



HAL
open science

Leon Festinger, 1949, The Analysis of Sociograms Using Matrix Algebra. Version bilingue et commentée

Leon Festinger, Françoise Bahoken, Laurent Beauguitte

► To cite this version:

Leon Festinger, Françoise Bahoken, Laurent Beauguitte. Leon Festinger, 1949, The Analysis of Sociograms Using Matrix Algebra. Version bilingue et commentée. 2023. hal-04053386

HAL Id: hal-04053386

<https://hal.science/hal-04053386>

Submitted on 31 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Leon Festinger, 1949, The Analysis of Sociograms
Using Matrix Algebra
Human Relations, vol. 2, n°2, pp. 153-158.
Version bilingue et commentée

groupe f.m.r.


Mars 2023

Version traduite et commentée par Françoise Bahoken (Université Gustave Eiffel) et Laurent Beauguitte (CNRS, UMR Géographie-cités).

Présentation

Leon Festinger (1919-1989) est un psychosociologue américain connu pour sa théorie de la dissonance cognitive. Ses premiers travaux, dont cet article est issu, sont fortement influencés par l'approche de Kurt Lewin qui dirige sa thèse à l'université d'État de l'Iowa (1942). En 1945, il entre au *Massachusetts Institute of Technology* (MIT) où Kurt Lewin fonde le *Research Center for Group Dynamics* ; ce centre est transféré à l'Université du Michigan en 1948, après le décès de Lewin. Leon Festinger enseigne ensuite à l'Université du Minnesota (1951-1954) puis à l'Université de Stanford (1955-1968).

Cet article propose de mobiliser le calcul matriciel pour l'analyse des sociogrammes ; il paraît dans la jeune revue *Human Relations*, créée en 1947 par le *Tavistock Institute of Human Relations* (Londres). Cette revue jouera un rôle important dans l'« école de Manchester » dans la mesure où elle accueillera l'article fondateur de John Barnes (1954) mais également des articles d'Elizabeth Bott (1954, 1955) (Edwards, 2016).

Paul Edwards, 2016, [Human Relations: The first 10 years, 1947–1956](#), *Human Relations*, 70(1), p. 1-4

Les numéros entre crochets dans la version anglaise correspondent à la pagination de l'article paru en 1949. Toutes les notes en bleu dans la traduction française sont des traductrices.

The Analysis of Sociograms Using Matrix Algebra

Leon Festinger

There is, at present, no adequate analytical device for handling data from one of the most popular measurement techniques in the field of social psychology. Sociometric questions such as "who are your best friends?" or "what people do you like most to be with?" are increasingly being used whenever an interest in the "structure" or "patterning" of relationships among a number of persons is present. The exact wordings of the questions used represent an almost infinite series of variations, depending upon the context of the study in which the question is asked. The character of the data, however, is always the same. The data obtained are the specific persons mentioned by each one in response to the question.

In the study of groups by means of sociometric data, much attention has been given to the exact pattern of interconnections among individuals. Of importance have been such questions as who are a particular person's friends, what are his relations with the friends of his friends and what tendencies to subgroup or clique formation exist in the group; how many paths of influence exist among members of a group, over how much of a group does a person's influence extend, what is the nature of the two or three step influence chains that may exist; if an item of information enters a group, how many people will eventually hear about it; who will hear it and from whom, and how far removed from the original source will it be by the time a specific person hears about it?

Without any adequate representational techniques for handling such data, the analysis of the exact patterns of interconnections among members of a group is virtually impossible, unless the group is very small. As the size of the group increases, the complexity of the pattern generally makes it extremely difficult to comprehend by mere inspection. The result has been the relative neglect of this kind of analysis. Investigators have, by and large, contented themselves with analyzing sociometric patterns in such terms as how many choices people receive, what kinds of people get most choices, what proportion of the choices are inside the group and other such summary measures, which serve to relate the sociometric choices to other variables.

Initial attempts at the analysis and description of the exact patterning of interconnections in a group took the form of drawing complicated diagrams where the relations between persons were represented by lines, with arrows on them. These diagrams could then be examined. It is readily understandable that such a diagram would be completely unwieldy if the number of members in the group was large or if the number of choices made by each member was appreciable. There are no rules for operating on such diagrams. One merely arranges them by trial and error so as to make the diagram look as simple

as possible, and then one further inspects and examines with the hope that one will be alert enough to see what is to be seen.

It was the recognition of these difficulties that led people like Northway (4) and Cook (2) to attempt to formulate a system which would order the sociometric diagram so as to make it more easily understandable. These attempts, however, still left the analysis of sociometric patterns in a vague and relatively unsystematic state. A large step forward was taken by Forsyth and Katz (3) in suggesting the use of a matrix and some of the manipulations of matrix algebra for the analysis of sociometric patterns. Their idea was to represent the sociometric pattern in a matrix form and then to rearrange this matrix, according to certain principles, in order to have it reveal the subgroupings that were present. The matrix presentation is simply performed by listing the individuals in the group along the rows and along the columns in the same order. The choices that any individual makes are then indicated by the number 1 in one of the squares, such that the row corresponds to the person making the choice and the column corresponds to the person receiving the choice. The squares along the main diagonal of the matrix are of course left blank, since the person is not considered as having chosen himself. Looking across any row reveals who was chosen by that person, and looking down any column reveals from whom that person received choices.

The presentation in this form offers little or no advantage over the more complicated sociometric diagram for direct inspection. Some manipulations of the matrix, however, tend to simplify it. The suggestion of Forsyth and Katz is to rearrange the order of the members of the group so that the numbers in the matrix cluster as closely as possible along the diagonal. More technically, they suggest rearranging the order of the rows and columns so as to minimize the square of the perpendicular deviations of the numbers from the diagonal of the matrix. This rearrangement, once accomplished, will show, clustered together, those people who choose each other frequently, and, relatively separated, those who do not choose each other. It does thus enable us to separate out subgroups which exist and would seem to be a plausible means of analyzing changes in the pattern of connections in a group from one time to another. Such changes would show themselves in different ordering of the members of the group in the final matrix arrangement from one time to another. It does not seem to be too helpful a method for comparing two different groups, and the labor involved in obtaining the ordering of the matrix which will cluster the choices most closely along the diagonal is quite tedious.

There are other manipulations of matrix algebra which would seem to be of use for the analysis of such patterns of connections among members of a group. When the sociometric pattern is presented in a matrix form, an analysis of some aspects of the structure of the group can be performed by

the relatively simple means of squaring and cubing this Matrix.¹ The [155] squared matrix is obtained rather readily in the following way : to obtain the number which goes into the cell corresponding to column c and row r of the squared matrix we multiply each cell in column c of the original matrix by the corresponding cell of row r and then add these products up. The general equation for this multiplication might be written as follows :

$$A_{rc}^2 = A_{1c}A_{r1} + A_{2c}A_{r2} + A_{3c}A_{r3} + \cdots + A_{nc}A_{rn}.$$

In this equation A_{rc}^2 refers to the number in the cell of the squared matrix in the r row of the c column ; $A_{1c}A_{r1}$ refers to the product of the number in the cell in the first row of the c column and the number in the r row of the first column of the unsquared matrix ; $A_{2c}A_{r2}$ refers to the product of the number in the cell in the second row of the c column and the number in the r row of the second column of the unsquared matrix, and so on.

This procedure is actually a rather simple one to carry out. The products of the corresponding cells of a row and a column will yield numbers other than zero only if a *one* appears in both cells which are being considered. Thus, if in a matrix individual 2 chooses individual 11, and 11 chooses individual 6 this will contribute a number in the row 2, column 6 cell of the squared matrix.

It may be seen from the way in which the squared matrix was obtained that each figure in this matrix represents the number of two-step connections that exist between the specified two members of the group. It is clear, of course, that this relationship need not be symmetrical, just as the one-step connections need not be symmetrical.

The numbers that appear in the diagonal of this squared matrix have a special meaning. They indicate the number of two step connections that exist from a person to himself, or, in other words, they indicate the number of mutual sociometric choices in which this person was involved.

The meaning of these two-step connections between different people can be important. For example, if the original sociometric choice indicates influence from one person to another, the squared matrix would indicate the extent of indirect influence which this person has and exactly which other people he influences indirectly. If the original sociometric data indicated channels of communication for information, the squared matrix would tell us who would hear things from whom with the information going through one intermediary.

1. The application of matrix multiplication to the analysis of sociometric patterns was developed together with Mr. Albert Perry and Mr. Duncan Luce of the Massachusetts Institute of Technology.

The cube of the matrix, which gives further information of this sort, is obtained by multiplying the original matrix by the squared matrix in the same manner as the original matrix was multiplied by itself. The formula for obtaining the values of the cells of the cubed matrix would be written similarly as :

$$A_{rc}^3 = A_{1c}A_{r1}^2 + A_{2c}A_{r2}^2 + A_{3c}A_{r3}^2 + \cdots + A_{nc}A_{rn}^2.$$

The actual calculation can be performed rather simply.

The meaning of the figures in this cubed matrix is similar to their meaning in the squared matrix. They indicate the number of three-step connections that exist between any two people. The numbers in the diagonal of the matrix now indicate the number of three-step connections from a person back to himself. The implications of these numbers in the diagonal of the cubed matrix will be elaborated on shortly.

It is apparent that these matrices may also be raised to higher powers to obtain the four-step or five-step or even more indirect connections among the members of a group. If we are concerned [156] with a question such as how many people will hear a given item of information in three or fewer steps if it is started with any particular person, the answer may be obtained by adding together the original, the squared and the cubed matrices. We can obtain such information as, who influences the greatest number of people in less than a specified number of steps, which people are influenced by the greatest number of people and which people are only subject to the influence of a few ; which people in the group are most indirectly connected to each other and how indirect is this connection ; what proportion of the possible connections among the various people actually exist. Being able to handle conveniently and efficiently these aspects of group structure and patterning of connections should make it feasible to study their effects on such processes as communication, influence, social pressures and many others.

The Determination of Cliques

The manipulation of matrices by means of raising them to the third power can, with complete accuracy, determine the existence of cliques of various sizes and with various degrees of "cliquishness." Let us begin by defining an extreme instance of clique formation within a group and then see how we may determine whether or not such cliques exist in any given structure. We will define this extreme type of clique as more than two individuals, all of whom choose each other mutually. In other words, direct one-step symmetrical connections exist between every possible pair of members of such a clique. Clearly, in order to determine the existence of such a clique

we would only concern ourselves with the symmetrical sub-matrix consisting only of mutual choices and not with the complete matrix of connections. If we raise such a symmetrical sub-matrix to the third power we will obtain all the three-step connections that exist between any two people which involve only mutual choices. What would then be the meaning of a three-step connection from a member back to himself which involves only symmetrical choices, that is, what will be the meaning of the numbers which appear in the main diagonal of this cubed symmetrical sub-matrix? Numbers will appear in the main diagonal of this cubed matrix if, and only if, there exists a clique, as defined above, within the group. If there does exist such a clique, then numbers will appear in those positions on the diagonal which correspond to those persons who are members of the clique. If only one clique exists in the group, or if more than one clique exists but they contain different members, then the number which appears in the diagonal for a particular individual will bear a given relationship to the number of people in the clique. If the clique is composed of n members, the number appearing in the diagonal for each of the members will be equal to $(n - 1)(n - 2)$. We may thus immediately determine from this cubed matrix whether or not there exist cliques, who belongs to these cliques if they exist, and how many members each clique has.

It is also possible to distinguish subgroups which are not extreme cliques. These subgroups may be defined on the basis of mutual choices or on the basis of the complete matrix. The best criteria for distinguishing subgroups of less than the extreme degree must still be determined, but it is likely that once defined, they may be relatively easily found by means of matrix multiplication.

This, of course, represents only the [157] beginning of the development of the application of matrix algebra to sociometric data. There are still a number of problems which exist and many indicated areas for extension of the development. I shall mention a few of those which we consider important.

1. Matrices which are cubed or raised to even higher powers will contain in them two kinds of chains or indirect connections among persons. One of these types of chains, which for certain purposes we should like to eliminate, we have called redundant chains. These are chains where one person appears more than once. An example would be : a-b-a-c. For coordinations of these matrices to such things as, say, influence patterns in a group, these redundant chains do not seem to have much meaning. These redundant chains can be eliminated very easily in the cubed matrix, but we have as yet found no way to eliminate them in matrices of higher power.

2. Thus far we have only concerned ourselves with matrices in which only zero or one can appear in a cell. Sociometric data are frequently gathered in two forms such as "whom do you like most?" and "whom do you dislike

most?" Such rejections are also an important aspect of social structure in a group. It would appear desirable to investigate the potentialities of matrix multiplication using matrices in which any cell can contain either a +1, 0 or -1.

3. The use of such a technique as matrix multiplication for analysis of a special kind of data immediately suggests extensions of its use. This method of analysis can clearly be applied to any data which consist essentially of relationships among a number of discrete points. In sociometric data the discrete points are people in a group, and the relationships are things like "liking" or "being together with." There are other areas in which the data have a similar form. Bavelas (1), for example, has recently published a mathematical model for dealing with topological structures such as those used by Professor Kurt Lewin. These structures consist of discrete regions having or not having a connectedness relationship. Bavelas deals with such concepts as diameter of such a structure and distance from one part of the structure to another. The technique of matrix multiplication can also clearly be used to operate on such problems.

It also becomes apparent from a consideration to what kinds of data this analytical device might be applicable that the relationships among the discrete parts need not be *either/or* relationships. Indeed, most sociometric data are in terms of *either/or* relationships only because the measurement technique is very imprecise. Certainly there are degrees of *friendship* and degrees of *liking* and degrees of *susceptibility to influence*. As yet we are completely unable to handle such data where the relationships among persons can assume continuous values. It would be useful to consider the possibilities for application of matrix algebra to such data.

It is to be hoped that empirical use of these techniques will clarify some of the problems concerning co-ordination to data and that it will stimulate sufficient interest to push forward further development rapidly.

[157]

REFERENCES

1. Bavelas, A. "A mathematical model for group structures." *Applied Anthropology*, 1948, 7, 16-30.
2. Cook, L. A. "An experimental sociographic study of a stratified 10th grade class," *Amer. Social. Rev.* 1945, 10, 250-261.
3. Forsyth, E., and Katz, L. "A matrix approach to the analysis of sociometric data : Preliminary report." *Sociometry*, 1946, 9, 340-347.
4. Northway, M. L. "A method for depicting social relationships obtained by sociometric testing." *Sociometry*, 1940, 3, 144-150.

Utiliser l'algèbre matriciel pour l'analyse des sociogrammes

Leon Festinger

Il n'existe pas, à l'heure actuelle, de dispositif analytique adéquat pour traiter les données issues de l'une des techniques de mesure les plus populaires dans le domaine de la psychologie sociale¹. Les questions sociométriques telles que « Qui sont vos meilleurs amis ? » ou « Quelles sont les personnes avec lesquelles vous aimez être le plus souvent ? » sont de plus en plus utilisées lorsqu'on s'intéresse à la « structure » ou au « schéma² » des relations entre un certain nombre de personnes. La formulation exacte des questions varie énormément selon le contexte de l'étude. Cependant, le type de données obtenues est toujours le même : la liste des personnes mentionnées en réponse à la question.

Dans l'étude sociométrique des groupes, une grande attention a été accordée à la structure exacte des interconnexions entre les individus. Des questions importantes ont été posées, telles que : qui sont les amis d'une personne donnée, quelles sont ses relations avec les amis de ses amis et quelles sont les tendances à la formation de sous-groupes ou de cliques³ dans le groupe ; combien de canaux d'information⁴ existent parmi les membres d'un groupe, sur quelle partie du groupe s'étend l'influence d'une personne, quelle est la nature des chaînes d'influence à deux ou trois étapes⁵ qui peuvent exister ; si une information entre dans un groupe, combien de personnes en entendront éventuellement parler, qui l'entendra et par qui, et à quelle distance de la source originale se trouvera-t-elle au moment où une personne spécifique en entendra parler ?

En l'absence de techniques de représentation adéquates pour traiter ces données, l'analyse des schémas exacts d'interconnexion entre les membres d'un groupe est pratiquement impossible, à moins que le groupe ne soit très

1. Cette affirmation est surprenante et passe sous silence l'article *Statistics of social configurations* de Moreno et Jennings (1938) ou celui de Bavelas, *A Mathematical Model for Group Structures*, (1948) que cite pourtant Festinger, Bavelas étant un de ses collègues au MIT.

2. Le terme *pattern* a été traduit ici et dans la suite du texte par schéma pour évoquer la configuration des liens du réseau social.

3. Le terme clique est utilisé dans ce texte au sens sociologique du terme et non au sens mathématique (sous-graphe maximal complet) ; cette dernière est définie cette même année dans l'article de Luce et Perry, *A method of matrix analysis of group structure*, *Psychometrika*, 14 (2) : 95–116 - article issu d'un travail dirigé par Festinger.

4. L'expression originale est *paths of influence*.

5. Il s'agit ici de la longueur des chemins au sein d'un groupe de personnes mais il est difficile de ne pas voir dans ce *two or three step influence chains* une allusion au *two-step flow model* proposé par Lazarsfeld *et al.* dans l'ouvrage *The People's Choice* paru l'année précédente.

petit. Plus la taille du groupe augmente, plus la complexité du schéma le rend généralement extrêmement difficile à comprendre par simple inspection visuelle. Il en résulte que ce type d'analyse a été relativement négligé. Les chercheurs se sont, dans l'ensemble, contentés d'analyser les structures sociométriques en calculant le nombre de choix offerts aux personnes, les types de personnes qui obtiennent le plus de choix, la proportion de choix à l'intérieur du groupe et d'autres mesures globales de ce genre, mesures servant à relier les choix sociométriques à d'autres variables.

Les premières tentatives d'analyse et de description du schéma exact des interconnexions au sein d'un groupe ont pris la forme de diagrammes compliqués où les relations entre les personnes étaient représentées par des lignes et des flèches. Ces diagrammes pouvaient ensuite être examinés. Il est facile de comprendre qu'un tel diagramme serait très difficile à manier si le nombre de membres du groupe était important ou si le nombre de choix effectués par chaque membre était considérable. Il n'existe pas de règles pour manipuler efficacement de tels diagrammes. On se contente de les arranger par tâtonnement afin que le diagramme paraisse aussi simple que possible puis on inspecte et on examine encore avec l'espoir d'être assez attentif pour voir ce qu'il faut voir ⁶.

C'est la prise en compte de ces difficultés qui a conduit des personnes comme Northway (4) et Cook (2) à tenter de formuler des règles de construction du sociogramme de manière à le rendre plus facilement compréhensible. Cependant, ces tentatives demeuraient vagues et relativement peu systématiques. Forsyth et Katz (3) ont fait un grand pas en avant en suggérant l'utilisation d'une matrice et de certaines manipulations de l'algèbre matricielle pour l'analyse des modèles sociométriques. Leur idée était de représenter le sociogramme sous forme de matrice puis de réarranger cette matrice selon certains principes afin qu'elle révèle les sous-groupes présents. La présentation matricielle s'effectue simplement en ordonnant de la même façon les individus du groupe en lignes et en colonnes. Les choix que fait un individu sont ensuite indiqués par le chiffre 1 dans l'une des cases, de sorte que la ligne corresponde à la personne qui fait le choix et la colonne à la personne qui reçoit le choix ⁷. Les cases situées le long de la diagonale principale de la matrice sont bien sûr laissées en blanc puisque la personne n'est pas considérée comme s'étant choisie elle-même. En regardant n'importe quelle ligne, on voit qui a été choisi par cette personne et en regardant n'importe quelle colonne, on voit qui a choisi cette personne.

6. La critique concernant tant la construction que la lisibilité des sociogrammes est beaucoup plus nette que celle de Forsyth et L. Katz en 1946 (cf *A Matrix Approach to the Analysis of Sociometric Data*); il est vrai que ces derniers publiaient dans la revue de Moreno et avaient donc une liberté de parole moindre.

7. Que ce soit dans cet article ou dans celui de Forsyth et Katz, le terme de matrice d'adjacence n'est pas utilisé.

La présentation sous cette forme n'offre que peu ou pas d'avantages par rapport au sociogramme pour une inspection visuelle. Certaines manipulations de la matrice permettent cependant de la simplifier. La suggestion de Forsyth et Katz est de réorganiser l'ordre des membres du groupe de façon à ce que les nombres dans la matrice se regroupent aussi précisément que possible le long de la diagonale. Plus concrètement, ils suggèrent de réorganiser l'ordre des lignes et des colonnes de manière à minimiser le carré des écarts perpendiculaires des nombres par rapport à la diagonale de la matrice. Ce réarrangement regroupe les personnes qui se choisissent fréquemment et sépare celles qui ne se choisissent pas. Il nous permet ainsi de séparer les sous-groupes existants et semble être un moyen pertinent d'analyser les changements dans le schéma des relations au sein d'un groupe. Ces changements se traduiraient par un classement différent des membres du groupe dans la matrice finale. Cette méthode ne semble pas être très utile pour comparer deux groupes différents et le travail nécessaire pour obtenir la disposition de la matrice qui regroupera les choix le plus précisément le long de la diagonale est assez fastidieux.

Il existe d'autres aspects de l'algèbre matricielle qui semblent utiles pour l'analyse de ces structures de relations entre les membres d'un groupe. Lorsque le sociogramme est présenté sous forme de matrice, une analyse de certains aspects de la structure du groupe peut être effectuée par les moyens relativement simples que sont l'élevation au carré et au cube de cette matrice⁸. La matrice élevée au carré s'obtient assez facilement de la manière suivante : pour obtenir le nombre qui entre dans la cellule correspondant à la colonne c et à la ligne r de la matrice au carré, nous multiplions chaque cellule de la colonne c de la matrice d'origine par la cellule correspondante de la ligne r puis nous additionnons ces produits. L'équation générale de cette multiplication peut s'écrire comme suit :

$$A_{rc}^2 = A_{1c}A_{r1} + A_{2c}A_{r2} + A_{3c}A_{r3} + \dots + A_{nc}A_{rn}.$$

Dans cette équation, A_{rc}^2 désigne le nombre dans la cellule de la matrice élevée au carré à la ligne r de la colonne c ; $A_{1c}A_{r1}$ désigne le produit du nombre dans la cellule à la première ligne de la colonne c et du nombre à la ligne r de la première colonne de la matrice non élevée au carré; $A_{2c}A_{r2}$ désigne le produit du nombre dans la cellule de la deuxième ligne de la colonne c et du nombre dans la ligne r de la deuxième colonne de la matrice non élevée au carré et ainsi de suite.

Cette procédure est en fait assez simple à mettre en œuvre. Les produits

8. L'application de la multiplication matricielle à l'analyse des schémas sociométriques a été développée en collaboration avec M. Albert Perry et M. Duncan Luce du Massachusetts Institute of Technology.

des cellules correspondantes d'une ligne et d'une colonne ne donneront des nombres différents de zéro que si un 1 apparaît dans les deux cellules considérées. Ainsi, si dans une matrice l'individu 2 choisit l'individu 11 et que 11 choisit l'individu 6, cela donnera un nombre dans la cellule de la ligne 2, colonne 6 de la matrice élevée au carré.

La façon dont la matrice au carré a été obtenue permet de constater que chaque chiffre de cette matrice représente le nombre de relations à deux étapes qui existent entre deux membres du groupe⁹. Il est évident, bien sûr, que cette relation ne doit pas nécessairement être symétrique, tout comme les relations à une étape ne doivent pas nécessairement être symétriques.

Les chiffres qui apparaissent dans la diagonale de cette matrice au carré ont une signification particulière. Ils indiquent le nombre de relations à deux étapes qui existent entre une personne et elle-même ou, en d'autres termes, ils indiquent le nombre de choix sociométriques mutuels dans lesquels cette personne a été impliquée.

La signification de ces relations en deux étapes entre différentes personnes peut être importante. Par exemple, si le choix sociométrique original indique l'influence d'une personne sur une autre, la matrice au carré mesure l'étendue de l'influence indirecte de cette personne et identifie les personnes influencées indirectement. Si les données sociométriques originales indiquent des canaux de communication pour l'information, la matrice au carré nous dit comment l'information circule *via* un intermédiaire.

Le cube de la matrice donne d'autres informations de ce type, il est obtenu en multipliant la matrice d'origine par la matrice au carré de la même manière que la matrice d'origine a été multipliée par elle-même. La formule permettant d'obtenir les valeurs des cellules de la matrice élevée au cube s'écrit de la manière suivante :

$$A_{rc}^3 = A_{1c}A_{r1}^2 + A_{2c}A_{r2}^2 + A_{3c}A_{r3}^2 + \dots + A_{nc}A_{rn}^2.$$

Le calcul peut être effectué assez simplement.

La signification des chiffres de cette matrice élevée au cube est similaire à leur signification dans la matrice élevée au carré. Ils indiquent le nombre de relations à trois étapes qui existent entre deux personnes. Les chiffres de la diagonale de la matrice indiquent maintenant le nombre de relations à trois étapes d'une personne vers elle-même. Les implications de ces nombres dans la diagonale de la matrice élevée au cube seront développées prochainement.

Il est évident que ces matrices peuvent également être élevées à des puissances plus grandes pour obtenir les quatre ou cinq étapes de relations encore

9. On parlerait aujourd'hui du nombre de chemins de longueur 2.

plus indirectes entre les membres d'un groupe. Si l'on cherche à savoir combien de personnes entendront un élément d'information donné en trois étapes ou moins si l'on commence par une personne donnée, la réponse peut être obtenue en additionnant la matrice d'origine, celle élevée au carré et celle élevée au cube. Nous pouvons obtenir des informations telles que : qui influence le plus grand nombre de personnes en moins d'un nombre déterminé d'étapes, quelles personnes sont influencées par le plus grand nombre de personnes et quelles personnes ne sont soumises qu'à l'influence de quelques-unes ; quelles personnes du groupe sont le plus indirectement reliées entre elles et dans quelle mesure cette relation est indirecte ; quelle proportion de relations possibles entre les différentes personnes existe réellement. Le fait de pouvoir traiter de manière concrète et efficace ces aspects de la structure du groupe et de la configuration des relations devrait permettre d'étudier leurs effets sur des processus tels que la communication, l'influence, les pressions sociales et bien d'autres.

La détermination des cliques

La manipulation des matrices élevées à la puissance trois permet de déterminer de manière précise l'existence de cliques de différentes tailles et avec différents degrés de « cliquitude ». Commençons par définir un cas extrême de formation de cliques au sein d'un groupe et voyons ensuite comment nous pouvons déterminer si de telles cliques existent ou non dans une structure donnée. Nous définirons ce type extrême de clique comme étant un ensemble de plus de deux individus qui se choisissent tous mutuellement. En d'autres termes, des relations symétriques directes à une étape existent entre chaque paire possible de membres d'une telle clique¹⁰. Il est clair que pour déterminer l'existence d'une telle clique, nous ne devons nous préoccuper que de la sous-matrice symétrique constituée uniquement de choix mutuels et non de la matrice complète des relations. Si nous élevons une telle sous-matrice symétrique à la puissance trois, nous obtiendrons toutes les relations à trois étapes qui existent entre deux personnes quelconques et qui n'impliquent que des choix mutuels. Quelle serait alors la signification d'une relation à trois étapes d'un membre vers lui-même qui n'implique que des choix symétriques, c'est-à-dire quelle serait la signification des nombres qui apparaissent dans la diagonale principale de cette sous-matrice symétrique élevée au cube ? Les nombres apparaîtront dans la diagonale principale de cette matrice élevée au cube si et seulement si il existe une clique, telle que définie ci-dessus, au sein du groupe. S'il existe une telle clique, les nombres apparaîtront dans les positions de la diagonale qui correspondent aux personnes qui sont membres de la clique. S'il n'existe qu'une seule clique dans le groupe, ou s'il en existe

10. Cette définition de clique « extrême » correspond de fait à la définition mathématique d'une clique.

plusieurs mais qu'elles contiennent des membres différents, le nombre qui apparaît dans la diagonale pour un individu donné aura une relation spécifique avec le nombre de personnes de la clique. Si la clique est composée de n membres, le nombre apparaissant dans la diagonale pour chacun des membres sera égal à $(n - 1)(n - 2)$. Nous pouvons donc immédiatement déterminer à partir de cette matrice élevée au cube s'il existe ou non des cliques, qui appartient à ces cliques si elles existent, et combien de membres compte chaque clique.

Il est également possible de distinguer des sous-groupes qui ne sont pas des cliques extrêmes. Ces sous-groupes peuvent être définis sur la base de choix mutuels ou sur la base de la matrice complète. Les meilleurs critères pour distinguer les sous-groupes de degré inférieur à l'extrême doivent encore être déterminés mais il est probable qu'une fois définis, ils puissent être trouvés relativement facilement au moyen de la multiplication de la matrice.

Ceci, bien sûr, ne représente que le début du développement de l'application de l'algèbre matricielle aux données sociométriques. Il y a encore un certain nombre de problèmes qui se posent et de nombreux développements sont envisageables. Je vais mentionner quelques-uns de ceux qui nous semblent importants.

1. Les matrices élevées au cube ou à des puissances plus élevées contiennent deux types de chaînes ou de relations indirectes entre les personnes. Le premier type de chaînes, que nous voudrions éliminer dans certaines cas, nous l'appelons chaînes redondantes. Il s'agit de chaînes où une personne apparaît plus d'une fois. Un exemple serait : a-b-a-c. Ces chaînes redondantes ne semblent pas avoir beaucoup de sens si l'on souhaite, par exemple, étudier les schémas d'influences dans un groupe. Ces chaînes redondantes peuvent être éliminées très facilement dans la matrice élevée au cube mais nous n'avons pas encore trouvé le moyen de les éliminer dans les matrices de puissance supérieure.

2. Jusqu'à présent, nous ne nous sommes intéressés qu'aux matrices dans lesquelles seul un 0 ou un 1 peut apparaître dans une cellule. Les données sociométriques sont fréquemment recueillies sous deux formes telles que « qui aimez-vous le plus ? » et « qui aimez-vous le moins ? ». Ces rejets constituent également un aspect important de la structure sociale d'un groupe. Il semble souhaitable d'étudier les potentialités de la multiplication matricielle en utilisant des matrices dans lesquelles toute cellule peut contenir +1, 0 ou -1.

3. L'utilisation d'une technique telle que la multiplication matricielle pour l'analyse d'un type particulier de données suggère immédiatement des extensions de son utilisation. Cette méthode d'analyse peut manifestement être appliquée à toutes les données qui consistent essentiellement en des relations

entre un certain nombre de points discrets. Dans les données sociométriques, les points discrets sont des personnes dans un groupe et les relations sont de type « aimer » ou « être ensemble ». Il existe d'autres domaines dans lesquels les données ont une forme similaire. Bavelas (1), par exemple, a récemment publié un modèle mathématique permettant de traiter des structures topologiques telles que celles utilisées par le professeur Kurt Lewin. Ces structures consistent en des régions discrètes ayant ou non une relation de connexité. Bavelas traite de concepts tels que le diamètre d'une telle structure et la distance d'une partie de la structure à une autre. La technique de la multiplication matricielle peut clairement être utilisée pour traiter de tels problèmes.

En examinant les types de données auxquelles ce dispositif analytique pourrait s'appliquer, il devient également évident que les relations entre les unités ne doivent pas nécessairement être des relations *soit/ou*. En effet, la plupart des données sociométriques sont exprimées en termes de relations *soit/ou* uniquement parce que la technique de mesure est très imprécise. Il existe certainement des degrés d'*amitié*, des degrés d'*sympathie* et des degrés d'*sensibilité à l'influence*. Jusqu'à présent, nous sommes totalement incapables de traiter de telles données où les relations entre les personnes peuvent prendre des valeurs continues. Il serait utile d'examiner les possibilités d'application de l'algèbre matricielle à de telles données ¹¹.

Nous espérons que l'utilisation empirique de ces techniques clarifiera certains des problèmes relatifs aux données et qu'elle suscitera un intérêt suffisant pour assurer leur développement rapide.

RÉFÉRENCES ¹²

1. Bavelas, A. "A mathematical model for group structures." *Applied Anthropology*, 1948, 7, 16-30.
2. Cook, L. A. "An experimental sociographic study of a stratified 10th grade class," *Amer. Social. Rev.* 1945, 10, 250-261.
3. Forsyth, E., and Katz, L. "A matrix approach to the analysis of sociometric data : Preliminary report." *Sociometry*, 1946, 9, 340-347.
4. Northway, M. L. "A method for depicting social relationships obtained by sociometric testing." *Sociometry*, 1940, 3, 144-150.

11. Aujourd'hui encore, les méthodes d'analyse de réseau les plus courantes s'appliquent à des relations de type 1-0.

12. Il n'échappera aux plus fidèles lecteurs et lectrices de la collection « textes » que trois de ces textes ont déjà été traduits ici.

groupe f.m.r.



La collection « [textes](#) » du groupe fmr (flux, matrices, réseaux) propose des rééditions bilingues d'articles consacrés à l'analyse de réseaux.

Disponibles en ligne

- L. Beauguitte, P. Beauguitte et P. Gourdon, 2021, « [William L. Garrison, 1960, Connectivity of the Interstate Highway System](#) ».
- L. Beauguitte et M. Maisonobe, 2021, « [Joseph B. Kruskal, 1956, On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem](#) ».
- L. Beauguitte et M. Maisonobe, 2021, « [E.W. Dijkstra, 1959, A Note on Two Problems in Connexion with Graphs](#) ».
- L. Beauguitte, 2021, « [Alfonso Shimbel, 1951, Application of Matrix Algebra to Communications Nets](#) ».
- L. Beauguitte, 2021, « [Alfonso Shimbel, 1953, Structural Parameters of Communication Networks](#) ».
- L. Beauguitte, 2021, « [Z. Prihar, 1956, Topological Properties of Telecommunication Networks](#) ».
- L. Beauguitte et M. Maisonobe, 2021, « [Alex Bavelas, 1948, A Mathematical Model for Group Structures](#) ».
- L. Beauguitte et J. Fen-Chong, 2021, « [Alex Bavelas, 1950, Communication Patterns in Task-Oriented Groups](#) ».
- L. Beauguitte, 2021, « [Forrest R. Pitts, 1965, A Graph Theoric Approach to Historical Geography](#) ».
- L. Beauguitte, 2021, « [M. L. Northway, 1940, A method for depicting social relationships obtained by sociometric testing](#) ».
- F. Bahoken et L. Beauguitte, 2021, « [John D. Nystuen and Michael F. Dacey, 1961, A Graph Theory Interpretation of Nodal Regions](#) ».
- F. Bahoken et L. Beauguitte, 2022, « [Edward L. Ullman, 1949, Mapping the world's ocean trade: A research proposal](#) ».
- L. et P. Beauguitte, 2022, « [F. L. Hitchcock, 1941, The Distribution of a Product from Several Sources to Numerous Localities](#) ».
- F. Bahoken, 2022, « [W. R. Tobler, 1969, Review of Sémiologie graphique: Les Diagrammes – Les réseaux – Les Cartes](#) ».
- F. Bahoken et L. Beauguitte, 2022, « [P. R. Gould, 1967, On the Geographical Interpretation of Eigenvalues](#) ».
- L. Beauguitte, F. Bahoken et J. Fen-Chong, 2022, « [E. Forsyth et L. Katz, 1946, A Matrix Approach to the Analysis of Sociometric Data](#) ».
- L. Beauguitte et F. Bahoken, 2023, « [J.L. Moreno et H.H. Jennings, 1938, Statistics of social configurations](#) ».

- L. Beauguitte, 2023, « [G. D. Dantzig, 1960, On the Shortest Route Through a Network](#) ».
- L. Beauguitte, 2023, « [K. J. Kansky, 1963, Structure of Transportation Networks](#) ».