NOUVELLES TECHNIQUES NEURONALES EN ANALYSE DES DONNEES APPLICATIONS A LA CLASSIFICATION, A LA RECHERCHE DE TYPOLOGIE ET A LA PREVISION

Marie Cottrell

SAMOS, Université Paris 1
90, rue de Tolbiac
75634 PARIS Cedex 13
FRANCE
E-Mail : cottrell@univ-paris1.fr

Résumé

Nous passons en revue ici différentes techniques d'analyse de données qui ont comme point commun d'être toutes inspirées de l'algorithme de Kohonen. Nous montrons comment cet algorithme permet de représenter des données multidimensionnelles tant quantitatives que qualitatives, de les classer, d'étudier les classes obtenues, de croiser les variables qualitatives supplémentaires, de mettre en évidence les relations existant entre les différentes modalités.

Ce qui suit reprend en tout ou en partie différents travaux des membres du SAMOS, en particulier de S.Bayomog, E. De Bodt, B.Girard, S.Ibbou, P.Letremy, P.Rousset.

Mots-clés

Statistique, Kohonen, analyse de données, data mining.

1 Introduction

Pour étudier, résumer, représenter des données multidimensionnelles comprenant à la fois des variables quantitatives (à valeurs continues réelles) et qualitatives (discrètes, ordinales ou nominales), les praticiens ont à leur disposition de très nombreuses méthodes performantes, éprouvées et déjà implantées dans la plupart des logiciels statistiques.

L'analyse de données consiste à construire des représentations simplifiées de données brutes, pour mettre en évidence les relations, les dominantes, la structure interne du *nuage* des observations. On peut distinguer deux grands groupes de techniques classiques : les *méthodes factorielles* et les *méthodes de classification*.

Les *méthodes factorielles* sont essentiellement linéaires ; elles consistent à chercher des sous-espaces vectoriels, des changements de repères, permettant de réduire les dimensions tout en perdant le moins d'information possible. Les plus connues sont *l'Analyse en Composantes Principales* (ACP) qui permet de projeter des données quantitatives sur les axes les plus significatifs et *l'Analyse des Correspondances* qui permet d'analyser les relations entre les différentes modalités de variables qualitatives croisées, (AFC pour deux variables, ACM pour plus de deux variables).

Les *méthodes de classification* sont très nombreuses et diverses. Elles permettent de grouper et de ranger les observations. Les plus utilisées sont la *Classification Hiérarchique* où le nombre de classes n'est pas fixé a priori et la *Méthode des Centres Mobiles* (ou K-means, ou nuées dynamiques) où on cherche à regrouper les données en un certain nombre de classes.

Il n'est pas question de donner ici une bibliographie complète sur ces sujets. On peut citer deux ouvrages récents par exemple : on trouvera dans Lebart et al. (1995) ou Saporta (1990) une présentation de ces méthodes avec de nombreux exemples.

Plus récemment, depuis les années 80, de nouvelles méthodes sont apparues, connues sous le nom de *méthodes neuronales*. Elles proviennent de travaux pluridisciplinaires où se sont retrouvés des biologistes, des physiciens, des informaticiens, des théoriciens du signal, des cogniticiens, et plus récemment encore des mathématiciens et notamment des statisticiens.

Outre le fait qu'elles sont partiellement issues d'une inspiration biologique ou cognitive, elles ont rencontré rapidement un certain succès en particulier à cause de leur caractère de « boîte noire », d'outil à tout faire, ayant de très nombreux domaines d'applications. Une fois dépassés un certain excès d'enthousiasme et certaines difficultés de mise en oeuvre, les chercheurs et utilisateurs disposent maintenant d'un arsenal de techniques alternatives, non-linéaires en général et algorithmiques. En particulier, les statisticiens commencent à intégrer ces méthodes parmi l'ensemble de leurs outils. Voir par exemple à ce sujet l'ouvrage de Ripley (1996), ou les nouveaux modules neuronaux des grands logiciels de Statistique.

Le plus connu des modèles neuronaux reste sans conteste de modèle du *Perceptron Multicouches* (voir par exemple Rumelhart et al., 1986), dont nous ne parlerons pas. Nous voulons présenter ici des méthodes de classification, de représentation, d'analyse des relations, toutes construites à partir du célèbre algorithme de Kohonen. Voir par exemple Blayo et Demartines (1991, 1992), Varfis et Versino (1992), Kaski (1997), ou une synthèse dans Cottrell et Rousset (1997).

Notre papier est structuré comme suit. Dans le paragraphe 2, nous définissons rapidement l'algorithme de Kohonen qui sert de base à tout le reste. Nous montrons dans le paragraphe 3 comment l'utiliser pour réaliser une classification des observations, basée sur les variables quantitatives et admettant une représentation analogue à une ACP, et comment étudier les classes obtenues. Dans le paragraphe 4, nous présentons une méthode de visualisation du croisement avec des variables qualitatives. Dans les paragraphes 5 et 6, nous définissons deux méthodes analogues à l'Analyse des Correspondances double (KORRESP) et multiple (KACM). Le paragraphe 7 est une conclusion (provisoire bien entendue).

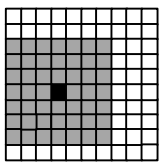
2 L'algorithme de Kohonen

Il s'agit d'un algorithme original de classement qui a été défini par Teuvo Kohonen, dans les années 80, à partir de motivations neuromimétiques (cf. Kohonen, 1984, 1995). Dans le contexte d'analyse des données qui nous intéresse ici, l'ensemble des entrées (ou inputs) est un ensemble fini et étant donnée une matrice de données, formée de N individus ou observations, décrits par un identificateur et p variables, l'algorithme regroupe les observations en classes, en *respectant la topologie de l'espace des observations*.

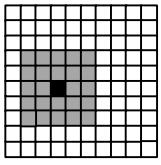
Cela veut dire qu'on définit a priori une notion de voisinage entre classes et que des observations voisines dans l'espace des variables (de dimension p) appartiennent (après classement) à la même classe ou à des classes voisines.

Les voisinages entre classes peuvent être choisis de manière variée, mais en général on suppose que les classes sont disposées sur une *grille* rectangulaire qui définit naturellement les voisins de chaque classe. On peut aussi considérer une topologie unidimensionnelle dite en *ficelle*, ou éventuellement un tore ou un cylindre.

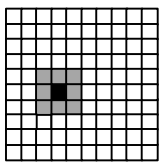
Par exemple sur une grille, on pourra prendre des voisinages de rayon 3 (49 voisins), de rayon 2 (25 voisins) ou de rayon 1 (9 voisins).



Voisinage de 49



Voisinage de 25



Voisinage de 9

Pour une ficelle, les mêmes rayons donnent 7, 5 et 3 voisins.


Voisinage de 7


Voisinage de 5


Voisinage de 3

Les classes situées sur les bords ont éventuellement moins de voisines.

Principe de l'algorithme de Kohonen

L'algorithme de classement est itératif. L'initialisation consiste à associer à chaque classe un *vecteur code* (ou *représentant*) de *p* dimensions choisi de manière aléatoire. Ensuite, à chaque étape, on choisit une observation au hasard, on la compare à tous les vecteurs codes, et on détermine la classe gagnante, c'est-à-dire celle dont le vecteur code est le plus proche au sens d'une distance donnée a priori. On rapproche alors de l'observation les codes de la classe gagnante et des classes voisines.

Cet algorithme est analogue à l'algorithme des nuées dynamiques ou des centres mobiles, mais dans ce dernier cas, il n'existe pas de notion de voisinage entre classes et on ne modifie à chaque étape que le *code* (ou *représentant*) de la classe gagnante.

Notations et définitions

On considère une table constituée de N observations $(x_1, x_2, ..., x_N)$, où chaque individu est décrit par p variables quantitatives. La matrice $(N \times p)$ ainsi formée est appelée *Matrice des données*. On rajoutera ensuite éventuellement des variables qualitatives.

On choisit une loi de probabilité $P = (P_1, P_2, ..., P_N)$ pour pondérer les N observations. Dans la plupart des cas, on pondérera les observations de façon uniforme, en prenant pour tout j de 1 à N, $P_j = 1/N$.

On se donne un réseau de Kohonen formés de n unités, rangées suivant une certaine topologie (en général une grille, ou bien une ficelle). Ces n unités sont munies d'un système de voisinages homogènes dans l'espace. Pour chaque unité i du réseau, on définit un voisinage de rayon r noté $V_r(i)$ et formé de l'ensemble des unités situées sur le réseau à une distance inférieure ou égale à r.

Chaque unité i est représentée dans l'espace R^p par un vecteur C_i (appelé vecteur poids par Kohonen) que nous désignerons par *vecteur code* ou par *représentant* de l'unité i (ou de la classe i). L'état du réseau au temps t est donné par $C(t) = (C_1(t), C_2(t), ..., C_n(t))$.

Pour un état donné C et une observation donnée x, l'unité (ou classe) *gagnante* $i_0(C,x)$ est l'unité dont le vecteur code $C_{i_0(C,x)}$ est le plus proche de l'observation x au sens d'une certaine distance. On a donc

$$i_0(C,x) = \underset{i}{Arg\,min} \left\| x - C_i \right\| .$$

Si l'on pose pour toute unité i, $G_i = \{x \in \{x_1, x_2, ..., x_N\} / i_0(C,x) = i\}$, on dira que G_i est la classe de numéro i et l'ensemble des classes $(G_1, G_2, ..., G_n)$ forme une partition de

l'ensemble des observations, appelée *Partition de Voronoï*. Chaque classe est représentée par le *vecteur code* correspondant. Chaque observation est représentée par le vecteur code le plus proche, exactement comme dans la méthode du plus proche voisin.

Alors pour un état C donné, le réseau définit une application Φ_C qui à chaque observation x associe l'unité gagnante correspondante, c'est-à-dire le numéro de sa classe. Après convergence de l'algorithme de Kohonen, l'application Φ_C *respecte la topologie* de l'espace des entrées, en ce sens que des observations voisines dans l'espace R^p se retrouvent associées à des unités voisines ou à la même unité.

L'algorithme de construction des vecteurs codes est défini de manière itérative comme suit :

- *Au temps 0*, les n vecteurs codes sont initialisés de manière aléatoire (on peut par exemple tirer au hasard n observations).

- *Au temps t*, l'état du réseau est $C(t)$, et on présente suivant la loi P une observation $x(t + 1)$, on a alors :

$$\begin{cases} i_0\big(C(t),\, x(t+1)\big) = & Arg\min\left\{\left\|x(t+1)-C_i(t)\right\|,\, 1\le i\le n\right\} & \\ C_i(t+1) = & C_i(t+1) - e(t)\left(C_i(t) - x(t+1)\right), & \forall i \in V_{r(t)}(i_0) \\ C_i(t+1) = & C_i(t+1) & \forall i \notin V_{r(t)}(i_0) \end{cases}$$

∗ où $e(t)$ est le *paramètre d'adaptation* ou *de gain,*
∗ et où $r(t)$ est le rayon des voisinages au temps t.

Les paramètres importants sont
- la dimension p de l'espace des entrées,
- la topologie du réseau (grille, ficelle, cylindre, tore, etc.),
- le paramètre d'adaptation, positif, compris entre 0 et 1, constant ou décroissant,
- le rayon des voisinages, en général décroissant,
- la loi de probabilité des observations P.

L'étude de la convergence de cet algorithme est incomplète et pose des problèmes mathématiques difficiles, voir par exemple Cottrell et Fort (1987) et Cottrell, Fort et Pagès (1995, 1997). Pour l'instant, l'essentiel des résultats correspond à la dimension 1 (topologie en ficelle et observations de dimension $p = 1$). Des résultats d'organisation et de convergence en loi sont disponibles quand le paramètre d'adaptation est constant. Pour les résultats de convergence presque sûre après réorganisation, il faut que la suite des paramètres $e(t)$ vérifient des conditions de Robbins-Monro, classiques pour les algorithmes stochastiques, qui s'énoncent :

$$\sum_t \varepsilon_t = +\infty \quad \text{et} \quad \sum_t \varepsilon_t^2 < +\infty .$$

En d'autres termes, le paramètre doit être « petit », mais « pas trop ».

Un des points qui rendent difficile l'étude théorique dans le cas général de l'algorithme de Kohonen, est qu'il n'existe pas de potentiel (ou énergie) associé lorsque les entrées sont munies d'une loi de probabilité *P* continue. Cf. Erwin et al. (1992). Mais au contraire, dans le cas qui nous intéresse ici, c'est-à-dire quand l'espace des entrées est fini, muni d'une probabilité discrète, Ritter et al. (1992) ont montré que l'algorithme de Kohonen pour un rayon constant *r* de voisinage, est alors un algorithme du gradient stochastique qui minimise le potentiel :

$$V(C_1, C_2, \ldots, C_N) = \sum_{i=1}^{n} \sum_{k \in V_r(i)} \sum_{x_j \in \Gamma_k} \left\| x_j - C_i \right\|$$

Ce potentiel généralise la variance intra-classes. Ici on calcule la somme des carrés des distances de chaque observation non seulement à son vecteur code mais aussi aux vecteurs codes des classes voisines.

L'étude mathématique de ce potentiel n'est pas simple non plus, car il n'est pas partout différentiable. En outre, il n'est pas possible d'éviter les minima locaux.

Un inconvénient de l'algorithme de base est que le nombre de classes doit être fixé a priori. Pour pallier cet inconvénient, on peut après la classification de Kohonen, pratiquer une classification de type hiérarchique sur les codes des classes de Kohonen, de manière à les regrouper en classes moins nombreuses. C'est ce que nous allons étudier dans le prochain paragraphe.

3 Classes et super-classes, étude des classes. KACP

Après convergence de l'algorithme, on a vu que les *N* observations sont classifiées en *n* classes selon la méthode du plus proche voisin, relativement à la distance choisie dans R^p.

On peut alors construire une représentation graphique selon la topologie du réseau. Dans chaque cellule du réseau (grille, ficelle, etc.), on dessine les observations associées ou on fait la liste de ces observations. Grâce à la propriété de conservation de la topologie, la représentation respecte les relations de voisinage. On obtient ainsi une carte de Kohonen où les grandes caractéristiques du nuage des données sont visibles après un petit nombre d'itérations (de l'ordre de 5 ou 6 *N*, en général). Cette carte est un instrument d'analyse des données qui fournit des informations analogues à celle que donne une Analyse en Composantes Principales (ACP). Bien sûr, il n'y a pas de projection à proprement parler, la carte est grossière. Mais elle est unique, ce qui évite de devoir combiner les différentes projections planes qu'on obtient en ACP. La continuité d'une classe à ses voisines permet de bien comprendre l'évolution le long de la grille, et est facile à interpréter. C'est pour ces raisons que l'ensemble des techniques de classification et représentation fournies par l'algorithme de Kohonen appliqué à une matrice de données est désigné par le sigle KACP.

On montrera pendant l'exposé des exemples de ces représentations (carte des pays repérés par des variables socio-économiques, courbes de consommation électrique, communes d'Ile-de-france, avec des variables immobilières, etc.). Cf. Blayo et Demartines (1991), Cottrell, Girard, Girard, Muller et Rousset (1995), Gaubert et al. (1995). Voir ci-dessous la représentation du contenu de 100 classes, pour un cylindre 10 par 10, et des données (transformées) de consommation demi-horaire. Cf. Cottrell et Rousset (1997). Dans ce cas, les observations qui sont des vecteurs de $p = 48$ points sont des courbes. Dans d'autres cas, on choisira des représentations en histogrammes, ou autres.

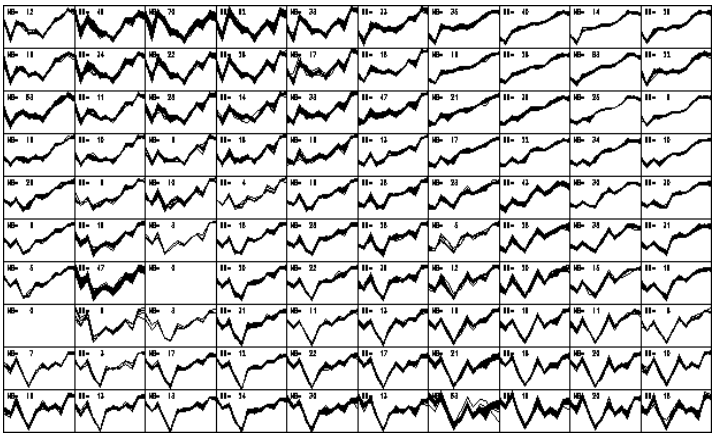


Fig.1. : *Contenu des classes*

Il est également intéressant de représenter sur le réseau les vecteurs codes, comme on le voit à la figure 2.

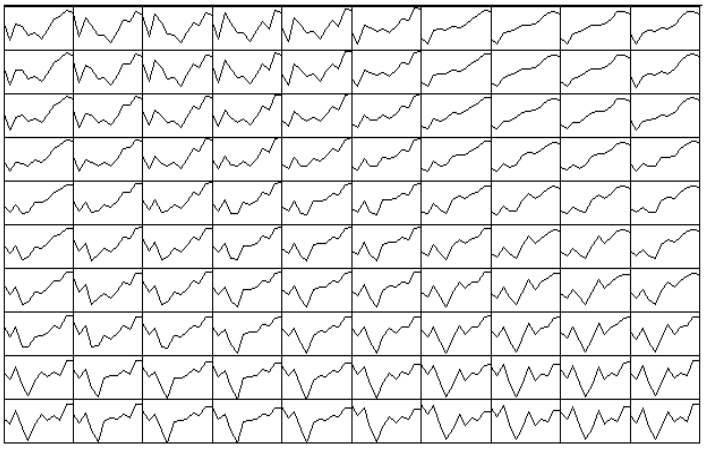


Fig.2 : *Vecteurs codes*

Distances entre les classes. Super-classes

On peut mettre en évidence visuellement les *distance*s entre les classes, qui sont artificiellement équidistantes dans les représentations ci-dessus. Pour cela, en suivant la méthode proposée par E. De Bodt et al. (1996), on dessine dans chaque cellule, un

octogone. Dans chacune des 8 directions principales, son sommet est d'autant plus proche du bord que la distance au voisin dans cette direction est petite. Ceci permet de faire apparaître les groupes de classes proches et donne une idée de la discrimination entre classes.

Comme le choix du nombre *n* de classes est arbitraire (et souvent élevé puisqu'on choisit couramment des grilles 8 par 8 ou 10 par 10), on peut réduire le nombre de classes, en les regroupant au moyen d'une *classification hiérarchique* classique sur les *n* vecteurs codes. On peut alors *colorier* les groupes de classes (appelés *super-classes*) pour les rendre visibles. On constate toujours que les super-classes ne regroupent que des classes contigües, ce qui s'explique par la propriété de respect de la topologie de l'algorithme de Kohonen. D'ailleurs, le non-respect de cette propriété serait un signe de manque de convergence de l'algorithme ou d'une structure particulièrement « repliée » du nuage des données.

Voir les figures 3 et 4, qui montrent les deux classements emboîtés (100 classes de Kohonen et 10 super-classes) et les distances inter-classes. Cf. Cottrell et Rousset (1997).

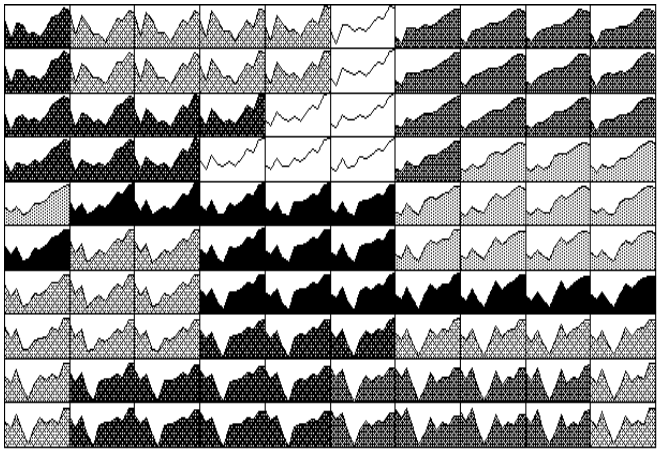


Fig. 3 : *Classes et super-classes coloriées.*

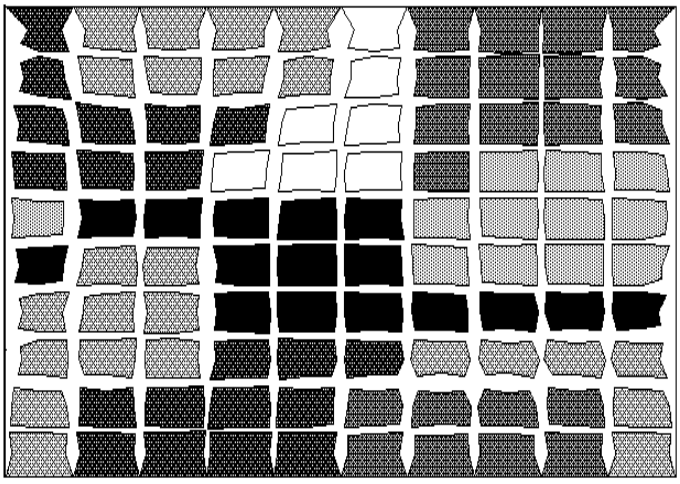


Fig. 4 *: Distances inter-classes*

L'avantage de cette double classification est la possibilité d'analyser les données à un niveau plus grossier où les caractéristiques principales peuvent apparaître, en facilitant l'interprétation.

4 Croisement avec des variables qualitatives

Pour interpréter les classes selon les variables qualitatives non utilisées dans l'algorithme de classement de Kohonen, il peut être intéressant d'étudier la répartition de leurs modalités dans chaque classe. Cf. Cottrell et Rousset (1997), Gaubert et al. (1995). Après avoir calculé des statistiques élémentaires dans chaque classe, on peut dessiner à l'intérieur de chaque cellule un camembert montrant comment sont réparties les modalités de chacune des variables qualitatives, comme le montrent les figures 5 et 6.

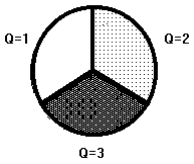


Fig. 5 : *Camembert, pour trois modalités équiréparties*

.

Dans la figure 6, il s'agit toujours des courbes de consommation et la variable qualitative est le jour avec 3 niveaux : Dimanche (noir), Samedi (gris) et jours de semaine (blanc).

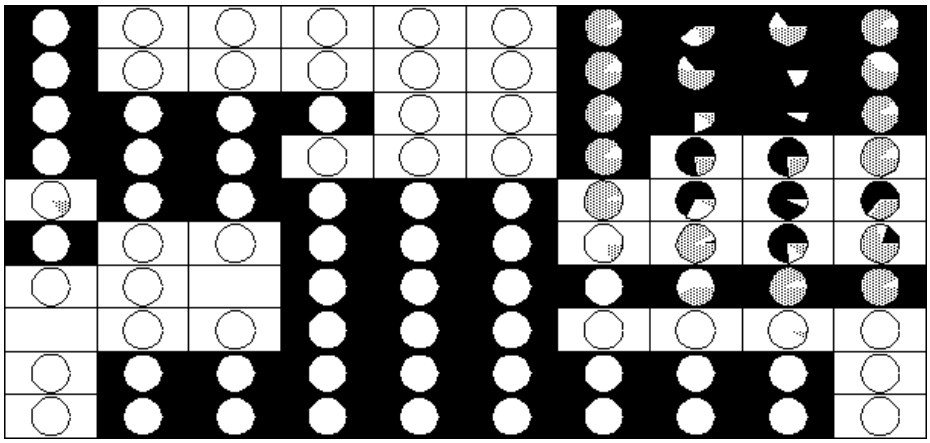


Fig. 6 : *Dans chaque cellule, la répartition de la variable Jour est représentée. Les unités 8, 9, 18, 19, 28, 29, 38, 39, 48, 49, 50, 59 contiennent principalement des dimanches. Les unités 7, 17, 27, 37, 47, 58, 59, 10, 20, 30, 40, 60, 70, contiennent surtout des samedis. Les autres unités ne contiennent que des jours de semaines.*

5 Analyse des relations entre variables qualitatives

On définit ici deux algorithmes originaux qui permettent d'analyser les relations entre variables qualitatives. Le premier est appelé KORRESP et n'est défini que pour l'analyse de deux variables qualitatives. Il est analogue à l'Analyse des Correspondances. Le second est dédié à l'analyse d'un nombre quelconque de variables qualitatives. Il est appelé KACM and ressemble à l'Analyse des Correspondances Multiples. Voir les références Cottrell, Letremy et Roy (1993), Cottrell et Letremy (1994), Cottrell et Ibbou (1995), où sont présentées ces méthodes accompagnées d'exemples.

Pour les deux algorithmes, on considère N individus et un certain nombre K de variables qualitatives. Chaque variable $k = 1,2, ..., K$ a m_k modalités. Chaque individu choisit une et une seule modalité pour chaque variable. Si M est le nombre total de modalités, chaque individu est représenté par un M-vecteur composé de 0 et de 1. Il n'y a qu'un 1 parmi les m_1 premières composantes, seulement un 1 entre la m_1+1-ième et la (m_1+m_2)-ième, etc.

Dans le cas général, quand $M > 2$, les données sont résumées dans une table de Burt qui est un tableau de contingence généralisé. C'est une matrice symétrique de dimension $M \times M$ composée de $K \times K$ blocs, tels que le bloc B_{kl} (pour $1 \le k \le l \le M$) est la table de contingence ($m_k \times m_l$) des variables k et l. Le bloc B_{kk} est une matrice diagonale, dont la diagonale est formée des nombres d'individus qui ont choisi les modalités 1, 2, ..., m_k , pour la variable k. Dans la suite, la table de Burt est notée B.

Dans le cas où M=2, on utilise seulement la table de contingence T qui croise les deux variables. Dans ce cas, on écrit p (resp. q) à la place de m_1 (resp. m_2).

L'algorithme KORRESP

Soit $M = 2$. Dans la table de contingence T, la première variable qualitative a p niveaux et correspond aux lignes. La seconde a q niveaux et correspond aux colonnes. Le terme n_{ij} est le nombre d'individus appartenant à la fois à la classe i et à la classe j. A partir de la table de contingence, on calcule la matrice des fréquences ($f_{ij} = n_{ij} / (\Sigma_{ij}\, n_{ij})$).

On calcule ensuite les profils lignes $r(i)$, $1 \le i \le p$ (le profil $r(i)$ est la distribution conditionnelle de la seconde variable quand la première variable vaut i) et les profils colonnes $c(j)$, $1 \le j \le q$ (le profil $c(j)$ est distribution conditionnelle de la première variable quand la seconde vaut j). L'Analyse des Correspondances est une Analyse en Composantes Principales simultanée et pondérée des profils lignes et des profils colonnes. La distance choisie est celle du χ^2. Dans les représentations simultanées on peut interpréter les proximités entre les modalités des deux variables en terme de corrélations.

Pour définir l'algorithme KORRESP, on construit une nouvelle matrice de données D : à chaque profil ligne $r(i)$, on associe le profil colonne $c(j(i))$ le plus probable sachant i, et réciproquement on associe à chaque profil colonne $c(j)$ le profil ligne $r(i(j))$ le plus probable sachant j. La matrice de données D est la matrice $((p+q) \times (q+p))$ dont les p premières lignes sont les vecteurs $(r(i),c(j(i)))$ et les q dernières sont les vecteurs $(r(i(j)),c(j))$. On applique l'algorithme de Kohonen aux lignes de la matrice D.

Il faut noter qu'on tire au hasard les entrées alternativement parmi les *p* premières lignes et parmi les *q* dernières, mais que l'unité gagnante est calculée seulement sur les *q* premières composantes dans le premier cas et sur les *p* dernières dans le second cas, et selon la distance du χ^2. Après convergence, chaque modalité des deux variables est classée dans une classe de Kohonen. Des modalités « proches » sont classées dans la même classe ou dans des classes voisines. Cet algorithme est une méthode très rapide et efficace d'analyse des relations entre deux variables qualitatives. Voir par exemple Cottrell et Letremy (1994) pour des applications sur données réelles.

6 L'algorithme KACM

Quand il y a plus de deux variables qualitatives, on prend comme matrice des données la table de Burt *B*. Les lignes sont normalisées avec somme 1. A chaque étape, on tire une ligne normalisée au hasard selon la loi de probabilité donnée par la distribution empirique de la modalité correspondante. On définit l'unité gagnante selon la distance du χ^2 et on met à jour les vecteurs codes comme d'habitude. Après convergence, on obtient une classification organisée des modalités, telle que des modalités « liées » appartiennent à la même classe ou à des classes voisines. Dans ce cas aussi, la méthode KACM fournit une technique alternative très intéressante à l'Analyse des Correspondances Multiples classique.

Les principaux avantages des méthodes KORRESP et KACM sont leur simplicité, leur rapidité et leur faible temps calcul. Elles produisent une seule carte alors que les analyses classiques fournissent plusieurs représentations d'information décroissante. Elles sont plus grossières, mais permettent une interprétation rapide. Voir Cottrell et Ibbou (1995), Cottrell, de Bodt et Henrion (1996) pour plus de détails et pour des applications à des données financières.

7 Conclusion et perspectives

Nous proposons un ensemble de méthodes pour analyser des données multi-dimensionnelles, en complément des méthodes linéaires classiques, quand les observations sont décrites par des variables quantitatives et qualitatives. Nous avons réalisé de nombreuses études à l'aide de ces méthodes (étude du fichier historique de l'ANPE, structure de consommation des ménages canadiens, étude comparée des communes de l'Ile-de-france, etc.). En fait, dans la pratique, il est nécessaire de combiner les différentes techniques entre elles. On peut par exemple se servir de la classification en super-classes réalisée à partir de l'algorithme de Kohonen pour définir une nouvelle variable qualitative et pratiquer une Analyse des Correspondances Multiples ou un KACM sur l'ensemble des variables qualitatives. Cela permet d'obtenir une typologie des classes et d'aider à l'interprétation.

Tous ces algorithmes ont été programmés dans le langage IML du logiciel SAS, et nous travaillons à compléter ces outils en y incorporant un certain nombre de calculs de

statistiques standards de manière à compléter les visualisations par des résultats quantitatifs.

Références bibliographiques

Blayo, F. & Demartines, P.(1991) : Data analysis : How to compare Kohonen neural networks to other techniques ? In *Proceedings of IWANN'91*, Ed. A.Prieto, Lecture Notes in Computer Science, Springer-Verlag, 469-476.

Blayo, F. & Demartines, P. (1992) : Algorithme de Kohonen: application à l'analyse de données économiques. *Bulletin des Schweizerischen Elektrotechnischen Vereins & des Verbandes Schweizerischer Elektrizitatswerke*, 83, 5, 23-26.

Cottrell, M. & Fort, J.C., (1987) : Etude d'un algorithme d'auto-organisation, *Annales de l'Institut Poincaré*, Vol. 23, 1, 1-20.

Cottrell, M., Letremy, P. & Roy, E. (1993) : Analyzing a contingency table with Kohonen maps : a Factorial Correspondence Analysis, *Proc. IWANN'93*, J.Cabestany, J.Mary, A.Prieto Eds., Lecture Notes in Computer Science, Springer-Verlag, 305-311.

Cottrell, M. & Letremy, P. (1994) : Classification et analyse des correspondances au moyen de l'algorithme de Kohonen : application à l'étude de données socio-économiques, *Proc. Neuro-Nîmes*, 74-83.

Cottrell, M. & Ibbou, S. (1995) : Multiple correspondence analysis of a crosstabulation matrix using the Kohonen algorithm, *Proc. ESANN'95*, M.Verleysen Ed., Editions D Facto, Bruxelles, 27-32.

Cottrell, M., Fort, J.C. & Pagès, G. (1995) : Two or three things that we know about the Kohonen algorithm, in *Proc of ESANN'94*, M. Verleysen Ed., D Facto, Bruxelles, p.235-244.

Cottrell, M., Girard, B., Girard, Y., Muller, C. & Rousset, P., (1995) : Daily Electrical Power Curves : Classification and Forecasting Using a Kohonen Map, *From Natural to Artificial Neural Computation, Proc. IWANN'95*, Springer, p. 1107-1113.

Cottrell, M., de Bodt, E. & Henrion, E.F. (1996) : Understanding the Leasing Decision with the Help of a Kohonen Map. An Empirical Study of the Belgian Market, *Proc. ICNN'96 International Conference*, Vol.4, 2027-2032.

Cottrell, M., Fort, J.C. & Pagès, G. (1997) : Theoretical aspects of the Kohonen Algorithm, *WSOM'97*, Helsinki 1997.

Cottrell, M. & Rousset, P. (1997) : The Kohonen algorithm : a powerful tool for analysing and representing multidimensional quantitative et qualitative data, *Proc. IWANN'97*, Lanzarote.

De Bodt, E. & Cottrell, M. (1996) : A Kohonen Map Representation to Avoid Misleading Interpretations, *Proc. ESANN'96*, M.Verleysen Ed., Editions D Facto, Bruxelles, 103-110.

Erwin, E., Obermayer, K. & Shulten, K., (1992) : Self-organizing maps : ordering, convergence properties and energy functions, *Biol. Cyb*., 67, p. 47-55.

Gaubert, P., Ibbou, S. & Tutin, C., (1995) : Housing market segmentation and price mechanisms in the Parisian metropolis, *International Journal of Urban and Regional Research*.

Kaski, S. (1997) : Data Exploration Using Self-Organizing Maps, *Acta Polytechnica Scandinavia*, 82.

Kohonen, T. (1984, 1993) : *Self-organization and Associative Memory*, 3°ed., Springer.

Kohonen, T. (1995) : *Self-Organizing Maps*, Springer Series in Information Sciences Vol 30, Springer.

Lebart, L., Morineau, A. & Piron, M. (1995) *: Statistique exploratoire multidimensionnelle*, Dunod.

Ripley, B.D. (1996) : *Pattern Recognition and Neural Networks*, Cambridge University Press.

Ritter, H., Martinetz, T. & Shulten, K., (1992) : *Neural computation and Self-Organizing Maps, an Introduction*, Addison-Wesley, Reading.

Rumelhart, D.E. & McClelland, J.L. (eds) (1986) : *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1. Foundations.* Cambridge, MA: The MIT Press.

Saporta, G. (1990) : *Probabilités, Analyse de données et Statistique*, Technip.

Varfis, A. & Versino, C. (1992) : Clustering of socio-economic data with Kohonen maps, *Neural Network World*, 2, 813-834.