European Journal of Operational Research 245 (2015) 779–788

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Decision Support

On the identification of the global reference set in data envelopment analysis

Mahmood Mehdiloozad[a,*], S. Morteza Mirdehghan[a,1], Biresh K. Sahoo[b,2], Israfil Roshdi[c,d,3]

[a] Department of Mathematics, College of Sciences, Shiraz University, Shiraz 71454, Iran
[b] Xavier Institute of Management, Xavier University, Bhubaneswar 751 013, India
[c] Department of Accounting and Finance, The University of Auckland, Auckland, New Zealand
[d] Department of Mathematics, Semnan Branch, Islamic Azad University, Semnan, Iran

ARTICLE INFO

Article history:
Received 29 October 2013
Accepted 24 March 2015
Available online 31 March 2015

Keywords:
Data envelopment analysis
Linear programming
Global reference set
Minimum face
Returns to scale

ABSTRACT

It is well established that multiple reference sets may occur for a decision making unit (DMU) in the non-radial DEA (data envelopment analysis) setting. As our first contribution, we differentiate between three types of reference set. First, we introduce the notion of *unary reference set* (URS) corresponding to a given projection of an evaluated DMU. The URS includes efficient DMUs that are active in a specific convex combination producing the projection. Because of the occurrence of multiple URSs, we introduce the notion of *maximal reference set* (MRS) and define it as the union of all the URSs associated with the given projection. Since multiple projections may occur in non-radial DEA models, we further define the union of the MRSs associated with all the projections as unique *global reference set* (GRS) of the evaluated DMU. As the second contribution, we propose and substantiate a general linear programming (LP) based approach to identify the GRS. Since our approach makes the identification through the execution of a single primal-based LP model, it is computationally more efficient than the existing methods for its easy implementation in practical applications. Our last contribution is to measure returns to scale using a non-radial DEA model. This method effectively deals with the occurrence of multiple supporting hyperplanes arising either from multiplicity of projections or from non-full dimensionality of minimum face. Finally, an empirical analysis is conducted based on a real-life data set to demonstrate the ready applicability of our approach.

1. Introduction

Data envelopment analysis (DEA), introduced by Charnes, Cooper, and Rhodes (1978, 1979) based on the seminal work of Farrell (1957), is a linear programming (LP) based method for measuring the relative efficiency of a homogeneous group of decision making units (DMUs) with multiple inputs and multiple outputs. Based on observed data and a set of postulates, DEA defines a reference technology set relative to which a DMU can be rated as *efficient* or *inefficient*. For an inefficient DMU, DEA recognizes a unique or multiple projection(s) on the efficient frontier of the technology set. Associated with each projection, it also identifies a set of observed efficient DMUs against which the DMU under evaluation is directly compared. Those efficient DMUs are called *reference DMUs*, and the corresponding set is referred to as a *reference set*.

The identification of *all* the possible reference DMUs for an inefficient unit is an important and interesting problem in DEA, on which we concentrate in this contribution by means of the non-radial range-adjusted model (RAM) of Cooper, Park, and Pastor (1999). This issue has received significant attention in the literature due to its wide range of potential applications in ranking (Jahanshahloo, Junior, Hosseinzadeh Lotfi, & Akbarian, 2007), benchmarking and target setting (Bergendahl, 1998; Camanho & Dyson, 1999), and measuring returns to scale (RTS) (Cooper, Seiford, & Tone, 2007; Krivonozhko, Førsund, & Lychev, 2014; Sueyoshi & Sekitani, 2007a; 2007b; Tone, 1996, 2005; Tone & Sahoo, 2006).

From a managerial point of view, the identification of all the reference DMUs is specifically important for two reasons. First, to improve the performance of an inefficient DMU, it may not be logical in practice to introduce an unobserved (virtual) projection as a benchmark. In such a situation, however, the identification provides the possibility to derive practical guidelines from benchmarking against the reference DMUs. Second, when some (but not all) reference DMUs are identified for an evaluated unit, the decision maker may be of the

* Corresponding author. Tel.: +98 9127431689.
E-mail addresses: m.mehdiloozad@gmail.com, m.mehdiloozad@shirazu.ac.ir (M. Mehdiloozad), mirdehghan@shirazu.ac.ir (S.M. Mirdehghan), biresh@ximb.ac.in (B.K. Sahoo), i.roshdi@gmail.com (I. Roshdi).
[1] Tel.: +98 912 4430612; fax: +98 711 2281335.
[2] Tel.: +91 674 6647735; fax: +91 674 2300995.
[3] Tel.: +98 936 8889198.

http://dx.doi.org/10.1016/j.ejor.2015.03.029
0377-2217/

opinion that the identified DMUs are not appropriate benchmarks and may wish to have more options in choosing targets. In such a case, the identification allows him/her to incorporate the preference information into analysis so as to yield a projection with the most preferred (i) closeness (Tone, 2010), (ii) values of inputs and outputs, and (iii) shares of reference units in its formation.

The pioneer attempt to find all the reference DMUs in non-radial DEA models was made by Sueyoshi and Sekitani (2007b). Based on strong complementary slackness conditions (SCSCs) of linear programming, they proposed a primal–dual based method using the RAM model. The proposed method in their impressive study is very interesting as a theoretical idea. However, as Krivonozhko, Førsund, and Lychev (2012b) have argued, not only the computational burden of Sueyoshi and Sekitani's (2007b) approach is high, but it also seems that the basic matrices defined in their approach are likely to be ill-conditioned, leading to erroneous and unacceptable results even for medium-size problems. Furthermore, the economic interpretation of some constraints of their proposed model does not make sense. In a more recent and conscious attempt to overcome these difficulties, Krivonozhko et al. (2014) have proposed a primal–dual based procedure based on solving several LP problems. Using computational experiments, they showed that their proposed method works reliably and efficiently on real-life data sets and outperforms Sueyoshi and Sekitani's (2007b) approach.

It is worth noting that the studies conducted by Sueyoshi and Sekitani (2007a, 2007b) and Krivonozhko et al. (2014) correctly found all the observed DMUs on *minimum face* – a face of minimum dimension on which all the projections are located – as a *unique* reference set of a given DMU. On the other hand, both of these studies pointed out that the occurrence of *multiple* reference sets was possible. However, neither of them explicitly made a clear distinction between the uniquely-found reference set and other types of reference set for which multipleness may occur. This lack of discrimination creates an ambiguity about the uniqueness and, consequently, about the mathematical well-definedness of the definition of reference set.

Therefore, we were motivated to eliminate this ambiguity effectively. To do so, we have proposed three types of reference set sequentially, as our first contribution. Corresponding to a given projection, we first introduce the notion of *unary reference set* (URS) including efficient DMUs that are active in a specific convex combination producing this projection. Since multiple URSs (hereafter referred to as problem Type I) may occur, we introduce the notion of *maximal reference set* (MRS) and define it as the union of all the URSs associated with the given projection. Since multiple projections may occur in the RAM model, we further define the union of the MRSs associated with all the projections as unique *global reference set* (GRS) of the evaluated DMU. We have had an interesting finding: the convex hull of the GRS is equal to the minimum face. The benefits of the introduced three types of reference set (i.e., URS, MRS and GRS) are outlined below.

- The introduced concepts are all mathematically well-defined.
- The URS and MRS help demonstrate the occurrence of multiple reference sets associated with a single and multiple projection(s), respectively.
- While the multipleness may occur for the URS and MRS, the GRS presents a unique reference set that contains all the possible reference DMUs.

As our second contribution, we have proposed an LP model that identifies the GRS, and provides a projection in the relative interior of the minimum face. The proposed approach has several important features. First, it can effectively deal with the simultaneous occurrence of problems Types I and II. Second, this approach involves solving a single LP problem, which makes this approach computationally more efficient than the existing ones for its easy implementation in practical applications. Third, the computational efficiency of our approach is higher than that of the previous primal–dual ones, since it is developed based on the primal (envelopment) form that is computationally more efficient than the dual (multiplier) form (Cooper et al., 2007). Fourth, since our proposed LP problem contains several upper-bounded variables, its computational efficiency can be enhanced by using the simplex algorithm adopted for solving the LP problems with upper-bounded variables, which is much more efficient than the ordinary simplex algorithm (Winston, 2003).

Fifth, our proposed approach is more general in the sense that it can be readily used without any change in both the 'additive model' (Charnes, Cooper, Golany, Seiford, & Stutz, 1985) and the 'BAM model' (Cooper, Pastor, Borras, Aparicio, & Pastor, 2011; Pastor, 1994; Pastor & Ruiz, 2007), because the difference between each of these two models and the RAM model lies only in the weights assigned to the input and output slacks in the objective function. With some minor changes, it can also be used in the 'RAM/BCC model' (Aida, Cooper, Pastor, & Sueyoshi, 1998), the 'DSBM model' of Jahanshahloo, Hosseinzadeh Lotfi, Mehdiloozad, and Roshdi (2012) and the 'GMDDF model' of Mehdiloozad, Sahoo, and Roshdi (2014). Furthermore, it can be easily implemented in any radial DEA model like the 'BCC model' of Banker, Charnes, and Cooper (1984), but with some minor changes. Finally, our proposed approach is free from the restricting assumption that the input–output data must be non-negative, so it can effectively deal with negative data. This can be very beneficial from a practical point of view since in many applications negative inputs or outputs could appear. See Pastor and Ruiz (2007) for various examples of applications with negative data.

The third contribution of this study is to measure the RTS in the non-radial DEA setting. As it is known, the concept of RTS is meaningful only when the relevant DMU lies on the frontier of the technology set. Hence, for an inefficient DMU, an efficient projection must be considered. In this case, the type and magnitude of the RTS is determined through the position(s) of the hyperplane(s) supporting the technology set at the projection used. The supporting hyperplane(s) passes/pass through the MRS associated with this projection and can be mathematically characterized via this MRS. Therefore, problem Type II causes the occurrence of multiple supporting hyperplanes (hereafter referred to as problem Type III), which makes the measurement of RTS difficult. Such a difficulty can be properly dealt with by using a relative interior point of the minimum face for the measurement of RTS. This is because the supporting hyperplane(s) binding at this point is/are characterized through the GRS, but not through a specific MRS. Nonetheless, the uniqueness of the characterized supporting hyperplane(s) cannot yet be guaranteed because the minimum face may not be a 'Full Dimensional Efficient Facet' (Olesen & Petersen, 1996, 2003).

To sum up, the difficulty raised by problem Type III in the measurement of RTS originates either from problem Type II or from the non-full dimensionality of the minimum face. To deal with this difficulty, we have developed a two-stage procedure for the measurement of RTS by exploiting the intensive study of Krivonozhko et al. (2014). In the first stage, we cope with the difficulty arising from problem Type II by finding a relative interior point of the minimum face via the LP problem proposed to identify the GRS. Then, for the obtained point,[1] we use the indirect method of Banker, Cooper, Seiford, Thrall, and Zhu (2004) or the direct method of Førsund, Hjalmarsson, Krivonozhko, and Utkin (2007) to resolve the difficulty resulted from the non-full dimensionality of the minimum face. To demonstrate the ready applicability of our approach in empirical works, we have conducted an illustrative empirical analysis based on a real-life data set of 70 public schools in the United States.

The remainder of this paper unfolds as follows. Section 2 deals with the description of the technology followed by a brief review

[1] Note that this point does not influence the RTS, since all the relative interior points of the minimum face have the same RTS (Krivonozhko, Førsund, & Lychev, 2012c).