ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Linear Algebra and its Applications xx (2004)xxx–xxx

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Typical rank and indscal dimensionality for symmetric three-way arrays of order $I \times 2 \times 2$ or $I \times 3 \times 3$

Jos M.F. ten Berge [a,*], Nikolaos D. Sidiropoulos [b,c], Roberto Rocci [d]

[a] *Department of Psychology, Heijmans Instituut, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands*

[b] *University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, USA*

[c] *Tech. Univ. of Crete, Chania 73100, Greece*

[d] *Department of SEFeMEQ, University of "Tor Vergata", Via Columbia 2, 00133 Roma, Italy*

Received 18 February 2002; accepted 2 March 2004

Submitted by G.P.H. Styan

Abstract

A peculiar property of three-way arrays is that the rank they typically have does not necessarily coincide with the maximum possible rank, given their order. Typical tensorial rank has much been studied over algebraically closed fields. However, very few results have been found pertaining to the typical rank of three-way arrays over the real field. These results refer to arrays sampled randomly from continuous distributions. Arrays that consist of symmetric slices do not fit into this sampling scheme. The present paper offers typical rank results (over the real field) for arrays, containing symmetric slices of order 2×2 and 3×3. Symmetric arrays often appear to have lower typical ranks than their asymmetric counterparts. This paper also examines whether or not the rank of a symmetric array coincides with the smallest number of dimensions that allow a perfect fit of INDSCAL. For all cases considered, this is indeed true. Thus, a full INDSCAL solution may require fewer dimensions for a symmetric array than a full CP decomposition applied to an asymmetric array of the same size. The reverse situation does not seem to arise. Next, we examine in which cases CP solutions inevitably are INDSCAL solutions. Finally, the rank-reducing impact of double standardizing the slices is discussed.

* Corresponding author. Tel.: +31-50-3636349; fax: +31-50-3636304.

E-mail addresses: j.m.f.ten.berge@ppsw.rug.nl (J.M.F. ten Berge), nikos@ece.umn.edu, nikos@telecom.tuc.gr (N.D. Sidiropoulos), roberto.rocci@uniroma2.it (R. Rocci).

doi:10.1016/j.laa.2004.03.009

Keywords: Three-way rank; Typical tensorial rank; Three-mode component analysis; Indscal; Candecomp; Parafac

The rank of a matrix $\mathbf{X}$ is defined as the smallest number of rank-one matrices (outer products of two vectors) that generate $\mathbf{X}$ as their sum. Equivalently, the rank of $\mathbf{X}$ is the smallest number of components that give a perfect fit in Principal Component Analysis. That is, when $\mathbf{X}$ can be decomposed as $\mathbf{X} = \mathbf{AB}'$, for matrices $\mathbf{A}$ and $\mathbf{B}$ with R columns, and when no such decomposition exists with less than R columns, the rank of $\mathbf{X}$ is R.

Similarly, the rank of a three-way array (over the real field) is defined as the smallest number of (real valued) rank-one three-way arrays (outer products of three vectors) that generate the array as their sum [5,6]. The rank of a three-way array is also the smallest number of components that give a perfect fit in CANDECOMP/PARAFAC [2,3]. Specifically, let the three-way array $\underline{\mathbf{X}}$ of order $I \times J \times K$ be composed of I slices $\mathbf{X}_1, \ldots, \mathbf{X}_I$, of order $J \times K$. Then a perfect fit in CANDECOMP/PARAFAC implies that there exist matrices $\mathbf{A}$ $(J \times R)$, $\mathbf{B}$ $(K \times R)$ and diagonal matrices $\mathbf{C}_1, \ldots, \mathbf{C}_I$ of order $R \times R$ such that, for $i = 1, \ldots, I$,

$$\mathbf{X}_i = \mathbf{AC}_i\mathbf{B}'. \tag{1}$$

The smallest value of R for which (1) can be solved is the (three-way) rank of the array $\underline{\mathbf{X}}$. It is well-known that non-singular transformations of the array in any direction do not affect the rank of that array.

A peculiarity of the rank of a three-way array is the distinction between the maximal rank an array may have, and its typical rank, that is, the rank it has with positive probability. For instance, a 3×2 matrix has rank 2 with probability 1, and 2 is also the maximal rank of such a matrix, but a $2 \times 4 \times 4$ array typically has rank 4 or 5, yet such arrays have maximal rank 6. That rank, however, will never be observed when the elements of the array are sampled randomly from a continuous distribution, also see Ten Berge [10].

Recently, some advances have been made in the study of the typical rank of three-way arrays over the real field [10,12]. The results rely on the assumption of random sampling of the entire array from a continuous distribution. Symmetric arrays, consisting of symmetric slices, do not fit into this sampling scheme. The present paper offers a first exploration of the typical rank of symmetric arrays. As has been mentioned above, the present paper is exclusively concerned with rank over the real field. Results for typical tensorial rank over algebraically closed fields can be found in [1].

As was mentioned above, the rank of a three-way array is equivalent to the smallest number of components that allows a perfect fit in CANDECOMP/PARAFAC (CP). For instance, the fact that a $7 \times 3 \times 3$ array has typical rank 7 [10] implies that the seven slices $\mathbf{X}_1, \ldots, \mathbf{X}_7$, of order 3×3, can almost surely be decomposed by CP as $\mathbf{X}_i = \mathbf{AC}_i\mathbf{B}'$, for a 3×7 matrix $\mathbf{A}$, a 3×7 matrix $\mathbf{B}$, and seven diagonal 7×7 matrices $\mathbf{C}_i$, $i = 1, \ldots, 7$, and that such a decomposition fails almost surely with less than seven components. Hence, results on typical rank imply the smallest number of components that will usually be enough for a full CP decomposition. Conversely, obtaining perfect fit with CP for a given value of R can be used to generate hypotheses about the typical rank.

When typical rank results for symmetric arrays (viz. three-way arrays with all slices symmetric) are considered, the connection to CP is less interesting, because such arrays are usually decomposed subject to the constraint that $\mathbf{A} = \mathbf{B}$, to satisfy the so-called INDSCAL model [2]. Therefore, the (typical) INDSCAL dimensionality would be of greater interest than the (typical) rank itself. Accordingly, the present paper is not confined to the rank of symmetric arrays, but also deals with the smallest number of dimensions that suffices for a perfect fit in INDSCAL. That number will be referred to as "dim". When, for instance, it is said that a symmetric $2 \times 3 \times 3$ array has "dim 4", there exists a 3×4 matrix $\mathbf{A}$ and diagonal 4×4 "salience matrices" $\mathbf{C}_1$ and $\mathbf{C}_2$ such that the two 3×3 slices $\mathbf{S}_1$ and $\mathbf{S}_2$ can be written as $\mathbf{S}_1 = \mathbf{A}\mathbf{C}_1\mathbf{A}'$, $\mathbf{S}_2 = \mathbf{A}\mathbf{C}_2\mathbf{A}'$, and such a decomposition fails in less than four dimensions.

Furthermore, the very relationship between the rank and the dim of a symmetric three-way array is of interest. Although it may be conjectured that they are the same, there is no formal proof for this, nor does a counterexample seem to exist. It is obvious that the rank can never exceed dim, because dim pertains to constrained CP fitting. But the reverse is in no way guaranteed. Therefore, the search for cases were dim might exceed rank has been one of the objectives of the present research. A variety of results is offered, showing that rank = dim in all cases considered. Nevertheless, this does not settle the issue. It has not been proven that rank and dim coincide for every symmetric array.

The organization of this paper is as follows. First, we start with two general results. Then, we deal with the typical rank of symmetric arrays, consisting of I symmetric $J \times J$ matrices $\mathbf{S}_1, \ldots, \mathbf{S}_I$, where $J = 2$ or 3. Next, to relate the rank values that arise with positive probability to dim values for the same arrays, the existence of CP solutions with $\mathbf{A}$ and $\mathbf{B}$ equal will be examined. In all cases considered, the existence of such a solution will be established, implying that dim equals rank for these cases. Next, we compare typical dim values for symmetric arrays to typical rank values of asymmetric arrays of the same size. Also, we determine when CP solutions will inevitably have $\mathbf{A}$ and $\mathbf{B}$ equal. Finally, the rank-reducing impact of double standardizing the slices will be discussed.

Result 1. *Symmetric* $2 \times J \times J$ *arrays have typical rank* $\{J, J+1\}$.

Proof. For any $2 \times J \times J$ array consisting of two $J \times J$ matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ (symmetric or non-symmetric), a CP solution of rank less than J has probability zero. A rank J solution exists if and only if the eigenvalues of $\mathbf{S}_1^{-1}\mathbf{S}_2$ are real. This happens with probability p, $0 < p < 1$ (e.g., [9]). Suppose that the eigenvalues are not real. Construct the matrix $\mathbf{X}$ of order $(J+1) \times 2J$ as $\mathbf{X} = [\mathbf{X}_1|\mathbf{X}_2] = \left[\begin{array}{c|c} \mathbf{S}_1 & \mathbf{S}_2 \\ \mathbf{x}' & \mathbf{y}' \end{array}\right]$, for random vectors $\mathbf{x}$ and $\mathbf{y}$. We shall now construct a rank $J+1$ decomposition $\mathbf{X}_i = \mathbf{AC}_i\mathbf{B}'$, $i = 1, 2$, as prescribed in (1).

To find a decomposition for $\mathbf{X}$, we need to solve $\mathbf{A}^{-1}\mathbf{X}_i = \mathbf{C}_i\mathbf{B}'$, $i = 1, 2$. Define $\mathbf{a}'_j$ as row j of $\mathbf{A}^{-1}$. These rows must satisfy the proportionality equation $\lambda_j\mathbf{a}'_j\mathbf{X}_1 = \mathbf{a}'_j\mathbf{X}_2$ for $J+1$ different values of λ_j. Because there obviously is a vector orthogonal to the columns of the $(J+1)\times J$ matrix $\lambda_j\mathbf{X}_1 - \mathbf{X}_2$, for every λ_j, and a set of $J+1$ such vectors will be linearly independent almost surely [10], a non-singular $\mathbf{A}^{-1}$ can be found almost surely. This proves that a rank $J+1$ solution is possible for $\mathbf{X}_1$ and $\mathbf{X}_2$, almost surely. Because adding a slice to an array cannot reduce the rank, the original array with slices $\mathbf{S}_1$ and $\mathbf{S}_2$ had rank $J+1$ at most, almost surely. □

Result 2. *When I is large compared to J, in the sense that $I \geqslant 0.5J(J+1)$, then symmetric $I\times J\times J$ arrays almost surely have $R = \dim = 0.5J(J+1)$, which is also the maximum rank.*

Proof. From Rocci and Ten Berge [8], we may adopt the explicit basis of $0.5J(J+1)$ rank one matrices which spans the space of symmetric $J\times J$ matrices. This shows that $R = 0.5J(J+1)$ is high enough. To prove that a lower rank is not enough almost surely, consider the I slices, strung out column-wise in a $J^2 \times I$ matrix $\mathbf{X} = [\mathrm{Vec}(\mathbf{S}_1)\ldots\mathrm{Vec}(\mathbf{S}_I)]$, which has $0.5J(J+1)$ unrepeated rows.Then a CP decomposition implies that $\mathbf{X} = \mathbf{HC}'$, where $\mathbf{H} = \mathbf{A}\bullet\mathbf{B} = [\mathbf{a}_1\otimes\mathbf{b}_1,\ldots,\mathbf{a}_R\otimes\mathbf{b}_R]$, $\bullet$ is the Khatri–Rao (column-wise Kronecker) product of the $I\times R$ matrix $\mathbf{A} = [\mathbf{a}_1\ldots\mathbf{a}_R]$ and the $J\times R$ matrix $\mathbf{B} = [\mathbf{b}_1\ldots\mathbf{b}_R]$, $\otimes$ is the Kronecker product, and $\mathbf{C}$ is the $I\times R$ matrix with rows containing the diagonal elements of $\mathbf{C}_1,\ldots,\mathbf{C}_I$. Note that, because of the repeated rows, the typical rank of $\mathbf{X}$ is $\min[I, J(J+1)/2] = J(J+1)/2$. When $R < 0.5J(J+1)$, rank($\mathbf{HC}'$) will be less than rank($\mathbf{X}$), and a solution for $\mathbf{X} = \mathbf{HC}'$ does not exist. □

Results 1 and 2 solve the typical rank issue for all symmetric $I\times 2\times 2$ arrays. When $I = 2$ we use Result 1, and Result 2 takes care of all cases where $I > 2$. For $I\times 3\times 3$ arrays, the situation is more complicated. Result 1 implies that the $2\times 3\times 3$ arrays have typical rank {3, 4}. Result 2 implies that all $I\times 3\times 3$ arrays with $I \geqslant 6$ have typical rank 6. The cases in between require separate treatment. These treatments are given in the next sections.

1. The typical rank of symmetric 3 × 3 × 3 arrays

Result 3. *The symmetric* $3\times 3\times 3$ *array has typical rank* 4.

Proof. It is obvious that a rank less than 3 has probability zero and requires no further treatment. When a rank 3 solution exists, $\mathbf{S}_i = \mathbf{AC}_i\mathbf{B}'$ implies $\mathbf{S}_1\mathbf{S}_2^{-1} = \mathbf{AC}_1\mathbf{C}_2^{-1}\mathbf{A}^{-1}$ and $\mathbf{S}_1\mathbf{S}_3^{-1} = \mathbf{AC}_1\mathbf{C}_3^{-1}\mathbf{A}^{-1}$, which means that the matrices $\mathbf{S}_1\mathbf{S}_2^{-1}$ and $\mathbf{S}_1\mathbf{S}_3^{-1}$

commute. This is also an event of probability zero [15]. Hence, the typical rank is above 3. To show that the typical rank is 4, we shall construct a rank 4 CP decomposition

$$\mathbf{S}_i = \mathbf{A}\mathbf{C}_i\mathbf{B}', \tag{2}$$

$i = 1, 2, 3$, where $\mathbf{A}$ and $\mathbf{B}$ are 3×4 matrices, $\mathbf{C}_i$ is diagonal.

To simplify this problem, we may mix (take linear combinations of) the slices to have each of them of rank 2. This is easy because a 3×3 real matrix always has at least one real eigenvalue. So if μ is a real eigenvalue of $\mathbf{S}_1^{-1}\mathbf{S}_2$, then $\mathbf{S}_2 - \mu\mathbf{S}_1$ will be of rank 2 almost surely. Let $\mathbf{S}_4, \mathbf{S}_5$ and $\mathbf{S}_6$ be symmetric rank-two matrices, obtained by taking linear combinations of $\mathbf{S}_1, \mathbf{S}_2$, and $\mathbf{S}_3$. Define $\mathbf{W}$ as the 3×3 matrix containing the right null of $\mathbf{S}_4$, $\mathbf{S}_5$, and $\mathbf{S}_6$, and replace $\mathbf{S}_i$ by $\mathbf{W}'\mathbf{S}_i\ \mathbf{W}$, $i = 4, 5, 6$. This method of simplification has been proposed by Rocci [7] in a different context. Note that symmetry is preserved by premultiplication by $\mathbf{W}'$ and postmultiplication by $\mathbf{W}$, and also by slab-mixing. The net result is that, without loss of generality, we may start from previously simplified symmetric matrices

$$\mathbf{S}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & a \\ 0 & a & b \end{bmatrix}, \quad \mathbf{S}_2 = \begin{bmatrix} 1 & 0 & c \\ 0 & 0 & 0 \\ c & 0 & d \end{bmatrix}, \quad \mathbf{S}_3 = \begin{bmatrix} 1 & e & 0 \\ e & f & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{3}$$

Upon dividing rows 3 and columns 3 by a/e, $\mathbf{S}_1$ and $\mathbf{S}_3$ have the same nonzero off-diagonal element. It follows that we may take $a = e$. We shall now create a basis of four symmetric rank-one matrices, which generates $\mathbf{S}_1, \mathbf{S}_2$, and $\mathbf{S}_3$ as linear combinations. Specifically, let

$$\begin{aligned} &\mathbf{K} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & a^2/b & a \\ 0 & a & b \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ &\mathbf{M} = \begin{bmatrix} 1 & a & 0 \\ a & a^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{N} = \begin{bmatrix} 1 & \lambda & \mu \\ \lambda & \lambda^2 & \lambda\mu \\ \mu & \lambda\mu & \mu^2 \end{bmatrix}. \end{aligned} \tag{4}$$

Clearly, $\mathbf{S}_1$ is a linear combination of $\mathbf{K}$ and $\mathbf{L}$, and $\mathbf{S}_3$ is a linear combination of $\mathbf{L}$ and $\mathbf{M}$. It remains to verify that $\mathbf{S}_2$ is a linear combination of all four matrices. Let $\mathbf{N}$ be determined by

$$\lambda = \frac{a(c^2 - d)}{c(b + c)} \quad \text{and} \quad \mu = \frac{bc + d}{b + c}. \tag{5}$$

Then it can be verified that $\mathbf{S}_2 = \alpha\mathbf{K} + \beta\mathbf{L} + \gamma\mathbf{M} + \delta\mathbf{N}$, with $\alpha = (d - c\mu)/b$, $\delta = c/\mu$, $\gamma = 1 - \delta$, and $\beta = -\alpha a^2/b - \gamma a^2 - \delta\lambda^2$. $\square$

2. The typical rank of symmetric 4 × 3 × 3 arrays

Result 4. *The symmetric* $4 \times 3 \times 3$ *array has typical rank* $\{4, 5\}$.

Proof. We try to construct a rank four solution and verify when that is (im)possible. Let the 9×4 matrix $\mathbf{X}$ contain the vecs of $\mathbf{S}_1, \ldots, \mathbf{S}_4$, let $\mathbf{X}_4$ contain rows 1, 2, 3, and 5 of $\mathbf{X}$, and let $\mathbf{G} = \mathbf{X}\mathbf{X}_4^{-1}$. A rank 4 solution exists if and only if there exists a rank 4 Khatri–Rao basis $\mathbf{A} \bullet \mathbf{B}$ which generates $\mathbf{X}$. Therefore, we need to solve $\mathbf{X} = (\mathbf{A} \bullet \mathbf{B})\mathbf{C}'$ by finding $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. Equivalently, we may solve $\mathbf{GW} = \mathbf{A} \bullet \mathbf{B}$, with $\mathbf{W} = \mathbf{X}_4(\mathbf{C}')^{-1}$. Then the problem is to find four linearly independent solutions for a vector $\mathbf{w}$, such that $\mathbf{Gw}$ is the Kronecker product of two vectors, which may be rescaled to be $[1\ b\ c]'$ and $[1\ d\ e]'$ respectively, for scalars b, c, d, and e. It can be verified that

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ f_1 & f_2 & f_3 & f_4 \\ 0 & 0 & 1 & 0 \\ f_1 & f_2 & f_3 & f_4 \\ g_1 & g_2 & g_3 & g_4 \end{bmatrix}, \quad \text{and} \quad \mathbf{Gw} = \begin{bmatrix} 1 \\ d \\ e \\ b \\ bd \\ be \\ c \\ cd \\ ce \end{bmatrix} \tag{6}$$

for fixed vectors $\mathbf{f} = [f_1\ f_2\ f_3\ f_4]'$ and $\mathbf{g} = [g_1\ g_2\ g_3\ g_4]'$. From rows 2 and 4 of (6) we have $w_2 = b = d$; and rows 3 and 7 imply that $w_3 = c = e$. We also have $w_1 = 1$ and $w_4 = d^2$ whence $\mathbf{w} = [1\ d\ e\ d^2]'$. The remaining problem is to solve the equations implicit in rows 8 and 9 of (6). This means that we want $f_1 + df_2 + e(f_3 - d) + d^2 f_4 = 0$ and $g_1 + dg_2 + e(g_3 - e) + d^2 g_4 = 0$. Writing $e = (f_1 + df_2 + d^2 f_4)/(d - f_3)$ and substituting this for e into the latter equation yields the fourth degree polynomial equation

$$\begin{aligned} &d^4(g_4 - f_4^2) + d^3(-2f_2f_4 + f_4g_3 - 2f_3g_4 + g_2) \\ &\quad + d^2(-f_2^2 - 2f_1f_4 + f_2g_3 - f_3f_4g_3 + g_1 - 2f_3g_2 + f_3^2g_4) \\ &\quad + d(-2f_1f_2 + f_1g_3 - f_2f_3g_3 - 2f_3g_1 + f_3^2g_2) \\ &\quad + g_1f_3^2 - f_1^2 - f_1f_3g_3 = 0. \end{aligned} \tag{7}$$

It can easily be verified that there are cases where (7) has four real distinct roots. For instance, when $\mathbf{f} = [2,\ 1,\ 1,\ -2]'$ and $\mathbf{g} = [1,\ 1,\ 1,\ 1]'$, the roots of $[-3,\ 1,\ 10,\ -4,\ -5]$ are real and distinct. Because the coefficients of (7) depend continuously on the data in $\mathbf{X}$, the existence of four distinct real roots, none of which are zero, has positive probability. In that case, $\mathbf{W}$ can be proven to be non-singular. Specifically, column i of $\mathbf{W}$ has the form $[1\ d_i\ e_i\ d_i^2]'$ $i = 1, 2, 3, 4$. Suppose we have $f_1 = 1$ and $f_3 = 0$. Then $e_i = (1 + d_i f_2 + d_i^2 f_4)/d_i$. Multiplying $\mathbf{W}$ by $\mathbf{D} = \mathrm{diag}(d_1,\ d_2,$

$d_3,\ d_4$) yields a matrix with columns of the form $[d_i\ d_i^2(1+d_if_2+d_i^2f_4)d_i^3]'$ Subtracting f_2 times row 1 plus f_4 times row 2 from $\mathbf{WD}$ yields another matrix with elements d_i, d_i^2, 1 and d_i^3 in each column. This is a Vandermonde matrix, known to be of rank four when d_1, d_2, d_3 and d_4 are distinct. Because $\mathbf{W}$ is non-singular for at least one case (when $f_1 = 1$ and $f_3 = 0$, that is), it is non-singular almost surely. Hence, the symmetric $4 \times 3 \times 3$ array has rank 4 with positive probability.

When some of the roots are imaginary, the rank exceeds 4. A solution of rank 5 can always be constructed. Specifically, let $\mathbf{X}$ be the 9×5 matrix containing the vecs of the four symmetric slices, the fifth column being the vector $[0\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ x]'$ with x "large enough" (details of what that means will soon follow). Define $\mathbf{X}_5$ as the 5×5 matrix containing rows 1, 2, 3, 5, and 6 of $\mathbf{X}$. Note that $\mathbf{X}_5$ is independent of x. Define $\mathbf{X}_4$ as the upper left 4×4 submatrix of $\mathbf{X}_5$. So $\mathbf{X}_5 = \begin{bmatrix} \mathbf{X}_4 & 0 \\ \mathbf{b}' & 1 \end{bmatrix}$, for some vector $\mathbf{b}$, with inverse $\mathbf{X}_5^{-1} = \begin{bmatrix} \mathbf{X}_4^{-1} & 0 \\ \mathbf{c}' & 1 \end{bmatrix}$, where row 5 is obtained as the vector orthogonal to the first four columns of $\mathbf{X}_5$, rescaled to have its fifth element 1. Define $\mathbf{G} = \mathbf{X}\mathbf{X}_5^{-1}$. We shall again construct a Khatri–Rao basis for the column space of $\mathbf{G}$. That is, we look for a basis of Kronecker products of vectors $[1\ b_i\ c_i]'$ and $[1\ d_i\ e_i]'$, $i = 1, \ldots, 5$. Ignoring the subscripts, we want a vector $\mathbf{w}$ such that $\mathbf{Gw}$ has Kronecker product form. It can be verified that $b_i = d_i$ and $c_i = e_i$ whence we have

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ g_1 & g_2 & g_3 & g_4 & x \end{bmatrix}, \quad \mathbf{Gw} = \begin{bmatrix} 1 \\ d \\ e \\ d \\ d^2 \\ de \\ e \\ de \\ e^2 \end{bmatrix}, \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 1 \\ d \\ e \\ d^2 \\ de \end{bmatrix}. \tag{8}$$

Clearly, the only problem that remains is solving the quadratic equation

$$g_1 + dg_2 + eg_3 + d^2g_4 + xde - e^2 = 0$$

for d and e, with x to be fixed in advance. Let $\mathbf{g} = [g_1\ g_2\ g_3\ g_4]'$ Note that $\mathbf{g} = \mathbf{f} + x\mathbf{c}$, where $\mathbf{f}'$ is row 9 of $\mathbf{X}$ without its fifth element x, postmultiplied by $\mathbf{X}_4^{-1}$. This shows that $g_1 \ldots, g_4$ do depend on x. To take that into account, we need to solve the quadratic equation

$$\begin{aligned} &-e^2 + e(f_3 + xc_3 + xd) \\ &\quad + (f_1 + xc_1 + df_2 + dxc_2 + d^2f_3 + d^2xc_3) = 0. \end{aligned} \tag{9}$$

Because we need a five dimensional Kronecker basis, we need to fix x in such a way that we can find five values for d where (9) has a real solution for e. This requires that the discriminant

$$(f_3 + xc_3 + xd)^2 + 4(f_1 + xc_1 + df_2 + dxc_2 + d^2 f_3 + d^2 xc_3) \tag{10}$$

be positive. Clearly, this discriminant will be positive when x is large enough. The solution can be obtained as follows:

1. Pick an arbitrary value for x.
2. Pick 5 different values of d, for instance $\{0.2, 0.4, 0.6, 0.8, 1.0\}$.
3. For each d, solve the quadratic (9). If any of the roots is imaginary, return to step 2 with a larger value of x.
4. From each pair $\{d, e\}$, construct $[1\ d\ e]'$ defining the five columns of a matrix $\mathbf{A}$. Also construct $\mathbf{w}$ by (8). This yields the columns of $\mathbf{W} = \mathbf{X}_4(\mathbf{C}')^{-1}$, such that $\mathbf{GW} = (\mathbf{A} \bullet \mathbf{A})$.

It can be concluded that $4 \times 3 \times 3$ arrays have typical rank $\{4, 5\}$. $\square$

3. The typical rank of symmetric 5 × 3 × 3 arrays

To find the typical rank of symmetric $5 \times 3 \times 3$ arrays, we start by assuming that a rank 5 solution is possible, and then construct that solution explicitly. In the process, it will become clear that the rank 5 solution does not always exist, and it will be shown that a rank 6 solution is always possible.

Result 5. *The symmetric* $5 \times 3 \times 3$ *array has typical rank* $\{5, 6\}$.

Proof. Let $\mathbf{X}$ be of order 9×5, with columns containing the vecs of the symmetric slices $\mathbf{S}_1, \ldots, \mathbf{S}_5$. Let $\mathbf{X}_5$ contain the rows 1, 2, 3, 5, and 6 of $\mathbf{X}$ and define $\mathbf{G} = \mathbf{X}\mathbf{X}_5^{-1}$. Note that $\mathbf{X}_5$ is non-singular almost surely. Proceeding as in the proof of Result 4, we now need five solutions for the equation $\mathbf{Gw} = \mathbf{g} \otimes \mathbf{h}$, with $\mathbf{g} = [1\ b\ c]'$, $\mathbf{h} = [1\ d\ e]'$. Note that

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ f_1 & f_2 & f_3 & f_4 & f_5 \end{bmatrix} \quad \text{and} \quad \mathbf{Gw} = \begin{bmatrix} 1 \\ d \\ e \\ b \\ bd \\ be \\ c \\ cd \\ ce \end{bmatrix}. \tag{11}$$

Clearly, rows $1, \ldots, 8$ imply $b = d$, $c = e$, and $\mathbf{w} = [1\ d\ e\ d^2\ de]'$. What remains is to solve the equation $\mathbf{f}'\mathbf{w} = e^2$, where $\mathbf{f}'$ is row 9 of $\mathbf{G}$. Therefore, we need a solution for the quadratic equation $-e^2 + e(f_3 + df_5) + (f_1 + df_2 + d^2 f_4) = 0$. Clearly, we can find e (two solutions) if and only if d has been chosen to render the discriminant

Table 1
Typical rank values for symmetric arrays

	2×2 Slices	3×3 Slices		$J \times J$ Slices
$I = 2$	{2, 3}	{3, 4}	. . .	$\{J, J + 1\}$
$I = 3$	3	4		
$I = 4$	3	{4, 5}		
$I = 5$	3	{5, 6}		
$I \geqslant 0.5J(J + 1)$	3	6	. . .	$.5J(J + 1)$

$(f_3 + df_5)^2 + 4(f_1 + df_2 + d^2 f_4)$ positive. Whenever this is possible, we solve for e given d (taking either of two roots) and find the vector $\mathbf{w} = [1\ d\ e\ d^2\ de]'$. We repeat this procedure five times, to create five different linearly independent solutions for $\mathbf{w}$ from five different choices of d.

There are cases where this method fails. For instance, when $\mathbf{f} = [-1, 2, 1, -2, -1]'$ or close to this vector, the discriminant will be (close to) $(1 - d)^2 - 4(1 - d)^2 - 4d^2 < 0$, and there is no solution. This means that, for a set of arrays with positive volume, we shall need at least a rank 6 solution. Result 2 guarantees that such a solution exists. □

Numerical experience indicates that the probability of rank 5, under random sampling from the uniform $[-0.5, 0.5]$ distribution, is very close to 1. Still, when $\mathbf{f}$ is $[-1, 2, 1, -2, -1]'$ or close to that vector, we do indeed find arrays of rank 6. This illustrates that the typical rank is indeed {5, 6}, as has been proven above.

Result 5 completes our treatment of typical rank for symmetric $I \times 3 \times 3$ arrays. Table 1 summarizes the typical ranks for $I \times 2 \times 2$ and $I \times 3 \times 3$ arrays of symmetric slices.

It is interesting to compare the values of Table 1 to their counterparts for asymmetric arrays of the same size [10]. There is no difference when $I = 2$, or $I = 3$ and $J = 2$. However, when $I > 3$, symmetric $I \times 2 \times 2$ arrays have typical rank 3, whereas their asymmetric counterparts have typical rank 4. When $I = 6, 7, 8$, or 9, $I \times 3 \times 3$ arrays have typical ranks 6 in case of symmetry, but the typical rank is I otherwise [10]. It seems that, whatever evidence is available, points to the conclusion that symmetry entails the same or lower typical rank. We shall now address the question of how many dimensions are typically needed for a full INDSCAL decomposition.

4. From typical rank to INDSCAL dimensionality

By themselves, the typical rank values of Table 1 do not imply anything about dim (the INDSCAL dimensionality), other than that they are lower bounds to the typical values of dim. INDSCAL solutions require that $\mathbf{A} = \mathbf{BD}$, for a diagonal matrix $\mathbf{D}$, and that constraint might increase the number of components required. The second part of this paper is concerned with the question to what extent the typical rank values of Table 1 are also typical dim values. Specifically, the question is whether or not CP solutions, using the typical rank as number of dimensions, display $\mathbf{A} = \mathbf{BD}$ at once, or, in case of multiple solutions, whether or not at least one solution exists that has $\mathbf{A} = \mathbf{BD}$, as is required for INDSCAL. We start with a generalization of a result by Ten Berge and Kiers [11, Result 2], based on uniqueness.

Definition. A CP solution for an array $\underline{\mathbf{X}}$ is said to be *unique* when every pair of solutions $\underline{\mathbf{X}} = (\mathbf{A} \bullet \mathbf{B})\mathbf{C}'$ and $\underline{\mathbf{X}} = (\overline{\mathbf{A}} \bullet \overline{\mathbf{B}})\overline{\mathbf{C}}'$, all matrices having R columns, are related by $\overline{\mathbf{A}} = \mathbf{APT}_a$, $\overline{\mathbf{B}} = \mathbf{BPT}_b$, and $\overline{\mathbf{C}} = \mathbf{CPT}_c$, for a permutation matrix $\mathbf{P}$ and diagonal matrices $\mathbf{T}_a$, $\mathbf{T}_b$ and $\mathbf{T}_c$, with $\mathbf{T}_a\mathbf{T}_b\mathbf{T}_c = \mathbf{I}$.

Result 6. *When a CP solution for a symmetric array is unique, it is an INDSCAL solution.*

Proof. When a given CP solution $\mathbf{X}_i = \mathbf{AC}_i\mathbf{B}' = \mathbf{X}_\mathbf{i}' = \mathbf{BC}_i\mathbf{A}'$, $i = 1, \ldots, I$, is unique, we have $\mathbf{B} = \mathbf{APT}_a$, $\mathbf{A} = \mathbf{BPT}_b$, and $\mathbf{C} = \mathbf{CPT}_c$, for a permutation matrix $\mathbf{P}$ and diagonal matrices $\mathbf{T}_a$, $\mathbf{T}_b$, and $\mathbf{T}_c$. Because $\mathbf{C}$ cannot have two proportional columns (that would contradict uniqueness at once), it follows from $\mathbf{C} = \mathbf{CPT}_c$ that $\mathbf{P}$ is the identity matrix. This implies that $\mathbf{A}$ and $\mathbf{B}$ are column-wise proportional. □

Below, we shall invoke Result 6 to prove that dim = rank in all cases where an array consisting of two symmetric $J \times J$ slices has a rank as low as J. When a symmetric $2 \times 2 \times 2$ array has rank 3, or a symmetric $2 \times 3 \times 3$ array has rank 4, we can use:

Result 7. *For symmetric* $2 \times 2 \times 2$ *arrays of rank* 3, *and* $2 \times 3 \times 3$ *arrays of rank* 4, *a CP solution with columns of* $\mathbf{A}$ *proportional to those of* $\mathbf{B}$ *can be constructed with dim = rank.*

Proof. When one of the slices is positive or negative definite, the two slices can be diagonalized simultaneously. This situation can always be achieved by using one additional CP component. For the $2 \times 2 \times 2$ case, when $\mathbf{S}_1\mathbf{k} = \lambda\mathbf{k}$, $\lambda < 0$, add $\mu\mathbf{kk}'$ to $\mathbf{S}_1$, for some $\mu > -\lambda$, and diagonalize the resulting positive definite matrix simultaneously with $\mathbf{S}_2$. For the $2 \times 3 \times 3$ case, do the same when $\mathbf{S}_1$ has only one negative eigenvalue. Otherwise, when $\mathbf{S}_1\mathbf{k} = \lambda\mathbf{k}$, λ being the only *positive* eigenvalue, subtract $\mu\mathbf{kk}'$ from $\mathbf{S}_1$, $\mu > \lambda$, to get a negative definite matrix, and diagonalize that matrix simultaneously with $\mathbf{S}_2$. This means that an explicit INDSCAL solution using one extra component can always be found when the two slices cannot be diagonalized simultaneously. □

We have now dealt with all cases pertaining to $I = 2$. The rest follows at once. First, the $I \times 2 \times 2$ cases with $I > 2$ are covered by Result 2, which gave typical rank and typical dim at once. The same goes for the $I \times 3 \times 3$ arrays with $I > 5$. For the

$3 \times 3 \times 3$ case, our proof has offered a rank 4 CP solution which is also an INDSCAL solution. For the $4 \times 3 \times 3$ case of rank 4, the solution is again necessarily an INDSCAL solution (note that we found $b = d$ and $c = e$). When such arrays have rank 5, the solution we constructed again was an INDSCAL solution. Finally, for $5 \times 3 \times 3$ arrays, the rank 5 case is necessarily one with $\mathbf{A} = \mathbf{B}$, and the rank 6 case is covered by Result 2, again implying $\mathbf{A} = \mathbf{B}$. It can be concluded that typical rank = typical dim in all cases examined.

5. Non-uniqueness of INDSCAL solutions

Above, the typical rank values of Table 1 have been proven to be typical dim values. It can thus be concluded that the INDSCAL constraint $\mathbf{A} = \mathbf{BD}$, with $\mathbf{D}$ a diagonal matrix, does not require additional components when $J = 2$ or 3. On the other hand, we also have seen that symmetry of the slices does seem to entail a lower rank = dim by itself. It can be concluded that, for all cases treated in Table 1, a full INDSCAL decomposition takes as many, or fewer components than does a full CP decomposition of asymmetric arrays of the same size.

We have examined typical rank by constructing closed-form CP solutions. In the process, having $\mathbf{A} = \mathbf{BD}$ sometimes appeared as a bonus, even though it was not imposed. Result 6 implies that this must happen in cases where CP has a unique solution. Sufficient conditions for uniqueness were first formulated by Harshman [4]. Kruskal [5] has proven that it is sufficient for uniqueness, when $R > 1$, that $k_a + k_b + k_c \geqslant 2R + 2$, where k_a is the so-called k-rank of $\mathbf{A}$, etc. [5,6]. Ten Berge and Sidiropoulos [14] have shown that, when R is 2 or 3, this condition is also necessary, and that rank and k-rank coincide in these cases when the condition is met. The condition implies that we usually have CP uniqueness when a $2 \times 2 \times 2$ array has rank 2, a $2 \times 3 \times 3$ has rank 3, and a $4 \times 3 \times 3$ array has rank 4, but in none of the remaining cases in Table 1.

Interestingly, in certain non-unique cases, we have also found that the CP solution is necessarily an INDSCAL solution. For instance, the $5 \times 3 \times 3$ case of rank 5 can only produce INDSCAL solutions. Apparently, there are cases where CP, applied to symmetric arrays, has non-unique solutions, all of which are INDSCAL solutions. The following results give an explanation for this remarkable phenomenon:

Result 8. *When $I \geqslant R + 2 - J$, $\mathbf{A}$, $\mathbf{B}$ have full row rank, and every set of $(R + 2 - J)$ columns of $\mathbf{C}$, containing the diagonal elements of $\mathbf{C}_i$ in its rows, is linearly independent, a CP solution is an INDSCAL solution.*

Proof. Let there be a CP solution $\mathbf{S}_i = \mathbf{AC}_i\mathbf{B}'$, $i = 1, \ldots, I$, in R dimensions. Premultiply $\mathbf{A}$ and $\mathbf{B}$, both of order $J \times R$, with $\mathbf{A}_J^{-1}$, the inverse of the leftmost $J \times J$ submatrix of $\mathbf{A}$, to obtain $\mathbf{A}^\sim = [\mathbf{I}_J|\mathbf{A}_+]$ and $\mathbf{B}^\sim = \mathbf{A}_J^{-1}\mathbf{B}$. Reparameterize $\mathbf{A}_J^{-1}\mathbf{S}_i(\mathbf{A}_J^{-1})'$ to $\mathbf{S}_i$, which is still symmetric, and reparameterize $\mathbf{A}^\sim$ and $\mathbf{B}^\sim$ to $\mathbf{A}$

and **B**, respectively. Let $\mathbf{U}(J^2 \times R)$ and $\mathbf{V}(J^2 \times R)$ contain the Khatri–Rao products $\mathbf{B} \bullet \mathbf{A}$ and $\mathbf{A} \bullet \mathbf{B}$, respectively. Note that $\mathbf{C}'$ is an $R \times I$ matrix. Let $\mathbf{X}(J^2 \times I)$ contain $\mathrm{vec}(\mathbf{S}_1), \ldots, \mathrm{Vec}(\mathbf{S}_I)$. Then $\mathbf{S}_i = \mathbf{A}\mathbf{C}_i\mathbf{B}' = \mathbf{B}\mathbf{C}_i\mathbf{A}'$, $i = 1, \ldots, I$, is equivalent to $\mathbf{X} = \mathbf{U}\mathbf{C}' = \mathbf{V}\mathbf{C}'$, hence $(\mathbf{U} - \mathbf{V})\mathbf{C}' = \mathbf{O}$. By construction, J rows of $\mathbf{U} - \mathbf{V}$ vanish, and the non-zero rows come in proportional pairs. Every non-zero row of $\mathbf{U} - \mathbf{V}$ contains at least $J - 2$ zeros, due to the fact that the first J columns of **A** form the identity matrix. When the $J - 2$ zeros are removed from any of the non-zero rows, the remaining vector is orthogonal to the columns of a particular $(R - J + 2) \times I$ submatrix of $\mathbf{C}'$. It follows that these remaining vectors vanish also if the corresponding submatrices of $\mathbf{C}'$ are of rank $R - J + 2$. Then all rows of $\mathbf{U} - \mathbf{V}$ vanish, whence $\mathbf{U} = \mathbf{V}$, which implies that $\mathbf{A} = \mathbf{B}\mathbf{D}$ for some diagonal matrix **D**. $\square$

It may be noted that the condition that every set of $R - J + 2$ rows of $\mathbf{C}'$ correspond to a matrix of rank $R - J + 2$ cannot be satisfied unless $I \geqslant R - J + 2$. The requirement that the submatrices have rank $R - J + 2$ means that the k-rank [5,6] of **C** must be at least $R - J + 2$.

An example may be instructive: Let **A** be 3×4, **B** be 3×4 and **C** be 3×4. After the preliminary transformation we have

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & a_{14} \\ 0 & 1 & 0 & a_{24} \\ 0 & 0 & 1 & a_{34} \end{bmatrix}, \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{bmatrix} \tag{12}$$

hence

$$\begin{aligned} \mathbf{U} - \mathbf{V} &= \begin{bmatrix} b_{11} & 0 & 0 & b_{14}a_{14} \\ 0 & b_{12} & 0 & b_{14}a_{24} \\ 0 & 0 & b_{13} & b_{14}a_{34} \\ b_{21} & 0 & 0 & b_{24}a_{14} \\ 0 & b_{22} & 0 & b_{24}a_{24} \\ 0 & 0 & b_{23} & b_{24}a_{34} \\ b_{31} & 0 & 0 & b_{34}a_{14} \\ 0 & b_{32} & 0 & b_{34}a_{24} \\ 0 & 0 & b_{33} & b_{34}a_{34} \end{bmatrix} - \begin{bmatrix} b_{11} & 0 & 0 & a_{14}b_{14} \\ b_{21} & 0 & 0 & a_{14}b_{24} \\ b_{31} & 0 & 0 & a_{14}b_{34} \\ 0 & b_{12} & 0 & a_{24}b_{14} \\ 0 & b_{22} & 0 & a_{24}b_{24} \\ 0 & b_{32} & 0 & a_{24}b_{34} \\ 0 & 0 & b_{13} & a_{34}b_{14} \\ 0 & 0 & b_{23} & a_{34}b_{24} \\ 0 & 0 & b_{33} & a_{34}b_{34} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ -b_{21} & b_{12} & 0 & a_{24}b_{14} - a_{14}b_{24} \\ -b_{31} & 0 & b_{13} & a_{34}b_{14} - a_{14}b_{34} \\ b_{21} & -b_{12} & 0 & a_{14}b_{24} - a_{24}b_{14} \\ 0 & 0 & 0 & 0 \\ 0 & -b_{32} & b_{23} & a_{34}b_{24} - a_{24}b_{34} \\ b_{31} & 0 & -b_{13} & a_{14}b_{34} - a_{34}b_{14} \\ 0 & b_{32} & -b_{23} & a_{24}b_{34} - a_{34}b_{24} \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned} \tag{13}$$

When each 3×3 submatrix of $\mathbf{C}'$ has rank 3, it is clear that all rows of $\mathbf{U} - \mathbf{V}$ vanish, because each row contains at least one zero. Then every column of $\mathbf{A}$ is proportional to the corresponding column in $\mathbf{B}$. Result 8 can be further qualified when $\mathbf{C}$ is square matrix.

Result 9. *When $I = R$, a CP solution is almost surely an INDSCAL solution.*

Proof. It is readily verified that $\mathbf{A}$ and $\mathbf{B}$ are of full row rank almost surely, because they must satisfy $\mathbf{S}_i = \mathbf{A}\mathbf{C}_i\mathbf{B}'$. Likewise, $\mathbf{C}$ will be of full row rank almost surely, because it must satisfy $\mathbf{U}\mathbf{C}' = \mathbf{X}$ with $\mathbf{X}$ of rank I almost surely. When $\mathbf{C}$ is square and non-singular, all subsets of its columns are linearly independent. It follows from Result 8 that a CP solution is inevitably an INDSCAL solution. □

From Results 8 and 9, we can draw the following inferences as to CP solutions being INDSCAL solutions almost surely. First, consider the $J = 2$ cases. When $I = 2$, and $R = 2$, Result 6 guarantees that there is a (unique) solution with $\mathbf{A} = \mathbf{B}$. When $I = 2$ and $R = 3$, however, CP tends to give non-INDSCAL solutions. Still, Result 7 explains how to arrive at an INDSCAL solution. When $I = 3$, we have $R = 3$ almost surely, and Result 9 applies. When $I > 3$, CP solutions also must be INDSCAL solutions, because three slices are already enough to enforce that, and adding slices beyond the first three means adding more constraints on the solution.

For the $J = 3$ cases with $I = 2$ and $R = 3$, CP must give a (unique) INDSCAL solution (Result 6). When $I = 2$ and $R = 4$, CP tends to give non-INDSCAL solutions, but Result 7 tells us how to arrive at an INDSCAL solution. When $I = 3$, R will be 4, and the solution of Result 3 is an INDSCAL solution. Other solutions, however, do exist, and some of these may be non-INDSCAL solutions, with linear dependence in at least one submatrix of $\mathbf{C}$, as indicated in Result 8. When $I = 4$, and $R = 4$, it follows from the proof of Result 4 that CP will yield a unique solution, which by virtue of Result 6 will be an INDSCAL solution. When $I = 4$ and $R = 5$, Result 8 applies. We often observe CP yielding INDSCAL solutions, but occasionally, non-INDSCAL solutions also arise. When they do, the right null of $\mathbf{C}$ does indeed reflect that not every 4×4 submatrix of $\mathbf{C}$ is non-singular.

When $J = 3$, $I = 5$ and $R = 5$, we are back at Result 9 (non-unique INDSCAL solutions); but when $R = 6$, we may have non-INDSCAL solutions, again with linear dependence in at least one square submatrix of $\mathbf{C}$, see Result 8. Finally, when $I = 6$, we have $R = 6$ almost surely, and Result 9 applies. It can be concluded that, even when CP solutions are non-unique, they will necessarily be INDSCAL solutions in a surprisingly large variety of cases.

6. Discussion

Most results of this paper are limited to arrays having slices of order 2×2 and 3×3. Throughout the cases examined, the same picture emerges: Firstly, typical rank of symmetric three-way arrays seems to be the same or smaller than the typical rank of non-symmetric counterpart arrays. Secondly, the typical rank of symmetric arrays seems to coincide with the typical "INDSCAL dimensionality", being the smallest number of components that allows a full INDSCAL decomposition. In general, a CP solution for a symmetric array is an INDSCAL solutions when it is unique. However, even when CP solutions are not unique, CP solutions are INDSCAL solutions in quite a number of cases.

Although the proofs used in this paper rely on similar principles, they are essentially ad hoc. More general proofs, that might also cover symmetric slices of higher order than 3×3, are sorely needed.

It should be noted that we have considered INDSCAL without constraints other than that $\mathbf{A} = \mathbf{B}$. Our analysis does not account for constraints imposed by various types of preprocessing or "natural" application-specific constraints. For instance, we have ignored sign constraints on the elements of the diagonal matrices of saliences. Fitting INDSCAL subject to non-negativity constraints for the saliences is possible [13] but has been ignored in the present paper. Such constraints are likely to increase the number of dimensions of INDSCAL, whence typical rank and dim may no longer coincide.

As far as preprocessing is concerned, it may be noted that estimated distance matrices are often converted to double-centered scalar product matrices [2, p. 286], before CANDECOMP is applied. This has a rank-reducing impact on the symmetric array. Specifically, let $\mathbf{T}$ be a non-singular matrix with last column the vector of ones. Upon premultiplying all slices with $\mathbf{T}$, and postmultiplying by $\mathbf{T}'$, all slices have zeros in the last rows and columns. Because (appending or) deleting zero slices does not change the rank, and double centering preserves symmetry, a double-centered symmetric $I \times J \times J$ array has the same rank as the reduced $I \times (J-1) \times (J-1)$ version of it. This means that double centering has a strong rank-reducing impact. For instance, a symmetric $6 \times 3 \times 3$ array has typical rank 6 without centering, but the typical rank will be as small as 3 upon double centering, see Table 1. The implication is that INDSCAL will need no more than three components in a case like this.

Acknowledgements

The authors are obliged to Henk Kiers for helpful comments.

References

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

[2] 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.

[3] R.L. Harshman, Foundations of the PARAFAC procedure: models and conditions for an “explanatory” multi-modal factor analysis, UCLA Working Papers in Phonetics, vol. 16, 1970, pp. 1–84.
[4] R.A. Harshman, Determination and proof of minimum uniqueness conditions for PARAFAC1, UCLA Working Papers in Phonetics, vol. 22, 1972, pp. 111–117.
[5] J.B. Kruskal, Three-way arrays: rank and uniqueness of trilinear decompositions with applications to arithmetic complexity and statistics, Linear Algebra Appl. 18 (1977) 95–138.
[6] J.B. Kruskal, Rank, decomposition, and uniqueness for 3-way and N-way arrays, in: R. Coppi, S. Bolasco (Eds.), Multiway Data Analysis, North-Holland, Amsterdam, 1989, pp. 7–18.
[7] R. Rocci, On the maximal rank of $3 \times 3 \times 3$, Università la Sapienza, Rome, 1993.
[8] R. Rocci, J.M.F. Ten Berge, A simplification of a result by Zellini on the maximal rank of a symmetric three-way array, Psychometrika 59 (1994) 377–380.
[9] 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.
[10] J.M.F. Ten Berge, The typical rank of tall three-way arrays, Psychometrika 65 (2000) 525–532.
[11] J.M.F. Ten Berge, H.A.L. Kiers, Some clarifications of the CANDECOMP algorithm applied to INDSCAL, Psychometrika 56 (1991) 317–326.
[12] J.M.F. Ten Berge, H.A.L. Kiers, Simplicity of core arrays in three-way principal component analysis and the typical rank of $P \times Q \times 2$ arrays, Linear Algebra Appl. 294 (1999) 169–179.
[13] J.M.F. Ten Berge, H.A.L. Kiers, W.P. Krijnen, Computational solutions for the problem of negative saliences and nonsymmetry in INDSCAL, J. Classification 10 (1993) 115–124.
[14] J.M.F. Ten Berge, N.D. Sidiropoulos, On uniqueness in Candecomp/Parafac, Psychometrika 67 (2002) 399–409.
[15] J.M.F. Ten Berge, A. Smilde, Non triviality and identification of a constrained Tuckals-3 analysis, J. Chemometr. 16 (2002) 609–612.