

APPORT DES MODELES GRAPHIQUES GAUSSIENS EN ANALYSE DE SIMILITUDE

NICOLAS ROUSSIAU

Université de Toulouse II, France

SAID JMEL

Université de Nancy II, France

JOSEPH SAINT PIERRE

Université de Toulouse III, France

Résumé: L'analyse de similitude, issue de la théorie des graphes (Flament, 1981; Vergès, 1985), est une technique amplement utilisée dans l'étude des représentations sociales. Elle présente néanmoins un certain nombre d'inconvénients, nous en avons retenu principalement un: elle ne permet pas de tester la significativité du modèle obtenu. Le présent travail se propose, dans le cadre de la théorie des modèles graphiques d'associations, d'utiliser le modèle graphique gaussien (Whittaker, 1990; Edwards, 1991, 1995), permettant ainsi de lever la limite précédemment évoquée. Ce modèle présente deux caractéristiques principales qui le distinguent de l'analyse de similitude. La première spécificité provient d'un principe d'élimination des arêtes qui est basé sur la corrélation partielle, c'est-à-dire que c'est par rapport aux autres variables que sont calculés les coefficients de corrélation des arêtes. L'interprétation du graphe se fait alors en terme d'indépendance conditionnelle. La seconde spécificité est la réalisation de tests d'hypothèses concernant la significativité du modèle graphique obtenu. Pour tester la validité de ce modèle dans le cadre théorique des représentations sociales, nous nous sommes appuyés sur une recherche réalisée avec 553 étudiants sur la représentation sociale de la politique. Nous avons formulé à partir des résultats des graphes des propositions de lecture par groupement d'items (comparaison inter et intra cliques) et des propositions d'hypothèses de centralité sur des items se retrouvant simultanément dans plusieurs cliques.

Abstract: The similitude analysis which is rooted in the graphs theory (Flament, 1981; Vergès, 1985), is a technique thoroughly used in social representations studies. Nevertheless, it has several inefficiencies one of them which we particularly retained is that it cannot be used to test the meaningfulness of the obtained model. Our task consists in using graphical gaussian models (Whittaker, 1990; Edwards, 1991, 1995) which allows to lift the limits previously mentioned. The graphical model has two main qualities which singles it out from

the similitude analysis. The first particularity comes from an elimination principle sides founded on partial correlation, which means: correlation coefficient of sides are calculated with links to other variables. The analysis of the graph is then made in conditionnal independence. The second particularity is the achievement of hypothesis tests as far as the obtained graphics model's meaningfulness is concerned. To test the validity of this model in the theory frame of social representation, we founded our research on a study on social representation of politics among 553 students. We have formulated from the graph results some reading proposals thanks to items groups: between intra and inner items gatherings, and proposal of centrality's hypotheses on items which are found simultaneonsly in several gathering.

I. INTRODUCTION ET POSITION DU PROBLEME

L'analyse de similitude est une technique majeure et amplement utilisée dans l'étude de la structure des représentations sociales. Elle suscite des réflexions méthodologiques (Flament, 1962, 1981; Degenne et Vergès, 1973; Di Giacomo, 1981; Vergès, 1985 etc.) aussi bien que des réflexions théoriques plus générales (Doraï, 1988 etc.). Plus quotidiennement on l'utilise, dans le cadre de l'analyse de situations concrètes, pour décrire une représentation sociale à un moment donné (Abric et Vacherot, 1976; Abric, 1984; Neyrand, 1985; Guimelli, 1989; Yapo, 1993 etc.) et éventuellement, pour suivre son évolution, par exemple dans des études de type longitudinal (Aïssani, 1991; Roussiau, 1996).

C'est en 1962 dans « les Cahiers du Centre de Recherche Opérationnelle » que Flament propose cette méthode pour la première fois. Elle dérive de la théorie des graphes et en utilise la terminologie: rappelons que la structure cognitive de la représentation y est présentée sous la forme d'un graphe arborescent; que les items, c'est-à-dire les éléments qui la composent, sont reliés entre eux par des arêtes; qu'il s'agit d'une méthode combinatoire et structurale. Il n'est pas nécessaire de revenir ici à ses fondements mathématiques qui ont été développés à l'origine (Degenne et Vergès, 1973; Vergès, 1985). Mais on retiendra sa remarquable adaptation à l'étude structurale des représentations.

Néanmoins elle présente un certain nombre d'inconvénients, ou plus exactement de limites. En premier lieu, en ce qui concerne l'analyse de la structure représentationnelle, et plus précisément l'identification de son système central, elle n'autorise pas à dépasser le stade des hypothèses de centralité (Moliner, 1994). En second lieu, parce qu'elle n'est qu'une technique descriptive, elle ne répond pas pleinement aux objectifs des recherches de type comparatif et/ou diachronique: que l'on ait à rendre compte de l'évolution d'une représentation par comparaison entre groupe (Guimelli, 1989), ou à travers des coupes successives au sein d'une même population dans des études longitudinales (Aïssani, 1991), on est toujours amené à effectuer des comparaisons entre deux graphes de similitude sans être en mesure de diagnostiquer la réalité et la portée des variations que l'on croit y discerner. Ce qui nous a amenés à utiliser des méthodes statistiques complémentaires de type inférentiel comme l'analyse de la variance, pour des études longitudinales (Roussiau et Soubiale, 1995-1996; 1996). A signaler que l'analyse de la variance et l'analyse de similitude sur un même corpus, ne mesurent pas la même

organisation des données, dans un cas (anova) ce sont des effets de taille et dans l'autre (ads) des corrélations.

Face à ces limites, et avec la volonté de rester dans ce courant d'analyse particulièrement pertinent pour appréhender l'organisation interne des représentations sociales, nous nous sommes interrogés sur la possible existence de modèles graphiques qui permettraient, par des tests d'hypothèses, de se prononcer sur la validité d'un modèle. Le modèle graphique gaussien¹ (voir: Whittaker, 1990; Edwards, 1991, 1995; Jmel, 1992; Fine, 1992 et Lauritzen, 1996) nous est apparu comme susceptible de répondre à cette exigence. Par rapport à l'analyse de similitude il présente une double spécificité.

1) La première spécificité du modèle, provient d'un principe d'élimination des arêtes qui est basé sur l'indépendance conditionnelle. En d'autres termes, on élimine une arête entre deux sommets du graphe si l'on peut accepter l'indépendance entre les deux variables correspondantes conditionnellement aux autres variables. Le graphe ainsi obtenu est dit d'indépendance conditionnelle. La loi conjointe des variables étant supposée gaussienne (ce qui implique la normalité de chaque variable), l'indépendance conditionnelle est alors équivalente à la nullité du coefficient de corrélation partielle². Rappelons que dans le cas de l'analyse de similitude, les différents indices de similitude sont calculés entre les variables deux à deux sans tenir compte des autres variables.

2) La seconde spécificité tient à la réalisation de tests d'hypothèse concernant la significativité de chaque arête du graphe et l'adéquation de celui-ci aux données. La recherche du « meilleur » modèle (meilleur en ce sens qu'il est le modèle le plus simple, c'est-à-dire avec le minimum d'arêtes et dont l'ajustement aux données est acceptable) peut être effectuée à l'aide du logiciel MIM (voir Edward, 1991, 1995). Il comporte une procédure de sélection automatique (la procédure STEPWISE dans la version 2.3). Il s'agit d'une procédure d'élimination d'arêtes pas à pas. Dans sa version la plus usuelle (le pas à pas descendant), on part d'un modèle de base (par exemple le modèle saturé associé au graphe complet), on enlève l'arête la moins significative correspondant à la plus petite corrélation partielle (en valeur absolue). L'adéquation du modèle obtenu par rapport au modèle de base est alors testée par calcul de la déviance (statistique du test du rapport de vraisemblance), pour un niveau de signification donné (en général, 5%). L'étape suivante utilise le modèle issu de l'étape précédente comme nouveau modèle de base, à partir duquel se répètent les opérations qui viennent d'être décrites. Et ainsi de suite jusqu'au moment où aucune des arêtes restantes ne peut être exclue, c'est-à-dire jusqu'au moment où l'exclusion d'une arête supplémentaire entraînerait le rejet du modèle engendré par cette opération.

¹ Ce modèle suppose la normalité conjointe des variables.

² Le coefficient de corrélation partielle mesure l'intensité de la relation entre deux variables après élimination de l'influence des autres variables (en maintenant constantes les autres variables). Ainsi pour trois variables X_1 , X_2 et X_3 , le coefficient de corrélation partielle entre X_1 et X_2 étant donné X_3 est, en faisant apparaître les coefficients de corrélation usuels (voir: Lebart, Morineau et Piron, 1995):

$$r_{12|3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}} .$$

Notons qu'il existe une autre procédure particulièrement adaptée quand le nombre de variables est très important³. Elle consiste à éliminer en bloc et dès la première étape de l'analyse toutes les arêtes correspondant à des petites corrélations partielles (en valeur absolue) par rapport à un seuil donné⁴. Cette procédure présente néanmoins des inconvénients: en effet elle peut éliminer dès le début, des arêtes qui peuvent être acceptées, c'est la raison pour laquelle on lui préfère les procédures de sélection pas à pas.

La présentation d'une application va éclairer cette démarche, tout en nous fournissant une occasion de réfléchir à son utilité dans le champ théorique des représentations sociales et notamment dans la perspective de la théorie dite du noyau central. Rappelons qu'ici une représentation est définie comme composée de deux systèmes, l'un central et l'autre périphérique. Le système central a la particularité de donner sens et cohérence à l'ensemble de la structure représentationnelle, il est le principe organisateur de la représentation. La périphérie, quant à elle, constitue une sorte « d'interface » entre la réalité extérieure et le noyau central; suivant le principe d'économie cognitive elle est plus facilement modifiable que le système central (Abric, 1994 et Flament, 1994).

II. UN EXEMPLE D'APPLICATION

L'application a porté sur 17 variables⁵ extraites des données recueillies dans le cadre de l'étude de la représentation sociale de la politique dans une population étudiante (N = 553). Pour l'analyse de ces données par modèles graphiques, nous avons utilisé les deux

³ Le logiciel MIM utilisé ici ne permet pas, dans sa version actuelle, de traiter plus de 26 variables.

⁴ La justification théorique de cette procédure repose sur la propriété suivante: la déviance du modèle d'indépendance de deux variables X_i et X_j conditionnellement aux variables de $X_{K-\{i,j\}}$ par rapport au modèle saturé, est donnée par:

$$D = -N \log(1 - r_{ij|K-\{i,j\}}^2)$$

où N désigne la taille de l'échantillon et $r_{ij|K-\{i,j\}}$ désigne la corrélation partielle de X_i et X_j dégagée de l'influence du reste des variables $X_{K-\{i,j\}}$. On montre que cette déviance D suit un χ^2 à 1 degré de liberté (voir Whittaker, 1990) et on accepte l'exclusion de l'arête (i,j) si cette quantité est petite par rapport à un seuil donné par la loi du χ^2 pour un certain niveau de signification α . Les plus petites corrélations partielles suggèrent donc d'accepter les indépendances conditionnelles des variables correspondantes et donc d'éliminer les arêtes associées.

⁵ Les opinions du questionnaire sont les suivantes: A « La Constitution de la V République fonctionne bien »; B « Le fossé entre la société et la classe politique tend à se réduire »; C « Les jeunes qui ont leur avenir devant eux sont les premiers concernés par la politique »; D « Nos institutions ont été pensées en vue de faire progresser la démocratie »; E « Si les institutions se sont maintenues jusqu'ici, c'est que ce sont de bonnes institutions »; F « Les partis agissent conformément aux vœux de leur électorat »; G « La politique suscite actuellement un renouveau d'intérêt chez les français »; H « Les hommes politiques mentent plus souvent qu'ils ne disent la vérité »; I « Les partis sont devenus des organisations où la communication est bloquée »; J « Bien des hommes politiques sont animés par le souci du service collectif »; K « C'est le goût du pouvoir qui pousse à se présenter à des élections »; L « L'Assemblée Nationale est un organe vital dans la vie démocratique »; M « Réaliser les promesses faites aux électeurs est l'objectif que poursuivent les élus »; N « Les gens ne se sentent plus correctement représentés par les élus »; O « La politique ne peut intéresser qu'une minorité de gens »; P « S'abstenir c'est ce qu'il y a de mieux à faire un jour d'élection »; Q « Beaucoup de politiciens deviennent malhonnêtes ».

procédures décrites au paragraphe précédent, la procédure d'élimination des arêtes pas-à-pas et la procédure d'élimination en bloc pour un seuil donné.

II-1. PREMIERE PROCEDURE.

Il s'agit de la procédure de sélection pas à pas décrite ci-dessus, cependant compte tenu du nombre relativement important de variables (17 variables génèrent 2^{136} modèles), nous avons procédé en deux phases pour la recherche d'un modèle acceptable ⁶.

Dans une première phase, nous avons lancé la procédure STEPWISE de MIM en fixant le niveau de signification à 5% et en restreignant la recherche à une sous-classe de modèles graphiques, modèles décomposables caractérisés par des graphes triangulés, c'est-à-dire ne comportant aucun cycle de longueur supérieure ou égale à 4 sans corde ⁷. Cette première phase nous a donné le modèle suivant:

(ABCGHJMO); (ABCJMOPQ); (ABCHJMOQ); (ABEJMOPQ); (ABELMOP); (ADEMP); (ACHIN); (ACGHN); (EFM); (CHJKQ).

Ce modèle graphique, constitué à partir de ses cliques, a une déviance de 59,9249 pour 63 degrés de liberté (d.d.l.) avec un niveau de signification ⁸ de 0,5866. Il est acceptable, mais son graphe comporte encore plusieurs arêtes. Il convient alors de poursuivre la recherche jusqu'à l'obtention d'un modèle minimal, c'est-à-dire ayant le minimum d'arêtes et facilement interprétable.

Par conséquent, la deuxième phase a consisté à améliorer ce premier modèle en rejetant tour à tour les arêtes les moins significatives (correspondant aux faibles corrélations partielles estimées) jusqu'à ce qu'aucune arête restante ne puisse être supprimée. Cette deuxième phase nous a permis d'obtenir le modèle suivant:

(ABM); (ACGHJM); (ACGHN); (ACHIN); (ACJMOQ); (ACHJMQ); (ADEM); (AEJMOQ); (EFM); (HJKQ); (ELMO); (COP).

Ce nouveau modèle graphique a une déviance de 104,4409 pour 83 degrés de liberté avec un niveau de significativité de 0,0559. Il est donc acceptable. Son graphe comporte moins d'arêtes que le précédent, ce qui rend son interprétation plus aisée. Et au seuil de 5%, aucune arête ne peut plus être exclue. Mais le nombre d'items par clique rend une présentation de ce modèle par graphe, illisible. En effet dans chaque clique tous les items sont corrélés les uns aux autres. L'interprétation ne s'effectuera donc qu'à partir des cliques telles qu'elles nous sont présentées dans le modèle ci-dessus.

La procédure ici décrite, et les résultats auxquels elle aboutit, nous conduisent à des propositions de lecture par groupement d'items (comparaison inter et intra cliques) et à la formulation d'hypothèses de centralité sur des items figurant dans plusieurs cliques.

a. Les propositions de lecture par groupement(s) d'items. La lecture centrée sur un groupement permet de constater son homogénéité sémantique, ou son hétérogénéité, voire son caractère contradictoire. Dans ce dernier cas, il est bon de ne pas conclure trop hâtivement que l'on se trouve en présence d'un agrégat d'opinions dépourvu

⁶ Remarquons que la stratégie utilisée ici n'est pas la seule possible et que le modèle final dépend de la stratégie adoptée (deux stratégies différentes ne donnent pas forcément le même modèle).

⁷ Une corde d'un cycle du graphe est une arête reliant deux sommets non consécutifs.

⁸ Le niveau de signification est donné ici par $P(\chi^2_{63} > 59,9249)$, le modèle est significatif à 5% si ce niveau dépasse 0,05.

de sens, sachant que des incohérences peuvent renvoyer à la complexité du champ de représentation (Capozza, Robusto, Squarza & De Carlo, 1995). Pour une grande majorité des cliques on constate une remarquable homogénéité sémantique. Prenons à titre d'exemple tout d'abord la clique composée des opinions ⁹ A-D-E-M qui fait apparaître un espace sémantique centré essentiellement autour de problèmes institutionnels: A. « La Constitution de la V République fonctionne bien » (m = 2,97); D. « Nos institutions ont été pensées en vue de faire progresser la démocratie » (m = 4,13); E. « Si les institutions se sont maintenues jusqu'ici, c'est que ce sont de bonnes institutions » (m = 2,84) et M « Réaliser les promesses faites aux électeurs est l'objectif que poursuivent les élus » (m = 2,17). Autre exemple la clique H-J-K-Q qui recouvre très clairement un espace sémantique lié aux problèmes de moralité: H. « Les hommes politiques mentent plus souvent qu'ils ne disent la vérité » (m = 4,75); J. « Bien des hommes politiques sont animés par le souci du service collectif » (m = 2,85); K. « C'est le goût du pouvoir qui pousse à se présenter à des élections » (m = 4,72) et Q. « Beaucoup de politiciens deviennent malhonnêtes » (m = 4,76). Le phénomène de l'émergence d'une crise de la moralité dans le monde politique n'est pas nouveau et rejoint les résultats d'autres recherches (Quadrio, Cattellani et Sala, 1988; Cattellani et Quadrio, 1991; Bonardi, De Piccoli, Larrue et Soubiale, 1994) où les problèmes de moralité apparaissent comme des éléments « fédérateurs » ou « organisateurs » d'une représentation négative de la politique caractérisée avant tout par ses dysfonctionnements. Les opinions regroupées dans cet espace sémantique comme dans le précédent prennent sens les unes par rapport aux autres. Ce type de lecture que nous présentons s'apparente à celui proposé pour l'arbre maximum, puisque, rappelons-le, l'interprétation de chaque opinion se fait en fonction des opinions qui lui sont proches. En effet, « Seul(s) compte(nt) la place des items, les uns par rapport aux autres (représentés par les sommets du graphe), (...). C'est la lecture de similitude de proche en proche, sur l'arbre, qui permet une interprétation de la signification d'un item en fonction de son voisinage » (Abric et Mardellat, 1983). C'est donc la lecture de proche en proche à l'intérieur d'une clique qui donne sens aux différentes opinions la constituant. Signalons simplement que chez ces auteurs, le fait que les items prennent sens en fonction de leur proximité les uns par rapport aux autres a un fondement mathématique, puisque les items connexes sur l'arbre ont été choisis un peu de la même façon par les sujets. Chez nous cette proposition de lecture correspond plus à un résultat mathématique qui n'est pas lié directement au choix associatif d'opinions des sujets.

La lecture en simultané de plusieurs groupements déchiffre leur sens les uns par rapport aux autres. A titre d'exemple comparons la clique précédente H-J-K-Q et la clique A-B-M: A. « La Constitution de la V République fonctionne bien » (m = 2,97); B. « Le fossé entre la société et la classe politique tend à se réduire » (m = 2,30) et M. « Réaliser les promesses faites aux électeurs est l'objectif que poursuivent les élus » (m = 2,17). On retrouve ici les thèmes « politique et institutions » (opinion A) et « politique et moralité » (opinion M) associés à un thème qui serait plus proche des problèmes qui lient

⁹ Les sujets répondent à chaque opinion sur une échelle en six points. "1" correspond à « désaccord total »; "2" correspond à « désaccord »; "3" correspond à « plus ou moins en désaccord »; "4" correspond à « plus ou moins en accord »; "5" correspond à « accord » et "6" correspond à « accord total ».

« la politique et la société » (opinion B). Une interprétation possible serait que l'écart décrété entre les hommes politiques et la société (opinion B) s'expliquerait par leur malhonnêteté (opinion M), qui elle-même serait à l'origine du jugement négatif porté sur le fonctionnement constitutionnel (opinion A). Dans cette clique, les problèmes de moralité apparaissent comme éléments organisateurs d'une perception négative du fonctionnement global de la politique (manque de confiance, etc.).

b. Les hypothèses de centralité. Elles s'inscrivent dans le champ théorique du noyau central, dont il s'agit de repérer et identifier les éléments. La présence répétée de certains items dans la suite des cliques, associée à une forte connexité avec d'autres items, fait émerger des hypothèses de centralité, susceptibles d'être vérifiées par la suite par la technique de la mise en cause (M-E-C) (Moliner, 1992). Dans les exemples ci-dessus, deux opinions reviennent avec une constance remarquable dans les cliques: A. « La Constitution de la V République fonctionne bien » ($m = 2,97$) et M. « Réaliser les promesses faites aux électeurs est l'objectif que poursuivent les élus » ($m = 2,17$). Ces deux opinions nous laissent entrevoir d'un côté (item A) plutôt un constat, un dysfonctionnement constitutionnel de la V république et de l'autre (item M) un autre constat qui serait une cause possible du précédent. En effet les réponses à cette deuxième opinion traduisent l'idée que les élus n'ont pas pour objectif de réaliser les promesses faites à leurs concitoyens, donc qu'ils manquent d'éthique et de moralité. Pour éviter toute généralisation abusive, nous en resterons là concernant l'interprétation des résultats obtenus par cette procédure. Les résultats du deuxième type de traitement nous aideront peut-être à nous faire une idée plus précise du sens à accorder aux items présents fréquemment dans plusieurs cliques.

II-2. DEUXIEME PROCEDURE

La procédure utilisée ici consiste tout d'abord à déterminer la matrice des corrélations partielles¹⁰ et ensuite à enlever en bloc l'ensemble des arêtes correspondant à des corrélations partielles petites par rapport à un seuil donné. Nous avons dans un premier temps fixé α à 1%, ce qui correspond à un seuil de 0.109 pour les corrélations partielles. L'élimination de l'ensemble des arêtes pour lesquelles les corrélations partielles sont inférieures ou égales à ce seuil aboutit au modèle:

(ADE); (AI); (B); (CG); (CI); (CO); (CP); (EF); (FM); (GJ); (HJK); (HM); (HKQ); (EJ); (LO); (N); (EQ)

dont la déviance est 286.6830 pour 116 degrés de liberté et un niveau de signification nul. Ce modèle n'est donc pas acceptable. D'autres arêtes doivent donc être prises en compte.

¹⁰ La matrice des corrélations partielles, R_p est simplement déduite de l'inverse de la matrice des corrélations R par la transformation

$$R_p = 2I - \left[\text{diag}(R^{-1}) \right]^{1/2} R^{-1} \left[\text{diag}(R^{-1}) \right]^{1/2}$$

où $\text{diag}(R^{-1})$ désigne la matrice diagonale ayant les mêmes éléments diagonaux que R^{-1} (voir par exemple Whittaker 1990, chap. 1) et I la matrice identité.

Nous avons relevé, par la suite, la valeur de α à 5%, ce qui correspond à un seuil de 0.083 pour les corrélations partielles. L'élimination en bloc des arêtes correspondant à des corrélations partielles inférieures ou égales à ce seuil nous a conduit au modèle suivant:

(ABGH); (ADE); (AEJ); (AHI); (AGHJ); (CG); (CI); (COP); (EF); (FM); (HJK); (HKQ); (BLM); (EL); (LMO); (BGHM); (DM); (GHJM); (GHN); (HIN); (EQ), dont la déviance est de 109,84 pour 96 degrés de liberté et un niveau de signification de 0,1581. Ce modèle est d'emblée acceptable au seuil de 5%. Nous nous attacherons donc, dans la suite de cette étude, à la lecture de ce modèle.

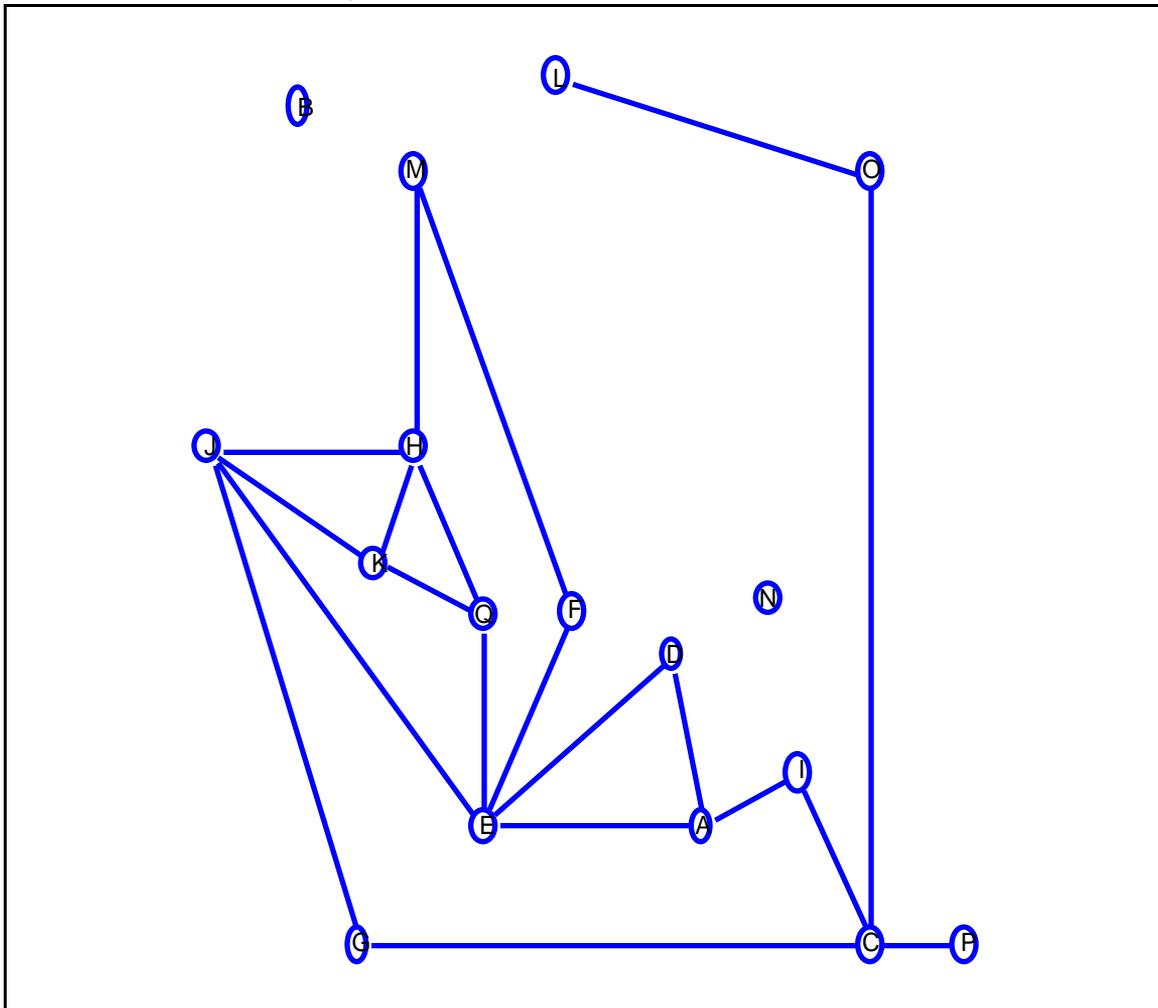


Figure 1

Graphe d'indépendance: les corrélations partielles supérieures à 0.109

a. Les propositions de lecture par groupement(s) d'items. Comme pour la première procédure l'on retrouve des associations d'items dont la cohérence sémantique est forte. Exemple la clique H-J-K qui associe des items traitant des problèmes de moralité. H. « Les hommes politiques mentent plus souvent qu'ils ne disent la vérité » ($m = 4,75$); J. « Bien des hommes politiques sont animés par le souci du service collectif » ($m = 2,85$) et K. « C'est le goût du pouvoir qui pousse à se présenter à des élections » ($m = 4,72$). Ou encore une clique centrée plutôt sur les problèmes

institutionnels A-D-E. A. « La Constitution de la V République fonctionne bien » ($m = 2,97$); D. « Nos institutions ont été pensées en vue de faire progresser la démocratie » ($m = 4,13$) et E. « Si les institutions se sont maintenues jusqu'ici, c'est que ce sont de bonnes institutions » ($m = 2,84$). Ces associations sémantiques sont très proches de celles obtenues dans le cadre de la première procédure d'élimination d'arêtes. Nous pouvons constater que cette seconde procédure diminue le nombre des items par clique. L'intérêt est une lecture plus ciblée et moins diffuse des espaces sémantiques. Si l'on s'attache maintenant à une analyse comparative des groupements d'items on peut repérer des ensembles de cliques traitant de problèmes similaires. En effet la clique H-J-K dont nous venons de parler est particulièrement proche de la clique H-K-Q, tout d'abord parce que deux items sur trois sont identiques, mais surtout parce que le troisième partage le champ sémantique des deux cliques: il s'agit en l'occurrence de l'item Q. « Beaucoup de politiciens deviennent malhonnêtes » ($m = 4,76$). Nous retrouvons très clairement identifiés les problèmes liés à l'éthique et la moralité dans le monde politique.

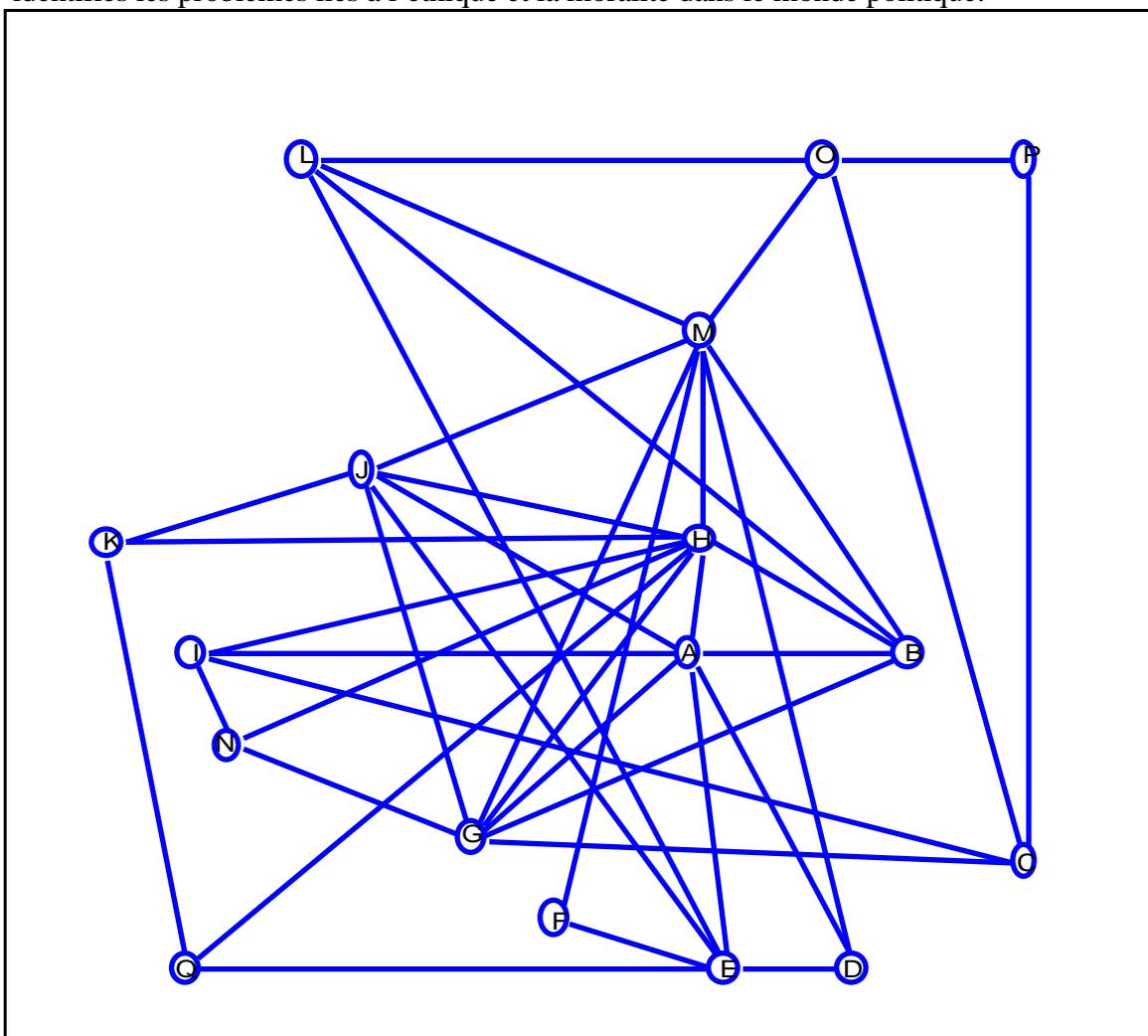


Figure 2
 Graphe d'indépendance: les corrélations partielles supérieures à 0.083

b. Les hypothèses de centralité. Nous l'avons vu, la fréquence d'apparition nous permet de repérer des items plus présents que d'autres dans les différentes cliques ou liaisons d'items deux à deux proposées par le modèle. Nous retrouvons comme pour le modèle analysé lors de la première procédure, l'importance de l'item M, qui comptabilise huit connexités sur le graphe et six fréquences dans les cliques. Un autre item a retenu notre attention puisqu'il présente un nombre plus élevé de connexités (neuf) et de fréquences d'apparition (neuf) que l'item M; il ne s'agit pas comme précédemment de l'item A, mais de l'item H. « Les hommes politiques mentent plus souvent qu'ils ne disent la vérité » ($m = 4,75$). Or ces deux items renvoient à l'une des deux modalités du système central de la représentation de la politique (problèmes de moralité et pouvoir) mis en évidence dans la recherche dont sont extraites les données reprises ici (Roussiau, 1996) et cela par plusieurs autres techniques: analyse de similitude, mots associés (Vergès, 1992, 1994) et induction par scénario ambigu (Moliner, 1993). La représentation graphique permet de voir très clairement que l'organisation structurale est « pilotée » par ces deux items qui comptabilisent un nombre conséquent de connexités.

On aura néanmoins noté la « pauvreté » du critère quantitatif qui fonde la formulation de ces hypothèses, puisqu'il ne s'agit présentement de rien d'autre que de la fréquence d'apparition d'une opinion dans la suite des cliques et de sa connexité avec d'autres items. C'est donc avec prudence que nous avancerons sur ce terrain. Et de même, l'avenir dira si le découpage des « espaces sémantiques » opéré par cette méthode est pertinent, et adapté à l'analyse structurale des représentations sociales. Des recherches plus spécifiques, c'est-à-dire centrées sur l'analyse de l'organisation interne des représentations à l'aide du modèle proposé, devront être réalisées dans le but d'affiner et de préciser des procédures de lecture encore totalement exploratoires.

CONCLUSIONS ET PERSPECTIVES

Soulignons en terminant un point important dans la comparaison entre l'approche que nous avons jusque-là utilisée (analyse de similitude) et l'approche par le modèle graphique gaussien « des corrélations partielles »: la première est purement exploratoire, alors que dans la deuxième, on « injecte » de l'inférentiel au cours de la recherche exploratoire de structure.

L'analyse par modèles graphiques gaussiens requiert l'hypothèse de normalité, cette hypothèse peut ne pas être satisfaite par les données analysées (ce qui semble le cas de l'exemple traité), on peut alors s'orienter vers les modèles graphiques loglinéaires pour variables qualitatives. Dans notre exemple le nombre important de modalités ne nous a pas permis d'utiliser le logiciel MIM dont nous disposons, sur ce point. Les différentes déviations ne sont donc données qu'à titre indicatif. Cette étude devrait être confirmée ultérieurement par des analyses sur données simulées.

Notons cependant que l'idée fondamentale de cet article est d'utiliser des coefficients partiels pour tenir compte des variables cachées et que la matrice de corrélations partielles peut être utilisée chaque fois qu'on peut utiliser la matrice de corrélations.

L'analyse par modèles graphiques gaussiens ne cherche pas à donner de meilleurs résultats que l'analyse de similitude, mais le modèle graphique peut détecter des corrélations superflues entre variables et éviter des interprétations erronées. Les deux approches peuvent donc être utilisées de façon complémentaire. En effet, l'utilisation de la

matrice de corrélations partielles comme matrice de similitude dans un programme d'analyse de similitude permettrait, sous réserve que l'hypothèse de normalité soit respectée, d'associer au graphe final un modèle graphique dont on peut tester l'adéquation aux données, de tester l'importance des différentes associations dans une optique confirmatoire mais aussi de déterminer des éléments tels que l'arbre maximum ou le filtrant.

Enfin il serait intéressant de considérer les modèles graphiques gaussiens à graphe orienté afin de mettre en évidence des structures causales -ce qui n'est pas dans notre propos- Mais c'est une étape qu'il nous paraît indispensable d'explorer.

BIBLIOGRAPHIE

- Abric, J.C. (1984). L'artisan et L'artisanat: analyse du contenu et de la structure d'une représentation. *Bulletin de Psychologie*, 366, 861-876.
- Abric, J.C. (1994). L'organisation interne des représentations sociales: système central et système périphérique. In C. Guimelli (Ed.), *Structure et transformation des représentations sociales*. Neuchâtel: Delachaux et Niestlé.
- Abric, J.C., Vacherot, G. (1976). Méthodologie et étude expérimentale des représentations sociales. Tâche, partenaire et comportement en situation de jeu. *Bulletin de Psychologie*, 29, 735-736.
- Abric, J.C., Mardellat, R. (1983). Les représentations sociales de l'artisan et de l'artisanat. Etude réalisée pour le Centre d'Etude et de Perfectionnement de l'artisanat et des Métiers. (CEPAM). *Etudes et Recherches du GIFRESH*.
- Aïssani, Y. (1991). Etude expérimentale de la transformation d'une représentation sociale sous influence majoritaire et minoritaire. Thèse de Doctorat: Université de Toulouse le Mirail.
- Bonardi, C., De Piccoli, N., Larrue, J., Soubiale, N. (1994). Dipendenza e interdipendenza delle rappresentazioni sociali: la rappresentazione dell'europa e quella della politica. *Giornale Italiano di Psicologia*, XXI, 3, 399-419.
- Capozza, D., Robusto, E., Squarza, R., De Carlo, N.A. (1995). La représentation de l'argent. *Textes sur les Représentations Sociales*, 4, 1, 185-104.
- Catellani, P., Quadrio, A. (1991). Ideal and real in the representation of politics. *Revue Internationale de Psychologie Sociale*, 4, 3/4, 233-256.
- Degenne, A., Vergès, P. (1973). Introduction à l'analyse de similitude. *Revue Française de Sociologie*, 14, 471-512.
- Di Giacomo, J.P. (1981). Commentaire à propos de l'analyse de similitude ... de Cl. Flament. *Cahiers de Psychologie Cognitive*, 1, 429-432.
- Doraï, M. (1988). Représentation sociale et stéréotypie. In J.L. Beauvois, R.V. Joule, J.M. Monteil (Eds.), *Perspectives Cognitives et Conduites Sociales*. Tome 2. Fribourg: Delval.
- Edwards, D. (1991). A guide to MIM 2.0, Statistical Research Unit. University of Copenhagen. Danemark.
- Edwards, D. (1995). *Introduction to Graphical Modelling*. Springer-Verlag.
- Fine, J. (1992). Modèles graphiques d'associations. Modèles pour l'analyse des données multidimensionnelles. *Economica*, 267-313.

- Flament, C. (1962). L'analyse de similitude. Cahier du Centre de Recherche Opérationnel, 4, 63-97.
- Flament, C. (1981). L'analyse de similitude: une technique pour les recherches sur les représentations sociales. Cahiers de Psychologie Cognitive, 1, 375-395.
- Flament, C. (1994). Aspects périphériques des représentations sociales. In C. Guimelli (Ed.), Structure et transformation des représentations sociales. Neuchâtel: Delachaux et Niestlé.
- Guimelli, C. (1989). Pratiques nouvelles et transformation sans rupture d'une représentation sociale: la représentation de la chasse et de la nature. In J.L. Beauvois, R.V. Joule & J.M. Monteil (Eds.), Perspectives Cognitives et Conduites Sociales. Tome 2. Fribourg: Delval.
- Jmel, S. (1992). Application des modèles graphiques au choix de variables et à l'analyse des interactions dans une table de contingence multiple. Thèse de Doctorat. Université Paul Sabatier: Toulouse III.
- Lauritzen, S.L. (1996). Graphical Models. Oxford University Press.
- Lebart, L., Morineau, A., Piron, M. (1995). Statistique exploratoire multidimensionnelle. Dunod: Paris.
- Moliner, P. (1992). La représentation sociale comme grille de lecture. Presses Universitaires de Provence: Aix-en-Provence.
- Moliner, P. (1993). L'induction par scénario ambigu: une méthode pour l'étude des représentations sociales. Revue Internationale de Psychologie Sociale, 6, 2, 7-21.
- Moliner, P. (1994). Les méthodes de repérage et d'identification du noyau des représentations sociales. In C. Guimelli (Ed.), Structure et transformation des représentations sociales. Neuchâtel: Delachaux et Niestlé.
- Neyrand, G. (1985). Femme privé/homme public: les tendances de la représentation publicitaire. Un exemple d'application de l'analyse de similitude. Informatique et Sciences Humaines, 66, 121-135.
- Quadrio, A., Catellani, P., Sala, V. (1988). La rappresentazione sociale della politica. Archivio di Psicologia, Neurologia e Psichiatria, 49, 5-27.
- Roussiau, N. (1996). Représentation sociale et théorie de l'engagement. La représentation sociale de la politique: étude expérimentale de ses transformations. Thèse de Doctorat. Université de Toulouse le Mirail.
- Roussiau, N., Soubiale, N. (1995-1996). Approche expérimentale de la modification d'une représentation sociale sous l'effet de la communication d'un message. Bulletin de Psychologie, Tome XLIX, N°422, 88-99.
- Roussiau, N., Soubiale, N. (1996). Etude de la transformation de la représentation sociale de l'Europe sous impact majoritaire et minoritaire. Anuario de Psicologia, Tome 70, 19-43.
- Vergès, P. (1985). Interprétation au premier degré. L'analyse de similitude au plus près des propriétés mathématiques. Informatique et Sciences Humaines, 15, 67, 27-40.
- Vergès, P. (1992). L'évocation de l'argent: une méthode pour la définition du noyau central d'une représentation. Bulletin de Psychologie, XLV, 405, 203-209.
- Vergès, P. (1994). Approche du noyau central: propriétés quantitatives et structurales. In C. Guimelli (Ed). Structure et transformations des représentations sociales. Neuchâtel: Delachaux et Niestlé.

- Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Wiley.
- Yapo, Y. (1993). La ruralité comme objet de représentation: noyaux structurants, fondements cognitifs et sociaux. *Les Cahiers Internationaux de Psychologie Sociale*, 20, 15-38.

Nicolas Roussiau
Maison de la recherche
Psychologie sociale
Université de Toulouse II
France