PSYCHOMETRIKA—VOL. 65, NO. 4, 525–532
DECEMBER 2000

THE TYPICAL RANK OF TALL THREE-WAY ARRAYS

JOS M.F. TEN BERGE

UNIVERSITY OF GRONINGEN

The rank of a three-way array refers to the smallest number of rank-one arrays (outer products of three vectors) that generate the array as their sum. It is also the number of components required for a full decomposition of a three-way array by CANDECOMP/PARAFAC. The typical rank of a three-way array refers to the rank a three-way array has almost surely. The present paper deals with typical rank, and generalizes existing results on the typical rank of $I \times J \times K$ arrays with $K = 2$ to a particular class of arrays with $K \geq 2$. It is shown that the typical rank is I when the array is tall in the sense that $JK - J < I < JK$. In addition, typical rank results are given for the case where I equals $JK - J$.

Key words: three-way rank, tensorial rank, CANDECOMP, PARAFAC, three-way component analysis.

The rank of a matrix $\mathbf{X}$ is defined as the smallest number of rank-one matrices 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 in $\mathbf{A}$ and $\mathbf{B}$, then the rank of $\mathbf{X}$ is r.

Similarly, the rank of a three-way array is defined as the smallest number of (real valued) rank-one arrays (outer products of three vectors) that generate the array as their sum (Kruskal, 1977, 1989). Equivalently, the rank of a three-way array is the smallest number of components that give a perfect fit in CANDECOMP/PARAFAC (Carroll & Chang, 1970; Harshman, 1970). Specifically, let the three-way array $\underline{\mathbf{X}}$ of order $I \times J \times K$ be composed of K frontal slices $\mathbf{X}_1, \ldots, \mathbf{X}_K$, of order $I \times J$. Then a perfect fit in CANDECOMP/PARAFAC implies that there exist matrices $\mathbf{A}(I \times R)$, $\mathbf{B}(J \times R)$ and diagonal matrices $\mathbf{D}_1, \ldots, \mathbf{D}_K$ of order $R \times R$ such that, for $k = 1, \ldots, K$,

$$\mathbf{X}_k = \mathbf{AD}_k\mathbf{B}'. \tag{1}$$

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 nonsingular transformations do not affect the rank of a three-way array.

One of the complications involved in the study of three-way rank is that, unlike the situation of two-way arrays (matrices), the maximum rank of three-way arrays may exceed the so-called typical rank of these arrays, which is the rank three-way arrays of a particular order have almost surely (with probability one). To illustrate the difference between typical rank and maximal rank, consider the class of $4 \times 4 \times 2$ arrays, consisting of frontal slabs of the form

$$\mathbf{X}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{X}_2 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -b \\ 0 & 0 & b & 0 \end{bmatrix},$$

for some value of $b \neq 0$. It is well-known (ten Berge, 1991) that the rank of an array that consists of two 4×4 matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ cannot be 4 unless the eigenvalues of $\mathbf{X}_1^{-1}\mathbf{X}_2$ are real. In the

The author is obliged to Henk Kiers, Tom Snijders, and Philip Thijsse for helpful comments.

Requests for reprints should be sent to Jos ten Berge, Heijmans Institute of Psychological Research, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, THE NETHERLANDS. E-mail: j.m.f.ten.berge@ppsw.rug.nl

0033-3123/2000-4/1999-0731-A $00.75/0

present case, the eigenvalues are complex, so the rank is at least 5. Now suppose that $b^2 \neq 1$. Define $a = (2b^2 + 2)/(b^2 - 1)$ and let

$$x = \begin{bmatrix} a \\ 0 \\ ab \\ 0 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}.$$

If we subtract $\mathbf{xy}'$ from $\mathbf{X}_2$, we get the following matrix:

$$\begin{bmatrix} 0 & -(a+1) & 0 & a \\ 1 & 0 & 0 & 0 \\ 0 & -ab & 0 & -b(1-a) \\ 0 & 0 & b & 0 \end{bmatrix}.$$

It has characteristic equation $\lambda^4 + \lambda^2(b^2 - ab^2 + a + 1) + b^2 = 0$, hence, from $a = (2b^2 + 2)/(b^2 - 1)$, we have $\lambda^4 - \lambda^2(b^2+1) + b^2 = 0$, which is the same as $(\lambda+1)(\lambda-1)((\lambda+b)(\lambda-b) = 0$. It follows that $\mathbf{X}_2 - \mathbf{xy}'$ has eigenvalues 1, -1, b, and $-b$. These are real and distinct, and hence the matrix has real eigenvalues and eigenvectors. Therefore, the array that has frontal slabs $\mathbf{X}_1$ and $\mathbf{X}_2 - \mathbf{xy}'$ can be decomposed in 4 dimensions (ten Berge, 1991). Considering that subtracting $\mathbf{xy}'$ requires one dimension in CANDECOMP/PARAFAC, the original array can be been decomposed in five dimensions, and its rank is five. It follows that, when b is sampled randomly from a continuous distribution, the probability that the array has rank 5 is 1. In other words, the array has rank five almost surely, or, equivalently, the array has typical rank 5. This does not mean that the maximum rank is also 5: When $b = 1$ or $b = -1$, the rank five solution cannot be constructed, and it turns out that the rank is 6, as can be verified by running CANDECOMP/PARAFAC to convergence with 5 and 6 components, respectively (a formal but complicated proof can be found in Thijsse, 1994). But 6 is also the maximum rank for the class of arrays under consideration, because a $4 \times 4 \times 2$ array has maximum rank 6 (Kruskal, 1989, p. 10).

The maximum rank 6 will never be encountered in practice, but can be constructed by forcing b to be 1 or -1 . For the array under consideration, rank 6 is impossible from a probabilistic point of view, yet it is logically possible because arrays that have rank six can be contrived. This pinpoints the difference between typical and maximal rank.

In this example, only one parameter of the array was determined by random sampling from a continuous distribution. In the sequel, the randomness typically will pertain to *all* elements of the array. However, in that context, the distinction between typical and maximum rank remains the same.

It should be noted that typical rank need not be single-valued. For instance, Kruskal (1983), also see ten Berge (1991), has shown that $3 \times 3 \times 2$ arrays have rank 3 or rank 4 (the maximum rank) with positive probability, and that the other possible rank values (0 or 1) arise with probability zero. In this case we shall say that the typical rank is {3, 4}.

From the perspective of data analysis, the typical rank is more interesting than the maximal rank because, in the analysis of empirical data, arrays with a rank other than the typical rank will not be encountered. The concept of (typical) rank of three-way arrays is as fundamental to CANDECOMP/PARAFAC as is the idea that a full decomposition of a two-way data matrix by Principal Components Analysis usually requires as many components as there are variables.

The study of maximum three-way rank is far from straightforward, and the literature shows some piecemeal results for specific cases (e.g., Kruskal, 1989, p. 10; Franc, 1992, pp. 214–215). The only case that has been generally solved is that of $K = 2$. For the study of typical rank of three-way arrays, the situation is more or less the same. The typical rank of all $I \times J \times K$ arrays with $K = 2$ has been determined by ten Berge and Kiers (1999). Specifically, they have shown that an $I \times J \times 2$ array has typical rank I when $J < I < 2J$. In addition, they have shown that every $I \times I \times 2$ array has typical rank $\{I, I + 1\}$, which means that both rank I and rank $I + 1$

arise with positive probability, and all other ranks have probability zero. Also, it is trivial that the typical rank of an $I \times J \times 2$ array is $2J$ when $I \geq 2J$. The present paper is aimed at generalizing these typical rank results to arrays with $I \geq J \geq K > 2$, when I (the largest of I, J, and K) is relatively large compared with J and K, in a sense to be specified below.

As a matter of convenience, let the three-way arrays be classified according to the sizes of I, J and K, as follows: An array with $I \geq J \geq K$ will be called "very tall" when $I \geq JK$; it will be called "tall" when $JK - J < I < JK$, and it will be called "compact" when $I \leq JK - J$. It may be noted that those $I \times J \times 2$ arrays which have typical rank I (ten Berge & Kiers, 1999) are tall. In the present paper, all tall three-way arrays are shown to have typical rank I. In addition, some results are given for the typical rank of the tallest among the compact arrays, that is, those with $I = JK - J$. It will be shown that these arrays have either typical rank I or typical rank $\{I, I+1\}$. The latter result generalizes the typical rank result for $I \times I \times 2$ arrays (ten Berge & Kiers, 1999).

This paper starts with a general treatment of the tall arrays. Then the case $I = JK - J$ will be dealt with. The compact arrays will not be discussed (except those with $I = JK - J$) because their typical rank is unknown. The very tall arrays will not be discussed either, but for a different reason: These arrays trivially have typical rank and maximum rank JK. For instance, a $12 \times 3 \times 3$ array has typical rank and maximum rank 9. To verify this, first note that the matrix $\mathbf{A}$ in CANDECOMP/PARAFAC must almost surely have at least nine columns to account for the 9 linearly independent columns of $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$. But 9 columns will also suffice: Factor $\mathbf{X}_k$ as $\mathbf{X}_k = \mathbf{A}_k\mathbf{B}_k'$, $k = 1, 2, 3$, define $\mathbf{A} = [\mathbf{A}_1 \mid \mathbf{A}_2 \mid \mathbf{A}_3]$ and $\mathbf{B} = [\mathbf{B}_1 \mid \mathbf{B}_2 \mid \mathbf{B}_3]$ and set $\mathbf{D}_1 = \text{diag}\{1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\}$, $\mathbf{D}_2 = \text{diag}\{0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\}$, and $\mathbf{D}_3 = \text{diag}\{0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\}$. This yields a decomposition $\mathbf{X}_k = \mathbf{A}\mathbf{D}_k\mathbf{B}'$, $k = 1, 2, 3$, from which it is immediate that the rank cannot be larger than 9.

1. The Typical Rank of Tall Arrays

For tall arrays with $K = 2$, ten Berge and Kiers (1999) have shown that the typical rank is I. They used a transformation to simplicity to render the typical rank "almost visible." However, when $K > 2$, it is not obvious how to transform arrays to simplicity, except for arrays with $I = JK - 1$ (the tallest among the tall arrays), see Murakami, ten Berge & Kiers (1998). In the present paper, typical rank results will be obtained in full disregard of simplicity.

The starting point taken here is the observation that an array which has all I horizontal slices of rank one has at most rank I. For instance, when $I = 5$ and $K = 2$ and the two frontal slices $\mathbf{X}_1$ and $\mathbf{X}_2$ of $\underline{\mathbf{X}}$ are of the form

$$\mathbf{X}_1 = \begin{bmatrix} \mathbf{a}' \\ \mathbf{b}' \\ \mathbf{c}' \\ \mathbf{d}' \\ \mathbf{e}' \end{bmatrix} \quad \text{and} \quad \mathbf{X}_2 = \begin{bmatrix} \mathbf{a}' \\ 2\mathbf{b}' \\ 3\mathbf{c}' \\ 4\mathbf{d}' \\ 5\mathbf{e}' \end{bmatrix} \tag{2}$$

for certain vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ and $\mathbf{e}$, then it is clear that the rank of $\underline{\mathbf{X}}$ is 5 at most because each horizontal slice can be accounted for by one rank-one array. In terms of (1), the array can be decomposed by $\mathbf{A} = \mathbf{I}_5$, $\mathbf{B}' = \mathbf{X}_1$, $\mathbf{D}_1 = \mathbf{I}_5$, and $\mathbf{D}_2 = \text{diag}\{1\ 2\ 3\ 4\ 5\}$. The next observation is also essential.

Result 1. When the rank of an $I \times J \times K$ array with $JK > I$ is I, then there exists almost surely a nonsingular matrix $\mathbf{S}$ such that $\mathbf{S}\mathbf{X}_1, \ldots, \mathbf{S}\mathbf{X}_K$ all have proportional rows, as is exemplified in (2).

Proof. Let $\mathbf{X}_k = \mathbf{A}\mathbf{D}_k\mathbf{B}'$, $k = 1, \ldots, K$. Let the rank of $\underline{\mathbf{X}}$ be I. Then $\mathbf{A}$ is usually an $I \times I$ matrix of rank I. It is possible that $\mathbf{A}$ is of rank less than I, but that would imply that the

I rows of $\mathbf{X}_1, \ldots, \mathbf{X}_K$ would not span a I-dimensional subspace, an event of probability zero. Premultiplying by $\mathbf{S} = \mathbf{A}^{-1}$ yields $\mathbf{S}\mathbf{X}_k = \mathbf{D}_k\mathbf{B}'$. Clearly, these matrices have the same rows (those of $\mathbf{B}'$), up to differential scaling by $\mathbf{D}_k$, $k = 1 \ldots, K$. □

It is instructive to consider the case $I = 9$, $J = 4$ and $K = 3$ as an example. Define the 9×12 matrix $\mathbf{X} = [\mathbf{X}_1 \mid \mathbf{X}_2 \mid \mathbf{X}_3]$. Suppose that it has rank 9. Then there is a nonsingular matrix $\mathbf{S}$ such that row i of $\mathbf{S}\mathbf{X}_1$ is proportional to row i of $\mathbf{S}\mathbf{X}_2$ and also to row i of $\mathbf{S}\mathbf{X}_3$, $i = 1, \ldots, 9$. As will be shown below, we can, for instance, solve for each row $\mathbf{s}_i'$ of $\mathbf{S}$ separately, by solving $\mathbf{s}_i' i\mathbf{X}_1 = \mathbf{s}_i'\mathbf{X}_2$ and $\mathbf{s}_i' i\mathbf{X}_2 = \mathbf{s}_i'\mathbf{X}_3$, which can also be written as $\mathbf{s}_i'(i\mathbf{X}_1 - \mathbf{X}_2) = \mathbf{s}_i'(i^2\mathbf{X}_1 - \mathbf{X}_3) = \mathbf{0}'$. This requires that the vector $\mathbf{s}_i$ is orthogonal to the columns of a 9×8 matrix $\mathbf{W}_i$ of the form

$$\mathbf{W}_i = [i\mathbf{X}_1 - \mathbf{X}_2 \mid i^2\mathbf{X}_1 - \mathbf{X}_3]. \tag{3}$$

Clearly, such vectors always exist. If the resulting matrix $\mathbf{S}$ is nonsingular, we have shown indeed that $\underline{\mathbf{X}}$ has at most rank 9. This example demonstrates the logic behind the following result:

Result 2. Tall arrays have typical rank I.

Proof. Construct a transformation matrix $\mathbf{S}$ with rows $\mathbf{s}_i'(i = 1, \ldots, I)$ by solving, for $i = 1, \ldots, I$,

$$\mathbf{s}_i'(i\mathbf{X}_1 - \mathbf{X}_2) = \mathbf{s}_i'(i^2\mathbf{X}_1 - \mathbf{X}_3) = \cdots = \mathbf{s}_i'(i^{K-1}\mathbf{X}_1 - \mathbf{X}_K) = \mathbf{0}'. \tag{4}$$

Hence the vector $\mathbf{s}_i$, $i = 1, \ldots, I$, has to be orthogonal to the columns of an $I \times (JK - J)$ matrix $\mathbf{W}_i$ of the form

$$\mathbf{W}_i = [i^1\mathbf{X}_1 - \mathbf{X}_2 \mid i^2\mathbf{X}_1 - \mathbf{X}_3 \mid \cdots \mid i^{K-1}\mathbf{X}_1 - \mathbf{X}_K]. \tag{5}$$

Because $I > JK - J$, we have enough parameters to solve (4) for each $\mathbf{s}_i$. An explicit way to solve it is by constructing, as was done in ten Berge and Kiers (1999), the matrix $\mathbf{H}_i = (\mathbf{I}_I - \mathbf{W}_i(\mathbf{W}_i'\mathbf{W}_i)^+\mathbf{W}_i')$, where the superscript $^+$ refers to the Moore–Penrose inverse, and taking $\mathbf{s}_i = \mathbf{H}_i\mathbf{e}_i$, where $\mathbf{e}_i$ is column i of $\mathbf{I}_I$. It is important to note that this solution will return $\mathbf{S} = \mathbf{I}_I$ whenever that is a possible solution. Specifically, let $\mathbf{e}_i$ be orthogonal to $\mathbf{W}_i$, $i = 1, \ldots, I$. Then $\mathbf{e}_i$ is in the column space of $\mathbf{H}_i$, and $\mathbf{H}_i\mathbf{e}_i = \mathbf{e}_i$. Clearly, we now obtain $\mathbf{s}_i = \mathbf{H}_i\mathbf{e}_i = \mathbf{e}_i$, $i = 1, \ldots, I$, hence the solution for $\mathbf{S}$ will indeed be $\mathbf{I}_I$.

It follows from Fisher (1966, Theorem 5.A.2), also see ten Berge and Kiers (1999), that we can almost surely solve for a nonsingular $\mathbf{S}$ (its determinant is an analytical function of the elements of $\underline{\mathbf{X}}$) if we can solve for a nonsingular $\mathbf{S}$ at least once. For every array size, it is easy to construct a case where the proportionality requirement of Result 1 is already satisfied in such a way that $\mathbf{S} = \mathbf{I}_I$. The fact that a nonsingular transformation arises at least once shows that the transformation will be nonsingular for almost every $\underline{\mathbf{X}}$. □

Result 2 generalizes the result of ten Berge and Kiers for tall $K = 2$ arrays to all tall arrays. It also includes the typical rank of the particular case where $I = JK - 1$, studied by Murakami, ten Berge and Kiers (1998). In fact, these are the tallest among the tall arrays. But also less tall arrays are covered by Result 2. For instance, Atkinson and Stephens (1979), also see Franc (1992, p. 214), have reported that the maximum rank of a $7 \times 3 \times 3$ array is 8. Result 2 implies that the typical rank is 7.

It should be noted that the solution to (4), as constructed above, implies an explicit solution for a full rank CANDECOMP/PARAFAC decomposition of tall arrays. However, this is just one solution among others. To ensure linearly independent rows of $\mathbf{S}$, elements from the columns $2, \ldots, K$ of the nonsingular $I \times I$ Vandermonde matrix, defined by

$$\mathbf{V}(I) = \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 4 & 8 & \cdots & 2^{I-1} \\ 1 & 3 & 9 & 27 & \cdots & 3^{I-1} \\ 1 & 4 & 16 & 64 & \cdots & 4^{I-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & I & I^2 & I^3 & \cdots & I^{I-1} \end{bmatrix} \tag{6}$$

have been inserted as weights into (4) to construct the elements of the matrices $\mathbf{D}_k$, $k = 1, \ldots, K$. However, every other set of weights will also work, as long as linear independence is preserved. It follows that the solution for a rank I CANDECOMP/PARAFAC decomposition of a tall array is indeterminate, in the sense that the weights in $\mathbf{D}_1, \ldots, \mathbf{D}_K$ can be fixed to be almost anything. The choice of weights in turn affects $\mathbf{A}$ and $\mathbf{B}$.

2. The Typical Rank When $I = JK - J$

ten Berge and Kiers (1999) have also obtained a typical rank result for a class of compact arrays with $K = 2$, namely, the $I \times I \times 2$ arrays. This result will now be generalized by considering "the tallest among the compact arrays," that is, when $I = JK - J$. The following result is immediate.

Result 3. Arrays with $I = JK - J$, have almost surely rank I or rank $I + 1$.

Proof. Rank less than I would imply linear dependence among the I rows of $\mathbf{X}$, which has probability zero. When a random horizontal slice is added to the array, a $(JK - J + 1) \times J \times K$ array results. The latter array is tall and therefore has rank $I + 1$ almost surely. Because adding a slice cannot reduce the rank, the original array had at most rank $I + 1$, almost surely. □

The next issue to clarify is whether or not rank I arises with positive probability. We consider the $6 \times 3 \times 3$ case as an example. Transform the first six columns of $\mathbf{X}$ to $\mathbf{I}_6$, and note that the last three columns of the transformed matrix contain two 3×3 matrices, $\mathbf{Y}_{13}$ and $\mathbf{Y}_{23}$ say. So the matrix containing the three frontal slabs of the transformed array next to each other has the form

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 & \mathbf{Y}_2 & \mathbf{Y}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{O} & \mathbf{Y}_{13} \\ \mathbf{O} & \mathbf{I}_3 & \mathbf{Y}_{23} \end{bmatrix} \tag{7}$$

Now suppose that both $\mathbf{Y}_{13}$ and $\mathbf{Y}_{23}$ have a real-valued eigendecomposition $\mathbf{Y}_{13} = \mathbf{K}_1\mathbf{L}_1\mathbf{K}_1^{-1}$ and $\mathbf{Y}_{23} = \mathbf{K}_2\mathbf{L}_2\mathbf{K}_2^{-1}$. Then both matrices can be diagonalized simultaneously with $\mathbf{I}_3$ in the corresponding rows of $\mathbf{Y}_1$ and $\mathbf{Y}_2$, respectively. An explicit CANDECOMP/PARAFAC decomposition can be constructed as follows: Let

$$\mathbf{A} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{K}_2 \end{bmatrix}$$

and

$$\mathbf{B}' = \begin{bmatrix} \mathbf{K}_1^{-1} \\ \mathbf{K}_2^{-1} \end{bmatrix},$$

and let

$$\mathbf{D}_1 = \begin{bmatrix} \mathbf{I}_3 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix},$$

$$\mathbf{D}_2 = \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_3 \end{bmatrix},$$

and

$$\mathbf{D}_3 = \begin{bmatrix} \mathbf{L}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{L}_2 \end{bmatrix}.$$

Then it is easy to verify that $\mathbf{Y}_k = \mathbf{A}\mathbf{D}_k\mathbf{B}'$, $k = 1, 2, 3$. This shows that a rank 6 decomposition is possible, so the array $\underline{\mathbf{Y}}$ has rank 6, and hence $\underline{\mathbf{X}}$ has rank 6. The set of arrays that can be transformed such that $\mathbf{Y}_{13}$ and $\mathbf{Y}_{23}$ simultaneously have real eigendecompositions has positive volume. Hence, the set of $6 \times 3 \times 3$ arrays that have rank 6 has positive volume. This will now be proven in full generality as Result 4.

Result 4. Arrays with $I = JK - J$ have rank I with positive probability.

Proof. Upon transforming the first I columns to $\mathbf{I}_I$, and noting that the $K - 1$ submatrices, of order $J \times J$, of the last frontal slice that results have real eigendecompositions with positive probability, the result is immediate. □

At this point, it is tempting to believe that the existence of real eigendecompositions for the submatrices of the last frontal slice (after transforming the first I columns to $\mathbf{I}_I$) is not only sufficient but also necessary for rank I. This is not the case. Before stating the underlying principle in full generality (amounting to Result 5) we shall use the $6 \times 3 \times 3$ case once again as an example.

Consider the transformed array (7) again, but this time let the assumption of real eigenvalues of $\mathbf{Y}_{13}$ and $\mathbf{Y}_{23}$ be dropped. To better explain the motivation for the method to be developed, it will be instructive to assume that the array still has rank 6, and then construct a CANDECOMP/PARAFAC solution in 6 dimensions. In the process, it will appear that the method to construct such a solution almost surely works for $6 \times 3 \times 3$ arays. In this heuristic way, it will be shown the typical rank of a $6 \times 3 \times 3$ array is 6, even when the assumption of real eigenvalues is not met.

So let, for $k = 1, 2, 3$, $\mathbf{Y}_k = \mathbf{A}\mathbf{D}_k\mathbf{B}'$, with $\mathbf{A}$ of order 6×6. Define $\mathbf{S} = \mathbf{A}^{-1}$. From Result 1 it follows that we can solve $\mathbf{S}\mathbf{Y}_k = \mathbf{D}_k\mathbf{B}'$, $k = 1, 2, 3$. Using the special form of (7) we see that row i of $\mathbf{S}$ can be written as $[d_{i1}\mathbf{b}_i' \mid d_{i2}\mathbf{b}_i']$, where $\mathbf{b}_i'$ is row i of $\mathbf{B}'$, and d_{ik} is the ith diagonal element of $\mathbf{D}_k$. It follows that we need to determine $\mathbf{S}$ of the form $\mathbf{S} = [\mathbf{D}_1\mathbf{B}' \mid \mathbf{D}_2\mathbf{B}']$. This does narrow down the choice of $\mathbf{S}$, but in no way constrains the elements of $\mathbf{B}$, $\mathbf{D}_1$, and $\mathbf{D}_2$. What remains is to solve $\mathbf{S}\mathbf{Y}_3 = \mathbf{D}_3\mathbf{B}'$, which can be written as $[\mathbf{D}_1\mathbf{B}' \mid \mathbf{D}_2\mathbf{B}']\mathbf{Y}_3 = \mathbf{D}_3\mathbf{B}'$, so $\mathbf{D}_1\mathbf{B}'\mathbf{Y}_{13} + \mathbf{D}_2\mathbf{B}'\mathbf{Y}_{23} = \mathbf{D}_3\mathbf{B}'$. This yields six equations for the rows of $\mathbf{B}'$. For each row $\mathbf{b}_i'$, $i = 1, \ldots, I$, of $\mathbf{B}$ we need to solve the equation

$$d_{i1}\mathbf{b}_i'\mathbf{Y}_{13} + d_{i2}\mathbf{b}_i'\mathbf{Y}_{23} - d_{i3}\mathbf{b}_i'\mathbf{I}_3 = \mathbf{0}',$$

or, equivalently,

$$\mathbf{b}_i'[d_{i1}\mathbf{Y}_{13} + d_{i2}\mathbf{Y}_{23} - d_{i3}\mathbf{I}_3] = \mathbf{0}'.$$

Clearly, this is an eigenvalue problem, and, for arbitrary values of d_{i1} and d_{i2}, the solution is found at once by letting d_{i3} be the largest (or, in fact, the only) real eigenvalue of $(d_{i1}\mathbf{Y}_{13} + d_{i2}\mathbf{Y}_{23})'$, with $\mathbf{b}_i$ the associated eigenvector. The only problem that might arise is when the resulting $\mathbf{S}$ becomes singular. Again, this can be avoided, for instance, by letting d_{i1} and d_{i2} be the second and third element, respectively, of row i of the $I \times I$ Vandermonde matrix, see (6).

Because every $J \times J$ matrix, with J odd, has at least one real eigenvalue, the method described here always produces a solution, and the resulting $\mathbf{S}$ will be nonsingular almost surely. Therefore, the $6 \times 3 \times 3$ array has typical rank 6, regardless of the assumption that $\mathbf{Y}_{13}$ and $\mathbf{Y}_{23}$ have real valued eigendecompositions. This property generalizes to the following result.

Result 5. When $I = JK - J$ and $K > 2$ and J is odd, the typical rank is I.

Proof. Transform the array to $\mathbf{Y} = [\mathbf{I}_I \mid \mathbf{Y}_K]$, as in the previous example. Partition $\mathbf{Y}_K$ in $K-1$ submatrices$\mathbf{Y}_{1K}, \ldots, \mathbf{Y}_{K-1,K}$ of order $J \times J$. In order to attain a transformed array $\mathbf{SY}$, with each row consisting of proportional subvectors, see Result 1, we need to solve the equations $\mathbf{SY}_K = \mathbf{D}_K\mathbf{B}'$ for $\mathbf{D}_k$ and $\mathbf{B}$, $k = 1, \ldots, K$, with row i of $\mathbf{S}$ defined as $[d_{i1}\mathbf{b}_i' | \ldots | d_{i,K-1}\mathbf{b}_i']$, where $\mathbf{b}_i'$ is row i of $\mathbf{B}'$, $i = 1, \ldots, I$. This amounts to solving the I equations

$$d_{i1}\mathbf{b}_i'\mathbf{Y}_{1K} + d_{i2}\mathbf{b}_i'\mathbf{Y}_{2K} + \cdots + d_{i,K-1}\mathbf{b}_i'\mathbf{Y}_{K-1,K} = d_{iK}\mathbf{b}_i'. \tag{8}$$

Upon defining $d_{ik} = i^k$, thus using elements from the Vandermonde matrix again, (8) can be written as

$$\mathbf{b}_i'[i\mathbf{Y}_{1K} + i^2\mathbf{Y}_{2K} + \cdots + i^{K-1}\mathbf{Y}_{K-1,K} - d_{iK}\mathbf{I}_J] = \mathbf{0}'. \tag{9}$$

Clearly, this amounts to solving I eigenvalue problems. Because, for K odd, there always is a real solution for d_{iK}, we can solve for each $\mathbf{b}_i$ independently, for example, by setting d_{iK} equal to the largest real eigenvalue. When we solve (9) for J odd and $K > 2$, a nonsingular transformation $\mathbf{S}$, which renders all rows of $\mathbf{SX}$ proportional, will arise almost surely, by virtue of a similar analytical function argument as was used in the proof of Result 2. It follows that the typical rank is indeed I, when J is odd and $K > 2$. □

It should be noted that Result 5 does not apply when $I = JK - J$ and $K = 2$. This is not surprising considering that the method given in the proof of Result 5 breaks down in that case, because, when $K = 2$, there is no differential weighting in (8): Different weights for $\mathbf{Y}_{1K}'$ will all produce the same eigenvector.

Neither does Result 5 apply when $I = JK - J$ and J is even. The reason for this is less obvious. Suppose there is a linear combination of $\mathbf{Y}_{1K}', \mathbf{Y}_{2K}', \ldots, \mathbf{Y}_{K-1,K}'$ that has at least one (hence at least two) real eigenvalues. Then small perturbations of these weights will produce different real eigenvectors, and a nonsingular $\mathbf{S}$, solving $\mathbf{SY}_K = \mathbf{D}_K\mathbf{B}'$ for $k = 1, \ldots, K$, can be obtained, which means that the rank of the array is I, see Result 1. However, when J is even, there exist matrix pairs such that every linear combination of these matrices has all eigenvalues complex, which means that the rank is larger than I. Specifically, for $8 \times 4 \times 3$ arrays, it may happen that no linear combination of $\mathbf{Y}_{13}'$ and $\mathbf{Y}_{23}'$ has a real eigenvalue, implying that a real valued solution for $\mathbf{b}_i$ in (9) does not exist. As a matter of fact, such cases can be encountered by random sampling, albeit that they are extremely rare. It follows that we encounter both rank 8 and rank 9 with positive probability in the $8 \times 4 \times 3$ case. This and other typical rank results are summarized in Table 1.

TABLE 1.
Typical rank results for some arrays with $K = 2$ and $K = 3$

	$K=2$				$K=3$		
	$J=2$	$J=3$	$J=4$		$J=3$	$J=4$	$J=5$
$I=2$	{2, 3}	3	4	$I=5$	?	?	?
$I=3$	3	{3, 4}	4	$I=6$	6	?	?
$I=4$	4	4	{4, 5}	$I=7$	7	?	?
$I=5$	4	5	5	$I=8$	8	{8, 9}	?
$I=6$	4	6	6	$I=9$	9	9	?
$I=7$	4	6	7	$I=10$	9	10	10
$I=8$	4	6	8	$I=11$	9	11	11
$I=9$	4	6	8	$I=12$	9	12	12

3. Discussion

Because I is obviously a lower bound to the typical rank of a tall array, it comes as a surprise that it is also an upper bound to the typical rank. An important implication is that a full three-way CANDECOMP/PARAFAC decomposition of a tall $I \times J \times K$ array $\underline{\mathbf{X}}$ takes no more components than would a full two-way decomposition of the associated $I \times JK$ matrix $\mathbf{X}$ by ordinary principal components analysis. In this sense, the three-way CANDECOMP/PARAFAC structure of a tall array comes with the same rank as its two-way structure.

The results of this paper also shed light on the behavior of the residual sum of squares in the fitting of CANDECOMP/PARAFAC: When any array is fitted in the least squares sense by CANDECOMP/PARAFAC, using the number of components specified by the typical rank, perfect fit will be obtained to any degree of accuracy. This also holds when the specific array under consideration does have a three-way rank above the typical rank. Arrays with higher rank than their typical rank are singular points, which are surrounded by non-singular points with ranks equal to the typical rank. Therefore, we can get the fit as close to zero as we like when we use the "typical" number of components in CANDECOMP/PARAFAC. This implies that the (numerical) study of maximum rank by running CANDECOMP/PARAFAC to "perfect fit" is of no avail: It could not distinguish maximum rank from typical rank.

CANDECOMP/PARAFAC is often applied to fit the INDSCAL model in the least-squares sense (Carroll & Chang, 1970). In that case the array must consist of symmetric slabs and the matrices $\mathbf{A}$ and $\mathbf{B}$ in (1) are constrained to be equal. It is important to note that the typical rank results of the present paper do not carry over to the number of dimensions required for perfect fit in INDSCAL. For instance, when seven symmetric 3×3 matrices are strung out as row vectors in a 7×9 matrix $\mathbf{X}$, we can verify that the array is tall. Nevertheless, both the number of dimensions needed for perfect INDSCAL fit and the typical rank are 6 rather than 7, see Rocci & ten Berge (1994). This discrepancy can be explained by the fact that symmetry does not arise with random sampling. To establish the typical dimensionality of INDSCAL, the results of the present paper are of no avail.

Establishing the typical dimensionality in INDSCAL will be a topic for future research. In addition, the typical rank of compact arrays is yet to be determined.

References

Atkinson, M.D., & Stevens, N.M. (1979). On the multiplicative complexity of a family of bilinear forms. *Linear Algebra and its Applications*, *27*, 1–8.

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

Fisher, F.M. (1966). *The identification problem in Econometrics*. New York, NY: McGraw-Hill.

Franc, A. (1992). *Etudes algebriques des multitableaux: Apports de l'algebre tensorielle*. Unpublished doctoral dissertation, University of Montpellier II.

Harshman, R.L. (1970). Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis. *UCLA Working Papers in Phonetics*, *16*, 1–84.

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

Kruskal, J.B. (1983). *Statement of some current results about three-way arrays*. Unpublished manuscript, AT&T Bell Laboratories, Murray Hill, NJ.

Kruskal, J.B. (1989). Rank, decomposition, and uniqueness for 3-way and N-way arrays. In R. Coppi & S. Bolasco (Eds.), *Multiway data analysis* (pp. 7–18). Amsterdam: North-Holland.

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

Rocci, R., & ten Berge, J.M.F. (1994). A simplification of a result by Zellini on the maximal rank of symmetric three-way arrays. *Psychometrika*, *59*, 377–380.

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

ten Berge, J.M.F., & Kiers, H.A.L. (1999). Simplicity of core arrays in three-way principal component analysis and the typical rank of $P \times Q \times 2$ arrays. *Linear Algebra and its Applications*, *294*, 169–179.

Thijsse, G.P.A. (1994). *Simultaneous diagonal forms for pairs of matrices* (Report 9450/B). Erasmus University, Rotterdam: Econometric Institute.

Manuscript received 23 APR 1999
Final version received 22 NOV 1999