Available online at http://ijim.srbiau.ac.ir/
Int. J. Industrial Mathematics (ISSN 2008-5621)
Vol. 7, No. 1, 2015 Article ID IJIM-00564, 6 pages
Research Article

Project selection with limited resources in data envelopment analysis

M. Jahantighi [*†], Z. Moghaddas [‡], M. Vaez-ghasemi [§]

Abstract

In this paper allocating a fixed resource for producing finite projects in order to obtaining a desired level of efficiency will be discussed. Note that it is assumed that a vector of limited sources is at hand. This vector of resources can be contained human resource, budget, equipment, and facilities. In any firm there exist different suggestions from subunits for running a new projects in line with the organization's objectives. Implementation of all the suggested projects need high level of resources. In accordance to this fact that resources are limited thus it is not possible to run all of the projects. Thus, selecting high quality projects or those with high efficiency is more desirable for implementation. In this paper a method for selecting projects will be proposed which has high performance.

Keywords : Data envelopment analysis; Efficiency; Error analysis.

1 Introduction

In daily routines it is of great importance to allocate budget in a proper way. As it seems to be necessary to perform this task in an ideal manner this paper mainly deals with budget allocation problem. For this purpose different methods presents in literature, some of them mentioned in this paper.
Dutra et al. [4] in their paper provided an economic probabilistic model for project selection. They presented a model three steps. For showing the applicability of the model their considered a portfolio of investment projects at a power distribution company. Tavana et al. [6] presented mathematical model based on the data envelopment analysis technique with ambiguity. They considered multi-objective fuzzy linear programming for modelling the vagueness of the objective functions and modelled the ambiguity of the input and output data considering the fuzzy sets and a new a-cut based method. The authors considered a high-technology project selection at NASA in order to show the applicability of the proposed models. The main purpose of this research is to propose an effective hybrid process for evaluating district development directions. They used the fuzzy Delphi method (FDM), the interpretive structural modelling (ISM), and the analytic network process (ANP) with benefits, opportunities, costs, and risks (BOCR) are integrated to construct a project selection model. A case study in in Taiwan is considered for evaluation of the provided model. As the complexity of the project selection problem is due to the high number of projects thus Khalili-Damghani et al [5] presented a hybrid fuzzy rule-based multi-objective framework for sustainable project portfolio selection. They noted that the proposed framework simultaneously considers the accuracy maximization and the

[*]Corresponding author. Mjahantighi@yahoo.com
[†]Department of Mathematics, Islamic Azad University, Zahedan Branch, Zahedan, Iran.
[‡]Department of Electrical, Computer and Biomedical Engineering, Islamic Azad University, Qazvin Branch, Qazvin, Iran.
[§]Department of Mathematics, Islamic Azad University, Rasht Branch, Rasht, Iran.

complexity minimization objectives As stated in literature DEA is a mathematical programming technique for performance assessment of a set of homogeneous decision making units (DMUs). The model of this mathematical technique are on basis of some fundamental axioms which can consider different types of production technology. As project selection is a common task in daily routines it is of importance to formulate a model for it in a proper way. Thus some models, based on DEA technique, introduced for selecting the proper project from among the presented projects from different organizations and a model for maximizing the total output of the system.
In the following section three mathematical programming problems are proposed for project selection each of which considered this aim from different aspects.

2 Preliminaries

Data envelopment analysis is a non-parametric approach for measuring relative efficiency of a set of decision making units. This technique is based upon mathematical programming. One of the major advantages in DEA analysis is that it is a linear programming method deals with multiple inputs (X) and multiple outputs (Y) with no pre-assumption about data. Noted that in this method it is assumed that input and output vectors are considered to be seropositive. In this technique in accordance to the observations an envelope constructed which surrounds all of them and this leads to generate a frontier which is called as production frontier. It also makes it possible to consider different types of technologies.
The most general way to characterize production technology is production possibility set T, which is defined with a set of semipositive (x, y) as:

$$T = \{(x,y)| \quad x \geq \sum_{j=1}^{n} \lambda_j x_j, \quad y \leq \sum_{j=1}^{n} \lambda_j y_j, \quad \lambda_j \geq 0, \quad j = 1,...,n\}$$

Those DMUs located onto this frontier is called best practice which performs efficiently, and those which are far away from this frontier is called inefficient. One important concerns of managers is policy making and guiding these inefficient units. One great key feature of DEA technique is that it also can be considered as benchmarks tool, as well. According to the constructed production frontier with efficiently performed units and the fact that relative efficiency of each unit derived from the comparison process to this frontier, this frontier can be considered for benchmarks. Thus for inefficient unit a proper unit, located onto this frontier, can be accounted for as a suitable and achievable target. Charnes et al. [3] presented CCR model for relative efficiency assessment of a set of homogeneous DMUs. In this model constant returns to scale form of technology is assumed. This model is written in input orientation which means inputs are contracted, while keeping the projected point at least at the same output levels. In order to efficiency evaluation and benchmarking one should performed two-stage optimization procedure. Which means at first, the optimal value of the objective function of the CCR model is obtained. This scaler shows the possible uttermost radial increment for inputs. Note that simultaneously the shortfall output or excess input usage also need to be considered. Thus for finding Pareto- efficient targets, the second stage should be solved. This stage seeks for the nonradial improvements for inputs and outputs by maximizing the sum of input and output slacks. Consider this two-stage model as following.

$$\begin{aligned} \min \quad & \theta - \varepsilon(1s^- + 1s^+) \\ s.t. \quad & \sum_{j=1}^{n} \lambda_j x_{ij} + s^- = \theta x_{io}, \quad \forall i, \\ & \sum_{j=1}^{n} \lambda_j y_{rj} - s^+ = y_{ro}, \quad \forall r, \\ & \lambda_j \geq 0, \quad \forall j. \end{aligned} \tag{2.1}$$

The same idea can be performed for output orientation which explores higher outputs while keeping the projected point at least at the same input levels. As Banker et al. [1] presented in a paper, variable returns to scale technology form of the productu=ion function can also be considered. In doing so, after performing the analysis constraint $\sum_{j=1}^{n} \lambda_j = 1$ needs to be added to model (2.1). Replacing this constraint with

$\sum_{j=1}^{n}\lambda_j \leq 1$ or $\sum_{j=1}^{n}\lambda_j \geq 1$ shows non-increasing and non-decreasing returns to scale technologies, respectively. As mentioned formerly, DEA models are linear programming problem thus it is possible to consider a dual problem for them.
The dual of the above model which is called multiplier form in input orientation is as follows. As stated in literature

$$\sum_{r=1}^{s} u_r y_{rD} - \sum_{i=1}^{m} v_i x_{iD} = 0, D \in \{1, ...n\}$$

forms a supporting hyperplane passing through efficient DMU_D.

$$\begin{aligned} \max \quad & \sum_{r=1}^{s} u_r y_{ro} \\ s.t. \quad & \sum_{i=1}^{m} v_i x_{io} = 1, \\ & \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \quad \forall j, \\ & u_r \geq 0, \quad v_i \geq 0, \quad \forall r, \quad \forall i. \end{aligned} \tag{2.2}$$

In the above- mentioned model v and u are the input and output weight vectors.
Both of the aforesaid models are in input orientation where the input reduction is due to maximized. The above models can be written in output orientation where the output shortfall is due to minimized.

3 Formulating project selection with limited resources

In this section different method of project selection with limited resources will be discussed. At first, a method is introduced which considerers all the possible selections along with the limited resources the project with high performance will be selected.
Let $b = (b_1, ...b_m)^t$ be the vector of limited resources at managers hand. Note that $b_i, b \neq 0, b \geq 0$ shows the extend of the i th resource for allocation and running projects.
Moreover, assume that $Z_1, ..., Z_n$ are the proposed projects. Each of these projects uses portions of the resources b and produces the output Y. If $Z_j = (X_j, Y_j)^t$ be the coordination of the j th project this means that X_j is proposed from the resources b for producing Y_j and $X_j \geq 0, X_j \neq 0, Y_j \geq 0, Y_j \neq 0$.

Note that in accordance with the variety of the proposed projects we should have:

$$\sum_{j=1}^{n} X_j \nleq b$$

The aim is to find a bundle of projects like j that;

$$P_1 : \sum_{j=1}^{n} X_j \leq b$$

$$P_2 : X_j \nleq b - \sum_{j=1}^{n} X_j, \quad j \notin J$$

$$P_3 : \{(X_j, Y_j | j \in J\} has \quad the \quad highest \quad rank.$$

In this case it is assumed that all the cases of selecting different projects from source of b, considering P_1 and P_2 can be investigated and calculated.
As an example consider:

$$b = (15, 12)^t, Z_1 = (5, 3, 7)^t,$$
$$Z_2 = (4, 7, 3)^t, Z_3(3, 5, 6)^t, Z_4(4, 1, 3)^t$$

and

$$C_1 = \{Z_1, Z_2, Z_4\}, C_2 = \{Z_1, Z_3, Z_4\},$$
$$C_3 = \{Z_2, Z_3\}$$

All the cases are these three state.
As an instance in the first case, selecting Z_1, Z_2, Z_3 we have the following relations which satisfy the conditions P_1, P_2, respectively:

$$\sum_{j \in C_i} X_j \leq b, \forall i$$

$$X_j \nleq b - \sum_{j \in C_i} X_j, \quad i \notin C_i$$

Like the above all the other cases can be verified. Let $C_1, ...C_k$ are all the possible bundles of the selections, according to the above explanation. This means:

$$\sum_{j \in C_i} X_j \leq b, \forall i$$

$$X_j \nleq \ b - \sum_{j \in C_i} X_j, \quad i \notin \ C_i$$

The aim is to selecting a set of C_t from among $C_1, ...C_k$ that have the highest performance. In doing so the following formulating ca be proposed.
Assume that efficiency of each project is calculated in its own bundle. In doing this, at first, considering all of the members of each set like C_i corresponding PPS is being constructed and then all the members of C_i are being evaluated in this PPS and their efficiencies calculated. Let $\theta^i = (\theta^i_j : j \in C_j)$. In this case each vector of θ^i is the criterion of selecting the bundle of C_i. In doing so the following suggestions can be presented.

$$1 - \alpha^i = \frac{\sum_{j \in C_i} \theta^i_j}{|C_i|}$$

$$\alpha^i = (\prod j \in i\theta^i_j)^{\frac{i}{|C_i|}}$$

$$\alpha^i = Min\{\theta^i_j : j \in C_i\}$$

3.1 Arithmetic Mean of efficiency

In this case each bundle which has the highest average efficiency is selected. This bundle, actually, contains those projects with high efficiency regarding to its own efficiency. For this aim, it is needed that projects have homogenous and similar behavior. In other words, in each selected bundle, there should not exist projects with efficiency much more than other projects. This can be happened by solving a mathematical programming problem. Thus, let y_j is a binary variable corresponding to project j.

$$y_j = \begin{cases} 1, & j \ \ th \ \ project \ \ is \ \ selected \\ 0, & otherwise \end{cases}$$

Let $\frac{Uy_j}{Vx_j}$ be the efficiency of project j, in this case corresponding problem for maximizing average of efficiency is as follows:

$$\begin{aligned} \max \quad & \frac{\sum_{j=1}^{n} y_j(Uy_j/Vx_j)}{\sum_{j=1}^{n} y_j} \\ s.t. \quad & y_j \frac{Uy_j}{Vx_j} \leq 1, \quad \forall j, \\ & U \geq 1\varepsilon, V \geq 1\varepsilon, \\ & \sum_{j=1}^{n} y_j x_j \leq b, \\ & (1-y_j)x_{ij} \geq (b - \sum_{j=1}^{n} y_j x_{ij}) \\ & +\varepsilon - (1-\delta_{ij})M - M' y_j \\ & \delta_{ij} \geq 1, \quad \forall j, \forall i, \\ & y_j \in \{0,1\}, \quad \forall j. \end{aligned} \tag{3.3}$$

where M, M' are the big M with the condition that $M' > M$. In the above model (3.3), there exist two bundles of binary variable. The first bundle, y_j, shows project selection according to the P_1 which satisfy that the selected projects can be performed utilizing from the resource vector of b. The second bundle of binary variables are δ_{ij} which satisfies P_2 that means non of the projects can be performed from remaining of the resources.

3.2 Geometrical average of efficiency

In this case the criterion is the geometrical average efficiency of the selected projects. The greater geometrical average is the criterion is better. In doing so, the following model is proposed:

$$\begin{aligned} \max \quad & [\textstyle\prod_{j=1}^{n} y_j(Uy_j/Vx_j)] \\ s.t. \quad & y_j \frac{Uy_j}{Vx_j} \leq 1, \quad \forall j, \\ & U \geq 1\varepsilon, V \geq 1\varepsilon, \\ & \sum_{j=1}^{n} y_j x_j \leq b, \\ & (1-y_j)x_{ij} \geq (b - \sum_{j=1}^{n} y_j x_{ij}) \\ & +\varepsilon - (1-\delta_{ij})M - M' y_j \\ & \delta_{ij} \geq 1, \quad \forall j, \forall i, \\ & y_j \in \{0,1\}, \quad \forall j. \end{aligned} \tag{3.4}$$

It is clear that as $f(x) = Lnx$ is a increasing function, thus the following expression can be replaced for the objective function of (3.4).

$$Max \frac{1}{\sum_{j=1}^{n} y_j} \left(\sum_{j=1}^{n} Ln(y_j + \varepsilon) + \sum_{j=1}^{n} Ln \frac{Uy_j}{Vx_j} \right)$$

Adding ε to the Y prevents Ln to equal zero values.

3.3 Maximum the Minimum of the efficiency

In this case a bundle is selected which has the maximum of the minimum efficiency. In other word, norm chebyshev the efficiency vector is maximized. Thus, the proposed efficiency is as follows:

$$\begin{array}{ll} \max & (Min\{y_j \frac{Uy_j}{Vx_j} : j = 1, ..., n\}) \\ s.t. & y_j \frac{Uy_j}{Vx_j} \leq 1, \quad \forall j, \\ & U \geq 1\varepsilon, V \geq 1\varepsilon, \\ & \sum_{j=1}^{n} y_j x_j \leq b, \\ & (1 - y_j)x_{ij} \geq (b - \sum_{j=1}^{n} y_j x_{ij}) \\ & +\varepsilon - (1 - \delta_{ij})M - M' y_j \\ & \delta_{ij} \geq 1, \quad \forall j, \forall i, \\ & y_j \in \{0, 1\}, \quad \forall j. \end{array} \tag{3.5}$$

This problem, (3.5), considering the following variable transformation,

$$w = Min\{y_j \frac{Uy_j}{Vx_j} : j = 1, ..., n\}$$

can be written as follows:

$$\begin{array}{ll} \max & w \\ s.t. & w \leq y_j \frac{Uy_j}{Vx_j}, \forall j, \\ & y_j \frac{Uy_j}{Vx_j} \leq 1, \quad \forall j, \\ & U \geq 1\varepsilon, V \geq 1\varepsilon, \\ & \sum_{j=1}^{n} y_j x_j \leq b, \\ & (1 - y_j)x_{ij} \geq (b - \sum_{j=1}^{n} y_j x_{ij}) \\ & +\varepsilon - (1 - \delta_{ij})M - M' y_j \\ & \delta_{ij} \geq 1, \quad \forall j, \forall i, \\ & y_j \in \{0, 1\}, \quad \forall j. \end{array} \tag{3.6}$$

Note that all the proposed problems, (3.3), (3.4), (3.5) and (3.6), are mixed nonlinear programming problems. Considering each of these problems it is possible to select projects bundles in accordance to the selected aims.

References

[1] R. D. Banker, A. Charnes, *Some models for estimating technical and scale efficiencies in data envelopment analysis*, Management Sience 30 (1984) 1078-1092.

[2] R. D. Banker, W. W. Cooper, M. Lawrence, R. Seiford, M. Thrall, J. Zhu, *Returns to scale in different DEA models*, European Journal of Operational Research 154 (2004) 345-362.

[3] A. Charnes, W.W. Cooper, E. Rhodes, *Measureing the efficiency of decision making units*, European journal of operational reaserch 2 (1978) 429-444.

[4] C. Costa Dutra, J. Luis Duarte Ribeiro, M. Monteiro de Carvalho, *An economic probabilistic model for project selection and prioritization*, International Journal of Project Management 32 (2014) 1042-1055.

[5] K. Khalili-Damghani, S. Sadi-Nezhad, F. Hosseinzadeh Lotfi, M. Tavana, *A hybrid fuzzy rule-based multi-criteria framework for sustainable project portfolio selection*, Information Sciences 220 (2013) 442-462.

[6] M. Tavana, K. Khalili-Damghani, S. Sadi-Nezhad, *A fuzzy group data envelopment analysis model for high-technology project selection: A case study at NASA*, Computers & Industrial Engineering, 66 (2013) 10-23.

[7] M. Wang, H.I. Lee, L. Peng, Z. Wu, *An integrated decision making model for district revitalization and regeneration project selection*, Decision Support Systems 54 (2013) 1092-1103.

Mohammad Jahantighi is a candidate for PhD at the Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran. His research interests include data envelopment analysis and operation research.

Zohreh Moghaddas is an assistant professor at the Department of Mathematics, Qazvin Branch, Islamic Azad University, Qazvin, Iran. Her research interests include operation research and data envelopment analysis.

Mohsen Vaez-Ghasemi is an assistant professor at the Department of Mathematics, Rasht Branch, Islamic Azad University, Rasht, Iran. His research interests include operation research and data envelopment analysis.