Applied Mathematical Sciences, Vol. 6, 2012, no. 72, 3579 - 3590

Procrustes Analysis and the Goodness-of-fit of Biplots: Some Thoughts and Findings

Siswadi, Toni Bakhtiar and Ririh Maharsi

Department of Mathematics, Bogor Agricultural University (IPB)
Jl. Meranti, Kampus IPB Darmaga, Bogor 16680, Indonesia
siswadi@inrr.org, tbakhtiar@ipb.ac.id, ririh12@gmail.com

Abstract

In this paper we provide some thoughts and findings on the application of Procrustes analysis in assessing the goodness-of-fit of biplots. Our first interest deals with the choice of object and variable matrices employed in biplots analysis, where we offer more general matrices in term of flexibility in selecting distance function. Next we pay our concern to the interconnection between Gabriel and Procrustes based goodness-of-fits in logratio biplots which widely used in compositional data analysis. Our last finding lies on the very fundamental matter of Procrustes analysis practice. It is well-known that in matching of two configuration matrices with different number of columns, common procedure then places a number of columns of zeros in the last such that the dimension of matrices are the same. We supply a justification that such columns of zeros can be placed not only in the last, but also anywhere.

Mathematics Subject Classification: 03F03, 15A04, 62A01

Keywords: Procrustes analysis, logratio biplots, goodness-of-fit

1 Biplots and Procrustes Analyses

The biplot is a low-dimensional display, usually only two dimensions, of a large data matrix X depicting a point for each of the n observation vectors (rows of X) along with a point for each of the p variables (columns of X). As the generalization of the well-known scatterplots, biplots were introduced by Gabriel [4] and have been discussed at length by [7].

Let X is a matrix of rank $r \leq \min\{n, p\}$ and its singular value decomposition (SVD) is provided by

$$X = ULW', \tag{1}$$

where U and W are $n \times r$ and $p \times r$ orthogonal matrices, and L is an $r \times r$ diagonal matrix whose diagonal elements are the singular values of X. To be more precise, we may write

$$U = \begin{pmatrix} u_1 & u_2 & \dots & u_r \end{pmatrix}, \quad (2)$$

$$L = \operatorname{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_r}), \quad (3)$$

$$W = \begin{pmatrix} w_1 & w_2 & \dots & w_r \end{pmatrix}, \quad (4)$$

where $w_i\ (i = 1, \dots, r)$ are eigenvectors correspond to positive eigenvalues of $X'X$ and $u_i\ (i = 1, \dots, r)$ are those of XX'. Clearly, U and W are constructed by orthonormal columns and we may define either $u_i := \frac{Xw_i}{\sqrt{\lambda_i}}$ or $w_i := \frac{X'u_i}{\sqrt{\lambda_i}}$. It is known that $X'X$ and XX' posses the same set of positive eigenvalues. Eigenvalues of $X'X$ can be ordered as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_p = 0$ with the corresponding eigenvectors $w_i\ (i = 1, \dots, r, r+1, \dots, p)$ and those of XX' are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$ with the corresponding eigenvectors $u_i\ (i = 1, \dots, r, r+1, \dots, n)$. If zero eigenvalues are included along with their corresponding eigenvectors, then we have a complete SVD. Anyhow, the SVD of X (1) can then be written as

$$X = \sum_{i=1}^{r} \sqrt{\lambda_i} u_i w_i'. \quad (5)$$

If X is approximated by using the first s singular values, with $s \leq r$, then $X \approx AB'$, where $A := U_s L_s^{\alpha}$, $B := W_s L_s^{1-\alpha}$, and α is a scalar in $[0, 1]$. Matrices X, XX', and $X'X$ are respectively known as data, object, and variable matrices [5]; while AB', AA', and BB' are their respective approximations.

According to Gabriel [5], the biplot measure of goodness of proportional fit between X and its approximation Y, notated by $\mathrm{GF}_G(X, Y)$, is defined as

$$\mathrm{GF}_G(X, Y) := 1 - \frac{\min_c \|X - cY\|^2}{\|X\|^2} = \frac{\mathrm{tr}^2(X'Y)}{\mathrm{tr}(X'X)\,\mathrm{tr}(Y'Y)}, \quad (6)$$

where c is a scaling factor which should be optimally picked. The goodness-of-fit coefficients related to data, object, and variable matrices are then given by $\mathrm{GF}_G(X, AB')$, $\mathrm{GF}_G(XX', AA')$, and $\mathrm{GF}_G(X'X, BB')$, respectively following last expression of (6).

Procrustes analysis was developed for comparison an observed data matrix to a target matrix, where both matrices have same number of samples (as rows) and may different number of variables (as columns). For gaining maximal matching, Procrustes analysis employs data and configuration adjustments in order to eliminate possible incommensurability of variables within the individual data sets and size differences between data sets [6]. Basically, translation, rotation, and dilation are the kinds of adjustments that may be deemed desirable before embarking on the actual Procrustes matching [3]. It is already proved in [2] that a series of adjustments in standard command, i.e., translation-rotation-dilation, does provide the smallest matching distance. To measure the difference between two $n \times p$ configurations $X = (x_{ij})$ and $Y = (y_{ij})$, Procrustes analysis exploits the sum of the squared distances E between the points in the Y space and the corresponding points in the X space, known as Procrustes distance, given by

$$E(X,Y) = \sum_{i=1}^{n} \sum_{j=1}^{p} (x_{ij} - y_{ij})^2 = \mathrm{tr}(X-Y)'(X-Y). \tag{7}$$

Subject to an optimal translation-rotation-dilation adjustment, the lowest possible Procrustes distance E^* is then provided by

$$E^*(X,Y) = \mathrm{tr}(X_{\mathrm{T}} X_{\mathrm{T}}') - \frac{\mathrm{tr}^2(X_{\mathrm{T}} Q' Y_{\mathrm{T}}')}{\mathrm{tr}(Y_{\mathrm{T}} Y_{\mathrm{T}}')}, \tag{8}$$

where X_{T} and Y_{T}, respectively, are translated configurations, i.e., $X_{\mathrm{T}} = X - \frac{1}{n} 1_n 1_n' X$ and $Y_{\mathrm{T}} = Y - \frac{1}{n} 1_n 1_n' Y$, and Q is an orthogonal matrix obtained from a certain SVD, i.e., $Q = VU'$ with $U \Sigma V'$ is the complete SVD of $X_{\mathrm{T}} Y_{\mathrm{T}}'$. Readers interested on how optimal transformation is performed, are suggested to refer [2].

Our work in [10] provides a direct derivation of the goodness-of-fit of biplots via Procrustes analysis and, in particular, suggests a new expression for goodness-of-fit measure of variable matrix to strengthen the accuracy of the current form in [5]. We propose a goodness-of-fit measure based on Procrustes analysis GF_P as follows

$$\mathrm{GF}_P(X,Y) := 1 - \frac{E^*(X,Y)}{\mathrm{tr}(XX')}. \tag{9}$$

2 Object and Variable Matrices

As stated in the previous section, it is implied the following approximation for object and variable matrices:

$$XX' \approx AA' = U_s L_s^{2\alpha} U_s' = \sum_{i=1}^{s} \lambda_i^{\alpha} u_i u_i', \tag{10}$$

$$X'X \approx BB' = W_s L_s^{2(1-\alpha)} W_s' = \sum_{i=1}^{s} \lambda_i^{1-\alpha} w_i w_i'. \tag{11}$$

In our belief, approximations (10) and (11) are not compatible since approximation matrices in the right-hand side depend heavily on α, while approximated object and variable matrices on the left-hand side do not. In fact, parameter α plays important role as it determines the distance function used. To ensure a fair association, note that (1) can also be written as

$$X = UL^{\alpha}L^{1-\alpha}W',$$

from which by defining $G := UL^{\alpha}$ and $H := WL^{1-\alpha}$ we then have

$$X = GH', \tag{12}$$

and furthermore,

$$GG' = UL^{2\alpha}U' = \sum_{i=1}^{r} \lambda_i^{\alpha} u_i u_i', \tag{13}$$

$$HH' = WL^{2(1-\alpha)}W' = \sum_{i=1}^{r} \lambda_i^{1-\alpha} w_i w_i'. \tag{14}$$

For $\alpha = 0$, which refers to *covariance biplots* [1], it is immediate that $G = U$ and $H = WL$, and thus $HH' = WL^2W' = X'X$, showing that variable matrix $X'X$ introduced by [5] can be recovered by HH'. To show more implication we write $X = (x_1\ x_2\ \dots\ x_n)'$, $G = (g_1\ g_2\ \dots\ g_n)'$, and $H = (h_1\ h_2\ \dots\ h_p)'$, where vectors x_i and g_i denote the i-th objects in the matrices, and h_i denotes the i-th column of X which corresponds to the i-th variable. From (12) we have $y_{ij} = g_i'h_j$; by which g_i represents the i-th object of X and h_j corresponds to the j-th variable. Next, by recalling covariance matrix $S = \frac{1}{n-1}X'X$, the squared Mahalanobis distance d_M^2 between x_i and x_j is given by

$$\begin{aligned}
d_M^2(x_i, x_j) &= (x_i - x_j)'S^{-1}(x_i - x_j) \\
&= (g_i - g_j)'H'S^{-1}H(g_i - g_j) \\
&= (g_i - g_j)'LW'(n-1)(X'X)^{-1}WL(g_i - g_j) \\
&= (n-1)(g_i - g_j)'LW'(WL^{-2}W')WL(g_i - g_j) \\
&= (n-1)(g_i - g_j)'(g_i - g_j) \\
&= (n-1)d_E^2(g_i, g_j),
\end{aligned}$$

where d_E^2 denotes the squared Euclidean distance. We just show that the Mahalanobis distance between x_i and x_j is proportional to the Euclidean distance between g_i and g_j.

Likewise, for $\alpha = 1$, which refers to *form biplots* [1], we have $G = UL$ and $H = W$, accordingly $GG' = UL^2U' = XX'$, showing that object matrix XX' introduced by [5] can be claimed by GG'. Further,

$$\begin{aligned}
d_E^2(x_i, x_j) &= (x_i - x_j)'(x_i - x_j) \\
&= (g_i - g_j)'H'H(g_i - g_j) \\
&= (g_i - g_j)'W'W(g_i - g_j) \\
&= d_E^2(g_i, g_j),
\end{aligned}$$

which shows that Euclidean distance between x_i and x_j is exactly identical with that between g_i and g_j.

These facts, however, suggest us to devote a more general conduct in selecting object and variable matrices. By proposing GG' as object matrix and HH' as variable matrix we offer a wider perspective in utilizing distance functions, not only limited to Euclidean and Mahalanobis distance functions, but also those in between by selecting $\alpha \in (0, 1)$. In addition, the fairness and compatibility of approximations $GG' \approx AA'$ and $HH' \approx BB'$ are now preserved, since all matrices depend directly on α.

3 Goodness-of-fit of Logratio Biplots

In this section we discuss the interconnection between goodness-of-fit of biplots proposed by Gabriel [5] and one derived by using Procrustes analysis. While in [10] we dealt with ordinary biplots by considering XX' and $X'X$ as object and variable matrices, we here proceed toward logratio biplots along with GG' and HH' as the matrices.

3.1 Logratio Biplots for Compositional Data

Biplots analysis commonly deals with interval scale data, where it is meant that when we compare two values, we look at their (interval) differences. On the other hand, many variables are measured on ratio scales where we would express the comparison as a multiplicative, or percentage, difference [7]. Compositional data which consists of row vectors whose elements are positive and have row-sum of 1 is a kind of ratio scale data. To tackle such data, ordinary biplots should be adapted by logarithmic transformation and double centering leads to the so-called logratio biplots.

Suppose $Z = (z_{ij})$ is an $n \times p$ compositional data matrix where $z_{ij} \geq 0$ and $\sum_{j=1}^{p} z_{ij} = 1$ for $i = 1, 2, \ldots, n$. Logarithmic transformation provides $Y = (y_{ij})$ where $y_{ij} := \log z_{ij}$. If Z contains essential and no-data zeros, then they can be handled by variables aggregation and very small number replacement such as $0.001 \leq z_{ij} \leq 0.01$ [9]. Subsequently, double centering, i.e., row-and-column mean centering, procedure gives

$$X := Y - \tfrac{1}{p} Y 1_p 1_p' - \tfrac{1}{n} 1_n 1_n' Y + \tfrac{1}{np} 1_n 1_n' Y 1_p 1_p'.$$

We can then easily verify that the following properties hold

$$1_n' X = 0_p', \quad 1_p' X' = 0_n', \tag{15}$$

suggesting that row-sum and column-sum of the double centering matrix X are all zero.

3.2 Goodness-of-fit of Data Matrix

Logratio biplots requires that the data matrix X is initially row-and-column mean centered. It means that translation process is by default already applied, as we can immediately verified. We denote by C_X and $C_{AB'}$ the centroids of data matrix X and its approximation AB', respectively. We have by (15),

$$\begin{aligned} C_X &= \tfrac{1}{n}1_n'X = 0_p', \\ C_{AB'} &= \tfrac{1}{n}1_n'AB' = \tfrac{1}{n}1_n'\sum_{i=1}^{s}\sqrt{\lambda_i}u_iw_i' = \tfrac{1}{n}1_n'X\sum_{i=1}^{s}w_iw_i' = 0_p', \end{aligned}$$

which show that the centroids of X and AB' are coincident at the origin.

Optimal rotation on AB' can be carried-out by multiplying AB' by an orthogonal matrix Q, where $Q = MK'$ with $K\Theta M'$ is the complete SVD of $X'AB'$. Since

$$X'AB' = \left(\sum_{i=1}^{r}\sqrt{\lambda_i}u_iw_i'\right)'\sum_{i=1}^{s}\sqrt{\lambda_i}u_iw_i' = \sum_{i=1}^{s}\lambda_iw_iw_i' = \bar{W}\Theta\bar{W}',$$

where $\bar{W} = \begin{pmatrix} w_1 & \dots & w_s & w_{s+1} & \dots & w_p \end{pmatrix}$ and

$$\Theta := \begin{pmatrix} \operatorname{diag}(\lambda_1, \dots, \lambda_s) & 0_{s\times(p-s)} \\ 0_{(p-s)\times s} & 0_{(p-s)\times(p-s)} \end{pmatrix},$$

the above decomposition suggests that $K = M = \bar{W}$ and thus an identity rotation matrix $Q = MK' = \bar{W}\bar{W}' = I$ is obtained, indicates that rotation was intrinsically employed.

Dilation adjustment is implemented by choosing an optimal scaling factor $c^* = \frac{\operatorname{tr}(X'AB')}{\operatorname{tr}(BA'AB')}$ such that

$$E^*(X, AB') = E(X, c^*AB') = \operatorname{tr}(X'X) - \frac{\operatorname{tr}^2(X'AB')}{\operatorname{tr}(BA'AB')}.$$

Substituting the above optimal Procrustes distance into (9) we then obtain

$$GF_P(X, AB') = \frac{\operatorname{tr}^2(X'AB')}{\operatorname{tr}(X'X)\operatorname{tr}(BA'AB')} = \frac{\sum_{i=1}^{s}\lambda_i}{\sum_{i=1}^{r}\lambda_i}, \qquad (16)$$

which establishes $GF_P(X, AB') = GF_G(X, AB')$. Note that the goodness-of-fit of data matrix is independent of α.

3.3 Goodness-of-fit of Object Matrix

Now we inspect the goodness-of-fit of object matrix GG' and its approximation AA'. Using (10), (13), and (15), we found the centroids of both matrices are

$$\begin{aligned} C_{GG'} &= \tfrac{1}{n}1_n'GG' = \tfrac{1}{n}1_n'\sum_{i=1}^{r}\lambda_i^{\alpha}u_iu_i' = \tfrac{1}{n}1_n'X\sum_{i=1}^{r}\lambda_i^{\alpha-1}w_iw_i'X' = 0_p', \\ C_{AA'} &= \tfrac{1}{n}1_n'AA' = \tfrac{1}{n}1_n'\sum_{i=1}^{s}\lambda_i^{\alpha}u_iu_i' = \tfrac{1}{n}1_n'X\sum_{i=1}^{s}\lambda_i^{\alpha-1}w_iw_i'X' = 0_p'. \end{aligned}$$

Since the centroids are coincident at the origin, then an adjustment by translation is not required. Suppose $P\Lambda R'$ is the complete SVD of $GG'AA'$, then

$$GG'AA' = \sum_{i=1}^{s}\lambda_i^{2\alpha}u_iu_i' := \bar{U}\Lambda\bar{U}',$$

where $\bar{U} = \begin{pmatrix} u_1 & \dots & u_s & u_{s+1} & \dots & u_n \end{pmatrix}$ and

$$\Lambda := \begin{pmatrix} \text{diag}(\lambda_1^{2\alpha},\dots,\lambda_s^{2\alpha}) & 0_{s\times(n-s)} \\ 0_{(n-s)\times s} & 0_{(n-s)\times(n-s)} \end{pmatrix}.$$

As a consequence, $Q := RP' = \bar{U}\bar{U}' = I$ evokes that rotation process is also unnecessary. The best Procrustes distance can then solely be obtained by dilation to give

$$E^*(GG',AA') = E(GG',c^*AA') = \text{tr}(GG'GG') - \frac{\text{tr}^2(GG'AA')}{\text{tr}(AA'AA')}, \tag{17}$$

where an optimal scaling factor $c^* = \frac{\text{tr}(GG'AA')}{\text{tr}(AA'AA')}$ has been applied. Accordingly, by (9), the goodness-of-fit for object matrix is stated as follows

$$\text{GF}_P(GG',AA') = \frac{\text{tr}^2(GG'AA')}{\text{tr}(GG'GG')\,\text{tr}(AA'AA')} = \frac{\sum_{i=1}^{s}\lambda_i^{2\alpha}}{\sum_{i=1}^{r}\lambda_i^{2\alpha}}. \tag{18}$$

Obviously, we have $\text{GF}_P(GG',AA') = \text{GF}_G(GG',AA')$.

3.4 Goodness-of-fit of Variable Matrix

Variable matrix of a data set is represented by HH' and its approximation by BB'. Both matrices are coincident at origin since their centroids are given by

$$\begin{aligned} C_{HH'} &= \tfrac{1}{n}1_p'HH' = \tfrac{1}{n}1_p'\sum_{i=1}^{r}\lambda_i^{1-\alpha}w_iw_i' = \tfrac{1}{n}1_p'X'\sum_{i=1}^{r}\lambda_i^{-\alpha}u_iu_i'X = 0_n', \\ C_{BB'} &= \tfrac{1}{n}1_p'BB' = \tfrac{1}{n}1_p'\sum_{i=1}^{s}\lambda_i^{1-\alpha}w_iw_i' = \tfrac{1}{n}1_p'X'\sum_{i=1}^{s}\lambda_i^{-\alpha}u_iu_i'X = 0_n'. \end{aligned}$$

Translation is then unnecessary. Next suppose the CFSVD of $HH'BB'$ is provided by $M\Omega N'$. Then we have

$$HH'BB' = \sum_{i=1}^{s} \lambda_i^{2(1-\alpha)} w_i w_i' := \bar{W}\Omega\bar{W}',$$

where

$$\Omega := \begin{pmatrix} \operatorname{diag}(\lambda_1^{2(1-\alpha)}, \dots, \lambda_s^{2(1-\alpha)}) & 0_{s\times(n-s)} \\ 0_{(n-s)\times s} & 0_{(n-s)\times(n-s)} \end{pmatrix}.$$

It suggests that an adjustment by rotation is also unnecessary since the optimal rotation matrix is an identity matrix, i.e., $Q := NM' = \bar{W}\bar{W}' = I$. These facts, however, reveal a central difference on the choice of matrix. When Gabriel [5] suggests $(X'X, BB')$, matrix $X'X$ is, in general, not column centered. Thus, translation and rotation should precede the calculation of goodness-of-fit to gain maximal agreement between them [10]. This is not the case of (HH', BB'). Consequently, under the later choice, goodness-of-fit by Gabriel (6) and that by Procrustes (9) are consistent. For variable matrix, it is given by

$$\mathrm{GF}_P(HH', BB') = \frac{\mathrm{tr}^2(HH'BB')}{\mathrm{tr}(HH'HH')\,\mathrm{tr}(BB'BB')} = \frac{\sum_{i=1}^{s} \lambda_i^{2(1-\alpha)}}{\sum_{i=1}^{r} \lambda_i^{2(1-\alpha)}}. \qquad (19)$$

One interesting fact can be drawn in the case of $\alpha = 0.5$, where we have

$$\mathrm{GF}(X, AB') = \mathrm{GF}(GG', AA') = \mathrm{GF}(HH', BB') = \frac{\sum_{i=1}^{s} \lambda_i}{\sum_{i=1}^{r} \lambda_i},$$

pointing-out that all goodness-of-fit curves intersect in one point at $\alpha = 0.5$. This will be verified in the following illustrative example.

3.5 Illustrative Example

To illustrate our result we consider a compositional data set of *jamu* (medicinal herbs) for menstruation control, taken from [11]. Ten samples of different brands of *jamu* are analyzed to identify their composition. Samples are then characterized according to five indications: (i) uterus stimulant (US), (ii) estrogenic (ES), (iii) both estrogenic and uterus stimulant (EU), (iv) abortive or teratogenic (AT), and (v) others, such as haemostatic, taste and aroma (OT). The data are presented in Table 1, from which we have a 10×5 data matrix of rank 4. The goodness-of-fit coefficients for data, object, and variable matrices are calculated according to (16), (18) and (19) with $s = 2$. The results with respect to α are then depicted in Figure 1. It is shown that the goodness-of-fit

Table 1: Compositional data of *jamu* (in percent).

Brand	US	ES	EU	AT	OT
1	30.00	15.00	20.00	10.00	25.00
2	18.00	16.00	10.00	30.00	26.00
3	0.50	14.92	24.88	24.88	34.82
4	0.50	39.60	34.65	24.75	0.50
5	20.00	10.00	45.00	10.00	15.00
6	4.99	0.10	29.97	44.96	19.98
7	32.00	18.00	10.00	24.00	16.00
8	40.00	16.00	8.00	12.00	24.00
9	39.80	9.95	29.85	0.50	19.90
10	59.70	0.50	9.95	19.90	9.95

of data matrix is constant at 80.24 percent. That of object matrix is increasing with respect to α, ranging from 50 percent when $\alpha = 0$ to 94.36 percent at $\alpha = 1$, while that of variable matrix shows an opposite trend. All curves pass across one point at $\alpha = 0.5$ as indicated.

4 Placing Columns of Zeros

In this last section we provide a minor note on Procrustes analysis practice. Suppose X_0 is a configuration of n points in a q dimensional Euclidean space. Configuration X_0 needs to be optimally matched to another configuration of n points in a p dimensional Euclidean space with coordinate matrix Y. It is assumed that the r-th point in the first configuration is in a one-to-one correspondence with the r-th point in the second configuration. If $p > q$ then $k := p - q$ columns of zeros are placed at the end of matrix X_0 so that both configurations are placed in the same space. We here supply a justification that such k columns of zeros can be placed not only in the last, but also anywhere within the configuration matrix.

Let initially we put only one column of zeros into j-th column of X_0. Indeed, adjustment by translation will preserve this zero column. Interchanging columns j and p will move the column of zeros into the last column. This procedure is known as type 1 column elementary operation, where interchanging

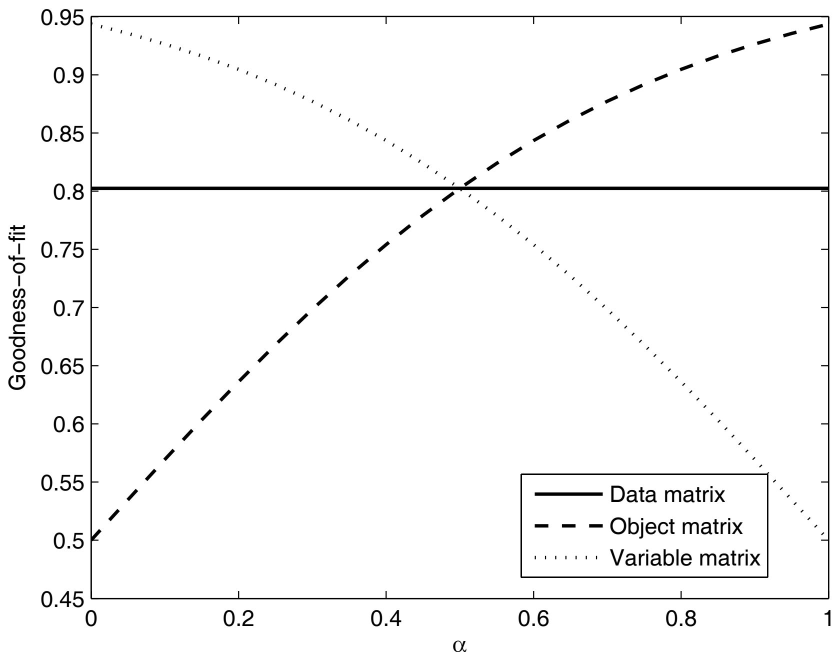

Figure 1: Goodness-of-fit coefficients for data, object, and variable matrices with respect to α.

columns i and j can be effected via right multiplication by the matrix

$$F = \begin{bmatrix} 1 & & & | & & & & | & & \\ & \ddots & & | & & & & | & & \\ & & 1 & | & & & & | & & \\ -- & -- & -- & 0 & -- & -- & -- & 1 & -- & --_i \\ & & & | & 1 & & & | & & \\ & & & | & & \ddots & & | & & \\ & & & | & & & 1 & | & & \\ -- & -- & -- & 1 & -- & -- & -- & 0 & -- & --_j \\ & & & | & & & & | & \ddots & \\ & & & |_i & & & & |_j & & 1 \end{bmatrix},$$

in which the two off-diagonal 1's are in the (i,j) and (j,i) positions and all unspecified entries are 0. Matrix F can also be obtained by doing the same column operation to identity matrix. We name matrix F as transposition matrix, a special type of permutation matrix [8]. In general, if we place k columns of zeros somewhere in X_0, then we need a number of column interchanges to move them all to the last columns. Consequently, we gain also a same number of transposition matrices, which have all to be right multiplied. Since the product of two permutation matrices is again a permutation matrix, they can be represented by a single transposition matrix F. We may write $X_0 = XF$, where X is now $n \times p$ matrix whose k last columns are zeros. Adjustment by rotation demands right multiplication of XF by an orthogonal matrix Q_0. As a permutation matrix, F is certainly orthogonal. Since the product of two orthogonal matrices is an orthogonal matrix, we can write $(XF)Q_0 = XQ$, where orthogonal matrix $Q := FQ_0$ should optimally be sought. It is then concluded that, without loss of generality, additional columns of zeros can be placed in the last.

References

[1] J. Aitchison and M. Greenacre, Biplots for compositional data, *Applied Statistics*, **51** (2002), 375–392.

[2] T. Bakhtiar and Siswadi, Orthogonal procrustes analysis: its transformation ordering and minimal distance, *International Journal of Applied Mathematics and Statistics*, **20** (2011), 16–24.

[3] P.G.N. Digby and R.A. Kempton, *Multivariate analysis of ecological communities*, Chapman & Hall, New York, 1987.

[4] K.R. Gabriel, The biplot graphic display of matrices with application to principal component analysis, *Biometrika*, **58** (1971), 453–467.

[5] K.R. Gabriel, Goodness-of-fit of biplots and correspondence analysis, *Biometrika*, **89** (2002), 423–436.

[6] J.C. Gower and G.B. Dijksterhuis, *Procrustes problems*, Oxford University Press, New York, 2004.

[7] M.J. Greenacre, *Biplots in practice*, BBVA Foundation, Madrid, 2010.

[8] R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge University Press, New York, 1985.

[9] J.A. Martin-Fernandez, C. Barcelo-Vidal and V. Pawlowsky-Glahn, Dealing with zero and missing values in compositional data sets using nonparametric inputation, *Mathematical Geology*, **35** (2003), 256–278.

[10] Siswadi and T. Bakhtiar, Goodness-of-fit of biplots via procrustes analysis, *Far East Journal of Mathematical Sciences*, **52** (2011), 191–201.

[11] N.P. Subanu, Praswanto, B. Dzulkarnain and D. Sundari, Informasi penelitian komponen jamu pengatur haid, *Cermin Dunia Kedokteran*, **108** (1996), 10–16.

Received: March, 2012