Lundy and Kruskal, 1985; Kruskal, Harshman and Lundy, 1989; Lundy, Harshman and Kruskal, 1989).

Krijnen and Ten Berge (1991) developed variants of the basic PARAFAC algorithm to put nonnegativity constraints on the solution by using special least squares regression algorithms from Lawson and Hanson (1974). Durrell et al. (1990) refer to programs for three-way and four-way PARAFAC models (Lee, 1988) which also included nonnegativity constraints.

3.6 Additional Issues

In the above sections, the general focus has been on models and algorithms, but there are several issues in connection with these models which have not been mentioned so far. Very prominent, for instance, in Harshman's work, has been the question of preprocessing (i.e., centering and standardisation) of the data before the three-way analysis. Harshman and Lundy (1984b) discuss this issue in great detail touching on both algebraic and practical aspects (see also Kroonenberg, 1983). Ten Berge and Kiers (1989) and Ten Berge (1989) provide some theoretical results with respect to the iterative centering and standardisation proposed by Harshman and Lundy.

Another issue in this context is the postprocessing of output, that is, representation, graphing and transformations of the basic output of the programs to enhance interpretability (see especially Harshman and Lundy, 1984b; Kroonenberg, 1983).

Smilde (1992) raises the issue of variable selection for three-way data, as well as the problem of nonlinearities in the data and their effect on the solutions. These issues can also be seen as a serious concern in such areas like ecology where nonlinearities are the rule rather than the exception (see, e.g., Faith, Minchin and Belbin, 1987).

A final point is that within the framework of the analysis of covariance structures, McDonald (1984) has discussed the PARAFAC model, cited its limitations and proposed an altogether different (stochastic) approach to the kind of three-way data psychologists often encounter.

4. CONCLUSION

With the above comments, I have attempted to give a rough outline of research on the PARAFAC model. The model itself is only one of several conceivable models for three-way data, but a fully fledged exposé is not feasible here. What makes the PARAFAC model special is that it has a unique solution, a situation which is fairly unique in three-way land.

Comment

Donald S. Burdick

Multilinear models are fascinating because of the richness of their mathematical structure and the usefulness of their applications. The authors have done a fine job of presenting both of these features. I welcome their paper and hope that it has the effect of stimulating interest in this important topic.

Having said that, I must add my opinion that it is a mistake to shy away from tensors. The geometry of tensor products can be a source of valuable insight when struggling with the complicated details of multilinear algebra. The geometric perspective is especially useful when trying to make sense out of the nonuniqueness that occurs when model parameters are not identifiable.

For example, the concept of tensor products of vector spaces can shed light on the structure of the T3 model. Let $\boldsymbol{Y}$ denote an $I \times J \times K$ data array and write

$$\boldsymbol{Y} = \mu + \mathrm{e}$$

where μ is given by (19). The data array $\boldsymbol{Y}$ is unconstrained, which is tantamount to saying that $\boldsymbol{Y}$ is an arbitrary vector in $R^I \otimes R^J \otimes \mathrm{R}^K$, the tensor product of real Euclidean spaces of dimensions I, J and K, respectively. The array μ, however, is constrained by expression (19). What is the nature of that constraint? Expression (19) stipulates that μ lie in $\mathcal{A} \otimes \mathcal{B} \otimes \mathcal{C}$, where $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are the respective subspaces of R^I, R^J and R^K spanned by the columns of $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{\Gamma}$, respectively. The least squares fit of μ to $\boldsymbol{Y}$ is the projection of $\boldsymbol{Y}$ on $\mathcal{A} \otimes \mathcal{B} \otimes \mathcal{C}$. From a geometric perspective, the nonidentifiability is obvious, because the projection of a data vector on a subspace is unaffected by changes in the basis spanning the subspace. Replacing $\boldsymbol{A}$ by $\boldsymbol{AM}$ amounts to no more than a change of basis for $\mathcal{A}$.

Donald S. Burdick is Associate Professor, Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina 27706.

The perspective of tensor geometry is even more helpful in acquiring insight into the T2 model. The constraint implied by (25) or (26) is that μ lie in $\mathcal{A} \otimes \mathcal{B} \otimes R^K$. Other examples of the value of the geometric perspective could be cited, but perhaps these will suffice to make the point.

There is, I admit, some validity to the authors' stated reason for avoiding tensor terminology. Much of the extant literature on tensors is written either for physicists or for mathematicians, and neither slant is particularly well suited for the statistical applications at hand. There is a tendency for the treatments to be on the one hand too abstract and on the other too specialized because of the focus on tensor powers of R^N. The concepts of covariance and contravariance, for example, arise when a tensor is regarded as an abstraction whose numerical representation is the result of an arbitrarily chosen coordinate system. The chemometrician need not struggle to understand these concepts. The numbers in his or her arrays are real data and not just coordinates with respect to some arbitrary basis.

Contrary to the impression one might get from books, the theoretical underpinnings of basic tensor product geometry are not difficult to comprehend. A key idea is the tensor product mapping, of which the outer product is a canonical example. The feature which distinguishes a tensor product mapping from other bilinear functions with two vector arguments, for example, the inner product, is its preservation of linear independence. Thus, if the F_1 columns of A are linearly independent and the F_2 columns of B are linearly independent, the F_1F_2 outer products of columns of A with columns of B are linearly independent. The span of these F_1F_2 outer products is by definition the tensor product $\mathcal{A} \otimes \mathcal{B}$, where $\mathcal{A}$ is the span of the columns of A and $\mathcal{B}$ is the span of the columns of B.

This presentation omits some mopping up details that are required for a fully rigorous definition of the tensor product of vector spaces. The essential features of the concept, however, are there. Anyone who can follow the development in the previous paragraph has a handle on a set of conceptual tools that can provide a valuable framework for interpreting multilinear models. These conceptual tools require names. We can use the existing tensor terminology or invent new terms. Others may have added bells and whistles that we don't want or need, but the terminology of tensors exists for the tools we do need, and it has been around for years. I say we should use it.

Rejoinder

Sue Leurgans and Robert T. Ross

We thank the editors for securing the comments of Kroonenberg and of deLeeuw, both of whom have contributed to the development and application of multiway methods in psychometrics, and of Burdick, a statistician with chemistry collaborators. Much of the work on multilinear models is deeply embedded in subject matter, and many contributions have been made outside the single application (spectroscopy) we have emphasized here. We thank the discussants for adding their views of important contributions in a variety of areas.

Our reply consists of a section with comments on mathematical issues raised by all discussants and a section of specific responses to selected points raised by each discussant.

1. GENERAL COMMENTS

We think one reason for the increased interest in arrays in recent decades is that arrays are no longer theoretical abstractions, but can be defined and manipulated in many high-level languages. The statistical package S (Becker, Chambers and Wilks, 1988) is just one of the programs that permits array calculations. Besides changing the nature of the questions that are important, the ability to calculate easily enables one to check conjectures.

The mathematical theory of arrays that needs to be applied to these questions requires some features that are more general than most descriptions of tensors. The most basic requirements are that the number of levels in each way of the array be arbitrary and not necessarily equal and that the elements of the array not be required to satisfy any symmetry assumptions. The different ways of the array need to be treated symmetrically in mathematics, although different applications may treat the ways differently. The notation employed will vary with the purpose of the exposition; see our reply to Burdick below. Other examples of publications on array results are Knuth (1965) and Lickteig (1985) [from the Geladi (1989) paper cited by Kroonenberg] and Lickteig's references.

The basic question about arrays is how to approximate an array by a simpler array or how to decompose