Iranaian J. Numerical Analysis and Optimization
Vol. 3, No. 1, (2013), 45-54

An alternative 2-phase method for evaluating of DMUs using DEA

Mohammadreza Alirezaee

Abstract

Computationally, selection of a proper numerical value for infinitesimal non Archimedean epsilon in DEA models has some difficulties. Although there are several algorithms for selecting the proper non-Archimedean epsilon, it is important to introduce methods in order to calculate the efficiency of DMUs without using epsilon. One of these methods is a two-phase method, which obtains the efficiency of each DMU through solving two LPs, which the second LP is depended to the first. This paper proposes a method, which is able to compute the efficiency of DMUs by two LPs, which are not depended to each other and computationally can solve in a parallel computation. The major of this method is to find two references for each unit and combine them to obtain actual reference.

Keywords: Data Envelopment Analysis (DEA); Decision Making Units (DMUs); Non-Archimedean; Two-phase method; Reference point.

1 Introduction

Since the first mathematical model of Data Envelopment Analysis (DEA) by Charnes et al. (1978), (known as CCR), and Banker et al. (1984) (known as BCC), there have been many theoretical and applied researches in DEA (Emrouznejad ea al., 2008). In 1979 the first version of DEA model has been updated by adding the non-Archimedean ε as a lower bound for weights of inputs and outputs of the corresponding DMUs(Charnes et al. 1979). Different methods have been proposed for computing a suitable value for ε. Ali and Seiford (1993) introduce a method to find an acceptable value for ε. Mehrabian et al. (2000) modify this method and propose an LP to select a proper value for ε. Up to now, some researchers have published methods and discussions about the non-Archimedean ε such as Amin and Toloo (2004), MirHassani and Alirezaee (2004), Alirezaee and Khalili (2006).

As the first and most important substitution method for epsilon-based DEA solving methods, Cooper et al. (1999) introduce the two-phase method

Mohammadreza Alirezaee
School of Mathematics, Iran University of Sciences and Technology, Tehran, Iran. e-mail: mralirez@iust.ac.ir

for evaluating the efficiency in DEA without using ε. In this method , two LPs must be solved respectively for each DMU.

In this paper, we introduce an alternative algorithm for evaluating the efficiency of DMUs without using ε. In the proposed method, firstly, we find the references of DMUs and then inherited the references in a way that we can find out if the reference point of the unit is located on the weak frontier. The most advantage of this method is reducing the overall running time, because we can use parallel computation for our independent LPs in the algorithm.

2 Classification of DMUs

In DEA a set of DMUs are partitioned into two main classes: efficient and inefficient. The efficient units make the efficiency frontier. Figure 1 shows the classification of DMUs, based on the position of their reference on the efficiency frontier (Charnes et al., 1991).

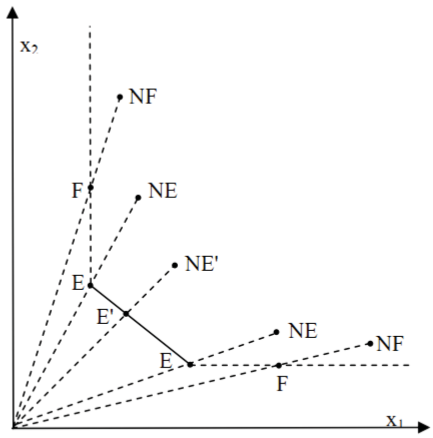

Fig. 1: Classifying DMUs.

In this classification, E and $E^{'}$ are efficient and NE and $NE^{'}$ are inefficient DMUs. In addition, F and NF are weakly efficient and weakly inefficient, respectively.

Suppose that there are n Decision Making Units (DMUs) each consumes m inputs to produce s outputs. Let $x_j = (x_{1j}, x_{2j}, \ldots, x_{mj})$ and $y_j = (y_{1j}, y_{2j}, \ldots, y_{sj})$ are input vector and output vector of DMUj $(j = 1, \ldots, n)$, respectively. Hence, the CCR model corresponding to DMUp is as follow:

CCR Model:

$$\min z = \theta - \varepsilon(\Sigma_{i=1}^{m}(s_i^-) + \Sigma_{r=1}^{s}(s_r^+))$$
$$s.t.$$
$$x_{ip}\theta - s_i^- - \Sigma_{j=1}^{n} x_{ij}\lambda_j = 0, \forall i$$
$$- s_r^+ + \Sigma_{j=1}^{n} y_{rj}\lambda_j = y_{rp}, \forall r$$
$$\lambda_j, s_i^-, s_r^+ \geq 0, \forall i, r, j$$

To introduce the new method, all of the values of variables for all DMUs are needed. So consider the following integrated model.

CCR$_P$ Model:

$$\min z = \theta_p - \varepsilon(\Sigma_{i=1}^{m} s_{ip}^- + \Sigma_{r=1}^{s} s_{rp}^+)$$
$$s.t.$$
$$x_{ip}\theta_p - s_{ip}^- - \Sigma_{j=1}^{n} x_{ij}\lambda_{jp} = 0, \quad \forall i$$
$$- s_{rp}^+ + \Sigma_{j=1}^{n} y_{rj}\lambda_{jp} = y_{rp}, \quad \forall r$$
$$\lambda_{jp}, s_{ip}^-, s_{rp}^+ \geq 0, \quad \forall i, r, j$$

In the above model, for each variable an index has been added. To simplify the notation, let $u_j = (y_j, x_j)$ be DMUj. Hence one can present members of productivity possibility set as $u = (y, -x)$. Clearly, $u_j \in PPS$,$(j = 1, \ldots, n)$. Let $(\theta_p^*, \lambda_p^*, s_p^{+^*}, s_p^{-^*})$ is an optimal solution of CCR_p, then efficiency of u_p is equal to θ_p^*. We also denote the efficiency of u by θ_u^*. The dominate space and reference of u_p are denoted by DS_p and $u(p)$, respectively, and the set of reference indices of u_p is denoted by E(p).

$$DS_p = \{u \in PPS | u \geq u_p\}, u(p) = \Sigma_{j=1}^{n} l_{jp}^* u_j, E(p) = \{j : \lambda_{jp}^* > 0\}$$

It is clear that $u(p) \in DS_p$ and $u(p) = \Sigma_{j \in E(p)} \lambda_{jp}^* u_j$. These concepts are illustrated in Figure 2.
The shading pattern in this figure and other figure in the paper represents the dominant space of u_p, so every point in that region has inputs less than or equal and outputs more than or equal to u_p. And dashed line in the frontier is weakly frontier.

Therefore, we have:

1. u_p is efficient $\Leftrightarrow \theta_p^* = 1, \Sigma_{i=1}^{m} s_{ip}^{-^*} + \Sigma_{r=1}^{s} s_{rp}^{+^*} = 0$.
2. u_p is weak efficient $\Leftrightarrow \theta_p^* = 1, \Sigma_{i=1}^{m} s_{ip}^{-^*} + \Sigma_{r=1}^{s} s_{rp}^{+^*} > 0$.
3. u_p is inefficient $\Leftrightarrow \theta_p^* < 1, \Sigma_{i=1}^{m} s_{iu(p)}^{-^*} + \Sigma_{r=1}^{s} s_{ru(p)}^{+^*} = 0$.
4. u_p is weak inefficient $\Leftrightarrow \theta_p^* < 1, \Sigma_{i=1}^{m} s_{iu(p)}^{-^*} + \Sigma_{r=1}^{s} s_{ru(p)}^{+^*} > 0$.

Notice that $u(p)$ lies on the efficiency frontier, and so $\theta_{u(p)}^* = 1$.

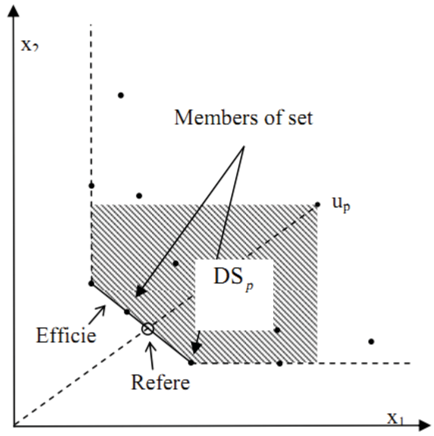

Fig. 2: Efficient frontier, dominate space, reference point and set of reference index.

3 New Method

As explained in previous section, for classifying the DMUs, the values of θ and slacks of reference point must be used. For each u_p, two reference points are determined, the first reference belongs to DS_p which minimize θ_p. The second reference belongs to DS_p which maximize $\Sigma_{i=1}^{m} s_{ip}^{-*} + \Sigma_{r=1}^{s} s_{rp}^{+*}$. Consider the following models:

Model P1:

$$\begin{aligned} \min z = \theta_p s.t.& \\ & x_{ip}\theta_p - \Sigma_{j=1}^{n} x_{ij}\lambda_{jp}^{l} \geq 0, \quad \forall i \\ & \Sigma_{j=1}^{n} y_{rj}\lambda_{jp}^{l} \geq y_{rp}, \quad \forall r \\ & \lambda_{jp}^{l} \geq 0, \quad \forall j \end{aligned}$$

Model P2:

$$\begin{aligned} \min z = -(\Sigma_{i=1}^{m} s_{ip}^{-} + \Sigma_{r=1}^{s} s_{rp}^{+})& \\ s.t.& \\ & s_{ip}^{-} + \Sigma_{j=1}^{n} x_{ij}\lambda_{jp}^{2} = x_{ip}, \quad \forall i \\ & - s_{rp}^{+} + \Sigma_{j=1}^{n} y_{ij}\lambda_{jp}^{2} = y_{rp}, \quad \forall i \\ & \lambda_{jp}^{2}, s_{ip}^{-}, s_{rp}^{+} \geq 0, \quad \forall j, i, r \end{aligned}$$

The first reference for u_p is determined by Model P1 as $u_1(P) = \Sigma_{j=1}^{n} \lambda_{jp}^{l^*} u_j$. Similarly, the second reference for u_p is determined by the model p2 as $u_2(p) = \Sigma_{j=1}^{n} \lambda_{jp}^{2^*} u_j$. These references are illustrated in Figure 3:

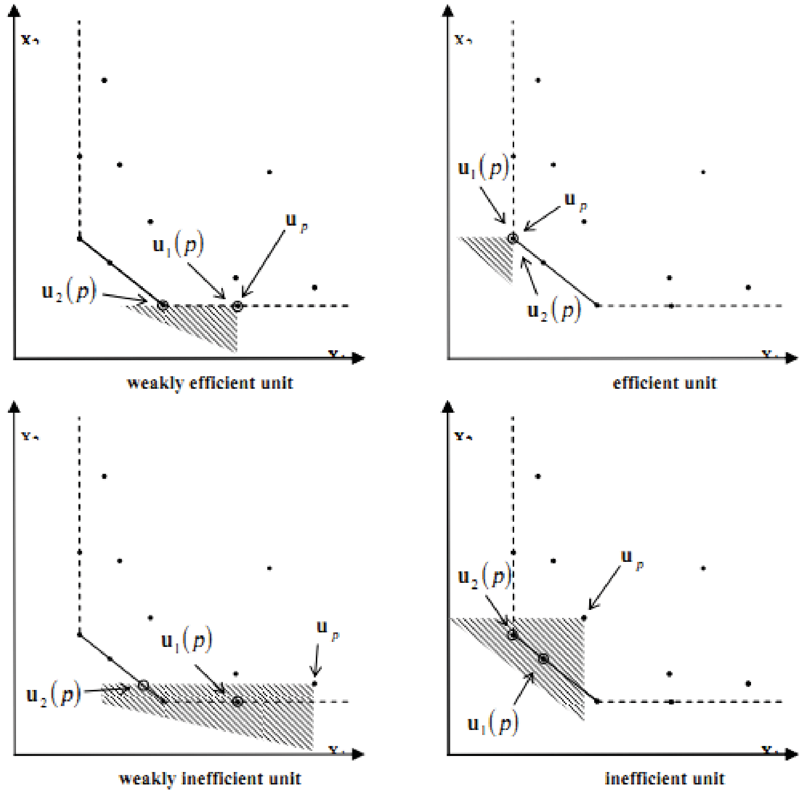

Fig. 3: u_p and its first and second references.

$u_1(p)$ is not always the same as the reference point which is defined as $(\theta^* x_j, y_j)$ for (x_j, y_j). If u_p be an efficient or inefficient unit then its reference point and $u_1(p)$ are the same and if it is a weakly efficient or inefficient unit then probably its reference probably its reference point and $u_1(p)$ are not equal to each other. For example, in the Figure 4(a), $u_1(p)$ and reference point of u_p are the same, but if the weakly efficient unit $u_1(p)$ is removed as illustrated in the Figure 4(b), the two definitions become different. In this case the reference point can be improved to the $u_1(p)$, therefore we must improve the reference point, too. Reference of each unit must be on the (strong) frontier and not on the weakly frontier. Weakly frontier is shown by dashed line in the figures in the paper.

Definition 3.1. The revised reference for u_p is defined as bellow:

$$\hat{u}(p) = \Sigma_{k=1}^{n}(\Sigma_{j=1}^{n}\lambda_{jp}^{1^*}\lambda_{kj}^{2^*})$$

If u_p be an efficient or an inefficient unit then the revised reference is $u_1(p)$, but if it is a weekly efficient or inefficient unit then the reference for u_p on the frontier is moved by $\hat{u}(p)$ using the weights belonging to E(P). Revised reference is simply the second reference of the first reference of the *pth* unit. We introduce a simple algorithm to identify the revised reference of u_p and describe the idea of revised reference. It woks as follows:

1. Find the first reference of u_p. There are one or more units that create the first reference of u_p. They are efficient or weakly efficient units.
2. For each unit that participating the construction of the first reference of u_p, find the second reference. For each of them, there are one or more units that are efficient (and not weakly efficient).
3. The second reference of the first reference of u_p is the actual reference of it, so calculate values of the new variable that create the second reference of the first reference of u_p. The revised reference is on the efficient frontier and is the actual reference of u_p.
   This algorithm computed the revised reference of u_p, in other word we move of u_p to $u_1(p)$ and then move to $u_2(u_1(p))$, as illustrated in Figure 4.

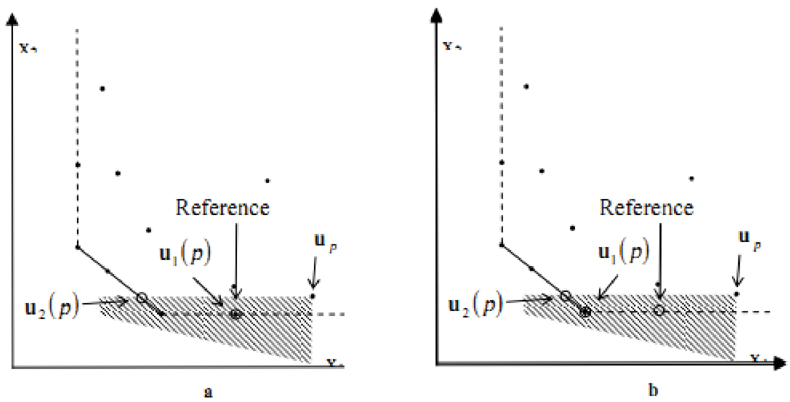

Fig. 4: Reference point and $u_1(p)$ may be different.

These concepts lead to the next theorems 1 proves that only efficient (not weakly efficient) units participating in construction of revised reference. And theorem 2 proves that comparing u_p to the revised reference results the actual efficiency.

Theorem 3.1. *If u_k is not an efficient unit then* $\sum_{j=1}^{n}\lambda_{jp}^{1^*}\lambda_{kj}^{2^*} = 0$.

Proof. Since $\lambda_{kj}^{2^*}$ is the optimum weight of u_k in the model P2 for evaluating u_j, if u_k is a weakly efficient, an inefficient or a weakly inefficient unit then for all j, $\lambda_{kj}^{2^*} = 0$ and $\sum_{j=1}^{n} \lambda_{jp}^{1^*} \lambda_{kj}^{2^*} = 0$. □

Theorem 3.2. *If we select $\hat{u}(p)$ as a reference for u_p in the model P1, then efficiency value of u_p is equal to θ_p^*.*

Proof. We knew that $u_1(p)$ is on the efficient frontier and θ_p^* is the minimum value of θ_p, so we must prove that $\hat{u}(p) \geq u_1(p)$. We rewrite the $\hat{u}(p)$ as follow: $\hat{u}(p) = \sum_{j=1}^{n} \lambda_{jp}^{1^*} \left(\sum_{j=1}^{n} \lambda_{kj}^{2^*} u_k \right) = \sum_{j=1}^{n} \lambda_{jp}^{1^*} u_2(j)$, where $u_2(j)$ is the reference of u_j in the model P2.

There are three possibilities for s_j:

1. If u_j is an inefficient or a weakly inefficient unit, then $\lambda_{jp}^{1^*} = 0$.
2. If u_j is an inefficient unit, then $\lambda_{jp}^{1^*} > 0$ and $u_2(j) = u_j$.
3. If u_j is an weakly unit, then $\lambda_{jp}^{1^*} > 0$ and $u_2(j) \geq u_j$.

In all cases we have $\hat{u}(p) = \sum_{j=1}^{n} \lambda_{jp}^{1^*} u_2(j) \geq \sum_{j=1}^{n} \lambda_{jp}^{1^*} u_j = u(p)$.

Since θ_p^* is the minimum value of θ_p, therefore, if we select $\hat{u}(p)$ as the reference of u_p in model P1, the efficiency value of u_p is equal to θ_p^*. □

This shows that $\hat{u}(p)$ is on the efficient frontier, and $\sum_{j=1}^{n} \lambda_{jp}^{1^*} \lambda_{kj}^{2*} > 0$ if and only if u_i is an efficient unit.

Based on Theorems 1 and 2, $\hat{u}(p)$ is a combination of (only) efficient units and its corresponding efficiency value is the same as in model P1. After applying models P1 and P2 for all units, it is possible to compute the revised references of all units, which are on the efficient (and not on the weakly efficient) frontier. It sounds that this method is similar two phase method, but in the new method, the two models are independent.

Based on theorems 1 and 2, $\hat{u}(p)$ is on the efficient frontier and as a reference for u_p, the efficiency value is not changed. $\hat{\lambda}_{kp}^* = \sum_{j=1}^{n} \lambda_{jp}^{1^*} \lambda_{kj}^{2*}$, where the related slacks are computed as follows:

$$\hat{s}_{ip}^{-*} = \theta_p^* x_{ip} - \sum_{j=1}^{n} \hat{\lambda}_{jp}^* x_{ij}, \ \hat{s}_{rp}^{+*} = \sum_{j=1}^{n} \hat{\lambda}_{jp}^* y_{rj} - y_{rp}.$$

Therefore, after applying P1 and P2 models for all units, we can compute the revised references, which are on the efficiency frontier, and the related slacks. According to classifying the DMUs, we have:

1. If $\theta_p^* = 1$ and $\sum_{i=1}^{m} \hat{s}_{ip}^{-*} + \sum_{r=1}^{s} \hat{s}_{rp}^{+*} = 0$, then u_p is efficient,
2. If $\theta_p^* = 1$ and $\sum_{i=1}^{m} \hat{s}_{ip}^{-*} + \sum_{r=1}^{s} \hat{s}_{rp}^{+*} > 0$, then u_p is weakly efficient,
3. If $\theta_p^* < 1$ and $\sum_{i=1}^{m} \hat{s}_{ip}^{-*} + \sum_{r=1}^{s} \hat{s}_{rp}^{+*} = 0$, then u_p is inefficient,
4. If $\theta_p^* < 1$ and $\sum_{i=1}^{m} \hat{s}_{ip}^{-*} + \sum_{r=1}^{s} \hat{s}_{rp}^{+*} > 0$, then u_p is weakly inefficient.

The reference for u_p is $\hat{u}(p) = \sum_{j=1}^{n} \hat{\lambda}_{jk}^{*} u_j$.

The basic role of this method is removing the non-Archimedean epsilon of models and using models that have no dependency to each others and could be solved separately while the models of two-phase method needs to be solved respectively, because second model uses of result of first model.

The basis of traditional two-phase method is to find the optimum value of θ in the first stage and fix it to the second LP and solve it to find the maximum value for sum of slacks.

4 Numerical Example

In this section we solve a simple numerical example. We add the constraints $\sum_{j=1}^{n} \lambda_{jp}^{1} = 1$ and $\sum_{j=1}^{n} \lambda_{jp}^{2} = 1$ to models P1 and P2 respectively and create the BCC versions of DEA models.

Consider the following example:

Table 1: Data for the numerical example

	Output y	*Input* x
DMU1	1	1
DMU2	2	1
DMU3	0.2	2
DMU4	3	3
DMU5	4	3

These data are illustrated in Figure 5.

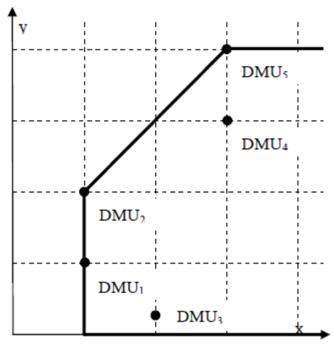

Fig. 5: DMUs of the example.

Table 2 shows the optimum solutions of models P1 and P2 in the BCC format:

Table 2: Results of the example using model P1 and P2 with variable returns to scale

DMU	θ^*	λ^{1^*}					λ^{2^*}				
		1	2	3	4	5	1	2	3	4	5
DMU1	1	1	0	0	0	0	0	1	0	0	0
DMU2	1	0	1	0	0	0	0	1	0	0	0
DMU3	0.50	1	0	0	0	0	0	1	0	0	0
DMU4	0.6667	0	0.5	0	0	0.5	0	0	0	0	1
DMU5	1	0	0	0	0	1	0	0	0	0	1

For example for DMU4, we have $\lambda^{2^*}_{24} = 0.5$, $\lambda^{2^*}_{54} = 0.5$, and $\lambda^{2^*}_{54} = 1$. The following table presents $\hat{s}^{-^*}$, $\hat{s}^{+^*}$ and $\hat{\lambda}^*$:

Table 3: $\hat{s}^{-^*}$, $\hat{s}^{+^*}$ and $\hat{\lambda}^*$

DMU	θ^*	$\hat{s}^{-^*}$	$\hat{s}^{+^*}$	$\hat{\lambda}^*$				
				1	2	3	4	5
1	1	0.0	1.0	0	1	0	0	0
2	1	0.0	0.0	0	1	0	0	0
3	0.50	0.0	1.80	0	1	0	0	0
4	0.6667	0.0	0.0	0	0.5	0	0	0.5
5	1	0.0	0.0	0	0	0	0	1

Thus, the revised references of DMUs are as follow:

Table 4: The revised references of DMUs

DMU	u_j	$u(j)$	$E(j)$
1	(1, -1)	$1 \times u_2$	$\{2\}$
2	(2, -1)	$1 \times u_2$	$\{2\}$
3	(0.2, -2)	$1 \times u_2$	$\{2\}$
4	(3, -3)	$0.5 \times u_2 + 0.5 \times u_5$	$\{2, 5\}$
5	(4, -3)	$1 \times u_5$	$\{5\}$

It is concluded that we can compute the results of the efficiency evaluation of DMUs by applying a linear programming for each unit.

5 Conclusion

There are two major methods for solving the basic DEA models: the epsilon based method, which selects a real number for epsilon, and the two phases method, which is used two LPs for each DMU. In this paper, we presented another method that determines two references for each DMU and then combines them and computes new lambdas and slack variables. Solving models without using non-Archimedean epsilon is an advantage of the new method and ability of computing the models in parallel can reduce the overall running time.

References

1. Ali, A.I. and Seiford, L. M., *Computational Accuracy and Infinitesimals in Data Envelopment Analysis.* INFOR, *37*, (1993), 290-297.
2. Alirezaee, M. R. and Khalili, M., *Recognizing the efficiency, weak efficiency and inefficiency of DMUs with an epsilon independent linear program.* Applied Mathematics and Computation, **183**, (2006), 1323-1327.
3. Amin, G. R. and Toloo, M., *A polynomial-time algorithm for finding Epsilon in DEA models.* Computers Operations Research, **31**, (2004), 803-805.
4. Banker, R. D., Charnes, A., and Cooper, W.W., *Some models for estimating technical and scale inefficiencies in data envelopment analysis*, Management Science, **30**, (1984), 1078-1092.
5. Charnes, A., Cooper, W. and Rhodes, E., *Measuring the Efficiency of Decision Making Units.* European Journal of Operational Research, **2**, (1978), 429-444.
6. Charnes, A., Cooper, W. and Thrall, R. M., *A Structure for Classifying and Characterizing Efficiency and Inefficiency in Data Envelopment Analysis.* The Journal of Productivity Analysis, **2**, (1991), 197-237.
7. Charnes, A. and Rhodes, E., *Short Communication Measuring the Efficiency of DMU's.* EJOR (1979), 339-339.
8. Cooper, W. W., Seiford, L .M. and Tone, K., *Data EnvelopmentAnalysis A Comprehensive Text with Models.*, Applications, References and DEA-Solver Software, Springer Science, (1999).
9. Emrouznejad, A., Parer, B. R. and Tavarese, G., *Evaluation of research in efficiency and productivity: A survery and analysis of the first 30 years of scholarly literature in DEA.* Socio-Economic Planning Sciences, **42**, (2008), 151-157.
10. Mehrabian, S., Jahanshahloo, G. R., Alirezaee, M. R. and Amin, G. R., *An Assurance Interval for the Non-Archimedean Epsilon in DEA Models.* Operations Research, **48**, (2000), 344-347.
11. MirHassani, S. A. and Alirezaee, M. R., *An efficient approach for computing non-Archimedean (Epsilon) in DEA based on integrated models.* Applied Mathematics and Computation, **166**, (2004), 449-456 .