JOURNAL OF CHEMOMETRICS
J. Chemometrics 2000; **14**: 229–239

On the uniqueness of multilinear decomposition of *N*-way arrays

Nicholas D. Sidiropoulos[1]*† and Rasmus Bro[2]

[1]*Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA*
[2]*Food Technology, Department of Dairy and Food Science, Royal Veterinary and Agricultural University, Frederiksberg, Denmark*

SUMMARY

We generalize Kruskal's fundamental result on the uniqueness of trilinear decomposition of three-way arrays to the case of multilinear decomposition of four- and higher-way arrays. The result is surprisingly general and simple and has several interesting ramifications. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: PARAFAC; multiway analysis; multilinear algebra; multilinear decomposition; rank; uniqueness; Khatri–Rao product; identifiability; CANDECOMP; INDSCAL; multidimensional scaling

1. INTRODUCTION

Consider an $I \times J$ matrix $\mathbf{X}$ and suppose that rank $(\mathbf{X}) = 3$. Let $x_{i,j}$ denote the (i, j)th entry of $\mathbf{X}$. Then it holds that $x_{i,j}$ admits a *three-component bilinear decomposition*

$$x_{i,j} = \sum_{f=1}^{3} a_{i,f} b_{j,f} \tag{1}$$

for all $i = 1,\ldots,I$ and $j = 1,\ldots,J$. Equivalently, letting $\mathbf{a}_f := [a_{1,f},\ldots,a_{I,f}]^{\mathrm{T}}$ and similarly for $\mathbf{b}_f$,

$$\mathbf{X} = \mathbf{a}_1\mathbf{b}_1^{\mathrm{T}} + \mathbf{a}_2\mathbf{b}_2^{\mathrm{T}} + \mathbf{a}_3\mathbf{b}_3^{\mathrm{T}} \tag{2}$$

i.e. $\mathbf{X}$ can be written as a sum of three outer products of the respective component vectors, which constitute rank-1 matrices. Let $\mathbf{A} := [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]$ and $\mathbf{B} := [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$, and $\mathbf{X}$ can be expressed as $\mathbf{X} = \mathbf{A}\mathbf{B}^{\mathrm{T}}$. Note the trivial fact that $\mathbf{A}\mathbf{T}\mathbf{T}^{-1}\mathbf{B}^{\mathrm{T}} = \mathbf{X}$ for any invertible $\mathbf{T}$; hence, given $\mathbf{X}$, infinitely many $(\mathbf{A}, \mathbf{B})$ pairs can potentially give rise to $\mathbf{X}$. Unless some additional structure is assumed or

* Correspondence to: N. D. Sidiropoulos, Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA.
† E-mail: nikos@ece.umn.edu
Contract/grant sponsor: National Science Foundation; Contract/grant number: NSF/CAREER CCR-9733540; Contract/grant number: NSF/Wireless CCR-9979295
Contract/grant sponsor: ONR/CSTAP
Contract/grant sponsor: LMC (Center for Advanced Food Studies).
Contract/grant sponsor: PMP (Center for Predictive Multivariate Process Analysis).

Received 29 August 1999
Accepted 2 November 1999

imposed on **A** and/or **B**, it is impossible to uniquely unravel the component profiles $\mathbf{a}_f$ and $\mathbf{b}_f$ from **X**.

Let us now consider an $I \times J \times K$ *three-way array* $\underline{\mathbf{X}}$ with typical element $x_{i,j,k}$ and the *F-component trilinear decomposition*

$$x_{i,j,k} = \sum_{f=1}^{F} a_{i,f} b_{j,f} c_{k,f} \tag{3}$$

for all $i = 1,\ldots,I$, $j = 1,\ldots,J$ and $k = 1,\ldots,K$. Equation (3) expresses the three-way array $\underline{\mathbf{X}}$ as a sum of F rank-1 three-way factors. Analogous to the definition of matrix (two-way array) rank, the rank of a three-way array $\underline{\mathbf{X}}$ can be defined [1,2] as the minimum number of rank-1 (three-way) components needed to decompose $\underline{\mathbf{X}}$.

Define an $I \times F$ matrix **A** with typical element $\mathbf{A}(i,f) := a_{i,f}$, $J \times F$ matrix **B** with typical element $\mathbf{B}(j,f) := b_{j,f}$ and a $K \times F$ matrix **C** with typical element $\mathbf{C}(k,f) := c_{k,f}$. A distinguishing feature of the trilinear model is its uniqueness. Under mild conditions the trilinear model is essentially unique; that is, given $\underline{\mathbf{X}}$, **A, B** and **C** are unique up to permutation and scaling of columns— both mostly trivial and inherently unresolvable ambiguities.

The first published uniqueness results are due to Harshman [3,4], who actually developed PARAFAC based on a principle put forth by Cattell [5] (see also Carroll & Chang [6]). Other results soon followed (e.g. Reference [7]), culminating in Kruskal's [1,2] result, which is the deepest to date. An unusual variety of applications in diverse disciplines has always been behind the quest for improved understanding of the trilinear model and extensions, generalizations and even restrictions (N. D. Sidiropoulos, X. Liu, submitted manuscript) thereof.

Any four-way array can be 'unfolded' into a three-way array, much like a matrix can be unfolded into a vector via the standard vec(·) operation; hence uniqueness of quadrilinear decomposition of four-way arrays follows from uniqueness of trilinear decomposition of three-way arrays. This simplistic view, however, does not shed light on the exact conditions under which a unique quadrilinear decomposition can be guaranteed, nor does it provide any clues as to the rigidness of the required conditions as one moves to higher dimensions. In this paper we generalize Kruskal's result to the case of multilinear decomposition of four- and higher-way arrays, providing a precise and simple sufficient condition for uniqueness which becomes less restrictive as one moves to higher dimensions.

The rest of this paper is structured as follows. Section 2 contains necessary preliminaries, including Kruskal's uniqueness result (Theorem 1) and a basic lemma (Lemma 1) that is key in extending it to N-way arrays. Proof of the basic lemma is included. Section 3 contains the main result and its proof. Conclusions are drawn in Section 4. A compact proof of Kruskal's uniqueness result for the three-way case is included in the Appendix for completeness.

2. PRELIMINARIES

The k-rank of a matrix is a rank-like concept that plays a key role in multilinear algebra. The concept of k-rank is implicit in the seminal work of Kruskal [1], but the term was later coined by Harshman and Lundy [8] (k-rank stands for *Kruskal rank*).

Definition 1

Given $\mathbf{A} \in \mathbb{C}^{I\times F}$, $r_{\mathbf{A}} := \text{rank}\ (\mathbf{A}) = r$ iff it contains *at least* a collection of r linearly independent columns, and this fails for $r+1$ columns. $k_{\mathbf{A}}$ (the k-rank of **A**) $= r$ iff *every* r columns are linearly independent, and this fails for at least one set of $r+1$ columns ($k_{\mathbf{A}} \le r_{\mathbf{A}} \le \min(I,F)$ for all **A**).

We are now ready to state Kruskal's uniqueness result [1].

Theorem 1 (uniqueness of trilinear decomposition)

Consider the *F*-component *trilinear* model

$$x_{i,j,l,m} = \sum_{f=1}^{F} a_{i,f} b_{j,f} c_{k,f}$$

for $i = 1,\dots,I$, $j = 1,\dots,J$ and $k = 1,\dots,K$, with $a_{i,f}, b_{j,f}, c_{k,f} \in \mathbb{C}$. Define an $I \times F$ matrix $\mathbf{A}$ with typical element $\mathbf{A}(i,f) := a_{i,f}$ and similarly $J \times F$ and $K \times F$ matrices $\mathbf{B}$ and $\mathbf{C}$. Given $x_{i,j,k}$ for $i = 1,\dots,I$, $j = 1,\dots,J$ and $k = 1,\dots,K$, $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are unique up to permutation and scaling of columns provided that

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2F + 2 \tag{4}$$

Note that Theorem 1 holds for real- or complex-valued parameters and array elements. The original proof of Kruskal was targeted to the real-valued case (all matrix and array elements drawn from $\mathbb{R}$) owing to some subtleties involved in the definition of three-way array rank [1,2]. In an attempt to investigate if the result carries over to the complex case, a compact proof that checks for the complex case has been derived in Reference [9]. That proof is also provided in the Appendix herein for completeness.

The following lemma is key in extending Kruskal's result to *N*-way arrays.

Lemma 1 (k-rank of Khatri–Rao product)

Consider $\mathbf{A} = [\mathbf{a}_1,\dots,\mathbf{a}_F] \in \mathbb{C}^{I \times F}$, $\mathbf{B} = [\mathbf{b}_1,\dots,\mathbf{b}_F] \in \mathbb{C}^{J \times F}$ and

$$\mathbf{B} \odot \mathbf{A} := \begin{bmatrix} \mathbf{A}\mathbf{D}_1(\mathbf{B}) \\ \mathbf{A}\mathbf{D}_2(\mathbf{B}) \\ \vdots \\ \mathbf{A}\mathbf{D}_J(\mathbf{B}) \end{bmatrix} = [\mathbf{b}_1 \otimes \mathbf{a}_1, \ \cdots, \ \mathbf{b}_F \otimes \mathbf{a}_F]$$

where $\otimes$ stands for the Kronecker product, $\odot$ stands for the Khatri–Rao (column-wise Kronecker) product and $\mathbf{D}_j(\mathbf{B})$ is a diagonal matrix containing the *j*th row of $\mathbf{B}$ on its diagonal. If $k_{\mathbf{A}} \geq 1$ and $k_{\mathbf{B}} \geq 1$, then it holds that

$$k_{\mathbf{B}\odot\mathbf{A}} \geq \min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, F)$$

whereas if $k_{\mathbf{A}} = 0$ or $k_{\mathbf{B}} = 0$,

$$k_{\mathbf{B}\odot\mathbf{A}} = 0$$

Lemma 1 is due to Sidiropoulos and Liu (submitted manuscript) which builds on an earlier result by Sidiropoulos *et al.*[10] on the *rank* of the Khatri–Rao product. We include a more compact stand-alone proof below, both for the sake of completeness and because Lemma 1 is key in proving our main result.

Proof

The proof is by contradiction. Let *S* be the *smallest* number of linearly dependent columns that can be drawn from $\mathbf{B} \odot \mathbf{A}$, denoted by $\mathbf{b}_{f_1} \otimes \mathbf{a}_{f_1},\ldots,\mathbf{b}_{f_S} \otimes \mathbf{a}_{f_S}$ (if a collection of linearly dependent columns cannot be found, then, by convention, $S = F + 1$; with this convention, $k_{\mathrm{B}\odot\mathrm{A}} = S - 1$). Then it holds that there exist $\mu_1 \in \mathbb{C},\ldots,\mu_S \in \mathbb{C}$, with $\mu_1 \neq 0,\ldots,\mu_S \neq 0$, such that

$$\mu_1 \mathbf{b}_{f_1} \otimes \mathbf{a}_{f_1} + \cdots + \mu_S \mathbf{b}_{f_S} \otimes \mathbf{a}_{f_S} = \mathbf{0}_{IJ\times 1}$$

or, equivalently,

$$\widetilde{\mathbf{A}}\mathrm{diag}([\mu_1, \cdots, \mu_S])\widetilde{\mathbf{B}}^{\mathrm{T}} = \mathbf{0}_{I\times J} \tag{5}$$

with

$$\widetilde{\mathbf{A}} := [\mathbf{a}_{f_1}, \cdots, \mathbf{a}_{f_S}], \qquad \widetilde{\mathbf{B}} := [\mathbf{b}_{f_1}, \cdots, \mathbf{b}_{f_S}]$$

Invoking Sylvester's inequality,

$$r := \mathrm{rank}\left(\widetilde{\mathbf{A}}\mathrm{diag}([\mu_1, \cdots, \mu_S])\widetilde{\mathbf{B}}^{\mathrm{T}}\right) \geq \mathrm{rank}(\widetilde{\mathbf{A}}) + \mathrm{rank}(\widetilde{\mathbf{B}}) - S$$

However, by definition of k-rank,

$$\mathrm{rank}(\widetilde{\mathbf{A}}) \geq \min(k_{\mathrm{A}}, S), \qquad \mathrm{rank}(\widetilde{\mathbf{B}}) \geq \min(k_{\mathbf{B}}, S)$$

and thus

$$r \geq \min(k_{\mathbf{A}}, S) + \min(k_{\mathbf{B}}, S) - S \tag{6}$$

Notice that $\min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, F) \geq \max(k_{\mathbf{A}}, k_{\mathbf{B}})$, because:

- if $k_{\mathbf{A}} + k_{\mathbf{B}} - 1 > F$, then $\min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, F) = F \geq \max(k_{\mathbf{A}}, k_{\mathbf{B}})$;
- else if $k_{\mathbf{A}} + k_{\mathbf{B}} - 1 \leq F$, then $\min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, F) = k_{\mathbf{A}} + k_{\mathbf{B}} - 1 = \max(k_{\mathbf{A}}, k_{\mathbf{B}}) + \min(k_{\mathbf{A}}, k_{\mathbf{B}}) - 1 \geq \max(k_{\mathbf{A}}, k_{\mathbf{B}})$, since $k_{\mathbf{A}}$, $k_{\mathbf{B}}$ and hence also $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq 1$.

Now consider the following cases for (6):

- if $1 \leq S \leq \min(k_{\mathbf{A}}, k_{\mathbf{B}})$, then (6) gives $r \geq S \geq 1$;
- else if $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) < S < \max(k_{\mathbf{A}}, k_{\mathbf{B}})$, then inequality (6) gives $r \geq \min(k_{\mathbf{A}}, k_{\mathbf{B}}) + S - S = \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq 1$;
- else if $\max(k_{\mathbf{A}}, k_{\mathbf{B}}) \leq S \leq \min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, F)$, then (6) gives $r \geq k_{\mathbf{A}} + k_{\mathbf{B}} - S$; but $S \leq \min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, F)$ implies that $k_{\mathbf{A}} + k_{\mathbf{B}} - 1 \geq S$, i.e. $k_{\mathbf{A}} + k_{\mathbf{B}} - S \geq 1$, and hence $r \geq 1$.

The conclusion is that $r \geq 1$ for $S \leq \min(k_{\mathbf{A}}+k_{\mathbf{B}}-1, F)$; but r is the rank of $\tilde{\mathbf{A}}$ diag $([\mu_1,\ldots,\mu_S])$ $\tilde{\mathbf{B}}^{\mathrm{T}}$, which according to Equation (5) is a zero matrix, hence its rank should be 0. We have therefore arrived at a contradiction, which shows that the smallest number of linearly dependent columns that can be drawn from $\mathbf{B} \odot \mathbf{A}$ must strictly exceed $\min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, F)$. It remains to patch the proof for the case where one or both of $k_{\mathbf{A}}$, $k_{\mathbf{B}}$ is 0. Note that $k_{\mathbf{A}} = 0$ if and only if $\mathbf{A}$ contains at least one identically zero column, in which case the column-wise Kronecker product will have at least one identically zero column, hence its k-rank will be 0. This completes the proof of the lemma. ■

We are now ready to extend Theorem 1 to higher-way arrays.

3. MAIN RESULT

Let us first consider uniqueness of quadrilinear decomposition of four-way arrays.

Theorem 2 (uniqueness of quadrilinear decomposition)

Consider the F-component *quadrilinear* model

$$x_{i,j,l,m} = \sum_{f=1}^{F} a_{i,f} b_{j,f} g_{l,f} h_{m,f}$$

for $i = 1,\ldots,I$, $j = 1,\ldots,J$, $l = 1,\ldots,L$ and $m = 1,\ldots,M$, with $a_{i,f},b_{j,f},g_{l,f},h_{m,f} \in \mathbb{C}$, and suppose that the model is *irreducible* in the sense that $x_{i,j,l,m}$ cannot be represented using fewer than F components (this is equivalent to saying that the four-way array with typical element $x_{i,j,l,m}$ is of rank F). Define an $I \times F$ matrix $\mathbf{A}$ with typical element $\mathbf{A}(i,f) := a_{i,f}$ and similarly $J \times F$, $L \times F$ and $M \times F$ matrices $\mathbf{B}$, $\mathbf{G}$ and $\mathbf{H}$. Given $x_{i,j,l,m}$ for $i = 1,\ldots,I$, $j = 1,\ldots,J$, $l = 1,\ldots,L$ and $m = 1,\ldots,M$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{G}$ and $\mathbf{H}$ are unique up to permutation and complex scaling of columns provided that

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{G}} + k_{\mathbf{H}} \geq 2F + 3$$

Proof

Define a three-way array with typical element

$$y_{i,j,k} := x_{i,j,\lceil \frac{k}{M} \rceil, k-(\lceil \frac{k}{M} \rceil -1)M}$$

for $i = 1,\ldots,I$, $j = 1,\ldots,J$ and $k = 1,\ldots,ML$, where $\lceil\cdot\rceil$ stands for the ceiling operator (smallest integer $\geq$ its argument). $y_{i,j,k}$ is simply the four-way array $x_{i,j,l,m}$ rearranged into an ordinary three-way array by concatenating the third and fourth modes such that the fourth mode is nested within the third mode. Further define an $ML \times F$ matrix $\mathbf{C}$ with typical element

$$c_{k,f} := g_{\lceil \frac{k}{M} \rceil, f} h_{k-(\lceil \frac{k}{M} \rceil -1)M, f}$$

It is then easy to see that

$$\mathbf{C} = \begin{bmatrix} \mathbf{HD}_1(\mathbf{G}) \\ \mathbf{HD}_2(\mathbf{G}) \\ \vdots \\ \mathbf{HD}_L(\mathbf{G}) \end{bmatrix} = \mathbf{G} \odot \mathbf{H}$$

Note that the assumed irreducibility of the quadrilinear model for $x_{i,j,l,m}$ (equivalently, rank of the four-way array) implies that $k_{\mathbf{A}},k_{\mathbf{B}},k_{\mathbf{G}},k_{\mathbf{H}}$ are all ≥ 1. In order to see this, suppose that $k_{\mathbf{A}} = 0$, in which case $\mathbf{A}$ contains at least one column that is identically zero. Let $\mathbf{a}_f$ be this column. Then the fth factor has identically zero contribution to the data $x_{i,j,l,m}$, and thus the model can be represented using $F - 1$ components, which contradicts irreducibility. Then, invoking Lemma 1 regarding the k-rank of a Khatri–Rao product, and Theorem 1 applied to $y_{i,j,k}$, we conclude that $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C} = \mathbf{G} \odot \mathbf{H}$ (and hence $\mathbf{G}$ and $\mathbf{H}$) are unique up to permutation and scaling of columns provided that

$$k_{\mathbf{A}} + k_{\mathbf{B}} + \min(k_{\mathbf{G}} + k_{\mathbf{H}} - 1, F) \geq 2F + 2 \quad (7)$$

We may assume without loss of generality that $k_{\mathbf{A}} \geq k_{\mathbf{B}} \geq k_{\mathbf{G}} \geq k_{\mathbf{H}}$. If $k_{\mathbf{G}} + k_{\mathbf{H}} > F + 1$, then $k_{\mathbf{A}} + k_{\mathbf{B}} \geq k_{\mathbf{G}} + k_{\mathbf{H}} > F + 1 \Rightarrow k_{\mathbf{A}} + k_{\mathbf{B}} \geq F + 2$, and hence

$$k_{\mathbf{A}} + k_{\mathbf{B}} + \min(k_{\mathbf{G}} + k_{\mathbf{H}} - 1, F) \geq F + 2 + F = 2F + 2$$

in which case inequality (7) is always satisfied. If, on the other hand, $k_{\mathbf{G}} + k_{\mathbf{H}} \leq F + 1$, then condition (7) becomes

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{G}} + k_{\mathbf{H}} \geq 2F + 3$$

and the proof is complete. ■

Remark 1

One may wonder if the sufficient condition derived herein is any better than the obvious one, i.e. applying Kruskal's condition to individual three-way subarrays. Consider the following situation: $k_{\mathbf{A}} = k_{\mathbf{B}} = k_{\mathbf{G}} = 2$, $k_{\mathbf{H}} = 3$ and $F = 3$. None of the individual three-way subarrays satisfies Kruskal's condition, because $2 + 2 + 3 = 7 < 2 \times 3 + 2 = 8$; however, $2 + 2 + 2 + 3 = 9 = 2 \times 3 + 3$, and hence the sufficient condition derived herein indicates that the model is unique.

The result further generalizes as follows.

Theorem 3 (uniqueness of multilinear decomposition)

Consider the F-component N-linear model

$$x_{i_1,\ldots,i_N} = \sum_{f=1}^{F} \prod_{\delta=1}^{N} a_{i_\delta,f}^{(\delta)}$$

for $i_\delta = 1, \ldots, I_\delta$ and $\delta = 1, \ldots, N$, with $a_{i_\delta,f}^{(\delta)} \in \mathbb{C}$, and suppose that it is irreducible ($\Leftrightarrow$ the rank of the N-way array with typical element $x_{i_1,\ldots,i_N}$ is F). Then, with obvious notation, given $x_{i_1,\ldots,i_N}$ for $i_\delta = 1, \ldots, I_\delta$ and $\delta = 1, \ldots, N$, $\mathbf{A}^{(\delta)}$ for $\delta = 1, \ldots, N$ are unique up to permutation and scaling of columns provided that

$$\sum_{\delta=1}^{N} k_{\mathbf{A}^{(\delta)}} \geq 2F + (N - 1)$$

Proof

Starting from the $N = 5$ case, assume without loss of generality that $k_{\mathbf{A}^{(1)}} \geq k_{\mathbf{A}^{(2)}} \ldots \geq k_{\mathbf{A}^{(5)}}$, concatenate modes 4 and 5, and apply Lemma 1 and Theorem 2 to conclude that $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(5)}$ are unique up to permutation and scaling of columns provided that

$$k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} + \min(k_{\mathbf{A}^{(4)}} + k_{\mathbf{A}^{(5)}} - 1, F) \geq 2F + 3 \quad (8)$$

If $k_{\mathbf{A}^{(4)}} + k_{\mathbf{A}^{(5)}} > F + 1$, then $k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} \geq k_{\mathbf{A}^{(4)}} + k_{\mathbf{A}^{(5)}} > F + 1$, hence $k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} \geq F + 2$, and thus

$$k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} + \min(k_{\mathbf{A}^{(4)}} + k_{\mathbf{A}^{(5)}} - 1, F) \geq k_{\mathbf{A}^{(1)}} + F + 2 + F \geq 2F + 3$$

since irreducibility implies that $k_{\mathbf{A}^{(1)}} \geq 1$. Therefore inequality (8) is always satisfied for $k_{\mathbf{A}^{(4)}} + k_{\mathbf{A}^{(5)}} > F + 1$. If, on the other hand, $k_{\mathbf{A}^{(4)}} + k_{\mathbf{A}^{(5)}} \leq F + 1$, then condition (8) becomes

$$k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} + k_{\mathbf{A}^{(4)}} + k_{\mathbf{A}^{(5)}} \geq 2F + 4$$

and the proof for $N = 5$ is complete. Similarly, for $N = 6$ one uses the condition for $N = 5$, and so on and so forth ■

Remark 2

Notice how the bilinear case ($N = 2$) is generically excluded: the k-rank of any matrix with F columns is at most F, hence the left-hand side of the inequality is at most $2F$, while the right-hand side is $2F + 1$ for $N = 2$. Also note that 'extra' modes of k-rank 1 (containing collinear profiles) neither contribute to nor take away from uniqueness. As a special case, it is well known that a three-way array where one way is made up of replicates will not provide sufficient variation for a unique decomposition, because the k-rank of the replicate loadings will be 1. This, however, is not the case for four- and higher-way arrays. Even if the loadings in one way are of k-rank 1, unique decomposition is still possible by virtue of Theorem 2, hence replicates may be unproblematically treated as an additional way if so desired.

Remark 3

If all matrices involved are full k-rank (a matrix whose columns are drawn independently from absolutely continuous distributions is full k-rank with probability one), then

$$\sum_{\delta=1}^{N} \min(I_\delta, F) \geq 2F + (N - 1)$$

is sufficient. Notice how going to higher dimensions improves things: for true N-dimensional data sets, $\min_{\delta=1,\ldots,N}(I_\delta) = 2$, meaning that one extra dimension increases the left-hand side by at least 2 but the right-hand side by only 1. It follows that for $2N \geq 2F + (N - 1)$ ($\Leftrightarrow$ in $N \geq 2F - 1$ dimensions), F-component multilinear models are generically unique by sheer dimensionality alone.

4. CONCLUSIONS

Given the complexity in the line of argument behind the proof of Theorem 1, the extension to N-way arrays is refreshingly simple. *A posteriori* the reason is clear: the three-way case is the first instance of multilinearity for which uniqueness holds, and from which uniqueness propagates by virtue of Khatri–Rao structure (induced by unfolding/matricizing a multilinear model) and Lemma 1.

ACKNOWLEDGEMENTS

N. Sidiropoulos gratefully acknowledges support provided by the National Science Foundation through NSF/CAREER CCR-9733540 and NSF/Wireless CCR-9979295, and ONR via the CSTAP program. R. Bro gratefully acknowledges support provided by LMC (Center for Advanced Food Studies) and PMP (Center for Predictive Multivariate Process Analysis) supported by the ministries of research and industry. It is our pleasure to acknowledge the pioneering and inspiring work of J. B.

Kruskal and R. A. Harshman, whose efforts paved our way. The authors would also like to thank the anonymous referees for helpful comments.

APPENDIX: Proof of Theorem 1

In proving Theorem 1, an important lemma is first needed. This lemma is interesting in itself and is stated and proved separately by Kruskal [1] without reference to uniqueness of triple product decomposition. Kruskal proves it assuming real-valued matrices, but, in contrast to the uniqueness of triple product decomposition results, it only involves rank (instead of k-rank) and span arguments for a pair of matrices. It therefore readily generalizes to the complex case.

Lemma 2 (permutation lemma)

Let w ($\mathbf{v}$) denote the number of non-zero elements of $\mathbf{v} \in \mathbb{C}^K$. Given two matrices $\mathbf{C}$ and $\overline{\mathbf{C}}$ with the same number of columns (F), suppose that $\mathbf{C}$ has no identically zero columns, and assume that the following implication holds for all v:

$$\mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \leq F - \mathrm{rank}(\overline{\mathbf{C}}) + 1 \quad \Rightarrow \quad \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{F}) \leq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}})$$

We then have that $\overline{\mathbf{C}} = \mathbf{C\Pi\Lambda}$, where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{\Lambda}$ is a non-singular complex diagonal scaling matrix.

With this lemma we are ready to give the main proof.

Main proof

Consider initially uniqueness of $\mathbf{C}$. Uniqueness of $\mathbf{A}$ and $\mathbf{B}$ will follow automatically from uniqueness of $\mathbf{C}$ by the symmetry of both the complex trilinear model and the k-rank condition. Suppose there exist $\overline{\mathbf{A}} \in \mathbb{C}^{I\times F}$, $\overline{\mathbf{C}} \in \mathbb{C}^{K\times F}$ and $\overline{\mathbf{B}} \in \mathbb{C}^{J\times F}$ such that $\mathbf{X_k} = \mathbf{AD_k}\,(\mathbf{C})\mathbf{B}^{\mathrm{T}} = \overline{\mathbf{A}}\mathbf{D_k}(\overline{\mathbf{C}})\overline{\mathbf{B}}^{\mathrm{T}}$ for $k = 1,2,\ldots,K$. We wish to show that

$$\text{IF } \mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \leq F - \mathrm{rank}(\overline{\mathbf{C}}) + 1$$
$$\text{THEN } \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}})$$

for all $\mathbf{v} \in \mathbb{C}^K$. Thus we are concerned in the following with those $\mathbf{v} \in \mathbb{C}^K$ for which $\mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \leq F - \mathrm{rank}(\overline{\mathbf{C}}) + 1$. We wish to show that for these $\mathbf{v} \in \mathbb{C}^K$ it always holds that $\mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq \mathrm{w}\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}})$. This is investigated under the premise of the k-rank condition (4), and the proof will be accomplished by first deriving an inequality for $\mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}})$ with respect to properties of the model $\mathbf{X}_k = \mathbf{AD}_k\mathbf{(C)B}^{\mathrm{T}}$. This inequality will then be proven to imply the above.

Taking linear combinations $\sum_{k=1}^{K} v_k\mathbf{X}_k$, it follows that

$$\mathbf{A}\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{C})\mathbf{B}^{\mathrm{T}} = \overline{\mathbf{A}}\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}})\overline{\mathbf{B}}^{\mathrm{T}} \quad \forall\mathbf{v} := [\upsilon_1, \cdots, \upsilon_K]^{\mathrm{T}} \in \mathbb{C}^K \tag{9}$$

The rank of a matrix product is always less than or equal to the rank of any factor, and thus

$$\mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) = \mathrm{rank}(\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}})) \geq \mathrm{rank}(\overline{\mathbf{A}}\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}})\overline{\mathbf{B}}^{\mathrm{T}}) = \mathrm{rank}(\mathbf{A}\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{C})\mathbf{B}^{\mathrm{T}}) \tag{10}$$

Let $\gamma := \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C})$ be the number of non-zero elements in $\mathbf{v}^{\mathrm{T}}\mathbf{C}$ and exclude those columns of $\mathbf{A}$ and $\mathbf{B}$ corresponding to the zeros of $\mathbf{v}^{\mathrm{T}}\mathbf{C}$. The resulting truncated matrices $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ have γ columns. Define $\mathbf{t}$

$\in \mathbb{C}^{\gamma\times 1}$ as the corresponding non-zero part of $\mathbf{v}^{\mathrm{T}}\mathbf{C}$. Sylvester's inequality results in

$$\mathrm{rank}(\mathbf{A}\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{C})\mathbf{B}^{\mathrm{T}}) = \mathrm{rank}(\widetilde{\mathbf{A}}\mathrm{diag}(\mathbf{t})\widetilde{\mathbf{B}}^{\mathrm{T}}) \geq \mathrm{rank}(\widetilde{\mathbf{A}}) + \mathrm{rank}(\widetilde{\mathbf{B}}\mathrm{diag}(\mathbf{t})) - \gamma \tag{11}$$

As all elements of $\mathbf{t}$ are non-zero, it follows that

$$\mathrm{rank}(\widetilde{\mathbf{A}}) + \mathrm{rank}(\widetilde{\mathbf{B}}\mathrm{diag}(\mathbf{t})) - \gamma = \mathrm{rank}(\widetilde{\mathbf{A}}) + \mathrm{rank}(\widetilde{\mathbf{B}}) - \gamma \tag{12}$$

The matrix $\tilde{\mathbf{A}}$ has γ columns of the original $\mathbf{A}$ and similarly $\tilde{\mathbf{B}}$ has γ columns of $\mathbf{B}$, and from the definition of k-rank it holds that

$$\mathrm{rank}(\widetilde{\mathbf{A}}) \geq \min(\gamma, k_{\mathbf{A}}), \qquad \mathrm{rank}(\widetilde{\mathbf{B}}) \geq \min(\gamma, k_{\mathbf{B}}) \tag{13}$$

Thus, from (10)–(13),

$$\mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \geq \min(\gamma, k_{\mathbf{A}}) + \min(\gamma, k_{\mathbf{B}}) - \gamma \tag{14}$$

For different values of $\gamma = \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C})$, this implies that

$$\mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \geq \begin{cases} \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}), & \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \quad \text{(i)} \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}), & \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \leq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq \max(k_{\mathbf{A}}, k_{\mathbf{B}}) \quad \text{(ii)} \\ k_{\mathbf{A}} + k_{\mathbf{B}} - \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}), & \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \geq \max(k_{\mathbf{A}}, k_{\mathbf{B}}) \quad \text{(iii)} \end{cases} \tag{15}$$

Observe that in order to establish the implication required by the *permutation lemma*, it suffices to show that the k-rank condition (4) and the condition that $\mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \leq F - \mathrm{rank}\ \overline{\mathbf{C}}) + 1$ jointly exclude the latter two ((ii) and (iii)) possibilities. This will be proven by contradiction.

First we need to prove that (4) implies $\mathrm{rank}(\mathbf{C}) \leq \mathrm{rank}(\overline{\mathbf{C}})$. From (9) it follows that

$$\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}} = \mathbf{0}^{\mathrm{T}}_{1\times F} \quad \Rightarrow \quad \mathbf{A}\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{C})\mathbf{B}^{\mathrm{T}} = \mathbf{0}_{I\times J}$$

Consider $\mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C})$, the number of non-zero elements in $\mathbf{v}^{\mathrm{T}}\mathbf{C}$. We will show that $\mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) = 0$. Suppose the opposite, namely that $1 \leq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq F$. As part of the derivation leading to (15), we have shown that

$$\mathrm{rank}(\mathbf{A}\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{C})\mathbf{B}^{\mathrm{T}}) \geq \begin{cases} \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}), & \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}), & \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \leq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq \max(k_{\mathbf{A}}, k_{\mathbf{B}}) \\ k_{\mathbf{A}} + k_{\mathbf{B}} - \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}), & \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \geq \max(k_{\mathbf{A}}, k_{\mathbf{B}}) \end{cases} \tag{16}$$

Since $k_{\mathbf{B}} \leq F$, condition (4) leads to

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2(F+1) \quad \Rightarrow \quad \begin{cases} k_{\mathbf{A}} + k_{\mathbf{B}} \geq F + 2 > F + 1 \\ \min(k_{\mathbf{A}}, k_{\mathbf{B}}) \geq 2 \end{cases}$$

Equation (16) therefore shows that $\mathrm{rank}(\mathbf{A}\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{C})\mathbf{B}^{\mathrm{T}}) \geq 1$ if $1 \leq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq F$. However, $\mathbf{A}\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{C})\mathbf{B}^{\mathrm{T}} = \mathbf{0}_{I\times J}$, so its rank should be $\mathbf{0}$. Therefore it holds that

$$\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}} = \mathbf{0}_{1\times F} \quad \Rightarrow \quad \mathbf{v}^{\mathrm{T}}\mathbf{C} = \mathbf{0}_{1\times F} \quad \Rightarrow \quad \mathrm{rank}(\mathbf{C}) \leq \mathrm{rank}(\overline{\mathbf{C}})$$

Recall the assumption of the *permutation lemma* and work leftwards:

$$F - k_{\mathbf{C}} + 1 \geq F - \mathrm{rank}(\mathbf{C}) + 1 \geq F - \mathrm{rank}(\overline{\mathbf{C}}) + 1 \geq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \tag{17}$$

where the leftmost inequality follows from the fact that k-rank $\leq$ rank. From the k-rank condition (4) it follows that

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2(F+1) \quad \Rightarrow \quad k_{\mathbf{A}} + k_{\mathbf{B}} - F - 1 \geq F - k_{\mathbf{C}} + 1 \tag{18}$$

The inequalities (17) and (18) are combined to give the second key inequality

$$k_{\mathbf{A}} + k_{\mathbf{B}} - F - 1 \geq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \tag{19}$$

Consider now inequalities (15) and (19) jointly. Suppose that the third leg (iii) of (15) is in effect, i.e. $\mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \geq \max(k_{\mathbf{A}}, k_{\mathbf{B}})$; then

$$k_{\mathbf{A}} + k_{\mathbf{B}} - F - 1 \geq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \geq k_{\mathbf{A}} + k_{\mathbf{B}} - \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C})$$

which is not possible, because $\mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq F$. Similarly, suppose that the second leg (ii) of (15) is in effect, i.e. $\min(k_{\mathbf{A}}, k_{\mathbf{B}}) \leq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq \max(k_{\mathbf{A}}, k_{\mathbf{B}})$; then we obtain

$$k_{\mathbf{A}} + k_{\mathbf{B}} - F - 1 \geq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \geq \min(k_{\mathbf{A}}, k_{\mathbf{B}})$$

which is impossible, as

$$k_{\mathbf{A}} + k_{\mathbf{B}} - F - 1 = \min(k_{\mathbf{A}}, k_{\mathbf{B}}) + \max(k_{\mathbf{A}}, k_{\mathbf{B}}) - F - 1 \leq \min(k_{\mathbf{A}}, k_{\mathbf{B}}) - 1$$

because $\max(k_{\mathbf{A}}, k_{\mathbf{B}}) \leq F$. The only remaining option is leg (i), i.e. $\mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C}) \leq \min(k_{\mathbf{A}}, k_{\mathbf{B}})$, and thus

$$\mathrm{w}(\mathbf{v}^{\mathrm{T}}\overline{\mathbf{C}}) \geq \mathrm{w}(\mathbf{v}^{\mathrm{T}}\mathbf{C})$$

which is exactly what is required by the *permutation lemma* (note the trivial fact that $k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2(F+1) \Rightarrow \min(k_{\mathbf{A}}, k_{\mathbf{B}}, k_{\mathbf{C}}) \geq 2$, which implies that no one of **A**, **B**, **C** has all-zero columns). Thus it has been shown that $\overline{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}$, where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{\Lambda}$ is a non-singular complex diagonal scaling matrix. The trilinear model is completely symmetric in the sense that any one of the three matrices can be put in the middle of the decomposition. The k-rank condition is also symmetric (*sum* of k-ranks). It therefore follows that **A** and **B** are also unique up to permutation and scale. This completes the proof. ■

Remark 4

It can be shown that the permutation is common to all three matrices, and the product of the respective scales is identity; however, we skip this for space considerations.

REFERENCES

1. Kruskal JB. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Appl.* 1977; **18**: 95–138.

2. Kruskal JB. Rank decomposition, and uniqueness for 3-way and N-way arrays. In *Multiway Data Analysis*, Coppi R, Bolasco S (eds). Elsevier/North-Holland: Amsterdam, 1988; 8–18.
3. Harshman RA. Foundations of the PARAFAC procedure: model and conditions for an 'explanatory' multi-mode factor analysis. *UCLA Working Papers Phonet.* 1970; **16**: 1–84.
4. Harshman RA. Determination and proof of minimum uniqueness conditions for PARAFAC. *UCLA Working Papers Phonet.* 1972; **22**: 111–117.
5. Cattell RB. Parallel proportional profiles and other principles for determining the choice of factors by rotation. *Psychometrika* 1944; **9**: 267–283.
6. Carroll JD, Chang J. Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart–Young' decomposition. *Psychometrika* 1970; **35**: 283–319.
7. Schonemann PH. An algebraic solution for a class of subjective metrics models. *Psychometrika* 1972; **37**: 441.
8. Harshman RA, Lundy ME. The PARAFAC model for three-way factor analysis and multidimensional scaling. In *Research Methods for Multimode Data Analysis*, Law HG, Snyder Jr CW, Hattie J, McDonald RP (eds). Praeger: New York, 1984; 122–215.
9. Sidiropoulos ND, Giannakis GB, Bro R. Blind PARAFAC receivers for DS-CDMA systems. *IEEE Trans. Signal Processing* 2000; **48(3)**: 810–823.
10. Sidiropoulos ND, Bro R, Giannakis GB. Parallel factor analysis in sensor array processing. *IEEE Trans. Signal Processing* in press.

