Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: https://www.tandfonline.com/loi/hmbr20

Scaling Sociomatrices by Optimizing an Explicit Function: Correspondence Analysis of Binary Single Response Sociomatrices

Elliot Noma & D. Randall Smith

Multivariate Behavioral Research, 1985, 20, 179–197

Scaling Sociomatrices by Optimizing an Explicit Function: Correspondence Analysis of Binary Single Response Sociomatrices

Elliot Noma and D. Randall Smith
Rutgers—the State University of New Jersey

Most methods for detecting structure in sociometric data involve either continuous spatial representations (e.g. MDS) or discrete hierarchical clustering analysis (e.g. CONCOR). By producing either spatial or clustering representations, these methods can highlight only some of the theoretically interesting group structures. Correspondence analysis, in contrast, can provide either spatial or clustering representations by assigning spatial coordinates to minimize the distance between individuals linked by a sociometric relationship. These scales may then be used to identify individuals' locations in a multidimensional representation of a group's structure or to reorder the rows and columns of a sociomatrix. Unlike many other methods of sociometric analysis, the numerical methods of correspondence analysis also are well understood and the optimization of the goodness-of-fit measure allows an evaluation of a particular model of group structure.

Inferring social structure (e.g. cliques or hierarchies) from individual level relationships is a long standing problem in network analysis. In general, the goal is the manipulation of a *sociomatrix*, A, (in which $a_{ij} = 1$ if i is operationally linked to j) to reveal the underlying structure. In some methods, the revealed structure is presented geometrically in a *sociogram* (in which individuals are identified as points in a space and the links between individuals are represented by arrows connecting these points). Other methods permute or partition the rows and columns of the sociomatrix to highlight the underlying structure. As Roistacher (1974) notes, these are substantively different approaches to social structure—sociograms emphasize spatial aspects of the structure while sociomatrices emphasize the person-to-person links.

In this paper we examine the strengths and weaknesses of three families of sociometric analysis: spatial models, cluster analysis, and the row-column reordering of the sociomatrix. To facilitate this discussion we have chosen representative techniques from each of these families: multidimensional scaling of the transformed sociomatrix (spatial models), blockmodeling (a cluster analysis, White, Boorman, & Breiger, 1976), and the row-column permutation of the sociomatrix

We would like to thank Joel Levine and several anonymous reviewers for helpful criticisms of an earlier draft. Send all correspondence to: Elliot Noma, Department of Psychology, 619 Tillett Hall, Rutgers—the State University of New Jersey, New Brunswick, N.J. 08903.

(Katz, 1947). Next we compare these methods to correspondence analysis, a method that both creates sociograms *and* orders the rows and columns in a sociomatrix. We argue that methods specific to spatial-modeling, clustering, or row-column reordering may occasionally fail to highlight some theoretically-interesting social structures. By contrast, correspondence analysis has the advantages of all three approaches, but does not share their conceptual limitations. The procedure optimizes an explicit criterion and is capable of yielding either spatial coordinates for individuals or partitions on each dimension. The algorithm is computationally easy to implement and does not implicitly treat nonchoices ($a_{ij} = 0$) as rejections. Initially, only the case of binary positive affect choices will be scaled. Later, the procedure will be applied to non-positive affect data.

Methods for Representing Sociometric Structure

Spatial Models

Placing individuals as points in space can highlight theoretically important structures such as subgroups, hierarchies, liaisons, and isolates. The representation of individuals as points in two or more dimensions not only permits the visual identification of structure, but also suggests a multilevel interpretation of the structure (e.g. hierarchies within subgroups). Originally, sociograms were an ad hoc placement of individuals in space (Moreno, 1934). With the advent of more systematic placement methods, such as multidimensional scaling (e.g., Kruskal, 1964a,b), two questions arise about the use of spatial modeling methods. First, are spatial representations the best method for highlighting important social structure, and second, what assumptions are made when applying any particular scaling algorithm.

For the identification of structure, Breiger, Boorman, and Arabie (1975) find that scaling "leads to far more interpretable results than do classical techniques like factor analysis," but conclude that scaling procedures are not "distinctive" in their ability to "feedback to underlying relational data." Multidimensional scaling could feedback to the underlying relational data if it were possible to use the scale values to reorder the rows and columns of the sociomatrix. A scaling in one dimension clearly can be used in this way, even though this one-dimensional solution usually will not nest within a multidimensional scaling of the same data using two or more dimensions. However, for solutions in more than one dimension, there is no uniquely specified ordering (without further assumptions). In fact, in a Euclidean space, there are an infinite number of orderings, each corresponding to a projection of all points onto one-dimensional lines through the space. Any of these orderings could then be used to permute the rows and columns of the matrix and there is little basis to choose one ordering over another. Even in a general Minkowski space, one still must decide which axis through the space is to be used to reorder the rows and columns of the matrix. Kendall (1971) has proposed a method around this problem by using a minimal spanning path to define an ordering of points in a two dimensional space. However, Kendall's HORSHO analysis may not generate a unique ordering of points. Therefore the multidimensional scaling results cannot be used straightforwardly to rewrite the sociomatrix.

The application of multidimensional scaling also requires additional assumptions about social interaction. Multidimensional scaling is a general method of analysis and therefore requires that the sociometric data be suitably prepared prior to analysis: for example, Breiger et al. (1975) scale the columnwise correlations of the sociomatrix, while Laumann and Pappi (1973) scale a matrix of path distances between individuals. Unfortunately, this preprocessing of the data is based on additional assumptions about the nature of the social interactions (Sailer, 1978). For example, using the columnwise correlation as a measure of individual-to-individual similarity means that zeroes (lack of a relation between row and column individuals) are treated as indicators of dissimilarity and thus the scaling will tend to separate unlinked individuals. However, Coombs (1964) and Levine (1979) argue that a zero represents only a non-chosen alternative and not a rejection. This asymmetry between choices and non-choices (as opposed to rejections) argues against the symmetric handling of zeroes and ones. In addition, preprocessing of the data causes an inevitable loss of information (Noma, 1984). This loss of information obscures the relationship between the criterion to be optimized and row-column clustering and reordering of the sociomatrix. Therefore, the necessary preprocessing of the data prior to scaling may introduce implicit (and often undesirable) assumptions.

Cluster Analysis

The clustering of individuals is a commonly used method for sociometric analysis. One popular approach to clustering is the blockmodeling methods developed by White and his colleagues (Boorman & White, 1976; Breiger et al., 1975; Heil & White, 1976; White et al., 1976). Blockmodeling's strength is its strong theoretical foundation in the identification of individuals with identical patterns of interactions (Lorrain & White, 1976). Blockmodeling attempts to cluster such structurally equivalent individuals. The confirmatory technique (BLOCKER) is based on the optimization of zero blocks (see White et al., 1976) and allows for strong tests of the theory of structural equivalence. The exploratory analog to BLOCKER is the CONCOR algorithm. CONCOR computes the product moment correlations between columns (or rows) of the sociomatrix, and iteratively creates a series of matrices by correlating columns of the preceding correlation matrix. CONCOR usually converges to a matrix containing submatrices entirely of +1.0's and −1.0's and thus identifies a bifurcation of the individuals defining the rows and columns. Partitions may be further subdivided by reapplying CONCOR to each part of the blocked sociometric data.

While CONCOR has been used successfully in numerous applications, the mathematical foundation is not well understood. In practice, convergence is usually achieved, but Schwartz (1977) proves that convergence does not occur for such substantively interesting structures as triangles and rectangles. Also, our lack of knowledge of the optimization criterion makes a complete analysis of the method impossible.

Row-Column Reordering of the Sociomatrix

A third approach to detecting structure has the combined advantages of optimizing a specific criterion, operating directly on the sociomatrix and conceptual simplicity. The matrix is systematically permuted to "maximize the concentration of positive choices about the main diagonal" (Katz, 1947) so that existing subgroups are more evident. Forsythe and Katz (1946) do this by minimizing the distance between individuals linked by friendship choice:

$$[1] \qquad \sum_{i} \sum_{i} t_{ij} (i - j)^2$$

where T is the reordered sociomatrix that minimizes equation [1] and i and j are the ranks of individuals. This approach has the important attribute of treating zeroes as 'missing' data and not as the rejection of an individual.

Methods were later developed for optimally permuting rows and columns of the sociomatrix to minimize equation [1] (Beum & Brundage, 1950; Coleman & MacRae, 1960; Spilerman, 1966). Beum and Brundage's (1950) procedure is the most systematic and seemed to offer the greatest chance of consistently obtaining "good" solutions, but Deutsch and Martin (1971) show that the method may arrive at different stable solutions for different initial configurations of the row-column order. Also, there are matrices in which the algorithm cycles through several orderings rather than converging to a single optimal order. Thus the consistent minimization of equation [1] is unresolved.

Each of the three families of methods has particular strengths for highlighting social structure. Each also has certain weaknesses. Multidimensional scaling requires additional assumptions during data preparation; without further assumptions the method cannot be used to redraw the sociomatrix. Blockmodeling is inherently nonspatial and therefore may have difficulty revealing the global aspects of the data. CONCOR, the most common clustering method for sociometric analysis, has computational properties that are not well understood. Row-column reordering also is unable to reveal global aspects of the data and the methods for deriving solutions can be computationally inadequate.

Correspondence analysis overcomes many of the disadvantages inherent in Katz's approach, MDS, and CONCOR, while maintaining the advantages of each. The method is based on an explicit model of social interaction which locates individuals as points in space (sociogram) as well as reordering the rows and columns of the matrix. On each dimension, the technique bifurcates the points of the sociomatrix, thus producing subgroups in the same manner as CONCOR. (Schwartz, 1977 notes a possible connection between the initial bifurcation using CONCOR and the first eigenvector of the sociomatrix.) Also, the eigenvector-eigenvalue computations required for correspondence analysis are well understood.

Correspondence Analysis

One can interpret the Forsythe and Katz (1946) method as an optimal assignment of consecutive integers to the rows and columns of a sociomatrix. However, restricting the scale values of each individual to consecutive integers may not be the best representation of the data. Alternatively, the consecutive integer requirement can be relaxed to allow scale values to be any real number. That is, one can determine values x_i and x_j for individuals i and j to minimize the squared distance from the main diagonal to each link in the reordered matrix:

$$[2] \qquad \sum_i \sum_j a_{ij} (x_i - x_j)^2$$

where A is the original sociomatrix. Note that equation [2] is nearly identical to equation [1]. Once one obtains scale values, x_i, the rows and columns of the original sociomatrix can be permuted by a rank order of the scale values or the sociomatrix can be redrawn by spacing the rows and columns according to the scale values. The range of values in the representation is important, so the following constraint is added:

$$[3] \qquad \sum_i \sum_j a_{ij} (x_i^2 + x_j^2) = 1$$

This constraint fixes the size of the picture by specifying that the sum of the squared distance of each item (linked by a sociometric choice) to the zero point is a constant. Nishisato (1980) and Heiser (1981, chapter 3) prove that equation [2] may be minimized with the constraint (equation [3]) by applying standard eigenvector-eigenvalue decomposition to a matrix derived from the original sociomatrix. (In the analysis of square sociomatrices, the eigenvalues range from -1 to $+1$.)

The value of equation [2] is equal to one minus the eigenvalue for the particular eigenvector. Therefore, the eigenvalues can be interpreted as goodness-of-fit measures: eigenvectors with eigenvalues near one (equation [2] equal to 0) are the best-fit solutions. Aside from the trivial first eigenvalue of 1, which places all actors in one coordinate (see Nishisato, 1980), eigenvectors with the largest eigenvalues minimize the distance between individuals linked by a sociometric relationship and thus yield the best spatial representation of the group structure. For attracting relations (e.g. friendship) the most representative eigenvectors are associated with the largest eigenvalues.

Note that this derivation of correspondence analysis is based on minimizing interpoint distances, rather than the more standard maximization of a correlation ratio (see Guttman, 1941). Distance minimizing and correlation maximizing are computationally identical and permit the same graphic and tabular representations of the scale values. However, in distance minimization, the eigenvalues are equal to one minus the sum of the squared interpoint distances for linked individuals (one minus the value of equation [2]). Since the method processes the sociomatrix as a square, on-diagonal matrix, the range of eigenvalues is -1 to $+1$ (see Noma, 1982b), making the values for equation [2] range from 0 to 2. When equation [2] equals 0, all linked individuals are placed at a distance zero from each other (corresponding to an eigenvalue of $+1$). When equation [2] equals 2, linked individuals are maximally separated (an eigenvalue of -1).

Correspondence analysis provides a model for sociometric data that is derived directly from the adjacency matrix, A, and does not require preprocessing of this matrix. Unlike MDS, there is a unique orientation and ranking of nested dimensions. That is, the second dimension of a three dimensional solution is identical to the second dimension of a four dimensional solution, and so forth. (As in a principal components analysis, the axes may be rotated [Noma, 1982a], but the original, function-minimizing eigenvectors point in specific directions through the space—unlike Euclidean distance MDS and classical factor analysis, there is no rotational indeterminancy problem.) This permits a sequential analysis of dimensions starting with those associated with the largest eigenvalues. Also, the existence of a fixed zero point on each dimension (save the trivial first dimension) is particularly useful. Each dimension bifurcates the set of actors in the same manner that CONCOR converges to a matrix of plus and minus 1.0's. This permits a unique partitioning of points according to their positive or negative values on a dimension from which a blocking may be derived.

Sampson's Monastery Study

Scaling a Positive Affect Sociomatrix

Sampson's (1969) study of the social structure within an American monastery is achieving the status of an illustrative data set because of his detailed account of the interpersonal relationships which provides a benchmark for evaluating the results of statistical techniques. As a result, the several sociomatrices in his study have been blockmodeled (White et al., 1976), scaled (Breiger et al., 1975), and used for the investigation of the underlying role structure (Boorman & White, 1976).

The 12-month study was divided into five time periods with each respondent retrospectively judging all others on affect, esteem, influence, and sanctioning at the end of each period. Respondents ranked their first three choices on each of the four classes of relations. For instance, they were asked to "list those three brothers whom you personally *liked the most*." The fourth time period is of special interest since these were the last data collected before the 18 members of the monastery were reduced to six due to expulsions and resignations.

A binary matrix was created from the four positive choice sociomatrices (as reported in White et al., 1976) of the fourth time period. In this matrix, $a_{ij} = 1$ if individual i ranks j as first or second for at least one of the four classes of relations; $a_{ij} = 0$ otherwise. The matrix resulting from this procedure is shown in Table 1.

This matrix of relationships was analyzed in the present study by using correspondence analysis. As noted previously, each eigenvector permits a redrawing of the sociomatrix by placing the rows and columns at the individual scale values (coordinates on that dimension). Figure 1 shows the sociomatrix with the rows and columns reordered and spaced by the scale values for the second eigenvector ($\lambda = .793$)

Table 1

Positive Affect Matrix from Period four of Sampson's Monastery Study: $a_{ij}=1$ if individual i ranks j as first or second for at least one of the four classes of positive relations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1		1	1					1				1		1				
2	1							1				1						
3	1												1					1
4					1					1	1							
5				1							1							
6				1					1		1							
7		1														1		
8	1			1		1					1							
9	1	1		1				1				1						
10				1									1					
11				1	1			1	1									
12	1	1																
13					1		1				1							1
14	1	1													1			
15		1										1		1				
16		1					1								1			
17			1															1
18		1	1														1	

Table 2

Negative Affect Matrix from Period four of Sampson's Monastery Study: $a_{ij}=1$ if individual i ranks j as first or second for at least one of the four classes of negative relations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1			1				1			1	1	1						1
2			1	1				1					1					
3				1		1						1						
4	1											1		1				
5																	1	1
6			1				1										1	
7			1	1		1		1										
8			1									1	1			1	1	1
9			1														1	1
10																		
11			1				1					1					1	1
12														1				1
13						1						1					1	1
14				1				1			1		1					
15			1	1									1					
16			1	1							1						1	1
17				1		1		1										
18				1		1		1		1	1							

with a 1 indicating that the row individual picked the column individual and numbers greater than 1 indicating several links between multiple individuals located on the same row or column (e.g. individuals 18 and 12 both picked individual 2 in Figure 1). The second dimension bifurcates the group by setting off brothers 13, 9, 8, 6, 11,

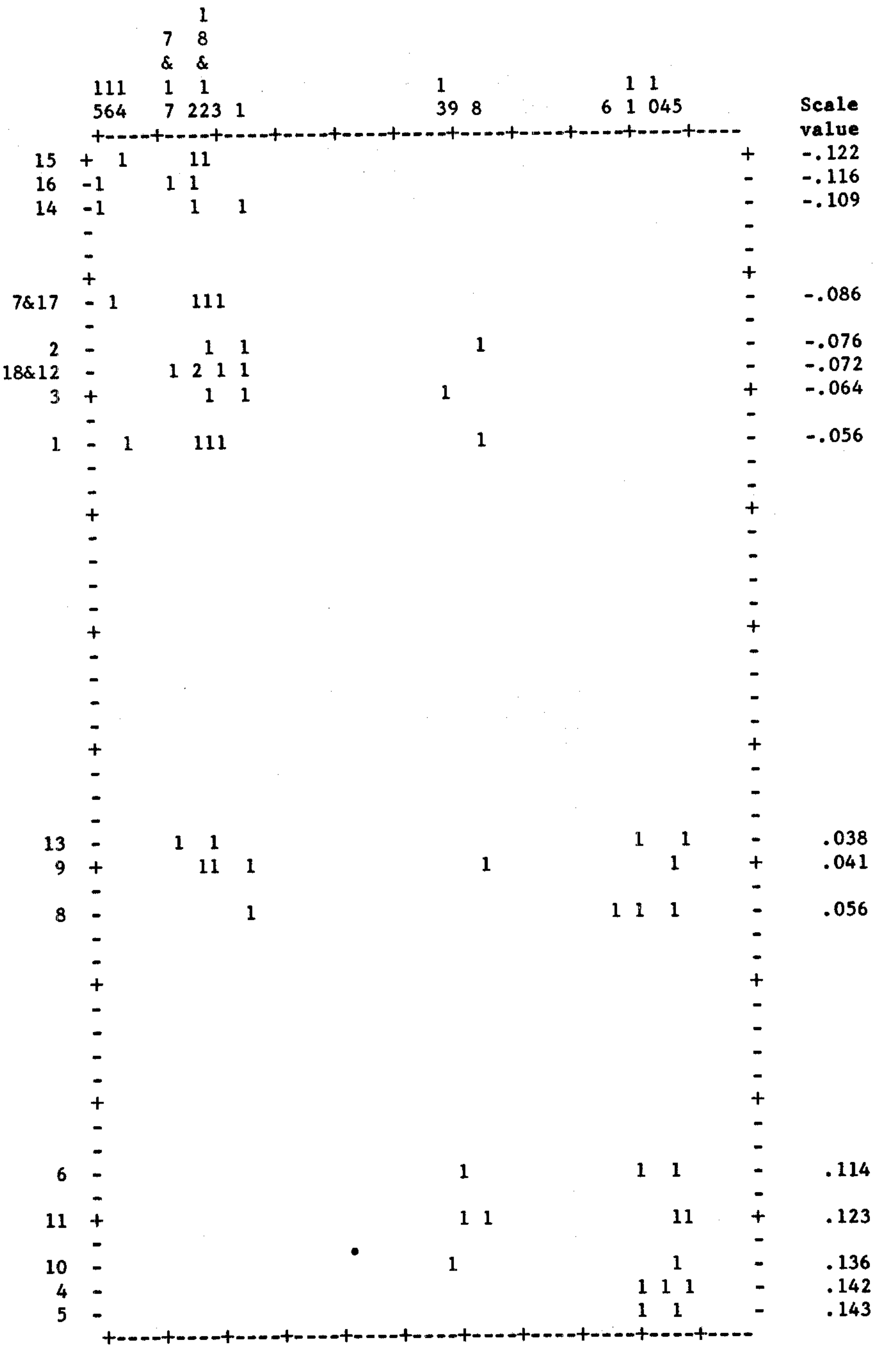

Figure 1.
Positive sociomatrix from period four of Sampson's Monastery Study with rows and columns spaced according to the *second* eigenvector of the correspondence analysis solution.

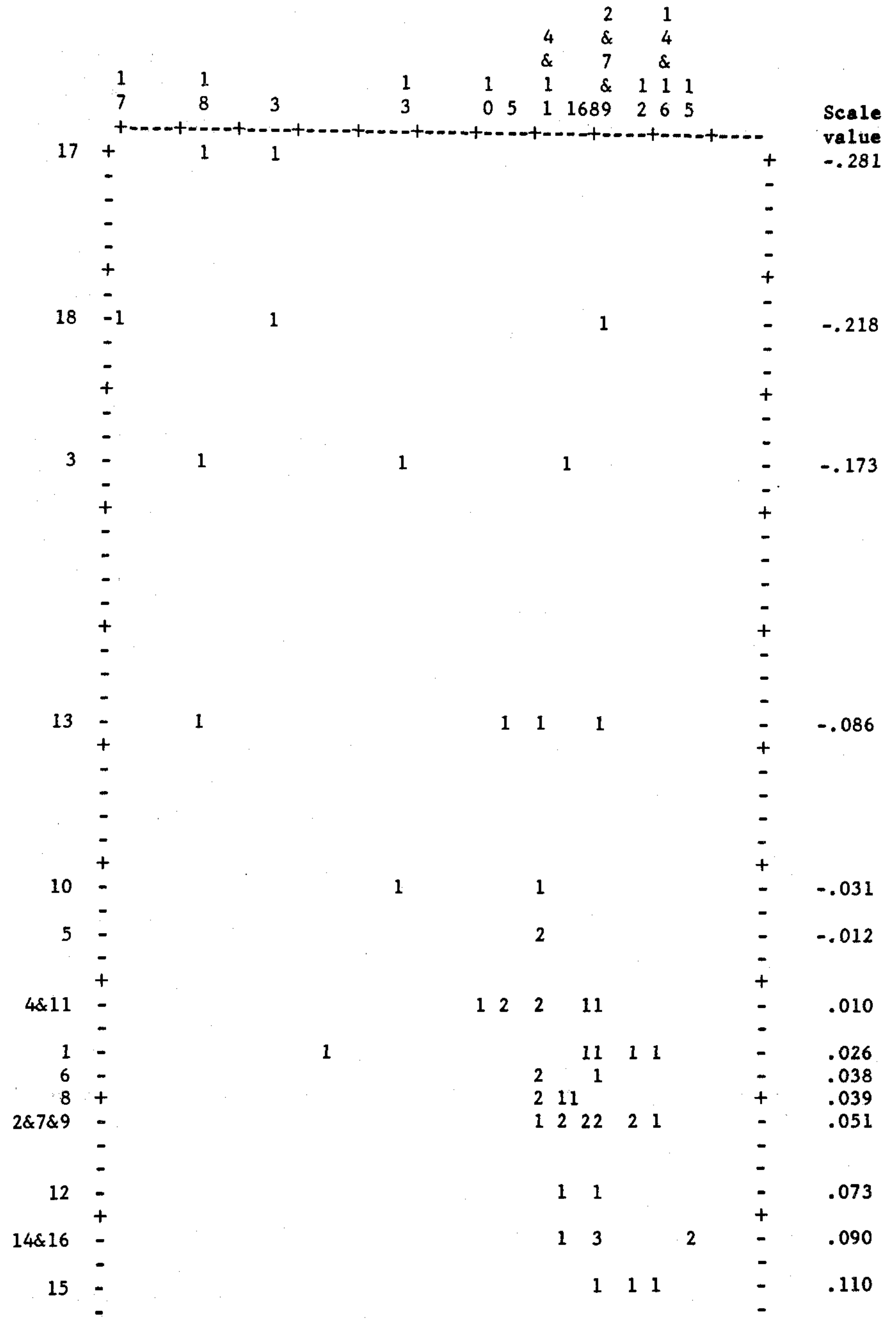

Figure 2.
Positive sociomatrix from period four of Sampson's Monastery Study with rows and columns spaced according to the *third* eigenvector of the correspondence analysis solution.

10, 4, and 5 from the others. (All eight have positive values on the second dimension.) Sampson labeled these individuals the "Loyal Opposition" due to their conservative approach to authority. Within this group, individuals 4, 5, 6, 10, and 11 are placed together at the bottom of Figure 1. Only brother 10 originates ties to others outside the Loyal Opposition, and he was the only member of this subgroup to leave at the end of Sampson's study.

Reordering the sociomatrix using the third eigenvector (λ = .722, Figure 2) separates brothers 17, 18, 3, 13, 10 and 5 from the others, with brothers 17, 18, and 3 having scale values that are considerably more negative than 13, 10 and 5. Sampson labeled the former three the "Outcasts". All three were expelled from the monastery, but otherwise they had very little in common and were peripheral to the main conflict between the "Young Turks" and the Loyal Opposition. (The results in the next section allow for a fuller interpretation of the relationships between the Outcasts and the Young Turks and Loyal Opposition.) The Outcasts relationships are reflected in the distance separating the three (the top of Figure 2). Unlike the Loyal Opposition, the Outcasts are not a group of "close" individuals, but rather an unconnected set that do not cluster tightly.

The division of individuals into Loyal Opposition, Young Turks, and Outcasts is further supported by comparing the second and third eigenvalues (λ = .793 and .722, respectively) against the substantially smaller values of the fourth and fifth eigenvalues (.544 and .314 respectively). Chi-squares (Nishisato, 1980) of 59.05 (df = 33, $p < .01$) and 47.95 (df = 31, $p < .05$) for the second and third eigenvalues and 30.29 (df = 29, $p > .30$) for the fourth eigenvalue indicate that the major break occurs between the third and four eigenvalues.

To locate the main group, against which the Loyal Opposition and Outcasts are contrasted, the correspondence analysis solution is used to construct a sociogram by plotting each individual at the coordinates of the second and third eigenvectors. In Figure 3, the three groups—Young Turks, Loyal Opposition, and Outcasts—are evident. As noted before, the second eigenvector (vertical axis) extracts the Loyal Opposition and the third eigenvector extracts the Outcasts.

The position of brother 13 in Figure 3 deserves note; this individual has developed a "history" in the analyses of these data. Breiger et al. (1975) blocked four composite matrices and classified him with the Loyal Opposition. White et al. (1976) blocked the eight stacked matrices and placed him with the Outcasts. Breiger et al.'s (1975; Figure 7, p. 366) MDSCAL analysis of the first-correlation matrix placed him much closer to the Outcasts, which is in agreement with a

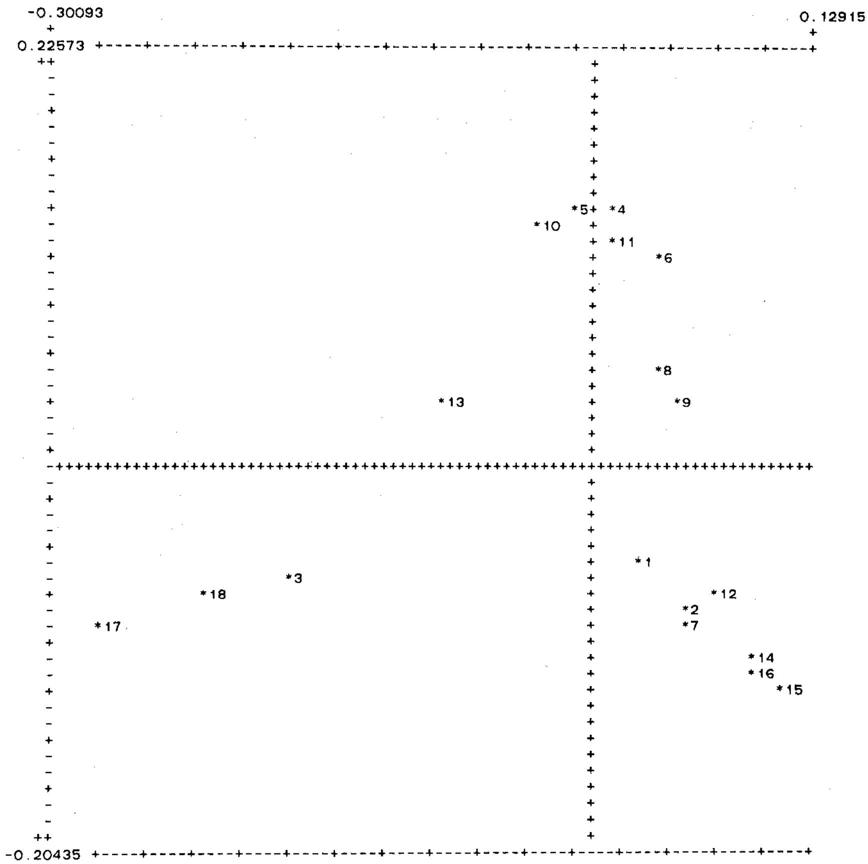

Figure 3.
Positive sociogram from period four of Sampson's Monastery Study with individuals located according to their scale of values from the *second* and *third* eigenvectors of the correspondence analysis solution.

HICLUS (Johnson, 1967) analysis of the same matrix. Sampson's original account has brother 13 as a "Waverer" (along with 8 and 10) who was not strongly aligned with any subgroup. (Brother 13 was the only person to vote for an Outcast, 3, as chair of an important meeting.) The present analysis, based upon the two largest eigenvectors, places him squarely between the Outcasts and the Loyal Opposition, thus suggesting a reason for some of the ambiguity in the earlier analyses: if indeed his location is between the two subgroups, then his placement would be sensitive to the manner in which the sociomatrix was preprocessed. Our findings are reinforced by Sampson's observation that brother 13 was interstitial to the two main subgroups while ranking low, along with the Outcasts, on esteem. Note, too, that brother 13 is the only Outcast who originates positive ties to both the

Loyal Opposition and the Young Turks. Similarly, Figure 3 shows brothers 8 and 9 to be located between the Loyal Opposition and the Young Turks. Brother 8's position is consistent with the waverer status accorded to him by Sampson. Brother 9's position is due to the ties he has to brothers 1, 2, and 12 of the Young Turks. (Again, a fuller understanding of the positions of these individuals can be obtained from the analysis in the next section.)

Scaling Non-attracting Relations

With respect to both the blockmodeling results and the correspondence analysis of positive affect, brother 12 is a Young Turk. But White et al. noted, "A puzzle in both Sampson's picture and the blockmodel is why monk 12 remained" (1976, p. 754) after all other Young Turks and Outcasts had resigned or been expelled. The anomaly of brother 12 may be due to differences between the structures of positive and negative affect. We can test for any differences by a correspondence analysis that retains the basic distinction between positive interpersonal ties and negative relationships.

The correspondence analysis of positive affect is a straightforward minimization of the interpoint distances of linked individuals. Distance minimizing the linked individuals in negative sociometric data would place individuals with dislike, disesteem, and so forth *closer* to each other in the solution. But, to the extent that negative affect sociometric data indicate the "repulsion" between individuals, zero-one matrices of negative choices require that linked individuals be maximally separated. Over all individuals, this is equivalent to a maximizing of equation [2]. Since the value of equation [2] is one minus the eigenvalue, negative choices can be represented by the eigenvectors associated with the most negative eigenvalues (eigenvalues near -1, corresponding to values of equation [2] near 2).

Sampson collected negative relational data in his study (e.g. "list in order those three brothers whom you esteemed the least"). As in the case of positive choices (Table 1), these data were taken from the four negative matrices of the fourth reporting period. The matrix in Table 2 contains a one if the row individual ranked the column individual as the lowest or second lowest on at least one of the four classes of relations: affect, esteem, influence, or sanctioning; $a_{ij} = 0$ otherwise.

The correspondence analysis of these negative-choice data resulted in solutions with non-trivial eigenvalues ranging from .393 to $-.760$. The low value (.393) of the most positive eigenvalue demonstrates that the brothers are not split into disjoint subgroups of mutual dislike. In contrast, an eigenvalue equal to -1 indicates a grouping

```
                                                    1
                 4                                  7
                 &                                  &
       1         1    11          1        1        1   1
       2 5     9 18 6 03        1 6    3   5 2 7    8   4         Scale
       +----+----+----+----+----+----+----+----+----+----+----    value
   12 +                                             1   1     +   -.159
      -                                                       -
    5 -                                             2         -   -.147
      -                                                       -
      -                                                       -
      +                                                       +
      -                                                       -
      -                                                       -
    9 -                                2            2         -   -.111
      -                                                       -
 4&11 +                         1      1    2 1     2   1     +   -.094
    8 -                1          1    1    1       2         -   -.092
      -                                                       -
    6 -                                1      1     1         -   -.081
      -                                                       -
   10 +                                                       +   -.069
   13 -             1                       1       2         -   -.062
      -                                                       -
      -                                                       -
      -                                                       -
      +                                                       +
      -                                                       -
      -                                                       -
      -                                                       -
      -                                                       -
    1 +          1    1                1     1 1    1         +   -.006
      -                                                       -
   16 -          2                     1            2         -    .006
      -                                                       -
      -                                                       -
      +                                                       +
      -                                                       -
    3 -          1       1                   1                -    .030
      -                                                       -
      -                                                       -
      +                                                       +
   15 -          1     1               1                      -    .055
      -                                                       -
    2 -          11    1               1                      -    .071
      -                                                       -
    7 +          11 1                  1                      +    .078
      -                                                       -
      -                                                       -
      -                                                       -
      -                                                       -
17&18 +          32 2 1                                       +    .111
      -                                                       -
      -                                                       -
   14 -          21    1                                      -    .131
      -                                                       -
       +----+----+----+----+----+----+----+----+----+----+----
```

Figure 4.
Negative affect sociomatrix from period four of Sampson's Monastery Study with rows and columns spaced according to the *eighteenth* eigenvector of the correspondence analysis solution.

with no negative affect within groups and only negative affect between groups. The eigenvalue near -1 ($-.760$) indicates a relatively good fit to such a model. Therefore, we use the eigenvector with the most negative eigenvalue to reorder the rows and columns of the sociomatrix (Figure 4). The ordering along the eighteenth eigenvector reiterates the antagonism between the Young Turks and the Loyal Opposition. In particular, this dimension identifies those individuals (with the exception of brother 8, the last to leave voluntarily) who remained after the turmoil. The division of the two groups at the origin of the axis, however, is not totally consistent with that in the second eigenvector of the positive affect sociomatrix (Figure 1): individual 12 is now clearly classified with the Loyal Opposition instead of the Young Turks. Correspondence analysis of the negative affect matrix thus suggests an explanation for brother 12's anomalous behavior. The main reason for his survival may have been his lack of antagonism toward those in the Loyal Opposition since he is the only Young Turk without negative links to the opposition.

The eigenvectors associated with the two most negative eigenvalues were also used to plot the individuals in a sociogram (Figure 5). ($\lambda = -.561$ for the seventeenth eigenvector.) Though less tightly grouped than in Figure 3, individuals may still be divided into three groups: Outcasts (3, 17, 18) plus 7, Young Turks (1, 16, 2, 15, 14) excluding 12 and 7, and Loyal Opposition (13, 4, 10, 11, 8, 5, 9, 6) plus 12. Brother 7's location in Figure 4 reflects a number of factors noted by Sampson. After the departure of several novices (not in the present analysis), brother 7 was "left a complete isolate" and the "least liked of all." He also had a long standing hatred of brother 5 and was unique in his attachment to the departed brothers. As noted before, the movement of brother 12 from one cluster to another illuminates an important factor determining his survival in the monastery. The other five survivors (4, 11, 5, 9, 6) also tend toward the left of the plot. Overall, our findings suggest that survival was more a function of whom one disliked (Figure 5), as opposed to whom one liked (Figure 3).

Discussion

Structure within groups has been highlighted using sociograms and sociomatrices. These are complimentary methods of representation. A permuted sociomatrix emphasizes the clustering of individuals and the structural equivalence of their positions. By contrast, sociograms place fewer strictures on the output and allow for a spatial

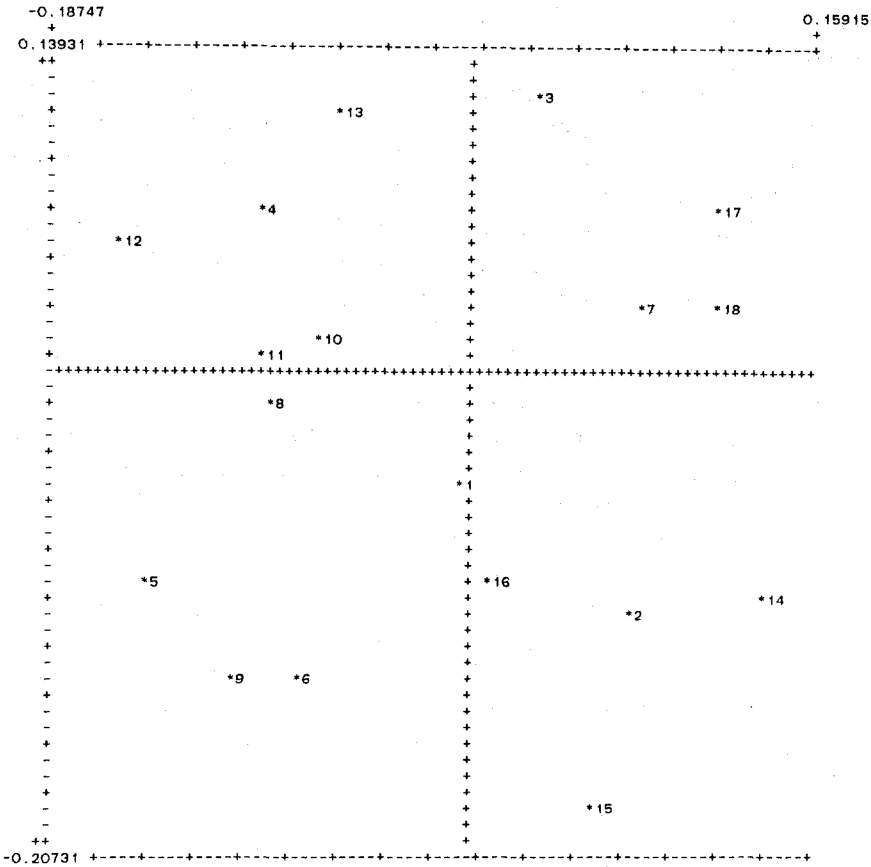

Figure 5.
Negative affect sociomatrix from period four of Sampson's Monastery Study with individuals located according to their scale values from the *seventeenth* and *eighteenth* eigenvectors of the correspondence analysis solution.

interpretation of structure. We have shown that correspondence analysis can be used to both reorder the rows and columns of a sociomatrix and place individuals as points in a sociogram, thereby exploiting the advantages of both representations.

Correspondence analysis is based on a model of within-group interaction which can maintain the distinction between positive and negative choices. Positive choices, such as "liking," should indicate close proximity between the chooser and the chosen. Therefore, a scaling method should try to group the givers and receivers as close to each other as possible. Correspondence analysis does this. Similarly, negative choices indicate that the giver and receiver should be maximally separated. Correspondence analysis also can be used to derive this solution.

Correspondence analysis can be an exploratory tool for determining blocks for a blockmodel. That is, the method is capable of clustering individuals without the numeric uncertainty associated with CONCOR. Along with the computational uncertainties, CONCOR provides no information on the adequacy of blockmodeling to fit the data, and does not suggest the number of blocks to be used. For each dimension, correspondence analysis provides two indicators of bifurcation: the eigenvalue associated with each dimension and the sociomatrix with the rows and columns reordered by the scale values. In correspondence analysis, the number of eigenvalues equal to one is the number of totally disjoint sets of individuals—sets that will have interblock densities of zero in a blockmodel. The number of eigenvalues near one shows the number of nearly isolated subgroups, thus indicating low interblock densities. Therefore, the spectrum of eigenvalues suggests the number of blocks to be used in the blockmodeling. The information contained in the sociomatrix, when reordered and spaced by the high eigenvalues can then be used to verify the bifurcation of the groups. Simultaneous consideration of more than one dimension—the spatial representation of the sociogram—is a further check for individuals, such as Sampson's brother 13 (Figure 3), who do not fit cleanly into any block.

Many of the competing methods (e.g. Beum & Brundage, 1950, and CONCOR) are based on computational procedures with unknown numerical properties (see McQuitty & Clark, 1968). They may not converge or they may converge to non-optimal solutions (see Deutsch & Martin, 1971). In this way, their behavior is similar to many of the nonmetric multidimensional scaling routines. By contrast, the computational heart of correspondence analysis has been extensively studied (see Wilkinson, 1965) so there is a large body of alternative computational procedures to circumvent problems arising in an eigenvalue-eigenvector analysis.

In summary, the scale values from correspondence analysis are a flexible tool for organizing and representing sociometric data. Using these individual scale values to reorganize the configuration of points in a sociogram is one representation. Spacing the rows and columns of a sociomatrix is another representation. Correspondence analysis maintains some of the best features of multidimensional scaling (a spatial representation of the sociogram), principal components analysis (dimensional interpretation of the sociometric structure using fixed dimensions), and exploratory blocking procedures (a display of sociometric choices placing structurally equivalent individuals either maximally near to, or distant from, each other).

References

Beum, C. O., & Brundage, E. G. (1950). A method for analyzing the sociomatrix. *Sociometry, 13,* 141–145.

Boorman, S. A., & White, H. C. (1976). Social structure from multiple networks: II. Role structures. *American Journal of Sociology, 81,* 1384–1446.

Breiger, R. L., Boorman, S. A., & Arabie, P. (1975). An algorithm for clustering relational data, with application to social network analysis and comparison with multidimensional scaling. *Journal of Mathematical Psychology, 12,* 328–383.

Coleman, S., & MacRae, D. (1960). Electronic processing of sociometric data from groups up to 1,000 in size. *American Sociological Review, 25,* 722–726.

Coombs, C. H. (1964). *A Theory of Data.* New York: Wiley.

Deutsch, B., & Martin, J. J. (1971). An ordering algorithm for analysis of data arrays. *Operations Research, 19,* 1350–1362.

Forsythe, E., & Katz, L. (1946). A matrix approach to the analysis of sociometric data: Preliminary report. *Sociometry, 9,* 340–347.

Guttman, L. (1941). The quantification of a class of attributes: A theory and method of scale construction. In P. Horst (Ed.), *The Prediction of Personal Adjustment,* pp. 251–364. New York: Social Science Research Council.

Heil, G. H., & White, H. C. (1976). An algorithm for finding simultaneous homomorphic correspondences between graphs and their image graphs. *Behavioral Science, 21,* 26–35.

Heiser, W. J. (1981). *Unfolding Analysis of Proximity Data.* Leiden.

Johnson, S. C. (1963). Hierarchical clustering schemes. *Psychometrica, 32,* 241–254.

Katz, L. (1947). On the matrix analysis of sociometric data. *Sociometry, 10,* 233–241.

Kendall, D. G. (1971). Seriation from abundance matrices. In F. R. Hodson, D. G. Kendall, & P. Tautu (Eds.), *Mathematics in the Archaeological and Historical Sciences,* pp. 213–252. Edinburgh: University of Edinburgh Press.

Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika, 29,* 1–27.

Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika, 29,* 115–129.

Laumann, E. O., & Pappi, F. (1973). New directions in the study of community elites. *American Sociological Review, 38,* 212–230.

Levine, J. H. (1979). Joint-space analysis of 'pick-any' data: Analysis of choices from an unconstrained set of alternatives. *Psychometrika, 44,* 85–97.

Lorrain, F. P., & White, H. C. (1971). Structural equivalence of individuals in social networks. *The Journal of Mathematical Sociology, 1,* 49–80.

McQuitty, L. L., & Clark, J. A. (1968). Clusters from iterative, intercolumnar correlational analysis. *Educational and Psychological Measurement, 28,* 211–238.

Moreno, J. L. (1934). *Who Shall Survive? A New Approach to the Problem of Human Inter-relations.* New York: Beacon House.

Nishisato, S. (1980). *Analysis of Categoric Data: Dual Scaling and Its Applications.* Toronto: University of Toronto Press.

Noma, E. (1982a). The simultaneous scaling of cited and citing articles in a common space. *Scientometrics, 4,* 205–231.

Noma, E. (1982b). Untangling citation networks. *Information Processing and Management, 18,* 43–53.

Noma, E. (1984). Co-citation and the invisible college. *Journal of the American Society for Information Science, 35,* 9–33.

Roistacher, R. C. (1974). A review of mathematical methods in sociometry. *Sociological Methods & Research, 3,* 123–171.

Sailer, L. D. (1978). Structural equivalence: Meaning, and definition, computation and application. *Social Networks, 1,* 73–90.

Sampson, S. F. (1969). *Crisis in a Cloister.* Ann Arbor, Michigan: University Microfilms, No. 69-5775.

Schwartz, J. E. (1977). An examination of CONCOR and related methods for blocking sociometric data. In D. R. Heise (Ed.), *Sociological Methodology 1977,* pp. 255–282. San Francisco: Jossey Bass.

Spilerman, S. (1966). Structural analysis and the generation of sociograms. *Behavioral Science, 11,* 312–318.

White, H. C., Boorman, S. A., & Breiger, R. L. (1976). Social structure from multiple networks. I. Blockmodels of roles and positions. *American Journal of Sociology, 81,* 730–780.

Wilkinson, J. H. (1965). *The Algebraic Eigenvalue Problem.* Oxford: Oxford University Press.