Computers & Industrial Engineering 66 (2013) 1163–1170

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Common weights data envelopment analysis with uncertain data: A robust optimization approach ☆

H. Omrani *

Department of Industrial Engineering, Urmia University of Technology, Urmia, Iran

A R T I C L E I N F O

Article history:
Received 12 March 2012
Received in revised form 1 April 2013
Accepted 28 July 2013
Available online 2 August 2013

Keywords:
Data envelopment analysis
Common weights
Robust optimization
Uncertainty

A B S T R A C T

One of the primary issues on data envelopment analysis (DEA) models is the reduction of weights flexibility. There are literally several studies to determine common weights in DEA but none of them considers uncertainty in data. This paper introduces a robust optimization approach to find common weights in DEA with uncertain data. The uncertainty is considered in both inputs and outputs and a suitable robust counterpart of DEA model is developed. The proposed robust DEA model is solved and the ideal solution is found for each decision making units (DMUs). Then, the common weights are found for all DMUs by utilizing the goal programming technique. To illustrate the performance of the proposed model, a numerical example is solved. Also, the proposed model of this paper is implemented by using some actual data from provincial gas companies in Iran.

1. Introduction

DEA was introduced by Charnes, Cooper, and Rhodes (1978) to estimate the efficiency scores of the DMUs. The original DEA model was named as DEA-CCR to honor the authors. DEA is a well known performance assessment model to calculate the relative efficiency of homogeneous DMUs that use the similar inputs and outputs. In DEA, to calculate the relative efficiency scores, a linear programming model is separately executed for each DMU. Therefore, each unit is free to set its weights to reach the efficient frontier. Several researchers have studied weights and proposed various models to restrict flexibility in input and output weights. The latter models are classified as weight restriction models (Dyson & Thanassoulis, 1988), assurance region models (Thompson, Langemeier, Lee, & Thrall, 1990), cone ratio models (Charnes, Cooper, Huang, & Sun, 1990) and common weights models (Kao & Hung, 2005). Common weights models were first introduced by Roll, Cook, and Golany (1991) and then were developed by Kao and Hung (2005), Kao (2010), and Zohrehbandian, Makui, and Alinezhad (2010). In common weights models, no flexibility is allowed and all DMUs are ranked by the same weights. The DEA model allows the DMUs to select the input–output weights to try to reach the efficiency frontier. This is the strength and weakness of the DEA model. If a DMU cannot be placed on frontier by selecting free weights for inputs and outputs, the inefficiency of DMU is very meaningful. Unlike, it is possible for the different DMUs to select very small weights (close to zero) for the inputs and outputs which will not be acceptable for the decision maker. In addition, various DMUs tend to give very different weights to similar inputs and outputs. To avoid these problems and to find out the common weights, Kao and Hung (2005) proposed a common weight DEA (CWDEA) model. The proposed DEA model was a nonlinear programming model which calculated the efficiency scores through using the common weights.

Trough a compromise solution approach, Makui, Alinezhad, Kiani Mavi, and Zohrehbandian (2008), and Zohrehbandian et al. (2010) improved Kao and Hung (2005) model and showed that the efficiency scores obtained from two approaches had high correlation. The latter models are linear and they calculate common weights by using the goal programming technique.

The previous common weight DEA models assume that there is no uncertainty in input–output data. In real world applications, it is almost impossible to have accurate data for the inputs and outputs (Sadjadi & Omrani, 2010). Therefore, to reach reliable results, it is necessary to consider the perturbation in data. There are many approaches for handling uncertain data in DEA model which can be classified as chance constraint DEA (CCDEA), imprecise DEA (IDEA), bootstrap DEA, fuzzy DEA and robust DEA (RDEA). The CCDEA was developed based on chance constraints approach and it assumes that the data have the stochastic nature and the probability distribution function of the data is known. In this approach, the DEA constraints are considered as chance constraints and finally, the DEA is transferred to a deterministic model and the latter model is solved by the quadratic programming technique. It is noteworthy that when we consider different probability distribution functions for data, we usually get different efficiency scores. However, it is difficult to find out a suitable distribution function for the data

☆ This manuscript was processed by Area Editor Imed Kacem, Pr.

* Tel.:+98 4453554181; fax: +98 4413554184.
E-mail address: h.omrani@uut.ac.ir

0360-8352/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cie.2013.07.023

(Cooper, Deng, Huang, & Li, 2002; Land, Lovell, & Thore, 1993; Olesen & Petersen, 1995). In IDEA approach, originally developed by Cooper, Park, and Yu (1999), different ranges and bounds are considered for input and output data and since the data are specified in the form of different ranges, the estimated efficiency scores are also interval. In addition, in this case, the DEA model is changed to a nonlinear programming where it may be difficult to find out the optimal solution. In the bootstrap DEA, the confidence interval for efficiency scores is constructed. In this approach, it is assumed that there is a potentially infinite population of units corresponding to a data generating process and the current set is simply a sample of the population. Therefore, the bootstrap technique is applied for re-sampling and re-producing the new units. The bootstrap DEA was extended by Simar and Wilson (1998), Simar and Wilson (2000). Another way of modeling uncertainty in the data is considering the input and the output values as fuzzy numbers. Sengupta (1992) was first who introduced fuzzy DEA with fuzzy objective function and constraints. Kao and Liu (2000) estimated the fuzzy efficiency scores by using the membership function of fuzzy degrees of efficiency. Guo and Tanaka (2001) developed a fuzzy DEA model with symmetrical triangular fuzzy numbers. They applied the α-cut technique on the constraints and calculated the efficiency scores. For more details about fuzzy DEA the readers can see Wen and Li (2009). In this paper, a robust optimization approach is proposed to incorporate the uncertainty associated with the data. Robust optimization is an alternative for the stochastic programming and sensitivity analysis. Robust optimization ensures that a planning constraint is violated through a very low probability given by the decision/design vector. Robust optimization was originally introduced by Soyster (1973) and later it was extended based on the studies of Ben-Tal and Nemirovski (1998), Ben-Tal and Nemirovski (1999, Ben-Tal and Nemirovski (2000), Bertsimas and Sim (2003), Bertsimas and Sim (2004, Bertsimas and Sim (2006), and Bertsimas, Pachamanova, and Sim (2004). This technique is used to model optimization problems with data uncertainty to obtain a solution that is guaranteed to be good for all or most possible realizations of the uncertain parameters. Ben-Tal and Nemirovski (1998), Ben-Tal and Nemirovski (1999, Ben-Tal and Nemirovski (2000), El-Ghaoui and Lebret (1997), and El-Ghaoui, Oustry, and Lebret (1998) have introduced a new idea for modeling the uncertainty in data based on ellipsoidal uncertainty sets. Also, Bertsimas and Sim (2004) proposed a robust optimization approach based on polyhedral uncertainty sets which preserves the class of problems under analysis. Sadjadi and Omrani (2008) developed a robust DEA model and suggested a new formulation of DEA which is suitable in the uncertainty environments. Also, they showed that the robust DEA (RDEA) based on Bertsimas and Sim (2004) is easier to solve compared to the robust DEA based on Ben-Tal and Nemirovski (2000) approach. In RDEA, the data are assumed to be uncertain and the probability distribution function of the data is unknown. In summary, if available data have no stochastic nature and the probability distribution function for data not to be clear, the RDEA model will be recommended.

As mentioned above, in DEA model the DMUs are free to select the input–output weights to attempt reaching the efficient frontier. However, in several cases, top manager and decision maker want to evaluate DMUs by using the common set of weights for the inputs and outputs. For example, in the banking industry, the general manager would like to measure the performance assessment of the branches by using the common weights. The objective of this paper is to seek a common set of weights to create the best efficiency score of the DMUs. Also, it is assumed that there is uncertainty in input and output data. Since the data do not have stochastic nature or it is difficult to find out a suitable probability distribution function, the robust optimization approach is suitable for dealing with uncertainty in data. If uncertainties are not considered in the data, then the DEA constraints will not be immune against violation and the ranks and efficiency scores will not be reliable, too (Wang & Wei, 2010). Unfortunately, there is an equality constraint in the multiplier DEA-CCR model that allows no uncertainty considered in inputs. Hence, this paper introduces an appropriate DEA model with inequality constraints which can consider the uncertainty in both inputs and outputs. Therefore, the final purpose of the paper is to find out the common set ofweights with considering uncertainty in data. This common set of weights is applied to evaluate the absolute efficiency of each efficient DMUs in order to rank them. Briefly, in this paper, first an appropriate RDEA model is developed and the efficiency score is calculated for each DMU. The recent scores are considered as the ideal solution. Then, the goal programming approach is used to minimize the amount of deviation from the ideal solution of the RDEA model.

The rest of this paper is organized as follows: First, the robust optimization technique is presented in Section 2. In Sections 3, the counterpart robust of common weight DEA model and the goal programming approach to find out the common weights are proposed. In Section 4, a numerical example based on data used by Kao and Hung (2005) is solved. Then, the common weights for the provincial gas companies in Iran are found in Section 5. Finally, the conclusion remarks of the paper are given in Section 6 to summarize the contribution of the paper.

2. Robust optimization

In this section, the robust optimization approach, an alternative for the stochastic programming and sensitivity analysis, will be explained. In robust optimization, the consideration is to ensure that a planning constraint is violated through a very low probability given by the decision/design vector. To present the robust modeling, consider the following LP problem:

$$\begin{aligned}&\text{minimize } c'x\\&\text{subject to:}\\&Ex = d\\&Ax \geqslant b\\&x \in X\end{aligned} \tag{1}$$

In the model (1), c' is the row vector of the corresponding costs, and x is the column vector of variables. The constraints are separated into two parts: equality and inequality constraints. E and A are the coefficient matrices of the equality and inequality constraints, respectively. Also, d and b are the column vectors of right-hand side values of the equality and inequality constraints, respectively. It is assumed that the coefficients in A are uncertain and X is a polyhedron. In robust optimization technique, consider a particular row $i(i = 1, \ldots, m)$ of the matrix A and let J_i represents the set of coefficients in the row i that are subject to uncertainty. Assume that the true values $\tilde{a}_{ij}(j = 1, \ldots, n)$ of uncertain data entries in ith inequality constraint are obtained from the nominal values a_{ij} of the entries by random perturbations:

$$\tilde{a}_{ij} = (1 + e_{ij}\zeta_{ij})a_{ij} = a_{ij} + \zeta_{ij}\hat{a}_{ij} \tag{2}$$

where $e_{ij} > 0$ is a given uncertainty level (percentage of perturbation) and $\hat{a}_{ij}$ measures the precision of the estimation. Also, ζ_{ij} ($\zeta_{ij} = 0$ for $j \notin J_j$) is the scaled deviation from nominal value and has an unknown but symmetric distribution which takes values in $[-1,1]$. Therefore, each entry $\tilde{a}_{ij}, j \in J_i$ is modeled as a symmetric and bounded random variable which takes values in $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$. Although the aggregated scaled deviation for constraint i could take any value between $-n_i$ and n_i, but it is limited to $\sum_{j=1}^{n}\zeta_{ij} \leqslant \Gamma_i, \forall i$. Therefore, for each constraint i, a parameter Γ_i, not necessarily integer, is introduced taking values in the interval $[0, n_i]$