PSYCHOMETRIKA—VOL. 51, NO. 2, 269–275
JUNE 1986

AN APPROACH TO n-MODE COMPONENTS ANALYSIS

ARIE KAPTEYN

TILBURG UNIVERSITY

HEINZ NEUDECKER

UNIVERSITY OF AMSTERDAM

TOM WANSBEEK

GRONINGEN UNIVERSITY

As an extension of Lastovicka's four-mode components analysis an n-mode components analysis is developed. Using a convenient notation, both a canonical and a least squares solution are derived. The relation between both solutions and their computational aspects are discussed.

Key words: principal components, multidimensional matrices.

Introduction

In a recent article, Lastovicka (1981) extends Tucker's three-mode factor analysis model to four modes and presents a "canonical" solution. In this paper we extend his research in three respects:

(i) We extend his results to an arbitrary number of modes;
(ii) We introduce some new notation, fit for handling multidimensional matrices, and show how the calculus involved works;
(iii) We derive a least squares (LS) solution and compare it to Lastovicka's solution.

The problem and the notation are introduced first. We then present the LS solution and compare it with Lastovicka's solution, which we will denote by the term "canonical." An iteration method to compute LS is given next. We compare the canonical and the LS solution by means of an empirical example.

Apart from Lastovicka's work cited above, various other related references can be mentioned. Carroll and Chang (1970) present methods for analyzing multidimensional matrices. This work is restrictive in the sense that the component dimensionality of each mode is taken equal. Carroll, Pruzansky and Kruskal (1980) extend these results to include linear constraints on the parameters. The present paper can be considered to be an n-mode extension of the work of Kroonenberg and de Leeuw (1980) on three-mode data, which itself is a further development of Tucker's (1966) work on three-mode principal components. They also give an algorithm (TUCKALS3) based on what they call alternating least squares. Hence, Kroonenberg (1983), in his comprehensive monograph on

Requests for reprints should be sent to Tom Wansbeek, Econometrics Institute, Groningen University, PO Box 800, 9700 AV Groningen, THE NETHERLANDS.

The first draft was written while Wansbeek was with the Netherlands Central Bureau of Statistics. We thank Jaap Verhees for performing the computations and for many discussions on the subject, John Lastovicka for kindly making available his data to us, and the Editor, the referees, Jeroen Weesie and Pieter Kroonenberg for their useful comments.

0033-3123/86/0600-5022$00.75/0

three-mode data, describes our model as a "Tucker$n$ model", and our iteration method as a "TUCKALSn algorithm."

We agree with the observation by Kroonenberg (1983): "In principle, the extension to n-mode data is straightforward, but it becomes increasingly complex to keep track of the summations over the proper indices. Moreover, the description of such an n-mode procedure becomes exceedingly cumbersome without new notation" (p. 73). These are precisely the points where the present paper might offer a useful contribution.

The Problem and Some Notation

Let $\mathbf{A}_i$ be an unknown $\ell_i \times m_i$ matrix, $\ell_i \geq m_i$, satisfying $\mathbf{A}_i'\mathbf{A}_i = \mathbf{I}_{m_i}$, $i = 1, \ldots, n$, and define

$$\mathbf{A} \equiv \mathbf{A}_1 \otimes \cdots \otimes \mathbf{A}_n. \tag{1}$$

The symbol "$\otimes$" refers to the (right) Kronecker product, with, for example, $\mathbf{C} \otimes \mathbf{D} = [c_{ij}\mathbf{D}]$. We do not introduce special notation for the multiple Kronecker product like the one appearing in (1). Let

$$\ell \equiv \prod_{i=1}^{n} \ell_i \tag{2}$$

$$m \equiv \prod_{i=1}^{n} m_i. \tag{3}$$

So, $\ell \geq m$ and $\mathbf{A}$ is of order $\ell \times m$. Let $\mathbf{y}$ be a (known) ℓ-vector, whose elements are identified by n indices, with the i-th index ($i = 1, \ldots, n$) assuming values from 1 to ℓ_i. The elements are arranged in such a way that the first index runs slowly and the last index runs fast. Analogously, let $\mathbf{b}$ be an (unknown) m-vector, also with elements identified by n indices, with the i-th index running from 1 to m_i ($i = 1, \ldots, n$). The elements are arranged in the same way as the elements of $\mathbf{y}$. (The vectors $\mathbf{y}$ and $\mathbf{b}$ can be considered as "stacked" versions of n-dimensional matrices.) It is our aim to choose $\mathbf{A}_i$, $i = 1, \ldots, n$ and $\mathbf{b}$ in such a way that $\mathbf{A}b$ represents y "as well as possible."

Let $\mathbf{C}_i$ be the $\ell \times \ell$ commutation matrix that changes the running order of the indices in such a way that when applied to $\mathbf{y}$ the i-th index runs fastest and the running order of the other indices remains unaffected. Let further $\mathbf{D}_i$ be the $m \times m$ commutation matrix that reshuffles the elements of $\mathbf{b}$ in the same way as does $\mathbf{C}_i$ with the elements of $\mathbf{y}$. The sizes of $\mathbf{C}_i$ and $\mathbf{D}_i$ differ. As $\mathbf{C}_i$ and $\mathbf{D}_i$ are commutation matrices they have the property

$$\mathbf{C}_i\mathbf{C}_i' = \mathbf{C}_i'\mathbf{C}_i = \mathbf{I}_\ell, \tag{4}$$

$$\mathbf{D}_i\mathbf{D}_i' = \mathbf{D}_i'\mathbf{D}_i = \mathbf{I}_m. \tag{5}$$

The $\mathbf{C}_i$'s and $\mathbf{D}_i$'s are generalizations of the well-known commutation matrix for $n = 2$, which has been extensively studied by, among many others, Balestra (1976), Magnus and Neudecker (1979), and Henderson and Searle (1981).

It will be convenient to introduce the matrix $\mathbf{A}^i \equiv \mathbf{A}_1 \otimes \cdots \otimes \mathbf{A}_{i-1} \otimes \mathbf{A}_{i+1} \otimes \cdots \otimes \mathbf{A}_n$. The following definitional relationship then emerges:

$$\mathbf{A}^i \otimes \mathbf{A}_i = \mathbf{C}_i(\mathbf{A}_1 \otimes \cdots \otimes \mathbf{A}_n)\mathbf{D}_i' = \mathbf{C}_i\mathbf{A}\mathbf{D}_i'. \tag{6}$$

For reasons to become clear later on we define the matrices $\mathbf{Y}_i$ and $\mathbf{B}_i$ by

$$\text{vec } \mathbf{Y}_i' \equiv \mathbf{C}_i\mathbf{y} \quad \text{and} \quad \text{vec } \mathbf{B}_i' \equiv \mathbf{D}_i\mathbf{b}, \tag{7}$$

where $\mathbf{Y}_i$ is of order $(\ell^i \times \ell_i)$ and $\mathbf{B}_i$ is of order $(m^i \times m_i)$, with $\ell^i \equiv \ell/\ell_i$ and $m^i \equiv m/m_i$.

It will be clear that in spite of the normalization $\mathbf{A}_i'\mathbf{A}_i = \mathbf{I}_{m_i}$, the representation $(\mathbf{A}_1 \otimes \cdots \otimes \mathbf{A}_n)\mathbf{b}$ is not unique, because if $\mathbf{T}_i$, $i = 1, \ldots, n$, are orthogonal m_i, m_i matrices there holds

$$(\mathbf{A}_1\mathbf{T}_1 \otimes \cdots \otimes \mathbf{A}_n\mathbf{T}_n)(\mathbf{T}_i' \otimes \cdots \otimes \mathbf{T}_n')\mathbf{b} = (\mathbf{A}_1 \otimes \cdots \otimes \mathbf{A}_n)\mathbf{b}. \tag{8}$$

This indeterminacy will be used to pick a convenient solution when discussing the representation problem.

The Least Squares Solution

The LS solution amounts to finding $\mathbf{A}_1, \ldots, \mathbf{A}_n$ and $\mathbf{b}$ such that

$$\underline{Q} \equiv (\mathbf{y} - \mathbf{A}\mathbf{b})'(\mathbf{y} - \mathbf{A}\mathbf{b}) \tag{9}$$

is minimal, subject to $\mathbf{A}_i'\mathbf{A}_i = \mathbf{I}_{m_i}$ and (1). The solution to this problem is readily obtained. First, it follows from standard LS theory that for any choice of $\mathbf{A}_1, \ldots, \mathbf{A}_n$, the solution for b is

$$\hat{\mathbf{b}} = \mathbf{A}'\mathbf{y} = (\mathbf{A}_1 \otimes \cdots \otimes \mathbf{A}_n)'\mathbf{y}. \tag{10}$$

Notice that

$$\begin{aligned} \mathbf{y}'(\mathbf{A}_1 \otimes \cdots \otimes \mathbf{A}_n)\hat{\mathbf{b}} &= \mathbf{y}'(\mathbf{A}_1\mathbf{A}_1' \otimes \cdots \otimes \mathbf{A}_n\mathbf{A}_n')\mathbf{y} \\ &= \mathbf{y}'\mathbf{C}_i'\mathbf{C}_i(\mathbf{A}_1\mathbf{A}_1' \otimes \cdots \mathbf{A}_n\mathbf{A}_n')\mathbf{C}_i'\mathbf{C}_i\mathbf{y} \\ &= (\operatorname{vec} \mathbf{Y}_i')'(\mathbf{A}^i\mathbf{A}^{i\prime} \otimes \mathbf{A}_i\mathbf{A}_i')\operatorname{vec} \mathbf{Y}_i' \\ &= \operatorname{tr} \mathbf{A}_i'\mathbf{Y}_i'\mathbf{A}^i\mathbf{A}^{i\prime}\mathbf{Y}_i\mathbf{A}_i, \end{aligned} \tag{11}$$

for any one index i. It is convenient to define

$$\mathbf{S}_i \equiv \mathbf{Y}_i'\mathbf{A}^i\mathbf{A}^{i\prime}\mathbf{Y}_i. \tag{12}$$

Then

$$Q = \mathbf{y}'\mathbf{y} - \operatorname{tr} \mathbf{A}_i'\mathbf{S}_i\mathbf{A}_i. \tag{13}$$

Minimization of Q subject to $\mathbf{A}_i'\mathbf{A}_i = \mathbf{I}_{m_i}$, $i = 1, \ldots, n$, can be done by differentiating the Lagrangean function:

$$L = \mathbf{y}'\mathbf{y} - \operatorname{tr} \mathbf{A}_i'\mathbf{S}_i\mathbf{A}_i + \operatorname{tr} \sum_{i=1}^{n} \mathbf{F}_i(\mathbf{A}_i'\mathbf{A}_i - \mathbf{I}_{m_i}), \tag{14}$$

with $\mathbf{F}_i$ a symmetric $m_i \times m_i$ matrix of Lagrange multipliers, with respect to $\mathbf{A}_i$ and setting the result equal to zero. This yields:

$$\mathbf{S}_i\mathbf{A}_i - \mathbf{A}_i\mathbf{F}_i = \mathbf{0}. \tag{15}$$

So a solution is to choose for $\mathbf{F}_i$ the diagonal matrix $\hat{\mathbf{F}}_i$ containing the m_i largest eigenvalues of $\mathbf{S}_i$, and for $\mathbf{A}_i$ the corresponding orthonormal eigenvectors.

Since $\mathbf{S}_i$ ($\equiv \mathbf{Y}_i'\mathbf{A}^i\mathbf{A}^{i\prime}\mathbf{Y}_i$) depends on the unknown parameters in $\mathbf{A}^i$ the solution to the minimization problem has to be obtained iteratively. An iteration process and its convergence properties are discussed below.

The fit of the LS solution is assessed as follows. (Carets indicate LS solutions.) Define the ℓ-vector of residuals:

$$\underline{\hat{\mathbf{e}}} \equiv \mathbf{y} - \hat{\mathbf{A}}\hat{\mathbf{b}}, \tag{16}$$

then, analogous to (13) we have:

$$\hat{e}'\hat{e} = y'y - \hat{b}'\hat{b} = y'y - \text{tr } \hat{\mathbf{A}}_i'\hat{\mathbf{S}}_i\hat{\mathbf{A}}_i = y'y - \text{tr } \hat{\mathbf{F}}_i, \tag{17}$$

for any one i. So the sum of the m_i largest eigenvalues in the LS solution is the same for all i. It is obvious to define

$$\underline{R}^2 \equiv \frac{\hat{\mathbf{b}}'\hat{\mathbf{b}}}{\mathbf{y}'\mathbf{y}} = \text{tr } \frac{\hat{\mathbf{F}}_i}{\mathbf{y}'\mathbf{y}} \tag{18}$$

as the coefficient of determination.

The Canonical Solution

Lastovicka (1981) proposes a canonical solution $\tilde{\mathbf{A}}_i$, $i = 1, \ldots, n$, where the columns of $\tilde{\mathbf{A}}_i$ are the m_i orthonormal eigenvectors corresponding to the m_i largest eigenvalues of $\mathbf{Y}_i'\mathbf{Y}_i$. The solution for b is $\tilde{b} = (\tilde{\mathbf{A}}_1' \otimes \cdots \otimes \tilde{\mathbf{A}}_n')y$. The LS and the canonical solutions differ only with respect to the $\mathbf{A}_i$. Given $\mathbf{A}_i$, b is the same in both cases. With LS, we have

$$\begin{aligned}\text{vec } \hat{\mathbf{B}}_i' &= \mathbf{D}_i\hat{b} = \mathbf{D}_i(\hat{\mathbf{A}}_1' \otimes \cdots \otimes \hat{\mathbf{A}}_n')y = \mathbf{D}_i(\hat{\mathbf{A}}_1' \otimes \cdots \otimes \hat{\mathbf{A}}_n')\mathbf{C}_i'\mathbf{C}_i y \\ &= (\hat{\mathbf{A}}^{i\prime} \otimes \hat{\mathbf{A}}_i')\mathbf{C}_i y = (\hat{\mathbf{A}}^{i\prime} \otimes \hat{\mathbf{A}}_i')\text{vec } \mathbf{Y}_i' = \text{vec } \hat{\mathbf{A}}_i'\mathbf{Y}_i'\hat{\mathbf{A}}^i.\end{aligned} \tag{19}$$

So

$$\hat{\mathbf{B}}_i = \hat{\mathbf{A}}^{i\prime}\mathbf{Y}_i\hat{\mathbf{A}}_i. \tag{20}$$

By the same procedure the canonical solution is

$$\tilde{\mathbf{B}}_i = \tilde{\mathbf{A}}^{i\prime}\mathbf{Y}_i\tilde{\mathbf{A}}_i. \tag{21}$$

As $\mathbf{Y}_i'\hat{\mathbf{A}}^i\hat{\mathbf{A}}^{i\prime}\mathbf{Y}_i\hat{\mathbf{A}}_i = \hat{\mathbf{A}}_i\hat{\mathbf{F}}_i$ (cf. (15)), $\hat{\mathbf{B}}_i = \hat{\mathbf{A}}^{i\prime}\mathbf{Y}_i\hat{\mathbf{A}}_i$ contains the first m_i principal components of $\hat{\mathbf{A}}^{i\prime}\mathbf{Y}_i$, provided that $m_i \leq m^i$. This (necessary) condition follows from the fact that $\hat{\mathbf{A}}^{i\prime}\mathbf{Y}_i\hat{\mathbf{A}}_i$ is an $m^i \times m_i$ matrix. (It is clear that this condition cannot be satisfied when $n = 2$, unless $m_1 = m_2$.)

Let us write the canonical equations as:

$$\mathbf{Y}_i'\mathbf{Y}_i\tilde{\mathbf{A}}_i = \tilde{\mathbf{A}}_i\tilde{\mathbf{F}}_i, \quad \text{where} \quad \tilde{\mathbf{A}}_i'\tilde{\mathbf{A}}_i = \mathbf{I}_{m_i}. \tag{22}$$

Now, $\tilde{\mathbf{A}}_i$ does not contain the eigenvectors of $\mathbf{Y}_i'\tilde{\mathbf{A}}^i\tilde{\mathbf{A}}^{i\prime}\mathbf{Y}_i$, but of $\mathbf{Y}_i'\mathbf{Y}_i$. Hence if for all i except one $\mathbf{A}_i = \mathbf{I}_{\ell_i}$, the two solutions coincide, as $\mathbf{A}^i = \mathbf{I}_{\ell^i}$ in that case. Thus, when considering index i, the canonical solution does not involve the data reduction with respect to any other index.

Still another interpretation of the LS solution is obtained by observing that $\mathbf{A}^i\mathbf{A}^{i\prime}$ is idempotent and hence a projection matrix. It projects onto the space spanned by $\mathbf{A}^i$. So the LS solution amounts to a search for principal components, not of $\mathbf{Y}_i$ itself, but of its projection onto the space spanned by $\mathbf{A}^i$.

An Iteration Method

An obvious iterative procedure to obtain the LS solution is as follows. Take $\tilde{\mathbf{A}}_2, \ldots, \tilde{\mathbf{A}}_n$ as starting values for $\mathbf{A}_2, \ldots, \mathbf{A}_n$. Use these values to form a first estimate of $\mathbf{A}^1$, $\hat{\mathbf{A}}^1_{(1)}$ say. Compute the first estimate of $\mathbf{A}_1$, say $\hat{\mathbf{A}}_{1(1)}$, as the m_1 eigenvectors of $\mathbf{Y}_1'\hat{\mathbf{A}}^1_{(1)}\hat{\mathbf{A}}^{1\prime}_{(1)}\mathbf{Y}_1$ corresponding to the m_1 largest eigenvalues. Using $\hat{\mathbf{A}}_{1(1)}$ and $\tilde{\mathbf{A}}_3, \ldots, \tilde{\mathbf{A}}_n$, we can form $\hat{\mathbf{A}}^2_{(1)}$ and estimate $\hat{\mathbf{A}}_{2(1)}$ in a similar manner. Having computed $\hat{\mathbf{A}}_{1(1)}$ through $\hat{\mathbf{A}}_{n(1)}$, we start with $\mathbf{A}_1$ again and form estimates $\hat{\mathbf{A}}_{1(2)}, \ldots, \hat{\mathbf{A}}_{n(2)}$. The process is continued until convergence.

Convergence of the iterations follows from the following considerations. Q defined in (9) is quadratic and consequently nonnegative. For each i, the solution for $\mathbf{A}_i$ corresponding to (15) minimizes Q, for any value of $\mathbf{A}^i$. So, in the above iterative process, each newly computed $\mathbf{A}_i$ lowers the values of Q or leaves it unaffected. Thus, we obtain a nonincreasing sequence of values of Q which is bounded from below by zero. As a result, the sequence converges.

Some Empirical Results

In order to get some practical experience with the LS approach we applied it to the same data set as was used by Lastovicka (1981). This data set consists of a four-dimensional matrix of order $27 \times 6 \times 5 \times 16$, containing the scores (ranging from 1: "strongly disagree" to 6: "strongly agree") given by $\ell_1 = 27$ individuals to each of $\ell_2 = 6$ TV commercials after each of $\ell_3 = 5$ exposures for each item from a list of $\ell_4 = 16$. This data set was standardized such that the 16×16 matrix $\mathbf{Y}_4'\mathbf{Y}_4$ is a correlation matrix; see Lastovicka (1981) for details. (The factor $1/(27)^{1/2}$ mentioned there on page 51 should read $1/(810)^{1/2}$.)

We first recomputed Lastovicka's results, which are based on $m_1 = 4$, $m_2 = 2$, $m_3 = 3$ and $m_4 = 3$. In general, we were able to reproduce his results, with one notable exception (as was brought to our attention by Jaap Verhees, see also Verhees, 1985, the columns headed "I_{p*}" and "II_{p*}" in his Table 3, containing the varimax rotated eigenvectors of $\mathbf{Y}_3'\mathbf{Y}_3$ corresponding to the largest and second largest eigenvalues, respectively, should be interchanged). As a result, most of the entries in his Table 5, which gives the "core" matrix, a two-dimensional display of $\tilde{\mathbf{b}}$, are incorrect. (But we were also unable to reproduce the entries that are unaffected by the error, i.e., those in the two columns headed "Emotive Response" in his Table 5. The overall magnitudes of the entries in this table are too low to be correct. A rough check is provided by computing the sum of squares of the entries, i.e., $\tilde{\mathbf{b}}'\tilde{\mathbf{b}}$ in our notation. This yields .069 or, cf. (18), $R^2 = .004$, an improbably low value. In fact, $\tilde{\mathbf{b}}'\tilde{\mathbf{b}} = 4.634$ according to our computations.)

We next analyzed the same data by the LS method. We used the same values for the m_i's, and followed the iteration method discussed above. The computer programming is straightforward, with the exception of the procedure to build up the matrices of the type $\mathbf{S}_i$. In order to avoid storage capacity and time limits, one has to break down the construction of $\mathbf{S}_i$ into a set of nested "do loops" in a scalar fashion rather than to construct all intermediate matrices in full.

As expected, convergence towards the optimum was rapid. The iteration was terminated as soon as tr $\mathbf{F}_i$, for some i, showed a relative difference from tr $\mathbf{F}_i$ in the previous round of less than 5×10^{-6}. This occurred with tr $\mathbf{F}_4$ at the end of the tenth round of iterations. (Notice that tr $\mathbf{F}_i$ depends on all $\mathbf{A}_i$. In general it will therefore only remain constant between iterations if all $\mathbf{A}_i$ remain constant.)

The most interesting question is, of course, how the LS method compares to the canonical one with respect to the fit; see Table 1, where we show how $\mathbf{y}'\mathbf{y}$ ($= 16$, as the 16×16 matrix $\mathbf{Y}_4'\mathbf{Y}_4$ is a correlation matrix, and $\mathbf{y}'\mathbf{y} = \mathrm{tr}\ \mathbf{Y}_i'\mathbf{Y}_i$ for all i) is split up in an "unexplained" part ($\hat{\mathbf{e}}'\hat{\mathbf{e}}$) and an "explained" part ($\hat{\mathbf{b}}'\hat{\mathbf{b}}$), see (17). It appears that R^2 goes up from .290 to .323 when replacing the canonical method by ours. The overall level of explanation may not seem to be very high, but one has to keep in mind that the job of reproducing the $27 \times 6 \times 5 \times 16 = 12{,}960$ elements of $\mathbf{y}$ has to be borne by the $4 \times 2 \times 3 \times 3 = 72$ elements of the "core matrix" $\hat{\mathbf{b}}$.

As to the results on the varimax-rotated $\mathbf{A}_i$-matrices, there were some differences between both methods. The results on $\mathbf{A}_2$ (ads) and $\mathbf{A}_3$ (exposures) did not differ very much, but those on $\mathbf{A}_1$ (individuals) and $\mathbf{A}_4$ (items) did. In these two cases, there was a marked difference in that (a) the principal loadings per row were often in different columns (for 21 of the 27 rows of $\mathbf{A}_1$, for 4 of the 16 rows of $\mathbf{A}_4$), and (b) the LS solution

TABLE 1.

LS versus the canonical method: some results

Method	$y'y$	$\hat{e}'\hat{e}$	$\hat{b}'\hat{b}$	R^2
LS	16.000	10.835	5.165	.323
Canonical	16.000	11.366	4.634	.290

shows a tendency to concentrate the principal loadings (i.e., the largest factor scores) in the columns corresponding with the largest eigenvalues. For example, $\hat{\mathbf{A}}_1$ has 11 of the 27 row maxima (absolute values) in the first column and $\tilde{\mathbf{A}}_1$ only 7; for $\hat{\mathbf{A}}_4$ these figures are 9 out of 16 and for $\tilde{\mathbf{A}}_4$ 6 out of 16.

As can be expected from these differences, there were also considerable differences between $\hat{\mathbf{b}}$ and $\tilde{\mathbf{b}}$.

Conclusion

In this paper we presented a least squares approach to multidimensional component analysis. A possibly fruitful byproduct of the paper is the introduction of a notation, based on permutation matrices, that allows for easy handling of the algebra involved, making derivations simple and transparent.

As we have LS as our loss function, our method scores, of course, higher on this criterion than the canonical one. A troublesome aspect of multidimensional component analysis appears to be the apparent sensitivity of at least part of the results to the criterion used. A possible way to study this phenomenon is to consider the individuals in the data set as a random sample from some population and to set up a statistical model to explain the observations. A model which would justify the LS procedure is the following one:

$$\mathbf{y} = \mathbf{A}\mathbf{b} + \mathbf{u}, \tag{23}$$

where $\mathbf{u}$ is an ℓ-vector of i.i.d. errors with mean zero. If we assume the elements of $\mathbf{u}$ to be normally distributed, the LS procedure provides maximum likelihood estimators of $\mathbf{A}$ and $\mathbf{b}$.

Given this model, we can carry out an asymptotic F-test to see whether Lastovicka's results are consistent with the data (e.g., Goldfeld & Quandt, 1972). The only ingredients we need for this test are the R^2-values for the canonical and the LS solutions. We find that F (145, 12743) = 4.28, which is highly significant. Thus, assuming the statistical model, (23), the canonical solution is rejected by the data.

References

Balestra, P. (1976). *La dérivation matricielle* [Matrix derivatives]. Paris: Sirey.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika, 35*, 283–319.

Carroll, J. D., Pruzansky, S., & Kruskal, J. B. (1980). CANDELINC: A general approach to multidimensional analysis of many-way arrays with linear constraints on parameters. *Psychometrika, 45*, 3–23.

Goldfeld, S. M., & Quandt, R. E. (1972). *Nonlinear methods in econometrics*. Amsterdam: North-Holland.

Henderson, H. V., & Searle, S. R. (1981). The Vec-permutation matrix, the Vec-operator and Kronecker products: A review. *Linear and Multilinear Algebra, 9*, 271–288.

Kroonenberg, P. M. (1983). *Three-mode principal component analysis*. Leyden: DSWO-Press.

Kroonenberg, P. M., & de Leeuw, J. (1980), Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika, 45*, 69–97.

Lastovicka, J. L. (1981). The extension of component analysis to four-mode matrices. *Psychometrika, 46*, 47–57.

Magnus, J. R., & Neudecker, H. (1979). The commutation matrix: Some properties and applications. *The Annals of Statistics, 7*, 381–394.

Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika, 31*, 279–311.

Verhees, J. (1985). *Four-mode principal component analysis; evaluation of solutions from two alternative algorithms* (Report). Groningen: Groningen University, Econometrics Institute.

Manuscript received 9/14/82
Final version received 9/26/85