Classification and Related Methods of Data Analysis
H.H. Bock (Editor)

BEYOND CORRESPONDENCE ANALYSIS

Y. ESCOUFIER

Unité de Biométrie, ENSA-INRA-USTL, Place Pierre Viala,
34060 Montpellier Cedex. France

Correspondence Analysis is based on three choices : observed frequencies are compared with independence model by a least squares criterion. Modifying these choices, we meet a three dimensional set of problems. Taking more than two qualitative variables into consideration, we enter in a fourth dimension.

1. INTRODUCTION

This paper is related to the study of a contingency table (P_{ij}, $i \in I$, $j \in J$) with the usual notations for the margins ($P_{i.}$, $i \in I$) and ($P_{.j}$, $j \in J$).

Correspondence Analysis is a method developped for the exploration of this kind of data ; it is based on the three following choices :

a) The calculus are made on the frequencies whereas it is wellknown that other approaches use the logarithms of the frequencies.

b) The observed P_{ij} are compared with expected values which come from the independence hypothesis made for the variables defining rows and columns of the data array. Alternative reference hypothesis can be formulated.

c) The results are given by a singular values decomposition which is a consequence of the chosen least squares criterion. Alternative criteria can be taken into account. Thus, if we accept to modify these choices, we meet a three dimensional set of problems which will be explored in this paper.

When more than two variables are available, the usual extension of Correspondence Analysis is Multiple Correspondence Analysis. In that method, if we have three variables, the P_{ijk} are never introduced in the calculus but only $P_{ij.}$, $P_{i.k}$, and $P_{.jk}$. We will try to go beyond that limit.

2. LEAST SQUARES CRITERION IN INDEPENDENCE MODEL CONTEXT (Lebart and al. [10] ; Greenacre [6])

2.1. Let $D_I = \begin{pmatrix} P_{1.} & & \\ & \ddots & \\ & & P_{I.} \end{pmatrix}$ and $D_J = \begin{pmatrix} P_{.1} & & \\ & \ddots & \\ & & P_{.J} \end{pmatrix}$

and let P the matrix, IxJ, with element (i,j) equal to P_{ij}.
Let $X = D_I^{-1}(P - D_I \, \underset{\sim}{1}_I \, \underset{\sim}{1}'_J \, D_J) \, D_J^{-1}$ where $\underset{\sim}{1}'_I = (1,\dots,1)$ is the I-dimensional vector with its elements all unity.

The k order Correspondence Analysis of P can be defined as the search of the rank k matrix $\tilde{X}^{(k)}$ with I rows and J columns which minimizes :

$$\sum_{i\in I}\sum_{j\in J}(X_{ij}-\tilde{X}_{ij}^{(k)})^2\, P_{i.}\, P_{.j} \tag{1}$$

As a consequence of the singular values decomposition of $D_I^{1/2}\, X\, D_J^{1/2}$ we have

$$\tilde{X}_{ij}^{(k)} = \sum_{\alpha=1}^{k}\sqrt{\lambda_\alpha}\,\psi_{\alpha i}\,\phi_{\alpha j}$$

and $$\sum_{i\in I}\sum_{j\in J}(X_{ij}-\tilde{X}_{ij}^{(k)})^2\, P_{i.}\, P_{.j} = \sum_{\alpha=k+1}^{\nu}\lambda_\alpha$$

where $X\, D_J\, X'\, D_I\, \psi_\alpha = \lambda_\alpha\, \psi_\alpha$ with $\psi'_\alpha\, D_I\, \psi_\alpha = 1$

$X'D_I\, X\, D_J\ \ \phi_\alpha = \lambda_\alpha\, \phi_\alpha$ with $\phi'_\alpha\, D_J\, \phi_\alpha = 1$

and $\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_k$

ν is the rank of X, that is to say the minimum of (I-1) and (J-1). Moreover, the property $\underset{\sim}{1}'_I\, D_I\, X = X\, D_J\, \underset{\sim}{1}_J = 0$ implies $\underset{\sim}{1}'_I\, D_I\, \psi_\alpha = \underset{\sim}{1}'_J\, D_J\, \phi_\alpha = 0$.

Introducing the values of X_{ij} and $\tilde{X}_{ij}^{(k)}$ in (1) we get

$$\sum_{i\in I}\sum_{j\in J}\left(\left(\frac{P_{ij}}{P_{i.}P_{.j}}-1\right)-\sum_{\alpha=1}^{k}\sqrt{\lambda_\alpha}\,\psi_{\alpha i}\,\phi_{\alpha j}\right)^2 P_{i.}\, P_{.j} = \sum_{\alpha=k+1}^{\nu}\lambda_\alpha \tag{2}$$

with the particular cases

$$\sum_{i\in I}\sum_{j\in J}\left(\frac{P_{ij}}{P_{i.}P_{.j}}-1\right)^2 P_{i.}\, P_{.j} = \sum_{\alpha=1}^{\nu}\lambda_\alpha \tag{3}$$

and

$$\sum_{i\in I}\sum_{j\in J}\left(\left(\frac{P_{kj}}{P_{i.}P_{.j}}-1\right)-\sum_{\alpha=1}^{\nu}\sqrt{\lambda_\alpha}\,\psi_{\alpha i}\,\phi_{\alpha j}\right)^2 P_{i.}\, P_{.j} = 0 \tag{4}$$

In these formulas Correspondence Analysis appears as looking for the model

$$\left(\frac{P_{ij}}{P_{i.}P_{.j}}-1\right) = \sum_{\alpha=1}^{k}\sqrt{\lambda_\alpha}\,\psi_{\alpha 1}\,\phi_{\alpha j} \tag{5}$$

when the parameters λ_α, $\psi_{\alpha i}$, $\phi_{\alpha j}$ are solutions of a generalized least squares approximation in which item (i,j) receives weight $P_{i.}P_{.j}$.

2.2. Goodman [5] considers the model

$$P_{ij} = a_i\, b_j\, \exp\left[\sum_{\alpha=1}^{k}\sqrt{\lambda_\alpha}\,\psi_{\alpha 1}\,\phi_{\alpha j}\right] \tag{6}$$

where a_i and b_j must be positive numbers and the vectors ψ_α and ϕ_α must satisfy the equality

$$\underset{\sim}{1}'_I\, D_I\, \psi_\alpha = \underset{\sim}{1}'_J\, D_J\, \phi_\alpha = 0$$

Developping that constraint, we can choose

$$\log(a_i) = \sum_{j\in J} (\log(P_{ij}))\, P_{.j}$$

and $$\log(b_j) = \sum_{i\in I} (\log(P_{ij}))\, P_{i.} - \sum_{i\in I}\sum_{j\in J} (\log(P_{ij}))\, P_{i.}\, P_{.j}$$

Let us then consider the array Y defined by its elements :

$$Y_{ij} = \log(P_{ij}) - \log(a_i) - \log(b_j).$$

It can be shown that $\underset{\sim}{1}'_I\, D_I\, Y = Y\, D_J\, \underset{\sim}{1}_J = 0$

A k order least squares approximation of Y (Escoufier and al. [4]) by a rank k matrix $\tilde{Y}^{(k)}$ with I rows and J columns can be searched which minimizes

$$\sum_{i\in I}\sum_{j\in J} (Y_{ij} - Y_{ij}^{(k)})^2\, P_{i.}\, P_{.j}$$

The singular values decomposition of $D_I^{1/2}\, Y\, D_J^{1/2}$ will give us a generalized least squares solution for the log-bilinear model and we will get

$$\sum_{i\in I}\sum_{j\in J} ((\log(P_{ij})-\log(a_i)-\log(b_j)) - \sum_{\alpha=1}^{k} \sqrt{\lambda_\alpha}\, \psi_{\alpha i}\, \phi_{\alpha j})^2\, P_{i.}\, P_{.j} = \sum_{\alpha=k+1}^{\nu} \lambda_\alpha \quad (7)$$

2.3. The gaussian model studied by Ihm and al. [7] [8] can be studied in this least squares framework.

For any $I\in I$ and $j\in J$, we suppose

$$P_{ij} = a_i\, b_j \exp\left[- \frac{(\psi_i-\phi_j)^2}{2\,\sigma^2}\right] \quad (8)$$

It is evident that if the IxJ couples (ψ_i, ϕ_j) are solutions then for every constant c, the IxJ couples (ψ_i+c, ϕ_j+c) are solutions too. So it is possible to choose, for instance, a condition such as $\sum_{i\in I} \psi_i\, P_{i.} = \sum_{j\in J} \phi_j\, P_{.j} = 0$ (9)

From (8), we have

$$\mathrm{Log}(P_{ij}) = \mathrm{Log}(a_i) - \frac{\psi_i^2}{2\sigma^2} + \mathrm{Log}(b_j) - \frac{\phi_j^2}{2\sigma^2} + \frac{\psi_i\, \phi_j}{\sigma^2}$$

$$= A_i + B_j + \sqrt{\lambda}\, \psi_i\, \phi_j \quad (10)$$

The properties (9) give

$$\sum_{i\in I} \mathrm{Log}(P_{ij})\, P_{i.} = \sum_{i\in I} A_i P_{i.} + B_j$$

and $$\sum_{j\in J} \mathrm{Log}(P_{ij})\, P_{.j} = \sum_{j\in J} B_j P_{.j} + A_i$$

If the couples (A_i, B_j) are solutions then for every constant c, the couples (A_i+c, B_j-c) are solutions too. So we introduce the condition $\sum_{j\in J} B_j P_{.j} = 0$ which determines the A_i and B_j.

Let us now consider the array Z defined by its elements

$$Z_{ij} = \mathrm{Log}(P_{ij}) - A_i - B_j$$

It comes from the previous sections that if ψ and ϕ are respectively the first eigenvectors of $Z\,D_J\,Z'\,D_I$ and $Z'\,D_I\,Z\,D_J$ and λ the first eigenvalue, the matrix $\tilde{Z}^{(1)}$ defined by its elements $\tilde{Z}^{(1)}_{ij} = \sqrt{\lambda}\,\psi_i\,\phi_j$ is the first order least squares approximation of Z. We have :

$$\sum_{i\in I}\sum_{j\in J}(Z_{ij} - \sqrt{\lambda}\,\psi_i\,\phi_j)^2\,P_{i.}\,P_{.j} = \sum_{\alpha=2}^{\nu}\lambda_\alpha \tag{11}$$

It is evident that this approach can be extended to models

$$P_{ij} = a_i\,b_j\,\exp\left[-\sum_{\alpha=1}^{k}\frac{(\psi_{\alpha i}-\phi_{\alpha j})^2}{2\,\sigma_\alpha^2}\right]$$

3. LEAST SQUARES CRITERION FOR ASYMMETRIC MODELS.

3.1. In the previous sections, the observed P_{ij} or their logarithms are compared with models in which I and J play a symmetrical role. Lauro and al. [9] have introduced an asymmetric method with the aim to replace the χ^2 criterion (3) by the Goodman-Kruskal criterion

$$\sum_{i\in I}\sum_{j\in J}\left(\frac{P_{ij}}{P_{i.}} - P_{.j}\right)^2 P_{i.} = \sum_{\alpha=1}^{\nu}\lambda_\alpha \tag{12}$$

To do that, consider the array T with element $T_{ij} = \dfrac{P_{ij}}{P_{i.}} - P_{.j}$

It can be seen that $\underset{\sim}{1}'_I\,D_I\,T = T\,\underset{\sim}{1}_J = 0$ so that $\nu = \min\,((I-1),(J-1))$.

A k order least squares approximation of T is

$$\tilde{T}^{(k)} = \sum_{\alpha=1}^{k}\sqrt{\lambda_\alpha}\,\psi_{\alpha i}\,\phi_{\alpha j}$$

where $T\,T'\,D_I\,\psi_\alpha = \lambda_\alpha\,\psi_\alpha$ with $\psi'_\alpha\,D_I\,\psi_\alpha = 1$

$T'\,D_I\,Z\,\phi_\alpha = \lambda_\alpha\,\phi_\alpha$ with $\phi'_\alpha\,\phi_\alpha = 1$

and $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_k$

We have $\underset{\sim}{1}'_I\,D_I\,\psi_\alpha = \underset{\sim}{1}'_J\,\phi_\alpha = 0$ and

$$\sum_{i\in I}\sum_{j\in J}\left(\left(\frac{P_{ij}}{P_{i.}} - P_{.j}\right) - \sum_{\alpha=1}^{k}\sqrt{\lambda_\alpha}\,\psi_{\alpha i}\,\phi_{\alpha j}\right)^2 P_{i.} = \sum_{\alpha=k+1}^{\nu}\lambda_\alpha \tag{13}$$

3.2. Following this point of view, it is possible to consider an asymmetric log-model :

$$\frac{P_{ij}}{P_{i.}} = b_j\,\exp\left(\sum_{\alpha=1}^{k}\sqrt{\lambda_\alpha}\,\psi_{\alpha i}\,\phi_{\alpha j}\right) \tag{14}$$

with the constraints $\underset{\sim}{1}'_I\,D_I\,\psi_\alpha = \underset{\sim}{1}'_J\,\phi_\alpha = 0$

The calculus give $\log(b_j) = \sum_{i\in I} (\log(\frac{P_{ij}}{P_{i.}}))\, P_{i.}$

and we are led to look at the array U with elements

$$U_{ij} = \log(\frac{P_{ij}}{P_{i.}}) - \log(b_j)$$

In this case, we will get

$$\sum_{i\in I} \sum_{j\in J} ((\log(\frac{P_{ij}}{P_{i.}}) - \log(b_j)) - \sum_{\alpha=1}^{k} \sqrt{\lambda_\alpha}\, \psi_{\alpha i}\, \phi_{\alpha j})^2\, P_{i.} = \sum_{\alpha=k+1}^{\nu} \lambda_\alpha \qquad (15)$$

which is a generalized least squares solution for the asymmetric log-model in which item (i,j) receives the weights $P_{i.}$ whatever the value of j.

4. ALTERNATIVE CRITERIA FOR ANY MODELS : CLUSTERING APPROACH

4.1. The five above cases can be summarized in a sole framework. Let R and S be diagonal positive matrices with diagonal elements ($r_{i.}$; $i \in I$) and ($s_{.j}$; $j \in J$). Let A be a I×J matrix such that $\underset{\sim}{1}'_I\, R\, A = A\, S\, \underset{\sim}{1}_J = 0$. The rank k matrix $\tilde{A}^{(k)}$ which minimizes $\sum_{i\in I} \sum_{j\in J} (A_{ij} - A_{ij}^{(k)})^2\, r_{i.}\, s_{.j}$ is

$$\tilde{A}_{ij}^{(k)} = \sum_{\alpha=1}^{k} \sqrt{\lambda_\alpha}\, \psi_{\alpha i}\, \phi_{\alpha j} \qquad (16)$$

where

$$A\, S\, A'\, R\, \psi_\alpha = \lambda_\alpha\, \psi_\alpha \quad \text{with} \quad \psi'_\alpha\, R\, \psi_\alpha = 1 \qquad (17)$$

$$A'\, R\, A\, S\, \phi_\alpha = \lambda_\alpha\, \phi_\alpha \quad \text{with} \quad \phi'_\alpha\, S\, \phi_\alpha = 1 \qquad (18)$$

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_k \qquad k \leqslant \nu = \min((I-1),(J-1))$$

Rao [13] has shown that $\tilde{A}^{(k)}$ is the best approximation of A not only for the least squares criterion but also for all the measures of the discrepancy between A and $\tilde{A}^{(k)}$ which depend only of the singular values of $A-\tilde{A}^{(k)}$. In that sense, $\tilde{A}^{(k)}$ is a very good and important approximation of A. However, we cannot forget that others criteria can be conceived.

For instànce, it is well known that for the classical log-models approximation the criterion coming from the probabilistic environment is the maximum likelihood criterion. We can also consider non-probabilistic criteria such that

$$\min_{\tilde{A}^{(k)}} \sum_{i\in I} \sum_{j\in J} |A_{ij} - \tilde{A}_{ij}^{(k)}|\, P_{i.}\, P_{.j}$$

or

$$\min_{\tilde{A}^{(k)}} \max_{(i,j)} |A_{ij} - \tilde{A}_{ij}|\, P_{i.}\, P_{.j}$$

4.2. Going back to (17) and (18), it comes that

$$(\psi'_\alpha\, R\, A)\, S\, (A'\, R\, \psi_\alpha) = (\phi'_\alpha\, S\, A')\, R\, (A\, S\, \phi_\alpha) = \lambda_\alpha$$

or, in another writing :

$$\sum_{j\in J} s_{.j}\, (\psi'_\alpha\, R\, A^j)^2 = \sum_{i\in I} r_{i.}\, (A_i\, S\, \phi_\alpha)^2 = \lambda_\alpha \qquad (19)$$

where A^j and A_i are respectively the j^{th} column and the i^{th} row of A.

From (19), it can be deduced (Okamoto [11]) that whatever $(\phi_1^*, \phi_2^*, \ldots, \phi_k^*)$ such that $\phi_\alpha^{*'} S \phi_{\alpha'}^* = \delta_{\alpha\alpha'}$

$$\sum_{\alpha=1}^{k} \sum_{i\in I} r_{i.} (A_i S \phi_\alpha)^2 \geq \sum_{\alpha=1}^{k} \sum_{i\in I} r_{i.} (A_i S \phi_\alpha^*)^2 \qquad (20)$$

This optimality property can suggest a new family of criteria substituing $(A_i S \phi_\alpha)^2$ by any judicious function of ϕ_α, S and A_i.

We can use criteria based on the ranks of $\phi_{\alpha j}$ and A_{ij} or, removing the orthogonality condition $\phi'_\alpha S \phi_{\alpha'} = \delta_{\alpha\alpha'}$ we can use criteria suggested by projection pursuit works.

4.3. I propose to consider a partition approach in which we look for k vectors $\phi_\alpha \in R^J$ such that $\phi'_\alpha S \phi_\alpha = 1$ and kxI numbers $w_{\alpha i}$ with values 0 or 1 such that for every $i\in I$, $\sum_{\alpha=1}^{k} w_{\alpha i} = 1$ which maximize

$$\sum_{i\in I} \sum_{\alpha=1}^{k} w_{\alpha i} r_i (A_i S \phi_\alpha)^2 \qquad (21)$$

The unknown parameters ϕ_α and $w_{\alpha i}$ can be found as the solutions of the following reallocation procedure :

i) choose k arbitrary ϕ_α. For every $i\in I$, if $(A_i S \phi_\alpha)^2 = \max_{\alpha'} (A_i S \phi_{\alpha'})^2$ define $w_{\alpha i} = 1$ and $w_{\alpha' i} = 0$ for $\alpha' \neq \alpha$

ii) Let I_α be the set of $i\in I$ which realize $w_{\alpha i} = 1$. Let R the diagonal matrix, IxI, with element $(R_\alpha)_{ii} = r_{i.}$ if $i \in I_\alpha$ and zero if $i \notin I_\alpha$. Define ϕ_α as the first eigenvector of $A' R_\alpha A S$ and go back to i) with the new vectors $(\phi_1, \ldots, \phi_k)$.

Because of the optimality property of the first eigenvector, it can be shown that this algorithm give an increasing sequence of values $\sum_{i\in I} \sum_{j\in J} w_{\alpha i} r_{i.} (A_i S \phi_\alpha)^2$ with upper bound $\sum_{i\in I} r_{i.} (A_i S A'_{i.})$. So the algorithm will be stopped when two successive values would be sufficiently near.

Formula (16) leads to an interpretation of the $\sqrt{\lambda_\alpha}\, \phi_\alpha$ as typical components of A's rows. In Principal Component Analysis Rao [12] calls the $\sqrt{\lambda_\alpha}\, \phi_\alpha$ "typical points" of R^J. Any row A_i is approximated by a linear combination of rows $\sqrt{\lambda_\alpha}\, \phi'_\alpha$ in which the coefficients are the $\psi_{\alpha i}$. But $\phi'_\alpha S \phi_{\alpha'} = \delta_{\alpha\alpha'}$. So, the question is whether it is judicious in a pratical problem to think that a row can be viewed as the weighted sum of orthogonal ϕ_α. In the present approach, any row is attached to only one typical component ϕ_α ; the meaning of the results becomes more obvious.

4.3. Example (Artificial Data)

The following data array has been constructed to give a caricatured view of the differences between the least square approach and the clustering one.

Transposed data array

	Variable I									
Variable J	140	137	131	128	132	73	72	68	61	66
	115	114	88	84	86	111	112	116	87	86
	101	100	99	102	99	98	101	99	101	99

We consider the asymmetric model ($\frac{P_{ij}}{P_{i.}} - P_{.j}, r_{i.} = P_{i.}, s_{.j} = 1$). The least squares criterion leads to the following results :

eigenvalue	0.00875	0.00100
percentage	89.76	10.24
cumulative percentage	89.76	100.00

	Coordinates of the rows									
ψ_{1i}	0.07	0.06	0.10	0.09	0.10	-0.10	-0.10	-0.12	-0.11	-0.08
ψ_{2i}	0.04	0.04	-0.01	-0.03	-0.02	0.02	0.02	0.03	-0.05	-0.05

The graphical representations (figure 1) shows four groups of individuals. The groups 6,7,8 and 3,4,5 are the negative and positive aspect of a same behavior and this can be said also for the groups 9,10 and 1,2. It is obvious that the principal components are not typical existing behaviors. They are combinations of behaviors.

The clustering approach gives the two following axes (typical rows to be compared with the $\frac{P_{ij}}{P_{i.}} - P_{.j}$) with $\cos(\phi_1, \phi_2) = - 0.7698$

ϕ_1	-0.7750	0.6100	0.1651
ϕ_2	0.7606	-0.1232	-0.6374

The coordinates of the rows (orthogonal projections on the subspace of R^3 spanned by ϕ_1 and ϕ_2) on these axes are :

0.00	0.00	-0.09	-0.11	-0.10	0.10	0.10	0.13	0.01	-0.00
0.08	0.08	0.01	-0.2	0.01	0.00	-0.01	0.01	-0.11	-0.10

The graphical representations (figure 2) obviously shows that the method has recognized typical behaviors.

Figure 1 (least squares criterion)

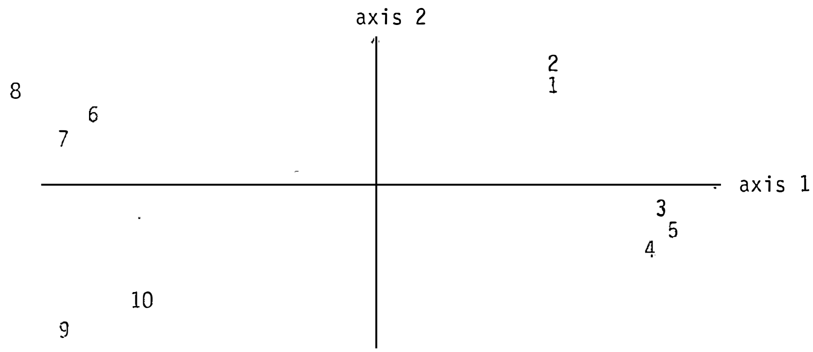
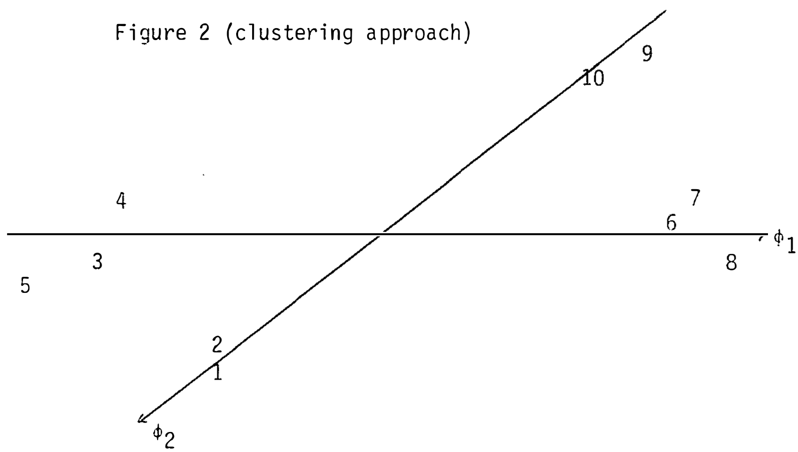


Figure 2 (clustering approach)



5. BEYOND THE P_{ij}

5.1. Suppose now that we have three qualitative variables studied on the same n individuals provided with weights given by the diagonal of a nxn diagonal positive matrix C. We suppose that $\underset{\sim}{1}_n' D \underset{\sim}{1}_n = 1$. Let X, Y and Z be the three dummy matrices nxI, nxJ, nxK associated with the three variables. We introduce the following notations.

$$D_I = X'D\,X \quad ; \quad D_J = Y'\,D\,Y \quad ; \quad D_K = Z'\,D\,Z$$

and $$P_{IJ} = X'D\,Y \; ; \quad P_{IK} = X'D\,Z \quad ; \quad P_{JK} = Y'D\,Z.$$

$P_{ij.}$ will denote item (i,j) of P_{IJ} and $P_{i..}$ item (i,i) of D_I. We will use the analogous $P_{i.k}$ for P_{IK}, $P_{.jk}$ for P_{JK}, $P_{.j.}$ for D_J and $P_{..k}$ for D_K.

Let $\pi_Z = Z(Z'DZ)^{-1} Z'D$ the D-orthogonal projector onto the subspace S(Z) of R^n spanned by the columns vectors of Z.

$Q_Z = \mathcal{I}_{nxn} - \pi_Z$ is the D-orthogonal projector onto $S(Z)^{\perp D}$ and we have $\pi_Z' \, D \, Q_Z = 0$.

5.2. The projections of X and Y onto S(Z) are $\pi_Z(X) = Z \, D_K^{-1} \, P_{KI}$ and $\pi_Z(Y) = Z \, D_K^{-1} \, P_{KJ}$. Following Daudin [1], we will denote by M_{IJ} the IxJ matrix $(\pi_Z(X))'D(\pi_Z(Y)) = P_{IK} \, D_K^{-1} \, P_{KJ}$. It can be easily verified that

$M_{ij} = \sum_{k\in K} \frac{P_{i.k}P_{.jk}}{P_{..k}}$. M_{IJ} represents the part of the link between X and Y which is explained by Z.

We see that $\sum_{i\in I} M_{ij} = P_{.j.}$ and $\sum_{j\in J} M_{ij} = P_{i..}$

So following section 2.1, we consider the matrix A with element

$A_{ij} = \frac{M_{ij}}{P_{i..} P_{.j.}} - 1$ which realises the identities $A \, D_J \, 1_J = 1_I' \, D_I \, A = 0$.

The k order Correspondence Analysis of M will provide us with the best least squares approximation of A by a matrix of rank k and we will be able to study the discrepancy between the part of the link of X and Y which is explained by Z and the independence model for X and Y.

5.3. Looking now to the projections of X and Y onto $S(X)^{\perp D}$, we obtain $(Q_Z(X))'D \, (Q_Z(Y)) = P_{IJ} - M_{IJ}$. Matrices P_{IJ} and M_{IJ} have the same margins D_I and D_J. So following the work made by Escofier [2] on Correspondence Analysis with respect to a model (see also Escoufier [3]) we consider the matrix A with element $A_{ij} = \frac{P_{ij}-M_{ij}}{P_{i..} P_{.j.}}$.

We know that $\tilde{A}^{(k)}$ which minimizes $\sum_{i\in I} \sum_{j\in J} (A_{ij}-A_{ij}^{(k)})^2 \, P_{i..} \, P_{.j.}$ is given by $\tilde{A}_{ij}^{(k)} = \sum_{\alpha=1}^{k} \sqrt{\lambda_\alpha} \, \psi_{\alpha i} \, \phi_{\alpha j}$ where

$$A \, D_J A' \, D_I \, \psi_\alpha = D_I^{-1}(P_{IJ}-M_{IJ}) \, D_J^{-1}(P_{IJ}-M_{IJ})' \, \psi_\alpha = \lambda_\alpha \, \psi_\alpha \tag{22}$$

and

$$A' D_I A \, D_J \, \phi_\alpha = D_J^{-1}(P_{IJ}-M_{IJ})' \, D_I^{-1}(P_{IJ}-M_{IJ}) \, \phi_\alpha = \lambda_\alpha \, \phi_\alpha$$

Daudin [1] leads us to consider a matrix M^* with element $P_{i..} P_{.j.} + (P_{ij}-M_{ij})$, the margins of which are D_I and D_J. So following section 2.1, we will look to $\frac{P_{i..} P_{.j.} + (P_{ij}-M_{ij})}{P_{i..} P_{.j.}} - 1 = A_{ij}$ and Daudin and Escofier do the same study.

Yanai [14] reminds that, for two qualitative variables, Corrspondence Analysis is a Canonical Correlation Analysis for the two dummy matrices. He proposes to consider the Canonical Correlation Analysis of $Q_Z(X)$ and $Q_Z(Y)$ and call this technic "Partial Correspondence Analysis". In this approach, the λ_α , ψ_α and ϕ_α are given by the usual equations. For instance the ψ_α verify :

$$[(Q_Z(X))'D(Q_Z(X))]^{-1}(P_{IJ}-M_{IJ})\ [(Q_Z(Y))'D(Q_Z(Y))]^{-1}\ (P_{IJ}-M_{IJ})'\ \psi_\alpha = \lambda_\alpha\ \psi_\alpha \quad (23)$$

Going back to (22) we see that, as quoted by Daudin, (22) is in fact a simplication of (23).

Remark : The approach presented in section 4 can be considered for exploring matrices M_{IJ} and $P_{IJ}-M_{IJ}$. The log-model can be used for M_{IJ} but it is not easy to use for $P_{IJ}-M_{IJ}$ and for M^*_{IJ} because these matrices can have negative items.

6. REFERENCES

[1] Daudin, J.C., Analyse factorielle des dépendances partielles, Rev. Stat. Appli., XXIX n° 2 (1981) 15-29.

[2] Escofier, B., Analyse factorielle en référence à un modèle. Application à l'analyse de tableaux d'échanges. R.T. 337, INRIA, Domaine de Voluceau, Rocquencourt, BP. 105, 78153 Le Chesnay, France (1984).

[3] Escoufier, Y., L'analyse des Correspondances : ses propriétés et ses extensions, Proceedings of the 45th Session ISI, 1985, 28.2.1-28.2.16.

[4] Escoufier, Y. and JUNCA, S., Least-squares approximation of frequencies or their logarithms, Int. Stat. Rev., 54 n°3 (1986) 279-283.

[5] Goodman, L., Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables, Int. Stat. Rev., 54 n° 3 (1986) 243-278.

[6] Greenacre, M.J., Theory and applications of correspondence analysis. Academic Press, 1984.

[7] Ihm, P. and Van Groenewoud, H., A multivariate ordering of vegetation data based on gaussian type gradiant response curves, J. Ecol., 63 (1975) 767-778.

[8] Ihm, P. and Van Groenewoud, H., Correspondence Analysis and Gaussian Ordination, Compstat Lectures 3. J.M. Chambers and al. (eds), Physica Verlag, Vienna, 1984, 5-60.

[9] Lauro, N., D'Ambra, L., L'analyse non symétrique des correspondances, in Data Analysis and Informatics III, E. Diday et al., (eds), Elsevia Science Publishers B.V. North-Holland, 1983, pp. 433-446.

[10] Lebart, L., Morineau, A., Warwick, Multivariable descriptive statistical analysis. John Wiley, 1984.

[11] Okamoto, M., Optimality of principals components, in Multivariate Analysis II, P.K. Krishnaiah, (eds.). Academic Press, 1969, 673-685.

[12] Rao, R., The use and interpretation of principal component analysis in applied research, Sankhya A, 26 (1965) 329-358.

[13] Rao, C.R., Matrix approximation and reduction of dimensionality in Multivariate Statistical Analysis, in Multivariate Analysis 5, P.R. Krishnaiah, (eds.), North-Holland, Amsterdam, 1980, pp. 3-22.

[14] Yanai, H., Partial Correspondence Analysis and its properties. Japanese-French Seminar on "Recent Developments in Clustering and Data Analysis". Academic Press, Boston, 1987 (to be published).