Available online at www.ijo.ir

Iranian Journal of Optimization 1(2009) 199-210

Iranian Journal of Optimization

Common weights determination in data envelopment analysis

Alireza Amirteimoori[a*], Sohrab Kordrostami[b]

[a]*Department of Mathematics, Islamic Azad University, Rasht-Iran*
[b]*Department of Mathematics, Islamic Azad University, Lahijan-Iran*

Abstract

In models of data envelopment analysis (*DEA*), an optimal set of weights is generally assumed to represent the assessed decision making unit(*DMU*) in the best light in comparison to all the other *DMU*s, and so there is an optimal set of weights corresponding to each *DMU*. The present paper, proposes a three stage method to determine one common set of weights for decision making units. Then, we use these weights to rank efficient units. We demonstrate the approach by applying it to rank gas companies.

***Keywords**:* Data envelopment analysis; common set of weights; Ranking.

* *Corresponding author. E-mail: teimoori@guilan.ac.ir*

1. Introduction

Data envelopment analysis (*DEA*) is a non-parametric technique for evaluating the performance of many activities. *DEA* evaluates the relative efficiency of a set of homogeneous decision making units(*DMU*) by using a ratio of the weighted sum of outputs to the weighted sum of inputs. Specifically, it determines a set of weights such that the efficiency of a target *DMU* relative to the other *DMU*s is maximized. So, there is an optimal set of weights corresponding to each *DMU*.
We would like to select one common set of weights for all *DMU*s. The idea of common weights in DEA is not new. It was first introduced by Cook et al.[5] and Roll et al.[10]. Cook and Kress[4] gave a subjective or dual preference ranking by developing common weights through a series of bounded runs by closing the gap between the upper and lower limits of the weights. Ganley et al.[8] considered the common weights for all *DMU*s by maximizing the sum of efficiency ratios of all *DMU*s. Liu and Peng [9] proposed a method to determine one common set of weights to create an efficiency score of a set of efficient *DMU*s. They have used these weights to ranking efficient units. Cook and Zhu [6] have developed a goal programming model to derive a common-multiplier set. The important feature of their multiplier set is that it minimizes the maximum discrepancy among the within-group scores from their ideal levels.
As it is, the selection of weights for a set of *DMU*s is connected with the efficient facet of production possibility set (see [7]). In the present paper, we aim to search one common set of weights to determine the efficiency score of a set of *DMU*s. In the procedure we propose, common weights will be associated with an efficient facet of the frontier. This might be equivalently stated as a selection of weights associated with hyperplane that not only maximize contact with the production possibility set but also maximize the technical efficiency of all *DMU*s. This will be done by using a mixed binary linear programming problem. By using these common weights, *DMU*s will be ranked. The rest of the paper is organized as follows: section two begins with the basic *DEA* models. In the next section, section 3, we describe a method to determine one common set of weights using a simple example. We then draw the general approach in section 4. In section six we introduce an empirical example which uses evaluations of the gas companies. Conclusions appear in section 7.

2. Preliminaries

To describe the *DEA* efficiency measurement, let there are n *DMUs* and the performance of each *DMU* is characterized by a production process of m inputs $(x_{ij};\ i=1,\dots,m)$ to yields s outputs $(y_{rj}; r=1,\dots,s)$. We can represent the efficiency of DMU_O (output per unit of input) by the expression

$$e_o = \frac{\sum_{r=1}^{s} u_{ro} y_{ro}}{\sum_{i=1}^{m} v_{io} x_{io}}$$

where u_{ro}, r=1,...,s and v_{io}, i=1,...,m are vectors of multipliers for outputs and inputs, respectively. Charnes et al. [2] proposed measuring the relative efficiency of a set of n *DMUs* by solving, for each DMU_O, the following linear fractional programming problem:

$$e_o^{(CCR)} = Max \frac{\sum_{r=1}^{s} \bar{u}_{ro} y_{ro}}{\sum_{i=1}^{m} \bar{v}_{io} x_{io}}$$

$$\text{s.t.}$$

$$\frac{\sum_{r=1}^{s} \bar{u}_{ro} y_{rj}}{\sum_{i=1}^{m} \bar{v}_{io} x_{ij}} \le 1, \quad j=1,...,n, \qquad (1)$$

$$\bar{u}_{ro} \ge \varepsilon, \quad r=1,...,s,$$

$$\bar{v}_{io} \ge \varepsilon, \quad i=1,...,m$$

where ε is a very small number $(0 < \varepsilon \ll 1)$. The essential idea behind this model is to afford each unit o the opportunity to present its efficiency picture in most favorable light possible. Hence, each *DMU* is allowed to choose multipliers that maximize its efficiency score. Since (1) is a linear fractional programming problem, we can transform it into a linear format using the manner of Charnes and Cooper [3]. Toward this end, let $\left[\sum_{i=1}^{m} \bar{v}_{io} x_{io}\right]^{-1} = t$, $u_{ro} = t\bar{u}_{ro}$ and $v_{io} = t\bar{v}_{io}$. Then, (1) can be written as:

$$e_o^{(CCR)} = Max \sum_{r=1}^{s} u_{ro} y_{ro}$$

$$\text{s.t.}$$

$$\sum_{i=1}^{m} v_{io} x_{io} = 1$$

$$\sum_{r=1}^{s} u_{ro} y_{rj} - \sum_{i=1}^{m} v_{io} x_{ij} \le 0, \ j=1,...,n, \qquad (2)$$

$$u_{ro} \ge \varepsilon, \ r=1,...,s,$$

$$v_{io} \ge \varepsilon, \ i=1,...,m$$

3. Common weights

As we have seen in foregoing section, in *DEA*, each *DMU* maximizes its efficiency score, under the constrains that none of *DMUs'* efficiency scores is allowed to exceed 1. Decision makers always take the maximum efficiency score 1 as the common benchmark level for *DMUs*.

Liu and Peng [9] have used this benchmark level to describe and determine common weights. In their model, the goal is to maximize the efficiency of the aggregate *DMU*, under the conditions that the efficiency score of *DMUs* in E, the set of all *CCR*-extreme efficient *DMUs*, cannot exceed the benchmark level.

In our approach, we believe that there should be a common set of weights which ensure that, if possible, all *DMUs* have efficiency one. In order to motivate our approach to common weights determination, consider the simple example shown in table 1(This example is taken from Liu and Peng[9]). In this example, we have four *DMUs* with two inputs and two outputs. Hence, we need to determine weights u_r, r=1, 2 and v_i, i=1, 2 such that:

$$
\begin{aligned}
&e_1 = \frac{6\,u_1 + 18\,u_2}{3\,v_1 + 5\,v_2} = 1,\ e_2 = \frac{5\,u_1 + 22\,u_2}{4\,v_1 + 3\,v_2} = 1,\\
&e_3 = \frac{14\,u_1 + 9\,u_2}{2\,v_1 + 6\,v_2} = 1,\ e_4 = \frac{13\,u_1 + 15\,u_2}{3\,v_1 + 2\,v_2} = 1, \qquad (3)\\
&u_1,\ u_2,\ v_1,\ v_2 \geq \varepsilon
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
&3\,v_1 + 5\,v_2 - 6\,u_1 - 18\,u_2 = 0,\\
&4\,v_1 + 3\,v_2 - 5\,u_1 - 22\,u_2 = 0,\\
&2\,v_1 + 6\,v_2 - 14\,u_1 - 9\,u_2 = 0, \qquad (4)\\
&3\,v_1 + 2\,v_2 - 13\,u_1 - 15\,u_2 = 0,\\
&u_1,\ u_2,\ v_1,\ v_2 \geq \varepsilon
\end{aligned}
$$

In other word, we need to determine weights u_r and v_i such that for each *DMU*, the weighted sum of outputs is equal to the weighted sum of inputs. Here, we have four equality constraints and four non negativity variables (in general n equality constraints and $m+s$ non negativity variables), and hence, in general, we cannot guarantee the consistency of the system of equations.

We define ρ_j as the deviation of efficiency of DMU_j as follows:

$$\rho_j = 1 - \frac{u_1 y_{1j} + u_2 y_{2j}}{v_1 x_{1j} + v_2 x_{2j}} \geq 0, \quad \text{j=1,...,4.}$$

In our approach to determine common weights, we believe that common weights should be determined such that $\sum_{j=1}^{4}\left[\sum_{r=1}^{2} u_r y_{rj} - \sum_{i=1}^{2} v_i x_{ij}\right]$ is minimized. Toward this end, we define $P_{\min}$ as a lower bound of $\{u_1 y_{rj} + u_2 y_{2j} - v_1 x_{1j} - v_2 x_{2j},\ \text{j=1,...,4}\}$. Then, at the first stage, consider the following:

$$\begin{array}{ll} Max & P_{\min} \\ s.t. & \\ & P_{\min} \leq 6\,u_1 + 18\,u_2 - 3\,v_1 - 5\,v_2 \leq 0, \\ & P_{\min} \leq 5\,u_1 + 22\,u_2 - 4\,v_1 - 3\,v_2 \leq 0, \\ & P_{\min} \leq 14\,u_1 + 9\,u_2 - 2\,v_1 - 6\,v_2 \leq 0, \\ & P_{\min} \leq 13\,u_1 + 15\,u_2 - 3\,v_1 - 2\,v_2 \leq 0, \\ & u_1,\ u_2,\ v_1,\ v_2 \geq \varepsilon. \end{array} \qquad (5)$$

The left hand side inequalities of (5) ensure that $P_{\min}$ is at least as small as the smallest difference between the weighted sum of outputs and the weighted sum of inputs of any *DMU*, while the inequalities of right hand side ensure that zero is at least as large as the largest difference between the weighted sum of outputs and weighted sum of inputs of any *DMU*. Maximizing $P_{\min}$ in (5) means that we minimize the dispersion of pairs $\left(\sum_{j=1}^{2} v_i x_{ij}, \sum_{r=1}^{2} u_r y_{rj}\right)$; j=1,...,4 in the space of (weighted inputs - weighted outputs). In other word, we attempt to bring the *DMUs* close to benchmark level. Hence, if it is possible to find a set of weights such that each *DMU* has efficiency one i.e. to have $P_{\min} = 0$, then this will automatically be found when we come to solve the above linear program.

Let $P_{\min}^*$ be the optimal solution value to (5). In this stage, we maximize contact with the production possibility set. Toward this end, we use the slack variables s_j and binary variables b_j and rewrite the right hand side inequalities of (5) as equality. We attempt to force $s_j = 0$ as many as possible.

So, at the second stage, our approach to common weights determination solves the following mixed binary linear program:

$$
\begin{aligned}
&Min \quad \sum_{j=1}^{4} b_j \\
&s.t. \\
&\qquad 6u_1 + 18u_2 - 3v_1 - 5v_2 + s_1 = 0, \\
&\qquad 5u_1 + 22u_2 - 4v_1 - 3v_2 + s_2 = 0, \\
&\qquad 14u_1 + 9u_2 - 2v_1 - 6v_2 + s_3 = 0, \\
&\qquad 13u_1 + 15u_2 - 3v_1 - 2v_2 + s_4 = 0, \\
&\qquad P_{\min}^{*} \leq 6u_1 + 18u_2 - 3v_1 - 5v_2, \\
&\qquad P_{\min}^{*} \leq 5u_1 + 22u_2 - 4v_1 - 3v_2, \\
&\qquad P_{\min}^{*} \leq 14u_1 + 9u_2 - 2v_1 - 6v_2, \\
&\qquad P_{\min}^{*} \leq 13u_1 + 15u_2 - 3v_1 - 2v_2, \\
&\qquad s_j \leq Mb_j, \; j = 1,\ldots,4, \\
&\qquad b_j \in \{0,1\}, \; j = 1,\ldots,4, \\
&\qquad u_1, \; u_2, \; v_1, \; v_2 \geq \varepsilon.
\end{aligned} \tag{6}
$$

M is a large positive number. Clearly, selecting $b_t = 0$ forces the $s_t = 0$.

Finally, we refine the selection of common weights made in second stage by choosing those that are positive. Toward this end, let $\phi = Min\{u_1, \; u_2, \; v_1, \; v_2\}$ and b^* be the optimal value to (6). We now solve the following:

$$
\begin{aligned}
&Max \quad \phi \\
&\text{s.t.} \\
&\quad 6u_1+18u_2-3v_1-5v_2+s_1=0, \\
&\quad 5u_1+22u_2-4v_1-3v_2+s_2=0, \\
&\quad 14u_1+9u_2-2v_1-6v_2+s_3=0, \\
&\quad 13u_1+15u_2-3v_1-2v_2+s_4=0, \\
&\quad P_{\min}^{*} \le 6u_1+18u_2-3v_1-5v_2, \\
&\quad P_{\min}^{*} \le 5u_1+22u_2-4v_1-3v_2, \\
&\quad P_{\min}^{*} \le 14u_1+9u_2-2v_1-6v_2, \\
&\quad P_{\min}^{*} \le 13u_1+15u_2-3v_1-2v_2, \\
&\quad s_j \le Mb_j,\ j=1,\dots,4, \\
&\quad b_j \in \{0,1\},\ j=1,\dots,4, \\
&\quad \phi \le \mathrm{u}_{\mathrm{r}},\ \mathrm{r}=1,2, \\
&\quad \phi \le \mathrm{v}_{\mathrm{i}},\ \mathrm{i}=1,2, \\
&\quad u_1,\ u_2,\ v_1,\ v_2 \ge \varepsilon.
\end{aligned}
\tag{7}
$$

Model (7) is restricted to maintain $\sum_{j\in E} b_j = b^*$. In this program, we select between the alternative optimal solutions (if any) provided by (6) by maximizing the minimum value of the weights. For our simple four *DMU* example, running the proposed approach yields to the results that are listed in table 1. As the 7-th column of the table indicates, units 1, 3 and 4 belong to the efficient facet

$$0.0309y_1+0.0885y_2-0.5655x_1-0.0164x_2=0.$$

The total deviation of efficiency is 0.0907, while this index is 0.1296 in Liu and Peng approach.

4. General approach

We can now present our general approach for common weights determination. As before, let m be the number of input measures, s be the number of output measures, n be the number of *DMUs*, y_{rp} be the value of output measure r for DMU_p, x_{ip} be the value of input measure i for DMU_p, ε be a very small number $(0 < \varepsilon \quad 1)$ and E be the set of all *CCR*- extreme efficient DMUs.

Let $P_{\min}$ be a lower bound of $\left\{ \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \; : j \in E \right\}$. Then our approach to common weights determination is as follows:

Stage 1: We first need to minimize the dispersion of *DMUs*, in set E, in the space of (weighted inputs-weighted outputs) using the linear program:

$$
\begin{aligned}
& Max \quad p_{\min} \\
& s.t. \\
& \qquad P_{\min} \le \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \; j \in E, \\
& \qquad u_r, \; v_i \ge \varepsilon, \; for \; all \; i, r.
\end{aligned} \tag{8}
$$

Maximizing $P_{\min}$ in (8) means that we bring the *DMUs* close to the benchmark level. Let $P_{\min}{}^*$ be the optimal solution value associated with (8).

Stage 2: We now need to maximize contact with the production possibility set by solving:

$$
\begin{aligned}
& Min \quad \sum_{j \in E} b_j \\
& s.t. \\
& \qquad P_{\min}{}^* \le \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij}, \; j \in E, \\
& \qquad 0 \le s_j \le M b_j, \; j \in E, \\
& \qquad b_j \in \{0,1\}, \; j \in E, \\
& \qquad u_r, \; v_i \ge \varepsilon, \; for \; all \; i, r.
\end{aligned} \tag{9}
$$

The purpose of (9) is to find from among the alternate optima, a supporting hyperplane with a maximal number of efficient units in the support set. Let b^* be the optimal solution value associated with (9).

Stage 3: We now select between alternative optimal solutions (if any) provided by (9) by maximizing the minimum value of the weights. Toward this end, let $\phi = Min\left\{u_1,...,u_s,v_1,...,v_m\right\}$. To determine one common set of weights, solve the following:

$$
\begin{aligned}
& Max \quad \phi \\
& s.t. \\
& \qquad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + s_j = 0, \; j \in E \\
& \qquad p_{\min}^{\;*} \le \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij}, \; j \in E \\
& \qquad 0 \le s_j \le M b_j, \; j \in E, \\
& \qquad \sum_{j \in E} b_j = b^*, \\
& \qquad b_j \in \{0,1\}, \; j \in E, \\
& \qquad \phi \le u_r, \; r = 1,...,s, \\
& \qquad \phi \le v_i, \; i = 1,...,m, \\
& \qquad u_r, \; v_i \ge \varepsilon, \; for \; all \; i, \; r.
\end{aligned}
\qquad (10)
$$

As can be seen, the objective in (10) is to maximize the minimum value of the weights.

5. An empirical study

This section illustrates the common weights determination discussed in this paper with the analysis of gas companies activities. The data set consists of 25 gas companies located in 24 regions in Iran. The data for this analysis are derived from operations during 2006. We use six variables from the data set as inputs and outputs. Inputs include capital, number of staff and operational costs(excluding staff costs), and outputs include number of subscribers, amount of pipe-laying(kilometers) and length of gas network(kilometers). In table 2 we have recorded the data set (All monetary variables are stated in ten millions of current Iranian Rial). Using *CCR* model (1) we have found that five companies #2, #3, #5, #8 and #12 are extreme efficient and so *E = {2, 3, 5, 8, 12}*. Applying our general approach given above we have that:

Stage 1: We find that $P_{\min}{}^{*} = -1162.97,$

Stage 2:

$$b_j = \begin{cases} 0 & j = 2,3,4,5 \\ 1 & elsewhere \end{cases}$$

$$b^{*} = 4,$$

Stage 3:

$$u_1^{*} = 0.0097, \qquad v_1^{*} = 0.0043,$$

$$u_2^{*} = 0.0043, \qquad v_2^{*} = 0.0043,$$

$$u_3^{*} = 0.0067, \qquad v_3^{*} = 0.0043,$$

$$\sum_{j \in E} \rho_j^{*} = 1.0647 \quad \text{and} \quad \sum_{j=1}^{25} \rho_j^{*} = 12.7132$$

It is to be noted that these weights are associated with an efficient facet of the frontier on which companies #2 and #8 are located. In table 2 we have recorded the results.

Table 1: data and results

#j	x_1	x_2	x_3	y_1	y_2	y_3	e_j	$e_j^{(CCR)}$	ρ_j
1#	377430	1401	1528325	27564	501	201529	0.1974(24)	0.2681	0.8026
2#	221338	1094	1186905	44136	803	840445	1(1)	1	0
3#	267806	1079	1323325	27690	251	832616	0.8539(2)	1	0.1461
4#	160912	444	648685	45882	816	251770	0.6125(7)	0.9700	0.3875
5#	177214	801	909539	72676	654	443507	0.7861(4)	1	0.2139
6#	146325	686	545115	19839	177	341585	0.8336(3)	0.8926	0.1664
7#	195138	687	790348	40154	695	233822	0.4616(12)	0.6926	0.5384
8#	108146	152	236722	37770	606	118943	0.7849(5)	1	0.2151
9#	165663	494	523899	28402	652	179315	0.4983(11)	0.7378	0.5017
10#	195728	503	428566	63701	959	195303	0.7177(6)	0.9318	0.2823
11#	87050	343	298696	17334	221	16037	0.1662(25)	0.0.5311	0.8338
12#	124313	129	198598	30242	565	61836	0.5104(10)	1	0.4896
13#	67545	117	131649	14139	153	46233	0.5216(9)	0.6900	0.4784
14#	47208	165	228730	13505	211	42094	0.3482(19)	0.8119	0.6518
15#	43494	106	165470	8508	114	44195	0.4216(14)	0.5843	0.5787
16#	48308	141	180866	7478	248	45841	0.3858(16)	0.9162	0.6141
17#	55959	146	194470	1818	230	36513	0.2442(22)	0.7335	0.7558
18#	40605	145	179650	6422	127	70380	0.5636(8)	0.6861	0.4364
19#	61402	87	94226	1860	182	36592	0.3941(15)	0.7613	0.6059
20#	87950	104	91461	2900	170	47650	0.4508(13)	0.9258	0.5492
21#	33707	114	88640	3326	85	13410	0.2323(23)	0.2401	0.7677
22#	100304	254	292995	1478	318	79883	0.3255(20)	0.5857	0.6745
23#	94286	105	98302	9105	273	32553	0.3708(17)	0.9884	0.6292
24#	67322	224	287042	5332	241	72316	0.3522(18)	0.6813	0.6478
25#	102045	104	155514	8082	441	30004	0.2536(21)	0.9968	0.7464

The 8-th column of the table gives the value of $e_j = \frac{\sum_{r=1}^{3} u_r^* y_{rj}}{\sum_{i=1}^{3} v_i^* x_{ij}}, \ j=1,...,25.$ We have shown the *CCR* efficiency of each company in column 9. The value of ρ js are listed in column 10. Returning to table 2, we see that only one company #2 has zero deviation. As can be seen, the total value of deviation of efficiency is 12.7132.

6. Conclusion

In the current paper, we have presented a *DEA* based approach to determine one common set of weights for all *DMUs*. Common weights are selected simultaneously for all *DMUs* so as to minimize the deviation of efficiency of *DMUs*. They also, maximize contact with the production possibility set. We believe that the contribution of this paper is to present a *DEA*-based approach that determines one common set of weights that is associated with an efficient facet of the production possibility set and that these weights can be used for ranking efficient units.

References

[1] Andersen P., and Petersen N. C., A procedure for ranking efficient units in data envelopment analysis, Management Science 39, 1261-1264, 1993.

[2] Charnes A., Cooper W. W., and Rhodes E., Measuring the efficiency of decision making units, European Journal of Operational Research 2, 429-444, 1978.

[3] Charnes A., and Cooper W. W., Programming with linear fractional functions, Naval Research Logistics Quarterly 9, 181-186, 1962.

[4] Cook W., and Kress M., Data envelopment model for aggregating performance ranking, Management Science 36(11), 1302-1310, 1990.

[5] Cook W., Roll Y., and Kazakov A., A DEA model for measuring the relative efficiencies of highway maintenace patrols, INFOR 28(2), 113-124, 1990.

[6] Cook W., and Zhu J., Within-group common weights in DEA: an analysis of power plant efficiency, European Journal of Operational Research 178, 207-216, 2007.

[7] Cooper W. W., Ruiz J., and Sirvant I., Choosing weights from alternative optimal solutions of dual multiplier models in DEA, European Journal of Operational Research 30, 91-107, 2007.

[8] Ganley JA., and Cubbin JS., Sector measurement: application of data envelopment analysis, Amsterdam: Nort-Holland, 1992.

[9] Liu F., and Peng H. H., Ranking of units on the DEA frontier with common weights, Computers & Operations Research 35 (5), 1624-1637., 2008.

[10] Roll Y., Cook W., and Golany B., Controlling factor weights in data envelopment analysis, IIE Transactions 24, 1-9, 1991.