Applied Multivariate Research, Volume 13, No. 1, 2009, 1-3

MULTIVARIATE MODELS FOR THREE-DIMENSIONAL DATA

James W. Grice
Oklahoma State University

As early as the publication of Raymond Cattell's (1966) 'basic data relation matrix', or 'data box' researchers have been interested in data structures that extend beyond two dimensions. The development of statistical and mathematical models as well as accessible computer software for the analysis of three-dimensional data structures, particularly, has accelerated in recent years. A three-dimensional data structure might, for example, be comprised of 500 persons who rated themselves on 50 items from a personality questionnaire on 10 different occasions. The resulting $500 \times 50 \times 10$ data cube could be analyzed to explore the multivariate relationships between the three *facets*: persons, items, occasions. One general strategy would entail applying traditional factor analyses to different pairwise combinations of the facets by collapsing over the third facet. Gorsuch (1983, pp. 322-327) describes a number of variations on this approach. Another strategy, originating with Tucker (1964), would entail factoring all three facets simultaneously, yielding a 'core matrix' that describes their interrelations.

In this special issue of *Applied Multivariate Research* Kroonenberg, Harshman, and Murakami offer an excellent review of Tucker's original model and also compare it to Harshman's (Harshman & Lundy, 1984) popular Parafac model for analyzing three-dimensional data matrices. They furthermore provide clear guidance regarding practical decisions that must be made when employing either model, and they use a genuine parenting styles data set to exemplify the issues. While the models may appear complex at first glance, Kroonenberg and his colleagues show how the analyses parsimoniously uncover common styles of parenting while simultaneously revealing individual variability between families.

Leenen and Ceulemans also compare and contrast two different models for analyzing thee-dimensional data matrices. Their paper, however, is centered around the HICLAS method introduced by De Boeck and Rosenberg (1988) for modeling the hierarchical relations among binary variables. HICLAS essentially weds a form of binary factor analysis with set theory, and in this issue Leenen and Ceulemans compare and contrast two different models: INDCLAS and Tucker-3 HICLAS. Like Kroonenber's paper, their contribution reveals Tucker's legacy to analyzing three-dimensional data matrices. Leenen and Ceulemans analysis of two genuine data sets suggest that the Tucker-3 HICLAS model may yield more parsimonious results than the INDCLAS model. In an interesting example that is near and dear to my own heart, given my interests in person-centered statistics (see Grice, 2007), the authors also show how their technique can be applied to a single case. The intra-personal perceptions (or 'object relations') of a young woman are explored, and some interesting findings regarding the stability of her emotional relations to her father and mother are revealed through the analysis.

As another strategy for analyzing three-dimensional binary data in this special issue, González, Tuerlinckx, and De Boeck discuss the 2sSEM model, which is essentially a log-linear approach with Bayesian estimation. They demonstrate how the 2sSEM model can be used to model the structure of a *sociomatrix*, which is a binary matrix that records the links between different people in a social network. After pointing out the limitations of aggregating across dimensions to analyze three-dimensional data and then presenting the formal model, they demonstrate their technique using data from secondary students. González et al.'s results reveal very interesting asymmetries in the students' perceptions of themselves regarding their popularity. The results furthermore confirm the loss of information that can occur when three-dimensional data structures are aggregated and analyzed with traditional two-dimensional approaches.

Finally, the similarities between Cattell's original data box and Generalizability Theory are obvious, including the use of the term 'facets' to refer to the dimensions that constitute the box. In their landmark book on Generalizability Theory, Cronbach, Gleser, Nanda, and Rajaratnam (1972) viewed analysis of variance (ANOVA) as a method for wedding data boxes to psychometric theory. Similarly, Gower (1975) introduced Generalized Procrustes Analysis (GPA) as a technique for exploring the structure of three-dimensional data matrices using ANOVA. In this special issue Grice and Assad explain the fundamental features of GPA, including the centrality of the 'consensus matrix' which is essentially the average of a series of rotated two-dimensional matrices. ANOVA is used to assess the variability of the different two-dimensional data matrices about the consensus matrix along the third facet in the data matrix. Grice and Assad describe the results of a study in which a small group of students were asked to rate themselves and each of their peers on Big Five trait scales. Results from the GPA revealed a moderate degree of similarity among the trait profiles even though the students had known each other for only seven days.

This special issue thus brings together a variety of different methods and statistical techniques for analyzing three-dimensional data matrices. They are common to the extent they permit researchers to extend beyond examining only two dimensions and to discover multivariate relationships that would otherwise be masked in aggregates or untapped by simpler analyses. In other words, these models open up new avenues of research. Of course with similarity, one must also find difference, and the methods in this special issue are not exempt from this dialectic. It could be argued that they differ primarily with regard to the statistical models underlying the data analysis and secondarily with regard to their technical features (e.g., examining classes with the HICLAS models). Perhaps this statement is an oversimplification, but a thorough and insightful comparison of the techniques is beyond the scope of this brief introduction. Such comparisons would certainly be beneficial, much like the comparisons that have been made between Multidimensional Scaling and Factor Analysis over the years. It is our hope that this special issue will spur such comparisons as well as novel investigations in psychology, marketing, education, and other domains of empirical thought where three-dimensional data structures are encountered.

REFERENCES

Cattell, R. B. (1966). The data box: Its ordering of total resources in terms of possible relational systems. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology*. Chicago: Rand-McNally.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

De Boeck, P., & Rosenberg, S. (1988). Hierarchical classes: Model and data analysis. *Psychometrika, 53*, 361-381.

Gorsuch, R. L. (1983). *Factor analysis*. New Jersey: Erlbaum.

Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika, 40*, 33-51.

Grice, J. W. (2007). Person-centered structural analyses. In R. Robins, C. Fraley, and R. Krueger (Eds.) *Handbook of research methods in personality psychology* (pp. 557-572). New York: Guilford Press.

Harshman, R. A., & Lundy, M. E. (1984). The PARAFAC model for three-way factor analysis and multidimensional scaling. In H. G. Law, C. W. Snyder Jr., J. A. Hattie, & R. P. McDonald (Eds.), *Research methods for multimode data analysis* (pp. 122-215). New York: Praeger.

Tucker, L. R. (1964). The extension of factor analysis to three-dimensional matrices. In N. Frederiksen & H. Gulliksen (Eds.), *Contributions to mathematical psychology*. New York: Holt, Rinehart & Winston.