Some DEA based methods for measuring the efficiency of decision making units with negative and positive input output

Reenu Kumari
Maharaja Surajmal Institute of Technology,Delhi

Abstract

The aim of this paper is to provide an insight of the existing methods in literature based on DEA (Data Envelopment Analysis) which can handle the positive as well as negative data (input,output).DEA methods are applied for the purpose of measuring the efficiency of a set of homogeneous Decision making units (DMUs).The main problem of the DEA methods was that all the values of input and output should be strictly positive,but this is not possible for all the practical and real life problems.We find many real life problems in which input and output can take positive as well as negative values.There have been many work focusing on this problem and different methods have been introduced to overcome this weakness of the DEA models. This study review the existing techniques and the effective methods which can be used in the presence of negative variables also.A well known numerical example is solved by different methods and the obtained results are compared.The discussion and the findings of this paper can be used as a guideline to managerial and analyst which method can be applied to the real life problems where both (positive as well as negative) type of variables appears.

1 Introduction

Data envelopment analysis (DEA) introduced by Charnes et al. (1978) is a linear programming technique to evaluate the relative efficiency for each member of a set of homogeneous decision making units (DMUs) responsible for producing multiple output using multiple inputs.It has been widely and effectively used to measure the efficiency in many areas like government organizations,hospitals,schools,universities etc. The main weakness of the traditional DEA models is that it cannot deal with negative input or output values.Many researcher worked on this problem,However there is no standard method for handling negative data and this issue is still a hot research problem in the recent times. The simplest solution of the problem of negative data was to interchange the roles of input and output.Scheel (2001) consider the all negative input as positive output to increased the absolute value to obtain the decrements in negative input.And vice versa they treated all negative output as positive input to obtain an increment in negative output.The drawback of this method was that this method can not applied in the case where the data set can have both positive and negative values.Seiford and Zhu (2002) Pastor (1996) used translation in-variance property of DEA models to handle negative variables in DEA .Under variable return to scale the Additive models are translation invariant and used in case of negative data.One of the drawback of additive models is that they do not provide any information about an inefficient DMU.

Portela et al. (2004) introduced a modified form of the generic directional distance model (referChambers et al. (1998)). They introduced two form of RDM (Range Directional Model),one form called RDM$^+$ and another is RDM$^-$.The efficiency score obtained by RDM models similar to those obtained by radial models.Sharp et al. (2007) introduced MSBM (modified slack based model) which can deal positive as well as negative data set.Emrouznejad et al. (2010) introduced semi-oriented radial measure (SORM) to deal with negative variables in dataset. In SORM each variables is breaks down into two variable, one assigned the negative values and the other the positive values of the original variable. Cheng et al. (2013) propose a variant of radial measure (VRM) which yields a measure of efficiency and also is able to handle variables consisting of positive values for some and negative values for other sample DMUs.
This paper unfolds as follows.Next section includes a brief introduction to the existing models used to handle the negative data in DEA.A numerical example has been solved by using different methods discussed above in Section 2.Finally the conclusion is presented in last section.

2 Some effective models to handle negative variables in DEA

In this section different DEA based models are discussed which can be used in case of positive as well as negative input output data.The advantage and disadvantage of the methods also discussed in this section. Consider a set of n homogeneous DMUs which use m inputs x_{ij} to produce s outputs y_{rj}

2.1 Range Directional Measure

In order to deal with the negative data in the conventional DEA model, Portela et al. (2004) proposed a Range Directional Model (RDM) based on directional distance model Chambers et al. (1998) as follows:

$$\theta^{RDM} = \max \theta$$

$$s.t. \begin{cases} \sum_{j=1}^{n} \lambda_j \cdot y_{kj} \geq y_{r0} + \theta \cdot e_{k0}, & k = 1,2,\ldots,s, \\ \sum_{j=1}^{n} \lambda_j \cdot x_{ij} \leq x_{i0} - \theta \cdot d_{i0}, & i = 1,2,\ldots,m, \\ \sum_{j=1}^{n} \lambda_j = 1 \ \ \lambda_j \geq 0, & j = 1,2,\ldots,n, \end{cases} \quad (1)$$

where the directional vectors are computed as

$$e_{k0} = \max_{j=1,2,\ldots,n} y_{kj} - y_{k0}, \qquad k = 1,2,\ldots,s, \quad (2)$$

$$d_{i0} = x_{i0} - \min_{j=1,2,\ldots,n} x_{ij}, \qquad i = 1,2,\ldots,m. \quad (3)$$

The zero direction can be identified before solving the DEA model. If at least one zero direction appears, the target unit is classified as efficient because no

improvement in inputs and outputs is possible simultaneously.
The optimal score θ^{RDM} measures the inefficiency, thus the lower value of optimal score or the zero value is preferred corresponding to efficiency. Hence, either $\widetilde{\theta}^{RDM} = \left(1 - \theta^{RDM}\right)$ or $\widetilde{\theta}^{RDM} = \left(\frac{1-\theta^{RDM}}{1+\theta^{RDM}}\right)$ represents the efficiency of the target unit. The main advantage of RDM method is that RDM models provides us the efficiency measure similiar to those obtained by radial models and the efficiency score lies between 0 and 1. The drawback of the RDM model that it may yield unbounded solution when the DMU under evaluation has the maximum values for every output variable or the minimum levels for all the inputs.

2.2 Modified slack based measure

The SBM model given by Tone (2001) was an effective method to deal with input excess and output shortfall.By generalizing the SBM model Sharp et al. (2007) proposed a modified slack based measure (MSBM) which handle 'natural' negative inputs and outputs and given as

$$\theta^{MSBM} = \min \frac{1 + \sum_{i=1}^{m} w_i s_i^- / d_{i0}}{1 - \sum_{k=1}^{s} \nu_k s_k^+ / e_{k0}}$$

$$s.t. \begin{cases} \sum_{j=1}^{n} \lambda_j \cdot y_{kj} - s_k^+ = y_{k0} & k = 1, 2, \ldots, s \\ \sum_{j=1}^{n} \lambda_j \cdot x_{ij} + s_i^- = x_{i0} & i = 1, 2, \ldots, m \\ \sum_{j=1}^{n} \lambda_j = 1, \ \sum_{i=1}^{m} w_i = 1, \ \sum_{k=1}^{s} \nu_k = 1 \\ \lambda_j, w_i, \nu_k, s_k^+, s_i^- \geq 0, & \forall\, j = 1, 2, \ldots, n; i = 1, 2, \ldots, m; \ k = 1, 2, \ldots, s, \end{cases} \tag{4}$$

where the directional vectors d_{k0} and e_{k0} have the same meaning as defined in previous model in (RDM). The optimal value measures efficiency, so larger values are preferred to lower ones. The MSBM model overcomes the lack of translation invariance in the slacks-based measure model by drawing on the ideas from the range directional model (RDM). The MSBM model takes into account individual input and output slacks, which provides more precise evaluation of inefficient decision-making units (DMUs). The MSBM model have limited application as compared to RDM and SORM because it was devised for naturally negative inputs.

2.3 Semi Oriented Radial Measure Model

Emrouznejad et al. proposed a model to deal with negative data. They used a partitioning approach in modeling negative data, and proposed a semi oriented radial measure (SORM) model for performance evaluation of the observed production units. Let

$$I = \{i \in \{1, 2, \ldots, m\} : x_{ij} \geq 0, \forall\, j = 1, 2, \ldots, n\}$$

$$L = \{l \in \{1, 2, \ldots, m\} : \exists\, j \in \{1, 2, \ldots, n\} \text{ for which } x_{lj} < 0\}$$

$$R = \{k \in \{1,2,\ldots,s\} : y_{kj} \geq 0, \forall\, j = 1,2,\ldots,n\}$$

and

$$K = \{p \in \{1,2,\ldots,s\} : \exists\, j \in \{1,2,\ldots,n\} \text{for which } y_{pj} < 0\}.$$

Let $x_{lj} = x^1_{lj} - x^2_{lj}$ for each $l \in L$, where $x^1_{lj}, x^2_{lj} \geq 0\ \forall j = 1,2,\ldots,n$ and

$$x^1_{lj} = \begin{cases} x_{lj}, & \text{if } \ x_{lj} \geq 0 \\ 0, & \text{if } \ x_{lj} < 0 \end{cases}$$

and

$$x^2_{lj} = \begin{cases} 0, & \text{if } \ x_{lj} \geq 0 \\ -x_{lj}, & \text{if } \ x_{lj} < 0. \end{cases}$$

Also, $y_{pj} = y^1_{pj} - y^2_{pj}$ for each $p \in K$, where $y^1_{pj}, y^2_{pj} \geq 0\ \ \forall j = 1,2,\ldots,n$ and

$$y^1_{pj} = \begin{cases} y_{pj}, & \text{if } \ y_{pj} \geq 0 \\ 0, & \text{if } \ y_{pj} < 0 \end{cases}$$

and

$$y^2_{pj} = \begin{cases} 0, & \text{if } \ y_{pj} \geq 0 \\ -y_{pj} & \text{if } \ y_{pj} < 0. \end{cases}$$

To assess DMUs, Emrouznejad et al. proposed the following input oriented variable return to scale SORM model as follows.

$$\theta^{SORM} = \min \theta$$

$$s.t. \begin{cases} \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{i0}, & \forall i \in I, \\ \sum_{j=1}^n \lambda_j x^1_{lj} \leq \theta x^1_{l0}, & \forall l \in L, \\ \sum_{j=1}^n \lambda_j x^2_{lj} \geq \theta x^2_{l0}, & \forall l \in L, \\ \sum_{j=1}^n \lambda_j Y_{kj} \geq h Y_{r0}, & \forall k \in R, \\ \sum_{j=1}^n \lambda_j y^1_{pj} \geq y^1_{p0}, & \forall p \in K, \\ \sum_{j=1}^n \lambda_j y^2_{pj} \leq y^2_{p0}, & \forall p \in K, \\ \sum_{j=1}^n \lambda_j = 1 \ \ \lambda_j \geq 0, & j = 1,2,\ldots,n, \end{cases} \tag{5}$$

The optimal solution of the model, θ^{SORM} represents the efficiency of DMU_0 in the presence of negative data. The optimal value $\widetilde{\theta}^{SORM} = \theta^{SORM}$ measures the efficiency, so larger values are preferred to lower ones. The disadvantage is that the increase in dimensionality of the problem, consequent on treating negative parts of a variable as a distinct variable, means that part of the original production possibility set is deleted and the method may not necessarily determine Pareto efficient targets. However, the method generally cannot lead to targets that are worse than the observed input–output levels of the unit

2.4 Variant of radial measure

Cheng et al. (2013) introduced The input oriented VRM model under variable return to scale is follows

$\theta^{VRM} = \max\beta$ subject to

$$\begin{aligned} &\sum_{j=1}^{n} \lambda_j \cdot x_{ij}^{-} + \beta \left|x_{i0}^{-}\right| \leq x_{i0}^{-}, \qquad i = 1,2,\ldots,m, \\ &\sum_{j=1}^{n} \lambda_j \cdot y_{kj}^{-} \geq y_{k0}^{-}, \qquad k = 1,2,\ldots,s, \\ &\sum_{j=1}^{n} \lambda_j = 1,\ \lambda_j \quad j = 1,2,\ldots,n,. \end{aligned} \tag{6}$$

The output oriented VRM model under variable return to scale is as follows

$\theta^{VRM} = \max\beta$ subject to

$$\begin{aligned} &\sum_{j=1}^{n} \lambda_j \cdot y_{kj}^{-} - \beta \left|y_{k0}^{-}\right| \geq y_{k0}^{-}, \qquad i = 1,2,\ldots,m, \\ &\sum_{j=1}^{n} \lambda_j \cdot x_{ij}^{-} \leq x_{i0}^{-}, \qquad k = 1,2,\ldots,s, \\ &\sum_{j=1}^{n} \lambda_j = 1,\ \lambda_j \quad j = 1,2,\ldots,n,. \end{aligned} \tag{7}$$

The VRM model can violate the monotonicity rule in case a variable is with mixed sign (positive and negative both).The VRM model always yields improved targets for the inefficient DMUs in cases where positive and negative input (output) data appear in a particular variable, either by input shrinkage or by output expansion.

2.5 Base point slack based measure

Tone et al. (2020) proposed a Base Point Slack Based Measure (BP-SBM) which can be used in case of negative data by modifying the base point which was origin originally. In this model the modified input output x_{ij}^{-} and y_{kj}^{-} is used in place of original input output x_{ij} and y_{kj}. The BPSBM model is as follows:

$$\theta^{BPSBM} = \min \ \frac{1 - \frac{1}{m}\sum_{i=1}^{m} s_i^-/x_{i0}^-}{1 + \frac{1}{s}\sum_{k=1}^{s} s_k^+/y_{k0}^-}$$

subject to

$$\begin{aligned} &\sum_{j=1}^{n} \lambda_j \cdot y_{kj}^- - s_k^+ \geq y_{k0}^-, \qquad k = 1, 2, \ldots, s, \\ &\sum_{j=1}^{n} \lambda_j \cdot x_{ij}^- + s_i^- \leq x_{i0}^-, \qquad i = 1, 2, \ldots, m, \\ &\sum_{j=1}^{n} \lambda_j = 1, \ \lambda_j, s_i^-, s_k^+ \geq 0, \quad j = 1, 2, \ldots, n,. \end{aligned} \tag{8}$$

where $\lambda_j, s_i^-, s_k^+ \geq 0 forall j, i, k$
the formula for setting x_{ij}^- and y_{kj}^- are given as

$$x_{ij}^- = x_{ij} - x_i^{min} > 0, y_{kj}^- = y_{kj} - y_k^{min} > 0 textrm forall(i, k, j)$$

1. To find x_i^{min}
let $\delta_i = min(x_{i1}, x_{i2}, \ldots x_{in}) \quad (i = 1, 2 \ldots n)$.To decide the value of x_i^{min} the following transformation is used

$$\begin{aligned} \text{if} \quad \delta_i > 0 \quad \text{then} \quad x_i^{min} = 0 \\ \text{if} \quad \delta_i = 0 \quad \text{then} \quad x_i^{min} = -\sigma_i \\ \text{if} \quad \delta_i < 0 \quad \text{then} \quad x_i^{min} = \delta_i(1 + \tau_i) \end{aligned} \tag{9}$$

where σ_i and τ_i are positive numbers.
2.similarly to find the value of y_k^{min}
$\omega_i = min(y_{k1}, y_{k2}, \ldots y_{kn}) \quad (k = 1, 2 \ldots s)$.To decide the value of y_k^{min} the following transformation is used

$$\begin{aligned} \text{if} \quad \omega_k > 0 \quad \text{then} \quad y_k^{min} = 0 \\ \text{if} \quad \omega_k = 0 \quad \text{then} \quad y_k^{min} = -\rho_k \\ \text{if} \quad \omega_k < 0 \quad \text{then} \quad y_k^{min} = \delta_k(1 + \gamma_k) \end{aligned} \tag{10}$$

where ρ_k and γ_k are positive numbers.
The optimal value measures efficiency, so larger values are preferred to lower ones. The BP-SBM method can also be applied to the different return to scale conditions (RTS) and also satisfy the units invariance property of DEA.

3 Numerical Example

This section shows a numerical example with data set from Sharp et al.(2006).In this examples there are 13 DMUs with positive as well as negative input output.All the methods applied to this example are used in output oriented form.Table

Table 1: Input-output data

DMU	Input 1	Input 2	output1	output 2	output3
1	1.03	-0.05	0.56	-0.09	-0.44
2	1.75	-0.17	0.74	-0.24	-0.31
3	1.44	-0.56	1.37	-0.35	-0.21
4	10.80	-0.22	5.61	-0.98	-3.79
5	1.30	-0.07	0.49	-1.08	-0.34
6	1.98	-0.10	1.61	-0.44	-0.34
7	0.97	-0.17	0.82	-0.08	-0.43
8	9.82	-2.32	5.61	-1.42	-1.94
9	1.59	0.00	0.52	0.00	-0.37
10	5.96	-0.15	2.14	-0.52	-0.18
11	1.29	-0.11	0.57	0.00	-0.24
12	2.38	-0.25	0.57	-0.67	-0.43
13	10.30	-0.16	9.56	-0.58	0.00

1 shows the input output data and table 2 shows the efficiency score and ranking obtained by different methods in output oriented measure .RDM model used in output oriented form by taking $\beta = 0$ for input-related constraints.From table 2 we can conclude that all methods agreed on the efficiency score for efficient DMU.We ALSO observe from table 2 that MSBM,VRM,SORM model efficiency scores of inefficient units are less than or equal to those generated by the RDM.The MSBM model efficiencies take into account individual input and output slacks, which are ignored in RDM. This generally leads to lower efficiencies for inefficient DMUs than with the RDM model.

4 conclusion

The traditional radial DEA models gives incorrect results in the presence of negative input or output data .In this paper we review the methods existing in the literature that can be effectively used in the presence of negative and positive input output data in order to increase the power of this analytical technique.We discuss Range directional model (RDM), Modified Slack based measure (MSBM),Semi-oriented radial model (SORM) and Variant of radial model (VRM).The paper also describes the advantages and disadvantages of each of these methods.Many of the DEA based method discussed in this paper can be modified and extended further to remove the disadvantages of these methods. The discussion and the findings of this paper can be used as a guideline to analysts and managers to decide the best method which can applied to the data set having both type positive as well as negative variables.

Table 2: Ranking Related Scores Assigned to portfolios by different Ranking Methods

DMUs	RDM (rank)	SORM (rank)	MSBM(rank)	VRM(rank)
1	0.97(3)	0.63(4)	0.88(3)	0.906(3)
2	0.91(5)	0.45(6)	0.74(5)	0.77(5)
3	1.00(1)	1.00(1)	1.00(1)	1.00(1)
4	0.50(8)	0.59(5)	0.56(9)	0.68(7)
5	0.92(4)	0.41(7)	0.70(7)	0.77(5)
6	0.97(3)	0.86(3)	0.78(4)	0.86(4)
7	1.00(1)	1.00(1)	1.00(1)	1.00(1)
8	1.00(1)	1.00(1)	1.00(1)	1.00(1)
9	0.99(2)	0.91(2)	0.89(2)	0.91(2)
10	0.65(7)	0.39(8)	0.72(6)	0.73(6)
11	1.00(1)	1.00(1)	1.00(1)	1.00(1)
12	0.81(6)	0.25(9)	0.68(8)	0.65(8)
13	1.00(1)	1.00(1)	1.00(1)	1.00(1)

References

Chambers, R. G., Chung, Y., and Färe, R. (1998). Profit, directional distance functions, and nerlovian efficiency. *Journal of optimization theory and applications*, 98(2):351–364.

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.

Cheng, G., Zervopoulos, P., and Qian, Z. (2013). A variant of radial measure capable of dealing with negative inputs and outputs in data envelopment analysis. *European Journal of Operational Research*, 225(1):100–105.

Emrouznejad, A., Anouze, A. L., and Thanassoulis, E. (2010). A semi-oriented radial measure for measuring the efficiency of decision making units with negative data, using dea. *European Journal of Operational Research*, 200(1):297–304.

Pastor, J. T. (1996). Translation invariance in data envelopment analysis: A generalization. *Annals of Operations Research*, 66(2):91–102.

Portela, M., Thanassoulis, E., and Simpson, G. (2004). A directional distance approach to deal with negative data in dea: An application to bank branches. *Journal of Operational Research Society*, 55(10):1111–1121.

Scheel, H. (2001). Undesirable outputs in efficiency valuations. *European journal of operational research*, 132(2):400–410.

Seiford, L. M. and Zhu, J. (2002). Modeling undesirable factors in efficiency evaluation. *European journal of operational research*, 142(1):16–20.

Sharp, J. A., Meng, W., and Liu, W. (2007). A modified slacks-based measure model for data envelopment analysis with 'natural'negative outputs and inputs. *Journal of the Operational Research Society*, 58(12):1672–1677.

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European journal of operational research*, 130(3):498–509.

Tone, K., Chang, T.-S., and Wu, C.-H. (2020). Handling negative data in slacks-based measure data envelopment analysis models. *European Journal of Operational Research*, 282(3):926–935.