

Structured DEA Model Considering Relation among Input and Output Elements

Shingo Yamasaki

Department of Computer Science and Intelligent Systems
Osaka Prefecture University
Osaka, Japan
yamasaki@mis.cs.osakafu-u.ac.jp

Shingo Aoki

Department of Computer Science and Intelligent Systems
Osaka Prefecture University
Osaka, Japan
aoki@cs.osakafu-u.ac.jp

Abstract- Data Envelop Analysis (DEA) is method for evaluating management efficiency of DMU (Decision Making Unit). DEA can analyze DMU's characteristic of data by assigning variable weight to each input and output elements. However the more the number of input and output elements is, the more complex assigning weight is. So, there is a problem that when the number of input and output elements is so large, it is difficult to analyze DMU's characteristics. This paper proposed new DEA model considering relation among input and output elements in order to analyze DMU's characteristics more easily. This model can limit assigned weights. Therefore, above problem can be solved. The power of the proposed method is examined by applying data set of Japanese professional baseball player's result in 2010.

Keywords. Data Envelop Analysis, Linear Programming, Decision Making Unit,

I. INTRODUCTION

Data Envelopment Analysis (DEA) is a method for evaluating management efficiency of DMU (Decision Making Unit). Now Application of DEA is spreading, for example Universities, hospital, professional sport player and so on[1]. And it is expected that DEA will develop further. DEA has two characteristic as follows. (1) evaluates DMU by using multi-input and multi-output data. (2) can analyze DMU's characteristic of data by assigning variable weights to each element of input and output elements. However, the more the number of input and output elements is, the more non-zero weights are[2]. So, if the number of input and output elements are so large, it is difficult to analyze characteristics of DMU[3]. There is another problem of DEA. DMU is evaluated without considering relation among input and output elements. Therefore, the efficiency value of DMU is not necessarily correct.

This study proposes structured DEA method that is considering relation among input and output elements to solve above problems. Then the power of the proposed method is examined by applying data set of Japanese professional baseball player's result in 2010.

II. DATA ENVELOPMENT ANALYSIS

A. Outline of DEA

DEA was proposed by A. Charnes et al. in 1978 as a

method for management analysis. The applicable field of DEA is wide such as data mining or prospect of bankruptcy[4][5].

DEA regards each DMU as production function that produces output by input. Then the efficiency of DMU is calculated compared with other DMUs relatively[6]. There are two characteristics DEA has; (1) weights are assigned to each input and output data and virtual input and output are generated. These weights are not fix but variable so that each DMU can employ suitable weight to be evaluated better. (2) common index for evaluation is shown as efficiency value. The efficient DMU gets efficiency value as one. On the other hand, inefficient DMU gets efficiency value less than one.

Assuming that there are n DMUs and each DMU is characterized by m input and s output. Then the DMU K has input expressed as x_{ik}, K, x_{mk} and output expressed as y_{ik}, K, y_{sk} . Here the efficiency value is shown by calculating following linear programming.

$$\begin{aligned} \max \quad & \sum_{r=1}^s u_r y_{rk} \\ \text{s.t} \quad & - \sum_{r=1}^s v_i x_{ij} + \sum_{r=1}^s u_r y_{rj} \leq 0 \quad (j=1,2,3,\dots,n) \\ & \sum_{r=1}^s v_i x_{ij} = 1 \\ & v_i \geq 0 \quad (i=1,2,3,\dots,n) \quad u_r \geq 0 \quad (r=1,2,3,\dots,n) \end{aligned} \quad (1)$$

Above formula signifies that weights are assigned to input of DMU K to make it as one. Then DEA controls the efficiency value of other DMUs does not exceed over one based on the weights for DMU K . Objective function has the role for maximizing output of remarkable DMU. Moreover, it is possible to analyze strong points of each DMU by weights. That is because input and output are considered in evaluation when the weights have value.

B. Problem of DEA

1) Problem of increasing input and output elements

DEA can analyze characteristics of DMU more multipronged from optimal weights as the number of input and output elements increase. However, increase of them causes

problem about analyzing characteristics of DMU. Generally, if input and output elements are so large, the number of DMU evaluated efficiency is also large. In this case, it is difficult to compare DMUs from their efficiency values. Moreover, non-zero weights which are actually used for evaluation increases. In other words, input and output elements which are characteristic of DMU increase. Then, it is difficult to understand strong characteristics of DMU.

Table I shows example of a result when input and output elements are so large (DMU:10 input:6 output:9). All DMU in this result are evaluated efficiency. So, it is impossible to compare DMUs by efficiency values. Moreover, Non-zero weights of DMU10 are 14 in 15. Because non-zero weights are too large, analyst can't understand strong characteristics of DMU3.

2) Problem of relation among inputs and outputs elements

Generally, analyst considers relation among input and output elements when he evaluates DMUs. However, DEA evaluates DMU regarding input and output elements as not having relation each other at all. So, the relation doesn't reflect in calculation even if there are strong relations among input and output elements. efficiency value of DMU is not necessarily correct.

III. PROPOSED METHOD

This paper points out two problems of DEA in section II. Then, this paper proposes structured DEA for solving them. This model makes structure of DMU and defined component of

B. Calculating efficiency

Efficiency value of component is calculated by formula (1). It is impossible to calculate efficiency value of middle layer and total efficiency, because they don't have input elements. Then their efficiency value is calculated by formula (2) as follows.

$$\begin{aligned} \max \quad & \sum_{r=1}^s u_r y_{rk} \\ \text{s.t} \quad & \sum_{r=1}^s u_r y_{rj} \leq 0 \quad (j=1,2,3,\dots,n) \\ & x_i = 1 \quad (i=1,2,3,\dots,n) \\ & v_i \geq 0 \quad (i=1,2,3,\dots,n) \quad u_r \geq 0 \quad (r=1,2,3,\dots,n) \end{aligned} \quad (2)$$

structure as group of related input and output elements. Structure shows relation among components and total efficiency of DMU. It is possible to evaluate each component by calculating efficiency value. And total efficiency value is calculated by efficiency value of components.

A. Structure of DMU

The structure of this study is constructed according to the following procedure.

Step1 is "Making component". Component is evaluated aspect of DMU and made by making input and output elements which related to an evaluated aspect to same group. It is defined that input elements set of Component K is X_k , output elements set is Y_k . And its efficiency value is θ_k . θ_k is calculating by (1) formula.

Step2 is "Adding middle layer and connecting components". If there is evaluated aspect that is composed of some component which are made in Step1, new component that is called middle layer added to the structure. Middle layer has some outputs which are efficient values of some component that compose middle layer. At the end, connect middle layer and total efficiency are connected based on their relation.

The example of the structure is shown as Figure 1. Now, there are four components, Component A ~ Component D and Middle Layer1. Arrow in Figure 1 shows that a middle layer at arrow's head is composed of component at arrow's tail. For example, Middle Layer 2 is composed of Component C, and Component D.

This formula is based on formula (1) and different point is that formula(2) assumes input value is 1. Therefore, all DMU's input value is equal. So, formula (2) calculates efficiency value by only relative comparison of output. Output elements of the middle layer are efficiency values of components that compose the middle layer. Equally, Efficiency value of total efficiency is calculated by formula (2).

Proposed method can overcome the problem of relation among inputs and outputs elements by calculating efficiency value in each component. Efficiency value of each component is calculated with only inputs and outputs which have relation. And, inputs and outputs which don't have relation aren't used for calculation together. Therefore, proposed method can evaluate as considering relation among inputs and outputs elements.

Table I Example of result when inputs and outputs elements are so large

DMU	efficiency value	weights						weights								
		input1	input2	input3	input4	input5	input6	output1	output2	output3	output4	output5	output6	output7	output8	output9
DMU1	1	0.779	0.802	0.095	0.401	0	0.275	0.584	0	0.153	0	0.014	0.181	0.000	0	0
DMU2	1	0.990	0	0	0.509	0	0.656	1.244	0	0	0	0	0	0	0	0
DMU3	0.9242011	1.291	0.036	0.007	0	0.142	0.696	0.962	0.038	0.157	0.214	0.254	0.657	0.081	0.000	0
DMU4	1	0.662	0.892	0.041	0	0.173	0.378	0.705	0.202	0.043	0	0.007	0.367	0.000	0.00104	0.001
DMU5	1	0.701	1.125	0	0.225	0.041	0.240	0.753	0.253	0	0.059	0	0.216	0.000	0	0
DMU6	1	1.209	0.644	0	0.108	0.022	0.314	0.894	0.278	0	0.231	0.103	0.267	0.083	0.000	0
DMU7	1	0.625	0.191	2.546	0	0	0.026	1.373	0.571	0	0	0	0	0	0	0
DMU8	1	0.269	0.400	1.966	0	0.164	0	0.829	0.546	0.079	0	0.216	0	0	0	0
DMU9	1	0.979	0.400	0	0	0	0	1.024	0	0	0	0	0	0	0	0
DMU10	1	0.524	0.680	0.030	0.071	0.188	0.710	0.583	0.201	0	0.037	0.168	0.548	0.20663	0.140	0.140

Now, calculation example by using Figure1 is shown as follow. First, efficiency value of Component A ~ Component D is calculated by formula (1). Next, efficiency value of Middle Layer1 is calculated by formula (2) using $\theta_A, \theta_B, \theta_C$ as outputs. Middle Layer2 calculated as well as Middle Layer1 using θ_C, θ_D . At the end, Total Efficiency value is calculated by formula (2) using θ_1, θ_2 .

C. Analyzing result

Analyzing result in traditional method is done with optimal weights. On the other hand, proposed method can overcome the problem that it is difficult to understand strong characteristics of DMU by using efficiency values of components and middle layers for analyzing result.

In proposed method, if efficiency value of Component K is 1, the evaluation of Component K is superior to the other DMUs. So, Component K is considered to be characteristic of the DMU. Oppositely, if efficiency value of Component K is low, Component K isn't considered to be characteristic of the DMU.

In analyzing DMUs as stated above, analyst can get characteristics of each DEA without using optimal weights. Therefore, it isn't caused problem in proposed method, when the input and output elements are so large. Moreover, proposed method can compare DMUs whose efficiency values are 1 by efficiency values of each component.

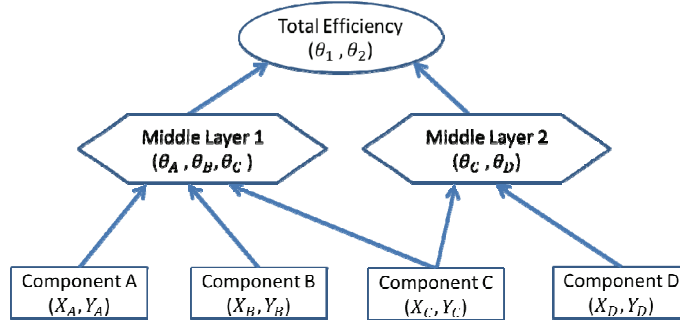


Fig. 1 Example of the structure

IV. NUMERICAL EXPERIMENTS

A. Experimental conditions and data

In order to examine utility of proposed method, Proposed method is applied to data set of 61 Japanese professional baseball player's result in 2010. There are 6 input and 9 output elements as follows.

• Input

- (1) Total plate appearance
- (2) Number of strikeout
- (3) Number of failing stolen bases
- (4) Number of ground into double plays
- (5) Number of error
- (6) Total Defense opportunity

• Output

- (1) Number of hits
- (2) Number of caught stealing
- (3) Number of homerun
- (4) Number of run batted in
- (5) Number of Stolen Bases
- (6) Number of assists
- (7) Number of putouts
- (8) Number of double plays
- (9) Number of sacrifice bunts

From above input and output elements, considering input and output elements, we made five components and two middle layers as follows.

<Component 1> Bat Control

Input: (1),(2),(4) Output:(1)

<Component 2> Stolen Bases

Input: (3) Output:(5)

<Component 3> Sacrifice Bunts

Input: (1),(2),(4) Output:(9)

<Component 4> Slugging

Input: (1),(2),(4) Output:(3) (4)

<Component 5> Defense

Input: (5),(6) Output:(2) (6) (7) (8)

<Middle Layer 1>Technique of hitting

Input : 1 Output: efficiency of Component_1,2,3.

<Middle Layer 2>Offense

Input : 1 Output: efficiency of Middle Layer_1,

efficiency of Component_4

From above component and middle layer, the structure of Baseball player is constructed as Fig. 2.

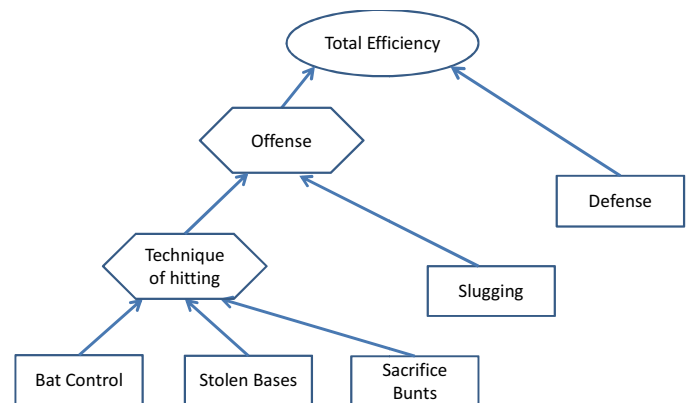


Fig. 2 Structure of Baseball Player

B. Results

The results is shown in Table II . For example, DMU1 has efficiency as 0.918 in Total Efficiency, 0.918 in Technique of hitting and 0.537 in Defense.

C. Disucussion

1) Result of Traditional and Proposed methods

The efficiency value of the traditional method is calculated by formula(1). Table III shows a part of the result of traditional method. It is shown that all DMU exclude only DMU60 have the efficiency value as 1 by traditional method. In this case, it is difficult for analyst to understand difference of DMUs by the efficiency value.

One of the reasons of it is increasing input and output elements. The other is assigning weights without considering relation among input and output elements. For example, DMU1's weight of input(3) is non-zero value, but weight of output(5) is zero. Input(3) and output(5) should be calculated together, because they have relation. Thus, calculation without considering relations makes DMU efficiency in the appearance, but analyst doesn't know whether DMU1 is real efficiency or inefficiency.

In proposed method, assigning weights are limited by making component. So, proposed method can evaluate considering relation among input and output elements. In the result of proposed method, number of DMU which has the efficiency value as 1 is fewer than the result of the traditional method. DMU1 has efficiency as 1 in traditional method but in proposed method efficiency as 0.918. Proposed method could find DMU1 is inefficiency, though analyst doesn't know in traditional method.

Weights have value 13 of 15 input and output elements in DMU3. Therefore, DMU3 is analyzed that it has characteristics overall. However, it is difficult for analyst to understand detail of characteristics, for example which elements is most superior or inferior. Proposed method can

analyze detail of characteristics by the efficiency of each Component or Middle Layer. If DMU has high efficiency value of Component or Middle Layer, it is superiority. Oppositely, if DMU has low efficiency value of Component or Middle Layer , it is inferiority.

In this case, it is easy for analyst to understand characteristics of DMU3 by proposed method. Fig. 3 shows the structure and efficiency value of DMU3. The result based on

proposed method writes in structure figure like Fig.3, it is possible to understand characteristics of DMU just one look.

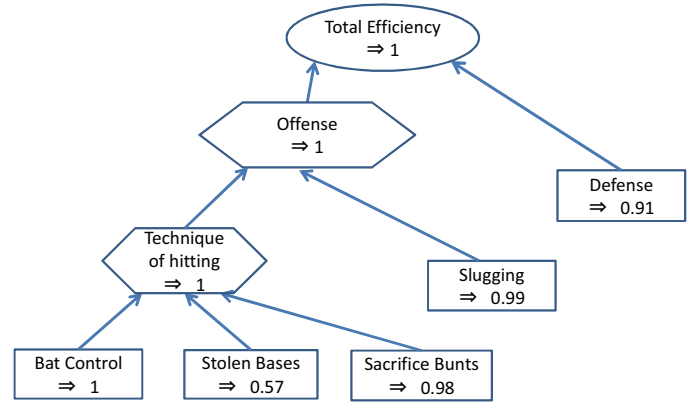


Fig. 3 Structure and Efficiency Value of DMU 3

For instance in Fig. 3, the efficiency value of Offense is 1, so DMU3 has superiority on Offense. The reason why Offense is superiority is found by the efficiency of its lower component. In this case, the reason is that the efficiency value of Technique of hitting is 1. Moreover, DMU3 has high efficiency value of Sacrifice Bunts and Slugging. So, Sacrifice Bunts and Slugging is also characteristics of DMU3. Like this, proposed method can analyzes characteristics of DMU which is difficult to understand by weights based traditional methods.

Table II a part of the result of proposed method

DMU	Total efficiency value	efficiency value of component						
		Offense	Slugging	Technique of hitting	Stolen Bases	Sacrifice Bunts	Bat Control	Deffense
DMU1	0.918	0.918	0.894	0.918	0.476	0.761	0.918	0.537
DMU2	1	1	1	0.94	0.500	0.743	0.940	0.905
DMU3	1	1	0.987	1.000	0.571	0.985	1	0.905
DMU4	1	1	1	0.887	0.476	0.793	0.887	1
DMU5	1	1	1	0.863	0.643	0.680	0.863	0.748
DMU6	1	1	1	1	1	0.993	1	0.834
DMU7	1	1	0.875	1	0.5	0.905	1	0.666
DMU8	1	0.885	0.706	0.885	0.286	0.777	0.885	1.000
DMU9	1	1	0.686	1	0.841	0.675	1	0.879
DMU10	1	0.991	0.961	0.991	0.476	0.962	0.991	1

Table III a part of the result of traditional method

DMU	efficiency value	weights						weights								
		Input(1)	Input(2)	Input(3)	Input(4)	Input(5)	Input(6)	Output(1)	Output(2)	Output(3)	Output(4)	Output(5)	Output(6)	Output(7)	Output(8)	Output(9)
DMU1	1	0.779	0.802	0.095	0.401	0	0.275	0.584	0	0.153	0	0.014	0.181	0.197	0	0
DMU2	1	0.990	0	0	0.509	0	0.656	1.244	0	0	0	0	0	0	0	0
DMU3	1	1.291	0.036	0.007	0	0.142	0.696	0.962	0.038	0.157	0.214	0.254	0.657	0.078	0.081	0
DMU4	1	0.662	0.892	0.041	0	0.173	0.378	0.705	0.202	0.043	0	0.007	0.367	0.235	0	0.001
DMU5	1	0.701	1.125	0	0.225	0.041	0.240	0.753	0.253	0	0.059	0	0.216	0.178	0	0
DMU6	1	1.209	0.644	0	0.108	0.022	0.314	0.894	0.278	0	0.231	0.103	0.267	0.060	0.083	0
DMU7	1	0.625	0.191	2.546	0	0	0.026	1.373	0.571	0	0	0	0	0	0	0
DMU8	1	0.269	0.400	1.966	0	0.164	0	0.829	0.546	0.079	0	0.216	0	0	0	0
DMU9	1	0.979	0.400	0	0	0	0	1.024	0	0	0	0	0	0	0	0
DMU10	1	0.524	0.680	0.030	0.071	0.188	0.710	0.583	0.201	0	0.037	0.168	0.548	0	0.207	0.140

2) Comparison characteristics between DMUs

When the number of DMU which has efficiency value as 1 is so large, it is difficult for analyst to compare DMUs by efficiency value based traditional method. Proposed method can compare the efficiency values of same component between DMUs. So, proposed method can compare DMUs that have efficiency value as 1. For example, it is shown Comparison between DMU2 and DMU4. DMU2 and DMU4 have total efficiency value as 1. Therefore, it is impossible to compare total efficiency value. DMU4 has the efficiency value as 1 in Defense, on the contrary DMU2 has efficiency value as 0.905. From this, DMU2 is superior to DMU4 in Defense. Similarly, it is compared the efficiency value in Technique of hitting, it is found that DMU2 is inferior to DMU4 in Technique of hitting. It is possible to compare each component or middle layer. As a result, it is shown that the proposed method can compare DMUs which have efficiency value as 1.

V. CONCLUSION

This study pointed out two problems of DEA. In order to solve them, this paper proposes Structured DEA. Two problems are as follows: (1) Difficulty to understand characteristics of DMU from the optimal weights calculated by the traditional DEA, (2) DEA can't consider relation among input and output elements. Structured DEA calculates partial evaluation of DMU by input and output elements which have relation. Total evaluation is calculated by the result of partial evaluation.

The proposed method is examined by applying 61 data of Japanese baseball players. From the numerical experiments, it is found that proposed method can analyzes characteristics of DMU which is difficult to understand by weights based traditional methods and can evaluate considering relation among input and output elements.

REFERENCES

- [1] K. Inoue et al. "Data envelopment analysis for evaluating Japanese Universities", *Proc. of The Fifteenth International Symposium on Artificial Life and Robotics 2010*, pp.110-114,(2010)
- [2] Charnes, A., Cooper, W. W. and Rhodes et al. (1999), "Measuring the Efficiency of Decision Making Units," *Proc. of European Journal of Operational Research*, Vol.2, pp.429-444 ,(1978)
- [3] T. Sueyoshi, Y. Kinase and S. Aoki, "DEA Duality on Returns to Scale in Production and Cost Analysis", *Proceedings of the Sixth Asia Pacific Management Conference 2000*, pp.1-7, (2000)
- [4] T. Sueyoshi and S. Aoki, "A use of a nonparametric statistic for DEA frontier shift: the Kruskal and Wallis rank test", *Proc. of The International Journal of Management Science*, Vol.29, No.1, pp.1-18, (2001)
- [5] T.Sueyoshi, K.Onishi, and Y.Kinase, "A Bench Mark Approach for Baseball Evaluation", *Proc. of European Journal of Operational Research*, Vol.115, pp.429-428,(1999)
- [6] D. K. Despotis and Y. G. Smirlis, "Data Envelopment analysis with imprecise data", *Proc. of European Journal for Operational Research*, Vol. 140, pp. 24-36,(2002)