Practical common weights scalarizing function approach for efficiency analysis

A. Alinezhad
Department of Industrial Engineering, Islamic Azad University of Qazvin, Qazvin, Iran
Corresponding author, Alimizad_ir@yahoo.com

R. Kiani Mavi
Department of Industrial Management, Islamic Azad University of Qazvin, Qazvin, Iran
rezakianimavi@yahoo.com

M. Zohrehbandian
Department of Mathematics, Islamic Azad University of Karaj, Karaj, Iran
zohrebandian@yahoo.com

A. Makui
Department of Industrial Engineering, Iran University of Science & Technology, Tehran, Iran
amakui@iust.ac.ir

Abstract
A characteristic of Data Envelopment Analysis (DEA) is to allow individual decision making units (DMUs) to select the factor weights which are the most advantageous for them in calculating their efficiency scores. This flexibility in selecting the weights, on the other hand, deters the comparison among DMUs on a common base. For dealing with this difficulty and assessing all the DMUs on the same scale, this paper proposes using a multiple objective linear programming (MOLP) approach based on scalarizing function for generating common set of weights under the DEA framework. This is an advantage of the proposed approach against general approaches in the literature which are based on multiple objective nonlinear programming.

Keywords: MOLP, Scalarizing function, DEA,

1. Introduction
Data Envelopment Analysis (DEA) has been widely applied to measure the relative efficiency for a group of homogeneous decision making units (DMUs) with multiple inputs and outputs. Its characteristic is to focus on each individual DMU to select the weights attached to the inputs and outputs, and to show each DMU in its most favorable light as long as the efficiency scores of all DMUs calculated from the same set of weights do not exceed one.

As the models of DEA are run separately for each DMU, the set of weights will typically be different for the various DMUs, and in some cases, this is unacceptable that the same factor is accorded widely differing weights. This flexibility in selecting the weights, on the other hand, deters the comparison among DMUs on a common base. A possible answer to this difficulty lies on the specification of a common set of weights, which was first introduced by Roll et al.(1991). In other words, the major purpose for generating common set of weights is to provide a common base for ranking the DMUs, both the efficient and inefficient ones.

Research about the idea of common set of weights and rankings has developed gradually in recent years. Kao and Hung (2005) based on multiple objective nonlinear programming and by using compromise solution approach, proposed a method to generate a common set of weights for all DMUs which are able to produce a vector of efficiency scores closest to the efficiency scores calculated from the standard DEA model (ideal solution). Likewise, Jahanshahloo et al.(2005) based on multiple objective nonlinear programming and Maximization of the minimum value of the efficiency scores, proposed a method to generate a common set of weights for all DMUs. Some of the other studies in this field are Doyle and Green (1994), Karsak and Ahiska (2005), Roll and Golany (1993).

The plan for the rest of this paper is as follows. In section 2 we present a brief discussion about DEA models and the multiple objective linear programming (MOLP). The mathematical foundation of our method for finding a common set of weights accompany with the method itself is discussed in Section 3. Numerical example is presented in section 4 and finally, section 5 draws the conclusive remarks.

2. DEA and MOLP Preliminaries

Thirty years after the publication of the founding paper of Charnes et al.(1978), DEA can safely be considered as one of the recent success stories in Operations Research and
several hundreds of papers have been published since then. Interestingly, Charnes and Cooper have also had a significant impact on the development of multiple objective linear programming through the development of Goal Programming (GP); Charnes and Cooper (1961). Since the 1970s, MOLP has become a popular approach for modelling and analyzing certain types of multiple criteria decision making (MCDM) problems. Although Charnes and Cooper have played a significant role in the development of DEA and MOLP, researchers in these two camps have generally not paid much attention to research performed in the other camp. Some work on the interactions between MCDM and DEA, are as follows: Bouyssou (1999), Estellita et al. (2004), Giokas (1997), Golany (1988), Joro et al. (1998), Stewart (1996), Xiao and Reeves (1999).

2-1. Data Envelopment Analysis

Consider $n$ production units, or DMUs, each of them consume varying amount of $m$ inputs to produce $s$ outputs. Suppose $x_i \geq 0$ denotes the amount consumed of the $i$-th input and $y_j \geq 0$ denotes the amount produced of the $j$-th output by the $i$-th decision making unit. Then, the following set is the production possibility set (PPS) of obviously most widely used DEA model, CCR with constant returns to scale characteristic:

$$T = \left\{ (x, y) \bigg| \begin{array}{l} x \geq \sum_{j=1}^{n} \lambda_j x_j, \quad y \leq \sum_{j=1}^{n} \lambda_j y_j, \\ \lambda \geq 0, \quad j = 1, 2, \ldots, n \end{array} \right\}$$

**Definition 1:** DMU $j$, $j = 1, 2, \ldots, n$ is called efficient iff there does not exist another $(x, y) \in T$ such that $x < x_j$ and $y > y_j$, and is called pareto efficient iff there does not exist another $(x, y) \in T$ such that $x \leq x_j$ and $y \geq y_j$ and $(x, y) \neq (x_j, y_j)$.

In DEA, the measure of efficiency of a DMU is defined as a ratio of a weighted sum of outputs to a weighted sum of inputs subject to the condition that corresponding ratios for each DMU be less than or equal to one. The model chooses nonnegative weights for a DMU in a way that is most favorable for it. The original model proposed by Charnes et al. (1978), for measuring the efficiency of unit ‘p’, was a fractional linear program as follows:

$$\text{Max} \quad \frac{\sum_{j=1}^{s} u_j y_{jp}}{\sum_{i=1}^{m} v_i x_{ip}}$$

**subject to**

$$\sum_{j=1}^{s} u_j y_{jp} \leq 1, \quad j = 1, 2, \ldots, n$$

$$u_j \geq 0, \quad r = 1, 2, \ldots, s$$

$$v_i \geq 0, \quad i = 1, 2, \ldots, m$$

The above model can be transformed to a linear program by setting the denominator in the objective function equal to an arbitrary constant (e.g., unity) and maximizing the numerator. The obtained model, called input oriented CCR multiplier model, is as follows:

$$\text{CCR}_m \quad \text{Max} \quad \sum_{j=1}^{s} u_j y_{jp}$$

**subject to**

$$\sum_{j=1}^{s} u_j y_{jp} - \sum_{i=1}^{m} v_i x_{ip} \leq 0, \quad j = 1, 2, \ldots, n$$

$$\sum_{i=1}^{m} v_i x_{ip} = 1$$

$$u_j \geq 0, \quad r = 1, 2, \ldots, s$$

$$v_i \geq 0, \quad i = 1, 2, \ldots, m$$

where $u_j$ and $v_i$ are the weights to be applied to the outputs and inputs, respectively and optimum solution of the problem is associated to a normalized coefficient $\left( -v^T u \right)$ of a supporting hyperplane (a hyperplane that contains the PPS in only one of the halfspaces and pass among at least one of the points of it). The dual problem of model $\text{CCR}_m$ will also be used afterwards, called input oriented CCR envelopment model, which has a strong intuitive appeal and is typically the one used to explain and visualize DEA.

$$\text{CCR}_d \quad \text{Min} \quad \theta_p$$

**subject to**

$$\sum_{j=1}^{s} \lambda_j x_{jp} - \theta_p x_{ip} \leq 0, \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{s} \lambda_j y_{jp} \geq y_{op}, \quad r = 1, 2, \ldots, s$$

$$\lambda_j \geq 0, \quad j = 1, 2, \ldots, n$$

A DMU is efficient if and only if the objective function values associated with the optimal solutions to the above problems are one; otherwise it is inefficient. Moreover, if in the former model all variables take a strictly positive value or as in counterpart in the later model all slack variables be equal to zero, it is pareto efficient.
Lemma 1: If $\theta_p^*$ be the optimum solution of model (3), then $\left(\theta_p^*, x_p^*, y_p^*\right)$, called projection of DMU$_p$ on the efficient frontier, is an efficient virtual DMU.

Lemma 2: DMU$_p$ is efficient iff there exist a nonnegative vector $(v, u) \in \mathbb{R}^n \times \mathbb{R}^m$ associated to the gradient vector of a supporting hyperplane where

$$\sum_{i=1}^n u_i y_{ip} - \sum_{i=1}^m v_i x_{ip} = 0$$

2-2. Multiple Objective Linear Programming

The MOLP problem can be written in the general form as follows:

$$\text{Max } f(x) = Cx$$

subject to:

$$x \in X = \left\{ \| g_i (x) \| \leq 0, \; i = 1, 2, ..., m \right\}$$

Where $x \in \mathbb{R}^n$, the objective function matrix $C \in \mathbb{R}^{1 \times n}$ and $g_i (x) \leq 0, \; i = 1, 2, ..., m$, are linear functions.

In MOLP, efficient solution is introduced as follows:

Definition 2: $x^* \in X$ is called an efficient solution (or non-dominated solution) iff there does not exist another $x \in X$ such that $Cx \geq Cx^*$ and $Cx \neq Cx^*$.

One of the methods for finding an efficient solution of model (4) is scalarizing function method. Wierzbicki suggests a scalarizing function to characterize the non-dominated set in the decision space for multiple objective linear programming problem as follow:

$$\min \left\{ \max_{1 \leq i \leq m} \left( b_i - f_j (x) \right) + \delta \sum_{j=1}^m \left( b_j - f_j (x) \right) \right\}$$

subject to

$$x \in X$$

where, $b_j$ is the aspiration level for $f_j$, and $\delta$ is a very small positive number.

The decision maker can give different aspiration levels to each objective function to generate the non-dominated set of objective functions.

In special case if $b_j = f_j^* (x), \; j = 1, 2, ..., n$ and $f_j^* (x)$ is optimum value for each objective function and calculated from the following formulation:

$$\max f_j (x), \; j = 1, 2, ..., n$$

subject to

$$x \in X$$

Therefore, formulation (7) is rewritten as:

$$\min \left\{ \max_{1 \leq j \leq m} \left( b_j - f_j (x) \right) + \delta \sum_{j=1}^m \left( b_j - f_j (x) \right) \right\}$$

subject to

$$x \in X$$

3. A Method for Finding Common Set of Weights

In DEA for calculating the efficiency of different DMUs, different set of weights are obtained, which seems to be unacceptable in reality. Kornbluth (1991) noticed that the DEA model could be expressed as a multi-objective linear fractional programming problem. The objective function of the model is the same as in the CCR model (1), but applied to maximize efficiency of all DMUs, instead of one at a time, and the restrictions remaining unchanged. However, the proposed model was nonlinear. Some other methods also proposed in the literature which were based on Kornbluth’s approach, and unfortunately, all of them are nonlinear.

In this section, we present an improvement to Kornbluth’s approach by introducing an MOLP for finding common weights in DEA. But, firstly, the following model is introduced to find efficiency value of DMU$_p$ which has the same results as the CCR multiplier model. However, it has some advantages compared to foregoing models that will be discussed later on.

$$\text{Max } \sum_{i=1}^n u_i y_{ip} - \theta_p \left( \sum_{i=1}^m v_i x_{ip} \right)$$

subject to

$$\sum_{i=1}^n u_i y_{ip} - \theta_p \left( \sum_{i=1}^m v_i x_{ip} \right) \leq 0, \; j = 1, 2, ..., n$$

$$\sum_{i=1}^n u_i + \sum_{i=1}^m v_i = 1$$

$$u_i \geq 0, \quad r = 1, 2, ..., s$$

$$v_i \geq 0, \quad i = 1, 2, ..., m$$

Where $\theta_j, \; j = 1, 2, ..., n$ is optimum value obtained from the model CCR$_e$, when DMU$_j$ is under consideration. In other words, $\left(\theta_j^*, x_j^*, y_j^*\right)$ is input oriented projection of DMU$_j$ so it is efficient and is settle on the efficient frontier.

Theorem 1: The optimum value of the above model is zero and for its optimal solution, say $\left(u^*, v^*\right)$, we have

$$\frac{\sum_{i=1}^n u_i y_{ip}}{\sum_{i=1}^m v_i x_{ip}} = \theta_p.$$
by Kornbluth (1991). The idea behind the identification of the common weights is formulated as the simultaneously maximizing the ratio of outputs to inputs for all projected DMUs. So we present the following MOLP problem.

\[
\begin{align*}
\text{Max} & \quad \frac{\sum_{i} u_{i} y_{i}}{\sum_{i} v_{i} \theta_{i} x_{i}} \\
\text{subject to} & \quad \sum_{i} u_{i} y_{i} - \sum_{i} v_{i} \theta_{i} x_{i} \leq 0, \quad j = 1,2,...,n \\
& \quad \sum_{i} u_{i} + \sum_{i} v_{i} = 1 \\
& \quad u_{i} \geq 0, \quad r = 1,2,...,s \\
& \quad v_{i} \geq 0, \quad i = 1,2,...,m
\end{align*}
\]

Furthermore, in order to solve the above MOLP model, we use the scalarizing function approach according to the model (5) and theorem 1.

\[
\begin{align*}
\min & \quad \max_{1 \leq j \leq n} \left( \frac{\sum_{i} u_{i} y_{i}}{\sum_{i} v_{i} \theta_{i} x_{i}} \right) \\
\text{subject to} & \quad \sum_{i} u_{i} y_{i} - \sum_{i} v_{i} \theta_{i} x_{i} \leq 0, \quad j = 1,2,...,n \\
& \quad \sum_{i} u_{i} + \sum_{i} v_{i} = 1 \\
& \quad u_{i} \geq 0, \quad r = 1,2,...,s \\
& \quad v_{i} \geq 0, \quad i = 1,2,...,m
\end{align*}
\]

By introducing variable \( w \) to transform the above formulation to linear form, we have:

\[
\begin{align*}
\min & \quad w + \delta \sum_{j} \left( \sum_{i} v_{i} \theta_{i} x_{i} - \sum_{i} u_{i} y_{i} \right) \\
\text{subject to} & \quad \sum_{i} v_{i} \theta_{i} x_{i} - \sum_{i} u_{i} y_{i} \leq w, \quad j = 1,2,...,n \\
& \quad \sum_{i} u_{i} y_{i} - \sum_{i} v_{i} \theta_{i} x_{i} \leq 0, \quad j = 1,2,...,n \\
& \quad \sum_{i} u_{i} + \sum_{i} v_{i} = 1 \\
& \quad u_{i} \geq 0, \quad r = 1,2,...,s \\
& \quad v_{i} \geq 0, \quad i = 1,2,...,m \\
& \quad w \geq 0
\end{align*}
\]

Solving the above model gives us a common set of weights and then efficiency score of DMU \( j \), \( j=1,...,n \), can be obtained by using these common weights as

\[
\frac{\sum_{i} u_{i} y_{i}}{\sum_{i} v_{i} x_{i}}.
\]

If for \( \left( u^{*}, v^{*} \right) \) we have \( \sum_{i} u_{i}^{*} y_{i} = 1 \), then DMU \( p \) is called efficient. Furthermore, by defining the set \( A=\{j; \text{DMU}_{j} \text{ is efficient in model (8)}\} \), and using the same approach as in Jahanshahloo et al. (2005) a complete ranking of DMUs will be obtained.

**Theorem 2:** Such a DMU \( p \) which is indicated efficient by model (8), also is efficient in input oriented CCR model.

**Proof:** According to the first inequalities we have:

\[
\sum_{i} u_{i} y_{i} = \theta_{p}^{*} \leq 1.
\]

Therefore, if \( \sum_{i} u_{i} y_{i} = 1 \) then \( \theta_{p}^{*} = 1 \) and DMU \( p \) is CCR efficient.

**Theorem 3:** There is a DMU \( j \), \( j=1,...,n \) for which model (8) characterize DMU \( j \) as efficient DMU.

**Proof:** Let \( M=m1 \) and \( S=s1 \) where \( 1 \) is a vector of \( m \) and \( s \) ones, \( [1, -1, 1, 1, 1] \), respectively. There is a DMU \( p \), \( P \in \{1,2,...,n\} \) for which the first inequality in (8) is binding. Because, if it is not the case, there is a sufficiently small value \( \varepsilon > 0 \) where \( \left( u^{*}, v^{*} \right) = \left( u + \varepsilon M, v - \varepsilon S \right) \) satisfy the first and the last set of restrictions in (8). On the other hand, the value of \( z \) associated with \( \left( u^{*}, v^{*} \right) \) and the second restrictions will be increased and this is contradicted with optimum value of \( z \). Therefore, there is a DMU \( p \), \( P \in \{1,2,...,n\} \) for which we have:

\[
\sum_{i} u_{i} y_{i} - \sum_{i} v_{i} \theta_{i} x_{i} = 0
\]

However, \( \left( \theta_{p}^{*} x_{p}, y_{p}^{*} \right) \) is efficient. Therefore, \( (u,v) \) is associated to the gradient vector of a supporting hyperplane. Furthermore, this supporting hyperplane must support the PPS at some extreme efficient DMUs. Therefore, such a DMUs are indicated efficient by the model (8).

Roll et al. (1991) and Golany and Yu (1995) indicate that a general requirement for the common set of weights is that it explains as high a portion as possible of DMU performance. This requirement implies that at least one DMU must attain efficiency 1 with the common weights. If there is no DMU with efficiency score 1, then it is obvious that the efficiencies are under-estimated in the sense of relative comparison. More importantly, there is no way of knowing whether the production frontier appropriately represents the sampled DMUs. In this sense, the efficiency scores obtained from the proposed method are not underestimated and satisfied the general requirement.

**4. Numerical Example**

To illustrate the idea of the proposed approach, an example from Kao and Hung (2005) is utilized. In that example, there are 17 forest districts (DMUs). Four inputs (11-14): budget (in US dollars), initial stocking (in cubic meters), labor (in number of employees), and land (in hectares), and
three outputs (O1-O3): main product (in cubic meters), soil conservation (in cubic meters), and recreation (in number of visits) are considered for measuring the efficiency. Table (1) contains the original data, while table (2) shows the common set of weights generated by proposed method (scalarizing function method), respect to inputs and outputs. Furthermore, table (3) shows the efficiency scores of the 17 forest districts calculated from the CCR Model, efficiency scores of the compromise solution approach by Kao and Hung (2005), and efficiency scores of the scalarizing function approach in this paper, respectively.

The CCR efficiency scores are the highest values that the districts can attain, and there are nine efficient units which cannot be differentiated. Regarding the compromise solution approach, Kao and Hung (2005), three values of \( p \), viz., 1, 2, and \( \infty \), have been considered and the results are referred to as MAD, MSE and MAX. The common sets of weights generated from these four models, from which the efficiency scores of every district are calculated, are different sets of weights due to the fact that they are obtained from different viewpoints. Therefore, it is inappropriate to say which weights are correct and which are not. In general, the rankings of these four methods, as shown in parentheses in Table 3, are consistent with those of the CCR model, indicating that the results are reasonable. In addition to this, they are more informative. They not only differentiate the efficient units, but also detect some abnormal efficiency scores calculated from the CCR model. The efficiency scores obtained for districts 9 and 11 are two of such examples. Empirically, in this example the spearman’s correlation between the set of efficiency scores of the scalarizing function method and \( \text{MAX}^* \), which are obtained from similar viewpoints, is greater than 88%.

However, scalarizing function approach solved a linear problem and this is an advantageous of it against Kao and Hung’s approach, which solved a nonlinear problem.

5. Conclusion

The flexibility in the choice of weights is both a weakness and strength of DEA approach. It is a weakness because it deteres the comparison among DMUs on a common base. This flexibility is also a strength, however, if a unit turns out to be inefficient even when the most favorable weights have been incorporated in its efficiency measure, then this is a strong statement and in particular the argument that, the weights are incorrect is not tunable. For dealing with this difficulty and assessment of all DMUs on the same scale, this paper is proposed to examine the application of MOLP approach (scalarizing function) for generating common set of weights under the DEA framework. There are other methods in the literature which are also able to generate common weights. A case taken from Kao and Hung (2005) is solved to investigate the differences among these methods and some conclusions are derived. However, the model proposed in this paper has the following advantages. Firstly, against general approaches in the literature which are based on solving nonlinear problems.

![Table 1](image1)

<table>
<thead>
<tr>
<th>DMUs</th>
<th>I1</th>
<th>I2</th>
<th>I3</th>
<th>I4</th>
<th>O1</th>
<th>O2</th>
<th>O3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>51.62</td>
<td>11.23</td>
<td>49.22</td>
<td>33.52</td>
<td>40.49</td>
<td>14.89</td>
<td>3166.7</td>
</tr>
<tr>
<td>2</td>
<td>85.78</td>
<td>123.98</td>
<td>55.13</td>
<td>108.46</td>
<td>43.51</td>
<td>173.93</td>
<td>6.45</td>
</tr>
<tr>
<td>3</td>
<td>66.65</td>
<td>104.18</td>
<td>257.09</td>
<td>13.65</td>
<td>139.74</td>
<td>115.96</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>27.87</td>
<td>107.6</td>
<td>14</td>
<td>146.43</td>
<td>25.47</td>
<td>131.79</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>51.28</td>
<td>117.51</td>
<td>32.07</td>
<td>84.5</td>
<td>46.2</td>
<td>144.99</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>36.05</td>
<td>193.32</td>
<td>59.52</td>
<td>8.23</td>
<td>46.88</td>
<td>190.77</td>
<td>822.92</td>
</tr>
<tr>
<td>7</td>
<td>25.83</td>
<td>105.8</td>
<td>9.51</td>
<td>227.2</td>
<td>19.4</td>
<td>120.09</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>123.02</td>
<td>82.44</td>
<td>87.35</td>
<td>98.8</td>
<td>43.33</td>
<td>125.84</td>
<td>404.69</td>
</tr>
<tr>
<td>9</td>
<td>61.95</td>
<td>99.77</td>
<td>33</td>
<td>86.37</td>
<td>45.43</td>
<td>79.6</td>
<td>1252.6</td>
</tr>
<tr>
<td>10</td>
<td>80.33</td>
<td>104.65</td>
<td>53.3</td>
<td>79.06</td>
<td>27.28</td>
<td>132.49</td>
<td>42.67</td>
</tr>
<tr>
<td>11</td>
<td>205.92</td>
<td>183.49</td>
<td>144.16</td>
<td>59.66</td>
<td>14.09</td>
<td>196.29</td>
<td>16.15</td>
</tr>
<tr>
<td>12</td>
<td>82.09</td>
<td>104.94</td>
<td>46.51</td>
<td>127.28</td>
<td>44.87</td>
<td>108.53</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>202.21</td>
<td>187.74</td>
<td>149.39</td>
<td>93.65</td>
<td>44.97</td>
<td>184.77</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>67.55</td>
<td>82.83</td>
<td>44.37</td>
<td>60.85</td>
<td>26.04</td>
<td>85</td>
<td>23.95</td>
</tr>
<tr>
<td>15</td>
<td>72.6</td>
<td>132.73</td>
<td>44.67</td>
<td>173.48</td>
<td>5.55</td>
<td>135.65</td>
<td>24.13</td>
</tr>
<tr>
<td>16</td>
<td>84.83</td>
<td>104.28</td>
<td>159.12</td>
<td>171.11</td>
<td>11.53</td>
<td>110.22</td>
<td>49.09</td>
</tr>
<tr>
<td>17</td>
<td>71.77</td>
<td>88.16</td>
<td>69.19</td>
<td>123.14</td>
<td>44.83</td>
<td>74.54</td>
<td>6.14</td>
</tr>
</tbody>
</table>

![Table 2](image2)

<table>
<thead>
<tr>
<th>( v_1 )</th>
<th>( v_2 )</th>
<th>( v_3 )</th>
<th>( v_4 )</th>
<th>( u_1 )</th>
<th>( u_2 )</th>
<th>( u_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00003</td>
<td>0.0011</td>
<td>0.00002</td>
<td>0.00003</td>
<td>0.00001</td>
<td>0.0012</td>
<td>0.00001</td>
</tr>
</tbody>
</table>

45/45
problems the proposed approach solve a linear problem (an MOLP model). Secondly, when common weights of the input/output factors are available, efficiency scores can be measured on the same scale. Moreover, all DMUs can be ranked in terms of a common base. Finally, the proposed method, simply and with appropriate modifications, can be generalized to the other DEA models.

6. References


