Annals of **F**uzzy **M**athematics and **I**nformatics
Volume 5, No. 2, (March 2013), pp. 387–397
ISSN: 2093–9310 (print version)
ISSN: 2287–6235 (electronic version)
http://www.afmi.or.kr

Qualitative factors in data envelopment analysis

Faranak H. Saljooghi, Najme E. Giski

Received 27 Novemer 2011; Revised 12 April 2012; Accepted 29 May 2012

Abstract. In data envelopment analysis (DEA), performance evaluation is generally assumed to be based on a set of quantitative data. But in reality, many factors cannot be measured in a precise manner. In recent years, in different applications of DEA, data have been observed whose values are imprecise. Imprecise data can be fuzzy, interval, qualitative, ordinal or probabilistic. In many real world problem the qualitative factors have an important roles. So attention to qualitative factors is essential. The rankings are often provided from best to worst relative to particular attributes. Such rank positions might better be presented in an ordinal, rather than numerical sense. This paper introduces methods for solving DEA models with qualitative factors and investigation the properties of these models. Also we construct a two-level mathematical programming model, that its optimal value represents the upper-bound of the efficiency score.

2010 AMS Classification: 62–07, 90C99, 97R20

Keywords: Data envelopment analysis(DEA), Efficiency, Qualitative factor, Ordinal data.

Corresponding Author: Faranak H. Saljooghi (f_h_saljooghi@yahoo.com)

1. Introduction

In the data envelopment analysis (DEA) model of Charnes et al. [2], each member of a set of decision making units (DMUs) can be evaluated relativly to its peers. The standard data envelopment analysis method requires that the values for all inputs and outputs be known exactly. However, this assumption may not be true. When some outputs and inputs are unknown decision variables such as fuzzy, quantitative, bounded data, ordinal data, and ratio bounded data, the DEA model becomes a nonlinear programming problem and is called imprecise DEA (IDEA). In real world applications, however, there are cases where the data is described by qualitative terms, which have only ordinal relations, without exact numerical values. Typical example is using the words excellent, good, fair, unsatisfactory, and poor to represent performance. Linguistically, excellent is better than good, good is better than fair, and so forth; however, to what degree the former is better than the latter is not specified. Very often it is the case that for a factor such as management competence, one can, at most, provide a ranking of the DMUs from best to worst relative to this attribute. The capability of providing a more precise, quantitative measure reflecting such a factor is generally beyond the realm of reality. In some situations such factors can be legitimately 'quantified', but very often such quantification may be superficially forced as a modeling convenience. In such circumstances, 'data' for some certain factors (inputs and outputs) might be better represented as rank positions in an ordinal data, rather than numerical values. In this regard, assigning exact values to ordinal data distorts the original concept. The information available may permit one only to put each DMU into one of L categories or groups (e.g. 'high', 'medium' and 'low' competence) [4].

How to properly represent qualitative data has been widely discussed in the literature. The research started with the work of Cook et al. [3]. with the basic idea of using ordinal numbers to represent the precedence relationships of the qualitative data. Cooper et al. [5] coined the name imprecise DEA (IDEA) for mixtures of interval and ordinal data. Cooper et al. [5] apply a scale transformation technique accompanied with variable alterations and successfully transform the nonlinear model to a linear one. Zhu [12] simplifies their approach to reduce the computational burden. It will be shown in this paper that the efficiency scores calculated from their method are the upper bound of the efficiency intervals. Here a problem arise from Zhu's approach, he used of too many zero inputs and outputs in the process of efficiency evaluation. Cook et al.[4] presented a unified structure for embedding ordinal (rank order) data into the DEA framework.

The structure of this paper is as follows. Section 2 presents standard DEA model and IDEA approach for converting the ordinal data to interval data, also the determination suitable interval based on the point of decision maker is proposed. The scale-transformation and variable-alternative approach is shown in section 3. The two-level mathematical programming model indicate in Section 4, then the model for calculating the upper-bound of efficiency score is proposed. Section 5 explains a numerical example. The final section is allocated to conclusions.

2. Converting qualitative data to interval data in DEA

Data envelopment analysis (DEA) has been proved to be an useful tool in evaluating relative performance of homogeneous decision-making units (DMUs) in a multiple-input and multiple-output setting. The traditional DEA estimates the efficiency index by calculating the ratio of weighted outputs to weighted inputs, and the input and output weights are decided according to the best interests of the DMU being evaluated.

Consider n decision making units, one DMU receives m different inputs to produce s different outputs. Let x_{ij} and y_{rj} denote the ith input and rth output, respectively, of the jth decision making unit. Consider the multiplier form of the input-oriented

CCR model (Charnes et al. [2]). The efficiency of DMU_k, E_k, is calculated as:

$$
\begin{aligned}
\max \quad E_k = \quad & \sum_{r=1}^{s} \mu_r y_{rk} \\
s.t. \quad & \sum_{i=1}^{m} v_i x_{ik} = 1, \\
& \sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \ all\ j, \\
& \mu_r, v_i \geq \varepsilon, \ all\ r, i.
\end{aligned} \tag{2.1}
$$

where μ_r and v_i are virtual multipliers and ε is a small non-Archimedean value. It is assumed that all the input and output data x_{ij} and y_{rj} (i=1,...,m; r=1,...,s; j=1,...,n) cannot be exactly obtained due to the existence of uncertainty. They are only known to lie within the upper and lower bounds represented by the interval $[x_{ij}^L, x_{ij}^U]$ and $[y_{rj}^L, y_{rj}^U]$, where $x_{ij}^L > 0$ and $y_{rj}^L > 0$.

In order to deal with such an uncertain situation, the following pair of linear programming models has been developed to generate the upper and lower bounds of interval efficiency for each DMU [11]:

$$
\begin{aligned}
\max \quad E_k^U = \quad & \sum_{r=1}^{s} \mu_r y_{rk}^U \\
s.t. \quad & \sum_{i=1}^{m} v_i x_{ik}^L = 1, \\
& \sum_{r=1}^{s} \mu_r y_{rj}^U - \sum_{i=1}^{m} v_i x_{ij}^L \leq 0, \ all\ j, \\
& \mu_r, v_i \geq \varepsilon, \quad all\ r, i.
\end{aligned} \tag{2.2}
$$

$$
\begin{aligned}
\max \quad E_k^L = \quad & \sum_{r=1}^{s} \mu_r y_{rk}^L \\
s.t. \quad & \sum_{i=1}^{m} v_i x_{ik}^U = 1, \\
& \sum_{r=1}^{s} \mu_r y_{rj}^U - \sum_{i=1}^{m} v_i x_{ij}^L \leq 0, \ all\ j, \\
& \mu_r, v_i \geq \varepsilon, \quad all\ r, i.
\end{aligned} \tag{2.3}
$$

where E_k^U stands for the best possible relative efficiency achieved by DMU_k when all the DMUs are in state of best production activity, while E_k^L stands for the lower bound of the best possible relative efficiency of DMU_k.

Without loss of generality, assume rth output is qualitative factor, and a DMU_j can be assigned to one of $L(L \leq n)$ rank position. rank position values $y_r(\ell)\,(\ell = 1, ..., L)$ must satisfy in $y_r(\ell) > y_r(\ell + 1)$. Therefore, the strong ordinal preference information must satisfy as

$$y_r(1) > y_r(2) > ... > y_r(L) \tag{2.4}$$

Since DEA model has the property of unit-invariance, the use of scale transformation to ordinal preference information does not change the original ordinal relationships and has no effect on the efficiencies of DMUs. Therefore, it is possible to conduct a scale transformation to every ordinal input and output index so that its best ordinal datum is less than or equal to unity and then give an interval estimate for each ordinal datum.

For strong ordinal preference information $y_r(1) > y_r(2) > ... > y_r(L)$, we have the following ordinal relationships of after scale transformation:

$$1 \geq \tilde{y}_r(1) > \tilde{y}_r(2) > ... > \tilde{y}_r(L) \geq \alpha_r \tag{2.5}$$

where α_r is a small positive number reflecting the ratio of the possible minimum of $\{y_r(\ell)|\, \ell = 1, ..., L\}$ to its possible maximum. It can be approximately estimated by the decision maker. It is referred as the ratio parameter for convenience.

Strong ordinal preference information $\tilde{y}_r(\ell) > \tilde{y}_r(\ell+1)$ can be expressed as $\tilde{y}_r(\ell) \geq \chi_r \tilde{y}_r(\ell + 1)$, where $\chi_r > 1$ is the parameter on the degree of preference intensity provided by decision maker and satisfying $(\chi_r)^{1-L} \geq \alpha_r$. Therefore, for strong ordinal preference information $y_r(1) > y_r(2) > ... > y_r(L)$, there is the following ordinal relationships:

$$1 \geq \tilde{y}_r(1) \geq \chi_r \tilde{y}_r(2) \geq ... \geq (\chi_r)^{\ell-1} \tilde{y}_r(\ell) \geq ... \geq (\chi_r)^{L-1} \tilde{y}_r(L) \geq (\chi_r)^{L-1} \alpha_r \tag{2.6}$$

The resultant permissible interval for each $\tilde{y}_r(\ell)$ can be derived as follows:

$$\tilde{y}_r(\ell) \in [\alpha_r (\chi_r)^{L-\ell}, (\chi_r)^{1-\ell}], \quad \ell = 1, .., L \tag{2.7}$$

Through the scale transformation above and the estimation of permissible intervals, all the ordinal preference information is converted into interval data and can thus be incorporated interval DEA models [11].

Note that wang et al. [11] utilized the following data for each $\tilde{y}_r(\ell)$ $(\ell = 1, ..., L)$ to evaluate the worst possible relative efficiency of DMU_k:

$$\tilde{y}_r(\ell^1_{rk}) = \alpha_r (\chi_r)^{L-\ell^1_{rk}},$$

$$\tilde{y}_r(\ell) = (\chi_r)^{1-\ell}, \qquad (\ell \neq \ell^1_{rk})$$

Lemma 2.1. *Let $\tilde{y}_r(\ell^1_{rk}) = \alpha_r (\chi_r)^{L-\ell^1_{rk}}$, $\tilde{y}_r(\ell) = (\chi_r)^{1-\ell}$ $(\ell \neq \ell^1_{rk})$. The ordinal relationships (2.6) satisfies if and only if $\alpha_r = (\chi_r)^{1-L}$. (Note that $(\chi_r)^{1-L} \geq \alpha_r, \chi_r > 1$)*

Proof. Assume that the ordinal relationships (2.6) holds. Then $\tilde{y}_r(\ell^1_{rk}) \geq \chi_r \tilde{y}_r(\ell^1_{rk} + 1)$, in other words $\alpha_r (\chi_r)^{L-\ell^1_{rk}} \geq \chi_r (\chi_r)^{1-(\ell^1_{rk}+1)}$ therefore $\alpha_r \geq (\chi_r)^{1-L}$. Since $(\chi_r)^{1-L} \geq \alpha_r$, it follows that $\alpha_r = (\chi_r)^{1-L}$.

Conversely, suppose that $\alpha_r = (\chi_r)^{1-L}$. Then $\tilde{y}_r(\ell) = (\chi_r)^{1-\ell}\, (\forall \ell)$, that obviously the ordinal relationships (2.6) holds. □

3. Scale-transformation and variable-alternation approach

Consider the situation in which a set of n decision making units (DMUs), $j = 1, ..., n$ are to be evaluated in terms of R_1 quantitative outputs, R_2 qualitative outputs, I_1 quantitative inputs, I_2 qualitative inputs. Let $Y^1_j = (y^1_{rj})$, $Y^2_j = (y^2_{rj})$ denote the R_1-dimensional and R_2-dimensional vectors of outputs, respectively. Similarly, let $X^1_j = (x^1_{ij})$ and $X^2_j = (x^2_{ij})$ be the I_1-and I_2-dimensional vectors of inputs, respectively.

In the situation where all factors are quantitative and $DMU_k(k = 1, ..., n)$ is under evaluation by the CCR model (Charnes et al. [2]), we have

$$(3.1)\qquad \begin{aligned} E_k = max \quad & \sum_{r\in R_1} \mu_r^1 y_{rk}^1 + \sum_{r\in R_2} \mu_r^2 y_{rk}^2 \\ s.t. \quad & \sum_{i\in I_1} v_i^1 x_{ik}^1 + \sum_{i\in I_2} v_i^2 x_{ik}^2 = 1, \\ & \sum_{r\in R_1} \mu_r^1 y_{rj}^1 + \sum_{r\in R_2} \mu_r^2 y_{rj}^2 - \sum_{i\in I_1} v_i^1 x_{ij}^1 - \sum_{i\in I_2} v_i^2 x_{ij}^2 \le 0, all\ j, \\ & \mu_r^1, \mu_r^2, v_i^1, v_i^2 \ge \varepsilon, \text{ all } r, i. \end{aligned}$$

where E_k is the efficiency score of DMU_k , μ_r^1, μ_r^2, v_i^1 and v_i^2 are virtual multipliers and ε is a small non-Archimedean value.

To place the problem in a general framework, assume that for each qualitative factor $(r \in R_2,\ i \in I_2)$, a DMU_j can be assigned to one of $L\,(L \le N)$ rank position. We use the convention that for both outputs and inputs, a rating 1 is 'best', and L 'worst'. For outputs, this means that a DMU ranked at position 1 generates more output than a DMU in position 2, and so on. For inputs, a DMU in position 1 consumes less input than one in position 2.

One can view the allocation of a DMU to a rank position ℓ on an output r, for example, as having assigned that DMU an output value or worth $y_r^2(\ell)$. The implementation of the DEA model (3.1) thus involves determining two things:

(1) multiplier values μ_r^2, v_i^2 for outputs $r \in R_2$ and inputs $i \in I_2$;

(2) rank position values $y_r^2(\ell), r \in R_2$, and $x_i^2(\ell), i \in I_2$, all ℓ.

To facilitate development herein, define the L-dimensional unit vectors $\gamma_{rj} = (\gamma_{rj}(\ell))$, and $\delta_{ij} = (\delta_{ij}(\ell))$ where

$$\gamma_{rj}(\ell) = \begin{cases} 1 & \text{if } DMU_j \text{ is ranked in } \ell\text{th position on output } r, \\ 0, & \text{otherwise,} \end{cases}$$
$$\delta_{ij}(\ell) = \begin{cases} 1 & \text{if } DMU_j \text{ is ranked in } \ell\text{th position on input } i, \\ 0, & \text{otherwise.} \end{cases}$$

It is noted that y_{rj}^2, x_{ij}^2 can be represented in the form

$$y_{rj}^2 = y_r^2(\ell_{rj}^1) = \sum_{\ell=1}^{L} y_r^2(\ell)\gamma_{rj}(\ell),$$
$$x_{ij}^2 = x_i^2(\ell_{rj}^2) = \sum_{\ell=1}^{L} x_i^2(\ell)\delta_{ij}(\ell),$$

where ℓ_{rj}^1 and ℓ_{ij}^2 are the rank position occupied by DMU_j on output r and input i, respectively. Hence, model (3.1) can be rewritten in the more representative format:

$$(3.2)\qquad \begin{aligned} E_k = max \quad & \sum_{r\in R_1} \mu_r^1 y_{rk}^1 + \sum_{r\in R_2}\sum_{\ell=1}^{L} \mu_r^2 y_r^2(\ell)\gamma_{rk}(\ell) \\ s.t. \quad & \sum_{i\in I_1} v_i^1 x_{ik}^1 + \sum_{i\in I_2}\sum_{\ell=1}^{L} v_i^2 x_i^2(\ell)\delta_{ik}(\ell) = 1, \\ & \sum_{r\in R_1} \mu_r^1 y_{rj}^1 + \sum_{r\in R_2}\sum_{\ell=1}^{L} \mu_r^2 y_r^2(\ell)\gamma_{rj}(\ell) - \\ & \sum_{i\in I_1} v_i^1 x_{ij}^1 + \sum_{i\in I_2}\sum_{\ell=1}^{L} v_i^2 x_i^2(\ell)\delta_{ij}(\ell) \le 0, all\, j, \end{aligned}$$

$$\{Y_r^2 = (y_r^2(\ell)),\ X_i^2 = (x_i^2(\ell))\} \in \Psi, \mu_r^1, \mu_r^2, v_i^1, v_i^2 \geq \varepsilon,\ all\ r, i.$$

The values or worths $\{y_r^2(\ell)\}$, $\{x_i^2(\ell)\}$, attached to the ordinal rank position for outputs r and inputs i, respectively, must satisfy the minimal requirement that it is more important to be ranked in ℓth position than in the $(\ell+1)$th position on any such qualitative factor. Specifically, $y_r^2(\ell) > y_r^2(\ell+1)$ and $x_i^2(\ell) < x_i^2(\ell+1)$. That is, for outputs, one places a higher weight on being ranked in ℓth place than in $\ell+1$th place. For inputs, the opposite is true. Cook and Zhu [4] a set of linear conditions that produce this realization is defined by the set Ψ, where

$$\Psi = \{(Y_r^2, X_i^2)|\quad \begin{array}{l} y_r^2(\ell) - y_r^2(\ell+1) \geq \delta,\ \ell = 1, ..., L-1, y_r^2(L) \geq \delta, \delta > 0\,, \\ x_i^2(\ell+1) - x_i^2(\ell) \geq \sigma,\ \ell = 1, ..., L-1,\ x_i^2(1) \geq \sigma, \sigma > 0\} \end{array} \tag{3.3}$$

Arguably, δ and σ could be made dependent upon ℓ (i.e. replace δ by δ_ℓ, σ by σ_ℓ).

It is noted that with the change of variables $w_{r\ell}^1 = \mu_r^2 y_r^2(\ell)$ and $w_{i\ell}^2 = v_i^2 x_i^2(\ell)$ the model (3.2) can be converted to the following linear program:

$$\begin{array}{lll} E_k = max & \sum\limits_{r\in R_1} \mu_r^1 y_{rk}^1 + \sum\limits_{r\in R_2} \sum\limits_{\ell=1}^{L} w_{r\ell}^1 \gamma_{rk}(\ell) & \\ s.t. & \sum\limits_{i\in I_1} v_i^1 x_{ik}^1 + \sum\limits_{i\in I_2} \sum\limits_{\ell=1}^{L} w_{i\ell}^2 \delta_{ik}(\ell) = 1, & \\ & \sum\limits_{r\in R_1} \mu_r^1 y_{rj}^1 + \sum\limits_{r\in R_2} \sum\limits_{\ell=1}^{L} w_{r\ell}^1 \gamma_{rj}(\ell) - & \\ & \sum\limits_{i\in I_1} v_i^1 x_{ij}^1 + \sum\limits_{i\in I_2} \sum\limits_{\ell=1}^{L} w_{i\ell}^2 \delta_{ij}(\ell) \leq 0,\ \ all\, j, & \\ & w_{r\ell}^1 - w_{r,\ell+1}^1 \geq \mu_r^2 \delta, \quad \ell = 1, ..., L-1, & all\ r \in R_2, \\ & w_{rL}^1 \geq \mu_r^2 \delta, & all\ r \in R_2, \\ & w_{i\ell}^2 - w_{i,\ell+1}^2 \geq v_i^2 \sigma, \quad \ell = 1, ..., L-1, & all\ i \in I_2, \\ & w_{i1}^2 \geq v_i^2 \sigma, & all\ i \in I_2, \\ & \mu_r^1, \mu_r^2, v_i^1, v_i^2 \geq \varepsilon,\ all\ r, i. & \end{array} \tag{3.4}$$

Theorem 3.1. *For any positive values of δ, σ and adding (3.3) to model (3.2), this model is equivalent with adding (12) to model (3.2). where*

$$\Psi' = \{(Y_r^2, X_i^2)|\quad \begin{array}{l} y_r^2(\ell) - y_r^2(\ell+1) \geq \delta',\ \ell = 1, ..., L-1, y_r^2(L) \geq \delta', \\ x_i^2(\ell+1) - x_i^2(\ell) \geq \sigma',\ \ell = 1, ..., L-1,\ x_i^2(1) \geq \sigma'\} \end{array} \tag{3.5}$$

and σ', $\delta' \approx 0$.

Proof. For any positive values δ, σ select a large enough number ρ, ω such that $\frac{\delta}{\rho} = \delta'$, $\frac{\sigma}{\omega} = \sigma'$. Now, for $r \in R_2$ and $i \in I_2$, define

$$\tilde{y}_r^2(\ell) = \frac{y_r^2(\ell)}{\rho},\ \tilde{x}_i^2(\ell) = \frac{x_i^2(\ell)}{\omega},\ \tilde{\mu}_r^2 = \rho\mu_r^2\,, \tilde{v}_i^2 = \omega v_i^2.$$

Then model (3.2) with (3.3) is equivalent to model (3.2) with (3.5). This completes the proof. □

Theorem 3.1 indicates that when (3.3) is imposed, model (3.2) is unable to discriminate efficiency values based on $y_r^2(\ell) > y_r^2(\ell+1)$, $x_i^2(\ell) < x_i^2(\ell+1)$ from the ones based on $y_r^2(\ell) \geq y_r^2(\ell+1)$, $x_i^2(\ell) \leq x_i^2(\ell+1)$, i.e., expression (3.3) cannot replace $y_r^2(\ell) > y_r^2(\ell+1)$, $x_i^2(\ell) < x_i^2(\ell+1)$ in computation. In other words, expression (3.3) is not a valid and functional modification to $y_r^2(\ell) > y_r^2(\ell+1)$, $x_i^2(\ell) < x_i^2(\ell+1)$. Thus we define

(3.6)

$$\Psi = \{(Y_r^2, X_i^2)| \quad \begin{array}{l} y_r^2(\ell) \geq \chi_r y_r^2(\ell+1),\ \ell = 1,...,L-1, y_r^2(L) \geq \chi_r, \\ \eta_i x_i^2(\ell) \leq x_i^2(\ell+1),\ \ell = 1,...,L-1,\ x_i^2(1) \geq \eta_i, \quad \chi_r\ , \eta_i > 1\} \end{array}$$

where χ_r, η_i can be approximately estimated by the decision maker. Hence, model (3.2) rewritten as follows:

$$\begin{array}{lll} E_k = max & \sum\limits_{r\in R_1} \mu_r^1 y_{rk}^1 + \sum\limits_{r\in R_2} \sum\limits_{\ell=1}^{L} w_{r\ell}^1 \gamma_{rk}(\ell) & \\ s.t. & \sum\limits_{i\in I_1} v_i^1 x_{ik}^1 + \sum\limits_{i\in I_2} \sum\limits_{\ell=1}^{L} w_{i\ell}^2 \delta_{ik}(\ell) = 1, & \\ & \sum\limits_{r\in R_1} \mu_r^1 y_{rj}^1 + \sum\limits_{r\in R_2} \sum\limits_{\ell=1}^{L} w_{r\ell}^1 \gamma_{rj}(\ell) - & \\ & \sum\limits_{i\in I_1} v_i^1 x_{ij}^1 + \sum\limits_{i\in I_2} \sum\limits_{\ell=1}^{L} w_{i\ell}^2 \delta_{ij}(\ell) \leq 0, \quad all\ j, & \\ & w_{r\ell}^1 - \chi_r w_{r,\ell+1}^1 \geq 0, \qquad \ell = 1,...,L-1, \quad all\ r \in R_2, & \\ & w_{rL}^1 \geq \chi_r \mu_r^2, \qquad all\ r \in R_2, & \\ & w_{i,\ell+1}^2 - \eta_i w_{i\ell}^2 \geq 0, \qquad \ell = 1,...,L-1, \quad all\ i \in I_2, & \\ & w_{i1}^2 \geq \eta_i v_i^2, \qquad all\ i \in I_2, & \\ & \mu_r^1, \mu_r^2, v_i^1, v_i^2 \geq \varepsilon,\ all\ r, i. & \end{array} \tag{3.7}$$

4. The two-level mathematical programming model

Consider the situation in which a set of n decision making units are to be evaluated in terms of R_1 quantitative outputs, R_2 qualitative outputs, I_1 quantitative inputs, I_2 qualitative inputs. Without loss of generality, we assume

$$\begin{array}{l} 1 \geq y_r^2(1) \geq y_r^2(2) \geq ... \geq y_r^2(L) \geq \alpha_r, \quad (r \in R_2) \\ \beta_i \leq x_i^2(1) \leq x_i^2(2) \leq ... \leq x_i^2(L) \leq 1, \quad (i \in I_2) \end{array} \tag{4.1}$$

where α_r, β_i are small positive numbers. Each set of pair $(y_r^2(\ell), x_i^2(\ell))$ $(\ell = 1,...,L; r \in R_2; i \in I_2)$ that satisfies the ordinal relationships (4.1), with other exact values, can apply model (3.1) to calculate an efficiency score. Different $(y_r^2(\ell), x_i^2(\ell))$ sets have different efficiency scores. The set of $(y_r^2(\ell), x_i^2(\ell))$ that attains the highest efficiency score for DMU_k can be determined from the following two-level mathematical programming model:

(4.2)

$$E_k^U = \underset{\substack{1\geq y_r^2(1)\geq y_r^2(2)\geq ...\geq y_r^2(L)\geq b_r, \\ a_i\leq x_i^2(1)\leq x_i^2(2)\leq ...\leq x_i^2(L)\leq 1,}}{\max imize} \left\{ \begin{array}{l} \tilde{E}_k = \max \sum\limits_{r\in R_1} \mu_r^1 y_{rk}^1 + \sum\limits_{r\in R_2} \mu_r^2 y_r^2(\ell_{rk}^1) \\ \quad s.t.\ \sum\limits_{i\in I_1} v_i^1 x_{ik}^1 + \sum\limits_{i\in I_2} v_i^2 x_i^2(\ell_{ik}^2) = 1, \\ \qquad \sum\limits_{r\in R_1} \mu_r^1 y_{rj}^1 + \sum\limits_{r\in R_2} \mu_r^2 y_r^2(\ell_{rj}^1) - \\ \qquad \sum\limits_{i\in I_1} v_i^1 x_{ij}^1 - \sum\limits_{i\in I_2} v_i^2 x_i^2(\ell_{ij}^2) \leq 0,\ all\ j, \\ \qquad \mu_r^1, \mu_r^2, v_i^1, v_i^2 \geq \varepsilon,\ all\ r, i. \end{array} \right.$$

The inner program, i.e., the second-level program, calculates the efficiency score for each set of $(y_r^2(\ell), x_i^2(\ell))$ defined by the outer program, i.e., the first-level program, while the outer program determines the set of $(y_r^2(\ell), x_i^2(\ell))$ that produces the highest efficiency score. The optimal value E_k^U is the upper bound of the efficiency score for DMU_k.

Proposition 4.1. *The highest efficiency score for DMU_k is attained by setting its qualitative outputs at the upper bounds and the qualitative inputs at the lower bounds, meanwhile, the qualitative outputs of all other DMUs at their corresponding lowest levels and the qualitative inputs at their corresponding highest levels. in other words, the optimal value of following model is equals E_k^U.*

$$\begin{aligned} E_k^* = max \quad & \sum_{r\in R_1} \mu_r^1 y_{rk}^1 + \sum_{r\in R_2} \mu_r^2 \zeta_{rk}^1 \\ s.t. \quad & \sum_{i\in I_1} v_i^1 x_{ik}^1 + \sum_{i\in I_2} v_i^2 \zeta_{ik}^2 = 1, \\ & \sum_{r\in R_1} \mu_r^1 y_{rj}^1 + \sum_{r\in R_2} \mu_r^2 \zeta_{rj}^1 - \sum_{i\in I_1} v_i^1 x_{ij}^1 - \sum_{i\in I_2} v_i^2 \zeta_{ij}^2 \le 0, \ all\ j, \\ & \mu_r^1, \mu_r^2, v_i^1, v_i^2 \ge \varepsilon, \ all\ r, i. \end{aligned} \tag{4.3}$$

where $\zeta_{rj}^1, \zeta_{ij}^2$ define as follows:

$$\zeta_{rj}^1 = \begin{cases} 1 & \ell_{rj}^1 = 1, ..., \ell_{rk}^1, \\ b_r & \ell_{rj}^1 = \ell_{rk}^1 + 1, ..., L, \end{cases}$$

$$\zeta_{ij}^2 = \begin{cases} a_i & \ell_{ij}^2 = 1, ..., \ell_{ik}^2, \\ 1 & \ell_{ij}^2 = \ell_{ik}^2 + 1, ..., L. \end{cases}$$

Proof. Obviously, $E_k^* \le E_k^U$. Let set of $(y_r^2(\ell), x_i^2(\ell))$ $(\ell = 1, ..., L; r \in R_2; i \in I_2)$ is arbitrary and satisfies the ordinal relationships (4.1). Denote $(\tilde{\mu}_r^1, \tilde{\mu}_r^2, \tilde{v}_i^1, \tilde{v}_i^2)$ and $\tilde{E}_k$ as the optimal solution and the optimal value for the second-level program (4.2), respectively. Also suppose E_k^* is optimal value to model (4.3), then can decrease $\tilde{\mu}_r^2$ to $\bar{\mu}_r^2$ so that $\tilde{\mu}_r^2 y_r^2(\ell_{rk}^1) = \bar{\mu}_r^2$ and increase $\tilde{v}_i^2$ to $\bar{v}_i^2$ so that $\tilde{v}_i^2 x_i^2(\ell_{ik}^2) = \bar{v}_i^2 a_i$. So

$$\sum_{i\in I_1} \tilde{v}_i^1 x_{ik}^1 + \sum_{i\in I_2} \bar{v}_i^2 \zeta_{ik}^2 = \sum_{i\in I_1} \tilde{v}_i^1 x_{ik}^1 + \sum_{i\in I_2} \tilde{v}_i^2 x_i^2(\ell_{ik}^2) = 1, \tag{4.4}$$

$$\tilde{\mu}_r^2 y_r^2(\ell_{rj}^1) \ge \tilde{\mu}_r^2 y_r^2(\ell_{rk}^1) = \bar{\mu}_r^2 = \bar{\mu}_r^2 \zeta_{rj}^1, \ \ (\ell_{rj}^1 = 1, ..., \ell_{rk}^1;\ r \in R_2) \tag{4.5}$$

$$\tilde{\mu}_r^2 y_r^2(\ell_{rj}^1) \ge \bar{\mu}_r^2 b_r = \bar{\mu}_r^2 \zeta_{rj}^1, \ \ (\ell_{rj}^1 = \ell_{rk}^1 + 1, ..., L;\ r \in R_2) \tag{4.6}$$

$$\tilde{v}_i^2 x_i^2(\ell_{ij}^2) \le \tilde{v}_i^2 x_i^2(\ell_{ik}^2) = \bar{v}_i^2 a_i = \bar{v}_i^2 \zeta_{ij}^2, \ \ (\ell_{ij}^2 = 1, ..., \ell_{ik}^2;\ i \in I_2) \tag{4.7}$$

$$\tilde{v}_i^2 x_i^2(\ell_{ij}^2) \le \bar{v}_i^2 = \bar{v}_i^2 \zeta_{ij}^2, \ \ (\ell_{ij}^2 = \ell_{ik}^2 + 1, ..., L;\ i \in I_2) \tag{4.8}$$

The conditions (4.5)-(4.8) imply that:

$$\begin{aligned} & \sum_{r\in R_1} \tilde{\mu}_r^1 y_{rj}^1 + \sum_{r\in R_2} \bar{\mu}_r^2 \zeta_{rj}^1 - \sum_{i\in I_1} \tilde{v}_i^1 x_{ij}^1 - \sum_{i\in I_2} \bar{v}_i^2 \zeta_{ij}^2 \le \\ & \sum_{r\in R_1} \tilde{\mu}_r^1 y_{rj}^1 + \sum_{r\in R_2} \tilde{\mu}_r^2 y_r^2(\ell_{rj}^1) - \sum_{i\in I_1} \tilde{v}_i^1 x_{ij}^1 - \sum_{i\in I_2} \tilde{v}_i^2 x_i^2(\ell_{ij}^2) \le 0, \ \ all\ j \end{aligned} \tag{4.9}$$

The conditions (4.4) and (4.9) imply that $(\tilde{\mu}_r^1, \bar{\mu}_r^2, \tilde{v}_i^1, \bar{v}_i^2)$ is a feasible solution to model (4.3). Also

$$\tilde{E}_k = \sum_{r\in R_1} \tilde{\mu}_r^1 y_{rk}^1 + \sum_{r\in R_2} \tilde{\mu}_r^2 y_r^2(\ell_{rk}^1) = \sum_{r\in R_1} \tilde{\mu}_r^1 y_{rk}^1 + \sum_{r\in R_2} \bar{\mu}_r^2 \zeta_{rk}^1,$$

Then it holds that $E_k^* \geq \tilde{E}_k$ therefore $E_k^* \geq E_k^U$. This completes the proof. □

In this study, the basic idea is to find the values for the qualitative data and the associated weights which yield the most favorable efficiency score for the DMU in concern. As a result, in calculating the efficiency of a DMU, different values may be assigned to the same ordinal rank position, e.g. ℓth position, from different DMUs. Another case is that a smaller value is assigned to the ℓth position for output r from one DMU than that assigned to the $(\ell+1)$th position from another DMU. Therefore we propose to adopt variable production frontiers to measure the efficiency of different DMUs. Moreover, in this method the values for two consecutive ordinal variables distinguished by a small value of δ. Mathematically, the small difference of δ reflects the precedence relationship. In reality, however, this small difference is not meaningful. Therefore this method only superficially reflect the precedence relationship of the ordinal data. Kao and Lin [8] representing the qualitative levels by fuzzy numbers and uses the opinion obtained from the DMUs being evaluated to construct the membership function of the qualitative level. To deal quantitatively with imprecise data, Bellman and Zadeh [1] introduced the notion of fuzziness, and this approach has been successfully applied to solving many real world problems. Under the framework of DEA, different models have also been developed for measuring fuzzy efficiencies (Dia [6]; Guo and Tanaka [7]; Kao and Liu [9]; Leon et al. [10]). However, these approaches are not appropriate for dealing with qualitative data, because they treat all fuzzy observations independently in calculating efficiency. Therefore Kao and Lin [8], developed a new model. For more discussions on their method, please refer to Kao and Lin [8]. Consider 12 DMUs with two inputs and two outputs as shown by Table 1. In this table, each DMU has two inputs x_1, x_2 (columns 2 and 3) and two outputs y_1, y_2 (columns 4 and 5). Note that y_2 is qualitative indicator. We consider the qualitative indicator (y_2) with "1" for the best and "5" for the worst.

5. Numerical example

Consider 12 DMUs with two inputs and two outputs as shown by Table 1. In this table, each DMU has two inputs x_1, x_2 (columns 2 and 3) and two outputs y_1, y_2 (columns 4 and 5). Note that y_2 is qualitative indicator. We consider the qualitative indicator (y_2) with "1" for the best and "5" for the worst.

The columns 2 and 3 in Table 2 indicated the efficiency scores by means of model (3.4) when δ equal to 0.01 and 200, respectively. It can be seen that approach yields the same efficiency scores. (according to Theorem 3.1)

Table 1. The data for 12 DMU

DMUs	*Input* x_1	*Input* x_2	*Output* y_1	*Output* y_2
1	4	51	90	1
2	31	46	97	3
3	152	24	54	2
4	83	56	83	2
5	74	75	104	3
6	26	25	36	4
7	1	3	2	3
8	2	33	49	5
9	51	7	70	3
10	28	289	26	4
11	401	179	630	2
12	241	233	384	3

We set $\chi = 1.5$ and $\alpha = 0.01$ for output y_2. Based upon (2.7), we obtain the set of interval data $y_2(\ell) \in [0.01(1.5)^{5-\ell}, (1.5)^{1-\ell}]$ $(\ell = 1, ..., 5)$. The interval efficiency shown in column 4. The efficiency scores obtained from (4.3) are presented in final column.

Table 2
The result of hypothetical data for 12 DMU

DMUs	Efficiency score for $\delta = 0.01$	Efficiency score for $\delta = 200$	Efficiency interval	Highest efficiency score
1	1	1	[1,1]	1
2	0.862	0.862	[0.862,0.862]	0.862
3	1	1	[0.252,0.312]	0.258
4	0.747	0.747	[0.438,0.439]	0.439
5	0.495	0.495	[0.495,0.495]	0.495
6	0.535	0.535	[0.503,0.503]	0.503
7	1	1	[0.327,1]	0.443
8	1	1	[1,1]	1
9	1	1	[1,1]	1
10	0.065	0.065	[0.0503,0.0504]	0.0504
11	0.831	0.831	[0.831,0.831]	0.831
12	0.577	0.577	[0.577,0.577]	0.577

There are the highest efficiency score in efficiency interval, but their scores aren't always the most value in efficiency interval for all DMUs. In the results of model (4.3), DMU_3 has efficiency score equal 0.258 but this unit is efficient in model (2.4), it can be seen that DMU_3 uses large inputs , therefore, this unit can not be efficient.

6. Conclusions

This paper reviews the two approaches for solving DEA with ordinal data. One is based on converting ordinal data to interval data and the other is based on scale-transformation and variable-alternation approach. The correction of intervals in converting ordinal data to interval data is proposed on the point of decision maker. It is also shown that the strong ordinal relations are not correctly dealt in the approach of scale-transformation and variable-alternation. We provide the improved and corrected approach for imposing strong ordinal relations. In this paper is used a two-level mathematical programming model, such as the calculated optimal value represents the upper-bound of the efficiency; Also its equivalent model as a linear programming is presented. The numerical example illustrate the weaknesses of the old approaches, also shows the results of corrected approaches.

References

[1] R. Bellman and L. A. Zadeh, Decision making in a fuzzy environment, Management Sci. 17 (1970/71) B141–B164.
[2] A. Charnes, W. W. Cooper and E. Rhodes, Measuring the efficiency of decision making units, European J. Oper. Res. 2 (1978) 429–444.
[3] W. D. Cook, M. Kress and L. Seiford, On the use of ordinal data in data envelopment analysis, European J. Oper. Res. 44 (1993) 133–140.
[4] W. D. Cook and J. Zhu, Rank order data in DEA: A general framework, European J. Oper. Res. 174 (2006) 1021–1038.
[5] W. W. Cooper, K. S. Park and G. Yu, IDEA and AR-DEA: Models for dealing with imprecise data in DEA, Manage. Sci. 45 (1999) 597–607.
[6] M. Dia, A model of fuzzy data envelopment analysis, INFOR. Inf. Syst. Oper. Res. 42 (2004) 267–279.
[7] P. Gao and H. Tanaka, Fuzzy DEA: A perceptual evaluation method, Fuzzy Sets and Systems 119 (2000) 149–160.
[8] C. Kao and P. H. Lin, Qualitative factors in data envelopment analysis: A fuzzy number approach, European J. Oper. Res. 211 (2011) 586–593.
[9] C. Kao and S. T. Liu, Fuzzy efficiency measures in data envelopment analysis, Fuzzy Sets and Systems 113 (2000) 427–437.
[10] T. Leon, V. Liern J. L. Ruiz and I. Sirvent, A fuzzy mathematical programming approach to the assessment of efficiency with DEA models, Fuzzy Sets and Systems 139 (2003) 407–419.
[11] Y. M. Wang, R. Greatbanks and J. B. Yang, Interval efficiency assessment using data envelopment analysis, Fuzzy Sets and Systems 153 (3) (2005) 347–370.
[12] J. Zhu, Imprecise data envelopment analysis (IDEA): A review and improvement with an application, European J. Oper. Res. 144 (2003) 513–529.

Faranak Hosseinzadeh Saljooghi (f_h_saljooghi@yahoo.com)
Department of Mathematics, University of Sistan and Baluchestan, Iran

Najme Eslami Giski (eslamig.najme@yahoo.com)
Department of Mathematics, University of Sistan and Baluchestan, Iran