Les Cahiers de l'Analyse des Données
Vol. XIX - 1994 - n°3 - pp.343-352.

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

[REC. BLOC. IV]

*R. ROUSSEAU**

1 Introduction

1.1 Position du problème

Par structure de blocs, nous entendrons ici, la décomposition la plus fine d'un tableau rectangulaire en blocs diagonaux; décomposition qui, comme on le sait, existe et est unique pour tout tableau.

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,

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

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 une classe de la hiérarchie créée; auquel cas, toutes les classes créées sont, soit des parties d'un sous-ensemble I_b convenable, soit une réunion de tels sous-ensembles.

(*) Vice-Recteur de l'Université Catholique de l'Ouest; B.P. 80849008 Angers CEDEX 01.

1.2 Rappel des résultats déjà obtenus

Au problème de la reconnaissance de la structure de blocs, a déjà été consacrée une suite de quatre articles dont les références sont rappelées à la fin de celui-ci.

L'article [4] portant sur la reconnaissance de la structure de blocs diagonaux d'un tableau par quelques méthodes de Classification Ascendante Hiérarchique a montré que, dans le cadre de ces méthodes, la structure en blocs n'est reconnue systématiquement que dans certains cas qui sont délimités à la fois par la nature du tableau, la formule de distance et le critère d'agrégation employés:

A) tableau logique en présence-absence:

indices de distance définis à partir d'indices de similitudes n'ayant que les coïncidences au numérateur;

critères soit du saut minimum, soit de la distance moyenne, soit de la moyenne pondérée;

B) tableau de nombres positifs:

distance euclidienne quelconque;

critère de la distance angulaire minimum; et donc, par extension, tout critère basé sur le produit scalaire;

C) tableau de nombres positifs, dans lequel chacun des blocs diagonaux est de cardinalité inférieure ou égale à 3 ou d'inertie inférieure ou égale à 1:

distance euclidienne du χ^2;

critère d'agrégation suivant la variance (dit encore de WARD).

Hormis le cas B), parmi les méthodes classiques de CAH s'appliquant à un tableau de nombres positifs, la plus efficace apparaît être C): conjonction de la distance du χ^2 et du critère de WARD.

Dans l'article [4], une autre méthode de CAH a été définie pour un tableau de nombres positifs : distance du χ^2 et critère de WARD modifié en ajoutant la contrainte suivante “on agrègera à une étape donnée les deux éléments à écart minimum à condition que leur produit scalaire ne soit pas nul si leur écart est strictement inférieur à 1”.

Cette contrainte supplémentaire permet de reconnaître, en toute généralité, la structure en blocs diagonaux. Cependant, la méthode ne vérifie pas l'axiome de la médiane; lequel, comme on le sait, assure que la CAH ne présente pas d'inversion dans la suite des niveaux d'agrégation.

1.3 Objet du présent article

Une question restait donc ouverte: la CAH fondée sur distance du χ^2 et critère de WARD modifié est-elle, sur tout tableau de correspondance non décomposable en blocs diagonaux, identique à la CAH fondée sur distance du χ^2 et critère de WARD usuel ?

en d'autres termes: la CAH, fondée sur distance du χ^2 et critère de WARD modifié, donne-t-elle, pour un tableau de correspondance décomposé en blocs diagonaux, la décomposition en blocs; puis, pour chaque bloc, les résultats que donnerait la CAH fondée sur distance du χ^2 et critère de WARD usuel ?

ou encore: la CAH fondée sur distance du χ^2 et critère de WARD modifié peut-elle être considérée comme une variante de la CAH fondée sur distance du χ^2 et critère de WARD usuel; ne différant de celle-ci qu'en ce qu'elle reconnaîtrait à coup sûr toute décomposition en blocs diagonaux?

L'objet du présent article est d'apporter une réponse négative à cette question ouverte.

2 Cas particuliers : cardI = 3 ou cardI = 4 ; ou: inertie totale inférieure ou égale à 1

N.B. Au lieu de: CAH fondée sur la distance du χ^2 avec pour critère l'agrégation suivant la variance, nous écrirons, en bref: CAH{χ^2, WARD}; et, de même, s'il s'agit du critère modifié: CAH{χ^2, WARD modifié}.

Soit k_{IJ} un tableau de correspondance non décomposable en blocs diagonaux tel que cardI = 3 ou cardI = 4 ; ou tel que son inertie totale soit inférieure ou égale à 1.

Alors, s'il existe deux éléments de I à produit scalaire nul pour la distance du χ^2, ils ne seront jamais agrégés entre eux lors de la première étape de la CAH{χ^2, WARD}.

Preuve pour cardI = 3

Le tableau k_{IJ} n'étant pas décomposable en blocs diagonaux, le troisième élément de I, noté i_3, a un produit scalaire non nul avec chacun des deux autres éléments de I, de produit scalaire nul, notés i_1 et i_2. On a la relation suivante :

$$\begin{aligned}\mathrm{Crit}(i_1, i_2) &= ((f_{i_1} \,.\, f_{i_2} / (f_{i_1} + f_{i_2}))\,.\, d^2(i_1, i_2)\\ &= ((1/f_{i_1}) + (1/f_{i_2}))\,.\,\, (d^2(i_1, 0_J) + d^2(i_2, 0_J)) \;\;;\end{aligned}$$

où 0_J désigne l'origine de l'espace vectoriel R_J des mesures portées par l'ensemble J.

Reprenant l'idée de la démonstration donnée au §3 *in* [1], on peut montrer que, sous l'hypothèse:

$$f_{i_1} . d^2(i_1, 0_J) \leq f_{i_2} . d^2(i_2, 0_J) \quad ;$$

est vraie l'appartenance:

$$Crit(i_1, i_2) \in [f_{i_1} . d^2(i_1, 0_J) , f_{i_2} . d^2(i_2, 0_J)] \; .$$

Or:

$$f_i . d^2(i, 0_J) = f_i . (1+d^2(i, I)) = f_i^{\,i} \quad ; \text{ (cf §1 } in \text{ [2]).}$$

Donc:

$$Crit(i_1, i_2) \in [f_{i_1}^{\,i_1}, f_{i_2}^{\,i_2}].$$

Or la propriété suivante est vraie quand cardI = 3, ou quand l'inertie totale est inférieure à 1 (cf §3 *in* [2]) :

$$\forall i \in I, \; \exists i' \in I : Crit(i, i') < f_i^{\,i} .$$

Donc, pour $i = i_1$, il existe $i' \in I$ tel que $Crit(i_1, i') < f_{i_1}^{\,i_1} \leq Crit(i_1, i_2)$. L'inégalité stricte précédente montre que i' ne peut être i_2 ; et donc est i_3. Donc $Crit(i_1, i_3) < Crit(i_1, i_2)$.

Et ainsi, quelle que soit la valeur de $Crit(i_2, i_3)$, i_1 et i_2 ne s'agrégeront jamais lors de la première étape de la CAH lorsque cardI = 3.

Preuve pour une inertie totale ≤ 1 et cardI quelconque

Notons encore i_1 et i_2 deux éléments dont le produit scalaire est nul. Supposons de même que $f_{i_1}^{\,i_1} \leq f_{i_2}^{\,i_2}$. Alors $Crit(i_1, i_2) \in [f_{i_1}^{\,i_1}, f_{i_2}^{\,i_2}]$ et il existe $i_3 \in I$ distinct de i_2 et i_1, tel que $Crit(i_1, i_3) < f_{i_1}^{\,i_1} \leq f_{i_2}^{\,i_2}$.

Deux cas se présentent : ou bien le produit scalaire de i_1 et i_3 est non nul et alors i_1 et i_2 ne s'agrègeront pas ensemble lors de la première étape de la CAH; ou bien le produit scalaire de i_1 et i_3 est nul.

Dans ce dernier cas, $Crit(i_1, i_3) < f_{i_1}^{\,i_1}$ implique: $f_{i_3}^{\,i_3} < f_{i_1}^{\,i_1}$; et donc, par le même raisonnement, on obtient l'existence de $i_4 \in I$ tel que $Crit(i_3, i_4) < f_{i_3}^{\,i_3} < f_{i_1}^{\,i_1} \leq f_{i_2}^{\,i_2}$. Deux cas se présentent encore: le produit scalaire de i_3 et i_4 est non nul et alors i_1 et i_2 ne s'agrègeront pas ensemble lors de la première étape de la CAH; ou bien le produit scalaire de i_3 et i_4 est nul, et alors i_4 est distinct de i_3, i_2 et i_1 car on a: $f_{i_4}^{\,i_4} < f_{i_3}^{\,i_3} < f_{i_1}^{\,i_1} \leq f_{i_2}^{\,i_2}$ ...

Supposons que l'on poursuive ainsi cette chaine $i_3, i_4, i_5, \ldots, i'_x, i_x$ jusqu'au dernier élément i_x de l'ensemble I; avec deux éléments consécutifs à produit scalaire nul, on aura la suite d'inégalités suivantes :

$$f_{i_x}{}^{i_x} \leq \mathrm{Crit}(i_x, i'_x) < f_{i'_x}{}^{i'_x} < \ldots < f_{i_5}{}^{i_5} < f_{i_4}{}^{i_4} < f_{i_3}{}^{i_3} < f_{i_1}{}^{i_1} \leq f_{i_2}{}^{i_2} \;;$$

De plus, il existe $i \in I$ tel que: $\mathrm{Crit}(i, i_x) < f_{i_x}{}^{i_x}$. Supposons que le produit scalaire de i et i_x soit nul, alors, l'élément i vérifierait $f_i{}^1 \leq \mathrm{Crit}(i, i_x) < f_{i_x}{}^{i_x}$; et donc :

$$f_i{}^i < f_{i_x}{}^{i_x} < f_{i'_x}{}^{i'_x} < \ldots < f_{i_5}{}^{i_5} < f_{i_4}{}^{i_4} < f_{i_3}{}^{i_3} < f_{i_1}{}^{i_1} \leq f_{i_2}{}^{i_2} \;;$$

alors même que i est l'un des éléments de $\{i_x, i'_x, \ldots, i_5, i_4, i_3, i_2, i_1\}$. D'où une contradiction.

Donc, i et i_x ont un produit scalaire non nul et vérifient $\mathrm{Crit}(i, i_x) < \mathrm{Crit}(i_1, i_2)$. Ainsi, i_2 et i_1 ne peuvent s'agréger lors de la première étape de la CAH.

Preuve pour cardI = 4

Posons: $I = \{1, 2, 3, 4\}$. Supposons que la paire (1, 2) à écart minimum lors de la première étape de la CAH soit telle que 1 et 2 soient à produit scalaire nul (ce qu'on notera: sc(1, 2) = 0).

On a, pour i et i' $\in$ I, les relations suivantes :

$$\begin{aligned} \mathrm{Crit}(i, i') &= ((f_i.f_{i'} / (f_i+f_{i'})) \,.\, d^2(i, i') \\ &= ((f_i.f_{i'} / (f_i+f_{i'})) \,.\, (d^2(i, 0_J) + d^2(i', 0_J) - 2.\mathrm{sc}(i, i')) \;; \\ \mathrm{Crit}(i, i') &= (1/(f_i+f_{i'})) \,.\, (f_{i'}.f_i{}^i + f_i.f_{i'}{}^{i'} - 2.f_i.f_{i'}.\mathrm{sc}(i, i')] \;. \end{aligned}$$

Pour i = 1 et i' = 2, on a donc $\mathrm{Crit}(1, 2) = (1/(f_1+f_2)).\ (f_2.f_1{}^1+ f_1.f_2{}^2)$; d'où en supposant de plus que $f_1{}^1 \leq f_2{}^2$, on obtient (d'une autre manière que dans la preuve pour CardI = 3) que $\mathrm{Crit}(1, 2) \in [f_1{}^1; f_2{}^2]$. On a de plus les relations suivantes :

$$\begin{aligned} \mathrm{Crit}(1, 3) &= (1/(f_1+f_3)) \,.\, (f_3.f_1{}^1+ f_1.f_3{}^3 - 2.f_1.f_3.\mathrm{sc}(1,3)) \;; \\ \mathrm{Crit}(1, 3) - f_1{}^1 &= (f_1 /(f_1+f_3)) \,.\, (f_3{}^3 - f_1{}^1 - 2.f_3.\mathrm{sc}(1,3)) \;; \\ \mathrm{Crit}(1, 4) - f_1{}^1 &= (f_1 /(f_1+f_4)) \,.\, (f_4{}^4 - f_1{}^1 - 2.f_4.\mathrm{sc}(1,4)) \;; \end{aligned}$$

La paire (1, 2) étant à écart minimum, $\mathrm{Crit}(1, 3) \geq \mathrm{Crit}(1, 2) \geq f_1{}^1$, donc $\mathrm{Crit}(1, 3) - f_1{}^1 \geq 0$, donc on a l'inégalité $f_3{}^3 \geq f_1{}^1 + 2.f_3.\mathrm{sc}(1,3)$. De même, on a: $f_4{}^4 \geq f_1{}^1 + 2.f_4.\mathrm{sc}(1,4)$. Ainsi:

$$f_3{}^3 + f_4{}^4 \;\geq\; 2.f_1{}^1 + 2.f_3.\mathrm{sc}(1, 3) + 2.f_4.\mathrm{sc}(1, 4) \;.$$

Or: $\quad 1 = \mathrm{sc}(f_J{}^1, f_J)$

$$= f_1.\mathrm{sc}(f_J{}^1, f_J{}^1) + f_2.\mathrm{sc}(f_J{}^1, f_J{}^2) + f_3.\mathrm{sc}(f_J{}^1, f_J{}^3) + f_4.\mathrm{sc}(f_J{}^1, f_J{}^4) \;;$$

donc on a l'égalité: $1 = f_1{}^1 + f_3.sc(1, 3) + f_4.sc(1, 4)$.

Ainsi, $f_3{}^3 + f_4{}^4 \geq 2$, donc $f_3{}^3 = f_4{}^4 = 1$, donc $\forall\ i \neq 3 : f_3{}^i = 0$, donc 3 est orthogonal à 1, 2 et 4. De même, 4 est orthogonal à 1, 2 et 3. Sachant que 1 et 2 sont orthogonaux entre eux, on a donc les quatre éléments de I orthogonaux deux à deux et une décomposition du tableau en 4 blocs diagonaux.

On en conclut donc que si le tableau pour cardI=4 n'est pas décomposable en blocs, on ne peut agréger lors de la première étape de la CAH deux éléments dont le produit scalaire est nul.

3 Corollaire des résultats démontrés au §2

Pour tout tableau k_{IJ} de correspondance non décomposable en blocs diagonaux tel que cardI ≤ 4 ou tel que son inertie totale soit inférieure ou égale à 1, la CAH$\{\chi^2$, WARD$\}$ n'agrège jamais ensemble deux éléments de produit scalaire nul.

Preuve

Le corollaire est évident pour cardI = 3 d'après le résultat obtenu au §2.

Lorsque l'inertie totale est inférieure ou égale à 1, deux éléments à produit scalaire nul ne sont pas agrégés à la première étape de la CAH; donc, en considérant le tableau déduit du tableau initial par cumul des deux éléments agrégés (les deux lignes correspondantes sont remplacées par une seule ligne qui est la somme de celles-ci), les deux éléments dont le produit scalaire est nul conservent cette propriété dans un tableau d'inertie totale inférieure ou égale à 1 (car inférieure à l'inertie du tableau initial) et non décomposable en blocs (car sinon le tableau initial le serait). Donc les deux éléments dont le produit scalaire est nul ne seront pas agrégés ensemble à la deuxième étape de la CAH, etc.

Le cas cardI = 4 se résoud de la même façon.

4 Cas Général, cardI ≥ 8

Pour un tableau de correspondance non décomposable en blocs diagonaux tel que cardI ≥ 8, deux éléments dont le produit scalaire est nul peuvent être agrégés ensemble par la CAH$\{\chi^2$, WARD$\}$.

Autrement dit, lorsque cardI ≥ 8, il n'est pas vrai que la CAH$\{\chi^2$, WARD modifié$\}$ soit identique à la CAH$\{\chi^2$, WARD$\}$ sur tout tableau de correspondance non décomposable en blocs diagonaux.

Preuve

En reprenant l'exemple 6.2 *in* [2] et en ajoutant un nombre positif η arbitrairement petit dans la case (1, n+p-2), on obtient le tableau suivant (où les cases laissées vides sont, en fait, nulles) :

	1	2	3	4	5	.	.	.	n-1	n	n+1	.	.	.	n+p-2
1	1														η
2		1													
3			1												
4				1											
5					1										
.						1									
.							1								
.								1							
n-1									1						
n	ε	ε	ε	ε	ε	ε	ε	ε	ε						
n+1										1					
.											1				
.												1			
.													1		
.														1	
n+p-1															1
n+p										ε	ε	ε	ε	ε	ε

On suppose $n \geq p \geq 4$ et $\varepsilon < (p-3)/(p-1)$. Ce tableau est décomposable en deux blocs diagonaux si $\eta = 0$ et n'est pas décomposable en blocs diagonaux si $\eta \neq 0$. Pour $\eta = 0$, on sait que l'écart minimum interblocs (obtenu entre n et n+p) est strictement inférieur à l'écart minimum intrablocs (obtenu entre n+p et un élément quelconque *i* de [n+1, n+p-1]).

La démonstration suivante est basée sur les approximations par développements limités, car il est pratiquement impossible de calculer les écarts réels entre points de I pour $\eta \neq 0$ et donc d'en déterminer exactement les écarts les plus faibles.

En notant par Crit le critère de WARD usuel, on a les égalités :

$$\mathrm{Crit}(i, i') = ((f_i.f_{i'} / (f_i+f_{i'})) . d^2(i, i')$$

$$\mathrm{Crit}(i, i') = ((k(i).k(i') / (k(i)+k(i'))) \times \ldots$$

$$\Sigma\{ (1/k(j)).((k(i, j)/k(i)) - (k(i', j)/k(i')))^2 \mid j \in J \} .$$

Notons k_{IJ} le tableau avec $\eta = 0$ et k^*_{IJ} le tableau avec $\eta \neq 0$. On marquera d'une ' *' les termes calculés sur k^*_{IJ}.

En prenant $\eta > 0$ négligeable, on obtient les caractéristiques suivantes pour les CAH sur I =]n+p] :

A) la marge sur J = [1; n+p-2] est inchangée sauf pour j = n+p-2 :

$$\forall\, j \in [1; n+p-3] \;:\; k^*(j) = k(j) = 1 + \varepsilon \;;$$

$$k^*(n+p-2) \;=\; 1+\varepsilon+\eta \;=\; k(n+p-2)+\eta \;=\; k(n+p-2)+o(\eta)\;;$$

B) la marge sur I est inchangée sauf pour $i = 1$:

$$\forall\, i \in [2;\, n+p] \;:\; k^*(i) = k(i) \quad ;$$

$$k^*(1) = 1+\eta = k(1)+\eta = k(1)+o(\eta) \quad ;$$

C) les écarts entre deux éléments quelconques de $[2;\, n+p-2]$ sont inchangés :

$$\forall\, i \in [2;\, n+p-2],\ \forall\, i' \in [2;\, n+p-2] \;:\; \mathrm{Crit}^*(i, i') = \mathrm{Crit}(i, i') \quad ;$$

D) les écarts entre $n+p-1$ et un élément quelconque $i \in [2;\, n+p-2]$ diminuent, quand on passe de Crit* à Crit, de :

$$(k(i)\,/\,(1+k(i)))\,.\,((1/(1+\varepsilon)) - (1/(1+\varepsilon+\eta)))$$
$$= ((k(i)\,/\,(1+k(i)))\,.\,(\eta\,/\,(1+\varepsilon).(1+\varepsilon+\eta)) \quad ;$$

soit: $o(\eta)$;

E) les écarts entre $n+p$ et un élément quelconque $i \in [2;\, n+p-2]$ diminuent, quand on passe de Crit* à Crit, de :

$$((p-1).\varepsilon.k(i)/((p-1).\varepsilon+k(i))).((1/(1+\varepsilon)) - (1/(1+\varepsilon+\eta))).\ (1/(p-1)^2)$$
$$((p-1).\varepsilon.k(i)/((p-1).\varepsilon+k(i))).(\eta/(1+\varepsilon)(1+\varepsilon+\eta)).(1/(p-1)^2)$$

soit: $o(\eta)$;

F) l'écart entre $n+p$ et $n+p-1$ diminue, entre Crit* et Crit de :

$$((p-1)\varepsilon\,/\,(1+(p-1)\varepsilon))\,.\,(\,\eta\,/\,(1+\varepsilon)(1+\varepsilon+\eta))\,.\,((p-2)/(p-1))^2 \quad ;$$

soit: $o(\eta)$;

G) l'écart entre 1 et un autre élément i quelconque de $[2;\, n+p]$ subit une variation plus difficile à cerner quant à son sens, augmentation ou diminution:

le coefficient de masse devient $[(1+\eta)k(i)/(1+\eta+k(i)] = [k(i)/(1+k(i)] + o(\eta)$;

sous le signe Σ, les termes correspondant à $j \in [2;\, n+p-3]$ restent inchangés ;

le terme $(1/(1+\varepsilon))\,.\,(1 - (k(i,1)/k(i)))^2$, correspondant à $j = 1$ sous le signe Σ, est remplacé par:

$$(1/(1+\varepsilon))\,.\,((1/(1+\eta)) - (k(i,1)/k(i)))^2$$
$$= (1/(1+\varepsilon))\,.\,(1 - (k(i,1)/k(i)) + o(\eta))^2 \quad ;$$

qui est égal à celui-ci augmenté de $o(\eta)$;

le terme: $(1/(1+\varepsilon)) \cdot (k(i, n+p-2)/k(i)))^2$, correspondant à j=n+p-2 sous le signe Σ, est remplacé par :

$$(1/(1+\varepsilon+\eta)) \cdot ((\eta/(1+\eta)) - (k(i, n+p-2)/k(i)))^2$$
$$= (1/(1+\varepsilon)) \cdot (1 + o(\eta)) \cdot ((k(i, n+p-2) / k(i)) + o(\eta))^2 ;$$

qui est égal à celui-ci plus $o(\eta)$; donc:

$$\mathrm{Crit}^*(1, i) = \mathrm{Crit}(1, i) + o(\eta)$$

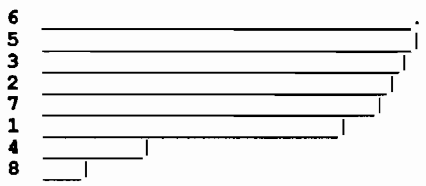
Ainsi, en prenant η suffisamment petit, les écarts entre points de I varieront insuffisamment pour changer la paire d'éléments à écart minimum, à savoir la paire (n, n+p). Donc, à la première étape de la CAH{χ^2, WARD} , on agrègera n et n+p, soit deux éléments à produit scalaire nul; tandis que ceux-ci ne pourront s'agréger ensemble pour la CAH{χ^2, WARD modifié}.

Vérification

Soit le tableau correspondant à: n = p= 4 ; ε = 0,1 ; η = 0,01 :

	1	2	3	4	5	6
1	1	0	0	0	0	0,01
2	0	1	0	0	0	0
3	0	0	1	0	0	0
4	0,1	0,1	0,1	0	0	0
5	0	0	0	1	0	0
6	0	0	0	0	1	0
7	0	0	0	0	0	1
8	0	0	0	0,1	0,1	0,1

L' arbre de la CAH sur I =]8] est :



On constate que les éléments 4 et 8 à produit scalaire nul s'agrègent bien ensemble à la première étape de la CAH{χ^2, WARD}.

5 Cas particuliers cardI = 5, 6, 7

En modifiant les contre-exemples à la non-reconnaissance de la structures en blocs pour la CAH{χ^2, WARD} donnés dans [3] (de la même façon que dans le contre-exemple ci-dessus par ajoût de η dans une case d'un bloc non diagonal), on vérifie aisément que la CAH{χ^2, WARD modifié} et la CAH{χ^2, WARD} ne sont pas identiques sur des tableaux de correspondance non décomposables en blocs diagonaux de cardinalité 5, 6 ou 7; car pour la CAH{χ^2, WARD} deux éléments à produit scalaire nul sont agrégés ensemble.

6 Conclusion

La CAH{χ^2, WARD modifié} et la CAH{χ^2, WARD} usuelle sont nécessairement identiques pour tout tableau de correspondance non décomposable en blocs diagonaux lorsque celui-ci est de cardinalité ≤ 4 ou d'inertie ≤ 1; en dehors de ces conditions, on trouve des exemples de tableaux pour lesquels ces deux méthodes fournissent des résultats différents.

Ou, en d'autres termes:

si un tableau non décomposable en blocs diagonaux est de cardinalité ≤ 4, ou d'inertie ≤ 1, la CAH{χ^2, WARD} n'agrège jamais deux éléments à produit scalaire nul; en dehors de ces conditions, une telle agrégation peut se produire.

Références bibliographiques

[1] 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).

[2] 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).

[3] 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).

[4] R. ROUSSEAU: "Reconnaissance de la structure de blocs d'un tableau par la Classification Ascendante Hiérarchique"; in *C.A.D.*, Vol XVI, n°2, pp. 237-248; (1991).