Revue de statistique appliquée

Yves Escoufier

Le positionnement multidimensionnel

Revue de statistique appliquée, tome 23, nº 4 (1975), p. 5-14
<http://www.numdam.org/item?id=RSA_1975__23_4_5_0>

LE POSITIONNEMENT MULTIDIMENSIONNEL (1)

Yves ESCOUFIER
Centre de Recherche en Informatique et Gestion MONTPELLIER

Après avoir rappelé les bases des méthodes de positionnement multidimensionnel qui conduisent au calcul des vecteurs propres d'une matrice semi-définie positive, l'article présente les procédés qui sont utilisés pour se ramener à ce cas lorsque la condition de positivité n'est pas remplie. Il présente ensuite la méthode proposée par Kruskal.

I – INTRODUCTION

I.1 – Soit I un ensemble de n objets. Si p caractères quantitatifs ont été observés sur chacun des objets, il est aisé de représenter l'ensemble des objets par n points d'un espace euclidien E de dimension p : on peut par exemple choisir un repère de p axes orthonormés de R^p et identifier chacun des axes à l'un des caractères.

Pour mesurer la ressemblance des objets, l'espace E est muni d'une métrique euclidienne M. Le choix de cette métrique dépend de la nature de l'ensemble I et de celle des caractères qui ont été observés : le but recherché est que les points représentant les objets soient d'autant plus proches au sens de la métrique M qu'ils sont plus ressemblants pour l'expérimentateur.

I.2 – Le problème que les auteurs américains abordent sous le titre de "multidimensionnal scaling" (que nous traduisons "positionnement multidimensionnel") correspond à une problématique rigoureusement opposée à la précédente : disposant d'informations sur les ressemblances mutuelles de n objets, on cherche une configuration de n points (dans R^2 en général) qui soit telle que si les objets i et j se ressemblent plus que les objets ℓ et k, on ait $d_{ij} < d_{k\ell}$ (où d_{ij} est la longueur du segment qui joint les points représentant les objets i et j).

Ce problème se rencontre lorsque l'étude ne conduit pas à des données quantitatives individuelles mais directement à une matrice $n \times n$ dont l'élément (i , j) mesure la ressemblance (ou la dissemblance) des objets i et j.

On peut également remarquer que, lorsque p caractères quantitatifs ont été observés sur n objets, la recherche d'une représentation utilisable des n objets dans le plan des deux premières composantes principales est sans aucun doute une technique de positionnement multidimensionnel.

---

(1) Article remis en Mai 1974, révisé en Mars 1975.

Le but de l'exposé est de montrer les possibilités et les limites des différentes méthodes qui sont proposées pour résoudre ce problème ainsi que les liens qui existent entre elles.

II – LES METHODES QUI CONDUISENT A CALCULER LES VECTEURS PROPRES D'UNE MATRICE SEMI-DEFINIE-POSITIVE

II.1 – Pour situer le problème, nous reprenons les notations de [3]. Soit X la matrice pxn des p caractères quantitatifs observés sur chacun des n objets. Des poids p_i $(p_i > 0 \,,\ \sum_{i=i}^{n} p_i = 1)$ étant affectés aux objets, la matrice X est supposée centrée.

X_i^j est l'élément situé à l'intersection de la colonne i et de la ligne j. Si $(\underline{e}_1 , \ldots \underline{e}_p)$ désigne la base canonique de $E = R^p$, à l'individu i est associé le vecteur de E : $\underline{X}_i = \sum_{i=i}^{p} X_i^k \, \underline{e}_k$. De la même manière si $(\underline{f}_1 , \ldots \underline{f}_n)$ est la base canonique de $F = R^n$, au caractère j est associé le vecteur

$$F : \underline{X}^j = \sum_{k=1}^{n} X_k^j \, \underline{f}_k \,.$$

A l'objet i est également associé le vecteur $\underline{f}_i^*$ de F^*, dual de F, défini par :

$$< \underline{f}_i^* , \underline{f}_j > = \begin{cases} 1 & \text{si} \quad j = i \\ 0 & \text{sinon} \end{cases}$$

Il apparaît alors que la transformation linéaire de F^* dans E qui à $\underline{f}_i^*$ fait correspondre $\underline{X}_i$ a pour matrice associée X lorsqu'on rapporte E à la base $(\underline{e}_1 , \ldots , \underline{e}_p)$ et F^* à la base $(\underline{f}_1^* , \ldots , \underline{f}_n^*)$.

De façon analogue, au caractère j est associé le vecteur $\underline{e}_j^*$ de E^*, dual de E, défini par :

$$< \underline{e}_j^* , \underline{e}_i > = \begin{cases} 1 & \text{si} \quad i = j \\ 0 & \text{sinon} \end{cases}$$

et X' est la matrice associée à la transformation linéaire qui fait correspondre $\underline{e}_j^*$ de E^* à $\underline{X}^j$ de F, lorsqu'on rapporte F à la base $(\underline{f}_1 , \ldots , \underline{f}_n)$ et E^* à la base $(\underline{e}_1^* , \ldots , \underline{e}_p^*)$.

Si on choisit, pour mesurer la ressemblance des objets, la métrique euclidienne M sur E et pour mesurer la ressemblance entre caractères la métrique euclidienne N sur F, on est conduit au schéma suivant dit "schéma de dualité" :

$$\begin{tikzcd}[row sep=large, column sep=huge]
E = R^p \arrow[d, shift right, "M"'] & F^* \arrow[l, "X"'] \arrow[d, shift right, "W"'] \\
E^* \arrow[u, shift right, "V"'] \arrow[r, "X'"'] & F = R^n \arrow[u, shift right, "N"']
\end{tikzcd}$$

$W = X' \circ M \circ X$ est l'écart euclidien qui s'impose sur F^*

si on veut que les distances entre objets soient les mêmes calculées dans F^* et dans E. ($V = X \circ N \circ X'$ joue un rôle analogue sur E^*). Si la distance entre les objets i et j est notée d_{ij}, on a :

$$d_{ij}^2 = M(\underline{X}_i - \underline{X}_j , \underline{X}_i - \underline{X}_j) = W(\underline{f}_i^* - \underline{f}_j^* , \underline{f}_i^* \ \underline{f}_j^*).$$

En particulier

$$\begin{aligned} d_{ij}^2 &= W(f_i^* , f_i^*) + W(f_j^* , f_j^*) - 2\, W(f_i^* , f_j^*) \\ &= W_{ii} + W_{jj} - 2\, W_{ij} \end{aligned}$$

où W_{ij} est l'élément (i , j) de la matrice associée à la métrique euclidienne W quand ont rapporte F^* et F à leurs bases canoniques respectives.

Remarques : On prend en général pour N la métrique euclidienne diagonale D_p (D_p (i , i) = p_i) ; ce choix sera fait dans la suite de l'exposé.

– Soit $L_2(\Omega , \mathcal{A} , P)$ l'ensemble des variables aléatoires centrées de variance finie sur $(\Omega , \mathcal{A} , P)$. A tout vecteur dont les composantes appartiennent à $L_2(\Omega , \mathcal{A}, P)$ l'auteur a associé un opérateur $\mathcal{U}$ de $L_2(\Omega , \mathcal{A}, P)$ en lui-même [5]. On doit a J.M. BRAUN [2] d'avoir remarqué que $W \circ Dp$ joue sur F, le rôle que joue $\mathcal{U}$ sur $L_2(\Omega , \mathcal{A}, P)$.

II.2 – Analyse en composantes principales

Si p est supérieur à 2 et si une représentation utilisable des n objets est souhaitée, l'analyse en composantes principales fournit le sous-espace de E de dimension $s < p$ (en général $s = 2$) par rapport auquel le moment d'inertie du nuage des n individus est minimum. Nous ne reviendrons pas sur cette méthode bien connue. Il est important toutefois de noter que, si $(\lambda_1 , \dots , \lambda_s)$ sont les s plus grandes valeurs propres de l'opérateur $W \circ D_p$, la solution consiste à déterminer s nouveaux caractères c^j , D_p orthonormaux, vérifiant $W \circ D_p\, \underline{c}^j = \lambda_j\, \underline{c}^j$. La méthode doit son sens à la positivité de l'opérateur $W \circ D_p$: les valeurs propres étant positives, peuvent être interprétées comme des parts d'inertie expliquée par les caractères c^j auxquels elles sont associées. La qualité de la solution trouvée est alors mesurée par la quantité

$$\sum_{i=1}^{s} \lambda_s / \mathrm{Tr}(W \circ D_p)\ .$$

II.3 – Analyse factorielle d'une matrice de similarité semi-définie positive

Dans certaines applications tout ou partie des caractères observés sont qualitatifs ; il est possible alors de calculer une matrice de similarités ou de dissimilarités entre les objets en utilisant l'une quelconque des formules de coefficients de dissimilarité ou de similarité proposés dans la littérature. D'autres situations expérimentales peuvent également se rencontrer dans lesquelles les données initiales sont une matrice de similarités ou de dissimilarités.

L'usage courant des statisticiens est alors le suivant [6]. Si une matrice S de similarités a été calculée, et si cette matrice est semi-définie positive, elle prend la place de la matrice W du schéma de dualité. Une représentation utilisable des individus peut alors être obtenue par une analyse en composantes principales ; sa qualité est appréciée comme en II.2. Il est clair, d'après II. 1, que cette procédure revient à définir la distance entre les objets par $(S_{ii} + S_{jj} - 2\,S_{ij})^{1/2}$.

Lorsque les données du problème consistent en une matrice δ de dissimilarités entre les objets, la méthode proposée par TORGERSON [15] est de calculer une matrice S par la formule :

$$S_{ij} = \frac{1}{2}(\delta_{i.}^2 + \delta_{j.}^2 - \delta_{ij}^2 - \delta_{..}^2)$$

avec $$\delta_{i.}^2 = \sum_{k=1}^{n} p_k\, \delta_{ik}^2 \quad \text{et} \quad \delta_{..}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} p_i\, p_j\, \delta_{ij}^2$$

Si S est semi-définie positive, elle peut prendre la place de W dans le schéma de dualité et il est facile de voir que la distance utilisée entre les objets, qui d'après ce qui précède est égale à $(S_{ii} + S_{jj} - 2S_{ij})^{1/2}$ est bien δ_{ij}.

Pour l'utilisation de ces procédés, il est intéressant de noter les deux résultats suivants :

a) lorsque la matrice S est semi-définie positive, on peut affirmer que les δ_{ij} réalisent les axiomes des distances mais il ne suffit pas que les δ_{ij} réalisent les axiomes des distances pour que S soit semi-définie positive.

La première partie du résultat est évidente puisque S semi-définie positive est une métrique euclidienne sur F* et que

$$\delta_{ij}^2 = S_{ii} + S_{jj} - 2\,S_{ij} = S(\underline{f}_i^* - \underline{f}_j^*\,,\underline{f}_i^* - \underline{f}_j^*)$$

Pour démontrer la seconde partie, il suffit d'expliciter un ensemble de valeurs réalisant les axiomes des distances mais non réalisable dans un espace réel. Prenons par exemple quatre points dont les distances mutuelles sont données par la matrice suivante :

	A	B	C
B	$\sqrt{2}$		
C	$\sqrt{3}$	1	
D	1	2	1

Ayant placé A, B et C dans un plan, il est facile de voir que les sphères centrées en A, B, et C de rayons respectifs 1, 2 et 1 n'ont pas de points communs.

b) Si les δ_{ij} réalisent les axiomes des distances ultramétriques alors S est semi-définie positive. La démonstration qui suit est due à [7]. On trouve une démonstration très voisine de ce résultat en [8].

Soit I un ensemble d'objets sur lequel est donnée une matrice de distances ultramétriques δ_{ij}. Mis à part le cas trivial où tous les δ_{ij} sont égaux entre eux, on peut trouver une partition de I en deux sous ensembles I_1 et I_2, tels que $\delta_{kl} = \alpha = \max\limits_{I \times I} \delta_{ij}$ pour tout couple $(k , l) \in I_1 \times I_2$.

Supposons alors que I_1 puisse être représenté par une configuration de points placés sur une hypersphère de R^{p_1} de diamètre $2r_1 < \sqrt{2} \max\limits_{I_1 \times I_1} \delta_{ij}$. Si l'origine des axes est placée au centre de la sphère, les coordonnées des points de I_1 sont données par une matrice $\underset{\sim}{X}_1$, $(p_1 \times |I_1|)$. Une hypothèse analogue est faite pour I_2.

Pour représenter I choisissons la configuration de coordonnées

$$\left(\begin{array}{c|c} \underset{\sim}{X}_1 & 0 \\ \hline \underset{\sim}{Y}_1 & \underset{\sim}{Y}_2 \\ \hline 0 & \underset{\sim}{X}_2 \end{array}\right)$$

où $\underset{\sim}{Y}_i$ est un vecteur dont les p_i composantes sont égales à Y_i. Pour tout couple $(k , l) \in I_1 \times I_2$, on a :

$$\delta^2_{kl} = r_1^2 + r_2^2 + (Y_1 - Y_2)^2$$

et on peut choisir Y_1 et Y_2 de telle sorte que

$$\delta^2_{kl} = \alpha^2$$

Montrons que les points de l'ensemble I peuvent être placés sur une hypersphère de $R^{p_1 + p_2 + 1}$ de diamètre $2\, r < \sqrt{2} \max\limits_{I \times I} \delta_{ij} = \sqrt{2}\, \alpha$.

Pour que l'origine soit à l'origine des axes, il faut que :

$$r_1^2 + Y_1^2 = r_2^2 + Y_2^2 = r^2$$

D'où :

$$\begin{aligned} \alpha^2 - r_1^2 - r_2^2 = (Y_1 - Y_2)^2 &= [(r^2 - r_1^2)^{1/2} - (r^2 - r_2^2)^{1/2}]^2 \\ &= 2\, r^2 - 2(r^2 - r_1^2)^{1/2} (r^2 - r_2^2)^{1/2} - r_1^2 - r_2^2 \end{aligned}$$

Soit

$$\alpha^2 - 2\, r^2 = - 2(r^2 - r_1^2)^{1/2} (r^2 - r_2^2)^{1/2}$$

En élevant au carré on a :

$$\alpha^4 - 4\, \alpha^2\, r^2 = 4\, r_1^2\, r_2^2 - 4\, r^2\, (r_1^2 + r_2^2)$$

soit

$$\alpha^4 - 4 r_1^2 r_2^2 = 4 r^2 (\alpha^2 - r_1^2 - r_2^2)$$

et

$$r^2 = \frac{\alpha^4 - 4 r_1^2 r_2^2}{4 (\alpha^2 - r_1^2 - r_2^2)}$$

Par hypothèse $r_i < \alpha/\sqrt{2}$ donc le numérateur est bien positif. Le dénominateur l'est également puisqu'égal à $(Y_1 - Y_2)^2$.

Mais on a aussi :

$$(\alpha^2 - 2 r_1^2) (\alpha^2 - 2 r_1^2) > 0$$

D'où :

$$2 \alpha^4 - 2 \alpha^2 r_1^2 2 \alpha^2 r_2^2 > \alpha^4 - 4 r_1^2 r_2^2$$

Soit :

$$2 \alpha^2 (\alpha^2 - r_1^2 - r_2^2) > \alpha^4 - 4 r_1^2 r_2^2$$

et

$$\frac{\alpha^2}{2} > \frac{\alpha^4 - 4 r_1^2 r_2^2}{4 (\alpha^2 - r_1^2 - r_2^2)} = r^2$$

Il en découle

$$2 r < \sqrt{2} \alpha$$

III – LES METHODES QUI RAMENENT A L'ANALYSE FACTORIELLE D'UNE MATRICE DE SIMILARITES SEMI-DEFINIE POSITIVE

Dans de nombreuses applications la matrice S des similarités n'est pas semi-définie positive. La méthode précédente qui repose sur la positivité des valeurs propres de S n'est donc pas applicable.

III.1. – Pour tourner la difficulté sans changer de méthode, les utilisateurs font souvent une analyse en composantes principales de S en ne retenant que les vecteurs propres associés aux valeurs propres positives. Il leur serait bon dans ce cas d'être conscients que cette technique revient à remplacer S par S*, semi-définie positive, approximation au sens des moindres carrés de S et que $\mathrm{Tr}(S - S^*)^2$ qui mesure l'écart entre S et S* est égale à la somme des carrés des valeurs propres négatives de S [13]. En particulier, les distances réellement prises en compte entre les individus sont ici $(S_{ii}^* + S_{jj}^* - 2 S_{ij}^*)^{1/2}$.

Le résultat b du paragraphe précédent permet d'envisager de substituer à une matrice de dissimilarités, une matrice de distances ultramétriques. N'importe quelle méthode de classification hiérarchique peut être utilisée dans ce but. On calcule alors S par la méthode de Torgerson à partir de la matrice des distances ultramétriques. Quelques expérimentations simples nous ont montré les imperfections de ce procédé.

III.2 – Une idée introduite par Shepard [14] est de substituer aux δ_{ij} des quantités δ_{ij}^* telles que

c1) $\delta_{ij} < \delta_{k\ell} \Rightarrow \delta_{ij}^* < \delta_{k\ell}^*$

c2) la matrice δ^* conduit à une matrice S semi-définie positive.

Prenant comme représentation initiale des n objets les points du simplexe de R^{n-1}, l'algorithme de Shepard construit la solution par des déformations successives de ce simplexe. De plus, il tend à diminuer la dimension de l'espace dans lequel la configuration se trouve en utilisant le principe intuitif selon lequel la dimension de l'espace diminuera si la variance des distances augmente. A ce stade de son traitement Shepard dispose donc d'une représentation des objets dans un espace de dimension p. Une analyse en composantes principales lui permet d'obtenir une représentation approchée utilisable.

Reprenant ces travaux, Benzécri [1] a établi ce que Shepard avait conjecturé à savoir qu'il était toujours possible de trouver des distances δ^*_{ij} qui conduisent à une matrice S semi-définie positive et qui soient rangées dans le même ordre que les δ_{ij}. L'argument utilisé est un argument de continuité : si tous les δ^*_{ij} sont égaux à l'unité, S est semi-définie positive ; par continuité S sera semi-définie positive si les δ^*_{ij} sont suffisamment voisins de l'unité, condition compatible avec le fait que les δ^*_{ij} doivent être rangées dans le même ordre que les δ_{ij}.

III.3 – Cette idée de continuité peut être rattachée à celle antérieure de "constante additive" introduite par Torgerson dans le cas unidimensionnel puis généralisée au cas multidimensionnel par Messick et Abelson [voir Torgerson (15) p. 271 et sqs]. Il s'agit de trouver une constante c telle que les $\delta^*_{ij} = \delta_{ij} + c$ conduisent à une matrice S semi-définie positive. On peut toujours choisir c assez grand pour que les $(\delta_{ij} + c)/c$ soient aussi voisins de l'unité qu'on le veut et donc conduisent à une matrice S semi-définie positive par l'argument de Benzécri. Il est aisé de voir que le résultat reste valable pour les $\delta^*_{ij} = \delta_{ij} + c$. Afin de ne pas dénaturer les distances, on cherchera bien sûr une valeur de c aussi petite que possible. L'algorithme proposé par L.G. Cooper [4] semble moins sujet à caution que celui de Messick et Abelson. Il faut toutefois souligner le danger de cette méthode : si la valeur de c obtenue est grande par rapport aux distances, la configuration des points est très peu représentative des objets. De plus, dans ce cas, les δ^*_{ij} sont sensiblement égales, si bien que l'analyse en composantes principales qui sera faite pour obtenir une représentation utilisable ne peut que donner une mauvaise approximation dans le plan de deux premiers vecteurs propres.

IV – LA METHODE DE J.B. KRUSKAL [9 , 10]

Plutôt que de chercher une matrice S semi-définie positive puis une représentation utilisable des objets dans un espace de dimensions réduites par une analyse factorielle de S, J.B. Kruskal cherche directement une configuration dans un espace de dimensions limitées qui remplisse la condition c1) . La procédure est la suivante.

IV.1 – Considérons une configuration de n points dans un espace de dimensions $k\,(k = 2$ en général) et soit $\{d_{ij}/i = 1, \ldots n\ ;\ j = 1, \ldots n\}$ les distances mutuelles des points. S'il existe une fonction f monotone croissante sur $[0, +\infty]$ telle que $f(\delta_{ij}) = d_{ij}$ alors la configuration est une bonne représentation des objets. Si F est l'ensemble des fonctions monotones croissantes sur $[0, +\infty]$ et Z l'ensemble des couples d'objets différents, on est conduit à chercher $g \in F$ telle que :

$$\sum_{(i,j)\in Z} (g(\delta_{ij}) - d_{ij})^2 = \min_{f\in F} \sum_{(i,j)\in Z} (f(\delta_{ij}) - d_{ij})^2$$

En fait, la procédure consiste à chercher des valeurs numériques δ^*_{ij} telles que $\sum_{(i,j)\in Z} (\delta^*_{ij} - d_{ij})^2$

soit minimum sous les contraintes

$$(\delta^*_{ij} - \delta^*_{k\ell})(\delta_{ij} - \delta_{k\ell}) \geqslant 0 \qquad \text{pour tout} \qquad ((i,j),(k,\ell)) \in Z \times Z$$

L'existence d'une solution pour ce problème a été clairement discutée par Kruskal [11] : les contraintes définissent un polyèdre convexe fermé et la minimisation de la fonction revient à chercher la projection du vecteur des d_{ij} sur ce polyèdre. On sait que la solution existe et est unique. De plus de nombreux algorithmes sont proposés dans la littérature traitant de programmation quadratique convexe pour résoudre ce problème. En fait Kruskal a proposé un algorithme spécifique [10] qui a l'avantage d'être particulièrement efficace.

Il faut noter que cette approximation au sens des moindres carrés de quantités d_{ij} par des quantités δ^*_{ij} rangées dans le même ordre que des δ_{ij} données est un problème de régression maintenant connu sous le terme de régression monotone : au lieu de chercher l'approximation des d_{ij} par une droite de la forme $a\,\delta_{ij} + b$, on cherche l'approximation des d_{ij} par une fonction $f(\delta_{ij})$ monotone croissante.

IV.2 – Connaissant les δ^*_{ij}, la seconde étape consiste à mouvoir les points de la configuration, c'est-à-dire à modifier les d_{ij} de façon à minimiser la quantité $\sum_{(i,j)\in Z} (\delta^*_{ij} - d_{ij})^2$. Cette quantité est une fonction des $n \times k$ coordonnées des points de la configuration. Kruskal propose d'en rechercher le minimum par application de la méthode du gradient. D'autres méthodes peuvent être retenues. Quoiqu'il en soit, nous ne sommes pas en présence d'une fonction des coordonnées qui soit convexe si bien qu'on peut craindre de ne pas trouver le véritable minimum, mais seulement un minimum local : on ne peut échapper à ce genre de difficulté.

La solution finale sera obtenue en utilisant en alternance les deux algorithmes précédents. On s'arrêtera soit lorsqu'il ne sera plus possible d'améliorer la solution, soit lorsque la quantité

$$\sum_{(i,j)\in Z} (\delta^*_{ij} - d_{ij})^2 / \sum_{(i,j)\in Z} d^2_{ij}$$

appelée "le stress" par Kruskal sera suffisamment petite.

V – CONCLUSION

Sur quatre objets, munis des poids 1/4, deux caractères quantitatifs ont été observés. La matrice des données est la suivante :

	A	B	C	D
X^1	1	1	−1	−1
X^2	1,25	−1,25	0,01	−0,01

Si on munit E de la métrique identité, la première composante principale est X^1. C'est pourtant X^2, deuxième composante principale qui donne une configuration des points respectant l'ordre des distances initiales ! Faut-il pour autant renoncer à l'emploi des méthodes factorielles ? Certainement pas car le contexte mathématique dans lequel elles se situent permet d'obtenir des informations que ne fournit pas l'analyse proposée par Kruskal. Alors, que faire ? Peut-être faut-il simplement rester conscient du fait qu'analyser des données ne se réduit pas à les traiter par un programme standard mais plutôt à les explorer par des méthodes complémentaires. On ne dira jamais assez qu'une méthode, quelle qu'elle soit, ne donne qu'un point de vue particulier sur des données. Pour approcher la réalité d'un ensemble de données, le chercheur doit comprendre l'intérêt qu'il trouve à chercher plusieurs points de vue.

REMERCIEMENTS

Je tiens à remercier très chaleureusement M. J.P. Pagès pour les remarques très pertinentes qu'il m'a faites au nom du comité de lecture de la revue. Les lecteurs qui apprécieraient cet article devront penser qu'une bonne part de ses qualités provient des suggestions qui m'ont été faites.

REFERENCES

[1] BENZECRI J.P. (1964) – Sur l'analyse factorielle des proximités *pub. Inst. Stat. Univ. Paris,* vol. XIII, p. 235-282.

[2] BRAUN J.M. (1973) – Séries chronologiques multiples : recherche d'indicateurs *R.S. A,* vol. XX1, n° 1, p. 81, 106.

[3] C.E.E.E. – Analyse des données multidimensionnelles.

[4] COOPER L.G. (1972) – A new solution to the additive constant problem in metric multidimensional scaling. *Psychometrika,* vol. 37, n° 3, p. 311-322

[5] ESCOUFIER Y. (1970) – Echantillonnage dans une population de variables aléatoires réelles. *Pub. Inst. Stat. Univ. Paris,* vol. 19, p. 1 à 47.

[6] GOWER J.C. (1966) – Some distances properties of latent root and vector methods used in multivariate analysis. Biometrika. vol. 53, p. 325-338.

[7] GOWER J.C. (1974) – Communication privée

[8] HOLMAN E.W. (1972) – The relation between hierarchical and euclidean models for psychological distances. *Psychometrika,* vol. 37, n° 4, p. 417-423.

[9] KRUSKAL J.B. (1964) – Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika,* vol. 29, p. 1-27.

[10] KRUSKAL J.B. (1964) – Nonmetric multidimensional scaling : a numerical method *Psychometrika,* vol. 29, p. 115-129.

[11] KRUSKAL J.B. (1971) – Monotone regression : Continuity and differentiability properties. *Psychometrika,* vol. 36, p. 57-62.

[12] RAO C.R. (1965) – The use and interpretation of principal component analysis in applied research Sankhya A, vol. 26, p. 329-58.

[13] SCHWERTMAN N.C. and ALLEN D.M. (1973) – The smoothing of an indefinite matrix with applications to growth curve analysis with missing observations Technical Report n° 56 – Dèpartement of Statistics – University of Kentucky

[14] SHEPARD R.N. (1962) – The analysis of proximities : multidimensional scaling with an unknown distance function I, vol. 27, p. 125-140 II, vol. 27, p. 219-246.

[15] TORGERSON W.S. (1958) – Theory and Methods of Scaling John Wiley and Sons.