Int. J. Contemp. Math. Sciences, Vol. 2, 2007, no. 4, 159 - 166

Ranking of DMUs on Interval Data by DEA

F. Hosseinzadeh Lotfi

Department of Mathematics, Science and Research Branch
Islamic Azad University
P.O. Box 14155/775 and 14155/4933, Post code: 1477893855, Tehran, Iran
hosseinzadeh_lotfi@yahoo.com

G. R. Jahanshahloo

Department of Mathematics, Teacher Training University
Tehran, Iran

F. Rezai Balf

Department of Mathematics, Islamic Azad University
Qaemshahr, Iran

H. Zhiani Rezai

Department of Mathematics, Islamic Azad University
Mashhad, Iran

Abstract

The paper considers the problem ranking interval data in data envelopment analysis (IDEA), different of the other methods. In this approach, we use the statistical concepts. The end DEA models are modified and for illustration a numerical example is proposed.

Mathematics Subject Classification: Operations Research, No. 90

Keywords: Data Envelopment Analysis, Ranking, Interval Data, Variance-Covariance Matrices

1. Introduction

Using DEA models for evaluating relative efficiency of Decision Making Units (DMUs), usually give score 1 for a set of DMUs with cardinality greater

than or equal one [1,2]. This raises a logical question: which is of these efficient DMUs better than others, when DMUs are as interval data. In this paper, we are concerned on the mean and generalized variance of evaluated decision making interval unit.
The paper is organized as follows: in section 2, we give a concept of interval data, section 3 is a description on variance-covariance matrix and generalized variance that uses to assess rank, in section 4 we introduce a approach (criteria) for ranking of under evaluate unit, also a numerical example are given in section 5, and in section 6 conclusion is put forward.

2. Background

Consider n, DMUs each using m input to produce s output. We denote by x_{ij} the level of the $i-th$ input $(i = 1, ..., m)$ and the level of the $r-th$ output $(r = 1, ..., s)$ from the $j-th$ unit, $(j \in J = \{1, ..., n\})$. The input-output data are known to lie within bounded intervals, i.e. $x_{ij} \in [x_{ij}^L, x_{ij}^U]$ and $y_{rj} \in [y_{rj}^L, y_{rj}^U]$, with the upper and lower of intervals given as constants and assumed positive. We use a model which is defined as follows [5]:

$$\begin{array}{lll} Max & \sum_{r=1}^{s} u_r[y_{rp}^L, y_{rp}^U] & \\ S.t. & \sum_{i=1}^{m} v_i[x_{ip}^L, x_{ip}^U] = 1 & \\ \multicolumn{2}{l}{\sum_{r=1}^{s} u_r[y_{rj}^L, y_{rj}^U] - \sum_{i=1}^{m} v_i[x_{ij}^L, x_{ij}^U] \leq 0,} & j = 1, ..., n \\ u_r \geq 0, & & r = 1, ..., s \\ v_i \geq 0, & & i = 1, ..., m. \end{array} \tag{1}$$

Obviously, above model is non-linear. Now, the following models [3], are obtained upper and lower bounds of the efficiency scores for DMU_p $(p \in J)$.

$$\begin{array}{lll} Max & z_p^U = \sum_{r=1}^{s} u_r y_{rp}^U & \\ S.t. & \sum_{i=1}^{m} v_i x_{ip}^L = 1 & \\ & \sum_{r=1}^{s} u_r y_{rp}^U - \sum_{i=1}^{m} v_i x_{ip}^L \leq 0, & \\ & \sum_{r=1}^{s} u_r y_{rj}^L - \sum_{i=1}^{m} v_i x_{ij}^U \leq 0, & j = 1, ..., n, j \neq p \\ & u_r \geq 0, & r = 1, ..., s \\ & v_i \geq 0, & i = 1, ..., m. \end{array} \tag{2}$$

and

$$\begin{array}{lll} Max & z_p^L = \sum_{r=1}^{s} u_r y_{rp}^L & \\ S.t. & \sum_{i=1}^{m} v_i x_{ip}^U = 1 & \\ & \sum_{r=1}^{s} u_r y_{rp}^L - \sum_{i=1}^{m} v_i x_{ip}^U \leq 0, & \\ & \sum_{r=1}^{s} u_r y_{rj}^U - \sum_{i=1}^{m} v_i x_{ij}^L \leq 0, & j = 1, ..., n, j \neq p \\ & u_r \geq 0, & r = 1, ..., s \\ & v_i \geq 0, & i = 1, ..., m. \end{array} \tag{3}$$

Suppose that the efficiency z_p^{L*} and z_p^{U*} be attained by evaluating DMU_p in model (2) and model (3) respectively. Suppose that $z_p^*(x, y)$ be optimal objective value in model (1), then for each $x_{ij} \in [x_{ij}^L, x_{ij}^U]$ and $y_{rj} \in [y_{rj}^L, y_{rj}^U]$, where x, y are inputs and outputs matrices respectively, it will prove $z_p^*(x, y) \in [z_p^{L*}, z_p^{U*}]$ [3].

2.1. Classification of DMUs

On the basis of the obtained efficiency scores intervals, the units can be first classified in three subsets as follows [4,5,6]:

$E^{++} = \{j \in J \ \mid \ z_j^{L*} = 1\}$
$E^{+} = \{j \in J \ \mid \ z_j^{L*} < 1, \ \ z_j^{U*} = 1\}$
$E^{-} = \{j \in J \ \mid \ z_j^{U*} < 1\}$

Where J shows for the index set $\{1, ..., n\}$ of the units. The set E^{++} contains of the units that are efficient in any situation (the best and worst position). The set E^{+} contains of units are efficient in the best position, but they are not efficient in worst position. Finally, the set E^{-} contains of inefficient units in any situation. In Fig. 1 the sets of DMUs is shown with one interval input and one interval output in T_v, which T_v is production Possibility Set (PPS) of model BCC [2]. Regarding to Fig. 1 we obtain:
$E^{++} = \emptyset, \quad E^{+} = \{A, B, C, D, E\}, \quad E^{-} = \emptyset$

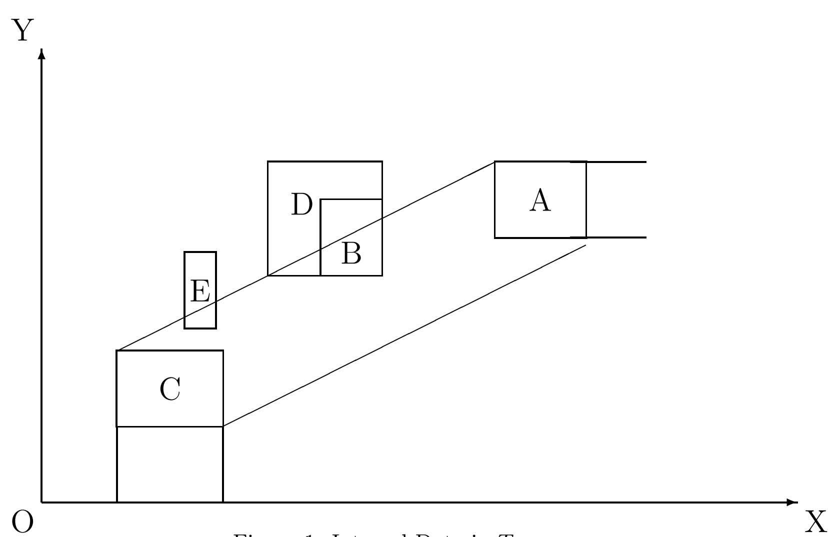

Figure 1. Interval Data in T_v

3. Variance-Covariance Matrix and Generalized Variance

3.1. Definition
Matrix $S = (s_{ij})_{m\times m}$ is called variance-covariance matrix if and only if

$$s_{ij} = \begin{cases} var(x_i), & i = j \\ covar(x_i, x_j), & i \neq j \end{cases}$$

Where $var(.)$ and $covar(.)$ are variance and covariance, respectively. [7]

3.2. Definition
Suppose that $S = (s_{ij})_{m\times m}$ be variance-covariance matrix. Then, generalized variance (gvar) is defined as follow:

$$gvar(S) = tr(S) = \sum_{i=1}^{m} s_{ii}$$

The generalized variance provide for away for writing information in all variance-covariance as a solely number. For example with regarding above definitions, we obtain generalized variance of the matrix $X = [x^L \quad x^U]^t$ that is $2 \times m$ matrix, where $x^L = (x_1^L, ..., x_m^L)$ and $x^U = (x_1^U, ..., x_m^U)$(the sign t is used for Transpose).
In this purpose, we will perform the following operations.

First step, we assess
$\bar{X} = \frac{1}{m} X \overrightarrow{1}$
Where, 1 is a column m-vector with all components equal to one.

Second step, we obtain

$\bar{X}\overrightarrow{1}^t = \frac{1}{m} X \overrightarrow{1}\,\overrightarrow{1}^t = \begin{bmatrix} \bar{x}^L & ... & \bar{x}^L \\ \bar{x}^U & ... & \bar{x}^U \end{bmatrix}$ and $X - \bar{X}\overrightarrow{1}^t = \begin{bmatrix} x_1^L - \bar{x}^L & ... & x_m^L - \bar{x}^L \\ x_1^U - \bar{x}^U & ... & x_m^U - \bar{x}^U \end{bmatrix}$
Last step, we get
$(m-1)S = \begin{bmatrix} x_1^L - \bar{x}^L & ... & x_m^L - \bar{x}^L \\ x_1^U - \bar{x}^U & ... & x_m^U - \bar{x}^U \end{bmatrix} \begin{bmatrix} x_1^L - \bar{x}^L & ... & x_m^L - \bar{x}^L \\ x_1^U - \bar{x}^U & ... & x_m^U - \bar{x}^U \end{bmatrix}^t = (X - \frac{1}{m} X \overrightarrow{1}\,\overrightarrow{1}^t)(X - \frac{1}{m} X \overrightarrow{1}\,\overrightarrow{1}^t)^t = X(I - \frac{1}{m}\overrightarrow{1}\,\overrightarrow{1}^t)X^t$
Because;
$(I - \frac{1}{m}\overrightarrow{1}\,\overrightarrow{1}^t)(I - \frac{1}{m}\overrightarrow{1}\,\overrightarrow{1}^t) = I - \frac{1}{m}\overrightarrow{1}\,\overrightarrow{1}^t - \frac{1}{m}\overrightarrow{1}\,\overrightarrow{1}^t + \frac{1}{m^2}\overrightarrow{1}\,\overrightarrow{1}^t\overrightarrow{1}\,\overrightarrow{1}^t = I - \frac{1}{m}\overrightarrow{1}\,\overrightarrow{1}^t$
Hence;

$S = \frac{1}{m-1} X(I - \frac{1}{m}\overrightarrow{1}\,\overrightarrow{1}^t)X^t$, and finally we get: $gvar(X) = tr(S) = \sum_{i=1}^{m} s_{ii}$

3.3. Example

Let $X = \begin{bmatrix} 0.2 & 1 & 0.6 \\ 1 & 0.2 & 0.3 \end{bmatrix}$ be a 2×3 matrix, where $x_1 = (0.2 \quad 1 \quad 0.6)$ and $x_2 = (1 \quad 0.2 \quad 0.3)$ are two vectors in R^3_+.
then $\bar{X} = \begin{bmatrix} 0.6 \\ 0.5 \end{bmatrix}$ and $X - \bar{X}\overrightarrow{1}^t = \begin{bmatrix} -0.4 & 0.4 & 0 \\ 0.5 & -0.3 & -0.2 \end{bmatrix}$. Therefore;
$S = \frac{1}{2}\begin{bmatrix} -0.4 & 0.4 & 0 \\ 0.5 & -0.3 & -0.2 \end{bmatrix}\begin{bmatrix} -0.4 & 0.4 & 0 \\ 0.5 & -0.3 & -0.2 \end{bmatrix}^t = \begin{bmatrix} 0.16 & -0.16 \\ -0.16 & 0.127 \end{bmatrix}$
$gvar(X) = tr(S) = 0.16 + 0.127 = 0.287$

4. Ranking of DMUs with interval data in DEA

With attentively to previous section classification, the units in E^{++} are better than of the units in E^+ and the units in E^+ are better than of the units E^-. Here, only DMUs belonging to E^+ are ranked. The method can be applied for ranking all DMUs belong to E^{++} as well.
Now consider DMU_p belong to E^+, then we use the following model for obtain the efficiency measure of $(\bar{x}_p, \bar{y}_p)$ [8].

$$\begin{array}{lll} Max & \bar{\theta}_p = \sum_{r=1}^{s} u_r \bar{y}_{rp} & \\ S.t. & \sum_{i=1}^{m} v_i \bar{x}_{ip} = 1 & \\ & \sum_{r=1}^{s} u_r y_{rj}^U - \sum_{i=1}^{m} v_i x_{ij}^L \leq 0, & j = 1, ..., n, j \neq p \\ & u_r \geq 0, & r = 1, ..., s \\ & v_i \geq 0, & i = 1, ..., m. \end{array} \qquad (4)$$

where $(\bar{x}_p, \bar{y}_p)$ is the mean (x_p^L, y_p^U) and (x_p^U, y_p^L), so $\bar{x}_p = \frac{x_p^L + x_p^U}{2}$, $\bar{y}_p = \frac{y_p^L + y_p^U}{2}$.
Now, let $G_p = \begin{bmatrix} x_p^L \\ x_p^U \end{bmatrix}$ and $H_p = \begin{bmatrix} y_p^L \\ y_p^U \end{bmatrix}$, then, we introduce Δ_p as a criteria for ranking DMU_p as bellow:

$$\Delta_p = \frac{\bar{\theta}_p^*}{\sqrt{1 + gvar(G_p) + gvar(H_p)}}$$

where, $\bar{\theta}_p^*$ is the optimal solution model (4), $gvar(G_p)$ and $gvar(H_p)$ are the generalized variance of input matrix and output matrix, respectively. This means that
$gvar(G_p) = tr(S_{G_p}) = \sum_{i=1}^{m} s_{ii} \quad gvar(H_p) = tr(S_{H_p}) = \sum_{r=1}^{m} s_{rr}$

Note, the presented criteria for ranking of interval data based upon two concepts, one to enlarge of the value efficiency and second relaxing in generalized variance.
It is seems that, it is logical that each interval unit as DMU_p which it have been higher efficiency in mean of (x_p^L, y_p^U) and (x_p^U, y_p^L), also lower generalized variance of $gvar(G_p)$ and $gvar(H_p)$, it will have high rank.

4.1. Lemma. If all of interval units be exact units, then $\Delta_{\bar{p}}, (p = \bar{p})$ is same efficient value from basic DEA models.

Proof: proof is obviously, because $gvar(G_p) = 0$, $gvar(H_p) = 0$.

5. Numerical example

To illustrate the above consider the interval data of Table1 (5 units with 2 inputs and 2 outputs) and the efficiency scores obtained by applying models (2) and (3).

Table 1. Interval Data, Efficiency Scores

DMU	Inputs		Outputs		Efficiency and Classification		
j	x_{1j}	x_{2j}	y_{1j}	y_{2j}	z_j^{L*}	z_J^{U*}	
1	[12,15]	[0.21,0.48]	[138,144]	[21,22]	0.224	1	E^+
2	[10,17]	[0.1,0.7]	[143,159]	[28,35]	0.227	1	E^+
3	[4,12]	[0.16,0.35]	[157,198]	[21,29]	0.823	1	E^+
4	[19,22]	[0.12,0.19]	[158,181]	[21,25]	0.445	0.907	E^-
5	[14,15]	[0.06,0.09]	[157,161]	[28,40]	1	1	E^{++}

With respect to Table 1, we have $E^+ = \{1, 2, 3\}$. For the next stage, consider Table 2, consist the center of gravity of DMUs in E^+ and their efficiency.

Table 2. The average of DMUs in E^+ and their efficiency

DMU	Inputs		Outputs		Efficiency
j	$\bar{x}_{1j}$	$\bar{x}_{2j}$	$\bar{y}_{1j}$	$\bar{y}_{2j}$	$\bar{\theta}_j^*$
1	13.5	0.345	141	21.5	0.306
2	13.5	0.4	151	31.5	0.391
3	8	0.255	177.5	25	1.395

Ranking of the units in E^+ according the presented approach, are given in Table 3.

Table 3. Ranking of DMUs in E^+

DMU	Generalized Variance		Δ_j
j	$gvar(G_j)$	$gvar(H_j)$	Δ_j
1	174.917	14286	0.002544
2	181.850	14300.5	0.003248
3	75.234	23528.5	0.009079

According to the values Δ_j in Table 3, the efficient units in E^+ are ranked in following order units 3, unit 2 and unit 1. Ranking of units is based on higher efficiency of mean and lower variance in their inputs and outputs. In other words, Δ_p is a function of efficiency value and variance so, each DMU which it has highest efficiency value and lowest variance, it is better than the other DMUs in E^+.

6. Conclusion

In this paper, we present criteria for ranking of interval data based upon two concepts, one increasing of efficiency of the mean two points (worst and best points) of evaluated decision making interval unit and second decreasing of variance in inputs and outputs it. It is seems that, ranking by this method be more superiority of other methods, say, method presented in Despotis et al [5], because they have ranked base on $z^{*U} - z^{*L}$ regarding two models (2) and (3), or in other words $1 - z^{*L}$, for each unit in E^+, so they only have respected to different value z^{*L} of 1. Meanwhile, presented method in this paper consider, firstly decreasing different z^{*U} and z^{*L}, that is, $(z^{*U} - z^{*L})$ based on the concept of variance inputs and outputs and furthermore, increasing efficiency of the mean two points (worst and best points) of evaluated decision making interval unit.

For example, in Fig. 1, the units A and B belong to set E^+ , so $z_A^{*L} = z_D^{*L}$, that is, rank of units A and B are equal $(1 - z_A^{*L} = 1 - z_D^{*L})$ with respect to Despotis's approach, meanwhile our proposed method ranks unit A greater than unit B. In finally, it is mentioned that this approach can be used for ranking of units in E^{++}.

References

[1] A .Charnes, W. W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, European Journal of Operational Research 2 (6) (1978)429$_4$44.

[2] R. D. Banker, A. Charnes, W. W. Cooper, Some model for estimating technical and scale inefficiencies in data envelopment analysis, Management

Science 30 (1984) 1078$_1$092.
[3] L. M. Seiford, J. Zhu, Stability regions for maintaining efficiency in data envelopment analysis, European Journal of Operational Research 108 (1998)127$_1$39.
[4] W. W. Cooper, K. S. Park, G. Yu, IDEA and AR_IDEA: models for dealing with imprecise data in DEA, Management Science 45 (1999) 597$_6$07.
[5] D. K. Despotis, Y. G. Smirlis, data envelopment analysis with imprecise data, European Journal of Operational Research 140 (2002)24$_3$6.
[6] G.R. Jahanshahloo,F. Hosseinzadeh Lotfi, M. Moradi, Sensitivity and stability analysis in DEA with interval data, Applied Mathematics and Computation, 156 (2004) 463$_4$77.
[7] R. A. Johnson, D. W. Wichern, Applied Multivariate Statistical Analysis, 1942
[8] P. Andersen, N.C. Petersen, A procedure for ranking efficient units in data envelopment analysis. Managemen science, 39, (1993)$1261 - 1264$.

Received: July 7, 2006