

J.-P. BENZÉCRI

Sur la généralisation du tableau de Burt et son analyse par bandes

Les cahiers de l'analyse des données, tome 7, n° 1 (1982),
p. 33-43

http://www.numdam.org/item?id=CAD_1982__7_1_33_0

© Les cahiers de l'analyse des données, Dunod, 1982, 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/>

SUR LA GÉNÉRALISATION DU TABLEAU DE BURT ET SON ANALYSE PAR BANDES [BANDES BURT]

par J.-P. Benzécri (1)

0 Rappel : Le tableau de Burt associé à un tableau sous forme disjonctive complète

Classiquement, on part du tableau k_{IJ} en $(0,1)$ recensant les réponses fournies par un ensemble I de sujets (ou individus) à un ensemble Q de questions avec :

$$J = \cup \{J_q | q \in Q\} ;$$

on a $k(i,j) = 1$ si le sujet i adopte la modalité j , chaque sujet adoptant une modalité et une seule de l'ensemble J_q des modalités de réponses admises par chaque question q .

Le tableau de Burt k'_{JJ} , ou tableau de cooccurrence des modalités est alors défini par l'une ou l'autre des formules suivantes :

$$\begin{aligned} k'(j,j') &= \text{Card}\{i | i \in I ; k(i,j) = k(i,j') = 1\} \\ &= \Sigma\{k(i,j) k(i,j') | i \in I\} \\ &= \Sigma\{k(i,j) k(i,j')/k(i) | i \in I\} \cdot \text{Card } Q ; \end{aligned}$$

sur la première de ces formules on lit immédiatement que $k'(j,j')$ est le nombre de sujets (ou individus) i ayant adopté simultanément les modalités j et j' ; d'après la dernière formule, il apparaît que l'analyse de k'_{JJ} fournit sur J les mêmes facteurs normalisés φ^J que l'analyse du tableau initial k_{IJ} (avec toutefois des valeurs propres élevées au carré) car la transition f^J_J associée au tableau de correspondance k'_{JJ} , n'est autre que le composé π^J_J des deux transitions f^I_J et f^J_I associées au tableau k_{IJ} : $f^J_J = \pi^J_J = f^I_J \circ f^J_I$, ou :

$$f^{j,j'}_j = \Sigma\{f^i_j \cdot f^{j'}_i | i \in I\} ; \text{ car :}$$

$$k'(j,j')/k'(j') = \Sigma\{(k(i,j)/k(i))(k(i,j')/k(j')) | i \in I\},$$

parce que $k'(j') = \text{Card } Q k(j')$; (d'où $f^j_j = f_j$, ou $f^J_J = f_J$).

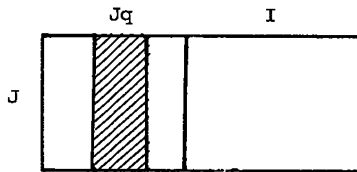
(1) Professeur de statistique. Université Pierre et Marie Curie.

Pour placer les individus i dans l'analyse du tableau de Burt k'_{JJ} , on adjoint simplement à celui-ci, en lignes supplémentaires les lignes du tableau initial k_{IJ} . On obtient ainsi sur l'ensemble I les mêmes facteurs $F(i)$ (coordonnées sur les axes factoriels) que par l'analyse du tableau initial k_{IJ} . La raison essentielle de cette identité est que dans l'analyse de k'_{JJ} et dans celle de k_{IJ} on opère dans un même espace R_J des profils sur J , avec pour métrique la métrique du χ^2 de centre $f'_J = f_J$, définie par la loi marginale commune aux deux tableaux k'_{JJ} et k_{IJ} . Il s'agit donc dans les deux cas d'un même nuage $N(I)$; qui de plus se trouve projeté sur les mêmes axes.

On peut regarder de ce même point de vue la structure en blocs du tableau de Burt, ou plus précisément sa structure en bandes. Appelons bande $k'_{J \times Jq}$ (ou en bref k'_q ou bande q) le sous tableau du tableau de Burt formé de l'ensemble des colonnes afférentes à une question q . Le tableau $J \times Jq$ ainsi construit a même marge sur J que le tableau de Burt (ou le tableau initial) car :

$$k'_q(j) = k(j) = k'(j)/\text{Card } Q.$$

On a dès le début (cf. [BIN. MULT.]) remarqué qu'il résulte de cette propriété que dans l'analyse du tableau de Burt (ou dans celle du tableau initial) le sous nuage $N(Jq)$ des modalités de réponse à la question q a son centre de gravité à l'origine (si chaque j de Jq est muni de sa masse naturelle, ou fréquence $k(j)$). Nous voulons souligner ici que par l'analyse de la bande $J \times Jq$, le nuage $N(Jq)$ des colonnes se trouve placé dans le même espace R_J , muni de la même métrique de centre f_J que lors de l'analyse du tableau de Burt complet. Il n'en est pas de même pour le nuage des lignes (à chacune desquelles est associé un profil sur Jq) mais il est facile, par la technique des éléments supplémentaires de déduire de l'analyse de la bande q une vue nouvelle des nuages $N'(J)$ (associé au tableau de Burt) et $N(I)$ (associé au tableau initial), en projection sur les axes et plans issus de cette analyse. Le schéma ci-joint précise la disposition des



BURT tableau en $(0,1)$

N.B. seule la bande hachurée est en principal ; le reste est en supplémentaire.

tableaux. Ce qu'on a dit de la bande q afférente à une question unique q , vaut aussi bien pour une bande plus large, afférente à un sous ensemble quelconque Q' de l'ensemble Q des questions. On doit alors considérer le tableau $J \times JQ'$, où

$$JQ' = \cup \{Jq \mid q \in Q'\} ;$$

il importe seulement de noter que le format disjonctif complet jouant

ici un rôle essentiel, le bloc JQ' des colonnes principales doit nécessairement comprendre toutes les modalités de réponse à une question dès qu'il en comprend une seule.

L'analyse des bandes du tableau de Burt est intéressante à divers titres parce qu'elle reprend et coordonne des pratiques déjà reçues en analyse des données. Nous soulignerons trois aspects.

Allègement des calculs : Le coût d'une analyse factorielle dépendant principalement de la plus petite dimension du tableau principal analysé, il est clair que l'analyse d'une bande est moins coûteuse que celle du tableau entier. D'ailleurs pour une partition du tableau en quelques bandes de largeur acceptable, on peut en retenant les premiers facteurs de l'analyse de chacune de ces bandes, obtenir l'essentiel de l'information que recèle l'ensemble du tableau carré de Burt (donc aussi du tableau initial) ; et B. Escofier a montré dans sa thèse qu'on peut même, par combinaison linéaire des facteurs retenus, obtenir une excellente approximation des facteurs de l'analyse globale.

Relations entre groupes de variables : Il est classique d'analyser des sous tableaux du tableau de Burt, croisant (e.g.) les ensembles JQ1 et JQ2 des modalités de réponses à deux parties Q1 et Q2 du questionnaire Q. Ainsi JQ1 est caractérisé par ses associations avec JQ2 et réciproquement.

Cependant les nuages $N(JQ1)$ et $N(JQ2)$ construits par l'analyse d'un sous tableau ne sont pas dans le même espace que les sous nuages JQ1 et JQ2 du nuage général $N(J)$. En analysant une bande $J \times JQ1$, on obtient une représentation de JQ1 (ensembles des colonnes principales) qui peut être regardée d'une part comme rendant compte de l'interaction de JQ1 avec l'ensemble J de toutes les modalités (y compris celles de JQ1 lui-même), et d'autre part comme une vue globale de J par rapport à des axes qui sont ajustés non à $N(J)$ tout entier, mais à son sous-nuage $N(JQ1)$. On perd, il est vrai, la spécificité d'une image fondée sur $JQ1 \times JQ2$ (à l'exclusion des autres variables), mais vuela redondance fréquente des informations, l'image résultant de $J \times JQ1$ peut n'être pas différente de celle issue de $JQ2 \times JQ1$; et on gagne, en tout cas, quant à la cohérence des diverses présentations qu'on offre des données.

Interprétation des facteurs : On lit communément des discours tels que le suivant : "le plan 1×3 est dominé par la variable taille du ménage, dont les modalités se succèdent dans leur ordre naturel sur une courbe parabolique exprimant un fort effet Guttman "ou encore" : l'axe 2 est un axe de ruralité, dominé par la variable taille de la commune" etc. . Il est de règle alors, qu'on interprète le plan ou l'axe, en rapportant à la variable dominante, l'ensemble des modalités des autres variables. Dans ces conditions, n'est-il pas opportun de considérer du nuage $N(J)$, une projection sur un axe ou un plan choisi expressément afin que les modalités de telle variable (ou de tel groupe de variables) y soient représentées au mieux. Par rapport à l'analyse globale, la représentation de la variable choisie ne pourra qu'être améliorée (quant à l'inertie du sous nuage projeté) ; et par le fait même, les corrélations de cette variable avec les autres apparaîtront au plus clair. On perd, il est vrai, l'intérêt qu'il y a, à voir une variable se détacher spontanément de l'analyse globale ; mais on peut ici faire trois remarques. D'abord l'émergence d'une variable peut souvent résulter de la redondance du questionnaire autant et plus que l'importance de la variable elle-même. Ensuite l'interprétation d'un axe de rang élevé est souvent rendue incertaine parce que la variable la mieux corrélée à l'axe est déjà en partie représentée sur les axes précédents. Enfin l'importance relativement au nuage global $N(J)$ d'axes construits par ajustement au sous nuage $N(Jq)$ (ou $N(JQ1)$), s'apprécie

numériquement par le calcul de l'inertie du nuage global en projection sur ces axes : d'où des généralisations des valeurs propres et des taux, dont le calcul est aisé grâce aux colonnes CTR du listage des éléments supplémentaires. (Nous y reviendrons au § 2).

En somme, tout ce qui précède fait grandement désirer qu'on applique à tout tableau de données les constructions auxquelles se prêtent les tableaux sous forme disjonctive complète. Le découpage des variables en classes, permet en effet de réduire au format disjonctif complet les données les plus diverses et les plus hétérogènes. Mais dans l'analyse même des questionnaires, le format disjonctif complet est parfois gênant : on souhaiterait mettre en supplémentaire des modalités rares (notamment les modalités d'omission) mais par là on sort du format disjonctif complet ; il n'est plus vrai que les bandes du tableau principal ainsi élaguées aient toutes exactement la même loi marginale. D'ailleurs, il y a de nombreuses données qui ne sont pas sous forme logique (en 0,1) et qu'il convient d'analyser sans recodage. Or aussi bien la construction du tableau de Burt que l'analyse des bandes s'étendent à des tableaux de correspondance $I \times J$ quelconques, au prix d'adaptations minimales : c'est ce qu'on exposera dans la suite.

1 Construction générale du tableau de Burt

Etant donné un tableau de correspondance quelconque $k(i,j)$ sur $I \times J$, nous poserons :

$$k''(j,j') = \Sigma \{k(i,j) k(i,j') / k(i) \mid i \in I\} ;$$

au coefficient Card Q près (lequel en analyse des correspondances est sans effet), c'est la définition du tableau k' donnée au § 0. Nous appellerons le tableau $k''_{J \times J}$ tableau de Burt généralisé associé à $k_{I \times J}$, ou en bref tableau de Burt. Par construction le tableau de Burt k'' est symétrique ($k''(j,j') = k''(j',j)$), et il a même marge sur J que le tableau de base k :

$$\begin{aligned} k''(j) &= \Sigma \{k''(j,j') \mid j' \in J\} \\ &= \Sigma \{k(i,j) k(i,j') / k(i) \mid i \in I ; j' \in J\} \\ &= \Sigma \{k(i,j) \Sigma \{k(i,j') \mid j' \in J\} / k(i) \mid i \in I\} \\ &= \Sigma \{k(i,j) (k(i) / k(i)) \mid i \in I\} = k(j). \end{aligned}$$

D'où il résulte qu'on a (comme au § 0) :

$$f''^J_J = \pi^J_J = f^I_J \circ f^J_I.$$

(formule où f'' et f désignent respectivement des transitions associées aux tableaux k'' et k). Le tableau de Burt k'' produit les mêmes facteurs normalisés φ^J sur J que le tableau de base k , avec seulement des valeurs propres élevées au carré : $\lambda'' = (\lambda)^2$. On a les facteurs $F(i)$ issus de k en adjoignant ce tableau en éléments supplémentaires au tableau de Burt k'' .

Il importe de regarder le rôle du diviseur $k(i)$ dans la définition du tableau k'' . Supposons que chaque ligne du tableau de base k soit multipliée par son total $k(i)$ et notons

$$K(i,j) = k(i,j) \times k(i).$$

Au tableau K est associé un tableau de Burt K'' ; par la formule :

$$K''(j, j') = \Sigma \{k(i)k(i, j)k(i)k(i, j')\} / (k(i))^2 | i \in I\} \\ = \Sigma \{k(i, j) k(i, j') | i \in I\}$$

(en effet $K(i)$ total de la ligne i du tableau K est le carré de $k(i)$). Donc si on omet le diviseur $k(i)$ dans la définition de k'' , on obtient exactement le tableau de Burt K'' associé à un tableau K qui ne diffère de k que par une repondération de chacune de ses lignes par $k(i)$. (Et il serait facile d'introduire dans la définition de k'' des coefficients de pondération quelconques).

Supprimer le diviseur $k(i)$, c'est suivre la définition usuelle du tableau de Burt (§ 0) : cela ne change rien si toutes les lignes ont même total, ce qui est le cas non seulement pour un tableau en $(0,1)$ sous forme disjonctive complète, mais aussi pour un tableau codé sous forme floue (i.e. avec pour tout individu i et toute question q un ensemble $\{k(i, j) | j \in J_q\}$ de nombres quelconques astreints seulement à avoir pour total 1) ! mais dans ce dernier cas, P. Cazes a montré (cf. [ANA. BLOCS II], §§ 4.2.3, 4.2.4, *Cahiers*, Vol V n° 4 (1980)) qu'il était préférable de réduire à leur diagonale principale les sous tableaux carrés diagonaux $J_q \times J_q$ du tableau de Burt.

Soit J' un sous ensemble quelconque de l'ensemble J des modalités d'un tableau de base en $(0,1)$ k_{IJ} . Le sous tableau carré $J' \times J'$, du tableau de Burt (usuel : k') issu de k_{IJ} apparaît comme ne différant du tableau de Burt (généralisé : k'') associé au tableau k_{IJ} , (en $(0,1)$); mais, généralement, non sous forme disjonctive complète ; du fait de l'élimination de diverses modalités) que par des repondérations : le poids de chaque individu étant multiplié, en bref, par le nombre de questions auxquelles il a fourni une réponse j rentrant dans l'ensemble J' des modalités conservées ; modification qui peut même paraître souhaitable, ce qui justifie pleinement l'analyse de sous tableaux $J' \times J'$ du tableau de Burt usuel k' à un questionnaire.

2 Bandes du tableau de Burt et leur généralisation

Ainsi qu'on l'a dit au § 0 l'analyse des bandes $J \times JQ'$ utilise le fait que le sous tableau a même profil de marge sur J que le tableau carré tout entier $J \times J$; et elle requiert des calculs d'inertie du nuage total projeté sur les axes ajustés au sous nuage JQ' . Il faut généraliser tout cela en partant du cas d'un tableau usuel sous forme disjonctive complète.

Soit donc $Q' \subset Q$; $JQ' = \cup \{Jq | q \in Q'\} \subset J$. Prenons en main le listage de l'analyse de la bande $J \times JQ'$ du tableau de Burt, avec en éléments supplémentaires l'ensemble $(J - JQ')$ des modalités restantes. La trace, ou inertie totale associée au tableau $J \times JQ'$ se calcule à partir d'une valeur propre et du taux correspondant écrit sur le listage.

$$\text{trace } b = \lambda b_1 / \tau b_1 = \text{trace } b(JQ')$$

(où l'initiale "b" rappelle qu'il s'agit de la bande). Cette trace est aussi l'inertie du nuage $N_b(JQ')$ (des colonnes principales). On a l'inertie du nuage $N_b(J - JQ')$ des colonnes supplémentaires en multipliant trace b par le total de la colonne INR des éléments supplémentaires (plus exactement on doit diviser par 1000 le total, exprimé en millièmes, que porte le listage). Notons :

$$\text{trace } b(J) = \text{trace } b(JQ') (1 + \Sigma \{ \text{INR}(j) | j \in J - JQ' \}) .$$

A un coefficient près qu'on calculera $\text{trace } b(J)$ est aussi l'inertie totale du nuage $N(J)$ telle qu'elle résulte de l'analyse du tableau de Burt carré $J \times J$. D'où vient le coefficient ? Dans les deux analyses ($J \setminus JQ'$ ou $J \times J$) il s'agit du même nuage $N(J)$ mais dans l'analyse de la bande le poids total des modalités afférentes à une question est $1/\text{Card } Q'$ (afin que le poids total pour JQ' soit 1) ; tandis que dans l'analyse du tableau carré, ce poids est $1/\text{Card } Q$. On a donc.

$$\text{trace}(J) = (\text{Card } Q' / \text{Card } Q) \text{trace } b(J)$$

(où $\text{trace}(J)$ est la trace calculée sur l'analyse du tableau carré).

Un calcul analogue, mais fondé sur la colonne CTR des éléments supplémentaires donne l'inertie $\lambda_\alpha(J)$ du nuage $N(J)$ (avec sa masse totale 1) sur l'axe α issu de l'analyse de la bande $J \times JQ'$. On a :

$$\lambda_\alpha(J) = \lambda_\alpha(1 + \Sigma\{\text{CTR}\alpha(j) \mid j \in J - JQ'\}) (\text{Card } Q' / \text{Card } Q) ;$$

où λ_α est la valeur propre lue sur le listage de l'analyse de $J \times JQ'$.

Eventuellement, on calculera $\tau_\alpha(J) = \lambda_\alpha(J) / \text{trace}(J)$.

Il importe de noter que, puisque les axes issus de $J \times JQ'$, sont ajustés au sous nuage $N(JQ')$ et non à $N(J)$, la suite des inerties $\lambda_\alpha(J)$ ne décroît pas nécessairement quand α augmente. En revanche on est assuré que la valeur maxima de $\lambda_\alpha(J)$ est inférieure (ou, au plus, égale) à la première valeur propre λ_1 issue de l'analyse du tableau carré $J \times J$. En effet λ_1 réalise un maximum absolu. De même le maximum de la somme $\lambda_\alpha(J) + \lambda_\beta(J)$ calculée pour un plan (α, β) quelconque ($\alpha \neq \beta$) est inférieure à $\lambda_1 + \lambda_2$; et en général le maximum de la somme :

$$\lambda_{\alpha_1}(J) + \lambda_{\alpha_2}(J) + \dots + \lambda_{\alpha_r}(J)$$

(où $\alpha_1, \alpha_2, \dots, \alpha_r$ désignent des indices tous distincts entre eux) est inférieure à la somme $\lambda_1 + \dots + \lambda_r$ des r premières valeurs propres issues de l'analyse du tableau carré de Burt $J \times J$.

Les considérations ci-dessus s'appliquent *sans aucune modification* au cas d'un tableau k sur $J \times J$ (qu'on supposera symétrique, bien qu'à la vérité la nature de l'ensemble des lignes n'intervienne pas) partagé en bandes ayant chacune même profil de marge au sens suivant

$$\begin{aligned} J &= \cup\{Jq \mid q \in Q\}; \\ \forall j \in J, \forall q \in Q : \Sigma\{k(j, jq) \mid jq \in Jq\} / kq \\ &= \Sigma\{k(j, j') \mid j' \in J\} / k, \end{aligned}$$

où on a noté kq et k les totaux généraux respectifs de la bande et du carré :

$$kq = \Sigma\{k(j, j') \mid j \in J, j' \in Jq\} ; k = \Sigma\{k(j, j') \mid j \in J, j' \in J\}.$$

Si les kq ne sont pas tous égaux entre eux, il faut seulement substituer au coefficient $(\text{Card } Q' / \text{Card } Q)$ des calculs ci-dessus le quotient $\Sigma\{kq \mid q \in Q\} / k$. En particulier est partagé en bandes de cette sorte le tableau k' suivant issu du recensement de Dijon et analysé par M. Essadaoui (cf. thèse et [DIJON] { à paraître}).

On a :

I un ensemble d'ilots d'une ville ;

Q un ensemble de variables pouvant décrire un logement

$k(i, j_q)$ = nombre de logements de l'îlot i rentrant dans la modalité j_q de la variable q .

$k'(j, j') = \Sigma \{k(i, j) k(i, j') / k(i) \mid i \in I\}$; tableau de Burt généralisé construit comme il est expliqué au § 1.

Il est clair en effet que le tableau k hérite du tableau (non explicitement construit) décrivant tous les logements, la propriété d'avoir même profil vertical de marge pour chaque bloc J_q de colonnes.

Mais que faire pour les bandes qu'on peut être amené à considérer dans un tableau de Burt généralisé $J \times J$, qui n'est pas donné *a priori* avec une structure de blocs satisfaisant aux conditions ci-dessus ? Pour traiter ce cas nous serons amené à une généralisation qui améliorera même en un certain sens le traitement du cas particulier où est satisfaite la condition de constance du profil des bandes.

Soit donc $k_{J'}^J$ un tableau de Burt, construit comme on l'a dit au § 1 ; et soit J' une partie quelconque de J (le choix de J' est guidé par la signification des données, par des analyses préalables etc. ; mais les constructions qui suivent valent quel que soit $J' \subset J$). On note :

$$J'r = J' \cup \{\text{reste}(J')\}$$

et on définit un tableau $J \times J'r$ en adjoignant à la bande $J \times J'$ du tableau de Burt $J \times J$, une colonne "reste" obtenue en cumulant l'ensemble $J - J'$ des colonnes restantes. Le tableau $J \times J'r$ a, par construction, même colonne de marge que le tableau de Burt. Le nuage $N(J'r)$, associé au tableau $J \times J'r$, comprend d'une part le sous-nuage $N(J')$ du nuage $N(J)$ associé au tableau $J \times J$, avec non seulement les mêmes distances dans le même espace R_J , mais aussi les mêmes masses, et d'autre part le point "reste" qui n'est autre que le sous nuage $N(J - J')$ concentré en son centre de gravité avec sa masse. Ainsi avec le sous ensemble J' des colonnes principales et l'ensemble $J - J'$ des colonnes supplémentaires on a une représentation de $N(J)$ sur des axes adaptés à $J'r$.

Il importe d'abord de voir que dans le cas où la bande $J \times J'$ du tableau de Burt k a même colonne de marge que le tableau carré, l'analyse du tableau $J \times J'r$ ne diffère de celle de la bande $J \times J$ que par un changement des masses des colonnes, visible seulement dans les λ , les INR et les CTR du listage. En effet quant aux nuages la seule différence est le point "reste", qui étant situé alors à l'origine, i.e. au centre de gravité de $N(J')$ aussi bien que de $N(J)$ ou de $N(J - J')$ ne change rien aux axes.

Ceci dit le calcul de l'inertie totale du nuage $N(J)$ et de son inertie en projection sur un axe se fait même plus simplement que précédemment car le coefficient $(\text{Card} Q' / \text{Card} Q)$ a disparu. On a :

$$\text{trace}(J) = \text{trace}(J'r) \times (1 - \text{INR}(\text{reste})) + \Sigma \{\text{INR}(j) \mid j \in J - J'\}$$

où (cf. *supra*), $\text{trace} J'r = \lambda_1 / \tau_1$; et le terme retranché $\text{INR}(\text{reste})$ est nul si $J \times J'$ a même colonne de marge que $J \times J$; pour un facteur il vient de même :

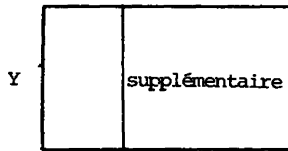
$$\lambda_\alpha(J) = \lambda_\alpha(1 - \text{CTR}_\alpha(\text{reste})) + \Sigma \{\text{CTR}_\alpha(j) \mid j \in J - J'\}.$$

Inutile de dire que ces calculs se font sans peine en dépouillant le listage. La seule particularité que présente l'analyse d'une bande est de calculer la colonne "reste" qui doit précéder l'application du programme usuel d'analyse de correspondance.

3 Calcul approché des facteurs issus du tableau de Burt après des analyses de bandes

Le principe de ce calcul (qui vient de la thèse d'état de B. Escofier (Paris 6) (1979) est simple. Dans l'espace R_J des profils sur J muni de la métrique du χ^2 de centre f_J , on fait choix d'un sous-espace E ajusté approximativement au nuage $N(J)$. On cherche alors les axes principaux d'inertie du nuage $N_E(J)$ obtenu en projetant orthogonalement $N(J)$ sur E : les coordonnées sur ces axes sont les facteurs approchés.

Ici le tableau découpé en bandes est un tableau carré symétrique $J \times J$: mais cette propriété particulière n'est pas requise pour les calculs que nous développons ci-dessous. Donc, pour plus de généralité, nous supposons que le tableau k'' sur lequel on opère est un tableau de correspondance quelconque dont on note Y l'ensemble des lignes et J l'ensemble des colonnes. Ainsi le nuage $N(J)$ est dans l'espace R_Y des profils sur Y . Le sous-espace E étant dans R_Y est noté E_Y . L'analyse des bandes se fait comme au § 2 : à une partie J_r de J on adjoint une colonne "reste" obtenue en cumulant les colonnes de $J - J'$. Quant à l'ensemble I des individus à partir desquels au § 1 est construit le tableau de Burt k''_{JJ} , il nous suffit ici de le considérer comme un ensemble de colonnes supplémentaires quelconques adjointes au tableau $Y \times J$.



Dans le calcul approché des facteurs, nous distinguerons ci-dessous trois étapes.

3.1 Choix du sous-espace E_Y : Ce sous-espace passe par le centre de gravité de $N(J)$ et est engendré par des axes factoriels issus d'analyses de bandes $Y \times J'$. Les parties J' de J sont choisies en se basant sur la signification des données et sur les résultats des analyses partielles déjà faites. Supposons qu'on commence par une première partie J'_1 de J : on retient les premiers facteurs issus de $Y \times J'_1$ de façon à avoir une qualité de représentation satisfaisante pour J'_1 ; par le fait même, certains des éléments de $J - J'_1$ (fortement corrélés à J'_1) se trouvent eux aussi avoir une bonne qualité de représentation : mais s'il reste un résidu d'éléments de J dont la qualité de représentation est jugée insuffisante, on analyse une bande $Y \times J'_2$; etc. le choix des bandes successives J'_2, J'_3 restant guidé à la fois par la qualité QLT, par le sens des données (faire des blocs homogènes de variables) et la contrainte d'espace qui limite Card J' ...

3.2 Coordonnées dans le sous-espace E_Y : Supposons qu'on ait retenu un facteur α issu de l'analyse d'une bande $Y \times J'$. Ce facteur est défini par un vecteur axial unitaire $u_{\alpha Y}$ de R_Y dont les composantes sont données par :

$$u_{\alpha Y} = \{u_{\alpha Y}\} = \{\lambda_{\alpha}^{-1/2} F_{\alpha}(y) f_y | y \in Y\}$$

Dans cette formule, λ_α est la v.p. issue de l'analyse de $Y \times J'r$; et $F_\alpha(y)$ est le facteur α calculé sur les lignes, toutes principales, du tableau $Y \times J'r$. Les coordonnées sur l'axe $u_{\alpha Y}$ des points du nuage $N(J)$ sont données pour $j \in J'$, par les valeurs du facteur $G_\alpha(j)$ pour les colonnes principales; et pour $j \in J - J'$, par les valeurs de $G_\alpha(j)$ pour les colonnes supplémentaires.

Les vecteurs $u_{\alpha Y}$ retenus dans les diverses analyses engendrent E_Y : mais ils ne constituent pas une base orthonormée: il convient de les orthonormaliser par la procédure usuelle de Schmidt et Hilbert (cf. [ALG. EUCL.] TIIB n° 12 § 1.3).

(Le produit scalaire comporte ici les coefficients $(1/f_Y)$ de la métrique du χ^2 :

$$\langle u_{\alpha Y}, u_{\alpha' Y} \rangle = \Sigma \{ u_{\alpha Y} u_{\alpha' Y} / f_Y | Y \in Y \};$$

si la procédure utilisée ne comporte pas de coefficients dans le produit scalaire, on changera de coordonnées en posant

$$v_{\alpha Y} = u_{\alpha Y} (f_Y)^{-1/2};$$

et on orthogonalisera les $v_{\alpha Y}$ pour lesquels le produit scalaire est la somme des produits deux à deux des composantes :

$$\langle v_{\alpha Y}, v_{\alpha' Y} \rangle = \Sigma \{ v_{\alpha Y} v_{\alpha' Y} | Y \in Y \}.$$

Nous supposons désormais que les vecteurs $u_{\alpha Y}$ après orthonormalisation sont désignés par la même lettre u , mais avec de nouveaux indices β , les coordonnées sur ces nouveaux axes étant notées $G_\beta(j)$; si donc :

$$u_{\beta Y} = \Sigma \{ h_{\beta \alpha} u_{\alpha Y} | \alpha \in A \}$$

(où A désigne l'ensemble des axes retenus de diverses analyses) on a aussi :

$$G_\beta(j) = \Sigma \{ h_{\beta \alpha} G_\alpha(j) | \alpha \in A \}$$

3.3 Recherche des axes principaux d'inertie : Dans E_Y muni du système orthonormé des axes $u_{\beta Y}$ cette recherche est très simple. On note :

$$\text{som}(\beta, \beta') = \Sigma \{ G_\beta(j) G_{\beta'}(j) f_j | j \in J \};$$

on diagonalise cette matrice d'inertie: soit donc (en notant B l'ensemble des vecteurs axiaux orthonormés) :

$$s_{\gamma B} = \{ s_{\gamma \beta} | \beta \in B \}; \quad \Sigma \{ s_{\gamma \beta}^2 | \beta \in B \} = 1,$$

un vecteur propre normé de la matrice som. La coordonnée sur l'axe correspondant est donnée par :

$$\begin{aligned} G_\gamma(j) &= \Sigma \{ s_{\gamma \beta} G_\beta(j) | \beta \in B \} \\ &= \Sigma \{ s_{\gamma \beta} h_{\beta \alpha} G_\alpha(j) | \beta \in B, \alpha \in A \}, \end{aligned}$$

où les $h_{\beta \alpha}$ sont les coefficients d'orthonormalisation du § 3.2, et les

$G_\alpha(j)$ sont les facteurs initiaux issus d'une analyse de bande.

On pourra lister outre les $G_Y(j)$ des CTR et des COR

$$\begin{aligned} \text{CTR}_Y(j) &= f_j(G_Y(j))^2 / \lambda_Y \\ \text{COR}_Y(j) &= G_Y(j)^2 / \|f_J - f_J^j\|^2 \end{aligned}$$

où λ_Y est la valeur propre (moment d'inertie) de la matrice som, et $\|f_Y - f_Y^j\|^2$ la distance du χ^2 de centre f_Y , distance du point j de $N(J)$ au centre f_Y de ce nuage.

L'analyse du nuage $N(J)$ (plus exactement de sa projection sur le sous espace E_Y de R_Y qui a été choisi au § 3.1) est ainsi complète. Reste à considérer (cf. figure au début du § 3) les ensembles Y et I .

Quant aux individus i de I , ils ont eux-mêmes (comme colonnes supplémentaires aux bandes $Y \times J$ 'r) des facteurs $G_\alpha(i)$ donc aussi des coordonnées $G_\beta(i)$ puis $G_Y(i)$ calculées avec les mêmes coefficients que pour les j , et on peut pour les i comme pour les j calculer des COR. Quant aux CTR, dans le cas où le tableau k analysé est le tableau de Burt (cf. § 1) associé au tableau de base k décrivant I par sa correspondance avec J , ($Y=J$) il faut prendre garde que (à supposer que l'on soit sur les axes exacts issus de $J \times J$ ou de $I \times J$) l'inertie de $N(I)$ est la racine carrée de la v.p. λ issue du tableau de Burt k''_{JJ} .

Dans le cas d'un tableau de Burt symétrique $J \times J$ (cas où $Y = J$), l'analyse est maintenant achevée. Sinon il reste à trouver les facteurs sur l'ensemble Y . Deux voies sont possibles :

a) Calculer explicitement les vecteurs de base u_{YY} par la formule :

$$u_{YY} = \sum \{s_{Y\beta} u_\beta \mid \beta \in B\} = \{u_{YY} \mid Y \in Y\}$$

et poser : $F_Y(y) = u_{YY}(\lambda_Y)^{1/2}$.

b) Transporter les facteurs de J vers Y en appliquant la formule de transition usuelle au tableau $Y \times J$.

Les deux voies ne donnent des résultats identiques que si l'on a atteint les facteurs exacts : avec des facteurs approchés, les résultats diffèrent quelque peu et ce peut être l'occasion d'apprécier la précision du calcul.

4 Autre généralisation de l'analyse des bandes, variance inter-classe et variance intraclasse

Dans ce § nous considérons un tableau de données quelconque k sur $Y \times J$ et proposons d'en déduire d'autres tableaux k sur $Y \times C$, par cumul des colonnes suivant les classes c d'une partition quelconque C de J . De façon précise soit :

$$J = \cup \{c \mid c \in C\} ;$$

$$\forall c, c' \in C : c \neq c' \Leftrightarrow c \cap c' = \emptyset$$

On définit le tableau k sur $Y \times C$ par :

$$k(y, c) = \sum \{k(y, j) \mid j \in c\}$$

Il est clair que le schéma proposé au § 3 rentre dans ce cadre : il suffit de poser :

$$C = \{\{j\} | j \in J'\} \cup \{J - J'\}$$

i.e. une classe par élément j de J' , et la classe "reste" = $J - J'$.

Comme au § 2 le tableau $Y \times C$ a même marge sur Y que le tableau initial $Y \times J$; et le nuage $N(J)$ peut être adjoint en élément supplémentaire à la représentation du nuage $N(C)$ que fournit l'analyse du tableau $Y \times C$. On a donc comme précédemment des vues différentes d'un même nuage $N(J)$; seuls changent les axes : ajustés à $N(J)$ dans l'analyse de $Y \times J$ et à $N(C)$ dans l'analyse de $Y \times C$. Eventuellement on peut utiliser les facteurs issus d'analyses de divers tableaux $Y \times C$ pour construire, comme au § 3 des facteurs approchés relatifs à l'analyse de $Y \times J$. En particulier on peut considérer d'une part le tableau $Y \times C$ associé à une partition déterminée C de J ; et d'autre part les bandes construites comme au § 2 en prenant successivement pour J' chacune des classes c de C .

Comme au § 2 encore, il est facile d'après un listage de l'analyse $Y \times C$, comportant $Y \times J$ en colonnes supplémentaires, de calculer l'inertie totale de $N(J)$ ainsi que son inertie en projection sur chacun des axes ajustés à $N(C)$. De façon précise on a :

$$\text{trace}(J) = \text{trace}(C) \times (\sum \{ \text{INR}(j) | j \in J \}) ;$$

$$\lambda_{\alpha}(J) = \lambda_{\alpha}(C) \times (\sum \{ \text{CTR}_{\alpha}(j) | j \in J \}) .$$

Puisque $\text{trace}(J)$ est l'inertie totale de $N(J)$ tandis que $\text{trace}(C)$ est l'inertie du nuage des centres des classes (classiquement appelée inertie *interclasse*) on peut par différence calculer l'inertie *intra-classe*.

$$\text{trace}(J + C) = \text{trace}(J) - \text{trace}(C) ;$$

et de même pour chaque axe ajusté à C .

Ces calculs servent en particulier quand J est muni d'une structure de produit : $J = C \times T$, chacun des j étant un *objet* c considéré à un *temps* t . Alors on peut aussi bien regarder C que T comme des partitions de J ; ce qui revient à analyser soit le tableau de marge $Y \times C$ soit le tableau de marge $Y \times T$, associé au tableau ternaire $Y \times C \times T$. On a alors avec l'inertie totale diverses inerties partielles :

$$\text{trace}(J) = \text{inertie totale}$$

$$\text{trace}(J) - \text{trace}(C) = \text{inertie intra objet}$$

$$\text{trace}(C) = \text{inertie inter objet}$$

$$\text{trace}(J) - \text{trace}(T) = \text{inertie intra temps}$$

$$\text{trace}(T) = \text{inertie inter temps}$$

Dans les cas usuels $\text{trace}(T)$ est très inférieure à $\text{trace}(C)$; au contraire $\text{trace}(J) - \text{trace}(C)$ est souvent faible relativement à $\text{trace}(J)$, ou tout au plus en est-elle la moitié.