Benchmarking: An International Journal

Assessment of performance using MPSS based DEA

Subhadip Sarkar

Received 4 February 2014
Revised 25 September 2014
Accepted 26 September 2014

Assessment of performance using MPSS based DEA

Subhadip Sarkar
Department of Management Studies, National Institute of Technology, Durgapur, India

Abstract

Purpose – Identification of the best school among other competitors is done using a new technique called most productive scale size based data envelopment analysis (DEA). The paper aims to discuss this issue.

Design/methodology/approach – A non-central principal component analysis is used here to create a new plane according to the constant return to scale. This plane contains only ultimate performers.

Findings – The new method has a complete discord with the results of CCR DEA. However, after incorporating the ultimate performers in the original data set this difference was eliminated.

Practical implications – The proposed frontier provides a way to identify those DMUs which follow cost strategy proposed by Porter.

Originality/value – A case study of six schools is incorporated here to identify the superior school and also to visualize gaps in their performances.

Keywords Benchmarking, Data envelopment analysis, Frontier function, Principal component analysis

Paper type Research paper

1. Introduction

The journey of data envelopment analysis (DEA), as proposed by Charnes *et al.* (1978) (the CCR model), commenced from the dissertation of Rhodes when the performance of students from participating and not participating schools were compared using a non-linear model and an equivalent data-oriented, linear programming-based, non-parametric approach. A DMU is called an efficient performer if it uses fewer quantities of each input to generate the same set of outputs or produces more outputs from the same set of input resources than its rivals. Thus, it makes a place in a production possibility set. Later on, the assumption of constant return on scale (CRS), was extended by Banker *et al.* (1984). The renowned BCC model of these researchers was able to administer variable scaling techniques. As a result, weak efficient and strong efficient DMUs, most productive scale size (MPSS) and, scale efficiency became prevalent. To estimate the CRS frontier function, the regression approach was modified by Winsten (1957) by using a corrected ordinary regression technique. It enabled the detection of CRS efficient DMU instead of classifying them into below average, average and above average units (Cooper and Seiford, 2011). Later on, the DEA estimators were found statistically consistent (Banker and Maindiratta, 1992). The detailed methodology of the frontier function estimation was done by Greene (1980) on a generalized form proposed by Aigner and Chu (1968). The exploration of stochastic DEA has proven to be highly effective for adapting this approach to abrupt changes. The experiment on "Program follow Through and Non-follow Through" school sites (originally considered by Charnes *et al.*, 1978) was revisited by Land *et al.* (1993) who, instead of taking average values for inputs and outputs, suggested a deterministic equivalent of the chance constrained model by assuming normally distributed output variables which were conditional on inputs. In an efficiency evaluation of the research activities in economic departments at Danish Universities, Olesen and Petersen (1995)

Benchmarking: An International Journal
Vol. 22 No. 4, 2015
pp. 624-642

1463-5771
DOI 10.1108/BIJ-02-2014-0012

developed a chance constrained programming model while distinguishing two reasons (true inefficiency and random disturbance) to remain inefficient.

Nicole Adler and Boaz Golani (Seiford, 1989), adopted a principal component analysis (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. Kard Yen and Örkcu (2006) prepared a new data set for the application of PCA by dividing each input by each output; this approach yielded an intuitive model that is capable of producing highly correlated weighted scores with the DEA productivity indexes of the DMUs.

In this paper, a single MPSS-based CRS frontier function is constructed using PCA (unlike corrected ordinary least squares (COLS) approach where CRS function is assumed linear and all unknown coefficients are determined from regression analysis). This PCA made function contains only those DMUs which are PCA efficient (technically efficient) in case of all outputs (thus all slacks are zero). The expected level of any output is derived using this function and the error of each DMU, due to inefficiency, is found from subtracting the observed output from it. A special type of slack-based optimization model (to maximize the minimum error) is used here to prepare a frontier made from a convex combination of all predicted errors. This new frontier satisfies the MPSS condition (Ray, 2004) and plays a major role to measure the performance of a DMU. A comparison with the regular CCR model is also included here to unearth any critical issues regarding the ranking of the DMUs.

2. Definitions and theorems

2.1 DEA with CCR model

From an assumption of CRS, Charnes *et al.* (1978) found proportional changes in weighted output that derive from the alterations in weighted inputs. The algebraic models of CRS for c DMUs (each of which consumes v inputs given by the matrix $R=[R_{ij}]_{c,v}$ to generate m outputs given by a matrix $Y=[y_{ij}]_{c,m}$) are as follows.

Primal form	Dual form
Max $h_0 = \left(\sum_{j=1}^{m} q_j.y_{rj}\right)$	Min θ_0
Subjected to:	Subjected to:
$\sum_{j=1}^{v} v_j.R_{rj} = 1$	$R_{ri} \leqslant \sum_{r=1}^{c} \lambda_r.R_{ri}; \lambda_r \geqslant 0$
$\sum_{i=1}^{v} u_i.R_{ri} \geqslant \sum_{j=1}^{m} q_j.y_{rj};$	For any jth input $i=1, 2, \ldots, v$
$u_i, q_j \geqslant 0$	$\theta_0.y_{rj} \geqslant \sum_{r=1}^{c} \lambda_r.y_{rj}$
For any DMU r	for $j=1,2, \ldots, m$

2.2 Size of a DEA to remain effective

According to Dyson *et al.* (2001), to be effective, a multiple input-output (v, m) DEA model must contain be at least $2mv$ cases.

2.3 A CCR-efficient unit

A DMU is called CCR-efficient if $\boldsymbol{\theta^*} = \mathbf{1}$, and if there exists at least one optimal solution $(\boldsymbol{u^*}, \boldsymbol{q^*})$, for which $\boldsymbol{u^*} > \mathbf{0}$ and $\boldsymbol{q^*} > \mathbf{0}$, otherwise, the DMU in question is considered to be CCR-inefficient. A solution $(\boldsymbol{u^*}, \boldsymbol{q^*})$ from CCR-inefficient units $(\boldsymbol{\theta^*} < \mathbf{1})$, 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 follows:

$$E'_0 = \left\{ r : \sum\nolimits_{j=1}^{m} q_j y_{rj} = \sum\nolimits_{i=1}^{v} u_i R_{ri} \right\}$$

2.4 Production possibility set

The set of all technically feasible combinations of inputs and outputs, representing the technology of a firm. According to Cooper *et al.* (2002/2011), in case of a CCR model, any production possibility set, is defined as follows.

(A1) If an activity (R_{ri}, y_{rj}) belongs to P, then the activity (tR_{ri}, ty_{rj}) belongs to P for any positive scalar t.

(A2) For an activity (R_{ri}, y_{rj}) in P, any semi-positive activity (R_{ki}, y_{kj}) with $(R_{ki} \geqslant R_{rj})$ and $(y_{ki} \leqslant y_{ri})$ is included in P. That is, any activity with input no less than in any component and with output no greater than in any component is feasible.

A3 The non-negative combination of the DMUs in the set J as:

$$P = \left\{ \left((R_0, y_0) \middle| R_0 \leqslant \sum_{r=1}^{c} \lambda_r R_{ri};\; y_0 \leqslant \sum_{r=1}^{c} \lambda_j y_{rj};\; \lambda_r \geqslant 0; \text{for } r = 1, 2, \ldots, c \right) \right\}$$

2.5 COLS

With an initial assumption of similar structure of the production technology among the central tendency and the best practice, Winsten (1957) suggested a two steps estimation procedure to derive a production frontier which was able to lie on and above the data. In the first step ordinary least squares (OLS) is used to obtain consistent and unbiased estimates of the slope parameters and a consistent but biased estimate of the intercept parameter. In the second step the biased OLS intercept (β_0) is shifted up ("corrected") to ensure that the estimated frontier bounds the data from above. The COLS intercept is estimated consistently by:

$$\widehat{\beta}_0^* = \widehat{\beta}_0 + \max_i\{\widehat{u}_i\}$$

where the $\widehat{u}_i$ are the OLS residuals. The OLS residuals are corrected in the opposite direction, and so:

$$-\widehat{u}_i^* = u_i - \max_i\{\widehat{u}_i\}$$

The COLS residuals $\widehat{u}_i^*$ are non-negative, with at least one being zero, and can be used to provide consistent estimates of the technical efficiency of each producer by means of the expression:

$$TE_i = \operatorname{Exp}\left(-\widehat{u}_i{}^*\right)$$

2.6 PCA

PCA can be defined as the orthogonal projection of the data onto a lower dimensional linear space, known as the principal subspace, such that the variance of the projected data is maximized in the subspace. According to Rencher (2002), PCA deals with a single sample of n observation vectors $\boldsymbol{y_1}, \boldsymbol{y_2}, \ldots, \boldsymbol{y_n}$ that form an ellipsoidal swarm of points in a p-dimensional space. If the variables $y_1, y_2, \ldots, y_p$ in 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 ($\boldsymbol{y^*}$) of $\boldsymbol{y_1}, \boldsymbol{y_2}, \ldots, \boldsymbol{y_n}$. 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 $[S]_{p\times p}$ (or the correlation matrix $[R]_{p\times p}$) of the observed variables which also minimizes the mean squared distance between the data points and their projections (shown below):

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

2.7 Specific consumption matrix T and specific covariance matrix S

Under the conditions of $m = 1$ and $v < c$ in a primal-model of the DEA (CCR), there exist a positive definite covariance matrix S derived from the origin (having with a non-zero determinant) with dimensions of ($v \times v$) that can be defined as follows:

$$S_r = {T_r}^T T_r = \{s_{ij}\}_{v\times v} \text{where } s_{ij} > 0, \text{ and } T_r = \{{t_{ij}}^r\}_{c\times v} \text{where } t_{ij}^r = \frac{R_{ri}}{y_{rj}}$$
$$T_r^T = \begin{bmatrix} T_1 & T_2 & T_c \end{bmatrix}; \quad \text{where } T_i = \begin{bmatrix} t_{i1}^r & t_{i2}^r & t_{iv}^r \end{bmatrix}^T \tag{1}$$

$$s_{ij} = \begin{cases} \sum_{r=1}^{C} \left(\frac{R_{ri}}{y_{rj}}\right)^2 \ldots & i = j \\ \sum_{r=1}^{C} \left(\frac{R_{ri}}{y_{ri}}\right)\left(\frac{R_{rj}}{y_{rj}}\right) \ldots & i \neq j \end{cases} \tag{2}$$

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

2.8 A non-central PCA and its application on specific covariance matrix S_v

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:

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

The optimal solution of this problem gives rise to eigenvectors of S_v which are orthogonal to each other.

2.9 Economic interpretation of principal components of the matrix S_v

Being a square matrix of size ($v \times v$), S_v, 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 Appendix 1 and Appendix 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 eigen value) 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 reduction of "cost" (which focusses on decreasing the utilization of resources) and reduction of cost from the remaining dimensions (by manipulating proper resources).

2.10 Definition of inefficiency error

2.10.1 Inefficiency error in case of a single output. The predicted amount of any rth output from any jth DMU, can be given by the dot product of the resource vector ($\boldsymbol{R_j}$) of the same DMU and the eigenvector ($\boldsymbol{E_r}$) of the first principal component of a specific consumption matrix S_r which is derived from any rth output:

$$y_{rj}^{Pre} = \frac{1}{p_{i\min}} E_r.R_j;$$

where $p_{r\min} = \min[(E_r.T_1), (E_r.T_2), \ \ldots \ , (E_r.T_c)]$

Thus, error (Pe_{rj}) on any rth output made by any jth DMU can be determined by subtracting the observed output $\left(y_{rj}^{Obs}\right)$ from the predicted output given by y_{rj}^{Pre}:

$$Pe_{rj} = \left(y_{rj}^{Pre} - y_{rj}^{Obs}\right)$$

2.10.2 Inefficiency error in case of a multiple outputs. The joint representation of an error is derived from the linear convex combination of all errors due to individual outputs:

$$E_j = \sum_{i=1}^{m} a_i.Pe_{rj} = \sum_{i=1}^{m} a_i.\left(y_{rj}{}^{Pre} - y_{rj}{}^{Obs}\right) = Z_j{}^{Pre} - Z_j{}^{Obs}; \quad \text{where } \sum_{r=1}^{m} a_r = 1$$

Here, $Z_j^{Pre}\left\{= \sum_{r=1}^{m} a_r.\left(y_{rj}{}^{Pre}\right)\right\}$ and $Z_j^{Obs}\left\{= \sum_{r=1}^{m} a_r.\left(y_{rj}^{Obs}\right)\right\}$, are the indicators of the performance expected and actual performance from the jth DMU, respectively. The unknown value of $\boldsymbol{a_i}$ is determined by using the following LPP:

$$\text{Maximize } S;$$

$$\text{subjected to}: \sum_{i=1}^{m} a_i.Pe_{rj} \geqslant S[\mathbf{1}]^T$$

$$\sum_{r=1}^{m} a_r = 1 \text{ where}, \ [\mathbf{1}]^T = \begin{bmatrix} 1 & 1, & \ldots, & 1 \end{bmatrix}^T$$

2.11 Technical efficiency or performance index

The performance index of any jth DMU is given by the ratio of actual performance and expected performance as follows:

$$PF_j = \frac{Z_j^{Obs}}{Z_j^{pre}}$$

2.12 PCA measure of efficiency for DMUs

If $T=[t_{ij}]$, for $\{t_{ij}\ (>0)\}$, is the specific consumption matrix consisting of elements t_{ij}, which represent the specific consumption of the ith type of input (for $i=1, 2, \ldots, v$) by the rth DMU (for $r=1, 2, \ldots, c$) then the PCA measure of efficiency for any DMU r is given by [$\mathbf{min}\ (\boldsymbol{T.U})/(\boldsymbol{T_J.U})$], where U is the eigenvector that directs the major axis of the embedded PCA and $\boldsymbol{T_J}$ is the specific consumption vector of the rth DMU. This Eigenvector describes the direction of maximum variation in case of a specific consumption under a particular type of output. The magnitude of projection taken in this direction represents the keenness toward the production of the same output. A DMU is considered as keen to toward an output if the value of the projection is less.

2.13 Axiomatic definition of the MPSS frontier

(1) According to Starrett (Ray, 2004), any MPSS-based transformation function can be represented as $K(R, Y)=0$ which has a $\left(\frac{\alpha}{\beta}\right)$ ratio of 1. With an assumption of an explicit form of this function, $z=F(Y)=P(R)$ is used here instead. The differential form of this model is displayed as follows:

$$z.\frac{\partial z}{z}=\sum_{j=1}^{m}\left(\frac{\partial F}{\partial y_j}\right)(y_j).\left(\frac{\partial y_j}{y_j}\right)=\sum_{i=1}^{v}\left(\frac{\partial P}{\partial R_i}\right)R_i.\left(\frac{\partial R_i}{R_i}\right)$$

$$\text{as, } \beta=\left(\frac{\partial y_j}{y_j}\right)=\frac{\partial z}{z}; \quad \text{for } j=1,2,\ \ldots,\ m \text{ and } \alpha=\left(\frac{\partial R_i}{R_i}\right) \quad \text{for}$$

$$i=1,2,\ \ldots,v; \quad \text{and } \ \alpha=\beta; \ \text{ thus,}$$

$$z=\sum_{j=1}^{m}\left(\frac{\partial F}{\partial y_j}\right)(y_j)=\sum_{i=1}^{v}\left(\frac{\partial P}{\partial R_i}\right)R_i$$

The later relationship of $\beta=\partial z/z$ can be made if $z=F(Y)$ becomes a linear function of all individual outputs, y_j, for $j=1,2, \ldots, m$. This proposition is also valid due to the following equivalence and for a convex combination:

$$\partial z/z=\sum_{i=1}^{u}a_i\partial y_i/\sum_{i=1}^{u}a_iy_i=\partial y_i/y_i \text{ for all values of } i \text{ where } \sum_{i=1}^{u}a_i=1.$$

(2) The ultimate performer: the MPSS frontier contains those DMUs which remain PCA efficient (and thus strongly efficient) in each arena of output (efficient in all outputs).

(3) Basic elements within the set: if (R_p, Y_p) is an element in this pseudo production possibility set, then, the pairs of (R'_p, Y_p) and (R_p, Y'_p) will also be contained by the same set for the conditions of $(R'_p \geqslant R_p)$ and $(Y_p \geqslant Y'_p)$.

(4) Members on the frontier: if (R_p, Y_p) is an efficient combination according to the PCA, then, for any non-negative value t, the pair of (tR_p, tY_p) will be on the same plane.

(5) Unlike CCR model the proposed model assumes that any member in the production possibility set should abide by the following relation:

$$E_0{}'=\left\{r:\sum_{i=1}^{v}u_iR_{ri}=\sum_{j=1}^{m}q_j\,y_{rj}^{Pre}\right\}$$

where $u_i,\ q_j \geqslant 0$ for all values of i and j;

y_{rj}^{Pre} is the maximum amount of any jth output for any rth DMU using the proposed model. The set is comprised with those DMUs which are PCA efficient in each output. If such ultimate performer is not present in the data set then E_0' will become a null set. In that case, it will not contain any technically feasible combinations of inputs and outputs.

The predicted value of any output from all possible inputs is determined from a PCA-based linear function. This production function satisfies the following postulates.

(P1) $g(R)$ is monotonic in R. Since, $z_r = f(Y_{pre,r}) = AY_{pre,r} = g(R_{rj}) = BR_{rj}$, then, for $R_{1j} \geqslant R_{2j}$, and $R_{ij} \geqslant 0$, the inequality of $g(R_{1j}) \geqslant g(R_{2j})$ has to be true.

(P2) $g(R)$ is concave. Hence, if, R_1, $R_2 \in R$, and $R' = \alpha R_{r1} + (1-\alpha)R_{r2}$; such that $0 < \alpha < 1$, then $g(R') = \alpha g(R_{r1}) + (1-\alpha)g(R_{r2})$.

This property is also followed by the above proposed function (shown below):

$$g(R') = \alpha g(R_{r1}) + (1-\alpha)g(R_{r2}) = \alpha BR_{r1} + (1-\alpha)BR_{r2} = BR'$$

(P3) For each observation, (R_{rj}, Y_{rj}), $g(R_{rj}) \geqslant AY_{rj}$; for $j = 1, 2, \ldots, m$. Owing to the relationship of $Y_{Pre,j} \geqslant Y_{rj}$ the stated relationship can be proved. $g(R_j) = AY_{pre,j} \geqslant AY_{rj}$.

3. The proposed model

The whole process has been subdivided into three sections such as conversion of DEA (Section 3.1), which is essential for making a resemblance between DEA and embedded PCA, construction of CRS frontier using embedded PCA (Section 3.2) while measuring the PCA efficiencies and lastly, derivation of the MPSS-based frontier function (Section 3.3) with multiple output using the CRS frontier function.

3.1 Conversion of DEA

The converted form of the DEA can be provided with one additional constraint, as shown below.

The remodeling of the constraints of the CCR DEA:

$$T_{1j}^{r}U_1 + T_{2j}^{r}U_2 +, \cdots, + T_{vj}^{r}U_v \geqslant \frac{1}{\phi} \quad \text{for } r = 1, 2, \ldots, v \text{ and } \phi > 0$$

where $T_{ij}{}^{r} = R_{ri}/y_{rj}$ and $U_i = u_i/\phi d$

$$R_{1j}U_1 + R_{2j}U_2 +, \cdots, + R_{vj}U_v = \frac{1}{\phi d}$$

where $U_1{}^2 + U_2{}^2 +, \ldots, + U_v{}^2 = 1$

$$T_r = \left[T_{1j}^{r}, \quad T_{2j}^{r} \ldots \quad T_{vj}^{r} \right]^T \quad \text{for } r = 1, 2, \ldots, c$$

The matrix form of the above new set can be produced as follows:

$$T^T U \geqslant [1] \cdot \left(\frac{1}{\phi}\right), \quad U^T \left(TT^T\right) U = U^T (S) U \geqslant \left(\frac{1}{\phi^2}\right) [1]^T [1] = \frac{c}{\phi^2}$$

where $[1] = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}^T$ and $S = T_r T_r^T$

Therefore the objective function can be reiterated in the following form:

$$\text{minimization of } \frac{1}{d.y_j} = \phi.\left(T_{1j}{}^r U_1 + T_{2j}{}^r U_2 +, \ldots, + T_{vj}{}^r U_v\right) = \phi T_j U \qquad (4)$$

The inequality constraint, $U^T(S)U \geqslant c/\phi^2$, has an impact from the perspective of the embedded PCA, to verify that whether U remains same as the direction vector of the covariance matrix S or not. Due to the non-negativity property of the decision variables in DEA, U as a direction vector must possess entirely positive or semi-positive elements. Because of the property of a unit vector in addition to non-negativity property ($U > [0]$ and $U < [1]$), the constraint, $[1]^T U > U^T\, U = 1$ will be true:

Theorem 3.1.1. Only the major principal direction vector of the embedded PCA can have entirely non-negative elements (see the Appendix).

3.2 Construction of CRS frontier using embedded PCA

This section addresses the basic reason of a PCA efficient DMU to behave like a DEA efficient DMU (under CRS) while this is not true (the reverse is not true) for others who are inefficient. Moreover, it also clarifies why an embedded PCA is able to produce a CRS frontier:

Theorem 3.2.1. A PCA efficient DMU under any jth output is efficient under multiple output-oriented DEA.

Proof. The mathematical model of an output-oriented CCR DEA is given as follows:

maximize θ;

$$R^T\mu \leqslant R^T_{PCAj}; \text{where } R = \begin{bmatrix} R_{11} & R_{12} & R_{1m} \\ \vdots & \ddots & \vdots \\ R_{c1} & \ldots & R_{cm} \end{bmatrix}$$

$$Y^T\mu \geqslant \theta Y^T_{PCAj}; \text{where } Y = \begin{bmatrix} y_{11} & y_{12} & y_{1v} \\ \vdots & \ddots & \vdots \\ y_{c1} & \ldots & y_{cv} \end{bmatrix} \text{ and } \mu = \begin{bmatrix} \mu_1 & \mu_2\ldots & \mu_c \end{bmatrix}^T$$

If the kth member in the list of DMUs is a PCA efficient DMU under any jth output, then dividing the first set of input equations by y_{kj} the following equation set can be found:

$$St^T_j\mu \leqslant S^T_{PCAj} \text{ where } St^T_j = \frac{1}{y_{kj}} R^T \text{ and } St^T_{PCAj} = \frac{1}{y_{kj}} R^T_{PCAj}$$

Due to the properties of a first principal eigenvector E^T_j the scalar product will be same as given below:

$$E^T_j.St^T_j\mu \leqslant E^T_j.St^T_{PCAj}$$

$P^T_j\mu \leqslant P_{PCAj}$; where P^T_j is the projection in the direction of E^T_j given by $E^T_j St^T_j$.

Since, the kth DMU is PCA efficient, so, its projection in the direction of E^T_j should be minimum. This condition can only be satisfied if the vector μ possesses 1 in the kth place and contains 0s in the other places. The second set of constraints is rearranged as follows:

$$M^T\mu \geqslant \theta[\mathbf{1}]^T; \quad \text{where } M = \begin{bmatrix} \left(\frac{y_{11}}{y_{k1}}\right) & \left(\frac{y_{12}}{y_{k2}}\right)\cdots & \left(\frac{y_{1i}}{y_{ki}}\right)\cdots & \left(\frac{y_{1v}}{y_{kv}}\right) \\ \vdots & \ddots & \cdots & \vdots \\ 1 & 1 & 1\ \cdots & 1 \\ \left(\frac{y_{c1}}{y_{k1}}\right) & \cdots & \left(\frac{y_{ci}}{y_{ki}}\right)\cdots & \left(\frac{y_{cv}}{y_{kv}}\right) \end{bmatrix}$$

Putting the value of μ in the above equation the following relationship can be established:

$$[\mathbf{1}]^T \geqslant \theta[\mathbf{1}]^T$$

This expression clearly indicates the maximum value of θ as 1. ■

Theorem 3.2.2. The embedded PCA-defined CRS frontier function is given by a plane, orthogonal to the first principal eigenvector, passing through an embedded PCA efficient DMU (which is also a CCR DEA efficient DMU).

Proof. Let, the eigenvector of the major axis is $E_1 = \begin{bmatrix} e_1 & e_2 & \cdots & e_v \end{bmatrix}^T$ where $e_i > 0$, then, the plane orthogonal to it will be same as $x_1e_1 + x_2e_2,\ \ldots,\ +x_ve_v = p$. In order to find the unknown value of p it is assumed that the frontier will pass through the PCA efficient DMU. Therefore, replacing the values of x_i with the corresponding element in T_{MIN} or $X_1 = \begin{bmatrix} x_1 & x_2 & \cdots & x_v \end{bmatrix}^T = T_{\text{MIN}}$, the value of p is given as:

$$p = T_{\text{MIN}1}e_1 + T_{\text{MIN}2}e_2,\ \ldots,\ + T_{\text{MIN}v}e_v > 0; \quad \text{as } T_{\text{MIN}i} > 0 \tag{4}$$

This equation has positive intercept for each useful resource and can be reiterated in terms of any actual output level, y produced at the expense of actual inputs:

$$E_1^T.R = R_1e_1 + R_2e_2,\ \ldots,\ + R_ve_v = py = F(R) \quad \text{where } R_i = y.T_{\text{MIN}i} \text{ for } i = 1, 2,\ \ldots,\ v \tag{5}$$

According to Starrett (Ray, 2004), any MPSS, in case of a multiple input and multiple output problem, has a (α/β) ratio 1. Thus, for any transformation function $K(R,y) = 0$ or $y = P(R)$, passing through all efficient pairs (specifically through the point $((R)^*, y^*)$ where MPSS holds), the relationship shown below will be true:

$$\left(y \cdot \frac{\partial y}{y}\right)_{y=y*} = \sum_{i=1}^{n} \left(\frac{\partial P}{\partial R_i}\right)_{R*} (R_i)^*.\left(\frac{\partial R_i}{(R_i)^*}\right) \tag{6}$$

Applying the condition given by Starrett, $(\alpha/\beta) = 1$ or $\left(\partial R_i/(R_i)^*\right) = \left(\partial P/y\right)_{y=y*}$ for $i = 1, 2, \ldots$, the following equality can be shown:

$$(y)_{y=y*} = \sum_{i=1}^{n} \left(\frac{\partial P}{\partial R_i}\right)_{R*} (R_i)^* = \text{GRAD}\left((P)_{R^*}\right).(R)^* \tag{7}$$

Rearranging the equation, $E_1^T.R = py$, as $(E_1^T/p).R = y$, and making a comparison with the above equation, it can be seen that at $R = (R)^*$, $y = P(R)$ must need a gradient (which is non-negative) to be equal with $(E_1{}^T/p)$, (> 0). Banker (Ray, 2004) has earlier proved that, a MPSS can only occur at a CCR DEA efficient DMU (which is an embedded PCA efficient as well). So, the relation $R = (R)^*$ must be true here and for a function, $y = P(R)$, which satisfies $\text{GRAD}\left((P)_{R^*}\right) = (E_1^T/p)$, the condition stated by Starrett is fulfilled absolutely. Hence, it is proved that with such function which satisfies these conditions, can obey the rule of MPSS and can become a CRS frontier function. ■

Theorem 3.2.3. Any linear combination of outputs produced from embedded PCA-defined CRS frontier function, obeys Starrett defined MPSS condition.

Proof. Considering the presence of more than one output, the performance measurement of a DMU can be judged through the definition 2.10.2. If z^{Pre} comprises the collection of all outputs then it can be expressed as follows:

$$z^{Pre} = \sum_{i=1}^{m} a_i y_i{}^{Pre} = \sum_{i=1}^{m}\left(\frac{a_i}{p_{i\min}}\right)(E_i R) = \sum_{i=1}^{m}\sum_{j=1}^{v}\left(\frac{a_i}{p_{i\min}}\right)(e_{ij}R_j)$$

$$\text{or, } \left(z^{Pre}.\frac{\partial z^{Pre}}{z^{Pre}}\right)_{y=y*} = \sum_{i=1}^{m}\sum_{j=1}^{v}\left(\frac{a_i}{p_{i\min}}\right)(e_{ij}R_j)\left(\frac{\partial R_j}{R_j}\right)$$

$$\text{or, } \left(z^{Pre}\right)_{y=y*} = \sum_{i=1}^{m}\sum_{j=1}^{v}\left(\frac{a_i}{p_{i\min}}\right)(e_{ij}R_j) \quad (8)$$

$$\text{since, } \frac{\partial z^{Pre}}{z^{Pre}} = \left(\frac{\partial R_j}{R_j}\right); \quad \text{for all } j \text{ (from the condition of MPSS)}$$

$$\text{but, } \frac{\partial z^{Pre}}{z^{Pre}} = \frac{\sum_{i=1}^{m} a_i \partial y_i^{Pre}}{\sum_{i=1}^{m} a_i y_i^{Pre}} = \frac{\partial y_i^{Pre}}{y_i^{Pre}} \quad \text{for all values of } i \quad (9)$$

Thus, a convex linear combination of predicted outputs obeys the MPSS condition. ■

4. A mathematical example

To rank according to the proposed model six schools are considered in Table I. The quality of a school is judged based on the average writing score per student (O1) and science score per student (O2). Two inputs, spending per student (I1) and the financial condition of a student represented in terms of percent not from low income (I2), are also recorded here.

A school is recognized as a quality producer if it is capable of producing output scores by spending lower amount per pupil and also giving opportunities to the poorer sections. In this context, two DEA models like CCR DEA and DEA for checking super efficiency are applied here. Tables II and III contain the outputs of CCR DEA. Scores shown in Table II clearly discriminates the inefficient schools B-D from the efficient schools A, E and F.

The weight vector defined by (***u****, ***q****) for each school is displayed in Table III. The weight vector for any school is represented as *W* (name of the school, input/output).

Although, efficient schools like A, E and F have weight vectors along with few zeroes, but the reduced cost in each cases remain absolutely zeroes (which is a must be condition for becoming efficient). Table IV also explains the ranking among efficient ones. F is super-efficient followed by A and E.

The weight vector of schools, given in Table V, shows that the school A remains super-efficient due to the second input and both type of outputs. On the other hand, E and F are in the same category due to the good use of different set of inputs and outputs.

The specific consumption patterns, in Table VI, show that A assumes minimum value in input 2 under both outputs. Thus, it can be counted under the list of efficient DMUs.

Table I.
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

Table II.
CCR DEA output

Productivity	Value	Productivity	Value	Productivity	Value
Score (A)	1	Score (C)	0.9635	Score (E)	1
Score (B)	0.9096	Score (D)	0.9143	Score (F)	1

Table III.
Values of input and output weights

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	0.01555	0.03968	0	0.0000172	0.00861	0	0.00317
Reduced cost	0	0	0	0	0	0	0.21214	0
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	0.0000165	0.00825	0	0.00304	0.000017	0.00853	0	0.00314
Reduced cost	0	0	3.90828	0	0	0	4.35459	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	0.0000184	0.0092	0	0.00339	0.000212	0	0.03922	0
Reduced cost	0	0	0	0	0	0	0	0

Table IV.
DEA model for super efficiency

Productivity	Value	Productivity	Value	Productivity	Value
Score (A)	1.2348	Score (C)	0.9635	Score (E)	1.0841
Score (B)	0.9096	Score (D)	0.9143	Score (F)	1.2396

It also explains the reason that E and F achieve minimum specific consumption scores in input 1 under output 2 and in input 2 under output 1, respectively.

The covariance matrix, eigenvalues and eigenvectors, pertaining to the embedded PCA, are shown in Table VII. These eigenvectors assume largest degree of explanation (>90 percent) and reflects the usual practice of schools. First input has a higher impact than the second. Table VII is important for the derivation of the expected amount of outputs. These MPSS-based CRS frontiers, for each output, are shown below:

$$0.999949298.R_1 + 0.010069824R_2 = 185.08.y_1$$

$$0.9999R_1 + 0.01R_2 = 21.155y_2$$

Spending has higher impact on both outputs than the later one. An efficient school must produce output according to these equations. Inefficiency creeps in if any deviation exists among the observed output and the derived output. Table VIII shows the magnitude of inefficiency errors for each DMU in each output.

The important aspect of this table is that school A, which has been considered as an efficient DMU, is scoring errors on both occasions. However, E and F are able to keep their errors very close to zero and hence can be counted under the list of efficient DMUs.

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	0.0155	0.03214	0.00191	1.7164E-05	0.0086	0	0.0032
Reduced cost	6057.23	0	0	0	0	0	0.21214	0
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	1.646E-05	0.0083	0	0.003	1.702E-05	0.0085	0	0.0031
Reduced cost	0	0	3.9083	0	0	0	4.355	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	3.17E-05	0.0084	0	0.0037	0.000212	0	0.049	0
Reduced cost	0	0	6.636	0	0	8.862	0	54.56

Table V. Values of input and output weights

Schools	I1/O1	I2/O1	I1/O2	I2/O2
A	354.7222222	*2.551587302*	40.08520179	*0.288340807*
B	305.8510638	3.510638298	30.05226481	0.344947735
C	367.7891156	3.387755102	34.11041009	0.314195584
D	402.9545455	3.636363636	36.55670103	0.329896907
E	229.4117647	3.536764706	*21.15254237*	0.326101695
F	*185.0588235*	3.133333333	21.25675676	0.35990991

Table VI. Specific consumption matrix of two outputs

Contents	From output 1		From output 2	
S matrix	603,890.4534	6,081.332298	5,909.2	59.2503
	6,081.332298	65.86168651	59.25	0.6455
Eigenvalue	603,951.6945		38.22645	
Eigenvector	0.999949298	0.010069824	0.9999	0.01

Table VII. The eigenvalue and eigenvector of the covariance matrix

Table IX displays the MPSS-based optimization model for problem considered above. The output of this LPP (shown in Table X) depicts the proportions for mixing two scores. Three constraints which are considered for first three schools yield positive slack values are unable to reach up to the desired level of output.

The condition of the remaining last two schools is somewhat better in this regard. Though, the school F gets higher importance in this table and Table XI clarifies its position from the column of ranking. It ranks 2nd among others due to the ability of its students in the domain of language group. Having a positive dual price and first rank among the competitors, school E, sets a bench mark in the arena of science group. An extended output-oriented CCR model is applied here for resolving the issue of contradictions stated before. Six more schools are adopted for analysis apart from

Table VIII. Predicted output level and inefficiency error

Schools	Predicted output 1	Observed output 1	Error in output 1	Predicted output 2	Observed output 2	Error in output 2
A	48.3	25.2	23.1	422.56	223	199.6
B	46.6	28.2	18.4	407.73	287	120.7
C	58.4	29.4	29.0	511.15	317	194.16
D	57.5	26.4	31.1	502.88	291	211.89
E	33.7	27.2	6.5	295	295	0
F	25.5	25.5	0	223.10	222	1.1

Table IX. Linear model of MPSS DEA

1. Maximize S
   Subject to
2. $25.2 \times a_1 + 223 \times a_2 + S <= a_1 \times 48.2988238265909 + a_2 \times 422.562114693711$
3. $28.2 \times a_1 + 287 \times a_2 + S <= a_1 \times 46.6042432434978 + a_2 \times 407.736304094862$
4. $29.4 \times a_1 + 317 \times a_2 + S <= a_1 \times 58.4255279303645 + a_2 \times 511.159703649781$
5. $26.4 \times a_1 + 291 \times a_2 + S <= a_1 \times 57.4798480100451 + a_2 \times 502.886038920934$
6. $27.2 \times a_1 + 295 \times a_2 + S <= a_1 \times 33.718493954692 + a_2 \times 295$
7. $25.5 \times a_1 + 222 \times a_2 + S <= a_1 \times 25.5 + a_2 \times 223.097138205249$
8. $a_1 + a_2 = 1$

Table X. Output of MPSS-based DEA

Variable	*Value*	*Reduced cost*
S	0.9390801	0
a_1	0.144064	0
a_2	0.855936	0
Constr	*Slack or surplus*	*Dual price*
1	173.201	0
2	105.0549	0
3	169.4307	0
4	184.8993	0
5	0	0.144064
6	0	0.855936
7	*0*	*0.93908*

the original DMUs. Those new schools are assumed to be using the same resources like the original schools but producing outputs similar to the PCA-efficient DMU (shown in Table XII).

The equation below incorporates the LPP of the Extended CCR Model:

$$\begin{aligned} &\text{Max}\theta \\ &\text{Subjected to}: R_o \geqslant R^{Ori}.\beta_{Ori} + R^{Der}.\beta_{Der} \\ &\theta.y_o \leqslant Y^{Ori}.\beta_{Ori} + Y^{Der}.\beta_{Der}; \text{and} \qquad \beta \geqslant 0 \end{aligned}$$

The summary of these optimizations is displayed in Tables XIII and XIV.

It is quite evident from here that school E is the best DMU among the other competitors followed by F, B, etc. School A stays at last according to this model.

The ranking in this model is same as shown by the proposed model but the productivity scores offered by this model is larger than the proposed one.

In these tables only the results of original six schools are included. The productivity score of E is highest among others. Table XIV contains values of weights on each DMU during each optimization. The weight vector for original schools remains zero after each optimization and thus is not included in the above table. A very important

Schools	Error in output 1 (weight = 0.144)	Error in output 2 (weight = 0.856)	Combined error	Performance ratio	Ranking
A	23.1	199.6	174.15	0.52762	6
B	18.4	120.7	106.00	0.702023	3
C	29.0	194.16	170.38	0.617952	4
D	31.1	211.89	185.85	0.576408	5
E	6.5	0	0.9387	0.996353	1
F	0	1.1	0.9391	0.995175	2

Table XI. Combined error

Old schools	[(I1, I2), (O1, O2)]	New schools	[(I1, I2), (O1, O2)]
A	[(8,939, 64.3), (25.2, 223)]	AA	[(8,939, 64.3), (48.3, 422.56)]
B	[(8,625, 99), (28.2, 287)]	BB	[(8,625, 99), (46.6, 407.73)]
C	[(10,813, 99.6), (29.4, 317)]	CC	[(10,813, 99.6), (58.4, 511.15)]
D	[(10,638, 96), (26.4, 291)]	DD	[(10,638, 96), (57.5, 502.88)]
E	[(6,240, 96.2), (27.2, 295)]	EE	[(6,240, 96.2), (33.7, 295)]
F	[(4,719, 79.9), (25.5, 222)]	FF	[(4,719, 79.9), (25.5, 223.10)]

Table XII. Data set for new CCR model

Schools	θ	Productivity ($1/\theta$)	Rank
A	1.89	0.5291	6
B	1.42	0.70423	3
C	1.61	0.62112	4
D	1.73	0.57803	5
E	1	1	1
F	1.00027	0.99973	2

Table XIII. Ranking of schools using extended CCR model

	School A		*School E*		*School F*	
Weights	Value	Reduced cost	Value	Reduced cost	Value	Reduced cost
X1	1	0	0.54031	0	0.958987	0
X2	0	3.35E-05	0	2.6E-05	0	2.35E-05
X3	0	4.05E-05	0	3.15E-05	0	2.85E-05
X4	0	2.34E-05	0	1.81E-05	0	1.64E-05
X5	0	1.33E-05	0	1.03E-05	0	9.36E-06
X6	0	0	0.804231	0	0.474808	0
Constrs.	Slack or surplus	Dual price	Slack or surplus	Dual price	Slack or surplus	Dual price
1	0	0.000212	0	0.000165	0	0.000149
2	0	2.51E-06	0	1.95E-06	0	1.77E-06
3	0.548825	0	6.541476	0	11.0195	0
4	0	−0.00448	0	−0.00348	0	−0.00315
	School D		*School E*		*School F*	
Weights	Value	Reduced cost	Value	Reduced cost	Value	Reduced cost
X1	1.7281	0	0.1086	0	0	0.000653
X2	0.9663	0	0	2.53E-05	0	0.000762
X3	0	2.56E-05	0	3.06E-05	0	0.0018
X4	0	3.1E-05	0	1.77E-05	0.4436	0
X5	0	1.79E-05	0	1.01E-05	0	0.001104
X6	0	1.02E-05	1.1166	0	0	0.000271
Constrs.	Slack or surplus	Dual price	Slack or surplus	Dual price	Slack or surplus	Dual price
1	0	0.000162	0	0.000165	0	0.000212
2	0	1.93E-06	0	1.95E-06	37.315	0
3	11.858	0	6.5415	0	0	−0.03922
4	0	−0.00344	0	−0.00348	1.0166	0

Table XIV. Summary of extended CCR DEA

observation can be made from this table. Although, new schools were brought in from the original schools, but, apart from school A, no other original schools are containing their respective new schools under their production possibility set (like school B does not have new school BB in the production possibility set).

5. Conclusion

To address the question "why this new DEA model" it is referred here that the traditional CCR DEA model is very useful to define a kinked frontier which discriminates the efficient DMUs from the rests. However, it hardly mentions anything about the internal dimensions of resource utilization. The cost (collective consumption) frontier is derived from the first principal component of a non-central covariance matrix. The corresponding eigenvector shows positive direction of the joint variation of all resources. Any DMU which has lowest intercept will be counted as economic user of resources. As a result it remains close to the cost frontier to become cost efficient.

The proposed method offers ranking to the DMU based on their performance index. Although, it does not have any resemblance with the ranking found in case of CCR DEA or from their super-efficiencies, but, it clearly supports the claim of these two models that the schools E and F are very close to be referred as efficient performers. However, deficits are found high in case of schools like B-E in both outputs. Apart from this, the magnitudes of these errors are much less than whatever is seen in Table XI.

As per the proposed model, school A, which is among the super-efficient group, has never been put under the list of efficient performers as the errors due to inefficiency in each output are close to inefficient schools like C and D. This contradiction is settled by using the new CCR DEA model which provides the same type of ranking as Table XI. The productivity scores, however, are not identical with the former one. The reason of this difference can be realized by the fact that the proposed model is based on a pessimistic view which locates the MPSS frontier through points where the model maximizes the minimum error. Thus, the performance measured from this plane will always be less than whatever is found in case of new CCR model. It is, thus, capable of identifying a superior point which is situated closer to the point generated from the proposed method to generate better productivity. The specialty of this work is that it can measure inefficiency related errors and also can be extended to the stochastic analysis of errors.

References

Aigner, D.J. and Chu, S. (1968), "On estimating the industry production function", *American Economic Review*, Vol. 58 No. 4, pp. 826-239.

Banker, R. and Maindiratta, A. (1992), "Maximum likelihood estimation of monotone and concave production frontiers", *Journal of Productivity Analysis*, Vol. 3 No. 4, pp. 401-415.

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, M.L. and Zhu, J. (2011), *Hand Book on Data Envelopment Analysis, International Series in Operation Research & Management Science*, NY, Vol. 168, Springer.

Cooper, W.W., Seiford, M.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.

Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V. and Shale, E.A. (2001), "Pitfalls and protocols in DEA", *European Journal of Operational Research*, Vol. 132 No. 2, pp. 245-259.

Greene, W.H. (1980), "Maximum likelihood estimation of econometric frontier functions", *Journal of Econometrics*, Vol. 13 No. 1, pp. 27-56.

Kard Yen, F. and Örkcu, H.H. (2006), "The comparison of principal component analysis and data envelopment analysis in ranking of decision making units", *G.U. Journal of Science*, Vol. 19 No. 2, pp. 127-133.

Land, C.K., Lovel, C.A.K. and Thore, S. (1993), "Chance constrained data envelopment analysis", *Managerial and Decision Economics*, Vol. 14 No. 6, pp. 541-554.

Olesen, O.B. and Petersen, N.C. (1995), "Chance constrained efficiency evaluation", *Management Science*, Vol. 41 No. 3, pp. 442-457.

Ray, S.C. (2004), *Data Envelopment Analysis Theory & Techniques for Economics & Operation Research*, Cambridge University 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, Department of Industrial Engineering and Operations Research, University of Amherst, Massachussets.

Winsten, C.B. (1957), "Discussion on Mr. Farrell's paper", *Journal of the Royal Statistical Society Series a-Statistics in Society*, Vol. 120 No. 3, pp. 282-284.

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:

$$|A - \gamma.I| = \begin{bmatrix} a_{11} - \gamma & a_{12} & \ldots a_{1n} \\ a_{12} & a_{22} - \gamma & \ldots a_{2n} \\ a_{1n} & a_{2n} & a_{nn} - \gamma \end{bmatrix} = 0; \quad \text{(A1.1)}$$

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

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

This can also be expressed as follows:

$$\gamma V_1 = \gamma \begin{bmatrix} x_1 \\ X_1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{1p} \\ a_{p1} & A_1 \end{bmatrix} \begin{bmatrix} x_1 \\ X_1 \end{bmatrix} \quad \text{(A1.2)}$$

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 $(\gamma > \max(a_{11}, a_{22}, \ldots, a_{nn}))$ 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:

$$\gamma V_2 = \gamma \begin{bmatrix} -x_2 \\ X_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{1p} \\ a_{p1} & A_1 \end{bmatrix} \begin{bmatrix} -x_2 \\ X_2 \end{bmatrix} \quad \text{(A1.3)}$$

$$\begin{bmatrix} x_1 & {X_1}^T \end{bmatrix} \begin{bmatrix} -x_2 \\ X_2 \end{bmatrix} = 0 \quad \text{(A1.4)}$$

However, this will violate the condition $(\gamma > \max(a_{11}, a_{22}, \ldots, a_{nn}))$. 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:

$${X_1}^T(\gamma I - A_1).X_1 = \frac{{X_1}^T(a_{p1}a_{1p}).X_1}{(\gamma - a_{11})} x_1 \quad \text{(A1.5)}$$

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, will maintain a special structure

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

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

Any ith column of the matrix, $[Ej]_{p\times p}$, 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, γi. Using these concepts the following three equations can be derived:

$$a_{11}.(-\boldsymbol{e}_{12}) + a_{12}.(\boldsymbol{e}_{22}) +, \ldots, + a_{1i}.(\boldsymbol{e}_{\boldsymbol{i}2}), \ldots, + a_{1p}.(\boldsymbol{e}_{\boldsymbol{p}\mathbf{2}}) = \gamma_2.(-\boldsymbol{e}_{12}) \quad \text{(A2.1)}$$

$$a_{11}.(\boldsymbol{e}_{13}) + a_{12}.(-\boldsymbol{e}_{23}) +, \ldots, + a_{1i}.(\boldsymbol{e}_{\boldsymbol{i}3}), \ldots, + a_{1p}.(\boldsymbol{e}_{\boldsymbol{p}3}) = \gamma_3.(\boldsymbol{e}_{13}) \quad \text{(A2.2)}$$

Using former two Equations (A2.1) and (A2.2) 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:

$$a_{12}.\left(\frac{\boldsymbol{e}_{22}}{\boldsymbol{e}_{12}} - \frac{\boldsymbol{e}_{23}}{\boldsymbol{e}_{13}}\right) +, \ldots, + a_{1i}.\left(\frac{\boldsymbol{e}_{\boldsymbol{i}2}}{\boldsymbol{e}_{12}} + \frac{\boldsymbol{e}_{\boldsymbol{i}3}}{\boldsymbol{e}_{13}}\right), \ldots, + a_{1p}.\left(\frac{\boldsymbol{e}_{\boldsymbol{p}2}}{\boldsymbol{e}_{12}} + \frac{\boldsymbol{e}_{\boldsymbol{p}3}}{\boldsymbol{e}_{13}}\right) = \gamma_3 - \gamma_2 \quad \text{(A2.3)}$$

$$\gamma_2 - \gamma_3 \geqslant a_{12}.\left(\frac{(-\boldsymbol{e}_{12}) + (\boldsymbol{e}_{22}) +, \ldots, + (\boldsymbol{e}_{\boldsymbol{i}2}), \ldots, + (\boldsymbol{e}_{\boldsymbol{p}\mathbf{2}})}{(\boldsymbol{e}_{12})} + \frac{(\boldsymbol{e}_{13}) + (-\boldsymbol{e}_{23}) +, \ldots, + (\boldsymbol{e}_{\boldsymbol{i}3}), \ldots, + (\boldsymbol{e}_{\boldsymbol{p}3})}{(\boldsymbol{e}_{13})}\right) \quad \text{(A2.4)}$$

But, to make the relationship of $\gamma_2 \geqslant \gamma_3$ to happen in A2.4, there must be two inequalities to be satisfied always (as $a_{12} > 0$):

$$[(-\boldsymbol{e}_{12}) + (\boldsymbol{e}_{22}) +, \ldots, + (\boldsymbol{e}_{\boldsymbol{i}2}), \ldots, + (\boldsymbol{e}_{\boldsymbol{p}2})] \geqslant 0 \text{ or } \boldsymbol{e}_{12} \leqslant [(\boldsymbol{e}_{22}) +, \ldots, + (\boldsymbol{e}_{\boldsymbol{i}2}), \ldots, + (\boldsymbol{e}_{\boldsymbol{p}2})]$$

$$[(e_{13})+(-e_{23})+, \ldots, +(e_{i3}), \ldots, +(e_{p3})] \geqslant 0 \text{ or } e_{23} \leqslant [(e_{13})+, \ldots, +(e_{i3}), \ldots, +(e_{p3})] \tag{A2.5}$$

On the contrary, the reversal of inequality signs stated in A2.5 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 A2.5 are sufficient. ■

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