E. JOLIVET et al., éd.
Biométrie et Temps (1980) 59-76

L'ANALYSE CONJOINTE DE PLUSIEURS MATRICES DE DONNEES.

ESCOUFIER Yves

Centre de Recherche en Informatique de Gestion,
34075 MONTPELLIER CEDEX.

RESUME

On étudie les différentes techniques bien connues pour l'analyse de plusieurs matrices de données ; elles s'appliquent en particulier quand chaque matrice correspond à un instant.

On reconnaît en général trois étapes : représentation globale, recherche d'un compromis et représentation détaillée.

Les limites de l'unification sont liées à la nature des données.

SUMMARY

Different well known approaches to the analysis of several data matrices are studied ; they can be suitable, of course, in the case of time as index. Three stages may be recognized : global representation, compromise solution and detailed representation. The limits of this unified presentation are due to the kind of data.

I.- INTRODUCTION

1.1. Le but de ce travail est de présenter de manière unifiée différentes techniques proposées dans la littérature statistique française récente pour l'analyse conjointe de plusieurs matrices de données. L'unification des techniques est faite au travers des grandes étapes qui les constituent. On reconnaîtra en général une étape de représentation globale (interstructure), une étape de recherche de compromis, une étape de représentation détaillée (intrastructures). On verra que les limites de l'unification sont dues aux différences de nature des différentes données dont l'analyse est envisagée.
Pour être compris, ce travail devrait être lu en confrontation constante avec les références [1], [4], [7], [8], [9], [10], [12], [13], dont il n'est qu'un commentaire.
En particulier, on ne trouvera ici aucun exemple. Le lecteur devra se rapporter aux références ci-dessus pour en trouver. Cette absence d'exemples et bien sûr regrettable, mais elle est due à la nature de la présentation : un seul exemple aurait nécessairement péché par son manque de généralité ; plusieurs exemples posaient un problème de place. Cette imperfection devra être une incitation supplémentaire à lire les références ci-dessus.
La compréhension du texte sera également facilitée par une bonne connaissance des notations assez généralement employées depuis le livre de CAILLIEZ et PAGES [3] .

1.2. Les différentes situations envisagées dans ce travail sont les suivantes :

Situation 1 - K matrices de données quantitatives concernant les mêmes n individus.
Le point de départ est constitué de K matrices $\{ X_k/k = 1, \dots , K \}$ de données quantitatives centrées pour des poids rangés dans la diagonale d'une matrice D. La matrice X_k est de dimension $p_k \times n$. Chacune des matrices traduit un point de vue sur l'ensemble des individus. Ce point de vue est caractérisé par
$W_k D = X'_k X_k D$, opérateur associé à l'étude (X_k, I_k, D) où I_k est la matrice identité d'ordre k .

Situation 2 - K matrices de données quantitatives résultant de l'observation des mêmes variables.
Dans cette situation, les matrices centrées X_k sont de dimensions $p \times n_k$. Si D_k est la matrice diagonale des poids qui ont permis de centrer X_k, $S_k = X_k D_k X'_k$ est la matrice des variances et covariances associée à l'étude (X_k, I_p, D_k)

Situation 3 - K matrices de données quantitatives résultant de l'observation des mêmes variables sur les mêmes individus. Cette situation englobe les deux situations précédentes. A l'étude (X_k, I_p, D) peuvent être associés $W_k D$ et S_k. La comparaison des études peut donc être envisagée soit au travers de la comparaison des $W_k D$, soit au travers de celle des S_k.

Situation 4 - K variables qualitatives concernant les mêmes individus.
Soit U_k, $p_k \times n$, la matrice des variables indicatrices de la variable k. On pose $D_k = U_k D U'_k$, matrice diagonale des poids des modalités de la variable k.
Si $U^*_k = U_k (I - D \underset{\sim}{1} \underset{\sim}{1}')$ où $\underset{\sim}{1}$ est le vecteur $n \times 1$ dont toutes les composantes sont égales à l'unité, $A_k D = U^{*'}_k D_k U^*_k D$ est le projecteur associé à la variable k. $A_k D$ caractérise la variable k.

Situation 5 - K tableaux de contingence concernant les deux mêmes variables qualitatives.
X_k est le tableau $p \times q$ de contingence pour les deux variables associé à l'indice k. Soient $D_p(k)$ et $D_q(k)$, les matrices diagonales des poids marginaux déduits de X_k. On considère les tableaux de BURT :

$$B_k = \left|\begin{array}{c|c} D_p(k) & X_k \\ \hline X'_k & D_q(k) \end{array}\right|$$

La comparaison des B_k permettra celle des X_k.

<u>Remarque 1</u> : On peut envisager également le cas de K matrices de similarités $\{ C_k \ / \ k = 1, \ldots, K \}$

Sous réserve de substituer aux C_k des matrices C_k^* qui en sont de bonnes approximations semi-définies positives, on se ramène à la situation 1 avec $W_k \, D = C_k^* \, D$.

I.- ETAPE DE L'INTERSTRUCTURE

2.1. L'idée directrice dans cette étape est de construire une matrice C, $K \times K$, de produits scalaires entre les objets caractéristiques, que ces objets soient des opérateurs, des matrices de variances et covariances, des projecteurs on des tableaux de BURT. Les vecteurs propres de C permettent une représentation euclidienne des K objets dans laquelle la proximité de deux points est significative de la proximité des objets au sens de la métrique induite par le produit scalaire considéré.
Les produits scalaires sont systématiquement construits à partir de l'opérateur " trace ", extension naturelle de la métrique usuelle aux matrices.
On distingue des produits scalaires de type " Covv " et des produits scalaires de type " Rv ". Ces derniers ont la propriété de considérer comme identiques des objets proportionnels.

	produits scalaires Type " Covv "	Type " Rv "
Situation 1	$Tr\ (W_k\ D\ W_l\ D)$	$\dfrac{Tr\ (W_k\ D\ W_l\ D)}{\sqrt{Tr\ (W_kD)^2\ Tr(W_lD)^2}}$
Situation 2	$Tr\ (S_k\ S_l)$	$\dfrac{Tr\ (S_k\ S_l)}{\sqrt{Tr\ (S_k)^2\ Tr\ (S_l)^2}}$
Situation 4	$Tr\ (A_kD\ A_lD)$ $= \dfrac{\chi^2_{kl}}{n}$	$\dfrac{Tr\ (A_kD\ A_l\ D)}{\sqrt{Tr\ (A_kD)^2\ Tr\ (A_lD)^2}}$ $= \dfrac{\chi^2_{kl}}{n\ \sqrt{(p-1)\ (q-1)}}$
Situation 5	$Tr\ (B_k\ B_l)$	

Pour la situation 3 , on peut choisir de comparer les études au travers des W_kD ou des S_k . Un calcul simple montre que les deux matrices C ne sont pas les mêmes et, à ce jour, on ne connaît pas les relations qui existent entre les deux approches. Pour ce cas, JAFFRENOU [10] propose un produit scalaire qui dans la situation la plus simple est $Tr\ (X_k X'_1)$.

Pour lui la matrice C ne dépend donc pas d'une approche privilégiant les individus (comparaison aux travers des W_kD) ou les variables (comparaison aux travers des S_k).

2.2. L'interprétation qui est faite de la représentation fournie par les vecteurs propres de C doit bien sûr faire référence au produit scalaire utilisé.
Les produits scalaires de type Rv confondent les objets proportionnels puisqu'ils reviennent pratiquement à considérer les objets $W_kD \,/\, \|W_kD\|$, $S_k \,/\, \|S_k\|$ et $A_k \,/\, \|A_k\|$ qui sont de norme unité.
Les distances des points à l'origine s'interprètent alors exactement comme sur un cercle des corrélations dans une analyse en composantes principales sur matrice de corrélations : une étude est bien représentée si son point représentatif est proche du cercle de rayon unité ; les angles entre les vecteurs joignant les points à l'origine ont des cosinus qui, dans le cas de représentation parfaite sont égaux aux coefficients Rv .
Pour les produits scalaires de type Covv , la représentation obtenue est analogue à celle des variables dans une analyse en composantes principales sur matrice des covariances. L'interprétation est à faire en fonction des normes des vecteurs issus de l'origine et ayant pour extrémité les points représentant les études. On reconnaîtra des études associées à une grande norme, des études associées à une petite norme. L'interprétation est à faire également en angle, en remarquant que si la colinéarité des vecteurs est représentative de la proportionnalité des objets qu'ils représentent, l'écart entre deux vecteurs est synonyme d'un écart à la proportionnalité.
On comprend mieux de telle représentation si on a bien en mémoire, qu'utiliser le produit scalaire $Tr\ (W_kD\ W_1D)$ par exemple est équivalent à considérer la distance de carré $Tr\ (W_kD - W_1D)^2$.

Pour le cas particulier évoqué précédemment où le produit scalaire est $Tr\ (X_k X_l')$, on voit que la distance est $Tr\ (X_k - X_l)\ (X_k - X_l)'$ ce qui veut dire que la distance est nulle quand $X_k = X_l$ et que les vecteurs sont colinéaires quand $X_k = Cte\ X_l$.

<u>Remarque 2</u> : La matrice de produits scalaires C peut être soumise à d'autres études que celle évoquée dans ce paragraphe. Elle peut être point de départ pour des classifications hiérarchiques ou non hiérarchiques, pour des programmes de représentation non métrique (Indscal). L'intérêt de l'approche évoquée ici est de conduire naturellement à la définition du compromis qui fait l'objet du paragraphe suivant.

<u>Remarque 3</u> : Soit D_k une matrice diagonale de poids associés aux différentes études. On note encore $\underset{\sim}{1}$, le vecteur $K \times 1$ dont toutes les composantes sont égales à l'unité. La matrice centrée en ligne et en colonne $(I_k - \underset{\sim}{1}\,\underset{\sim}{1}'D_k)\ C\ (I_k - D_k\underset{\sim}{1}\,\underset{\sim}{1}')$ engendre les mêmes distances que C et fournit des représentations centrées des K études, c'est à dire des représentations analogues aux représentations des individus en analyse en composantes principales. Ces représentations méritent d'être regardées.

- ETAPE DU COMPROMIS

Dans cette étape, l'objectif est de trouver un objet fictif (ou éventuellement plusieurs), de même nature que les objets étudiés et qui ait de bonnes propriétés de compromis par rapport à l'ensemble des objets. Nous écrirons ce paragraphe pour les objets $\{W_k D \ / \ k = 1, \ldots, K\}$ mais tout ce qui est dit s'applique sans changement aux autres types d'objets à l'exception des $\{B_k \ / \ k = 1, \ldots, K\}$ qui seront évoqués dans la remarque 6.

Soit L, le vecteur propre de C associé à la plus grande valeur propre que nous noterons δ. On suppose que $L' = (L_1, \ldots, L_k)$ et $L'L = 1$.

Théorème 1 : Soit α un vecteur réel tel que $\alpha' = (\alpha_1, \ldots, \alpha_k)$ et $\alpha'\alpha = 1$. On pose $WD = \sum_{k=1}^{K} L_k W_k D$ et $W_\alpha D = \sum_{k=1}^{K} \alpha_k W_k D$. Alors :

i) $$\sum_{k=1}^{K} \left[Tr\ (W_\alpha D\ W_k D) \right]^2 \leqslant \sum_{k=1}^{K} \left[Tr\ (W\ D\ W_k D) \right]^2 = \delta^2$$

ii) $$Tr\ (W_\alpha D)^2 \leqslant Tr\ (W\ D)^2 = \delta$$

Démonstration :

i) $$\sum_{k=1}^{K} \left[Tr\ (W_\alpha D\ W_k D) \right]^2 = \sum_{K=1}^{K} \left[\sum_{l=1}^{K} \alpha_l\ Tr\ (W_l D\ W_k D) \right]^2$$

$$= \sum_{k=1}^{K} \sum_{l_1=1}^{K} \sum_{l_2=1}^{K} \alpha_{l_1}\ \alpha_{l_2}\ Tr(W_{l_1} D\ W_k D)\ Tr\ (W_{l_2} D\ W_k D)$$

$$= \sum_{l_1=1}^{K} \sum_{l_2=1}^{K} \alpha_{l_1} \alpha_{l_2}\ (C^2)_{l_1 l_2}$$

où $(C^2)_{l_1 l_2}$ est l'élément (l_1, l_2) du carré de la matrice C.

Le maximum sera atteint en prenant $\alpha = L$, premier vecteur propre de C et donc de C^2. La valeur du maximum est la valeur de la première valeur propre de C^2 donc δ^2 .

ii) $Tr\,(W_\alpha D)^2 = Tr\,(\sum_{l=1}^{K} \alpha_l\, W_l\, D)^2$

$$= \sum_{l_1=1}^{K} \sum_{l_2=1}^{K} \alpha_{l_1}\, \alpha_{l_2}\, T_r(W_{l_1}\, D\, W_{l_2}\, D)$$

$$= \sum_{l_1=1}^{K} \sum_{l_2=1}^{K} \alpha_{l_1}\, \alpha_{l_2}\, C_{l_1\, l_2}$$

Le maximum est donc atteint pour $\alpha = L$ et vaut δ .

La D- symétrie des $W_k\,D$ permet d'assurer que tous les éléments de C sont positifs. Il en résulte que les composantes de L sont toutes de même signe et peuvent donc être choisies positives. Il s'ensuit que W est une combinaison linéaire positive de matrices semi-définies positives. W est donc semi-définie positive. Il en résulte que $W\,D$ est de même nature que les $W_k D$ et peut être regardé comme l'opérateur d'une étude fictive. La matrice Y dont les lignes sont les vecteurs propres de $W\,D$ fournira les coordonnées d'une représentation des n individus qui est un compromis entre les représentations fournies par les différentes études. $W\,D$ est " le meilleur " compromis au sens des énoncés i) et ii) du théorème

<u>Théorème 2</u> : Soit α un vecteur réel tel que

$$\alpha' = (\alpha_1, \ldots, \alpha_K) \quad \text{et} \quad \sum_{i=1}^{n} \alpha_i = 1$$

On définit $\overline{W}\,D = \frac{1}{K} \sum_{k=1}^{K} W_i\, D$ et $W_\alpha\, D = \sum_{k=1}^{K} \alpha_k\, W_k\, D$

Alors $\sum_{k=1}^{K} Tr\,(\overline{W}\,D - W_k D)^2 \leqslant \sum_{k=1}^{K} Tr\,(W_\alpha\, D - W_k\, D)^2$

Démonstration - On peut écrire :

$$\sum_{k=1}^{K} \mathrm{Tr}\,(W_\alpha D - W_k D)^2 = \sum_{k=1}^{K} \Big[\mathrm{Tr}\,(W_\alpha D - \overline{W} D)^2 + \mathrm{Tr}\,(\overline{W} D - W_k D)^2$$

$$- 2\,\mathrm{Tr}\,(W_\alpha D - \overline{W} D)\,(\overline{W} D - W_k D)\Big]$$

$$= \sum_{k=1}^{K} \mathrm{Tr}\,(W_\alpha D - \overline{W} D)^2 + \sum_{k=1}^{K} \mathrm{Tr}(\overline{W} D - W_k D)^2$$

Le résultat s'obtient en remarquant que $\overline{W} D - W_\alpha D$ est D - symétrique ; ses valeurs propres sont donc réelles ce qui entraîne $\mathrm{Tr}\,(W_\alpha D - \overline{W} D)^2 \geq 0$

$\overline{W} D$ est donc un autre compromis possible, " le meilleur " au sens du théorème 2.

Remarque 4 : Les vecteurs propres de C autres que L ont nécessairement des composantes négatives puisque L n'en a pas. Les combinaisons linéaires qu'ils permettent de définir ne sont donc pas en général définies positives et ne conduisent pas à des compromis utilisables.

Remarque 5 : La matrice C mise en place sur la base du produit scalaire $\mathrm{Tr}\,(X_k X'_l)$ n'a aucune raison d'avoir des éléments tous positifs. Ses vecteurs propres ont donc des composantes de signes quelconques. Ceci ne gêne en rien ici car toutes les combinaisons linéaires des matrices X_k ont la même nature que les matrices X_k : " elles sont $p \times n$ ". Dans ce cas, les vecteurs propres de C fournissent donc une hiérarchie de compromis.

Remarque 6 : Travaillant sur les B_k , FOUCARD est guidé par le souci d'avoir un compromis qui puisse être considéré comme un tableau de BURT.

Il pose $\quad L^* = L \,/ \sum_{k=1}^{K} L_k$

Et choisit $\quad B = \sum_{k=1}^{K} L^*_k\, B_k$

Ce compromis n'a aucune des propriétés qui font l'objet des théorèmes 1 et 2 . Il reste intéressant car il pondère de manière non uniforme les B_k : de même qu'un point aberrant intervient peu dans une estimation robuste de la moyenne d'une distribution, un B_k aberrant (= très différent de l'ensemble des autres B_k) intervient peu dans la définition de B .

Ayant choisi un compromis, nous déduisons de ses vecteurs propres une représentation unique des individus dans les cas où nous travaillons avec les W_kD et les A_kD ; une représentation unique des variables lorsque nous travaillons sur les S_k . L'analyse des correspondances de B fournira une représentation unique des modalités des variables lorsque nous travaillons sur les B_k .

IV.- ETAPE DES INTRASTRUCTURES

C'est dans cette étape que la nature des données traitées intervient le plus, jusqu'à rendre impossible une présentation qui puisse être unique pour l'ensemble des situations. Nous sommes donc conduits à étudier cette étape en référence à chacune des situations envisagées.

Situation 1 -

Au point où nous en sommes, nous disposons des K visions des n individus tels qu'ils sont vus par les K études et de la vision du compromis traduite par la matrice Y dont les lignes sont les vecteurs propres de $W\,D$. Pour comparer ces différentes visions, il est logique de chercher à les représenter toutes dans un même espace. Pour cela, on calcule les matrices

$$\{\hat{Y}_k = Y_k\, D\, Y'(Y\, D\, Y')^{-1}\, Y \ /\ k = 1, \ldots , K\}$$

$\hat{Y}_k$ est l'approximation au sens des moindres carrés de Y_k fournie par Y. Se limitant aux deux premières lignes de Y, on peut représenter les n individus tels qu'ils sont vus (à peu près) par le compromis. En superposant les différentes représentations que fournissent les deux premières lignes des $\hat{Y}_k$, on peut comparer visuellement les différentes visions des individus.

On a donc ici une représentation de $(K + 1) \times n$ points dans laquelle chaque individu est représenté une fois pour le compromis et une fois pour chacune des K études. Lorsque les études sont indicées par le temps, cette représentation permet d'étudier les trajectoires suivies par les individus.

On peut faire plus, en s'appuyant sur le fait que les lignes de Y sont les composantes principales de l'opérateur compromis. Ceci conduit à réaliser la représentation des X_k et des Y_k sur le principe qui permet de représenter des variables supplémentaires en analyse en composantes principales.

On calcule donc $X_k\, D\, Y'(Y\, D\, Y')^{-1/2}$

et $Y_k\, D\, Y'(Y\, D\, Y')^{-1/2}$

On a ainsi un moyen de visualiser les variables ce qui permet de nommer les Y comme en composantes principales. La visualisation des Y_k permet éventuellement de détecter des interversions dans l'ordre des vecteurs propres de l'étude k par rapport à l'ordre du compromis.

Remarque 7 : Dans le traitement de K matrices de similarités (cf. Remarque 1), tout ce qui précède reste vrai pour les matrices Y_k dont les lignes sont les vecteurs propres des $C_k^* D$.

Situation 2 -

Le traitement de K matrices S_k de variances et covariances des mêmes p variables peut être considéré comme un cas particulier de celui de K matrices de similarités. Les variables jouent ici le rôle des individus. Y permet la représentation compromis des variables.

Les $\hat{Y}_k$ permettent de superposer les représentations des variables telles qu'elles sont vues au travers de chaque S_k .

Situation 3 -

Ici les variables et les individus sont les mêmes tout au long des études. On peut donc considérer cette situation soit comme un cas particulier de la situation 1, soit comme un cas particulier de la situation 2.
Dans la première hypothèse, toutes les représentations se font sur des bases bien connues du praticien de l'analyse des données : projections orthogonales pour les individus, variables supplémentaires. La difficulté vient dans la pratique du nombre n des individus puisque le compromis est de dimension $n \times n$, ce qui pose des problèmes de réalisation informatique non négligeable.
On pourrait pour éviter cette difficulté, tenter d'expliciter le fait que $W D$ est l'opérateur associé à l'étude de matrice transposée

$$(\sqrt{L_1}\, X'_1 \,|\, \sqrt{L_2}\, X'_2 \,|\, \ldots\ldots \,|\, \sqrt{L_K}\, X'_K)$$

Ceci n'a pas été fait jusqu'ici.

Dans la seconde hypothèse, le compromis, disons S, est de dimension $p \times p$. On évite donc les difficultés de réalisation informatique. La représentation des variables est fondée sur les outils habituels de l'analyse des données. Pour représenter les individus nous calculons les matrices $Y X_k$; chaque individu est donc représenté autant de fois qu'il y a d'études. Les propriétés mathématiques de cette représentation sont encore mal connues.

Remarque 8 -

Ne cachons pas que si les deux hypothèses conduisent à des représentations des variables et des individus, nous ne savons pas encore s'il existe des liens entre ces représentations

Remarque 9 -

Dans l'approche décrite dans (10), un compromis est une matrice X, $p \times n$. On peut donc représenter chacune des variables de la matrice X_k comme variable supplémentaire et obtenir une représentation des individus par la formule

$$\hat{Y}_k = Y_k D Y (Y' D Y)^{-1} Y$$

où Y est la matrice des composantes principales de X et Y_k celle de X_k.

Situation 4

Dans l'étude qu'ils ont faite de cette situation, CAZES et Coll.[4] ont remarqué que la représentation des individus fournie par la diagonalisation du compromis était la même que celle de l'analyse des correspondances du tableau

$$n \times (\sum_{k-1}^{K} P_k) \qquad (\sqrt{L_1}\, U'_1 \,|\, \sqrt{L_2}\, U'_2 \,|\, ----- \,|\, \sqrt{L_K}\, U'_K)$$

Cette remarque leur permet d'obtenir en même temps que la représentation (compromis) des individus, la représentation des modalités des différentes variables. Ces auteurs n'ont pas envisagé la représentation des individus tels qu'ils sont vus par les différentes études.

Situation 5

Pour FOUCARD [7] , le compromis a consiste en un tableau de BURT :

$$B = \left(\begin{array}{c|c} D_p & X \\ \hline X' & D_q \end{array}\right)$$

La représentation compromis des modalités des deux variables est fournie par l'analyse des correspondances de B ou plus simplement de X . La représentation des modalités telles qu'elles sont vues dans les différentes études est alors obtenue, dans l'analyse des correspondances de X , en projetant en éléments supplémentaires les lois conditionnelles de chacun des tableaux de contingence X_k .

V.- CONCLUSION

Ce survol rapide de différentes méthodes proposées récemment pour l'analyse conjointe de plusieurs matrices montre sans aucun doute que les statisticiens ont acquis dans les dernières années de nouveaux outils puissants.

Nous avons souligné dans le texte des points qui méritent encore éclaircissement. Remarquons également que lorsque les différentes matrices correspondent à des instants d'observations différents, les résultats fournis ne dépendent en aucune manière de la chronologie des observations. Le temps sera un outil d'interprétation, rien de plus. Dans sa thèse, BOUROCHE [2] a essayé d'utiliser le temps plus explicitement en le faisant intervenir sous la forme d'un effet moyen qui pourra être extrait de chacune des matrices. Bien que limitée (effet moyen, mêmes variables et mêmes individus, nature algorithmique des solutions) sa proposition reste intéressante quant à l'idée de faire intervenir le temps à un autre stade que celui de l'interprétation finale. Nous ne pouvons à ce jour que constater qu'il y a là un problème ouvert.

REFERENCES

(1) BERNARD M.L. , DIAZ LLANOS F.J. , ESCOUFIER Y. - (1979) - La méthode Statis : une application à l'évolution des campagnes Languedociennes.
C.R.I.G. Av. D'Occitanie - 34 075 Montpellier Cedex.

(2) BOUROCHE J.M. (1975) - *Analyse des données ternaires : la double analyse en composantes principales* -
Thèse Université Pierre et Marie Curie - Paris VI

(3) CAILLIEZ F. , PAGES J.P. - (1976) - *Introduction à l'Analyse des données.*
SMASH - 9 Rue Duban - 75 016 Paris.

(4) CAZES P. , BONNEFOUS S. , BAUMERDER A. , PAGES J.P (1976) - Description cohérente des variables qualitatives prises globalement et de leurs modalités.
S.A.D. 2, 48-62.

(5) ESCOUFIER Y. - (1976) - Opérateur associé à un tableau de données.
Annales de l'INSEE n° 22-23, 165-179 .

(6) ESCOUFIER Y. - (1979) - *Cours d'Analyse des Données* -
C.R.I.G. Av. D'Occitanie - 34 075 Montpellier Cedex.

(7) FOUCARD T. - (1978) - Sur les suites de tableaux contingence indexés par le temps.
S.A.D. - 2, 67-85.

(8) FOUCARD T. - (1979) - Prévision d'une suite de tableaux de probabilité. Ajustement à des marges données.
S.A.D. 2, 51-71.

(9) FOUCARD T. - (1979) - *Structure des tableaux de probabilités. Description et Prévision.*
Thèse Université des Sciences et Techniques du Languedoc Montpellier.

(10) JAFFRENOU P.A. - (1978) - *Sur l'Analyse des familles finies de variables vectorielles.*
Thèse - Université Claude Bernard - Lyon I.

(11) L'HERMIER DES PLANTES H. - (1976) - *Structuration des tableaux à trois indices de la Statistique.*
Thèse - Université des Sciences et Techniques du Languedoc - Montpellier II.

(12) L'HERMIER DES PLANTES H., ESCOUFIER Y. - (1978)
A propos de la comparaison graphique des matrices de variance.
Biom. J. 20 - 5, 477-483.

(13) L'HERMIER DES PLANTES (H.) ; THIEBAUT (B.) - 1977
Etude de la pluviosité au moyen de la méthode STATIS
R.S.A. 25, 2, 57-81

(14) MAILLES J.P. - 1978 - *Analyse factorielle des tableaux de dissimilarités.*
Thèse - Université Pierre et Marie Curie - Paris VI

(15) STENEMELEN J. - 1977 - Prévision des tableaux d'échanges les marges étant connues.
Cahier du Buro n° 28.