

Analyse géométrique des données pour l'affinement de la connaissance : cas des données EPGY (Education Program for Gifted Students, Stanford University)

Brigitte Le ROUX

Université René Descartes, 45 Rue des Saints-Pères, 75270 Paris cedex

lerb@math-info.univ-paris5.fr,

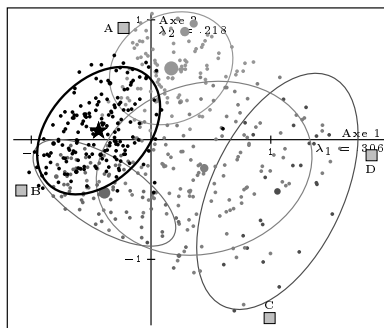
<http://www.math-info.univ-paris5.fr/~lerb>

Nous nous proposons de montrer sur un exemple comment l'utilisation conjointe des méthodes géométriques d'analyse des données, à savoir l'analyse des correspondances multiples (ACM) et la classification ascendante hiérarchique (CAH), permet de mettre en évidence des "patterns" qui sont potentiellement utiles et, en tout cas, explicables.

L'exemple concerne les enfants Surdoués qui suivent le programme EPGY d'enseignement du 3ème niveau de Mathématiques. Le but principal était de trouver comment se structurent les différences individuelles entre les élèves surdoués et de mettre en évidence des groupes d'enfants surdoués présentant des patterns d'attitudes bien différenciés, certains méritant d'être encouragés.

L'analyse géométrique des données, effectuée en collaboration avec le Pr P. Supes (Stanford University), a montré que pour chaque indicateur de performance (taux d'erreurs, temps de latence et nombre d'exercices pour maîtriser une notion), il y a *homogénéité des matières* (Entiers, Fractions, Géométrie, Logique et Mesurage), à l'exception de la Géométrie, et que les différences individuelles s'articulent autour de *deux échelles* : l'une est celle des taux d'erreurs et du nombre d'exercices, l'autre est celle des temps de latence.

Au partir de la représentation des 6 classes de la CAH dans le plan des axes 1 et 2 (ACM) sous forme de sous-nuages avec leurs ellipses de concentration, on a construit une partition en 5 classes qui met en évidence une classe (points noirs sur la figure) particulièrement intéressante du point de vue de l'éducation : c'est une classe de bonnes performances avec des taux d'erreurs faibles et des temps de latence relativement longs. Cette classe constitue un profil peu encouragé par les standards habituels d'évaluation des connaissances dans l'éducation aux USA.



Références :

Le Roux B. et Rouanet H. (2004), Geometric Data Analysis of Individual Differences, <http://epgy.stanford.edu/research/GeometricDataAnalysis.pdf>

Le Roux B., Rouanet H. (2004), *Geometric Data Analysis : From Correspondence Analysis to Structured Data Analysis*, Dordrecht, Kluwer Academic Press.