American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-7, Issue-5, pp-264-271
www.ajer.org
Research Paper Open Access

The Accuracy Level Of Biplot Analysis Based On The Variance-Covariance Matrix

Georgina Maria Tinungki
Department of Mathematics Faculty of Mathematics and Natural Science Hasanuddin University Makassar 90245, Indonesia
Corresponding Author: Georgina Maria Tinungki

ABSTRACT : *Biplot analysis is categorized into a double variable exploration analysis, which intended to present multiple variable data in two-dimensional map, so that the data behavior is easy to be seen and interpreted. Biplot is a descriptive statistical technique that can be presented visually in order to simultaneously present the (n)observational object and the (p)variable in plane space, so that the characteristics of the variables and the observation objects, and relative positions between observed objects and variables can be analyzed.The accuracy of Biplot analysis in explicating the total variance of matrix in the original data is formulated as follows:* $\eta = \frac{(\lambda_1 + \lambda_2)}{\sum_{i=1}^{r}\lambda_i}$. *If the real dimension of X is2, then* η *=1. If r* $\geq$ *3, then* η *<1. If the value of* η *approaches the value of one,means that Biplot obtained from a square approached matrix will provide an improved presentation of the information contained in the original data.*

KEYWORDS : *- Accuracy of Biplot analysis , Biplot analysis, multiple variable.*

Date of Submission: 27-04-2018 Date of acceptance: 12-05-2018

I INTRODUCTION

Biplot is an analysis which is applicable both for positioning and perceptual mapping from a set of observational points (product, service, or company).In the process, Biplot analysis requires data from a number of observational points with variables (Gower, 2011). The result of this analysis will be displayed in the form of a two-dimensional display, which contains the information: 1. The relative position between objects. According to this information, the two objects, which have the closest distance, are inferred to have a high degree of similarity based on the observed attributes, compared to longer-distance objects (Greenacre, 2010). This information can be used as the basis of segmentation.2. Relationship between attributes. From this information, the linear relationship (correlation) between attributes and the importance of an attribute based on its diversity (variance) will be known.3. Through the merging of information 1 and 2 known as "bi-plot", the characteristics of each object (group of objects or segments) based on observed attributes will be known. Another use of this analysis is that it can be used in segmentation or translating segments formed from other analysis results, such as cluster analysis. Biplot was introduced by Gabriel in 1971. Biplot is categorized into the multivariate exploration in low-dimensional space, usually two (or three), so that the data behavior is easy to be seen and interpreted. In the depiction in two-dimensional space, Biplot displays a relative position plot of the *n* observation with *p* variant simultaneously. In other words, Biplot is a data display in two-dimensional form using *n* observation, and simultaneously with plots of *p* variables so that the information about the relationship between variables and observations can be obtained (Mevik, 2007). The data, which is displayed visually using Biplot, can observe the relationships between variables, relative similarities between observations, and relative positions between observations and variables (Gunarto, 2014).

II LITERATURE REVIEW

II.1 Multivariat Analysis :

*Multivariate Analysis*is a statistical method that simultaneously include many variables in analyzing data (Gabrie, 1971). In multivariate data built by an investigator, it involves *p*variable which is simultaneously measured at each *n* item, individualor experimental unit (Jollife, 1986). To facilitate the data management, multivariate data is shown in matrix form with *n* row and *p* column.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nj} & \cdots & x_{np} \end{bmatrix} = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_i \\ \vdots \\ x'_n \end{bmatrix}$$

For example, x_{ij} is used to express the measurement result ofthe-*i*individualfor the-*j*variable, where $x'_i = \lfloor x_{i1}, x_{i2}, \ldots, x_{ip} \rfloor$isa measurement on the same unit so that it is correlated. Meanwhile, $x_1, x_2, \ldots, x_n$ are able to be correlated or not (Johnson & Wichern, 2002).

Oyedele (2014) suggests that multivariate analysis can be divided into two groups:
1) Dependence analysis, aims to explain or forecast the value of dependent variables based on more than one independent variable which affecting.
2) Interdependence analysis, aims to analyze the interrelationships of all variables without distinction made over the types of variables. Moreover, ittries to explain the underlying structure of the data by simplifying the complexity of the data.

II.2 Biplot Analysis :

Biplot is a dual dimensional descriptive statistical technique based on Singular Value Decomposition (SVD (Mattjik,, 2011)). For example, an nxp-sized data matrix X containing n observation and p variable corrected to their mean and r square, can be written in the form X = U L A '.
Description:
Matrix U (nxr) and A(pxr). So, U'U = A'A = Ir
L is a (rxr) diagonal matrix whose diagonal element is the square root of the characteristic root X'X or XX'. So, ' 1λ 2λ ≥ rλ ≥....≥ .
The matrix column A is the characteristic vectorwhich is commensurate with the matrix X'X or XX'.λand characteristic roots. The rows of Matrix U can be calculated by using the following formula:
$Ui = (1/\sqrt{\lambda_i})\ x\ a_i$
Where λ_i is the characteristic root of the-i of the matrix X'X and a_i is the-i row of matrix A.
Interpretation and Information obtained from Biplot
1. The proximity between objects. Two objects with similar characteristics will be described as two factors whose positions are close together.
2. The diversity of variables. Variables with small diversity are described as short vectors, and vice versa.
3. The Relationship between variables: If the angle of two variables <900, then the correlation is positive. If the angle of two variables > 900, then the correlation is negative. The smaller the angle is, the stronger the correlation.
4. The Value of variable on an object. Characteristics of an object can be inferred from its closest relative position to a variable.

Information can be obtained from Biplot
1. The relationship between variables
2. The relative similarity between objects of observation
3. The relative position between objects of observation with variables

Noteworthy facts in Biplot .Representing the reduction of space vwith large dimension into two-dimensional space.A consequence of the information reduction contained in Biplot is minimally 70% of information contained in Biplot.

III RESULT AND DISCUSSION

For example, a multivariate observation consists of *p*variable vector $X_1, X_2, \ldots, X_p$, with an average sample of the*n*measurement at the-*j* variable is given by

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij} = \frac{(x_{1j}.1 + x_{2j}.1 + \cdots + x_{nj}.1)}{n} \tag{2.1}$$

So,

$$\bar{x} = \frac{1}{n}\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \frac{1}{n}\mathbf{X}'\mathbf{1} \tag{2.2}$$

1 shows column vector $n \times 1$.

The average matrix of the sample is written in the form below:

$$\overline{\mathbf{X}} = \mathbf{1}\bar{x}' = \frac{1}{n}\mathbf{1}\mathbf{1}'\mathbf{X} = \begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \\ \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \end{bmatrix} \tag{2.3}$$

The sample variance is often used to explain the magnitude of variance in the measurement of a variable. When the *p*variable is observed in each unit, the variance can be described by variance-covariance matrix.

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{12} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{1p} & s_{2p} & \cdots & s_{pp} \end{bmatrix}$$

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})' \tag{2.4}$$

$$= \frac{1}{n-1}\begin{bmatrix} x_1 - \bar{x}_1 \\ x_2 - \bar{x}_2 \\ \vdots \\ x_p - \bar{x}_p \end{bmatrix}\begin{bmatrix} x_1 - \bar{x}_1 & x_2 - \bar{x}_2 & \cdots & x_p - \bar{x}_p \end{bmatrix}$$

$$= \frac{1}{n-1}\begin{bmatrix} x_{11} - \bar{x}_1 & x_{21} - \bar{x}_1 & \cdots & x_{n1} - \bar{x}_1 \\ x_{12} - \bar{x}_2 & x_{22} - \bar{x}_2 & \cdots & x_{n2} - \bar{x}_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_{1p} - \bar{x}_p & x_{2p} - \bar{x}_p & \cdots & x_{np} - \bar{x}_p \end{bmatrix} \times \begin{bmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \cdots & x_{1p} - \bar{x}_p \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \cdots & x_{2p} - \bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & \cdots & x_{np} - \bar{x}_p \end{bmatrix}$$

$$= \frac{1}{n-1}(\mathbf{X} - \overline{\mathbf{X}})'(\mathbf{X} - \overline{\mathbf{X}}) \tag{2.5}$$

For example, to know the diversity of data variable *p*, it can be discovered by using the main component *k*, where $k < p$. If the informationabout the diversity of data from the initial variables in the main component *k* is wanted to be acquired as much as possible, then the main component *k* can substitute the initial variables, and the initial data set containing the *n*-sized variable can be reduced into the*k* main component sized *n*.
The main component is a linear combination of the observed variables. The main component depends on the variance-covariance matrix as shown in the equation (2.5). If the main component depends on this matrix, it is said that the determination of the main component, which will then be used in graphical representation using Biplot Analysis, is based on the **variance-covariance matrix S.**
For example, a random variable $X = (X_1, X_2, \cdots, X_p)$ has a variance-covariance matrix**S** with pairs of eigenvalues and eigenvectors (λ_1, e_1), (λ_2, e_2), (λ_p, e_p), where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$. Then, the main component of the-*k* is denoted by

$$Y_k = e_k'\mathbf{X} = e_{k1}X_1 + e_{k2}X_2 + \cdots + e_{kp}X_p, \qquad k = 1,2,\ldots,p \qquad (2.6)$$

The eigenvalue of the matrix **S** is derived from a characteristic equation

$$|\mathbf{S} - \lambda\mathbf{I}| = 0 \qquad (2.7)$$

Where, the eigenvector**x**corresponding to the eigen value, λ,if

$$\mathbf{Sx} = \lambda\mathbf{x} \qquad (2.8)$$

From (2.6), $e_k = \dfrac{x_k}{\sqrt{x_k' x_k}}$ is a normalized eigenvector, where x_k is an eigenvector corresponding to the eigen value λ_k.
Furthermore, if the main component is obtained from the result of standardization of the produced main component, then the producedmain component is based on the standardized observations with the correlation matrix **R**.The main component of the-*k*is shown below:

$$Y_k = e_k'\mathbf{z} = e_{k1}z_1 + e_{k2}z_2 + \cdots + e_{kp}z_p, \qquad k = 1,2,\ldots,p \qquad (2.9)$$

Where,$(\lambda_k, e_k)$is a pair of eigen values and eigenvectors of **R** with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$.
Basically the correlation matrix is a standardized variance-covariance matrix. The correlation matrix can be constructed by substituting each x_{ij} observationwith $(x_{ij} - \bar{x}_j)/\sqrt{s_{jj}}$. The matrix obtained from the standardized observations is as follows

$$\mathbf{Z} = \begin{bmatrix} z_1' \\ z_2' \\ \vdots \\ z_n' \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1p} \\ z_{21} & z_{22} & \cdots & z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{np} \end{bmatrix} = \begin{bmatrix} \dfrac{x_{11}-\bar{x}_1}{\sqrt{s_{11}}} & \dfrac{x_{12}-\bar{x}_1}{\sqrt{s_{12}}} & \cdots & \dfrac{x_{1p}-\bar{x}_1}{\sqrt{s_{1p}}} \\ \dfrac{x_{21}-\bar{x}_2}{\sqrt{s_{12}}} & \dfrac{x_{22}-\bar{x}_2}{\sqrt{s_{22}}} & \cdots & \dfrac{x_{2p}-\bar{x}_2}{\sqrt{s_{2p}}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{x_{n1}-\bar{x}_p}{\sqrt{s_{1p}}} & \dfrac{x_{n2}-\bar{x}_p}{\sqrt{s_{2p}}} & \cdots & \dfrac{x_{np}-\bar{x}_p}{\sqrt{s_{pp}}} \end{bmatrix} \qquad (2.10)$$

With average vector

$$\bar{z} = \frac{1}{n}\mathbf{Z}'\mathbf{1} = \frac{1}{n}\begin{bmatrix} \sum_{i=1}^{n} \dfrac{x_{i1}-\bar{x}_1}{\sqrt{s_{11}}} \\ \sum_{i=1}^{n} \dfrac{x_{i2}-\bar{x}_2}{\sqrt{s_{22}}} \\ \\ \sum_{i=1}^{n} \dfrac{x_{ip}-\bar{x}_p}{\sqrt{s_{pp}}} \end{bmatrix} = 0 \qquad (2.11)$$

Thus, the standardized variance-covariance matrix is written as

$$\mathbf{S}_z = \frac{1}{n-1}\left(\mathbf{Z}-\frac{1}{n}\mathbf{11}'\mathbf{Z}\right)'\left(\mathbf{Z}-\frac{1}{n}\mathbf{11}'\mathbf{Z}\right)$$

$$= \frac{1}{n-1}\left(\mathbf{Z}-\mathbf{1}\bar{z}'\right)'\left(\mathbf{Z}-\mathbf{1}\bar{z}'\right)$$

$$\mathbf{S}_z = \frac{1}{n-1}\mathbf{Z}'\mathbf{Z} \tag{2.12}$$

$$= \frac{1}{n-1}\begin{bmatrix} \frac{(n-1)s_{11}}{s_{11}} & \frac{(n-1)s_{12}}{\sqrt{s_{11}}\sqrt{s_{22}}} & \cdots & \frac{(n-1)s_{1p}}{\sqrt{s_{11}}\sqrt{s_{pp}}} \\ \frac{(n-1)s_{12}}{\sqrt{s_{11}}\sqrt{s_{22}}} & \frac{(n-1)s_{22}}{s_{22}} & \cdots & \frac{(n-1)s_{2p}}{\sqrt{s_{22}}\sqrt{s_{pp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{(n-1)s_{1p}}{\sqrt{s_{11}}\sqrt{s_{pp}}} & \frac{(n-1)s_{2p}}{\sqrt{s_{22}}\sqrt{s_{pp}}} & \cdots & \frac{(n-1)s_{pp}}{s_{pp}} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{s_{12}}{\sqrt{s_{11}}\sqrt{s_{22}}} & \cdots & \frac{s_{1p}}{\sqrt{s_{11}}\sqrt{s_{pp}}} \\ \frac{s_{12}}{\sqrt{s_{11}}\sqrt{s_{22}}} & 1 & \cdots & \frac{s_{2p}}{\sqrt{s_{22}}\sqrt{s_{pp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{s_{1p}}{\sqrt{s_{11}}\sqrt{s_{pp}}} & \frac{s_{2p}}{\sqrt{s_{22}}\sqrt{s_{pp}}} & \cdots & 1 \end{bmatrix}$$

$$= \mathbf{D}^{-1/2}\mathbf{S}\mathbf{D}^{-1/2} = \mathbf{R} \tag{2.13}$$

$\mathbf{D}^{-1/2}$ is diagonal matrix, where $diag\left(\mathbf{D}^{-1/2}\right) = \left(1/\sqrt{s_{11}}, 1/\sqrt{s_{22}}, \ldots, 1/\sqrt{s_{pp}}\right)$.

Biplot Construction

Biplot is a dual dimensional descriptive statistical technique that can be displayed visually by simultaneously displaying a set of observed objects and variables in a graph on a plane figure so that the characteristics of the variable and the object of observation and the relative position between the observed object and the variable can be analyzed.

The fundamental of this analysis is that each matrix $n\times p$with square$r$ [min $\{n,p\}$] can be depicted exactly in a dimensionspace r. For matrixwith square$r$and want to be well depicted in a dimensional space k, where$k \le r$, then an optimum approximation with a square matrix $k$is applied. From the best approached matrix, then will be described the object configuration and variables in the dimensional space$k$. In Biplot analysis, k = 2,thus, the approach can be described in a two-dimensional space. Biplot analysis is based on the Singular Value Decomposition (SVD) of a matrix. It is found that SVD are able to reorient the coordinate axis whichcause the matrix data is closer to the pattern created from the point of the matrix itself.

The general form of SVD, for instance, an$n\times p$sized data matrix X containing $n$observations and $p$variables, then the matrix X can be written in the form:

$$\underset{(n\times p)}{\mathbf{X}} = \underset{(n\times n)}{U}\,\underset{(n\times p)}{\Lambda}\,\underset{(p\times p)}{V'} \tag{2.14}$$

Where

U : The orthogonal matrix of XX'

Λ : The diagonal matrix whose elements are singular value $\ell_1 \geq \ell_2 \geq \dots \geq \ell_r \geq 0$, r = min (n,p) which is the square root of the eigenvalue $\lambda_1, \lambda_2, \dots, \lambda_r$ matrix XX' or $X'X$

V : The orthogonal matrix of $X'X$

Specifically,there is r constants $\ell_1, \ell_2, \dots, \ell_r$ so, the equation (2.14) can be written as

$$\mathbf{X} = \sum_{i=1}^{r} \ell_i u_i v_i' = U_r \Lambda_r V_r' \tag{2.15}$$

With $U_r = [u_1, u_2, \dots, u_r]$ and $V_r = [v_1, v_2, \dots, v_r]$, where, these two vectors are normalized eigenvectors. Equation (2.14) can be written as follows:

$$\mathbf{X} = U\Lambda^{\alpha} - \Lambda^{1-\alpha}V' \tag{2.16}$$

For example,$\mathbf{G} = U\Lambda^{\alpha}$ and $\mathbf{H}' = \Lambda^{1-\alpha}V'$, so then the-($i$,$j$) element of matrix**X**can be written in the form $x_{ij} = g_i' h_j$, where g_i' is the-i row from**G**and h_j isthe-$j$column from**H**. If **X**is a square, then the effect vector of the row g_i and h_j can be described exactly in a 2-dimensional space.If the matrix **X** value is more than square, itis usually approached by a matrix with square root.

So, the element of the-(i, j) matrix X is written as follows:

$$_2 x_{ij} = g_{i^*}' h_{j^*} \tag{2.17}$$

For each g_i^* dan h_j^*,contain the first two elements of the vector g_i and h_j. So, for r = 2 the plot consists of points $g_1, \dots, g_n$ and $h_1, \dots, h_p$, where $G'_{2\times n} = \left(g_{1_{2\times1}}, \dots, g_{n_{2\times1}}\right)$and $H'_{2\times p} = \left(h_{1_{2\times1}}, \dots, h_{p_{2\times1}}\right)$. It is the same with factorization of JK'.

Constructing Biplot of the main component of the sample, in which the main component of the sample can be obtained from the variance-covariance matrix or correlation matrix.

Biplot Based on Variance-Covariance Matrix

Matrix**S**on the equation (2.5) can be rewritten in the form $X_c' X_c$ = (n-1) **S**, where in the previous equation(2.17), it is known that X_c isthe data matrix corrected against its mean, $X_c = (x_i - \bar{x})$.

Biplot construction is started by describing the matrix X_c into the SVD form

$$\underset{(n\times p)}{X_c} = \underset{(n\times p)}{U}\,\underset{(p\times p)}{\Lambda}\,\underset{(p\times p)}{V'} \tag{2.14}$$

Λ is a diagonal matrix where, diag(Λ)=($\ell_1, \ell_2, \dots, \ell_p$), $U$and$V$is an orthogonal matrix whose each column being the eigenvector of $X_c X_c'$ and $X_c' X_c$ = (n-1)**S.**

In interpreting Biplot, Khattree dan Naik (2000) choose the value α = 0, α = 1/2, and α = 1. If α = 0, then **G** = U and **H** = $V\Lambda$ .Since**G** = $U$is orthogonal matrix, then $\mathbf{G}'\mathbf{G} = \mathbf{I}$, as its consequence.

$$\begin{aligned}\underset{(p\times p)}{X'X} &= \left(\underset{(n\times r)}{G}\,\underset{(r\times p)}{H'}\right)'\left(\underset{(n\times r)}{G}\,\underset{(r\times p)}{H'}\right)\\ &= \underset{(p\times r)}{H}\,\underset{(r\times n)}{G'}\,\underset{(n\times r)}{G}\,\underset{(r\times p)}{H'}\\ &= \underset{(p\times r)}{H}\,\underset{(r\times p)}{H'}\end{aligned} \tag{2.15}$$

Due to the fact that $X'X = HH'$ = (n-1) **S**, the multiplication result of $h_i' h_j$ will be equal to (n-1) s_{ij} and the length of vector h_i describes the diversity of variables the-i. The correlation between the-i and the-j variable is shown by the angle cosine value between the vectors h_i and h_j .Variables that have a high negative correlation will be displayed in the form of two rows in opposite directions, or form a wide angle (blunt), while the uncorrelated one will be displayed with two rows with an acute angle.

If α = 1/2, then $\mathbf{G} = U\Lambda^{1/2}$ and $\mathbf{H} = V\Lambda^{1/2}$. By having the option, the observations and variables can be directly interpreted. Therefore, Biplot based on the option is usually used in practice.

If α = 1, then $\mathbf{G} = U\Lambda$ and $\mathbf{H} = V$. Since $\mathbf{H} = V$, where V is an orthogonal matrix, so then $\mathbf{H}'\mathbf{H} = \mathbf{I}$. Thus it is obtained

$$\underset{(n\times n)}{XX'} = \left(\underset{(n\times r)}{\mathbf{G}}\underset{(r\times p)}{\mathbf{H}'}\right)\left(\underset{(n\times r)}{\mathbf{G}}\underset{(r\times p)}{\mathbf{H}'}\right)'$$
$$= \underset{(n\times r)}{\mathbf{G}}\underset{(r\times p)}{\mathbf{H}'}\underset{(p\times r)}{\mathbf{H}}\underset{(r\times n)}{\mathbf{G}'}$$
$$= \underset{(n\times r)}{\mathbf{G}}\underset{(r\times n)}{\mathbf{G}'} \tag{2.16}$$

At the state of the Euclid distance between g_i and g_j will approach the Euclid distance between the -$i$andthe-j on the rows of the data matrix X .

Biplot Based on The Correlation Matrix

Biplot analysis based on the correlation matrix basically uses the same procedure as Biplot based on the variance-covariance matrix. The difference lies in the input data matrix which is used. Based on the equation (2.8), it can be written $X_z' X_z$ = (n-1)**R**where X_z is a standardized data matrix

$$X_z = (x_i - \bar{x})D^{-1/2} \tag{2.17}$$

By using the matrix X_z in the equation (2.17) as the input data matrix, and in the same way as the Biplot based on the variance-covariance matrix, the Biplot coordinates can be obtained.

Biplot analysis based on these two different types of matrix, will provide different interpretations. However, the type of information obtained about the observed variables and objects will be the same.

The Accuracy Level of Biplot Analysis

The accuracy of Biplot analysis in explaining the total variance of the original data matrix, formulated as follows

$$\eta = \frac{(\lambda_1 + \lambda_2)}{\sum_{i=1}^{r} \lambda_i} \tag{2.18}$$

Where, λ_1 = The first largest eigenvalue

λ_2 = The second largest eigenvalue

λ_i = The-Ilargest eigenvalue

IV CONCLUSION

Based on the results and discussion above, the conclusion is as follows:

1. If the actual dimension of Xis2, then η =1. If $r \geq 3$, then η <1. So, if the value is closer to the value of one, it means that Biplot obtained from the matrix of the squareapproach will provide a better presentation of the information contained in the originaldata.
2. The reduction of dimension makes the information contained in the Biplot decrease.Thus, Biplot which is able to provide 70% of all information isquite accurate.

REFERENCES

[1]. Gabriel, K.R., (1971), The Biplot Graphic Display of Matrices With Application to Principal Analysis, Biometrika, 58(3), hal.453-467.

[2]. Gower JC, Lubbe S, Le Roux (2011) Understanding Biplots. John Wiley & Sons: Chicester, UK.

[3]. Greenacre MJ (2010) Biplots in Practice. Fundación BBVA: Barcelona, Spain.

[4]. Gunarto, Muji dan Syarif, Muhammad Amiruddin. 2014.Penggunaan Analisis Biplot pada Pemetaan Perguruan Tinggi Swasta di Kota Palembang, Forum Manajemen Indonesia 6 Medan.

[5]. Hawkins, D.M., Liu, L., Young, S.S., (2001), Robust Singular Value Decomposition, National Institute of Statistical Science, Vol. 122, hal 1-12.
[6]. Jollife and Rawlings. (1986), Primcipal Component Analysis, Springer-Verlag, New York.
[7]. Mattjik, Ahmad Ansori dan Sumertajaya, I Made. (2011). Sidik Peubah Ganda dengan menggunakan SAS. Institut Pertanian Bogor (IPB) : Departemen Statistika.
[8]. Mevik BH, Wehrens R (2007) The pls Package: Principal Component and Partial Least Squares Regression in R. Journal of Statistical Software 2(18): 1-24.
[9]. Oyedele OF, Gardner L, Sugnet (2014) The Covariance Biplot to Reveal Relationships between Different Sets of Variables. Int Sci Technol J Namibia 3(1): 107-113.

Georgina Maria Tinungki. " The Accuracy Level Of Biplot Analysis Based On The Variance-Covariance Matrix"American Journal of Engineering Research (AJER), vol. 7, no. 5, 2018, pp.264-271.