Computers & Industrial Engineering 82 (2015) 78–81

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Negative data in data envelopment analysis: Efficiency analysis and estimating returns to scale ☆

Maryam Allahyar, Mohsen Rostamy-Malkhalifeh *

Department of Mathematics, Science and Research Branches, Islamic Azad University, Tehran, Iran

ARTICLE INFO

Article history:
Received 22 May 2014
Received in revised form 2 December 2014
Accepted 21 January 2015
Available online 30 January 2015

Keywords:
Data envelopment analysis
Negative data
Efficiency
Returns to scale

ABSTRACT

The classic Data Envelopment Analysis (DEA) models developed with the assumption that all inputs and outputs are non-negative, whereas, we may face a case with negative data in the actual business world. So, the need to adapt the DEA models so that they are applicable to cases includes inputs and outputs which can take both negative and non-negative values has been an issue. It can be readily demonstrated that the assumption of constant returns to scale (CRS) is not possible in technologies under negative data. So, one of the interesting and challenge questions is how to determine the state of RTS in the presence of negative data under variable returns to scale (VRS) technology. Accordingly, in this contribution, we first address the efficiency measure and then suggest a method to discover the state of returns to scale (RTS) in the presence of negative input and output values which has not been discussed much enough so far in DEA literature. Finally, the main results are elaborated by some illustrative examples.

1. Introduction

Data Envelopment Analysis (DEA) is a non-parametric technique for the performance assessment of a set of Decision Making Units (DMUs), each of which consumes multiple inputs to produce multiple outputs. The CCR (Charnes, Cooper, & Rhodes, 1978) and BCC (Banker, Charnes, & Cooper, 1984) models are two basic radial models under constant returns to scale (RTS) and variable RTS, respectively. Actually, the classic DEA models developed with the assumption that all inputs and outputs are non-negative, whereas, this assumption is very restrictive and is not always hold in the real world. For example, the net profit as one of the components of output vector can take a negative value at some DMUs. So, the need to handle negative data and adapt the DEA models has been an attractive issue in DEA literature.

Among the commonly suggested methods facing with negative data, two points of view are highly desirable. One is that the negative inputs are treated as positive outputs and the negative outputs are treated as positive inputs (Scheel, 2001; Zhu, 2009). The other is applying DEA models with translation invariance to make all negative data positive by adding a big enough scalar to the negative variables (Charnes, Cooper, Seiford, & Stutz, 1983; Ali & Seiford, 1990; Lovell & Pastor, 1995; Seiford & Zhu, 2002). The additive model (Charnes, Cooper, Golany, Seiford, & Stutz, 1985) is translation invariance and the BCC model is input – translation invariance in output oriented and output – translation invariance in input oriented. However, these two models have some drawbacks. For instance, the additive model is not able to give the efficiency score of inefficient units and the efficiency score obtained through the BCC model is not translation invariance. Owing to these weaknesses, several methods have been presented which can handle negative variables without translation of the original data, e. g. RDM (Portela, Thanassoulis, & Simpson, 2004), MSBM (Sharp, Meng, & Liu, 2007), SORM (Emrouznejad, Anouze, & Thanassoulis, 2010) and VRM (Cheng, Zervopoulos, & Qian (2013)). In a recent paper, Kazemi Matin, Amin, and Emrouznejad (2014) surveyed the target setting with SORM models and presented a modified SORM approach.

Indeed, determining the RTS property (Constant RTS, Increasing RTS and Decreasing RTS) provides important information by which the manager is able to assess the optimal size of each individual DMU by resizing the scale of their operations. It is well known that RTS is a local phenomenon for each DMU. Focusing on this fact, the right and left RTS notion was introduced (Golany & Yu, 1997). In a related study, Allahyar and Rostamy-Malkhalifeh (2014) overcome the shortcoming of Golany and Yu's method and proposed two LP models for estimating the right and left RTS.

Reviewing the existing various attempts for dealing with negative data, we find that study on the estimation of RTS has not received attention much enough. Hence, the interesting and challenge question is how to determine the state of RTS in the presence

☆ This manuscript was processed by Area Editor Imed Kacem.

* Corresponding author. Tel.: +98 2144865030.
E-mail addresses: Allahyar_Maryam@yahoo.com (M. Allahyar), Mohsen_Rostamy@yahoo.com (M. Rostamy-Malkhalifeh).

http://dx.doi.org/10.1016/j.cie.2015.01.022
0360-8352/

of negative data. Accordingly, the main contribution of this current paper is to address the efficiency measure and also discover the state of RTS corresponding to each efficient DMU when all input and output values are allowed to take both negative and non-negative values.

The paper is organized in five sections. Section 2 introduces a new method for measuring the efficiency under negative data. In Section 3, we develop two LP modes to discover the state of RTS in the presence of negative data. To validate our suggested methods, Section 4 illustrates two examples. Finally, Section 5 gives conclusions.

2. The relative efficiency measure of negative data

As noted earlier, some models have been suggested for measuring the efficiency under negative data in DEA literature. However, most of them are in input- (output-) oriented models. In this section, inspired by the generic directional distance model (Chambers, Chung, & Fare, 1998), our aim is to introduce a non-oriented model which permits the inputs and outputs to take both positive and negative values and yields an efficiency measure.

Consider a set of n DMUs, where the input and output vector of each DMU$_j$ ($j = 1, \ldots, n$) is $X_j = (x_{1j}, \ldots, x_{mj})$ and $Y_j = (y_{1j}, \ldots, y_{sj})$ respectively. Also, we assume that the data set is allowed to be both positive and negative. Let the following sets for DMU$_j$ ($j = 1, \ldots, n$):

$$I_j^+ = \{i \in \{1, \ldots, m\} | x_{ij} \geqslant 0\}$$
$$I_j^- = \{i \in \{1, \ldots, m\} | x_{ij} < 0\}$$
$$O_j^+ = \{r \in \{1, \ldots, s\} | y_{rj} \geqslant 0\}$$
$$O_j^- = \{r \in \{1, \ldots, s\} | y_{rj} < 0\}$$

To assess DMU$_k$ ($k \in \{1, \ldots, n\}$) we consider the direction vector $(-|X_k|, |Y_k|) = (-|x_{1k}|, \ldots, -|x_{mk}|, |y_{1k}|, \ldots, |y_{sk}|)$ by using the absolute values of data and formulate the following model under variable returns to scale (VRS).

$$\begin{aligned} \theta^* = Max \quad & \theta \\ s.t. \quad & \sum_{j=1}^{n} \lambda_j X_j \leqslant X_k - \theta |X_k| \\ & \sum\nolimits_{j=1}^{n} \lambda_j Y_j \geqslant Y_k + \theta |Y_k| \\ & \sum\nolimits_{j=1}^{n} \lambda_j = 1 \\ & \lambda_j \geqslant 0 \ j = 1, \ldots, n \end{aligned} \tag{1}$$

Model (1) yields the efficiency rating θ^* corresponding each DMU. It is evident that $\theta^* \geqslant 0$. To have an efficiency score between 0 and 1 for inefficient units, we consider $E_k = 1 - \frac{\theta^*}{\bar{\theta}_k}$ ($k = 1, \ldots, n$) where $\bar{\theta}_k$ ($k = 1, \ldots, n$) is defined as:

$$\bar{\theta}_k = Min\left\{Min_i\left\{\frac{x_{ik} - x_{il}}{|x_{ik}|}, x_{ik} \neq 0\right\}, Min_r\left\{\frac{y_{rl} - y_{rk}}{|y_{rk}|}, y_{rk} \neq 0\right\}, \right.$$
$$\left. i = 1, \ldots, m, r = 1, \ldots, s\right\} \tag{2}$$

In fact, by defining $I = (Min_j\{x_{ij}, i = 1, \ldots, m\}, Max_j\{y_{rj}, r = 1, \ldots, s\}) = (X_I, Y_I)$ and $AI = (Max_j\{x_{ij}, i = 1, \ldots, m\}, Min_j\{y_{rj}, r = 1, \ldots, s\}) = (X_{AI}, Y_{AI})$ as the ideal and anti-ideal point respectively, we have

$$\begin{aligned} & x_{il} \leqslant x_{ik} - \theta |x_{ik}| \leqslant x_{iAI} \qquad i = 1, \ \ldots, \ m \\ & y_{rAI} \leqslant y_{rk} + \theta |y_{rk}| \leqslant y_{rl} \qquad r = 1, \ \ldots, \ s \end{aligned} \tag{3}$$

So $\bar{\theta}_k$ is obtained by considering the following expression:

$$\begin{aligned} & \frac{x_{ik} - x_{iAI}}{|x_{ik}|} \leqslant \theta \leqslant \frac{x_{ik} - x_{il}}{|x_{ik}|} \qquad i = 1, \ \ldots, \ m \\ & \frac{y_{rAI} + y_{rk}}{|y_{rk}|} \leqslant \theta \leqslant \frac{y_{rl} - y_{rk}}{|y_{rk}|} \qquad r = 1, \ \ldots, \ s. \end{aligned} \tag{4}$$

Expression (4) and $\theta^* \geqslant 0$ together imply that $0 \leqslant \theta^* \leqslant \bar{\theta}$. If $\bar{\theta} = 0$ then $\theta^* = 0$ otherwise $0 \leqslant \frac{\theta^*}{\bar{\theta}} \leqslant 1$. DMU$_k$ is an efficient DMU if $E_k = 1$.

3. Estimating the right and left RTS in the presence of negative data

When DMUs are allowed to take both negative and non-negative input and output, it is easy to verify that the assumption of CRS is not possible in Production technology. Hence, the interesting and challenge question is how to determine the RTS status for such units under VRS technology.

It is well known that RTS is a local phenomenon for each efficient DMU. So, in order to detect RTS more precisely, it is fair to be examined locally. So, the focus of this section is on estimating RTS to the right and left side of the efficient DMUs.

Consider an efficient unit identified by Model (1), say, DMU$_k$ whose coordinate is (X_k, Y_k).

In our proposed approach, we formulate and solve for each efficient unit the two following models to estimate the right and left RTS respectively, in the presence of negative data:

$$\begin{aligned} \beta^* = Max \quad & \beta \\ s.t. \quad & \sum_{j=1}^{n} \lambda_j x_j \leqslant x_k + \delta |x_k| \\ & \sum\nolimits_{j=1}^{n} \lambda_j y_j \geqslant y_k + \beta |y_k| \\ & \sum\nolimits_{j=1}^{n} \lambda_j = 1 \\ & \lambda_j \geqslant 0 \ j = 1, \ \ldots, \ n \end{aligned} \tag{5}$$

and

$$\begin{aligned} \alpha^* = Max \quad & \alpha \\ s.t. \quad & \sum_{j=1}^{n} \lambda_j x_j \leqslant x_k - \alpha |x_k| \\ & \sum\nolimits_{j=1}^{n} \lambda_j y_j \geqslant y_k - \gamma |y_k| \\ & \sum\nolimits_{j=1}^{n} \lambda_j = 1 \\ & \lambda_j \geqslant 0 \ j = 1, \ \ldots, \ n \end{aligned} \tag{6}$$

where the parameters δ and γ assume a positive small arbitrary value. To be more precise, by allocating the positive small enough value to parameters and solving Models (5) and (6), we reach the points $(\bar{X}, \bar{Y})$ and $(\tilde{X}, \tilde{Y})$ in the immediate neighborhood to the right and left of DMU$_k$ along the efficient frontier respectively; and then, we will be able to determine the right and left RTS in the presence of negative data, comparing the performance of these points to that of the DMU$_k$.

The projection points $(\bar{X}, \bar{Y})$ and $(\tilde{X}, \tilde{Y})$ are defined as follows:

$$\bar{x}_i = \begin{cases} (1+\delta)x_{ik} & i \in I_k^+ \\ (1-\delta)x_{ik} & i \in I_k^- \end{cases}, \quad \bar{y}_r = \begin{cases} (1+\beta^*)y_{rk} & r \in O_k^+ \\ (1-\beta^*)y_{rk} & r \in O_k^- \end{cases}$$
$$\tilde{x}_i = \begin{cases} (1-\alpha^*)x_{ik} & i \in I_k^+ \\ (1+\alpha^*)x_{ik} & i \in I_k^- \end{cases}, \quad \tilde{y}_r = \begin{cases} (1-\gamma)y_{rk} & r \in O_k^+ \\ (1+\gamma)y_{rk} & r \in O_k^- \end{cases}$$