Applied Mathematical Sciences, Vol. 6, 2012, no. 116, 5769 - 5774

A Comment on Theory of Integer-Valued Data Envelopment Analysis

Dariush Khezrimotlagh*, Shaharuddin Salleh and Zahra Mohsenpour

Department of Mathematics, Faculty of Science, UTM, Johor, Malaysia

Abstract

The mixed-integer linear programming data envelopment analysis models were proposed due to this reason that the conventional DEA models may suggest some non-integer values for the inputs and/or outputs of decision making units (DMUs) which those data can only be specified with integer values such as the number of employees and books. There are a few researches on integer-valued DEA and this paper focuses on the three previous ones. The paper characterizes some shortcomings on those researches with some mathematic logics.

Mathematics Subject Classification: 90

Keywords: Data envelopment analysis, Mixed-integer linear programming.

1. Introduction

Data envelopment analysis (DEA) is a nonparametric technique in operations research which is proposed to measure the efficiency and productivity of decision making units (DMUs) with multiple inputs and outputs. Each DMU has usually some input and output values which some of those values can only be integer numbers such as the number of magazines and passengers. However, the conventional DEA models may suggest some targets for inefficient DMUs which those corresponding components have non-integer values. Therefore, Lozano and Villa [5] proposed a model to improve the conventional DEA models. Next, Kuosmanen and Kazemi Matin [4] claimed that there are some problems on Lozano and Villa's work and proposed some new axioms and a model to remove those problems. This paper illustrates the shortcomings in some of those claims and it is organized in four sections. Section 2 is the background. The critiques on those papers and the conclusions are illustrated in Sections 3 and 4, respectively.

* Corresponding author e-mail address: *khezrimotlagh@gmail.com*

2. Background

DEA was proposed by Charnes et al. [2] based on the following axioms where there are n DMUs ($\text{DMU}_i, i = 1,2,\dots,n$) with m nonnegative inputs ($x_{ij}, j = 1,2,\dots,m$) and p nonnegative outputs ($y_{ik}, k = 1,2,\dots,p$) for each DMU which at least one of its inputs and one of its outputs are not zero.

a) *Inclusion of observation*: each observed DMU_i belongs to T, $i = 1,2,\dots,n$.
b) *Free disposability of inputs and outputs*: if $(x,y) \in T$, $x' \geq x$, then $(x',y) \in T$, and if $(x,y) \in T$, $y' \leq y$, then $(x,y') \in T$.
c) *Convexity*: if $(x,y) \in T$ and $(x',y') \in T$, then $[\lambda(x,y) + (1-\lambda)(x',y')] \in T$, for $\lambda \in [0,1]$.
d) *Constant returns to scale*: if $(x,y) \in T$ then $(\lambda x, \lambda y) \in T$, for $\lambda \geq 0$.
e) *Minimum extrapolation*: T is the intersection of all sets satisfying (a)-(d).

Lozano and Villa [5] assumed that some observed x_{ij} and y_{ik} of DMUs are integer for $j \in I_I \subseteq I = \{1,2,\dots,m\}$ and $k \in O_I \subseteq O = \{1,2,\dots,s\}$ where I and O are two sets of inputs and outputs dimensions and I_I and O_I the subsets of the corresponding dimensions that must be integer; and supposed the following production possibility set (PPS):

$$T = \{(\hat{x},\hat{y}): \hat{x}_j \geq \textstyle\sum_{i=1}^{n} \lambda_i x_{ij}\ \forall j, \hat{y}_k \leq \sum_{i=1}^{n} \lambda_i y_{ik}\ \forall k, \hat{x}_j, \hat{y}_k \in \mathbb{Z}_{\geq 0}\ \forall j \in I_I, k \in O_I\}$$

Next, Kuosmanen and Kazemi Matin [4] proposed the following axioms for making the PPS with integer values directly.

A) *Natural disposability*: $(x,y) \in T, (u,v) \in \mathbb{Z}_+^{m+s}, y \geq v \Rightarrow (x+u, y-v) \in T$.
B) *Natural convexity*: $(x',y'),(x'',y'') \in T, (x,y) = \lambda(x',y') + (1-\lambda)(x'',y'')$, $0 \leq \lambda \leq 1$ and $(x,y) \in \mathbb{Z}_+^{m+s} \Rightarrow (x,y) \in T$.
C) *Natural divisibility*: $(x,y) \in T, \exists\lambda \in [0,1]; (\lambda x, \lambda y) \in \mathbb{Z}_+^{m+s} \Rightarrow (\lambda x, \lambda y) \in T$.
D) *Natural augment-ability*: $(x,y) \in T, \exists\lambda \geq 1; (\lambda x, \lambda y) \in \mathbb{Z}_+^{m+s} \Rightarrow (\lambda x, \lambda y) \in T$.
E) *Minimum extrapolation*: T is the intersection of all sets satisfying (A)-(D).

The Lozano and Villa's MILP and Kuosmanen and Kazemi Matin's MILP are as follows where both of them have two below stages:

1) Minimizing θ_l i.e., considering only possible input that decreases while keeping at least the present output levels.

2) Maximizing the summation of slacks (s_j^- and s_k^+). The slacks are non-negative numbers which shows the potential reduction on inputs and potential increase of outputs.

Lozano and Villa's model:

$min\,\theta_l - \varepsilon(\sum_{j=1}^{m} s_j^- + \sum_{k=1}^{p} s_k^+)$,
Subject to
$\sum_{i=1}^{n} \lambda_i x_{ij} = x_j \quad \forall j \in I$,
$x_j = \theta_l x_{lj} - s_j^- \ \ \forall j \in I$,
$\sum_{i=1}^{n} \lambda_i y_{ik} = y_k \quad \forall k \in O$,
$y_k = y_{lk} + s_k^+ \ \ \forall k \in O$,
$\lambda_i \geq 0 \ \ \forall i$,
$s_j^-, x_j \geq 0 \ \ \forall j \in I$,
$s_k^+, y_k \geq 0 \ \ \forall k \in O$,
θ_l free,
x_j integer $\forall j \in I_I$,
y_k integer $\forall k \in O_I$.

Kuosmanen and Kazemi Matin's model:

$min\,\theta_l - \varepsilon(\sum_{j=1}^{m} s_j^- + \sum_{k=1}^{p} s_k^+ + \sum_{j=1}^{p} s_j^I)$,
Subject to
$\sum_{i=1}^{n} \lambda_i x_{ij} + s_j^- = \theta_l x_{lj} \quad \forall j \in I - I_I$,
$\sum_{i=1}^{n} \lambda_i y_{ik} - s_k^+ = y_{lk} \quad \forall k \in O$,
$\sum_{i=1}^{n} \lambda_i x_{ij} + s_j^- = x_j \quad \forall j \in I_I$,
$x_j = \theta_l x_{lj} - s_j^I \ \ \forall j \in I_I$,
x_j integer $\forall j \in I_I$,
$\lambda_i \geq 0 \ \ \forall i$,
$s_j^- \geq 0 \ \ \forall j \in I$,
$s_k^+ \geq 0 \ \ \forall k \in O$,
$s_j^I, x_j \geq 0 \ \ \forall j \in I_I$,
θ_l free.

3. The shortcomings in the theory of integer-valued DEA

Although, there is no any problem to propose some axioms for integer values in DEA, the conventional DEA axioms are enough to operate with both integer and real values. Indeed, when the domain of data are restricted in the integer numbers set, it is only enough to restrict the DEA axioms by the integer numbers set. For instance, as it is supposed that some data are restricted to be integer, the convexity axiom is also able to restrict to be integer i.e., $\{(x,y): (x,y) = \lambda(x',y') + (1-\lambda)(x'',y''), (x',y') \in T, (x'',y'') \in T, 0 \leq \lambda \leq 1\} \cap \mathbb{Z}_{\geq 0}^{m+s}$. This is a simple rule of the Sets Theory to restrict a domain or an axiom. However, Kuosmanen and Kazemi Matin in the first sentence of the first paragraph in section 7 of page 665 in their paper [4] claimed that "*The conventional axioms of convexity and free disposability fail if DMUs are restricted to operate with integer-valued input and output quantities*." In fact, they would want to make the PPS directly from the integer numbers set and there are no any weaknesses for this purpose, but the conventional DEA axioms are never failed where the data of DMUs are restricted to integer values, because they should just restrict to the integer numbers set similar to restrict data.

On the other side, as it can be seen in their proposed axioms, they used the real number variable λ to write the axioms (B), (C) and (D). Now, if it has been supposed that only the integer numbers set is considered, then it should not have been used the real number variable in the integer axioms! In fact, a new axiom must not have any doubts or parallel affects with those previous axioms. In other words, an axiom is an evident premise as to be accepted as true without controversy. If the real numbers set can be selected, it is enough to make the PPS from the real numbers set and then it should be restricted to the integer numbers set. In other words, suppose that there are n DMUs ($\text{DMU}_i, i = 1,2,\dots,n$) with these integer/real inputs and outputs components $(x_{j\in I_I}, x_{j\in I-I_I}, y_{k\in O_I}, y_{k\in O-O_I})$. Since $\mathbb{Z} \subset \mathbb{R}$, then the production possibility set T to those DMUs can be made by conventional DEA axioms. Now, the corresponding PPS to those DMUs can simply be restricted in $\mathbb{Z}_{\geq 0}^{|I_I|} \times \mathbb{R}_{\geq 0}^{m-|I_I|} \times \mathbb{Z}_{\geq 0}^{|O_I|} \times \mathbb{R}_{\geq 0}^{p-|O_I|}$. This making of a production possibility set is exactly the intersection of all the PPS which includes those DMUs with the proposed axioms and it is also the same as the proposed PPS by Lozano and Villa. Therefore, if the aim of proposing the integer axioms allows using the real numbers, there is no any worth for introducing an extra axiom in DEA.

In order to remove the above shortcoming in the integer DEA axioms it is proposed to replace λ with two integer numbers u and v, and define $\mathbb{Z}_{+} = \mathbb{N} \cup \{0\} = \{0,1,2,\ldots\}$. The corrected axioms are as following where $\frac{a}{b}c \in \mathbb{Z}$, for $a, b, c \in \mathbb{Z}$, if $b|c$, i.e., b divides c, that is, there is an integer t such that $c = bt$.

B. *Integer convexity*: $(x', y'), (x'', y'') \in T,\ \exists u, v \in \mathbb{Z}_{+},\ u \leq v, (x, y) \in \mathbb{Z}_{+}^{m+s}$ and $(x, y) = \frac{u}{v}(x', y') + \left(1 - \frac{u}{v}\right)(x'', y'') \Longrightarrow (x, y) \in T$,

C. *Integer divisibility*:

$$(x, y) \in T, \exists u, v \in \mathbb{Z}_{+}: u \leq v, \left(\frac{u}{v}x, \frac{u}{v}y\right) \in \mathbb{Z}_{+}^{m+s} \Rightarrow \left(\frac{u}{v}x, \frac{u}{v}y\right) \in T,$$

D. Integer augment-ability:

$$(x, y) \in T, \exists u, v \in \mathbb{Z}_{+}: u \geq v, \left(\frac{u}{v}x, \frac{u}{v}y\right) \in \mathbb{Z}_{+}^{m+s} \Rightarrow \left(\frac{u}{v}x, \frac{u}{v}y\right) \in T.$$

Furthermore, Kuosmanen and Kazemi Matin in the second sentence in the third line of page 659 in their paper [4] mentioned that "*We develop a new axiomatic foundation for integer-valued DEA models, and show that the production possibility set proposed by Lozano and Villa (2006) is consistent with the proposed set of axioms*." However, in the end sentence of page 658 [4], they contrary claimed that "*Lozano and Villa's model is not consistent with the minimum extrapolation principle (Banker et al., 1984)* [1]*, which is the foundation of all DEA models.*" Although, the word 'model' may not be interpreted for the axioms, the earlier sentences in that address undoubtedly illustrate that they mentioned to the proposed PPS of Lozano and Villa. Moreover, if it is also supposed that they mentioned to the model of Lozano and Villa, it is easy to prove that the constraints of Lozano and Villa's model make the same production possibility set as the PPS which is made with those proposed axioms. It seems that they would just want to mention that the Lozano and Villa did not illustrate how they defined the PPS, because the proposed PPS by Lozano and Villa is a set of the intersection of all those sets which are satisfied by axioms (A) to (D) and it has precisely the minimum extrapolation principle property.

On the other hand, Kuosmanen and Kazemi Matin in the last sentence in the second paragraph in the page 664 in their paper [4] claimed that for DMU 13: "*Lozano and Villa's MILP formulation yields input target* $(0{,}874{,}0)$*, with two units higher target value for input 2. The result is due to the fact that it is impossible to find intensity weights* λ *that satisfy both* $\sum_{i=1}^{42} x_i\lambda_i = (0{,}872{,}0)$ *and constraints* $\sum_{i=1}^{42} y_i\lambda_i \in \mathbb{Z}_{+}^{4}$, $\sum_{j=1}^{42} y_j\lambda_j \geq (812, 8, 10, 2)$ *simultaneously*".

This claim has two major shortcomings. First, since the point $(0,872,0; 812, 8, 10, 2)$ belongs to the PPS and $\sum_{j=1}^{42} x_j\lambda_j = (0,872,0)$ and $\sum_{j=1}^{42} y_j\lambda_j \geq (812, 8, 10, 2)$, therefore, there is absolutely a $\mu \in \mathbb{R}_{\geq 0}^{42}$ which yields that $\sum_{j=1}^{42} x_j\mu_j = (0,872,0)$ and $\sum_{j=1}^{42} y_j\lambda_j = (812, 8, 10, 2)$. Because, if there is no any $\mu \in \mathbb{R}_{\geq 0}^{42}$ which yields those equalities, it means that the point $(0,872,0; 812, 8, 10, 2)$ does not belong to the PPS. Now, the constraints of Lozano and Villa's model clearly illustrate that it is undoubtedly possible to find an intensity to satisfy those equalities, however, Kuosmanen and Kazemi Matin claimed that it is impossible unfortunately.

Second, the Lozano and Villa's MILP never suggests input target $(0,874,0)$ for DMU 13 where the hypothesis of output is considered the same as what it was considered for Kuosmanen and Kazemi Matin's MILP. In fact, before the applying the models Kuosmanen and Kazemi Matin should have characterized whether the outputs are restricted in integer vales or not. Although, from their proposed MILP it is easy to know that they supposed the outputs as real, they did not unfortunately suppose the same hypothesis for outputs to apply Lozano and Villa's MILP. In other words, they assumed that the outputs are as integer and then applied Lozano and Villa's model. This is a simple reason for that mistake, but it unfairly leads the weaknesses of Lozano and Villa's model, whereas it suggests the same targets for DMU 13 i.e., $(0,872,0; 812, 8, 10, 2)$, where the output are not restricted to integer values like the proposed output for Kuosmanen and Kazemi Matin's model. In fact, in this case $O_I = \{\,\}$ and from the simple rule in Sets Theory and Mathematics Logic the corresponding constraints should be removed in Lozano and Villa's MILP, because those constraints do not have any effects in optimization. Therefore, both output constraints of those models are the same where $O_I = \{\,\}$. In fact, the results of Lozano and Villa's MILP in those tables absolutely depict that Kuosmanen and Kazemi Matin supposed different hypothesis for outputs and applied the models which shows the big shortcoming in comparing the models.

Unfortunately, not only the results for DMU 13 are the same by applying both MILPs, but also the targets and efficiency scores of all 42 universities are the same by applying those models and there is no any difference between those MILP in that numerical example. In other words, although, those MILP are different, their proposed numerical example is never able to characterize the distinction between the models.

The above shortcomings are also in the tables in the paper of Kazemi Matin and Kuosmanen [3]. Indeed, for their numerical example there is no any difference between Lozano and Villa's model and Kuosmanen and Kazemi Matin' model under alternative returns to scale. In other words, the results for Kuosmanen and Kazemi Matin's MILP in those tables should have written for Lozano and Villa's MILP, too. Moreover, the result for Lozano and Villa's model in those tables are correct when the outputs are integer. In this case, their proposed model should have improved and its results will also be the same as Lozano and Villa's model for the proposed numerical example.

4. Conclusion

This paper illustrates some of the most important shortcomings in Kuosmanen and Kazemi Matin's papers and improves their proposed axioms for integer-valued in DEA. The paper clearly illustrates that the results of Tables in Kuosmanen and Kazemi Matin's papers are not significant. In fact, the results of Lozano and Villa's MILP and Kuosmanen and Kazemi Matin's MILP are the same for the 42 universities in the numerical example of Kuosmanen and Kazemi Matin's papers.

References

[1] R.J. Banker, A. Charnes, W.W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science **30** (1984), 1078-1092.

[2] A. Charnes, W.W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, European Journal Operational Research, **2** (1978), 429-444.

[3] R. Kazemi Matin, T. Kuosmanen, Theory of integer-valued data envelopment analysis under alternative returns to scale axioms, Omega **37** (2009), 988-995.

[4] T. Kuosmanen, R. Kazemi Matin, Theory of integer-valued data envelopment analysis, European Journal of Operational Research, **192** (2009), 658-667.

[5] S. Lozano, G. Villa, Data envelopment analysis of integer-valued inputs and outputs, Computer and Operations Research, **33** (2006), 3004-3014.

Received: June, 2012