

# Dimensionality Reduction Methods

Luigi D'Ambra,<sup>1</sup> Pietro Amenta,<sup>2</sup> and Michele Gallo<sup>3</sup>

## Abstract

In case one or more sets of variables are available, the use of dimensional reduction methods could be necessary. In this contest, after a review on the link between the Shrinkage Regression Methods and Dimensional Reduction Methods, authors provide a different multivariate extension of the Garthwaite's PLS approach (1994) where a simple linear regression coefficients framework could be given for several dimensional reduction methods.

## 1 Introduction

When the number of variables is very large, as well as, in presence of more than one sets of them playing a logical asymmetrical role (explanatory and response variables), it may be advantageous to find for each set a linear combination of variables (latent variables) having some properties in terms of correlation, covariance or variance. The criteria for an appropriate new basis depends, of course, on the application. One way of approaching this problem is to project the data on the maximum data variation subspace, i.e. the subspace spanned by the largest principal components (Principal Component Analysis – PCA). Nevertheless, the study of multivariate predictions could be, also, faced with several approaches, for example, Constrained Principal Component Analysis (CPCA) (D'Ambra and Lauro, 1982). In customer satisfaction evaluation where the relationships between expectations and perceptions are taken in account, an analysis could be developed by looking for the subspace, maximizing the covariance between the projected scores of both sets. This subspace provides the largest singular values of the covariance matrix between expectation and perception data (D'Ambra et al., 1999). Finally, when the goal is to predict a dependent variable as well as possible in terms of least square error, an appropriate model is Reduced Rank Regression (RRR). In general, when the goal is to predict more dependent variables by substituting the set of observed explanatory variables with a fewer sequence of orthogonal latent variables, Dimensional Reduction Methods (DRM) should be applied. The commonly used DRM methods are Principal Component Regression (PCR), Canonical Correlation Regression (CCR), RRR and Partial Least Squares (PLS; Wold, 1966). These methods, together with the shrinkage ones, play an important role in order to overcome the collinearity problem.

---

<sup>1</sup> Department of Mathematics and Statistics, University of Naples "Federico II", Italy; dambra@unina.it

<sup>2</sup> Department of Analysis of Economic and Social Systems, University of Sannio, Italy; amenta@unisannio.it

<sup>3</sup> Department of Social Science, University of Naples "L'Orientale", Italy; mgallo@iuo.it

The paper is organized into 5 sections. In Section 2 the basic notation is given. Section 3 briefly presents the linkage between the shrinkage regression methods and the dimensional reduction methods. In this Section we also propose an extension of the Principal Covariates Regression (de Jong and Kiers, 1992) in order to find a continuum among the DRM method. Main focus of this paper is in Section 4. Following the Garthwaites' PLS approach (1994), we show how a simple linear regression coefficients framework could be given for the considered DRM methods. Last Section includes some conclusive remarks on the methodology proposed, as well as topics for further research.

## 2 Notation

Let  $Q_1, \dots, Q_k, \dots, Q_K$  be  $K$  response variable groups observed on  $N$  statistical units and collected in a matrix  $Y^* = [Y^1 | \dots | Y^k | \dots | Y^K]$ , of order  $(N, \sum_{k=1}^K Q_k)$ , where  $Y_{(NxQ_1)}^1, \dots, Y_{(NxQ_k)}^k, \dots, Y_{(NxQ_K)}^K$  are  $K$  different matrices. The  $k$ -th matrix with generic element  $y_{iq}^k$  ( $i = 1, \dots, N$ ;  $q = 1, \dots, Q_k$ ) denotes the value of the  $q$ -th criteria variable observed on the  $i$ -th statistical unit for the  $k$ -th response variable groups. Moreover, let  $X_{(N \times J)}$  be a matrix of independent variables with  $rank(X) = S < \min(N, J)$ . The generic element  $x_{ij}$  ( $i = 1, \dots, N$ ;  $j = 1, \dots, J$ ) is the value of the  $j$ -th independent variable observed on the same  $i$ -th statistical unit. In this paper we assume that all variables have zero mean as regards the weight diagonal metric  $D$  whose general term is  $1/N$ . Let  $P_X = X(X^T X)^{-1} X^T$  orthogonal projector onto the subspace spanned by the columns of  $X$  with  $X^T$  the transpose of matrix  $X$ . Finally, let  $T_{(S)}$  be an orthogonal matrix of order  $(N \times S)$  containing  $S$  latent variables so as to obtain the fitted response matrix by  $\hat{Y}_{(S)} = T_{(S)}(T_{(S)}^T T_{(S)})^{-1} T_{(S)}^T Y = X B_{(S)}$  with  $L_S = X^T T_{(S-1)}(T_{(S-1)}^T X X^T T_{(S-1)})^{-1} T_{(S-1)}^T X$ . Let denote  $\tilde{X}$  the standardized  $X$  matrix.

## 3 Shrinkage regression and dimensional reduction methods for multivariate analysis

In literature many shrinkage regression methods have been proposed. PCR, PLS, RRR and Continuum Regression (CR) are only some among the most famous ones (Stone and Brooks, 1990; Frank and Friedman, 1993; Brown, 1993; Brooks and Stone, 1994). These methods should be used when a large singular value is associated to two or more independent variables with "large" variance decomposition portions. These variables may determine collinearity problems with unrealistic and shaky ordinary least square coefficients  $b^{OLS} = (X^T X)^{-1} X^T y_q^k$  ( $k = 1, \dots, K$ ,  $q = 1, \dots, Q_k$ ).

An approach to solve the collinearity problem consists in replacing the factor  $(X^T X)^{-1}$  in expression of  $b^{OLS}$  with a better-conditioned matrix  $G$ . In the PCR, the matrix  $G$  is given from a spectral decomposition of  $X^T X$ :  $X^T X = \sum_{j=1}^s \lambda_j v_j v_j^T$  where  $S < \min(N, J)$  is the rank of  $X$ . Differently, PLS looks for a vector  $c$  ( $\|c\| = 1$ ) such that the scalar product  $y^T X c$  is maximal and  $b \propto c$ . This leads to consider the predictor  $b^{PLS} \propto X^T y_q^k$  replacing  $(X^T X)^{-1}$  with a better-conditioned matrix  $G \propto I_p$ . Finally, Hoerl (1962) and Hoerl and Kennard (1970) recommend the use of the ridge regression

with  $b^{RR} = (X^T X + \delta I_p)^{-1} X^T y_q^k$  and  $\delta \geq 0$ . In Table 1 all the conditioned matrices for the different techniques are given.

**Table 1:** Several conditioned matrices  $G$ .

General solution $b = GX^T y_q^k$				
OLS	PCR	PLS	RRR	$\hat{Y}$
Conditioned matrix $G$				
$G = (X^T X)^{-1}$	$G = \sum_j \lambda_j^{-1} v_j v_j^T$	$G \propto I_p$	$G \propto (X^T X + \delta I_p)^{-1}$	$G = A$
Predictor $b$				
$(X^T X)^{-1} X^T y_q^k$	$(\sum_j \lambda_j^{-1} v_j v_j^T) X^T y_q^k$	$\propto X^T y_q^k$	$(X^T X + \delta I_p)^{-1} X^T y_q^k$	$A X^T y_q^k$

When there is only one dependent variable ( $y_q^k$  for each  $k = q = 1$ ) the OLS, PLS and PCR could be considered like a particular case of the CR (Stone and Brooke, 1990). The coefficient  $b$  is determined by simple regression of  $y$  on a one dimensional  $Xc$ , where the coefficient vector  $c$  is chosen by maximising different criteria: the squared correlation coefficient  $r^2(y, Xc)$ , the covariance  $Cov(y, Xc)$  and the variance  $Var(Xc)$ , respectively.

Stone and Brooke (1990) suggest a general principle to determine the coefficient vector  $c$ , for a fixed continuum solution parameter  $\gamma \geq 0$ . The coefficient  $c$  is obtained by the maximization of  $T(\gamma, c) = (y^T Xc)^2 |Xc|^{2(\gamma-1)} \propto r^2(y, Xc) |Xc|^2$  subject to the constrain  $\|c\| = 1$ . Where for  $\gamma = 0$ ,  $\gamma = 1$  and  $\gamma \rightarrow \infty$  we have the continuum solution among OLS, PLS and PCR, respectively. Many of these shrinkage regression methods can be seen in a more general multivariate framework based on a common objective function for the DRMs (Abraham and Merola, 2001). All the DMRs objective functions are measures of association between couples of unit norm latent variables, which are linear combinations of the dependent variables ( $u_j = Y^k d_j$ ) and of the independent ones ( $t_j = X a_j$ ).

These measures are expressed in term of squared covariance between the latent variables  $t_j$  and  $u_j$  as well as their variance, respectively (Table 2).

When  $X^T X$  is almost singular, it is possible to highlight that the ‘‘PCR smooth’’ criteria of this matrix can be used in other approaches obtaining mixed DRMs. In same time the ‘‘PCR smooth’’ criteria can be obtained by mixed DRMs approaches (i.e. in CPCA we can obtain as solution matrix  $Y^{kT} X (\sum_j \lambda_j^{-1} v_j v_j^T) X^T Y^k$  which is equivalent to the PCR one).

**Table 2:** Objective functions of the DRMs.

Method	Object function	Solution matrix
PCA	$\max(a_j^T X^T X a_j)$	$X^T X$
CCR	$\max[(a_j^T X^T Y^k d_j)^2 / (\ t_j\ ^2 \ u_j\ ^2)]$	$(X^T X)^{-1} X^T Y^k (Y^{kT} Y^k)^{-1} Y^{kT} X$
RRR	$\max[(a_j^T X^T Y^k d_j)^2 / \ t_j\ ^2]$	$(X^T X)^{-1} X^T Y^k Y^{kT} X$
CPCA*	$\max(d_j^T Y^T P_X Y d_j)$	$Y^{kT} X (X^T X)^{-1} X^T Y^k$
SIMPLS	$\max(a_j^T X^T Y^k d_j)^2$	$(I - L_j)^{-1} X^T Y^k Y^{kT} X$
*with the constraints $a_j^T a_j = d_j^T d_j = 1$ , $a_j^T X^T X a_i = 0$ , $j > i$ .		

### 3.1 A different approach to Principal Covariates Regression

In literature there is a trade-off between the RRR and the PCR aims: the former tries to maximize the variance of the criterion variables retained by the predictors latent subspace while the latter tries to maximize only the variance of the predictors with PLS considered as a compromise. A similar continuum can be obtained with an extension of the Principal Covariates Regression (PCovR) or “Weighted maximum overall redundancy” (de Jong and Kiers, 1992; Abraham and Merola, 2001). In order to find a low-dimensional subspace of the predictor space spanned by the columns of  $X$  accounting for the maximum variation of  $X$  and  $Y^k$ , we propose to consider the model

$$\begin{cases} T = XW \\ X = TZ_X + E_X \\ Y^k = TZ_{Y^k} + E_{Y^k} \end{cases} \quad (3.1)$$

where  $T$  contains scores on  $S$  components,  $W$  is the  $J \times S$  matrix of component weights with  $Z_X$  and  $Z_{Y^k}$  loading matrices, of order  $(S \times J)$  and  $(S \times Q_k)$ , containing the regression parameters that relate the predictors and the response variables to the components in  $T$ , respectively. Following de Jong and Kiers (1992), we propose to maximise the following least-squares loss function

$$\alpha \|X - TZ_X\|^2 + \mu \|X^T Y^k - Z_X^T Z_{Y^k}\|^2 + (1 - \alpha - \mu) \|Y^k - TZ_{Y^k}\|^2 \quad (3.2)$$

with  $T^T T = I$  and  $T^T E_X = T^T E_{Y^k} = 0$ . The least-squares solutions are given by the first  $S$  eigenvectors of matrix  $[\alpha X X^T + (1 - \alpha - \mu) \hat{Y}^k \hat{Y}^{kT} + \mu X X^T Y^k Y^{kT}]$  if  $X$  spans the complete space and  $T$  contains scores on all components with  $\hat{Y}^k = X(X^T X)^{-1} X^T Y^k$ .  $W$  may be computed by regression of  $T$  on  $X$ , if  $X^T X$  has not full rank, otherwise, with  $W = X^- T$  where  $X^-$  is any generalized inverse of  $X$ . We introduce two parameters ( $\alpha$  and  $\mu$ ), both varying between 0 and 1, so that  $\mu$  tells how much the model is PLS like and  $(1 - \alpha - \mu)$  determines its Multiple Linear Regression (MLR) nature. We highlight some special cases:

- for  $\alpha = 0$  and  $\mu = 0$  if  $S = \min[\text{rank}(X), \text{rank}(Y^k)]$  than the solution leads to MLR, with an emphasis on fitting  $Y^k$ , otherwise to RRR if

$$S < \min[\text{rank}(X), \text{rank}(Y^k)]$$

- for  $\alpha = 1$  and  $\mu = 0$  the solution puts an emphasis on reconstructing  $X$  with a PCA of  $X$  or with PCR if we use the principal components as predictors for  $Y^k$ ;
- for  $\alpha = 0$  and  $\mu = 1$  the solution leads to Partial Least Squares of  $X$  and  $Y^k$ ;
- finally, for  $\mu = 0$  and for any admissible value for  $\alpha$ , we have the original PCovR solution. In case of  $\alpha = 1/2$ , the authors find a compromise situation comparable to PLS regression (de Jong and Kiers, 1992).

## 4 Simple Linear Regression Coefficients approach to DRM

In order to investigate the dependence structure between  $X$  and the  $Y^k$ , we define the matrix  $\hat{Y}^* = [\hat{Y}^1 | \dots | \hat{Y}^k | \dots | \hat{Y}^K]$  of order  $(N, \sum_{k=1}^K Q_k)$ . The generic  $q$ -th column of the  $k$ -th matrix  $Y^k$  is given by  $\hat{y}_q^k = \sum_{j=1}^J f_j x_j b_{jq}^k$  where  $\hat{y}_q^k$  is given by the weighted sum of simple linear regression considering slope coefficient  $b_{jq}^k = f_j (x_j^T x_j)^{-1} x_j^T y_q^k$  with weights  $f_j$  and intercept equal to zero. For this weight Garthwaite (1994) suggests  $f_j = 1/J$  or  $f_j = x_j^T x_j$  according to different weighting policies.

Matrix  $\hat{Y}^* = [\hat{Y}^1 | \dots | \hat{Y}^k | \dots | \hat{Y}^K]$  can be also expressed as

$$\hat{Y}^* = XFB = \sum_{j=1}^J f_j P_{x_j} Y^*$$

with  $M_X = \text{diag}(x_1^T x_1, \dots, x_J^T x_J)$ ,  $F = \text{diag}(f_1, \dots, f_J)$ ,  $B = M_X^{-1} X^T Y^*$  and  $P_{x_j} = x_j (x_j^T x_j)^{-1} x_j^T$ . The dependence structure between  $X$  and  $Y^*$ , in a best approximation subspace, could be displayed on the principal axe  $t_s$  so as

$$\min_{t_s} \sum_{j=1}^J \sum_{k=1}^K \sum_{q=1}^{Q_k} \|f_j P_{x_j} y_q^k - f_j t_s^T P_{x_j} y_q^k\|^2 \quad (4.1)$$

subject to constraints  $t_s^T t_s = 1$  and  $t_{s'}^T t_s = 0$  for  $s' \neq s$ . This leads us to the extraction of the eigenvalues  $\lambda_s$  and eigenvectors  $t_s$  associated to the eigen-system  $\hat{Y}^* \hat{Y}^{*T} t_s = \lambda_s t_s$ .

**Table 3:** Special cases of the proposed approach. <sup>(†)</sup> First solution.

Variance Criteria	Covariance Criteria
<ul style="list-style-type: none"> <li>• PCA(<math>\hat{Y}^*</math>) is equivalent</li> <li>Multiple PCR(<math>X</math>)</li> <li>(†) MCOA (<math>\hat{Y}^1, \dots, \hat{Y}^K</math>)</li> <li>(†) COA (<math>\hat{Y}^1, \dots, \hat{Y}^K, \hat{Y}^*</math>)</li> <li>(†) OMCOA-PLS(<math>\hat{Y}^1, \dots, \hat{Y}^K, \hat{Y}^*</math>)</li> </ul>	<ul style="list-style-type: none"> <li>• <math>\text{Cov}(Y^* a, \hat{Y}^* b)</math> is equivalent to</li> <li>PLS (<math>Y^*, \tilde{X}</math>)</li> <li>PLS(<math>X, Y^*</math>) with <math>X</math> metric equal to <math>M</math></li> <li>• <math>\sum_{k=1}^K \text{Cov}^2(\hat{Y}^* a, \hat{Y}^k d_k)</math></li> <li>is equivalent to OMCOA-PLS</li> </ul>

The analysis of  $\hat{Y}^k$  and  $X$ , based on the above mentioned criteria, lead to well known techniques and interesting properties (Table 3), where MCOA stands for Multiple Coinertia Analysis (Chessel and Hanafi, 1996); COA stands for Concordance Analysis (Lafosse and Hanafi, 1997); OMCOA stands for Orthogonal Multiple Coinertia Analysis (Vivien, 1999), and finally OMCOA-PLS is the acronym for Orthogonal Multiple Coinertia Analysis – Partial Least Squares (Vivien and Sabatier, 2000).

This approach highlights an equivalence between the variance and covariance criteria in Table 3. Moreover, this can be also showed following two different approaches: the former (*B matrix approach*) is based on the matrix  $B$  of regression coefficients. An uncentred PCA on matrix  $B$  is equivalent to PLS ( $Y^*, \tilde{X}$ ) as well as the uncentred PCA on  $B'$

leads to COA, OMCOA, OMCOA-PLS, Multiblock-PLS (Wangen and Kowalski, 1988) and Generalized Constraint Principal Component Analysis (Generalized CPCA; Amenta and D'Ambra, 2001). The latter (*Crossed regression approach*) can be performed by using the  $(\sum_{k=1}^K Q_k) \times J$  simple linear regressions of each generic  $q$ -th column of the  $k$ -th matrix  $Y^k$  against each  $x_j$  (D'Ambra et al., 1998, 2001). We can write  $\sum_k Q_k$  matrices  $XB_g$  ( $g = 1, \dots, \sum_k Q_k$ ) with  $B_g$  diagonal matrix containing the  $J$  weighted regression coefficients  $b_j^g$ . In order to analyze the common structure of these  $\sum_k Q_k$  matrices we consider the MCOA approach with generic metric  $M_g$ .

Briefly, MCOA is a technique that enables the simultaneous analysis of  $Z$  tables. According to the  $Z$  subsets of  $p_g$  variables ( $g = 1, \dots, Z$ ), MCOA considers  $Z$  statistical triplets:  $(X_g, M_g, D)$  with  $M_g$  positive defined symmetrical matrix (metric) and  $X_g$  of dimensions  $(p_g \times p_g)$  and  $(N \times p_g)$ , respectively. It optimizes the variance within each table and the correlation between the scores of each individual table and synthetic scores providing a reference structure. MCOA first searches for a set of  $M_g$ -normalized  $u_g^{(1)}$  vectors, maximizing the projected variance of  $X_g$  on  $u_g^{(1)}$  and an auxiliary  $D$ -normalized vector  $v^{(1)}$ , maximizing the projected variance of  $X_g^T$  on  $v^{(1)}$ , such that the squared covariance between them is optimized,  $\max \sum_{g=1}^Z \pi_g \left( X_g M_g u_g^{(1)} | v^{(1)} \right)_D^2$ , where  $\pi_g$  represents a weight assigned to each  $X_g$ . This weight can be uniform, the inverse of global inertia or the inverse of the greatest eigenvalue of each table.

The first order solutions  $u_g^{(1)}$ 's and  $v^{(1)}$  are given by a PCA of the weighted table  $\hat{X}^{(1)} = [\pi_1^{1/2} X_1 | \dots | \pi_Z^{1/2} X_Z]$  according to the eigen decomposition of the matrix  $\hat{X}^{(1)} \tilde{Q} \hat{X}^{(1)T}$  with  $\tilde{Q} = \text{diag}(M_1, \dots, M_Z)$ . In similar way, for the solution of order 2, MCOA searches for  $M_g$ -normalized  $u_g^{(2)}$  vectors and an auxiliary  $D$ -normalized vector  $v^{(2)}$  by using the same optimization criterion with the additional orthogonal constraints  $u_g^{(1)T} M_g u_g^{(2)} = v^{(1)T} D v^{(2)} = 0$ . Solutions of order 2 are given by the first order PCA solution of the juxtaposed residual matrix  $[X_1 - X_1 P_1^{(1)T} | \dots | X_Z - X_Z P_Z^{(1)T}]$  with  $P_g^{(1)}$  the  $M_g$  orthogonal projection operator onto the subspace spanned by the vector  $u_g^{(1)}$ . The successive solutions are found in similar way.

By applying the MCOA approach to the  $Z = \sum_k Q_k$  matrices ( $g = 1, \dots, \sum_k Q_k$ ), first order solutions  $u_g^{(1)}$ 's and  $v^{(1)}$  are then given by a PCA of the weighted table  $\hat{X}^{(1)} = [\pi_1^{1/2} X B_1 | \dots | \pi_{\sum_k Q_k}^{1/2} X B_{\sum_k Q_k}] = X \tilde{M}$  with  $\tilde{M} = [\pi_1^{1/2} B_1 | \dots | \pi_{\sum_k Q_k}^{1/2} B_{\sum_k Q_k}]$ . The first order solutions are given by the eigen decomposition of the matrix

$$\hat{X}^{(1)} \tilde{Q} \hat{X}^{(1)T} = X \tilde{M} \tilde{Q} \tilde{M}^T X^T$$

with  $\tilde{Q} = \text{diag}(M_1, \dots, M_{\sum_k Q_k})$ . Solutions of order 2 are given by the first order PCA solution of the juxtaposed residual matrix

$$\hat{X}^{(2)} = [X B_1 - X B_1 P_1^{(1)T} | \dots | X B_{\sum_k Q_k} - X B_{\sum_k Q_k} P_{\sum_k Q_k}^{(1)T}]$$

We remark that if  $M_g = I$  then the first solution of PCA of  $\hat{X}^{(1)}$  is equivalent to the same solution of a PCA of matrix  $X$  with diagonal metric containing the weighted sums of the explained variances by each  $x_j$ . If  $M_g = \text{diag}(1/y_g^T y_g)$  and  $f_j = x_j^T x_j$  then this approach is equivalent to a PCA on the matrix  $X$  with diagonal metric  $(\sum_g \pi_g B_g M_g B_g^T)$

of the weighted sums of the coefficients of determination  $r_g^2$ :  $X(\sum_g \pi_g B_g M_g B_g^T)X^T$ . We highlight that this approach can be considered as an asymmetrical extension of MCOA of  $K$  response variable groups  $Y^k$  ( $k = 1, \dots, K$ ) respect to a set of predictive variables  $X$ .

Moreover, the weighted sum of the explained variances by each  $x_j$  can be used as weight within the Garthwaite's univariate approach as well as within the Multiple Coinertia Analysis. In this sense, it is interesting to note the role played by the coefficient regression  $b_j^g$  within the different proposed approaches as well as it's easy to show that all the proposals are linked by transition formula. Obviously, this approach works also with a single dependent variable  $y$  as well as with a single group of variable ( $K = 1$ ).

This proposal provides a suitable conditioned matrices  $G$  within the shrinkage regression methods too (see Table 1). The approach based on the  $\hat{y}_q^k$  as sum of orthogonal projections onto single rank subspaces spanned by the  $x_j$ 's, leads also to consider the covariance between the  $x_j$ 's and the  $\hat{y}_q^k$ 's. In this case, we have  $cov(X, \hat{Y}^*) = AX^T Y^*$  where  $A$  is a matrix of order  $(J \times J)$  whose general element is the weighted paired regression coefficient among the  $x_j$ 's:  $a_{j,j'} = f_j cov(x_j, x_{j'}) / var(x_{j'})$ , ( $j, j' = 1, \dots, J$ ).

If we refer to the  $q$ -th column of  $Y^k$ , we obtained the predictor  $b^{\hat{Y}^k} = AX^T y_q^k$ . In this way we can consider the matrix  $A$  as an alternative conditioned matrix for collinearity problem in Table 1. We remark that this approach tries to get back the relationships among the predictor variables which are loosed in simple linear regression.

## 5 Conclusions

The main aim of this paper is to find the linkage between several multidimensional techniques like MCOA, PLS, OMCOA-PLS, COA, OMCOA, Multiblock - PLS and Generalized CPCA, within a simple linear regression framework. At the same time new methodological proposals are done.

These results are particularly important when the matrix of explicative variables has a rank lower than  $\min(N, J)$  that could lead to problems of stability. Another advantage of this approach is that it can be performed without specialized software.

An extension of this framework, to several matrices of explicative and dependent variables, will appear in a next paper. An extension to categorical variables is also under investigation.

## Acknowledgements

The present paper is financially supported by the Cofin04 fund (responsible prof. L. D'Ambra) and Cofin04 (responsible prof. P. Amenta).

## References

- [1] Abraham, B. and Merola, G. (2001): Dimensionality reduction approach to multivariate prediction. In Esposito Vinzi V. et al. (Eds.): *PLS and related methods*, CISIA, 3-17.

- 
- [2] Amenta, P. and D'Ambra, L. (2001): Generalized constrained principal component analysis. In Borra, S., Rocci, R., Vichi, M., and Schader, M. (Eds.): *Advances in Classification and Data Analysis*, Springer, 137-144.
- [3] Brooks, R. and Stone, M. (1994): Joint continuum regression for multiple predictands, *JASA*, **89**, 1374-1377.
- [4] Brown, P.J. (1993): *Measurement, Regression and Calibration*. Oxford: Oxford Univ. Press.
- [5] Chessel, D. and Hanafi, M. (1996): Analyse de la co-inertie de  $K$  nuages de points. *Revue Statistique Appliquée*, XLIV, 35-60.
- [6] D'Ambra, L. and Lauro, N.C. (1982): Analisi in componenti principali in rapporto ad un sottospazio di riferimento, *Rivista di Statistica Applicata*, **15**, 1-25.
- [7] D'Ambra, L., Amenta, P., Rubinacci, F., Gallo, M., and Sarnacchiaro, P. (1999): *Multidimensional Statistical Methods based on Co-Inertia for the Customer Satisfaction Evaluation*. International Quality Conference IQC, 8-11/12/99, Bangkok, Thailandia.
- [8] D'Ambra, L., Sabatier, R., and Amenta, P. (1998): Analisi fattoriale delle matrici a tre vie: sintesi e nuovi approcci. (Invited lecture) Atti XXXIX Riunione SIS. Sorrento.
- [9] D'Ambra, L., Sabatier, R. and Amenta, P. (2001): Three way factorial analysis: synthesis and new approaches. *Italian Journal of Applied Statistics*, **13**, 101-117.
- [10] de Jong, S. and Kiers, H.A.L. (1992): Principal covariates regression. Part I. Theory, *Chemometrics and Intelligent Laboratory Systems*, **14**, 155-164.
- [11] Frank, I.R. and Friedman, J.H. (1993): A statistical view of some chemometrics regression tools. *Technometrics*, **35**, 109-148.
- [12] Garthwaite, P.H. (1994): An interpretation of partial least squares, *JASA*, **89**, 122-127.
- [13] Hoerl, A.E. (1962): Application of ridge analysis to regression problems, *Chemical Engineering Progress*, **58**, 54-59.
- [14] Hoerl, A.E. and Kennard, R.W. (1970): Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, **12**, 55-67.
- [15] Lafosse, R. and Hanafi, M. (1997): Concordance d'un tableau avec  $K$  tableaux: définition de  $K+1$  uples synthétiques. *Revue de Statistique Appliquée*, 45, **4**, 111-126.
- [16] Stone, M. and Brooks, R.J. (1990): Continuum regression: cross validated sequentially constructed prediction embracing ordinary least squares and principal component regression. *J. Royal Stat. Soc.*, B, **52**, 237-269.

- 
- [17] Vivien, M. (1999): *Nouvelles approches en Analyse Multitableaux*. Mémoire de stage de DEA, Université Montpellier II.
- [18] Vivien, M. and Sabatier R. (2000): Une extension multi-tableaux de la régression PLS, *Revue Statistique Appliquée*, **49**, 31-54.
- [19] Wangen, L.E. and Kowalski, B.R. (1988): A multiblock partial least squares algorithm for investigating complex chemical system, *Journal of Chemometrics*, **3**, 3-20.
- [20] Wold, H. (1966): *Estimation of principal components and related models by iterative least squares*. In Krishnaiah, P.R. (Ed.): *Multivariate Analysis*, NY: Ac. Press.