Proceedings of the International MultiConference of Engineers and Computer Scientists 2016 Vol II,
IMECS 2016, March 16 - 18, 2016, Hong Kong

A New Weighted Restriction Method for Data Envelopment Analysis

Asama Supanimitcharoenporn, Siraprapa Manomat, and Orathai Polsen

Abstract—**Data Envelopment Analysis (DEA) is one method to evaluate the relative efficiency of Decision Making Units (DMUs). One of the drawbacks of this method is optimization becoming extremely weighted referring to advantaged inputs or outputs becoming more disadvantaged in terms of DMUs ranking. Since yielding many numbers of DMUs as efficient. The main purpose of this study is to overcome the above problem by using weight restriction. In this study, the author developed a new absolute weight restriction method using Chebyshev's inequality to bound input-output weights to achieve this goal. Also, the performance of the proposed method was examined and compared to the modified CVDEA.**

Keywords—**Data Envelopment Analysis, weight restriction, absolute weights restriction, Chebyshev's inequality**

I. Introduction

Data Envelopment Analysis (DEA) is a mathematical programming technique to evaluate the relative efficiency of Decision Making Units (DMUs) with multiple inputs and outputs. DEA has been widely used and applied in many areas. For example, in engineering, it is useful to compare suppliers for selection, see [1]-[4]. In logistics, it is interesting to evaluate the efficiency of enterprises, see [5], [6]. Moreover, in industry, it is often of interest to evaluate an efficiency of machinery for performance improvement, and rank assessment, see [7].

The DEA method was developed by [8], they used frontier analysis principles which are a concept of [9]. Efficiency scores calculated from the ratio of the weighted sums of inputs to the weighted sums of outputs. The DEA obtains the optimal weights for all inputs and outputs of each unit referring to disadvantage inputs or outputs with no weights and advantage inputs or outputs with extremely high weights, which does not reflect on actual operation of DMUs. This means by yielding many numbers of DMUs as efficient, which leads to DMUs cannot be ranked. This problem arises when applying techniques of weight restriction.

The assurance region approach was developed by [10]. This method attaches the constraint of lower and upper bound input-output weights in the DEA model. After that [11],[12] divided weight restriction into 3 groups as follows assurance region of type I, assurance region of type II, and absolute weights restriction which valuator must determine the bounds on weights. This method uses the bounds on weights considered by human value judgement.

Weight restriction using coefficient of variation of inputs and outputs for a nonlinear optimization model was suggested by [13]. Later [14] modified the model of [13] called Modified CVDEA (MO-CVDEA) which continued to use coefficient for variation to weight restriction. This method was developed for the betterment of dispersion of input-output weights based on the minimization of coefficient of variation. The drawback of this method was the nonlinear optimization model which cause the problem to the construction of dual problem.

In this study, the author has developed an absolute weights restriction method using Chebyshev's inequality. This method is conducted without human interaction referring to bounds on weights. As a result, the weights do not become zero, the relative efficiency of DMUs can be ranked. Moreover, this model suitable for linear optimization.

II. Data Envelopment Analysis

Assuming that there are n DMUs each with m inputs and s outputs, the relative efficiency of a particular DMU$_k$ ($k \in \{1,2, \dots n\}$) is obtained by solving the following fractional programing problem.

$$\begin{aligned}
\text{Max} \quad & \theta_k = \frac{\sum_{r=1}^{s} u_r Y_{rk}}{\sum_{i=1}^{m} v_i X_{ik}} \\
\text{s.t} \quad & \frac{\sum_{r=1}^{s} u_r Y_{rj}}{\sum_{i=1}^{m} v_i X_{ij}} \le 1 \qquad , j = 1,2,...,n \qquad (1) \\
& u_r \ge 0 \qquad , r = 1,2,...,s \\
& v_i \ge 0 \qquad , i = 1,2,...,m \ ,
\end{aligned}$$

where j is the DMU index, $j = 1,\dots,n$, r is the output index, $r = 1,\dots,s$, i is the input index, $i = 1,\dots,m$, Y_{rj} is the value of the rth output for the jth DMU, X_{ij} is the value of the ith input for the jth DMU, u_r is the weight given to the rth

Manuscript received December 5, 2015; revised December 22, 2015. This work was supported in part by graduate college of King Mongkut's University of Technology North Bangkok, Thailand.

Asama Supanimitcharoenporn is with the Applied Statistics Department, King Mongkut's University of Technology North Bangkok, Thailand; (e-mail: su.asama@hotmail.com).

Siraprapa Manomat is with the Applied Statistics Department, King Mongkut's University of Technology North Bangkok, Thailand; (e-mail: srm@kmutnb.ac.th).

Orathai Polsen is with the Applied Statistics Department, King Mongkut's University of Technology North Bangkok, Thailand; (e-mail: orathai.p@sci.kmutnb.ac.th).

output, v_i is the weight given to the ith input, and θ_k is the relative efficiency of DMU$_k$, the DMU under evaluation or the target DMU. In this model, $\theta_k = 1$ means DMU$_k$ is efficient.

This fractional program can be converted into a linear programing where the optimal value of the objective function indicates the relative efficiency of DMU$_k$. The reformulated linear programing problem, also known as the Constant Returns to Scale (CRS) model, is as follows:

$$\begin{aligned}
\text{Max} \quad & \theta_k = \sum_{r=1}^{s} u_r Y_{rk} \\
\text{s.t} \quad & \sum_{i=1}^{m} v_i X_{ik} = 1 \qquad\qquad (2)\\
& \sum_{r=1}^{s} u_r Y_{rj} - \sum_{i=1}^{m} v_i X_{ij} \le 0 \qquad , j = 1,2,\dots,n \\
& u_r \ge 0 \qquad , r = 1,2,\dots,s \\
& v_i \ge 0 \qquad , i = 1,2,\dots,m
\end{aligned}$$

III. Modified Models Based on Coefficient of Variation

A modified model was suggested by [14]. Basically, this method uses coefficient of variation of weights into optimization objective which minimizing coefficient of variation of weights.

Let $\bar{v}_i = \frac{1}{n}\sum_{j=1}^{n} v_{ij}$ and $\bar{u}_r = \frac{1}{n}\sum_{j=1}^{n} u_{rj}$, v_{ij} and u_{rj} are optimal weight of input and output variables respectively which are obtained from (2). The following model, which is a nonlinear optimization model, is based on the CRS model. In this model, they considered the differences of weight and its mean as follows:

$$\begin{aligned}
\text{Max} \quad & \sum_{r=1}^{s} u_{rk} Y_{rk} - \frac{\sqrt{\dfrac{\sum_{r=1}^{s}\left(u_{rk} - \bar{u}_r\right)^2}{s-1}}}{\sum_{r=1}^{s} \bar{u}_r} - \frac{\sqrt{\dfrac{\sum_{i=1}^{m}\left(v_{ik} - \bar{v}_i\right)^2}{m-1}}}{\sum_{i=1}^{m} \bar{v}_i} \\
\text{s.t} \quad & \sum_{i=1}^{m} v_{ik} X_{ik} = 1 \qquad\qquad (3)\\
& \sum_{r=1}^{s} u_{rk} Y_{rj} - \sum_{i=1}^{m} v_{ik} X_{ij} \le 0 \qquad , j = 1,2,\dots,n \\
& u_{rk} \ge 0 \qquad , r = 1,2,\dots,s \\
& v_{ik} \ge 0 \qquad , i = 1,2,\dots,m ,
\end{aligned}$$

where u_{rk} is the weight given to the rth output of kth DMU, v_{ik} is the weight given to the ith input of kth DMU.

The optimal weights from model (2) are used to evaluate the efficiency score (W$_k$) of DMU$_k$, as follows:

$$\mathrm{W}_k = \sum_{r=1}^{s} u_{rk} Y_{rk} ,$$

where W$_k$ = 1 means DMU$_k$ is efficient.

IV. Absolute Weights Restriction Using Chebyshev's Inequality

This study developed an absolute weights restriction by using Chebyshev's inequality to bound the weights, called CIDEA. The bounds of weights based on the CRS model defined as follows:

$$\begin{aligned}
\text{Max} \quad & Z_k = \sum_{r=1}^{s} u_r Y_{rk} \\
\text{s.t} \quad & \sum_{i=1}^{m} v_i X_{ik} = 1 \qquad\qquad (4)\\
& \sum_{r=1}^{s} u_r Y_{rj} - \sum_{i=1}^{m} v_i X_{ij} \le 0 \qquad , j = 1,2,\dots,n \\
& \delta_i \le v_i \le \tau_i \\
& \rho_r \le u_r \le \eta_r \\
& u_r \ge 0 \qquad , r = 1,2,\dots,s \\
& v_i \ge 0 \qquad , i = 1,2,\dots,m ,
\end{aligned}$$

where δ_i and τ_i are lower bounds, and upper bounds of weight of ith inputs respectively. ρ_r and η_r are lower bounds, and upper bounds of weight of rth outputs respectively.

The δ_i and τ_i are calculated from $\bar{v}_i \pm l \cdot sd\left(V_i\right)$, where $\bar{v}_i = \frac{1}{n}\sum_{j=1}^{n} v_{ij}$ and $sd\left(V_i\right) = \sqrt{\dfrac{\sum_{j=1}^{n}\left(v_{ij} - \bar{v}_i\right)}{n-1}}$. Also, the ρ_r and η_r are calculated from $\bar{u}_r \pm l \cdot sd\left(U_r\right)$, where $\bar{u}_r = \frac{1}{n}\sum_{j=1}^{n} u_{rj}$ and $sd\left(U_r\right) = \sqrt{\dfrac{\sum_{j=1}^{n}\left(u_{rj} - \bar{u}_r\right)}{n-1}}$, v_{ij} and u_{rj} are optimal weights of input and output variables respectively which are obtained from (2), and l is a constant from Chebyshev's inequality for each weight. Let V_i is a random variable represents the ith input weight, then

$$\mathrm{P}\left(\left|V_i - \bar{v}_i\right| \ge l \cdot sd\left(V_i\right)\right) \le \frac{1}{l^2} .$$

So that (5)

$$\mathrm{P}\left(\bar{v}_i - l \cdot sd\left(V_i\right) \le V_i \le \bar{v}_i + l \cdot sd\left(V_i\right)\right) \ge 1 - \frac{1}{l^2} .$$

The above expression is used to compute the bound of weights in the case of inputs. If the lower bound weight is less than zero, it will be defined as 10^{-5}, which was suggested by [15]. The output bounds can be calculated in the same way.

The optimal weights in model (4) as Z_k is the relative efficiency of DMU$_k$. In this model, Z_k = 1 means DMU$_k$ is efficient.

V. Numerical Example

In this study a numerical example was used to compare the CIDEA with the MO-CVDEA approach. The purpose is to study the dispersion of weights and find out the best method for ranking DMUs.

Numerical Example. The input and output variables are drawn randomly from Uniform [10,200] distribution with three inputs and three outputs variables for ten DMUs (see Table I).

TABLE I
DATA OF TEN DMUS

DMU	X_1	X_2	X_3	Y_1	Y_2	Y_3
1	52	122	69	16	69	26
2	89	58	41	84	93	48
3	58	195	51	128	120	116
4	135	125	64	67	22	198
5	182	72	91	112	176	76
6	132	178	173	127	125	123
7	98	127	158	67	158	123
8	152	130	13	182	76	17
9	162	141	69	20	135	19
10	178	93	182	103	156	179

The data in table I was analyzed by using the DEA model (2) and the results are shown in table II and table III. It can be seen in table II that there are seven DMUs as efficient identifies DMU 2, DMU 3, DMU 4, DMU 5, DMU 7, DMU 8, and DMU 10. Table III presents that the DEA model obtains a high number of zero optimal weights, showing that it does not emphasize input and output variables.

TABLE II
EFFICIENCY SCORES OF THE DEA MODEL

DMU	1	2	3	4	5	6	7	8	9	10
Efficiency	0.7194	1	1	1	1	0.8341	1	1	0.8148	1

In Table IV, the efficiency scores of CIDEA models (the bounds on weight at l = 0.4, 0.45, 0.5, 0.6, 0.7 and 1) and MO-CVDEA models are shown. Table IV shows that four DMUs are efficient in the bounds at l = 0.4 of CIDEA model and six DMUs are efficient in the MO-CVDEA model. Therefore, the CIDEA model (l = 0.4) is relatively better in ranking DMUs than the MO-DEA model and the worst case of the CIDEA model (l = 1) is still able to show indifferent ranking with the MO-CVDEA model. The CIDEA model at l = 1 was the worst case and at l = 0.4 was the best case of the CIDEA model.

In Table V, it is seen that the weight of the 3th input variable of DMU 6, DMU 7, and DMU 10 in MO-CVDEA model are zero. It means this model still does not emphasize that input.

TABLE III
THE WEIGHTS OF THE DEA MODEL

DMU	v_1	v_2	v_3	u_1	u_2	u_3
1	0.01382	0.00230	0.00000	0.00000	0.01043	0.00000
2	0.00114	0.01531	0.00026	0.01190	0.00000	0.00000
3	0.00385	0.00398	0.00000	0.00515	0.00000	0.00293
4	0.00000	0.00680	0.00235	0.00366	0.00000	0.00381
5	0.00000	0.01389	0.00000	0.00000	0.00568	0.00000
6	0.00316	0.00327	0.00000	0.00423	0.00000	0.00241
7	0.00430	0.00221	0.00188	0.00000	0.00633	0.00000
8	0.00469	0.00221	0.00000	0.00549	0.00000	0.00000
9	0.00150	0.00085	0.00923	0.00000	0.00604	0.00000
10	0.00000	0.00817	0.00132	0.00000	0.00000	0.00559

TABLE V
THE WEIGHTS OF THE MO-CVDEA MODEL

DMU	v_1	v_2	v_3	u_1	u_2	u_3
1	0.00326	0.00594	0.00153	0.00305	0.00285	0.00147
2	0.00535	0.00728	0.00247	0.00495	0.00496	0.00256
3	0.00435	0.00344	0.00152	0.00331	0.00305	0.00180
4	0.00243	0.00438	0.00194	0.00299	0.00119	0.00390
5	0.00261	0.00583	0.00116	0.00312	0.00303	0.00154
6	0.00373	0.00285	0.00000	0.00257	0.00225	0.00150
7	0.00382	0.00492	0.00000	0.00294	0.00313	0.00182
8	0.00246	0.00455	0.00269	0.00465	0.00192	0.00051
9	0.00224	0.00390	0.00125	0.00186	0.00300	0.00088
10	0.00287	0.00525	0.00000	0.00285	0.00257	0.00171

TABLE IV
EFFICIENCY SCORES OF THE CIDEA AND MO-CVDEA MODELS

DMU	MO-CVDEA	CIDEA					
		l =1	l = 0.7	l = 0.6	l = 0.5	l = 0.45	l = 0.4
1	0.2835	0.5663	0.5033	0.4760	0.4484	0.4331	0.4194
2	1	1	1	1	0.9859	0.9442	0.9070
3	1	1	1	1	1	1	0.9695
4	1	1	0.9136	0.8630	0.8116	0.7804	0.7504
5	1	1	1	1	1	1	1
6	0.7920	0.8334	0.8318	0.8264	0.7873	0.7326	0.6836
7	0.9164	1	1	1	1	1	1
8	1	1	1	1	1	1	1
9	0.4590	0.7732	0.7141	0.6482	0.5807	0.5434	0.5100
10	1	1	1	1	1	1	1

TABLE VI
THE WEIGHTS OF THE CIDEA MODEL AS $l = 1$

DMU	v_1	v_2	v_3	u_1	u_2	u_3
1	0.00739	0.00391	0.00202	0.00058	0.00674	0.00354
2	0.00399	0.01101	0.00016	0.00695	0.00265	0.00354
3	0.00739	0.00278	0.00056	0.00148	0.00674	0.00001
4	0.00274	0.00366	0.00269	0.00477	0.00001	0.00343
5	0.00295	0.00333	0.00245	0.00001	0.00415	0.00354
6	0.00316	0.00327	0.00001	0.00422	0.00001	0.00241
7	0.00495	0.00404	0.00001	0.00335	0.00291	0.00257
8	0.00399	0.00259	0.00436	0.00389	0.00384	0.00001
9	0.00301	0.00150	0.00436	0.00001	0.00572	0.00001
10	0.00001	0.01071	0.00001	0.00510	0.00001	0.00264

TABLE VII
THE WEIGHTS OF THE CIDEA MODEL AS $l = 0.4$

DMU	v_1	v_2	v_3	u_1	u_2	u_3
1	0.00490	0.00461	0.00265	0.00348	0.00441	0.00230
2	0.00491	0.00794	0.00251	0.00461	0.00441	0.00230
3	0.00395	0.00386	0.00036	0.00428	0.00129	0.00230
4	0.00366	0.00386	0.00036	0.00400	0.00129	0.00230
5	0.00238	0.00454	0.00265	0.00148	0.00374	0.00230
6	0.00190	0.00386	0.00036	0.00243	0.00129	0.00173
7	0.00462	0.00386	0.00036	0.00263	0.00342	0.00230
8	0.00305	0.00386	0.00265	0.00460	0.00201	0.00065
9	0.00169	0.00386	0.00265	0.00148	0.00347	0.00065
10	0.00258	0.00386	0.00100	0.00148	0.00279	0.00230

Table VI and Table VII show the weights of input and output variables of the CIDEA models (l = 1 and l = 0.4), indicating the weights of input and output variables are not zero.

For each model, using the following hypotheses, it is proven that the two approaches can obtain the similar rankings to DMUs.

H_0 : The DMU efficiency scores for the CIDEA model are uncorrelated to the DMU efficiency scores for the DEA model.
H_1 : The DMU efficiency scores for the CIDEA model are correlated to the DMU efficiency scores for the DEA model.

In testing the hypotheses, the Spearman and Kendall correlations were used (see Table VIII and IX).

TABLE VIII
THE SPEARMAN'S CORRELATIONS BETWEEN THE EFFICIENCY SCORES OF THE CIDEA MODELS AND THE DEA MODEL

Model	l = 1	l = 0.7	l = 0.6	l = 0.5	l = 0.45	l = 0.4
Spearman Correlation	1	0.9157	0.9157	0.8670	0.8670	0.8386

TABLE IX
THE KENDALL'S CORRELATIONS BETWEEN THE EFFICIENCY SCORES OF THE CIDEA MODELS AND THE DEA MODEL

Model	l = 1	l = 0.7	l = 0.6	l = 0.5	l = 0.45	l = 0.4
Kendall Correlation	1	0.8944	0.8944	0.8281	0.8281	0.7845

Tables VIII and IX show that there are high correlations between the efficiency scores of DMUs and hence a similarity between the ranks obtained by the CIDEA and the DEA models for each case.

Next, the comparison of weight variances corresponding to the CIDEA models and the DEA models which the following hypotheses were tested at significance level 0.05 ($\alpha = 0.05$). (see Table X)

H_0 : There is no difference between the weight variances corresponding to the CIDEA models and the DEA models.
H_1 : The weight variances for the CIDEA models are less than that of the DEA model.

In this study, the sample sizes are large. By the law of large number, the F test can be used to test the hypotheses.

TABLE X
RESULTS OF TESTING THE HOMOGENEITY OF WEIGHT VARIANCES FROM THE CIDEA MODELS AND THE DEA MODEL

Model	l = 1	l = 0.7	l = 0.6	l = 0.5	l = 0.45	l = 0.4
P-value	0.0004	2.37E-05	8.98E-08	6.87E-08	8.34E-09	1.10E-12

The comparison of the weight variances corresponding to the CIDEA models and the MO-CVDEA models which the following hypotheses were then tested. (see Table XI)

H_0 : There is no difference between the weight variances corresponding to the CIDEA models and the MO-CVDEA models.
H_1 : The weight variances for the CIDEA models are different to that of the MO-CVDEA model.

TABLE XI
RESULT OF TESTING THE HOMOGENEITY OF WEIGHT VARIANCES FROM THE CIDEA MODEL AND THE MO-CVDEA MODEL

Model	l = 1	l = 0.7	l = 0.6	l = 0.5	l = 0.45	l = 0.4
P-value	0.0002	0.0026	0.0755	0.0849	0.1922	0.7546

Table X and XI show the results of the homogeneity test. The dispersion of input-output variable weights of the CIDEA model is less than that of the DEA model in all cases, however, the dispersion of weights of the CIDEA model is not different to the one of the MO-CVDEA model in cases l = 0.4, l = 0.45, l = 0.5, and l = 0.6.

ISBN: 978-988-14047-6-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

In summary, the CIDEA approach attaches importance to all input-output variables while the DEA and the MO-CVDEA ignored some variables. From the results of efficiency scores, it is clear to see the CIDEA approach can reduce the number of efficient DMUs better than the MO-CVDEA approach.

VI. Conclusion

The CIDEA models were used to overcome the problem of DEA. Since the DEA models yielded several zero input-output weights or provided extreme values, this lead to many efficient DMUs not becoming ranked. Several advantages of this model over the MO-CVDEA were discovered. Firstly, the efficiency scores can be calculated within 2 stages instead of 3 iteration stages. Secondly, since the input-output weights of this model do not become zero, this model emphasize all input-output variables, Moreover this model is linear optimization model so it can be successfully formulated to handle dual problems.

Acknowledgment

The authors are grateful to the anonymous referees for valuable comments and suggestions.

References

[1] Gencer, C., and Gürpinar, D. “Analytic network process in supplier selection: A case study in an electronic firm.” *Applied Mathematical Modelling*, vol. 31, No. 11, pp. 2475-2486, 2007.

[2] Hong G. H., Park S. C., Jang D. S., and Rho H. M. “An Effective Supplier Selection Method for Constructing a Competitive Supply-Relationship.” *Expert Systems with Applications*, vol. 28, No. 4, pp. 629-639, 2005.

[3] Liu J., Ding F. Y., and Lall V. “Using Data Envelopment Analysis to Compare Suppliers for Supplier Selection and Performance Improvement.” *Supply Chain Management: An International Journal*, vol. 5, No. 3, pp. 143-150, 2000.

[4] Youssef M. A., Zairi M., and Mohanty B. “Supplier Selection in an Advanced Manufacturing Technology Environment: an Optimization Model.” *Benchmarking for Quality Management & Technology*, vol. 3, No. 4, pp. 60-72, 1996.

[5] Ghodsypour S. H., and O’Brien C. “The Total Cost of Logistics in Supplier Selection, Under Conditions of Multiple Sourcing, Multiple Criteria and Capacity Constraint.” *International Journal of Production Economics*, vol. 73, No. 1, pp. 15-27, 2001.

[6] Huaqing W., Jie W., Liang L., and Yongjun L. “Efficiency assessment of Chinese logistics firms using DEA.” *International Journal of Shipping and Transport Logistics*, vol. 4, No. 3, pp. 212-234, 2012.

[7] Sten T., George K., and Fred P. “DEA of financial statements data: The U.S. computer industry.” *Journal of Productivity Analysis*, vol. 5, pp. 229-248, 1994

[8] Charnes, Cooper, and Rhodes. “Measuring the Efficiency of Decision Making Units.” *European Journal of Operational Research*, vol. 2, pp. 429-444, 1978.

[9] Farrell. “The Measurement of Productive Efficiency.” *Journal of Royal Statistical Society*, vol. A120, pp. 253-293, 1957.

[10] Thompson, G. Russell., F. D. Singleton, Jr., and Robert M. Thrall. “Comparative Site Evaluations for Locating a High-Energy Physics Lab in Texas.” *Interface*, vol. 16, pp. 35-49, 1986.

[11] Allen, R. Athanassopoulos, A. Dyson, R.G., and Thanassoulis, E. “Weights restrictions and value judgements in data envelopment analysis: Evolution, development and future directions.” *Annals of Operations Research*, vol. 73, pp. 13-34, 1997.

[12] Dyson, R.G., and Thanassoulis, E. “Reducing weight flexibility in data envelopment analysis.” *Journal of the Operational Research Society*, vol. 39, pp. 563-576, 1988.

[13] Bal, H. Orkcu, H.H., and Celebioglu, S. “A new method based on the dispersion of weights in data envelopment analysis.” *Computer & Industrial Engineering*, vol. 54, pp. 502-512, 2008.

[14] Jahanshahloo, G.R., and Shahmirzadi, P.F. “New methods for ranking decision making units based on the dispersion of weight and Norm 1 in Data Envelopment Analysis.” *Computer & Industrial Engineering*, vol. 65, pp. 187-193, 2013.

[15] Banker, Charnes, and Cooper "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis.” *Management Science*, vol. 30, pp. 1078–1092, 1984.