"Science Stays True Here"
Journal of Mathematics and Statistical Science (ISSN 2411-2518, USA), Vol.6, Issue 6, 187-194 | Science Signpost Publishing

An Inverse DEA Model for Input/Output Estimation with Integer Restriction

K G Shinto[1], C M Sushama[2]

1. Department of Mathematics, Christ College Irinjalakuda, India, shintomaths@gmail.com.

2. Department of Mathematics, National Institute of Technology Calicut, India.

Abstract

Basic DEA model considers real valued inputs and outputs and measures the relative efficiency of similar decision making units (DMUs). If a particular DMU is not efficient, a set of target values for the input/output can be suggested by projecting the DMU to the efficient frontier which will make the DMU efficient. Conventional inverse DEA model is used to estimate the input (or output) for a decision making unit (DMU) when some or all outputs (or inputs) of this DMU is changed. But in a DEA model with integer restriction in input and/or output values, these models are not acceptable if the estimated values are not integers. Rounding off these values may not provide the result in all cases. This paper presents a situation when the input values of an efficient DMU are increased. With these new input values and with the existing output values this DMU may become inefficient. The projection of this DMU into the efficient frontier may not preserve the integer restriction on the output values. The paper presents an iterative MILP model to handle the situation.

Keywords: Data Envelopment Analysis (DEA), Inverse DEA, Mixed integer linear programming (MILP), CRS DEA, Output oriented DEA.

1. Introduction

Data Envelopment Analysis (DEA) is a linear programming based technique used to evaluate the relative efficiency of similar decision making units (DMUs) by comparing it with the other DMUs that converts single/multiple inputs into single/multiple outputs. DEA originated from Farell's seminal work [6] and popularized by Charnes et al. [4]. Since its development in 1978 by Charnes et al. [4], DEA has become one of the core tools available for measuring the organizational performances. From 1978 onwards several books, articles and dissertations are published involving DEA and hundreds of applications of DEA methods have been reported [5]. Traditional DEA models assume real valued inputs and outputs. But there are many occasions in which some inputs and/or outputs can assume only integer values.

Wei et al. [13] introduced the inverse DEA for the input output estimation. Their work is based on the extension of the inverse optimization problem. After this work, Jahanshahloo et al .[7] extended the inverse DEA by considering extra input. But these models assume only real valued inputs and outputs. Lozano and Villa [11] are the first who handled integer data in DEA. Based on this work Kazemi Matin and Kuosmanen T. [10], Khezrimotlagh et al. [8,9] and Lozano and Villa [12] extended this integer DEA by considering their own models. All these models are developed based on the MILP formulation of the problem. In this paper, we discuss an inverse integer DEA problem and give an iterative MILP method to solve such problems.

The rest of the paper is organised as follows. In section 2 basic definitions and theoretical background of the existing integer DEA and inverse DEA models is explained. Section 3 explains the inverse integer DEA model and the computational algorithm for output estimation. Section 4 is devoted to explain the theoretical background of the algorithm. In section 5 example for the algorithm is explained.

2. Basic definitions for integer DEA and inverse DEA models

2.1 Integer DEA

Let $X = (\bar{X}_1, \bar{X}_2 \ldots \ldots \bar{X}_n)$ and $Y = (\bar{Y}_1, \bar{Y}_2 \ldots \ldots \bar{Y}_n)$ be the input and output matrices where $\bar{X}_j$ and $\bar{Y}_j$ denote the input and output vectors for DMU_j for $j = 1,2, \ldots \ldots . n$. The production possibility set corresponding to the CRS technology is given by

$$T_{CRS} = \{(\bar{X}, \bar{Y}): \exists \lambda_1, \lambda_1, \ldots . . \lambda_n; \lambda_j \geq 0 \forall j \quad \bar{X} \geq X\lambda, \bar{Y} \leq Y\lambda\}$$

The efficient frontier is a subset of the PPS formed by the set of all non dominating points. That is,

$$T_{CRS}^{eff} == \{(\bar{X}, \bar{Y}) \in T_{CRS}: \forall (\bar{X}', \bar{Y}') \in T_{CRS}, (\bar{X}' \leq \bar{X}) \cap (\bar{Y}' \geq \bar{Y}) \leftrightarrow (\bar{X}', \bar{Y}') = (\bar{X}, \bar{Y})\}$$

Now if some or all of the input and/or output components can take only integer values, then the PPS of such DEA models is the subset of T_{CRS} satisfying the integer restriction. Let $I = \{1,2, \ldots . m\}$ and $O = \{1,2, \ldots \ldots p\}$ be the set of input and output dimensions and let $O' \subseteq O$ and $I' \subseteq I$ be the dimensions for integer valued. Then the CRS integer PPS is given by

$$\begin{aligned} T'_{CRS} = \{(\bar{X}, \bar{Y}): \exists \lambda_1, \lambda_1, \ldots . . \lambda_n; \lambda_j \geq 0 \forall j \\ \bar{X} \geq X\lambda, \bar{Y} \leq Y\lambda; \bar{X}_i \in Z, \forall i \in I'; \\ \bar{Y}_i \in Z, \forall i \in O'\} \end{aligned}$$

DMU_j is said to be integer efficient if no other integer valued point of PPS dominates it. The CRS integer efficient frontier is therefore the set of non dominated integer valued points of the PPS.

Now if a particular DMU is CRS efficient then it will be in the efficient frontier T_{CRS}^{eff}. Again since its target values are the given values, it satisfies the integer restriction and so it will be in the integer efficient frontier. That is, if DMU_j is CRS efficient, then it is CRS integer efficient. The converse is obviously not true.

The classical input oriented CRS model is defined as follows

$$\begin{aligned} &\max \ \theta_j + \epsilon(S^- + S^+) \\ &\text{Subject to } X\lambda - \bar{X}_j + S^- = 0 \\ &\qquad Y\lambda - S^+ = \theta_j \bar{Y}_j \\ &\qquad \lambda, S^-, S^+ \geq 0 \end{aligned}$$

It is a two phase LPP algorithm. In phase I maximum $S^- + S^+$ is considered and using this maximum θ is considered in phase II. DMU_j is CRS efficient if $S^- = S^+ = 0$ and $\theta_j = 1$. If DMU_j is not CRS efficient, then either $\theta_j > 1$ or S^-or S^+ is non zero. The target values for such a DMU is given by

$$\hat{X}_j = \bar{X}_j - S^-$$
$$\hat{Y}_j = \theta_j \bar{Y}_j + S^+$$

to make the DMU efficient. With this transformation the DMU will project into the efficient frontier and will thus become efficient.

Now if there is integer restriction then the target $(\hat{X}, \hat{Y})$ should be integer vectors. Lozano and Villa [12] introduced a model to solve output oriented CRS integer valued problems. It is a two phase model and is as follows

$$\max\ \theta_j + \epsilon\left(\sum_{i} \mathrm{s}_{\mathrm{i}}^{-} + \sum_{k} \mathrm{s}_{\mathrm{k}}^{+}\right)$$

subject to

$$\sum_j \lambda_j x_{ij} = x_i \forall i \notin I'$$

$$\sum_j \lambda_j x_{ij} \leq x_i \forall i \in I'$$

$$x_i = x_{ij} - s_i^-$$

$$\sum_j \lambda_j\, y_{kj} = y_k \forall k \notin O'$$

$$\sum_j \lambda_j\, y_{kj} \geq y_k \forall k \in O'$$

$$y_k = \theta_j y_{kj} + s_k^+$$

$$\lambda_j, s_i^-, s_k^+, x_i, y_k \geq 0$$

x_i, y_k are integers $\forall i \in I' \& \forall k \in O'$

Where θ_j is the CRS integer efficiency score. An existing DMU, $\mathrm{DMU_j}$ is CRS integer efficient if and only if $\theta_j = 1, s_i^- = 0\ \forall i\ \&\ s_k^+ = 0\ \forall k.$

The classical additive CRS DEA model is defined as follows

$$\begin{aligned} &\max\ \ (S^- + S^+) \\ &\text{Subject to} \quad X\lambda + \mathrm{S}^- = \overline{\mathrm{X}}_{\mathrm{j}} \\ &\qquad\qquad\quad\ \mathrm{Y}\lambda - \mathrm{S}^+ = \overline{\mathrm{Y}}_{\mathrm{j}} \\ &\qquad\qquad\quad\ \lambda, \mathrm{S}^-, \mathrm{S}^+ \geq 0 \end{aligned}$$

An existing DMU, $\mathrm{DMU_j}$ is ADD efficient if and only if $s_i^- = 0\ \forall i\ \&\ s_k^+ = 0\ \forall k.$

Lozano and Villa's integer additive model [] is given by

$$\max\ \left(\sum_{i} \mathrm{s}_{\mathrm{i}}^{-} + \sum_{k} \mathrm{s}_{\mathrm{k}}^{+}\right)$$

subject to

$$\sum_j \lambda_j x_{ij} = x_i \forall i \notin I'$$

$$\sum_j \lambda_j x_{ij} \leq x_i \forall i \in I'$$

$$x_i = x_{ij} - s_i^-$$

$$\sum_j \lambda_j\, y_{kj} = y_k \forall k \notin O'$$

$$\sum_j \lambda_j\, y_{kj} \geq y_k \forall k \in O'$$

$$y_k = y_{kj} + s_k^+$$

$$\lambda_j, s_i^-, s_k^+, x_i, y_k \geq 0$$

x_i, y_k are integers $\forall i \in I' \& \forall k \in O'$

By this model a DMU is integer efficient if $s_i^- = 0\ \forall i$ & $s_k^+ = 0\ \forall k$.

2.2 Inverse DEA

The inverse DEA model proposed by Wei et al. [13] is based on a multi objective optimization problem. The algorithm becomes more complicated if there is more number input and output factors. They illustrated their algorithm for single input-single output case. Even this case is also very complicated for target evaluation. Also the algorithm is developed for real valued inputs and outputs. So this algorithm is not suitable for integer valued inputs and outputs.

The inverse DEA model proposed by Jahanshahloo et al. [7] is also based on multi objective optimization and is developed for input estimation. This paper is also developed for real valued inputs and outputs. So this algorithm is also not suitable for integer valued inputs and outputs.

So the existing models for inverse DEA models are developed for real valued inputs and outputs and its solution methodology is very complicated even for real valued factors. So imposing integer restriction on these models will definitely give more complicated algorithms. Hence we consider Lozano and Villa model for inverse DEA with integer restriction on input and output factors.

3. Inverse integer DEA model and Computational Algorithm

Given $X = (\bar{X}_1, \bar{X}_2 \dots \dots \bar{X}_n)$ and $Y = (\bar{Y}_1, \bar{Y}_2 \dots \dots \bar{Y}_n)$, respectively the input and output matrices where $\bar{X}_j$ and $\bar{Y}_j$ denote the input and output vectors for DMU_j for $j = 1,2, \dots \dots . n$. Assume that the input of DMUj, $\bar{X}_j$ is increased by a quantity $\delta\bar{X}_j$. That is the revised input of DMUj is $\bar{X}_j^* = \bar{X}_j + \delta\bar{X}_j$. We have to estimate the output corresponding to this input. Let X^*denote the input matrix so that $X^* = (\bar{X}_1, \bar{X}_2 \dots \dots \bar{X}_n)$ whose jth column is $\bar{X}_j^* = \bar{X}_j + \delta\bar{X}_j$.

The computational algorithm of the proposed iterative output oriented inverse integer DEA model to estimate the output $\bar{Y}_j^*$ of DMUj is as follows

Step 1:

Solve the mixed integer LP problem

$$
\begin{aligned}
&\max\ (S^- + S^+) \\
&\text{Subject to} \quad X^*\lambda + S^- = \bar{X}_j \\
&\qquad\qquad\quad Y\lambda - S^+ = \bar{Y}_j^* \\
&\qquad\qquad\quad \lambda, S^-, S^+ \geq 0 \\
&\qquad\qquad\quad S^-, S^+ \text{ are integers.}
\end{aligned}
$$

Let $(\lambda, S^{-*}, S^{+*})$ be the optimal solution

If $S^{-*} = S^{+*} = 0$, then stop.

$\bar{Y}_j^* = \bar{Y}_j$ is the target output so that with this output DMUj is efficient and $(\bar{X}_j^*, \bar{Y}_j^*)$ will be a point in the integer efficient frontier.

Otherwise if $S^{-*} \neq 0$ and $S^{+*} = 0$, then also stop.

$\bar{Y}_j^* = \bar{Y}_j$ is the target output. But for this output only $\bar{X}_j^* - S^{-*}$ of input is required and hence $(\bar{X}_j^* - S^{-*}, \bar{Y}_j^*)$ will be a point in the integer efficient frontier.

Otherwise go to step 2.

Step 2:

Set $\bar{Y}_j^* = \bar{Y}_j^* + S^{+*}$

Then replace the jth column of Y with this $\bar{Y}_j^*$ and then go to step 1.

4. Explanation of the Algorithm

The algorithm is designed so that each iteration of the algorithm tries to reduce the output shortfall present in the output factors so that after the final iteration the output shortfall becomes zero. The iterative algorithm estimates the output value for DMUj so that with this output, the DMU will place in the integer efficient frontier.

The theoretical validity of the algorithm is explained using Lozano and Villa [11] model. By this model an existing DMU, DMU_j is integer efficient if and only if $\theta_j = 1, s_i^- = 0\ \forall i\ \&\ s_k^+ = 0\ \forall k$. The proposed algorithm is designed so that $\theta_j = 1$. So we assume $\theta_j = 1$ in Lozano and Villa's model and the algorithm proceeds so as to get the condition $s_i^- = 0\ \forall i\ \&\ s_k^+ = 0\ \forall k$ in the final iteration. In Lozano and Villa's model the assumption of $\theta_j = 1$ will imply that the input excesses and output shortfalls of the considered DMUj should be integer vectors. So we get the following result.

Result 1

In the Lozano and Villa model if $\theta_j = 1$, then the slacks s_i^- and s_k^+ should be integers for all $i \in I'$ and for all $k \in O'$

The proof is obvious if we consider $\theta_j = 1$ in the algorithm of Lozano and Villa.

So with the help of this result we developed the model as an iterative MILP by a modification to the basic CCR model such that $\theta_j = 1$ and S^-, S^+ as integer vectors. In the final iteration we get $S^- = S^+ = 0$. Hence the corresponding DMU should become integer efficient with the estimated output vector. If either S^- or S^+ is non zero, then the corresponding DMU is not efficient and to achieve efficiency it is required to revise the input, output vectors of the DMU. For this first of all we consider the transformation $\hat{Y}_j = \hat{Y}_j + S^{+*}$ to reduce the output shortfall by increasing the output. After reducing all possible output shortfalls if there is any positive input excess, we will remove it by using the transformation $\hat{X}_j = \bar{X}_j - S^{-*}$.

As a consequence of the above discussion we get the following result which ensures the efficiency of the algorithm.

Result 2

The integer target value obtained using the algorithm is a point in the integer efficient frontier.

The convergence of the algorithm is obvious. The iterative algorithm proceeds to reduce the output shortfall present in the considered DMU. The output shortfall should be an integer vector less than $\bar{Y}_j$. So the convergence of the algorithm is very rapid. In most of the problems the algorithm terminates in the first iteration itself. The objective function value at each iteration gives a Linear Diophantine equation. By the reduction of the input excesses will reduce the value of the equation for the next iteration. The maximum value of the Linear Diophantine equation is $S^{-*} + S^{+*}$ and in the final iteration the value is zero. From the Diophantine equations we can see the convergence of the algorithm.

5. Illustration Using Examples

For illustration consider the situation of the evaluation of the relative efficiency of 12 DMUs in terms of two inputs and two outputs. Assume the input and output values can take only integer values. The data is given in table 1.

Table 1

	DMU	A	B	C	D	E	F	G	H	I	J	K	L
Input	Input 1	20	19	25	27	22	55	33	31	30	50	53	38
	Input 2	151	131	160	168	158	255	235	206	244	268	306	284
Output	Output 1	100	150	160	180	94	230	220	152	190	250	260	250
	Output 2	90	50	55	72	66	90	88	80	100	100	147	120

By using Lozano and Villa model we can see that DMU A is an integer efficient DMU. But if there is any increase in the input vector of this DMU may cause it to become inefficient.

Assume the vector has an increase $\delta\bar{X}_1 = (5,20)$, then the input vector will become $\bar{X}_1 = (25,171)$.

The MILP formulation of the problem using the proposed model can be defined as follows

Max $\quad f = (s_1^- + s_2^- + s_1^+ + s_2^+)$

Subject to

$$25\lambda_1 + 19\lambda_2 + 25\lambda_3 + 27\lambda_4 + 22\lambda_5 + 55\lambda_6 + 33\lambda_7 + 31\lambda_8 + 30\lambda_9 + 50\lambda_{10} + 53\lambda_{11} + 38\lambda_{12} + s_1^- = 25$$

$$171\lambda_1 + 131\lambda_2 + 160\lambda_3 + 168\lambda_4 + 158\lambda_5 + 255\lambda_6 + 235\lambda_7 + 206\lambda_8 + 244\lambda_9 + 268\lambda_{10} + 306\lambda_{11} + 284\lambda_{12} + s_1^- = 171$$

$$100\lambda_1 + 150\lambda_2 + 160\lambda_3 + 180\lambda_4 + 94\lambda_5 + 230\lambda_6 + 220\lambda_7 + 152\lambda_8 + 190\lambda_9 + 250\lambda_{10} + 260\lambda_{11} + 250\lambda_{12} + s_1^- = 100$$

$$90\lambda_1 + 50\lambda_2 + 55\lambda_3 + 72\lambda_4 + 66\lambda_5 + 90\lambda_6 + 88\lambda_7 + 80\lambda_8 + 100\lambda_9 + 100\lambda_{10} + 147\lambda_{11} + 120\lambda_{12} + s_1^- = 50$$

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8, \lambda_9, \lambda_{10}, \lambda_{11}, \lambda_{12}, s_1^-, s_2^-, s_1^+, s_2^+ \geq 0$$

$s_1^-, s_2^-, s_1^+, s_2^+$ are integers.

But on solving this MILP we can see that

$s_1^- = 0, s_2^- = 0, s_1^+ = 0, s_2^+ = 0$. Hence the DMU is integer efficient.

Now assume that $\delta \bar{X}_1 = (30,50)$, then the input vector will become $\bar{X}_1 = (50,201)$. Corresponding to this input the MILP becomes

Max $\quad f = (s_1^- + s_2^- + s_1^+ + s_2^+)$

Subject to

$$25\lambda_1 + 19\lambda_2 + 25\lambda_3 + 27\lambda_4 + 22\lambda_5 + 55\lambda_6 + 33\lambda_7 + 31\lambda_8 + 30\lambda_9 + 50\lambda_{10} + 53\lambda_{11} + 38\lambda_{12} + s_1^- = 25$$

$$171\lambda_1 + 131\lambda_2 + 160\lambda_3 + 168\lambda_4 + 158\lambda_5 + 255\lambda_6 + 235\lambda_7 + 206\lambda_8 + 244\lambda_9 + 268\lambda_{10} + 306\lambda_{11} + 284\lambda_{12} + s_1^- = 171$$

$$100\lambda_1 + 150\lambda_2 + 160\lambda_3 + 180\lambda_4 + 94\lambda_5 + 230\lambda_6 + 220\lambda_7 + 152\lambda_8 + 190\lambda_9 + 250\lambda_{10} + 260\lambda_{11} + 250\lambda_{12} + s_1^- = 100$$

$$90\lambda_1 + 50\lambda_2 + 55\lambda_3 + 72\lambda_4 + 66\lambda_5 + 90\lambda_6 + 88\lambda_7 + 80\lambda_8 + 100\lambda_9 + 100\lambda_{10} + 147\lambda_{11} + 120\lambda_{12} + s_1^- = 50$$

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8, \lambda_9, \lambda_{10}, \lambda_{11}, \lambda_{12}, s_1^-, s_2^-, s_1^+, s_2^+ \geq 0$$

$s_1^-, s_2^-, s_1^+, s_2^+$ are integers.

The model estimates the output $\bar{Y}_1 = (198,90)$. But for this output only $\bar{X}_1 = (34,201)$ is required. Hence by our algorithm we suggest input-output values for the DMU as $\bar{X}_1 = (50,201)$ and $\bar{Y}_1 = (198,90)$.

6. Conclusion

In this paper, an iterative MILP algorithm is proposed to set integer output values for an efficient DMU whose input value/values are increased. The algorithm can be used to estimate the output value for inefficient DMUs also. The example considered to explain the algorithm is has 12 DMUs with two inputs and two outputs. In the considered example the solution can estimate with two iterations. So the algorithm converges rapidly.

References

[1]. Amin, G. R. (2009). Comments on Finding the Most Efficient DMUs in DEA: An Improved Integrated Model. Computers and Industrial Engineering; 56(4); 1701-1702.

[2]. Banker, R. J., Charnes, A. and Cooper, W. W. (1984). Some Models for Estimating Technical and Scale inefficiencies in Data Envelopment Analysis. Management Science 30(9), 1078-1092.

[3]. Charnes, A., Cooper, W. W., Lewin, A. Y. and Seiford, L. M. (1994). Data Envelopment Analysis- Theory, Methodology and Applications. Kluwer Academic Publishers, Boston.

[4]. Charnes, A., Cooper, W. W. and Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units. European Journal of Operational Research; 2(6), 429-444.

[5]. Cook, W. D. and Seiford, L. M. (2009). Data Envelopment Analysis (DEA)-Thirty Years on. European Journal of Operational Research; 192 1-17.

[6]. Farell, M. (1957). The measurement of productive efficiency. Journal of the Royal Statistical Society, Series A(General); 120 253-281.

[7]. Jahanshahloo, G. R., Lotfi, F. H., Shoja, N., Tohidi, G. and Razavyan, S. (2004). Input estimation and identification of extra inputs in inverse DEA models. Applied Mathematics and Computation; 156 427-437.

[8]. Khezrimotlagh, D., Salleh, S. and Mohsenpour, Z. (2012). A comment on theory of Integer-valued Data Envelopment Analysis. Applied Mathematical Sciences; 6(116); 5769-5774.

[9]. Khezrimotlagh, D., Salleh, S. and Mohsenpour, Z. (2013). A note on Integer-valued Radial Model in DEA. Computers & Industrial Engineering; 66; 199-200 .

[10]. Kuosmanen, T. and Kazemi Matin, R. (2009). Theory of integer Integer –Valued Data Envelopment Analysis. European Journal of Operational Research; 192; 658-667.

[11]. Lozano, S. and Villa, G. (2006). Data Envelopment Analysis of Integer –valued inputs and outputs. Computers and Operations Research, 33; 3004-30014.

[12]. Lozano, S. and Villa, G. (2007). Integer DEA Models. In Modelling Data Irregularities and Structural Complexities in Data Envelopment Analysis. pp 271-289. Springer , New York.

[13]. Wei, Q., Zhang, J. and Zhang, X. (2000). An inverse DEA for input output estimate. . European Journal of Operational Research; 121 151-163.

[14]. Wu, J. and Zhou, Z. (2011). A Mixed-objective Integer DEA Model. Annals of Operations Research; 1-15.