



# Some Correlational Measures And Their Application In Principal Component

Annie Claret

## ► To cite this version:

Annie Claret. Some Correlational Measures And Their Application In Principal Component. Annales de l'ISUP, 2001, XXXXV (1), pp.15-29. <hal-03648334>

**HAL Id: hal-03648334**

**<https://hal.science/hal-03648334v1>**

Submitted on 21 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

*Pub. Inst. Stat. Univ. Paris*  
*XXXV, fasc. 1, 2001, 15 à 29*

# Some Correlational Measures And Their Application In Principal Component Analysis

Annie CLARET  
 Faculté des Sciences Economiques  
 Espace Richter  
 Avenue de la Mer  
 B.P. 9606  
 34054 MONTPELLIER CEDEX 1

## ABSTRACT

We are interested in the problem of matrix approximation and we propose some new correlational measures constructed from unitarily invariant norms. We prove that the solution of Principal Component Analysis with respect to Instrumental Variables given by Rao in 1964 is optimal for all these correlational measures.

Key-words : Correlational measure

Unitarily invariant norm

Principal Component Analysis with respect to Instrumental Variables



## INTRODUCTION

a) Consider two matrices  $X_1$  and  $X_2$  with  $p$  and  $q$  columns respectively both having  $n$  rows. They represent measures of  $p$  and  $q$  variables respectively observed on the same  $n$  individuals.

From these matrices geometrical representations of the  $n$  individuals are observed as two configurations of  $n$  points in  $\mathbb{R}^p$  and  $\mathbb{R}^q$  and our aim is to measure the similarity of these configurations. In order to measure such a similarity we first need to do some linear transformations of the data given in  $X_1$  and  $X_2$ ,  $L_1, p \times s$  and  $L_2, q \times s$ , to compare  $X_1 L_1$  and  $X_2 L_2$  through the configurations we obtain in  $\mathbb{R}^s$ , and we need then a correlational measure on  $\mathcal{M}_{n \times s}$ , space of the complex matrices of order  $n \times s$ , to compare them.

b) The general framework we use is the following one :

① Consider a matrix  $X$  of order  $n \times p$ , of measures of  $p$  variables observed on  $n$  individuals. We can affect weights to the individuals,  $p_i$ ,  $i \in \{1, \dots, n\}$ ,  $p_i \geq 0$  for each  $i$ , and  $\sum_{i=1}^n p_i = 1$  and define a diagonal matrix  $D$ ;  $D = \text{diag}(p_1, \dots, p_n)$  such that the distance between the  $j$ th and  $k$ th variable is equal to  $(X_j^i - X_k^i) D (X_j^i - X_k^i)$  where  $X_j^i$  represents the  $j$ th column of the matrix  $X$ .

We see that when the variables are centered for the matrix  $D$ , that is  $X' D U = 0$  with  $U = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ ,  $V = X' D X$  represents the matrix of variance between the variables.

② We often need to do some linear transformation of the data and replace  $X$  by  $XL$ ,  $L$   $p \times s$ .

The principal components of  $XL$ , when the weights given to the individuals are defined by the diagonal matrix  $D$ , are the vectors  $\psi$



that satisfy :

$$XL(L'X'DXL\varphi) = XL(\lambda\varphi)$$

$$(XLL'X'D)\psi = \lambda\psi$$

so the vectors  $\psi$  are eigen vectors of the matrix  $XLL'X'D$  ; we shall write :  $WD = XLL'X'D$ .

Given the linear transformation  $L$ ,  $LL'$  is a positive semi-definite matrix  $Q$ . Reciprocally if  $Q$  is a positive semi-definite matrix  $Q$ ,  $Q$  can be written  $Q = LL'$  where  $L$  is the matrix of a linear transformation,  $L$  being of order  $p \times s$ .

So there is an equivalence between the choice of a metric defined by  $Q$  on the space  $\mathbb{R}^p$  and a linear transformation of the data.

$WD = XLL'X'D = XQX'D$  is then characteristic of the study in the sense that its eigenvectors are the principal components.

③ We shall then associate to the matrix  $X$ , two metrics  $Q$  and  $D$  and we shall speak of the array  $(X, Q, D)$ .

Then we shall characterize the study  $(X, Q, D)$  by the operator  $WD$ .

c) The operators  $W_1D$  and  $W_2D$  characterize the studies  $(X_1, Q_1, D)$  and  $(X_2, Q_2, D)$  respectively. In order to measure the similarity of the two configurations of the  $n$  points, we want to compare the operators  $W_1D$  and  $W_2D$  and we need a correlational measure on  $\mathbb{R}^n$ .

Escoufier proposed the RV-coefficient constructed from the scalar product on  $H_n(\mathbb{R})$ , subspace of  $\mathcal{M}_n(\mathbb{R})$  formed by the D-symmetrical matrices :

$$H_n(\mathbb{R}) \times H_n(\mathbb{R}) \longrightarrow \mathbb{R} \quad \text{defined by} \quad RV(A,B) = \frac{\text{tr}(AB)}{[\text{tr}(A^2)\text{tr}(B^2)]^{1/2}} \text{ and}$$

$$(A,B) \longmapsto \text{tr}(AB)$$

computed  $RV(W_1D, W_2D)$ .

$$(H_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) ; A = WD, W' = W\})$$

d) Our aim in this paper is to propose some new correlational measures constructed from unitarily invariant norms and demonstrate that the solution of Principal Component Analysis with respect to Instrumental Variables given by Rao in 1964 is optimal for all those coefficients and not only for the RV-coefficient as it is proved in the paper of Robert and Escoufier (1979).



## I- Some useful results and definitions

### I-1. Correlational measures for matrices.

Définition 1 (Renyi, 1959). The mapping  $r : \mathcal{M}'_{n \times s} \times \mathcal{M}_{n \times s} \rightarrow [0,1]$  is called a correlation function in the metric  $D$  if for all non zero scalar  $a$  and  $b$  and when  $A$  and  $B$  are not both zero :

$$c_1 : r(aA, B) = r(A, bB) = r(A, B)$$

$$c_2 : r(A, B) = r(B, A)$$

$$c_3 : r(A, B) = 1 \text{ if } A = bB$$

$$c_4 : r(A, B) = 0 \text{ iff } A'DB = 0$$

Remark. ① The condition  $c_3$  is necessary to ensure  $r$  is a useful measure of correlation. We should prefer to obtain an equivalence. Remark the equivalence is obtained when  $r$  is constructed as a cosines, from a scalar product on  $\mathcal{M}_{n \times s}$  as the RV-coefficient.

② The condition  $c_4$  is interesting because we are concerned by the representation of individuals. In the dual problem of the representation of variables, we will replace  $c_4$  by  $c'_4$  : " $r(A, B) = 0$  iff  $AQB' = 0$ ."

### I-2. Unitarily invariant norm

1) Definition 2. A real valued function defined on the space  $\mathcal{M}_{n \times s}$  of  $n \times s$  complex matrices is called a unitarily invariant norm if it satisfies the following conditions :

1.  $\|A\| \geq 0$  if  $A \neq 0$
2.  $\|aA\| = |a| \|A\|$
3.  $\|A+B\| \leq \|A\| + \|B\|$
4.  $\|U_1 A U_2\| = \|A\|$  for any unitary matrices  $U_1$  and  $U_2$  of orders

$n$  and  $s$  respectively.

Remark : if  $A = U_1 \Sigma U_2^*$  is the singular value decomposition of the matrix  $A$   
 $\|A\| = \|\Sigma\|$

## 2) Theorem (Schatten, 1960)

Let  $A$  be a matrix of  $M_{n \times s}$  and  $s_i(A)$  the  $i^{\text{th}}$  singular value of  $A$ . For any  $r \in \mathbb{R}^{+*}$  :

$$M_{n \times s} \longrightarrow \mathbb{R}^+$$

$$A \longmapsto \sum_{i=1}^s [s_i(A)]^r \frac{1}{r}$$

is a unitarily invariant norm, called Schatten's  $r$ -norm.

## II- Some correlationnal measures on $M_{n \times s}$

Let us consider the array  $(A, Q, D)$

$A \in M_{n \times s}$  is considered as the matrix of a linear mapping from  $(\mathbb{R}^s, Q)$  to  $(\mathbb{R}^n, D)$ .  $A^*$ , the adjoint of  $A$  is given by :  $A^* = Q^{-1} A' D$



## II-1. Definitions

1. Definition 1. Let  $m \in \mathbb{N}^*$ 

$$\|\cdot\|_m : \mathcal{M}_{n \times s} \longrightarrow \mathbb{R}^+$$

$$A \longrightarrow \|A\|_m = \left[ \sum_{i=1}^s s_i^{4m}(A) \right]^{\frac{1}{4m}}$$

$\|\cdot\|_m$  is a unitarily invariant norm.

2. Definition 2. Let  $m \in \mathbb{N}^*$ 

$$\varphi_m : \mathcal{M}_{n \times s} \times \mathcal{M}_{n \times s} \longrightarrow \mathbb{R}$$

$$(A, B) \longmapsto \varphi_m(A, B) = \left[ \sum_{i=1}^s s_i^{2m}(A^*B) \right]^{\frac{1}{2m}}$$

We can see that  $\varphi_m(A, A) = \|A\|_m^2$  and then define :

3. Definition 3. Let  $m \in \mathbb{N}^*$ , for any  $A$  and  $B$  non both zero, we define :

$$\rho_m(A, B) = \frac{\varphi_m(A, B)}{\|A\|_m \|B\|_m} = \frac{\left[ \sum_{i=1}^s s_i^{2m}(A^*B) \right]^{\frac{1}{2m}}}{\left[ \left( \sum_{i=1}^s s_i^{4m}(A) \right) \left( \sum_{i=1}^s s_i^{4m}(B) \right) \right]^{\frac{1}{4m}}}$$

Remarks : ① For any  $m, m \in \mathbb{N}^*$ ,  $\rho_m(A, B) = \left[ \frac{(\text{tr}[A^*BB^*A]^m)^2}{\text{tr}[A^*A]^{2m} \text{tr}[B^*B]^{2m}} \right]^{\frac{1}{4m}}$

We are going to show that  $\rho_m$  takes its values in  $[0, 1]$ , using a lemma concerning the trace of matrix products and this is the reason of the choice of  $m \in \mathbb{N}^*$ .

② For any  $m, m \in \mathbb{N}^*$ ,  $\rho_m$  satisfies the conditions  $c_1$ - $c_4$  of the definition 1 and for any matrices  $A, B$ ,  $\rho_m(A, B) \geq 0$ .

③ If  $A$  and  $B$  belong to  $\mathbb{R}^n$  with  $A'DU = B'DU = 0$  and  $Q = I$ , for any  $m$ ,  $\rho_m(A, B)$  is equal to the absolute value of the linear correlation coefficient between  $A$  and  $B$ .



We have seen that, in the general case,  $\rho_m$  satisfies the conditions  $c_1$ - $c_4$  of Renyi's definition and is positive.

It remains to be proved that, for any  $m \in \mathbb{N}^*$ , for any matrices  $A$  and  $B$  in  $\mathcal{M}_{n \times s}$ ,  $\rho_m(A, B) \leq 1$ . We prove this result in the following theorem :

II-2. Theorem : For any  $m$ ,  $m \in \mathbb{N}^*$

For any  $A$  and  $B$  in  $\mathcal{M}_{n \times s}$ ,  $\rho_m(A, B) \in [0, 1]$ .

Proof : It is equivalent to show that  $[\text{tr}[(AA^*BB^*)^m]]^2 \leq \text{tr}[(AA^*)^{2m}] \text{tr}[(BB^*)^{2m}]$

① We suppose first that  $Q = I_s$  and  $D = I_n$ .

$A^*$  is then the conjugate-transpose of the matrix  $A$ .

For any square and diagonalisable matrix  $C$ , we denote  $\Lambda_C$ . The diagonal matrix with the eigenvalues of  $C$  arranged in decreasing order on the diagonal.

We obtain the following lemma :

Lemma 1 : Let  $C$  and  $D$  be two hermitean matrices of order  $n$ , with real and non negative eigenvalues.

For any  $m$ ,  $m \geq 1$  with  $\text{tr}[(CD)^m]$  real

$$\text{tr}[(CD)^m] \leq \text{tr}(\Lambda_C^m \Lambda_D^m)$$

Proof of the lemma :  $\text{tr}[(CD)^m] = \text{tr}(CD \dots CD) = \text{tr}(I_n C I_n D I_n \dots C I_n D)$

$I_n$  : identity matrix is unitary. A Kristof's theorem (Kristof, 1970)

proves :



$$\max \operatorname{Re}(\operatorname{tr} Z_1 C Z_2 D \dots Z_{2m-1} C Z_{2m} D) \leq \sum_{k=1}^n (\alpha_k \beta_k \dots \alpha_k \beta_k) \\ (Z_1, \dots, Z_{2m}) \in \mathcal{U}_n^{2m}, \mathcal{U}_n \text{ space of unitary matrices of order } n.$$

where  $\alpha_k$  and  $\beta_k$  denotes the  $k^{\text{th}}$  eigenvalue of  $C$  and  $D$  respectively.

By transitivity we obtain ; as  $\operatorname{tr}[(CD)^m]$  is a real number :

$$\operatorname{tr}[(CD)^m] \leq \sum_{k=1}^n (\alpha_k \beta_k)^m = \sum_{k=1}^n \alpha_k^m \beta_k^m = \operatorname{tr}(\Lambda_C^m \Lambda_D^m)$$

Then we choose  $C = AA^*$  and  $BB^*$ .  $C$  and  $D$  are hermitean matrices and  $\operatorname{tr}[(AA^*BB^*)^m]$  is real. So we obtain :

$$\operatorname{tr}[(AA^*BB^*)^m] \leq \operatorname{tr}[\Lambda_{(AA^*)^m} \Lambda_{(BB^*)^m}]$$

and  $(\operatorname{tr}[\Lambda_{(AA^*)^m} \Lambda_{(BB^*)^m}])^2 \leq \operatorname{tr}[(AA^*)^{2m}] \operatorname{tr}[(BB^*)^{2m}]$ , Cauchy-Schwarz inequality.

We can end :

$$[\operatorname{tr}[(AA^*BB^*)^m]]^2 \leq \operatorname{tr}[(AA^*)^{2m}] \operatorname{tr}[(BB^*)^{2m}]$$

2 We suppose now that there are no restrictive conditions upon  $Q$  and  $D$ . Weights defined by the diagonal matrix  $D$  are given to the individuals defined by the matrix  $A$ ,  $n \times s$ .

A linear transformation as been done on the variables defined by  $A$  that permits to define a positive semi-definite matrix  $Q$

$$Q^{-1} \text{ can be written : } Q^{-1} = M'M \text{ (Cholesky's-decomposition)}$$

Let  $\alpha$  be defined by the equality  $\alpha = D^{1/2} A M'$ .

Consider the array  $(\alpha, I_s, I_n)$ .



Let  $\alpha_*$  be the adjoint of  $\alpha$  in respect to the metrics  $I_S$  and  $I_n$

$$\alpha_* = I_S^{-1} \alpha' I_n = \alpha'.$$

The results found in (1) permit to assume that :

$$0 \leq \rho_m(\alpha, \beta) \leq 1$$

with  $\alpha = D^{1/2} A M'$  and  $\beta = D^{1/2} B M'$ .

Our aim is to prove that  $\rho_m(\alpha, \beta) = \rho_m(A, B)$ .

Lemme 2 : The positive eigenvalues of  $A^* A$  and those of  $\alpha_* \alpha$  are the same.

Proof of the lemma :  $A^* A = Q^{-1} A' D A = M' M A' D A = M' Z$  with  $Z = M A' D A$

$$\text{and} \quad \alpha_* \alpha = M A' D^{1/2} D^{1/2} A M' = M A' D A M' = Z M'$$

and the non zero eigenvalues of  $M' Z$  and  $Z M'$  are the same.

$$\text{So : } \{s_i(A) ; i = 1, \dots, s\} = \{s_i(\alpha) ; i = 1, \dots, s\}$$

$$\{s_i(B) ; i = 1, \dots, s\} = \{s_i(\beta) ; i = 1, \dots, s\}, \quad \beta = D^{1/2} B M', \quad Q^{-1} = M' M.$$

And  $A^* B = Q^{-1} A' D B = M' M A' D B = M' T$  with  $T = M A' D B$

$$\alpha_* \beta = \alpha' \beta = M A' D^{1/2} D^{1/2} B M' = M A' D B M' = T M'$$

proves that :

$$\{s_i(A^* B) ; i = 1, \dots, s\} = \{s_i(\alpha_* \beta) ; i = 1, \dots, s\}$$

$$\text{and} \quad \rho_m(A, B) = \rho_m(\alpha, \beta).$$

The result in (1) shows that  $0 \leq \rho_m(\alpha, \beta) \leq 1$  then  $0 \leq \rho_m(A, B) \leq 1$  ■

We are now interested in solving a problem of approximation using these coefficients.



### III- Principal components analysis with respect to instrumental variables.

#### III-1. Position of the problem and its solutions using the RV-coefficient.

The data are collected in two matrices  $X_1$  and  $X_2$  of order  $n \times p$  and  $n \times q$  respectively concerning  $p$  and  $q$  variables measured on the same individuals.

$X_1$  and  $X_2$  are centered for the metric  $D$  and  $X_2$  is often supposed to be a one to one transformation from  $\mathbb{R}^p$  to  $\text{Im}(X_2)$ .

The array  $(X_1, Q_1, D)$  and the matrix  $X_2$  are given.

Let us consider a  $q \times k$  matrix  $M$ , a linear transformation of the variables defined by  $X_2$ . We wish to replace  $X_2$  by  $X_2 M$  in such a way that the predictive efficiency of  $X_2 M$  for  $X_1$  is a maximum. It is equal to find a symmetrical semi-definite positive matrix  $Q_2 = MM'$  such that the operators  $W_1 D$  and  $W_2 D$  be as close as possible.

Then we call principal component analysis of the matrix  $X_2$  with respect to the instrumental variables  $X_1$ , of order  $k$ ,  $k \leq \min(p, q)$ , the determination of the matrix  $Q_2$ ,  $q \times q$ ,  $Q_2 = MM'$  of rank  $k$  such that :

$$\|X_1 Q_1 X_1' D - X_2 Q_2 X_2' D\| \text{ is a minimum,}$$

where  $\|\cdot\|$  is a suitably chosen norm.

Another way of interpretation of the problem is to search  $Q_2$ ,  $q \times q$ ,  $Q_2 = MM'$ , of rank  $k$   $k \leq \min(p, q)$ , such that  $r(W_1 D, W_2 D)$  is maximum, where  $r$  is a correlational measure.

Using the RV-coefficient the following result has been found  
[Escoufier, Robert, 1979].



Theorem 1 : Let  $X_2'DX_2 = LL'$  and let  $\Lambda$  be the diagonal matrix of the  $k$  largest eigenvalues of  $A = L^{-1}X_2'DW_1DX_2L'^{-1}$

If  $M$ ,  $q \times k$  satisfies the two following equations :

$$\begin{cases} X_2'DW_1DX_2M = X_2'DX_2M\Lambda \\ M'X_2'DX_2M = \Lambda \end{cases}$$

Then  $M$  realizes the maximum of  $Rv(W_1D, X_2MM'X_2'D)$  under the constraints  $M'X_2'DX_2M$  is a diagonal matrix and  $\text{rank}(MM') = k$

And the maximum value of  $Rv$  is equal to  $\frac{[\text{tr } \Lambda]^2]^{1/2}}{[\text{tr}(W_1D)^2]^{1/2}}$

We prove now that the same value of  $M$  maximizes all the coefficients  $\rho_m$  introduced in II.

III-2. Solution of the problem using the coefficients  $\rho_m$ ,  $m \in \mathbb{N}^*$ .

We are interested in the search of  $Q_2$ ,  $q \times q$ ,  $Q_2 = MM'$  of rank  $k$   $k \leq \min(p, q)$  such that :

$$\left[ \frac{(\text{tr}[(W_1DX_2Q_2X_2'D)^m])^2}{\text{tr}[(W_1D)^{2m}] \text{tr}[(X_2Q_2X_2'D)^{2m}]} \right]^{\frac{1}{4m}} \text{ be a maximum,}$$

we shall write this quantity  $\rho_m(W_1D, W_2D)$  (instead of  $\rho_m(X_1, X_2)$ )

We obtain the following result that is a generalization of theorem 1.

Theorem 2 : Given the array  $(X_1, Q_1, D)$ ;  $X_1 \in \mathbb{R}^{n \times p}$  and the matrix  $X_2 \in \mathbb{R}^{n \times q}$  Let  $X_2'DX_2 = LL'$  and let  $\Lambda$  be the diagonal matrix of the  $k$  largest eigenvalues of  $A = L^{-1}X_2'DW_1DX_2L'^{-1}$ ;  $M$  defined in the theorem 1 realizes

$$\text{the maximum, for any } m, m \geq 1, \text{ of } \rho_m(W_1D, X_2Q_2X_2'D) = \left[ \frac{(\text{tr}[(W_1DX_2Q_2X_2'D)^m])^2}{\text{tr}[(W_1D)^{2m}] \text{tr}[(X_2Q_2X_2'D)^{2m}]} \right]^{\frac{1}{4m}}$$

under the constraints  $M'X_2'DX_2M$  is a diagonal matrix and  $\text{rank}(MM') = k$



And the maximum value of  $\rho_m$  is equal to :  $\left[ \frac{[\text{tr}(\Lambda^{2m})]}{\text{tr}(W_1 D)^{2m}} \right]^{\frac{1}{4m}}$ .

Proof of the theorem : Remark it is equal to find a matrix  $M$ ,  $q \times k$ ,  $k \leq q$ , such that  $M'X_2'DX_2M$  is diagonal,  $\text{rank}(MM') = k$  that maximizes  $[\rho_m(W_1DX_2MM'X_2'D)]^{2m} = \beta_m$  because  $f_m : \mathbb{R}^+ \longrightarrow \mathbb{R}^m$  is an increasing function.  

$$x \longmapsto x^{1/2m}$$

Let  $X_2'DX_2 = LL'$ ,  $M$   $q \times k$ , of rank  $k$  with  $M'X_2'DX_2M$  is diagonal.

Let  $N = L'M$ . Then  $N'N$  is a diagonal matrix and the columns of  $N$  are orthogonal.

$$MM' = L'^{-1}NN'L^{-1}$$

$$\text{and } \beta_m = \frac{\text{tr}[(X_2'DW_1DX_2MM')^m]}{[\text{tr}(W_1D)^{2m}]\text{tr}(X_2'DX_2MM')^{2m}]^{1/2}}$$

With :  $MM' = L'^{-1}NN'L^{-1}$

$$\text{and } \text{tr}[(X_2'DX_2MM')^{2m}] = \text{tr}[(M'X_2'DX_2M)^{2m}] = \text{tr}[(N'N)^{2m}]$$

$$\beta_m = \frac{\text{tr}[L^{-1}X_2'DW_1DX_2L'^{-1}NN']^m}{[\text{tr}(W_1D)^{2m}]\text{tr}(N'N)^{2m}]^{1/2}}$$

We now use the lemma 1 proved in II with  $C = L^{-1}X_2'DW_1DX_2L'^{-1} = A$  and

$D = NN'$ , and obtain :

$$\text{tr}[L^{-1}X_2'DW_1DX_2L'^{-1}NN']^m \leq \text{tr}(\Lambda_{(L^{-1}X_2'DW_1DX_2L'^{-1})^m} \Lambda_{(NN')^m})$$

and  $\text{rank}(NN') = k$ ,  $k \leq q$ . Then  $\Lambda_{(NN')^m}$  is the diagonal matrix with the eigenvalues of  $NN'$  ranged in decreasing order, so the last  $q-k$  terms on the diagonal are equal to zero.

$$\text{So : } \text{tr}(\Lambda_{A^m} \Lambda_{(NN')^m}) = \text{tr}([\Lambda_{A^m}]_k \Lambda_{(NN')^m})$$

where  $[\Lambda_{A^m}]_k$  is the diagonal matrix having the  $k$  largest eigenvalues of  $A^m$  as its  $k$  first terms on its diagonal, all the other terms being equal to zero.



Cauchy-Schwarz inequality proves :  $\text{tr}([\Lambda_{A^m}]_k \Lambda_{(NN')^m}) \leq [\text{tr}(\Lambda_{A^m})_k^2 \text{tr}(\Lambda_{(NN')^m})^2]^{\frac{1}{2}}$

And by transitivity we obtain :

$$\text{tr}[(ANN')^m] \leq [\text{tr}(\Lambda_{A^m})_k^2 \text{tr}(\Lambda_{(NN')^m})^2]^{\frac{1}{2}}$$

and  $NN'$  and  $N'N$  have the same non zero eigenvalues. Then :

$$\text{tr}(\Lambda_{(NN')^m})^2 = \text{tr}[(N'N)^{2m}]$$

then :

$$\beta_m \leq \left[ \frac{\text{tr}(\Lambda_{A^m})_k^2 \text{tr}[(N'N)^{2m}]}{\text{tr}[(W_1 D)^{2m}] \text{tr}[(N'N)^{2m}]} \right]^{\frac{1}{2}} = \left[ \frac{\text{tr}(\Lambda^{2m})}{\text{tr}[(W_1 D)^{2m}]} \right]^{\frac{1}{2}}$$

Now let  $M$  be the  $q \times k$  matrix defined in the theorem

$$\text{Then } \beta_m = \frac{\text{tr}[(M'X_2'DX_2M\Lambda)^m]}{[\text{tr}[(W_1 D)^{2m}] \text{tr}[(M'X_2'DX_2M)^{2m}]]^{1/2}} = \frac{\text{tr}(\Lambda^{2m})^m}{[\text{tr}[(W_1 D)^{2m}] \text{tr}(\Lambda^{2m})]^{1/2}}$$

$$\beta_m = \left[ \frac{\text{tr}(\Lambda^{2m})}{\text{tr}[(W_1 D)^{2m}]} \right]^{1/2} \text{ and the maximum is attained}$$

So the following result is proved :

For any integer  $m$ ,  $m \geq 1$ , for any  $M$   $q \times k$  of rank  $k$ , with  $M'X_2'DX_2M$  diagonal, .

$$\rho_m(W_1 D, X_2 M M' X_2' D) \leq \left[ \frac{\text{tr}(\Lambda^{2m})}{\text{tr}[(W_1 D)^{2m}]} \right]^{\frac{1}{4m}}$$

and the bound is attained for  $M$  defined in the theorem.

It shows that the solution of Principal Component Analysis with respect to Instrumental Variables given by Rao, [Rao, 1964], is optimum for all the coefficients  $\rho_m$ , and we can express the value of this optimum.



# Bibliography

CLARET, A. 1987, Contribution au problème de l'approximation factorielle d'un tableau de données, Doctoral thesis, Montpellier, France.

ESCOUFIER, Y. ; ROBERT, P. ; 1979, Choosing variables and metrics optimizing the Rv-coefficient, Academic Press, INC. Rustagi.

KRISTOF, W., 1970, A theorem on the Trace of certain Matrix Products and some Applications, *Journal of Mathematical psychology*, 7, 515-530.

RAMSAY, J.O. ; ten BERGE J. ; STYAN, G.P.H., 1984, Matrix correlation *Psychometrika*, 49, n°3, 403-423.

RAO, C.R., 1964, The use and interpretation of principal component analysis in applied research, *Sankya*, ser. A, 26, 329-359.

RAO, C.R., 1973, Linear statistical inference and its applications, John Wiley, New-York.

ROBERT, P. ; ESCOUFIER, Y., 1976, A unifying tool for linear multivariate statistical methods : the RV-coefficient, *Applied Statistics*, 25, 257-265.

SCHATTEN, N., 1960, Norms ideals of completely continuous operators, Springer Verlag.

VON NEUMANN, J., 1937, Some matrix inequalities and metrization of metric-space, *Collected works*, Pergamon, Oxford, vol. IV, 205-218.