R. ROUSSEAU

Reconnaissance de la structure de blocs d'un tableau par la classification ascendante hiérarchique

Les cahiers de l'analyse des données, tome 16, nº 2 (1991), p. 237-248

<http://www.numdam.org/item?id=CAD_1991__16_2_237_0>

Les Cahiers de l'Analyse des Données
Vol XVI - 1991 - n° 2 - pp. 237-248.

RECONNAISSANCE DE LA STRUCTURE DE BLOCS D'UN TABLEAU PAR LA CLASSIFICATION ASCENDANTE HIÉRARCHIQUE

[REC. BLOCS III]

*R. ROUSSEAU**

1 Introduction

1.1 Position du problème

Par structure de blocs, nous entendrons ici, la décomposition en blocs diagonaux la plus fine, laquelle, comme on le sait, existe et est unique. De façon précise, pour un tableau k_{IJ} sur $I \times J$, on dit que deux partitions sur I et J, indicées par un même ensemble B,

$$I = \cup \{ I_b \mid b \in B \} \quad ; \quad J = \cup \{ J_b \mid b \in B \} ;$$

définissent une partition en blocs diagonaux compatible avec le tableau k_{IJ} donné si :

$$\forall \ (i,j) \in I \times J : (k(i,j) \neq 0) \Rightarrow \exists \, b : (i \in I_b ; \ j \in J_b) .$$

Si on effectue sur l'ensemble I une classification ascendante hiérarchique, nous dirons que la structure en blocs a été reconnue par la CAH, si chacun des sous-ensembles I_b de la partition associée à la décomposition en blocs la plus fine est toujours une classe de la hiérarchie créée; auquel cas, toutes les classes créées sont, soit des parties d'un ensemble I_b convenable, soit une réunion de tels sous-ensembles.

Il semblerait assez souhaitable que la CAH reconnût toute structure de blocs; car cette structure est la plus typique, les individus étant déjà partitionnés en classes homogènes. Cependant, le fait d'agréger les individus deux par deux en commençant par le bas n'implique pas, comme nous le verrons, que l'on agrège entre eux systématiquement les éléments d'un même bloc.

(*) Vice-Recteur de l'Université Catholique de l'Ouest; Angers.

Cet article étudie la reconnaissance de la structure en blocs par quelques méthodes de classification ascendante hiérarchique. Le cas d'une structure en blocs dont chacun des blocs est réduit à un seul élément est évidemment éliminé. Les contre-exemples sont donnés par référence à des articles publiés ou se trouvent dans cet article.

1.2 Les méthodes de CAH étudiées

Une méthode de CAH dépend de l'indice de distance initial et du critère d'agrégation, eux-mêmes liés aux types de tableau traité. L'article présent étudie les méthodes dont l'indice de distance initial est choisi parmi les suivants (cf. [1] et [2]) :

A indices de distance sur tableau logique de présence-absence définis à partir d'indices de similitude n'ayant que les présences communes au numérateur;

B indices de distance sur tableau logique de présence-absence définis à partir d'indices de similitude n'ayant que les coïncidences au numérateur;

C distance euclidienne usuelle sur tableau quelconque;

D distance du χ^2 sur tableau de correspondance;

le critère d'agrégation étant choisi parmi les suivants (cf. [1] et [2]) :

1 critère du saut minimum;
2 critère du diamètre;
3 distance moyenne ou moyenne pondérée;
4a critère flexible $x > 0$;
4b critère flexible $x < 0$;
5 critère d'agrégation suivant la variance (ou de Ward);
6 moment centré d'ordre 2 minimum de la réunion de deux classes;
7 variance minimum de la réunion de deux classes;
8 distance minimum entre centres de gravité;
9 méthode minimum (médiane);
10 distance angulaire minimum .

Nous ne reprendrons pas ici l'explication de ces indices et critères.

2 Reconnaissance dans le cas des indices de distance classiques et d'un tableau logique en présence-absence

Considérons d'abord les indices de distance définis à partir d'indices de similitude n'ayant que les présences communes au numérateur (Jaccard, Dice, ...).

Proposition : Pour un tel indice, les CAH avec comme critère le saut minimum, la distance moyenne et la moyenne pondérée reconnaissent toute structure en blocs; tandis que les CAH avec comme critère le diamètre et le critère flexible avec x<0 ne reconnaissent que les structures en blocs dont les cardinaux de tous les blocs sont inférieurs ou égaux à 2; et que la CAH avec comme critère le critère flexible avec x>0 ne reconnaît aucune structure en blocs.

Preuve: en effet, pour deux individus de deux blocs différents, leur similitude est égale à 0 et leur distance vaut 1 et est maximale. Ainsi, lors de la première étape de la CAH, on agrègera ensemble deux individus d'un même bloc, sauf si le tableau est décomposé en blocs d'un seul individu.

Pour l'étape suivante, la distance entre le nœud constitué formé de deux éléments d'un même bloc et un élément d'un autre bloc sera aussi égale à 1 pour le saut minimum, le diamètre, la distance moyenne et la moyenne pondérée; mais la distance entre le nouveau nœud et un individu du même bloc peut aussi être égale à 1, ceci se produisant dans le cas où un des deux individus agrégés n'a pas de présence commune avec l'autre individu pour le diamètre et dans le cas où les deux individus agrégés n'ont pas de présence commune avec l'autre individu pour le saut minimum, la distance moyenne et la moyenne pondérée.

Ainsi toutes les distances entre le nouveau nœud et les autres individus du bloc associé à ce nœud seront égales à 1 si et seulement si pour le diamètre, tout élément du bloc associé au nœud n'a pas de présence commune avec l'un des éléments du nœud, et si et seulement si ,pour le saut minimum, la distance moyenne et la moyenne pondérée, tout élément du bloc associé au nœud n'a pas de présence commune avec les deux éléments du nœud ,i.e., si et seulement si les deux éléments du nœud forment à eux seuls un bloc ce qui est exclu.

$$\begin{array}{c|cccc|} 1 & 1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 1 & 0 \\ 3 & 0 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 1 \end{array}$$

contre-exemple 1 *pour l'indice de Jaccard*

Donc, pour le saut minimum, la distance moyenne et la moyenne pondérée, à chaque étape de la CAH, il existera une distance strictement inférieure à 1 entre nœuds d'un même bloc. La CAH reconnaît la structure en blocs pour ces trois critères et le contre-exemple est facile à trouver pour le diamètre (cf. contre-exemple 1 avec l'indice de Jaccard) dès que l'un des blocs est de cardinalité supérieure ou égale à 3.

Pour le critère flexible de Lance et Williams avec x>0, la distance entre le nœud de la première étape de la CAH et un élément d'un autre bloc sera strictement inférieure à 1. Il est alors facile de construire le contre-exemple (cf. contre-exemple 2 valable pour l'indice de Jaccard et pour x>1/9, et généralisable à x>0 quelconque).

```
1   | 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 |
2   | 0 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 |
3   | 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 |
4   | 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 |
```

contre-exemple 2 *pour l'indice de Jaccard*

La résolution du critère flexible avec x=0 est déjà faite puisque le critère est la moyenne pondérée dans ce cas particulier. Pour x<0, le fait qu'à la seconde étape, la distance entre le nouveau nœud et un élément d'un autre bloc soit supérieure strictement à 1 ne compense pas nécessairement l'augmentation des distances entre le nouveau nœud et les éléments de son propre bloc, sauf dans le cas évident où ce bloc n'a que deux éléments.

```
1   | 1 1 0 0 0 0 0 0 0 0 0 0 |
2   | 0 1 1 1 1 1 1 1 1 1 0 0 |
3   | 0 1 1 1 1 1 1 1 1 0 1 0 |
4   | 0 0 0 0 0 0 0 0 0 0 0 1 |
```

contre-exemple 3 *pour l'indice de Jaccard*

Le contre-exemple existe (cf. contre-exemple 3 valable pour l'indice de Jaccard et pour x=-(1/4), et généralisable à x<0 quelconque) et ainsi la décomposition en blocs diagonaux n'est pas reconnue par le critère flexible pour x < 0 si la cardinalité d'un bloc est supérieure ou égale à 3.

Proposition : Pour les indices de distance définis à partir d'indices de similitude n'ayant que les coïncidences au numérateur (Sokal-Michener, Rogers-Tanimoto), présence et absence jouent des rôles symétriques, donc, *a priori*, les CAH utilisant ces indices ne reconnaîtront pas la structure en blocs.

```
1   | 1 0 1 0 0 |
2   | 0 1 1 1 0 |
3   | 0 0 0 0 1 |
```

contre-exemple 4

Preuve : le contre-exemple est facile à trouver (cf. contre-exemple 4) avec un bloc de deux éléments et un autre bloc d'un élément, pour lequel on pourra agréger en premier, quel que soit le critère, deux éléments de deux blocs différents.

3 Tableaux de nombres réels

3.1 La distance euclidienne usuelle sur tableau quelconque

Le seul résultat général est le suivant.

Proposition : La seule CAH avec le critère distance angulaire minimum reconnaît la structure en blocs sur tableau positif.

Preuve : si le tableau est positif, tous les cosinus seront positifs et ainsi les distances entre nœuds seront comprises entre 0 et 1. D'autre part, pour deux éléments de blocs distincts, leur cosinus sera égal à 0 et leur distance égale à 1. On agrègera donc deux éléments d'un même bloc, sauf dans le cas particulier d'un tableau décomposable en autant de blocs qu'il y a d'individus. Les distances entre le nouveau nœud et les individus d'un bloc différent seront égales à 1 car le produit scalaire est nul. Ainsi, les blocs sont conservés et on se retrouve dans la situation précédente après agrégation de deux éléments d'un même bloc. Donc, la CAH reconnaîtra toute structure en blocs. Ce résultat n'est plus vrai si le tableau n'est pas positif comme le montre le contre-exemple 5.

```
1  | 1 0 |        1  | 1 0 |
2  |-3 0 |        2  | 3 0 |
3  | 0 1 |        3  | 0 1 |
```

contre-exemple 5 *contre-exemple* 6

D'autre part, le contre-exemple 6 montre que l'on agrègera en premier les éléments 1 et 3 de deux blocs différents pour tous les critères autres que la distance angulaire minimum.

3.2 La distance du χ^2 sur tableau de correspondance

La seule CAH reconnaissant toute structure en blocs est celle du critère distance angulaire minimum. En effet, vaut ici la démonstration générale donnée pour la distance euclidienne usuelle et le critère de la distance angulaire minimum. La non-reconnaissance de la structure en blocs est donnée ci-après pour tous les autres critères.

Proposition : La CAH avec comme critère celui de la variance (Ward) reconnaît la structure en blocs si chacun des blocs est de cardinalité inférieure ou égale à 3, ou, d'inertie inférieure ou égale à 1. D'autre part, si cette CAH donne K niveaux d'agrégation égaux à 1, alors le tableau est décomposable en au moins (K+1) blocs diagonaux.

Preuve : le fait que l'existence de K nœuds de niveau d'agrégation égal à 1 implique la décomposition en au moins K+1 blocs a déjà été démontré par J.-P. BENZÉCRI et P. CAZES en 1978 (cf. [3]) en utilisant le fait que la somme des K premières valeurs propres de l'AFC est supérieure ou égale à la somme des K derniers nœuds. Ce résultat peut se démontrer directement en utilisant la définition même de la distance entre nœuds à l'aide de la propriété suivante (cf. [4]) :

$$nv(I_1, I_2) = 1 \iff \text{le tableau } k_{IJ} \text{ se décompose en au moins deux blocs}$$

avec la partition sur I: $\{ I_1 , I_2 , I - (I_1 \cup I_2) \}$, où I_1 et I_2 sont deux parties disjointes de I (le troisième bloc pouvant être vide).

D'autre part, on démontre, (cf. [5] et [6]), que le problème de la reconnaissance de la structure en blocs se ramène à la propriété suivante d'un tableau de correspondance non décomposable en blocs:

$$\forall\ i \in I,\ \exists\, i' \in I\ :\ nv(i, i') < f_i\,(1 + d^2(f_J^{\,i}, f_J)) = f_i^{\,i}\ ;$$

équivalente à : $\forall\ i \in I,\ \exists\ i' \in I\ :\ f_{i'}^{\,i'} < f_i^{\,i} + 2 f_{i'}^{\,i}$

Mise sous cette forme, il est facile de démontrer que cette propriété est vraie pour tout tableau de correspondance non décomposable en blocs d'inertie inférieure ou égale à 1; ainsi que pour l'ensemble I d'un tableau de correspondance k_{IJ} tel que cardI soit inférieur ou égal à 3 quel que soit l'ensemble J (cf. [6]). Cependant, la propriété n'est plus nécessairement vraie si cardI est supérieur ou égal à 4 à l'aide d'un contre-exemple général (cf. [6]).

Ainsi, la reconnaissance de la structure en blocs par la CAH avec distance du χ^2 et critère de Ward est vérifiée si chacun des blocs satisfait à l'une au moins des deux conditions d'avoir une cardinalité inférieure ou égale à 3 ou une inertie inférieure ou égale à 1; mais n'est pas nécessairement vérifiée si l'un des blocs est de cardinalité supérieure ou égale à 4 (cf. contre-exemple général donné dans [6] et contre-exemples particuliers donnés dans [7] lorsque l'un des blocs est de cardinalité supérieure ou égale à 4 et lorsque l'autre bloc a une cardinalité inférieure ou égale à 3).

Proposition : La CAH avec comme critère le moment centré d'ordre 2 minimum de la réunion de deux classes reconnaît la structure en blocs si chacun des blocs est de cardinalité inférieure ou égale à 2.

Preuve : Les distances initiales étant les mêmes que pour le critère de Ward, les contre-exemples précédents du critère de Ward sont également valables. Pour deux blocs $I_1 = \{1,2\}$ et $I_2 = \{3,4\}$ de cardinal 2, à la première étape de la CAH, on agrègera comme pour le critère de Ward deux éléments d'un même bloc, soit par exemple 1 avec 2. Par des considérations d'inertie intra-classe et d'inertie inter-classe, on montre, dans le cas où I_1 et I_2 ne se décomposent pas en blocs, que:

$$nv(1 \cup 2, 3) = nv(1, 2) + nv(3, 4) + (1 - crit(1 \cup 2 \cup 3, 4))\ ;$$

où crit est l'écart selon le critère de Ward. Comme celui-ci est inférieur strict à 1, on a donc le résultat $nv(1 \cup 2,3) > nv(3,4)$. Par symétrie $nv(1 \cup 2,4) > nv(3,4)$. Donc, à la seconde étape de la CAH, on agrègera 3 avec 4. Le cas de deux blocs $I_1=\{1,2\}$ et $I_2=\{3\}$ étant déjà résolu (cf. critère de Ward), la CAH reconnaît la structure en blocs si chacun des blocs est de cardinalité inférieure ou égale à 2.

1	200	100	0	0	
2	100	200	0	0	
3	1	1	100	0	
4	0	0	0	100	

Contre-exemple 7

1	2000	1000	0	0	0
2	1000	2000	0	0	0
3	10	10	1000	0	0
4	0	0	0	1000	0
5	0	0	0	1	1000

Contre-exemple 8

Quant aux cas $1 \leq \text{card } I_1 \leq \text{card } I_2 = 3$, les contre-exemples 7, 8 et 9 prouvent que la structure en blocs n'est pas reconnue par cette CAH.

1	2000	1000	0	0	0	0
2	1000	2000	0	0	0	0
3	10	10	1000	0	0	0
4	0	0	0	1000	0	0
5	0	0	0	1	1000	0
6	0	0	0	0	1	1000

Contre-exemple 9

1	1990	10	0
2	1	0	0
3	0	0	30

Contre-exemple 10

Proposition : Les CAH, avec comme critères le saut minimum, le diamètre, la distance moyenne, la moyenne pondérée, le critère flexible, la distance minimum entre centres de gravité et la méthode minimum ne reconnaissent pas la structure en blocs.

Preuve : le contre-exemple 10 montre que la distance minimum entre éléments de I est atteint par le couple (1, 3) formé d'éléments de deux blocs différents. Donc, on agrègera à la première étape de la CAH deux éléments de deux blocs différents, et ainsi la structure en blocs n'est pas reconnue par ces CAH. Il est à noter que le contre-exemple général du critère de Ward (cf. [6]) l'est aussi pour les critères pré-cités dès que l'on a $n \geq p \geq 4$ et avec ε strictement positif quelconque.

Enfin, on a la proposition suivante:

Proposition : la CAH avec le critère de la variance minimum de la réunion de deux classes ne reconnaît pas la structure en blocs.

Preuve : Le contre-exemple général au critère de Ward (cf. [6]) vaut aussi pour le critère de la variance minimum de la réunion de deux classes mais avec des restrictions supplémentaires : $n \geq p \geq 6$ et ε strictement compris entre $(4(n+p-2)^{-1}(p-2)^{-1})$ et $(n+1)^{-1}$. De plus, le contre-exemple 11 avec $5=p \leq n=6$ (qui permet de construire le cas général $1 \leq p \leq n \leq 6$ en supprimant autant d'individus et de variables correspondantes qu'il le faut dans chacun des blocs tout en conservant les individus 1 et 7) montre que la structure en blocs n'est jamais reconnue par cette CAH.

1	100	0	0	0	0	0	0	0	0	0	0
2	1	100	0	0	0	0	0	0	0	0	0
3	0	1	100	0	0	0	0	0	0	0	0
4	0	0	1	100	0	0	0	0	0	0	0
5	0	0	0	1	100	0	0	0	0	0	0
6	0	0	0	0	1	100	0	0	0	0	0
7	0	0	0	0	0	0	300	0	0	0	0
8	0	0	0	0	0	0	1	300	0	0	0
9	0	0	0	0	0	0	0	1	300	0	0
10	0	0	0	0	0	0	0	0	1	300	0
11	0	0	0	0	0	0	0	0	0	1	300

Contre-exemple 11

3.3 Tableau récapitulatif

Le tableau croisé suivant indique si toute structure en blocs est reconnue par la méthode de CAH indice × critère (cf. notations du §1.2: lettres pour les indices, chiffres pour les critères) :

	1	2	3	4a	4b	5	6	7	8	9	10
A	V	*	V	F	*						
B	F	F	F	F	F						
C	F	F	F	F	F	F	F	F	F	F	*
D	F	F	F	F	F	*	*	F	F	F	V

(V) signifie que n'importe quelle structure en blocs est recxonnue par la CAH;
(F) signifie qu'il existe des structures en blocs non reconnues par la CAH;
(*) signifie que certaines structures en blocs particulières sont reconnues systématiquement par la CAH.

La méthode C × 10 (distance usuelle et distance angulaire minimum) reconnait la structure en blocs si le tableau est positif, et ne la reconnaît pas si le tableau est quelconque.

La méthode D × 5 (distance du χ^2 et critère de Ward) reconnaît la structure en blocs si chacun des blocs est de cardinalité inférieure ou égale à 3 et-ou d'inertie inférieure ou égale à 1. D'autre part, si cette méthode donne K nœuds de niveau d'agrégation égal à 1, alors le tableau est décomposable en au moins (K+1) blocs diagonaux.

Les méthodes D × 6 (distance du Chi2 et moment centré d'ordre 2 minimum de la réunion de deux classes), A × 2 et A × 4b (indices de distance sur tableau logique de présence-absence définis à partir d'indices de similitude n'ayant que les présences communes au numérateur; et le critère du diamètre ou le critère flexible avec x<0) reconnaissent la structure en blocs si chacun des blocs est de cardinalité inférieure ou égale à 2.

Ainsi, sur tableau non logique, il n'existe, à part la distance angulaire minimum sur tableau positif, parmi les méthodes de CAH étudiées aucune méthode reconnaissant la structure en blocs diagonaux, si elle existe. La seule s'en rapprochant quelque peu est la méthode distance du χ^2 et critère de Ward (le coefficient de masse permettant d'avoir la propriété pour des blocs de cardinalité inférieure ou égale à 3, mais étant dissuasif au-delà de 3). Deux possiblités s'offrent pour l'amélioration de cette méthode : soit définir un indicateur permettant de savoir si en cas d'existence la décomposition en blocs diagonaux est reconnue, soit modifier l'algorithme de façon à ce qu'il reconnaisse toute structure en blocs.

3.4 Note : Indicateur de reconnaissance de la décomposition en blocs (χ^2 - Ward)

Dans l'algorithme de la CAH distance du χ^2 et critère de Ward, on peut mettre un indicateur permettant de savoir si, à une étape k quelconque de la CAH, on a agrégé ensemble deux éléments orthogonaux (à produit scalaire nul) et tels que leur écart soit strictement inférieur à 1. On posera Ind(k) égal à 1 s'il y a agrégation à l'étape k de deux éléments à produit scalaire nul et à écart strictement inférieur à 1, et égal à 0 sinon. Si, à la fin de la CAH, l'indicateur $\mathrm{Ind}=\sup\{\mathrm{Ind}(k) / 1 \leq k \leq \mathrm{card}I-1\}$ est nul, alors la structure en blocs est reconnue si elle existe (cf. [4]) .

4 Algorithme modifié χ^2 - Ward

On peut modifier légèrement l'algorithme général de la CAH χ^2-Ward de façon à ce qu'il reconnaisse toute structure en blocs du tableau de correspondance, si elle existe, en empêchant l'agrégation directe entre individus orthogonaux en ajoutant la contrainte suivante :

On agrègera, à une étape donnée, les deux éléments i et i' à distance minimum et tels que leur produit scalaire soit non nul si leur écart est strictement inférieur à 1.

Ceci est justifié par le fait que la distance entre deux parties disjointes de I pour le critère de Ward et distance du χ^2 est égale à 1 si et seulement si ces deux parties et le complémentaire de leur réunion forment une partition de I associée à une décomposition en blocs diagonaux du tableau de correspondance (cf. [4]).

Tout se passe comme si on donnait pour écart entre deux éléments la valeur 1 s'ils sont orthogonaux et l'écart normal du critère de Ward s'ils ne le sont pas.

1	1	1/2	0	0	0
2	0	1/2	1/2	1/2	0
3	0	0	0	1/2	1
4	5	5	1/2	5	5

Contre-exemple 12

Sous cette forme, il est facile de voir que l'algorithme ne vérifie pas la formule de la médiane: nous le ferons à partir du contre-exemple 12 .

Notons crit' le critère de Ward modifié. On a donc crit'(1,3) = 1, et les deux relations suivantes sont vraies :

crit'(2,3) < crit'(1,2) = inf(crit'(1,2),crit'(1,3))

inf(crit'(1,2),crit'(1,3)) = crit'(1,2) > crit'(2u3,1)

et ces relations montrent que la formule de la médiane n'est pas vérifiée pour le critère crit'.

Sous la première forme (CAH avec contrainte), l'algorithme peut être accéléré et ainsi utiliser les méthodes de voisinages réductibles, de voisins réciproques et de recherche en chaîne de voisins réciproques, avec l'adjonction de la contrainte et la mise à jour simple du tableau des produits scalaires sc entre nœuds :

pour i et i' éléments de I :

sc(i, i') = si (produit scalaire entre i et i' $\neq$ 0) alors 1; sinon 0.

pour n, n' et n" nœuds de la hiérarchie :

sc(n $\cup$ n', n") = si (sc(n, n") = 1 ou sc(n', n") = 1) alors 1; sinon 0.

Le problème d'existence ou non d'inversions pour cet algorithme n'est pas encore résolu. L'inversion ne peut se produire que dans le cas d'agrégation, à une étape donnée, de deux nœuds orthogonaux (car sinon ce critère est celui de Ward qui n'admet pas d'inversions). Du fait que le centre de gravité d'une classe est le cumul des individus de cette classe, le cas d'inversion revient au problème d'agrégation à la première étape de la CAH de deux éléments orthogonaux.

Deux cas peuvent se présenter : soit il est possible d'agréger deux éléments orthogonaux et l'inversion est peut-être possible, soit cette agrégation est impossible et alors l'algorithme modifié revient à agglomérer dans ces nœuds les plus hauts les différentes CAH (distance du Chi2 et critère de Ward) des différents blocs, ce qui correspondrait à la situation idéale.

5 Conclusion

Donc, les seules CAH parmi celles étudiées, reconnaissant toute structure en blocs, sont caractériss comme suit :

- indices de distance sur tableau logique de présence-absence définis à partir d'indice de similitude n'ayant que les présences communes au numérateur avec le critère de saut minimum, de distance moyenne ou de moyenne pondérée.

- distance euclidienne usuelle sur tableau positif avec le critère de distance angulaire minimum

- distance du χ^2 sur tableau de correspondance avec le critère de distance angulaire minimum ou le critère de Ward modifié

Un problème reste encore ouvert : la CAH distance du χ^2 et critère de Ward modifié est-elle identique à la CAH distance du χ^2 et critère de Ward usuel, sur tout tableau de correspondance non décomposable en blocs ? Si la réponse est positive, la CAH distance du χ^2 et critère de Ward modifié ne possède pas d'inversions sans vérifier la formule de la médiane, et de plus admet, dans le cas d'une structure en blocs, comme sous-hiérarchies les hiérarchies associées aux différentes CAH des différents blocs.

Références bibliographiques

[1] J.L. CHANDON, S. PINSON: *Analyse typologique*; Masson, Paris; (1981).

[2] R.R. SOKAL, P.H.A. SNEATH: *Principles of Animal Taxonomy*; (1963); repris dans: J.-P. BENZÉCRI et coll.: *L'analyse Des Données*; (1973); et M. JAMBU, M.O. LEBEAUX: *Classification automatique*; (1978) .

[3] J.P. BENZÉCRI, P. CAZES: "Problème sur la classification"; in *C.A.D.*, Vol III, n°1, pp. 95-101; (1978).

[4] R. ROUSSEAU: "CAH distance du Chi2 - critère de Ward et tableau de correspondance décomposé en blocs"; in *Les Cahiers de l'IMA* , Cl 87-3, Juin 1987.

[5] I. KHARCHAF, R. ROUSSEAU: "Reconnaissance de la structure de blocs d'un tableau de correspondance par la CAH"; in *C.A.D.*, Vol XIII, n°4, pp. 439-443; (1988).

[6] R. ROUSSEAU: "Reconnaissance de la structure de blocs d'un tableau de correspondance par la CAH (suite)"; in *C.A.D.*, Vol XIV, n°3, pp. 257-266; (1989).

[7] R. ROUSSEAU: "Reconnaissance de la structure de blocs d'un tableau de correspondance par la CAH: contre-exemple complémentaire"; in *C.A.D.*, Vol XIV, n°3, pp. 377-378; (1989).