

# Uncovering Organizational Hierarchies

LINTON C. FREEMAN

University of California, Irvine School of Social Sciences, Irvine, CA 92697-5100

email: lin@aris.ss.uci.edu

## **Abstract**

This paper adapts a procedure, canonical analysis of asymmetry, originally defined by Gower (1977) to the analysis of hierarchical properties in organizational structures. Its application is demonstrated using two available data sets, Freeman and Freeman's (1980) data on computer communication and Krackhardt's (1987) data on advice-giving and getting in an organization.

**Keywords:** hierarchy, dominance, asymmetry, social networks

## **1. Introduction to the Problem**

Organizational data usually come in the form of square matrices that contain records of social connections between pairs of actors. If the research is concerned with reciprocated connections—like mutual friendship or mutual enmity—and if the data in the matrices are symmetric, then analysis is straightforward. Several powerful procedures for uncovering structural patterning in symmetric data, like multidimensional scaling, principal components and correspondence analysis, are available (Weller and Romney 1990; Wasserman and Faust 1994). Even if their data depart from symmetry, analysts who are focused on reciprocated connections typically symmetrize their data and use these standard procedures for their analysis.

But some investigators are directly concerned with connections—like dominance or power or influence—that are by definition asymmetric. In such cases it would be useful to have a collection of standard procedures, like those used to analyze symmetric data, that could be used to uncover the patterning of asymmetric data. Unfortunately, we do not have a similar collection of standard procedures that can be used for the analysis of asymmetric relations.

The aim of the present paper is to show the potential of one such procedure for analyzing asymmetries in organizational structures. I will describe a procedure that was originally introduced by Gower (1977), who called it *canonical analysis of asymmetry*. In the section below I will review Gower's approach. Following that, I will examine its potential for the study of organizations by applying it to two network data sets: first Freeman and Freeman's (1980) data on the impact of computer facilitated communication among social network analysts, and second Krackhardt's (1987) data on a perceived advice network in a small corporation.

## 2. Gower's Canonical Analysis of Asymmetry

Suppose we have data that come in the form of an asymmetric matrix

$$X = \begin{bmatrix} - & 32 & 29 & 16 \\ 28 & - & 9 & 29 \\ 21 & 11 & - & 32 \\ 4 & 21 & 28 & - \end{bmatrix}.$$

Gower showed that such a matrix can be partitioned into two parts: (1) a symmetric matrix  $Y$ —that captures the proximity or overlap of each pair of objects  $i$  and  $j$ ,

$$y_{ij} = \frac{x_{ij} + x_{ji}}{2},$$

and (2) a skew-symmetric matrix  $Z$ —that captures the asymmetry or the extent to which each object  $i$  dominates each other  $j$ ,

$$z_{ij} = \frac{x_{ij} - x_{ji}}{2}.$$

Of course,

$$x_{ij} = \frac{x_{ij} + x_{ji}}{2} + \frac{x_{ij} - x_{ji}}{2},$$

so, in matrix terms,

$$X = Y + Z.$$

Thus, when we partition a data matrix like  $X$  we end up with a symmetric matrix

$$Y = \begin{bmatrix} - & 30 & 25 & 10 \\ 30 & - & 10 & 25 \\ 25 & 10 & - & 30 \\ 10 & 25 & 30 & - \end{bmatrix},$$

and a skew-symmetric matrix

$$Z = \begin{bmatrix} - & 2 & 4 & 6 \\ -2 & - & -1 & 4 \\ -4 & 1 & - & 2 \\ -6 & -4 & -2 & - \end{bmatrix}.$$

It is easy to uncover the patterning of the linkages among the objects for symmetric data of  $Y$ . We can use any of the standard procedures listed above. But, since all those procedures require symmetry, the skew-symmetric data of  $Z$  require a different approach.

Gower has shown that ordinary singular-value decomposition can be used to uncover the structural patterning of skew-symmetric matrices. Thus, matrix  $Z$  can be defined

as a product,

$$Z = USV',$$

in which  $U$  is the matrix of eigenvectors of  $ZZ'$ ,  $V$  is the matrix of eigenvectors of  $Z'Z$ , and the diagonal matrix  $S$  contains the eigenvalues (the squared singular values of  $ZZ'$ ).<sup>1</sup> In standard applications of singular-value decomposition, successive eigenvalues are monotone decreasing, so less variance is associated with each successive dimension (eigenvector). When we use singular value decomposition, then, we seek a kind of simplification in which most or all the variance is associated with the first few dimensions. When that occurs, a relatively small number of dimensions can provide an acceptably accurate depiction of the overall structure.

But the decomposition of a skew-symmetric matrix results in a special kind of non-standard structure. In this case, because of the balance between positive and negative cell values associated with skew-symmetry, the first dimension is associated with exactly the same amount of variance in the original data as is the second. The same is true of succeeding pairs of dimensions (3 and 4, 5 and 6 and so on). Therefore, when skew-symmetric matrices are used as input, successive pairs of eigenvalues (1 and 2, 3 and 4 and so on) will always turn out to be equal in value and successive pairs of dimensions that share the same eigenvalue must be taken together and treated as units. So, in this case, our hope is to capture the structural patterning in a small number of *pairs* of dimensions.

The singular value decomposition of the skew-symmetric matrix  $Z$  yields a diagonal matrix of squared singular values, or eigenvalues:

$$S = \begin{bmatrix} 8.52 & 0 & 0 & 0 \\ 0 & 8.52 & 0 & 0 \\ 0 & 0 & 2.11 & 0 \\ 0 & 0 & 0 & 2.11 \end{bmatrix}.$$

And the two sets of eigenvectors associated with  $S$  are

$$U = \begin{bmatrix} .371 & .787 & .139 & -.473 \\ -.181 & .459 & .439 & .751 \\ -.290 & .398 & -.850 & .185 \\ -.863 & .109 & .254 & -.422 \end{bmatrix}$$

and

$$V = \begin{bmatrix} .787 & .371 & .473 & .139 \\ .459 & -.181 & -.751 & .439 \\ .398 & -.290 & -.185 & -.850 \\ .109 & -.863 & .422 & .254 \end{bmatrix}.$$

Here, the first two eigenvectors account for  $17.04/21.26 = 80.1\%$  of the variance in  $Z$ . The four objects, therefore, can be represented in two dimensions with little distortion. To construct such a representation, we simply multiply the elements in the first two columns

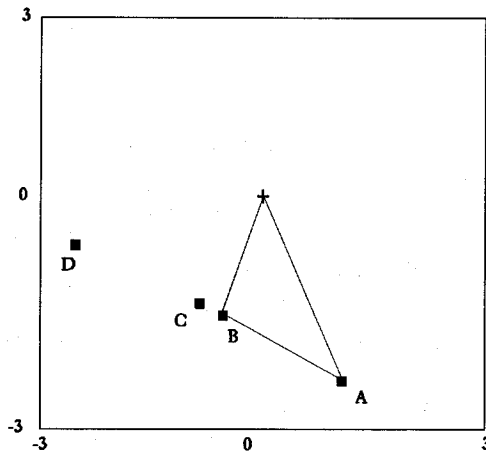


Figure 1. Plot of the first two dimensions of hypothetical data.

of  $U$  by the square root of the first eigenvalue,  $s_{11} = \sqrt{8.517} = 2.92$ , and we can plot the points in the first two dimensions,  $x$  and  $y$ , as follows:

Object	$x$	$y$
A	1.084	2.296
B	-.527	1.338
C	-.847	1.162
D	-2.519	.317

The resulting plot is shown in Figure 1.

Distances between points in this figure cannot be calculated in the standard Euclidean way. The asymmetries between objects in  $Z$  are captured here, not by the linear distances between them, but by the planar areas they enclose in the figure. The amount of asymmetry between object  $A$  and object  $B$ , for example, is represented by the area of the triangle enclosing point  $A$ , point  $B$ , and the origin (the  $+$  falling at  $0, 0$ ) shown in Figure 1.

Typically, what we are looking for when we analyze any asymmetric structure is a way of ordering the objects in question along a single linear dimension. And it turns out that singular-value decomposition does precisely that when the data permit such an arrangement. Consider, for example, an order relation that *is* linear, one that is that is irreflexive, asymmetric, transitive and complete. In such a relation, one actor  $A$  in some sense dominates everyone, the next  $B$  dominates everyone but  $A$ , and so on, until  $E$  is dominated by all the other actors. The matrix shown in Table 1 is a linear dominance matrix. And the matrix of Table 1 yields the skew-symmetric matrix shown in Table 2.

When this kind of matrix is subjected to singular value decomposition the results are dramatic. The points are equidistant from one another and they are all arranged in an arc of a circle that is centered at the origin. That circle has a radius of  $2\sqrt{(s_{11}/n)}$  where  $n$  is the number of points in the linear order. And the first pair of eigenvalues is determined by  $n$  so that

$$s_{11} = s_{22} = \frac{1}{2} \tan\{(n-1)\pi/2n\}.$$

Table 1. Matrix representing a linear dominance relation.

	A	B	C	D	E
A	0	1	1	1	1
B	0	0	1	1	1
C	0	0	0	1	1
D	0	0	0	0	1
E	0	0	0	0	0

Table 2. The dominance matrix of Table 1 in skew-symmetric form.

	A	B	C	D	E
A	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
B	$-\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
C	$-\frac{1}{2}$	$-\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$
D	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	0	$\frac{1}{2}$
E	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	0

Figure 2 illustrates the result of applying singular value decomposition to the objects in Table 2.

When the arrangement is this orderly, of course, we do not have to treat differences in terms of the areas of planar figures; instead we can use the ordinary Euclidean distances between points to reconstruct the order in the original data matrix. As a matter of fact, we can simply straighten out our arc and produce an image that captures the linearity of the dominance data. See Figure 3.

Most collections of actual data, of course, are something less than perfectly linear. If their departures are small, they will be projected close to the arc defined above and we will still be able to order the points in a single dimension. As their departures from the linear pattern grow, points will be projected farther and farther from the arc. In the next section I will explore the potential of this procedure by applying it to a small set of actual organizational data on an e-mail type of communication linking a collection of social network analysts.

### 3. Canonical Analysis of Asymmetry of Communication among Network Analysts

For 18 months, beginning in 1978, Sue Freeman and I were involved in a project examining the impact of an e-mail kind of communication on a collection of 50 specialists in the study of social networks (Freeman and Freeman 1980; Freeman 1980, 1984). During that period the message traffic among these specialists was monitored and the result was a  $50 \times 50$ , who-to-whom matrix of communication frequency.

For the purpose of illustrating the present procedure, it would be useful to have a much smaller data matrix, so I abstracted out the communication frequencies among the subset of seven network specialists who were the most frequent communicators.<sup>2</sup> This gives us the

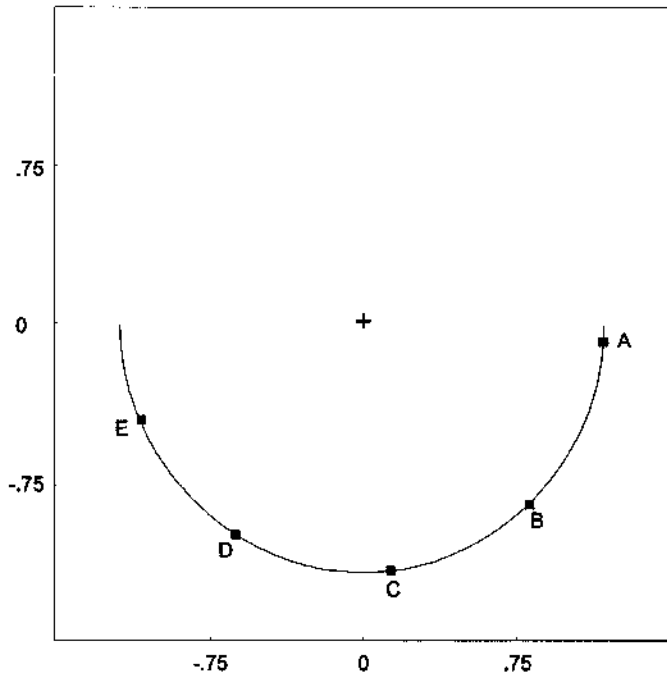


Figure 2. The first two dimensions of a linear order.

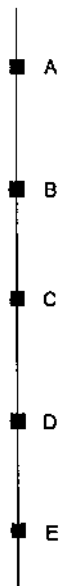


Figure 3. A straight-line projection of the order in figure 2.

*Table 3.* Number of messages flowing between pairs of network analysts.

	Fre	Whi	Alb	Ber	Dor	Mul	Wel
Freeman	—	115	17	93	53	33	84
White	84	—	4	5	5	0	15
Alba	16	10	—	15	3	3	4
Bernard	127	22	17	—	57	12	34
Doreian	57	9	4	57	—	8	10
Mullins	23	4	3	9	8	—	33
Wellman	118	24	5	35	15	45	—

$7 \times 7$  matrix shown in Table 3. Clearly the data of Table 3 are not symmetric. Wellman, for example sent 45 messages to Mullins, but Mullins sent only 33 to Wellman. Indeed, for most pairs of individuals, either one or the other sent more messages than he received. The asymmetry in this table, then, stems from differences in garrulousness, or propensity to initiate computer-based communication.

First, much of the information in Table 3 was thrown away by reducing the data to dichotomous form. I do not recommend throwing data away as a general practice, but for the present illustrative purposes there are three good reasons to do so. (1) As it is usually conceived, dominance is a binary relation (Kendall and Gibbons 1990, Chapter 11). As it is modeled, then, any individual either dominates a particular other, or does not. (2) The results of singular value decomposition on dichotomous skew symmetric data are simpler and easier to interpret than are those produced when frequencies are used. Any non-linearity in frequency data disturbs the results by generating ambiguities where the order of individuals is masked. (3) Almost all of the relational data we collect come to us in dichotomous—yes/no—form. For all these reasons, then, dichotomous data are preferred for the present illustration.

So the data were converted to binary form by applying the rule that, for any pair of individuals, whichever individual sent the most messages to the other was considered to be the more garrulous of the two. This rule was applied to the data of Table 3 to construct the garrulousness matrix of Table 4.

*Table 4.* Garrulousness relation among network analysts.

	Fre	Whi	Alb	Ber	Dor	Mul	Wel
Freeman	—	1	1	0	0	1	0
White	0	—	0	0	0	0	0
Alba	0	1	—	0	0	0	0
Bernard	1	1	1	—	0	1	0
Doreian	1	1	1	0	—	0	0
Mullins	0	1	0	0	0	—	0
Wellman	1	1	1	1	1	1	—

Table 5. Skew-symmetric transformation of the data of Table 4.

	Fre	Whi	Alb	Ber	Dor	Mul	Wel
Freeman	—	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$
White	$-\frac{1}{2}$	—	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$
Alba	$-\frac{1}{2}$	$\frac{1}{2}$	—	$-\frac{1}{2}$	$-\frac{1}{2}$	0	$-\frac{1}{2}$
Bernard	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	—	0	$\frac{1}{2}$	$-\frac{1}{2}$
Doreian	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	—	0	$-\frac{1}{2}$
Mullins	$-\frac{1}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	0	—	$-\frac{1}{2}$
Wellman	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	—

The data in Table 4 do not reflect the presence of a strict linear order. There are, for example, some pairs of individuals, like Bernard and Doreian and Doreian and Mullins, who each sent the same number of messages to each other. These, pairs cannot be directly ordered in terms of their interaction together.

In order to produce a linear order from this partial order, we then transform the zero-one matrix of Table 4 into the skew-symmetric matrix of Table 5. And the next step was to subject the skew-symmetric data of Table 5 to singular value decomposition. I did that, and the projection of the locations of the seven individuals in the first two dimensions is shown in Figure 4 and they are projected on a straight-line image in Figure 5.

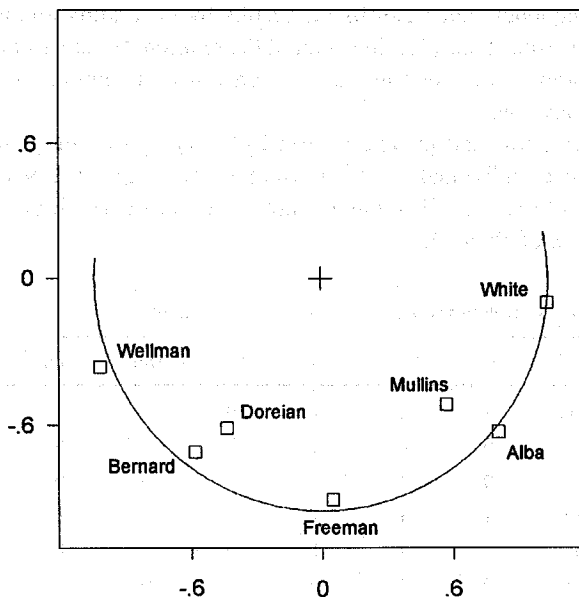


Figure 4. The first two dimensions of the communication data.



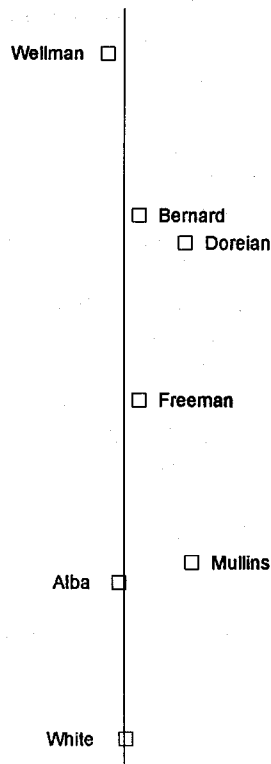


Figure 5. A straight-line projection of the order in Figure 4.

Unlike the idealized case presented above, these points do not fall exactly on the arc. But they are close enough to permit us to order those pairs of individuals that were in ambiguous positions *vis a vis* one another. In this case, then, singular value decomposition is able to uncover an explicit linear order in the data<sup>3</sup>. Its success results from the fact that, in ordering a pair, singular value decomposition does not focus only on which person, of that pair, sent more messages to the other. Instead, it focuses on the overall pattern of message sending among all of the individuals. The position of each individual is determined in terms both of how many others it dominates and who those others are—in terms of how many they dominate. And the process is continued; the individuals at the second level are rated in terms of how many others they, in turn, dominate, and so on. Thus, the computation is based on the *total* patterning of who dominates which others.

How can we evaluate the order in the present case?<sup>4</sup> I certainly have no explicit external criterion against which to map these results. But those of us in the social networks research community would certainly agree that Barry Wellman takes the initiative. He was then, and still is, the most active organizer—the organizational leader—of that community!

In the next section, I will examine the potential of this approach by drawing on a more elaborate data set.



and Faust 1982; Freeman, Romney and Freeman 1987; Kumbasar, Romney and Batchelder 1994). These positional biases, however, tend to cancel one another out when the responses of all the participants are pooled. Therefore, in order to estimate who had more interpersonal influence than whom, I simply summed the data matrices of the 21 managers and dichotomized as I did in the case of network specialists. Here a matrix entry of 1 means that the row person had more votes for influencing the column person than the column person had for influencing the row person<sup>5</sup>. This dominance matrix is shown in Table 6.

The dichotomized data were then converted to a skew-symmetric form and subjected to singular value decomposition. The vertical projection of the plot of the first two axes is shown in Figure 6. The hierarchical structure of the data was emphasized by Krackhardt (1987, p. 119) who commented that the data showed “a considerable degree of hierarchy.” This hierarchy is apparent in the arrangement in Figure 5. The arrangement, moreover, is consistent with Krackhardt’s general impressions about that hierarchical structure<sup>6</sup>.

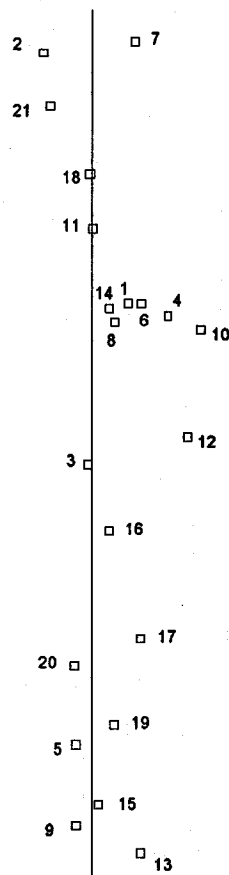


Figure 6. A straight-line projection of the order of the advice data.

Individual #7, at the top, is the president of the company. The next three individuals are three of the four vice-presidents. The other vice president is #14. On the basis of his ethnographic impressions, Krackhardt characterized these individuals "centers of advice."

Krackhardt also asked his subjects to rate one another on a seven point scale in terms of their "power." He observed that the bottom two individuals on the present scale are "the only ones attributed with (average) power levels below 3.0." Overall, then the order uncovered by the canonical analysis, seems to fit quite well with what we know about the organizational roles played by these individuals.

## 5. Summary and Conclusions

The present paper has shown how Gower's procedure for canonical analysis of asymmetry can be applied in the analysis of organizations to uncover various kinds of dominance orders. I have illustrated its potential by applying it to two data sets, one by Freeman and Freeman (1980) on communication frequency among social network analysts, and another by Krackhardt (1987) on a perceived advice network in an organization. In both cases, the orders it produced were consistent with intuition and, for the Krackhardt data, they were consistent with external criteria.

Gower's canonical analysis has two additional features that makes it particularly attractive for applications of this sort. First, it is based on an adaptation of a well known and widely used computational device, singular value decomposition. The formal properties of singular value decomposition have been explicitly spelled out and it has been applied widely enough to allow its users to understand clearly how it organizes different kinds of data. Gower's adaptation of singular value decomposition to asymmetric analysis does not change these features. Each individual's position is determined by taking the *total* pattern into account.

Second, Gower's procedure does not require that elaborate steps be taken in order to collect dominance data. Indeed, the very simplest yes/no observations that are typically used to determine who does what to whom among the participants in an organization are ideal for input into canonical analysis.

There are a number of other procedures that might be used to impose order on asymmetrical data. Most involve one, or some combination, of several seemingly *ad hoc* transformations to order a set of individuals. Typically these procedures depend on assigning each individual an index based on some function of the number of others dominated by—and/or the number of others dominating—that individual (Beilharz, Butcher and Freeman 1966; Brantas 1968; Barner-Barry 1977; Clutton-Brock, Guinness and Albon 1982; Barrette and Vandall, 1985). Their lack of formal foundation diminishes their general utility.

Some procedures based on Thurstone's (1927) paired comparisons have also been used to order individuals (Bradley and Terry 1952; Batchelder and Bershada 1979; Boyd and Silk 1983; McMahan and Morris 1984). Like Gower's canonical analysis, these procedures embody explicit models. Thus they are clearly specified, and they can be used in many areas of application. The only problem in the present context is that all of them rest on probabilistic assumptions—like continuity and independence—that may not be appropriate in the study of dominance.

## Notes

1. A good introduction to the fundamentals of singular value decomposition can be found in Green and Carroll (1976).
2. These were the present author, Doug White, Richard Alba, Russ Bernard, Pat Doreian, Nick Mullins and Barry Wellman.
3. It should be noted here that the model was able to impose order on all of these points because the patterns of who they dominated and who they were dominated by were *distinct*. But if the data embody an organizational tree—where collections of points are regularly equivalent (dominating and dominated by others at equivalent levels), singular value decomposition will place all the members of each set of equivalent actors in the same spot. Thus the result will show that they all fall at the same level.
4. Note that in this context “dominance” is used in a very special sense. Here it means simply “took the greater initiative in electronic communication.” That might indicate that the initiative taker was, as initiator, the “leader” or it might mean that the non-initiator was, as the person approached, the leader.
5. Krackhardt (1987) argues that there are advantages of using the “locally aggregated” matrix in which the reports of each individual are used in assigning that individual’s position. Future research should compare the results of the method used here with that proposed by Krackhardt.
6. In a personal communication, Krackhardt commented on the overall arrangement and provided additional data that are used in this description.

## References

- Barner-Barry, C. (1977), “An Observational Study of Authority in a Preschool Peer Group,” *Political Methodology*, 4, 415–449.
- Barrette, C. and D. Vandal (1985), “Social Rank, Dominance, Antler Size, and Access to Food in Snow-Bound Wild Woodland Caribou,” *Behaviour*, 97, 118–146.
- Batchelder, W.H. and N.J. Bershad (1979), “The Statistical Analysis of a Thurstonian Model for Rating Chess Players,” *Journal of Mathematical Psychology*, 19, 39–60.
- Beilharz, R.G., D.F. Butcher and A.E. Freeman (1966), “Social Dominance and Milk Production in Holsteins,” *Journal of Dairy Science*, 49, 887–892.
- Boyd, Robert and Joan B. Silk (1983), “A Method for Assigning Cardinal Dominance Ranks,” *Animal Behaviour*, 31, 45–58.
- Bradley, R.A. and M.E. Terry (1952), “The Rank Analysis of Incomplete Block Designs, I. The Method of Paired Comparisons,” *Biometrika*, 39, 324–345.
- Brantas, G.C. (1968), “On the Dominance Order in Friesian-Dutch Dairy Cows,” *Zeitschrift Tierzucht Zuchtungsbiologie*, 84, 127–151.
- Clutton-Brock, T.H., F.E. Guinness and S.D. Albon (1982), *Red Deer: Behavior and Ecology of Two Sexes*. University of Chicago Press, Chicago, IL.
- Freeman, L.C. (1980), “Q-Analysis and the Structure of Friendship Networks,” *International Journal of Man-Machine Studies*, 12, 367–378.
- Freeman, L.C. (1984), “The Impact of Computer Based Communication on the Social Structure of an Emerging Scientific Specialty,” *Social Networks*, 6, 201–221.
- Freeman, L.C. and S.C. Freeman (1980), “A Semi-Visible College: Structural Effects on a Social Networks Group,” in M.M. Henderson and M.J. MacNaughton (Eds.), *Electronic Communication: Technology and Impacts*, Westview Press, Inc., Boulder, CO.
- Freeman, L.C., A.K. Romney and S.C. Freeman (1987), “Cognitive Structure and Informant Accuracy,” *American Anthropologist*, 89, 311–325.
- Gower, J.C. (1977), “The Analysis of Asymmetry and Orthogonality,” in J. Barra, F. Brodéau and G. Romier (Eds.), *Recent Developments in Statistics*, North Holland, Amsterdam.
- Green, P.E. and J.D. Carroll (1976), *Mathematical Tools for Applied Multivariate Analysis*. New York: Academic Press.
- Kendall, M. and J.D. Gibbons (1990), *Rank Correlation Methods*. Edward Arnold, London, UK.

- Krackhardt, D. (1987), "Cognitive Social Structures," *Social Networks*, 9, 109–134.
- Kumbasar, E., A.K. Romney and W.H. Batchelder (1994), "Systematic Biases in Social Perception," *American Journal of Sociology*, 100, 477–505.
- McMahan, C.A. and M.D. Morris (1984), "Application of Maximum Likelihood Paired Comparison Ranking to Estimation of a Linear Dominance Hierarchy in Animal Societies," *Animal Behaviour*, 32, 374–378.
- Romney, A.K. and K. Faust (1982), "Predicting the Structure of a Communications Network from Recalled Data," *Social Networks*, 4, 285–304.
- Thurstone, L.L. (1927), "A Law of Comparative Judgment," *Psychological Review*, 34, 273–286.
- Wasserman, S. and K. Faust (1994), *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, MA.
- Weller, S.C. and A.K. Romney (1990), *Metric Scaling: Correspondence Analysis*, Sage Publications, Newbury Park, CA.

**Linton Freeman** is Research Professor in the Institute for Mathematical Behavioral Sciences at the University of California, Irvine. There he conducts research on social and cognitive structure and contributes tools for visualizing social networks.

<http://eclectic.ss.uci.edu/~lin/lin.html>