

Sur la comparaison et la visualisation des partitions floues

François Bavaud, Université de Lausanne

in *Méthodes et perspectives en classification*, pp. 71-74, Y.Dodge et G.Melfi (éditeurs), Presses Académiques, Neuchâtel(2003)

Résumé

Une partition floue assigne à chaque objet (parmi n objets) une distribution sur a catégories. Par des méthodes d’algèbre linéaire élémentaire, on définit et étudie des propriétés telles que l’emboîtement des partitions, leur itération, ou leur stabilité en relation avec une autre partition. L’introduction de mesures de similarités “naturelles” entre objets, non pondérées (R) ou pondérées (T, P) permet de définir des distances euclidiennes entre objets, mais aussi entre partitions, lesquelles peuvent alors être représentées comme des points dans un espace factoriel de basse dimensionalité par MDS classique. Les versions pondérées T et P diffèrent pour les partitions floues, et engendrent diverses constructions formelles n’ayant pas d’équivalent au niveau des partitions ordinaires (=déterministes). Ce travail suggère une certaine vue de l’analyse multivariée de variables catégorielles floues, autrement dit de l’*analyse des correspondances floues multiples*¹.

1 Matrices d’appartenance

Définition 1 Une partition *floue* \mathcal{A} de n objets en a groupes est définie par une matrice ($n \times a$) d’appartenance $Z^{\mathcal{A}} = (z_{ij})$ telle que $z_{ij} \geq 0$, $\sum_{j=1}^a z_{ij} = 1$ ($\forall i = 1, \dots, n$) et $n_j^{\mathcal{A}} := \sum_{i=1}^n z_{ij}^{\mathcal{A}} > 0$ ($\forall j = 1, \dots, a$). La notion d’appartenance en jeu peut s’interpréter comme $z_{ij} =$ “probabilité que l’objet i appartienne au groupe j ”.

La partition est *déterministe* si $z_{ij} = 1$ ou $z_{ij} = 0$ pour tous i, j , i.e. si $z_{ij}^2 = z_{ij}$. La partition est *pleine* si $\text{rang}(Z) = a$, et *défective* si $\text{rang}(Z) < a$.

Une partition déterministe est pleine. En général, Z est formé de $c(\mathcal{A}) \leq a$ sous-blocs irréductibles indicés par $J = 1, \dots, c(\mathcal{A})$. Chaque composante J est constituée (en lignes) de groupes j tels que $z_{ij'} = 0$ si $z_{ij} > 0$, pour $j \in J$ et $j' \notin J$; de même, chaque composante J est constituée (en colonnes) d’objets i tels que $z_{ij} = 0$ si $j \notin J$. On définit les matrices ($a \times a$)

$$B := Z'Z \quad \text{i.e.} \quad b_{jj'} := \sum_i z_{ij} z_{ij'} \quad N := \text{diag}(\mathbf{1}'Z) \quad \text{i.e.} \quad n_{jj'} := \delta_{jj'} n_j \quad (1)$$

¹Le travail a bénéficié de discussions stimulantes avec M. Rajman dans le cadre du projet UNIL-EPFL “Clavis” (2001).

On a $b_{jj'} = 0$ si j et j' appartiennent à des composantes J et J' différentes. Aussi

$$\begin{aligned} c(\mathcal{A}) = m &\Leftrightarrow \mathcal{A} \text{ est déterministe} \Leftrightarrow B = N \\ \text{rang}(Z) = m &\Leftrightarrow \mathcal{A} \text{ est pleine} \Leftrightarrow B^{-1} \text{ existe} \end{aligned}$$

Lorsque \mathcal{A} n'est pas déterministe, $G := N^{-1}B$ diffère de l'identité et génère des *partitions itérées* d'ordre r définies comme $Z^{(r)} := Z G^{r-1}$, avec limite $z_{ij}^{(\infty)} = n_j I(i \in J(j))/n_{J(j)}$ (où $I(E)$ est la fonction caractéristique de l'événement E). La partition $Z^{(\infty)}$ est pleine ssi \mathcal{A} est déterministe.

Exemple 1 : on considère la partition floue \mathcal{A} de $n = 5$ objets dans $a = 4$ groupes avec

$$Z^{\mathcal{A}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.2 & 0.8 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix} \quad N = \begin{pmatrix} 1.2 & 0 & 0 & 0 \\ 0 & 1.8 & 0 & 0 \\ 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 1.2 \end{pmatrix} \quad B = \begin{pmatrix} 1.04 & 0.16 & 0 & 0 \\ 0.16 & 1.64 & 0 & 0 \\ 0 & 0 & 0.4 & 0.4 \\ 0 & 0 & 0.4 & 0.8 \end{pmatrix}$$

2 Similarités et distances entre objets

Soit $S = (s_{ii'})$ une matrice générale de similarités entre objets, définie non négative, et telle que $s_{ii'} \geq 0$, $s_{ii'} = s_{i'i}$ et $s_{ii'} \leq \sqrt{s_{ii} s_{i'i'}}$. Les matrices $R := ZZ'$, $T := ZN^{-1}Z'$ et (pour une partition pleine) $P := ZB^{-1}Z'$ constituent trois candidats naturels pour S :

$$r_{ii'} := \sum_{j=1}^a z_{ij} z_{i'j} \quad t_{ii'} := \sum_{j=1}^a \frac{z_{ij} z_{i'j}}{n_j} \quad p_{ii'} := \sum_{j,j'=1}^a z_{ij} b_{jj'}^{(-1)} z_{i'j'} \quad (2)$$

$R = (r_{ii'})$ définit, pour une partition déterministe, la relation d'équivalence “ i et i' appartiennent au même groupe”. $T = (t_{ii'})$ est une matrice de transition markovienne de distribution stationnaire $\pi_i = 1/n$, vérifiant $\sum_{i'} t_{ii'} = 1$; $P = (p_{ii'})$ (pour lequel $p_{ii'} \geq 0$ peut être violé, $|p_{ii'}| \leq \sqrt{p_{ii} p_{i'i'}}$ restant valide) est une matrice de projection vérifiant $P^2 = P$. On a $T = P$ ssi la partition est déterministe. Aussi, les itérés de T et P construits à partir de $Z^{(r)} := Z G^{r-1}$ vérifient $T^{(r)} = T^{2r-1}$ et $P^{(r)} = P$.

Par le théorème de Schoenberg² (MDS classique), les quantités $D_{ii'}^S := (d_{jj'}^S)^2 = s_{ii} + s_{i'i'} - 2s_{ii'}$ constituent des (carrés de) distances euclidiennes, pour lesquelles une configuration exacte de dimension $\leq n - 1$ peut être reconstituée par MDS classique (diagonalisation de S). Explicitement

$$D_{ii'}^R = \sum_j (z_{ij} - z_{i'j})^2 \quad D_{ii'}^T = \sum_j \frac{(z_{ij} - z_{i'j})^2}{n_j} \quad D_{ii'}^P = \sum_{j,j'} (z_{ij} - z_{i'j}) b_{jj'}^{(-1)} (z_{ij'} - z_{i'j'}) \quad (3)$$

²SCHOENBERG, I.J. (1935) : “Remarks to Maurice Fréchet's article “Sur la définition axiomatique d'une classe d'espaces vectoriels distancés applicables vectoriellement sur l'espace de Hilbert”” *Annals of Mathematics*, 36, 724-732

Ainsi, pour une partition déterministe, on a

$$\begin{aligned}
r_{ii'} = 1 \quad t_{ii'} = p_{ii'} = \frac{1}{n_j} & \quad D_{ii'}^R = D_{ii'}^T = D_{ii'}^P = 0 & \quad \text{pour } i, i' \in j \\
r_{ii'} = t_{ii'} = p_{ii'} = 0 & \quad D_{ii'}^R = 2 \quad D_{ii'}^T = D_{ii'}^P = \frac{1}{n_j} + \frac{1}{n_{j'}} & \quad \text{pour } i \in j, i' \in j' \text{ avec } j \neq j'
\end{aligned}$$

Exemple 1, suite : les matrices de similarité et distances *entre objets* associées sont

$$R = \begin{pmatrix} 1 & .2 & 0 & 0 & 0 \\ .2 & .68 & .8 & 0 & 0 \\ 0 & .8 & 1 & 0 & 0 \\ 0 & 0 & 0 & .52 & .44 \\ 0 & 0 & 0 & .44 & .68 \end{pmatrix} \quad T = \begin{pmatrix} .83 & .17 & 0 & 0 & 0 \\ .17 & .39 & .44 & 0 & 0 \\ 0 & .44 & .56 & 0 & 0 \\ 0 & 0 & 0 & .58 & .42 \\ 0 & 0 & 0 & .42 & .58 \end{pmatrix} \quad P = \begin{pmatrix} .98 & .12 & -.10 & 0 & 0 \\ .12 & .40 & .48 & 0 & 0 \\ -.10 & .48 & .62 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$D^R = \begin{pmatrix} 0 & 1.28 & 2 & 1.52 & 1.68 \\ 1.28 & 0 & .08 & 1.2 & 1.36 \\ 2 & .08 & 0 & 1.52 & 1.68 \\ 1.52 & 1.2 & 1.52 & 0 & .32 \\ 1.68 & 1.36 & 1.68 & .32 & 0 \end{pmatrix} \quad D^T = \begin{pmatrix} 0 & .89 & 1.39 & 1.42 & 1.42 \\ .89 & 0 & .06 & .97 & .97 \\ 1.39 & .06 & 0 & 1.14 & 1.14 \\ 1.42 & .97 & 1.14 & 0 & .33 \\ 1.42 & .97 & 1.14 & .33 & 0 \end{pmatrix} \quad D^P = \begin{pmatrix} 0 & 1.14 & 1.79 & 1.98 & 1.98 \\ 1.14 & 0 & .07 & 1.40 & 1.40 \\ 1.79 & .07 & 0 & 1.62 & 1.62 \\ 1.98 & 1.40 & 1.62 & 0 & 2 \\ 1.98 & 1.40 & 1.62 & 2 & 0 \end{pmatrix}$$

3 Partitions emboîtées

Définition 2 La partition \mathcal{B} (définie par la matrice $(n \times b)$ d'appartenance $Z^{\mathcal{B}}$) est *plus grossière* que la partition \mathcal{A} (définie par la matrice $(n \times a)$ d'appartenance $Z^{\mathcal{A}}$), i.e. \mathcal{A} est *plus fine* que \mathcal{B} , noté $\mathcal{B} \leq \mathcal{A}$, si $Z^{\mathcal{B}} = Z^{\mathcal{A}} W^{\mathcal{A}\mathcal{B}}$ où $W^{\mathcal{A}\mathcal{B}} = (w_{jk}^{\mathcal{A}\mathcal{B}})$ est une matrice $(a \times b)$ avec $w_{jk}^{\mathcal{A}\mathcal{B}} \geq 0$ et $\sum_{k=1}^b w_{jk}^{\mathcal{A}\mathcal{B}} = 1$.

La relation “ $\mathcal{B} \leq \mathcal{A}$ ” est un ordre partiel, d'élément minimal \mathcal{O} (avec $z_{ij}^{\mathcal{O}} \equiv 1$: partition à un groupe) et d'élément maximal \mathcal{N} (avec $z_{ij}^{\mathcal{N}} = \delta_{ij}$: partition à n groupes). Aussi, si $\mathcal{B} \leq \mathcal{A}$ (partitions pleines), alors $P^{\mathcal{A}} P^{\mathcal{B}} = P^{\mathcal{B}} P^{\mathcal{A}} = P^{\mathcal{B}}$. La séquence $\mathcal{A}^{(r)}$ de partitions associées à $Z^{(r)}$ est décroissante : $\mathcal{A}^{(r+1)} \leq \mathcal{A}^{(r)}$.

4 Distances entre partitions

$S^{\mathcal{A}} = (s_{ii'}^{\mathcal{A}})$ et $S^{\mathcal{B}} = (s_{ii'}^{\mathcal{B}})$ étant les similarités associées aux partitions \mathcal{A} et \mathcal{B} , une distance euclidienne (quadratique) entre ces dernières peut être définie comme

$$D_{\mathcal{A},\mathcal{B}}^S := \sum_{ii'} (s_{ii'}^{\mathcal{A}} - s_{ii'}^{\mathcal{B}})^2 = \text{Tr}(S^{\mathcal{A}} - S^{\mathcal{B}})^2 = \text{Tr}((S^{\mathcal{A}})^2) + \text{Tr}((S^{\mathcal{B}})^2) - 2\text{Tr}(S^{\mathcal{A}} S^{\mathcal{B}}) \quad (4)$$

Par construction, le MDS classique appliqué à un ensemble de partitions munies de la distance $D_{\mathcal{A},\mathcal{B}}^S$ permet de visualiser exactement la configuration de façon euclidienne, chaque partition étant représentée par un point (voir figure 1).

Exemple 2 Soient $n = 5$ objets. On définit \mathcal{A} comme dans l'exemple 1 ; \mathcal{B} comme la partition (123; 45) ; \mathcal{C} comme (12; 345) ; $\mathcal{D} \equiv \mathcal{N}$ comme (1; 2; 3; 4; 5) ; $\mathcal{E} \equiv \mathcal{O}$ comme (12345), et $\mathcal{F} \equiv \mathcal{A}^{(\infty)}$ comme la partition itérée limite :

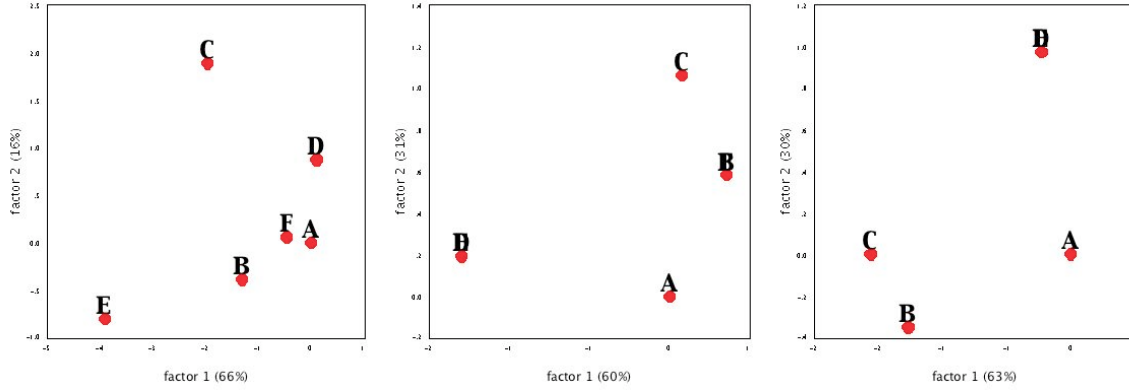


FIG. 1 – Visualisation euclidienne des distances entre partitions $D_{\mathcal{A},\mathcal{B}}^S$, pour les 6 partitions de l'exemple (2), dans les versions $S = R$ (gauche), $S = T$ (milieu) and $S = P$ (droite). \mathcal{D} and \mathcal{E} sont confondues dans les versions T et P ; \mathcal{B} and \mathcal{F} sont confondues dans la version T . Enfin, \mathcal{F} est déficiente et donc non représentable dans la version P .

$$Z^{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} Z^{\mathcal{C}} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} Z^{\mathcal{D}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} Z^{\mathcal{E}} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} Z^{\mathcal{F}} = \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.4 & 0.6 & 0 & 0 \\ 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0.4 & 0.6 \\ 0 & 0 & 0.4 & 0.6 \end{pmatrix}$$

qui définissent les dissimilarités *entre partitions* suivantes (dans l'ordre alphabétique) :

$$D^T = \begin{pmatrix} 0 & 0.63 & 1.37 & 1.74 & 1.74 & 0.63 \\ 0.63 & 0 & 1.11 & 3 & 3 & 0 \\ 1.37 & 1.11 & 0 & 3 & 3 & 1.11 \\ 1.74 & 3 & 3 & 0 & 0 & 3 \\ 1.74 & 3 & 3 & 0 & 0 & 3 \\ 0.63 & 0 & 1.11 & 3 & 3 & 0 \end{pmatrix} \quad D^P = \begin{pmatrix} 0 & 2 & 2.63 & 1 & 1 \\ 2 & 0 & 1.11 & 3 & 3 \\ 2.63 & 1.11 & 0 & 3 & 3 \\ 1 & 3 & 3 & 0 & 0 \\ 1 & 3 & 3 & 0 & 0 \end{pmatrix}$$

En général, pour des partitions floues et pleines, $\mathcal{B} \leq \mathcal{A}$ entraîne $D_{\mathcal{A},\mathcal{B}}^P = \text{Tr}(P^{\mathcal{A}}) - \text{Tr}(P^{\mathcal{B}}) = a - b$, et donc $D_{\mathcal{A},\mathcal{C}}^P = D_{\mathcal{A},\mathcal{B}}^P + D_{\mathcal{B},\mathcal{C}}^P$ pour $\mathcal{C} \leq \mathcal{B} \leq \mathcal{A}$ (ou $\mathcal{A} \leq \mathcal{B} \leq \mathcal{C}$).

Dans le cas de partitions \mathcal{A} et \mathcal{B} déterministes à a et b groupes, on a $D_{\mathcal{A},\mathcal{B}}^R = N_{\mathcal{A},\mathcal{A}} + N_{\mathcal{B},\mathcal{B}} - 2N_{\mathcal{A},\mathcal{B}}$ et $D_{\mathcal{A},\mathcal{B}}^T = D_{\mathcal{A},\mathcal{B}}^P = (a-1) + (b-1) - \frac{2}{n} \chi_{\mathcal{A},\mathcal{B}}^2$, où $N_{\mathcal{A},\mathcal{B}}$ est le nombre de paires d'objets classés dans le même groupe j de \mathcal{A} et k de \mathcal{B} , et $\chi_{\mathcal{A},\mathcal{B}}^2$ est le chi2 usuel associé à la table de contingence n_{jk} (=nombre d'objets dans le groupe j de \mathcal{A} et k de \mathcal{B}). De plus, $D_{\mathcal{A},\mathcal{B}}^R = \sum_{j=1}^a \rho_j^{\mathcal{B}}$ et $D_{\mathcal{A},\mathcal{B}}^T = D_{\mathcal{A},\mathcal{B}}^P = \sum_{j=1}^a \tau_j^{\mathcal{B}}$, où $\rho_j^{\mathcal{B}} := n_j^2 - 2 \sum_{i,i' \in j} r_{ii'}^{\mathcal{B}} + \sum_{i \in j; i'} r_{ii'}^{\mathcal{B}} \geq 0$ et $\tau_j^{\mathcal{B}} := 1 - \frac{2}{n_j} \sum_{i,i' \in j} p_{ii'}^{\mathcal{B}} + \sum_{i \in j} p_{ii}^{\mathcal{B}} \geq 0$. Les quantités $\rho_j^{\mathcal{B}}$ et $\tau_j^{\mathcal{B}}$ s'interprètent comme mesures d'instabilité du groupe j de \mathcal{A} relativement à la partition \mathcal{B} . En particulier, $\rho_j^{\mathcal{N}} = n_j(n_j - 1)$ et $\tau_j^{\mathcal{N}} = n_j - 1$, qui souligne l'instabilité des grands groupes face à \mathcal{N} ; aussi, $\rho_j^{\mathcal{O}} = (n - n_j)n_j$ (groupes moyens instables face à \mathcal{O}) et $\tau_j^{\mathcal{O}} = \frac{n - n_j}{n}$ (petits groupes instables face à \mathcal{O}).