

## A new algorithm for benchmarking in integer data envelopment analysis

F. H. Saljooghi<sup>a</sup>, A. R. Pishgooy<sup>b</sup> and M. M. Omran<sup>c</sup>

<sup>a</sup>Department of Mathematics, University of Sistan and Baluchestan, Zahedan, Iran

<sup>b</sup>Department of Mathematics, Islamic Azad University, Zahedan Branch, Zahedan, Iran

<sup>c</sup>Department of Industrial Engineering, Iran University of Science & Technology, Tehran, Iran

### ARTICLE INFO

#### Article history:

Received October 1, 2011  
Received in Revised form  
November, 14, 2011  
Accepted 30 January 2012  
Available online  
20 February 2012

#### Keywords:

Data Envelopment Analysis (DEA)  
Inefficient units  
Benchmarking

### ABSTRACT

The aim of this study is to investigate the effect of integer data in data envelopment analysis (DEA). The inputs and outputs in different types of DEA are considered to be continuous. In most application-oriented problems, some or all data are integers; and subsequently, the continuous condition of the values is omitted. For example, situations in which the inputs/outputs are representatives of the number of cars, people, etc. In fact, the benchmark unit is artificial and does not contain integer inputs/outputs after projection on the efficiency frontier. By rounding off the projection point, we may lose the feasibility or end up having inefficient DMU. In such cases, it is required to provide a benchmark unit such that the considered unit reaches the efficiency. In the present short communication, by proposing a novel algorithm, the projecting of an inefficient DMU is carried out in such a way that produced benchmarking takes values with fully integer inputs/outputs.

## 1. Introduction

One of the robust tools in evaluating of decision making units (DMUs) is Data Envelopment Analysis (DEA). DEA is a nonparametric method that has been used to evaluate the efficiency of entities, which are responsible for utilizing resources to obtain outputs of interest; also this method does not require prior information about the relationship among multiple performance measures, it estimates the empirical tradeoff curve from the observations. DEA technique has allocated to itself a wide variety of research in operations research field. In fact, DEA has become increasingly popular for efficiency analysis in the practical projects in field of management, industrial, economy, education, sport, etc. DEA has different models which are created according to application, radial models as CCR (Charnes et al., 1978), BCC (Banker et al., 1985) or non-radial models as SBM (Tone, 2011) and so on. DEA models take the inputs/outputs as real value.

\* Corresponding author.

E-mail addresses: [omran@b-just.ac.ir](mailto:omran@b-just.ac.ir) (M. M. Omran)

In many situations, it is needed to have them in integer formats. Then, by applying the DEA model, the under evaluation DEA are divided into two efficient sub categories. On the other hand, one of the most important features of DEA is to have benchmarks for inefficient DMUs.

Since the values of the variables in ordinary DEA models are not taken as integers and the produced values are not always integer, as well as the convexity condition with pure integer inputs/outputs are not taken into consideration; thus, finding integer reference units is a significant task. As an illustration, let us assume having three professors; the considered approach is inefficient and the analysis of suggesting DEA number for attaining the efficiency is 2.4. If the number of professors is considered as 2, then there is no evidence to show that 2 professors can achieve the educational and scientific goals. Moreover, rounding off 2.4 to 3 increases the cost of education, significantly and the efficiency of the team will be 0.8.

For the first time, Lozano and Villa (2006), proposed a Mixed Integer Linear Programming (MILP) for integer evaluation of computationally aims. Next, Kuosmanen and KazemiMatin (2009) investigated a new MILP for this purpose.

In this paper, after projecting of the inefficient DMU on the efficient frontier; we give the efficient benchmark unit for the inefficient unit by using the new algorithms and the concepts of hyperplanes and integer part. The rest of this work is summarizing as follows. Section 2 discusses the basic models of DEA and some of its definitions. In Section 3, the suggested algorithm is given and then its implementation is also illustrated by numerical example in Section 4.

## 2. Data Envelopment Analysis (DEA)

The method of DEA was first coined by Charnes et al. (1978), which is also known as CCR model. Organizations subject to evaluation in the DEA literature are called DMUs. Consider  $n$  decision making units, One DMU receives inputs and changes them to outputs. The  $y_{rj}, x_{ij} > 0$  in the model are constant which represent amounts of  $r^{\text{th}}$  output ( $r = 1, 2, \dots, s$ ) and the  $i^{\text{th}}$  input ( $i = 1, 2, \dots, m$ ) of the  $j^{\text{th}}$  Decision Making Unit ( $DMU_j$ ),  $j = 1, 2, \dots, n$ . Therefore, each observed DMU is characterized by a pair of non-negative input and output vectors  $(X_j, Y_j) \in \mathbb{R}^{m+s}$ ,  $j \in \{1, 2, \dots, n\}$ .

The standard (traditional) DEA approach determines the technical efficiency ( $\theta_o$ ) of any  $DMU_o$ . The enveloped form of CCR model is defined as

min  $\theta$

subject to

$$\sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io} \quad i=1, \dots, m$$

$$\sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro} \quad r=1, \dots, s$$

$$\lambda_j \geq 0 \quad j=1, \dots, n$$

**Definition 1.** Productivity-possibility set contains units, which can produce the output  $Y$  by using input  $X$ . Based on this concept and the CCR model, the possibility set is as comes next:

$$T = \left\{ \begin{pmatrix} X \\ Y \end{pmatrix} \left| X \geq \sum_{j=1}^n \lambda_j x_j \text{ \& } Y \leq \lambda_j Y_j \text{ \& } \lambda_j \geq 0, j = 1, \dots, n \right. \right\}. \quad (1)$$

DEA model assumes that the underlying production possibility set, denoted by  $T$ , satisfies the following axioms:

(A1) Envelopment:  $(x_j, y_j) \in T \forall j = 1, 2, \dots, n$ .

(A2) Possibility: if  $x$  can produce  $y$ , then  $\bar{x} \geq x$  can produce  $y$ , also  $\bar{y} \leq y$  can be produced by  $x$ .

(A3) Constant returns to scale:  $\forall (x, y) \in T, \forall \lambda \geq 0 \Rightarrow (\lambda x, \lambda y) \in T$ .

(A4) Convexity: if  $(x, y)$  and  $(x', y') \in T, \forall \lambda \geq 0$  then  $\lambda(x, y) + (1 - \lambda)(x', y') \in T$

**Definition 2.** If in enveloped form of CCR model is input oriented:

- a.  $\theta^* = 1$  and values of the slack variables in all optimal solutions be zero, then  $DMU_o$  is strongly efficient (or Parato efficient).
- b. If  $\theta^* < 1$ , then  $DMU_o$  is inefficient.

**Definition 3.** Let  $DMU_j$  use the input  $x_j$  for producing  $y_j$ . It is said that  $DMU_k$  dominates on  $DMU_f$  when

$$(-x_{1k}, \dots, -x_{mk} \text{ and } y_{1k}, \dots, y_{sk}) \geq (-x_{1t}, \dots, -x_{mt} \text{ and } y_{1t}, \dots, y_{st}). \quad (2)$$

Note that the strong inequality is true for at least component.

**Definition 4.** Efficient benchmarks are units that with the same benchmark units always dominant on inefficient DMU which are also called as reference units.

Note that in the envelopment form of CCR model for an inefficient DMU; the reference set ( $E_o$ ) can be defined as follows:

$$E_o = \{j | \lambda_j^* > 0\}$$

It is the existence of this collection of efficient DMUs that forces the  $DMU_o$  are inefficient. The set spanned by  $E_o$  is called the *efficient frontier* of  $DMU_o$ .

### 3. An algorithm for finding benchmark units

Suppose we have a decision-making unit with inefficient integer input and output. In order to increase the efficiency, we try to introduce the efficient benchmark with integer values to the inefficient units. Thus, in viewpoint of DEA, this unit is projected on the efficiency boundary, as done in Fig 1. Therefore, the proposed algorithm for determining integer benchmark is as follows:

**Step 1.** First, we find the intersection of hyperplane equation including the under evaluation DMU with the benchmark DMU hyperplane equation. If the intersection contains integer inputs/outputs; then we consider that the point as the benchmark DMU and if not, we go to Step 2.

**Step 2.** If the model is input-oriented; we consider the input point as: the integer part of the point plus 1. In case of output-oriented the integer part of the intersection point is taken into consideration as the integer point.

**Step 3.** The new integer point is a feasible point. To avoid dominating of this point by other integer points, we satisfy each of the integers input/output in the hyperplane equation including DMU. In case of being non-integer, go to Step 2.

At the end, if the inefficient unit be projected on this obtained point (which was attained through the above steps of our algorithm), then there will be an integer benchmark.

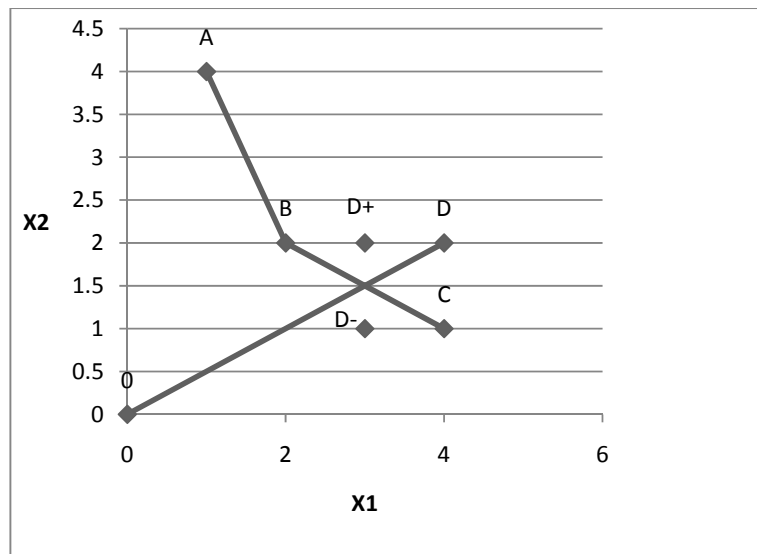
#### 4. A numerical example

Consider four decision-making units with two inputs and one output, according to Table 1. (The first input is the number of workers; the second input is the number of productive equipments while the output is considered to be the number of produced goods in four factories). Note that all data in table 1 are in base 100.

**Table 1**

Input/output data for four decision-making units

DMU	A	B	C	D
$I_1$	1	2	4	4
$I_2$	4	2	1	2
O	1	1	1	1



**Fig. 1.** The way of projecting for two fixed inputs and one out put

A, B, C are integer units while D is an inefficiency unit. For increasing the efficiency of unit D, we project this DMU on the efficiency boundary. It is observed from Fig 1. that the projection of unit D on the efficiency boundary is not integer. If we round this point to  $D^+$ , then it will be dominated by DMU B. If we round off to  $D^-$ , which is not in the feasible region. To remedy this, we use the proposed algorithm as comes next

*Step 1.* We have

The hyperplane binding in DMU D :  $\alpha = \frac{1}{2}\beta$

BC hyperplane:  $\alpha = -\frac{1}{2}\beta + 3$

$$\begin{cases} \alpha - \frac{1}{2}\beta = 0 \\ \alpha + \frac{1}{2}\beta = 3 \end{cases} \Rightarrow \beta = 3, \alpha = \frac{3}{2} \quad (3)$$

*Step 2.* Since the model is input-oriented then we have  $[\alpha]^+ = 2$ , and thus  $\tilde{D} \approx \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ . The point  $\tilde{D} \approx \begin{pmatrix} 3 \\ 2 \end{pmatrix}$  is a feasible point, which is dominated by DMU B. Thus, we go to Step 3.

*Step 3.* Let us satisfy the point  $\tilde{D} \approx \begin{pmatrix} 3 \\ 2 \end{pmatrix}$  in the hyperplane equation of BC as follows:

$$\text{BC hyperplane: } \alpha = -\frac{1}{2}\beta + 3 \Rightarrow \alpha = -\frac{1}{2} \times 3 + 3 = \frac{3}{2}$$

$$\text{then } \alpha = \left\lceil \frac{3}{2} \right\rceil + 1 = 2$$

$$\beta = 6 - 2\alpha \Rightarrow \beta = 6 - 4 = 2$$

The point  $D^* = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$  is an efficient integer benchmark that with input  $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$  will produce the output 1 which by projecting on the unit B; it will have the same output by the input  $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ . Such algorithms are more economic, due to producing same output by considering lower input. Thus, the best projection for DMU D; is projecting on DMU B.

## 5. Conclusion

The problem considered in this research was to find an integer benchmark for inefficient units by integer input and outputs. The proposed model of this paper first project the inefficient units on the efficient boundary and in to increase the efficiency; the integer efficiency benchmark has been compared for the inefficient DMU. If for evaluating of inefficient decision-making units, the first projecting point is on the efficiency frontier and be non-integer point, then we do the steps of the proposed algorithm. The implementation of the proposed model of this paper has been demonstrated using some simple example.

## Acknowledgment

The authors would like to thank the anonymous referees for their constructive comments on earlier version of this work.

## References

- Banker, R.D., Charnes, A., & Cooper, W.W. (1985). Some model for measuring technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30, 1078-1092.
- Charnes, A., Cooper W.W., & Rhodes E. (1978). Measuring the efficiency of decision making units. *European Journal of Operations Research*, 2(6), 429– 444.
- Charnes, A., Cooper, W.W., Golany, B., & Seiford, L.M. (1985). Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions. *Journal of Econometrics*, 30, 91-107.

- Kuosmanen, T., & KazemiMatin, R. (2009). Theory of integer valued data envelopment analysis. *European Journal of Operational Research*, 192, 658–667.
- Lozano S., & Villa G., (2006). Data envelopment analysis of integer-valued inputs and outputs. *Computers & Operations Research*, 33(10), 3004–3014.
- Tone, K. (2001). A Slack-Based Measure of Efficiency in Data Envelopment Analysis. *European Journal of Operational Research*, 130, 498-509.