ELSEVIER Linear Algebra and its Applications 294 (1999) 169–179

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Simplicity of core arrays in three-way principal component analysis and the typical rank of $p \times q \times 2$ arrays ☆

Jos M.F. Ten Berge *, Henk A.L. Kiers

Heijmans Institute, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands

Received 19 June 1998; accepted 4 March 1999

Submitted by R.A. Brualdi

Abstract

Interpreting the solution of a Principal Component Analysis of a three-way array is greatly simplified when the core array has a large number of zero elements. The possibility of achieving this has recently been explored by rotations to simplicity or to simple targets on the one hand, and by mathematical analysis on the other. In the present paper, it is shown that a $p \times q \times 2$ array, with $p > q \geqslant 2$, can almost surely be transformed to have all but $2q$ elements zero. It is also shown that arrays of that form have three-way rank p at most. This has direct implications for the typical rank of $p \times q \times 2$ arrays, also when $p = q$. When $p \geqslant 2q$, the typical rank is $2q$; when $q < p < 2q$ it is p, and when $p = q$, the rank is typically (almost surely) p or $p + 1$. These typical rank results pertain to the decomposition of real valued three-way arrays in terms of real valued rank one arrays, and do not apply in the complex setting, where the typical rank of $p \times q \times 2$ arrays is also min[p,$2q$] when $p > q$, but it is p when $p = q$.

Keywords: Three-way rank; Typical tensorial rank; Candecomp; Parafac; Three-mode principal component analysis; Core arrays; Simple structure

☆ The authors are obliged to Alexander Shapiro for pointing out Result 2 and its corollary.

* Corresponding author. Tel.: +31-50-3636349; fax: +31-50-3636304; e-mail: j.m.f.ten.berge@ppsw.rug.nl

0024-3795/99/$ – see front matter
PII: S0024-3795(99)00057-9

1. Introduction

Three-mode principal component analysis [21,13] represents a three-way data array by means of component matrices for each of the three modes (viz. individuals, variables and occasions) and a three-way array called "the core", describing the interactions between these components. Specifically, when $\underline{\mathbf{X}}$ is an $I \times J \times K$ data array, then the three-mode principal component analysis model is

$$\hat{x}_{ijk} = \sum_{\alpha=1}^{p}\sum_{\beta=1}^{q}\sum_{\gamma=1}^{r} a_{i\alpha} b_{j\beta} c_{k\gamma} g_{\alpha\beta\gamma},$$

where $\hat{x}_{ijk}$ denotes the estimate for the element (i,j,k) of $\underline{\mathbf{X}}$; $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ (with elements $a_{i\alpha}, b_{i\beta}$, and $c_{k\gamma}$) are component matrices of orders $I \times p$, $J \times q$ and $K \times r$, respectively, and $\underline{\mathbf{G}}$ is a $p \times q \times r$ three-way array denoted as the *core*, with elements

$$g_{\alpha\beta\gamma}, \quad i = 1,\ldots,I,\ j = 1,\ldots,J,\ k = 1,\ldots,K,$$
$$\alpha = 1,\ldots,p,\ \beta = 1,\ldots,q,\ \text{and}\ \gamma = 1,\ldots,r.$$

The matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ can be considered component score or loading matrices for the A-, B- and C-mode-entries, respectively. The elements of the core indicate how components from different modes interact. Three-mode principal component analysis consists of fitting the above model to a data array by minimizing the sum of squared residuals $\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(x_{ijk} - \hat{x}_{ijk}\right)^2$ over $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\underline{\mathbf{G}}$.

As already noted by Tucker [21], the three-mode principal component analysis model is not uniquely determined: Nonsingular transformations of the component matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ do not affect the model provided that they are compensated for in the core. Specifically, the component matrices $\tilde{\mathbf{A}} = \mathbf{A}(\mathbf{S}')^{-1}$, $\tilde{\mathbf{B}} = \mathbf{B}(\mathbf{T}')^{-1}$, and $\tilde{\mathbf{C}} = \mathbf{C}(\mathbf{U}')^{-1}$ and the core $\tilde{\underline{\mathbf{G}}}$ the elements of which are defined as

$$\tilde{g}_{ijk} = \sum_{\alpha=1}^{p}\sum_{\beta=1}^{q}\sum_{\gamma=1}^{r} s_{\alpha i} t_{\beta j} u_{\gamma k} g_{\alpha\beta\gamma}, \quad i = 1,\ldots,p,\ j = 1,\ldots,q,\ k = 1,\ldots,r,$$

give the same model estimates for $\hat{\mathbf{X}}$ as do $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\underline{\mathbf{G}}$. As a consequence, when we have obtained a solution, we may always transform this in three directions to get a, in some respects, more attractive solution.

Until recently, a major obstacle to practical applications of three-mode principal component analysis has been the lack of a procedure for enhancing simplicity of the core array. Simplifying the core array in three-mode PCA by nonsingular transformations has become a topic of increasing interest, see Refs. [8–11,19]. Experience with some of these procedures has revealed that nonsingular transformations of $p \times q \times r$ $(p \geqslant q \geqslant r)$ core arrays, when $p = qr - 1$, can produce a vast majority of zeroes. A direct, noniterative expression for the simplifying transformations has been given by Murakami et al. [18]. This means that a high degree of simplicity can be obtained at once in the restricted class of arrays with $p = qr - 1$.

In the present paper, another class of arrays is treated for which high simplicity is feasible: The set of $p \times q \times 2$ arrays with $p > q$. A method will be given to transform these arrays into a form of simultaneous quasi-diagonality. Specifically, the two $p \times q$ slices $(p > q)$ $\mathbf{X}_1$ and $\mathbf{X}_2$ of the $p \times q \times 2$ array $\underline{\mathbf{X}}$ (denoting any three-way array from now on) are transformed into $\mathbf{Y}_1$ and $\mathbf{Y}_2$, respectively, such that $\mathbf{Y}_1$ has $\mathbf{I}_q$ (the $q \times q$ identity matrix) in the first q rows and zeroes elsewhere, and $\mathbf{Y}_2$ has $\mathbf{I}_q$ in the last q rows, and zeroes elsewhere. The transformation relies on nonsingular matrices $\mathbf{S}(p \times p)$ and $\mathbf{T}(q \times q)$ such that

$$\mathbf{S}'\mathbf{X}_1\mathbf{T} = \mathbf{Y}_1 = \begin{bmatrix} \mathbf{I}_q \\ \mathbf{O} \end{bmatrix} \quad \text{and} \quad \mathbf{S}'\mathbf{X}_2\mathbf{T} = \mathbf{Y}_2 = \begin{bmatrix} \mathbf{O} \\ \mathbf{I}_q \end{bmatrix}. \tag{1}$$

A remarkable feature of (1) is that the transformations consist only of a premultiplication of the slices by a $p \times p$ matrix $\mathbf{S}'$ and postmultiplication by a $q \times q$ matrix $\mathbf{T}$: Multiplication in the third direction (which means taking linear combinations of $\mathbf{X}_1$ and $\mathbf{X}_2$) is conspicuously absent. The proposed method of simplifying $p \times q \times 2$ arrays will be shown to work almost surely.

The simplicity result of the present paper has important implications for the rank of three-way arrays [14,16]. To define the rank of a three-way array, Kruskal used the concept of a rank one array. A three-way array is of rank one when it is the outer product of three vectors [16, p. 8]. Equivalently, a $p \times q \times r$ array is of rank one when its r slices are proportional to the same $p \times q$ matrix of rank one. When proportionality to the same rank one matrix holds for the slices in one direction, it also holds in the other two directions. The rank of a three-way array $\underline{\mathbf{X}}$ is defined as the smallest number of rank-one arrays that generate $\underline{\mathbf{X}}$ as their sum. Equivalently, the rank of a three-way array is the smallest value of s for which the slices $\mathbf{X}_1, \ldots, \mathbf{X}_r$ of the array can be decomposed as $\mathbf{X}_j = \mathbf{A}\mathbf{D}_j\mathbf{B}'$, $j = 1, \ldots, r$, for some $p \times s$ matrix $\mathbf{A}$, some $q \times s$ matrix $\mathbf{B}$, and some diagonal $s \times s$ matrix $\mathbf{D}_j$, $j = 1, \ldots, r$. This decomposition is well-known as the CANDECOMP/PARAFAC decomposition in s dimensions [3,6]. In other words, the smallest dimensionality that allows a CANDECOMP/PARAFAC decomposition equals the three-way rank of the data array.

In ordinary matrix algebra, the maximum rank of a $p \times q$ matrix $(p > q)$ is q, and the maximum rank is also the *typical* rank of the matrix, that is, the rank a matrix has almost surely. Specifically, when the elements of the $p \times q$ matrix are sampled from a continuous distribution, the matrix will have rank q almost surely (with probability one), so its typical rank is q. In three-way analysis, however, typical rank and maximum rank no longer coincide. For instance, Kruskal [15] has reported that a $2 \times 2 \times 2$ array has maximum rank 3, but the rank is typically (almost surely) either 2 or 3. For a $4 \times 3 \times 2$ array, the maximum rank is five, yet such arrays have rank 4 almost surely [15, p. 26].

The problem of how to determine either the maximum rank or the typical rank of a $p \times q \times r$ array $(p \geqslant q \geqslant r)$ from p, q, and r has not been solved. Apart from some miscellaneous results (e.g., see [16, p. 10] and [5, pp. 214, 215]) hardly anything has been achieved for cases with $r > 2$. However, for the special case of $p \times q \times 2$ arrays, $p \geqslant q \geqslant 2$, two results of some generality have been established. Kruskal ([16, p. 10] also see [7]) has given an explicit expression for the *maximum* rank of $p \times q \times 2$ arrays. When $p \geqslant q \geqslant 2$,

$$\text{max rank}\{p, q, 2\} = q + \min[q, \text{floor}(p/2)],$$

where floor(x) is the largest integer equal to or below x. For instance, max rank {4,3,2} is 5; max rank {5,4,2} is 6, and max rank {6,5,2} is 8.

The second result to be mentioned here pertains to the typical rank, for cases where $r = 2$ and $p = q$. Generalizing earlier work by Kruskal [16], Ten Berge [20] has shown that $p \times p \times 2$ arrays $(p \geqslant 2)$ have rank p with some probability P, $0 < P < 1$, and a rank higher than p with some probability $1 - P$. In the present paper, a more general result for the typical rank of $p \times q \times 2$ arrays will be established: The issue of typical rank will be solved for all $p \times q \times 2$ arrays, those with $p = q$ included.

The organization of this paper is as follows. First, we deal with transformations to simplicity of $p \times q \times 2$ arrays with $p > q$. Rank-preserving transformations will be used that bring these arrays into the simple form described in (1). Next, the transformations will be shown to exist almost surely and it will be explained that the three-way array $\underline{\mathbf{Y}}$, consisting of the slices $\mathbf{Y}_1$ and $\mathbf{Y}_2$ of the form given in (1), has rank min[p,2q]. The implication is that, almost surely, the $p \times q \times 2$ arrays have rank p when $q < p < 2q$, and rank $2q$ when $p \geqslant 2q$.

Finally, we turn to $p \times p \times 2$ arrays. Although a transformation analogous to (1) is not generally available here, any such array is easily extended to a $(p + 1) \times p \times 2$ array by adding a random slice to the array. Applying the typical rank result, obtained for cases with $p > q$, to the extended array, we can show that the $p \times p \times 2$ arrays have rank p or $p + 1$, almost surely.

The net result is twofold. First, we may, for all practical purposes, use the typical rank instead of the (often much higher) maximum rank (see Kruskal's formula given above) for $p \times q \times 2$ arrays. This implies, for instance, that almost every $p \times q \times 2$ array has a CANDECOMP/PARAFAC decomposition in min[p,2q] dimensions. In addition, we generalize the result on rank volumes for $p \times p \times 2$ arrays. Instead of knowing that, almost surely, the rank is p or larger than p, this paper establishes that the rank is either p or p+1, almost surely.

It should be pointed out that the present paper is exclusively concerned with the decomposition of real valued arrays in terms of real valued parameters. Generalized approaches, involving complex values, have been the subject of investigation in computational complexity theory, e.g., see [22,1,2].

2. Transformation to simplicity when $p > q$

When $p \geqslant 2q$, obtaining simplicity as in (1) is trivial: When the $2q$ columns of $\mathbf{X}_1$ and $\mathbf{X}_2$ are linearly independent, we can simply insert $p - 2q$ columns in between $\mathbf{X}_1$ and $\mathbf{X}_2$ to construct a nonsingular $p \times p$ matrix. Premultiplying $\mathbf{X}_1$ and $\mathbf{X}_2$ with the inverse of that matrix already yields $\mathbf{Y}_1$ and $\mathbf{Y}_2$ in agreement with (1). It is also evident that usually the rank is at least $2q$, because every column of $\mathbf{Y}_1$ and $\mathbf{Y}_2$ has to be in the column space of a $p \times s$ matrix $\mathbf{A}$, see the previous section, to obtain a CANDECOMP/PARAFAC decomposition in s dimensions. The rank is also at most $2q$, because we may trivially take $\mathbf{A} = [\mathbf{Y}_1 | \mathbf{Y}_2]$, $\mathbf{B} = [\mathbf{I}_q | \mathbf{I}_q]$, and let $\mathbf{D}_1$ and $\mathbf{D}_2$ have $\mathbf{I}_q$ in their upper left and lower right hand corners, respectively, and zeroes elsewhere, to get a tautological CANDECOMP/PARAFAC decomposition. Accordingly, when $p \geqslant 2q$, the typical rank is trivially $2q$. So we shall only have to determine $\mathbf{S}$ and $\mathbf{T}$ for the case $q < p < 2q$.

We start with the observation that every nonzero row of $\mathbf{Y}_1$ repeats itself as a row of $\mathbf{Y}_2$. Hence, when $\mathbf{S}$ and $\mathbf{T}$ satisfy (1) for a given pair $\{\mathbf{X}_1,\mathbf{X}_2\}$, every nonzero row of $\mathbf{S}'\mathbf{X}_1 = \mathbf{Y}_1\mathbf{T}^{-1}$ is also a row of $\mathbf{S}'\mathbf{X}_2 = \mathbf{Y}_2\mathbf{T}^{-1}$. Specifically, we have, for $i = 1, \ldots, q$, that

$$\mathbf{s}_i'\mathbf{X}_1 = \mathbf{s}_{p-q+i}'\mathbf{X}_2, \tag{2}$$

where $\mathbf{s}_j$ is column j of $\mathbf{S}$. Also, the first $p-q$ rows of $\mathbf{S}'\mathbf{X}_2$ must vanish, and so do the last $p-q$ rows of $\mathbf{S}'\mathbf{X}_1$, which implies that, for $i = 1, \ldots, p-q$,

$$\mathbf{s}_i'\mathbf{X}_2 = \mathbf{0}', \tag{3}$$

and for $i = q+1, \ldots, p$,

$$\mathbf{s}_i'\mathbf{X}_1 = \mathbf{0}'. \tag{4}$$

It should be noted that, when the relations (2)–(4), hold and $\mathbf{S}$ is nonsingular, $\mathbf{S}'\mathbf{X}_1$ usually has rank q, and we get $\mathbf{T}$ at once as $\mathbf{T} = (\mathbf{X}_1'\mathbf{S}\mathbf{S}'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{S}\mathbf{Y}_1$. Accordingly, our main concern is with finding a nonsingular matrix $\mathbf{S}$, the columns of which satisfy (2)–(4).

To solve these equations, it is convenient to express them in matrix form. Define $\mathbf{s} = \mathrm{Vec}(\mathbf{S})$, so $\mathbf{s}' = [\mathbf{s}_1' | \ldots | \mathbf{s}_p']$. Eqs. (2)–(4) are equivalent to orthogonality of $\mathbf{s}$ to the columns of a certain matrix. That matrix can be constructed from $\mathbf{X}_1$ and $\mathbf{X}_2$, as follows. Let $t=p-q$ and let $\mathbf{O}$ be a $p^2 \times qt$ matrix of zeroes. Define $\mathbf{W}_1$ and $\mathbf{W}_2$ as the $p^2 \times (pq+qt)$ matrices $\mathbf{W}_1 = [\mathbf{O} | \mathbf{I}_p \otimes \mathbf{X}_1]$ and $\mathbf{W}_2 = [\mathbf{I}_p \otimes \mathbf{X}_2 | \mathbf{O}]$, where $\otimes$ is the Kronecker product, and define $\mathbf{W} \equiv \mathbf{W}_2 - \mathbf{W}_1$. Observing that $\mathbf{s}'(\mathbf{I}_p \otimes \mathbf{X}_j) = [\mathbf{s}_1'\mathbf{X}_j| \ldots |\mathbf{s}_p'\mathbf{X}_j]$, $j=1,2$, it is readily verified that (2)–(4) are equivalent to

$$\mathbf{s}'\mathbf{W} = \mathbf{0}', \tag{5}$$

the first qt equations in (5) representing (3), the last qt equations representing (4), and the equations in between representing (2). Once a solution for (5) has been found, we construct $\mathbf{S}$, and when $\mathbf{S}$ is nonsingular, $\mathbf{T}$ is easily derived and a solution of the form (1) will be obtained.

The null-space of $\mathbf{W}$ has at least dimension

$$d \equiv p^2 - (pq + qt) = p^2 - q(2p - q) = (p - q)^2 = t^2.$$

When $d = 1$, the set of vectors orthogonal to the columns of $\mathbf{W}$ is of dimension 1, and $\mathbf{s}$ is determined up to a scalar. That scalar does not matter because it will be compensated for by $\mathbf{T}$. However, when $d > 1$, $\mathbf{s}$ can be any vector in the d-dimensional null-space of $\mathbf{W}$, which includes vectors that would generate a singular matrix $\mathbf{S}$, for instance, having a zero column. To obtain a solution for $\mathbf{s}$ that entails a nonsingular $\mathbf{S}$, when that is possible, we suggest taking $\mathbf{s}$ as the unique vector

$$\mathbf{s} = [\mathbf{I}_{p^2} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{+}\mathbf{W}']\mathrm{Vec}(\mathbf{I}_p), \tag{6}$$

where $(\mathbf{W}'\mathbf{W})^{+}$ is the Moore–Penrose inverse of $(\mathbf{W}'\mathbf{W})$. This choice has the property, to be used below, that it yields $\mathbf{S} = \mathbf{I}_p$ in the trivial case where $\underline{\mathbf{X}}$ itself is simple, i.e., when $\underline{\mathbf{X}} = \underline{\mathbf{Y}}$, regardless of p and q. This is Result 1.

Result 1. *When* $\underline{\mathbf{X}} = \underline{\mathbf{Y}}$, *taking* s *according to* (6) *yields* $\mathbf{S} = \mathbf{I}_p$.

Proof. Assume that $\underline{\mathbf{X}} = \underline{\mathbf{Y}}$. To show that $\mathbf{s}$ defined by (6) is equal to $\mathrm{Vec}(\mathbf{I}_p)$, it is sufficient to show that $\mathrm{Vec}(\mathbf{I}_p) = [\mathbf{I}_{p^2} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{+}\mathbf{W}']\mathrm{Vec}(\mathbf{I}_p)$. Hence, defining $\mathbf{v}_p \equiv \mathrm{Vec}(\mathbf{I}_p)$, it is sufficient to show that $\mathbf{v}_p$ is orthogonal to the columns of $\mathbf{W}$, which is the same as showing that $\mathbf{v}_p'\mathbf{W}_1 = \mathbf{v}_p'\mathbf{W}_2$. The latter vectors have the simple form $[\mathbf{0}_{qt}'|\mathbf{v}_q'|\mathbf{0}_{qt}']$, where $\mathbf{0}_{qt}'$ is a row vector of qt zeroes, and $\mathbf{v}_q$ is $\mathrm{Vec}(\mathbf{I}_q)$. To verify this, write $\mathbf{v}_p'\mathbf{W}_2$, with $\mathbf{X}_2 = \mathbf{Y}_2$, as $\mathbf{v}_p'[\mathbf{I}_p \otimes \mathbf{X}_2|\mathbf{O}] = \mathbf{v}_p'[\mathbf{I}_p \otimes \mathbf{Y}_2|\mathbf{O}] = [\mathbf{e}_1'\mathbf{Y}_2| \ldots |\mathbf{e}_p'\mathbf{Y}_2|\mathbf{0}_{qt}'] = [\mathbf{0}_{qt}'|\mathbf{f}_1'| \ldots |\mathbf{f}_q'|\mathbf{0}_{qt}'] = [\mathbf{0}_{qt}'|\mathbf{v}_q'|\mathbf{0}_{qt}']$, where $\mathbf{e}_i$ is column i of $\mathbf{I}_p, i = 1, \ldots, p$, and $\mathbf{f}_j$ is column j of $\mathbf{I}_q, j = 1, \ldots, q$. In an analogous fashion, it can be shown that $\mathbf{v}_p'\mathbf{W}_1 = [\mathbf{0}_{qt}'|\mathbf{v}_q'|\mathbf{0}_{qt}']$. □

A Matlab program [17] to solve for $\mathbf{S}$ and $\mathbf{T}$ using (6) is given in Appendix A. With random data, it never fails to yield simplicity, with $\mathbf{S}$ and $\mathbf{T}$ nonsingular. However, it is possible to contrive data where $\mathbf{S}$ will be singular. In Section 3, it will be proven that, when $\mathbf{X}_1$ and $\mathbf{X}_2$ are randomly sampled from a continuous distribution, the $\mathbf{S}$ constructed by (6) will indeed be nonsingular.

3. Nonsingularity of S

To show that a singular solution for $\mathbf{S}$ will never be encountered in practice, we shall use the following result:

Result 2. *Let $f(\mathbf{x})$ be a real valued analytic function defined on $\mathbb{R}^n$. Suppose that $f(\mathbf{x})$ is not identically zero. Then the set $\{\mathbf{x}\colon f(\mathbf{x}) = 0\}$ is of Lebesgue measure zero in $\mathbb{R}^n$.*

Proof. See [4, Theorem 5.A.2]. □

From Result 2 the following corollary is immediate.

Corollary 1. *Consider an analytic mapping F from $\mathbb{R}^n$ into the space of $p \times p$ matrices. Let $f(\mathbf{x}) \equiv \det(F(\mathbf{x}))$. Then $f(\mathbf{x})$ is an analytic real valued function. Suppose that $f(\mathbf{x})$ is not identically zero. Then, by virtue of Result 2, $f(\mathbf{x})$ is nonzero for almost every $\mathbf{x}$ in $\mathbb{R}^n$ and hence $F(\mathbf{x})$ has rank p for almost every $\mathbf{x}$ in $\mathbb{R}^n$.*

In the present context, let $F(\mathbf{X}_1,\mathbf{X}_2) \equiv \mathbf{S}$, with $\mathbf{s} = \mathrm{Vec}(\mathbf{S})$ defined by (6). This is an analytic mapping from $\mathbb{R}^{2pq}$ into the space of $p \times p$ matrices. Therefore, $\mathbf{S}$ is almost surely nonsingular (Corollary 1) if, for at least one pair $\{\mathbf{X}_1,\mathbf{X}_2\}$, a nonsingular matrix $\mathbf{S}$ satisfying (2)–(4) can be found. Result 1 guarantees the existence of such a pair. We have thus proven Result 3.

Result 3. *A $p \times q \times 2$ array $(p > q \geqslant 2)$ can be brought into the simple form* (4) *by nonsingular matrices* $\mathbf{S}$ *and* $\mathbf{T}$ *almost surely.*

4. The rank of simplified arrays

Having explored the generic possibility of solving (1), we are now in a position to determine the typical rank of such arrays. This will be accomplished by using the Vandermonde matrix $\mathbf{V}$ of order p. That matrix is the $p \times p$ matrix, the ith column of which, for $i = 1, \ldots, p$, contains the integers $1, \ldots, p$, raised to the power $i - 1$. The Vandermonde matrix is nonsingular. When we select the first q columns of $\mathbf{V}$, and the last q columns of $\mathbf{V}$, respectively, two matrices arise with proportionality in each of the p rows, which means that the three-way array consisting of these two matrices has visibly rank p at most, see the example to be given shortly. This reasoning is at the heart of Result 4.

Result 4. *When* $\mathbf{X} = [\mathbf{X}_1 | \mathbf{X}_2]$ *can be transformed to* $\mathbf{Y} = [\mathbf{Y}_1 | \mathbf{Y}_2]$ *as in* (1), *the rank of* $\underline{\mathbf{X}}$ *is* $\min[p, 2q]$.

Proof. Premultiply $\mathbf{Y}$ with the Vandermonde matrix $\mathbf{V}$ of order p. Clearly, $\mathbf{VY}_1$ contains the first q columns of $\mathbf{V}$, and $\mathbf{VY}_2$ contains the last q columns of $\mathbf{V}$. It is readily verified that row i of $\mathbf{VY}_1$ is $[i^0|i|\ldots|i^{q-1}]$ and row i of $\mathbf{VY}_2$ is $[i^t| i^{t+1}| \ldots | i^{p-1}]$. Clearly, the latter row is i^t times the former. It follows that each row of $\mathbf{V}[\mathbf{Y}_1| \mathbf{Y}_2]$, interpreted as a horizontal slice of the corresponding three-way array, can be accounted for by a rank one array. Therefore, the rank of that array is at most p. The rank is also at least p (unless $p > 2q$), because $\mathbf{Y}$ contains p linearly independent columns. So when $\underline{\mathbf{Y}}$ has the form in (1) with $p < 2q$, it has rank p. Because $\underline{\mathbf{Y}}$ arose from nonsingular transformations of $\underline{\mathbf{X}}$, $\underline{\mathbf{X}}$ has also rank p. When $p \geqslant 2q$, the rank is obviously $2q$. $\square$

As an example, consider the $4 \times 3 \times 2$ case. Let $\mathbf{X}_1$ and $\mathbf{X}_2$ be matrices of order 4×3, transformed to the simple form

$$\mathbf{Y} = [\mathbf{Y}_1|\mathbf{Y}_2] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \quad \text{with} \quad \mathbf{V} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix}.$$

Then

$$\mathbf{VY}_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix} \quad \text{and} \quad \mathbf{VY}_2 = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 8 \\ 3 & 9 & 27 \\ 4 & 16 & 64 \end{bmatrix}.$$

Clearly, each row of $\mathbf{VY}_1$ is proportional to the corresponding row of $\mathbf{VY}_2$, which shows that the array $\underline{\mathbf{Y}}$ has at most rank 4.

Combining Results 3 and 4, and noting that the transformation to simplicity preserves the rank of the array involved, we have proven

Result 5. *A* $p \times q \times 2$ *array* ($2 \leqslant q < p \leqslant 2q$) *has typical rank* p.

For $p = 3$, the typical rank coincides with the maximum rank, see Kruskal's formula, but for $p > 3$, the maximum rank is attained with probability zero, because all rank values above $p + 1$ will have probability zero. For instance, the $6 \times 5 \times 2$ array has maximum rank 8, and typical rank 6; the $10 \times 8 \times 2$

array has maximum rank 13 and typical rank 10. Having dealt with cases where $p > q$, it remains to consider the cases with $p = q$.

5. The typical rank of $p \times p \times 2$ arrays

Above, we have used the transformation to simplicity as a means to arrive at the typical rank when $p > q$. For $p \times p \times 2$ arrays, the transformation to simplicity is not generally available. However, we can settle the issue of typical rank for the latter arrays at once, by using Result 5 in a different context.

Result 6. *Almost surely, the rank of a $p \times p \times 2$ array is either p or $p + 1$.*

Proof. Every $p \times p \times 2$ array can be embedded in a $(p + 1) \times p \times 2$ array, by adding one random row to $\mathbf{X}_1$ and another to $\mathbf{X}_2$. The resulting array is of order $(p + 1) \times p \times 2$, and has therefore almost surely rank $p + 1$. Because embedding cannot decrease the rank, the rank of a $p \times p \times 2$ array is at most $p + 1$, almost surely. Also, the $p \times p \times 2$ array has at least rank p almost surely, and it has rank p with positive probability [20]. It follows that, almost surely, its rank is either p or $p + 1$, so its typical rank is $\{p, p + 1\}$. □

To appreciate the implications of Result 6, one may consider, for example, the case $p = q = 10$. The maximal rank is 15, yet the typical rank is {10,11}. Arrays of rank higher than 11 can be contrived, but arise with probability zero.

6. Discussion

The proof of Result 4 has been derived from the transformation to simplicity, followed by premultiplication by the Vandermonde matrix. However, the row-wise proportionality of transformed versions of $\mathbf{X}_1$ and $\mathbf{X}_2$, essential for the proof of Result 4, could also have been obtained from premultiplication of $\mathbf{X}$ by a $p \times p$ matrix $\mathbf{G}$, with rows $\mathbf{g}'_1, \ldots, \mathbf{g}'_p$ that can be found from solving the equations $\mathbf{g}'_i(i\mathbf{X}_1 - \mathbf{X}_2) = 0', i = 1, \ldots, p$. Along these lines, an alternative proof for the typical rank of $p \times q \times 2$ arrays could have been developed. However, we have used the transformation to simplicity, because simplicity results are interesting in their own right.

The simplicity result of the present paper implies that a $p \times q \times 2$ array $(p > q)$ can usually be simplified to have at least $2pq - 2q$ zero elements. Accordingly, at least $100 \times (p - 1)/p$ % of the elements can artificially be set to zero. This implies that a $p \times q \times 2$ core array in three-mode principal component analysis, when $p > q \geqslant 2$, can be greatly simplified. Also, we have here another demonstration, in addition to the one given by Murakami et al. [18], of the high degree of overparameterization of three-mode PCA.

Starting from simplicity, results for the typical rank of $p \times q \times 2$ arrays have been obtained, including the case where $p = q$. However, we have not discussed the issue of simplicity for the latter type of arrays. Although transformations to simplicity for these arrays do seem possible, their relation to three-way rank is far from clear.

As we have noted before, the transformation to simplicity (1) does not involve a mixing of the slices $\mathbf{X}_1$and $\mathbf{X}_2$. Conceivably, this third direction of transformation will be needed to obtain simplicity for $p \times q \times r$ arrays with $p \geqslant q \geqslant r \geqslant 3$. Computer simulations indicate that, for such arrays, transformation to simplicity is not possible in general. The only exceptions encountered so far are the case $p = q = r = 3$, where 18 of the 27 elements can usually be transformed to be zero (also see [12]), and the cases where $p = qr - 1$, see [18].

A special point of interest, raised by an anonymous referee, is the comparison of typical rank of $p \times q \times 2$ arrays in the real setting, treated in the paper, with that in the complex setting. Specifically, Bürgisser et al. [2, Exercise 20.4] report that, when $p > q$, the typical tensorial rank is $\min[p,2q]$, precisely as it is in the real setting; when, however, $p = q$, the typical tensorial rank is p in the complex setting whereas it is $\{p,p+1\}$ in the real setting.

Appendix A. A MATLAB program for simplicity as in (1)

```
% input: X = [X1 X2], of order p×2q, and p > q.
% output: S, T, producing Y = [S′* X1*T S′*X2*T]
[p,qq] = size(X);q = qq/2;t = p–q;X1 = X(:,1:q);X2 = X(:,q+1:qq);
O = zeroes(p*p,t*q);W1 = [kron(eye(p),X2) O];W2 = [O kron(eye(p),X1)];
W = [W1-W2]; v = [];E = eye(p);
for i = 1:p
v = [v;E(:,i)];
end
s = (eye(p*p)-W*pinv(W′*W)*W′)*v;S = zeroes(p,p);
for i = 1:p
S(:,i) = s(1+(i-1)*p:i*p);
end
H = S'*X1;T = inv(H(1:q,:));Y = S'*[X1*T X2*T]
```

References

[1] R.A Bailey, C.A. Rowley, Maximal rank of an element of a tensor product, Linear Algebra and Appl. 182 (1993) 1–7.

[2] P. Bürgisser, M. Clausen, M.A. Shokrollahi, Algebraic Complexity Theory (Grundlehren der mathematischen Wissenschaften 315), Springer, Berlin, 1997.

[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 (1970) 283–319.

[4] F.M. Fisher, The Identification Problem in Econometrics, McGraw-Hill, New York, 1966.

[5] A. Franc, Etudes algebriques des multitableaux: Apports de l'algebre tensorielle, Ph.D. Thesis, University of Montpellier II, 1992.

[6] R.A. Harshman, Foundations of the PARAFAC procedure: models and conditions for an explanatory multi-modal factor analysis. University of California at Los Angeles, UCLA Working Papers in Phonetics 16 (1970) 1–84.

[7] J. Ja'Ja', Optimal evaluation of a pair of bilinear forms, SIAM J. on Computing 8 (3) (1979) 281–293.

[8] H.A.L. Kiers, Tuckals core rotations and constrained TUCKALS modelling, Statistica Applicata 4 (1992) 659–667.

[9] H.A.L. Kiers, Three-mode orthomax rotation, Psychometrika 62 (1997) 579–598.

[10] H.A.L. Kiers, 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, Springer, Tokyo, 1998, pp. 563–574.

[11] H.A.L. Kiers, Three-way SIMPLIMAX for oblique rotation of the three-mode factor analysis core to simple structure, Computational Statistics and Data Analysis 28 (1998) 307–324.

[12] H.A.L. Kiers, J.M.F. Ten Berge, R. Rocci, Uniqueness of three-mode factor models with sparse cores: The $3 \times 3 \times 3$ case, Psychometrika 62 (1997) 349–374.

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

[14] J.B. Kruskal, Three-way arrays: Rank and uniqueness of trilinear decompositions with applications to arithmetic complexity and statistics, Linear Algebra and Appl. 18 (1977) 95–138.

[15] J.B. Kruskal, Statement of some current results about three-way arrays, unpublished report, 1983.

[16] J.B. Kruskal, Rank, decomposition, and uniqueness for 3-way and N-way arrays. In: R. Coppi and S. Bolasco (Eds.), Multiway Data Analysis, North-Holland, Amsterdam, 1989.

[17] Mathworks, 1994. Matlab. Natick (Mass.), Mathworks.

[18] T. Murakami, J.M.F. Ten Berge, H.A.L. Kiers, A case of extreme simplicity of the core matrix in three-mode principal components analysis, Psychometrika 63 (1998) 255–261.

[19] R. Rocci, Three-mode factor analysis with binary core and orthonormality constraints, J. Italian Stat. Society 2 (1992) 413–422.

[20] J.M.F. Ten Berge, Kruskal's polynomial for $2 \times 2 \times 2$ arrays and a generalization to $2 \times n \times n$ arrays, Psychometrika 56 (1991) 631–636.

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

[22] J. Von zur Gathen, Maximal bilinear complexity and codes, Linear Algebra and Appl. 144 (1991) 49–61.