



# GH Biplot in Reduced-Rank Regression Based on Partial Least Squares

Willin Alvarez <sup>1,\*</sup>, Victor Griffin <sup>2</sup>

<sup>1</sup> *Universidad Regional Amazonica Ikiam, Country Ecuador/Departamento de Matematicas, Facultad de Ciencia y Tecnologia. Universidad de Carabobo, Venezuela*

<sup>2</sup> *Departamento de Matematicas, Facultad de Ciencia y Tecnologia. Universidad de Carabobo, Venezuela*

**Abstract** One of the challenges facing statisticians is to provide tools to enable researchers to interpret and present their data and conclusions in ways easily understood by the scientific community. One of the tools available for this purpose is a multivariate graphical representation called reduced rank regression biplot. This biplot describes how to construct a graphical representation in nonsymmetric contexts such as approximations by least squares in multivariate linear regression models of reduced rank. However multicollinearity invalidates the interpretation of a regression coefficient as the conditional effect of a regressor, given the values of the other regressors, and hence makes biplots of regression coefficients useless. So it was, in the search to overcome this problem, Alvarez and Griffin [1], presented a procedure for coefficient estimation in a multivariate regression model of reduced rank in the presence of multicollinearity based on PLS (Partial Least Squares) and generalized singular value decomposition. Based on these same procedures, a biplot construction is now presented for a multivariate regression model of reduced rank in the presence of multicollinearity. This procedure, called PLSSVD GH biplot, provides a useful data analysis tool which allows the visual appraisal of the structure of the dependent and independent variables. This paper defines the procedure and shows several of its properties. It also provides an implementation of the routines in R and presents a real life application involving data from the FAO food database to illustrate the procedure and its properties.

**Keywords** GH biplot, reduced-rank regression, partial least squares, singular value decomposition.

**AMS 2010 subject classifications** 62H25, 65F22

**DOI:** 10.19139/soic-2310-5070-1112

## 1. Introduction

One of the challenges facing statisticians is to provide tools to enable researchers to interpret and present their data and conclusions in ways easily understood by the scientific community. Among the many tools now available for this purpose are multivariate graphical representations arising from factorial families such as principal component analysis, correspondence analysis and biplot analysis. These tools are highly appreciated and often used by researchers. The biplot, for example, was introduced originally by Gabriel in 1971 as a graphical representation of the rows and columns of a data matrix based on the singular value decomposition (SVD) [7], and is widely used having been incorporated in the majority of statistical packages. The biplot analysis approximates a data matrix  $Y$  by another matrix  $Y_r = G_r H_r'$ , where  $G_r$  and  $H_r$  are matrices of complete column rank obtained from the first  $r$  terms of its singular value decomposition. The row vectors of  $G_r$  are called row markers, the row vectors of  $H_r$  are called column markers and the elements of the original data matrix are approximated by the internal product of these markers. Gabriel called this factorization a biplot factorization. Three classical biplots are the GH biplot (column metric preserving biplot), the JK biplot (row metric preserving biplot) and the SQRT biplot (symmetric biplot). Examples of biplot applications can be found in many different areas including medicine ([16], [10]) and

\*Correspondence to: Willin Álvarez (Email: willin.alvarez@ikiam.edu.ec/willingabriel@gmail.com). Universidad Regional Amazónica Ikiam.

meteorology, [20]. A review of biplot applications can be found in [5].

For Gabriel, the variables involved in the data matrix to be represented graphically were in a symmetric context, that is, the variables were not classified as dependant or independent (see [13]), however the non-symmetric context is very necessary when wanting to explain relations between variables such as in regression models. Work continues in this context. Ter Braak and Looman [18] describe how to construct the graphical biplot representation in nonsymmetric contexts such as approximations by least squares in multivariate linear regression models of reduced rank and justify their work in the following way "...In our view, reduced-rank regression is particularly useful, because it offers the possibility to produce a plot that easily summarizes the major aspects of the regressions. This possibility has so far not been exploited...The proposed biplot represents geometrically the fitted values of the regressions, the estimated regression coefficients and their associated t-ratios. The two-dimensional biplot is exact for the rank 2 mode..."

These authors demonstrated the usefulness of the reduced rank regression (RRR) biplot with an example using public health data. However they themselves made clear that "The reduced-rank regression was proposed by Davies & Tso [6] in a context with many predictor variables compared to the number of statistical units, so that regression coefficients are underdetermined due to multicollinearity . . . However, reduced-rank regression is no solution to the multicollinearity problem. Multicollinearity invalidates the interpretation of a regression coefficient as the conditional effect of a regressor, given the values of the other regressors, and hence makes biplots of regression coefficients useless". So it was, in the search to overcome this problem, Alvarez and Griffin ([1]), presented a procedure for coefficient estimation in a multivariate regression model of reduced rank in the presence of multicollinearity based on PLS (Partial Least Squares) and generalized singular value decomposition.

The objective of this paper is to continue the work proposed by Ter Braak and Looman in 1994 [18] in which the authors emphasized the importance of reduced rank regression and showed the possibility of using a biplot (see [8],[9]) to visualize the reduced rank model. They called it a Biplot in Reduced Rank Regression. However many questions immediately arise when considering the presence of multicollinearity among the independent variables. For example,

1. Is it possible to use the theory of partial least squares (PLS) to construct biplots in Reduced Rank Regression, If so,
2. What properties characterize this graphical representation,
3. Is it necessary to take the multicollinearity into account when constructing the biplot or is it possible to construct and interpret the biplot in the classical way.

We will now look at some preliminary facts which will help to provide the answers to these questions.

### 1.1. Preliminary ideas

As done by many authors, for example [3], [12], [6], [2] and [1], the study of multivariate regression in reduced Rank (RRR) can be considered a special case of the classic multivariate regression model:

$$Y = X\Theta + E \quad (1)$$

in which  $Y$  is a  $(n \times q)$  matrix where each row of  $Y$  contains the values of the  $q$  dependent variables measured on  $n$  subjects and  $X$  is a  $(n \times p)$  matrix where each row of  $X$  contains the values of the  $p$  independent variables measured on  $n$  subjects. We assume that  $X$  is fixed from sample to sample.  $\Theta$  is a  $(p \times q)$  matrix of parameters or regression coefficients and  $E$  is the  $(n \times q)$  matrix of errors or random perturbations.

In total agreement with the ideas of Ter Braak and Looman [18], Alvarez and Griffin [1] proceeded to factor the coefficient matrix as  $\Theta = AB'$  so that the multivariate reduced rank regression model became:  $Y = XAB' + E$ . The estimation of this model consists in minimizing  $\|(Y - XAB')\Gamma\|^2 = \|(Y - \hat{Y})\Gamma\|^2 + \|(\hat{Y} - XAB')\Gamma\|^2$

using as an initial estimation for  $\hat{Y}$  that obtained by SVD as:

$$\hat{Y}\Gamma = QDV' \quad (2)$$

where  $Q$  and  $V$  are orthogonal matrices of order  $(n \times t)$  and  $(q \times t)$  respectively with  $t = \min(n, q)$ , that contain the eigenvectors of  $(\hat{Y}\Gamma)(\hat{Y}\Gamma)'$  and  $(\hat{Y}\Gamma)'(\hat{Y}\Gamma)$  respectively;  $D$  is a diagonal matrix of order  $q$  containing the eigenvalues of these matrices in non-increasing order ( $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_t \geq 0$ ); and  $\Gamma$  is a positive definite matrix. Up to this point, the steps of the estimation procedures proposed by [18] and [1] coincide. However when wanting to estimate  $A = (X'X)^{-1}X'Q$ , the inconvenience of multicollinearity in the variables  $X$  arises. Now that  $Q = XA$ , it is possible to make a comparison with the regression model with  $Q$  representing the matrix of dependent variables,  $X$  the matrix of independent variables and  $A$  the matrix of regression coefficients. This motivated Álvarez and Griffin [1] to apply Partial Least Squares Regression (PLSR) to estimate  $A$ . PLSR is a technique that generalizes and combines the characteristics of principal component analysis and multivariate regression. It is particularly useful when it is necessary to predict a set of dependent variables in terms of a set of independent variables in the presence of severe or strong multicollinearity in which case the classical multivariate regression procedures have serious stability problems. PLSR is a good alternative in such circumstances since it uses basic iterative concepts to adjust bilinear models in several blocks of variables giving a solution to multivariate statistics in the presence of many collinear variables in the data and a modest number of observations. The solution of each stage permits the construction of the following matrices (see [1] for details):

$$T_h = [t_1, \dots, t_h]$$

$$W_h = [w_1, \dots, w_h]$$

$$P_h = [p_1, \dots, p_h]$$

$$\tilde{W}_h = [\tilde{w}_1, \dots, \tilde{w}_h]$$

where the subscript  $h$  indicates that the matrix comprises the sequence of the first  $h$  vectors.  $\tilde{W}$  spans the loading weights of the  $X$ -variables for a certain number of latent variables. Given the  $\tilde{w}_i = \prod_{i=1}^{h-1} (I - w_i p_i')$  (see [17], p.114), the two sets of weights  $w_i$  and  $\tilde{w}_i$ ,  $i = 1, 2, \dots, h$ , span the same space. The matrix  $T_h$  contains the scores of the  $X$ -variables and  $P_h$  is called the  $X$ -loadings. The relations between the matrices determine the estimation PLS of the model (1) as follows:

$$T_h = X\tilde{W}_h \quad (3)$$

$$P_h = X'T_h (T_h'T_h)^{-1} \quad (4)$$

$$\tilde{W}_h = W_h (P_h'W_h)^{-1} \quad (5)$$

Phatak and De Jong in 1997 [15], found the following very useful expression for the regression coefficient matrix obtained by the PLS regression:

$$\hat{\Theta}_{PLS} = \tilde{W}_h P_h' \hat{\Theta}_{OLS} = W_h (P_h'W_h)^{-1} P_h' \hat{\Theta}_{OLS} \quad (6)$$

This finding is of utmost importance since, as they themselves state, the matrix  $W_h (P_h'W_h)^{-1} P_h'$  is idempotent which implies that it is a projection matrix. Since it is not symmetric, it defines an oblique projection, which projects  $\hat{\Theta}_{OLS}$  in the space generated by the columns  $W_h$ . If we now apply this theory to the estimation of  $A$  using partial least squares where  $Q = XA$  is the model to be estimated, then

$$\hat{A}_{PLS} = \tilde{W}_h P_h' \hat{A}_{OLS} = W_h (P_h'W_h)^{-1} P_h' \hat{A}_{OLS} \quad (7)$$

which permits us to see the row markers ( $\hat{A}_{PLS}$ ) obtained from the PLS estimation as an oblique projection of the row markers obtained from a biplot in reduced rank regression using the OLS estimation ( $\hat{A}_{OLS}$ ). In addition,

using this initial estimate, the matrix of regression coefficients is obtained by combining the PLS estimation of  $A$  with the estimation of  $B$  given by SVD so that the reduced rank regression coefficient matrix given by PLSSVD is:

$$\hat{\Theta}_{PLS-SVD} = \hat{A}_{PLS} \hat{B}'_{SVD} \quad (8)$$

Finally the estimations of  $Y$  proposed by [1] are given by:

$$\hat{Y}_{PLS-SVD} = X \hat{\Theta}_{PLS-SVD} = X \hat{A}_{PLS} \hat{B}'_{SVD} \quad (9)$$

These procedures described in [1] provide the basis for this article which describes how to construct a biplot in reduced-rank regression in the presence of multicollinearity. We focus on describing the properties of 8 and its application to an example using data from the FAO data base.

The rest of this article is divided in three parts. The first describes the construction of the GH biplot in Reduced-Rank Regression based on Partial Least Squares that we have named PLSSVD GH biplot. The following core section describes properties of this biplot and provides their proofs. A previous reading of the articles [18], [15] and [1] is important for understanding this section. In the last part, we provide a real life example using food data from the FAO data base and routines in R provided in an appendix to apply the techniques and illustrate the properties of this biplot.

## 2. Biplot in reduced-rank regression

The chief cornerstone for the construction of a biplot in reduced-rank regression in the presence of multicollinearity using PLS is found in the paper of Ter Braak and Looman [18]. They showed how to factor the regression coefficient matrix using the techniques of [8] with OLS and SVD to obtain the biplot in reduced rank regression (RRR). This procedure for the RRR biplot can be summarized as follows:

1. Centralize or standardize the matrices  $X$  ( $n \times p$ ) and  $Y$  ( $n \times q$ ).
2. Express the multivariate regression model of (1) as:  $Y = XAB' + E$ , where  $\Theta = AB'$ .
3. Minimize  $\left\| (X'X)^{1/2} (\hat{\Theta}_{OLS} - AB') \Gamma \right\|^2$  with  $\Gamma$  a given weight matrix. The minimum for a rank  $r$  approximation is obtained by first applying the decomposition in singular values:

$$(X'X)^{-1/2} \Theta_{OLS} \Gamma = UDV' \quad (10)$$

and the minimum is then obtained by setting

$$\hat{A}_{OLS} = n^{1/2} (X'X)^{-1/2} [U]_r \text{ and } \hat{B}_{SVD} = n^{-1/2} \Gamma^{-1} [VD]_r \quad (11)$$

It is important to point out that the decomposition in singular values in 2 and 10 differ only in the matrices  $Q$  and  $U$  (see [18], p.991).

4. Finally, the estimation of the regression coefficient matrix is:  $\hat{\Theta}_{OLS-SVD} = \hat{A}_{OLS} \hat{B}'_{SVD}$ .

Remember that  $\Gamma$  is a given weight matrix. The results of the estimation procedure is affected by this weight matrix whose importance is amply explained in [1], [19] and [18]. In this paper  $\Gamma = I$  will be used.

## 3. Properties of the GH BIPILOT in Reduced-Rank Regression based on PLS (PLSSVD GH Biplot)

As we previously established, a biplot for the reduced rank regression coefficients begins by minimizing  $\left\| (X'X)^{1/2} (\hat{\Theta}_{OLS} - AB') \Gamma \right\|^2$ , from which we obtain  $A_{OLS} = (X'X)^{-1/2} U_r$  and  $B_{SVD} = \Gamma^{-1} [VD]_r$ . If we

assume that  $M = X'X$ , then:

$$\begin{aligned} A'_{OLS}MA_{OLS} &= \left[ (X'X)^{-1/2} U_r \right]' (X'X) (X'X)^{-1/2} U_r \\ &= U'_r (X'X)^{-1/2} (X'X)^{1/2} U_r \end{aligned}$$

This means that

$$A'_{OLS}MA_{OLS} = I_r \quad (12)$$

In the introduction we formulated some questions regarding the biplot in reduced-rank regression based on PLS in relation to its properties and interpretation. Two additional pertinent questions are

1. How is the geometry of PLS involved in the properties and
2. How simple or complicated is the mathematics needed to prove the properties

One of the reasons for writing this paper is to show the simplicity of the properties and how to apply this very useful data analysis tool. We now show how to use the property (12) and the geometric properties of PLS to establish the properties of the PLSSVD GH Biplot.

### 3.1. Properties of the column markers of the PLSSVD GH Biplot

**property 1.** In the PLSSVD GH Biplot the metric  $M = X'X$  satisfies :

$$A'_{PLS}MA_{PLS} = I_r \triangleleft$$

In essence:

$$\begin{aligned} A'_{PLS}MA_{PLS} &= \left( \tilde{W}_h P'_h \hat{A}_{OLS} \right)' (X'X) \left( \tilde{W}_h P'_h \hat{A}_{OLS} \right) \\ A'_{PLS}MA_{PLS} &= \hat{A}'_{OLS} P_h \tilde{W}'_h X' X \tilde{W}_h P'_h \hat{A}_{OLS} \end{aligned}$$

Associating conveniently:

$$\begin{aligned} A'_{PLS}MA_{PLS} &= \hat{A}'_{OLS} P_h (\tilde{W}'_h X') (X \tilde{W}_h) P'_h \hat{A}_{OLS} \\ A'_{PLS}MA_{PLS} &= \hat{A}'_{OLS} P_h (X \tilde{W}_h)' (X \tilde{W}_h) P'_h \hat{A}_{OLS} \end{aligned}$$

Since  $T_h = X \tilde{W}_h$  we can substitute this in the previous expression to obtain:

$$A'_{PLS}MA_{PLS} = \hat{A}'_{OLS} P_h T'_h T_h P'_h \hat{A}_{OLS}$$

Associating and transposing:

$$A'_{PLS}MA_{PLS} = \hat{A}'_{OLS} (T_h P'_h)' (T_h P'_h) \hat{A}_{OLS}$$

We also know that  $X = T_h P'_h$ , so that:

$$A'_{PLS}MA_{PLS} = \hat{A}'_{OLS} X' X \hat{A}_{OLS} = \hat{A}'_{OLS} M \hat{A}_{OLS} = I \quad (13)$$

**property 2.** The scalar product of the columns of  $\hat{\Theta}_{PLS-SVD}$  using the metric  $M$  approximates the scalar product of the column markers.  $\triangleleft$

In effect:

$$\begin{aligned} \left( \hat{\Theta}_{PLS-SVD} \right)' M \left( \hat{\Theta}_{PLS-SVD} \right) &= (A_{PLS} B'_{SVD})' M (A_{PLS} B'_{SVD}) \\ \left( \hat{\Theta}_{PLS-SVD} \right)' M \left( \hat{\Theta}_{PLS-SVD} \right) &= B_{SVD} A'_{PLS} M A_{PLS} B'_{SVD} \end{aligned}$$

Associating conveniently and applying property 1:

$$\left(\hat{\Theta}_{PLS-SVD}\right)' M \left(\hat{\Theta}_{PLS-SVD}\right) = B_{SVD} I B'_{SVD} = B_{SVD} B'_{SVD}$$

from which we obtain:

$$\hat{\theta}'_j M \hat{\theta}_k = b'_j b_k \text{ for } j \neq k \quad (14)$$

**corollary 1.** *The scalar product of the columns of  $\hat{Y}_{PLS-SVD}$  approximates the scalar product of the columns of  $\hat{\Theta}_{PLS-SVD}$  under the metric  $M$   $\triangleleft$*

In effect:

$$\hat{Y}'_{PLS-SVD} \hat{Y}_{PLS-SVD} = \left(X \hat{\Theta}_{PLS-SVD}\right)' \left(X \hat{\Theta}_{PLS-SVD}\right) = \hat{\Theta}'_{PLS-SVD} M \hat{\Theta}_{PLS-SVD} = B_{SVD} B'_{SVD}$$

from which we deduce that:

$$\hat{y}'_j \hat{y}_k = b'_j b_k \text{ for } j \neq k \quad (15)$$

**property 3.** *When the variables are centered, the squared length of the column markers approximates the sample covariance between the dependent variables.  $\triangleleft$*

In effect:

Let  $S$  be the sample covariance matrix and  $\hat{Y}_{PLS-SVD}$  centered at its mean so that:

$$S = \frac{\hat{Y}'_{PLS-SVD} \hat{Y}_{PLS-SVD}}{n-1} \quad (16)$$

By corollary 1:

$$S = \frac{B_{SVD} B'_{SVD}}{n-1} \quad (17)$$

From the equation 10 we know that:

$$S = \frac{\hat{Y}'_{PLS-SVD} \hat{Y}_{PLS-SVD}}{(n-1)^{1/2} (n-1)^{1/2}}$$

from which:

$$s_{jk} = \frac{\hat{y}'_j \hat{y}_k}{(n-1)^{1/2} (n-1)^{1/2}} \quad (18)$$

From corollary 1 and the equations (11) and (12), we have:

$$s_{jk} = \frac{b'_j b_k}{n-1} \quad (19)$$

From this we derive that the sample variance of a dependent variable can be expressed as:

$$s_{jj} = \frac{b'_j b_j}{n-1} = \frac{\|b_j\|^2}{n-1} = \frac{\|\hat{y}_j\|^2}{n-1}$$

and we can see that the standard deviation of any variable can be estimated by the associated column marker. This can be appreciated in the expression:

$$[(n-1) s_{jj}]^{1/2} = \|b_j\| = \|\hat{y}_j\| = (n-1)^{1/2} s_j \quad (20)$$

**property 4.** *The cosine of the angle between two column markers approximates the correlation between the dependant variables centered at their mean.  $\triangleleft$*

In effect:

Let  $b_j$  and  $b_k$  be two distinct column markers and  $\alpha$  the angle between them. By definition:

$$\begin{aligned} \cos(\alpha) &= \langle b_j \cdot b_k \rangle / (\|b_j\| \|b_k\|) \\ &= b'_j b_k / (\|b_j\| \|b_k\|) \end{aligned}$$

By corollary 1 and property 3

$$\cos(\alpha) = \hat{y}'_j \hat{y}_k / (\|\hat{y}_j\| \|\hat{y}_k\|)$$

The variables are centered at their mean so that by applying property 2 we have:

$$\cos(\alpha) = s_{jk} / (s_j s_k) \tag{21}$$

where  $s_{jk}$  is the sample covariance of the variables  $\hat{y}_j$  and  $\hat{y}_k$ ,  $s_j$  and  $s_k$  are the standard deviations of  $\hat{y}_j$  and  $\hat{y}_k$ , respectively.

**property 5.** *The quality of the column representation (QCR) under the metric  $M = X'X$  and the weight matrix  $\Gamma = I$  can be measured by*

$$QCR = \frac{\sum_{i=1}^r \lambda_i^2}{\sum_{i=1}^q \lambda_i^2} * 100 \triangleleft$$

We know that the sample covariance matrix  $S$  is given by:

$$S = \frac{\hat{Y}'_{PLS-SVD} \hat{Y}_{PLS-SVD}}{n - 1}$$

as  $B_{SVD} = \Gamma^{-1}VD$ , so that:

$$B_{SVD}B'_{SVD} = (\Gamma^{-1}VD) (\Gamma^{-1}VD)' = \Gamma^{-1}VD^2V'\Gamma^{-1}$$

By hypothesis  $\Gamma^{-1} = I$ , so that;

$$B_{SVD}B'_{SVD} = VD^2V'$$

which means that the total variability can be expressed as:

$$Trace \left( \hat{Y}'_{PLS-SVD} \hat{Y}_{PLS-SVD} \right) = Trace (VD^2V') = Trace (D^2V'V) = Trace (D^2) = \sum_{i=1}^q \lambda_i^2$$

thus,

$$QCR = \frac{\sum_{i=1}^r \lambda_i^2}{\sum_{i=1}^q \lambda_i^2} * 100 \tag{22}$$

### 3.2. Properties of the Row Markers

**property 6.** *The scalar product of the rows of  $\hat{\Theta}_{PLS-SVD}$  under the metric  $M = X'X$  and weight matrix  $\Gamma = I_r$  approximates the scalar product of the row markers.  $\triangleleft$*

Before proving this property we should take into account the property 6, where it was established that if  $\Gamma^{-1} = I_r$ , then  $B_{SVD}B'_{SVD} = VD^2V'$ . Now since the inverse of  $B_{SVD}B'_{SVD}$  is given by  $(B_{SVD}B'_{SVD})^{-1} = VD^{-2}V'$ , we see that:

$$\hat{\Theta}_{PLS-SVD} \left( \hat{\Theta}'_{PLS-SVD} M \hat{\Theta}_{PLS-SVD} \right)^{-1} \hat{\Theta}'_{PLS-SVD} = (A_{PLS} B'_{SVD}) (B_{SVD} B'_{SVD})^{-1} (A_{PLS} B'_{SVD})'$$

$$\hat{\Theta}_{PLS-SVD} \left( \hat{\Theta}'_{PLS-SVD} M \hat{\Theta}_{PLS-SVD} \right)^{-1} \hat{\Theta}'_{PLS-SVD} = A_{PLS} B'_{SVD} V D^{-2} V' B_{SVD} A'_{PLS}$$

Since  $B_{SVD} = VD$ :

$$\hat{\Theta}_{PLS-SVD} \left( \hat{\Theta}'_{PLS-SVD} M \hat{\Theta}_{PLS-SVD} \right)^{-1} \hat{\Theta}'_{PLS-SVD} = A_{PLS} D V' V D^{-2} V' V D A'_{PLS}$$

Taking into account that  $V$  is orthonormal, we have that  $V'V = I_r$ , and so:

$$\hat{\Theta}_{PLS-SVD} \left( \hat{\Theta}'_{PLS-SVD} M \hat{\Theta}_{PLS-SVD} \right)^{-1} \hat{\Theta}'_{PLS-SVD} = A_{PLS} D I_r D^{-2} I_r D A'_{PLS} = A_{PLS} D D^{-2} D A'_{PLS}$$

Additionally, since  $DD^{-2}D = I_r$ :

$$\hat{\Theta}_{PLS-SVD} \left( \hat{\Theta}'_{PLS-SVD} M \hat{\Theta}_{PLS-SVD} \right)^{-1} \hat{\Theta}'_{PLS-SVD} = A_{PLS} A'_{PLS} \quad (23)$$

This result indicates that the introduction of the metric  $M = X'X$  in the space generated by the columns of the matrix  $\hat{\Theta}_{PLS-SVD}$ , induces automatically an inner product with the metric  $\left( \hat{\Theta}'_{PLS-SVD} M \hat{\Theta}_{PLS-SVD} \right)^{-1}$  in the row space.

In addition, the PLSSVD regression coefficients of the matrix  $\hat{\Theta}_{PLS-SVD} = A_{PLS} B'_{SVD}$ , following the contributions of Gabriel [8], are approximated by the internal product of the row and columns markers using the expression:

$$\hat{\theta}_{ij} = \langle \vec{a}_{PLS_i} \cdot \vec{b}_{SVD_j} \rangle \quad (24)$$

$$= (\vec{a}_{PLS_i})' \vec{b}_{SVD_j} \quad (25)$$

$$= \|\vec{a}_{PLS_i}\| \|\vec{b}_{SVD_j}\| \cos \alpha \quad (26)$$

$$= \text{sign} \|\vec{b}_{SVD_j}\| \|\text{proy}_{\vec{b}_{SVD_j}} \vec{a}_{PLS_i}\| \quad (27)$$

$$= \text{sign} \|\vec{b}_{SVD_j}\| \|(\vec{a}_{PLS_i})_j^*\| \quad (28)$$

where  $\alpha$  is the angle between  $\vec{a}_{PLS_i}$  ( $i$ -th row marker) and  $\vec{b}_{SVD_j}$  ( $j$ -th column marker),  $\text{sign}$  is equal to 1 or -1 depending if the angle is acute or obtuse, respectively. We also have adopted the notation  $(\vec{a}_{PLS_i})_j^*$  to represent the orthogonal projection of the row marker  $\vec{a}_{PLS_i}$  on the column marker  $\vec{b}_{SVD_j}$ . If the rank of  $\hat{\Theta}_{PLS-SVD}$  is  $r = 2$ , the biplot is defined as the graphical representation in a plane of the row markers  $\vec{a}_{PLS_i} = \vec{a}_i = (a_{i1}, a_{i2})'$  and the columns markers  $\vec{b}_{SVD_j} = \vec{b}_j = (b_{j1}, b_{j2})'$ , both vectors having two elements, such that for each  $i$  and each  $j$ :

$$\hat{\theta}_{ij} = (\vec{a}_i)' \vec{b}_j \quad (29)$$

$$\hat{\theta}_{ij} = a_{i1} b_{j1} + a_{i2} b_{j2} \quad (30)$$



Considering the above, to approximate by means of a biplot  $\hat{\theta}_{13}$ , a hypothetical PLSSVD regression matrix of order  $(4 \times 3)$  and rank 2, we should set

$$\hat{\theta}_{13} = a_{11}b_{31} + a_{12}b_{32}$$

where  $\vec{a}_i = (a_{11}, a_{12})'$  and  $\vec{b}_j = (b_{31}, b_{32})'$ .

In Figure 1 it is seen how  $\hat{\theta}_{ij}$  can be interpreted in the plane for the hypothetical PLSSVD regression coefficient matrix of order  $(4 \times 3)$  and rank 2. The figure (1.a), shows how each row marker  $\vec{a}_{PLS_i} = \vec{a}_i$  should be orthogonally projected on the column marker  $\vec{b}_{SVD_j} = \vec{b}_j$ ; while in the Figure (1.b) one can visualize in the plane, the biplot of the PLSSVD regression coefficient matrix, its row and column markers as well as the projections of the row markers on the column marker  $\vec{b}_1$ . On the other hand, the figure (1.c) shows the ordering of the projections  $(\mathbf{a}_i)_j^*$  (in the graph bold letters represent vectors, that is,  $\vec{a} = \mathbf{a}$ ) on the column marker  $\mathbf{b}_1$ , so that:  $sign\|(\mathbf{a}_1)_1^*\| \geq sign\|(\mathbf{a}_3)_1^*\| \geq sign\|(\mathbf{a}_2)_1^*\| \geq sign\|(\mathbf{a}_4)_1^*\|$ . From this information and taking into account that  $\|\vec{b}_1\|$  is the same for the approximation of each element of the first column of the  $(4 \times 3)$  regression coefficient matrix, it is possible to approximate the ordering of the regression coefficients of the first columns of the matrix as:  $\hat{\theta}_{11} > \hat{\theta}_{31} > \hat{\theta}_{21} > \hat{\theta}_{41}$ . For more detail we refer the reader to [4](p. 145).

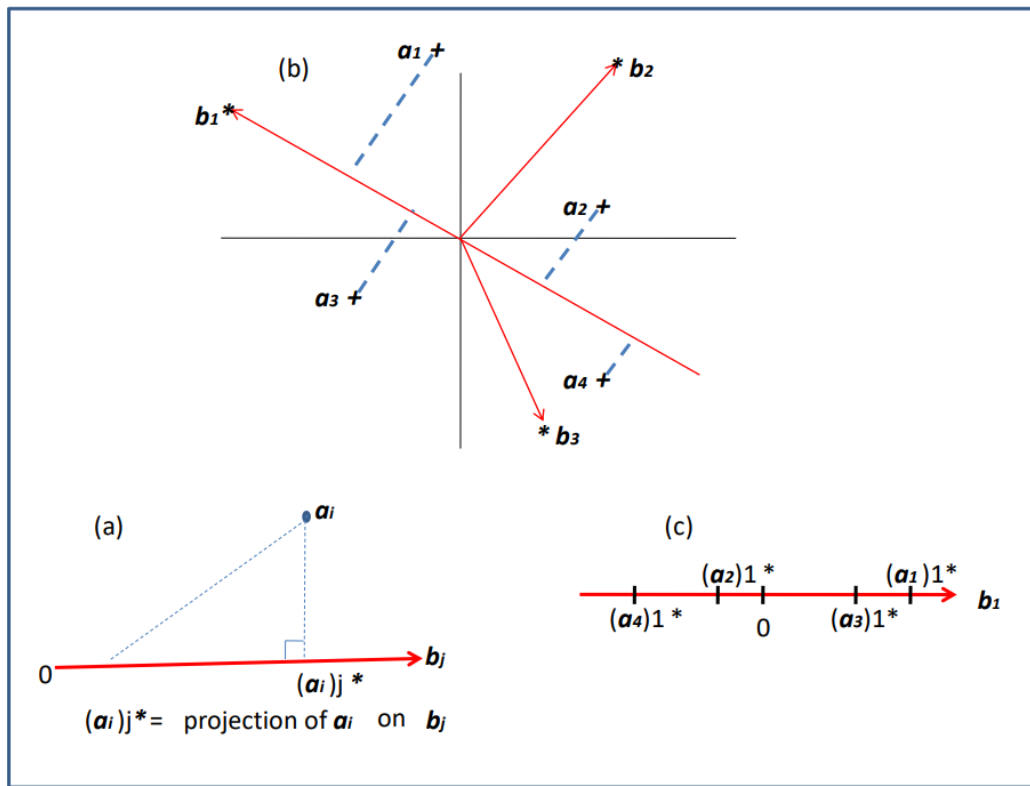


Figure 1. (a) projection of a row marker  $\vec{a}_i$  on the column marker  $\vec{b}_j$  denoted as  $(\vec{a}_i)_j^*$ , (b) biplot of a hypothetical PLSSVD regression of order  $(4 \times 3)$ , (c) ordering of the projections of the  $\vec{a}_i$  on  $\vec{b}_1$ .

#### 4. The Algorithm

In the appendix we have included the pls-svd algorithm which is based on the ideas of [17], [11], [21] and [1]. The algorithm begins by assigning the two sets under study: the set of dependent variables  $Y$ , and the set of independent variables  $X$ . One of the requirements of the algorithm is that the columns of  $X$  and  $Y$  be standardized with mean zero and variance 1. To perform this, the function *standardize()* was implemented in R. The standardizations of  $Y$  and  $X$  are called  $Fo$  and  $EO$  respectively. In [1], we established that if  $\Gamma = I$ , then  $B_{SVD} = VD$ . The initial estimation of  $AX = Q$  uses  $Q$  obtained from the decomposition in singular values of  $Fo$  using  $Q = L\$u$  decomposition. It is important to remember that the method described in this article is useful when the set of independent variables  $X$  has "Severe multicollinearity". In this study the diagnostic of collinearity is performed by the eigensystem  $X'X$  proposed by [14] using the function *CollinDiag(X)* to calculate the condition number of  $X'X$ . When this number is less than 100, there does not exist a serious problem of multicollinearity and the routine outputs the message "there isn't multicollinearity". If, on the other hand, the condition number is between 100 and 1000, the multicollinearity is moderate to strong so the routine outputs the message "Moderate to strong multicollinearity". Lastly, if the condition number is greater than 1000, the multicollinearity is strong and to avoid confusion, the routine outputs the message "Severe multicollinearity".

The implementation used here is slightly different from the algorithm presented in [1] regarding the calculation of the vectors  $w_h$  and  $r_h$ . Using the suggestions of [21] we here begin with the autoequations:

$$Eo'_{h-1}Q_{h-1}Q'_{h-1}Eo_{h-1}w_h = \lambda_h w_h$$

$$Q'_{h-1}Eo_{h-1}Eo'_{h-1}Q_{h-1}r_h = \lambda_h r_h$$

After executing the algorithm pls-svd, we calculate the estimation of the PLSSVD regression coefficient matrix ( $\hat{\Theta}_{PLS-SVD}$ ) as  $\hat{\Theta}_{PLS-SVD} = Apls \% * \%t(B)$  and with the obtained elements we proceed to declare the labels of the biplot using the instructions

```
biplot(Apls,B, var.axes = TRUE, xlab = "Comp1",ylab = "Comp2",
xlabs =labell, ylabs = label2,
main = "BILOT PLS-SVD");abline(h=0,v=0)
```

#### 5. Illustration using Food composition database for biodiversity

The FAO is an agency of the United Nations which designs strategies to fight hunger. The principal objective of this agency is to assure food supply for all peoples so that access to sufficient good quality food is guaranteed on a regular basis in order to lead an active and healthy life. This organization has more than 194 members in more than 130 countries. Among its different functions is that of providing data to the nations to help them achieve their own specific objectives.

In this section we wish to apply the previously described biplot to FAO data. Specifically we use the Food composition database for biodiversity, BioFoodComp 4.0, which can be accessed by the link <http://www.fao.org/infoods/infoods/tables-and-databases/faoinfoodsdbases>. The FAO states in its own words, "The Biodiversity Database is a global repository of analytical data on food biodiversity of acceptable data quality". For our analysis we have selected fruit from the Biodiversity Database BioFoodComp 4.0 with the objective to construct a graphical representation biplot to explain the content of Vitamins and Carotenoids (dependent variables  $Y$ ) of two fruit groups, banana and papaya, by means of mineral content and trace elements (independent variables  $X$ ). For that purpose we included rows 16 through 36 of that original Excel database. The resulting dependent variables to be explained are shown in Table 1; while the explaining (independent) variables are shown in Table 2.

Table 1. Dependent Variable  $Y$ : Vitamins/Carotenoids

Component ID/Tagname	Component name	Unit
VITA.RAE	vitamin A retinol activity equivalent (RAE); calculated by summation of the vitamin A activities of retinol and the active carotenoids	mcg
VITC	vitamin C	mg
CARTB	beta-carotene	mcg
LUTN	lutein	mcg

Table 2. Independent Variable  $X$ : Minerals and trace elements

Component ID/Tagname	Component name	Unit
WATER	water	g
B	boron	mcg
CA	calcium	mg
CU	copper	mg
FE	iron, total	mg
K	potassium	mg
MN	manganese	mg
MG	magnesium	mg
NA	sodium	mg
ZN	zink	mg

In Table 3 is the correlation matrix for the data. We can there appreciate the correlations between the dependent variables ( $R_{yy}$ ), those of the independent variables ( $R_{xx}$ ), as well as the correlations between the dependent and independent variables ( $R_{xy}$ ). The dependent variables made up of Vitamins and Carotenoids, show high positive correlations between vitamin A retinol activity equivalent (VITA.RAE) and the beta-carotene (CARTB). These have moderate correlation with lutein (LUTN) and low correlation with vitamin C (VITC). The lutein (LUTN) and vitamin C (VITC) present low correlation between themselves. On the other hand, in the set of independent variables (Minerals and trace elements), the amount of water (WATER) has high correlations with almost all of the other independent variables except boron (B), Water (WATER) shows high positive correlation with calcium (CA) and high negative correlations with the rest. Boron (B) has low to moderate correlation with the rest of the independent variables. We also can see that the other independent variables have a high correlation with at least one of the remaining variables which is indicative of the existence of multicollinearity in this set of variables. Regarding the correlations between Minerals and trace elements with Vitamins and Carotenoids, we see that vitamin A retinol activity equivalent (VITA.RAE) and the beta-carotene (CARTB) are the variables that present highest correlation; the lutein (LUTN) has almost all moderate correlations and vitamin C (VITC) presents highest correlation with zink (ZN) and low to moderate relations with the rest.

To explain the relation between the variables we use the model described in (1). We begin with the diagnosis of multicollinearity using the function `ColinDiag(X)` described previously and provided in the appendix. The independent variables  $X$  present Severe multicollinearity, which means that the classic RRR biplot ([18]) cannot be applied to this data.

```
> ColinDiag(X)
[1] "Severe multicollinearity"
```

In Figure 2 is shown the biplot for the regression coefficient matrix of Minerals and trace elements vs Vitamins/Carotenoids. Applying Property 5, we see that the first component comprises 68,70% of the total

Table 3. Correlations of the data

	WATER	B	CA	CU	FE	K	MG	MN	NA.	P	ZN	VITA.RAE	CARTB	LUTN	VITC
WATER	1.00	-0.15	0.67	-0.61	-0.69	-0.92	-0.87	-0.72	-0.52	-0.98	-0.61	0.69	0.58	0.27	0.20
B	-0.15	1.00	0.21	0.48	0.32	0.23	0.40	0.25	0.29	0.23	0.15	0.17	0.20	0.01	0.05
CA	0.67	0.21	1.00	-0.45	-0.37	-0.53	-0.45	-0.45	-0.29	-0.68	-0.43	0.89	0.84	0.27	0.28
CU	-0.61	0.48	-0.45	1.00	0.65	0.55	0.53	0.52	0.40	0.69	0.49	-0.47	-0.41	-0.24	-0.12
FE	-0.69	0.32	-0.37	0.65	1.00	0.79	0.66	0.57	0.72	0.71	0.56	-0.49	-0.43	-0.46	-0.02
K	-0.92	0.23	-0.53	0.55	0.79	1.00	0.86	0.61	0.53	0.89	0.46	-0.61	-0.52	-0.49	-0.29
MG	-0.87	0.40	-0.45	0.53	0.66	0.86	1.00	0.73	0.58	0.84	0.56	-0.44	-0.31	-0.18	-0.03
MN	-0.72	0.25	-0.45	0.52	0.57	0.61	0.73	1.00	0.40	0.77	0.62	-0.51	-0.44	-0.17	0.12
NA	-0.52	0.29	-0.29	0.40	0.72	0.53	0.58	0.40	1.00	0.50	0.68	-0.27	-0.21	-0.17	0.30
P	-0.98	0.23	-0.68	0.69	0.71	0.89	0.84	0.77	0.50	1.00	0.59	-0.71	-0.61	-0.33	-0.24
ZN	-0.61	0.15	-0.43	0.49	0.56	0.46	0.56	0.62	0.68	0.59	1.00	-0.46	-0.42	-0.01	0.59
VITA.RAE	0.69	0.17	0.89	-0.47	-0.49	-0.61	-0.44	-0.51	-0.27	-0.71	-0.46	1.00	0.99	0.49	0.27
CARTB	0.58	0.20	0.84	-0.41	-0.43	-0.52	-0.31	-0.44	-0.21	-0.61	-0.42	0.99	1.00	0.50	0.24
LUTN	0.27	0.01	0.27	-0.24	-0.46	-0.49	-0.18	-0.17	-0.17	-0.33	-0.01	0.49	0.50	1.00	0.36
VITC	0.20	0.05	0.28	-0.12	-0.02	-0.29	-0.03	0.12	0.30	-0.24	0.59	0.27	0.24	0.36	1.00

variability and the second component 24.81%, which means that the first factorial plane has the capacity to represent 93.51% of the total. Remember that Property 4 shows that the correlation between two dependent variables is approximated by the cosine of the angle that exists between the respective column markers of the variables so we see that the column markers of the variables VITA.RAE and CARTB are almost coincident so that the relation between the two variables is close to unity. Verifying in Table 3 we see that the correlation between these two variables is 0.99. Additionally, in Figure 2, we can see that the markers of these two variables are closer to the variable LUTN than to VITC, which shows that the relations of these two variables with LUTN are greater than with VITC. We can also observe that the markers for LUTN and VITC manifest moderate relations between their respective variables.

The most important objective of this biplot is to provide a graphical representation of  $\hat{\Theta}_{PLS-SVD}$ , the regression coefficient matrix which can be seen in Table 4. To highlight the usefulness of this graphical representation we have constructed Figure 3 which has a slight change from Figure 2: the biplot axis corresponding to the variable VITC has been lengthened and painted in blue.

As a second step, each of the row markers corresponding to the Minerals and trace elements have been projected orthogonally on this column marker. Note that on this blue axis an order to the projections has been established. The order is

$$sign \|(\vec{a}_{PLSZN})_{VITC}*\| \geq sign \|(\vec{a}_{PLSMN})_{VITC}*\| \geq sign \|(\vec{a}_{PLSCA})_{VITC}*\| \geq sign \|(\vec{a}_{PLSMG})_{VITC}*\| \geq sign \|(\vec{a}_{PLSWATER})_{VITC}*\| \geq sign \|(\vec{a}_{PLSB})_{VITC}*\| \geq sign \|(\vec{a}_{PLSCU})_{VITC}*\| \geq sign \|(\vec{a}_{PLSP})_{VITC}*\| \geq sign \|(\vec{a}_{PLSFE})_{VITC}*\| \geq sign \|(\vec{a}_{PLSK})_{VITC}*\|.$$

Table 4. PLSSVD REGRESSION COEFFICIENT MATRIX

	VITA.RAE	CARTB	LUTN	VITC
WATER	0.10	0.06	-0.05	0.14
B	0.26	0.28	0.01	-0.01
CA	0.49	0.50	0.18	0.25
CU	-0.14	-0.14	-0.19	-0.16
FE	-0.16	-0.21	-0.46	-0.16
K	-0.14	-0.13	-0.30	-0.39
MG	0.20	0.28	0.34	0.09
MN	-0.04	-0.02	0.13	0.23
NA	0.09	0.08	0.03	0.33
P	-0.17	-0.14	-0.10	-0.24
ZN	-0.05	-0.07	0.29	0.83

This order is very close, as can be noted, to the order found in Table 4, which shows that we can give an approximation of the multiple regression model of the variable VITC with respect to the predictor variables. Lastly, if we perform a similar procedure for the other column markers we will note that the biplot provides a good approximation to the other multiple regression models.

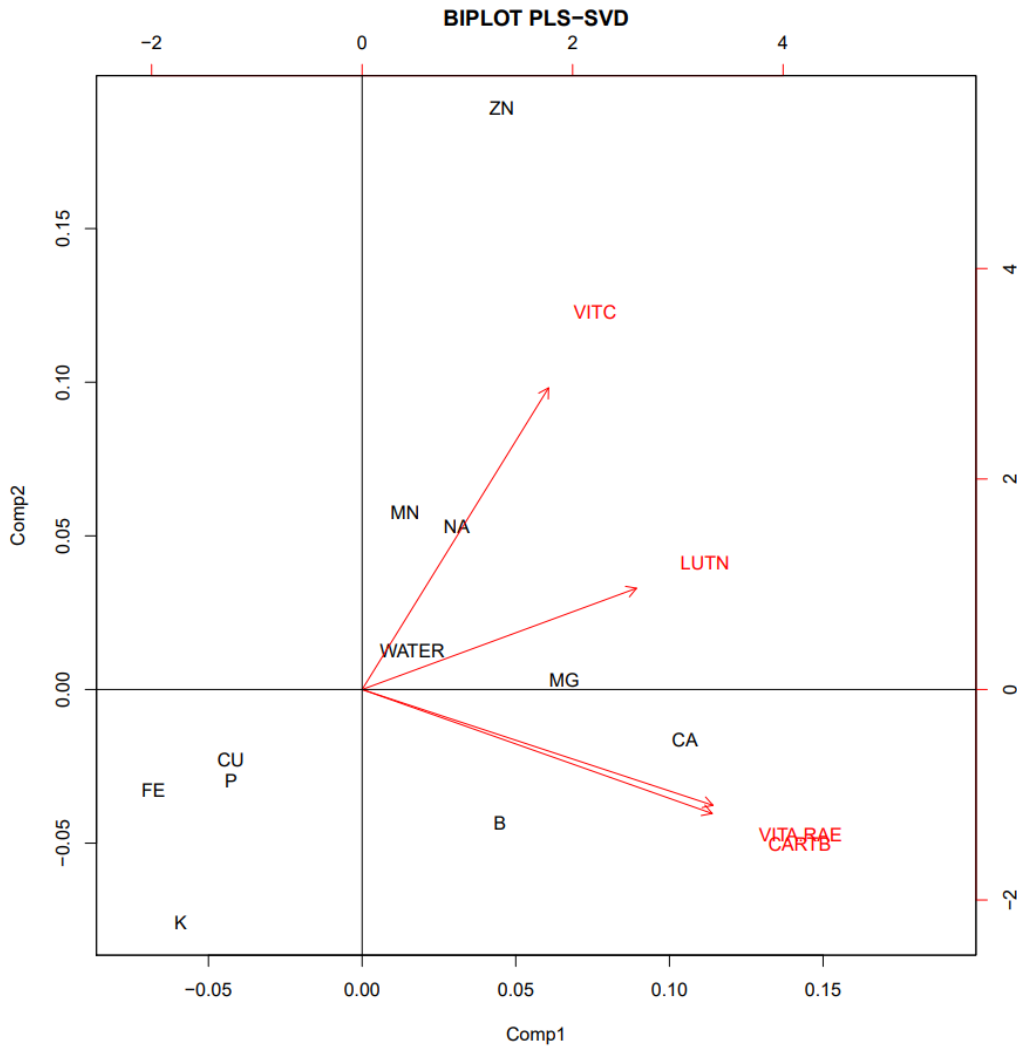


Figure 2. PLSSVD Biplot Minerals and trace elements vs Vitamins/Carotenoids

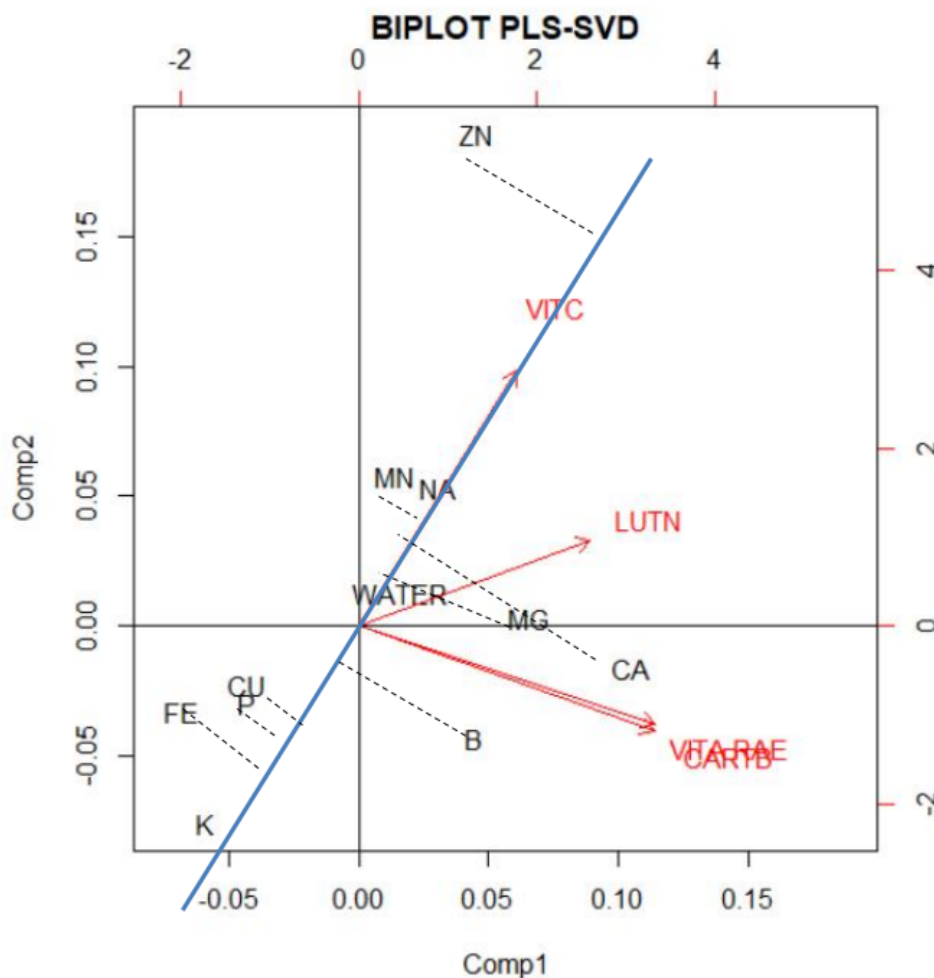


Figure 3. PLSSVD Biplot Minerals and trace elements vs Vitamins/Carotenoids modified by extending the biplot axis corresponding to VITC in order to visualize the orthogonal projections of Minerals and trace elements

## 6. Conclusions

We have shown how to prepare a biplot of reduced rank regression when the model shows multicollinearity. In the example it is evident that this occurs and that the use of partial least squares provides a good alternative when we desire a representation of the coefficient matrix. We have seen that generalized singular value decomposition and the use of a weight matrix both play a fundamental role in achieving this goal. Thanks to the geometry of PLS, we were able to show that the row markers in the PLSSVD GH Biplot are an oblique projection of the row markers of the RRR biplot.

A summary of the properties derived using this methodology and described in this article is as follows.

1. In the PLSSVD GH Biplot under the metric  $M = X'X$  we have shown that  $A'_{PLS}MA_{PLS} = I_r$

2. The scalar product of two columns of  $\hat{\Theta}_{PLS-SVD}$  under the metric  $M$  approximates the scalar product of the corresponding column markers.
3. The scalar product of two columns of  $\hat{Y}_{PLS-SVD}$  approximates the scalar product of the corresponding columns of  $\hat{\Theta}_{PLS-SVD}$  under the metric  $M$ .
4. When the variables are centered, the squared length of the column markers approximate the sample covariance between the dependent variables.
5. The cosine of the angle between two column markers approximates the correlation between the dependent variables centered on their mean.
6. The quality of the representation of the columns (QRC), under the metric  $M = X'X$  and weight matrix  $\Gamma = I$  can be measured by

$$QRC = \frac{\sum_{i=1}^r \lambda_i^2}{\sum_{i=1}^q \lambda_i^2} * 100$$

7. The scalar product of two rows of  $\hat{\Theta}_{PLS-SVD}$  under the metric  $M = X'X$  and weight matrix  $\Gamma = I_r$  approximates the scalar product of the corresponding row markers.

All these properties are subject to the given weight matrix  $\Gamma = I$  and even for these weights we are sure that there remain many more properties to state and prove. It is important to say that research remains open for other weight matrices such as, for example,  $\Gamma = R_{yy}^{-1/2}$ . This study has been based only on the GH biplot (Column Metric Preserving biplot) so that research remains to be done for the other classical biplots including the JK biplot and the SQRT Biplot. Another important aspect to be emphasized is that this research was done in response to the question in [18] and repeated in [1]... Why can one expect PLS regression methods to perform better than multiple linear regression, ridge regression and other well known regression techniques? In spite of this, we believe that there exist further opportunities for research in this area. For example, investigate the use of the ridge regression technique instead of PLS, comparing the biplots obtained through the different techniques. Many authors coincide that in order to improve the approximations of the biplots, scaling should be done. Scaling was not performed in this article so we invite other researchers to dive deeper in this and other biplot themes.

## 7. APPENDIX

```
#-----
#ALGORITHM PLS-SVD
#-----
rm(list=ls())
#-----
#PRESENTATION OF DATA
#-----
Y # DEFINE THE DEPENDENT VARIABLES
X # DEFINE THE DEPENDENT VARIABLES
#-----

#-----
# INICIAL EXPLORATION
#-----
#1) FUNTION COLLINEARITY DIAGNOSIS
#-----
```

```

CollinDiag=function(x)
{R=cor(x)
r1=length(svd(R, 0, 0)$d)# matrix rank
p1=ncol(R)
UDV=svd(R)
d1=diag(UDV$d)[1:r1,1:r1]
lamda=diag(d1)
lamda1=lamda[1:1]
lamdap=lamda[p1:p1]
coefi=lamda1/lamdap
if (coefi <= 100) {
print("there isn't multicollinearity")
} else if (coefi > 100 & coefi <=1000) {
print("Moderate to strong multicollinearity")
} else
print("Severe multicollinearity")
return()
}

```

```

#-----
#COLLINEARITY DIAGNOSIS
#-----
CollinDiag(X)
#-----

```

One of the requirements of this algorithm is that the columns of  $X$  and  $Y$  be standardized with mean zero and variance 1. The following function ensures this.

```

#-----
standardize=function(x)
{x=data.matrix(x)
x=x-t(matrix(rep(apply(x,2,mean),times=nrow(x)),nrow=ncol(x)))
x=x*%solve(diag(sqrt(apply(x,2,var))))
return(x)}
#-----

```

The algorithm pls is now presented.

```

#-----
#ALGORITHM PLS-SVD
#-----
#PROGRAM INITIALILIZATION
#-----
Eo=standardize(X)
Fo=standardize(Y)
Eoi=Eo
Foi=Fo
L<-svd(Fo)
Q<-L$u# Q=AX
B=L$v*%diag(L$d)
#-----
#Output declarations

```



```

#-----
R<-matrix(NA, ncol(Q), ncol(Q)); W<-matrix(NA, ncol(Eo), ncol(Q));
P<-matrix(NA, ncol(Eo), ncol(Q)); T<-matrix(NA, nrow(Q), ncol(Q));
U<-matrix(NA, nrow(Q), ncol(Q)); C<-matrix(NA, ncol(Q), ncol(Q))
for (i in 1:ncol(Q)){
Da=t(Eo)%*%Q%*%t(Q)%*%Eo
Db=t(Q)%*%Eo%*%t(Eo)%*%Q
w=eigen(Da)$vectors[,1]
r=eigen(Db)$vectors[,1]
t=(Eo%*%w)
u=(Q%*%r)
p=t(Eo)%*%t*as.numeric(1/(t(t)%*%t))
c=t(Q)%*%t*as.numeric(1/(t(t)%*%t))
Eo<-Eo-t%*%t(p)
Q<-Q-t%*%t(c)
W[, i] <- w
C[, i] <- c
P[, i] <- p
T[, i] <- t
U[, i] <- u
R[, i] <- r
}
#-----

```

#### Calculation of the estimation of the coefficient matrix

```

#-----
#CALCULATION OF THE REGRESSION COEFFICIENT MATRIX
#-----
We=W%*%solve(t(P)%*%W)
Apls=We%*%t(C)
THETApls=We%*%t(B)
#-----

```

#### biplot in Reduced-Rank Regression based on PLS

```

#-----
##BIPLOT PLS-SVD
#-----
label1=c()# DECLARE THE LABELS OF THE DEPENDENT VARIABLES
label2=c()# DECLARE THE LABELS OF THE INDEPENDENT VARIABLES
biplot(Apls,B, var.axes = TRUE, xlab = "Comp1",ylab = "Comp2",
xlabs =label1, ylabs = label2,
main = "BIPLOT PLS-SVD");abline(h=0,v=0)
#-----

```

#### REFERENCES

1. Álvarez W. and Griffin V. (2016). Estimation procedure for reduced rank regression, PLSSVD. *Stat., Optim. Inf. Comput.*, Vol. 4, pp 107-117.
2. Anderson T., (1999). Asymptotic Distribution of the Reduced Rank Regression estimator under general conditions. *The Annals of Statistics*, 27, 1141-1154.

3. Anderson T., (1951). Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions. *Ann. Math. Statist*, 22, 327-351
4. Cárdenas, O. and Galindo, P. (2004). Biplots con información externa basados en modelos bilineales generalizados. Universidad Central de Venezuela-Consejo de Desarrollo Científico y Humanístico.
5. Cárdenas, O., Galindo, P., Vicente-Villardón J.L. (2007). Los métodos Biplot: Evolución y aplicaciones. *Revista Venezolana de Análisis de Coyuntura*. Vol. XIII, num. 1, enero-julio, pp. 279 - 303. Caracas, Venezuela.
6. Davies, P. and Tso M., (1982). Procedures for reduced rank regression. *Applied Statistic* 31,244-255
7. Eckart, C. and Young, G., (1936). The Approximation de One Matrix for Another of Lower Rank. *Psychometrika* 1, 211-218.
8. Gabriel, K., (1971). The Biplot-graphic display of matrices with applications to principal component analysis. *Biometrika*, 58, 453-467.
9. Gabriel (1981), Biplot display of multivariate matrices for inspection of data and diagnosis, Barnett V. (ed.), *Interpreting Multivariate Data*, 147-173, Wiley, London.
10. Gabriel, K. R., Odoroff, C. L. (1990), Biplots in biomedical research, *Statistics in Medicine* 9 (5): 469-485.
11. Höskuldsson, A., (1988). PLS regression methods. *Journal of Chemometrics* 2, 211- 228.
12. Izenman, A. J., (1975). Reduced rank regression for the multivariate linear model. *Journal of Multivariate Analysis* 5, 248-264
13. Lauro, N. C. (2004) .Non Symmetrical Correspondence Analysis and Related Methods. Department of Mathematics and Statistics. University of Naples Federico II. Italy
14. Montgomery D., Peck E. and Vining G. (2001). *Introducción al análisis de regresión*. John Wiley and Sons
15. Phatak A. and De Jong S. (1997): The geometry of partial least squares. *Journal of Chemometrics*, Vol. 11:311-338
16. Strauss, J., Gabriel, K. R. (1979), Do psychiatric patients fit their diagnosis, pattern of symptomatology as described with the biplot, *Journal of Nervous and Mental Disease*. 167: 105-113.
17. Tenenhaus, M. (1998). *La régression pls, théorie et pratique*. Éditions Technip. Paris.
18. Ter Braak, C. and Looman, C. (1994). Biplots in reduced rank regression. *Biometrical Journal* 8, 983-1003.
19. Ter Braak, C. (1990). Interpreting canonical correlation analysis through biplots of structural correlations and weights. *Psychometrika*, 55, 519-531
20. Tsianco, M., Gabriel, K. R. (1984), Modeling temperature data: An illustration of the use of biplot and bimodels in non-linear modeling, *Journal of Climate and Applied Meteorology* 23: 787-799.
21. Valencia, J., Díaz-LLanos F. and Calleja S. (2004). *Métodos de predicción en situaciones límite*. Editorial La Muralla, S.A.