

E. YAGOLNITZER

## **Comparaison de deux correspondances entre les mêmes ensembles**

*Les cahiers de l'analyse des données*, tome 2, n° 3 (1977),  
p. 251-264

[http://www.numdam.org/item?id=CAD\\_1977\\_\\_2\\_3\\_251\\_0](http://www.numdam.org/item?id=CAD_1977__2_3_251_0)

© Les cahiers de l'analyse des données, Dunod, 1977, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

## COMPARAISON DE DEUX CORRESPONDANCES

### ENTRE LES MÊMES ENSEMBLES

#### [COMPAR. CORR.]

par E. Yagolnitzer (1)

1. Les données visées : Voici d'abord quelques exemples de données (proposés par des chercheurs avec lesquels nous travaillons), qui requièrent la comparaison de deux correspondances entre les mêmes ensembles.

1.1 Les jumeaux : (E. Yagolnitzer) : Soit I un ensemble de paires de jumeaux dont nous supposons qu'ils sont de sexe différent\* ; J un ensemble de variables qu'on peut déterminer sur chaque sujet, garçon ou fille. On peut constituer deux tableaux :  $k_{IJ}$  et  $kl_{IJ}$  :

$kO(i, j)$  = variable j déterminée sur le garçon de la paire i ;

$kl(i, j)$  = variable j déterminée sur la fille de la paire i .

1.2 Questionnaires aux ménages : (N. Tabard, CREDOC) : Soit I un ensemble de ménages ; J un ensemble de modalités de réponses proposées dans chaque ménage à chacun des deux conjoints : les réponses s'inscrivent en deux tableaux :  $k_{IJ}$  et  $kl_{IJ}$  :

$kO(i, j)$  = 1 si dans le ménage i, l'époux a adopté la réponse j, et zéro sinon.

$kl(i, j)$  = 1 si dans le ménage i, l'épouse a adopté la réponse j, et zéro sinon.

1.3 Questionnaire soumis à deux populations : (J. Goudard, Y. Grelet) : Soit Q un ensemble de questions q comportant chacune un ensemble  $M_q$  de modalités de réponse ; on note  $M = \cup \{M_q | q \in Q\}$  l'ensemble des modalités de réponse à toutes les questions de Q. Soit S un ensemble de sujets s dont chacun a répondu (éventuellement par l'abstention!, si cette attitude est comprise dans  $M_q$ ) à toutes les questions q de Q : on construit d'après ces réponses un tableau de contingence  $k_{MM}$  appelé tableau de Burt, et défini par

$k(m, m')$  = nombres des individus s de S ayant adopté simultanément les modalités de réponse m et m'.

On soumet communément à l'analyse des correspondances, (cf [Bin. Mult.], in Cahiers Vol II n°1), soit le tableau de Burt lui-même, soit divers sous-tableaux rectangulaires extraits de celui-ci : par exemple le sous-tableau  $k_{IJ}$ , où I est l'ensemble des modalités de réponse aux questions d'état (e.g. : âge, revenu, catégorie socioprofessionnelle) et J l'ensemble des modalités de réponse aux question d'opinion. Il est fréquent que le même questionnaire soit soumis à deux ensembles distincts de sujets  $S_0$  et  $S_1$ , entre lesquels peut exister une différence constante (e.g.  $S_0$  est un ensemble d'hommes ;  $S_1$  un ensemble de femmes ; mais sans qu'une correspondance biunivoque relie ces deux ensembles, comme il est le cas aux

\* Les jumeaux de même sexe ; et note sur les dermatoglyphes pp. 291-292 offrent aussi matière à semblable analyse ; cf infra § 7.2.3.

(1) E. Yagolnitzer, Centre d'Etudes et de Recherches sur les Qualifications (CEREQ), service informatique. UER de math. Univ. Paris V.

§§ 1.1 & 1.2). D'où deux tableaux de Burt  $k_0$  et  $k_1$  dont l'analyse peut révéler entre les modalités de réponses des corrélations distinctes qu'il importe de comparer systématiquement. Rendant compte d'une enquête sur les Lycéens du second cycle de l'agglomération parisienne, J. Goudard et Y. Grelet notent par exemple que : "jamais fatigué",..., se projette chez la fille au même niveau que "père : patron ou cadre supérieur" et "père : études supérieures", tandis que ce même trait se situe chez le garçon dans le quadrant des catégories socio-culturelles modestes ; cf infra § 7.2.4 et aussi l'article [Lycéens].

1.4 Tableau carré non symétrique : Avec tout tableau carré  $k_{II}$  non symétrique ( $k_0(i, i') \neq k_0(i', i)$ ) il est de règle de considérer son transposé  $k_{II}$  ( $k_1(i, i') = k_0(i', i)$ ).

Dans la suite, (§§ 2 à 6), sans nous interdire de faire parfois allusion à divers exemples, nous ferons quant à la structure des données la seule hypothèse que  $k_{IJ}$  et  $k_{IJ}$  sont deux tableaux de correspondance (i.e. tableaux de nombres positifs) sur le produit des mêmes ensembles finis I et J. Au § 7 nous considérerons de nouveaux exemples dont la structure particulière sera comprise à la lumière des résultats trouvés par E. Yagolnitzer.

2. Méthodes de comparaison : L'objet principal de la présente note est de montrer d'après E. Yagolnitzer que l'analyse de correspondance d'un tableau rectangulaire en quatre blocs (comprenant deux fois le tableau  $k_0$  et deux fois le tableau  $k_1$  ; cf § 3, sqq.) réalise à la fois l'analyse de correspondance du tableau somme  $k_{IJ} + k_{IJ}$  (où, ce qui revient au même, du tableau moyenne,  $(k_{IJ} + k_{IJ})/2$ ) et une expression du tableau différence  $k_{IJ} - k_{IJ}$ , en somme de produits de facteurs orthogonaux sur I et J (exactement : une réduction de  $k_{IJ} - k_{IJ}$ , considéré comme élément du produit tensoriel de deux espaces euclidiens, cf [Réd. Tens.] TII, B n° 6 ; et infra § 6). Auparavant rappelons que l'adjonction d'éléments supplémentaires fournie ici comme ailleurs, un moyen de comparaison suivant des calculs simples dont les résultats s'interprètent sans mirages.

Il suffit de prendre comme tableau de base le tableau moyenne  $(k_0 + k_1)/2$  et d'y adjoindre en éléments supplémentaires à la fois les lignes et les colonnes des deux tableaux  $k_0$  et  $k_1$  : soit schématiquement cinq blocs  $I \times J$ , disposés suivant la figure 1 :

	J	J0	J1
I	tableau principal ( $k_0 + k_1$ )/2	colonnes supplém. $k_0$	colonnes supplém. $k_1$
I0	lignes supplém. $k_0$		
I1	lignes supplém. $k_1$		

Figure 1 : schéma du tableau moyenne, avec cinq blocs supplémentaires.

Ainsi chaque  $i$  apparaît trois fois : une fois comme élément principal  $i$  (moyenne) et deux fois comme élément supplémentaire,  $i_0$  et  $i_1$ ; et de même pour chaque  $j$ . Si par exemple dans le cas du § 1.3 l'analyse étale sur une ligne l'échelle des catégories socio-culturelles, le point moyen "jamais fatigué" occupera sur cette ligne une position centrale ; mais les éléments supplémentaires "jamais fatigué ; fille" et "jamais fatigué ; garçon" s'y opposeront.

3. Le tableau en quatre blocs ; notations : Afin de démontrer les propriétés des facteurs issus du tableau en quatre blocs construit par E. Yagolnitzer (cf fig 2), posons des notations précises. Nous désignons par  $IO$  et  $I1$  deux exemplaires de l'ensemble  $I$  (dont les éléments seront affectés d'un indice numérique 0 ou 1)

$$IO = \{i_0 \mid i \in I\} ; I1 = \{i_1 \mid i \in I\} ;$$

et notons  $U$  la réunion :  $U = IO \cup I1$ . Un élément  $u$  de  $U$  étant de l'une des deux formes  $i_0$  et  $i_1$ , il est commode d'exprimer cette forme par des fonctions de  $u$  en posant :

$$u = i(u)_{n(u)} ; \text{ où } i(u) \text{ et } n(u) \text{ sont définies par :}$$

$$i(i_0) = i(i_1) = i ; n(i_0) = 0 ; n(i_1) = 1 ;$$

ainsi  $n(u)$  est l'un des deux nombres 0 ou 1, selon que  $u$  appartient à la moitié initiale  $IO$ , ou à la moitié finale  $I1$  de  $U$  ; et  $i(u)$  est cet  $i \in I$  qui se trouve dans  $u$ , affecté de l'indice numérique  $n(u)$ . On note de même pour l'ensemble des colonnes :

$$JO = \{j_0 \mid j \in J\} ; J1 = \{j_1 \mid j \in J\} ; V = JO \cup J1 ;$$

$$v = j(v)_{n(v)} ; j(j_0) = j(j_1) = j ; n(j_0) = 0 ; n(j_1) = 1 .$$

Maintenant, le tableau en quatre blocs  $k_{UV}$  peut être défini par la formule :

$$k_{UV}(u,v) = k p(n(u) + n(v)) (i(u), j(v)) ;$$

où l'on a noté  $p(n(u) + n(v))$  la parité de la somme  $n(u) + n(v)$  ; parité qui vaut 0 si  $n(u) = n(v)$  (cas des blocs diagonaux égaux au tableau  $k_0$  : soit  $n(u) = n(v) = 0$  ; soit  $n(u) = n(v) = 1$ ,  $n(u) + n(v) = 2$ ) ; et qui vaut 1 si  $n(u) \neq n(v)$  (cas des blocs hors de la diagonale descendante ; blocs égaux au tableau  $k1$ ). Cet exercice d'écriture du tableau  $k_{UV}$  peut sembler gratuit, car le contenu de ce tableau est clair d'après la figure 2 ; en revanche, il est essentiel pour la suite d'avoir mis en place

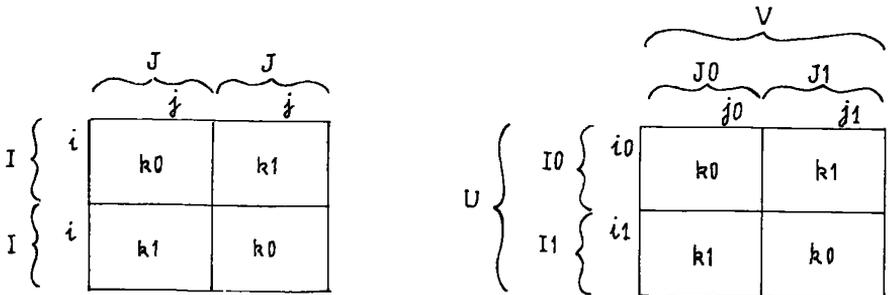


Figure 2 : à gauche, juxtaposition des quatre blocs ; à droite, notations adoptées pour désigner ceux-ci.

les ensembles  $i_0, i_1, U$  et  $j_0, j_1, V$  afin d'expliciter la symétrie du tableau qui se retrouve dans les résultats de l'analyse. La démonstration des résultats obtenus par E. Yagolnitzer occupe les §§ 4-5-6 qui suivent. Les formules utiles aux applications sont ensuite reprises au § 7 dont la lecture peut donc suffire au praticien.

#### 4. Symétrie des ensembles, du tableau, et des transitions.

4.1 Les formules de symétrie : Chacun des ensembles  $U$  et  $V$  est muni d'une symétrie évidente qui subsiste tout au long de l'analyse. Précisons cette symétrie que nous noterons  $\omega$  ; on pose :

$$\begin{aligned} \omega_U^U : \omega(i_0) &= i_1 ; \omega(i_1) = i_0 ; \\ \omega(u) &= i(u)_{1-n(u)} ; \omega \circ \omega = \text{identité} ; \\ \omega_V^V : \omega(j_0) &= j_1 ; \omega(j_1) = j_0 ; \\ \omega(v) &= j(v)_{1-n(v)} ; \omega \circ \omega = \text{identité} ; \end{aligned}$$

en bref, la symétrie  $\omega$  échange la moitié initiale et la moitié finale de  $U$  ou de  $V$  (du point de vue de la disposition des nombres, il s'agit donc de translations des deux moitiés : l'une vers la droite l'autre vers la gauche ; mais nous parlons de symétrie au sens abstrait, parce que  $\omega \circ \omega = \text{identité}$ ). Dans la suite, on écrira souvent  $\omega$ , sans préciser par des indices s'il s'agit de la symétrie  $\omega_U^U$  de  $U$ , ou de la symétrie  $\omega_V^V$  de  $V$  ; et on omettra les parenthèses, notant  $\omega u$  pour  $\omega(u)$ ,  $\omega v$  pour  $\omega(v)$ .

La symétrie  $\omega$  s'étend à la mesure  $f_{UV}$ , aux marges  $f_U, f_V$ , aux transitions  $f_V^U$  et  $f_U^V$ . Si comme il est d'usage en analyse des correspondances, on note  $k$  le total du tableau  $k_{UV}$ ,  $k(u)$  le total de la ligne  $u$ ,  $k(v)$  le total de la ligne  $v$ ,  $f_{uv} = k(u,v)/k$  ;  $f_u = k(u)/k$  ; etc., on a :

$$\begin{aligned} f_{\omega u} &= f_u & ; & & f_{\omega v} &= f_v & ; \\ f_{\omega u, \omega v} &= f_{u, v} & ; & & f_{\omega u, v} &= f_{u, \omega v} & ; \\ f_{\omega v}^{\omega u} &= f_v^u & ; & & f_{\omega v}^u &= f_v^{\omega u} & ; \\ f_{\omega u}^{\omega v} &= f_u^v & ; & & f_{\omega u}^v &= f_u^{\omega v} & . \end{aligned}$$

Démontrons brièvement ces formules. L'égalité  $f_{\omega u} = f_u$ , exprime que les deux moitiés consécutives de la marge sont identiques : en effet, les deux lignes  $i_0$  et  $i_1$  qui ne diffèrent que par l'échange de leurs deux moitiés (lignes  $i$  des tableaux  $k_0$  et  $k_1$  disposées successivement) ont même somme. L'application de la symétrie  $\omega$  aux deux indices  $u, v$  fait passer d'un bloc du tableau  $k_{UV}$  à un autre bloc qui lui est égal : e.g. :  $k(\omega i_0, \omega j_1) = k(i_1, j_0) = k_1(i, j)$  ; c'est pourquoi  $f_{\omega u, \omega v} = f_{u, v}$ . L'application de la symétrie  $\omega$  à un seul des deux indices, soit  $u$ , soit  $v$  fait au contraire passer d'un bloc du tableau à un bloc limitrophe qui ne lui est pas égal : e.g. :  $k(\omega i_0, j_1) = k(i_1, j_1) = k_0(i, j) = k(i_0, \omega j_1) = k(i_0, j_0)$  ; mais  $k(i_0, j_1) = k_1(i, j)$  ; et c'est pourquoi  $f_{\omega u, v} = f_{u, \omega v}$  ; etc.

4.2 Symétrie et transitions : En calcul tensoriel des transitions probabilistes, (cf [Note Lim.] TII B n° 1) les propriétés de symétrie s'écrivent sous une forme qui rend immédiate la démonstration des propriétés des facteurs, objet du § 5. Sans imposer au lecteur l'usage de ce calcul, récrivons dans ses notations les propriétés de symétrie déjà

énoncées. Les fonctions  $\omega_U^U$  et  $\omega_V^V$  sont considérées comme des transitions probabilistes particulières, (respectivement de U vers U et de V vers V) ; et l'on note comme d'usage  $\delta_U^U$ ,  $\delta_V^V$  l'application identité ( $\delta$  de Kronecker) ; on a :

$$\begin{aligned} \omega_U^U \circ f_U &= f_U \quad ; \quad \omega_V^V \circ f_V = f_V \quad ; \\ (\omega_U^U \circ \omega_V^V) \circ f_{UV} &= f_{UV} \quad ; \\ (\omega_U^U \circ \delta_V^V) \circ f_{UV} &= (\delta_U^U \circ \omega_V^V) \circ f_{UV} \quad ; \\ \omega_V^V \circ f_V^U \circ \omega_U^U &= f_V^U \quad ; \quad \omega_V^V \circ f_V^U = f_V^U \circ \omega_U^U \quad ; \\ \omega_U^U \circ f_U^V \circ \omega_V^V &= f_U^V \quad ; \quad \omega_U^U \circ f_U^V = f_U^V \circ \omega_V^V \quad ; \end{aligned}$$

la dernière formule exprime en bref que la transition  $f_U^V$  commute avec la symétrie  $\omega$  ; on passe de la formule  $\omega_U^U \circ f_U^V = f_U^V \circ \omega_V^V$  à la formule  $\omega_U^U \circ f_U^V \circ \omega_V^V = f_U^V$ , par composition (à droite) des deux membres avec  $\omega_V^V$ , si l'on se souvient que  $\omega_V^V \circ \omega_V^V$  est égale à l'identité  $\delta_V^V$  ; etc.

4.3 Somme et différence des tableaux donnés : Considérons le tableau  $f_{UV}$  et les tableaux construits par somme et différence à partir de  $k_0$  et  $k_1$  : leurs rapports, nous l'avons annoncé, sont la clé de l'interprétation algébrique des facteurs (§ 6).

Notons d'abord :

$$\begin{aligned} k_0 &= \text{total du tableau } k_{0IJ} \quad ; \quad k_1 = \text{total du tableau } k_{1IJ} \quad ; \\ k^+(i,j) &= k_0(i,j) + k_1(i,j) \quad ; \quad k^-(i,j) = k_0(i,j) - k_1(i,j) \quad ; \\ f_{i,j}^+ &= k^+(i,j)/(k_0 + k_1) \quad ; \quad f_{i,j}^- = k^-(i,j)/(k_0 + k_1) \quad ; \\ f_i^+ &= (k_0(i) + k_1(i)) / (k_0 + k_1) \quad ; \quad f_j^+ = (k_0(j) + k_1(j)) / (k_0 + k_1). \end{aligned}$$

Il importe de remarquer que  $f_{IJ}^+$  n'est autre que la loi de probabilité associée au tableau de correspondance  $k_{IJ}^+$ , tableau somme des tableaux donnés ; mais  $f_{IJ}^-$  n'est pas nécessairement un tableau de nombres positifs, le total de ses éléments n'est pas égal à 1, etc. ; son étude ne relève pas immédiatement de l'analyse des correspondances. (Quant à l'interprétation, elle semble *a priori*, d'autant plus désirable que les totaux  $k_0$  et  $k_1$  sont égaux et donc  $f_{IJ}^-$  a total nul : dans l'exemple du § 1.3 ce sera le cas si les deux populations S0 et S1 ont même effectif, ce à quoi on peut toujours se ramener par pondération<sup>(\*)</sup> ; sinon, à supposer e.g. que  $k_{0IJ}$  et  $k_{1IJ}$  aient des profils semblables et que  $k_0 = 2k_1$ , le profil  $f_{IJ}^-$  de la différence reproduira à peu près  $f_{0IJ} \# f_{1IJ}$ ).

Or  $f_{IJ}^+$  et  $f_{IJ}^-$  se retrouvent dans les parties paires et impaires de  $f_{UV}$  relativement à la symétrie  $\omega$ . Posons :

$$f_{u,v}^+ = (f_{u,v} + f_{\omega u,v})/2 \quad ; \quad f_{u,v}^- = (f_{u,v} - f_{\omega u,v})/2 \quad ;$$

(\*) On verra sur un exemple (cf infra § 7.2.4 et [Lycéens] § 4.2 pp 285-288) que cette précaution ne suffit pas : il faut encore que  $k_0$  et  $k_1$  aient mêmes marges.

(ici encore  $f_{UV}^+$  est une loi de probabilité sur  $U \times V$  ; mais non  $f_{UV}^-$ , qui comprend des nombres de tout signe) ; il vient :

$$f_{i,j}^+ = 4 f_{i0,j0}^+ = 4 f_{in,jn}^+$$

$$f_{i,j}^- = 4 f_{i0,j0}^- = 4 (-1)^{n+n'} f_{in,jn}^-$$

(où  $n$  et  $n'$  désignent des indices numériques valant 0 ou 1) ; ce qu'on peut encore écrire (en notant  $\epsilon = +$  ou  $-$ )

$$\forall u \in U, \forall v \in V : 4 f_{u,v}^\epsilon = \epsilon^{n(u)+n(v)} f_{i(u),j(v)}^\epsilon.$$

Quant aux marges on a simplement :

$$f_i^+ = 2 f_{i0} = 2 f_{i1} \quad ; \quad f_j^+ = 2 f_{j0} = 2 f_{j1}.$$

La vérification de ces formules nous paraît peu instructive ; expliquons seulement le coefficient 4 de  $f_{i,j}^+ = 4 f_{in,jn}^+$ , par ce que les tableaux  $f_{IJ}^+$  et  $f_{UV}^+$  ont tous deux total 1, et qu'à une case  $(i,j)$  de  $I \times J$ , il en correspond quatre  $(in,jn')$  de  $U \times V$ .

5. Symétrie des facteurs : Soit  $\psi^U$  une fonction sur  $U$  (e.g. une fonction à valeur réelle) ; la fonction composée  $\psi^U \circ \omega_U^U$  est aussi une fonction sur  $U$ , définie par :  $\forall u \in U : \psi \circ \omega(u) = \psi(\omega u)$ . Si l'on se représente la fonction  $\psi^U$  comme la suite de ses valeurs rangées dans l'ordre de l'ensemble  $U$  ( $U = I0 \cup I1$ ), la fonction  $\psi^U \circ \omega_U^U$  s'obtient en échangeant les deux moitiés (celle sur  $I0$  et celle sur  $I1$ ) de la fonction  $\psi^U$ . De même pour une fonction  $\psi^V$  sur  $V$  et  $\psi^V \circ \omega_V^V$ .

Ceci dit, on a la propriété fondamentale suivante : si  $(\varphi^U, \varphi^V)$  sont un couple de facteurs associés à la valeur propre  $\lambda$  (issus de  $f_{UV}$ ) il en est de même du couple symétrique  $(\varphi^U \circ \omega_U^U, \varphi^V \circ \omega_V^V)$ . Cette propriété nous permettra de partager les facteurs en deux classes : ceux qui sont symétriques relativement à  $\omega$  :  $\varphi^U \circ \omega_U^U = \varphi^U$  ;  $\varphi^V \circ \omega_V^V = \varphi^V$  ; et ceux qui sont antisymétriques :  $\varphi^U \circ \omega_U^U = -\varphi^U$  ;  $\varphi^V \circ \omega_V^V = -\varphi^V$ . Dans le premier cas, les deux segments successifs de  $\varphi$  sont identiques ( $\varphi^{i0} = \varphi^{i1}$  ;  $\varphi^{j0} = \varphi^{j1}$ ) ; dans le second ils sont opposés ( $\varphi^{i0} = -\varphi^{i1}$  ;  $\varphi^{j0} = -\varphi^{j1}$ ).

Pour démontrer la propriété fondamentale, rappelons une des définitions des couples de facteurs associés à une valeur propre :

$$\varphi^U \circ f_U^V = \lambda^{1/2} \varphi^V \quad ; \quad \text{i.e.} : \Sigma\{\varphi^U f_u^V \mid u \in U\} = \lambda^{1/2} \varphi^V ;$$

$$\varphi^V \circ f_V^U = \lambda^{1/2} \varphi^U \quad ; \quad \text{i.e.} : \Sigma\{\varphi^V f_v^U \mid v \in V\} = \lambda^{1/2} \varphi^U ;$$

Or des propriétés de symétrie énoncées au § 4 pour les transitions  $f_U^V$  et  $f_V^U$ , il résulte que l'on a :

$$(\varphi^U \circ \omega_U^U) \circ f_U^V = \lambda^{1/2} (\varphi^V \circ \omega_V^V) ;$$

$$(\varphi^V \circ \omega_V^V) \circ f_V^U = \lambda^{1/2} (\varphi^U \circ \omega_U^U) ;$$

ce qui exprime que  $(\varphi^U \circ \omega, \varphi^V \circ \omega)$  est un couple de facteurs associés à  $\lambda$ .

Démontrons, e.g., la première des égalités de ce système ; d'abord en développant des sommes, puis par le calcul des transitions. On a :

$$\begin{aligned} \Sigma \{ (\varphi \circ \omega)^u f_u^v \mid u \in U \} &= \Sigma \{ \varphi^{\omega u} f_u^v \mid u \in U \} \\ &= \Sigma \{ \varphi^{\omega u} f_{\omega u}^{v'} \mid u \in U \} = \Sigma \{ \varphi^{u'} f_{u'}^{v'} \mid u' \in U \} \\ &= \lambda^{1/2} \varphi^{\omega V} = \lambda^{1/2} (\varphi \circ \omega)^V ; \end{aligned}$$

(où l'on s'est souvenu que  $f_{\omega u}^{v'} = f_u^v$  ; et a fait dans la sommation le changement de variable  $\omega u = u'$ ). On peut encore écrire :

$$\varphi^U \circ \omega_U^U \circ f_U^V = \varphi^U \circ f_U^V \circ \omega_V^V = \lambda^{1/2} \varphi^V \circ \omega_V^V$$

(où l'on s'est souvenu que  $\omega \circ f = f \circ \omega$  ; et a fait usage de l'associativité de la composition des transitions).

Pour nous réduire à ne considérer que des facteurs soit symétriques, soit antisymétriques relativement à  $\omega$ , on applique la méthode générale de décomposition en parties paires et impaires (méthode déjà appliquée dans le § 4.3 au tableau  $f_{UV}$  pour définir  $f_{UV}^+$  et  $f_{UV}^-$ ) ; si  $E$  est un espace vectoriel possédant une symétrie  $\omega$  (i.e. une application de  $E$  dans  $E$  telle que  $\omega \circ \omega =$  identité),  $E$  est somme directe de deux sous-espaces  $E^+$ , sur lequel  $\omega$  agit comme l'identité ( $\forall x \in E^+ : \omega x = x$ ) et  $E^-$  sur lequel  $\omega$  agit en changeant le signe ( $\forall x \in E^- : \omega x = -x$ ) ; et la décomposition de tout  $x$  en ses composantes suivant  $E^+$  et  $E^-$  se calcule par somme et différence :  $x^+ = (x + \omega x)/2$  ;  $x^- = (x - \omega x)/2$  ;  $x^+ \in E^+$  ;  $x^- \in E^-$  ;  $x = x^+ + x^-$ . Ici, considérons un couple  $(\varphi^U, \varphi^V)$  de facteurs associés à la valeur propre  $\lambda$  : par linéarité, on voit que sont aussi associés à la valeur propre  $\lambda$  les deux couples :

$$((\varphi^U + \varphi^U \circ \omega)/2, (\varphi^V + \varphi^V \circ \omega)/2) : \text{ couple symétrique ;}$$

$$((\varphi^U - \varphi^U \circ \omega)/2, (\varphi^V - \varphi^V \circ \omega)/2) : \text{ couple antisymétrique.}$$

En général une valeur propre  $\lambda$  issue de l'analyse de  $f_{UV}$  est simple. Si  $(\varphi^U, \varphi^V)$  est couple associé à  $\lambda$ , l'un des deux couples  $((\varphi^U + \varepsilon \varphi^U \circ \omega)/2, (\varphi^V + \varepsilon \varphi^V \circ \omega)/2)$  est nul. Si c'est le couple des  $(\varphi - \varphi \circ \omega)/2$  qui est nul on a  $\varphi = (\varphi + \varphi \circ \omega)/2$  ; le couple  $(\varphi^U, \varphi^V)$  est symétrique.

Si c'est le couple des  $(\varphi + \varphi \circ \omega)/2$  qui est nul au contraire, on a  $\varphi = (\varphi - \varphi \circ \omega)/2$  ; le couple  $(\varphi^U, \varphi^V)$  est antisymétrique. Mais si la valeur propre  $\lambda$  est double (ou de multiplicité supérieure à 2), on peut

avoir à la fois couple symétrique et couple antisymétrique relatif à  $\lambda$  : le calcul pourra fournir des couples de facteurs  $(\varphi^U, \varphi^V)$  n'ayant pas de symétrie déterminée ; et dont la décomposition donnera un couple symétrique et un couple antisymétrique tous deux non-nuls.

Désormais on supposera faite la séparation des facteurs symétriques et antisymétriques. Ainsi l'ensemble  $A$  des facteurs relatifs à une valeur propre non nulle (ensemble où l'on comprendra le facteur trivial constant et égal à 1) est séparé en deux :  $A^+$  et  $A^-$ , facteurs symétriques et facteurs antisymétriques relativement à  $\omega$ . On notera  $\omega(\alpha) = 1$  si  $\alpha \in A^+$  ; et  $\omega(\alpha) = -1$  si  $\alpha \in A^-$ . Avec ces conventions on a pour tout facteur :

$$\varphi_{\alpha}^{i_0} = \omega(\alpha) \varphi_{\alpha}^{i_1} ; \quad \varphi_{\alpha}^{j_0} = \omega(\alpha) \varphi_{\alpha}^{j_1} ;$$

ou, si l'on considère non les facteurs  $\varphi$  (normalisés, de variance 1) , mais les facteurs  $F_{\alpha}$ ,  $G_{\alpha}$  (qui ont variance  $\lambda_{\alpha}$  ; et figurent sur les graphiques usuels) :

$$F_{\alpha}(i_0) = \omega(\alpha) F_{\alpha}(i_1) ; \quad G_{\alpha}(j_0) = \omega(\alpha) G_{\alpha}(j_1).$$

*Nota Bene* : il ne faut pas confondre la notion de symétrie (relative à  $\omega$ ) introduite pour les facteurs issus d'un tableau en quatre blocs égaux deux à deux ; et celle de parité, définie antérieurement pour les facteurs issus d'un tableau symétrique  $k_{II}$  (cf [Corr. Sym.], TII B n°9) : en particulier pour un tableau  $k_{UU}$  fait de quatre blocs carrés égaux deux à deux à deux blocs symétriques  $k_{0_{II}}$  et  $kl_{II}$ , les facteurs peuvent être classés à la fois quant à la parité et à l' $\omega$ -symétrie.

6. Interprétation analytique des deux classes de facteurs issus du tableau en quatre blocs.

En appliquant au tableau  $k_{UV}$  la formule de reconstitution des données en fonction des facteurs, on montre qu'en bref les facteurs de  $A^+$  sont ceux issus de l'analyse de correspondance usuelle du tableau somme  $k_{0_{IJ}} + kl_{IJ}$  (ou de la loi  $f_{IJ}^+$  ; cf § 4.3) ; tandis que ceux de  $A^-$  fournissent une reconstitution de  $k_{0_{IJ}} - kl_{IJ}$  (ou de  $f_{IJ}^-$  ; cf § 4.3) ; d'où il résulte que  $\text{Card } A^+$  et  $\text{Card } A^-$  sont tous deux bornés par  $\text{Card } I$  et  $\text{Card } J$ , (et généralement égaux au plus petit de ces nombres).

Ecrivons la formule de reconstitution pour  $f_{VV}$  puis pour ses parties paires et impaires relativement à  $\omega$  :  $f_{UV}^+$  et  $f_{UV}^-$  (cf § 4.3) :

$$f_{u,v} = f_u f_v \sum \{ \lambda_{\alpha}^{1/2} \varphi_{\alpha}^u \varphi_{\alpha}^v \mid \alpha \in A \} ;$$

$$f_{u,\omega v} = f_u f_v \sum \{ \lambda_{\alpha}^{1/2} \varphi_{\alpha}^u \varphi_{\alpha}^{\omega v} \mid \alpha \in A \} ;$$

d'où si l'on se souvient que (par définition de la parité des facteurs)  $\varphi_{\alpha}^{\omega v} = \omega(\alpha) \varphi_{\alpha}^v$

$$f_{u,v}^+ = (f_{u,v} + f_{u,\omega v})/2 = f_u f_v \sum \{ \lambda_{\alpha}^{1/2} \varphi_{\alpha}^u \varphi_{\alpha}^v \mid \alpha \in A^+ \} ;$$

$$f_{u,v}^- = (f_{u,v} - f_{u,\omega v})/2 = f_u f_v \sum \{ \lambda_{\alpha}^{1/2} \varphi_{\alpha}^u \varphi_{\alpha}^v \mid \alpha \in A^- \} .$$

Reste à rapporter ces formules à l'analyse des tableaux  $f_{IJ}^+$  et  $f_{IJ}^-$ , définis au § 4.3. Rappelons que l'on a :

$$f_{ij}^+ = 4 f_{i0,j0}^+ ; f_{i,j}^- = 4 f_{i0,j0}^- ; f_i^+ = 2 f_{i0}^+ ; f_j^+ = 2 f_{j0}^+ ;$$

De ces égalités et des formules de décomposition pour  $f_{U,V}^+$  et  $f_{U,V}^-$  il résulte que :

$$f_{i,j}^+ = f_i^+ f_j^+ \sum \{ \lambda_\alpha^{i0} \varphi_\alpha^{i0} \varphi_\alpha^{j0} \mid \alpha \in A^+ \} ;$$

$$f_{i,j}^- = f_i^- f_j^- \sum \{ \lambda_\alpha^{i0} \varphi_\alpha^{i0} \varphi_\alpha^{j0} \mid \alpha \in A^- \} ;$$

Pour montrer qu'on a dans ces formules l'analyse de la correspondance  $f_{IJ}^+$  et une décomposition (au sens de [Red. Tens.] TII B n° 6) de  $f_{IJ}^-$  considéré comme élément du produit tensoriel des deux espaces euclidiens  $R_I$  et  $R_J$  munis respectivement de la métrique du  $\chi^2$  de centre  $f_I$  et  $f_J$ , il faut considérer les facteurs  $\varphi_\alpha$  comme des fonctions sur I (ou J) et établir leur orthonormalité. Définissons  $\varphi_\alpha^I$  et  $\varphi_\alpha^J$  par les formules:

$$\varphi_\alpha^I = \{ \varphi_\alpha^i \mid i \in I \} ; \varphi_\alpha^i = \varphi_\alpha^{i0} ;$$

$$\varphi_\alpha^J = \{ \varphi_\alpha^j \mid j \in J \} ; \varphi_\alpha^j = \varphi_\alpha^{j0} .$$

Il est facile de montrer que sur I muni de la loi  $f_I^+$ , les  $\varphi_\alpha^I$  sont de variance 1 et que si  $\alpha$  et  $\alpha'$  sont de même symétrie ( $\omega(\alpha) = \omega(\alpha')$ )  $\varphi_\alpha^I$  et  $\varphi_{\alpha'}^I$  sont orthogonaux. En effet sur U on a les propriétés d'orthonormalité :

$$\sum \{ f_u \varphi_\alpha^u \varphi_{\alpha'}^u \mid u \in U \} = \delta_{\alpha, \alpha'} ; \text{ d'où :}$$

$$\sum \{ (f_i^+/2) (\varphi_\alpha^{i0} \varphi_{\alpha'}^{i0} + \varphi_\alpha^{i1} \varphi_{\alpha'}^{i1}) \mid i \in I \} = \delta_{\alpha, \alpha'} ;$$

$$\sum \{ (f_i^+/2) \varphi_\alpha^{i0} \varphi_{\alpha'}^{i0} (1 + \omega(\alpha) \omega(\alpha')) \mid i \in I \} = \delta_{\alpha, \alpha'} ;$$

$$\sum \{ f_i^+ ((1 + \omega(\alpha) \omega(\alpha'))/2) \varphi_\alpha^i \varphi_{\alpha'}^i \mid i \in I \} = \delta_{\alpha, \alpha'} ;$$

Si  $\alpha$  et  $\alpha'$  sont de symétries différentes (l'un symétrique l'autre antisymétrique),  $1 + \omega(\alpha) \omega(\alpha') = 0$  et l'on ne peut rien conclure de cette formule ; mais si  $\alpha, \alpha' \in A^+$  ou  $\alpha, \alpha' \in A^-$ , la formule donne le produit scalaire  $\langle \varphi_\alpha^I \mid \varphi_{\alpha'}^I \rangle$ . Nous concluons donc : sur I muni de la loi  $f_I^+$  les systèmes :

$$\{ f_\alpha^I \mid \alpha \in A^+ \} \text{ et } \{ f_\alpha^I \mid \alpha \in A^- \} ,$$

sont orthonormés (i.e. formés de fonctions de variance 1 deux à deux orthogonales). On montre de même que sur J muni de la loi  $f_J^+$ , sont orthonormés les systèmes :

$$\{ f_\alpha^J \mid \alpha \in A^+ \} \text{ et } \{ f_\alpha^J \mid \alpha \in A^- \} .$$

Ainsi il apparaît que la formule

$$f_{i,j}^+ = f_i^+ f_j^+ \Sigma \{ \lambda_\alpha^{1/2} \varphi_\alpha^i \varphi_\alpha^j \mid \alpha \in A^+ \},$$

fournit l'analyse de la correspondance  $f_{IJ}^+$  (car seuls les facteurs issus de l'a. des correspondances sont des systèmes orthonormés donnant une formule de reconstitution). Quant à :

$$f_{i,j}^- = f_i^+ f_j^+ \Sigma \{ \lambda_\alpha^{1/2} \varphi_\alpha^i \varphi_\alpha^j \mid \alpha \in A^- \}$$

c'est comme on l'a annoncé une décomposition de  $f_{I,J}^-$  généralisant celle d'Eckart et Young comme il est expliqué dans [Red. Tens.] TII B n° 6. A titre d'exercice on réécrira cette formule tensoriellement :

$$f_{IJ}^- = \Sigma \{ (\varphi_\alpha^I f_I^+) \otimes (\varphi_\alpha^J f_J^+) \mid \alpha \in A^+ \}$$

où l'on a noté  $(\varphi_\alpha^I f_I^+)$  la mesure sur I ayant par rapport à  $f_I^+$  la densité  $\varphi_\alpha^I$ . Pour la métrique du  $\chi^2$  de centre  $f_I^+$ , les mesures  $(\varphi_\alpha^I f_I^+)$  forment un système orthonormé ; et de même les  $(\varphi_\alpha^J f_J^+)$  sur J muni de  $f_J^+$ .

7. Formulaire et applications : On donne ici les résultats de l'analyse du tableau en blocs (§ 3) avec des formules aboutissant aux deux classes de facteurs distinguées ci-dessus (§§ 5-6). Puis on reprend en les complétant les applications annoncées au § 1.

7.1 Calcul des deux classes de facteurs : On a montré que l'analyse du tableau en quatre blocs dont la dimension est  $(2 \text{ Card } I) \times (2 \text{ Card } J)$  se réduit à celles de deux tableaux ayant tous deux pour dimension  $\text{Card } I \times \text{Card } J$  : du point de vue du calcul (durée et encombrement) ce sont ces dernières analyses qu'on doit effectuer. L'un des tableaux  $f^+$ , ne contient que des nombres positifs ; le programme d'analyse des correspondances s'y applique immédiatement ; l'autre  $f^-$  sera traité par une diagonalisation de matrice carrée symétrique qu'on précise ici. Pour plus de commodité rappelons les notations posées ci-dessus au § 4.3, puis les résultats des §§ 5-6 :

$$\begin{aligned} k_0 &= \text{total du tableau } k_{0IJ} ; & k_1 &= \text{total du tableau } k_{1IJ} ; \\ k^+(i,j) &= k_0(i,j) + k_1(i,j) ; & k^-(i,j) &= k_0(i,j) - k_1(i,j) ; \\ f_{ij}^+ &= k^+(i,j) / (k_0 + k_1) ; & f_{ij}^- &= k^-(i,j) / (k_0 + k_1) ; \\ f_i^+ &= (k_0(i) + k_1(i)) / (k_0 + k_1) ; & f_j^+ &= (k_0(j) + k_1(j)) / (k_0 + k_1) . \end{aligned}$$

Une première classe de facteurs, dits facteurs  $\omega$ -symétriques, s'obtiennent par l'analyse usuelle du tableau de correspondance  $k^+$ , ou de la loi associée  $f_{IJ}^+$ .

Si  $\varphi_\alpha^I, \varphi_\alpha^J$  est un couple de facteur (de variance 1) issu de  $f_{IJ}^+$ , on a pour le tableau en quatre blocs relativement à la même valeur propre les facteurs  $\varphi_\alpha^U, \varphi_\alpha^V$  obtenus en mettant bout à bout deux exemplaires de  $\varphi_\alpha^I$  et  $\varphi_\alpha^J$  respectivement ; i.e. avec les notations du § 4 :  
 $\varphi_\alpha(i_0) = \varphi_\alpha(i_1) = \varphi_\alpha(i)$  ;  $\varphi_\alpha(j_0) = \varphi_\alpha(j_1) = \varphi_\alpha(j)$ .

Quant aux facteurs  $\omega$ -antisymétriques  $\varphi_\beta^U, \varphi_\beta^V$ , ils sont définis sur U et V par :

$$\varphi_\beta(i_0) = -\varphi_\beta(i_1) = \varphi_\beta(i) \quad ;$$

$$\varphi_\beta(j_0) = -\varphi_\beta(j_1) = \varphi_\beta(j) \quad ;$$

i.e. en mettant bout à bout non deux vecteurs identiques, mais un vecteur  $\varphi_\beta^I$ , ou  $\varphi_\beta^J$  et son opposé ; et les  $\varphi_\beta^I, \varphi_\beta^J$  résultent de l'analyse de  $f_{IJ}^-$  avec pour coefficients de métrique non les sommes  $f_i^-, f_j^-$  de ces lignes et colonnes, sommes qui peuvent présenter des zéros, mais les mêmes  $f_i^+, f_j^+$  utilisés dans l'analyse de  $f_{IJ}^+$ . A cela près on peut reprendre les formules usuelles. De façon précise posons :

$$f_j^{-i} = f_{ij}^- / f_i^+ \quad ; \quad f_i^{-j} = f_{ij}^- / f_j^+ \quad ;$$

alors les valeurs propres et facteurs satisfont à la double équation de transition :

$$\lambda_\beta^{1/2} \varphi_\beta^i = \Sigma \{ \varphi_\beta^j \quad f_j^{-i} \mid j \in J \} ; \text{ ou : } \lambda_\beta^{1/2} \varphi_\beta^I = \varphi_\beta^J \circ f_J^{-I} \quad ;$$

$$\lambda_\beta^{1/2} \varphi_\beta^j = \Sigma \{ \varphi_\beta^i \quad f_i^{-j} \mid i \in I \} ; \text{ ou : } \lambda_\beta^{1/2} \varphi_\beta^J = \varphi_\beta^I \circ f_I^{-J} \quad ;$$

dans ces formules les facteurs  $\varphi_\alpha^I, \varphi_\alpha^J$  ne sont pas centrés en général, ils sont seulement orthonormalisés au sens suivant :

$$\Sigma \{ \varphi_\beta^i \quad \varphi_\beta^i \quad f_i^+ \mid i \in I \} = \lambda_\beta \delta_{\alpha\beta} \quad ; \quad F_\beta(i) = \lambda_\beta^{1/2} \varphi_\beta^i \quad ;$$

$$\Sigma \{ \varphi_\beta^j \quad \varphi_\beta^j \quad f_j^+ \mid j \in J \} = \lambda_\beta \delta_{\alpha\beta} \quad ; \quad G_\beta(j) = \lambda_\beta^{1/2} \varphi_\beta^j \quad ;$$

Pour le calcul, on se place sur celui des deux ensembles dont le cardinal est le plus faible, soit par exemple I ; et l'on pose :

$$\lambda_\beta \varphi_\beta^I = \varphi_\beta^I \circ f_I^{-J} \circ f_J^{-I} \quad ; \text{ i.e. le système :}$$

$$\lambda_\beta \varphi_\beta^i = \Sigma \{ \varphi_\beta^{i'} \quad f_{i'}^{-j} \quad f_j^{-i} \mid i' \in I, j \in J \} \quad ;$$

par un changement d'inconnue, on se ramène à diagonaliser une matrice carrée symétrique ; de façon précise si l'on pose :

$$\psi_\beta(i) = (f_i^+)^{1/2} \varphi_\beta^i \quad ; \text{ et pour la matrice symétrique } g :$$

$$g(i, i') = \Sigma \{ f_{ij}^- \quad f_{i'j}^- \quad (f_j^+)^{-1} \quad (f_i^+ \quad f_{i'}^+)^{-1/2} \mid j \in J \}$$

les  $\psi$  satisfont au système :

$$\lambda_\beta \psi_\beta(i) = \Sigma \{ \psi_\beta(i') \quad g(i, i') \mid i' \in I \} \quad ;$$

d'où par diagonalisation les  $\psi_\beta(i)$  ; puis les  $\varphi_\beta^i = (f_i^+)^{-1/2} \psi_\beta(i)$  ; puis par transition les  $\varphi_\beta^j = \lambda_\beta^{-1/2} \Sigma \{ \varphi_\beta^i \quad f_i^{-j} \mid i \in I \}$  ; et enfin les

$$F_\beta(i) = \lambda_\beta^{1/2} \varphi_\beta^i \quad , \quad G_\beta(j) = \lambda_\beta^{1/2} \varphi_\beta^j \quad .$$

7.2 Applications : Le format des résultats issus du tableau en quatre blocs étant connu, reprenons en les complétant les applications annoncées au § 1.

7.2.1 Tableaux carrés non symétriques : A un tableau carré non symétrique  $kO_{II}$  (e.g.  $kO(i, i')$  = valeur des exportations du pays  $i$  vers le pays  $i'$ ) il est commun d'associer le tableau transposé  ${}^t kO_{II} = kI_{II}$  (i.e.  $kI(i, i') = kO(i', i)$  = valeur des importations du pays  $i$  en provenance du pays  $i'$ ). Le tableau  $I \times U$  obtenu en juxtaposant les tableaux  $kO$  et  $kI$  contient dans chacune de ses lignes  $i$  toutes les informations relatives au pays  $i$  : exportations d'abord  $kO$ , puis importations  $kI$  : les facteurs  $F_{\alpha}(i)$  calculés d'après le tableau  $I \times U$  rendront compte de l'ensemble de l'activité commerciale du pays  $i$ . Mais parfois, on considère également le tableau en quatre blocs construit suivant la règle générale de E. Yagolnitzer en juxtaposant et superposant quatre blocs : il importe de savoir que les facteurs issus d'une telle analyse se rapportent soit au tableau symétrisé  $k_{II}^+$  (tableau qui dans l'exemple considéré mesure l'intensité des échanges entre pays, indépendamment du sens de ces échanges :  $k^+(i, i') = kO(i, i') + kO(i', i)$ ) soit au tableau différence  $k_{II}^-$  (i.e. dans l'exemple tableau antisymétrique du déséquilibre de la balance commerciale).

7.2.2 Tableaux doublement dédoublés : Chaque fois que les données contenues dans la colonne  $j$  d'un tableau  $kO_{IJ}$  sont des nombres positifs compris entre 0 et un maximum  $\max(j)$ , il est commun d'associer à  $kO_{IJ}$  le tableau  $kI_{IJ}$  des notes complémentaires définies par la formule (cf [DE-DOU] Ca Vol I n<sup>os</sup> 2-3) :

$$kI(i, j) = \max(j) - kO(i, j) ;$$

en particulier dans le cas d'un tableau booléen (en 0,1), si  $kO(i, j) = 1$  s'interprète "l'individu  $i$  possède la qualité  $j$ ",  $kI(i, j) = 1$  signifie au contraire "l'individu  $i$  ne possède pas la qualité  $j$ "; etc. Généralement, on soumet à l'analyse des correspondances le tableau à 2 Card  $J$  colonnes obtenu en juxtaposant  $kO$  et  $kI$  : ainsi (cf 7.2.1) chaque ligne contient toutes les informations relatives à un même individu  $i$ , sans donner aux qualités  $j$  de polarité privilégiée (e.g. si figure la note  $kO(i, j)$  = intensité de l'intérêt de l'individu  $i$  pour la télévision; figure aussi  $kI(i, j)$  = détachement de  $i$  vis à vis de la télévision). Mais parfois il y a des raisons de dédoubler non seulement les colonnes, mais les lignes ; et c'est pourquoi on a considéré les tableaux doublement dédoublés, qui ne sont autres que des tableaux en quatre blocs, au sens de E. Y.. Il importe de savoir que les facteurs issus de tels tableaux se rapportent tous au tableau différence  $k_{IJ}^-$  : en effet dans ce cas les lignes du tableau  $k^+$  sont toutes égales entre elles et l'on a :

$$k^+(i, j) = kO(i, j) + kI(i, j) = \max(j) ;$$

le tableau  $k_{IJ}^+$  ne fournit donc aucun facteur relatif à une valeur propre non nulle.

Dans le cas particulier d'un tableau de préférence, on pose d'ordinaire :

$$kO(i, j) = \text{Card } J - (\text{rang assigné par l'individu } i \text{ à l'objet } j)$$

ainsi les  $kO(i, j)$  deviennent des notes variant de 0 à  $\text{Card } J - 1$ , et d'autant plus élevées que l'individu  $i$  reconnaît plus de mérite à l'objet  $j$  ; et les notes complémentaires sont :

$$kl(i,j) = \text{Card } J - 1 - kO(i,j) ;$$

toutes les lignes du tableau somme  $k^+$  ont même total ( $\text{Card } J)(\text{Card } j - 1)$  ; toutes les lignes du tableau différence  $k^-$  ont zéro pour total (car en bref, tous les rangs figurent dans chaque ligne une fois et une seule) ; donc on a pour les profils la relation :

$$\Sigma\{f_i^{-j} f_j^+ \mid j \in J\} = \Sigma\{f_{ij}^- \mid j \in J\} = 0$$

et comme les facteurs  $\varphi_\alpha^J$  sur  $J$  sont d'après la formule de transition des combinaisons linéaires des  $f_i^{-J}$  (considérées comme fonctions sur  $J$ ), ces facteurs sont *centrés* sur  $J$  muni de la loi  $f_J^+$  (propriété qui on l'a dit n'est pas généralement vraie des facteurs issus de  $f_{IJ}^-$ ) ; i.e. on a :

$$\Sigma\{\varphi_\beta^J f_j^+ \mid j \in J\} = 0.$$

**7.2.3 Cas des jumeaux de même sexe :** On suppose donné un ensemble  $I$  de paires de jumeaux de même sexe ; l'ensemble  $J$  de mesure ayant été effectué sur chacun des individus. Supposons qu'on nomme  $i_0$  et  $i_1$  les deux individus formant la paire  $i$  : on a pour cette paire  $2 \text{ Card } J$  mesures ; mais si l'on écrit ces mesures sur une ligne, on peut soit placer d'abord les mesures relatives à  $i_0$ , puis ensuite celles relatives à  $i_1$  ; soit faire l'inverse ; car ici, en l'absence de différence de sexe, rien ne suggère un ordre dans la paire  $i = \{i_0, i_1\}$ . Si cependant d'une manière *arbitraire* un choix est fait au sein de chaque paire, d'un individu *zéro* et d'un individu *un*, on pourra définir deux tableaux  $kO_{IJ}$  et  $kl_{IJ}$  contenant le premier les mesures relatives aux individus *zéro*, le second les mesures relatives aux individus *un*. Il importe de voir qu'en définitive les résultats de l'analyse du tableau en quatre blocs ne dépendent pas du choix *arbitraire* fait.

Il est d'abord clair que le tableau  $f_{IJ}^+$  ne dépend pas de ce choix ; car  $k^+(i,j)$  est simplement la somme des mesures afférentes aux deux individus formant la paire  $i$ . Quant au tableau différence  $f_{IJ}^-$  chacune de ses lignes  $i$  n'est définie qu'au signe près, le signe dépendant de ce qu'on peut appeler l'orientation choisie sur la paire  $i$ , i.e. du choix de  $i_0$  et de  $i_1$  ; cependant si l'on modifie l'ensemble de la convention d'orientation, ce que nous ferons en forme par une fonction  $\eta^I$  définie sur  $I$  avec  $\forall i \in I, \eta^i \in \{+1, -1\}$  et  $\eta^i = -1$  si et seulement sur la paire  $i$ , il y a changement d'orientation, alors l'on retrouve les mêmes facteurs aux changements près suivants :

$$\varphi_\beta^j \rightarrow \varphi_\beta^j ; \varphi_\beta^i \rightarrow \eta^i \varphi_\beta^i$$

Finalement le facteur  $\varphi_\beta^I$  peut être regardé comme une fonction sur les paires *orientées* dont la valeur change de signe avec l'orientation ; le calcul des facteurs pouvant se faire comme pour tout tableau en quatre blocs (cf § 7.1), après qu'on ait choisi arbitrairement une orientation sur chaque paire (pour fixer  $kO$  et  $kl$ ). Pour un exemple d'application (cf note sur la structure des dermatoglyphes pp 291-292).

7.2.4 Les tableaux de Burt et leurs marges : Reprenons l'exemple du § 1.3 :  $k_0$  et  $k_1$  sont des tableaux de Burt remplis en passant au même questionnaire à deux populations ; par exemple d'une part des jeunes gens, d'autre part des jeunes filles. On a déjà dit au § 4.3 qu'il convenait d'amener les tableaux  $k_0$  et  $k_1$  à avoir même total, afin que la différence  $k^-$  ne fût pas en première approximation proportionnelle à  $k_0$ ,  $k_1$ ,  $k^+$  (qui *a priori* ne diffèrent pas grandement). Il est apparu dans une analyse de données réelles que cette précaution ne suffit pas : de façon précise si  $k_0$  et  $k_1$  ont des marges différentes, cette différence sera à l'origine de deux facteurs de la forme suivante :

$$d^i = (f_i^0 - f_i^1) / (f_i^0 + f_i^1) \quad ;$$

$$d^j = (f_j^0 - f_j^1) / (f_j^0 + f_j^1) \quad ;$$

$$F_1(i) \# d^i \quad ; \quad G_1(j) \# \lambda_1^{1/2} = (\text{var}(d^i))^{1/2} \quad ;$$

$$F_2(i) \# \lambda_2^{1/2} = (\text{var}(d^j))^{1/2} \quad ; \quad G_2(j) \# d^j \quad ;$$

i.e. sur I, le premier facteur est égal à la différence relative des marges de  $k_0$  et  $k_1$  sur I, et sur J il est constant et égal à l'écart-type de  $d^i$  (la variance de  $d^i$  étant la valeur propre  $\lambda_1$ ) ; et le deuxième facteur s'obtient de même *mutatis mutandis* à partir de  $d^j$  ; (évidemment savoir lequel de ces facteurs est effectivement le premier dépend de l'amplitude des variations des marges : soit  $d^I$  l'emporte, soit  $d^J$ ).

Nous ne développerons pas ici ce cas qui est étudié en détail sur un exemple au §4 de l'article [Lycéens].