Available online at www.sciencedirect.com
ScienceDirect
Computers & Operations Research 34 (2007) 2177–2190
computers & operations research
www.elsevier.com/locate/cor

Using lexicographic parametric programming for identifying efficient units in DEA ☆

Pekka J. Korhonen*, Pyry-Antti Siitari

Department of Economics and Management Science, Helsinki School of Economics P.O. Box 1210, 00101 Helsinki, Finland

Available online 3 October 2005

Abstract

In this paper, we propose the use of lexicographic parametric programming to recognize efficient units in data envelopment analysis (DEA). By using the parameterization of the right-hand side vector of the envelopment problem, we obtain the efficiency curve which is traversing through the efficient frontier from unit to unit. The units in the basis with any parameter value are efficient and the unit dominated by a point on an efficient facet is inefficient. Lexicographic parametric programming is needed to guarantee that all units to be considered are efficient—not only weakly efficient.

Keywords: Efficiency analysis; Data envelopment analysis; Lexicographic; Parametric programming; Computational aspects

1. Introduction

Charnes et al. [1] developed data envelopment analysis (DEA) for evaluating the relative efficiency of comparable units called decision making units (DMUs) essentially performing the same task using similar multiple inputs to produce similar multiple outputs. The units are assumed to operate under similar conditions. Based on information about existing data on the performance of the units and some preliminary assumptions, DEA forms an empirical efficient surface (frontier). If a DMU lies on the surface, it is referred to as an efficient unit, otherwise inefficient. DEA also provides efficiency scores and reference set for inefficient DMUs. The efficiency scores are used in practical applications as performance indicators of the DMUs. The reference set for inefficient units consists of efficient units and determines a virtual unit on the efficient surface. The virtual unit can be regarded as a target unit for the inefficient unit.

The target unit is found in DEA by projecting an inefficient DMU radially[1] to the efficient surface. To check the efficiency of a unit, and to find the reference set and the efficiency score for inefficient units requires the solving of an LP-model. The "standard" basic algorithm solves iteratively an LP-model for each unit separately. At each iteration, the right-hand side vector and one column (directional vector) in the coefficient matrix has to be updated. The optimal basis of the previous iteration is not valid for the next iteration as such. Actually, it is not necessary to solve an LP-model

☆ The research was supported by the Academy of Finland.
* Corresponding author. Tel.: +358 9 431 31; fax. +358 4313 8535.
E-mail address: pekka.korhonen@hkkk.fi (P.J. Korhonen).
[1] Term "radial" means that an efficient frontier is tried to reach by increasing the values of the current outputs or decreasing the values of the current inputs, or doing the both ones simultaneously in the same proportion.

0305-0548/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cor.2005.08.006

for each unit, because all units in an optimal basis for some unit are efficient, and those units are not needed for further investigation. The approach is usable in small problems,[2] but is computationally ineffective in large-scale problems.

Additional problems are caused by the only weakly efficient solutions. The original problem formulation [1] sometimes led to only weakly efficient solutions, but the authors recognized the problem and reformulated the model in [3] by using a so-called non-Archimedean infinitesimal in the model. There are two ways to deal with the infinitesimal in the model: (1) to replace it by a small number or (2) to use a lexicographic approach. When a small number is used, it is important that it is properly chosen. If it is too big, some efficient units are diagnosed inefficient. If it is too small, weakly efficient units are recognized efficient. The lexicographic approach is a better way. When the optimal solution of an LP-model is not unique, the sum of slack variables is used as a second objective to check whether the unit is efficient or not (see, e.g. Steuer [4], p. 445).

When the number of the units is large, let us say many ten thousands or even hundreds of thousands, computational aspects are important. Such problems appear when, for example, all high-schools or hospitals in Europe are evaluated, or when the efficiency analysis is made at an individual level. The straightforward approach to formulate an LP-model for each unit with an unknown status does not work. It is too time-consuming. Fortunately, the structure of the DEA-model makes it possible to develop special techniques for large-scale problems.

There are only few authors who have studied computational problems in DEA. In [5,6] the main idea was to restrict the basis entry. The basis always consists of a set of existing (efficient) units. When a unit is diagnosed inefficient, the corresponding column can be dropped from the set of potential basic vectors. In most cases, the technique clearly reduced computation time. Dulá and Helgason [7] proposed the solving of the problem in two phases. In phase I, the extreme point solutions of the polytope consisting of all units in the data set are defined. The efficiency scores of the other vectors are computed in phase II by using the minimal set of potential basic vectors, i.e. efficient extreme units. The idea was further developed in the paper by Dulá et al. [8]. The most recent developments by Dulá and his associates are presented in [2]. Because the computing time as the function of the units increases more than linearly, Barr and Durchholz [9] proposed the partition of the problem. The idea makes it possible to first identify the set of the efficient units in a small data set, and then to use those units to build a set of potential basic vectors. The union of those sets consists of all efficient units, but usually also inefficient units.

In this paper, we propose the use of lexicographic parametric programming [10] to classify the units as efficient and inefficient. Using that technique, we may move from unit to unit along an efficiency curve. The units entering the basis are recognized efficient and all units dominated by an efficient facet are inefficient. The move on the curve is terminated, when the end unit is diagnosed. The lexicographic parametric programming is needed to guarantee that the curve will stay on the efficient frontier also in case when it reaches the boundary.

After recognizing all efficient units, we may compute the scores of the inefficient units in the second phase as proposed by Barr and Durchholz [9] and Dulá et al. [8]. The number of the columns in these LP-problems is usually much smaller than in the original problem.

The paper is given in five sections. In the next section, necessary theory and the basic DEA models are represented. Section 3, describes the main principles of the procedure and illustrates a numerical example. Computational results are given and discussed in Section 4. Section 5 concludes the paper with some remarks.

2. Theoretical considerations

2.1. Lexicographic parametric programming

Consider the following problem ($p > 1$):

$$\begin{aligned} & lex\max && \{\boldsymbol{c}^1\boldsymbol{x}, \ldots, \boldsymbol{c}^p\boldsymbol{x}\} \\ & \text{s.t.} && \\ & && \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\ & && \boldsymbol{x} \geqslant \mathbf{0}, \end{aligned} \tag{2.1}$$

[2] Dulá and López [2] called small problems the ones consisted of less than 500 units.