Artificial Life and Robotics (2019) 24:415–421
https://doi.org/10.1007/s10015-019-00535-z

ORIGINAL ARTICLE

DEA cone ratio model based on a paired comparison

Yuto Kamitaka[1] · Shingo Aoki[1] · Kazushige Inoue[1]

Received: 9 April 2018 / Accepted: 20 February 2019 / Published online: 9 March 2019

Abstract
Data Envelopment Analysis (DEA) is well known as evaluation method to measure the efficiencies of decision making units (DMUs) relatively with multiple inputs and outputs items. One of the features of DEA is that DMUs are evaluated based on the Pareto optimal line composed of efficient DMUs (this line is called as "Efficiency Frontier" in DEA). As the efficiency value of the target DMU is calculated by the relative comparison between the current DMU and the point on the efficiency frontier which is the easiest to achieve, the compared points differ for every DMU. Therefore, it is possible to evaluate the efficiency of DMUs relatively considering features of them. On the other hand, there are many cases that a certain adjustment about the data is necessary to make some decisions from the objective results of DEA. To analyze the DEA results continuously, this paper proposed the DEA framework. A proposed framework consists of two following steps: (1) extracting subjectivity information based on a paired comparison, (2) extending the traditional DEA model. The proposed model does not add restrictions to a variable directly and it is formulated in the form where it corrects the search direction, "No solution" does not come out of it. Since the proposed model is correcting the search direction, an analyst's intention is incorporated without taking out an execution impossible solution.

Keywords Data Envelopment Analysis · AHP · Decision support system

1 Introduction

DEA (Data Envelopment Analysis) is well known as the method to measure efficiency of DMUs (decision making units) with multiple inputs and outputs [1, 2]. This method calculates efficiency score of each DMU based on Pareto optimal line which is called efficiency frontier. Then DEA shows a plan for improvement to inefficient DMUs. The procedure of this method consists of following four steps: (1) data item selection which shows the character of objective DMU's activity, (2) DMUs selection by which a relative comparison is carried out to measure the objective DMU's performance, (3) DEA model selection and calculation, and (4) adjustment of DEA parameters if analyst does not obtain the desirable results. This study focuses on step (4).

This work was presented in part at the 23rd International Symposium on Artificial Life and Robotics, Beppu, Oita, January 18–20, 2018.

✉ Kazushige Inoue
k.inoue.uv@it-hiroshima.ac.jp

Yuto Kamitaka
md17001@cc.it-hiroshima.ac.jp

Shingo Aoki
s.aoki.sm@it-hiroshima.ac.jp

[1] Hiroshima Institute of Technology University, Hiroshima, Japan

2 Traditional Data Envelopment Analysis

2.1 Output-based CCR model

To incorporate the subjective information, assurance region (AR) method is developed by DEA researchers [3, 4].

However, as analyst's experience or intuition is necessary in calculation and there are some infeasible cases, it is difficult to use. Therefore, this research expands CCR model which is the most basic DEA model [1] to a new DEA model which is embedded subjectivity information.

In explaining the CCR model, it is defined here as n DMUs ($\mathrm{DMU}_1, \mathrm{DMU}_2, \ldots, \mathrm{DMU}_k, \ldots, \mathrm{DMU}_n$), where each DMU is characterized by m inputs $(x_{1k}, x_{2k}, \ldots, x_{ik}, \ldots x_{mk})$ and s outputs $(y_{1k}, y_{2k}, \ldots, y_{rk}, y_{sk})$. Output-based CCR model can be mathematically formulated by

$$
\begin{aligned}
&\min \sum_{i=1}^{m} v_i^{(k)} x_{ik} = (\varphi_k) \\
&\text{s.t.} \; -\sum_{i=1}^{m} v_i^{(k)} x_{ij} + \sum_{r=1}^{s} u_r^{(k)} y_{rj} \le 0 \\
&\quad (j = 1, 2, \dots, n) \\
&\sum_{r=1}^{s} u_r^{(k)} y_{rk} = 1 \\
&\quad v_i^{(k)} \ge 0 \,, u_r^{(k)} \ge 0.
\end{aligned}
\tag{1}
$$

Here $u_{r^r}^{(k)}$ is multiplier weight given to the rth output, and $v_i^{(k)}$ is multiplier weight given to the ith input. Then, first restriction condition (or constraint) represents that the productivity of all DMU becomes 100% or less. And the objective function represents the minimization of the virtual inputs of DMU_k, setting that the virtual outputs of DMU_k is equal to 1 which is formulated in second restriction. Therefore, the optimal solution of $(v_i^{(k)}, u_r^{(k)})$ represents the convenient weight for DMU_k. Especially, the optimal objective function value indicates the evaluation value for DMU_k. This evaluation value used the convenient weight called "efficiency score" in the manner that $\varphi_k = 1\,(100\%)$ means the state of efficiency, while $\varphi_k < 1\,(100\%)$ means the state of inefficiency.

The dual form of (1) becomes

$$
\begin{aligned}
&\max \varphi_k \\
&s.t.\; x_{ik} \ge \sum_{j=1}^{n} x_{ij} \lambda_j^{(k)} \quad (i = 1, 2, \dots, m) \\
&\varphi_k y_{rk} \le \sum_{j=1}^{n} y_{rj} \lambda_j^{(k)} \quad (r = 1, 2, \dots, s) \\
&\quad \varphi_k \,:\, \text{free}, \; \lambda_j^{(k)} \ge 0.
\end{aligned}
\tag{2}
$$

Here, variable $\lambda_j^{(k)}$ is considered to make a convex combination of the data. φ_k is regarded as the ratio of maximized data and current data. Especially, $\lambda_j^{(k)}$ denotes weights which indicate efficiency frontier and the position of kth DMU is calculated by combining the weight of $\lambda_j^{(k)}$. A set of $\lambda_j^{(k)}$ $(\lambda_j^{(k)} > 0)$ are called "the reference set for kth DMU".

The current position of a DMU is indicated by its own reference set. In general, any DMU has difference reference set from others. This difference causes the difficulty of ranking DMUs. This research illustrates this problem using simple situation. In Fig. 1, the efficiency scores of K, L, and M are assumed to be 0.6, 0.7, and 0.8, respectively. It does not always mean M is superior to others even if its efficiency score is highest. It is because the efficiency score of K is based on B and C, while that of L is based on A and B, and that of M is based on C and D. Then, the rank among three DMUs is not clear if a researcher uses only efficiency score based on original DEA.

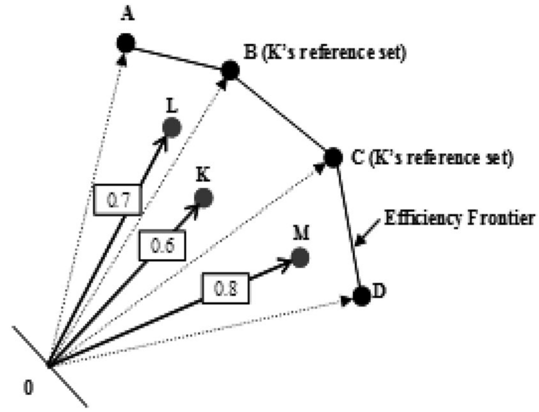

Fig. 1 Reference set

2.2 Assurance region (AR)

Although, there are some cases only specific input/output items such as extremely small input items or extremely large output items are assigned weights in (1) and weights of the rest input/output items are zero equally. As a result, it cannot be evaluated by (1) how much these items weighted as zero contribute to the efficiency. Although it is not a problem mathematically, it may be problems in practical application. To overcome these problems, assurance regions (AR) was developed as the method to adjust dataset [5, 6]. Reflecting subjective opinions of analysts to relations among input/output items, AR enables analysts to examine the influences of input/output items weighted as zero by DEA.

In particular, the following constraint was added to (1) [7].

$$
\begin{aligned}
&\underline{\alpha_i} \le \frac{v_i^{(k)}}{v_1^{(k)}} \le \overline{\alpha_i} (i = 2, \dots, m) \\
&\underline{\beta_r} \le \frac{u_r^{(k)} y_{rk}}{\sum_{r=1}^{s} u_r^{(k)} y_{rk}} \le \overline{\beta_r} (r = 1, \dots, s),
\end{aligned}
\tag{3}
$$

where α_i and β_r are parameters for reflecting subjective opinions of analysts. This approach is the most basic one and although there are other approaches, a lot of them are based on this approach. In AR, considering subjecting opinions by adding some importance of input/output items to their weights, it is attempted to resolve "zero weights". Although, depending on the parameters α_i and β_r, the second constraint of (1) is not satisfied and this model is infeasible.

3 DEA cone ratio model based on a paired comparison

3.1 Overview

To support the framework of the re-calculation incorporating an analyst's intention, this study proposed a new application of DEA to correct an efficient frontier, not operate a parameter directly like AR. The proposed method consists of the following steps: (1) measurement of the subjectivity value by a paired comparison, (2) correction of a specific efficient frontier based on the subjectivity value.

3.2 Measurement of the subjectivity value by a paired comparison

To take in an analyst's subjective information to DEA framework, this study considers the weight presumption problem of each I/O item to a paired comparison procession such as AHP (analytic hierarchy process). Here, this study aims at the paired comparison procession about the reference set (RS) about each output item. The paired comparison procession A about s output items becomes

$$A = (a_{ij})\left\{\in R^{s\times s} | a_{ij} > 0, a_{ji} = 1/a_{ij}\right\}. \tag{4}$$

Then, this method presumes the weight using a following eigenvalue problem:

$$A\mathbf{P} = \lambda_{\max}\mathbf{P}. \tag{5}$$

Let the vector $P = (p_1, p_2, \ldots, p_s)^{\mathrm{T}}$ be an importance vector of the analyst about an output item. Where, the vector P' which took the reciprocal of the element of the vector P is set as follows:

Similarly, let the vector $Q = (q_1, q_2, \ldots, q_n)^{\mathrm{T}}$ be an importance vector of analyst about n DMUs. Then, let us set the following vector:

$$Q' = (1/q_1, 1/q_2, \ldots, 1/q_n)^{\mathrm{T}}. \tag{6}$$

3.3 Correction of a specific efficient frontier

Setting an output data procession to $Y = (y_{rj}) \in R^{s\times n}$, this study considers the following corrected output data procession using the parameters P and Q.

$$Y^P = \begin{pmatrix} 1/P_1 & 0 & \cdot\cdot\cdot & 0 \\ 0 & 1/P_2 & 0\cdot\cdot & 0 \\ \cdot & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \cdot\cdot \\ 0 & 0 & \cdot\cdot\cdot & 1/P_s \end{pmatrix} Y, \tag{7}$$

$$Y^Q = Y \begin{pmatrix} 1/Q_1 & 0 & \cdot\cdot\cdot & 0 \\ 0 & 1/Q_2 & 0\cdot\cdot & 0 \\ \cdot & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \cdot\cdot \\ 0 & 0 & \cdot\cdot\cdot & 1/Q_n \end{pmatrix}. \tag{8}$$

Setting the element of the vector Y^P to $y^P_{rj} \in R^{s\times n}$, let us to correct second restriction of formula (2).

$$\varphi_k y_{rk} \le \sum_{j=1}^{n} y^P_{rj}\lambda^{(k)}_j \quad (r = 1, 2, \ldots, s). \tag{9}$$

Formula (9) has incorporated the data corrected to the coefficient of $\lambda^{(k)}_j$ which is a variable for forming the efficient frontier about an output. Formula (9) can be replaced by following formula (10) and (11) mathematically:

$$\varphi_k y_{rk} \le \sum_{j=1}^{n} y^P_{rj}\lambda^{(k)}_j = \sum_{j=1}^{n} (1/p_r) y_{rj}\lambda^{(k)}_j, \quad (r = 1, 2, \ldots, s) \tag{10}$$

$$(p_r\varphi_k) y_{rk} \le \sum_{j=1}^{n} y_{rj}\lambda^{(k)}_j. \quad (r = 1, 2, \ldots, s) \tag{11}$$

Formula (11) corrects the magnifying power of each output. Using Fig. 2, this study describes about a graphical interpretation. Figure 2 assumes that there is one inefficient

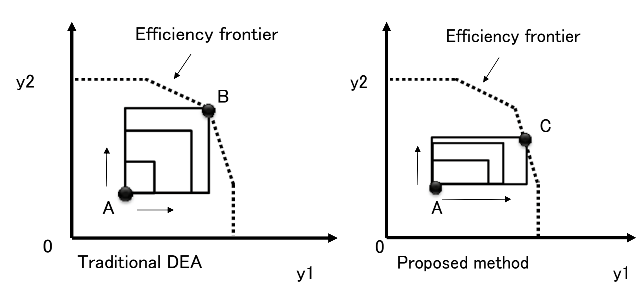

Fig. 2 Search direction

DMU (A) with two output element and an efficiency frontier which represents the best practice frontier. An evaluation value of DMU (A) is measured by an efficient frontier.

Because traditional DEA is treating the output element as the same rank at this time, ideal activity (A') which touches an efficiency frontier with search direction enlarged by the square is chosen.

On the other hand, because proposed method is used, corrected data by the importance vector ("$y1$ is more important than $y2$") about each output, ideal activity (A″) is chosen. It has the search direction which inclines to the output item $y1$ to think as important. Thus, the proposed method does not add restrictions to a variable directly, and since it is formulized in the form where it corrects to the search direction, "No solution" does not come out of it.

3.4 Correction for importance vector about the DMUs

Setting the element of the vector Y^Q to $y_{rj}^Q \in R^{s\times n}$ lets us to correct the first restriction of formula (1).

$$-\sum_{i=1}^{m} v_i^{(k)} x_{ij} + \sum_{r=1}^{s} w_r^{(k)} y_{rj}^Q < 0. \quad (j = 1, 2, \ldots, n) \tag{12}$$

Formula (12) can be replaced by following formulas (13), (14), and (15) mathematically:

$$\frac{\sum_{r=1}^{s} w_r^{(k)} y_{rj}^Q}{\sum_{i=1}^{m} v_i^{(k)} x_{ij}} < 1, \quad (j = 1, 2, \ldots, n) \tag{13}$$

$$\frac{\sum_{r=1}^{s} w_r^{(k)} y_{rj}^Q}{\sum_{i=1}^{m} v_i^{(k)} x_{ij}} = \frac{\sum_{r=1}^{s} w_r^{(k)} (1/q_j) y_{rj}}{\sum_{i=1}^{m} v_i^{(k)} x_{ij}} < 1, \quad (j = 1, 2, \ldots, n) \tag{14}$$

$$\frac{\sum_{r=1}^{s} w_r^{(k)} y_{rj}}{\sum_{i=1}^{m} v_i^{(k)} x_{ij}} < q_j. \quad (j = 1, 2, \ldots, n) \tag{15}$$

Formulas (13) and (14) represent the condition that productivity (virtual-corrected output/virtual input) of all DMU is made 100% or less.

On the other hand, formula (15) represents the condition that restricts the upper limit of the productivity in each DMU by the important vector about the DMUs. For example, if the analyst assumes that DMU "o" is more important than DMU "p" ($q_o > q_p$), then following formulation can be obtained:

$$\frac{\sum_{r=1}^{s} w_r^{(k)} y_{ro}}{\sum_{i=1}^{m} v_i^{(k)} x_{io}} > \frac{\sum_{r=1}^{s} w_r^{(k)} y_{rp}}{\sum_{i=1}^{m} v_i^{(k)} x_{ip}}. \tag{16}$$

Formula (16) represents that analyst's assumption is expressed correctly. Thus, proposed method based on the formula (8) can be interpreted with the model which adds restrictions to the order relation of each DMU.

Table 1 Data set

No.	Input		Output		
	$x1$	$x2$	$y1$	$y2$	$y3$
DMU1	27.37	28.20	10.47	97.37	799.7
DMU2	14.79	10.43	97.90	65.84	519.2
DMU3	32.25	37.01	125.20	75.25	610.2
DMU4	24.90	42.65	108.90	72.05	518.8
DMU5	39.00	48.20	115.53	102.94	718.2
DMU6	46.25	77.16	153.45	105.85	566.6
DMU7	30.55	50.86	120.15	79.25	589.7
DMU8	33.06	44.13	136.00	89.06	641.0
DMU9	38.50	85.68	109.30	70.95	441.5
DMU10	37.40	46.47	112.20	72.95	520.2
DMU11	46.40	88.25	114.33	92.53	559.4
DMU12	24.78	22.65	96.72	54.39	396.6
DMU13	36.22	75.44	107.17	82.17	540.4
DMU14	42.29	61.97	144.88	86.65	461.3
DMU15	31.95	55.81	109.42	56.26	342.3
DMU16	30.85	63.84	91.90	50.15	302.9
DMU17	35.10	28.33	108.70	69.65	465.5
DMU18	43.63	123.29	102.32	38.74	290.3
DMU19	66.05	78.77	173.25	105.10	578.5
DMU20	40.47	52.54	147.11	72.68	421.2
DMU21	48.56	94.95	70.78	52.06	227.7

Table 2 Experimental result

No.	Input		Output			Efficiency
	$v1$	$v2$	$u1$	$u2$	$u3$	θ
DMU1	0.037	0.000	0.000	0.000	0.001	0.832
DMU2	0.068	0.000	0.010	0.000	0.000	1.000
DMU3	0.031	0.000	0.005	0.000	0.000	0.586
DMU4	0.040	0.000	0.006	0.000	0.000	0.661
DMU5	0.026	0.000	0.000	0.006	0.000	0.593
DMU6	0.022	0.000	0.000	0.005	0.000	0.514
DMU7	0.033	0.000	0.005	0.000	0.000	0.594
DMU8	0.030	0.000	0.005	0.000	0.000	0.621
DMU9	0.026	0.000	0.004	0.000	0.000	0.429
DMU10	0.027	0.000	0.004	0.000	0.000	0.453
DMU11	0.022	0.000	0.000	0.005	0.000	0.448
DMU12	0.040	0.000	0.006	0.000	0.000	0.590
DMU13	0.028	0.000	0.000	0.006	0.000	0.510
DMU14	0.024	0.000	0.004	0.000	0.000	0.518
DMU15	0.031	0.000	0.005	0.000	0.000	0.517
DMU16	0.032	0.000	0.005	0.000	0.000	0.450
DMU17	0.028	0.000	0.004	0.000	0.000	0.468
DMU18	0.023	0.000	0.003	0.000	0.000	0.354
DMU19	0.015	0.000	0.002	0.000	0.000	0.396
DMU20	0.025	0.000	0.004	0.000	0.000	0.549
DMU21	0.021	0.000	0.000	0.005	0.000	0.241

Table 3 Experimental result of AR

No.	Input		Output			Efficiency
	$v1$	$v2$	$u1$	$u2$	$u3$	θ
DMU1	0.036	0.001	0.000	0.000	0.001	0.825
DMU2	0.067	0.001	0.005	0.000	0.001	1.000
DMU3			No solution			
DMU4	0.038	0.001	0.001	0.000	0.001	0.585
DMU5			No solution			
DMU6			No solution			
DMU7			No solution			
DMU8			No solution			
DMU9			No solution			
DMU10			No solution			
DMU11			No solution			
DMU12	0.039	0.001	0.001	0.000	0.001	0.470
DMU13			No solution			
DMU14			No solution			
DMU15			No solution			
DMU16			No solution			
DMU17			No solution			
DMU18			No solution			
DMU19			No solution			
DMU20			No solution			
DMU21			No solution			

Table 4 Correction of a specific efficient frontier

No.	Input		Output			Efficiency
	$v1$	$v2$	$u1$	$u2$	$u3$	θ
DMU1	0.037	0.000	0.000	0.000	0.002	1.665
DMU2	0.000	0.096	0.000	0.000	0.008	4.000
DMU3	0.000	0.027	0.000	0.000	0.002	1.325
DMU4	0.040	0.000	0.000	0.000	0.002	1.187
DMU5	0.000	0.021	0.000	0.000	0.002	1.197
DMU6	0.022	0.000	0.000	0.000	0.001	0.698
DMU7	0.033	0.000	0.000	0.000	0.002	1.100
DMU8	0.000	0.023	0.000	0.000	0.002	1.167
DMU9	0.026	0.000	0.000	0.000	0.001	0.653
DMU10	0.000	0.022	0.000	0.000	0.002	0.900
DMU11	0.022	0.000	0.000	0.000	0.001	0.687
DMU12	0.000	0.044	0.000	0.000	0.004	1.407
DMU13	0.028	0.000	0.000	0.000	0.002	0.850
DMU14	0.024	0.000	0.000	0.000	0.001	0.621
DMU15	0.031	0.000	0.000	0.000	0.002	0.610
DMU16	0.032	0.000	0.000	0.000	0.002	0.559
DMU17	0.000	0.035	0.000	0.000	0.003	1.320
DMU18	0.023	0.000	0.000	0.000	0.001	0.379
DMU19	0.000	0.013	0.000	0.000	0.001	0.590
DMU20	0.000	0.019	0.000	0.000	0.002	0.644
DMU21	0.021	0.000	0.000	0.000	0.001	0.267

4 Experimental result

4.1 Data set and DEA result

To confirm the effectiveness of the proposed method, experiments are conducted using the sample data in Table 1.

Each DMU has two inputs and three outputs.

Table 2 is the calculation result applied to Eq. (1).

For example, the evaluation value of DMU 1 is obtained as shown in formula (17).

$$\theta = \frac{0 \times 10.47 + 0 \times 97.37 + 0.001 \times 799.7}{0.037 \times 27.37 + 0 \times 28.20} = 0.832, \quad (17)$$

$v2$ and $u3$ are almost 0. An extreme value appears, causing problems that $v2$ and $u3$ are not reflected. It is impossible to find characteristic elements by not being able to analyze some items. To solve this problem, we will conduct experiments using AR of the conventional method and the proposed method.

4.2 Experimental result of AR

Table 3 shows the results obtained based on the constraint equation of AR. The constraint equation is shown in formula (18).

$$\begin{aligned} v2 &> 0.001 \\ u3 &> 0.001 \end{aligned} \quad (18)$$

With AR of the conventional method, there are some no solution cases. Therefore, the items that the analysts want to emphasize are not reflected. Also, the subjectivity of the analyst is necessary for setting.

4.3 Proposed method

Table 4 shows the results of fitted formula (19). $v2$ and $u3$ are multiplied by magnification, respectively. The efficiency value θ may exceed 1 in some cases.

As a result, calculation becomes impossible, value has come out.

$$P = (1/p_1, 2/p_2, 1/p_3, 1/p_4, 0.5/p_5)^{\mathrm{T}}. \quad (19)$$

5 Conclusion

We could expand the DEA model by extracting subjective information based on pair comparison. The proposed DEA model, developed by creating frameworks and modifying specific efficient frontiers, reflected analysts' intentions without taking out unfeasible solutions.

Moreover, it was able to confirm the effectiveness of the proposed method by numerical experiment.

References

1. Cooper WW, Seiford LM, Tone K (2007) Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software, 2nd edn. Springer, New York
2. Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2:429–444
3. Thompson RG, Singleton FD, Thrall RM, Smith BA (1986) Comparative site evaluations for locating a high-energy physics lab in Texas. Interfaces 16:35–49
4. Liu S (2008) A fuzzy DEA/AR approach to the selection of flexible manufacturing systems. Comput Ind Eng 54:66–76
5. Aoki S, Sato Y, Tsuji H (2007) Hierarchical search for improvement direction by data envelopment analysis. In: Proc of the 8th Asia Pacific industrial engineering and management system APIEMS, paper ID = 604, total 9 pages in CD-ROM
6. Sueyoshi T, Onishi K, Kinase Y (1999) A bench mark approach for baseball evaluation. Eur J Oper Res 115:429–428
7. Chang T, Wang H (2007) Application of DEA model to evaluate operating performance in Taiwan DRA industry. In: Proc of the 8th Asia Pacific industrial engineering and management system APIEMS, paper ID = 604, total 8 pages in CD-ROM