jk} / X_{ik}) \quad ,$$

$$= 1 - \min_{k:1\to m} (\sum_{j=1, j\neq i}^{n} \lambda_{jk}' X_{jk} / X_{ik}) = 1 - \theta_i^{''}$$

therefore $\theta = 1 - \theta_i^{''}$.

Proposition 4: When the new DMU in IV district, then $\theta_i^{'} \ge 1$, $\theta_i^{''} \le 1$, so $\theta = \theta_i^{'} - \theta_i^{''}$.

Proof: In this situation, the new DMU is at the

negative side of both the worst-practice frontier, so this distance should be taken into account. Its distance is $\max_{k\in\{1,...,m\}}(\sum_{j=1,j\neq i}^{n}\lambda_{jk}^{*}X_{jk}-X_{ik})/X_{ik}$.

At the same time, it is at the negative side of both the best-practice frontier, so this distance should also be taken into account. The relative distance of the new DMU to the worst-practice frontier is $\max_{k:1\to m}(X_{ik}-\sum_{j=1,j\neq i}^{n}\lambda'_{jk}X_{jk})/X_{ik}$.

Overall, the relative distance θ when DMU is classified as such category is

$$\max_{k\in\{1,...,m\}}(\sum_{j=1,j\neq i}^{n}\lambda_{jk}^{*}X_{jk}-X_{ik})/X_{ik} + \max_{k:1\to m}(X_{ik}-\sum_{j=1,j\neq i}^{n}\lambda'_{jk}X_{jk})/X_{ik}.$$

Since

$$\max_{k\in\{1,...,m\}}(\sum_{j=1,j\neq i}^{n}\lambda_{jk}^{*}X_{jk}-X_{ik})/X_{ik} = \max_{k\in\{1,...,m\}}(\sum_{j=1,j\neq i}^{n}\lambda_{jk}^{*}X_{jk}/X_{ik}-1) = \max_{k\in\{1,...,m\}}(\sum_{j=1,j\neq i}^{n}\lambda_{jk}^{*}X_{jk}/X_{ik})-1=\theta_i^{'}-1$$

and

$$\max_{k:1\to m}(X_{ik}-\sum_{j=1,j\neq i}^{n}\lambda'_{jk}X_{jk})/X_{ik} = \max_{k:1\to m}(1-\sum_{j=1,j\neq i}^{n}\lambda'_{jk}X_{jk}/X_{ik}) = 1-\min_{k:1\to m}(\sum_{j=1,j\neq i}^{n}\lambda'_{jk}X_{jk}/X_{ik})=1-\theta_i^{''},$$

therefore $\theta_i=\theta_i^{'}-\theta_i^{''}$.

3.2.2. Discriminant rules

When determining which category the new DMU belongs to, we firstly calculate $\theta_i^{'}$ and $\theta_i^{''}$, when the new DMU is assigned to the *k*th category, $k=1,2,...,l_1\times l_2\times...\times l_s$. Then we calculate θ_i corresponding to Proposition 1-4.

When the situation of proposition 1 only occurs once when classifying it into each certain category, then the new DMU belongs to this category. If the situation of Proposition 1 occurs several times when classifying it into each certain category, the new DMU should wait to be sentenced. Besides these, the new DMU shall belong to a type i_0 which has the minimum θ_i, that is,

$$\{i_0\mid\theta_{i_0}\triangleq min\{\theta_k\},k=1,2,....,l_1\times l_2\times...\times l_s\}$$

.

For verifying the rationality of our approach, we compare the results above with the geometrical results. A simple example of one input and one output which is shown in Figure 4 is applied to illustrate it.

(1) When $\theta_i^{'}\le 1$ and $\theta_i^{''}\ge 1$, $\theta_i=0$.

M_1' in Figure 4 belongs to this situation. Taking M_1' as the studied object, that is, I assume that it is classified into category I. The new DMU is at the positive side of both the best-practice frontier and the worst-practice frontier. So, it is in the inner of the category I. The relative distance of the new DMU is 0.

(2) When $\theta_i^{'}\ge 1$ and $\theta_i^{''}\ge 1$, $\theta_i=\theta'-1$.

M_2' in Figure 4 belongs to this situation. Taking M_2' as the studied object, that is, I assume that it is classified into category II. The new DMU is at the negative side of the best-practice frontier. At the same time, it is at the positive side of the worst-practice frontier, so this distance should be taken into account. The relative distance of the new DMU to the

best-practice frontier point A is $\frac{AM_2^{'}}{O^{'}M_2^{'}}$. The relative distance θ when DMU is classified as such category is $\theta = \frac{AM_2^{'}}{O^{'}M_2^{'}} = \frac{O^{'}A - O^{'}M_2^{'}}{O^{'}M_2^{'}} = \frac{O^{'}A}{O^{'}M_2^{'}} - 1 = \theta' - 1$.

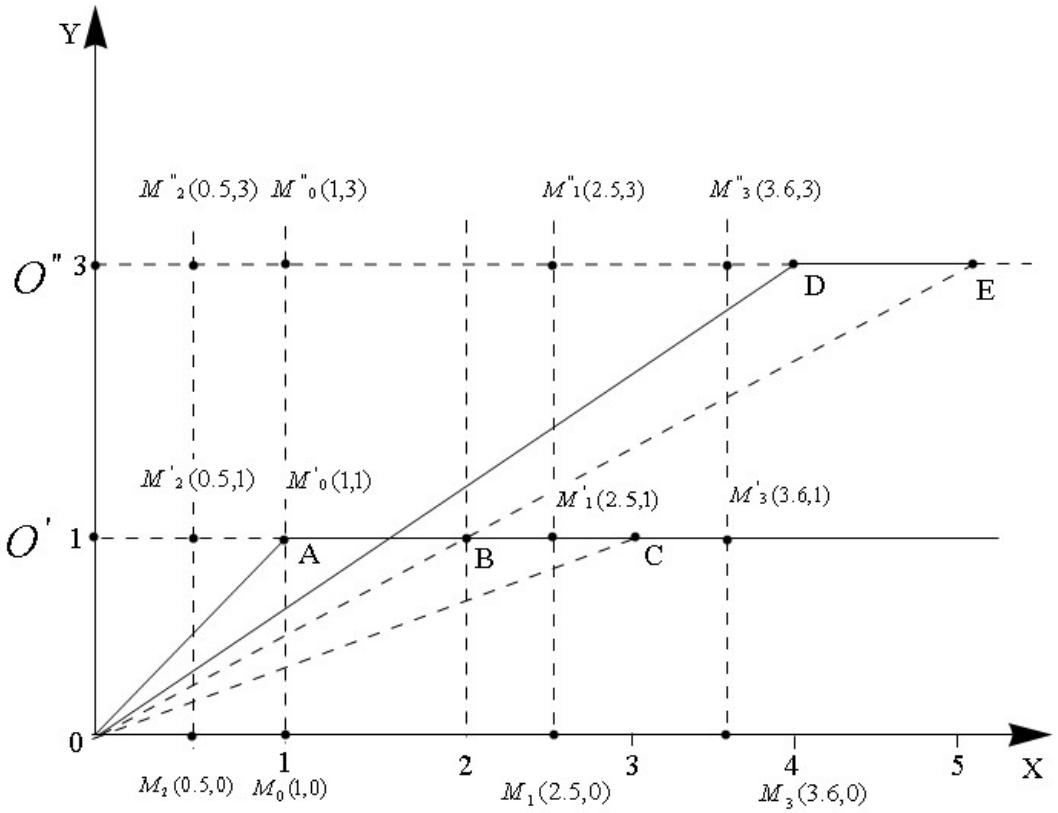

Fig. 4. Production frontier and analysis of various types of samples[1]

[1]Dotted lines represent the worst-practice frontier; solid lines represent the best-practice frontier.

(3) When $\theta_i^{'} \le 1$ and $\theta_i^{''} \le 1$, $\theta = 1 - \theta_i^{''}$.

M_3' in Figure 4 belongs to this situation.

Taking M_3' as the studied object, that is, I assume that it is classified into category III. The new DMU is at the positive side of the best-practice frontier. At the same time, it is at the negative side of the worst-practice frontier, so this distance should be taken into account. The relative distance of the new DMU to the worst-practice frontier point C is $\frac{CM_3^{'}}{O^{'}M_3^{'}}$. The relative distance θ when DMU is classified as such category is

$$\theta = \frac{CM_3^{'}}{O^{'}M_3^{'}} = \frac{O^{'}M_3^{'} - O^{'}C}{O^{'}M_3^{'}} = 1 - \frac{O^{'}C}{O^{'}M_3^{'}} = 1 - \theta^{''}.$$

(4) When $\theta_i^{'} \ge 1$ and $\theta_i^{''} \le 1$,

$\theta_i = \theta_i^{'} - \theta_i^{''}$.

This situation in single-input single-output case is almost impossible, unless the category is only one original sample. Therefore, icons can not explain it here.

From the above analysis, we can see the two methods have better consistency.

4. Solving steps

At the same time, we verify the Propositions in part 2 intuitionally by the Figure 1. The specific steps of discrimination are as follows:

1) Since the output variables are qualitative, the first treatment is quantification of them. Assuming each style as $w_i; i = 1, 2, ..., k$.

2) Grouping all the raw samples into *k* different groups, $G_i; i = 1, 2, ..., k$.

3) Assuming that the new sample's output is w_i, we obtain the new DMU and its results of $\theta_i^{'}$, $\theta_i^{''}$ by incorporating it into G_i and evaluating it based on the input of the best and the worst super-efficiency DEA evaluation. Because this paper studies the various decision-making units of different input values to bring out the difference between the outputs, it is more practical and significant based on the input than the output.

4) Determine the category of the new DMU by the results gained from step 3 and the discriminant rules in section 2.

5. Illustrations

In order to better demonstrate how the proposed approach works, we propose a example which takes thirty-five three-input two-output samples in Table 1 as an basic to distinguish the two new samples in Table 2.

Table 1. The raw samples

Input 1	Input 2	Input 3	Output 1	Output 2
39	6	20	1	1
39	12	20	1	1
47	6	12	1	1
47	12	12	1	1
32	19	75	1	1
6	28	30	1	1
113	18	75	1	1
52	12	40	1	1
52	6	40	1	1
113	35	180	1	2
172	14	45	1	2
172	15	45	1	2
32	4	75	1	2
30	10	70	1	2
32	12	75	1	2
11	3	15	1	2
30	9	25	1	2
8	4	30	2	1
8	1	30	2	1
161	4	70	2	1
161	1	70	2	1
6	12	30	2	1
6	3	30	2	1
6	5	30	2	1
6	7	18	2	1
113	6	75	2	1

113	8	75	2	2
52	6	40	2	2
52	8	40	2	2
97	5	180	2	2
97	5	180	2	2
89	10	180	2	2
56	13	180	2	2
172	6	45	2	2
283	6	45	2	2

Table 2. Inputs and outputs of the new samples

Sample	input1	input2	input3
1	3	12	60
2	100	20	50

The steps to determine the category of the two new samples are:

1) First, according to the outputs, we classify the samples into four categories. Category i: output1 is 1 and output2 is 1. Category ii: output1 is 1 and output2 is 2. Category iii: output1 is 2 and output2 is 1. Category iv: output1 is 2 and output2 is 2.

2) For the new samples 1 and 2, if they are classified as four categories in step 1) respectively, we can obtain their relative distance to best-practice frontier in Table 3 through model (3).

Table 3. The new sample's relative distance to the best-practice frontier

Sample	i	ii	iii	iv
1	2.2857	3.6667	2	17.3333
2	0.3819	0.3	0.36	0.8

3) For the new samples 1 and 2, if they are classified as four types in step 1) respectively, we can obtain their relative distance to the worst-practice frontier in Table 4 through model (5).

Table 4. The new sample's relative distance to the worst-practice frontier

Sample	i	ii	iii	iv
1	1.25	2.9167	0.8	1.0833
2	0.9732	1.3601	0.4892	0.638

4) On the basis of Proposition 1 to 4 in section 3, we obtain the new sample's relative distance to the raw samples as Table 5.

Table 5. The new sample's relative distance to the raw samples

Sample	i	ii	iii	iv
1	1.2857	2.6667	1.2	16.3333
2	0.0268	0	0.5108	0.362

5) According to the discrimination basic principles in section 2.3, we can see the new sample 1 should be grouped as Category iii; the new sample 2 should be grouped as Category ii.

6. Conclusions

In real life, we often encounter classification of a sample with multi-outputs. Many discriminant analysis methods have been proposed, such as statistics, econometrics, computer data mining technologies, mathematical programming and so on. The DEA method in DA has drawn more and more attention to the researchers. The modified super-efficiency models in this paper can effectively solve the classification problems and are especially easy to be understood and applied. This approach can deal with the raw data directly without the need for standardization; therefore we may avoid the difficulties of choosing standardized method. In addition, this approach may maintain discriminant capabilities by incorporating the non-parametric feature of DEA into DA.

Moreover, as one of the solutions, the proposed approach is only one way to combine DA and DEA. Adopting the idea of this new approach on classifying a new DMU, some extensions can be studied in the future, such as how to determine the relative distance using a simpler model, how to further classify the sample when the new sample have zero distance in several categories, and how to apply this idea to solve discriminant problem in more practical areas.

Acknowledgements

The research is supported by National Natural Science Funds of China for Innovative

Research Groups (No. 70821001), National Natural Science Funds of China (No. 70901069), Special Fund for the Gainers of Excellent Ph.D's Dissertations and Dean's Scholarships of Chinese Academy of Sciences, Research Fund for the Doctoral Program of Higher Education of China for New Teachers (No. 20093402120013), Research Fund for the Excellent Youth Scholars of Higher School of Anhui Province of China （No.2010SQRW001ZD）and Social Science Research Fund for Higher School of Anhui Province of China (No. 2010SK004).

References

1. P. Andersen and N.C. Petersen, A procedure for ranking efficient units in data envelopment analysis, *Manage. Sci.* 39(10) (1993) 1261–1264.
2. H. Bal and H. H. Örkcü, A new mathematical programming approach to multi-group classification problems, *Comput Oper Res*. 38(11) (2010) 105-111.
3. R. D. Banker, A. Charnes and W.W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Manage. Sci*. 30(9) (1984) 1078-1092.
4. D. S. Chang and Y. C. Kuo, An approach for two-group discriminant analysis: An approach of DEA, *Math Comput Model*. 47(9-10) (2008) 970-981.
5. A. Charnes, W. W. Cooper and E. Rhodes, Measuring the efficiency of decision making units, *Eur. J. of Oper. Res*. 2(6) (1978) 429–444.
6. W. D. Cook and L. M. Seiford, Data envelopment analysis (DEA)-Thirty years on, *Eur. J. of Oper. Res*. 192(1) (2008) 1-17.
7. W. W. Cooper, L. M. Seiford and J. Zhu (Eds.), *Data envelopment analysis*, (Kluwer Academic Publishers, London, 2004).
8. Z. L. Chen and G. Q. Chen, building an associative classifier based on fuzzy association rules, *Int. J. Comput. Intell. Syst*. 1(3) (2008) 262-273.
9. W. W. Cooper, K. S. Park and J. T. Pastor, RAM: A range adjusted measure of inefficiency for use with additive models and relations to other models and measures in DEA, *Journal of Productivity Analys*. 11(1) (1999) 5–42.
10. H. Doreswamy and M. N. Vanajaskhi, Similarity Measuring Approach for Engineering Materials Selection, *Int. J. Comput. Intell. Syst*. 3 (1) (2010) 115-122.
11. J. H. Dulá and B. L. Hickman, Effects of excluding the column being scored from the DEA envelopment LP technology matrix, *J Oper Res Soc*. 48(10) (1997) 1001–1012.
12. R. Färe and S. Grosskopf, A nonparametric cost approach to scale efficiency, *Journal of Economics*. 87(4) (1985) 594-604.
13. M. J. Farrell, The measurement of production efficiency, *J R Stat Soc A*, General, 120(3) (1957) 253-281.
14. R. A. Fisher, The use of multiple measurements in taxonomy problems, *Annals of Eugenics*. 7(2) (1936) 179–188.
15. F. H. F. Liu and C. L. Chen, The worst-practice DEA model with slack-based measurement, *Computers & Industrial Engineering*. 57(2) (2009) 496-505.
16. G. R. Jahanshahloo and M. Afzalinejad, A ranking method based on a full-inefficient frontier, *Appl Math Model*. 30(3) (2006) 248-260.
17. T. Kaya, Multi-attribute evaluation of website quality in E-business using an integrated fuzzy AHPTOPSIS methodology, *Int. J. Comput. Intell. Syst*. 3 (3) (2010) 301-314.
18. V. Lopez, M. Santos and J. Montero, Fuzzy Specification in Real Estate Market Decision Making, *Int. J. Comput. Intell. Syst*. 3 (1) (2010) 8-20.
19. A. Ozgen, G. Tuzkaya, U.R. Tuzkaya and D. Ozgen, A Multi-criteria decision making approach for machine tool selection problem in a fuzzy environment, *Int. J. Comput. Intell. Syst*. 4 (4) (2011) 431-445.
20. D. L. Retzlaff–Roberts, Relating discriminant analysis and data envelopment analysis to one another, *Comput Oper Res*. 23(4) (1996) 311-322.
21. L. M. Seiford and J. Zhu, An acceptance system decision rule with data envelopment analysis, *Comput Oper Res*. 25(4) (1998) 329–332.
22. L. M. Seiford and J. Zhu, Sensitivity analysis of DEA models for simultaneous changes in all the data, *J Oper Res Soc*. 49(10) (1998) 1060–1071.
23. L. M. Seiford and J. Zhu, Infeasibility of super-efficiency data envelopment analysis models, *Infor*. 37(2) (1999) 174–187.
24. C. A. B. Smith, Some examples of discrimination, *Annals of Eugenics*. 13(4) (1947) 272–282.
25. T. Sueyoshi, DEA-Discriminant Analysis: Methodological comparison among eight discriminant analysis approaches, *Eur. J. of Oper. Res*. 169(1) (2006) 247-272.

26. T. Sueyoshi, DEA-discriminant analysis in the view of goal programming, *Eur. J. of Oper. Res*. 115(3) (1999) 564–582.
27. T. Sueyoshi, Extended DEA-discriminant analysis, *Eur. J. of Oper. Res*. 131(2) (2001) 324–351.
28. T. Sueyoshi, Mixed interger programming approach of extended-discriminant analysis, *Eur. J. of Oper. Res*. 152(1) (2004) 45–55.
29. T. Sueyoshi, Financial ratio analysis of the electric power industry, *Asia Pac J Oper Res*. 22(3) (2005) 349-376.
30. T. Sueyoshi and S.N. Hwang, A use of nonparametric test for DEA-Discriminant analysis: A methodological comparison, *Asia Pac J Oper Res*. 21(2) (2004) 179-196.
31. T. Sueyoshi and Y. Kirihara, Efficiency measurement and strategic classification of Japanese banking institutions, *Int J Syst Sci*. 29(11) (1998) 1249–1263.
32. J. Wu, L. Liang and M.L. Song, Performance based clustering for benchmarking of container ports: an application of DEA and cluster analysis technique, *Int. J. Comput. Intell. Syst*. 3(6) (2010) 709-722.
33. X. Xu, R. Law, T. Wu, Support Vector Machines with Manifold Learning and Probabilistic Space Projection for Tourist Expenditure Analysis, *Int. J. Comput. Intell. Syst*. 2 (1) (2009) 17-26.
34. J. Zhu, Robustness of the efficient DMUs in data envelopment analysis, *Eur. J. of Oper. Res*. 90(3) (1996) 451–460.
35. E. Zio, P. Baraldi and I. C Popescu, From fuzzy clustering to a fuzzy rule-based fault classification model, *Int. J. Comput. Intell. Syst*. 1(1) (2008) 60-76.