

## Chapitre 7

# Comparaison des groupes de variables et modèle Indscal

Les chapitres précédents présentent des outils d'analyse détaillée des relations entre groupes de variables, tant du point de vue des variables (corrélations inter-groupes) que de celui des individus (comparaison des nuages partiels). Mais, dès lors que les données sont volumineuses (de nombreux groupes comportant chacun de nombreuses variables définies sur un grand nombre d'individus), il est nécessaire, au moins dans un premier temps, de disposer d'outils plus globaux. Ceci afin de répondre aux questions du type décrit ci-après.

- Deux groupes de variables étant donnés, peut-on considérer qu'ils sont liés ? Autrement dit : dans l'ensemble, les deux nuages partiels qui leur sont associés se ressemblent-ils ?
- A l'image des représentations des individus et des variables, peut-on disposer d'un graphique dans lequel chaque groupe est représenté par un point, les proximités entre points traduisant une liaison/ressemblance entre groupes ?

### 7.1 Nuage $N_J$ des groupes de variables

Les données brutes associées à un groupe de variables  $j$  constituent le tableau  $X_j$ . Dans le cas général, il n'y a pas de correspondances entre les colonnes des différents  $X_j$  et la comparaison directe des tableaux  $X_j$  entre eux n'est pas possible.

L'idée de comparer deux groupes de variables à travers leurs nuages partiels (point de vue du chapitre 5) suggère de représenter un groupe par la matrice des distances inter-individuelles. D'un groupe à l'autre, ces matrices ont les mêmes dimensions et leurs cases se correspondent deux à deux : on peut donc les comparer directement. Selon un autre point de vue, il est légitime de représenter un nuage d'individus par la matrice des produits scalaires (inter-individus ; notée  $XX'$  en 1.5.1) qui lui est associée : en effet, la diagonalisation de cette matrice permet de reconstituer