STATISTIQUE ET ANALYSE DES DONNÉES

B. FICHET

Sur des approximations d’indices de dissimilarité via les représentations euclidienne et hiérarchique

Statistique et analyse des données, tome 6, n° 2 (1981), p. 1-21
<http://www.numdam.org/item?id=SAD_1981__6_2_1_0>

Statistiques et Analyse de données
1981 - Vol. 6 n° 2 - pp. 1-21

SUR DES APPROXIMATIONS D'INDICES DE DISSIMILARITE VIA LES REPRESENTATIONS EUCLIDIENNE ET HIERARCHIQUE

B. FICHET

Laboratoire de Physique - Faculté de Médecine
Université d'Aix-Marseille II

Résumé : *Deux approximations d'indices de dissimilarité sont proposées , l'une pour la représentation euclidienne, l'autre pour la représentation hiérarchique. Pour la première, on opère dans un espace de formes quadratiques pour obtenir une forme positive définissant l'approximation. Pour la seconde, l'indice ultramétrique cherché répond aux deux critères d'optimalité suivants : moindres carrés sous une contrainte de meilleure conservation de l'ordre sur les données.*

Abstract : *It is here proposed two approximations of dissimilarity coefficients, respectively related to euclidian and hierarchical patterns. The first type of approximation is issued from the definition of a positive form in a given space of quadratic forms. In the second type of approximation it is given an ultrametric dissimilarity that satisfies the following two optimality criteria : least squares under a best order invariance constraint on data.*

Mots clés : *Dissimilarité, Approximation, Formes quadratiques, Ultramétrique, Moindres carrés.*

1 - INTRODUCTION

Tout au long de cet article, I est un ensemble fini de cardinalité n>1, et d est un indice de dissimilarité sur I. On note $\mathfrak{D}_I$ l'espace vectoriel des applications $\delta : I \times I \longmapsto \mathbb{R}$, vérifiant : $\forall i \in I$, $\delta(i,i) = 0$; $\forall (i,j) \in I \times I$: $\delta(i,j) = \delta(j,i)$; $\mathfrak{D}_I$ est de dimension $[n(n-1)]/2$ et un indice de dissimilarité est un élément de l'orthant positif de $\mathfrak{D}_I$.

Pour une représentation euclidienne du doublet (I,d), une méthodologie retenue, que l'on peut nommer A.C.P. de ce doublet, est la suivante : recherche d'une image euclidienne de (I,d) - i.e. recherche d'un espace (affine) euclidien E de dimension finie et d'une famille $\{M_i, i \in I\}$ de points de E, vérifiant : $\forall (i,j) \in I \times I$, $||\overrightarrow{M_i M_j}|| = d(i,j)$ - ; puis A.C.P. de cette image (ou du nuage $\{M_i, m_i, i \in I\}$ si, en outre, des masses $m_i, i \in I$ sont introduites au niveau des données premières).

Il est classique que certains doublets n'admettent pas d'image euclidienne même si d est une semi-distance. Alors dans ce cas, pour une méthodologie plus complète, on peut approcher d par un indice d_* tel que pour le doublet correspondant (I,d_*) existe une image euclidienne, et réaliser ensuite l'A.C.P. de ce doublet (ou du triplet $\{I, d_*, m_i, i \in I\}$).

Pour la représentation hiérarchique on se remémore le résultat établissant l'équivalence entre hiérarchie indicée sur I et semi-distance ultramétrique sur I. Dès lors, l'approximation d'un indice quelconque d par un indice ultramétrique d_* conduit à une représentation hiérarchique du doublet (I,d).

Nous proposons une approximation pour chacun des deux modes de représentation considérés ; toutes deux possèdent un trait technique commun : elles s'expriment en termes de projection sur un cône convexe.

Pour la représentation euclidienne, l'approximation au sens des moindres carrés n'est pas, à ce jour, résolue - le domaine est non convexe ! - Et c'est par une transposition du problème en termes de formes quadratiques que nous proposons une approximation. Pour la représentation hiérarchique, l'ultramétrique cherchée satisfait à un critère des moindres carrés sous des contraintes optimales, contraintes très fortes qui correspondent à une politique d'extrême prudence ; et, un algorithme simple, reposant éssentiellement sur un classement de nombres, permet de la construire.

2 - APPROXIMATION POUR LA REPRESENTATION EUCLIDIENNE

Considérons tout d'abord les deux figures suivantes, la deuxième étant un cas particulier de la première.

Figure 1

I est l'ensemble fini précédemment considéré.
E est un espace affine de dimension finie ; V est son espace vectoriel associé.
$\{M_i, i \in I\}$ est une famille de points de E.
F est le sous-espace affine (variété) engendré par les points $M_i, i \in I$; W est son espace vectoriel associé.

Figure 2

I est l'ensemble fini précédemment considéré.
E est un espace affine de dimension $n = \mathrm{card}(I)$; V est son espace vectoriel associé ; $\{0, \{\vec{e}_i, i \in I\}\}$ est un référentiel de E.
$\forall i \in I$, $M_i = 0 + \vec{e}_i$
F est le sous-espace affine (hyperplan) engendré par les points $M_i, i \in I$; W est son espace vectoriel associé.

Ainsi, dans cette dernière figure, les points $M_i, i \in I$, forment un $(n-1)$-simplexe non-dégénéré, engendrant l'espace F.

Par la suite, nous utiliserons les notations suivantes :

$\mathcal{F}(V)$ et $\mathcal{F}(W)$ sont respectivement les espaces vectoriels des formes bilinéaires symétriques sur V et W.

$\mathcal{F}^+(V)$, $\mathcal{F}^+(W)$ et $\mathcal{F}^+_V(W)$ sont respectivement les cônes convexes fermés des formes positives sur V, des formes positives sur W, des formes sur V dont la restriction à W est positive.

Rappelons alors un résultat fondamental, sur lequel repose notre approximation ; la démonstration dans le cas particulier de la figure 2 et pour une forme q particulière -[7] p. 253 - peut être aisément étendue au cas général de la figure 1 - voir aussi [4]-.

Théorème 1

Soit (I,d) le doublet précédemment considéré. Avec la figure 1, soit $q \in \mathcal{F}(V)$ satisfaisant à : $\forall (i,j) \in I \times I \quad q(\overrightarrow{M_iM_j}, \overrightarrow{M_iM_j}) = d^2(i,j)$. Alors :

i) Pour qu'existe une image euclidienne de (I,d) il faut et il suffit que $q \in \mathcal{F}^+_V(W)$.

ii) Si $q \in \mathcal{F}_V^+(W)$, la dimension de (I,d) -i.e. la dimension de toute image euclidienne de (I,d) - est égale au rang de la restriction, soit $\underline{q}$, de q à W.

Relativement aux données du théorème, que l'on suppose satisfaites, notre approximation $d \longmapsto d_*$, est alors définie de la façon suivante : munissant $\mathcal{F}(V)$ (ou $\mathcal{F}(W)$) d'une structure euclidienne, on projette q (ou $\underline{q}$) dans le cône convexe fermé $\mathcal{F}_V^+(W)$ (ou $\mathcal{F}^+(W)$) pour obtenir une forme q_* ; l'indice d_* est alors défini par : $\forall (i,j) \in I \times I$, $d_*^2(i,j) = q_*(\overrightarrow{M_i M_j}, \overrightarrow{M_i M_j})$. Il est naturellement souhaitable que les deux approximations, correspondant respectivement aux projections de q et $\underline{q}$, soient égales. On peut également projeter q dans $\mathcal{F}^+(V)$, ce qui devrait conduire à une solution numérique plus simple ; mais cette approximation, volontairement trop forte, ne saurait être retenue que dans la mesure où elle conduirait à l'un des résultats précédents.

La structure euclidienne considérée sur $\mathcal{F}(V)$ est classique. Faisant choix d'un produit scalaire sur V, soit p, et notant $Q = (Q_{ij})$ la matrice associée à l'élément q de $\mathcal{F}(V)$ dans une base p-orthonormée, la norme euclidienne classique de Q, induit une norme pour q, soit $||q||_{p}$, indépendante de la base p-orthonormée considérée. On a : $||q||_p^2 = \sum_i \sum_j Q_{ij}^2 = Tr(Q^2)$; c'est aussi la somme des carrés des valeurs propres de Q.

La structure euclidienne de $\mathcal{F}(W)$ est définie de façon analogue ; et, pour que les résultats puissent être comparables, on choisit comme produit scalaire sur W, la restriction, soit $\underline{p}$, de p à W.

On a alors les trois problèmes :

$$P1 : \min_{f \in \mathcal{F}^+(W)} ||\underline{q}-f||_{\underline{p}} \quad ; \quad P2 : \min_{f \in \mathcal{F}_V^+(W)} ||q-f||_p \quad ; \quad P3 : \min_{f \in \mathcal{F}^+(V)} ||q-f||_p$$

Le théorème suivant traite à la fois de la résolution numérique et de la comparaison des résultats. Signalons, que pour la résolution numérique, il a été obtenu conjointement par Critchley - [11] -, au moins dans le cas particulier où $W = V$ (problème P3).

<u>Théorème 2</u>

Avec les données et notations précédentes, soit $\{\vec{\varepsilon}_1, \ldots, \vec{\varepsilon}_r\}$ une base $\underline{p}$-orthonormée de W, pour laquelle $\underline{q}$ est diagonale. On note $W^\perp$ le supplémentaire p-orthogonal de W. Alors :

i) La solution, soit q_*, du problème P2 vérifie :

a) $\forall j,k = 1, \ldots r \quad q_*(\vec{\varepsilon}_j, \vec{\varepsilon}_k) = \max \left[q(\vec{\varepsilon}_j, \vec{\varepsilon}_k), 0 \right]$

b) $\forall \vec{x} \in V$, $\forall \vec{y} \in W^{\perp}$, $q_*(\vec{x},\vec{y}) = q(\vec{x},\vec{y})$

ii) La restriction, soit $\underline{q}_*$, de q_* à W, est la solution du problème P1.

iii) Si W et $W^{\perp}$ sont q-conjugués, $\underline{q}_*$ est la restriction à W, de la solution du problème P3.

Démonstration :

Si $\dim(V) = m > r$, complétons le système $\{\vec{\varepsilon}_1,..,\vec{\varepsilon}_r\}$ pour obtenir une base $\mathcal{B} = \{\vec{\varepsilon}_1,..,\vec{\varepsilon}_m\}$ p-orthonormée de V.

Pour $f \in \mathcal{F}^+_V(W)$, notons $F = (f_{jk})$ la matrice associée à f dans $\mathcal{B}$. Alors, puisque $\mathcal{B}$ est orthonormée :

$$||q-f||_p^2 = \sum_{j=1}^{m}\sum_{k=1}^{m}\left[q(\vec{\varepsilon}_j,\vec{\varepsilon}_k)-f_{jk}\right]^2$$

Si $J = \{j=1,..,r \mid q(\vec{\varepsilon}_j,\vec{\varepsilon}_j) < 0\}$, l'on a, puisque $\forall j=1,..,r \quad f_{jj} \geq 0$:

$$||q-f||_p^2 \geq \sum_{j \in J}\left[q(\vec{\varepsilon}_j,\vec{\varepsilon}_j)-f_{jj}\right]^2 \geq \sum_{j \in J}\left[q(\vec{\varepsilon}_j,\vec{\varepsilon}_j)\right]^2$$

(Si $J = \emptyset$, les deux dernières sommes sont prises égales à 0).

Et les inégalités ci-dessus ne sont des égalités que si nécessairement :

$$\begin{cases} (\text{si } m>r) \quad \forall j=1,..,m \ , \ \forall k=(r+1),..,m : \ f_{jk}=q(\vec{\varepsilon}_j,\vec{\varepsilon}_k) \\ \forall j,k=1,..,r \ , \ j \neq k : \ f_{jk}=q(\vec{\varepsilon}_j,\vec{\varepsilon}_k) = 0 \\ \forall j=1,..,r \ , \ j \notin J \ , \ f_{jj}=q(\vec{\varepsilon}_j,\vec{\varepsilon}_j) \ ; \ \forall j=1,..,r, j \in J : \ f_{jj} = 0 \end{cases}$$

Mais clairement, l'unique matrice soit F^*, vérifiant ces relations, définit, par rapport à $\mathcal{B}$, une forme de $\mathcal{F}^+_V(W)$; c'est donc la solution q_* du problème P2. Par construction même, a) de i) est prouvé. Et (si m>r), décomposant $\vec{x}$ et $\vec{y}$ respectivement sur $\mathcal{B}$ et $\{\vec{\varepsilon}_{r+1},..,\vec{\varepsilon}_m\}$, b) de i) découle encore de la construction de q_*.
Utilisant la même base $\{\vec{\varepsilon}_1,..,\vec{\varepsilon}_r\}$ de W, on en déduit que la solution, soit q' du problème P1 vérifie : $\forall j,k=1,..,r \quad q'(\vec{\varepsilon}_j,\vec{\varepsilon}_k) = \max\left[\underline{q}(\vec{\varepsilon}_j,\vec{\varepsilon}_k),0\right] = \underline{q}_*(\vec{\varepsilon}_j,\vec{\varepsilon}_k)$ et ii) est prouvé.

Enfin (si m>r), pour montrer iii), on peut toujours supposer que la base $\{\vec{\varepsilon}_{r+1},..,\vec{\varepsilon}_m\}$ p-orthonormée de $W^{\perp}$, vérifie de plus : $\forall j,k=(r+1),..,m$, $j \neq k$, $q(\vec{\varepsilon}_j,\vec{\varepsilon}_k)=0$. Comme W et $W^{\perp}$ sont q-conjugués, la base $\mathcal{B}$ satisfait aux conditions de l'énoncé pour le problème P2, avec W=V ; et, utilisant i) a), on en déduit iii).

□

Ainsi, P1 et P2 conduiront toujours à la même approximation d_* ; et si W et $W^{\perp}$ sont q-conjugués, P3 conduira encore à cette approximation.

Numériquement, si $\mathcal{B}' = \{\vec{\varepsilon}'_1,..,\vec{\varepsilon}'_m\}$ est une base p-orthonormée de V, telle que $\{\vec{\varepsilon}'_1,..,\vec{\varepsilon}'_r\}$ soit une base de W, et si $Q = \begin{pmatrix} Q_1 & Q_2 \\ Q_2' & Q_3 \end{pmatrix}$ est la matrice associée à q dans $\mathcal{B}'$, une diagonalisation de Q_1 donne r valeurs propres $\lambda_1,..,\lambda_r$ et r vecteurs propres orthonormés respectivement associés, $\{P_1,..,P_r\}$. De sorte que si $Q_1^* = \sum_{j=1}^{r} \max(\lambda_j,0)P_jP_j'$, $Q^* = \begin{pmatrix} Q_1^* & Q_2 \\ Q_2' & Q_3 \end{pmatrix}$ est la matrice associée à q_* dans $\mathcal{B}'$.

Pour utiliser ce résultat, il suffit de fixer une configuration de points dans un certain espace, et de rechercher une forme q satisfaisant à la condition du théorème 1. Pour la configuration dans l'espace F de dimension (n-1), donnée par la figure 2, rappelons alors le résultat suivant - [2], [4], [12] - .

Proposition 1

$\forall\ \delta \in \mathfrak{D}_I$, il existe une forme unique de $\mathcal{F}(W)$, soit $b=b_0(\delta)$, satisfaisant à :
$\forall (i,j)\in I\times I$: $b(\overrightarrow{M_iM_j},\overrightarrow{M_iM_j}) = \delta(i,j)$.
L'application b_0 de $\mathfrak{D}_I$ dans $\mathcal{F}(W)$ ainsi définie, est un isomorphisme.

Matriciellement pour une bijection donnée entre I et {0,..,(n-1)}, la matrice, soit B, associée à b, dans la base $\{\overrightarrow{M_0M_1},..,\overrightarrow{M_0M_{n-1}}\}$ de W, a pour terme général :

$$B_{ij} = \frac{1}{2}\left[\delta(0,i) + \delta(0,j) - \delta(i,j)\right] \qquad i,j=1,..,(n-1).$$

Si on désire oeuvrer dans l'espace E de dimension n, il existe une infinité de formes satisfaisant à la condition du théorème 1 : ce sont toutes les formes qui prolongent b, définie par la proposition 1, à V. Un prolongement simple consiste à fixer un point (quelconque) G dans F - Si G a pour coordonnées $m_i, i\in I$, dans le référentiel de E, G est le barycentre des points M_i affectés des masses m_i - et à prendre la forme qui rend les directions supplémentaires $\overrightarrow{OG}$ et W, conjuguées. On a ainsi la :

Proposition 2

$\forall G\in F$, $\forall k\in \mathbb{R}$, $\forall \delta\in\mathfrak{D}_I$, il existe une forme unique de $\mathcal{F}(V)$, soit $\tau = \tau_0(G,k,\delta)$, satisfaisant à :

i) $\forall (i,j) \in I \times I$ $\quad \tau(\overrightarrow{M_iM_j}, \overrightarrow{M_iM_j}) = \delta(i,j)$

ii) $\forall \vec{x} \in W$ $\quad \tau(\overrightarrow{OG}, \vec{x}) = 0$

iii) $\quad \tau(\overrightarrow{OG}, \overrightarrow{OG}) = k.$

Matriciellement, pour une bijection donnée entre I et $\{1,..,n\}$, la matrice, soit T, associée à $\tau_o(G,O,\delta)$ dans la base $\{\vec{e}_1,..,\vec{e}_n\}$ de V, a pour terme général :

$$t_{ij} = \frac{1}{2}\left[-\delta(i,j) + \delta(i,.) + \delta(j,.) - \delta(.,.)\right] \qquad i,j=1,..,n$$

(avec $\delta(i,.) = \sum_j m_j \delta(i,j)$, $\delta(.,.) = \sum_i m_i \delta(i,.)$).

Ce n'est rien d'autre que la matrice introduite par Torgerson - [19] -, que l'on retrouve dans le schéma de dualité.

d étant l'indice de dissimilarité sur I, nous considèrerons respectivement en dimension (n-1) et n, les formes $b=b_o(d^2)$ et $\tau=\tau_o(G,k,d^2)$, $G \in F$, $k \in \mathbb{R}$.

Pour utiliser le théorème 2, reste le choix de produits scalaires sur W et V, pour induire des structures euclidiennes sur $\mathcal{F}(W)$ et $\mathcal{F}(V)$; il est guidé par l'exemple suivant : si d_1 est la distance sur I vérifiant : $\forall (i,j) \in I \times I$, $i \neq j$, $d_1(i,j)=1$, la forme $b_1=b_o(d_1)$, et donc la forme $\tau_1=\tau_o(G_1,h,d_1)$; $G_1 \in F, h>0$, est définie-positive, assurant par là l'existence d'un (n-1)-simplexe régulier ; ce sont ces produits scalaires, qui ne priviligient a priori aucun point, que nous considèrerons. Alors, conformément au théorème 2, on a, en dimension (n-1) l'unique problème :

$$P4 : \min_{f \in \mathcal{F}^+(W)} \|b-f\|_{b_1}$$

et en dimension n, les troix problèmes :

$$P5 : \min_{f \in \mathcal{F}^+(W)} \|\underline{\tau}-f\|_{\underline{\tau}_1} \; ; \; P6 : \min_{f \in \mathcal{F}_V^+(W)} \|\tau-f\|_{\tau_1} \; ; \; P7 : \min_{f \in \mathcal{F}^+(V)} \|\tau-f\|_{\tau_1}$$

En dimension n, les points sont aux sommets des vecteurs de base du référentiel. Aussi, toujours dans le but de ne privilégier a priori aucun point, on peut considérer le produit scalaire sur V, noté encore p, rendant encore ces vecteurs de base orthonormés ; et, conformément au théorème 2, on a les trois problèmes :

$$P8 : \min_{f \in \mathcal{F}^+(W)} \|\underline{\tau}-f\|_{\underline{p}} \; ; \; P9 : \min_{f \in \mathcal{F}_V^+(W)} \|\tau-f\|_{p} \; ; \; P10 : \min_{f \in \mathcal{F}^+(V)} \|\tau-f\|_{p}$$

Pour comparer les résultats, on a le :

Corollaire 1

$\forall G \in F$, $\forall G_1 \in F$, $\forall k \in \mathbb{R}$, $\forall h>0$, les problèmes P4, P5 et P8 ont même solution, qui est restriction à W des solutions des problèmes P6 et P9.
C'est encore la restriction à W de la solution du problème P7 (resp. P10) si $G=G_1$ (resp. si G est l'isobarycentre des points $M_i, i \in I$).

Démonstration :
Par construction de τ et τ_1, $\underline{\tau}=b$ et $\underline{\tau}_1=b_1$; de sorte que les problèmes P4 et P5 sont identiques. Clairement, on a : $\forall(i,j) \in I\times I, i\neq j, \frac{1}{2}p(\overrightarrow{M_iM_j},\overrightarrow{M_iM_j})=1$; donc, par la proposition 1, $\underline{p}=2b_1$ (plus généralement on a : $p=2\tau_1$, pour G_1 isobarycentre des M_i, et $h=1/2n$) ; en conséquence, les problèmes P4 et P8 sont équivalents, et P4, P5 et P8 ont même solution, soit q_*. Par le théorème 2, q_* est restriction à W des solutions de P6 et P9. Enfin, si $G=G_1$ (resp. G est l'isobarycentre des M_i) $\overrightarrow{OG}$ définit la direction τ_1-orthogonale (resp. p-orthogonale) de W ; comme $\overrightarrow{OG}$ et W sont τ-conjugués, par le même théorème, q_* est restriction à W des solutions de P7 et P10.

□

Sur le plan numérique, et opérant en dimension $(n-1)$, il est nécessaire de diagonaliser b dans une base b_1-orthonormée ; cet handicap est toutefois minime, puisque, d'une part, pour une bijection entre I et $\{0,..,(n-1)\}$ b s'écrit très facilement dans la base $\{\overrightarrow{M_0M_1},..,\overrightarrow{M_0M_{n-1}}\}$, et que d'autre part, il existe une transformation analytique connue entre cette base et une certaine base b_1-orthonormée -voir [6]-

Si on opère en dimension n, c'est naturellement le problème P10, avec G isobarycentre des M_i, qui est le plus simple à résoudre ; pour $k=0$, une diagonalisation de la matrice de Torgerson T pour l'indice d, donne, en annulant les valeurs propres négatives, la matrice T^* de la forme définissant l'approximation ; cette dernière est encore centrée, de sorte que c'est la matrice de Torgerson pour l'indice d_*.

En conséquence, si on désire poursuivre par une A.C.P. avec masses constantes, une simple diagonalisation de T suffit : on ne conserve que les vecteurs propres F_j associés à des valeurs propres $\lambda_j>0$. Les facteurs sont ces F_j normalisés à λ_j $-\ F_j'F_j = \lambda_j\ -$; et si on ne retient que les r premiers facteurs $F_1,..,F_r$, le pourcentage d'inertie conservée (par rapport à l'image euclidienne approchée) est égal à : $(\lambda_1+..+\lambda_r)/\mathfrak{J}$ où $\mathfrak{J} = \Sigma\{\lambda_j | \lambda_j>0\}$. On remarque que l'algorithme est en tout point semblable à celui donné lorsque l'on sait qu'existe une image euclidienne, excepté, peut-être, pour la valeur de $\mathfrak{J}$, qui ne peut être obtenue par la trace de T.

On retrouve ainsi, dans le cas particulier du corollaire 1, et pour l'A.C.P. à masses constantes, la démarche heuristique de Benzecri-[1]-et le résultat, présenté sous une autre forme d'optimalité, par Le Calvé-[17]-.

Par contre, si on désire poursuivre par une A.C.P. avec des masses distinctes, on ne peut utiliser l'opérateur d'Escoufier à partir de T^* précédemment définie. Les facteurs sont toutefois les vecteurs propres μ-normés, à la valeur propre, de $T^*\mu$ avec $\mu=D(\mathbb{I} - \frac{1}{m}\mathbb{1}\,\mathbb{1}'D)$ -où D est la matrice diagonale des masses m_i, $\mathbb{I}$ la matrice unité, $\mathbb{1}$ la matrice colonne n'ayant que des 1 pour composantes, $m=\Sigma m_i$ -. Si on veut les obtenir par l'opérateur d'Escoufier, on peut, soit, résoudre le problème P7 avec $G=G_1$ barycentre des M_i pour les masses considérées, ce qui nécessite la recherche d'une base τ_1-orthonormée (non connue analytiquement !), soit, plus simplement, chercher la matrice T^* précédente et la recentrer (en$(\mathbb{I} - \frac{1}{m}\mathbb{1}\,\mathbb{1}'D)\, T^*(\mathbb{I} - \frac{1}{m} D\,\mathbb{1}\,\mathbb{1}')$).

Remarque 1

On peut faire ici l'analogie avec l'approximation classique $d \longmapsto d'_*$ telle que : $\forall (i,j) \in I\times I$, $i\neq j$, $d'^2_*(i,j) = d^2(i,j) + c$, $c \in \mathbb{R}_+$, c minimum. On a $c=-2\lambda_m=\max(0,-\mu_m)$, où λ_m et μ_m sont respectivement les plus petites valeurs propres de T, et de b dans une base b_1-orthonormée - on peut d'ailleurs montrer que si $\{\mu_2,..,\mu_n\}$ est le spectre de b dans cette base, $\{0,2\mu_2,..,2\mu_n\}$ est le spectre de T -.

Donc, comme précédemment, il convient, en dimension (n-1), de rechercher une base b_1-orthonormée, et en dimension n, de considérer la matrice de Torgerson T. Si on poursuit par une A.C.P. à masses constantes, les facteurs sont les vecteurs propres F_j de T, orthogonaux à $\mathbb{1}$, associés à $\lambda_j>\lambda_m$; ils vérifient $F_j'F_j=\lambda_j-\lambda_m$; le pourcentage d'inertie conservée est égal à :

$(\lambda_1+..+\lambda_r-r\lambda_m)/[\mathrm{Trace}(T)-(n-1)\lambda_m]$ -voir [7], [13] -. Et, si on poursuit par une A.C.P. avec des masses distinctes, il convient encore, soit d'user de la matrice μ précédente, soit, pour utiliser l'opérateur d'Escoufier, de considérer, après re-centrage, la matrice de Torgerson à masses constantes de l'approximation (ou de diagonaliser τ dans une base τ_1-orthonormée).

Remarque 2

Malgré l'avantage de l'obtention directe d'une image euclidienne pour l'approximation, on peut regretter que cette dernière s'exprime en termes de formes quadratiques plutôt qu'en termes d'indices de dissimilarité - de telles formes quadratiques pouvant même être interprétées comme des mesures de similarité ! -. Certes, dans le cas particulier du corollaire 1, et considérant le problème P4, l'isomorphisme introduit dans la proposition 1, permet de transporter le produit scalaire de $\mathcal{F}(W)$ sur $\mathcal{D}_I$, et transforme $\mathcal{F}^+(W)$ en un cône convexe fermé, soit C ; de sorte que, au sens de cette structure euclidienne, d_*^2 est la projection orthogonale de d^2 sur C. Mais que tout ceci est artificiel ! Notons enfin que C n'est rien d'autre que le transformé par l'application $\delta \longmapsto \delta^2$ du cône fermé (généralement non convexe !) des semi-distances pour lesquelles existe une image euclidienne.

Exemple 1

Soit $I = \{1,2,3\}$, et soit d l'indice de dissimilarité sur I vérifiant :

$$d(1,2) = 1 \quad ; \quad d(1,3) = 2 \quad ; \quad d(2,3) = 4$$

Formant la matrice de Torgerson, à masses constantes, on obtient les valeurs propres : $\lambda_1=0$; $\lambda_2=(7+2\sqrt{21})/2$; $\lambda_3=(7-2\sqrt{21})/2$

Pour l'approximation d'_*, définie à la remarque 1, on a donc $c = 2\sqrt{21}-7$, et ainsi :

$d'_*(1,2)=(2\sqrt{21}-6)^{1/2}=1,78\ldots$;$d'_*(1,3)=(2\sqrt{21}-3)^{1/2}=2,48\ldots$;$d'_*(2,3)=(2\sqrt{21}+9)^{1/2}=4,26\ldots$

Pour notre approximation d_*, il est nécessaire, en général, de rechercher les vecteurs propres pour les valeurs propres strictement positives (ici, $x_1=1$, $x_2=4+\sqrt{21}$, $x_1+x_2+x_3=0$ définit un vecteur propre pour λ_2). Mais, dans ce cas particulier (une valeur propre positive et une valeur propre négative, d'ordre quelconques), il est évident que d_* et d'_* sont proportionnels et vérifient :

$$\frac{d_*^2}{d'^2_*} = \frac{\lambda_2}{\lambda_2-\lambda_3} \text{ ; de sorte que l'on a :}$$

$d_*(1,2)=\frac{1}{2}(2\sqrt{21}+2)^{1/2}=1,67..$;$d_*(1,3)=\frac{1}{2}(3\sqrt{21}+8)^{1/2}=2,33..$;$d_*(2,3)=\frac{1}{2}(7\sqrt{21}+32)^{1/2}=4,00..$

L'intérêt de cet exemple est moins de comparer d_* et d'_* (ils sont proportionnels!), que de les comparer à une approximation idéale, celle au sens des moindres carrés, que l'on ne sait résoudre en général, mais que l'on peut résoudre ici. En effet, si card(I)=3, il est aisé de voir, qu'il existe une image euclidienne de (I,d) si et seulement si d est une semi-distance. De sorte, que l'approximation au sens des moindres carrés est un problème quadratique, sous six contraintes linéaires : trois pour la positivité de l'indice, trois pour les inégalités triangulaires. Et une résolution numérique simple donne une approximation d''_* vérifiant :

$d''_*(1,2)=4/3=1,33\ldots$; $d''_*(1,3)=7/3=2,33\ldots$; $d''_*(2,3)=11/3=3,66\ldots.$

3 - APPROXIMATION POUR LA REPRESENTATION HIERARCHIQUE

Pour approcher un indice de dissimilarité, de nombreuses ultramétriques ont été proposées, comme l'ultramétrique sous-dominante, les ultramétriques supérieures minimales, et, plus récemment, par Chandon, Lemaire et Pouget - [9], [10] - les ultramétriques les plus proches au sens des moindres carrés. Nous présentons ici une approximation $d \longmapsto d_*$ satisfaisant à un critère des moindres carrés, sous des contraintes traduites par l'appartenance de d_* à une certaine préordonnance. Cette dernière a été introduite sous forme géométrique par Bertraneu - [4], [5] - et sous forme algébrique par Schader - [18] -. En opérant dans l'espace $\mathcal{D}_I$, nous la présentons plus simplement que dans [5], où - mais cela est justifié par le contexte de [5] - l'on opère dans l'espace $\mathcal{F}(W)$ de la figure 2 ; et nous la complétons par une propriété d'optimalité - présente dans la construction algébrique de [18] -.
Rappelons brièvement quelques notations et résultats. $\tilde{I}$ est la classe des sous-ensembles à deux éléments de I. Si d et d' sont deux indices de dissimilarité équivalents -i.e. vérifiant : $\forall\{i,j\},\{k,\ell\} \in \tilde{I}$: $d(i,j)\leq d(k,\ell) \Longleftrightarrow d'(i,j)\leq d'(k,\ell)$ - on note $d \sim d'$. on peut alors identifier l'ensemble des classes d'équivalence d'indices de dissimilarité et l'ensemble des préordonnances sur I - i.e. l'ensemble des préordres totaux sur $\tilde{I}$ - : à une classe donnée, correspond la préordonnance associée à un élément (quelconque) d de cette classe - i.e. la préordonnance définie par $\{i,j\}R\{k,\ell\} \Longleftrightarrow d(i,j)\leq d(k,\ell)$ -. On appelle indice de rang d'une préordonnance R, l'indice de dissimilarité r, où pour tout $\{i,j\}$ de $\tilde{I}$, $r(i,j)$ est le nombre de paires strictement inférieures à $\{i,j\}$, augmenté de 1.
On a $r \in R$; $r(i,j)$ est appelé rang de $\{i,j\}$. Une préordonnance est ultramétrique si un (= tout) indice lui appartenant est ultramétrique. Un indice de dissimilarité est dit propre si : $d(i,j) = 0 \Longrightarrow i=j$.

Définition 1

Soient R une préordonnance sur I, r son indice de rang et $r_1 = 1<r_2<\ldots<r_p$ les p rangs distincts des éléments de $\tilde{I}$.
$\forall \ell \in \{1,..,p\}$, soit d_ℓ l'indice de dissimilarité sur I défini par :
$\forall i \in I$, $d_\ell(i,i) = 0$; $\forall \{i,j\} \in \tilde{I}$, $d_\ell(i,j) = d_\ell(j,i) = \{ \begin{smallmatrix} 1 \text{ si } r(i,j) \geq r_\ell \\ 0 \text{ sinon} \end{smallmatrix}$
$d_1,..,d_p$ sont appelés indices de décomposition de R (ou de tout indice $d \in R$).

La suite $\{d_1,..,d_p\}$ est strictement décroissante, et forme un système libre de $\mathfrak{D}_I$; c'est une base si et seulement si R est une ordonnance - ordre total sur $\tilde{I}$ -. On a alors les :

Proposition 3

Soient R une préordonnance sur I et $\{d_1,..,d_p\}$ la suite décroissante de ses indices de décomposition. Pour que l'on ait $d \in R$, il faut et il suffit que d soit de la forme : $d=\alpha_1 d_1+\ldots+\alpha_p d_p$, $\alpha_1 \geq 0$, $\alpha_2,\ldots,\alpha_p>0$.
Si $d \in R$, $\alpha_1,\ldots,\alpha_p$ sont uniques et d est propre si et seulement si : $\alpha_1>0$.

Proposition 4

Soit $\{d_1,\ldots,d_p\}$ une suite finie strictement décroissante d'indices de dissimilarité à valeurs dans $\{0,1\}$ et telle que : $\forall \{i,j\} \in \tilde{I}$, $d_1(i,j)=1$; $d_p \neq 0$. Alors : il existe une préordonnance unique telle que $\{d_1,\ldots,d_p\}$ en soit la suite décroissante des indices de décomposition.

A l'isomorphisme près donné par la proposition 1, les démonstrations sont en tout point, semblables à celles données dans [5] ; pour une démonstration directe, voir aussi [14]. On a aussi la :

Proposition 5

Une préordonnance est ultramétrique si et seulement si tous ses indices de décomposition sont ultramétriques.

On a dans [5] une propriété analogue, exprimée en termes de positivité de formes quadratiques ; nous donnons ici une démonstration beaucoup plus courte.

Démonstration :

Avec les notations de la définition 1, on a :

R non ultramétrique $\Longleftrightarrow$ r non ultramétrique

$\Longleftrightarrow \exists i,j,k$ distincts de I : $r(i,j) > \max[r(i,k), r(j,k)]$

$\Longleftrightarrow \exists i,j,k$ distincts de I, $\exists \ell \in \{1,..,p\}$: $r(i,j) \geq r_\ell$, $r(i,k) < r_\ell$, $r(j,k) < r_\ell$

$\Longleftrightarrow \exists \ell \in \{1,..,p\}$, $\exists i,j,k$ distincts de I : $d_\ell(i,j)=1$, $d_\ell(i,k)=d_\ell(j,k)=0$

$\Longleftrightarrow \exists \ell \in \{1,..,p\}$: d_ℓ non ultramétrique.

□

<u>Remarque 3</u>

Ces trois dernières propositions, associées à la proposition 1, permettent de démontrer très simplement le théorème de Holman - [15] - assurant que si d est une distance ultramétrique, alors d'une part il existe une image euclidienne de (I,d) et d'autre part la dimension de (I,d) est égale à (n-1) - comme cela apparait dans [7].

Pour introduire une propriété d'optimalité donnons la :

<u>Définition 2</u>

Soient R et S deux préordonnances sur I. S est dite plus fine que R, si et seulement si :

$\forall \{i,j\},\{k,\ell\} \in \tilde{I}$: $\{i,j\}\ R\ \{k,\ell\} \Longrightarrow \{i,j\}\ S\ \{k,\ell\}$

On a alors la :

<u>Proposition 6</u>

Une préordonnance S est plus fine qu'une préordonnance R, si et seulement si, la suite des indices de décomposition de S est une sous-suite de la suite des indices de décomposition de R.

Démonstration :

Notons respectivement r et s les indices de rang de R et S, et $\{d_1,..,d_p\}$ la suite décroissante des indices de décomposition de R.

Supposons que la suite des indices de décomposition de S soit une sous-suite de celle de R, soit $\{d_{\ell_1},..,d_{\ell_q}\}$ - avec nécessairement $d_{\ell_1}=d_1$ -. Comme $r \in R$, on a :

$\forall \{i,j\},\{i',j'\} \in \tilde{I}$: $\{i,j\}R\{i',j'\} \Longleftrightarrow r(i,j) \leq r(i',j') \Longrightarrow \forall \ell=1,..,p\ d_\ell(i,j) \leq d_\ell(i',j')$

$\Longrightarrow \forall m=1,..,q\ d_{\ell_m}(i,j) \leq d_{\ell_m}(i',j')$.

Et comme $s \in S$, par la proposition 3 :
$\forall m=1,..,q \ d_{\ell_m}(i,j) \leq d_{\ell_m}(i',j') \Longrightarrow s(i,j) \leq s(i',j') \Longleftrightarrow \{i,j\}S\{i',j'\}$
Ainsi S est plus fine que R.
Réciproquement, si S est plus fine que R, notons $s_1=1,..,s_q$ les q rangs distincts des éléments de $\tilde{I}$ (au sens de S) et $\{\delta_1=d_1,..,\delta_q\}$ la suite décroissante des indices de décomposition de S. Soit k (quelconque) de $\{1,..,q\}$.
Soient $\{i,j\} \in \tilde{I}$ réalisant $\min\{r(i',j') \mid s(i',j')=s_k\}$ et $\ell \in \{1,..,p\}$ tel que $r(i,j)=r_\ell$.
Alors $\forall\{i',j'\} \in \tilde{I}$ sont vraies les implications suivantes :
$\delta_k(i',j')=0 \Longrightarrow s(i',j')<s(i,j) \Longrightarrow \{i,j\} \not S \{i',j'\} \Longrightarrow \{i,j\} \not R \{i',j'\} \Longrightarrow r(i',j')<r(i,j)$
$\Longrightarrow d_\ell(i',j') = 0$
Si $s(i,j)<s(i',j')$, comme précédemment : $r(i,j)<r(i',j')$
et si $s(i,j)=s(i',j')=s_k$, par construction de $\{i,j\}$, $r(i,j)\leq r(i',j')$.
Dès lors, sont encore vraies les implications suivantes :
$\delta_k(i',j')=1 \Longrightarrow s(i,j)\leq s(i',j') \Longrightarrow r(i,j)\leq r(i',j') \Longrightarrow d_\ell(i',j') = 1.$
D'où $\delta_k = d_\ell$ et la démonstration est complète.

□

Les propositions 4, 5 et 6 assurent le :

Corollaire 2

Si R est une préordonnance sur I, il existe une moins fine, parmi les préordonnances ultramétriques plus fines que R.
Elle admet pour indices de décomposition, les indices de décomposition ultramétriques de R.

(I,d) étant le doublet considéré, nous noterons R la préordonnance associée à d, r son indice de rang, $r_1=1<...<r_p$ les p rangs distincts (au sens de R) des éléments de $\tilde{I}$, $\{d_1,..,d_p\}$ la suite décroissante des indices de décomposition de R, $\{d_{\ell_1}=d_1,..,d_{\ell_q}\}$ la sous-suite des indices de décomposition ultramétriques de R, S la moins fine des préordonnances ultramétriques plus fines que R.
Imposer à notre approximation d_* d'appartenir à S, revient à imposer une contrainte de conservation au sens large sur les données (au sens où $d(i,j)<d(i',j') \Longrightarrow d_*(i,j) \leq d_*(i',j')$) ; et la propriété d'optimalité de S, montre que l'on ne transformera qu'un minimum d'inégalités strictes en égalités.

On peut donc poser le problème :

$$P11 : \min_{d' \in S} \sum \left\{ \left[d(i,j) - d'(i,j) \right]^2 \mid \{i,j\} \in \tilde{I} \right\}$$

L'algorithme de Kruskal - [16]- ne peut être a priori utilisé ici : on projette dans S et non dans la fermeture $\bar{S}$ de S, dont chaque face est une préordonnance plus fine que S. En outre, la spécificité du problème - même ordre au sens large pour d et d' - conduit à une forme analytique directe de la solution.
Le lecteur pourra remarquer que la démonstration traduit la démarche suivante : on projette d dans le sous-espace vectoriel engendré par S, pour ensuite constater que cette projection appartient à S ; ce qui est bien naturel, puisque - proposition 3 - à la nullité près du coefficient α_1, le cône convexe simplicial S est ouvert relatif de $\mathfrak{D}_I$.

Proposition 7

Le problème P11 admet une solution, soit d_*.
Si $\beta_1,..,\beta_q$ $(0 \leq \beta_1 < ... < \beta_q)$ sont les q valeurs distinctes de d_* sur l'ensemble des éléments de $\tilde{I}$, on a (posant $r_{\ell_{q+1}} = [n(n-1)]/2 + 1$) :

$$\forall m=1,..,q \qquad \beta_m = \frac{1}{r_{\ell_{m+1}} - r_{\ell_m}} \sum \left\{ d(i,j) \mid \{i,j\} \in \tilde{I} \ , \ r_{\ell_m} \leq r(i,j) < r_{\ell_{m+1}} \right\}$$

Démonstration :
Notant $\gamma_1,..,\gamma_p$ les p valeurs distinctes de d sur l'ensemble des éléments de $\tilde{I}$, on doit minimiser :

$$\sum \left\{ \left[d(i,j) - d'(i,j) \right]^2 \mid \{i,j\} \in \tilde{I} \right\} = \sum_{m=1}^{q} \sum_{\substack{k \\ r_{\ell_m} \leq r_k < r_{\ell_{m+1}}}} \sum \left\{ (\beta_m - \gamma_k)^2 \mid r(i,j) = r_k \right\} \equiv A \text{ (posé)}$$

sous la contrainte $0 \leq \beta_1 < \beta_2 < ... < \beta_q$.
Sans cette contrainte, les problèmes suivants sont équivalents :

$$\min A \Longleftrightarrow \min \sum_{m=1}^{q} \left[(r_{\ell_{m+1}} - r_{\ell_m}) \beta_m^2 - 2\beta_m \sum_{\substack{k \\ r_{\ell_m} \leq r_k < r_{\ell_{m+1}}}} (r_{k+1} - r_k) \gamma_k \right]$$

$$\Longleftrightarrow \min \sum_{m=1}^{q} (r_{\ell_{m+1}} - r_{\ell_m}) \left[\beta_m - \frac{1}{(r_{\ell_{m+1}} - r_{\ell_m})} \sum_{\substack{k \\ r_{\ell_m} \leq r_k < r_{\ell_{m+1}}}}^{q} (r_{k+1} - r_k) \gamma_k \right]^2$$

(puisque les quantités à minimiser sont toutes égales à des constantes additives près).

Mais, clairement, la solution du dernier problème a la forme donnée dans l'énoncé ; et cette solution satisfait à la contrainte : $0 \leq \beta_1 < \ldots < \beta_q$. D'où l'existence et la forme analytique de d_*.

□

Remarque 4

Notre approximation satisfait trivialement à deux propriétés, que l'on peut considérer comme souhaitables pour toute approximation, et qui sont vérifiées par exemple par la sous-dominante -[3]-. Notant respectivement d_* et d'_* les approximations de deux indices de dissimilarité d et d', et R, R', R_*, R'_* les préordonnances associées respectivement à d, d', d_*, d'_*, on a :

a) $d \sim d'$ entraine $d_* \sim d'_*$

b) R plus fine que R' entraine R_* plus fine que R'_*.

Sur le plan numérique, il est nécessaire de connaître les indices ultramétriques $d_{\ell_1}, \ldots, d_{\ell_q}$. Bertraneu - [5] - propose un algorithme de construction de S. Nous proposons un autre algorithme permettant tout à la fois de trouver les indices de décomposition ultramétriques et de construire d_*, par simple classement de nombres. Il repose sur une propriété de caractérisation de l'ultramétricité d'un indice de décomposition ; c'est la :

Proposition 8

Aves les notations de la définition 1, soit d un élément (quelconque) de R et soient $\gamma_1 < \ldots < \gamma_p$ les p valeurs distinctes de d sur les éléments de $\mathcal{T}$.
Définissons : $\forall \ell = 1, \ldots, p \quad a_\ell = \min \{\min_k \max[d(i,k), d(j,k)] \mid r(i,j) \geq r_\ell\}$

Alors :
$$\begin{cases} a_\ell < \gamma_\ell \Longleftrightarrow d_\ell \text{ non ultramétrique} \\ a_\ell = \gamma_\ell \Longleftrightarrow d_\ell \text{ ultramétrique} \end{cases}$$

Démonstration :

Clairement $a_\ell \leqslant \gamma_\ell$ (en choisissant $\{i,j\}$ tel que $d(i,j) = \gamma_\ell$).

Et alors on a $a_\ell < \gamma_\ell$ si et seulement si :

$\exists i,j,k \in I : \max\left[d(i,k)\ ,\ d(j,k)\right] < \gamma_\ell\ ;\ r(i,j) \geqslant r_\ell$

Or cette condition est équivalente à :

$\exists i,j,k \in I : \max\left[r(i,k)\ ,\ r(j,k)\right] < r_\ell\ ;\ r(i,j) \geqslant r_\ell$ (car $d \sim r$)

elle-même équivalente à :

$\exists i,j,k \in I : d_\ell(i,k) = d_\ell(j,k) = 0\ ;\ d_\ell(i,j) = 1$

et cette dernière condition n'est rien d'autre que celle exprimant que d_ℓ n'est pas ultramétrique.

□

Algorithme

Pour une bijection entre I et $\{1,..,n\}$, on donne la matrice D (symétrique, à termes diagonaux nuls), de terme général $d(i,j)$.

a) Former la matrice F (symétrique, à termes diagonaux nuls), de terme général :

$$f(i,j) = \min_k \max\left[d(i,k),\ d(j,k)\right]$$

b) Ordonner par ordre croissant (ou décroissant) les $\left[n(n-1)\right]/2$ valeurs de D, correspondant aux éléments $\{i,j\}$ de $\tilde{I}$ - on note $\overline{D}$ ce tableau - et ordonner les valeurs correspondantes de F de telle sorte que, quel que soit $\{i,j\}$ de $\tilde{I}$, $f(i,j)$ et $d(i,j)$ aient même rang - on note $\overline{F}$ ce tableau-.

c) Par un balayage descendant (ou ascendant) chercher les rangs (au sens de S) des différents éléments de $\tilde{I}$ - si $\overline{D}$ est ordonné par ordre croissant, $\forall p=1,..,\left[n(n-1)\right]/2$, p est un tel rang si et seulement si : $\overline{D}(p-1)<\overline{D}(p)$ et $\overline{D}(p)=F^*(p)$ avec $F^*(p) = \min\limits_{q\geqslant p} \overline{F}(q) = \min\left[F^*(p+1),\ \overline{F}(p)\right]$ - et construire d_* par moyennage de $\overline{D}$ entre deux rangs ainsi obtenus.

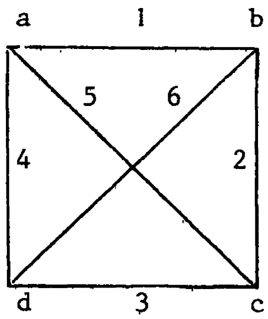
□

Le respect, au sens large, de l'ordre sur les données, est une contrainte très forte. Aussi, à moins que d soit proche d'une ultramétrique, la structure se dégageant de notre approximation ne peut qu'être faible - mais sûre ! -, voire inexistante, si d est très éloigné d'une ultramétrique - mais ne forcerait-on pas les choses à vouloir dégager une structure dans ce cas ? -.

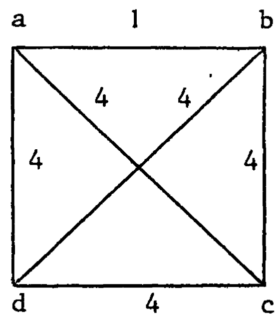
C'est ce que montrent les deux exemples suivants :

Exemple 2

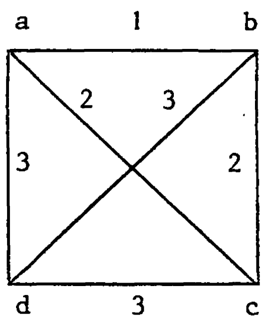
d, d_* et la sous dominante sont donnés par les diagrammes suivants :



indice d



indice d_*



indice sous-dominant

Exemple 3

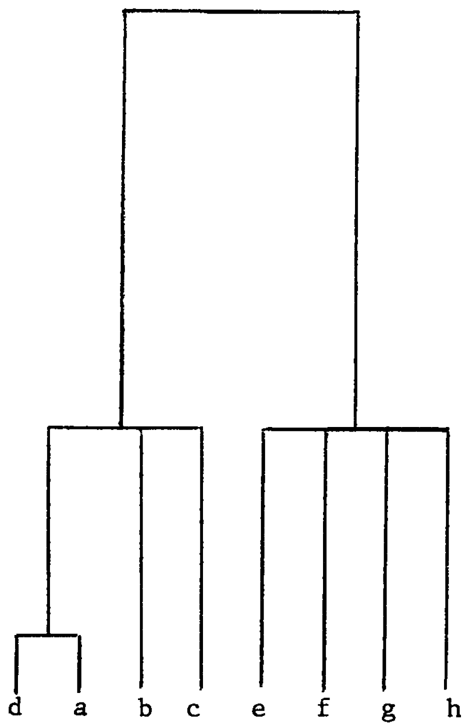
Les données sont celles d'un exemple de Chandon -[8] p. 22 - pour la recherche par un algorithme approximatif mais rapide, d'une ultramétrique la plus proche au sens des moindres carrés.

Triangle supérieur : d

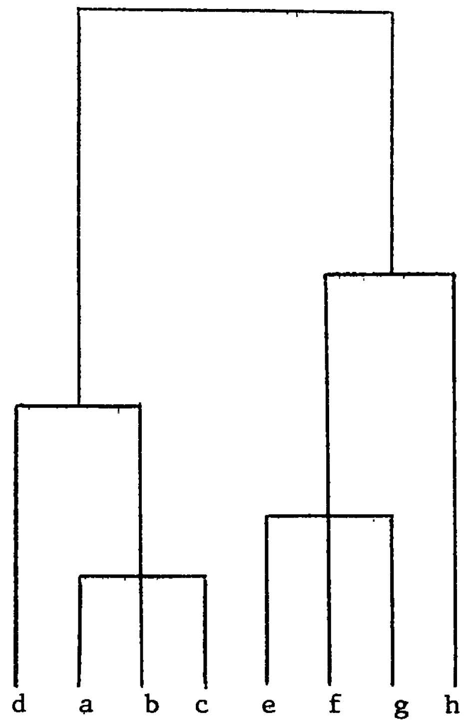
Triangle inférieur : d_*

	a	b	c	d	e	f	g	h
a	-	2	2	1	12	12	12	12
b	4,64	-	2	7	12	12	12	12
c	4,64	4,64	-	7	12	12	12	12
d	1	4,64	4,64	-	12	12	12	12
e	12	12	12	12	-	3	3	2
f	12	12	12	12	4,64	-	3	10
g	12	12	12	12	4,64	4,64	-	10
h	12	12	12	12	4,64	4,64	4,64	-

Ce qui donne pour hiérarchie, comparée à celle obtenue par Chandon :



hiérarchie pour d_*



hiérarchie proposée par Chandon

4 - REFERENCES

1 - BENZECRI J.P. *Sur l'analyse factorielle des proximités*. Publ. Inst. Stat. Univ. Paris. (1964), p. 235-282.

2 - BENZECRI J.P. *Représentation euclidienne d'un ensemble fini muni de masses et de distance*. ISUP (Sept. 1967).

3 - BENZECRI J.P. *L'analyse des données. Tome 1*. Dunod (1973).

4 - BERTRANEU A. *L'analyse des proximités : bases géométriques et méthode*. Thèse de 3ème cycle. Toulouse III (1979).

5 - BERTRANEU A. *Association univoque d'une préordonnance ultramétrique à une préordonnance donnée par une méthode géométrique*. Statistique et Analyse des données. Vol. 2 (1980).

6 - BERTRANEU, FICHET, FIESCHI. *Un algorithme d'analyse des proximités*. Secondes Journées Internationales Analyse des Données et Informatique. Versailles (Octobre 1979).

7 - CAILLIEZ et PAGES. *Introduction à l'Analyse des données*. Smash B.U.R.O. (1976).

8 - CHANDON J.L. *Construction de l'ultramétrique la plus proche au sens des moindres carrés : approximation versus optimisation*. I.A.E. Aix-Marseille (Mai 1978).

9 - CHANDON, LEMAIRE, POUGET. *Algorithme de branchement pour la détermination d'une ultramétrique la plus proche au sens des moindres carrés d'une dissimilarité*. Journées de Statistique. Nice (Mai 1978).

10 - CHANDON, LEMAIRE, POUGET. *Construction de l'ultramétrique la plus proche d'une dissimilarité au sens des moindres carrés*. R.A.I.R.O. Recherche Opérationnelle. Vol.14 n°2. (Mai 1980) p. 157-170.

11 - CRITCHLEY F. *Optimal norm characterisations of multidimensional scaling methods and some related data analysis problems*. Secondes Journées Internationales Analyse des Données et Informatique. Versailles (Octobre 1979).

12 - DROUET D'AUBIGNY G. *Description statistique des données ordinales : analyse multidimensionnelle*. Thèse 3ème cycle. Grenoble (1975).

13 - ESCOUFIER Y. *Cours d'analyse des données*. C.R.I.G. Montpellier (1978).

14 - GAUD E. *Sur la représentation hiérarchique en analyse des données*. D.E.A. Math. Appl. Université de Provence (Septembre 1980).

15 - HOLMAN W. *The relation between hierarchical and euclidian models for psychological distances*. Psychometrika - Vol. 37 n° 4 (Décembre 1972)

16 - KRUSKAL J.B. *Multidimensional scaling by optimizing goodness of fit to a non metric hypothesis*. Psychometrika Vol.29 (1964) p.1-7.

17 - LE CALVE G. *Quelques remarques sur certains aspects de l'analyse factorielle.* Université de Haute Bretagne, Rennes II. Cahier n° 2 (1976).

18 - SCHADER M. *Hierarchical Analysis : classification with ordinal object dissimilarities.* Metrika (1979).

19 - TORGERSON W.S. *Theory and methods of scaling.* New York. Wiley (1958).