LES CAHIERS DE L'ANALYSE DES DONNÉES

P. CAZES

Correspondance entre deux ensembles et partition de ces deux ensembles

Les cahiers de l'analyse des données, tome 11, nº 3 (1986), p. 335-340

<http://www.numdam.org/item?id=CAD_1986__11_3_335_0>

Les Cahiers de l'Analyse des Données
Vol. XI - 1986 - N° 3 - pp. 335-340

CORRESPONDANCE ENTRE DEUX ENSEMBLES ET PARTITION DE CES DEUX ENSEMBLES

[CORR. PART.]

par P. Cazes

On rencontre en analyse des correspondances des cas divers où des partitions des ensembles jouent un rôle essentiel : nous citerons l'analyse des questionnaires, les tableaux décomposés en une suite de blocs diagonaux en dehors desquels il n'y a que des zéros ; les modèles plus généraux de tableaux construits par blocs suivant des structures hiérarchiques. Dans de tels cas les facteurs sont souvent répartis en deux groupes : d'une part des couples de facteurs dont chacun est constant sur toute classe de la partition dont l'ensemble (I ou J) est muni ; d'autre part des couples de facteurs dont chacun a moyenne nulle sur toute classe de la partition dont I ou J est muni. L'objet du présent travail est de préciser les conditions les plus générales où une telle structure se réalise. Afin de simplifier les calculs, dans la mesure du possible, on se placera dans le cadre de la géométrie euclidienne multidimensionnelle, avec décomposition d'espaces en sommes directes, et produits tensoriels.

1 *Rappel de la définition des facteurs* : Une correspondance f_{IJ} étant donnée sur le produit $I \times J$, chacun des deux ensembles I et J est muni de la loi marginale f_I ou f_J. Ainsi l'ensemble R^I des fonctions à valeur réelle définies sur I devient un espace euclidien, noté L^I_2, l'espace des fonctions de carré sommable sur I : le produit scalaire de deux fonctions θ^I et ψ^I étant donné par la formule

$$<\theta^I , \psi^I> = \Sigma\{\theta^i \psi^i f_i | i \in I\}$$

L'espace L^I_2 contient d'une part la droite des constantes (définie par la fonction 1, souvent notée δ^I), et d'autre part l'hyperplan H^I des fonctions de moyenne nulle (lequel est orthogonal à la droite des constantes). De même pour J, on a L^J_2 et H^J.

L'analyse des correspondances peut être regardée comme une décomposition canonique de la fonction densité d^{IJ} de f_{IJ} par rapport au produit $f_I\, f_J$: la formule de reconstitution s'écrit en effet :

$$d^{ij} = f_{ij}/(f_i . f_j) = 1 + \Sigma_\alpha\{\lambda_\alpha^{1/2} \varphi_\alpha^{\ i} \ \varphi_\alpha^{\ j}\} ;$$

où les $\varphi_\alpha^{\ I}$, $\varphi_\alpha^{\ J}$ sont des fonctions de moyenne nulle, les facteurs

(*) *Professeur de statistique à l'Université Paris Dauphine (CEREMADE).*

Les cahiers de l'analyse des données — 0339-3097/86/03 335 6/$ 2.60/ © Gauthier-Villars

normalisés. Avant toute analyse, il est clair que la fonction $d^{IJ}-1$, est une somme de produits deux à deux d'une fonction sur I par une fonction sur J ; c'est dire que :

$$d^{IJ} - \delta^I \delta^J \in H^I \otimes H^J.$$

Il est également classique d'associer à f_{IJ} une structure euclidienne sur l'espace $R^{I\times J}$ des fonctions sur $I \times J$: on a alors un espace L_2^{IJ} avec le produit scalaire :

$$<\xi^{IJ}, \eta^{IJ}> = \Sigma\{\xi^{ij} \eta^{ij} f_{ij} | i \in I, j \in J\} ;$$

et L_2^I, L_2^J sont des sous-espaces de L_2^{IJ}, parce que, en bref une fonction d'une variable est un cas particulier de fonction de deux variales, la norme L_2 étant d'autre part respectée. L'analyse de f_{IJ} s'interprète alors comme l'étude dans l'espace euclidien L_2^{IJ} de la figure formée par les deux sous-espaces L_2^I et L_2^J, les facteurs φ_α^I, φ_α^J étant les couples de vecteurs pour lesquels est stationnaire le $\cos^2$ (c'est-à-dire en terme statistique le coefficient de corrélation au carré).

On peut encore remarquer (ce qui est à peu près classique, sinon dit explicitement dans les cours...) que toute fonction ξ^{IJ} sur $I \times J$, (telle que d^{IJ}) définit à la fois une application linéaire de L_2^I dans L_2^J et une application linéaire de L_2^J dans L_2^I, (ces deux applications étant transposées l'une de l'autre), suivant les formules :

$$\forall\, \varphi^I \in L_2^I \;:\; \xi(\varphi^I) = \alpha^J \;;\; \text{où} :$$

$$\forall\, j \in J \;:\; \alpha^j = \Sigma\{\varphi^i f_i \xi^{ij} | i \in I\} ;$$

et de même :

$$\forall\, \psi^J \in L_2^J \;:\; \xi(\psi^J) = \beta^I \;;\; \text{où} :$$

$$\forall\, i \in I \;:\; \beta^i = \Sigma\{\xi^{ij} f_j \psi^j | j \in J\}.$$

Pour comprendre ces formules, il suffit de considérer le cas où ξ^{IJ} est un produit de deux fonctions d'une variable, μ^I, ν^J :

$$\forall\, i \in I, \forall\, j \in J \;:\; \xi^{ij} = \mu^i \nu^j ;$$

on a alors :

$$\xi(\varphi^I) = <\varphi^I, \mu^I> \nu^J \;;\; \xi(\psi^J) = <\psi^J, \nu^J> \mu^I.$$

L'intérêt de ces applications est d'abord que pour la fonction d^{IJ}, elles coïncident avec les applications usuelles associées aux transitions f_I^J et f_J^I, c'est-à-dire à la projection orthogonale de L_2^I sur L_2^J et de L_2^J sur L_2^I, au sein de L_2^{IJ} ; on a en effet :

$$\forall\, \varphi^I \in L_2^I \;:\; d(\varphi^I) = \varphi^I \circ f_I^J = \alpha^J ;$$

(i.e. : $\forall\, j \in J : \alpha^j = \Sigma_i\{\varphi^i f_i^j\} = \Sigma_i\{\varphi^i f_i d^{ij}\})$;

$$\forall\, \psi^J \in L_2^J \;:\; d(\psi^J) = \psi^J \circ f_J^I.$$

C'est pourquoi, pour une fonction quelconque ξ^{IJ} il est commode de noter :

$$\xi^I_J = \{\xi^i_j\} ; \ \xi^i_j = \xi^{ij} f_j ;$$

$$\xi^J_I = \{\xi^j_i\} ; \ \xi^j_i = \xi^{ij} f_i .$$

Et les applications $\xi(\varphi^I)$, $\xi(\psi^J)$ s'écrivent alors comme des transitions :

$$\xi(\varphi^I) = \alpha^J = \varphi^I \circ \xi^J_I ;$$
$$\xi(\psi^J) = \beta^I = \psi^J \circ \xi^I_J .$$

Mais ce qui nous importe ici est qu'en associant des applications linéaires à toute fonction sur I.J, on peut facilement écrire toutes les décompositions de fonctions et de tableaux subordonnés à des partitions sur I et J.

2 *Décompositions orthogonales associées à des partitions*

Comme au § 1 on considère une correspondance f_{IJ}, I et J étant munis des lois f_I et f_J. De plus, on suppose données une partition X de I et une partition Y de J ; la classe d'un élément, i ou j, étant notée :

$$x(i) \in X \ ; \ y(j) \in Y.$$

Considérons d'abord l'un des ensembles, soit X.

On a sur X une loi f_X définie par :

$$\forall x \in X : f_x = \Sigma\{f_i | i \in x\}$$

(en effet, x est une partie de I). L'espace L^X_2 des fonctions de carré sommable sur X, avec la norme associée à f_X, s'identifie au sous-espace de L^I_2 formé des fonctions qui sont constantes sur chaque classe x:

$$(\varphi^I \in L^X_2) \Leftrightarrow (\forall\ i, i' \in I : x(i) = x(i') \Rightarrow \varphi^i = \varphi^{i'})$$

Dans L^I_2, le supplémentaire orthogonal de L^X_2 n'est autre que le sous-espace, noté $H^{I/X}$ des fonctions φ ayant moyenne nulle sur chaque x :

$$(\varphi^I \in H^{I/X}) \Leftrightarrow (\forall\ x \in X : \Sigma\{\varphi^i f_i | i \in x\} = 0)$$

La projection φ^X sur L^X_2 d'une fonction quelconque φ^I sur I, s'obtient en faisant la moyenne de φ^I sur chacune des classes x : on a:

$$\forall\ x \in X : \varphi^x = \Sigma\{\varphi^i f_i/f_x | i \in x\}.$$

Quant à la projection $\varphi^{I/X}$ de φ^I sur $H^{I/X}$, le plus simple est de la calculer comme la différence $\varphi^I - \varphi^X$; et il est clair que la fonction ainsi obtenue a bien moyenne nulle sur chaque x, puisque précisément on soustrait de φ^i la moyenne $\varphi^{x(i)}$ calculée sur la classe x(i) de i.

Le passage de φ^I à φ^X peut s'exprimer par un calcul de transition. Si on pose :

$\forall\, i, i' \in I$: $r_i^{i'} =$ si $(x(i) = x(i'))$ alors $f_i/f_{x(i)}$ sinon zéro ; on a $\varphi^X = \varphi^I \circ r_I^I$. Par différence, $\varphi^{I/X}$ se calcule aussi par une transition. Si on pose $t_I^I = \delta_I^I - r_I^I$, on a $\varphi^{I/X} = \varphi^I \circ t_I^I$;

Les mêmes calculs avec des notations semblables se font pour J et Y.

Considérons maintenant le produit tensoriel euclidien $L_2^I \otimes L_2^J$; c'est-à-dire les fonctions de deux variables i et j avec la métrique $f_I f_J$ (et *non* f_{IJ}). Ce produit tensoriel se décompose en somme directe orthogonale de 4 sous-espaces :

$$\begin{aligned} L_2^I \otimes L_2^J &= (L_2^X \oplus H^{I/X}) \otimes (L_2^Y \oplus H^{J/Y}) \\ &= (L_2^X \otimes L_2^Y) \oplus (H^{I/X} \otimes L_2^Y) \oplus (L_2^X \otimes H^{J/Y}) \oplus (H^{I/X} \otimes H^{J/Y}); \end{aligned}$$

Pour comprendre cette décomposition, il suffit de la considérer pour une fonction de i et de j qui est le produit de deux fonctions d'une seule variable ; soit (avec les notations déjà utilisées au § 1):

$$\xi^{IJ} = \mu^I \otimes \nu^J \; ; \; \xi^{ij} = \mu^i \nu^j$$

on a alors :

$$\xi^{IJ} = (\mu^X \otimes \nu^Y) + (\mu^{I/X} \otimes \nu^Y) + (\mu^X \otimes \nu^{J/Y}) + (\mu^{I/X} \otimes \nu^{J/Y}),$$

ce qu'on peut écrire :

$$\xi^{IJ} = \xi^{XY} + \xi^{I/X;Y} + \xi^{X;J/Y} + \xi^{I/X;J/Y} .$$

Dans le cas d'une fonction ξ^{IJ} quelconque la décomposition se calcule en faisant des moyennes partielles, par rapport à l'une ou l'autre des deux variables i et j (ou les deux) sur les classes de la partition correspondante. On note

$$\begin{aligned} \xi^{iy} &= \Sigma\{\xi^{ij}(f_j/f_y) \mid j \in y\} \; ; \\ \xi^{xj} &= \Sigma\{\xi^{ij}(f_i/f_x) \mid i \in x\} \\ \xi^{xy} &= \Sigma\{\xi^{ij}(f_i/f_x)(f_j/f_y) \mid i \in x \; ; \; j \in y\} \; ; \end{aligned}$$

D'où pour ξ^{ij} la décomposition en quatre termes :

$$\begin{aligned} \xi^{ij} = \xi^{x(i)y(j)} &+ (\xi^{i\,y(j)} - \xi^{x(i)y(j)}) + (\xi^{x(i)j} - \xi^{x(i)y(j)}) \\ &+ (\xi^{ij} - \xi^{i\,y(j)} - \xi^{x(i)j} + \xi^{x(i)\,y(j)}) \; ; \end{aligned}$$

on vérifie par exemple que le deuxième terme appartient à $H^{I/X} \otimes L_2^Y$, car en tant que fonction de i et j, il a moyenne nulle sur chaque classe x et est constant sur chaque classe y.

Pour les applications linéaires ξ_J^I et ξ_I^J on a les décompositions correspondantes en 4 termes ; par exemple pour ξ_J^I :

$$\begin{aligned} \xi_J^I &= (r_J^J + t_J^J) \circ \xi_J^I \circ (r_I^I + t_I^I) \\ &= (r_J^J \circ \xi_J^I \circ r_I^I) + (r_J^J \circ \xi_J^I \circ t_I^I) + (t_J^J \circ \xi_J^I \circ r_I^I) \\ &\quad + (t_J^J \circ \xi_J^I \circ t_I^I). \end{aligned}$$

(formule qui résulte de ce que $r^J_J + t^J_J = \delta^J_J$, l'identité ; et de même pour I).

3 *Décomposition orthogonale et analyse factorielle*

Revenons au problème considéré initialement : caractériser les correspondances pour lesquelles les facteurs vont par couples, soit constant sur les classes des partitions X et Y de I et J, soit de moyenne nulle sur ces classes. Avec les notations du § 2 cela implique, compte tenu de la formule de reconstitution rappelée au § 1 :

$$d^{IJ} = d^{XY} + d^{I/X;J/Y} \quad ; \quad \text{ou encore}$$

$$d^{I/X;Y} = 0 \; ; \; d^{X;J/Y} = 0 \; ; \; \text{c'est-à-dire :}$$

$$\forall\; i \in I,\; j \in J \;:\; d^{i\,y(j)} = d^{x(i)y(j)} = d^{x(i)j};$$

formule qu'on peut récrire non plus en terme de densité, mais en terme de masses :

$$\forall\; i \in I,\; j \in J \;:\; f_{i;y(j)}\, f_{x(i)} f_j = f_{x(i);y(j)}\, f_i\, f_j = f_{x(i);j}\, f_i\, f_{y(j)}.$$

Réciproquement si $d^{I/X;Y} = 0$ et $d^{X;J/Y} = 0$, les facteurs ont la forme particulière expliquée ci-dessus : car la décomposition canonique de d^{IJ}, s'obtient en faisant séparément la décomposition canonique des deux termes :

$$d^{XY} \in L^X_2 \otimes L^Y_2 \quad ; \quad d^{I/X;J/Y} \in H^{I/X} \otimes H^{J/Y} \; ;$$

et ces deux décompositions canoniques donnent des facteurs des deux types voulus.

Les facteurs du premier type peuvent être calculés en analysant la correspondance sur XY obtenue en cumulant la correspondance f_{IJ} sur chacun des blocs x y ; les facteurs du deuxième type peuvent aussi se calculer séparément (si on le juge bon) en analysant la correspondance sur $I \times J$ dont la densité est $(1+d^{I/X;J/Y})$ (correspondance dont le tableau peut comporter des nombres négatifs ; mais qui a bien pour marges f_I et f_J).

Les facteurs des deux types peuvent encore se définir dans l'espace L^{IJ}_2 muni de la métrique associée à f_{IJ} : pour les facteurs du 1-er type, on considère la figure formée par les deux sous-espaces L^X_2 et L^Y_2 :

$$L^X_2 \subset L^I_2 \quad ; \quad L^Y_2 \subset L^J_2 \; .$$

Les couples de facteurs (φ^I , φ^J) du 1-er type peuvent être caractérisés comme des couples (φ^X, φ^Y) réalisant la stationnarité de l'angle : car il est clair que s'il y a stationnarité si on se place dans L^I_2 et L^J_2, il y a *a fortiori* stationnarité si on restreint la variation des vecteurs à des sous-espaces L^X_2 et L^Y_2.

De même pour les facteurs du 2-ème type, on pourra étudier la figure formée par les deux sous-espaces $H^{I/X}$ et $H^{J/Y}$.

Dans le cas général où d est quelconque relativement aux deux partitions X et Y, il n'y a pas de facteurs des types 1 et 2 ; ou du moins, s'il y en a, ce ne sont pas tous les facteurs. Mais on peut avec les deux termes d^{XY} et $d^{I/X;J/Y}$ construire une correspondance de densité d' aussi voisine que possible de la correspondance donnée, et n'admettant que des facteurs des types 1 et 2 : il suffit de poser :

$$d'^{IJ} = d^{XY} + d^{I/X;J/Y} \quad ; \text{ i.e. (cf. } supra \text{ § 2) :}$$

$$d'^{ij} = d^{x(i)\,y(j)} + (d^{ij} - d^{i\,y(j)} - d^{x(i)\,j} + d^{x(i)\,y(j)})$$

$$= d^{ij} - d^{i\,y(j)} - d^{x(i)\,j} + 2d^{x(i)y(j)}.$$

L'analyse de cette correspondance fournit des facteurs qu'on peut interpréter dans L^{IJ}_2 par les figures formées par deux couples de sous-espaces :

$$(L^X_2\,,\ L^Y_2)\,,\ \text{et :}\ (H^{I/X}\ ,\ H^{J/Y})\ ;$$

car les angles formés par les vecteurs de l'un ou de l'autre couple sont les mêmes que L^{IJ}_2 soit muni de la métrique associée à f_{IJ} (donné: de densité d^{IJ}) ou à f'_{IJ} (construit : de densité d'^{IJ}) : en effet les termes en $d^{I/X;Y}$ et $d^{X;J/Y}$ ne contribuent pas au produit scalaire $(\varphi^X\ ,\psi^Y)$ ou $(\varphi^{I/X}\ ,\ \psi^{J/Y})$.

4 Remarque

La méthode du § 3 permet de résoudre un problème analogue, que nous exposons brièvement pour son intérêt géométrique, sans envisager d'application pratique.

Les ensembles I et J étant comme précédemment munis de partitions X et Y, on cherche la condition pour que tout couple de facteurs non triviaux soit formé ou bien d'un facteur sur I constant sur toute classe x, associé à un facteur sur J centré sur toute classe y ; ou bien d'un facteur sur I centré sur tout x, associé à un facteur sur J constant sur tout y.

En terme géométrique cela s'écrit globalement :

$$d^{IJ} = 1 + d^{X;J/Y} + d^{I/X;Y} \quad ; \quad \text{ou encore :}$$

$$d^{XY} = 1\ ;\ d^{I/X;J/Y} = 0$$

Ce qu'on peut récrire en termes de valeurs des densités par deux conditions dont la 1-ère résulte en fait de la 2-ème ;

$$\forall\, x \in X,\ \forall\, y \in Y\ :\ d^{xy} = 1\ ;$$

$$\forall\, i \in I,\ \forall\, j \in J\ :\ 1 + d^{ij} = d^{i\,y(j)} + d^{x(i)\,j}\ ;$$

ou encore en terme de masses :

$$\forall\, i \in I,\ \forall\, j \in J\ :\ f_{ij} = -f_i\, f_j + f_{iy(j)}\,(f_j/f_{y(j)}) + f_{x(i)\,j}\,(f_i/f_{x(i)}).$$

On peut dire, en bref, que ce cas est complémentaire de celui traité au § 3.