JOURNAL OF CHEMOMETRICS
J. Chemometrics 2000; **14**: 105–122

Towards a standardized notation and terminology in multiway analysis

Henk A. L. Kiers*†

Heymans Institute (PA), University of Groningen, Grote Kruisstraat 2/1, NL-9712 TS Groningen, The Netherlands

SUMMARY

This paper presents a standardized notation and terminology to be used for three- and multiway analyses, especially when these involve (variants of) the CANDECOMP/PARAFAC model and the Tucker model. The notation also deals with basic aspects such as symbols for different kinds of products, and terminology for three- and higher-way data. The choices for terminology and symbols to be used have to some extent been based on earlier (informal) conventions. Simplicity and reduction of the possibility of confusion have also played a role in the choices made. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: multiway data; tensors; three-way methods

1. INTRODUCTION

The first proposals for three- and higher-way generalizations of factor and principal component analysis date back to the 1960s [1,2] and early 1970s [3,4]. The model proposed by Tucker [1,2] generalized the principal component and factor analysis model in that it used one component matrix for all three 'ways' of a three-way data array; these component matrices are related to each other by a so-called core array. The model proposed independently by Carroll and Chang [3] and Harshman [4], called the 'CANDECOMP' and 'PARAFAC' model respectively, also uses component matrices for all three ways, but in their model each component is related to only one component of each of the other two ways. The latter aspect entails that its solution is (under mild conditions) unique. This uniqueness has made the model very popular for purposes of estimation of parameters in situations where two-way data would not allow for unique estimation (e.g. estimating concentrations of chemical analytes in a mixture). In certain situations, however, the model is too restrictive or the uniqueness conditions are not satisfied. In such situations the model proposed by Tucker offers a useful alternative.

The above two models can be considered the fundamental models underlying most currently

* Correspondence to: H. A. L. Kiers, Heymans Institute (PA), University of Groningen, Grote Kruisstraat 2/1, NL-9712 TS Groningen, The Netherlands.
† E-mail: h.a.l.kiers@ppsw.rug.nl

used three-way models and their multiway generalizations [3,5]. Unfortunately, the models, having been developed within psychometrics, have not always been widely known and have been reinvented under different names. Moreover, even when the models are properly attributed, the notation in which they are described differs widely over different papers and disciplines, and so does the terminology employed (even as far as the names of the models are concerned). Some informal attempts have been made towards a standard notation and terminology, but this has not been successful yet. The present paper is an attempt to formally standardize notation and terminology of the most important aspects in multiway analysis. The proposal takes into account any informal conventions the author is aware of; in cases of several competing choices encountered in the literature, mnemonic simplicity and conceptual clarity are the main criteria for adopting a standard.

2. DATA

2.1. General

In statistics, data are generally denoted by the symbol X. In the case of two-way data (i.e. a data matrix) the bold-face version $\mathbf{X}$ is used. To distinguish the more or less standard two-way data sets from the less common three- and higher-way data sets, the latter are indicated by an underlined bold-face $\underline{\mathbf{X}}$. When, in addition to $\underline{\mathbf{X}}$, there are other data sets, it is preferred to use subscripts or to denote them by other letters at the far end of the alphabet ($\underline{\mathbf{V}}$, $\underline{\mathbf{W}}$, $\underline{\mathbf{Y}}$, $\underline{\mathbf{Z}}$). For example, in the case of regression involving three-way data, the three-way array pertaining to the predictor variables is denoted by $\underline{\mathbf{X}}$ and that pertaining to the criterion variables by $\underline{\mathbf{Y}}$.

Data to be analysed are usually real numbers, but in some instances they are complex. In the present paper, only real-valued data are considered. Mostly, the notation and properties mentioned in the present paper readily generalize to complex-valued data, but one should be aware of exceptions (e.g. the complex conjugate differs from the simple transpose).

2.2. Indices

The elements of a three-way array $\underline{\mathbf{X}}$ are denoted by x_{ijk}, where the indices are taken to run from 1 to their capital version: $i = 1,\ldots,I$, $j = 1,\ldots,J$, $k = 1,\ldots,K$. For higher-way arrays, additional subscripts are used, chosen as the next letters in the alphabet (e.g. an element of a six-way array is denoted as x_{ijklmn}). In some higher-way cases one may thus run out of symbols (considering that we need other running indices as well). In that case one can resort to the more complex use of indices i_1, i_2, i_3, i_4, etc. running from 1 to I_1, I_2, I_3, I_4, etc.

2.3. Tensors

An N-way data array is sometimes generically called a *tensor* (see e.g. Reference [6], p. 10). Thus the notion 'tensor' captures arrays of different sizes: a vector of order I is a tensor in $\mathbb{R}^{\mathrm{I}}$, an $I \times J$ matrix is a tensor in $\mathbb{R}^{I \times J}$, an $I \times J \times K$ three-way array is a tensor in $\mathbb{R}^{I \times J \times K}$, etc. In general, an $I_1 \times I_2 \times \ldots \times I_N$ N-way array is a tensor in $\mathbb{R}^{I_1 \times I_2 \times \ldots \times IN}$.

2.4. Modes of an array

In Figure 1 a three-way array is depicted. The entities along the vertical axis are indicated by the first index (i), those along the horizontal axis by the second index (j) and those along the depth axis by the third index (k). Note that the use of the first and second indices is the same as what is common for matrices. The three sets of entities define the three 'ways' or three

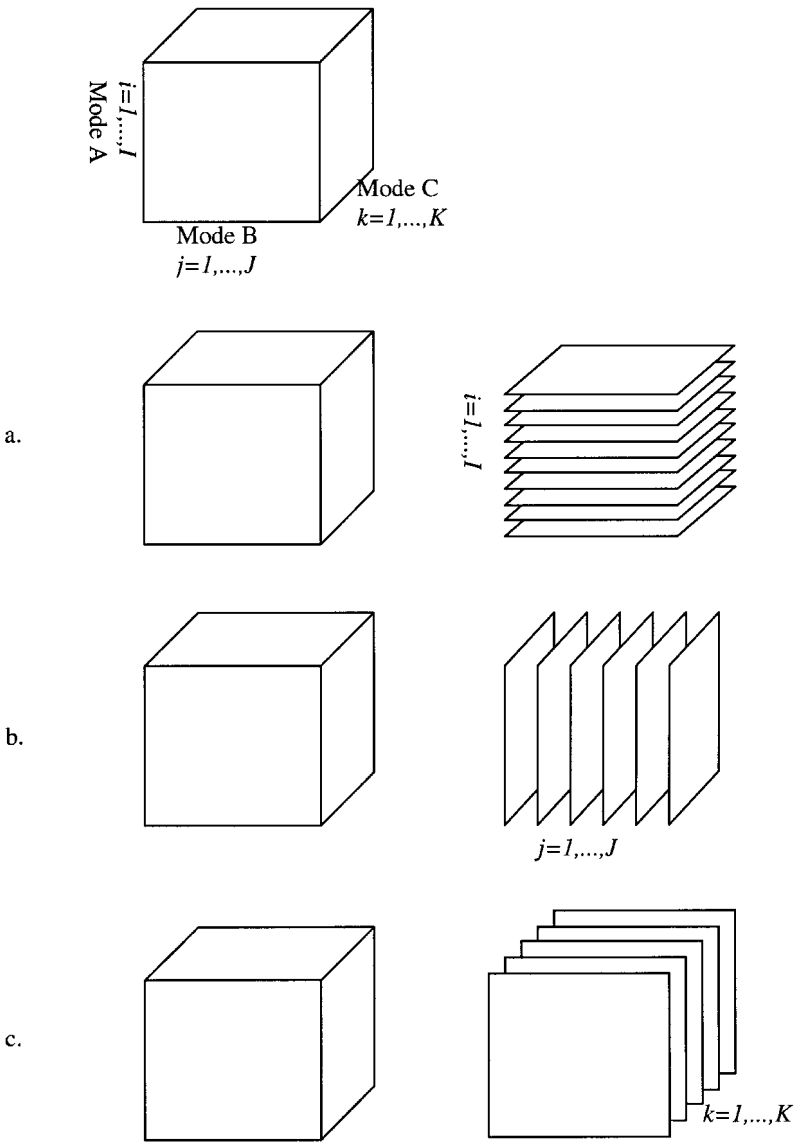

Figure 1. Three-way array, cut into (a) horizontal, (b) lateral and (c) frontal slices.

'dimensions' of the three-way array. Because the term 'way' is not very specific and the term 'dimension' may be confused with the concept of dimensions of a factor space in factor analysis, here, following Tucker [7], we use the term 'mode' to refer to a set of entities. In the case of three-way arrays the first mode is denoted as *mode A*, the second mode as *mode B* and the third mode as *mode C*. This terminology can be extended to higher-way data (leading to *modes D, E,* etc.), but often for higher-way data it is more convenient to denote the modes by numbers, hence as *mode 1, mode 2,…, mode N*.

2.5. One-mode, two-mode and three-mode three-way data

Carroll and Arabie [8] distinguished three-way data into three types as follows. If the three modes pertain to three different sets of entities, then the data are denoted as *three-mode three-way* data or simply *three-way* data. If two modes pertain to the same set of entities (e.g. when dealing with correlations between variables or (a)symmetric proximities between objects) and hence only two *different* modes are involved, this can be indicated by specifying such three-way data as *two-mode* three-way data. Finally, if all three modes pertain to the same set of entities (e.g. in the case of transitions between a number of states over three consecutive time points), the data can be denoted as *one-mode* three-way data. This terminology can easily be extended to *N*-way data by simply counting the number of different modes, but is not always fruitful then, because it does not distinguish, for instance, between four-way arrays with two modes that both occur twice and four-way arrays with one mode that occurs three times and one mode that occurs only once. There is one exception in which there is no doubt about the meaning: *one-mode N*-way data pertain to data in which all modes pertain to the same entities. Incidentally, it should be noted that even in cases where some of the modes are equal, we will still refer to them by different names (*mode A, mode B, mode C*, etc. or *mode 1, mode 2,…, mode N*), as described in Section 2·4.

2.6. Matrices, vectors and other subarrays of three- and higher-way arrays

A three-way array is frequently considered in terms of a set of matrices. These matrices form the *horizontal, lateral* and *frontal* slices of the three-way array, as visualized in Figures 1a–1c. Specifically, the I horizontal slices pertain to the entities $i = 1,\ldots,I$ of mode A, the J lateral slices pertain to the entities $j = 1,\ldots,J$ of mode B and the K frontal slices pertain to the entities $k = 1,\ldots,K$ of mode C. For higher-way arrays, submatrices can be defined as well, but it no longer makes sense to give them intuitive names as is done in the three-way case.

Sometimes it is useful to consider a three-way array as a set of vectors. Then three sets of such vectors (denoted as *fibers*) can be distinguished, namely vertical fibers, horizontal fibers and 'depth' fibers, which run over the mode A entities, mode B entities and mode C entities respectively, as visualized in Figures 2a–2c. To compare with matrices: horizontal fibers are rows and vertical fibers are columns. In general, these vectors are called 'mode n vectors' (compare Reference [6]), where n denotes the mode over which the vectors run. They are associated with particular vector spaces, to be called the 'mode n spaces'. Thus for a three-way array the mode A space is the (sub)space spanned by *all* vertical fibers, the mode B space is the (sub)space spanned by *all* horizontal fibers and the mode C space is the (sub)space spanned by *all* depth fibers. These spaces are subspaces of $\mathbb{R}^I$, $\mathbb{R}^J$ and $\mathbb{R}^K$ respectively.

To give a numerical example, consider the $4 \times 3 \times 2$ array with frontal planes

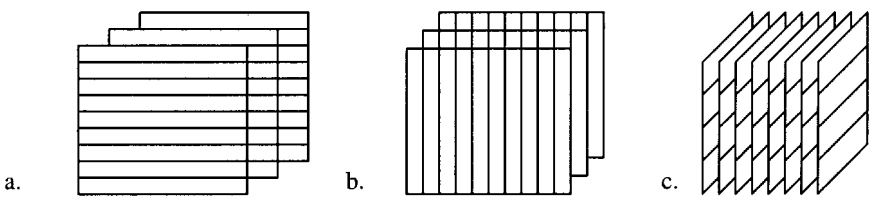

Figure 2. Three-way array, cut into (a) horizontal, (b) vertical and (c) depth fibers.

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The vectors

$$\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

are two vertical fibers (or mode A vectors) of the array, the vectors (1 2 0), (1 0 −1) and (1 2 0) are examples of three horizontal fibers (or mode B vectors) and the vectors (1 0), (2 1) and (0 1) are examples of depth fibers (or mode C vectors). The mode A space is the space spanned by the vectors

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

which can readily be seen to be the whole $\mathbb{R}^4$. The mode B and mode C spaces are defined analogously.

In general, for *N*-way arrays we can define subarrays of any size desired. For instance, one may define the three-way arrays that are a subset of a four-way array. Such arrays can hardly be given intuitive names as we did for matrices and vectors embedded within three-way arrays, but they can be specified using a notation system similar to that used in MATLAB, where we use subscript colons for the modes that remain intact and ordinary indices to indicate the entities with which the subarray is associated. For example, $\mathbf{X}_{i::}$, $\mathbf{X}_{:j:}$ and $\mathbf{X}_{::l}$ then denote the horizontal, lateral and frontal planes of a three-way array respectively, and examples of submatrices of a four-way array are $\mathbf{X}_{i::l}$ and $\mathbf{X}_{:jk:}$. Fibers (or mode *n* vectors) can be denoted by small bold letters using the same indexing system. For example, $\mathbf{x}_{:jk}$ denotes a mode A vector (vertical fiber) in a three-way array and $\mathbf{x}_{ij:l}$ denotes a mode C vector in a four-way array. Furthermore, three- or higher-way subarrays then are denoted by underlined bold capitals, again with the same indexing system. For example $\underline{\mathbf{X}}_{::kl:}$ denotes the three-way subarray of the five-way array $\underline{\mathbf{X}}$ associated with the *k*th entity of mode 3 and the *l*th entity of mode 4. Thus any subarray of an *N*-way array can be identified unambiguously. In practice, the colons will often be omitted, as in many cases there will be no reason for confusion. A common situation is that where a three-way array is only subdivided into its frontal planes. These are then simply denoted as $\mathbf{X}_1,\ldots,\mathbf{X}_K$.

2.7. Matricization: transforming a three-way or N-way array into a matrix

It is sometimes fruitful to collect all *mode n* vectors in a single matrix. The supermatrix with all vertical fibers of a three-way array collected next to each other in an $I \times JK$ matrix, with mode B entities ($j = 1,\ldots,J$) nested within mode C entities ($k = 1,\ldots,K$), is denoted as $\mathbf{X}_a$. This matrix simply contains all the frontal slices of the array next to each other (see Figure 3a). The

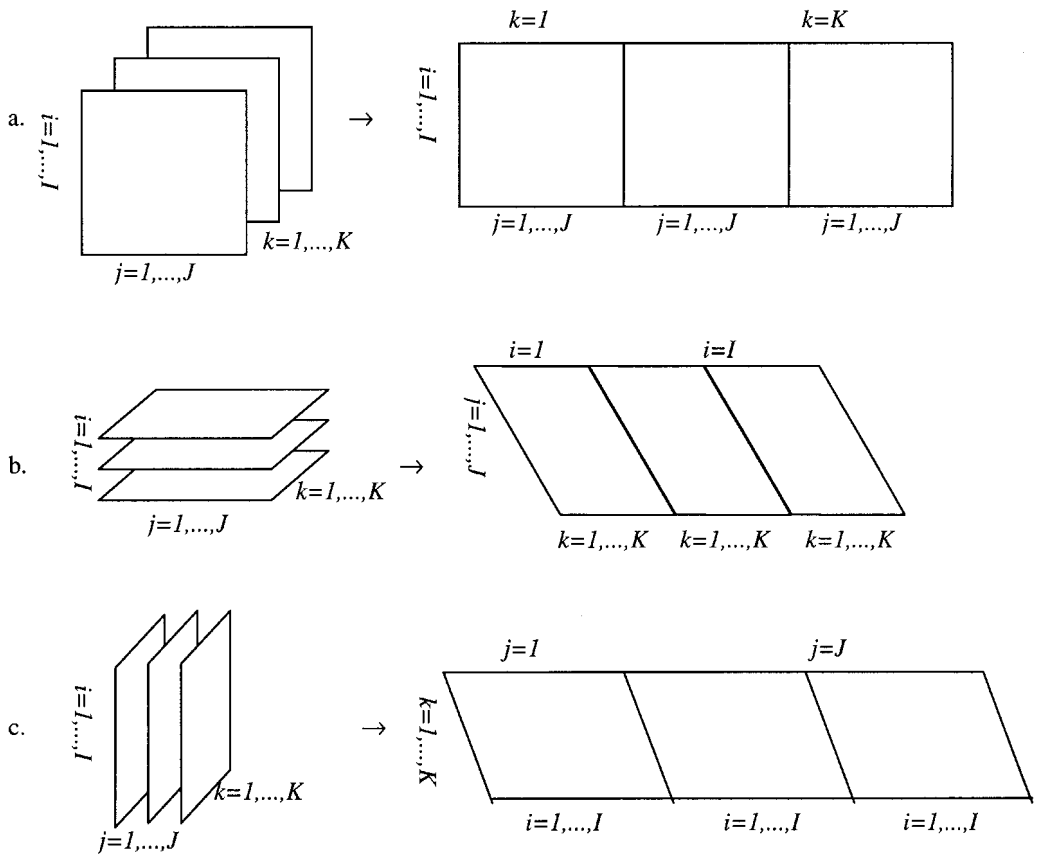

Figure 3. Matricizing a three-way array: (a) mode A matricization; (b) mode B matricization; (c) mode C matricization.

process of rearranging the elements of $\underline{\mathbf{X}}$ into $\mathbf{X}_a$ is often called 'unfolding' in chemometrics, but this term is confusing, because in psychometrics unfolding [9] is a particular technique for multidimensional scaling of data with distances between two sets of entities. Here this process is denoted as 'matricizing' a three-way array into a matrix (in analogy to the more common term 'vectorizing' for rearranging a matrix into a vector), and the reverse process is then called 'reshaping' a matrix into a three-way array. Other matricizations are those that form the supermatrices $\mathbf{X}_b$ (of order $J \times KI$, with mode C entities nested within mode A entities) and $\mathbf{X}_c$ (of order $K \times IJ$, with mode A entities nested within mode B entities) (see Figure 3b and 3c). Other nestings are possible, but without further specification, matricization pertains to one of the above procedures.

The above matricizations are related to each other by a simple cyclic permutation of the modes. Analogous to $\mathbf{X}_a$ containing frontal planes of $\underline{\mathbf{X}}$ next to each other, $\mathbf{X}_b$ contains frontal planes of the three-way array that is obtained upon once cyclically permuting the modes of $\underline{\mathbf{X}}$ (i.e. the second index becomes the first, the third index becomes the second and the first index becomes the third), and $\mathbf{X}_c$ contains frontal planes of the three-way array that is obtained upon twice permuting the modes of $\underline{\mathbf{X}}$. For the numerical example in Section 2·6 we have that

$$\mathbf{X}_a = \begin{pmatrix} 1 & 2 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 1 & 2 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{X}_b = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 & -1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{X}_c = \begin{pmatrix} 1 & 0 & 1 & -1 & 2 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Matricization is not only useful for three-way arrays. In fact, *N*-way data usually are to be read from file in a two-way structure. An *N*-way array is 'matricized' in essentially the same way as a three-way array. For instance, $\mathbf{X}_a$ (or $\mathbf{X}_1$ if it is preferred to denote modes by numbers rather than letters, as is the case when there are many modes) contains all vertical fibers collected next to each other in an $I \times JKLM\ldots$ matrix, with in the columns the mode B entities nested within mode C entities, mode C entities nested within mode D entities, etc. Likewise, $\mathbf{X}_b$ (or $\mathbf{X}_2$) is the $J \times KLM\ldots I$ matrix, with in the columns the mode C entities nested within mode D entities, mode D entities nested within mode E entities, etc., with mode A entities as the 'outermost' entities. Thus, again, $\mathbf{X}_b$ is obtained from the once cyclically permuted array $\underline{\mathbf{X}}$ in the same way as $\mathbf{X}_a$ is obtained from $\underline{\mathbf{X}}$ itself. This permutational equivalence makes programming with *N*-way arrays relatively straightforward, even when procedures are to employ matricized versions. If the matricized versions of arrays are to be given names in addition to symbols, we call $\mathbf{X}_a$ the mode A matricized version of $\underline{\mathbf{X}}$, $\mathbf{X}_b$ the mode B matricized version of $\underline{\mathbf{X}}$, etc., and $\mathbf{X}_a$ the mode *n* matricized version of $\underline{\mathbf{X}}$.

Obviously, other possibilities exist for transforming a tensor into a matrix. The above procedures should be seen as the standard and preferable ones, and the use of others should always be explicitly described. For example, the procedure of writing an $I \times J \times K \times L$ arrays as an $IJ \times KL$ matrix can be denoted as 'matricization by combining modes 1 and 2 (1 nested within 2) and modes 3 and 4 (3 nested within 4)'.

2.8. Vectorization: transforming a three-way or N-way array into a matrix

Sometimes it is useful to represent all the elements of a three- or *N*-way array as a vector. This can be done by 'vectorizing' the array. For matrices, 'vectorization' is defined as putting the successive columns of the matrix below each other in a single vector. The vectorization of the matrix $\mathbf{U}$ into a vector $\mathbf{u}$ is denoted as $\mathbf{u} = \mathrm{Vec}(\mathbf{U})$. To vectorize a three-way or higher-way array, we simply vectorize the mode A matricized version of it and obtain $\mathbf{x} = \mathrm{Vec}(\mathbf{X}_a)$ as the vectorized version of an array of arbitrary order.

2.9. Mode n rank and tensor rank

To define the 'rank' of a three- or *N*-way array is more complex than for a matrix. In fact, two types of rank have been defined. The simplest is the mode *n* rank [6], which is defined as the rank of the mode *n* space (see Section 2·6). For instance, the mode A rank (or mode 1 rank) of an $I \times J \times K \times L$ array $\underline{\mathbf{X}}$ is the rank of the space spanned by the *JKL* 'mode A vectors' $\mathbf{x}_{\cdot jkl}$. Clearly, the mode A rank is the rank of the matrix $\mathbf{X}_a$. The mode B rank (or mode 2 rank) of

this array is the rank of the space spanned by the *IKL* 'mode B vectors' $\mathbf{x}_{i:kl}$, hence it is the rank of $\mathbf{X}_b$. The mode C and mode D ranks of $\underline{\mathbf{X}}$ are defined analogously. For matrices the mode *n* rank can be viewed as the column rank (mode A rank) or row rank (mode B rank) of a matrix. Since the row and column ranks of matrices are known to be equal, for two-way arrays the mode A and mode B ranks are equal. For three- and higher-way arrays such an equality no longer holds.

The usual definition of (*tensor*) *rank* [10] uses a decomposition of a tensor as a sum of 'rank-1 tensors'. A rank-1 tensor is a tensor for which the elements can be written as $x_{ijkl\ldots} = a_i b_j c_k d_l \ldots$. The *rank* of a tensor then is the smallest number of rank-1 tensors sufficient to fully decompose the tensor additively. In the case of matrices this rank is equal to the row and column ranks, but for three- and higher-way arrays no such equivalence exists and the rank can actually be considerably higher than (and is never less than) any of the mode *n* ranks.

2.10. Special tensors

A tensor which has all elements zero is called the 'zero tensor' (or 'zero array'). For (*hyper*)*cubic* tensors, i.e. tensors with all modes of equal size, we have some special terminology. A tensor which has all elements zero except those for which all indices are the same (which are called the 'superdiagonal' elements) is called a 'superdiagonal' tensor. In the case where these superdiagonal elements all equal one, it is called the 'unit superdiagonal' tensor, denoted by the symbol $\underline{\mathbf{I}}$ (chosen because, as far as its elements are concerned, it resembles the identity matrix). Note that the unit superdiagonal tensor should not to be denoted as 'identity', because it does not perform a role similar to that of the identity element (1) or matrix (**I**) in ordinary (matrix) algebra (see Reference [6], p. 26).

3. SOME SPECIAL SYMBOLS AND PROPERTIES

In multiway analysis, certain specific matrix products and operators are often used. They will be summarized in this section. Here a_{ij} or $a_{i,j}$ is used to denote element (i,j) of a matrix **A**, and $\mathbf{a}_l$ denotes the lth column of **A**. Furthermore, a 'prime' denotes transposition of a matrix or vector.

The *Kronecker product* is denoted by the symbol $\otimes$ and is defined according to $(\mathbf{U}\otimes\mathbf{V})_{ik,jl} = u_{ij}v_{kl}$. Thus we have

$$\mathbf{U}\otimes\mathbf{V} = \begin{pmatrix} u_{11}\mathbf{V} & \ldots & u_{1J}\mathbf{V} \\ \ldots & \ldots & \ldots \\ u_{I1}\mathbf{V} & \ldots & u_{IJ}\mathbf{V} \end{pmatrix}$$

The *columnwise Kronecker product* (also denoted as the Khatri–Rao product; see Reference [11], p. 13), denoted by the symbol $\odot$, can be computed between matrices of the same column order and is defined according to $(\mathbf{U}\odot\mathbf{V})_{ik,l} = u_{il}v_{kl}$. Hence, if **U** and **V** both have L columns, we have $\mathbf{U}\odot\mathbf{V} = (\mathbf{u}_1\otimes\mathbf{v}_1|\ldots|\mathbf{u}_L\otimes\mathbf{v}_L)$.

The *elementwise* or *Hadamard product*, denoted by the symbol *, can be computed between matrices of the same order only and is defined according to $(\mathbf{U}*\mathbf{V})_{ij} = u_{ij}v_{ij}$. In multiway analysis the latter product is typically encountered upon matrix multiplication of two columnwise Kronecker products, because $(\mathbf{U}\odot\mathbf{V})'(\mathbf{U}\odot\mathbf{V}) = (\mathbf{U}'\mathbf{U})*(\mathbf{V}'\mathbf{V})$. This property can be derived as follows:

$$(\mathbf{U}\odot\mathbf{V})'(\mathbf{U}\odot\mathbf{V}) = \begin{pmatrix} \mathbf{u}_1'\otimes\mathbf{v}_1' \\ \dots \\ \mathbf{u}_L'\otimes\mathbf{v}_L' \end{pmatrix}(\mathbf{u}_1\otimes\mathbf{v}_1|\dots|\mathbf{u}_L\otimes\mathbf{v}_L)$$

$$= \begin{pmatrix} \mathbf{u}_1'\mathbf{u}_1\otimes\mathbf{v}_1'\mathbf{v}_1 & \dots & \mathbf{u}_1'\mathbf{u}_L\otimes\mathbf{v}_1'\mathbf{v}_L \\ \dots & \dots & \dots \\ \mathbf{u}_L'\mathbf{u}_1\otimes\mathbf{v}_L'\mathbf{v}_1 & \dots & \mathbf{u}_L'\mathbf{u}_L\otimes\mathbf{v}_L'\mathbf{v}_L \end{pmatrix} = (\mathbf{U}'\mathbf{U})*(\mathbf{V}'\mathbf{V})$$

The *mode n multiplication*, denoted by the symbol $\times_n$ (see Reference [6], p. 15; see also Reference [12] for a similar definition), defines the multiplication of an array by a matrix along mode n of the array. Specifically, multiplication of the array $\underline{\mathbf{G}}$ by the matrix $\mathbf{S}_n$ (with an arbitrary number of rows and I_n columns) along mode n is written as $\underline{\tilde{\mathbf{G}}} = \underline{\mathbf{G}} \times_n \mathrm{S}_n$ and the outcome is equivalent to the matrix multiplication $\tilde{\mathbf{G}}_n = \mathbf{S}_n\mathbf{G}_n$ involving the mode n matricized versions of $\underline{\mathbf{G}}$ and $\underline{\tilde{\mathbf{G}}}$, $\mathbf{G}_n$ and $\tilde{\mathbf{G}}_n$ (of order $I_n \times I_1I_2\ldots I_N/I_n$) respectively. For example, let $\underline{\mathbf{X}}$ be the $2\times2\times2$ array with frontal planes

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}$$

and let

$$\mathbf{S} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{U} = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}$$

Then $\underline{\mathbf{X}} \times_1 \mathbf{S}$ can be obtained in mode A matricized version as

$$\mathbf{S}\mathbf{X}_{\mathrm{a}} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 1 & -2 & -1 \end{pmatrix}$$

and hence has as frontal planes

$$\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix}$$

To find $\underline{\mathbf{X}} \times_3 \mathbf{U}$, we compute its mode C matricized version as

$$\mathbf{U}\mathbf{X}_{\mathrm{c}} = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 2 \\ 2 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 4 & 0 & 1 & 4 \end{pmatrix}$$

and from this recover its frontal planes as

$$\begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$$

An overview of properties of the above special products can be found, for instance, in Reference [13]. There (p. 263) the following not very well-known but important property relating vectorization of a product of matrices to the Kronecker product is given:

$$\mathrm{Vec}(\mathbf{UVW}) = (\mathbf{W}'\otimes\mathbf{U})\mathrm{Vec}(\mathbf{V}) \tag{1}$$

4. FUNDAMENTAL THREE-WAY MODELS

4.1. Names: Tucker models and CANDECOMP/PARAFAC

As already mentioned in Section 1, the fundamental models used in three-way analysis are the model proposed by Tucker [1] and that proposed by Carroll and Chang [3] and Harshman [4]. The former was termed the 'three-mode factor analysis' model by Tucker. Kroonenberg and De Leeuw [14], who proposed an algorithm for least squares fitting of this model, called their procedure 'three-mode principal component analysis' (to distinguish it from the more narrowly defined term factor analysis which usually is fitted by means of fitting covariances, using assumptions of uncorrelated unique factors). They also distinguished two special cases: the case where one mode is represented by as many components as there are entities (no reduction in this mode), and the case where no reduction takes place in two modes. Consequently, they denoted the original model as the *Tucker3 model* (reduction in all three modes), the model where only two modes were reduced was denoted as the *Tucker2 model*, and that where only one mode is reduced is called the *Tucker1 model*. This distinction between models is a useful one and is often made, albeit not necessarily by these names. Here it is proposed to generally follow this terminology for denoting the models, but we make two modifications. Firstly, in cases where confusion may arise, we add the term 'three-way' to indicate that these are three-way models, and thus we can analogously specify four- and higher-way models where not all modes are reduced; for instance, the 'four-way Tucker2 model' indicates a four-way model in which only two of the modes are reduced. Secondly, in the case where all modes are reduced, which we see as the 'default', we may simply drop the number indicating how many modes are reduced; thus the three-way Tucker3 model can be denoted as the *three-way Tucker model*, and more generally, the *N-way Tucker model* indicates the Tucker model for *N*-way data in which all modes are reduced. The methods fitting these models in the least squares sense are then called *N-way Tucker analysis* if a model with all modes reduced is considered, or *N-way Tucker1 analysis, N-way Tucker2 analysis, N-way Tucker3 analysis*, etc. if only some modes are reduced. In fact, *N-way Tucker1 analysis* comes down to a principal component analysis (PCA) of all fibers pertaining to one mode, and hence of a matricized version of the data array. For three-way data this method has also been denoted as *PCASUP* [15], and in chemometrics it is sometimes denoted as 'unfold PCA' or 'multiway PCA', but, as mentioned in Section 2·7, the term 'unfolding' is confusing and the term 'multiway PCA' is easily confused with other multiway generalizations of PCA.

The model proposed by Carroll and Chang [3] and Harshman [4] received entirely different names by its two proposers, referring to different features of the model. To give credit to both proposers and to both types of features, the model is referred to as the CANDECOMP/PARAFAC model or, abbreviated, CP model. The order of the constituent names in the name of the model has been chosen alphabetically and leads to the least confusing abbreviation (PC model resembles PC (personal computer) or PCA). The CP model can be written mathematically as a constrained variant of the three-way Tucker model. This property is often used, and therefore terminology and notation in the two models should be well adjusted.

4.2. Standard notation for models

The most general model is the three-way Tucker model. This is given by

$$x_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} a_{ip} b_{jq} c_{kr} g_{pqr} + e_{ijk} \tag{2}$$

In (2), a_{ip}, b_{jq} and c_{kr} denote elements of the component matrices **A** (for the mode A), **B** (for the mode B) and **C** (for the mode C) of orders $I \times P$, $J \times Q$ and $K \times R$ respectively. Thus for the mode **A** components the running index is p, which runs from 1 to P, for the mode **B** components the running index is q, which runs from 1 to Q, and for the mode **C** components the running index is r, which runs from 1 to R. Furthermore, g_{pqr} denotes the element (p,q,r) of the $P \times Q \times R$ core array $\underline{\mathbf{G}}$. Finally, e_{ijk} denotes the error term for element x_{ijk} and is an element of the $I \times J \times K$ array $\underline{\mathbf{E}}$. Up to six-way generalizations can be defined analogously using component matrices **A, B, C, D, E** and **F** and component indices p, q, r, s, t and u. For clarity, in cases where **E** denotes the E-mode component matrix, the error array may be denoted by a different symbol, even though no confusion is likely to arise, because the error array and the component matrix play entirely different roles and do not emerge in the same term. From seven-way on, however, we would need a component matrix **G**, which could become confusing with the core $\underline{\mathbf{G}}$. Therefore for N-way arrays with $N > 6$ (or, if preferred, even for $N > 3$) we use subscripts to distinguish the component matrices for the different modes: $\mathbf{A}_1$, $\mathbf{A}_2$,…, $\mathbf{A}_N$. The elements of $\mathbf{A}_n$ are then indexed by i_n and p_n and hence given by $a_{i_n p_n,n}$, where the matrix subscript is separated from the others by a comma. Note that in this general form the final subscript n is unnecessary, but in specific cases we need all three subscripts to identify a particular element of a particular component matrix (e.g. $a_{34,2}$ denotes element (3,4) of component matrix $\mathbf{A}_2$ and clearly a_{34} would be incomplete). With this general formulation the *N-way Tucker model* [5] can be written as

$$x_{i_1 i_2 \ldots i_N} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \cdots \sum_{p_N=1}^{P_N} \left(\prod_{n=1}^{N} a_{i_n p_n,n} \right) g_{p_1 \ldots p_N} + e_{i_1 i_2 \ldots i_N} \tag{3}$$

The main advantage of (3) is that it is fully general. It is, however, rather hard to read, hence this formulation should only be used if the specific notation, preferably using different letters for different modes, is unfeasible.

The (three-way) CP model is given by

$$x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} + e_{ijk} \tag{4}$$

Here a_{ir}, b_{jr} and c_{kr} again are elements of the component matrices **A** (for the mode A), **B** (for the mode B) and **C** (for the mode C), now of orders $I \times R$, $J \times R$ and $K \times R$ respectively. Thus all component matrices have the same number of columns (R). Analogously, the *N-way CP model* is given by

$$x_{i_1 i_2 \ldots i_N} = \sum_{r=1}^{R} \left(\prod_{n=1}^{N} a_{i_n r,n} \right) + e_{i_1 i_2 \ldots i_N} \tag{5}$$

For completeness we also give the three-way Tucker2 and Tucker1 models

$$x_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} a_{ip} b_{jq} g_{pqk} + e_{ijk} \tag{6a}$$

$$x_{ijk} = \sum_{p=1}^{P} a_{ip} g_{pjk} + e_{ijk} \tag{6b}$$

which are special cases of the three-way Tucker model with $\mathbf{C} = \mathbf{I}_K$ and $\underline{\mathbf{G}}$ of order $P \times Q \times K$ (Tucker2) and $\mathbf{B} = \mathbf{I}_J$, $\mathbf{C} = \mathbf{I}_K$ and $\underline{\mathbf{G}}$ of order $P \times J \times K$ (Tucker1). All models are usually fitted to a data set by minimizing the sum of squared error terms, $\sum_{ijk} {e_{ijk}}^2$, over $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\underline{\mathbf{G}}$.

4.3. Matrix formulations of models

Matrix formulations of the Tucker and CP models are often given. For the three-way Tucker model this formulation is

$$\mathbf{X}_a = \mathbf{A}\mathbf{G}_a(\mathbf{C}' \otimes \mathbf{B}') + \mathbf{E}_a \tag{7}$$

From (4) and (2) it is clear that the CP model is the constrained version of the three-way Tucker model with all elements in the core equal to zero, except the elements g_{111}, g_{222},…, g_{RRR} which are equal to one [3]. Thus, using the $R \times R \times R$ three-way unit superdiagonal array $\underline{\mathbf{I}}$ (see Section 2·10), which has unit elements in the positions (r,r,r), $r = 1,\ldots,R$, and zeros elsewhere, we can view the CP model as the special case of the three-way Tucker model with the core equal to the unit superdiagonal array. Hence the CP model is

$$\mathbf{X}_a = \mathbf{A}\mathbf{I}_a(\mathbf{C}' \otimes \mathbf{B}') + \mathbf{E}_a \tag{8}$$

An alternative notation can be based on the columnwise Kronecker product (see e.g. Reference [16], p. 25). Then we obtain

$$\mathbf{X}_a = \mathbf{A}(\mathbf{C} \odot \mathbf{B})' + \mathbf{E}_a \tag{9}$$

which expresses the structural part of the model only in terms of its three parameter matrices. Note that whereas $(\mathbf{C} \otimes \mathbf{B})' = (\mathbf{C}' \otimes \mathbf{B}')$, we do *not* have that $(\mathbf{C} \odot \mathbf{B})'$ equals $(\mathbf{C}' \odot \mathbf{B}')$.

For higher-way generalizations of the three-way Tucker model the matrix formulations are obtained straightforwardly. For instance, the four-way Tucker model given by

$$x_{ijkl} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} \sum_{s=1}^{S} a_{ip} b_{jq} c_{kr} d_{ls} g_{pqrs} + e_{ijkl} \tag{10}$$

can be written in matrix representation as

$$\mathbf{X}_a = \mathbf{A}\mathbf{G}_a(\mathbf{D}' \otimes \mathbf{C}' \otimes \mathbf{B}') + \mathbf{E}_a \tag{11}$$

More generally, the matrix formulation of the N-way Tucker model (3) is given by

$$\mathbf{X}_1 = \mathbf{A}_1\mathbf{G}_1(\mathbf{A}_N' \otimes \ldots \otimes \mathbf{A}_3' \otimes \mathbf{A}_2') + \mathbf{E}_1 \tag{12}$$

where $\mathbf{X}_1$ and $\mathbf{G}_1$ denote mode 1 matricized versions, hence with the row order equal to the order of mode 1. The *N*-way CP model (5) is similarly defined as

$$\mathbf{X}_1 = \mathbf{A}_1(\mathbf{A}_N \odot \ldots \odot \mathbf{A}_3 \odot \mathbf{A}_2)' + \mathbf{E}_1 \tag{13}$$

Models for cyclically permuted versions of the array are easily obtained by cyclic permutation of the letters or numbers that indicate the modes; for the three-way Tucker model we thus have

$$\mathbf{X}_b = \mathbf{B}\mathbf{G}_b(\mathbf{A}' \otimes \mathbf{C}') + \mathbf{E}_b \tag{14}$$

$$\mathbf{X}_c = \mathbf{C}\mathbf{G}_c(\mathbf{B}' \otimes \mathbf{A}') + \mathbf{E}_c \tag{15}$$

and it becomes clear that the component matrices play fully symmetric roles.

4.4. Vector formulations of models

By applying (1) (see Section 3) to (7), we can write the vectorized version of the three-way Tucker model as

$$\mathbf{x} = (\mathbf{C} \otimes \mathbf{B} \otimes \mathbf{A})\mathbf{g} + \mathbf{e} \tag{16}$$

where **x, g** and **e** denote the vectorized versions of $\underline{\mathbf{X}}$, $\underline{\mathbf{G}}$ and $\underline{\mathbf{E}}$ respectively. Expression (16) better displays the symmetry of the three-way Tucker model than does (7), and also gives insight into the role of the elements of the core (as regression weights for the columns of $\mathbf{C}\otimes\mathbf{B}\otimes\mathbf{A}$). Moreover, the *N*-way version of (16) is obtained by straightforward generalization as

$$\mathbf{x} = (\mathbf{A}_N \otimes \mathbf{A}_{N-1} \otimes \ldots \otimes \mathbf{A}_2 \otimes \mathbf{A}_1)\mathbf{g} + \mathbf{e} \tag{17}$$

Finally, the three- and *N*-way CP models are now given by

$$\mathbf{x} = (\mathbf{C} \odot \mathbf{B} \odot \mathbf{A})\mathbf{1}_R + \mathbf{e} \tag{18}$$

and

$$\mathbf{x} = (\mathbf{A}_N \odot \mathbf{A}_{N-1} \odot \ldots \odot \mathbf{A}_2 \odot \mathbf{A}_1)1_R + \mathbf{e} \tag{19}$$

respectively, where $\mathbf{1}_R$ denotes a vector of order *R* with unit elements only.

5. PRE- AND POSTPROCESSING

5.1. Centering and scaling 'within ' and 'across '

Before actually carrying out a multiway analysis, it is often useful to preprocess the data, just as in two-way analysis. In two-way analysis, data are often centered and/or normalized (the combination being called 'standardized' or 'autoscaled') across the rows to eliminate unwanted differences in level and scale. For three- and higher-way data one may similarly wish to eliminate such unwanted differences, but it is no longer obvious how each of the modes should be dealt with. Therefore it is most important to carefully explain how the data are centered and/or normalized prior to analysis. In their detailed account of preprocessing

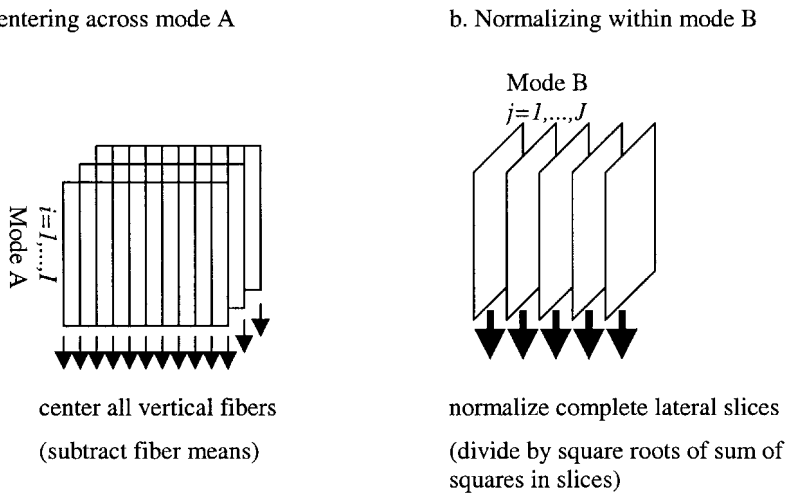

Figure 4. Visualization of 'centering across' and 'normalizing within'.

three-way data, Harshman and Lundy [17] and Ten Berge [18] used the following terminology.

The type of centering is indicated by specifying *across* which modes the data are centered. For instance, 'centering across (mode) A' is carried out by first averaging the data (only) over the entities of mode A and then subtracting each thus obtained average from all data elements that partake in it (see Figure 4a). Analogously, 'centering across the combination of modes A and B' is carried out by first averaging the data over the entities of modes A and B and then subtracting each such average from the data that partake in it. In formulae for three-way data, 'centering across A' leads to computing centered data as

$$\widetilde{x}_{ijk} = x_{ijk} - \overline{x}_{.jk} \tag{20}$$

where the subscript dot is used to indicate the mean across $i = 1,\ldots,I$; 'centering across the combination of modes A and B' leads to computing centered data as

$$\widetilde{x}_{ijk} = x_{ijk} - \overline{x}_{..k} \tag{21}$$

The type of normalization used is indicated by specifying the entities *within* which the data are normalized. Specifically, 'normalization within (the levels of mode) A' consists of, per entity, first computing the sum of squares of all data elements associated with this particular entity and then dividing all these elements by the square root of this sum of squares (see Figure 4b for 'normalization within B'). In formulae, for three-way data, normalizing within A leads to computing a normalization factor

$$\nu_i = \sqrt{\sum_{j=1}^{J}\sum_{k=1}^{K} x_{ijk}^2} \tag{22}$$

and computing the normalized data as

$$\widetilde{x}_{ijk} = x_{ijk}/\nu_i \tag{23}$$

Normalization 'within the combination of modes A and B' consists of, for each combination of an entity of mode A and one of mode B, first computing the sum of squares over all other modes and then dividing all associated elements by the square root of this sum of squares.

Why do we propose different terminology for centering and normalization? For instance, for four-way data, normalizing within A could also be denoted as normalizing across the combination of modes B, C and D. One reason for using *across* for centering and *within* for normalization is that in computing the centered data the subtracted average is best recognized by the mode(s) across which the average is computed (the ones for which dot subscripts are used: $\bar{x}_{.jk}$ contains the average across i), whereas in normalization the scaling factor usually carries the subscript for the entities of the mode(s) *within* which the normalization takes place (e.g. ν_j). A second reason is that the preferred procedures for centering and normalizing multiway data are centering *across* one mode at a time and normalizing *within* one mode at a time (see Reference [17], where it is argued that only these types of preprocessing leave the structure of the CP or three-way Tucker model intact); with the above terminology these procedures can be referred to in a simple way. If the 'inappropriate' preprocessing procedures are to be used, this (justly) requires rather complex terminology. In three-way analysis these preferred forms of centering and scaling are called *fiber* centering and *slab* scaling. In multiway analysis the term *fiber centering* can still be used for centering across one mode only; normalizing within a mode should then no longer be called slab scaling, however.

It can be useful to center across several different modes successively or to normalize within several modes successively. For instance, we may first center across A and then across B (or *vice versa*, which gives the same result). We denote this as '(fiber) centering across A and across B' (or, for short, 'centering across A and B'). Similarly, normalizing first within A and then within B can be denoted as normalization 'within A and within B respectively'. Here it is important to give the order of the normalizations, because one normalization affects the other, and the reversed order generally leads to a different outcome. Proposals have been made [17,18] for iterative normalizations within different modes. If such iterative procedures are used, this should be mentioned specifically. Also, when using combinations of centering and normalization in most cases, the order should be specified, and if the procedure is performed iteratively, this should be mentioned as well.

5.2. *Rotation: transformation of Tucker models*

The CP model typically gives unique solutions up to permutations and scalings. The three-way Tucker model, on the other hand, is by no means identifiable. Specifically, (7) can be written as

$$\mathbf{X}_a = \mathbf{A}\mathbf{G}_a(\mathbf{C}' \otimes \mathbf{B}') + \mathbf{E}_a = \widetilde{\mathbf{A}}\widetilde{\mathbf{G}}_a(\widetilde{\mathbf{C}}' \otimes \widetilde{\mathbf{B}}') + \mathbf{E}_a \tag{24}$$

with $\tilde{\mathbf{A}} = \mathbf{AS}$, $\tilde{\mathbf{B}} = \mathbf{BT}$, $\tilde{\mathbf{C}} = \mathbf{CU}$ and $\tilde{\mathbf{G}}_a = \mathbf{S}^{-1}\mathbf{G}_a(\mathbf{U}^{-1\prime} \otimes \mathbf{T}^{-1\prime})$. Here the matrices **S**, **T** and **U** are non-singular square matrices of orders $P \times P$, $Q \times Q$ and $R \times R$ respectively.* Sometimes

*Clearly, it would have been nicer if the symbols to denote the orders of the matrices and the symbols for the matrices were the same. However, using symbols **P, Q** and **R** for rotation matrices would be confusing, since **R** usually refers to a correlation matrix. Conversely, using *S, T* and *U* for the numbers of components would bypass the common choices for denoting the number of components as *Q* and *R*. Therefore it seems best to ignore this inconsistency.

they are taken orthonormal, in which case they are called *orthogonal rotation matrices*. When they are not orthonormal, we call them *oblique rotation matrices*.

For the *N*-way Tucker model a similar rotational freedom holds. However, for higher-way cases we will denote the rotation matrices as $\mathbf{S}_1$, $\mathbf{S}_2$, …, $\mathbf{S}_N$, of orders $P_1 \times P_1$, $P_2 \times P_2$,…, $P_N \times P_N$ respectively, and we have

$$\widetilde{\mathbf{A}}_1 = \mathbf{A}_1\mathbf{S}_1, \widetilde{\mathbf{A}}_2 = \mathbf{A}_2\mathbf{S}_2, \ldots, \widetilde{\mathbf{A}}_N = \mathbf{A}_N\mathbf{S}_N \quad \text{and} \quad \widetilde{\mathbf{G}}_a = \mathbf{S}_1{}^{-1}\mathbf{G}_a(\mathbf{S}_N{}^{-1'} \otimes \ldots \otimes \mathbf{S}_2{}^{-1'}) \qquad (25)$$

Obviously, (25) could also be used for the three-way case, but it has the disadvantage that subscripted matrices are needed. This renders formulae somewhat more difficult to read (and write) and is particularly embarrassing in cases where elements of such matrices are to be specified, as we then need three indices.

It frequently happens that rotation in only one 'direction' is to be performed. This comes down to multiplication of an *N*-way array with a matrix. Although such multiplications can easily be written in matrix algebra, as can be seen above, it is sometimes useful to use the mode *n* multiplication (see Section 3). Specifically, multiplication of the array $\underline{\mathbf{G}}$ by $\mathbf{S}_n{}^{-1}$ along the mode *n* is written as $\underline{\tilde{\mathbf{G}}} = \underline{\tilde{\mathbf{G}}} \times_n \mathbf{S}_n{}^{-1}$ and the outcome is equivalent to the matrix multiplication $\tilde{\mathbf{G}}_n = \mathbf{S}_n{}^{-1}\mathbf{G}_n$ involving the mode *n* matricized versions of $\underline{\mathbf{G}}$ and $\underline{\tilde{\mathbf{G}}}$, $\mathbf{G}_n$ and $\tilde{\mathbf{G}}_n$ respectively. Successive multiplications can then be written as $\underline{\tilde{\mathbf{G}}} = \underline{\mathbf{G}} \times_n \mathbf{S}_n{}^{-1} \times_m \mathbf{S}_m{}^{-1}$, etc. The full rotational freedom of the *N*-way Tucker model is represented by $\underline{\tilde{\mathbf{G}}} = \underline{\tilde{\mathbf{G}}} \times_1 \mathbf{S}_1{}^{-1} \times_2 \mathbf{S}_2{}^{-1} \ldots \times_N \mathbf{S}_N{}^{-1}$. The form has a nice symmetric treatment of all modes, but for practical purposes, where actual matrix multiplications are to be carried out, expression (25) is much more useful.

6. OTHER MULTIWAY MODELS AND CONCLUSION

6.1. *Extended models*

A first class of other multiway models that can be considered is that where the fundamental models (CP or Tucker) are extended by additive terms. Such models have been described in Reference [17] and various variants have popped up at several places in the literature. Adding additive terms implies combination with the general linear model, and it is therefore appropriate to follow common notation used there. Thus first-order additive terms are denoted by single-indexed α, β, γ, etc., with as indices simply those that we used for modes A, B, C, etc., leading to α_i, β_j, γ_k, etc. Second-order terms are denoted as double-subscripted $\alpha\beta$, $\alpha\gamma$, etc. Thus $\alpha\beta_{ij}$ denotes the interaction effect of mode A entity *i* with mode B entity *j*, etc. For higher-order terms this system of combining symbols is simply extended.

6.2. *Structurally different N-way models*

The present overview of notational conventions is based on only two models and hence is by no means complete. For other models it is suggested to choose notation as much as possible along the same scheme. Thus first-mode entities are denoted by the letter 'a' and index *i*, while their components are indexed by *p*, and the associated rotation is denoted by **S**, and analogously for the second mode, etc. Furthermore, as far as other models have parameter matrices in common with the 'fundamental' models, these should have the same symbol. If parameters play the same role as those in the fundamental model but differ in certain details, this could be expressed by using the same symbol but with a subscript indicating its special role.

6.3. Concluding remarks

The present paper provides a large number of suggestions for notation and terminology. These suggestions are based on notions and operations that are currently in use in *N*-way analysis. Some notions, e.g. symmetry of (hyper)cubic arrays, have been ignored given their rather high complexity and limited use. The same holds for operations that rearrange *N*-way arrays into *M*-way arrays. The two most common of these are vectorization and matricization (hence with $M = 1$ and $M = 2$ respectively), and these can indeed be expected to remain the two most common rearranging operations, because they are sufficient prerequisites for using matrix algebra. However, arranging *N*-way arrays into three-way arrays may become quite common too, and so may other operations. For arranging tensors into three-way arrays, the term 'ternarizing' would seem appropriate (as a ternary array is sometimes used to denote a three-way array), but we are not aware of similar terminological possibilities for arranging tensors into four- or higher-way arrays. It should be noted that, as for matricizing, for ternarizing it would be necessary to indicate specifically which modes are combined and how this is done (see Section 2·7).

Obviously, the present paper on notation and terminology is based on subjective choices that may not always be deemed most fortunate or useful. However, the aim of this paper is that the notation scheme set out here sets 'a default': when there are no reasons to choose a different notation scheme, use the present one. The use of a common notation scheme is expected to increase readability of multiway literature by enhancing recognizability of model formulations.

ACKNOWLEDGEMENTS

The author is obliged to Jos ten Berge, Claus Andersson, Rasmus Bro, Age Smilde and three anonymous reviewers for helpful comments on an earlier version of this paper.

APPENDIX: OVERVIEW OF NOTATION

$\underline{\mathbf{X}}$, $\underline{\mathbf{Y}}$	*N*-way arrays
$\underline{\mathbf{I}}$	'unit superdiagonal' tensor
$\mathbf{X}_a$, $\mathbf{X}_b$, $\mathbf{X}_c$,… or $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$, …	matricized versions of $\underline{\mathbf{X}}$
Vec( )	vectorized version of a matrix
$i, j, k, \ldots$ or $i_1, i_2, i_3, \ldots$	running indices for respective modes
$I, J, K, \ldots$ or $I_1, I_2, I_3, \ldots$	sizes of respective modes
$\underline{\mathbf{G}}$	core array in Tucker model
$\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$ or $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \ldots$	component matrices for respective modes
$p, q, r, \ldots$ or $p_1, p_2, p_3, \ldots$	running indices for components for respective modes
$P, Q, R, \ldots$ or $P_1, P_2, P_3, \ldots$	numbers of components for respective modes
$\mathbf{S}, \mathbf{T}, \mathbf{U} \ldots$ or $\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \ldots$	rotation matrices for respective modes
$\alpha_i, \beta_j, \gamma_k, \ldots$	additive terms for respective modes
$\alpha\beta_{ij}$, $\alpha\gamma_{ik}$, $\alpha\beta\gamma_{ijk}$, etc.	examples of additive interaction terms
$\otimes$	Kronecker product
$\odot$	columnwise Kronecker product (Khatri–Rao product)
$*$	elementwise or Hadamard product
$\times_n$	mode *n* multiplication
$\mathbf{X}_1 = \mathbf{A}_1\mathbf{G}_1(\mathbf{A}_N' \otimes \ldots \otimes \mathbf{A}_3' \otimes \mathbf{A}_2') + \mathbf{E}_1$	mode 1 matricized version of *N*-way Tucker model
$\mathbf{X}_1 = \mathbf{A}_1(\mathbf{A}_N \odot \ldots \odot \mathbf{A}_3 \odot \mathbf{A}_2)' + \mathbf{E}_1$	mode 1 matricized version of *N*-way CP model

REFERENCES

1. L. R. Tucker, Implications of factor analysis of three-way matrices for measurement of change in: C. W. Harris (Ed.) *Problems in measuring change*. The University of Wisconsin Press, Madison (1963), 122–137.
2. L. R. Tucker, Some mathematical notes on three-mode factor analysis, *Psychometrika*, **31**, 279–311 (1966).
3. J. D. Carroll, & J.-J. Chang, Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition, *Psychometrika* **35**, 283–319 (1970).
4. R. A. Harshman, Foundations of the PARAFAC procedure: models and conditions for an explanatory multi-mode factor analysis. University of California at Los Angeles, *UCLA Working Papers in Phonetics*. **16**, 1–84 (1970).
5. A. Kapteyn, H. Neudecker, & T. Wansbeek, An approach to n-mode components analysis, *Psychometrika* **51**, 269–275 (1986).
6. L. De Lathauwer, *Signal Processing based on Multilinear Algebra*, doctoral dissertation, University of Leuven, Belgium (1997).
7. L. R. Tucker, The extension of factor analysis to three-dimensional matrices in N. Frederiksen & H. Gulliksen (Eds.), *Contributions to mathematical psychology* (pp. 109–127), Holt, Rinehart, and Winston, New York (1980).
8. J. D. Carroll, & P. Arabie, Multidimensional scaling, *Annual Review of Psychology*, **31**, 438–457 (1980).
9. C. H. Coombs, Psychological scaling without a unit of measurement. *Psychological Review*, **57**, 145–158 (1950).
10. J. B. Kruskal, Three-way arrays: Rank and uniqueness of trilinear decompositions. *Linear Algebra and its Applications*, **18**, 95–138 (1977).
11. C. R. Rao, & S. K. Mitra, *Generalized Inverse of Matrices and its Applications*. Wiley, New York (1971).
12. J. D. Carroll, S. Pruzansky, & J. B. Kruskal, CANDELINC: a general approach to multidimensional analysis of many-way arrays with linear constraints on parameters. *Psychometrika*, **45**, 3–24 (1980).
13. J. R. Schott, *Matrix Analysis for Statistics*. Wiley: New York (1997).
14. P. M. Kroonenberg, & J. de Leeuw, Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika* **45**, 69–97 (1980).
15. H. A. L. Kiers, Hierarchical relations among three-way methods, *sychometrika* **56**, 449–470).
16. R. Bro, *Multiway analysis in the food industry. Models, algorithms and applications*, Doctoral dissertation, University of Amsterdam (1998).
17. R. A. Harshman, & M. E. Lundy, Data preprocessing and the extended PARAFAC model in H. G. Law, C. W. Snyder, J. A. Hattie, and R. P. McDonald (Eds.), *Research methods for multimode data analysis*, (pp. 216–284). Praeger, New York (1984).
18. J. M. F. ten Berge, Convergence of PARAFAC preprocessing procedures and the Deming-Stephan method of iterative proportional fitting in R. Coppi, & S. Bolasco *Multiway Data Analysis* (pp. 53–63). Elsevier, Amsterdam (1989).

