Benchmarking: An International Journal

Application of PCA and DEA to recognize the true expertise of a firm: a case with primary schools

Subhadip Sarkar

Received 6 July 2014
Accepted 16 November 2014

Application of PCA and DEA to recognize the true expertise of a firm: a case with primary schools

Subhadip Sarkar
Department of Management Studies, NIT Duragapur, West Bengal, India

Abstract

Purpose – The purpose of this paper is to identify important dimensions which are essential to remain competitive. To generate scores which will be as effective as the original outputs to determine the radial efficiency scores etc.

Design/methodology/approach – A non-central principal component analysis (PCA) were used to determine various dimensions for each output. The objective was set to identify those special schools which could minimize certain pre-prescribed scores.

Findings – Few schools were trying to concentrate on the students from the rich society and spending less per student. There were other schools which targeted to minimize the social loss by providing education to the poorer section and were funding more for them.

Research limitations/implications – Small group was considered. However, the number can be extended.

Practical implications – The valuable findings of Hillman and Jenkner (2002), stated that – "Children are entitled to a free, quality basic education. Many children who do attend school receive an inadequate education because of poorly trained, underpaid teachers, overcrowded classrooms, and a lack of basic teaching tools such as textbooks, blackboards, and pens and paper […]." "In an ideal world, primary education would be universal and publicly financed, and all children would be able to attend school regardless of their parents' ability or willingness to pay. The reason is simple: when any child fails to acquire the basic skills needed to function as a productive, responsible member of society, […] The cost of educating children is far outweighed by the cost of not educating them. Adults who lack basic skills have greater difficulty in finding well-paying jobs and escaping poverty […]." In order to understand which fact has been stressed more the proposed model is very useful.

Social implications – It is capable of describing the current standpoint of a group of homogenous schools or firms. Quality and cost cutting principals can be isolated quite easily.

Originality/value – Introduces concepts of non-central PCA. Provides alternative scores which are as important as the original output. Detects and analyze various important dimensions.

Keywords Benchmarking, Principal component analysis, Data envelopment analysis, Organizational performance, Properties of positive definite matrix

Paper type Research paper

Nomenclature

Symbols	Meaning
r	represents any rth DMU where $r = 1,2, \ldots, c$
i	represents any ith input $i = 1,2, \ldots, v$
j	represents any jth output $j = 1,2, \ldots, m$
$PROJ_{rij}$	projection of any rth DMU on E_{ij}
$[R_{ij}]_{c,v}$	resource matrix of c number of DMUs
$[y_{ij}{}^{obs}]_{c,m}$	output matrix of c number of DMUs
u_i, q_j	weight vector on input and output
S_j	slack variable
λ_r	weight used on any rth DMU
$x^r{}_{ij}$	specific usage (SU) of the ith input for any jth output of the rth DMU
Sp_r	specific consumption matrix

Benchmarking: An International Journal
Vol. 23 No. 3, 2016
pp. 740-751

1463-5771
DOI 10.1108/BIJ-11-2014-0100

E_{ij}	ith principal component of Sp_j derived from any jth output	T_{rj}	specific usage vector for any jth output of the rth DMU

1. Introduction

According to Porter (1985) the fundamental basis of above average performance in the long run is due to the sustainable competitive advantage which a firm can possess by low cost or differentiation. Appropriate strategies can provide strength to survive the stress of the five forces in a better way than its rivals in the long run. This practice makes a firm to remain efficient and to become a benchmark in future. This perpetual urge compels a DMU to assess its performance among its rivals and to remain an efficient performer while using fewer quantities of each input to generate the same set of outputs or producing more outputs from the same set of input resources than its. Data envelopment analysis (DEA) plays a major role for the assessment of performance. The journey of DEA commenced when the performance of students from participating and not participating schools were compared by Charnes *et al.* (1978), using a non-linear model and an equivalent data-oriented, linear programming-based approach. A CCR-efficient DMU was found to be scoring redial efficiency of one and did not have any kind of input or output slacks. Later on, the assumption of constant return on scale, was extended by Banker *et al.* (1984). The renowned BCC model of these researchers was able to administer variable scaling techniques. Till the year of 1989 the application of DEA was restricted to the non-negative data as DMU scoring negative value on a variable were eliminated.

The combined effort of Ali and Seiford (1990) (Cooper *et al.*, 2002/2011) added a new property in DEA called "translational invariance." However, they never made any comments on zero or on negative data. Pastor (1993) was the first who applied this Theorem to all three basic models of DEA for solving the problem of measuring the performances of 23 bank branches. A data transformation process was applied to turn the negative values into positive values. This new form of data was fed into those models which were invariant to translation. Moreover, he showed (Pastor, 1996), that a displacement does not alter the efficient frontier for certain DEA formulations (specifically, the additive model for both inputs and outputs and the BCC model for outputs (or inputs)) and thus these approaches are translation invariant. Though, additive models were found most efficient in this regard, the solution was not unit invariant and "it yields in respect of an inefficient unit the 'furthest' targets on the production frontier" (Portela *et al.*, 2004). A completely new thought came up as Portela *et al.* (2004), proposed a range directional model, based on directional distance functions from a so called Ideal point, for helping a Portuguese bank to manage the performance of its branches. The bank wanted to set targets for the branches on such variables as growth in number of clients, growth in funds deposited and so on. These variables could take positive and negative values, preventing the use of traditional DEA. They pointed out inefficiency of a firm in comparison to deviation seen from the Ideal point in the context of input as well as output.

A case of few primary schools, presented in this paper, considers two inputs like spending per student and percentage of income not from poor group. The reason for taking them into account is that in many developing countries, the governments' lack either the financial resources or the political will to meet their citizens' educational needs. The valuable findings of Hillman and Jenkner (2002), in this regard can be cited here as – "Children are entitled to a free, quality basic education. Many children who do attend school receive an inadequate education because of poorly trained, underpaid teachers, overcrowded classrooms, and a lack of basic teaching tools such as textbooks, blackboards, and pens and paper […].” The inclusion of the first input is due to the measurement of willingness of a primary school to impart education. Commenting on the ill-effects they mentioned “In an ideal world, primary education would be universal and publicly financed, and all children would be able to attend school regardless of their parents’ ability or willingness to pay. The reason is simple: when any child fails to acquire the basic skills needed to function as a productive, responsible member of society, […] The cost of educating children is far outweighed by the cost of not educating them. Adults who lack basic skills have greater difficulty in finding well-paying jobs and escaping poverty […].” Thus, the second input plays a key role to measure the intention of a primary school to serve for the social benefit.

The proposed model, rather, concentrates on the generating scores in various dimensions using principal component analysis (PCA). There is an instance of PCA application when Nicole Adler and Boaz Golani (L.M. Seiford, 1989), adopted a PCA DEA model in a case study of municipal solid waste, in the Oulu district of Finland, for curtailing the number of analyzed variables by grouping highly correlated variables within a factor. In their second model PCA was applied separately on the input and output variables for strengthening the power of DEA. However, they had never spoken of about the exploration of expertise of a firm using PCA. The first principal eigenvector of a specific consumption matrix of an output reflects the cost (as it has all positive elements), whereas, the remaining principal vectors denote several other attributes essential for gaining competitive advantage (as it contains one negative element and thus represent dexterity of a firm manage that essential dimension). These scores are treated as new inputs (which can assume positive and negative values) and used along with the original outputs to produce efficiency scores.

2. Definitions and theorems

2.1 Data envelopment analysis with BCC model

From an assumption of constant returns to scale, Banker *et al.* (1984) found proportional changes in weighted output that derive from the alterations in weighted inputs (Table I).

The algebraic models of variable return to scale for c DMUs (each of which consumes v inputs given by the $R=[R_{ij}]_{c,v}$ to generate m number of outputs (observed from the c number of homogeneous systems considered) given by $Y=[y_{ij}{}^{obs}]_{c,m}$) are shown above.

Primal form	Dual form
$Max\ h_0=\left(\sum_{j=1}^{m} q_j.y_{rj}\right)-w_0$	$Min\ \theta_0-\epsilon\left(\sum_{i=1}^{v} S_i+\sum_{j=1}^{m} S_j\right)$
Subjected to: $\sum_{j=1}^{v} v_j.R_{rj}=1$;	Subjected to: $\theta_0.R_{ri}+S_i=\sum_{r=1}^{c}\lambda_r.R_{ri};\ S_i,\lambda_r\geqslant 0$
$\sum_{i=1}^{v} u_i.R_{ri}\geqslant\sum_{j=1}^{m} q_j.y_{rj}{}^{obs}+w_0$;	For any jth input $i=1,2,\ldots,v$
$u_i, q_j\geqslant\in$; $w_0=$ unrestricted	$y_{rj}{}^{obs}=S_j+\sum_{r=1}^{c}\lambda_r.y_{rj}{}^{obs}$ for $j=1,2,\ \ldots,\ m$
Where $\in=$ non-Archemedean value	$S_i\geqslant 0$ for $j=1,2,\ \ldots,\ m;\ \sum_{r=1}^{c}\lambda_r=1;$
For any DMU r	

Table I. Primal and dual form of BCC DEA

2.2 A BCC-efficient unit

A DMU is called BCC-efficient if $\theta^* = 1$, and if there exists at least one optimal solution (u^*, q^*), for which $u^* > 0$ and $q^* > 0$, otherwise, the DMU in question is considered to be BCC-inefficient. A Solution (u^*, q^*) from BCC-inefficient units ($\theta^* < 1$), must necessarily involve at least one DMU (known as a peer group) within the given set that manages to yield weighted outputs that are equivalent to its weighted inputs. The set of peer groups is specified as: $E_0' = \{r : \sum_{j=1}^{m} q_j y_{rj}^{obs} = \sum_{i=1}^{v} u_i R_{ri}\}$

2.3 Specific consumption matrix T_j and specific covariance matrix S_j

Specific consumption matrix, T_j, for any jth output, is a column matrix, filled with all specific consumptions of resources. A positive definite covariance matrix S_j (having with a non-zero determinant) derived from T_j is defined as follows:

$$S_j = T_j^T . T_j = \{s_{ij}\}_{vxv} \text{ where } s_{ij} > 0;\ T_j = \{t_{ij}^r\}_{cxv} \text{ and } t_{ij}^r = \frac{R_{ri}}{y_{rj}} \tag{1}$$

$$T_j^T = \begin{bmatrix} T_{1j} & T_{2j} & , \ldots, & T_{cj} \end{bmatrix} \text{ where } T_{rj}^T = \begin{bmatrix} t_{1j}^r & t_{2j}^r & , \ldots, & t_{vj}^r \end{bmatrix}$$

$$s_{ij} = \begin{pmatrix} \sum_{r=1}^{c} \left(\frac{R_{ri}}{y_{rj}}\right)^2 & i = i' \\ \sum_{r=1}^{c} \left(\frac{R_{ri}}{y_{rj}}\right)\left(\frac{R_{ri'}}{y_{ri'}}\right) & i \neq i' \end{pmatrix} \tag{2}$$

t_{ij} is known as the specific usage of the ith input of the rth DMU.

2.4 A non-central PCA and its application on specific covariance matrix S_j

Economic use of resources of a DMU should be tested from the origin and not from any mean vector derived from a group of DMUs. To observe the mutually independent underlying characteristics of resource utilization the specific consumption matrix is projected on a unit vector so that the directions of maximum variance (from the origin vector and not from their mean vector) can be explored. This leads to the following optimization problem to be solved:

$$Max\, z = \gamma^T . T_j^T . T_j . \gamma = \gamma^T . S_j . \gamma; \text{ subjected to: } \gamma^T \gamma = 1;$$

The optimal solution of this problem gives rise to eigenvectors of S_j which are orthogonal to each other. Appendices 1 and 2 can be referred here, to display the properties of these eigenvectors.

This problem is similar to a non-central PCA. According to Rencher (2002), PCA deals with a single sample of n observation vectors $\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_n}$ that form an ellipsoidal swarm of points in a p-dimensional space. If the variables $y_1, y_2, \ldots, y_p$ in $\mathbf{y}$ are correlated, the natural axes of the swarm of points become identical to with the axes of the ellipsoid having an origin at the mean vector ($\mathbf{y}^*$) of $\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_n}$. The resulting natural axes of the ellipsoid, yield the new uncorrelated variables called (principal components). These resulting axes will be similar to the eigenvectors (E_r) derived from the covariance matrix $[S]_{pxp}$ (or the correlation matrix $([R]_{pxp})$ of the observed variables which also minimizes the mean squared distance between the data points and their projections (shown below):

$$S.E_r = \gamma_r.E_r \text{ for } r = 1, 2, \ldots, p \text{ such that } \gamma_1 > \gamma_2 > \ldots > \gamma_p$$

But, unlike PCA, the proposed model determines the covariance matrix from the origin.

2.5 Economic interpretation of principal components of the matrix S_j

Being a square matrix of size (v x v), S_j, has v number of eigenvectors (and eigenvalues). These vectors carry significant information about the usage of all ingredients. Other than the first vector none of the remaining ones assume all positive elements (shown in the Appendices 1 and 2). The first eigenvector acknowledges the cost consciousness of a firm as less projected value on this vector implies the lower combined consumption of inputs. The reason of calling it "cost" or "combined spending" is that, the firm in view of acquiring future benefits would like to concentrate on the current collective expenditure. Remaining dimensions (which reflect unique capacity of a firm) are indeed essential to gain various competitive advantages. Each of these vectors has its own priority level (equivalent to the corresponding eigenvalue) set by the industry. Baring this, they contain one negative element which is indicative of the worth of a particular resource over the rest for reducing the cost due to that dimension. Therefore, the firm has to be more decisive in managing the cost and the right dimension to sustain in the market. Therefore, the proposed model lies on the balance between first, reduction of "cost" (which focusses on decreasing the utilization of resources) and second, reduction of cost from the remaining dimensions (by manipulating proper resources).

2.6 Derivation of scores under various dimensions

Each eigenvector derived from a specific covariance matrix, due to any output, is assumed to be representing important orthogonal traits or dimensions to produce the same. Therefore, for any v number of inputs and m number of outputs, there will be mv number of traits (shown in Table II with their priority levels).

If the weight of any output j is assumed to be w_j, then the total score obtained by any rth DMU at any ith priority level due to its current consumption of resources (R_r) is given below:

$$SCORE_{ri} = \sum_{j=1}^{m} \sum_{i=1}^{v} w_j.R_r{}^T.E_{ij}$$

Table II.
Weight vector of output set

Eigenvector	Priority level	Output 1	Output 2	…	Output m
Eigenvector 1	1 (top)	E_{11}	E_{12}	–	E_{1m}
Eigenvector 2	2	E_{21}	E_{22}	–	E_{2m}
…				E_{ij}	
Eigenvector v	v (least)	E_{v1}	E_{v2}	–	E_{vm}
Output weight		w_1	w_2	–	w_m

Assigning equal weights to all outputs, a total score of an rth DMU on any ith dimension (or priority level) is determined by the following expression:

$$SCORE_{ri} = \sum_{j=1}^{m}\sum_{i=1}^{v} R_{ri}{}^{T}.E_{ij} = \sum_{j=1}^{m}\sum_{i=1}^{v} PROJ_{rij}$$

3. Correspondence between an output-oriented BCC DEA on both data file and the translated derived score matrix

Theorem. The radial efficiencies resulting from an output-oriented BCC DEA on the given data and the translated derived score matrix, are same.

Proof. Assuming θ_0, S_i, S_j, R_{ri} and λ_r as radial efficiency, slack variables due to ith input and jth output, ith input resource consumed by any rth DMU and weights ascribed on any rth DMU, respectively, the envelopment form of the BCC model is shown below.

Envelopment form:

$$Max = \theta_0 + \epsilon\left(\sum_{i=1}^{v} S_i + \sum_{j=1}^{m} S_j\right)$$

Subjected to:

$$R_{ri} + S_i = \sum_{r=1}^{c} \lambda_r.R_{ri};\ S_i, \lambda_r \geqslant 0$$

For any jth input $i = 1,2, \ldots, v$:

$$\theta_0.y_{rj}{}^{obs} = S_j + \sum_{r=1}^{c} \lambda_r.y_{rj}{}^{obs} \text{ for } j = 1, 2, \ldots, m$$

$$\sum_{r=1}^{c} \lambda_r = 1; S_i \geqslant 0 \text{ for } j = 1, 2, \ldots, m$$

The constraints pertaining to inputs in the envelopment form of banker's BCC model (shown in Section 2.1) is further multiplied with various eigenvectors signifying key dimensions of resource handling (shown below). The entire multiplication process gives rise to same number of constraints as before:

$$\sum_{j=1}^{m} \left[E_j{}^{T}\right]_{vxv} [R_{oi}]_{vx1} + \sum_{j=1}^{m} \left[E_j{}^{T}\right]_{vxv} [S]_{vx1_i} = \sum_{j=1}^{m} \left[E_j{}^{T}\right]_{vxv} \left[R^{T}\right]_{vxc}.[\lambda]_{cx1};$$

The properties defined in the Appendix 2 clarifies that there can be one solution or which at any value of j, the ith column of $[E_j]$ contains a negative element at (i−1)th row. Thus, the first and the third element of the above equation can contain few non-positive values. Apart from that, the new slacks originated from the second term now remain unrestricted (barring the first one). Being aware of the difficulties of handling the negative data shown by Portela *et al.* (2004), a translational process is carried out in order to turn all negative numbers into positive (applied only on inputs). Since, the input slacks present in the objective function (shown below) are of unrestricted type (barring the first one), the maximization process on the strictly non-negative data will force these negative slacks (S_i') to take zero values (since, according to the property of translational invariance, an output-oriented BCC DEA remains unchanged when subjected to a translational change on the input side):

$$Max = \theta_0 + \epsilon\left(\sum_{i=1}^{v}\left(S_i - S_i'\right) + \sum_{j=1}^{m} S_j\right) \text{such that} \left[S - S'\right] = \sum_{j=1}^{m}\left[E_j^{\,T}\right]_{vxv}[S]_{vx1}$$

$$\text{and } S = [S_i]_{vx1};\, S' = \left[S_i'\right]_{vx1} \text{ where } S_i, S_i' \geqslant 0$$

Hence, it is proved that in spite of adopting a combined process of transformation and translation the output-oriented BCC DEA will give identical results for the previously mentioned two cases. ■

4. A mathematical example

The research was initiated from a very fundamental question "How does one verify the cost leadership perspective of a primary school among many other primary schools." Performances of six schools are analyzed based on the effort given to the students (spending per student (I1) and the financial condition of a student represented in terms of average percent that he or she does not belong to low income group (I2)) and results of two tests taken on them (average writing score per student (O1) and average science score per student (O2)) (Table III). The objective of this study is to identify the list of efficient schools which remain competitive by cost minimization. Any economic performance is inclined to producing more output scores by spending lower amount per pupil and also giving more opportunities to the poor.

As an initial step of making discrimination, productivity scores from BCC DEA (Table IV) effectively isolate the Farrell inefficient schools B and D from the Farrell efficient schools A, C, E and F (considering the local scale of operation). The inefficient ones are dominated by respective hypothetical firms made from various linear compositions of A and E, C and E.

Table V reveals an important fact that being efficient schools (like A, C, E and F), the weight vectors (weight vector, (***u****, ***q****), for any school is denoted by W (name of the school, input/output) and derived from running the Primal BCC DEA model on the corresponding school) do possess few zeroes. However, in reality all of them are more than non-Archimedean values.

Table III.
Data

Schools	Input 1 (I1)	Input 2 (I2)	Output 1 (O1)	Output 2 (O2)
A	8,939	64.3	25.2	223
B	8,625	99	28.2	287
C	10,813	99.6	29.4	317
D	10,638	96	26.4	291
E	6,240	96.2	27.2	295
F	4,719	79.9	25.5	222

For testing of the mix inefficiencies among these DMUs, the output (Table VI) of a simple slack-based DEA model is analyzed. Owing to the zero scores of slacks for schools like A, C, E and F, they are classified under BCC (or technically efficient (locally)) efficient schools.

But, to fulfill the present need of the problem, the collective practice of schools is illustrated here from the covariance matrix, eigenvalues and eigenvectors, pertaining to the embedded PCA, (Table VII). These vectors are representative of two important dimensions and are used to generate scores (as described in Section 2.7) on them (Table VIII). The first

		Composition of the hypothetical peer firm					
Productivity	Value	A	B	C	D	E	F
Score (A)	1	1	0	0	0	0	0
Score (B)	0.9948	0	0	0.5215	0	0.4785	0
Score (C)	1	0	0	1	0	0	0
Score (D)	0.9466	0.1020	0	0	0	0.8980	0
Score (E)	1	0	0	0	0	1	0
Score (F)	1	0	0	0	0	0	1

Table IV. BCC DEA output

Weights	W(A,I1)	W(A,I2)	W(A,O1)	W(A,O2)	W(B,I1)	W(B,I2)	W(B,O1)	W(B,O2)
Value	0	0.00472	0.03968	0	1.70597E–05	0	0.03546	0
Reduced cost	0	0	0	0	0	1.02677	0	17.9739
Weights	W(C,I1)	W(C,I2)	W(C,O1)	W(C,O2)	W(D,I1)	W(D,I2)	W(D,O1)	W(D,O2)
Value	0	0	0.03401	0	0	0.00915	0	0.00344
Reduced cost	0	0	0	0	16.11615	0	1.0826	0
Weights	W(E,I1)	W(E,I2)	W(E,O1)	W(E,O2)	W(F,I1)	W(F,I2)	W(F,O1)	W(F,O2)
Value	1.76869E–05	0	0.03676	0	1.7E–05	0.0025	0.03922	0
Reduced cost	0	0	0	0	0	0	0	0

Table V. Results of ratio model of BCC (input and output weights)

Slacks	A	B	C	D	E	F
S1 for Input 1	0	0.0028	0	16.128	0	0
S2 for Input 2	0	1.0267	0	0	0	0
S3 for Output 1	0	0	0	1.0826	0	0
S4 for Output 2	0	17.974	0	0	0	0

Table VI. Output of slack-based DEA

Type (eigenvalue)	From Output 1 (603,951.6945)		From Output 2 (5,909.7949)	
S Matrix	603,890.4534	6,081.332298	5,909.2	59.2503
	6,081.332298	65.86168651	59.25	0.6455
Eigenvector 1	0.9999493	0.010069824	0.9999	0.01
Eigenvector 2	−0.01006982	0.9999493	−0.01	0.9999

Table VII. The eigenvalue and eigenvector of the covariance matrix one is referred to as cost (since it bears all positive elements) whereas uniqueness is termed for the second one (as it accepts one negative element). The dimension termed as "cost" (on both occasions) claims near about 100 times of spending per student than its counterpart. This fact points out the societal practice of paying more preference to spending than educating the poor people. A successful DMU may spend less and focus on the students from the richer-section. On the contrary, "uniqueness" points out those schools which may be pleased to impart education to poorer section while spending more. As a result, Score 1 has all positive values and few DMUs find negative values for Score 2. A negative value implies for more focus on the uniqueness dimension than the standard mark (mentioned as zero here).

The post translational values of Score 2 is fed to the output-oriented BCC DEA model. no significant discrepancy is observed while measuring radial output-oriented efficiency mentioned in the Tables IV and IX. There is no alteration found in the efficient list too.

To confirm that their strongly efficient position the slack-based model is run. Quite interestingly, the so called efficient DMUs scored zero slack values but the rests came up with slack values different from Table V (Table X).

5. Conclusion

The current model has few added advantages as well. Like a regular BCC DEA model it is capable of generating identical radial efficiency scores. Apart from that, it is

Table VIII. Projection in the direction of principal components

Schools	A	B	C	D	E	F
PROJ 11	8,939.19	8,625.56	10,813.5	10,638.4	6,240.65	4,719.57
PROJ 12	8,938.75	8,625.13	10,812.9	10,637.9	6,240.34	4,719.33
PROJ 21	−25.717	12.1428	−9.29	−11.128	33.3594	32.3765
PROJ 22	−25.096	12.7401	−8.54	−10.39	33.7904	32.702
Score 1	17,877.9	17,250.7	21,626.4	21,276.3	12,481	9,438.89
Score 2	−50.814	24.8829	−17.83	−21.517	67.1498	65.0785

Table IX. BCC DEA on translated data (composition of the hypothetical firm)

Productivity	Value	A	B	C	D	E	F
Score(A)	1	1	0	0	0	0	0
Score(B)	0.9948	0	0	0.5215	0	0.4785	0
Score(C)	1	0	0	1	0	0	0
Score(D)	*0.9494*	*0.1118*	*0*	*0*	*0*	*0.8882*	*0*
Score(E)	1	0	0	0	0	1	0
Score(F)	1	0	0	0	0	0	1

Table X. Output of slack-based DEA

Slacks	A	B	C	D	E	F
S1 for Input 1	0	0.0077	0	69.105	0	0
S2 for Input 2	0	2.0529	0	0	0	0
S3 for Output 1	0	0	0	1.1251	0	0
S4 for Output 2	0	17.974	0	0	0	0

beneficial for detecting weakly efficient and truly inefficient DMUs. The proposed model strives to explore few underlying dimensions emerging out of the industry practice. The first dimension, regarded as "cost," decides the position of a school according to the current industrial trend for the collective spending of inputs for a better society. A group of schools would like to minimize the "cost" (combined consumption) by reducing spending per student and taking inputs from the rich section. As a whole, they accept the fact that the spending has more weight than its counterpart. As a result, Score 1 values will be much low for them. On the contrary, the other group ensures their inclination toward educating more students from the poorer sections and spends more to educate them to protect from an immense societal loss in future. These schools would like to minimize Score 2.

References

Ali, A.I. and Seiford, L.M. (1990), "Translation invariance in data envelopment analysis", *Operations Research Letters*, Vol. 9 No. 6, pp. 403-405.

Banker, R.D., Charnes, A. and Cooper, W.W. (1984), "Some models for estimating technical and scale inefficiencies in data envelopment analysis", *Management Science*, Vol. 30 No. 9, pp. 1078-1092.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operational Research*, Vol. 2 No. 6, pp. 429-444.

Cooper, W.W, Seiford, L. and Tone, K. (2002/2011), *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Vol. 2, Kluwer Academic Publishers, New York, NY, pp. 42-43.

Hillman, A.L. and Jenkner, E. (2002), "User payments for basic education in low income country", International Monetary Fund Working Paper No. WP/02/182, Washington, DC.

Pastor, J.T. (1993), "Efficiency of bank branches through DEA: the attracting of liabilities", working paper, Universidad de Alicante, Alicante.

Pastor, J.T. (1996), "Translation invariance in data envelopment analysis: a generalization", *Annals of Operations Research*, Vol. 66 No. 2, pp. 93-102.

Portela, M.C.A.S., Thanassoulis, E. and Simpson, G. (2004), "Negative data in DEA: a directional distance approach applied to bank branches", *Journal of the Operational Research Society*, Vol. 55 No. 10, pp. 1111-1121.

Porter, M.E. (1985), *Competitive Advantage: Creating and Sustaining Superior Performance*, The Free Press, New York, NY.

Rencher, A.C. (2002), *Methods of Multivariate Analysis*, John Wiley & Sons, New York, NY.

Seiford, L.M. (1989), "A bibliography of data envelopment analysis", working paper, Dept. of Industrial Engineering and Operations Research, University of Amherst, MA.

Appendix 1. The highest eigenvalue of a positive definite matrix that contains entirely positive elements will always be greater than the highest diagonal element of that matrix

Let A be a positive definite matrix with all non-negative elements, and let x be the eigenvector corresponding to the eigenvalue, γ, then, from the definition of an eigenvalue, $[Ax - \gamma.Ix] = 0$ and therefore $det\ |A - \gamma.I| = 0$; must hold:

$$\left|A-\gamma.I\right| = \begin{bmatrix} a_{11}-\gamma & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22}-\gamma & \dots & a_{2n} \\ a_{1n} & a_{2n} & \dots & a_{nn}-\gamma \end{bmatrix} = 0; \tag{A1}$$

Thus, the linearized form of the first (n−1) rows and n columns are as follows:

$$\begin{array}{cccc} (a_{11}-\gamma).x_1 & a_{12}x_2 & \dots & a_{1n-1}x_{n-1} = -a_{1n}x_n \\ a_{12}.x_1 & (a_{22}-\gamma).x_2 & \dots & a_{2n-1}x_{n-1} = -a_{2n}x_n \\ a_{1n-1}x_1 & a_{2n}.x_2 & (a_{n-1n-1}-\gamma)x_{n-1} = -a_{n-1n}x_n & \end{array}$$

This can also be expressed as follows:

$$\gamma V_1 = \gamma \begin{bmatrix} x_1 \\ X_1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{1p} \\ a_{p1} & A_1 \end{bmatrix} \begin{bmatrix} x_1 \\ X_1 \end{bmatrix} \tag{A2}$$

The first set of linear equation represents $(\gamma-a_{11}).x_1 = a_{1p}.X_1 > 0$; which essentially refers to two conditions; $(\gamma > a_{11})$ when $x_1 > 0$ and $(\gamma < a_{11})$ when $x_1 < 0$. As a result, it can be interpreted that any ith element of an eigenvector will be positive if the corresponding eigenvalue is more than the ith diagonal element. Therefore, if an eigenvector contains all positive elements then the relationship $(\gamma > max(a_{11}, a_{22}, \dots, a_{nn}))$ must be true.

If another eigenvector V_2 (which is orthogonal to V_1) is considered with a negative element $-x_2$ where $x_2 > 0$. Then, the following equations will exist:

$$\gamma V_2 = \gamma \begin{bmatrix} -x_2 \\ X_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{1p} \\ a_{p1} & A_1 \end{bmatrix} \begin{bmatrix} -x_2 \\ X_2 \end{bmatrix} \tag{A3}$$

$$\begin{bmatrix} x_1 & {X_1}^T \end{bmatrix} \begin{bmatrix} -x_2 \\ X_2 \end{bmatrix} = 0 \tag{A4}$$

However, this will violate the condition $(\gamma > max(a_{11}, a_{22}, \dots, a_{nn}))$. Thus, an eigenvector with all positive elements can be generated only from the largest eigenvalue.

The second equation is given as $(\gamma I-A_1).X_1 = a_{p1}.x_1$. Using the first equation the following expression can be established:

$${X_1}^T(\gamma I-A_1).X_1 = \frac{{X_1}^T(a_{p1}a_{1p}).X_1}{(\gamma-a_{11})}x_1 \tag{A5}$$

For the largest eigenvalue, $\gamma-a_{11} > 0$; must be true. The eigenvector, corresponding to it, will necessarily make X_1, $x_1 > 0$ to happen and as a result it will also impose a positive definiteness to the $(\gamma I-A_1)$ matrix (as $a_{1p} > 0$).

Appendix 2. If A is a positive definite matrix (shown below), then the matrix of its eigenvectors, E_j, can have a special structure

$$[A]_{pxp} = \begin{bmatrix} a_{11} & a_{1p} \\ a_{p1} & A_1 \end{bmatrix} = \begin{bmatrix} a_{11} & \boldsymbol{a_{12}} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & & \ddots & \\ a_{p1} & a_{p2} & & a_{pp} \end{bmatrix}, \ \boldsymbol{a_{i,j}} \geqslant \boldsymbol{a_{k,l}} > 0 \quad \text{for } i \leqslant k,\ j \leqslant l;$$

$$[E_j]_{pxp} = \begin{bmatrix} E_{1j} & E_{2j} & \ldots E_{ij} & \ldots E_{pj} \end{bmatrix} = \begin{bmatrix} e_{11} & -\boldsymbol{e_{12}} & \ldots e_{1i} \ldots & e_{1p} \\ e_{21} & e_{22} & \ldots e_{2i} \ldots & e_{2p} \\ \vdots & \vdots & -\boldsymbol{e_{i-1,i}} \ddots & \\ e_{p-1,1} & e_{p-1,2} & e_{p-1,i} & -\boldsymbol{e_{p-1,p}} \\ e_{p1} & e_{p2} & \ldots e_{pi} \ldots & e_{pp} \end{bmatrix}$$

Any ith column of the matrix, $[E_j]_{pxp}$, which is due to the eigenvalue γ_i such that for $i = 1, 2, \ldots, p$; and $\gamma_i > \gamma_{i+1}$, must give a positive value for the sum of its elements or $1^T E_{ij} \geqslant 0$.

Proof. From the property of diagonalization any positive definite matrix can be expressed as $[A] = [E_j].[D].[E_j]'$ such that the matrix, $[E_j]$, remains orthogonal ($[E_j]'.[E_j] = [E_j].[E_j]' = I$). Moreover, from the *Theorem* of eigenvalue, $[A].[E_{ij}] = \gamma_i.[E_{ij}]$ will be satisfied for any ith eigenvalue, γ_i. Using these concepts the following three equations can be derived:

$$a_{11}.(-\boldsymbol{e_{12}}) + a_{12}.(\boldsymbol{e_{22}}) + \ldots + a_{1i}.(\boldsymbol{e_{i2}}) \ldots + a_{1p}.\left(\boldsymbol{e_{p2}}\right) = \gamma_2.(-\boldsymbol{e_{12}}) \tag{A6}$$

$$a_{11}.(\boldsymbol{e_{13}}) + a_{12}.(-\boldsymbol{e_{23}}) + \ldots + a_{1i}.(\boldsymbol{e_{i3}}) \ldots + a_{1p}.\left(\boldsymbol{e_{p3}}\right) = \gamma_3.(\boldsymbol{e_{13}}) \tag{A7}$$

Using former two Equations (A6) and (A7) along with the elemental properties of A, ($a_{i,j} \geqslant a_{k,l}$ for $i \leqslant k, j \leqslant l$), the subsequent relationships can be established:

$$a_{12}.\left(\frac{\boldsymbol{e_{22}}}{\boldsymbol{e_{12}}} - \frac{\boldsymbol{e_{23}}}{\boldsymbol{e_{13}}}\right) + \ldots + a_{1i}.\left(\frac{\boldsymbol{e_{i2}}}{\boldsymbol{e_{12}}} + \frac{\boldsymbol{e_{i3}}}{\boldsymbol{e_{13}}}\right) \ldots + a_{1p}.\left(\frac{\boldsymbol{e_{p2}}}{\boldsymbol{e_{12}}} + \frac{\boldsymbol{e_{p3}}}{\boldsymbol{e_{13}}}\right) = \gamma_3 - \gamma_2 \tag{A8}$$

$$\gamma_2 - \gamma_3 \geqslant a_{12}.\left(\frac{(-\boldsymbol{e_{12}}) + (\boldsymbol{e_{22}}) + \ldots + (\boldsymbol{e_{i2}}) \ldots + \left(\boldsymbol{e_{p2}}\right)}{(\boldsymbol{e_{12}})} + \frac{(\boldsymbol{e_{13}}) + (-\boldsymbol{e_{23}}) + \ldots + (\boldsymbol{e_{i3}}) \ldots + \left(\boldsymbol{e_{p3}}\right)}{(\boldsymbol{e_{13}})}\right) \tag{A9}$$

But, to make the relationship of $\gamma_2 \geqslant \gamma_3$ to happen in (A9), there must be two inequalities to be satisfied always (as $a_{12} > 0$):

$$\left[(-\boldsymbol{e_{12}}) + (\boldsymbol{e_{22}}) + \ldots + (\boldsymbol{e_{i2}}) \ldots + \left(\boldsymbol{e_{p2}}\right)\right] \geqslant 0 \; or \; \boldsymbol{e_{12}} \leqslant \left[(\boldsymbol{e_{22}}) + \ldots + (\boldsymbol{e_{i2}}) \ldots + \left(\boldsymbol{e_{p2}}\right)\right]$$

$$\left[(\boldsymbol{e_{13}}) + (-\boldsymbol{e_{23}}) + \ldots + (\boldsymbol{e_{i3}}) \ldots \left(\boldsymbol{e_{p3}}\right)\right] \geqslant 0 \text{ or } \boldsymbol{e_{23}} \leqslant \left[(\boldsymbol{e_{13}}) + \ldots + (\boldsymbol{e_{i3}}) \ldots + \left(\boldsymbol{e_{p3}}\right)\right] \tag{A10}$$

(A10) remains valid if the right hand side of it remains positive (as e_{12} and e_{23} are non-negative). One way of achieving it is by obtaining each term as non-negative values on the right hand side of (A10). On the contrary, the reversal of inequality signs stated in (A10) remains inconclusive. This proposition is not only true for all other eigenvectors, which possess one negative element. The first column of $[E_j]$ can also be categorized under the same set. Conversely, to prove the proposition, $\gamma_2 > \gamma_3$, to be true, the conditions shown in (A10) are sufficient. Hence, it is proved that $\boldsymbol{A}$ can have eigenvectors stated above. ■

Corresponding author
Subhadip Sarkar can be contacted at: rajsarkar77@yahoo.co.in