Available online at http://ijim.srbiau.ac.ir/

Int. J. Industrial Mathematics (ISSN 2008-5621)

Vol. 5, No. 1, 2013 Article ID IJIM-00254, 9 pages

Research Article

Directional closest-target based measures of efficiency: Hölder norms approach

G. R. Jahanshahloo *, M. Mehdiloozad †‡, I. Roshdi §

Abstract

Recently, an innovative single-stage approach was developed in [J. Aparicio, J. L. Ruiz, I. Sirvent, Closest targets and minimum distance to the Pareto-efficient frontier in DEA, Journal of Productivity Analysis 28 (2006) 209 – 218], to determine the closest Pareto-efficient targets for a given inefficient decision making unit (DMU). The purpose of this paper is to perfect this approach via integrating it with the concepts of Hölder norms and directional distance function. To this purpose, first, we introduce a furthest-target based directional, named Linear FDHDF, Hölder distance function. Then, we characterize the set of Pareto-efficient points of the production possibility set dominating directionally the assessed DMU. Finally, we develop a closest-target based directional, named Linear CDHDF, Hölder distance function that, as well as providing an efficiency index, determines the closest targets. Comparing to the earlier approach, our approach is more general and the decision maker's preference information can be appropriately incorporated into efficiency assessment and target setting. Furthermore, it is more flexible in computer programming.

Keywords : DEA; Efficiency; Pareto-efficient; Closest target; Directional distance function; Hölder norms.

1 Introduction

Data envelopment analysis (DEA), originally developed by Charnes et al. [3] and subsequently extended by Banker et al. [10], is a non-parametric linear programming based technique to evaluate the relative efficiency of a set of homogeneous DMUs. In addition to the efficiency score, as a practical outcome, the most powerful piece of the obtained information by a DEA analysis is the set of Pareto-efficient projections for the DMU under evaluation. Coordinates of a projection point can be interpreted as the "target" levels of operation of inputs and outputs where give a way of how the assessed DMU can be improved to perform efficiently. Obviously, the more the targets are "similar" to a given DMU, the less the DMU needs practical effort to be efficient. In this sense, the closest targets which the smallest modifications in inputs and outputs of the given DMU is required to reach them, are as much similar as possible to the inputs and outputs of the assessed DMU. However, the Pareto-efficient targets determined via the conventional DEA models (see [2, 3, 9, 10, 18]) have the maximum distance from the assessed DMU whereas this property is in contrast with the concept of similarity. In considering this, the problem of finding closest Pareto-efficient targets for a given DMU, attracted the attention of many researchers. A nice and brief review of the appeared researches in DEA literature concerning this problem can be found in [8]. There are other related papers such as [15, 17], which establishing a relation between the concepts of Hölder norms and directional distance function (DDF), recently introduced in [11, 12], try to obtain the minimum distance to the weak efficient frontier of the production possibility set (PPS). Briec and Leleu [16] tried to measure arbitrary normed distances and normed projections of the inefficient DMUs onto the efficient frontier. More recently, Amirteimoori and Kordrostami [1] have developed the Euclidean distance-based (EDB) measure of efficiency and interpreted it as the "easiest and shortest path to the efficient frontier". However, their approach suffers from several serious problems that we point out them

*Faculty of Mathematical Sciences and Computer, Kharazmi University, Tehran, Iran.

†Faculty of Mathematical Sciences and Computer, Kharazmi University, Tehran, Iran.

‡Corresponding author. m.mehdiloozad@gmail.com

§Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran.

in Appendix.
Among the developed approaches, the approach presented in [8] is an innovative single-stage one and is based on characterizing the Pareto-efficient points of PPS dominating the assessed DMU. Our aim in this paper is to perfect this approach. To this aim, we develop a closest-target based directional, named Linear CDHDF, Hölder distance function, by integrating the concepts of similarity, Hölder norms and DDF. The Linear CDHDF, as well as providing an efficiency index and determining the closest Pareto-efficient targets, is more flexible in computer programming. Furthermore, in comparison with the approach presented in [8], our approach is more general and proposes further modifications in it such as

1. Establishing a close relationship between the DDF and the Hölder norms for measuring efficiency.

2. Incorporating the decision maker's (DM's) preference information in both efficiency analysis and target setting.

3. Determining the set of closest observed reference DMUs to the under assessment DMU.

The reminder of this paper is organized as follows. In the next Section, we conduct a brief review of the DDF and its useful role in efficiency measurement. In Subsection *3.1*, we introduce a furthest-target based Hölder, named Linear FDHDF, distance function. In Subsection *3.2*, by characterizing the set of Pareto-efficient points of the PPS dominating directionally the assessed DMU, we develop the Linear CDFDH. After discussing about the practical advantages of the proposed distance functions in Subsection *3.3*, we derive special cases of them in Subsection *3.4*. In Section 4, we illustrate the developed approach. Finally, the last section summarizes the results and concludes the paper summary.

2 Preliminaries

First, we introduce the necessary notations and define the basic concepts used in this article. Throughout this paper, we deal with n DMUs with m inputs $(i = 1, ..., m)$ and s outputs $(r = 1, ..., s)$. The input and output vectors of DMU_j $(j = 1, ..., n)$, are $x_j = (x_{1j}, ..., x_{mj})^T$ and $y_j = (y_{1j}, ..., y_{sj})^T$ where $x_j \geq 0$, $x_j \neq 0$, $y_j \geq 0$ and $y_j \neq 0$. Further, we consider DMU_o as the DMU under evaluation.

2.1 Production possibility set

One of the first steps in DEA, after identifying inputs and outputs and gathering corresponding data, is choosing an appropriate technology, i.e., determining the PPS. The PPS, T, is the set of all feasible input-output vectors is given by the following production technology:

$$T = \{(x, y) : \ x \ can \ produce \ y\}. \tag{2.1}$$

Under the standard assumptions of inclusion of observations, convexity, constant returns to scale (CRS) and free disposability of inputs and outputs, the unique non-empty PPS spanned by n observed DMUs, $\text{DMU}_j = (x_j, y_j), j = 1, ..., n$, is as follows:

$$T_C = \{ (x, y) \in \mathbb{R}^{m+s}_{\geq 0} | \sum_{j=1}^{n} \lambda_j x_{ij} \leq x \ \ \sum_{j=1}^{n} \lambda_j y_{rj} \geq y \\ \lambda_j \geq 0, \ j = 1, ..., n \}. \tag{2.2}$$

Note: Our study is carried out under the CRS and the results can be recast for other types of returns to scale after some simple changes.

2.2 Directional distance function

The directional distance function, recently introduced in [11, 12] , is a version of Luenberger's shortage function (see [5, 6], which generalizes the traditional Shephard distance function (see [13]) and is well-suited to the task of providing a measure of technical efficiency in the full input-output space. This function projects a given input-output vector, (x, y), proportionally from itself to the frontier of PPS, T, in a pre-assigned direction vector $g = (-g^-, g^+) \in (-\mathbb{R}^m_+) \times \mathbb{R}^s_+$, and is defined as:

$$\begin{aligned} &\vec{D}_T\left(x, y; -g^-, g^+\right) \\ &\quad = Max \left\{\beta \left| \left(x - \beta g^-, y + \beta g^+\right) \in T \right.\right\}. \end{aligned} \tag{2.3}$$

The DEA formulation for the DDF, relative to (2.2), becomes

$$\begin{aligned} \beta^* = Max \ & \beta \\ s.t. \ & \textstyle\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{io} - \beta g_i^- , \quad i = 1, ..., m, \\ & \textstyle\sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro} + \beta g_r^+ , \quad r = 1, ..., s, \\ & \lambda_j \geq 0 \, , \ j = 1, ..., n. \end{aligned} \tag{2.4}$$

Although β is, in principle, unrestricted in sign, its optimal value will never be less than zero, i.e. $\beta^* \geq 0$. The optimal objective, β^*, in general, cannot be interpreted as an efficiency index for any arbitrary direction vector. A way of avoiding this shortcoming is imposing the following primary conditions on the direction vector g

$$\underset{i}{Max} \left\{x_{ij}/g_i^-\right\} \leq 1, \ j = 1, ..., n, \tag{2.5}$$

which guarantees that $\beta^* \leq 1$ and, thereby, $1 - \beta^*$ can be interpreted as an efficiency index. For instance, each of the following direction vectors satisfies the condition (2.5):

$$g_i^- = x_{io}, i = 1, ..., m,\; g_r^+ = y_{ro}, r = 1, ..., s. \quad (2.6)$$

$$g_i^- = \overline{x}_i = \underset{j}{Max}\{x_{ij}\}, i = 1, ..., m, \quad (2.7)$$
$$g_r^+ = \overline{y}_r = \underset{j}{Max}\{y_{rj}\}, r = 1, ..., s.$$

3 Our proposed distance functions

3.1 Linear furthest-target based directional Hölder distance function

Due to proportional adjustments of inputs and outputs in the model (2.4), the non-zero input and output slacks are omitted and, therefore, the DDF fails to take account them as sources of inefficiency. Thus, the model (2.4) does not necessarily lead to a Pareto-efficient projection point on the frontier. To remedy this shortcoming, extending the DDF to non-radial form and using the Hölder norms, we develop a new distance function.

Definition 3.1 °The *Hölder norms* L_P, $p \in [1, \infty]$, are defined over an $n-$dimensional real-normed space as follows:

$$\|\|_p : x \to \|x\|_p = \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}, & if\ p \in [1, \infty[\\ \underset{i}{Max}\{|x_i|\}, & if\ p = \infty. \end{cases} \quad (3.8)$$

The norm $\| \,.\, \|_p$ is called the *p-norm*. Note that for $p = 2$, we have the usual *Euclidean distance*. If $p = \infty$, then the Hölder norm is also called *infinity* or *Chebyshev norm*. Over $\mathbb{R}^n$, the most commonly used norms are $\| \,.\, \|_p$, $p = 1, 2, \infty$.
Now, we introduce a complete furthest-target based directional, named Linear FDHDF, Hölder distance function, relative to T_C, as

$$\begin{aligned} N_o^{p,\,q}(g) = Max \quad & \|\beta^-\|_p + \|\beta^-\|_q \\ s.t. \quad & \textstyle\sum_{j=1}^n \lambda_j x_{ij} = x_{io} - \beta_i^- g_i^-, \quad i = 1, ..., m, \\ & \textstyle\sum_{j=1}^n \lambda_j y_{rj} = y_{ro} + \beta_r^+ g_r^+, \quad r = 1, ..., s, \\ & \lambda_j \geq 0,\; \beta_i^- \geq 0,\; \beta_r^+ \geq 0, \forall j, \forall i, \forall r. \end{aligned} \quad (3.9)$$

In the above model, the vector g, which satisfies in (2.5), represents the pre-assigned direction vector along which DMU$_o$, if it is an inefficient unit, is projected onto the Pareto-efficient frontier of the PPS. The variables β_i^- $(i = 1, ..., m)$ and β_r^+ $(r = 1, ..., s)$ respectively represent the individual rates of contraction and expansion in the ith input and rth output of DMU$_o$, in the direction of g.

The Linear FDHDF simultaneously seeks to non-proportionally expand outputs and reduce inputs. Therefore, it takes account all non-zero slacks and has higher discriminatory power in evaluating DMUs, comparing with the model (2.4).
We define the optimal objective $N_o^{p,\,q}(g)$ as an inefficiency index. Further, corresponding to the model (3.9), we introduce a new efficiency index, named $E_o^{p,\,q}$, as follows:

$$E_o^{p,\,q}(g) := \left[1 - \frac{1}{m}\left\|\beta^{-*}\right\|_p\right] \times \left[1 + \frac{1}{s}\left\|\beta^{+*}\right\|_q\right]^{-1}.$$

It is worth to note that imposing the condition (2.5) on g makes the above index well-defined. This index can be interpreted as the product of two separate components of the input efficiency, $\theta_I = 1 + \frac{1}{m}\sum_{i=1}^m \tau_i^{-*}$, and the output efficiency, $\theta_O = \left(1 - \frac{1}{s}\sum_{r=1}^s \tau_r^{+*}\right)^{-1}$. This interpretation gives a better explanation of the efficiency of the under assessment DMU.

Definition 3.2 ° DMU$_o$ is said to be L-efficient if and only if $E_o^{p,\,q}(g) = 1$.

The above conditions is equivalent to $\beta_i^{-*} = \beta_r^{+*} = 0$, for all i, r in each optimal solution of the model (3.9), i.e., there is no input inefficiency (waste) and no output inefficiency (shortfall) in all inputs and outputs in any optimal solution. The Linear FDHDF satisfies the following properties that can be readily verified:

1. Efficiency requirement: $0 \leq E_o^{p,\,q}(g) \leq 1$. (see [10, 13, 18])
2. $N_o^{p,\,q}(g)$ and $E_o^{p,\,q}(g)$ are "complete" in the sense that are non-oriented and also take account all inefficiencies associated with the non-zero slacks. (see [18])
3. $N_o^{p,\,q}(g) = 0$ if and only if $E_o^{p,\,q}(g) = 1$.
4. $\beta^* \leq N_o^{p,\,q}(g)$ and $E_o^{p,\,q}(g) \leq 1 - \beta^*$.
5. $E_o^{p,\,q}(g) = 1$ if and only if DMU$_o$ is Pareto-efficient.
6. Homogeneity of minus one: $N_o^{p,\,q}(\alpha g) = \frac{1}{\alpha} N_o^{p,\,q}(g)$.
7. Input monotonicity:

$$x' \geq x \Rightarrow$$
$$N^{p,\,q}\left(x', y; -g^-, g^+\right) \geq N^{p,\,q}\left(x, y; -g^-, g^+\right).$$

8. Output monotonicity:

$$y' \leq y \Rightarrow$$
$$N^{p,\,q}\left(x', y; -g^-, g^+\right) \geq N^{p,\,q}\left(x, y; -g^-, g^+\right).$$

9. By choosing a direction vector such that the ith component of $g_i^-(i=1,...,m)$ and rth component of $g_r^+(r=1,...,s)$ respectively have the same units of measurement as the ith input and rth output, e.g., the vectors (2.6) and (2.7), $N_o^{p,\,q}(g)$ and $E_o^{p,\,q}(g)$ will be *unit invariant.* (see [4, 18])

10. Translation invariant: by adding the convexity constraint $\sum_{j=1}^{n}\lambda_j = 1$, the indices $N_o^{p,\,q}(g)$ and $E_o^{p,\,q}(g)$ will be translation invariant, as long as translation of data does not affect the pre-assigned direction vector. (see [7, 9, 13, 18])

3.2 Linear closest-target based directional Hölder distance function

A serious drawback of the traditional DEA models is that the obtained targets by them are the "furthest" Pareto-efficient targets from a given inefficient DMU, instead of the closest ones. This is a consequence of the common feature of maximizing slacks in these models (e.g., see the objective of the models presented in [2, 9, 18]. Similarly, the targets obtained by the Linear FDHDF are the furthest ones from the DMU under evaluation. Therefore, this model does not remove the above-mentioned shortcoming. As pointed out in [8], this is in contrast with the idea of similarity that is implemented as closeness between the values of the inputs and/or outputs of the evaluated DMU and the obtained target levels. To remedy this deficiency, similar to the work of Aparicio et al. [8], we introduce a new closest-target based directional Hölder distance function, which provides the closest Pareto-efficient targets for a given inefficient DMU. Before characterizing the set of Pareto-efficient points of T_C dominating DMU_o in the direction of g, denoted by $\partial_{PE}^{o}T_C$, consider the following model:

$$\begin{aligned} N_o^{1,\,1} = Max \quad & \sum_{i=1}^{m}\beta_i^- + \sum_{r=1}^{s}\beta_r^+ \\ s.t. \quad & \sum_{j\in E}\lambda_j x_{ij} = x_{io} - \beta_i^- g_i^-, \quad i=1,...,m, \\ & \sum_{j\in E}\lambda_j y_{rj} = y_{ro} + \beta_r^+ g_r^+, \quad r=1,...,s, \\ & \lambda_j \ge 0,\ \beta_i^- \ge 0,\ \beta_r^+ \ge 0, \forall j, \forall i, \forall r. \end{aligned} \tag{3.10}$$

which is a reduced form of the model (3.9) for $p=q=1$. The dual problem of the above model is formulated as follows:

$$\begin{aligned} Min \quad & \sum_{i=1}^{m} v_i x_{io} - \sum_{r=1}^{s} u_r y_{ro} \\ \text{s.t.} \quad & \sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} - d_j = 0, \quad j\in E, \\ & d_j \ge 0,\ v_i \ge \tfrac{1}{g_i^-},\ u_r \ge \tfrac{1}{g_r^+}, \forall j, \forall i, \forall r. \end{aligned} \tag{3.11}$$

In the following theorem, connecting the models (3.10) and (3.11), we characterize the set $\partial_{PE}^{o}T_C$.

Theorem 3.1 *$(X,Y)\in\partial_{PE}^{o}T_C$ if and only if there exist $\lambda_j \ge 0$, $d_j \ge 0$, $I_j \in \{0,1\}$, $j\in E$, $v_i \ge \frac{1}{g_i^-}$, $\beta_i^- \ge 0$, $u_r \ge \frac{1}{g_r^+}$, $\beta_r^+ \ge 0$, for all i, r, such that*

$$\begin{aligned} & X = \textstyle\sum_{j\in E}\lambda_j x_j,\ Y_r = \sum_{j\in E}\lambda_j y_j, \\ & \textstyle\sum_{j\in E}\lambda_j x_{ij} = x_{io} - \beta_i^- g_i^-, \quad i=1,...,m, \\ & \textstyle\sum_{j\in E}\lambda_j y_{rj} = y_{ro} + \beta_r^+ g_r^+, \quad r=1,...,s, \\ & \textstyle\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} - d_j = 0, \quad j\in E, \\ & d_j - M\,I_j \le 0, \quad j\in E, \\ & \lambda_j \le M\,(1-I_j), \quad j\in E. \end{aligned} \tag{3.12}$$

Proof. This theorem can be proved as the Theorem cited in [8] with slight modifications.

Since $g>0$, each optimal solution of the above system satisfies $(u,v)>0$ and for any j that $d_j=0$, DMU_j is Pareto-efficient, accordingly. Thus, the resulting projection point, $(X,Y)=(\sum_{j\in E}\lambda_j x_{ij}, \sum_{j\in E}\lambda_j y_{rj})$ would be Pareto-efficient.
Now, using minimization form of the Linear FDHDF on $\partial_{PE}^{o}T_C$, we develop a complete closest-target based directional, named Linear CDHDF, Hölder distance function as

$$\begin{aligned} CN_o^{p,\,q}(g) = Min \quad & \|\beta^-\|_p + \|\beta^-\|_q \\ s.t. \quad & \sum_{j\in E}\lambda_j x_{ij} = x_{io} - \beta_i^- g_i^-, \quad i=1,...,m, \\ & \sum_{j\in E}\lambda_j y_{rj} = y_{ro} + \beta_r^+ g_r^+, \quad r=1,...,s, \\ & \sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} - d_j = 0, \quad j\in E, \\ & d_j - M\,I_j \le 0, \quad j\in E, \\ & \lambda_j \le M\,(1-I_j), \quad j\in E, \\ & \lambda_j \ge 0,\ I_j\in\{0,1\}, \quad j\in E, \\ & v_i \ge \tfrac{1}{g_i^-},\ u_r \ge \tfrac{1}{g_r^+}, \quad \forall i, \forall r, \\ & \beta_i^- \ge 0,\ \beta_r^+ \ge 0, \quad \forall i, \forall r. \end{aligned} \tag{3.13}$$

where the direction vector g = satisfies in (2.5). The corresponding closest targets to the above model will be given by the optimal projections, $\left(\sum_{j=1}^{n}\lambda_j^* x_{ij},\ \sum_{j=1}^{n}\lambda_j^* y_{rj}\right)$. Meanwhile, E is the set of all extreme efficient DMUs founded in [8].
The Linear CDHDF measures the minimum linear distances from DMU_o to $\partial_{PE}^{o}T$ and, similar to the approach presented in [8], provides a single-stage procedures to identify the closest Pareto-efficient targets for DMU_o. Corresponding to the Linear CDHDF, we

introduce a new efficiency index, $CE_o^{p,\,q}$, as follows:

$$CE_o^{p,\,q}(g) := \left[1 - \frac{1}{m}\left\|\beta^{-*}\right\|_p\right] \times \left[1 + \frac{1}{s}\left\|\beta^{+*}\right\|_q\right]^{-1}.$$

Definition 3.3 $^\circ$ DMU$_o$ is said to be LC-efficient if and only if $CE_o^{p,\,q}(g) = 1$.

Because $\partial_{PE}^{o}\mathrm{T} \subseteq \mathrm{T}$ and the Linear FDHDF and Linear CDHDF respectively maximizes and minimizes the same function on the regions T and $\partial_{PE}^{o}\mathrm{T}$, so the following relation is held:

$$CN_o^{p,\,q}(g) \le N_o^{p,\,q}(g). \tag{3.14}$$

The above relation demonstrates that the projection onto the closest point on the efficient frontier leads to an inefficiency score that is not less than that obtained from the usual DEA projection.

The traditional DEA models often, choose a reference set that it contains the furthest Pareto-efficient observations from the inefficient DMU. By contrast, the Linear CDHDF has a most important property that its projection points for a given inefficient DMU$_o$ can be expressed in terms of such observed Pareto-efficient DMUs that they are as much similar as possible to DMU$_o$.

Definition 3.4 *(Closest-Reference Set) Let $(\lambda^*, \beta^{-*}, \beta^{+*})$ be an optimal solution of (3.13) corresponding to a given inefficient DMU$_o$. We define the set of all DMUs corresponding to positive λ_j^* as the closest-reference set to DMU$_o$ which is denoted by R_o i.e.,*

$$\mathrm{R_o} = \left\{DMU_j \,\middle|\, \lambda_j^* > 0\,,\ j = 1, ..., n\right\}. \tag{3.15}$$

Each member of the set R$_o$ is called a *closest reference DMU* to DMU$_o$. In fact, the inputs and outputs levels of each closest reference DMU to DMU$_o$ can be interpreted as the *closest observed targets* for DMU$_o$ and also it can be presented as a *benchmark*. By a technique similar to that presented in [14], we can determine all the observed closest reference DMUs to DMU$_o$.

3.3 Practical advantages of the Linear FDHDF and Linear CDHDF

The Linear FDHDF and the Linear CDHDF have several desirable and important properties that we discuss them here.

1. *Incorporating the DM's preference information into efficiency assessment and target setting*

   In some practical cases, if the DM does not equally prefer the efficient units, then it is necessary to incorporate the DM's judgments or a priori knowledge into the consideration. According to the preference orders of inputs and outputs given by the DM, we can flexibly modify the vector g. Indeed, the values of the modified direction vector, g','s components describe the relative importance of inputs and outputs given by the DM. Let the non-zero weights, w_i, $i = 1, ..., m$ and v_r, $r = 1, ..., s$, respectively are associated with the priorities given by the DM to the inputs and outputs such that the larger the w_i (v_r), the more important the ith input (rth output) is. After incorporating these weights in these models, the coefficients of the variables β_i^- and β_r^+, in the objective function will be w_i and v_r, respectively. Therefore, the components of the modified direction vector, g', should be $g_i^{-'} = \xi_i g_i^-$ and $g_r^{+'} = \psi_r g_r^+$, where $\xi_i = 1/w_i$ and $\psi_r = 1/v_r$. This shows that if an input (output) has a larger importance, it should be attached a larger weight or equivalently a small direction's component. By considering (2.5), we must have $\xi_i \ge 1$, $i = 1, ..., m$, equivalently $w_i \le 1$, $r = 1, ..., s$ [1]. We will clearly exemplify this property via an illustrative example in the next section.

2. *Flexibility in computer programming* A practical advantage of our measures is their computational flexibility i.e., by writing a computer code for one of them, changing only the direction vector's inputs in this program is enough to achieve new scores associated with a new direction vectors. This capability of our models greatly assists the DM to make a more accurate evaluation, by considering several direction vectors, when he cannot rely on an assessment depending only on a specific direction vector. In this case, by running the program for the DM's given direction vectors for example, the average of the obtained scores can be suggested to the DM as a final score for a given DMU.

3. *Finding closest Pareto-efficient targets*

   The levels of the Pareto-efficient targets obtained from the Linear CDHDF for a given inefficient DMU provide a way of how improving it to be efficient. Obviously, the more the targets are close to a DMU, the less DMU needs practical effort to be efficient. Therefore, the Linear CDHDF finding the closest observed targets suggests a way of for how improving the given DMU with the lowest effort to make it efficient.

[1] If the given weights do not satisfy these conditions, the normalized (dividing by $Max\{w_i : i = 1, ..., m\}$) form of them will satisfy these conditions.

Table 1: The data set for Example 4.1.

	A	*B*	*C*	*D*	*E*	*F*	*G*
I_1	1	2	4	7	4	5	7
I_2	6	4	2	1	7	4	5
O_1	1	1	1	1	1	1	1

Table 2: Efficiency scores, targets and reference sets obtained by the Linear FDHDF.

	Eff. Score	Tar.	Ref.
E	0.5455	1.67*B+0.17*C=(4,7,1.8333)	B, C
F	0.6667	0.5*B+C=(5,4,1.5)	B, C
G	0.5000	0.5*B+1.5*C=(7,5,2)	B, C

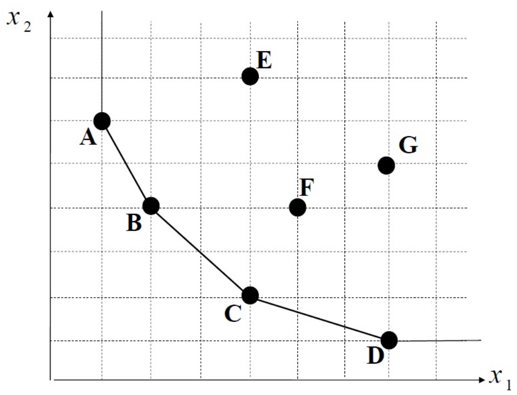

Figure 1: The data set for Example 4.1.

3.4 Specific furthest- and closest-target based directional distance functions

In this subsection, specifying which norm (commonly L_1, L_2 or L_∞) is used in the models (3.9) and (3.13), we derive the following specific distance functions.

The L_1-***distance case:***

By setting $p = q = 1$ in (3.9) and (3.13), they reduce to the Linear L_1-FDHDF and Linear L_1-CDHDF models defined as

$$N_o^{1,\,1}(g) = Max\ \sum_{i=1}^{m}\beta_i^- + \sum_{r=1}^{s}\beta_r^+$$

$$s.t.\ (x_o - \beta^- g^-, y_o + \beta^+ g^+) \in T_C.$$

$$CN_o^{1,\,1}(g) = Min\ \ \sum_{i=1}^{m}\beta_i^- + \sum_{r=1}^{s}\beta_r^+$$

$$s.t.\ \ (x_o - \beta^- g^-, y_o + \beta^+ g^+) \in \partial_{PE}^{o} T_C.$$

where, $\beta^- g^- = (\beta_1^- g_1^-, ..., \beta_m^- g_m^-)^T$ and $\beta^+ g^+ = (\beta_1^+ g_1^+, ..., \beta_s^+ g_s^+)^T$.

The L_2-***distance case:***

By setting $p = q = 2$ in (3.9) and (3.13), they reduce to the Linear L_2-FDHDF and Linear L_2-CDHDF models defined as

$$N_o^{2,\,2}(g) = Max\ \left(\sum_{i=1}^{m}(\beta_i^-)^2\right)^{1/2} + \left(\sum_{r=1}^{s}(\beta_r^+)^2\right)^{1/2}$$

$$s.t.\ (x_o - \beta^- g^-, y_o + \beta^+ g^+) \in T_C.$$

$$CN_o^{2,\,2}(g) = Min\ \left(\sum_{i=1}^{m}(\beta_i^-)^2\right)^{1/2} + \left(\sum_{r=1}^{s}(\beta_r^+)^2\right)^{1/2}$$

$$s.t.\ \ (x_o - \beta^- g^-, y_o + \beta^+ g^+) \in \partial_{PE}^{o} T_C.$$

The $L_\infty-$***distance case:***

By setting $p = q = \infty$ in (3.9) and (3.13), they reduce to the Linear L_∞-FDHDF and Linear L_∞-CDHDF models defined as

$$N_o^{\infty,\,\infty}(g) = Max\ \underset{i}{Max}\{\beta_i^-\} + \underset{r}{Max}\{\beta_r^+\}$$

$$s.t.\ (x_o - \beta^- g^-, y_o + \beta^+ g^+) \in T_C.$$

$$CN_o^{\infty,\,\infty}(g) = Min\ \ \underset{i}{Max}\{\beta_i^-\} + \underset{r}{Max}\{\beta_r^+\}$$

$$s.t.\ (x_o - \beta^- g^-, y_o + \beta^+ g^+) \in \partial_{PE}^{o} T_C.$$

Note that, if in the above models we set $z = \underset{i}{Max}\{\beta_i^-\}$ and $w = \underset{r}{Max}\{\beta_r^+\}$, then we can convert them to the following equivalent forms:

Table 3: Efficiency scores, targets and reference sets obtained by the Linear CDHDF.

	Eff. Score	Tar.	Ref.
E	0.7143	A	A
F	0.8333	0.67*C+0.33*D=(5,1.667,1)	C, D
G	0.7143	D	D

$$N_o^{\infty,\infty}(g) = Max \;\; z+w$$
$$s.t. \;\; (x_o-\beta^- g^-, y_o+\beta^+ g^+) \in T_C$$
$$z \geq \beta_i^-, \quad i=1,...,m,$$
$$w \geq \beta_r^+, \quad r=1,...,s.$$

$$CN_o^{\infty,\infty}(g) = Min \;\; z+w$$
$$s.t. \;\; (x_o-\beta^- g^-, y_o+\beta^+ g^+) \in \partial_{PE}^{o} T_C$$
$$z \geq \beta_i^-, \quad i=1,...,m,$$
$$w \geq \beta_r^+, \quad r=1,...,s.$$

4 Illustrative examples

In this section, two examples are given to provide numerical illustrations of our proposed measures and their features. By the first example, we draw direct comparisons between efficiency scores, target levels and reference sets obtained from the Linear FDHDF and the Linear CDHDF. By the second example, we elaborate on the effect of incorporating the DM's preference information into efficiency measurement and target setting.

Example 4.1 *Consider the hypothetical set of seven homogeneous DMUs, A, B, C, D, E, F and G, using two inputs to produce one output (Table 1).*

Fig 1 can be viewed as representing a section at a given output level, say y=1, of the PPS generated by the seven units. It is easy to see that the units A, B, C and D are extreme Pareto-efficient and so the Pareto-efficient frontier consists of the dark solid piecewise linear part ABCD. Moreover, the units E, F and G are inefficient. Tables 2 and 3 record the results obtained when we assess the inefficient units via the Linear FDHDF and the Linear CDHDF, with the direction vector (2.7). For each of the inefficient DMUs, E, F and G, we have reported the value of its efficiency scores, target levels and the corresponding reference set.
As it is evident from Tables 2 and 3, the efficiency scores and reference sets obtained from the Linear FDHDF is different from that obtained from Linear CDHDF. The efficiency scores obtained from the Linear CDHDF ($CE^{1,1}$) are not less than that obtained from the Linear FDHDF ($E^{1,1}$). This happens because the Linear CDHDF, unlike the FDHDF ones, seeks to minimize the distance from the efficient frontier and so minimize the value of the slack variables.
In evaluating the inefficient units by the Linear FDHDF, the units B and C are determined as the furthest observed reference DMUs for each of them whereas the closest observed reference DMUs determined by the Linear CDHDF for them are different.
In evaluating the units E and G by the Linear FDHDF, they have different score and the unit E performs better than G. However, in evaluating them by the Linear CDHDF, they have equal efficiency scores. Therefore, we can conclude that finding the closest Pareto-efficient targets the ranking orders may be changed.

Example 4.2 *Table 4 shows seven DMUs with two inputs and two outputs. The first four DMUs, A, B, C, D, are extreme Pareto-efficient and the last three ones are inefficient. Using this data set, we elaborate how our proposed models are capable to easily incorporate the DM's preference information into efficiency measurement and target setting. So the efficiency scores and the obtained targets will be based on both the PPS's characteristics together with the given information by the DM. To illustrate this, we consider a specific vector of weights corresponding to I_1, I_2 and O_1 and O_2. Obviously, in real world these weights depend on the context and the experience of the DM. So, only for illustrative purposes, we consider here the vector of weights $(w_1, w_2, v_1, v_2) = (1/3,\ 1,\ 1/10,\ 1)$. As explained in the preceding section, in order to take these information into account, the direction vector (here we use the direction (2.7)) used in the Linear FDHDF and the Linear CDHDF should be modified as follows:*

$$g = (7,7,6,4) \Rightarrow g' = (21,7,60,4\,).$$

Table 5 records the results of evaluating inefficient DMUs by the Linear CDHDF, before and after incorporating the DM's preference information.

Comparing the results of evaluating inefficient DMUs by the Linear CDHDF, before and after incorporating the DM's preference information, we observe significant differences between the obtained scores and the ranking orders. For instance, the rank of unit E in evaluating by the Linear CDHDF, relative to the units F and G, before and after taking account the

Table 4: The data set for Example 4.2.

	A	*B*	*C*	*D*	*E*	*F*	*G*
I_1	1	2	4	7	4	5	4
I_2	6	4	2	1	7	4	5
O_1	2	3	5	6	1	2	3
O_2	1	4	4	3	3	2	3

Table 5: The results for Example 4.2 before and after incorporating the DM's preference information in the Linear CDHDF.

	Before incorporating		After incorporating	
	Score	Ref.	Score	Ref.
E	0.7158	A, B	0.9325	A, B
F	0.7286	A, B	0.9095	A, B
G	0.8000	A, B	0.8941	A, B

DM's preference information are 3 and 1, respectively. Furthermore, there are noticeable increases in the efficiency scores obtained from the Linear CDHDF after taking account the DM's preference information. It is worth to note that the closest observed reference DMUs to the unit F before and after taking account the DM's preference information respectively are C and B.
Overall, our findings highlight the pivotal role of the direction vector in taking the DM's preference information into account and its effect on the efficiency, the resulting target levels.

5 Conclusion

In this paper, using the concepts of the DDF and the mathematical Hölder norms, we proposed new furthest- and closest-target based directional, named Linear FDHDF and Linear CDHDF, Hölder distance functions and, in this way, made a useful generalization of the approach presented in [8]. Our proposed distance functions provide complete efficiency indices, which have straightforward interpretations. Furthermore, the Linear CDHDF provides a more acceptable and realistic evaluation via comparing the performance of the inefficient DMUs with the closest Pareto-efficient targets. In addition, determining the levels of the closest targets for an inefficient DMU, this distance function suggests a way for how improving a given inefficient DMU with the lowest effort to make it efficient. From the practical viewpoint, our approach is very useful since it is more flexible in computer programming and appropriately incorporates the DM's preference information into efficiency assessment and target setting.

Appendix

The approach presented by Amirteimoori and Kordrostami [1] is based upon a mathematical programming, model (3.8) on page 990 in [1], which suffers from several serious problems such as:

1. If the number of *all* the extreme CCR-efficient DMUs in the data set is less than $m + s - 1$ *or* each of the defining hyperplanes of the PPS has this property that the number of extreme efficient DMUs lying on it is less than $m + s - 1$, then the model ((3.8)) will be infeasible.

2. Considering that the Euclidean distance from (x_o, y_o) to the hyperplane $H = \{(x, y)|-\alpha^t x + \beta^t y = 0\}$ is computed by $\frac{|-\alpha^{\,t} x_o + \beta^{\,t} y_o|}{\|(\alpha,\beta)\|_2}$. Thus, to determine such a hyperplane with minimum Euclidean distance from DMU_o, the objective $\frac{\alpha^{\,t} x_o - \beta^{\,t} y_o}{\|(\alpha,\beta)\|_2}$ must be minimized, instead of $\alpha^{\,t} x_o - \beta^{\,t} y_o$.

References

[1] A. Amirteimoori, S. Kordrostami, *A Euclidean distance-based measure of efficiency in data envelopment analysis*, Optimization 59 (2010) 985-996.

[2] A. Charnes, W. W. Cooper, B. Golany, L. Seiford, J. Stutz, *Foundations of data envelopment analysis and Pareto-Koopmans empirical production functions*, Journal of Econometrics 30 (1985) 91-107.

[3] A. Charnes, W. W. Cooper, E. Rodes, *Measuring the efficiency of decision making units*, European Journal of Operational Research 2 (1978) 429-444.

[4] C. A. K. Lovell, J. T. Pastor, *Units invariant and translations invariant DEA*, Operations Research Letters 18 (1995) 147-151.

[5] D. G. Luenberger, *Benefit functions and duality*, Journal of Mathematical Economics 21 (1992) 461-481.

[6] D. G. Luenberger, *Microeconomic Theory*, McGraw Hill, New York, 1995.

[7] I. Ali, L. Seiford, *Translation invariance in data envelopment analysis*, Operations Research Letters 9 (1990) 403-405.

[8] J. Aparicio, J. L. Ruiz, I. Sirvent, *Closest targets and minimum distance to the Pareto-efficient frontier in DEA*, Journal of Productivity Analysis 28 (2006) 209-218.

[9] J. T. Pastor, *Translation invariance in DEA: A generalization*, Annals of Operations Research 66 (1996) 93-102.

[10] K. Tone, *A slacks-based measure of efficiency in data envelopment analysis*, European Journal of Operational Research 130 (2001) 498-509.

[11] R. D. Banker, A. Charnes, W. W. Cooper, *Some models for estimating technical and scale inefficiencies in data envelopment analysis*, Management Science 30 (1984) 1078-1092.

[12] R. G. Chambers, Y. Chung, R. Fre, *Benefit and distance functions*, Journal of Economic Theory 70 (1996) 407-419.

[13] R. G. Chambers, Y. Chung, R. Fre, *Profit, directional distance functions, and Nerlovian efficiency*, Journal of Optimization Theory and Applications 98 (1998) 351-364.

[14] R. W. Shephard, *Theory of Cost and Production Function*, Princeton University Press, Princeton New Jersey, 1970.

[15] T. Sueyoshi, K. Sekitani, *The measurement of returns to scale under a simultaneous occurrence of multiple solutions in a reference set and a supporting hyperplane*, European Journal of Operational Research 181 (2007) 549-570.

[16] W. Briec, B. Lemaire, *Technical efficiency and distance to a reverse convex set*, European Journal of Operational Research 114 (1999) 178-187.

[17] W. Briec, H. Leleu, *Dual representations of non-parametric technologies and measurement of technical efficiency*, Journal of Productivity Analysis 20 (2003) 71-96.

[18] W. Briec, *Hölder distance function and measurement of technical efficiency*, Journal of Productivity Analysis 11 (1998) 111-131.

[19] W. W. Cooper, K. S. Park, J. T. Pastor, *RAM: A range measure of inefficiency for use with additive models, and relations to other models and measures in DEA*, Journal of Productivity Analysis 11 (1999) 5-42.

Golam Reza Jahanshahloo is a Professor of Applied Mathematics at Kharazmi (Tarbiat Moallem) University, Tehran, Iran. His research interests include linear programming, multi-objective linear programming problems, and data envelopment analysis. He has supervised many M.Sc and Ph.D students in these areas, and has published many papers in international journals.

Mahmood Mehdiloozad was born in the Iran, Zanjan in 1986. He has got M.Sc degree in Applied Mathematics, Operations Research field in 2011 from Kharazmi (Tarbiat Moallem) University as the top student. Main research interests include linear program- ming and data envelopment analysis.

Israfil Roshdi was born in the Iran, East Azerbaijan in 1982. He has got M.Sc degree in Ap- plied Mathematics, Operations Research field in 2007 from Kharazmi (Tarbiat Moallem) University as the top student. He is now a PhD student of Applied Mathematics at Sci- ence and Research Branch, Islamic Azad University. He is working on linear programming, data envelopment analysis, and multi-objective linear programming problems.