Journal of Mathematical Sociology

ISSN: 0022-250X (Print) 1545-5874 (Online) Journal homepage: https://www.tandfonline.com/loi/gmas20

Alternative approaches to correspondence analysis of sociomatrices

John M. Roberts Jr.

To cite this article: John M. Roberts Jr. (1996) Alternative approaches to correspondence analysis of sociomatrices, Journal of Mathematical Sociology, 21:4, 359-368, DOI: 10.1080/0022250X.1996.9990188

To link to this article: https://doi.org/10.1080/0022250X.1996.9990188

Journal of Mathematical Sociology
1996, Vol. 21(4), pp. 359–368
Reprints available directly from the publisher
Photocopying permitted by license only

ALTERNATIVE APPROACHES TO CORRESPONDENCE ANALYSIS OF SOCIOMATRICES

JOHN M. ROBERTS, JR.*

Department of Sociology, University of New Mexico, Albuquerque, NM 87131, USA

If the usual correspondence analysis is viewed as a decomposition of departures from the model of independence, there are problems in applications to sociomatrices with regard to the treatment of the diagonal and cell estimates under independence. Other related techniques from the literature are more appropriate for analysis of sociomatrices. The different approaches are used to analyze a familiar sociomatrix, and the results of the techniques are compared.

KEY WORDS: Social networks, correspondence analysis, generalized correspondence analysis, residual scaling.

In recent years correspondence analysis has gained popularity as a tool for the analysis of social network data of various types (Wasserman and Faust, 1989; Wasserman, Faust, and Galaskiewicz, 1990). The most common form of social network data is the zero-one sociomatrix (symmetric or asymmetric) in which row entries indicate ties sent and column entries indicate ties received. Typically the diagonal of the sociomatrix carries no information. Noma and Smith (1985) discuss correspondence analysis for sociomatrices, and compare it to some other scaling techniques. Kumbasar, Romney, and Batchelder (1994) apply correspondence analysis to compare sociomatrices generated by individuals to an aggregated view of the group structure. In applications of correspondence analysis to the sociomatrix the missing diagonal is typically filled with ones since a general practice for proximity data is to view each object's proximity to its self as at least as great as its proximity to any others (Weller and Romney, 1990). In practice this approach seems to yield interpretable results, but it is not completely satisfactory on two grounds.

MISSING DIAGONAL

As the sociomatrix generally has no information on the diagonal, different choices are possible for filling in the diagonal. Since the only reason for filling in the diag-

*The author is indebted to A. Kimball Romney for valuable help and advice. The Editor and anonymous reviewers made numerous suggestions throughout, particularly with regard to interpreting correspondence analysis without appealing to departures from a model and to issues of diagonal treatment.

onal is practical—to have a complete matrix for the singular value decomposition (SVD)—an alternative to a greater understanding of the effects of different choices for the diagonal entries would be to use a method which allowed for cells with missing entries. de Leeuw and van der Heijden (1988) describe correspondence analysis on tables with structural zeros or missing cells. They discuss situations in which this approach is preferable to the usual correspondence analysis and how programs for correspondence analysis may be adapted for the incomplete table case. If the usual correspondence analysis is interpreted as a decomposition of departures from the model of independence, in the case of missing cells correspondence analysis is interpreted as a decomposition of departures from quasi-independence, or independence for the included cells (Bishop, Fienberg, and Holland, 1975). Following the notation of de Leeuw and van der Heijden (1988) the usual correspondence analysis is based on the SVD of

$$D_r^{-1/2}(P-E)D_c^{-1/2}$$

where P is the observed matrix of cell proportions (the observed table divided by the sum of all cell entries), E is the matrix of estimated proportions under the model of independence, and the matrices D are diagonal matrices of the row (D_r) and column (D_c) margins, again in proportion form. The proportion form used in this notation is not required; van der Heijden and de Leeuw (1985) give the notation in terms of the observed cell entries and the cell estimates under independence. From the SVD $U\Lambda V'$ unweighted row and column scores (standard coordinates) are obtained by

$$R = D_r^{-1/2}U, \qquad C = D_c^{-1/2}V.$$

For the case of missing cells correspondence analysis is based on the SVD of

$$S_r^{-1/2}(P-Q)S_c^{-1/2}$$

where Q is the matrix of estimates of cell proportions under the model of quasi-independence. For the missing cells (the diagonal for the sociomatrix) the difference between P and Q is zero. The matrices S are analogous to the D matrices above and are used in the same way to form row and column scores, but are constructed from the "quasi-margins" estimated in fitting the quasi-independence model

$$\pi_{ij} = \alpha_i\beta_j$$

for included cells (i,j).

In a sense this replaces the convention of placing ones on the diagonal in the original data with the convention of placing zeros in the diagonal of the matrix to be submitted to the SVD. And it must be noted that in lower dimensional approximations to $(P-Q)$ obtained from the decomposition the diagonal entries will not be zero. But, in the perspective of decomposing departures from a model of independence, it is desirable to obtain appropriate model estimates for included cells, and to say that departures from the model are the same and zero for non-included cells. If one were simply fitting a model to a table with no information on the diagonal, quasi-independence would be unmistakably preferred to independence. Since the replacement of independence with quasi-independence does not introduce any practical difficulties, the incomplete table correspondence analysis is more satisfactory than the usual correspondence analysis for the sociomatrix.

CELL ESTIMATES GREATER THAN ONE

This is at least one approach to dealing with the missing diagonal in the sociomatrix. But there is another issue to consider. The notation above is consistent with the interpretation of correspondence analysis as decomposition of departures from the model of independence (or with missing cells, quasi-independence). In social network applications this is often desirable as a description of structure beyond that accounted for by the margins, although because row and column scores are selected so that their sum weighted by the margins is zero the margins still exert an influence (Clogg and Rao, 1991). Independence (quasi-independence) is taken as a baseline model. The problem is that with some sets of margins estimates of cell counts (rather than the proportions used in the notation above) may exceed one under this baseline model.

Different conceptions of the sociomatrix would lead to different degrees of concern with this issue. One perspective is that the sociomatrix represents one realization of a "true" probability model describing the probability of observing an arc between two nodes—this perspective is consistent with many statistical approaches to social network data (Holland and Leinhardt, 1981). In this perspective there could be many replications, and the long-run proportion of replications in which an arc is observed would be that probability given by the true model. Our models would attempt to estimate these true probabilities. If this perspective is taken a suitable baseline model might be required to prevent estimates of the count in a cell from exceeding one, since a sensible estimate of the probability of observing an arc cannot exceed one.

The importance of this requirement also depends on how important one finds the interpretation of correspondence analysis as a decomposition of departures from the model of independence. de Leeuw and van der Heijden (1988) show, for instance, that incomplete table correspondence analysis can be motivated without the image of departures from quasi-independence. The issue of inappropriateness of estimates under a model would arise in many applications outside social networks. Nishisato (1994) gives an extensive treatment of different types of "incidence data" which may be submitted to correspondence analysis. Certainly analysts obtain interpretable results with the usual correspondence analysis in settings in which the model of independence can give inappropriate estimates. Even so, straightforward alternatives to independence which give more appropriate estimates are worth exploring.

If one decides to choose a baseline model which keeps cell estimates between zero and one, generalized correspondence analysis (Escofier, 1983; van der Heijden and de Leeuw, 1985) can be applied to decompose departures from the baseline model. In the notation of van der Heijden and de Leeuw, for an appropriate choice of matrices S the SVD is applied to

$$S_r^{-1/2}(G_1 - G_2)S_c^{-1/2}.$$

In the present case G_1 would be the observed table and G_2 would be estimated under some model. Note that, unlike the notation given for the usual and incomplete table correspondence analysis above, the notation here refers to the table as composed of counts rather than proportions—this difference is also reflected in the matrices S. It does lead to a slight difference in forming row and column scores—the R and C given above are multiplied by the square root of the trace of S_r and S_c respectively. The incomplete table correspondence analysis is a special case of generalized correspondence analysis, in which the difference between the observed table and estimates under quasi-independence is decomposed.

For the sociomatrix the p1 model introduced by Holland and Leinhardt (1981) assumes that dyads are independent, and that for a given dyad the probability of observing both, one, or neither of the possible ties is determined by the propensities of the involved actors to send and receive ties, as well as a system wide propensity to reciprocate ties. In the notation of Holland and Leinhardt these are represented by α, β, and ρ. Under p1 the probability of observing sociomatrix x is given by

$$\exp\left(\rho m + \theta x_{++} + \sum_i \alpha_i x_{i+} + \sum_j \beta_j x_{+j}\right) \times \prod_{i<j} n_{ij}$$

where the n_{ij} simply make probabilities sum to one. Holland and Leinhardt develop a number of variations on p1, include omission of the mutuality parameter ρ. If ρ is omitted (set equal to zero) the model implies the stronger condition of independence of all arcs rather than dyadic independence, as there is no propensity to reciprocate ties. In this case the log odds of an arc from i to j may be written

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \theta + \alpha_i + \beta_j.$$

This seems an appropriate choice for a baseline, as α and β—the marginal effects—determine the structure of probabilities in the sociomatrix. Use of this baseline is thus closely analogous to the use of the model of independence in the usual correspondence analysis, or quasi-independence for the incomplete table, in which the expected values under the model are determined by the margins, but with the requirement that estimates of the probability of a tie do not exceed one. In the case of the sociomatrix it seems reasonable to use the explicitly sociometric p1 less mutuality specification rather than the independence specification.

This model fits in-degrees and out-degrees (row and column margins) exactly. Since in-degrees and out-degrees are fit exactly, the sociomatrix and the estimates under the p1 model less the mutuality parameter will have the same margins, and one of the conditions of van der Heijden and de Leeuw (1985, p. 444) for interpretation is met. To apply generalized correspondence analysis G_2 would be taken to be estimates under p1 less mutuality. With regard to the matrices S in this case there is a difficulty. The expression above gives the log odds of an arc, rather than the log of the probability of an arc. Then there is no way to write the cell probability itself, rather than the log odds, as a product of "quasi-margins" as in quasi-independence. This is needed for the usual inertia interpretation to hold (de Leeuw and van der Heijden, 1988). In general van der Heijden and de Leeuw (1985) advise use of the usual margins to form S matrices, attributing this instruction to Escofier (1983). This choice of S meets another of their conditions for interpretation. So while generalized correspondence analysis based on p1 less the mutuality parameter seems to address some concerns discussed above, it is still not entirely satisfactory.

RESIDUAL SCALING

Novak and Hoffman (1990) describe why generalized correspondence analysis may not always be appropriate, pointing out the particularly important difficulty that the squared singular values need not represent a partitioning of chi-square, thus disrupting the usual interpretation of inertia. This is the case for models which cannot be expressed as independence or quasi-independence on a two-way table: " ... the inertia in GCA is not expressible as a chi-square statistic unless the model is independence expressible, and could have been analyzed with ordinary CA (Novak and Hoffman, 1990, p. 355)". As noted above, for the sociomatrix expected values under p1 less mutuality cannot be written as the product of row and column margins or quasi-margins. Instead of generalized correspondence analysis Novak and Hoffman recommend a very straightforward approach called residual scaling. The SVD is simply applied to a matrix of residuals (standardized in some way) from a model. They make clear that for a complete table, the usual chi-square residual and the independence model, residual scaling is identical to the usual correspondence analysis. In that case standardizing of residuals has the same effect as pre- and post-multiplication by diagonal matrices involving the margins.

Novak and Hoffman (1990) note that if this procedure uses a given standardization, the corresponding fit statistic will be decomposed and inertia is interpretable. Novak and Hoffman recommend using the deviance residuals which compose the likelihood ratio chi-square (G^2). In the case of the sociomatrix the many zero cells seem to suggest that the usual standardized residuals, which compose Pearson's chi-squared (X^2). Again using the notation of G_1 for the observed table and G_2 for estimates under some model, residual scaling with Pearson's standardized residuals is an SVD of the matrix with elements

$$\frac{G_{1ij} - G_{2ij}}{G_{2ij}^{1/2}}.$$

It is natural to assign zero residuals to cells with structural zeros, if they are thought of as fit exactly. In the sociomatrix case, then, the diagonal would be given values of zero in the matrix of standardized residuals. For the sociomatrix the same models as were appropriate for generalized correspondence analysis are appropriate for residual scaling, and p1 less the mutuality parameter seems particularly appealing. Row and column scores are

$$R = U\Lambda, \qquad C = V\Lambda.$$

The three alternatives to the usual correspondence analysis all address the concern with the diagonal. If the issue of cell estimates greater than one is not deemed important the incomplete table correspondence analysis is satisfactory. Generalized correspondence analysis allows the use of the better baseline model p1 less the mutuality parameter, but because this model is not independence expressible the usual inertia interpretation is lost. Residual scaling based on p1 less mutuality addresses this difficulty and is appealingly simple.

TABLE 1
Information Exchange Sociomatrix, Knoke and Kuklinski (1982)

—	1	0	0	1	0	1	0	1	0
1	—	1	1	1	0	1	1	1	0
0	1	—	1	1	1	1	0	0	1
1	1	0	—	1	0	1	0	0	0
1	1	1	1	—	0	1	1	1	1
0	0	1	0	0	—	1	0	1	0
0	1	0	1	1	0	—	0	0	0
1	1	0	1	1	0	1	—	1	0
0	1	0	0	1	0	1	0	—	0
1	1	1	0	1	0	1	0	0	—

TABLE 2
Cell Estimates for Information Exchange Sociomatrix Under Quasi-Independence, With Estimated α and β

Quasi-Independence									
—	.762	.366	.438	.783	.085	.777	.182	.429	.178
.835	—	.697	.835	1.491	.162	1.479	.346	.817	.339
.646	1.122	—	.646	1.154	.126	1.144	.268	.632	.262
.438	.762	.366	—	.783	.085	.777	.182	.429	.178
.959	1.667	.801	.959	—	.187	1.700	.398	.939	.390
.302	.525	.252	.302	.540	—	.535	.125	.296	.123
.359	.624	.300	.359	.642	.070	—	.149	.351	.146
.618	1.074	.516	.618	1.104	.120	1.095	—	.605	.251
.328	.570	.274	.328	.586	.064	.581	.136	—	.133
.514	.894	.429	.514	.919	.100	.912	.213	.503	—
Estimated α									
.080	.152	.118	.080	.175	.055	.065	.113	.060	.094
Estimated β									
.112	.195	.093	.112	.200	.022	.198	.046	.110	.045

EXAMPLE

It is instructive to examine the results of these different approaches for a particular data set. Table 1 gives the 10 actor sociomatrix representing information exchange in a set of organizations from Knoke and Kuklinski (1982, p. 44). One feature of this sociomatrix is that one actor receives ties from all others. This means that in the p1 less mutuality baseline these cells are all fit exactly. Different results might obtain if this actor were treated differently or excluded. Here the whole sociomatrix is analyzed, although this leads to column scores of zero on all dimensions for this actor in the techniques based on p1 less mutuality. The usual correspondence analysis (CA) with the diagonal filled with ones, the incomplete table correspondence analysis with the missing diagonal (ICA), and both generalized correspondence analysis (GCAP1LM) and residual scaling (RSP1LM) based on p1 less mutuality were applied to this sociomatrix. Table 2 gives cell estimates for quasi-independence and the estimated α and β. Note that in keeping with the notation above the estimates for quasi-independence are not used directly but rather estimates of π_{ij}, or m_{ij} divided

TABLE 3
Cell Estimates for Information Sociomatrix Under p1 Less Mutuality

p1 Less Mutuality									
—	.898	.304	.396	.899	.017	1.000	.065	.362	.059
.925	—	.891	.925	.994	.245	1.000	.567	.914	.538
.799	.982	—	.799	.982	.095	1.000	.297	.774	.273
.396	.898	.304	—	.899	.017	1.000	.065	.362	.059
.976	.998	.964	.976	—	.514	1.000	.810	.972	.791
.153	.708	.107	.153	.709	—	1.000	.019	.135	.017
.320	.864	.239	.320	.865	.012	—	.048	.290	.043
.726	.973	.639	.726	.973	.065	1.000	—	.697	.201
.174	.740	.123	.174	.741	.006	1.000	.022	—	.020
.531	.938	.430	.531	.939	.029	1.000	.107	.495	—

TABLE 4
Cumulative Proportion of Sum of First Five Squared Singular Values for the Various Techniques in the Example

	Dimensions				
	1	2	3	4	5
Technique					
CA	.406	.625	.794	.921	.973
ICA	.349	.603	.766	.873	.928
GCAP1LM	.433	.719	.834	.916	.959
RSP1LM	.445	.783	.878	.945	.973

by the sum of all cells (here 49). The π_{ij} may also be obtained by multiplying α_i and β_j. Table 3 gives cell estimates for p1 less mutuality.

The first comparison of interest is in the dimensionality suggested by the different techniques. Table 4 gives the cumulative proportion of the sum squared singular values for the first five dimensions. As mentioned earlier, in this case the usual inertia interpretation is inappropriate for the generalized correspondence analysis. The overall impression is that differences are apparent but not enormous. The first two dimensions account for quite a bit more in the p1 less mutuality residual scaling and generalized correspondence analysis than in the usual and incomplete table correspondence analysis, but all approaches suggest that more than two dimensions are necessary to characterize the data. Residual scaling has the most and incomplete table correspondence analysis the least accounted for at each cumulation until the last in which the usual correspondence analysis and residual scaling each have a proportion of .973.

The other main area of interest is in the plots produced by the different techniques. In the correspondence analysis literature there is considerable controversy about methods of displaying results (Carroll, Green and Schaffer, 1986; Greenacre and Hastie, 1987). Different methods have different implications for the interpretation of distances in the plots and the possibility of interpreting row and column points in the same space. Kumbasar, Romney, and Batchelder (1994) use a modification of a display of row and column scores weighted by the square root of the appropriate singular value, in the symbols above $R\Lambda^{1/2}$ and $C\Lambda^{1/2}$. This is consistent with simultaneous representation (c) in van der Heijden and de Leeuw (1985) and simultaneous representation (e) in Carroll, Green, and Schaffer (1987). For the example here the possibility of displaying row and column points in the same plot will not be considered and row points and column points will be treated separately. When row scores are weighted by the appropriate singular value rather than its square root—as $R\Lambda$ and $C\Lambda$—distances between row points in the plot approximate chi-square distance between rows of the table. The equivalent holds for column points. This is consistent with simultaneous representations (a) and (b) in van der Heijden and de Leeuw (1985) and simultaneous representations (b), (c), and (d) in Carroll, Green and Schaffer (1987). Since here row and column points will only be considered separately, the choice of a particular simultaneous representation is not important.

For sociomatrix applications this may be thought of as placing actors close together which have similar patterns of ties to others (in the plot of row points) or ties from others (in the plot of column points). In the example here weighting of scores by Λ is used for correspondence analysis, incomplete table correspondence analysis, and generalized correspondence analysis. For residual scaling, with the row and column scores just as written above distances between row points (or between column points) are approximations of Euclidean distance between the rows (or columns) of the matrix of standardized residuals (Novak and Hoffman, 1990). For the sociomatrix, again actors are placed together who have similar patterns of ties to or from others.

For practical reasons most displays focus on the first two dimensions. So, although in this example two dimensions are not adequate to represent the structure in the data, comparisons between techniques will be based on the first and second dimensions. Comparison of the plots themselves is difficult. It is difficult to determine visually whether one pair of plots is more or less similar than another pair. Kumbasar, Romney, and Batchelder (1994) mention some issues in comparing plots. One simple way to compare the plots is to compute interpoint distances in the two dimensional representation produced by each technique, and then compute correlations among these sets of interpoint distances.

Table 5 presents row points and Table 6 column points (using the weighting described above) for the first two dimensions. Table 7 gives correlations among interpoint distances for the resulting plots in the different techniques. The correlations are all relatively high, and extremely high in some cases. The results for plots of row points seem to suggest that ICA, GCAP1LM, and RSP1LM give substantially similar images with CA somewhat different from the rest. For column points it seems that CA, ICA, and GCAP1LM give similar results with RSP1LM somewhat different from the rest. Column results are probably harder to interpret for this sociomatrix than row results because of the one actor which received ties from all others.

The different techniques do not seem to yield radically different views of structure in this sociomatrix. Of course this is not necessarily a general result, and is based on the crude method of correlating interpoint distances. Often interest is greatest in one actor or one set of actors. Particular interpretations of one actor's position and its proximity to specific other actors may be different even in the face of a high

TABLE 5
Row Points in First Two Dimensions from Different Techniques in Example

CA									
−.34	−.19	.72	−.46	−.07	1.43	−.36	−.43	−.20	.11
.37	.18	−.46	−.19	−.15	.70	−.20	.27	.50	−.59
ICA									
−.42	−.28	1.30	−.20	.04	−.34	.11	−.29	−.21	−.12
.10	−.15	−.08	.57	−.20	−1.31	.71	.22	.37	−.10
GCAP1LM									
−.23	−.42	1.34	−.00	−.31	−.55	.32	−.16	.06	−.05
.03	−.11	.27	−.43	−.09	1.50	−.49	−.30	−.17	.14
RSP1LM									
−.48	−.43	1.82	.67	−.24	−3.53	.86	.16	.22	−.24
−.33	−.71	3.02	−.72	−.47	1.48	−.23	−.77	−.16	.04

TABLE 6
Column Points in First Two Dimensions from Different Techniques in Example

CA									
−.44	−.26	.77	−.25	−.26	2.08	.04	−.45	.06	.49
−.05	−.08	−.17	−.25	−.08	.31	.11	.25	.81	−1.06
ICA									
−.38	−.04	.00	.29	−.01	2.40	−.09	−.28	−.51	1.18
.17	.34	−.92	.33	.32	−.34	.03	−.30	−.53	−.31
GCAP1LM									
−.44	.10	−.23	.22	.10	2.74	0	−.58	−.67	1.09
−.25	−.34	1.05	−.30	−.34	.80	0	−.16	.56	.37
RSP1LM									
−.08	.76	−2.54	.77	.76	1.34	0	−.24	−2.65	.70
−1.11	−.38	1.46	−.07	−.39	2.72	0	−.56	.18	1.39

TABLE 7
Correlations Among Interpoint Distances in Two Dimensions for the Different Techniques in Example

	CA	ICA	GCAP1LM	RSP1LM
Row Points				
CA	—	.77	.82	.81
ICA		—	.96	.95
GCAP1LM			—	.99
RSP1LM				—
Column Points				
CA	—	.91	.92	.69
ICA		—	.98	.73
GCAP1LM			—	.77
RSP1LM				—

overall correlation between interpoint distances in two plots. Also the singular values suggest that more than two dimensions need to be examined. If the results are taken at face value, at least two interpretations are possible. One is that because results of the usual correspondence analysis do not differ greatly from those of the other techniques, there is no reason to pursue the alternatives. Another interpretation is that the results from other techniques are similar enough that the substantive appeal of correspondence analysis is not lost in the other approaches. Since the underpinnings of the other techniques are more satisfactory with regard to the issues discussed above, and are not fundamentally harder to understand, the other approaches may be attractive alternatives to the usual correspondence analysis.

REFERENCES

Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975) *Discrete Multivariate Analysis*, Cambridge, MA: MIT.

Carroll, J. D., Green, P. E., and Schaffer, C. M. (1986) Interpoint distance comparisons in correspondence analysis. *Journal of Marketing Research* **23:** 271–280.

Carroll, J. D., Green, P. E., and Schaffer, C. M. (1987) Comparing interpoint distances in correspondence analysis: A clarification. *Journal of Marketing Research* **24:** 445–450.

Clogg, C. C., and Rao, C. R. (1991) Comment on Goodman. *Journal of the American Statistical Association* **86:** 1118–1120.

de Leeuw, J., and van der Heijden, P. G. M. (1988) Correspondence analysis of incomplete contingency tables. *Psychometrika* **53:** 223–233.

Escofier, B. (1983) Analyse de la difference entre deux mesures sur le produit de deux memes ensembles [Analysis of the difference between two measures on the product of two identical sets]. *Cahiers de l'Analyse des Donnees* **3:** 325–329.

Greenacre, M., and Hastie, T. (1987) The geometric interpretation of correspondence analysis. *Journal of the American Statistical Association* **82:** 437–447.

Holland, P. W., and Leinhardt, S. (1981) An exponential family of probability distributions for directed graphs (with discussion). *Journal of the American Statistical Association* **76:** 33–65.

Knoke, D., and Kuklinski, J. H. (1982) *Network Analysis*, Beverly Hills, CA: Sage.

Kumbasar, E., Romney, A. K., and Batchelder, W. H. (1994) Systematic biases in social perception. *American Journal of Sociology* **100:** 477–505.

Nishisato, S. (1994) *Elements of Dual Scaling: An Introduction to Practical Data Analysis*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Noma, E., and Smith, D. R. (1985) Scaling sociomatrices by optimizing an explicit function: Correspondence analysis of binary single response sociomatrices. *Multivariate Behavioral Research* **20:** 179–197.

Novak, T. P., and Hoffman, D. L. (1990) Residual scaling: An alternative to correspondence analysis for the graphical representation of residuals from log-linear models. *Multivariate Behavioral Research* **25:** 351–370.

van der Heijden, P. G. M., and de Leeuw, J. (1985) Correspondence analysis used complementary to loglinear analysis. *Psychometrika* **50:** 429–477.

Wasserman, S. S., and Faust, K. (1989) Canonical analysis of the composition and structure of social networks, in Clogg, C. C. (ed.) *Sociological Methodology*, Cambridge, MA: Basil Blackwell, 1–42.

Wasserman, S. S., Faust, K., and Galaskiewicz, J. (1990) Correspondence and canonical analysis of relational data. *Journal of Mathematical Sociology* **15:** 11–64.

Weller, S. C., and Romney, A. K. (1990) *Metric Scaling*, Newbury Park, CA: Sage.