International Journal of Applied Operational Research
Vol. 5, No. 2, pp. 61-65, Spring 2015

Journal homepage: ijorlu.liau.ac.ir

A New Clustering Technic by the Preferences of the Objective in Data Envelopment Analysis

S. R. Mirshafaei*, F. Keshavarz Gildeh, A. Bina Vajargahi

Received: 17 August 2014 ; **Accepted:** 28 February 2015

Abstract The ways of placing decision making units (DMUs) in certain clusters are found as a subject in statistics, these ways usually are heuristic. The proposed clustering approach in this article considers preferences of DMUs. This study applies **Data Envelopment Analysis (DEA)**; DMUs are clustered by solving multi-objective linear problem (MOLP) and by considering preferences of each DMU at production of each output or consumption of each input. All of DMUs are partitioned into k clusters based on their reference units. The models can be classic models in DEA; in the article clustering DMUs are in the base of CCR envelopment model in which reference units play an important role. The idea of axial solutions is used for solving MOLPs that considers preferences of DMUs. As result, clustering DMUs is done by optimization solutions that are most preference solution at point of view of the decision maker. Moreover, a numerical example is presented and the approach is compared with two other different methods.

Keywords: Clustering approach- DEA- Axial solution-Preferences of functions.

1 Introduction

The analysis of placing decision making units (DMUs) in certain clusters is a subject in statistics. Some clustering algorithms are procedures that maximize total dissimilarity [1]. Recently, some papers have been proposed regarding clustering models based on DEA. DEA is a mathematical programming approach for measuring the relative efficiency of DMus, particular with multiple outputs and multiple inputs. Paper [5] has proposed a clustering approach using DEA and has employed piecewise production functions for data clustering. Each piecewise frontier is considered as one cluster that a specific DMU belongs to it. Then [5] has presented a comment on [3], that employs CCR envelopment form instead of mazrabi form that is applied in paper [3], and by λ –factors and reference sets provides a more simplifying approach than [5]. [3] proposed a clustering method which applies an integer linear programming model. The number and size distribution of groups are criteria for

* **Corresponding Author.** (✉)
E-mail: mirshafaei@yahoo.com (S. R. Mirshafaei)

S. R. Mirshafaei
MSc, Department of Mathematics, Roudsar & Amlash Branch, Islamic Azad University, Roudsar, Iran.

F. Keshavarz Gildeh
MSc, Department of Mathematics, Roudsar & Amlash Branch, Islamic Azad University, Roudsar, Iran.

A. Bina Vajargahi
MSc, Department of Mathematics, Roudsar & Amlash Branch, Islamic Azad University, Roudsar, Iran.

group membership. Results of clustering are the same in both of them, for the example [3]. In fact we attempt to find groups of analogous data, while we were not able to forecast these groups before. A clustering is valuable if DMUs of each group is similar.

In the present study by considering relation between input and output and preferences of DMUs [2], a new clustering approach for DMUs is proposed. M.A. Hinojosa and A.M.Marmol, 2011, introduced axial solutions and by using it solved MOLP. We apply that method for clustering DMUs, in which partial information is available. We can use axial solution when DMUs have the same objective with different preferences and the most preferred solutions are sought among all solutions. In results an optimal solution set is reduced.

The rest of this paper is as follows: the following section provides a review on axial solution, new clustering approach is given in section 3, and section 4 presents a numerical example. This study ends with conclusion.

2 Solving a multiple objective linear problem [2]

If partial information is considered, multiple objective problem is presented by a triplet (Ω, f, Λ).

Multi-objective linear problem is as follows:

$$\begin{aligned} &Max \quad f(\lambda) = [f_1(\lambda),...,f_r(\lambda),...,f_s(\lambda)] \\ &s.t. \\ &\qquad \lambda \in \Omega. \end{aligned} \tag{1}$$

Ω is in decision space and feasible set and functions are $f_r : R^n \to R^1; r = 1,2,...,s$, are the real value and linear continuously differentiable function. The information about preferences is presented in a set, $\Lambda \subseteq \Delta^{s-1}$, That, Δ^{s-1} is $\{\alpha \in R^S_+, \sum_{r=1}^{s} \alpha_r = 1\}$. It is called of information

Λ is the set of weights of DMUs. These weights are admissible of point of view of the decision maker.

A. Definition of Axial Solution

$P \in R^S_{++}$ (R^S_{++} is s-fold Cartesian product of R_{++}, set of all positive real numbers), given improvement axial, $\lambda \in \Omega$ the feasible solution, is an axial solution to the problem (Ω, f, Λ), if $f(\lambda) \succ_\Lambda t^* P$ where $t^* = \max\{t \in R;\ \exists \lambda \in \Omega, f(\lambda) \succ_\Lambda t^* P\}$. The set of axial is denoted by $A^P(\Omega, f, \Lambda)$.

If $P \in R^S_{++}$, be the improvement axis and $\alpha^1, \alpha^2, ..., \alpha^k$ be the extreme points of the information set $\Lambda \subseteq \Delta^{s-1}$, $\lambda^* \in A^P(\Omega, f, \Lambda)$ if be optimal solution to the following linear problem:

$$\begin{aligned} &Max \quad t \\ &s.t. \\ &\qquad \alpha^h f(\lambda) \ge t \alpha^h . P, \qquad h = 1,2,...,k, \\ &\qquad \lambda \in \Omega. \end{aligned} \tag{2}$$

and vice versa. In fact, problem (1), using axial solution transforms to problem (2). Assume DMU_j,j=1,2,…,n, products outputs denoted by y_{rj}, r=1,2,…,s, and consume m inputs denoted by x_{ij} *,i=1,2,…,m.*

DMU is Λ -efficient in problem 2, if $t^* = 1$

3 Clustering approach

At first CCR model can be transformed to MOLP. The reason is that preferences of DMUs to consume input or to produce output are accounted.

$$
\begin{aligned}
&Max \quad [\sum_{j=1}^{n}\lambda_j y_{1j},\dots,\sum_{j=1}^{n}\lambda_j y_{rj},\dots,\sum_{j=1}^{n}\lambda_j y_{sj}] \\
&s.t. \\
&\qquad \lambda \in \Omega_{jo} = \{\lambda \mid \sum_{j=1}^{n}\lambda_j x_{rj}, i = 1,2,\dots,m; \lambda_j \geq 0, j = 1,2,\dots,n\}
\end{aligned} \tag{3}
$$

Multi-objective linear model (3) is transformed to model (4) by using linear model (2).

$$
\begin{aligned}
&Max \quad t_o \\
&s.t. \\
&\qquad \frac{\sum_{j=1}^{n}\left(\sum_{r=1}^{s}\alpha_r^h y_{rj}\right)\lambda_j}{\sum_{r=1}^{s}\alpha_r^h y_{rj_o}} \geq t_o, \quad h = 1,2,\dots,k, \\
&\qquad \lambda \in \Omega_{j_o} = \left\{\lambda; \sum_{j=1}^{n}\lambda_j x_{rj} \leq x_{ij_o}, i = 1,2,\dots,m;\ \lambda_j \geq 0, j = 1,2,\dots,n\right\}
\end{aligned} \tag{4}
$$

By solving model (4), DMUs that have the same reference set place in a cluster. But, now $\lambda's$ are factors that depend on to preferences. Also when each DMU has different preferences, model (4) changes to model (5). In model (4), α_r^h is same for all of DMUs, but in model (5) is different for each DMU.

$$
\begin{aligned}
&Max \quad t_o \\
&s.t. \\
&\qquad \frac{\sum_{j=1}^{n}\left(\sum_{r=1}^{s}\alpha_{r_j}^h y_{rj}\right)\lambda_j}{\sum_{r=1}^{s}\alpha_{r_{j_o}}^h y_{rj_o}} \geq t_o, \quad h = 1,2,\dots,k, \\
&\qquad \lambda \in \Omega_{j_o} = \left\{\lambda; \sum_{j=1}^{n}\lambda_j x_{ij} \leq x_{ij_o}, i = 1,2,\dots,m;\ \lambda_j \geq 0, j = 1,2,\dots,n\right\}
\end{aligned} \tag{5}
$$

4 Numerical example

Also a numerical example is examined. This example has twenty DMUs and each DMU produces one output by consuming two inputs. Approaches of [3] and [4] have the same clusters for this example. Both have three clusters and clusters are similar. The example has an output, therefore it is necessary that model is changed to the input –oriented one.

$$
\begin{aligned}
&Max \quad t_o \\
&s.t. \\
&\qquad \frac{\sum_{j=1}^{n}\left(\sum_{i=1}^{m} -\alpha_{i_j}^{h} x_{ij}\right)\lambda_j}{\sum_{i=1}^{m} \alpha_{i_{j_o}}^{h} x_{ij_o}} \ge t_o, \qquad h = 1,2,\dots,k, \\
&\qquad \lambda \in \Omega_{j_o} = \left\{\lambda; -\sum_{j=1}^{n} \lambda_j y_{rj} \le -y_{rj_o}, r = 1,2,\dots,s;\ \lambda_j \ge 0, j = 1,\dots,n\right\}
\end{aligned}
\tag{6}
$$

Model (6) results in different clusters than [3] and [5].
Set of preferences of DMUs are given as follows:

$\Lambda_1 = \Lambda_2 = \Lambda_3 = \{\alpha = (\alpha_1, \alpha_2) \mid \alpha_1 \ge \frac{1}{2}\alpha_2\}$

$\Lambda_4 = \Lambda_5 = \Lambda_6 = \Lambda_7 = \{\alpha = (\alpha_1, \alpha_2) \mid \frac{6}{10}\alpha_2 \ge \alpha_1, \alpha_1 \le \frac{8}{10}\alpha_2\}$

$\Lambda_8 = \Lambda_9 = \Lambda_{10} = \Lambda_{11} = \{\alpha = (\alpha_1, \alpha_2) \mid \frac{6}{10}\alpha_2 \ge \alpha_1\}$

$\Lambda_{12} = \Lambda_{13} = \{\alpha = (\alpha_1, \alpha_2) \mid \alpha_2 = \alpha_1\}$

$\Lambda_{14} = \Lambda_{15} = \Lambda_{16} = \{\alpha = (\alpha_1, \alpha_2) \mid \alpha_1 \ge \frac{7}{10}\alpha_2\}$

$\Lambda_{17} = \Lambda_{18} = \{\alpha = (\alpha_1, \alpha_2) \mid \alpha_1 \ge \frac{7}{10}\alpha_2\}$

$\Lambda_{19} = \Lambda_{20} = \{\alpha = (\alpha_1, \alpha_2) \mid \frac{3}{10}\alpha_1 \le \alpha_2 \le \frac{9}{10}\alpha_1\}$

For instance $\alpha_1 \ge \frac{1}{2}\alpha_2$ means, preference of output 1 is more than $\frac{1}{2}$ preference of output2 about DMU1, DMU2 and DMU3.
The result clusters are presented as follows:

Table 1 The clusters

Cluster1	DMU1
Cluster2	DMU2
Cluster3	DMU3,DMU4,DMU5,DMU6,DMU7,DMU8,DMU9, DMU10,DMU11,DMU14, DMU15,DMU16
Cluster4	DMU12,DMU13
Cluster5	DMU17,DMU18,DMU19, DMU20

5 Conclusions

This paper provided a clustering approach for DMUs with preferences of objective functions. A set of information is presented; it involves partial information about preferences of DMUs in consuming each input or producing each output. Different preferences were considered for each objective function of each DMU. CCR envelopment form is applied, and then transformed to multiple objective linear problem, so by improvement axis that is introduced in M.A. Hinojosa and A.M.Marmol reference (2011), MOLP changes to LP. Reference units of each DMU are attained by solving this LP. These reference units specified the clusters. All DMU with same reference units belong to a cluster. Also a numerical example is presented and the results compared to in two ways.

References

1. Duda, R. O., Hart, P. E., (1973). Pattern classification and sense analysis. Wiley, New York.
2. Hinojosa, M. A., Marmol, A. M., (2011). Axial solutions for multiple objective linear problems. An application to target setting in DEA models with preferences," *Omega*, vol. 39, pp.159-168.
3. Po, R. W., Gush, Y. Y., Yang, M. S., (2009). A new clustering approach using data envelopment analysis. *European Journal of Operational Research,* vol. 199 , pp. 276-284.
4. Jessop, A. A., (2009). Multi-criteria block model for performance assessment. Omega 37, pp. 204–214**.**
5. Kruger, J. J., (2010). Comment on "A new clustering approach using data envelopment analysis.European Journal of Operational Research, vol. 206, pp. 269-270.