A CASE OF EXTREME SIMPLICITY OF THE CORE MATRIX IN THREE-MODE PRINCIPAL COMPONENTS ANALYSIS

TAKASHI MURAKAMI

NAGOYA UNIVERSITY

Jos M. F. TEN BERGE AND HENK A. L. KIERS

UNIVERSITY OF GRONINGEN

In three-mode Principal Components Analysis, the $P \times Q \times R$ core matrix G can be transformed to simple structure before it is interpreted. It is well-known that, when P = QR, G can be transformed to the identity matrix, which implies that all elements become equal to values specified a priori. In the present paper it is shown that, when P = QR - 1, G can be transformed to have nearly all elements equal to values specified a priori. A closed-form solution for this transformation is offered. Theoretical and practical implications of this simple structure transformation of G are discussed.

Key words: three-mode principal components analysis, core matrix rotations, simple structure.

Let X be a three-way data array of order $I \times J \times K$. In three-mode PCA (Kroonenberg & de Leeuw, 1980; Tucker, 1966) X is decomposed as $X = \tilde{X} + E$, where \tilde{X} is the structural (explained) part of X, and E is the residual part. Let the K frontal slabs X_1, \ldots, X_K of X, each of order $I \times J$, be juxtaposed in the $I \times JK$ supermatrix $X_f = [X_1 | \ldots | X_K]$. Then the structural part of X can be written as

$$\tilde{\mathbf{X}}_f = \mathbf{A}\mathbf{G}_f(\mathbf{C}' \otimes \mathbf{B}'),\tag{1}$$

where **A**, **B**, and **C** are column-wise orthonormal component matrices of order $I \times P$, $J \times Q$, and $K \times R$, respectively, and G_f is a $P \times QR$ matrix containing the R frontal $P \times Q$ slabs of the core array **G**. The model in (1) is called the Tucker-3 model.

To see the rotational indeterminacy inherent to (1), let S, T and U be nonsingular matrices of order $P \times P$, $Q \times Q$ and $R \times R$, respectively. Then it is obvious that (1) can be written equivalently as

$$\tilde{\mathbf{X}}_f = (\mathbf{A}\mathbf{S}'^{-1})\mathbf{S}'\mathbf{G}_f(\mathbf{U}\otimes\mathbf{T})(\mathbf{U}^{-1}\mathbf{C}'\otimes\mathbf{T}^{-1}\mathbf{B}'), \tag{2}$$

which shows that G_f can be transformed to $S'G_f(U \otimes T)$ provided that A, B and C are postmultiplied by the inverse of S', T' and U', respectively.

Tucker (1966) already noticed that one may exploit the rotational indeterminacy to rotate the component matrices (A, B and C) or the core to simple structure. Rotation of component matrices to simple structure (e.g., by varimax) is relatively straightforward. Rotation of the core to simple structure is more involved, partly because the generalization of simple structure criteria is not straightforward. Some procedures (e.g., Cohen, 1974; Kiers, 1992; Kroonenberg, 1983; MacCallum, 1976) avoid this problem, by approximating an a priori specified pattern of zeros in the core, for example, diagonal frontal slabs or the super-diagonal form as appearing in CANDECOMP/PARAFAC. Others are based on

Requests for reprints should be sent to Takashi Murakami, Department of Educational Psychology, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-01, Japan. Constructive comments from anonymous reviewers are gratefully acknowledged.

specially designed three-way simple structure criteria, and seek a simple core exploratorily through orthogonal or oblique methods of rotation (e.g., Kruskal (1988); Kiers (1997, 1998b); also see Kiers, 1998a, for a review). To the extent that the latter type of rotation produces a sparse core, the Tucker-3 model is greatly simplified because only a few combinations of components from **A**, **B** and **C** need to be taken into account.

In the context of simplifying the core array, the question arises to what extent simplicity can be obtained. Specifically, it is important to ascertain how many elements of the core matrix can be set to zero by nonsingular transformations. A trivial example is the case where P = QR. In that case, G_f is a square matrix, and premultiplying G_f by $S' = G_f^{-1}$ produces a core array with $P^2 - P$ zero elements. However, experiments with analytic core rotation methods have demonstrated that similarly simple core arrays also can be obtained in certain non-trivial cases, where P < QR. In particular, when P = QR - 1, G_f can generally be transformed to have a vast majority of elements zero. The present paper is meant to clarify this case, and offers a closed form solution for non-singular transformations S, T and U such that only a few nonzero elements in $S'G_f(U\otimes T)$ remain, for any given G_f of order $P \times Q \times R$ with P = QR - 1.

Simple Cores When
$$P = QR - 1$$

To show how simple core arrays can be found when P = QR - 1, we shall first treat the special case P = 5, Q = 3, R = 2 in full detail. A closed-form solution will be presented, illustrated (Table 1), and proven. The general solution for simplicity when P = QR - 1 will be deferred till the next section.

The case $\{P=5, Q=3, R=2\}$ starts from a 5×6 core matrix G_f , consisting of two frontal 5×3 slabs G_1 and G_2 . It is essential (for reasons that will become clear in due course) to require that G_f has previously been orthonormalized rowwise. This can be done, for any G_f^* of full row rank, by premultiplying G_f^* by any orthonormalizer, such as $(G_f^*G_f^{*'})^{-1/2}$, which represents an oblique transformation. When the core array has been obtained from the standard TUCKALS3 algorithm (Kroonenberg & de Leeuw, 1980), it is orthogonal in every direction, whence the orthonormalization simplifies to unit-length rescaling. At any rate, let it be given that

$$\mathbf{G}_f \mathbf{G}_f' = \mathbf{I}_5. \tag{3}$$

It will now be explained how to find nonsingular, and, in fact, orthonormal matrices S, T and U, such that $S'G_f(U \otimes T)$ is simple, of the form

$$\mathbf{S'G}_{f}(\mathbf{U}\otimes\mathbf{T}) = \begin{pmatrix} 0 & 1 & 0 & | & 0 & 0 & 0 \\ 0 & 0 & 1 & | & 0 & 0 & 0 \\ 0 & 0 & 0 & | & 1 & 0 & 0 \\ 0 & 0 & 0 & | & 0 & 0 & 1 \\ -\delta_{2} & 0 & 0 & | & 0 & \delta_{1} & 0 \end{pmatrix}$$
(4)

for certain scalars δ_1 and δ_2 . To find the S, T, and U, the following procedure is to be followed, after making sure that (3) holds:

- 1. Compute the vector \mathbf{y}' which completes \mathbf{G}_f to a square orthonormal matrix, when it is appended to \mathbf{G}_f as a sixth row.
- 2. Rearrange the elements of y in a 3×2 matrix Y such that y = Vec (Y). So when

$$\mathbf{y'} = [y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5 \quad y_6], \text{ then } \mathbf{Y} = \begin{bmatrix} y_1 & y_4 \\ y_2 & y_5 \\ y_3 & y_6 \end{bmatrix}.$$

TABLE 1

An Exemplary 5-3-2 Array

$$\mathbf{G}_f = \begin{pmatrix} .58 & .43 & .44 & .38 & .04 & .37 \\ -.25 & -.04 & -.23 & .11 & .78 & .51 \\ -.48 & .13 & .33 & .63 & .17 & -.47 \\ .58 & -.56 & -.08 & .19 & .37 & -.41 \\ .16 & .45 & -.79 & .34 & -.10 & -.17 \end{pmatrix}$$

$$\mathbf{y}' = \begin{pmatrix} .11 & .53 & .13 & -.54 & .46 & -.43 \end{pmatrix}$$

$$\mathbf{Y} = \begin{pmatrix} .11 & .54 \\ .53 & .46 \\ .13 & -.43 \end{pmatrix} \quad \mathbf{U} = \begin{pmatrix} .30 & .95 \\ .95 & -.30 \end{pmatrix} \quad \mathbf{T} = \begin{pmatrix} -.56 & .00 & -65 \\ .70 & .71 & -.06 \\ .43 & .49 & .76 \end{pmatrix} \quad \begin{array}{c} \delta_1 = .85 \\ \delta_2 = .52 \end{array}$$

$$\mathbf{H} = \mathbf{G}_f(\mathbf{U} \otimes \mathbf{T}) = \begin{pmatrix} -.40 & .63 & .01 & -.10 & .65 & -.07 \\ .32 & .74 & .25 & .12 & -.52 & -.09 \\ .03 & .20 & -.58 & .22 & -.05 & .76 \\ .11 & .11 & -.56 & -.75 & -.18 & -.25 \\ -.01 & .02 & -.54 & .60 & .01 & -.59 \end{pmatrix} \quad \mathbf{S} = \begin{pmatrix} .63 & .01 & -.10 & -.07 & -.76 \\ .74 & .25 & .12 & -.09 & .61 \\ .20 & -.58 & .22 & .76 & .06 \\ .11 & .56 & -.75 & -.25 & .21 \\ .02 & -.54 & .60 & -.59 & -.01 \end{pmatrix}$$

$$\mathbf{S}'\mathbf{G}_f(\mathbf{U} \otimes \mathbf{T}) = \begin{pmatrix} 0 & 1 & 0 & | & 0 & 0 & 0 \\ & 0 & 0 & | & | & 0 & 0 & 0 \\ & 0 & 0 & | & 1 & 0 & 0 & 0 \\ & 0 & 0 & 0 & | & 0 & 0 & 1 \\ & -.52 & 0 & 0 & | & 0 & .85 & 0 \\ \end{array}$$

- 3. Compute the singular value decomposition (SVD) $\mathbf{Y} = \mathbf{TDU'}$, with \mathbf{T} an orthonormal 3×3 matrix, \mathbf{U} an orthonormal 2×2 matrix, and \mathbf{D} a 3×2 matrix of singular values, of the form $\mathbf{D} = \begin{bmatrix} \delta_1 & 0 \\ 0 & \delta_2 \\ 0 & 0 \end{bmatrix}$.
- 4. Compute the matrix $\mathbf{H} = \mathbf{G}_f(\mathbf{U} \otimes \mathbf{T})$. Take the columns 2, 3, 4 and 6 of \mathbf{H} , and the unit length version of the fifth column of \mathbf{H} as the five columns of \mathbf{S} .
- 5. Using the S, T and U as prescribed here produces the solution of the form (4). A fully worked out example of these computations is given in Table 1.

It remains to prove that the solution given above does indeed transform any rowwise orthonormal $5 \times 3 \times 2$ array G_f to the simple form with 24 zeros, as displayed in (4):

Proof. Define $\mathbf{H} = \mathbf{G}_f(\mathbf{U} \otimes \mathbf{T})$ and $\mathbf{z}' = \mathbf{y}'(\mathbf{U} \otimes \mathbf{T})$. Using the svd $\mathbf{Y} = \mathbf{T}\mathbf{D}\mathbf{U}'$, and noting that $\mathbf{y} = \mathrm{Vec}(\mathbf{Y})$, we have $\mathrm{Vec}(\mathbf{Y}) = \mathbf{y} = (\mathbf{U} \otimes \mathbf{T})$ Vec (D). Hence

$$\mathbf{z} = (\mathbf{U} \otimes \mathbf{T})' \mathbf{y} = (\mathbf{U} \otimes \mathbf{T})' (\mathbf{U} \otimes \mathbf{T}) \operatorname{Vec} (\mathbf{D}) = \operatorname{Vec} (\mathbf{D})$$
 (5)

because (U \otimes T) is orthonormal. Incidentally, it can also be seen that $\mathbf{z}'\mathbf{z} = 1$, whence $\delta_1^2 + \delta_2^2 = 1$.

It is readily verified that

$$\begin{bmatrix} \mathbf{G}_f \\ \mathbf{y}' \end{bmatrix} [\mathbf{U} \otimes \mathbf{T}] = \begin{bmatrix} \mathbf{H} \\ \mathbf{z}' \end{bmatrix}. \tag{6}$$

Because the two matrices in the left-hand side are orthonormal, so is their product, and hence the 5×6 matrix **H** satisfies

$$\mathbf{H}'\mathbf{H} = \mathbf{I}_6 - \mathbf{z}\mathbf{z}'. \tag{7}$$

Next, we note that $\mathbf{z}' = \begin{bmatrix} \delta_1 & 0 & 0 & \delta_2 & 0 \end{bmatrix}$, so

$$\mathbf{H'H} = \begin{pmatrix} 1 - \delta_1^2 & 0 & 0 & 0 & -\delta_1 \delta_2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -\delta_1 \delta_2 & 0 & 0 & 0 & 1 - \delta_2^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{8}$$

It follows that the columns \mathbf{h}_2 , \mathbf{h}_3 , \mathbf{h}_4 and \mathbf{h}_6 of \mathbf{H} form an orthonormal set, and that the fifth column of \mathbf{H} is orthogonal to this set. Therefore, upon constructing $\mathbf{S} = [\mathbf{h}_2 \ \mathbf{h}_3 \ \mathbf{h}_4 \ \mathbf{h}_6 \ (\mathbf{h}_5/||\mathbf{h}_5||)]$, \mathbf{S} is orthonormal and $\mathbf{S}'\mathbf{H}$ displays, up to a permutation, precisely the zeros and ones as we have in $\mathbf{H}'\mathbf{H}$, when the first row is deleted. As for the two nonzero entries of the fifth row of $\mathbf{S}'\mathbf{H}$, it is readily seen that $\mathbf{h}_5'\mathbf{h}_1/||\mathbf{h}_5|| = -\delta_1\delta_2/\delta_1 = -\delta_2$, and $\mathbf{h}_5'\mathbf{h}_5/||\mathbf{h}_5|| = \delta_1$.

A remarkable feature of the solution is the orthonormality of S, T and U. Admittedly, however, the orthonormality of S would be lost if G_f were not rowwise orthonormal. If the initial orthonormalization of G_f is subsumed under the solution, we have an oblique transformation S, and orthogonal rotations T and U to produce the simple core matrix.

The General Case Where
$$P = QR - 1$$

In the $5 \times 3 \times 2$ case treated above, 24 of the 30 core elements were transformed to zero. In general, when P = QR - 1, and the modes are defined such that $Q \ge R$ (as can be done without loss of generality), we can transform the array to have only R(Q + R - 2) nonzero elements, which implies that we have QR(QR - 1) - R(Q + R - 2) zeros. This can be obtained by the following generalized algorithm:

Let G_f be a rowwise orthonormal $P \times Q \times R$ array with P = QR - 1.

- 1. Compute the vector \mathbf{y}' which completes \mathbf{G}_f to a square orthonormal matrix, when it is appended to \mathbf{G}_f as QR-th row.
- 2. Rearrange the elements of y in a $Q \times R$ matrix Y such that y = Vec (Y). So when $y' = [y'_1| \dots |y'_R|]$, then $Y = [y_1| \dots |y_R|]$.
- 3. Compute the SVD Y = TDU', with T an orthonormal $Q \times Q$ matrix, U an orthonormal $R \times R$ matrix, and D a $Q \times R$ matrix containing the R singular values of Y, and zeros. Define the index set IND as the set of R row indices of Vec (D) corresponding to the

TABLE 2											
Properties	of	the	Simplified	Core	Matrix	in	a	Number	of	Cases	

P	Q	R	Total	Unity	Unspecified	Zero	% of zeros	***************************************
Orthog	onal l	Method						,
3	2	2	12	2	2	8	66.7	
5	3	2	30	4	2	24	80.0	
7	4	2	56	6	2	48	85.7	
8	3	3	72	6	5	61	84.7	
9	5	2	90	8	2	80	88.9	
11	4	3	132	9	5	118	89.4	
15	4	4	240	12	9	219	91.3	
Oblique	e Met	hod						
8	3	3	72	8	2	62	86.1	
11	4	3	132	11	2	119	90.2	
15	4	4	240	15	3	222	92.5	

singular values in Vec (**D**). So IND = $\{1, Q + 2, 2Q + 3, \dots\}$. Define the complementary set as IND^c, corresponding to the QR - R zero elements of Vec (**D**).

Using the S, T and U of the above generalized algorithm produces a simple array $S'G_f(U\otimes T)$. It can be seen as a rowwise and columnwise permuted diagonal supermatrix with the identity matrix of order QR - R in the upper left hand corner, and a nonzero submatrix W of order $(R-1) \times R$ in the lower right hand corner, yielding a total of QR - R + R(R-1) nonzero elements. We can (for R > 2) premultiply S' by a rotation matrix that puts W in a quasi-upper triangular form, which means that $\frac{1}{2}(R-2)(R-1)$ elements of W can be rotated to zero. When this rotation is subsumed under S, we end up with orthonormal matrices S, T and U that yield $QR - R + R(R-1) - \frac{1}{2}(R-2)(R-1) = \frac{1}{2}R(R-1) + QR - 1$ nonzero elements in $S'G_f(U\otimes T)$. What this amounts to in various specific cases can be found in Table 2 under "Orthogonal Method". Typically, over 80% of the core elements can be transformed to zero.

A formal proof for the generalized algorithm will not be given here, but is available from the authors upon request (The key to the proof is the observation that (5), (6) and (7) hold generally). A few technical comments, however, are in order. First of all, the submatrix **W** that arises in the simplified core array is directly related to the singular values of **Y**: When these are arranged in a vector **v**, it can be shown that $[\mathbf{W}'|\mathbf{v}]$ is an orthonormal $R \times R$ matrix. The rotational freedom in **W** results directly from the indeterminacy of S_2 .

^{4.} Compute the matrix $\mathbf{H} = \mathbf{G}_f(\mathbf{U} \otimes \mathbf{T})$. Put those columns of \mathbf{H} indexed by \mathbf{IND}^c in a $P \times (QR - R)$ matrix \mathbf{S}_1 . Compute any $P \times (R - 1)$ matrix \mathbf{S}_2 such that \mathbf{S} , defined as $[\mathbf{S}_1|\mathbf{S}_2]$, is an orthonormal $P \times P$ matrix.

There is also an indeterminacy in **T**, in that the last (Q - R) columns of **T** are determined up to an orthogonal rotation.

Finally, there is a way of further simplifying the core matrix by quasi-diagonalizing **W**. This can be done by postmultiplying **S** by a $P \times P$ matrix that is constructed as \mathbf{I}_P , with an $(R-1) \times (R-1)$ diagonal block replaced by the inverse of **W***, where **W*** is an arbitrary $(R-1) \times (R-1)$ submatrix of **W**. The number of nonzero elements will then be further reduced to $QR - R + R(R-1) - (R-1)^2 + (R-1) = R(Q+1) - 2$. Examples of what this amounts to are given in Table 2, under "Oblique Method".

Discussion

It is well-known (Tucker, 1966, p. 288) that there is no point in having P larger than QR in fitting the Tucker-3 model: When the model is fit in QR + Q + R components for the A, B, and C mode, respectively, and P > QR, the same fit will be obtained when QR instead of P components are determined for the A mode. In this sense, having P = QR components is the largest admissible case for Tucker-3 when Q and R are considered fixed. Transforming the core to the identity matrix trivially produces simplicity, with $100 \times (QR - 1)/QR$ per cent of elements zero. The present paper has dealt with the "second largest case", where P = QR - 1. Again, a vast majority of core elements can be transformed to zero, as can be seen from Table 2. The percentages of elements that become zero are smaller than in the case P = QR, yet still quite high. Unfortunately, attempts to find the same type of simplicity with P = QR - 2, P = QR - 3, ... have failed. The approach we have adopted crucially relies on the possibility of constructing a $Q \times R$ matrix Y, the columns of which can be stacked in a row-vector that completes G_f to a square orthonormal matrix. For values of P other than QR - 1, no such Y exists.

The transformation to simplicity of the core has various implications. It is well-known, also see (2), that three-mode PCA is grossly overparameterized. This paper shows how the overparameterization may manifest itself in the simplicity of the core. Specifically, in three-mode PCA with P=QR-1, simplicity of the core should not be mistaken for a property of the data: The ensuing simplified version of the three-mode component solution, leaving only a few combinations of the components to be considered, is merely an artifact. In practice, solutions obtained by simplicity rotation techniques will be preferred, especially if they also take into account the simplicity of the component matrices. The artificially obtained simplicity as described here can be used as a baseline to compare such solutions with: If a solution is found in which the core is not nearly as simple as it can be made artificially, this suggests searching for simpler solutions, unless that would have detrimental effects to the interpretation of the component matrices.

Secondly, the explicit baselines of simplicity that we have established can be used to test the efficiency of any iterative method of transforming a core to simplicity. In particular, the method by Kiers (1998b) can now be tested in non-trivial cases (the trivial ones referring to the case P = QR). Because the global minimum for Kiers' method is known in such cases, it can be ascertained how sensitive his method is to local minima.

Conceivably, the simplicity result of the present paper may also have implications beyond the realm of three-way PCA. The transformations to simplicity can be applied to any three-way array rather than just a core array in three-mode component analysis. Since the transformations are non-singular, they preserve the three-way rank as defined by Kruskal (see Kruskal, 1989, for a review). The number of nonzero elements in a three-way array cannot be less than its three-way rank. Accordingly, simplicity of an array may facilitate the evaluation of its three-way rank.

References

- Cohen, H. S. (1974, June). *Three-mode rotation to approximate INDSCAL structure (TRIAS)*. Paper presented at the Psychometric Society Meeting, Palo Alto, CA.
- Kiers, H. A. L. (1992). TUCKALS core rotations and constrained TUCKALS modelling. *Statistica Applicata*, 4, 659-667.
- Kiers, H. A. L. (1997). Three-mode Orthomax rotation. Psychometrika, 62, 579-598.
- Kiers, H. A. L. (1998a). Recent developments in three-mode factor analysis: Constrained three-mode factor analysis and core rotations. In C. Hayashi, N. Ohsumi, K. Yajima, Y. Tanaka, H.-H. Bock, & Y. Baba (Eds.), Data Science, classification and related methods. Tokyo: Springer Verlag.
- Kiers, H. A. L. (1998b). Three-way simplimax for oblique rotation of the three-mode factor analysis core to simple structure. Computational Statistics and Data Analysis.
- Kroonenberg, P. M. (1983). Three-mode principal component analysis. Leiden, DSWO.
- 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.
- Kruskal, J. B. (1988, June). Simple structure for three-way data: A new method intermediate between 3-mode factor analysis and PARAFAC-CANDECOMP. Paper presented at the 53rd annual Meeting of the Psychometric Society. Los Angeles.
- Kruskal, J. B. (1989). Rank, decomposition, and uniqueness for 3-way and N-way arrays. In R. Coppi & S. Bolasco (Eds.), *Multiway data analysis* (7–18). Amsterdam: Elsevier Science Publishers.
- MacCallum, R. C. (1976). Transformations of three-mode multidimensional scaling solution to INDSCAL form. *Psychometrika*, 41, 385–400.
- Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, 31, 279–311.

Manuscript received 1/22/96

Final version received 8/6/97