View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Kentucky

University of Kentucky
UKnowledge

University of Kentucky Doctoral Dissertations

Graduate School

2006

KHATRI-RAO PRODUCTS AND CONDITIONS FOR THE UNIQUENESS OF PARAFAC SOLUTIONS FOR IxJxK ARRAYS

Heather Michele Clyburn Bush
University of Kentucky, hbush@ms.uky.edu

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation

Bush, Heather Michele Clyburn, "KHATRI-RAO PRODUCTS AND CONDITIONS FOR THE UNIQUENESS OF PARAFAC SOLUTIONS FOR IxJxK ARRAYS" (2006). *University of Kentucky Doctoral Dissertations*. 462.
https://uknowledge.uky.edu/gradschool_diss/462

This Dissertation is brought to you for free and open access by the Graduate School at UKnowledge. It has been accepted for inclusion in University of Kentucky Doctoral Dissertations by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

ABSTRACT OF DISSERTATION

Heather Michele Clyburn Bush

The Graduate School

University of Kentucky

2006

KHATRI-RAO PRODUCTS AND CONDITIONS FOR THE UNIQUENESS OF PARAFAC SOLUTIONS FOR IxJxK ARRAYS

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky

By
Heather Michele Clyburn Bush

Lexington, KY

Director: Dr. William S. Rayens, Associate Professor of Statistics

Lexington, KY

2006

Copyright © Heather Michele Clyburn Bush 2006

ABSTRACT OF DISSERTATION

KHATRI-RAO PRODUCTS AND CONDITIONS FOR THE UNIQUENESS OF PARAFAC SOLUTIONS FOR IxJxK ARRAYS

One of the differentiating features of PARAFAC decompositions is that, under certain conditions, unique solutions are possible. The search for uniqueness conditions for the PARAFAC Decomposition has a limited past, spanning only three decades. The complex structure of the problem and the need for tensor algebras or other similarly abstract characterizations provided a roadblock to the development of uniqueness conditions. Theoretically, the PARAFAC decomposition surpasses its bilinear counterparts in that it is possible to obtain solutions that do not suffer from the rotational problem. However, not all PARAFAC solutions will be constrained sufficiently so that the resulting decomposition is unique. The work of Kruskal, 1977, provides the most in depth investigation into the conditions for uniqueness, so much so that many have assumed, without formal proof, that his sufficient conditions were also necessary.

Aided by the introduction of Khatri-Rao products to represent the PARAFAC decomposition, ten Berge and Sidiropoulos (2002) used the column spaces of Khatri-Rao products to provide the first evidence for countering the claim of necessity, identifying PARAFAC decompositions that were unique when Kruskal's condition was not met. Moreover, ten Berge and Sidiropoulos conjectured that, with additional k-rank restrictions, a class of decompositions could be formed where Kruskal's condition would be necessary and sufficient.

Unfortunately, the column space argument of ten Berge and Sidiropoulos was limited in its application and failed to provide an explanation of why uniqueness occurred. On the other hand, the use of orthogonal complement spaces provided an alternative approach to evaluate uniqueness that would provide a much richer return than the use of column spaces for the investigation of uniqueness. The Orthogonal Complement Space Approach (OCSA), adopted here, would provide: (1) the answers to lingering questions about the occurrence of uniqueness, (2) evidence that necessity would require more than a restriction on k-rank, and (3) an approach that could be extended to cases beyond those investigated by ten Berge and Sidiropoulos.

KEYWORDS: Khatri-Rao products, Orthogonal Complement Spaces, Multilinear, PARAFAC, Uniqueness

Heather Michele Clyburn Bush

March 10, 2006

KHATRI-RAO PRODUCTS AND CONDITIONS FOR THE UNIQUENESS OF PARAFAC SOLUTIONS FOR IxJxK ARRAYS

By

Heather Michele Clyburn Bush

William S. Rayens
(Director of Dissertation)

Arnold Stromberg
(Director of Graduate Studies)

March 10, 2006

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgements.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

<u>Name</u> <u>Date</u>

DISSERTATION

Heather Michele Clyburn Bush

The Graduate School

University of Kentucky

2006

KHATRI-RAO PRODUCTS AND CONDITIONS FOR THE UNIQUENESS OF PARAFAC SOLUTIONS FOR IxJxK ARRAYS

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky

By
Heather Michele Clyburn Bush

Lexington, KY

Director: Dr. William S. Rayens, Associate Professor of Statistics

Lexington, KY

2006

Copyright © Heather Michele Clyburn Bush 2006

To my husband, who always encourages me to chase my dreams.
Thanks for supporting me every step of the way – I never walked alone.

ACKNOWLEDGEMENTS

This work would not have been possible had it not been for the patience and guidance of my advisor, William S. Rayens. Thank you, for the time spent reviewing, the discussions, the hours in chalk dust, and for being there when I was excited as well as when I was frustrated. It is evident that you give so much of yourself in all you do, and I am very appreciative of how much you gave to me. Thanks, too, for allowing me the freedom to explore, and providing me a research atmosphere where I could develop the confidence to trust my own ideas.

I would also like to express my deepest gratitude for the contribution of my committee members: William Griffith, Arnold Stromberg, Arne Bathke, Ren-Cang Li, and Anders Andersen. Thank you for your encouragement, your time, and your support. Additional thanks to Richard Harshman, Rasmus Bro, Jos ten Berge, and Nikos Sidiropoulos who I met at TRICAP 2003. Thank you for your work in the area of multilinear methods and for answering questions and providing insights.

Additionally, I would like to thank the Department of Statistics and the College of Public Health at the University of Kentucky, for the financial support during my graduate career, with special thanks to Claudia Hopenhayn and those I worked with during my research and teaching assistantships. Beyond financial support, I thank you for the many ways in which you enriched my graduate experience.

Finally, to family, faculty, and friends who continually offered support, advice, and inspiration, I give the utmost thanks. Thank you to my parents, husband, and family for patiently listening to my trials and celebrations – for letting me know that you were always there for me. A special thanks belongs to the faculty in the Department of Statistics who became more than professors and are responsible for any success that I enjoy. Thank you, Dr. Wood, Dr. Griffith, Kert, Arny, Arne, Ziyad, and Bill. This has been a life changing experience, and I will be forever grateful to all who cheered me on, made me work harder, and gave me the encouragement I needed to make mistakes and see success.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. MULTILINEAR MODELS

1.1 An Introduction

When dealing with a set of observations for several variables, researchers in psychometrics and chemometrics have used Principal Component Analysis (PCA), Multidimensional Scaling (MDS) or Factor Analysis to attempt to identify underlying characteristics.

Multilinear models originated in psychometrics, a field of study concerned with psychological measurement. In particular, research in psychometrics involves measurements of knowledge, abilities, attitudes, and personality traits. Traditionally, the instruments used to provide such measurements involve many variables, variables which are analyzed in order to discover underlying characteristics that might better represent the data relationships of particular cases.

A simple example of a problem in psychometrics would be to determine what underlying characteristics contributed to a consumer preferring a particular product, bread for example. Several attributes may be involved in the preference of bread and it would be difficult to capture these in a single question. Therefore, it might be of interest to collect data on certain sensory attributes such as the texture, taste, density, color, and look of bread for various types of breads. For each consumer, the goal would be to explore the underlying characteristics for preferences in bread. Such a problem is known as *bilinear*, where two *w*ays (bread types and sensory attributes) are explored. However, the results obtained from evaluating this bilinear problem would only be useful for describing the underlying factors in preferring bread for an individual consumer. If this same "experiment" is carried out on several consumers or raters, the problem adds an additional *way* and becomes a *multilinear* problem.

In addition to psychometrics, the use of multilinear models has found increasing popularity in the area of chemometrics. Although the principles are the same, the application of multilinear models in chemometrics has a different perspective than that of psychometrics. In chemometrics, the multilinear data structure emerges naturally from physical science.

Fluorescence spectroscopy offers a primary example of the applicability of the multilinear model. By definition, fluorescence spectroscopy measures the absorption or emission of particles from a sample as a function of wavelengths of energy (Leurgans and Ross, 1992). Multilinear models are ideal for these scenarios as the data are known to naturally follow a multilinear (Parallel Factor) model. In these types of problems an analyte at different concentrations is evaluated (Smilde, Bro, Geladi, 2004). When the analyte is excited at different wavelengths, particular wavelengths of spectra are emitted. When chemist work with several samples or different concentrations, the typical fluorescence model is equivalent to the multilinear (Parallel Factor) decomposition. Thus, the three *ways* of spectroscopic data might be wavelengths, emission or absorbance, and concentration. The data structure found in chemometrics suggests PARAFAC modeling, and applications of these techniques can provide chemists with unique descriptions of the underlying chemical make-up of particular samples.

When dealing with bilinear data, a set of observations for several variables, researchers in psychometrics and chemometrics have used Principal Component Analysis (PCA), Multidimensional Scaling (MDS) or factor analysis to attempt to identify underlying characteristics. However, it is well known that these methods do not provide decompositions so that one can identify a unique set of underlying characteristics. In other words, taking for example bread preferences, two researchers could analyze a consumer’s opinions on bread preferences and obtain two equivalent decompositions with completely different interpretations about what characteristics lead to bread choice. This is commonly referred to as the *rotational problem*. As will be discussed in the following sections, the use of certain types of multilinear models avoids the *rotational problem* of traditional psychometric and chemometric analysis tools.

1.2 The Bilinear Problem

While it is well known that principal components analysis suffers from a problem with rotational invariance, insofar as the concept of “plane of closest fit” is uniquely defined, while the basis for that plane is not, this is not what is typically meant by the *rotation problem* when the term is used in the context of multilinear uniqueness research. In fact, it could even be conjectured that there is perhaps a fair amount of confusion on this point in the literature. The aim of this section is to dispel some of that confusion. Indeed, it is important that the nature of the much maligned bilinear inadequacy be articulated since the problem being addressed in this dissertation is, in essence, the analogous problem for multilinear decompositions. While there are many ways to frame this bilinear problem, being particularly cognizant of the forthcoming discussion, the perspective presented below was chosen.

The singular value decomposition of $\mathbf{X}_{nxp}$ can be written as $\mathbf{X}_{nxp} = \mathbf{A}_{nxp}\mathbf{C}_{pxp}\left[\mathbf{B}_{pxp}\right]^t \equiv \hat{\mathbf{A}}_{nxp}\left[\mathbf{B}_{pxp}\right]^t$ where $\mathbf{A}^t\mathbf{A} = \mathbf{B}^t\mathbf{B} = \mathbf{I}_{pxp}$, and $\mathbf{C}_{pxp} = \text{diag}(\mathbf{c}_{px1})$. Typically, determining the component directions matrix $\mathbf{B}$ is the goal, and this is commonly said to be the solution to the principal components problem. The connections to PCA are not hard to see.

Assuming $\mathbf{X}_{nxp}$ is mean centered, the sample covariance matrix, $\mathbf{S}_{pxp}$, associated with $\mathbf{X}_{nxp}$ is $\mathbf{S}_{pxp} = \mathbf{X}^t\mathbf{X} = \mathbf{B}\mathbf{C}^t\mathbf{A}^t\mathbf{A}\mathbf{C}\mathbf{B}^t = \mathbf{B}\mathbf{C}^2\mathbf{B}^t$, which implies, of course, that $\mathbf{S}\mathbf{B} = \mathbf{B}\mathbf{C}^2$. Hence, the diagonal entries of $\mathbf{C}^2$ (the square of the singular values) are the eigenvalues of $\mathbf{S}$ and the columns of $\mathbf{B}$ the corresponding eigenvectors and we have the traditional solutions to the PCA problem. Notice also, that $\mathbf{X}\mathbf{B} = \mathbf{A}\mathbf{C}\mathbf{B}^t\mathbf{B} = \mathbf{A}\mathbf{C} \equiv \hat{\mathbf{A}}$ fills the classical role of the score matrix.

The rotation problem associated with the decomposition of $\mathbf{X}_{nxp}$ can be better understood by noting that $\mathbf{X}_{nxp}$ can also be written as follows:

(1.1)
$$\begin{aligned}\mathbf{X}_{nxp} &= \mathbf{A}_{nxp}\mathbf{C}_{pxp}\left[\mathbf{B}_{pxp}\right]^t \\ &= \left[\mathbf{A}_{nxp}\mathbf{C}_{pxp}\right]\left[\mathbf{B}_{pxp}\right]^t \\ &= \left[\mathbf{A}_{nxp}\mathbf{C}_{pxp}\right]\mathbf{T}_{pxp}\mathbf{T}^{-1}_{pxp}\left[\mathbf{B}_{pxp}\right]^t \\ &= \left[\mathbf{ACT}\right]\left[\mathbf{B}\left(\mathbf{T}^t\right)^{-1}\right]^t \\ &= \left[\hat{\mathbf{A}}\mathbf{T}\right]\left[\mathbf{B}\mathbf{T}_B\right]^t\end{aligned}$$

for *any* invertible matrix $\mathbf{T}_{pxp}$. Hence, an alternative decomposition of $\mathbf{X}_{nxp}$ with transformed $\hat{\mathbf{A}}$ and $\mathbf{B}$ is obtained, and, therefore, $\mathbf{BT}_B$ emerges as an equally good candidate for the component directions. Of course, the decomposition of the covariance matrix $\mathbf{S}_{pxp}$ is trivially the same as before in spite of the fact that there are uncountably infinitely many ways to decompose $\mathbf{X}_{nxp}$. So, on the one hand, strictly speaking, there is not a problem with PCA in this context. However, at issue is the decomposition of $\mathbf{X}_{nxp}$, and the clear ambiguity associated with the solution to that bilinear problem. In the literature on multilinear uniqueness issues, the PCA rotation problem is somewhat inappropriately, or at the very least "loosely", used to refer to this decomposition ambiguity.

It is also very instructive to look at the nature of the alternative decomposition given in (1.1). Indeed a careful look will help tremendously with an understanding of logic of the subsequent work in this dissertation, as well as provide a nice preview of the history of this problem. In essence, $\mathbf{X}_{nxp}$ has been decomposed as follows:

(1.2) $$\mathbf{X}_{nxp} = \left[\mathbf{ACT}\right]\left[\mathbf{B}\left(\mathbf{T}^t\right)^{-1}\right]^t = \left[\mathbf{A}\left(\text{diag}(\mathbf{c})\right)\mathbf{T}\right]\left[\mathbf{BT}_B\right]^t.$$

Hence, nontrivial transformations have been applied to (the rows of) $\mathbf{B}$ and $\mathbf{C} = \left(\text{diag}(\mathbf{c})\right)$, and an equivalent decomposition resulted. With multilinear models there is more than one $\mathbf{X}_{nxp}$ matrix available, more than one rater, to use the consumer preferences example given earlier. Therefore, the analogous problem in multilinear analysis is to decompose all of the matrices:

(1.3) $$\mathbf{X}_k = \left[\mathbf{A}\left(\text{diag}(\mathbf{c}_k)\right)\right]\left[\mathbf{B}\right]^t, \text{ where } \mathbf{C} = \begin{pmatrix}\mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_k\end{pmatrix}.$$

One could argue that that there are two ways to frame the "uniqueness" issue in the context of (1.3). Perhaps the most direct analogy would be to place transformations on the rows of $\mathbf{A}$, $\mathbf{B}$, *and* $\left(\text{diag}(\mathbf{c}_k)\right)$ so that the transformations cancel as in bilinear decompositions:

(1.4) $$\mathbf{X}_k = \left[\mathbf{AT}_A\right]\left[\mathbf{T}_A^{-1}\left(\text{diag}(\mathbf{c}_k)\right)\left(\mathbf{T}_B^t\right)^{-1}\right]\left[\mathbf{BT}_B\right]^t$$

and ask the question as to whether such transformations exist, and under what conditions it can be shown that the only such transformations that exist are somehow only “trivial”. Likewise, one could argue that a rational analogy to (1.2) would be to first absorb a transformation on the $\mathbf{c}_k$ term:

(1.5) $$\mathbf{X}_k = [\mathbf{AT}_A][(\mathrm{diag}(\mathbf{c}_k\mathbf{T}_C))][\mathbf{BT}_B]^t .$$

Early work in the area of uniqueness would incorporate these representations of $\mathbf{X}_k$ and will be discussed in later sections. However, it turns out that the key to uniqueness would lie with (1.5) while those who employed (1.4) would be led down the wrong path.

Although there are certain multilinear models which offer the possibility of uniqueness, unlike bilinear methods where hopes of uniqueness are abandoned, this is not the only, or even main, benefit. One of the attractions of multilinear models is that it is possible to explore data as it naturally occurs, in a higher dimensional structure. Two main decompositions are available to those exploring higher dimensional data through multilinear modes: PARAFAC and Tucker3.

The main advantage of the PARAFAC decomposition is the possibility of uniqueness while the main advantage of the Tucker3 decomposition is the ability to describe rich interactions. PARAFAC decompositions, under certain conditions, can provide solutions that are unique or do not suffer from the rotational problems of its bilinear predecessors. However, not all multilinear data can be decomposed using PARAFAC. In some cases, it is necessary to employ the Tucker3 decompositions, which cannot claim uniqueness. Although the focus of this dissertation will be to investigate the PARAFAC decomposition and the constraints that impose uniqueness, for completeness, the Tucker3 models will be described in following the introductory sections.

1.3 Theoretical Background

Extensions of and analogies to the well-known bilinear paradigms have created a substantial literature on so-called *multiway structure-seeking* methods. The conceptual similarity to bilinear methods is indicated in the schematic below, wherein a cube of data is decomposed into two components, each of which consists of three “profiles” (e.g. raters, bread type, and sensory attributes), which, taken as triples, are assumed to characterize the cube (Figure 1.1).

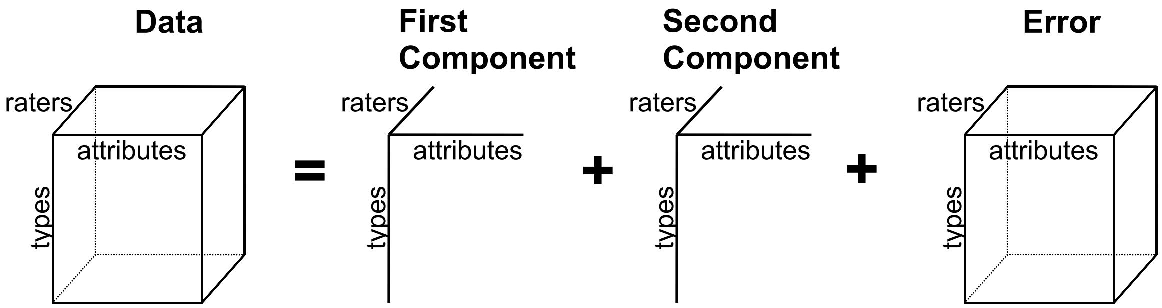

Figure 1.1 Example Trilinear Schematic

As was mentioned above, such techniques originated in psychology, but some of the most important contributions to understanding and application have come from within chemometrics, in particular from within florescence spectroscopy (see e.g. Bro, 1997; Burdick, 1995; Leurgans & Ross, 1992; Mitchell and Burdick, 1994; Rayens & Mitchell, 1997; Sanchez & Kowalski, 1988, 1990; Wold, et. al., 1987; Bro, 1999; Bro & Heimdal, 1996). An extensive reference list is available courtesy of Professor Rasmus Bro of the Chemometrics Group at the Royal Veterinary and Agricultural University in Denmark on the web at **http://www.optimax.dk/chemobro.html** and from the Three-Mode Company at **http://www.fsw.leidenuniv.nl/~kroonenb**. Although the field is far from unified, with several different multiway models and even different methods of implementing those models, the literature is growing in statistical sophistication and the successes of many applications are undeniable and intriguing.

1.3.1 Tensor Algebras

Traditionally, the theoretical description of multilinear models depended on a basic understanding of tensor algebras. Just as linear algebra provided clear definitions for bilinear methods, tensor algebra offered the means to clearly describe the structure of trilinear and higher-way models. However, tensor algebras are more abstract and lack the familiarity of linear algebra. Ironically, although tensor algebras provided clear definitions of higher-way models, users of multilinear models were slow to embrace this mathematical language that helped to provide a framework for interpreting multilinear models. The hesitancy to adopt the tensor algebra paradigm is in part responsible for the lack of progress in the investigation of properties and uniqueness characteristics with these decompositions.

Recent developments, however, have allowed the characterization of multilinear models without the use of tensor algebras. Although these newer representations will be used throughout the latter chapters of the dissertation, tensor algebras will be used to provide the first definitions and descriptions of multilinear models. Consistent with Burdick (1995), the description of these models is given elegantly with tensor algebras, providing details that may be missed with other techniques. Additionally, the use of tensor algebras to define multilinear models is fundamental to understanding the history and appreciating the progress of research in multilinear methods.

Hence, the technical overview of multiway methods will be presented by discussing trilinear methods from this abstract perspective. Extensions to higher-way arrays are straightforward but notationally cumbersome. Burdick's notation is used in the following.

First, a definition of the tensor product of two vectors and a tensor product of a vector and a matrix is given. Arrangement of coordinates in the definition is somewhat arbitrary. However, the following will suffice:

Definition 1.1 Let **a** be a vector in $\Re^{I}$ and **b** a vector in $\Re^{J}$.

- A *tensor product* of **a** and **b** is given by $\mathbf{a} \otimes \mathbf{b} = \mathbf{ab^t}$. Let $\mathbf{c} = (c_{11} \quad \dots \quad c_{1K})^t$ be a vector in $\Re^{K}$ and **X** be an $I \times J$ matrix.
- A *tensor product* of **X** and **c** is given by $\mathbf{X} \otimes \mathbf{c} = \left(c_{11}\mathbf{X} \quad c_{12}\mathbf{X} \quad \cdots \quad c_{1K}\mathbf{X}\right)_{I \times JK}$.
- If, in fact, $\mathbf{X} = \mathbf{a} \otimes \mathbf{b}$, then $\mathbf{X} \otimes \mathbf{c} = \left(c_{11}\mathbf{ab^t} \quad c_{12}\mathbf{ab^t} \quad \cdots \quad c_{1K}\mathbf{ab^t}\right)_{I \times JK}$

Definition 1.2 Let $U \subset \Re^I$ and $V \subset \Re^J$. A *tensor product of U and V,* denoted by $U \otimes V$, is the vector space consisting of all linear combinations of $\mathbf{a} \otimes \mathbf{b}$ where $\mathbf{a} \in U$ and $\mathbf{b} \in V$.

This idea is easily extendible to more than two vector spaces. It is not hard to check that if dim (U) = R and dim (V) = S, then dim ($U \otimes V$) = RS.

Typically, a bilinear errors-in-variables model employed to extract structure from $\mathbf{X}_{IxJ}$ has the form of $\mathbf{X}_{IxJ} = \mathbf{S}_{IxJ} + \mathbf{N}_{IxJ}$, interpreted as a signal matrix added to a noise, or error matrix. The issue becomes how one decides to model the structure in the signal matrix. Cosmetically different perspectives lead to identical bilinear models, but to very different multilinear models. To see this, assume that the **S** can be written as the sum of R rank one matrices. That is, one might assume that there exist vectors $\{\mathbf{a}_r\} \subset \Re^I$ and $\{\mathbf{b}_r\} \subset \Re^J$ such that:

(1.6) $$\mathbf{X} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r$$

If both $\{\mathbf{a}_r\} \subset \Re^I$ and $\{\mathbf{b}_r\} \subset \Re^J$ are linearly independent then **X** will have rank R. Similarly, one might adopt the perspective that there exist subspaces $U \subset \Re^I$ and $V \subset \Re^J$, with dim (U) = dim (V) = R, such that:

(1.7) $$\mathbf{X} \in U \otimes V$$

Decompositions (1.6) and (1.7) are equivalent, and each suffers equally from a well-known lack of uniqueness. For instance in (1.6), there are uncountably infinitely many vectors $\mathbf{a}_r$ and $\mathbf{b}_r$ that can describe **X** equally well from the point of view of decomposition. Hence, interpretations of the vectors $\mathbf{a}_r$ and $\mathbf{b}_r$ become as much an act of faith as an analytical exercise. This is analogous to the rotational problem defined in the previous section.

When (1.6) is extended to higher-way data structures, the so-called PARAllel FACtor model (*PARAFAC*) emerges. Throughout the dissertation, the noise matrix will be ignored when representing the PARAFAC solution. For the purposes of discussing uniqueness, the noise matrix is typically not included, and the PARAFAC solutions are presented as if the data array could be decomposed without error.

Definition 1.3 The *PARAFAC model* presumes there exist vectors $\{\mathbf{a}_r\} \subset \Re^I$, $\{\mathbf{b}_r\} \subset \Re^J$, and $\{\mathbf{c}_r\} \subset \Re^K$ such that:

(1.8) $$\boldsymbol{X} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r$$, where $\boldsymbol{X}$ represents the array.

One can think of these R tensor algebra products as representing the relative influences of R underlying latent characteristics that define the array. For instance, if the array was structured according to the bread example as bread types, sensory attributes, and raters, then the vectors $\mathbf{a}_r$ and $\mathbf{b}_r$ represent the relative influence of factor r on the type and rater modes, while the vector $\mathbf{c}_r$ contains the weights of the r^{th} factor for each of the *k* sensory attributes. From (1.8) it can easily be seen that each element of $\boldsymbol{X}$ can be written as the sum of the relative influences of each of the factors on the i^{th} bread type, the j^{th} rater, and the k^{th} sensory attribute, or $x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr}$. Notice that for each element, regardless of the attribute, $a_{ir} b_{jr}$ represents the contribution of the r^{th} factor to the i^{th} bread type and the j^{th} rater. For the k^{th} sensory attribute, this product is multiplied by c_{kr}. Thus, the whole influence of a factor on raters and types is proportionally adjusted for the influence on attributes. In other words, the influence of the factors is adjusted in parallel proportion by the elements of $\mathbf{c}_r$. The extension of (1.7) defines the so-called Tucker3 model.

Definition 1.4

The *Tucker3 model* presumes that there exist subspaces $U \subset \Re^I$, $V \subset \Re^J$, and $W \subset \Re^K$, with dim (*U*) = R_U, dim (*V*) = R_V, and dim (*W*) = R_W, such that $\boldsymbol{X} \in U \otimes V \otimes W$.

Notice, these two models are quite different. In particular, under rather general conditions, the PARAFAC model lays claim to a useful form of uniqueness (Kruskal, 1989), while the Tucker3 model cannot (Comments by Burdick on paper by Leurgans & Ross, 1992).

This representation of **S** is useful in the more complex case when $R_U = M$ factors can be extracted from the bread type mode, $R_V = P$ factors exist in the rater mode, and $R_W = Q$ factors can be found in the sensory attributes mode. As in the case of parallel factors, the factors in each of the modes contribute a relative influence on the elements. Unlike parallel factors, however, the existence of factors within each of the modes will necessarily require that the factors be interrelated. Continuing the representation as before, where $\mathbf{a}_m \in U$, $\mathbf{b}_p \in V$, and $\mathbf{c}_q \in W$. The vector $\mathbf{a}_m$ represents the relative influence of the m^{th} bread type factor on the elements of the types mode, $\mathbf{b}_p$ corresponds to the relative influence of the p^{th} rater on the elements of the raters mode, while the vector $\mathbf{c}_q$ contains the weights of the q^{th} sensory attribute factor for each of the *k* sensory attributes.

From (1.9), $\boldsymbol{X}$ is any linear combination of vectors of the form $\mathbf{a}_m \otimes \mathbf{b}_p \otimes \mathbf{c}_q$, and an element from this array can be written as $x_{ijk} = \sum_{m=1}^{M}\sum_{p=1}^{P}\sum_{q=1}^{Q} a_{im} b_{jp} c_{qk} g_{mpq}$, where the coefficient g_{mpq} represents the relative weights of the relationships among the factors. In this form, it is obvious that the whole influence of a particular factor will not merely change proportionally as sensory attributes vary, but will be dependent on the influences of factors from each of the other two modes. In general there are many analogous extensions that lead to slightly different mixed factors models.

1.3.2 The Matrix Representation of PARAFAC

Although the representation of PARAFAC using tensor algebras is beneficial from a conceptual perspective, most of the work in PARAFAC and uniqueness has used what is known as the matrix representation. Consider again the tensor representation, $\boldsymbol{X}_{I\times J\times K} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r$, where $\{\mathbf{a}_r\} \subset \Re^I$, $\{\mathbf{b}_r\} \subset \Re^J$, and $\{\mathbf{c}_r\} \subset \Re^K$.

Definition 1.5 The relative influences of the r^{th} factor for a particular way form a vector of weights. These weights can be referred to as *loadings* as a way to describe the variations of relative influence from one point to the next.

Definition 1.6 If the vectors of loadings are combined to form a matrix, the matrix is called a *loading matrix*.

Therefore, returning to the bread example, combining the loadings for the bread types could result in the loading matrix $\mathbf{A} = \{\mathbf{a}_1 \ \cdots \ \mathbf{a}_R\}$. Likewise, the vectors of $\mathbf{B} = \{\mathbf{b}_1 \ \cdots \ \mathbf{b}_R\}$ describe the variations of relative influence from one rater to another, and the vectors of $\mathbf{C} = \{\mathbf{c}_1 \ \cdots \ \mathbf{c}_R\}$ describe the variations of relative influence from one sensory attribute to the next.

If one thinks of $\boldsymbol{X}$ as composed of K I×J matrix slabs, $\mathbf{X}_k$, then it is easy to show that this representation of $\boldsymbol{X}$ is equivalent to the presumption that $\mathbf{X}_k = \mathbf{A}\mathbf{C}_k\mathbf{B}^t$, for $k=1,\ldots,K$, where $\mathbf{C}_k = \text{diag}(c_{k1} \ \cdots \ c_{kR})$ (Figure 1.2). The loading matrices $\mathbf{A}$ and $\mathbf{B}$ are common to every slab $\mathbf{X}_k$, while the diagonal $\mathbf{C}_k$ matrix is specific to each $\mathbf{X}_k$ slab. Thus, the $\mathbf{C}_k$ is known as the *core matrix*.

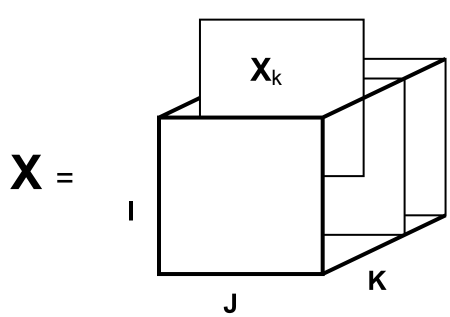

Figure 1.2 Matrix Representation of PARAFAC

An R-component decomposition amounts to the specification of the loading matrices, $\mathbf{A}_{I\times R}$, $\mathbf{B}_{J\times R}$ and $\mathbf{C}_{K\times R}$. Intuitively, using PCA-like language, one can think of the columns of **A** as the common *scores*, the columns of **B** as the common *directions*, and the columns of **C** as the relative weights that distinguish the slabs in the **C** direction. It is important to note that the model *presumes* that such matrices exist, that is, that the specified decomposition is possible. Typically, uniqueness results in the literature have been derived in the presence of this presumption.

1.4 An Example

Even though multilinear models can be viewed as an extension of more familiar bilinear methods, the paradigm of tensor algebras or weighted matrix representations may not be as familiar. Thus, it is useful to have an example to illustrate the types of questions that can be addressed by such models. In truth there are many perspectives on multilinear models that could be adopted in this example, much the same way as one could choose to emphasize variance summary or structure interpretation with an elementary introduction to principal components. However, for the purposes of this example, an approach that is similar to that utilized in multidimensional scaling (MDS) will be adopted. Of course this will make more or less sense depending on the familiarity with MDS.

In the beginning sections of this chapter, an example on consumer preferences for bread was introduced. This scenario will be further explored and the results of a PARAFAC decomposition will be provided. In this example five different breads were baked in replicates giving a total of ten samples, and eight different judges assessed the breads with respect to eleven different attributes. These data were kindly provided by Prof. Magni Martens, from the Royal Veterinary and Agricultural University (KVL) in Denmark and come from a student project in Sensory Science. They were popularized as an example of multilinear models by Prof. Rasmus Bro, also from KVL and a pioneer in the field. The schematic in Figure 1.3 illustrates how the data would be typically arranged into a 10 by 11 by 8 data cube. The ratings or assessment measures are the numerical entries in each slab or slice of the cube.

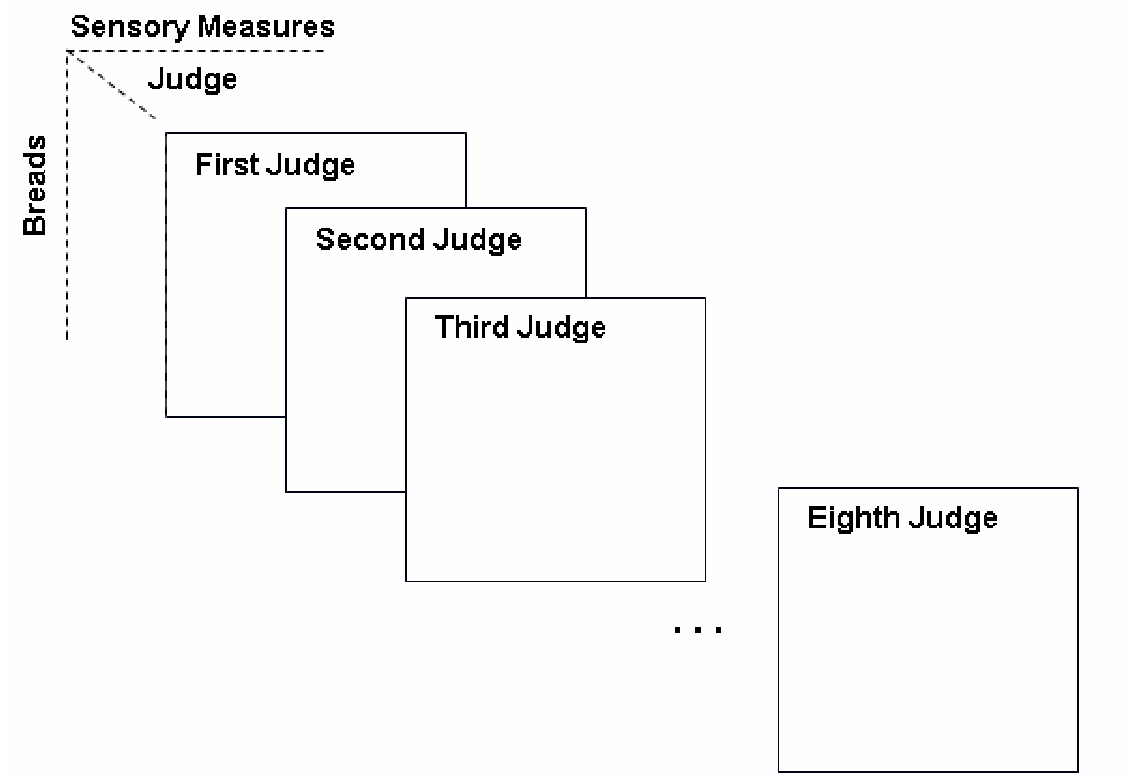

Figure 1.3 Bread Example Schematic

A typical question one might ask with such data is "Do the judges differ in their ratings depending on the sensory attribute being rated?" A common, and trivial, univariate approach might focus on a single sensory measure (or average all sensory measures) and then average across breads and use a simple ANOVA to address the question. Conversely, a multivariate approach might first employ a MANOVA. But it is worth noting that the Bread dimension does not exhibit 10 replications on the same bread (samples) but rather 2 replications on each of 5 different breads. Hence, that dimension is not the usual sample dimension and, hence, the two procedures mentioned above are not quite right. With the ANOVA one could argue that, in effect, two separate *dimensions* (sensory variable and bread type) have been collapsed in order to facilitate the univariate analysis, while with the MANOVA perhaps one dimension (bread type) has been unfortunately collapsed. Indeed, it is common in multilinear analysis for there to be *no* dimension that corresponds to a sampling dimension.

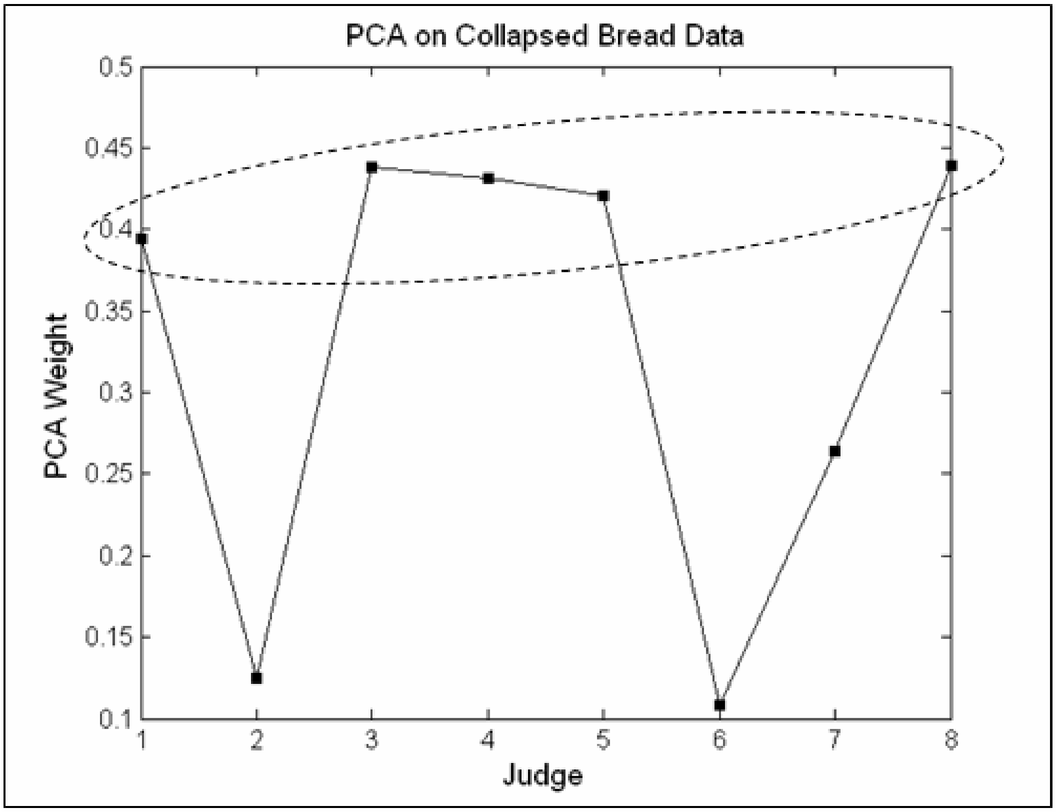

Figure 1.4 Principal Component Weights on Collapsed Bread Data Cube (1)

The absence of a sampling dimension does not preclude an analysis of structure by way of a typical bilinear method such as principal components (PCA). For instance, one could collapse the data across breads (say, by averaging) and then do a PCA on the resulting 11 (sensory measures) by 8 (judges) data matrix. Looked at as an 11 x 8 data matrix, one would naturally address the question above by extracting the first principal component *weight vector*, or corresponding *vector of loadings*, and a plot of these weights (loadings) might reveal suggested differences in the judges (Figure 1.4).

Conversely, looked at as an 8 x 11 data matrix, one could address the question above by, perhaps, extracting the first two PCA *scores* and plotting them for each of the 8 judges, being careful to look for groupings, similarities and differences among the judges in that score plot (Figure 1.5). Both the weights and the scores were generated from auto-scaled data and both suggest that Judges 1, 3, 4, 5 and 8 are somehow different from Judges 2 and 6.

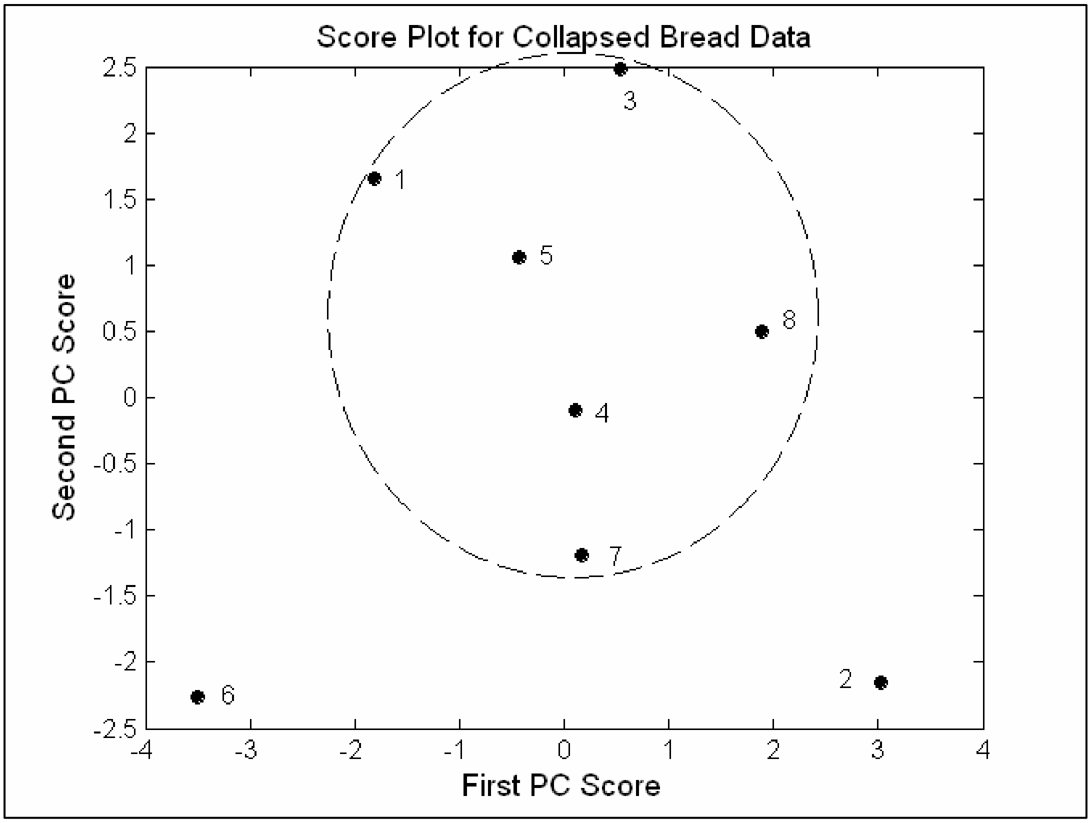

Figure 1.5 Principal Components Scores on Collapsed Bread Data Cube (2)

Fitting a simple multilinear (three-way, in this case) model yield similar constructs, but does not force the collapsing over dimensions. There are cosmetically different, but mathematically equivalent *ways* to motivate such a model. As mentioned above, this example attempts to capture the intuition provided by multidimensional scaling in general and "individual scaling" (INDSCAL) in particular. With this perspective on the analysis we would naturally presume a common underlying group score plot, say, a common score plot for all judges viewed as a group (Figure 1.6).

The actual group score plot from fitting the trilinear model is presented in that figure. But individual judges can have different weights applied to this common plot to produce their personal score plots. These weights shrink or stretch the dimensions for the particular individual (Figure 1.7).

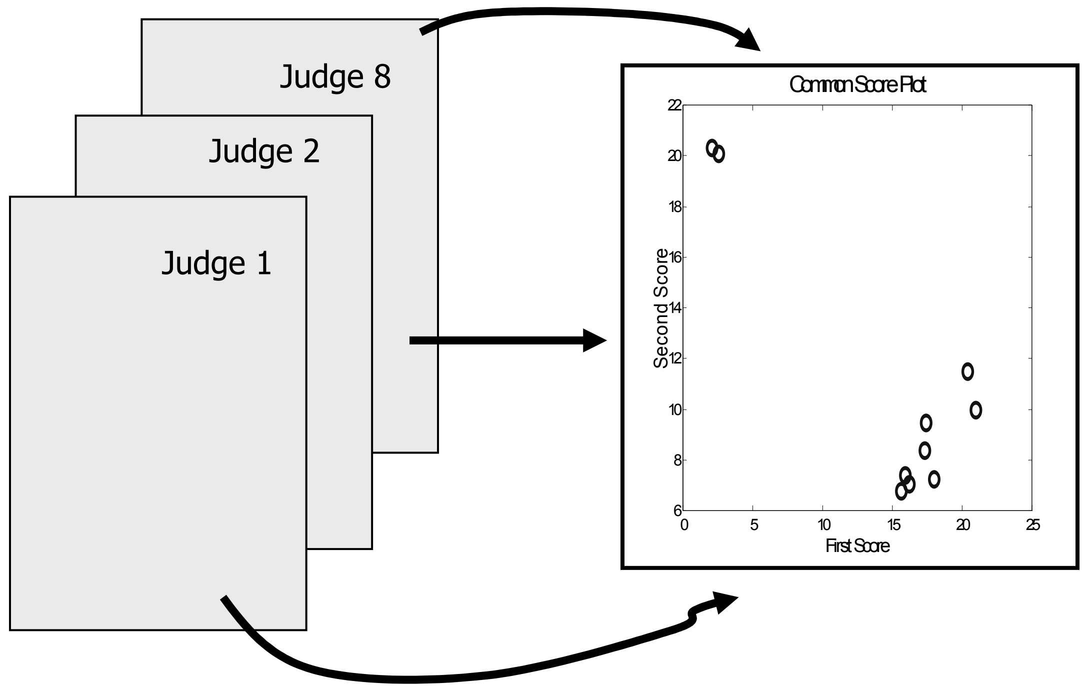

Figure 1.6 Depiction of Common Scoreplot for Judges as a Group

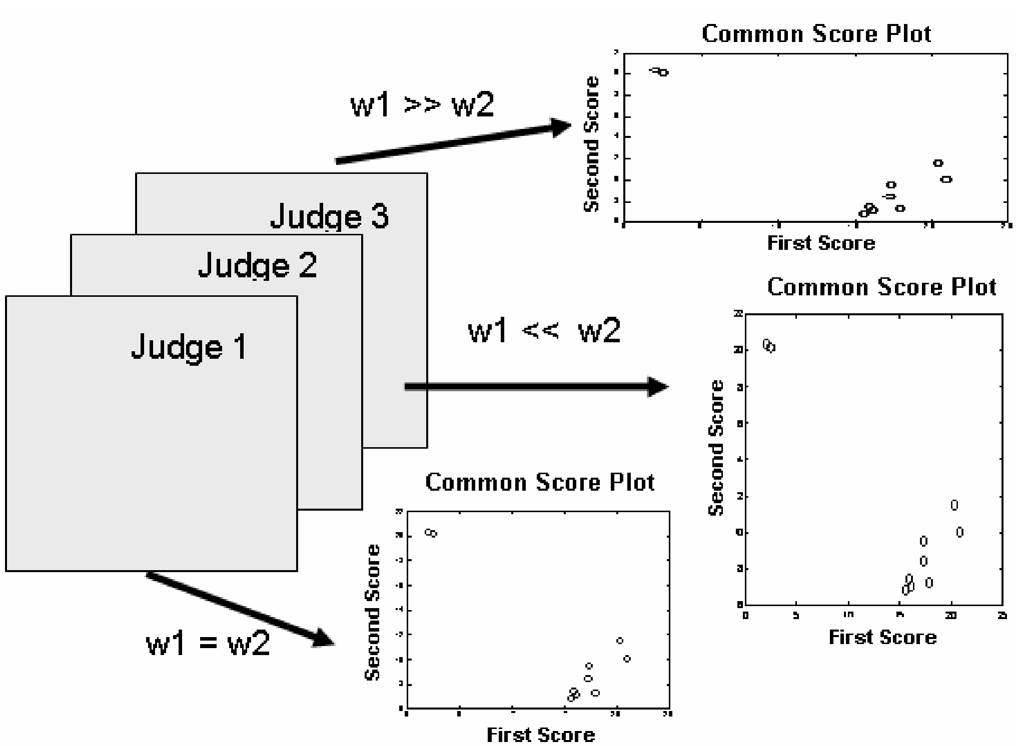

Figure 1.7 Depiction of Specific Scoreplots for Judges

In the plot of the 8 judges' weights (Figure 1.8), it appears that the weighting patterns for Judges 2, 6, and 7 are very similar. Additionally, it is possible to see the stretching and shrinking of groups maps that occur for the individual judges or raters. The diagonal line represents the common group map so that if a point fell on this line, the judge would have a group map that was the same as the overall group map (Figure 1.6). Here, it appears that Judge 8 tends to have a group map that is similar to common one. However, Judge 5 does not fall near this diagonal line, her group map would considerably stretch the common one.

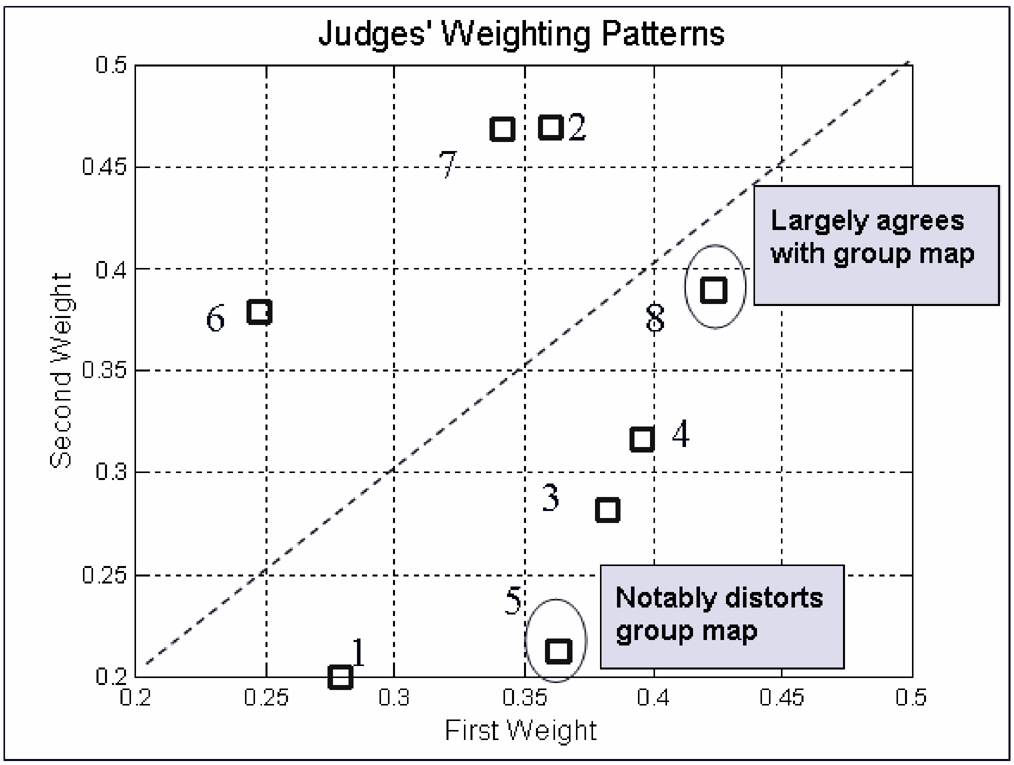

Figure 1.8 Judges' Weighting Patterns from Trilinear Model Fit

Figures 1.9 and 1.10 point to where these differences between the two groups of judges seem to lie with respect to the measured sensory variables. Indeed, it is reassuring that the broad-brush conclusions provided by the bilinear analysis are supported by the trilinear analysis where no collapsing was necessary.

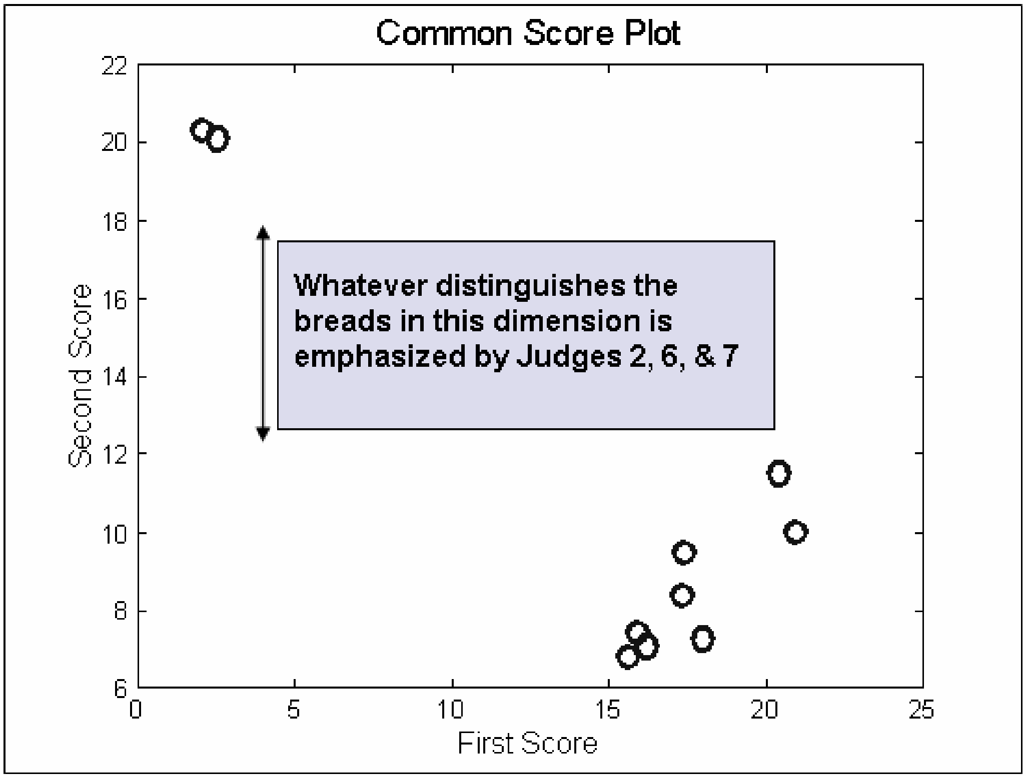

Figure 1.9 Interpretation of Judges' Weighting Patterns

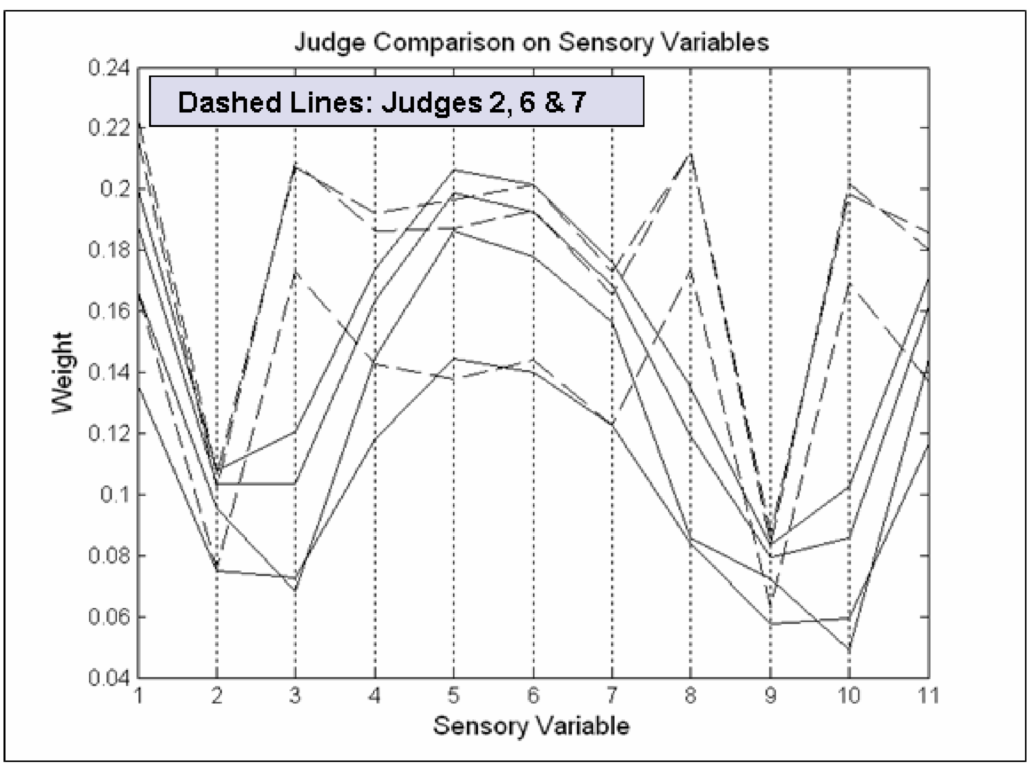

Figure 1.10 Grouping of Judges on Sensory Measures

This kind of application would be a very typical use of multilinear models in psychometrics, even though the example here is more slanted toward food science. Again, these models are preferred by many over bilinear alternatives because exploratory analysis is still possible without collapsing across dimensions. In addition, the claim to uniqueness enjoyed by multilinear models is also *perceived as a major advantage*. The issue of uniqueness is one that will be explored in the chapters that follow. However, before headway can be made into new perspectives it is important to understand how early uniqueness results evolved into the uniqueness results used today.

1.5 Literature Review: Uniqueness of the PARAFAC Decomposition

One of the differentiating features of PARAFAC decompositions is that, under certain conditions, unique solutions are possible. The search for uniqueness conditions for the PARAFAC Decomposition has a limited past. Until recently, the complex structure of the problem and the need for tensor algebras or other similarly abstract characterizations provided a roadblock to the development of uniqueness conditions. Theoretically, the PARAFAC decomposition surpasses its bilinear counterparts in that it is possible to obtain solutions that do not suffer from the rotational problem. However, not all PARAFAC solutions will be constrained sufficiently so that the resulting decomposition is unique. Although strides have been made to identify the types of PARAFAC solutions that will be unique, as reviewed in the following paragraphs, much about the conditions needed for uniqueness is still unknown.

Mathematical insights into the uniqueness properties of PARAFAC decompositions were originally discussed by Robert Jennrich and published in Harshman, 1970. In his proof, Jennrich showed that a unique solution would exist if an $I \times J \times K$ array was decomposed into square loading matrices, each with R measures. However, Harshman was able to find empirical evidence to suggest that although these conditions were sufficient, they were not minimal. Operating in the $I \times J \times 2$ case with full-column rank loading matrices, Harshman showed that if no two columns of the loading matrix **C** were proportional, the solution was unique without requiring the loading matrices to be square.

As evidenced by these first approaches, developing conditions for uniqueness would involve a discussion on the relationship of the rank of the loading matrices and properties of the columns. To that end, Kruskal (1977) developed the notion of *k-rank* (although the actual term was coined by Harshman and Lundy) as a means to investigate uniqueness.

Definition 1.7

Let **X** be a matrix. The *k-rank* of **X** is the largest value of k such that every collection of k columns in **X** is linearly independent.

Kruskal's key result, to be discussed in detail later, was that the loading matrices (**A**, **B**, and **C**) obtained from the parallel factor decomposition would be uniquely identified if $k_A + k_B + k_C \geq 2(R+1)$, where k_A was the *k-rank* of a matrix **A**, and so on. To date, his work remains the most extensive in the search for criteria; and his resulting condition is generally accepted as the climax of uniqueness research. Consequently, other uniqueness results have stemmed from his constraint on the *k-rank*s of the loading matrices in the hopes of developing a set of necessary conditions.

Developments in uniqueness conditions were slow to evolve in the years after Kruskal's theorem. A few results of note appeared in the early nineties. Leurgans and Ross (1993) and Krijnen (1993) made advances in terms of the linear independence of the columns. The work of Krijnen would prove particularly useful as he was able to prove that a necessary condition for uniqueness was that the k-rank of all loading matrices must be at least 2.

In the last few years, aided by the introduction of Khatri-Rao products by Bro (1998) discussed later, a resurgence of interest in uniqueness results and Kruskal's condition has ensued. In 2000, Sidiropoulos and Bro expanded Kruskal's result by $I \times J \times K$ generalizing the result to include multiway arrays. Liu and Sidiropoulos (2001) demonstrated that a necessary condition for uniqueness was that the columns of the Khatri-Rao products be linearly independent, a crucial element to the ideas and perspectives that will be presented in the chapters that follow. However, even with the advent of new interest and progress, Kruskal's result remains the optimal set of conditions for the uniqueness of PARAFAC decompositions.

To that point, in the absence of true necessary and sufficient conditions, many users of multilinear models have applied Kruskal's condition as if it were indeed necessary and sufficient, an idea that was deposed by ten Berge and Sidiropoulos in 2002. By producing alternative solutions when Kruskal's condition was not met, necessity was shown when the number of factors was two or three ($R = 2$ or $R = 3$). In the case of four factors ($R = 4$), however, uniqueness was achieved even when Kruskal's condition was not. Hence, the long assumed necessity of Kruskal's condition was disproved.

ten Berge and Sidiropoulos conjectured that more than k-rank was needed to define necessary conditions for uniqueness and suggested that Kruskal's condition might be made necessary if the loading matrices were constrained to have k-rank equivalent to rank. It is at this conjecture, the latest work in the search for uniqueness conditions, that this dissertation begins.

The remaining chapters of the dissertation will provide the necessary background to evaluate the work of ten Berge and Sidiropoulos, starting with a detailed review of the early work of Harshman and Kruskal given in Chapter 2. Chapter 3 will describe the new representation of PARAFAC solutions and the methodology provided by ten Berge and Sidiropoulos in 2002 to evaluate the necessity of Kruskal's condition. A critical investigation of the results from ten Berge and Sidiropoulos will also be included in Chapter 3. In Chapter 4, an alternative approach, stemming from the work of ten Berge and Sidiropoulos, will suggest that the identification of uniqueness would require additional measures other than rank and k-rank alone. Finally, Chapter 6 includes general comments (for all R) regarding uniqueness, suggesting different directions than current literature proposes.

Copyright © Heather Michele Clyburn Bush 2006

2. THE EARLY YEARS: HARSHMAN AND KRUSKAL

2.1 PARAFAC Uniqueness

The bilinear methods that preceded PARAFAC allowed transformations so that many alternative solutions provided the same decomposition but different interpretations, as described in Chapter 1. Thus, theoretical solutions achieved from bilinear methods were considered questionable and far from unique. The introduction of PARAFAC provided by Harshman (1970), however, not only provided an alternative method which preserved the structure of higher dimensional problems but also a means of analysis which did not fundamentally suffer from the rotational problem of the bilinear methods. In fact, for particular cases, PARAFAC solutions could be considered unique.

Imperative to the investigation of uniqueness is the definition of *alternative* PARAFAC solutions. It is critical to understand in what sense alternative PARAFAC solutions are different from the original. An alternative PARAFAC solution is obtained by transforming the loading matrices by post-multiplying by transformation matrices, synonymous to the rotational problem in the bilinear setting. However, not all transformations result in the same degree of modification. Transformation matrices that are *permutation-scale* do not transform the PARAFAC solution so that the alternative is truly different from the original.

Definition 2.1 (Harshman, 1972, paraphrased)

A matrix is *permutation-scale* if it can be written as the product of a nonsingular permutation matrix and diagonal matrix, and as such, has exactly one nonzero element in each row and column.

Permutation-scale transformations of the original loading matrices simply rearrange and scale the columns so that alternative loading matrices are not truly different from the original loading matrices. Thus, alternative PARAFAC solutions created from permutation-scale transformations are considered to be equivalent to the original PARAFAC solution. Consequently, if the only alternatives that can be formed are simply permutation-scale versions of the original, the PARAFAC solution is considered unique.

The history of PARAFAC and the search for uniqueness is a short one, spanning a little more than three decades. Even so, it is not a simple task to set out a chronology of the history. The progress of PARAFAC has not followed a sequence of linear events which shaped its history but has benefited from many different players with different emphases and perspectives. Thus, the evolution of PARAFAC is more of a tangled web of trial and error, failures and successes, making it difficult for those new to the area to simply start from the beginning. Much of the entanglement can be attributed to the complexity of the problem, making it difficult to set forth simple definitions and all-encompassing theorems. However, finding an appropriate definition for what it means to be *unique* is a critical first task and must be undertaken before attempting to search for which conditions would constrain a PARAFAC solution to be unique.

Both Professor Richard Harshman and Professor Joseph Kruskal were major contributors to the definition of uniqueness. Professor Richard Harshman developed the idea of parallel proportional profiles and presented the first theorems for uniqueness. Professor Joseph Kruskal provided a modification to Harshman's uniqueness definition and the most general set of conditions for uniqueness to date. His treatment of PARAFAC solution uniqueness in 1977 remains the gold standard in the field. Additionally, due to the success of Kruskal's results, many have applied his theorem as if it were a set of necessary and sufficient conditions (ten Berge and Sidiropoulos, 2002).

The interest for this dissertation began with a review of Harshman's 1972 paper on PARAFAC uniqueness. Harshman proposed that Jennrich's earlier conditions, presented in Harshman's 1970 work, were too restrictive and sought to provide evidence that uniqueness could be obtained with fewer conditions. Harshman suggested that when the array was $I \times J \times 2$ and two of the matrices had full column rank, the PARAFAC solutions would be unique when the columns of the loading matrices were not proportional. Although the conditions offered by Harshman were indeed sufficient for uniqueness, deeper investigation of his proof revealed an issue with the logic. Hence, as will be discussed later in the chapter, the 1972 work did not provide formal proof for Harshman's uniqueness conditions. While Harshman's proof did not provide formal evidence, his intuition was correct. As eluded to in Chapter 1, the application of the transformation to the diagonal matrix as in equation (1.4) would hinder the search for uniqueness conditions. Kruskal, however, would focus attention on alternatives where the transformation was multiplied to the vector before diagonalization as in (1.5) and would later provide a more encompassing set of conditions.

The work of both Harshman and Kruskal will be presented in the sections that follow. In addition to the presentation of the theorems and definitions from Harshman and Kruskal, the disadvantage and resulting problems of utilizing a representation which applied transformations to the diagonal matrix will be described. A description of Harshman and Kruskal in this context has not been found in the literature to date. Finally, it will be seen that Harshman's 1972 conditions could be evaluated for uniqueness using Kruskal's theorem. Thus, Kruskal's proof provided the formal evidence to support Harshman's intuition.

2.2 Harshman (1972)

Although the 1972 results of Harshman did not provide formal evidence for uniqueness, the magnitude of his contributions is enormous and cannot be ignored, and his conjectures for uniqueness were later shown to be true.

2.2.1 The Set-Up

Before providing a theorem for sufficient conditions for PARAFAC solution uniqueness, Harshman informally defined, what was meant by uniqueness.

Definition 2.2 (Harshman, 1972, paraphrased)

A PARAFAC solution is *unique* if every alternative PARAFAC solution is composed of loading matrices that are permutation-scale versions of the originals.

With a definition of PARAFAC uniqueness in place, Harshman proposed a theorem for uniqueness when the array had dimension $I \times J \times 2$. Additionally, Harshman required that two of the loading matrices, A and B, have full column rank. Thus, the matrix A with dimension $I \times R$ would have rank of R. Likewise, the loading matrix $\mathbf{B}_{\mathbf{J \times R}}$ would also have rank = R.

Definition 2.3 (Harshman, 1972):
Assume that **(A,B,C)** are the loading matrices of an R-component PARAFAC solution for the array ***X***, where **A** and **B** are full-column rank, and the matrix **C** has two rows, such that $\mathbf{X_1 = AC_1B^t}$ and $\mathbf{X}_2 = \mathbf{AC}_2\mathbf{B^t}$. Additionally, assume that $(\tilde{\mathbf{A}},\tilde{\mathbf{B}},\tilde{\mathbf{C}})$ are the loading matrices of an equivalent alternative R-component PARAFAC decomposition of the array ***X*** so that $\mathbf{X_1 = \tilde{A}\tilde{C}_1\tilde{B}^t}$ and $\mathbf{X}_2 = \mathbf{\tilde{A}\tilde{C}}_2\mathbf{\tilde{B}^t}$, where $\mathbf{\tilde{A} = AT_A}$ and $\mathbf{\tilde{B} = BT_B}$. $\mathbf{T}_A$ and $\mathbf{T}_B$ are square, nonsingular linear transformation matrices. If $\mathbf{C^P = C_1C_2^{-1}}$ has distinct diagonal elements, $\mathbf{T}_A$ and $\mathbf{T}_B$ are permutation-scale matrices.

The implication of Harshman's theorem was that imposing a condition on the rows of **C** restricted the nonsingular linear transformations to permutation-scale matrices, so that alternative solutions would, in a sense, be only trivially different from the original. By Definition 2.2, the theorem provided a sufficient condition for PARAFAC solution uniqueness when the array was of dimension $\mathbf{I \times J \times 2}$ and two of the loading matrices had full-column rank.

However, a thorough review of his proof revealed the disadvantage of applying the transformation as in (1.4), to the diagonal matrix. Transforming the diagonal matrix resulted in inadvertently assuming that $\mathbf{T}_A$ and $\mathbf{T}_B$ were permutation-scale matrices. Again, this was not to say that the theorem was untrue. Five years after the publication of Harshman's uniqueness theorem, the problem was approached from a slightly different perspective by Kruskal (1977). The conditions offered by Kruskal were much more global and encompassed any

PARAFAC solution of dimension $I \times J \times K$ and as such could be applied to Harshman's scenario to identify sufficient conditions when the array had dimension $I \times J \times 2$ and the solution had two component matrices with rank of R.

Under Harshman's constraints, Kruskal's condition was used and verified that the matrix with two rows would indeed need to have non-proportional columns in order for the PARAFAC solution to be unique. In other words, a sufficient condition for uniqueness would be that the elements of $\mathbf{C^P = C_1C_2^{-1}}$ were distinct as Harshman suspected. A detailed explanation of the problems of applying the transformation to the diagonal matrix will be presented, followed by description of Kruskal's sufficient conditions for uniqueness. Finally, it will be shown how the application of Kruskal's condition to the problem posed by Harshman provided formal evidence to Harshman's claim.

2.2.2 Issues in the proof

The fundamental problem with Harshman's proof was in the set-up of the alternative PARAFAC solution, which led him to inadvertently assume what he was trying to prove. The critical difference between Harshman's representation of the alternative PARAFAC solution and the representation of others that would follow was in the application of the transformation matrices. Similar to later approaches, the transformation matrices were applied by post-multiplying $\mathbf{T}_A$ and $\mathbf{T}_B$ to **A** and **B**, respectively. However, the transformation of the third matrix is complicated by the representation of the PARAFAC solution. In order to see where the problem occurred, it will be necessary to present a portion of his argument.

Consider the form of the alternative decomposition $\mathbf{X}_k = \mathbf{AC}_k\mathbf{B^t}$ for *k = 1, 2,* where $\mathbf{\tilde{A} = AT_A}$ and $\mathbf{\tilde{B} = BT_B}$. In the proof, Harshman used the equality of $\mathbf{X_1 = \tilde{A}\tilde{C}_1\tilde{B}^t}$ and $\mathbf{X_1 = AC_1B^t}$ to solve for $\mathbf{T}_A$ and $\mathbf{T}_B$, such that $\mathbf{T_A = C_1A^t(\tilde{A}^t)^{\mp}C_1^{-1}}$ and $\mathbf{T_B = C_2B^t(\tilde{B}^t)^{\mp}C_2^{-1}}$, had right inverses. Therefore, $\mathbf{T}_A\mathbf{T}_A^{-1} = \mathbf{I}$ and $\mathbf{T}_B\mathbf{T}_B^{-1} = \mathbf{I}$ were inserted into the original decomposition to obtain $\mathbf{X_k = AT_AT_A^{-1}C_k\left(T_B^t\right)^{-1}T_B^tB^t}$.

Substituting $\tilde{\mathbf{A}} = \mathbf{AT_A}$ and $\tilde{\mathbf{B}} = \mathbf{BT_B}$ into the above decomposition yielded $\mathbf{X_k} = \tilde{\mathbf{A}}\mathbf{T_A^{-1}C_k}\left(\mathbf{T_B^t}\right)^{-1}\tilde{\mathbf{B}}^\mathbf{t}$, which was then set equal to $\mathbf{X}_k = \tilde{\mathbf{A}}\tilde{\mathbf{C}}_k\tilde{\mathbf{B}}^\mathbf{t}$. Using the nonsingularity of $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$, Harshman found that $\tilde{\mathbf{C}}_\mathbf{k} = \mathbf{T_A^{-1}C_k}\left(\mathbf{T_B^t}\right)^{-1}$.

The traditional PARAFAC representation, described in detail in Chapter 1 involves a diagonal $R \times R$ matrix sandwiched between an $I \times R$ matrix and an $R \times J$ matrix. The diagonal matrix is formed from diagonalizing a row of the $K \times R$ matrix. Harshman applied the transformations directly to the diagonal matrix sandwiched between the two full column rank matrices. The transformations were placed in this way so that the alternative PARAFAC solution and the original PARAFAC solution would be the same. In order for the alternative PARAFAC solution to be equivalent to the original, these transformations would need to “cancel out” so that the two PARAFAC solutions would produce the same array.

The alternative PARAFAC solution, $\mathbf{X}_k = \tilde{\mathbf{A}}\tilde{\mathbf{C}}_k\tilde{\mathbf{B}}^\mathbf{t}$, was by definition a PARAFAC solution. Therefore, the PARAFAC structure would be mandated, and the matrix in the middle would necessarily be diagonal. Thus, $\tilde{\mathbf{C}}_\mathbf{k} = \mathbf{T_A^{-1}C_k}\left(\mathbf{T_B^t}\right)^{-1}$, the matrix in the middle, would be a diagonal matrix. Therefore, the pre-multiplication by $\mathbf{T}_A^{-1}$ and post-multiplication by the inverse of $\mathbf{T_B^t}$ would preserve the diagonal structure of $\mathbf{C_k}$. However, the diagonal structure of $\mathbf{C_k}$ would only be preserved if both $\mathbf{T}_A^{-1}$ and $\left(\mathbf{T_B^t}\right)^{-1}$ had exactly one nonzero element in each row and column. Consequently, by definition, $\mathbf{T}_A$ and $\mathbf{T}_B$ would have to be permutation-scale matrices. Although Harshman had supposed that these transformation matrices were arbitrary orthogonal matrices, he unknowingly required that they be permutation-scale, the conclusion of his uniqueness theorem.

Unfortunately as soon as $\tilde{\mathbf{C}}_\mathbf{k} = \mathbf{T_A^{-1}C_k}\left(\mathbf{T_B^t}\right)^{-1}$ was required to be diagonal, the form of $\mathbf{T}_A$ and $\mathbf{T}_B$ was implicitly assumed to be permutation-scale. Perhaps not mindful of this, Harshman proceeded to reason that the distinct elements of

$\mathbf{C^P = C_1C_2^{-1}}$ would force the transformation matrices to be permutation-scale. However, it was the diagonal form of $\tilde{\mathbf{C}}_1$ and $\tilde{\mathbf{C}}_2$ and not a condition on the columns of **C** that determined the structure of $\mathbf{T}_A$ and $\mathbf{T}_B$. Applying the transformations to the diagonal matrix $\mathbf{C_k}$ will always limit transformations to permutation-scale matrices so that the diagonal structure will be preserved. Accordingly, any approach to find uniqueness conditions that applied the transformations directly to $\mathbf{C_k}$ would suffer from this same problem.

2.3 Kruskal

The results offered by Kruskal in his 1977 work are very impressive, and the theorems and mathematical descriptions of PARAFAC solutions he developed form the foundation for essentially all treatments of PARAFAC uniqueness.

The first contribution of Kruskal was to re-formulate the definition of uniqueness provided first by Harshman. The definition posed by Kruskal defined alternatives where uniqueness occurred and was in essence the same as Harshman's definition. Both definitions had the premise that uniqueness occurred when alternative PARAFAC solutions were only formed from component matrices that were permutation-scale versions of the original. The difference, however, was that Kruskal applied the transformations to each of the loading matrices. Hence, Kruskal applied the transformation directly to the vector before it was diagonalized, preventing the diagonal structure from restricting the transformation.

Definition 2.4 (Kruskal, 1977 paraphrased)

Matrices **(A,B,C)** are *equivalent* to matrices $(\tilde{\mathbf{A}},\tilde{\mathbf{B}},\tilde{\mathbf{C}})$ if $\mathbf{A = A\Pi\Lambda_A}$, $\mathbf{B = B\Pi\Lambda_B}$, and $\mathbf{C = C\Pi\Lambda_C}$, where $\mathbf{\Lambda}_A\mathbf{\Lambda}_B\mathbf{\Lambda}_C = \mathbf{I}_{R\times R}$ and $\mathbf{\Pi}$ is a permutation matrix. When the only alternative representations (of the same rank) have the form $(\tilde{\mathbf{A}},\tilde{\mathbf{B}},\tilde{\mathbf{C}})$, where $\mathbf{\Lambda}_A\mathbf{\Lambda}_B\mathbf{\Lambda}_C = \mathbf{I}_{R\times R}$ and $\mathbf{\Pi}$ is a permutation matrix, the solution is *unique*.

Although the difference seems subtle, the ramifications are substantial. Harshman had applied the transformations to the two loading matrices and the diagonal matrix. Transforming the diagonal matrix instead of the loading matrix would always constrain the transformation matrices to be permutation-scale. However, by transforming each of the loading matrices, including the core matrix, more arbitrary transformation matrices could be considered and conditions for uniqueness could truly be investigated.

It should be noted that when the transformations are permutation-scale the two strategies are equivalent. Harshman and Lundy (1984) showed that when uniqueness occurred, it would not matter if the permutation-scale transformation was applied to the diagonal matrix or to the loading matrix. However, for investigating uniqueness, the use of Kruskal's definition and applying transformations to the loading matrix **C** would prevent the diagonal form of $\mathbf{C}_k$ from dictating the type of transformations allowed. Thus, with this more accurate definition of uniqueness it was possible for Kruskal to pursue and provide conditions for the uniqueness of PARAFAC solutions.

2.3.1 The Set-Up

As previously discussed, the definition of uniqueness posed by Kruskal considered transformations to the original component matrices; however, these transformations would essentially "cancel out" when considered in the solution. Hence, alternatives were only permuted and scaled versions of the original matrices, but the PARAFAC solutions were equivalent.

In addition to the definition of uniqueness, Kruskal needed to define a concept with the same flavor as Harshman's non-proportional columns. However, Kruskal was not only interested in the proportionality of the columns but the linear dependencies among the columns. Kruskal was able to characterize a measure of the linear independence of the columns of a matrix, Definition 1.7. The largest number of columns that can be grouped together so that the group forms a linearly independent set is the *k-rank* of a matrix. Harshman and Lundy (1984) later named this measure *k-rank*.

The idea of k-rank was that considering every combination of columns would result in a set with k columns such that the k columns were linearly independent, but the addition of one more column would create linear dependencies. Therefore, a matrix with full column rank would have a k-rank equal to the number of columns. The idea of k-rank is somewhat abstract and reduced row echelon forms can be used as a tool to better understand the concept. Additionally, reduced row echelon forms will be visited later and will prove to be valuable as Kruskal's condition is further investigated.

A matrix with four columns, where three of the columns are linearly independent and the fourth column is created by summing the three linearly independent columns, could be represented as

$$\begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 & \alpha_1+\beta_1+\gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 & \alpha_2+\beta_2+\gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 & \alpha_3+\beta_3+\gamma_3 \end{pmatrix}. \tag{2.9}$$

The reduced row echelon form of (2.1) would be

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \tag{2.10}$$

The k-rank of (2.2) is 3 or the number of nonzero elements in the column not involved in the identity partition.

Considering a numerical example, the matrix $\begin{pmatrix} 2 & -1 & 7 & 1 \\ 5 & 4 & 6 & 9 \\ 1 & 3 & 2 & 4 \end{pmatrix}$ has k-rank of 2. This can easily be seen by considering the reduced row echelon form of the matrix, $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$. The two 1's in the last column indicate that the last column was created by summing the first two columns.

It was suggested by Jennrich that the size of the matrix was essential to uniqueness, and Harshman suggested that the key was proportionalities among columns. It turned out that the key concept was k-rank, at least for providing conditions which would be sufficient for uniqueness.

Theorem 2.2 (Kruskal, 1977 paraphrased) Let the k-rank of a matrix **M** be represented by $\mathbf{k(M)}$. If $\mathbf{k(A)+k(B)+k(C)} \geq 2R+2$, where R is the number of factors, the PARAFAC solution is unique.

2.4 Kruskal's condition with Harshman's criteria

Harshman's proof and the problems with applying the transformations to the diagonal matrix instead of the entire loading matrix had not been explained; therefore, until now, there was not a need to verify the claims Harshman made in his 1972 work. However, the task of validating Harshman's assertions is made easy since Kruskal's theorem can simply be applied to the Harshman scenario, two loading matrices with rank = R and the third matrix with 2 rows.

Since two of the loading matrices have full-column rank, the k-rank of each matrix is R and the k-ranks sum to 2R. From Kruskal's theorem, a PARAFAC solution was unique when $\mathbf{k(A)+k(B)+k(C)} \geq 2R+2$. Applying Kruskal's condition to Harshman's scenario implied that $2R+\mathbf{k(C)} \geq 2R+2$ or that $\mathbf{k(C)} \geq 2$. Harshman required that the matrix **C** had only two rows. Therefore the maximum rank of the matrix **C** is 2. The k-rank of a matrix is at most the rank of the matrix. Hence, the k-rank of **C** is at most 2. Thus, according to Kruskal, Harshman's PARAFAC solution would be unique if $\mathbf{k(C)} = 2$.

Recall that Harshman concluded that the PARAFAC solution would be unique if $\mathbf{C^P = C_1C_2^{-1}}$ had distinct elements or the columns of **C** were non-proportional. The idea of non-proportionality of the columns was tantamount to assuming that the k-rank of **C** was at least 2. If two columns of a matrix were proportional, then the largest number of columns that could be grouped together so that the group of columns was linearly independent was one. Hence, the k-rank of a matrix with proportional columns was 1. Thus, under Harshman's scenario, his criteria of distinct element diagonal elements of $\mathbf{C^P = C_1C_2^{-1}}$ was equivalent to requiring the k-rank be equal to 2.

Hence, Harshman's ideas on uniqueness were verified by the application of Kruskal's theorem. Krijnen (1993) would later show that Harshman's intuition regarding the necessity of the non-proportionality of columns was correct by proving that $\mathbf{k(M)} \geq 2$ for all loading matrices **M** was a necessary condition for uniqueness.

2.5 Conclusions

In addition to providing the notion of proportional parallel profiles and PARAFAC, Harshman also laid the conceptual groundwork for uniqueness conditions. However, it was Kruskal who produced the most beneficial results for the pursuit of uniqueness conditions. The idea of k-rank and the subsequent k-rank conditions are considered to be the pinnacle of uniqueness research to date. No other work had been able to provide a more universal set of sufficient conditions or more valuable perspective of PARAFAC solution uniqueness.

As mentioned in Chapter 1, after the work of Kruskal, the work in the area of uniqueness was stymied for almost three decades. The exploration for additional or supplemental uniqueness conditions had been hindered by the algebraic subtleties found in Kruskal's proof technique. However, in recent years, the interest in pursuing uniqueness conditions has been renewed, most likely spawned from the introduction of a new representation of the PARAFAC solution. More contributions in the search for uniqueness conditions have emerged and provided novel approaches, approaches that begin with the application of Kruskal's theorem.

Copyright © Heather Michele Clyburn Bush 2006

3. A PARADIGM SHIFT IN UNIQUENESS RESEARCH

Since the work of Kruskal over twenty years ago, little had been accomplished in terms of providing a less restrictive set of conditions for uniqueness. The greatest obstacle to determining more liberal conditions for uniqueness was the form of the PARAFAC solution. Traditionally, there were two methods for representing a PARAFAC decomposition: tensor algebras and matrix notation. Both of these methods were described in Chapter 1. Tensor algebras have been avoided by most in describing PARAFAC decompositions, while matrix representations have been widely employed and dominate uniqueness literature. As explained in Chapter 1, the matrix representation describes the data arrays as a series of stacked matrices or slabs. The matrix representation of a PARAFAC solution consists of two loading matrices which are common to each slab and a diagonal matrix which contains the differentiating weights of each slab. Although intuitive, the matrix representation of the PARAFAC solution is mathematically awkward and suffocated efforts to finding necessary and sufficient conditions for uniqueness.

The greatest strides towards identifying conditions for uniqueness were made by Kruskal (1977). Offering the most rigorous and formal treatment of PARAFAC solution uniqueness, Kruskal was able to provide the least restrictive set of criteria for uniqueness to date, described in Chapter 2. However, Kruskal was only able to give formal proof that the condition was sufficient for uniqueness. Even so, in the absence of necessary and sufficient conditions, Kruskal’s sufficient uniqueness condition has been assumed to be and applied as if it were necessary as well as sufficient (ten Berge and Sidiropoulos, 2002). Again, the mathematical complexity that the PARAFAC solution structure imposed made it difficult to offer theoretical proof to deny or support the idea of necessity. However, a new perspective was provided in the representation of the PARAFAC decomposition through the application of Khatri-Rao products. With a new PARAFAC solution structure, the old complications in investigating uniqueness were mitigated.

3.1 The Set-up

3.1.1 Khatri-Rao Products

Previous uniqueness results used the matrix representation of the PARAFAC solution as repeated loading matrices, **A** and **C**, combined in various weights by the rows of **B**, or $\mathbf{X_j} = \mathbf{AB_jC^t}$. These $\mathbf{X_j}$ matrices were then stacked to compose the array ***X***. The Khatri-Rao (KR) product, introduced in the PARAFAC paradigm by Bro (1998), was a mathematical tool that allowed the PARAFAC solution to be discussed as the stacked array ***X*** instead of considering the individual $\mathbf{X_j} = \mathbf{AB_jC^t}$ matrices that composed the array.

Definition 3.1 (Bro 1998, McDonald 1980, Rao & Mitra 1971)

The *KR product* of two matrices **A** and **B**, represented as $(\mathbf{A} \circ \mathbf{B})$, was defined as the column-wise Kronecker product of **A** and **B** such that

$$\begin{aligned}\mathbf{A_{I\times R}} \circ \mathbf{B_{J\times R}} &= \left[\mathbf{vec(a_1 b_1^t)} \quad \cdots \quad \mathbf{vec(a_R b_R^t)}\right]_{\mathbf{IJ\times R}} \\ &= \left[\mathbf{(a_1 \otimes b_1)} \quad \cdots \quad \mathbf{(a_R \otimes b_R)}\right]_{\mathbf{IJ\times R}}\end{aligned}$$

Therefore, the array ***X*** could be specified as $\boldsymbol{X} = \mathbf{(A \circ B)C^t}$, and it could easily be seen that this specification of the PARAFAC solution was equivalent to the vertical stacking of the $\mathbf{X_j} = \mathbf{AB_jC^t}$ matrices to form the array ***X***. To demonstrate this, consider that the matrix $\mathbf{X_j}$ was defined as

$$\mathbf{X_j} = \begin{bmatrix} \sum_r a_{1r}b_{jr}c_{1r} & \cdots & \sum_r a_{1r}b_{jr}c_{Kr} \\ \vdots & \vdots & \vdots \\ \sum_r a_{Ir}b_{jr}c_{1r} & \cdots & \sum_r a_{Ir}b_{jr}c_{Kr} \end{bmatrix}_{\mathbf{I\times K}}$$

, so that the array would consist of a vertical stacking of these k matrices, $\boldsymbol{X} = \begin{bmatrix} \mathbf{X_1} \\ \vdots \\ \mathbf{X_J} \end{bmatrix}_{\mathbf{IJ\times K}}$.

Therefore, $$\boldsymbol{X} = \begin{bmatrix} \sum_r a_{1r}b_{1r}c_{1r} & \cdots & \sum_r a_{1r}b_{1r}c_{Kr} \\ & \vdots & \\ \sum_r a_{Ir}b_{Jr}c_{1r} & \cdots & \sum_r a_{Ir}b_{Jr}c_{Kr} \end{bmatrix}_{\mathbf{IJ\times K}}$$; and equivalently,

$$(\mathbf{A} \circ \mathbf{B})\mathbf{C}^t = \begin{bmatrix} \sum_r a_{1r}b_{1r}c_{1r} & \cdots & \sum_r a_{1r}b_{1r}c_{Kr} \\ & \vdots & \\ \sum_r a_{Ir}b_{Jr}c_{1r} & \cdots & \sum_r a_{Ir}b_{Jr}c_{Kr} \end{bmatrix}_{\mathbf{IJ\times K}}.$$

With the utilization of KR products, the PARAFAC solution was no longer encumbered by the awkwardness of loading matrices combined by weights from a third loading matrix. Instead, the PARAFAC solution could now be viewed as simply the outer product of two matrices. It was now possible to investigate uniqueness in terms of the properties and theorems provided by linear algebra.

3.1.2 The Application of Column Spaces (ten Berge and Sidiropoulos)

The KR product representation of the PARAFAC solution as $\boldsymbol{X} = (\mathbf{A} \circ \mathbf{B})\mathbf{C}^t$, suggested that the columns of $\boldsymbol{X}$ were linear combinations of the columns of the KR product $(\mathbf{A} \circ \mathbf{B})$ or $\boldsymbol{X} = \left[(\mathbf{A} \circ \mathbf{B})\mathbf{c}_1 \quad (\mathbf{A} \circ \mathbf{B})\mathbf{c}_2 \quad \cdots \quad (\mathbf{A} \circ \mathbf{B})\mathbf{c}_k\right]_{IJxK}$. Hence, the columns of $\boldsymbol{X}$ were generated as combinations of the columns of $(\mathbf{A} \circ \mathbf{B})$. Since the column space of a matrix is defined as the set of all combinations of the columns of the matrix, the column space of the KR product $(\mathbf{A} \circ \mathbf{B})$ consists of all combinations of the columns of $(\mathbf{A} \circ \mathbf{B})$. Hence, the array $\boldsymbol{X}$ would have columns that were elements of the column space of $(\mathbf{A} \circ \mathbf{B})$. A basis for this column space is, of course a set of linearly independent spanning vectors.

Liu and Sidiropoulos (2001) were able to show that a necessary condition for uniqueness was that the KR product had full-column rank, and so discussions of uniqueness conditions would be confined to the subset of PARAFAC solutions where the rank of the KR product was equal to the number of columns. Therefore, since a necessary condition for uniqueness was that the columns of the KR product were linearly independent, the columns of $(\mathbf{A} \circ \mathbf{B})$ form a linearly independent set of R vectors, or were a basis for the column space of $(\mathbf{A} \circ \mathbf{B})$.

Consequently, as explained by ten Berge and Sidiropoulos (2002), any other basis in KR product form that existed for the column space of $(\mathbf{A} \circ \mathbf{B})$ would form an alternative KR product that could be used to form an alternative PARAFAC solution. For a complete alternative PARAFAC solution that would reproduce the array $\boldsymbol{X}$, it was only necessary to choose the third loading matrix to be the projection matrix that resulted from projecting the columns of the alternative basis onto the column space of $\boldsymbol{X}$. It was important to note, however, that in order for this alternative basis to result in an alternative PARAFAC solution, the vectors of the basis had to be in KR product form. The alternative solution was non-unique if the matrices that composed the KR product were not permutation-scale versions of **A** and **B**. On the other hand, the alternative solution was unique if the only KR basis that would reproduce the array was comprised of permutation-scale versions of the original KR product. Thus, the non-uniqueness or uniqueness of a PARAFAC solution depended on whether or not a KR basis for the column space of $(\mathbf{A} \circ \mathbf{B})$, which was not composed of permutation-scale versions of the original loading matrices, could be found.

3.1.3 Symmetry

The symmetry of the PARAFAC decomposition had long been established as a property of PARAFAC. Although the idea of symmetry had only been applied in the matrix representation, and not Khatri-Rao product representation, of PARAFAC decompositions, the symmetry of solutions would follow from the equivalence of the two representations.

A PARAFAC decomposition could be represented as $\boldsymbol{X} = (\mathbf{A} \circ \mathbf{B})\mathbf{C}^{\mathbf{t}}$, where **A**, **B**, and **C** were loading matrices with dimensions $\mathrm{I} \times \mathrm{R}$, $\mathrm{J} \times \mathrm{R}$, and $\mathrm{K} \times \mathrm{R}$, respectively. Therefore, the resulting $\boldsymbol{X}$ was an array with dimension $\mathrm{IJ} \times \mathrm{K}$. However, it was possible to rearrange the elements of the array such that it had dimension $\mathrm{IK} \times \mathrm{J}$, resulting in a solution of the form $(\mathbf{A} \circ \mathbf{C})\mathbf{B}^{\mathbf{t}}$. This rearrangement could be continued for every permutation of the loading matrices **A**, **B**, and **C**. All such arrangements were considered equivalent. This ability to rearrange the loading matrices is popularly referred to as *symmetry of the PARAFAC decomposition*. Hence, by symmetry $\boldsymbol{X}_{IJ\times K} = (\mathbf{A} \circ \mathbf{B})\mathbf{C}^{\mathbf{t}}$ if and only if $\boldsymbol{X}_{IK\times J} = (\mathbf{A} \circ \mathbf{C})\mathbf{B}^{\mathbf{t}}$ if and only if $\boldsymbol{X}_{JK\times I} = (\mathbf{B} \circ \mathbf{C})\mathbf{A}^{\mathbf{t}}$. Therefore, the arrangement of loading matrices in a PARAFAC solution had no effect on whether the solution was unique or non-unique.

3.1.4 Simplification of Loading Matrices (ten Berge and Sidiropoulos)

ten Berge and Sidiropoulos (2002) sought to investigate the properties of the column spaces of the KR products to provide insights to Kruskal's uniqueness theorem (Theorem 2.2). Of course, the KR products of loading matrices offered a variety of possibilities. ten Berge and Sidiropoulos, however, were able to reduce the myriad of KR product options by considering loading matrices in simplified form or, more formally, Reduced Row Echelon Form (RREF). Loading matrices with the same rank and k-rank would share the same RREF, differing only in the values of the coefficients involved in the linear combination of the columns. By utilizing RREF, ten Berge and Sidiropoulos were able to classify matrices and KR products into classes and evaluated the uniqueness of these classes.

The justification for working with matrices in RREF was provided informally in the 2002 publication. Although the principles for simplification were simple ideas from linear algebra, they had only been colloquially applied to trilinear models. The following theorems regarding reduced forms are provided as formal support to the claims presented by ten Berge and Sidiropoulos. An underlying assumption to the utilization of RREF was that the matrix structure of an alternative loading matrix would be the same as the original loading matrix. In other words, a PARAFAC solution with loading matrices having a particular set of ranks and k-ranks would not be considered an alternative if the ranks and k-ranks were different from the original.

Theorem 3.1 $\mathbf{AC_kB^t} = \mathbf{\tilde{A}\tilde{C}_k\tilde{B}^t}$ if and only if $\mathbf{SAC_kB^t} = \mathbf{S\tilde{A}\tilde{C}_k\tilde{B}^t}$, for nonsingular matrices **S.**

Proof 3.1 By the definition of nonsingular matrices, $\mathbf{S^{-1}S} = \mathbf{I}$. Therefore, $\mathbf{AC_kB^t} = \mathbf{\tilde{A}\tilde{C}_k\tilde{B}^t} \Leftrightarrow \mathbf{S^{-1}SAC_kB^t} = \mathbf{\tilde{A}\tilde{C}_k\tilde{B}^t} \Leftrightarrow \mathbf{SAC_kB^t} = \mathbf{S\tilde{A}\tilde{C}_k\tilde{B}^t}$. □

Theorem 3.2 $\mathbf{AC_kB^t} = \mathbf{\tilde{A}\tilde{C}_k\tilde{B}^t}$ if and only if $\mathbf{SAC_k^*}\left(\mathbf{TB}\right)^{\mathbf{t}} = \mathbf{S\tilde{A}\tilde{C}_k^*}\left(\mathbf{T\tilde{B}}\right)^{\mathbf{t}}$, where $\mathbf{C_k^*}$ was a diagonalization of the k^{th} row of **UC,** for any invertible matrices **S**, **T**, and **U**.

Proof 3.2 By Theorem 3.1, $\mathbf{AC_kB^t} = \mathbf{\tilde{A}\tilde{C}_k\tilde{B}^t} \Leftrightarrow \mathbf{SAC_kB^t} = \mathbf{S\tilde{A}\tilde{C}_k\tilde{B}^t}$. Likewise, $\mathbf{AC_kB^t} = \mathbf{\tilde{A}\tilde{C}_k\tilde{B}^t} \Leftrightarrow \mathbf{SAC_kB^tT^t} = \mathbf{S\tilde{A}\tilde{C}_k\tilde{B}^tT^t}$. By the symmetry property of PARAFAC solutions, $\mathbf{SAC_kB^tT^t} = \mathbf{S\tilde{A}\tilde{C}_k\tilde{B}^tT^t} \Leftrightarrow \mathbf{C(SA)_kB^tT^t} = \mathbf{\tilde{C}(S\tilde{A})_k\tilde{B}^tT^t}$. By Theorem 3.1, $\mathbf{UC(SA)_kB^tT^t} = \mathbf{U\tilde{C}(S\tilde{A})_k\tilde{B}^tT^t} \Leftrightarrow \mathbf{UC(SA)_kB^tT^t} = \mathbf{U\tilde{C}(S\tilde{A})_k\tilde{B}^tT^t}$. By the symmetry property of PARAFAC solutions, $\mathbf{UC(SA)_kB^tT^t} = \mathbf{U\tilde{C}(S\tilde{A})_k\tilde{B}^tT^t} \Leftrightarrow \mathbf{SA(UC)_kB^tT^t} = \mathbf{S\tilde{A}(U\tilde{C})_k\tilde{B}^tT^t}$. Therefore, $\mathbf{AC_kB^t} = \mathbf{\tilde{A}\tilde{C}_k\tilde{B}^t} \Leftrightarrow \mathbf{SAC_k^*}\left(\mathbf{TB}\right)^{\mathbf{t}} = \mathbf{S\tilde{A}\tilde{C}_k^*}\left(\mathbf{T\tilde{B}}\right)^{\mathbf{t}}$. □

Thus, PARAFAC solutions could be multiplied by nonsingular matrices and the uniqueness of the solution would be unaffected. Reduced row echelon forms are created by pre-multiplying a matrix by a particular nonsingular matrix. The result of this pre-multiplication is a matrix where the linear combinations and dependencies are observable. Additionally, matrices with linearly dependent rows will have reduced row echelon forms with rows of zeroes. These linearly dependent rows are redundant and can be removed since they do not affect uniqueness or non-uniqueness. Therefore, since Reduced Row Echelon Forms were found by the pre-multiplication of nonsingular matrices, a discussion of uniqueness with simplified matrices, or those in RREF, was analogous to considering matrices that had been pre-multiplied by nonsingular matrices.

The idea of working with simplified forms afforded ten Berge and Sidiropoulos an approach to examining the KR products and the ensuing PARAFAC solutions. The realization that simplified forms would provide the essence of the KR products allowed ten Berge and Sidiropoulos to analyze the structure of PARAFAC solutions, an analysis that had been inhibited by the mathematical complexities of considering the many possibilities of loading matrices in an awkward structure.

3.1.5 Denying Kruskal to Show Necessity (ten Berge and Sidiropoulos)

The use of KR products to aid in representing the PARAFAC solutions was a pivotal change in the way uniqueness had been approached. Most importantly, KR products removed the mathematical "mystery" that had cloaked PARAFAC solutions so that simpler ideas from linear algebra could be used to investigate decompositions and alternatives. ten Berge and Sidiropoulos further facilitated the research of PARAFAC solution uniqueness in the application of simplified forms to summarize classes of loading matrices, KR products, and solutions. With the framework for discussing PARAFAC solution uniqueness in place, it was possible to evaluate the seminal theorem for uniqueness, Kruskal's k-rank sum condition.

The details of ten Berge and Sidiropoulos' approach using KR products and simplified forms will be outlined in the following sections. However, it is important to understand the general approach before delving into the details of column spaces and alternative KR products.

The approach of ten Berge and Sidiropoulos was similar to others before - begin with Kruskal's k-rank condition and expand from there. However, unlike previous uniqueness discussions, ten Berge and Sidiropoulos looked at cases where Kruskal's condition was not met, when $k(\mathbf{A}) + k(\mathbf{B}) + k(\mathbf{C}) < 2R + 2$. Kruskal's k-rank condition was a sufficient condition for uniqueness; meaning that when the condition was met, the PARAFAC solution was unique. Prior to the work of ten Berge and Sidiropoulos, it had been assumed by some, conjectured by others, always without formal proof, that Kruskal's condition was also necessary. The necessity of Kruskal's condition would mean that a PARAFAC solution was unique only if $k(\mathbf{A})+k(\mathbf{B})+k(\mathbf{C})\geq 2R+2$. ten Berge and Sidiropoulos suggested that it was possible to determine the necessity of Kruskal's condition by considering cases where the condition failed, or where $k(\mathbf{A})+k(\mathbf{B})+k(\mathbf{C})<2R+2$. Finding a unique PARAFAC solution with loading matrices having k-ranks such that $k(\mathbf{A})+k(\mathbf{B})+k(\mathbf{C})<2R+2$, would establish uniqueness when Kruskal's condition was not met. Such a PARAFAC solution would act as a counterexample to the belief that unique PARAFAC solutions existed only if Kruskal's condition was met, negating the claim that the sufficient condition was also necessary.

Hence, ten Berge and Sidiropoulos applied the ideas of KR products and column spaces to investigate whether or not alternative PARAFAC solutions would exist when Kruskal's k-rank condition failed. As described in Chapters 1 and 2, only alternatives that were not permutation-scale versions of the originals would show non-uniqueness. It should be noted that the work of ten Berge and Sidiropoulos was the first to investigate specific cases where Kruskal's condition failed. Until the introduction of KR products, moving outside of Kruskal's condition had been a daunting task and any attempts to do so did not see sufficient return. Thus, the fact that ten Berge and Sidiropoulos were able to describe specific cases where Kruskal's condition failed was an accomplishment in itself.

3.2 The ten Berge and Sidiropoulos Method (tBSM)

The ten Berge and Sidiropoulos Method (tBSM) sought to describe the uniqueness or non-uniqueness for a class of PARAFAC solutions where Kruskal's condition was not met. When alternative solutions that were not permutation-scale transformations of the original solutions were found, then non-uniqueness was confirmed. If all of the solutions for a particular class had alternatives that were non-unique, Kruskal's condition would be necessary and sufficient for that set of solutions. If, however, only permutation-scale alternatives could be found for a set of solutions, Kruskal's condition would not be necessary for that set of solutions. In order to determine the uniqueness or non-uniqueness of these sets of solutions, the tBSM utilized the KR product representation of the PARAFAC solution as well as the employment of simplified forms.

Thus, the solutions dealt with in the tBSM would be simplified PARAFAC solutions which used the KR product structure. Additionally, the tBSM incorporated two existing necessary conditions that had been found for uniqueness. The first was that no loading matrix could have k-rank less than 2 (Krijnen, 1993). The second, found by Liu and Sidiropoulos (2001), required that the columns of the KR product be linearly independent.

3.2.1 The KR Product with Matrices in Simplified Form

A matrix in simplified form, or reduced row echelon form, was shown to be partitioned into an identity matrix with dimension (Rank x Rank) and a matrix of elements of rank (Rank x (R-Rank)), where R represents the number of columns. Therefore, any two matrices in simplified form could both be partitioned into two matrices: the identity partition, denoted as **I*,** and the element partition, denoted as $\mathbf{E_m}$. When the matrix $\mathbf{E_m}$ had at least one column, the loading matrix would not be full-column rank. By definition, a matrix with deficient column rank would have one or more columns which were linear combinations of other columns in the matrix. The k-rank of a matrix, defined in Chapters 1 and 2 as the maximum number of columns which could be combined so that the set of columns were linearly independent. One of the properties of matrices with deficient column rank in reduced form was that $\mathbf{E_m}$ would be composed of the coefficients involved in the linear combinations of the columns. Hence, for a particular column of $\mathbf{E_m}$, the number of nonzero elements would indicate the number of columns that were combined to form the other columns. Therefore, the smallest number of nonzero elements in a column of $\mathbf{E_m}$ would indicate the k-rank of a matrix. Since the matrix $\mathbf{E_m}$ had the same number of rows as the rank of the unreduced matrix, when k-rank < rank there would exist zeroes in the columns of $\mathbf{E_m}$. In fact, there would be exactly rank – k-rank elements that were zero in at least one of the columns of $\mathbf{E_m}$.

The KR product of any two matrices is simply the column-wise Kronecker product of the two matrices. Therefore, the resulting KR product of any two matrices in simplified form will have some columns that were created from the Kronecker product of the identity partition of each of the two matrices. Let i^* = min ($dim(\mathbf{I}^*_1)$, $dim(\mathbf{I}^*_2)$). Hence, the KR product will have i^* columns with one element equal to 1 and the others 0. Therefore, the column space of the KR product would be partially defined by these i^* columns. The remaining $R - i^*$ columns will have elements in Kronecker product combination from the other partitions, and may or may not have an effect on the column space of the KR product depending on where the elements align in relation to the 1's in the i^* columns. Elements from the $\mathbf{E_m}$ partition that do align with the 1's in the i^* columns may be replaced without impact on the column space of the KR product.

This property would be utilized by ten Berge and Sidiropoulos to determine alternative forms. On the other hand, elements aligned with 0's in the i^* columns could not be replaced without affecting the column space of the KR product. These KR products would have to be treated differently in order to determine alternatives.

3.2.2 Strategies for Determining Alternative KR Products with Column Spaces

A fundamental component to these strategies and to the strategies described later in the dissertation is the idea of the symmetry of the decomposition. In order to de-emphasize whether the **A**, **B**, or **C** matrix is being discussed, loading matrices will simply be referred to as $\mathbf{M_1}$, $\mathbf{M_2}$, and $\mathbf{M_3}$. This convention will be used throughout the remainder of the dissertation.

The tBSM was to determine alternative KR products where the column space of the original KR product and the alternative KR product were the same. Hence, ten Berge and Sidiropoulos began with loading matrices in reduced form, found the KR product and sought to alter the KR product so that the column space would not be affected. The tBSM employed essentially two strategies for finding alternative KR products. The first was to swap out elements that would not change the column space and the second was to multiply transformation matrices to the KR product. Although these two strategies were effective for all the cases that arose for R = 3, they were not sufficient for describing all of R = 4. Hence another more ad-hoc method would have to be imposed. This alternative method will be described in a later section.

The first strategy of replacing elements in the KR product was employed when elements aligned with the 1's in the *i** columns and could be replaced without influencing the column space. In cases where replacing elements led to an alternative KR product that was not a permutation-scale transformation of the original, it was established that an alternative KR form with the same column space as the original existed. Hence, the KR product could then be separated into two loading matrices which were alternative, non-permutation-scale versions of the originals, violating claims of uniqueness.

In other cases where there was not alignment with 1's in the *i** columns, replacing the elements altered the column space, and it was necessary to employ the second strategy. The second strategy incorporated the matrix structure of the third matrix or the matrix not involved in the KR product. When the original third matrix, for example $\mathbf{M_3}$, was nonsingular, the simplified form would be the identity matrix with R rows and R columns. In this case, the columns of the alternative KR basis would also be combined by an alternative *nonsingular* matrix such that $(\mathbf{M_1} \circ \mathbf{M_2}) = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$. Since $\mathbf{M_3}$ was assumed to be nonsingular, $(\mathbf{M}_1 \circ \mathbf{M}_2)$ and $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ spanned the same space. Consequently, any nonsingular matrix $\mathbf{W}$ could transform $(\mathbf{M}_1 \circ \mathbf{M}_2)$ into $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$, or $(\mathbf{M_1} \circ \mathbf{M_2})\mathbf{W} = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$. A PARAFAC solution would be unique if the only matrix $\mathbf{W}$ that transformed $(\mathbf{M}_1 \circ \mathbf{M}_2)$ into $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ was permutation-scale. Hence, if a matrix $\mathbf{W}$ could be found that was not a permutation-scale matrix, an alternative KR product that was not a permutation-scale version of the original would have been discovered. Consequently, the loading matrices of this KR product could be identified and two alternative, non-permutation-scale loading matrices would result, yielding evidence that the PARAFAC decomposition was non-unique.

The tBSM was able to identify whether alternative, non-permutation scale KR products existed, answering the question of uniqueness for the PARAFAC solution in certain cases. In the tBSM, the column space of the KR product involves only two of the three loading matrices. If these two matrices can be transformed using non-permutation-scale matrices, then the PARAFAC decomposition cannot be unique. It only remains a linear algebra exercise to find the alternative third matrix needed to complete the alternative PARAFAC solution. However, if the investigation of the KR product results in transformations that can only be permutation-scale, uniqueness of the PARAFAC decomposition cannot be determined without the consideration of the third matrix. Even though the investigation of column spaces could only evaluate two matrices at a time, answering the question of uniqueness would require the consideration of all three loading matrices in a PARAFAC solution. However, the symmetry property would prove to be very helpful in using the findings from pairs of loading matrices and their KR products to identify unique PARAFAC solutions.

3.2.3 Using Alternative KR Products to Identify Uniqueness

For every three-way PARAFAC solution, there would be six possible KR products formed from six pairwise permutations of the loading matrices. From the symmetry of PARAFAC solutions, it would be sufficient to consider only the combinations, which resulted in three pairwise combinations of loading matrices and three subsequent KR products. The tBSM, using either strategy 1 or strategy 2, considered a single combination of matrices and once an alternative that was not a permutation-scale version of the original was found, all PARAFAC solutions having loading matrices with the same matrix properties as the ones in the pair were identified as non-unique. As will be discussed in the following sections, with the exception of two cases when the loading matrices had four columns, the evaluation of these pairs showed that non-permutation-scale transformations were possible in every case.

3.2.3.1 The tBSM when R = 3

The tBSM began with the case where a PARAFAC solution resulted in loading matrices with three columns, or R = 3. The only PARAFAC solutions considered were where Kruskal's condition failed, that is when $k(\mathbf{M_1}) + k(\mathbf{M}_2) + k(\mathbf{M}_3) < 2R + 2$. In this case, only two types of PARAFAC solutions were possible: one with matrices having rank and k-rank of 2 and the other where two of the matrices had k-rank and rank of 2 and the other had rank and k-rank of 3. In both of these cases, one of the pairwise combinations would involve the two matrices with rank and k-rank of 2. The tBSM utilized strategy 1 and found that KR products formed from two matrices with k-rank and rank of 2 would have an alternative that was not permutation-scale. Thus, when all of the loading matrices had rank and k-rank of 2 the resulting KR products would all have alternatives which preserved the column space but were not permutation-scale.

Hence, the PARAFAC solution with three loading matrices with rank and k-rank of 2 would be non-unique. When the third matrix had rank and k-rank of 3, or had full-column rank, one of the pairwise combinations would involve matrices with k-ranks of 2 and 3. However, one of the combinations of matrices would involve the matrices with k-ranks of 2. It was already established that the resulting KR product with two matrices with k-rank = 2 had an alternative KR product with the same column space. Thus, the PARAFAC solution with two loading matrices having ranks and k-ranks of 2 and one with k-rank = rank = 3 was also non-unique. In summary, the tBSM resulted in the following for R = 3:

- Limiting cases to $k(\mathbf{M_1}) + k(\mathbf{M}_2) + k(\mathbf{M}_3) < 2R + 2$ limited the size of k-rank (and rank);
- Every form of the KR product allowed alternatives that were not permutation-scale transformations; and,
- Only two RREFs needed to be considered.

In conclusion, the symmetry of PARAFAC solutions facilitated the investigation of only one of the KR product pairs. Since the arrangement of the loading matrices in a PARAFAC solution had no effect on uniqueness, it was only necessary to identify one pairwise combination of loading matrices that would result in a KR product with an alternative having the same column space as the original. Thus, the tBSM was used to identify all the types of PARAFAC solutions for R = 3 where Kruskal failed. Since every solution where $k(\mathbf{M_1}) + k(\mathbf{M}_2) + k(\mathbf{M}_3) < 2R + 2$ was found to have KR products that were not permutation-scale transformations, it was established for R = 3 that only PARAFAC solutions where $k(\mathbf{M_1}) + k(\mathbf{M}_2) + k(\mathbf{M}_3) \geq 2R + 2$ would be unique. Hence, the tBSM was able to demonstrate that Kruskal's condition was necessary and sufficient for R = 3.

3.2.3.2 The tBSM when R = 4

Since the tBSM was able to establish the necessity of Kruskal for R = 3, it was applied for R = 4 in order to further investigate if the assumption of Kruskal's condition as necessary and sufficient for all R was valid. The tBSM utilized the same strategies as in R = 3, considering KR products and determining whether or not alternatives which preserved the column space would exist.

The tBSM achieved similar results as in the case where R = 3, finding at least one KR product for every PARAFAC decomposition where the alternative loading matrices were not restricted to be permutation-scale versions of the originals. However, for two of the PARAFAC decompositions discussed by ten Berge and Sidiropoulos, non-uniqueness was not identified using column spaces. One of the decompositions was composed of one matrix with full-column rank and two with k-rank of 2 and rank of 3. The column space approach of ten Berge and Sidiropoulos demonstrated that decompositions of this form could be unique. The other decomposition was composed of three matrices which all had k-rank = rank = 3. In this case, the column space approach failed to offer evidence of uniqueness or non-uniqueness. These two decompositions will be investigated in detail in the remaining sections of this chapter.

Consider the class of PARAFAC solutions where two of the matrices had k-rank of 2 and rank of 3, and the 0 in the last column of both matrices occurred in different rows. In this case, when the third matrix had full column rank or k-rank = rank = 4, it was found that the alternatives were restricted to permutation-scale versions of the original. Although the k-ranks of these loading matrices summed to 8 which was less than 2R + 2 = 10, the PARAFAC solution was identified as unique. Hence, uniqueness was discovered when Kruskal's condition was not met. Therefore, by definition, Kruskal's condition was not necessary for R = 4. However, ten Berge and Sidiropoulos noticed that when the zero of the last column was in the same row position, non-uniqueness resulted.

From this discrepancy, ten Berge and Sidiropoulos hypothesized that the position of the zero had something to do with uniqueness or non-uniqueness of the PARAFAC solution. When k-rank and rank were equal, the non-Identity partition of the reduced form would have all non-zero elements. Thus, ten Berge and Sidiropoulos chose to focus on PARAFAC solutions where rank and k-rank were equivalent. Additionally, since there were no zero elements to consider, only one reduced form resulted for each k-rank, reducing the number of cases that would need analysis.

The strategies employed by ten Berge and Sidiropoulos were applied to the cases where k-rank = rank to determine whether solutions were unique or non-unique. Except for the case, mentioned earlier, where the k-ranks of all three loading matrices were 3, the application of the tBSM suggested that PARAFAC solutions were non-unique when k-rank = rank and R = 4. For the case where k-rank = rank = 3 for all three loading matrices, the strategies employed by the tBSM were not applicable.

The first strategy of replacing elements was not appropriate when the KR product was formed from two matrices with k-rank = rank = 3 because the column space was not preserved. ten Berge and Sidiropoulos explained that the second strategy would not apply because it required that one of the matrices be full column-rank or k-rank = rank = 4. Therefore, the tBSM could not be applied to evaluate uniqueness. Thus, ten Berge and Sidiropoulos offered another approach, presented in the appendix of the 2002 paper, to determine the uniqueness of the PARAFAC solution having three loading matrices with k-rank and rank of 3. It is important that their "appendix method" be addressed since it led to conjectures for all R.

3.2.4 Strategy 3 (The Appendix Method)

Since each of the strategies in the tBSM failed when every KR product pair combination resulted in a scenario where the third matrix must be singular and the elements in the KR product could not be replaced, ten Berge and Sidiropoulos had to construct an alternative approach to determining uniqueness. Although this method was presented without formal proof, the ten Berge and Sidiropoulos offered a numerical example to demonstrate that it could be used to show that PARAFAC solutions were non-unique in the case where R = 4 and the ranks and k-ranks of all three loading matrices were 3.

3.2.4.1 The Theory of the Appendix Method

The method presented in the appendix made use of a series of simplifications in order to reduce the complexities involved with the KR product and PARAFAC solutions. The first simplification that was made was to transform the loading matrices so that two of the loading matrices had identical forms.

Thus, $\mathbf{M_1}$ and $\mathbf{M}_2$ were transformed from Reduced Row Echelon Form to $\mathbf{M_1} = \mathbf{M_2} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$, where the weights from the Reduced Row Echelon Form of $\mathbf{M_1}$ and $\mathbf{M}_2$ were reassigned and incorporated into the weights in the Reduced Row Echelon Form of the third matrix, $\mathbf{M}_3$. Therefore, $M_3 = \begin{bmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & z \end{bmatrix}$, where *x*, *y*, and *z* were the combined weights of all three loading matrices. Finally, some elementary row operations were performed on $\mathbf{M}_3$ for continued simplification.

The next simplification that was made was to assume that two of the alternative loading matrices were equal, or that $\tilde{\mathbf{M}}_\mathbf{1} = \tilde{\mathbf{M}}_\mathbf{2}$. This simplification was a natural result from the simplification that $\mathbf{M_1}$ and $\mathbf{M}_2$ were transformed into two matrices with identical forms.

A further modification to the problem was in the creation of the general form for $\tilde{\mathbf{M}}_1$. The first row of the $\tilde{\mathbf{M}}_1$ matrix was composed of all 1's. This requirement of a row of 1's prevented the alternative loading matrix $\tilde{\mathbf{M}}_1$ from having columns that were permutation-scale versions of the columns of $\mathbf{M}_1$, and similarly for $\tilde{\mathbf{M}}_2$ and $\mathbf{M}_2$. Thus, the two alternative loading matrices were represented by

$$\tilde{\mathbf{M}}_1 = \tilde{\mathbf{M}}_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ g_1 & g_2 & g_3 & g_4 \\ h_1 & h_2 & h_3 & h_4 \end{bmatrix}.$$

The KR product was then created by taking the column-wise product of $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$. The array $\boldsymbol{X}$ was created from the loading matrices $\mathbf{M}_1$, $\mathbf{M}_2$, and $\mathbf{M}_3$, such that $\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_3^t$. Thus, an alternative PARAFAC solution demonstrating non-uniqueness would exist if $\boldsymbol{X} = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$, where $\tilde{\mathbf{M}}_3^t = [\mathbf{u} \,|\, \mathbf{v} \,|\, \mathbf{w}]_{4\times 3}$. Assuming the equality of $\boldsymbol{X}$ and $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$, the problem was further simplified by removing all redundant rows from $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ and the corresponding rows from the array $\boldsymbol{X}$. The resulting $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ had all distinct rows.

Additional simplifications were made by considering the linear dependencies of the rows. The first two columns of the array $\boldsymbol{X}$ were non-proportional; and, since $\boldsymbol{X} = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$, the column vectors u and v from the loading matrix $\tilde{\mathbf{M}}_3$ were non-proportional as well. The first two columns of the reduced array $\boldsymbol{X}$ had zero elements in the first three positions, indicating that the first three rows of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ were orthogonal to the column vectors u and v when $\boldsymbol{X} = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$. Thus, in order for $\boldsymbol{X} = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$, the first three rows would be linearly dependent. Using the definition of linear dependence, it was found that $g_i h_i = \lambda g_i + \delta h_i$, $i = 1, \ldots, 4$ for some scalar λ and δ. Therefore, another simplification was made by replacing the elements in the last row of $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$

by $h_i = \dfrac{\lambda g_i}{g_i + \lambda - 1}$, when $g_i \neq 1 - \lambda$ and removing the third row of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ which was linearly dependent on the first and second rows.

Thus, the KR product of $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$ could be represented in terms of 1's and g_i, $i = 1, \ldots, 4$, so that determining $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$ simply required determining g_i, $i = 1, \ldots, 4$. In order for $\boldsymbol{X} = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$, the g_i must be chosen so that *X* is in the column space of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$. ten Berge and Sidiropoulos found that the orthogonal complement space of $\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_3^t$ was spanned by vectors of the form $\mathbf{n} = \begin{bmatrix} -0.5w + \alpha & -0.5w - \alpha & yz & xz & xy \end{bmatrix}^{\mathbf{t}}$, where $w = xy + xz + yz + xyz$ and α was an arbitrary scalar. Therefore, the method to find g_i was to find where the columns of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ were orthogonal to the vector **n**.

After making another simplification by allowing $\lambda = 2$ and rewriting a column of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ as $(\mathbf{g} \circ \mathbf{h})_{\mathbf{i}} = \begin{bmatrix} g_i (g_i + 1)^2 & 2g_i (g_i + 1) & (g_i + 1)^2 & g_i^2 (g_i + 1)^2 & 4g_i^2 \end{bmatrix}^{\mathbf{t}}$, it was possible to solve for g_i such that $\mathbf{n}^{\mathbf{t}} (\mathbf{g} \circ \mathbf{h})_{\mathbf{i}} = 0$, $\forall i, i = 1, \ldots, 4$. Additionally, since each column of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ was identical except for the index *i*, it was possible to ignore the index *i* and simply find all roots to the quadratic equation that resulted from $\mathbf{n}^{\mathbf{t}} (\mathbf{g} \circ \mathbf{h}) = 0$ or to find the roots of $c_4 g^4 + c_3 g^3 + c_2 g^2 + c_1 g + c_0 = 0$, where

$$
\begin{aligned}
c_4 &= xz \\
c_3 &= -0.5w + \alpha + 2xz \\
c_2 &= -2w + xz + 4xy + yz\,. \\
c_1 &= -1.5w - \alpha + 2yz \\
c_0 &= yz
\end{aligned}
$$

Since α was allowed to be any arbitrary scalar, it was assumed that $\alpha \neq 0$, and the problem of solving for roots was again simplified by dividing the coefficients of the polynomial by α. This division resulted in new coefficients for the polynomial,

$$c_4 = \frac{xz}{\alpha}$$
$$c_3 = \frac{-0.5w + 2xz}{\alpha} + 1$$
$$c_2 = \frac{-2w + xz + 4xy + yz}{\alpha} \; .$$
$$c_1 = \frac{-1.5w + 2yz}{\alpha} - 1$$
$$c_0 = \frac{yz}{\alpha}$$

When $\alpha \to \infty$, these coefficients would tend to

$$c_4 = 0$$
$$c_3 = 1$$
$$c_2 = 0 \;,$$
$$c_1 = -1$$
$$c_0 = 0$$

so that the polynomial became $\left(g^3 - g\right) = 0$, which had roots 1, -1, and 0. For the fourth root, the well-known Vieta formula was used and as $\alpha \to \infty$, this root would tend to $\pm\infty$ depending on the sign of *xz*.

It was shown that for large α, the roots of the polynomial would be distinct nonzero real roots and could be used to complete the rows of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$. The resulting columns were then orthogonal to **n** and could be transformed by elementary row operations to form a Vandermonde matrix, which is nonsingular when the roots are distinct. Hence, the columns of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ were orthogonal to **n** and were linearly independent, and so the columns of $\boldsymbol{X}$ were contained in the column space of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$. It only remained to find a matrix $\tilde{\mathbf{M}}_3$ so that $\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_3^t = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$.

Since the KR product $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ was created from loading matrices that were not permutation-scale versions of $\mathbf{M}_1$ and $\mathbf{M}_2$, it appeared that a PARAFAC solution with k-rank = rank = 3 for all loading matrices was non-unique. In order to demonstrate this further, a numerical example was presented.

3.2.4.2 The Numerical Application of the Appendix Method

In the numerical example, $x = 1$, $y = 2$, and $z = 9$. The arbitrary scalar α was chosen large enough, 5.5, so that the roots of the polynomial were distinct, nonzero, and real. The roots that resulted were

$$\begin{bmatrix} 2.8080 & -1.9257 & -1.1925 & 0.3102 \end{bmatrix}.$$

Thus, the remaining rows of the matrices $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$ were completed using these roots so that

$$\tilde{\mathbf{M}}_1 = \tilde{\mathbf{M}}_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2.8080 & -1.9257 & -1.1925 & 0.3102 \\ 1.4748 & 4.1606 & 12.3911 & 0.4735 \end{bmatrix}.$$

Furthermore, using $\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_3^t$, the matrix $\tilde{\mathbf{M}}_3$ was found by solving $\tilde{\mathbf{M}}_3^t = \left[\left\{ (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)^t (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2) \right\}^{-1} (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)^t \boldsymbol{X} \right]_{4\times 3}$. With this $\tilde{\mathbf{M}}_1$, $\tilde{\mathbf{M}}_2$ and $\tilde{\mathbf{M}}_3$, a PARAFAC solution was found such that $\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_3^t = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$. Because the loading matrices $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$ were not permutation-scale versions of $\mathbf{M}_1$ and $\mathbf{M}_2$, the PARAFAC solution was deemed non-unique. Additionally, alternative roots for the polynomial with larger α were found to demonstrate that the roots of the polynomial did indeed tend to 1, -1, 0, and $\pm\infty$.

3.2.4.3 Problems with the Appendix Method

There were several issues within the appendix method that ten Berge and Sidiropoulos failed to address. Greater investigation of these issues revealed problems with the appendix method and numerical example that could not be ignored. In fact, these issues brought into question whether the appendix method was appropriate for solving the uniqueness question in the k-rank = rank = 3 for all loading matrices. Furthermore, it was also discovered that the numerical example suffered from numerical complexities that mislead ten Berge and Sidiropoulos into believing the question regarding uniqueness had been answered in this case.

The issues that arose upon further investigation will be discussed in detail below. In summary, the problem with the appendix method was that allowing $\alpha \to \infty$ caused inconsistencies in other aspects of the solution. Further, even with a small α, as in the numerical example, there was evidence to suggest that all of the loading matrices did not possess k-rank = 3. It will be shown that the appendix method actually suggested a solution for the case where at least one of the alternative loading matrices did not share the same matrix structure as the original(s). This resulting case was not the case intended; and, as a consequence, left the issue of uniqueness for solutions with k-rank = rank = 3 for all loading matrices unresolved.

3.2.4.4 Problems with the Underlying Theory of the Appendix Method

The first issue with the appendix method was in the choice of the vector **n**, the vector form that spanned the orthogonal complement space of the array. At this point, the array ***X*** was reduced to a $\mathbf{5 \times 3}$ matrix. The orthogonal complement space of such a matrix would have 2 basis vectors. For example, two linearly independent vectors that span the space could be $\mathbf{n}_1 = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \end{bmatrix}^{\mathbf{t}}$ and $\mathbf{n}_2 = \begin{bmatrix} \frac{-w}{xy} & 0 & \frac{z}{x} & \frac{z}{y} & 1 \end{bmatrix}^{\mathbf{t}}$. The authors created the vector form, **n**, as a linear combination of these two basis vectors, that is $-(0.5w + \alpha)\mathbf{n}_1 + (xy)\mathbf{n_2} = \mathbf{n}$. Thus, ten Berge and Sidiropoulos were able to find one vector, **n**, so that all vectors orthogonal to **n** would form a subspace that necessarily contained the columns of the array ***X***.

Thus, solving for g_i so that the columns of $(\tilde{\mathbf{M}}_\mathbf{1} \circ \tilde{\mathbf{M}}_\mathbf{2})$ were orthogonal to **n** provided a linear subspace that contained the columns of $\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_\mathbf{3}^\mathbf{t}$. Therefore, all that remained was to find $\tilde{\mathbf{M}}_3$ so that $\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_\mathbf{3}^\mathbf{t} = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_\mathbf{3}^\mathbf{t}$, which was done easily by projecting onto the column space of $(\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_\mathbf{3}^\mathbf{t}$. Hence it was possible to consider only one vector instead of two because the matrix $\tilde{\mathbf{M}}_3$ could be forced to make the equality,

$\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_3^t = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$, hold. The authors failed to note, however, that composing the matrix $\tilde{\mathbf{M}}_3$ such that $\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_3^t = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$ might have affected the matrix structure of $\tilde{\mathbf{M}}_3$.

The second issue requiring attention involved the arbitrary scalar α. The inclusion of an arbitrary α in the linear combination of the $\mathbf{n}_i$, $i = 1,2$, vectors provided ten Berge and Sidiropoulos with a dial, so to speak, so that real roots would result from the polynomial. The existence of a solution did not rely on the value of α to increase to infinity. In fact, as was shown in the numerical application, $\alpha = 5.5$ worked adequately to insure real roots. However, the concept of allowing $\alpha \to \infty$ resulted in several contradictions that should have been addressed.

First, with respect to the vector **n**, allowing $\alpha \to \infty$ would create larger and larger weights on the $\mathbf{n}_1$ vector. Consequently, the vector **n** would no longer be a linear combination of the two vectors $\mathbf{n}_1$ and $\mathbf{n}_2$, but a scaled version of the single vector $\mathbf{n}_1$. Second, ten Berge and Sidiropoulos noted that as $\alpha \to \infty$, the polynomial of interest became $\left(g^3 - g\right)$, with roots 1, -1, and 0. Therefore, $\exists i$ such that $g_i = -1$. ten Berge and Sidiropoulos had previously noted that $h_i = \dfrac{\lambda g_i}{g_i + \lambda - 1}$, when $g_i \neq 1 - \lambda$. With the choice of $\lambda = 2$, this suggested that $g_i \neq -1$. However, as $\alpha \to \infty$, $\exists i$ such that $g_i \to -1$. Additionally, the fourth root was found to increase (or decrease) to $\pm\infty$ when $\alpha \to \infty$. Regardless of the choice of λ, when the fourth root is substituted for g_i, the h_i that resulted was in indeterminate form. Therefore, the $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$ that resulted from $\alpha \to \infty$ would not be plausible loading matrices. In the same vein, ten Berge and Sidiropoulos used the idea that as long as the roots were real and nonzero, the columns of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ would be linearly independent. However, as $\alpha \to \infty$, one of the roots tended to 0, which contradicted the idea that the roots were nonzero.

These discrepancies contributed to a misleading alternative solution, which was evident after greater inspection of the numerical application of the appendix method.

3.2.4.5 Problems with the Numerical Example

In the numerical application, $\alpha = 5.5$. Obviously, the choice of this α was far from one that tended to infinity. However, it was chosen so that the roots of the polynomial would be real. The loading matrices that resulted from this choice of α were $\tilde{\mathbf{M}}_{\mathbf{1}} = \tilde{\mathbf{M}}_{\mathbf{2}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2.8080 & -1.9257 & -1.1925 & 0.3102 \\ 1.4748 & 4.1606 & 12.3911 & 0.4735 \end{bmatrix}$. Although the roots, found in the second row of $\tilde{\mathbf{M}}_{\mathbf{1}}$ and $\tilde{\mathbf{M}}_2$, were far from those that would be observed as α tended to infinity, the third row of $\tilde{\mathbf{M}}_{\mathbf{1}}$ and $\tilde{\mathbf{M}}_2$ already exhibited a pattern consistent with roots tending towards -1. As expected with roots of -1, the elements on the third row corresponding to the root of -1.1925 was the largest. The large value would yield larger values once the KR product of $\tilde{\mathbf{M}}_{\mathbf{1}}$ and $\tilde{\mathbf{M}}_2$ was calculated. In order for an alternative PARAFAC solution to exist using $(\tilde{\mathbf{M}}_{\mathbf{1}} \circ \tilde{\mathbf{M}}_{\mathbf{2}})$, an appropriate matrix $\tilde{\mathbf{M}}_3$ had to be found so that $\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_{\mathbf{3}}^{\mathbf{t}} = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_{\mathbf{3}}^{\mathbf{t}}$. The matrix that resulted from projecting the columns of $(\tilde{\mathbf{M}}_{\mathbf{1}} \circ \tilde{\mathbf{M}}_{\mathbf{2}})$ onto the column space of X was

$$\tilde{\mathbf{M}}_{\mathbf{3}}^{\mathbf{t}} = \begin{bmatrix} -0.0573 & 0.1610 & 0.1781 \\ 0.1239 & 0.2386 & -0.1631 \\ -0.0732 & -0.0873 & 0.0797 \\ 1.0067 & -0.3123 & 0.9053 \end{bmatrix}.$$

Although the values were not extreme, ten Berge and Sidiropoulos failed to consider that all of the elements in the third row of $\tilde{\mathbf{M}}_3^{t}$ (third column of $\tilde{\mathbf{M}}_3$) were approaching 0. Also overlooked was the position of the larger values that resulted from finding the KR product of $\tilde{\mathbf{M}}_{\mathbf{1}}$ and $\tilde{\mathbf{M}}_2$. Interestingly, the larger values in $(\tilde{\mathbf{M}}_{\mathbf{1}} \circ \tilde{\mathbf{M}}_{\mathbf{2}})$ aligned with the values in $\tilde{\mathbf{M}}_3^{t}$ that were approaching 0, a consequence of relying on the matrix $\tilde{\mathbf{M}}_3$ to obtain the columns of the array $\boldsymbol{X}$. Although the case of interest purported to have R = 4, or four columns, the way in which it was solved forced one column of $\tilde{\mathbf{M}}_3$ to have elements with very small values, bringing into question whether the number of columns was truly four.

Since the application of the appendix method was a numerical example, it was difficult to assess the matrix structure of the alternative loading matrices. In general, the idea of k-rank in a numerical setting was an abstract concept, and so verifying the k-rank of $\tilde{\mathbf{M}}_3$ would prove complicated. In theory, the Reduced Row Echelon form of a matrix should provide an indication of the linear dependencies among the columns. Although this was very useful when the matrices of interest were in symbolic form, considering the Reduced Row Echelon form of actual matrices was muddied by numerical intricacies.

Specifically, in this particular numerical application it was believed that the true matrix structure was masked by the complexity of working with values that were quite small. Ignoring that the matrix $\tilde{\mathbf{M}}_3$ might not have the appropriate number of columns, the Reduced Row Echelon Form of $\tilde{\mathbf{M}}_3$ was considered. Using the method suggested by ten Berge and Sidiropoulos, sub matrices of $\tilde{\mathbf{M}}_3$ were used to determine the Reduced Row Echelon form. In the explanation of simplified forms, ten Berge and Sidiropoulos noted that only matrices with nonzero determinants could be used to find the Reduced Row Echelon form of a matrix.

However, when considering sub matrices of $\tilde{\mathbf{M}}_3$, very small determinants resulted, (0.000142, 0.00116, -0.0118, 0.005296). Although these values were very close to 0, the reduction was continued and resulted in three forms with weights of [8.2, -37.3, -83.2], [4.57, 10.2, 0.1225], [0.098, -0.448, -0.012], and [0.219, -2.232, -0.027], respectively. In each of these cases, one of the weights was considerably smaller than the other two, and k-rank < 3 was suspected.

In order to investigate the k-rank of $\tilde{\mathbf{M}}_3$, MAPLE was employed. The weights obtained with the reduction suggested by ten Berge and Sidiropoulos were exactly those found using MAPLE and the ReducedRowEchelonForm function. It was possible, however, to use the ReducedRowEchelonForm function in MAPLE while leaving the *x*, *y*, and *z* in symbolic form. Hence, the final computations to calculate the weights was not performed, and it was possible to view the underlying fractions that composed the final weights that resulted when *x*, *y*, and *z* were assigned values. Using the submatrix with the largest determinant, the three weights that resulted from leaving the *x*, *y*, and *z* in symbolic form involved numerators consisting of a polynomial with very small coefficients. In the case of the first weight, the largest coefficient was on the order of 10^{-9}, which should have been treated as 0, resulting in a zero weight. The denominator for this weight was composed of coefficients with slightly larger values so that when the *x*, *y*, and *z* were assigned values, the numerical computation masked the fact that the numerator was essentially 0. The values are not as small in the second and third weight, and it could be argued that one of these might result in a truly nonzero weight.

Hence, in addition to fewer than four columns, the use of MAPLE suggested that the k-rank of $\tilde{\mathbf{M}}_3$ might be less than 3. Further evidence for less than four columns and k-rank < 3 could be seen when considering larger α. ten Berge and Sidiropoulos used a much larger α, $\alpha = 1005.5$, to demonstrate the asymptotic nature of the roots, finding that this α resulted in roots of

$$\begin{bmatrix} -111.1607 & 1.0362 & -1.0039 & 0.0173 \end{bmatrix}.$$

However, ten Berge and Sidiropoulos did not present what effect composing the $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$ matrices from these roots would have on $\tilde{\mathbf{M}}_3$, which would be needed to complete the equation $\boldsymbol{X} = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{M}_3^t = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)\tilde{\mathbf{M}}_3^t$. The resulting $\tilde{\mathbf{M}}_3^t$, in fact, would be

$$\tilde{\mathbf{M}}_3^t = \begin{bmatrix} 0.00001 & 0.00016 & -0.28577 \\ -0.01686 & 0.01746 & 0.96419 \\ -0.00003 & -0.00003 & 0.00003 \\ 1.01689 & -0.01759 & 0.03577 \end{bmatrix}.$$

The effect of the root closest to -1 was even greater here, as the elements of the third row of $\tilde{\mathbf{M}}_3^t$ (third column of $\tilde{\mathbf{M}}_3$) could be considered 0, suggesting that the matrix structure of $\tilde{\mathbf{M}}_3$ would have fewer than four columns and was not equivalent to $\mathbf{M}_3$. Again using MAPLE and unassigned *x*, *y*, and *z*, choosing the same submatrix as before, the weights that resulted had even smaller coefficients in the numerator than the case where $\alpha = 5.5$, indicating that the issue of matrix structure worsened as α tended to infinity.

As additional evidence to the problem with matrix structure as $\alpha \to \infty$, the numerical solution was ignored and a more generic solution with large α was considered. With very large α, the roots of the polynomial were 1,-1, 0 and *N,* where *N* was a large number. It was already assumed that

$$\tilde{\mathbf{M}}_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ g_1 & g_2 & g_3 & g_4 \\ \\ \dfrac{2g_1}{g_1+1} & \dfrac{2g_2}{g_2+1} & \dfrac{2g_3}{g_3+1} & \dfrac{2g_4}{g_4+1} \end{bmatrix}.$$

The Reduced Row Echelon Form of this matrix was

$$R := \begin{bmatrix} 1 & 0 & 0 & \dfrac{-g_4 g_1 g_2 + g_3 g_2 - g_4 g_2 + g_2 g_1 g_3 + g_4^2 g_1 + g_4^2 - g_4 g_1 g_3 - g_4 g_3}{(-g_2+g_1)(-g_3+g_1)(g_4+1)} \\ 0 & 1 & 0 & -\dfrac{g_4^2 g_2 - g_4 g_3 - g_4 g_1 g_2 + g_3 g_1 - g_4 g_3 g_2 + g_4^2 + g_2 g_1 g_3 - g_4 g_1}{(g_2-g_3)(g_4+1)(-g_2+g_1)} \\ 0 & 0 & 1 & \dfrac{(-g_4+g_1)(g_2-g_4)(g_3+1)}{(g_4+1)(-g_3+g_1)(g_2-g_3)} \end{bmatrix}$$

. Substitution of the roots 1, -1, 0, and N resulted in a Reduced Row Echelon form for $\tilde{\mathbf{M}}_1$ of

$$\begin{bmatrix} 1 & 0 & 0 & \dfrac{2N + 2N^2}{2(N+1)} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\dfrac{(-N+1)(-N-1)}{N+1} \end{bmatrix}.$$

The implication of this reduced form was that $\tilde{\mathbf{M}}_1$ had k-rank<3, indicating that as α tended to infinity, the matrix structure of $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$ tended to diverge from that of $\mathbf{M}_1$ and $\mathbf{M}_2$.

3.2.5 Conclusions and Lessons Learned from the tBSM

The issues discovered in investigating the appendix method of ten Berge and Sidiropoulos should call into question the validity of using this method to assess the uniqueness of a PARAFAC solution. Hence, the conclusions that ten Berge and Sidiropoulos made when the appendix method was applied to the case where all loading matrices had rank and k-rank of 3 were not justified. Therefore, the question of whether or not Kruskal's condition was necessary and sufficient in the case where R = 4 and k-rank = rank was left unanswered. Additionally, the case where all matrices had k-rank of 3 needed to be evaluated before any conclusions could be made regarding the class of PARAFAC solutions where k-rank = rank. Thus, the conclusion that Kruskal's condition was necessary and sufficient in the case where R = 4 and k-rank = rank was premature, and should not be made until the case where the rank = k-rank = 3 for all three loading matrices was decided. It would be necessary to employ a new method to truly determine whether PARAFAC solutions would be unique or non-unique.

3.3 A Strategy for Identifying Uniqueness

One of the essential challenges of understanding the literature, past and current, was extracting a clear and unambiguous definition of what is meant by the *uniqueness* of a PARAFAC solution. This clarity is critical, obviously, for a rigorous mathematical treatment of the topic. The structure of the PARAFAC solution had morphed from one which considered the individual "slabs" of the array to one that represented the entire array. Although the definition of uniqueness and equivalent matrices would be maintained, a new strategy for investigating uniqueness, which encompassed this new representation, would be needed.

Prior to this dissertation, the literature only hinted at extending the definitions of equivalent matrices and uniqueness to KR products. Whereas the original definitions focused on matrices, new definitions would need to focus on KR products. However, it is easy to see that the definitions are synonymous.

The definitions of equivalent matrices and uniqueness were given in the previous chapter. In summary, a PARAFAC decomposition was unique if and only if all alternative PARAFAC solutions were comprised of alternative loading matrices that were permutation-scale versions of the original loading matrices, where the scales were multiplicative inverses, i.e. the alternative loading matrices were defined as $\tilde{\mathbf{M}}_{\mathbf{i}} = \mathbf{M_i \Pi \Lambda_i}$, $\forall i, i = 1,\ldots,3$ and $\mathbf{\Pi}$ was a permutation matrix and $\mathbf{\Lambda}_i$ were diagonal matrices such that $\mathbf{\Lambda_1 \Lambda_2 \Lambda_3} = \mathbf{I_{R \times R}}$.

These definitions, and Kruskal's theorem for uniqueness, focus on the individual loading matrices and not the KR product. Even though it will be shown that there is no difference in considering individual matrices or KR products, for completeness a definition of uniqueness incorporating KR products will be given.

Definition 3.2 A PARAFAC solution was unique if and only if all possible KR products of the alternative loading matrices were permutation-scale versions of the respective KR products of the original loading matrices, i.e. $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$, $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_3)$, and $(\tilde{\mathbf{M}}_2 \circ \tilde{\mathbf{M}}_3)$ were permutation-scale versions of $(\mathbf{M}_1 \circ \mathbf{M}_2)$, $(\mathbf{M}_1 \circ \mathbf{M}_3)$, and $(\mathbf{M}_2 \circ \mathbf{M}_3)$, respectively.

In order to show that the KR product definition and the traditional loading matrix definition for uniqueness were equivalent, it would be necessary to show that permutation-scale alternative KR products would result if and only if alternative loading matrices were permutation-scale transformations.

Theorem 3.3 $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2) = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{\Pi}\mathbf{\Lambda}_{1,2}$ if and only if $\tilde{\mathbf{M}}_1 = \mathbf{M}_1\mathbf{\Pi}\tilde{\mathbf{\Lambda}}_1$ and $\tilde{\mathbf{M}}_2 = \mathbf{M}_2\mathbf{\Pi}\tilde{\mathbf{\Lambda}}_2$, where $\mathbf{\Pi}$ was a permutation matrix and $\mathbf{\Lambda}_{1,2}$ represents the diagonal matrix where the entries are the product of the respective diagonal elements from the diagonal matrices $\tilde{\mathbf{\Lambda}}_1$ and $\tilde{\mathbf{\Lambda}}_2$.

Proof 3.3 *Sufficiency:* Without loss of generality, consider the r^{th} column of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$, denoted as $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)_r$. By the definition of Khatri-Rao products, $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)_r$ was composed from the columnwise Kronecker product of $\tilde{\mathbf{M}}_1$ and $\tilde{\mathbf{M}}_2$, or $\text{vec}(\tilde{\mathbf{m}}_{(1)r}\tilde{\mathbf{m}}_{(2)r}^{\mathbf{t}}) = (\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)_r$, where $\tilde{\mathbf{m}}_{(n)r}$ represents the r^{th} column of the matrix $\tilde{\mathbf{M}}_n$. Therefore, consider $\text{vec}(\tilde{\mathbf{m}}_{(1)r}\tilde{\mathbf{m}}_{(2)r}^{\mathbf{t}})$. Since $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2) = (\mathbf{M}_1 \circ \mathbf{M}_2)\mathbf{\Pi}\mathbf{\Lambda}_{1,2}$, $\text{vec}(\tilde{\mathbf{m}}_{(1)r}\tilde{\mathbf{m}}_{(2)r}^{\mathbf{t}})$ was equivalent to a permuted and scaled column of $(\mathbf{M}_1 \circ \mathbf{M}_2)$. Without loss of generality, let this be the q^{th} column of $(\mathbf{M}_1 \circ \mathbf{M}_2)$ and scaled by the scalar θ. Hence, $\text{vec}(\tilde{\mathbf{m}}_{(1)r}\tilde{\mathbf{m}}_{(2)r}^{\mathbf{t}}) = \theta\,\text{vec}(\mathbf{m}_{(1)q}\mathbf{m}_{(2)q}^{\mathbf{t}})$. Therefore, J, the number of rows in $\mathbf{M}_2$, equivalencies of the form $\tilde{m}_{(1)ir}\tilde{m}_{(2)jr} = \theta m_{(1)iq}m_{(2)jq}$, would exist. Alternatively written, $\frac{\tilde{m}_{(1)ir}}{m_{(1)iq}} = \theta\frac{m_{(2)jq}}{\tilde{m}_{(2)jr}}$, $\forall j$. Hence, $\frac{\tilde{m}_{(1)ir}}{m_{(1)iq}} = \theta\frac{m_{(2)2q}}{\tilde{m}_{(2)2r}}, \cdots, \frac{\tilde{m}_{(1)ir}}{m_{(1)iq}} = \theta\frac{m_{(2)Jq}}{\tilde{m}_{(2)Jr}}$ or $\frac{m_{(2)1q}}{\tilde{m}_{(2)1r}} = \frac{m_{(2)2q}}{\tilde{m}_{(2)2r}} = \ldots = \frac{m_{(2)Jq}}{\tilde{m}_{(2)Jr}}$. Since all the ratios of the form $\frac{m_{(2)jq}}{\tilde{m}_{(2)jr}}$ were equivalent, without loss of generality, the elements of the q^{th} column of $\mathbf{M}_2$ could be represented as

$$\mathbf{m}_{(2)q}^{\mathbf{t}} = \begin{bmatrix} \tilde{m}_{(2)1r}\frac{m_{(2)2q}}{\tilde{m}_{(2)2r}} & m_{(2)2q} & \tilde{m}_{(2)3r}\frac{m_{(2)2q}}{\tilde{m}_{(2)2r}} & \cdots & \tilde{m}_{(2)Jr}\frac{m_{(2)2q}}{\tilde{m}_{(2)2r}} \end{bmatrix} = \left(\frac{m_{(2)2q}}{\tilde{m}_{(2)2r}}\right)\tilde{\mathbf{m}}_{(2)r}^{\mathbf{t}}.$$

Hence, this column of $\mathbf{M}_2$ was simply a scaled version of another column of $\tilde{\mathbf{M}}_2$.

This argument could be repeated for every column of $\mathbf{M}_2$ and $\tilde{\mathbf{M}}_2$, and similar arguments would follow for $\mathbf{M}_1$ and $\tilde{\mathbf{M}}_1$. Hence, by definition, $\tilde{\mathbf{M}}_1 = \mathbf{M}_1\mathbf{\Pi}\tilde{\mathbf{\Delta}}_1$ and $\tilde{\mathbf{M}}_2 = \mathbf{M}_2\mathbf{\Pi}\tilde{\mathbf{\Delta}}_2$, where $\mathbf{\Pi}$ was a permutation matrix and $\tilde{\mathbf{\Lambda}}_1$ and $\tilde{\mathbf{\Lambda}}_2$ were diagonal matrices that contained the appropriate column scaling scalars on the diagonal.

Necessity: The same permutation matrix $\mathbf{\Pi}$ was applied to both $\mathbf{M_1}$ and $\mathbf{M_2}$, rearranging the order of the columns. Although the r^{th} column of $\mathbf{M_1}$ and $\mathbf{M_2}$ would not necessarily be in the same column position in the permutated transformation, the new column position of the r^{th} column of $\mathbf{M_1}$ would be the same as the new position of the r^{th} column of $\mathbf{M_2}$. The r^{th} column of the Khatri-Rao product of $\mathbf{M_1}$ and $\mathbf{M_2}$ was obtained by taking the columnwise Kronecker product of the columns of r^{th} columns $\mathbf{M_1}$ and $\mathbf{M_2}$. Although the r^{th} column of $\mathbf{M_1}$ and r^{th} column of $\mathbf{M_2}$ were in new positions, these columns would still be combined by columnwise Kronecker to form a column in the Khatri-Rao product of the permutated matrices. In other words, when the permutation transformation was applied to $\mathbf{M_1}$ and $\mathbf{M_2}$, the column positions within each of the matrices changed, while the relative positions of the columns in $\mathbf{M_1}$ and $\mathbf{M_2}$ were preserved.

Thus, the Khatri-Rao product of $\mathbf{M_1}$ and $\mathbf{M_2}$ would still have a column that consisted of the columnwise Kronecker product of the r^{th} column of $\mathbf{M_1}$ and r^{th} column of $\mathbf{M_2}$, however it would not necessarily occur in the r^{th} column of $(\mathbf{M_1} \circ \mathbf{M_2})$. In fact, the new column position of the columnwise Kronecker product of the r^{th} column of $\mathbf{M_1}$ and r^{th} column of $\mathbf{M_2}$ would be located in the same new position as the permuted r^{th} column. Hence, the order of the columns of $(\mathbf{M_1} \circ \mathbf{M_2})$ would be permuted in the same way the columns of $\mathbf{M_1}$ and $\mathbf{M_2}$ were permuted.

Hence, $(\mathbf{M_1\Pi} \circ \mathbf{M_2\Pi}) = (\mathbf{M_1} \circ \mathbf{M_2})\mathbf{\Pi}$. Similarly, post-multiplying $\mathbf{M_1\Pi}$ and $\mathbf{M_2\Pi}$ by diagonal matrices simply scaled the columns of $\mathbf{M_1}$ and $\mathbf{M_2}$. Hence, the r^{th} column of $\mathbf{M_1}$ would be scaled by the r^{th} diagonal element of $\mathbf{\Lambda}_1$ and permuted to a new column position. Likewise, the r^{th} column of $\mathbf{M_2}$ would be scaled by the r^{th} diagonal element of $\mathbf{\Lambda}_2$ and permuted to the same new column position of the r^{th} column of $\mathbf{M_1}$. Hence, the resulting Khatri-Rao product of the scaled and permuted columns of $\mathbf{M_1}$ and $\mathbf{M_2}$ would consist of the columns of $(\mathbf{M_1} \circ \mathbf{M_2})$ that had been rearranged and scaled by the product of the respective diagonal entries of $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$, or $(\mathbf{M_1 \Pi \Lambda_1} \circ \mathbf{M_2 \Pi \Lambda_2}) = (\mathbf{M_1} \circ \mathbf{M_2})\mathbf{\Pi \Lambda}_{1,2}$, where $\mathbf{\Lambda}_{1,2}$ represents the diagonal matrix where the entries are the product of the respective diagonal elements of $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$. □

Theorem 3.4 The current definition of uniqueness is equivalent to a new definition based on KR products, Definition 3.2.

Proof 3.4 The traditional definition of uniqueness stated that every alternative loading matrix must be a permutation-scale version of the original. Therefore, the KR products formed from these alternative loading matrices would be permutation-scale versions of the KR products formed from the original loading matrices, by Theorem 3.3. Thus, every alternative KR product would be a permutation-scale version of the original KR product. Analogously, Definition 3.2 stated that every possible alternative KR product would be permutation-scale versions of the original KR products. Again, by Theorem 3.3, KR products that were permutation-scale transformations would be formed from loading matrices that were permutation-scale versions of the original loading matrices. □

Copyright © Heather Michele Clyburn Bush 2006

4. CONSTRAINTS FROM ORTHOGONAL COMPLEMENT SPACES

Although the tBSM offered novel insights to PARAFAC decompositions and uniqueness, the use of column spaces fell short in providing definitive uniqueness answers outside of R = 3, where decompositions with k-rank<rank were possible. The reasons the tBSM was unable to determine conditions for uniqueness in R = 4 were twofold. First, the method of employing column spaces was dependent on either requiring the third matrix to be full column rank or being able to substitute out elements for alternatives. Eventually, ten Berge and Sidiropoulos came to a scenario where both of these methods were unhelpful. Secondly, and more importantly, although the use of column spaces identified uniqueness or non-uniqueness, the method was unable to answer why. The suggestion that the position of the zero in the RREF was a factor was the only explanation that considering column spaces could offer. Also, ten Berge and Sidiropoulos were discouraged from considering k-rank < rank due to the many decompositions that could result.

4.1 Advantages of the Orthogonal Complement Space Approach (OCSA)

Although the use of column spaces in the tBSM supplied a venue to explore uniqueness that had previously been obstructed, the use of column spaces provided a partial glimpse into the story of uniqueness and PARAFAC. In fact, the use of column spaces to identify when a vector was in a subspace was somewhat imprecise. An alternative method to the tBSM's column space argument is to consider the orthogonal complement (OC) to the column space of the KR product. The use of OC spaces provides a clearer way to discuss whether or not a vector is in a particular subspace and affords a more obvious method for inducing constraints. The use of the OC to provide these constraints, as opposed to just reasoning directly about what it meant to be a linear combination of a spanning set for the KR column space, presented evidence for differences between uniqueness and non-uniqueness conclusions.

Using OC spaces over direct column space reasoning offers the additional advantage in that the method does not fail for larger k-rank, as in the decomposition with four columns and three loading matrices with rank and k-rank equal to 3. Finally, the utilization of the OC space to the column space of the KR product provided an opportunity to view patterns in the KR products, patterns which would lead to theorems and conclusions regarding OC spaces, the constraints imposed, and conditions for uniqueness. For convenience, the use of OCs to generate constraints will be referred to as the OCSA (Orthogonal Complement Space Approach).

4.2 The Logic of Using Orthogonal Complement

The relationship between the column space and the OC to the column space of a matrix, **M,** can be represented by

(4.11) $$\mathcal{C}(\mathbf{M}) = \left(\mathcal{N}(\mathbf{M}^{\mathbf{t}})\right)^{\perp},$$

where $\mathcal{C}(\mathbf{M})$ and $\mathcal{N}(\mathbf{M}^{\mathbf{t}})$ denote the column space of **M** and orthogonal complement to the column space of **M**, respectively. The null space of a matrix, represented by $\mathcal{N}(\mathbf{M}^{\mathbf{t}})$, consists of all vectors **y** such that $\mathbf{M}^{\mathbf{t}}\mathbf{y} = 0$. This identity implies that all vectors orthogonal to $\mathcal{N}(\mathbf{M}^{t})$ are elements of the $\mathcal{C}(\mathbf{M})$. The tBSM purported to use the ideas of column spaces to investigate the KR product, $(\mathbf{M_1} \circ \mathbf{M_2})$. The use of the OCSA will also investigate the column space of $(\mathbf{M_1} \circ \mathbf{M_2})$, but will do so by studying the *orthogonal complement* of the column space.

Similar to the tBSM, the utilization of OC spaces also incorporated the premise that an alternative KR product would have the same column space as the original, or $\mathrm{C}(\mathbf{M_1} \circ \mathbf{M_2}) = \mathrm{C}(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$. From equation (4.1), it is easy to see that the column space problem translates to one which considers OC spaces

(4.12) $$\mathcal{C}(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2) = \mathcal{C}(\mathbf{M_1} \circ \mathbf{M_2}) = \left(\mathcal{N}((\mathbf{M_1} \circ \mathbf{M_2})^{\mathbf{t}})\right)^{\perp}.$$

Therefore, it is possible to impose constraints on the columns of $(\tilde{\mathbf{M}}_1 \circ \tilde{\mathbf{M}}_2)$ by requiring that the columns be orthogonal to $\mathcal{N}((\mathbf{M_1} \circ \mathbf{M_2})^{\mathbf{t}})$. Any KR product with columns orthogonal to $\mathcal{N}((\mathbf{M_1} \circ \mathbf{M_2})^t)$ would possess the same column space as $(\mathbf{M_1} \circ \mathbf{M_2})$.

4.3 Applying Orthogonal Complement Spaces

The setup for applying OC spaces to discover constraints for alternative PARAFAC decompositions is similar to that of ten Berge and Sidiropoulos. Obviously, the structure of the PARAFAC solution will need to utilize the KR product form; and necessary conditions such as full-column rank of the KR products and at least k-rank of 2 for all loading matrices will be assumed. Additionally, the same symmetry properties of PARAFAC solutions would be employed. Finally, the idea of using simplified forms would be used to facilitate the investigation of the uniqueness and OC spaces.

4.3.1 Finding Constraints

As in the tBSM, the Reduced Row Echelon Forms (RREFs) for each original loading matrix were used, with arbitrary variables used whenever the entry was not a 0 or 1. KR products could then be formed as the columnwise Kronecker product of these generic RREF matrices, representing the original loading matrices from a PARAFAC solution. The tBSM sought to find alternative KR products that preserved the column space of the original KR product. Non-trivial transformations that preserved the column space implied non-uniqueness. The OCSA had the same intent, to identify alternatives with the same column space as the original KR product. However, the OCSA sought constraints that would restrict the columns of the alternative KR product to a particular subspace, the column space of the original KR product. As will be discussed later, these constraints will be used to identify whether non-trivial transformations are possible. The OC constraints, that restricted alternative KR products to have the same column space as the original, were obtained by finding basis vectors for the null space of the original KR product transposed.

Using applications from linear algebra, it was noted that finding basis vectors to the null space of the rows of the KR product would be equivalent to finding basis vectors orthogonal to the column space of the KR product. These basis vectors could then be used to form constraints on general alternative loading matrices. The alternative loading matrices are made general by considering matrices with symbolic entries that reflect the same rank, k-rank, and number of columns as the original loading matrices. The inner products of the OC basis vectors and the columns of the alternative KR product are then found, set to 0, and arbitrary elements of the loading matrices solved for, providing a set of constraints on the elements of alternative loading matrices.

Using the OCSA, it was possible to identify the types of constraints that would dictate permutation-scale transformations, a detail that had been alluded to in the switching elements strategy of the tBSM. Based on the structure of the constraints produced by the OCSA, it was possible to identify whether or not transformations other than permutation-scale were allowable for all classes of PARAFAC decompositions, unlike in the tBSM where the rule only applied to the switching of elements strategy. Two types of constraints resulted: constraints which combined elements of the loading matrices so that it was possible to separate the constraints according to the respective loading matrix and constraints which combined the elements in such a way that it was impossible to parse the elements out into individual constraints for each loading matrix. It turns out that the constraints which can be separated force permutation-scale transformations while those which cannot be separated allow non-trivial transformations.

4.3.2 Applying Constraints to Identify Uniqueness in PARAFAC solutions

Evaluating the constraints that resulted from considering a single KR product and the OC to the column space of the KR product did not supply the entire uniqueness story for a particular decomposition. The constraints which evolved from the application of the OCSA only provided insights to the alternatives for the KR product or two of the loading matrices. In order to discuss PARAFAC decomposition uniqueness, the other loading matrix would have to be incorporated as well, as will be seen in the following sections.

In Chapter 3, Definition 3.2, based on Kruskal's idea of equivalent matrices, offered a definition of uniqueness in the KR product setting. Consequently, a

PARAFAC decomposition will be unique if and only if every pairwise combination of loading matrices results in KR products that have alternatives which are restricted to permutation-scale transformations of the original. Hence, it would be necessary to evaluate each KR product combination before the determination of uniqueness could be made. Any PARAFAC decomposition where the constraints obtained from the OCSA suggested permutation-scale alternatives for every pairwise combination of loading matrices is defined as unique. Consequently, if at least one of the pairwise combinations of loading matrices had constraints such that alternatives were not restricted to be permutation-scale transformations, the PARAFAC decomposition would be non-unique.

As in the tBSM, if all decompositions in a class were found to be non-unique when Kruskal's k-rank sum condition was not met, Kruskal's condition would be necessary and sufficient for that class. However, with the application of the OCSA, it is possible to explore decompositions where ten Berge and Sidiropoulos could not. In the following sections it will be shown that the perspective of the OCSA and the new decompositions that could be evaluated are crucial to offering a justification for why decompositions are non-unique.

4.4 The Application of the OCSA for Specific R

The OCSA was used to investigate both PARAFAC decompositions where the tBSM had been applied and for PARAFAC decompositions where no determination of uniqueness had been made. For the cases where the tBSM had already been employed to evaluate PARAFAC decomposition uniqueness, the OCSA provided confirmation for most of the results obtained by ten Berge and Sidiropoulos. However, in the case where the tBSM was unable to address the question of uniqueness - $R = 4$ and the k-rank and rank of all three loading matrices were 3 - the OCSA contradicted the results of the appendix method.

Unlike the tBSM, the OCSA did not suffer from requiring substitution or needing the third matrix to have full-column rank. Thus, the OCSA was able to definitively answer the uniqueness question for $R = 4$ and k-rank = rank. Additionally, the systematic approach of considering pairwise combinations of loading matrices permitted the evaluation of cases where k-rank < rank. Where ten Berge and Sidiropoulos had been overwhelmed by variations of reduced forms, the OCSA was able to tabulate the results for each of the KR product cases and determined uniqueness for PARAFAC decompositions for R = 4 and k-rank < rank. Finally, the OCSA provided rationalization for when to expect alternative KR products that were permutation-scale versions of the original. While the tBSM was limited by the use of column spaces and unable to provide insights as to why uniqueness occurred for some KR products and not for others, the OCSA offered that uniqueness was rooted in the constraints of the OC space and not simply in the position of the zero.

Keeping with current strategies in the evaluation of uniqueness, the OCSA was employed for specific R. Cases (in R = 3 and R = 4) that had been investigated, at least in part are presented in this chapter. PARAFAC decompositions with R = 5 and R = 6, which had not been explored, are discussed in Chapter 5.

As in the tBSM, the evaluation of the KR products was performed when Kruskal's sufficient condition was not met. For each pairwise KR product resulting from loading matrices outside of Kruskal's condition, the resulting basis vectors of the OC space were described and used to impose constraints on the alternative forms of the loading matrices. The constraints were then studied to determine whether or not the restriction of permutation-scale alternatives would be imposed. In the following sections of this chapter and Chapter 5, the results for each value of R will be presented in tables listing the various combinations of the k-ranks and ranks for two of the three loading matrices involved in the KR product. Due to the symmetry of PARAFAC decompositions, it is not necessary to list every pairwise permutation of ranks and k-ranks. Thus, only the combinations needed to evaluate all PARAFAC decompositions for a given R are displayed in the tables.

After the constraints obtained by employing the OCSA were evaluated, it was possible to compose all possible PARAFAC decompositions where Kruskal's condition was not met. By the definition of uniqueness, a PARAFAC decomposition is unique if every pairwise combination of alternative loading matrices formed a KR product that was a permutation-scale version of the original. Hence, the tables are used to assess uniqueness by considering each of the pairwise combinations of the three-choose-two loading matrices. When every pairwise combination of the three loading matrices results in constraints that restrict alternatives to be permutation-scale versions of the original, uniqueness is established. On the other hand, when any of the combinations suggest that alternative loading matrices are allowed to be transformed by matrices other than permutation-scale, non-uniqueness is identified.

4.5 What the constraints showed for R = 3

The case where R = 3 was one of the cases evaluated by ten Berge and Sidiropoulos (2002). The use of column spaces to evaluate the potential alternative forms resulted in proving that Kruskal's sufficient condition was also necessary for R = 3. The application of the OCSA would confirm these results.

When considering the PARAFAC decompositions for R = 3, the options for loading matrices were restricted by k-rank, the number of columns, and the requirement that the sum of the k-ranks for all of the loading matrices had to be less than $2R + 2 = 6 + 2 = 8$. For all investigations of uniqueness, it is assumed that the k-ranks of all loading matrices are at least 2, a necessary condition for uniqueness. Since R was equal to 3, every loading matrix had exactly three columns. Given that there were only three columns and, by definition, k-rank ≤ rank, the loading matrices would be restricted to cases where k-rank was equivalent to rank. Consequently, the possibilities for the loading matrices when R = 3 were k-rank = 2 and k-rank = 3.

4.5.1 Evaluating the KR products (R = 3)

Each of the distinct pairwise combinations of loading matrices that could result when R was equal to 3 and Kruskal's condition was not met, were evaluated using the OCSA. The KR product for each of these combinations was found and the basis vectors of the orthogonal complement to the column space were obtained. The basis vectors were then applied to the columns of a KR product formed from alternative, arbitrary loading matrices with the same matrix properties of the original. The constraints that resulted were then examined to determine if alternative forms would be restricted to permutation-scale transformations.

4.5.1.1 KR Products: k-ranks of 2

For the first case, where the k-ranks were both 2, there was only one basis vector for the OC to the column space of the KR product, imposing only one constraint. Imposing this constraint did not require that alternative KR products be permutation-scale transformations (Table 4.1).

4.5.1.2 KR Products: k-rank of 3

However, in the case where one of the matrices had k-rank of 2 and the other 3, 3 basis vectors resulted. Only one of these vectors contained a term that was a constraint element consisting of values from the last column of the loading matrix with k-rank of 2. This constraint was sufficient to force alternative KR products to be permutation-scale versions of the original (Table 4.1).

Table 4.1 Pairwise Combinations of Loading Matrices

k-rank (M1)	k-rank (M2)	rank(M1)	rank(M2)	P/S Only?
2	2	2	2	No
2	3	2	3	Yes

In studying the constraints that resulted from the OCSA, it was observed that the only case that restricted alternatives to be permutation-scale transformations, had a sum of the k-ranks (and ranks) greater than or equal to $R+2=5$. The size of the k-ranks (and ranks) in this scenario would reduce the options for the third loading matrix in the PARAFAC solution. The premise for investigating these PARAFAC solutions was that Kruskal's condition was not met, or that the sum of the k-ranks for the three loading matrices was less than 8. To keep from obtaining Kruskal's condition when the sum of the k-ranks for two of the loading matrices was 5 or greater, the third matrix would need to have deficient column rank, i.e. a k-rank of 3 would invoke Kruskal's sufficient condition for uniqueness.

4.5.2 Evaluating the PARAFAC solutions (R = 3)

The definition for uniqueness required that all of the loading matrices be considered to determine the uniqueness of a PARAFAC decomposition. However, since only decompositions where the k-ranks of the three loading matrices summed to less than R+2 were allowed, the examination of the necessity of Kruskal's condition would restrict the options for these loading matrices. The limited number of columns for R = 3 also restricted the combinations of rank and k-rank. As mentioned in the previous section, options for the third matrix would be limited to cases where k-rank was equal to rank. As a result, two types of matrices were candidates for the third loading matrix in the decomposition, one with k-rank = rank of 2 and one with k-rank = rank of three.

4.5.2.1 The PARAFAC decomposition: k-rank of 2

For the first case, where two of the loading matrices had k-rank equal to 2, if the third matrix had k-rank = 2 as well, the PARAFAC solution would be non-unique since all permutations of two loading matrices would result in KR products that were not limited to permutation-scale transformations (Table 4.1).

4.5.2.2 The PARAFAC decomposition: k-ranks of 2, 2, & 3

When two of the loading matrices had k-rank of 2 and the third matrix had k-rank of 3, the KR product formed from matrices with k-rank of 2 and 3 would result in a KR product that was restricted to permutation-scale (Table 4.1). However, since there would be two matrices with k-rank and rank of 2, that combination would result in a KR product with alternatives that were not permutation-scale transformations (Table 4.1). Therefore, since one of the combinations allowed a non-permutation-scale transformation of the KR product, the PARAFAC solution would also be non-unique.

4.5.3 PARAFAC Decomposition Uniqueness Conclusions (R = 3)

When R = 3, the OCSA was able to identify that all PARAFAC solutions, where the k-ranks of the loading matrices summed to less than 8, were non-unique. The investigation of the constraints, found by considering the OC to the column space of the KR product, confirmed the results found by ten Berge and Sidiropoulos.

4.6 What the Constraints Showed for R = 4

Unlike the case when R was equal to 3, the options for loading matrices was much more varied for R = 4. In fact, for ease in evaluation the loading matrices of the PARAFAC decompositions were divided into 2 classes: all loading matrices had k-rank = rank and at least one loading matrix had k-rank < rank.

The first class was previously evaluated for uniqueness by ten Berge and Sidiropoulos, in part, by considering the column space of the KR products formed by two loading matrices. Even so, the case where all of the loading matrices had k-rank of 3 had not been satisfactorily evaluated. Additionally, with the exception of one rank/k-rank combination, conditions for uniqueness in the latter class remained unexplored.

For R = 4, each of the possible pairwise combinations for k-rank and rank, where Kruskal's sufficient condition for uniqueness failed, were evaluated in detail (Table 4.2). The determination of whether or not permutation-scale alternatives were allowed was possible in each of the combinations. In one of the combinations, different RREFs led to different conclusions about allowable transformations. In this case, also presented by ten Berge and Sidiropoulos, it will be shown that uniqueness will depend on a factor other than k-rank.

Table 4.2 Pairwise Combinations of Loading Matrices (R = 4)

k-rank (M1)	k-rank (M2)	rank(M1)	rank(M2)	P/S Only?
2	2	2	2	No
2	2	2	3	No
2	2	3	3	Yes/No[1]
2	3	2	3	No
2	3	3	3	Yes
2	4	2	4	Yes
2	4	3	4	Yes
3	3	3	3	Yes
3	4	3	4	Yes

[1]Only permutation-scale transformation depends on the number of OC constraints

4.6.1 Evaluating KR Products (R = 4, k-rank = rank)

The case of matrices where k-rank = rank was investigated by ten Berge and Sidiropoulos by manipulating the column spaces of the possible KR products. With the exception of one PARAFAC solution, the tBSM was able to identify non-uniqueness for this class of PARAFAC solutions. Although the tBSM was able to recognize non-uniqueness, it was unable to characterize what caused certain solutions to be unique and others not.

The application of OCSA for k-rank = rank when R = 4, would prove to be a better approach than the tBSM by

- Identifying uniqueness or non-uniqueness for the entire class of PARAFAC solutions where k-rank = rank;
- Discern whether Kruskal's condition was necessary and sufficient for k-rank=rank; and,
- Formulate necessary and sufficient conditions for uniqueness.

Every possible pairwise loading matrix combination, and the resulting KR products, were analyzed using the OCSA. For each combination, it was possible to examine whether the constraints would force alternatives to be permutation-scale transformations.

The descriptions that were developed attempt to evaluate the properties of the basis vectors of the OC to the column space of the KR products. Each KR product was described in terms of the following:

- The number of basis vectors in the OC space;
- The number of basis vectors that contain constraint elements; and,
- A description of the constraint elements.

Additionally, after considering the OC basis vectors associated with each pairwise combination of loading matrices, a third matrix was considered so that the uniqueness of a PARAFAC decomposition could be decided.

4.6.1.1 KR products: k-ranks of 2 and 2

When both loading matrices had k-rank and rank of 2, the dimension of the KR product was 4×4, resulting in no OC space basis vectors. With no constraints from the OC space, any alternatives were possible and were not limited to only permutation-scale transformations.

Result 4.1 For KR products with k-ranks of 2 and 2 the following was found:

- The OC space was empty
- No vectors with constraint elements resulted
- No constraints resulted.

4.6.1.2 KR products: k-ranks of 2 and 3

For loading matrices with k-ranks of 2 and 3, the number of basis vectors for the OC space of the columns of the KR product was 2, and each of the vectors contained constraint elements. One vector contained one element, a ratio of values from both loading matrices; while the other vector contained two elements, a ratio with elements from both and a ratio with elements from only the loading matrix with k-rank=2. The constraints that resulted from these two vectors did not force alternative KR products to be permutation-scale transformations.

Result 4.2 For KR products with k-ranks of 2 and 3 the following was found:

- Two OC space basis vectors resulted
- Two vectors with constraint elements resulted
- Two constraint elements formed from elements from both matrices and one constraint element formed from elements of only one matrix resulted.

4.6.1.3 KR products: k-ranks of 2 and 4

Since the matrices were in reduced row echelon form, the matrix with k-rank = rank = 4 would be the identity matrix. Hence, all of the constraint elements would be imposed from the matrix with k-rank of 2. In this case, 4 basis vectors for the OC space resulted. Of these four basis vectors, two contained one constraint element, ratios with values from the loading matrix with k-rank = 2.

Result 4.3 For KR products with k-ranks of 2 and 4 the following was found:

- Four OC space basis vectors resulted
- Two vectors with constraint elements resulted
- Two constraint elements from only one matrix (k-rank = 2) resulted.

The constraints imposed by these basis vectors were sufficient to require that alternative KR products and the corresponding alternative loading matrices be permutation-scale transformations.

4.6.1.4 KR products: k-ranks of 3 and 3

This KR product was of particular interest since ten Berge and Sidiropoulos were unable to employ their strategies, element replacement and use of a nonsingular third matrix, for investigating column spaces and had to rely on the more questionable appendix method. Therefore, this would be the first evaluation of the general KR product for loading matrices with k-ranks of 3.

In this case, there were 5 basis vectors for the OC to the column space of the KR product. All five of these vectors contained constraint elements. Three of the five vectors contained ratios comprised of elements from both loading matrices. The remaining vectors contained ratios as well, one vector containing a ratio from only the elements of one loading matrix and the other vector a ratio from the other loading matrix.

Result 4.4 For KR products with k-ranks of 3 and 3 the following was found:

- Five OC space basis vectors resulted
- Five vectors with constraint elements resulted
- Three constraint elements formed from elements from both matrices and two constraint elements formed from elements of only one matrix resulted.

The constraints offered from these basis vectors were enough to force alternative loading matrices and KR products to be permutation-scale transformations of the originals. Thus, the question of permutation-scale alternatives was answered for these KR products.

4.6.1.5 KR products: k-ranks of 3 and 4

For the last combination where k-rank = rank, the reduced form of the matrix with k-rank = 4 was the identity matrix, and only the matrix with k-rank of 3 would contribute constraint elements. Even so, the constraints that resulted were enough to limit alternative forms to be permutation-scale alternatives. Eight basis vectors resulted from considering the OC to the column space of this KR product, two of which contained constraint elements. Each of these two vectors contained exactly one ratio of elements from the loading matrix with k-rank of 3.

Result 4.5 For KR products with k-ranks of 3 and 4 the following was found:

- Eight OC space basis vectors resulted
- Two vectors with constraint elements resulted
- Two constraint elements from only one matrix (k-rank = 3) resulted.

4.6.2 Evaluating PARAFAC Decompositions (R = 4, k-rank = rank)

For R = 4, ten Berge and Sidiropoulos only presented a detailed evaluation for the cases were k-rank and rank were equivalent, and their method was not appropriate when all the ranks and k-ranks of the three loading matrices were 3. The OCSA, however, was able to analyze each of the KR products that resulted when k-rank = rank. Additionally, by considering the KR products that resulted from pairwise combinations of matrices, it was possible to employ Definition 3.2 for uniqueness and identify uniqueness or non-uniqueness for every PARAFAC decomposition having equivalent k-rank and rank.

When the k-rank = rank, the pairwise combinations of k-ranks for loading matrices $(\mathbf{M_1}, \mathbf{M_2})$, where Kruskal's sufficient condition would not be met, were (2,2), (2,3), (2,4), (3,3) and (3,4). In order to determine whether or not the PARAFAC decomposition was unique, the third matrix would also have to be considered. For each of the pairwise combinations, all of the options for the third matrix, where Kruskal's condition was not met, would need to be evaluated in order to determine whether or not the decomposition was unique. Including a third matrix would mean evaluating all of the three-choose-two KR products for a particular PARAFAC decomposition. Therefore, for a given PARAFAC solution, with three loading matrices, the three pairwise combinations (three KR products) would have to be explored to determine if all of the KR products were restricted to permutation-scale alternatives. When every KR product had alternatives restricted to permutation-scale transformations, the PARAFAC solution would be identified as unique.

4.6.2.1 PARAFAC Decompositions: A KR Product with k-ranks of 2 and 2

For a PARAFAC decomposition having two loading matrices with k-ranks of 2 and 2, any third loading matrix with k-rank equal to 2, 3, or 4 could be considered without causing the sum of the k-ranks for the matrices to exceed 10 or 2R+2. The evaluation of KR products in the previous section determined that a KR product formed from two loading matrices with k-rank = rank = 2 had alternatives that were not restricted to permutation-scale transformations (Table 4.2). By the definition of uniqueness, a PARAFAC solution was non-unique if any pairwise combination of loading matrices resulted in a KR product with alternatives that were not limited to permutation-scale transformations. Therefore, any PARAFAC solution where two of the loading matrices had k-rank = rank =2 would be non-unique. The details of identifying non-uniqueness for each of the PARAFAC solutions when two of the loading matrices have k-rank = rank = 2 are given in the following paragraphs.

4.6.2.1.1 The PARAFAC decomposition: k-ranks of 2, 2, and 2

In the case where, the third matrix had k-rank of 2, all three matrices have k-rank of 2 and the only KR product to consider was one composed of matrices with k-rank = rank = 2. There were alternative KR products for this case that were not permutation-scale versions, which implied that the PARAFAC solution was non-unique.

4.6.2.1.2 The PARAFAC decomposition: k-ranks of 2, 2, and 3

For a third matrix with k-rank of 3, the resulting pairwise combinations would be (2,3) and (2,2). Again, in both cases, the constraints obtained by considering the OC space did not impose permutation-scale as the only transformation for alternative KR products; and, so, the PARAFAC solution was non-unique.

4.6.2.1.3 The PARAFAC decomposition: k-ranks of 2, 2, and 4

For this final case, the resulting pairwise combinations of loading matrices k-ranks were (2,2) and (2,4). In Table 4.2 it can be seen that the KR product composed of loading matrices with k-ranks of 2 and 4 is constrained to be a permutation-scale transformation. However, since the KR product comprised of two loading matrices with rank and k-rank of 2 was not constrained to be a permutation-scale transformation, the PARAFAC solution was non-unique.

4.6.2.2 PARAFAC decompositions: A KR Product with k-ranks of 2 and 3

In this case, the third loading matrix could have k-ranks of 2, 3, or 4.The case where two of the loading matrices had k-ranks of 2 and 3 was also found to have alternatives that were not permutation-scale versions of the original in the previous section where KR products were evaluated. Therefore, by the definition of uniqueness, any PARAFAC solution which had two loading matrices with k-rank = rank and k-ranks of 2 and 3 would be non-unique. The process of determining PARAFAC solution uniqueness when two of the loading matrices had k-ranks of 2 and 3 are described in the paragraphs below.

4.6.2.2.1 The PARAFAC decomposition: k-ranks of 2, 2, and 3

The case where the third matrix had a k-rank of 2 was already investigated above; and, since both KR products formed from loading matrices with k-ranks of (2,2) and (2,3) had alternatives that were not constrained to be permutation-scale transformations, the PARAFAC solution was found to be non-unique.

4.6.2.2.2 The PARAFAC decomposition: k-ranks of 2, 3, and 3

When the third matrix had a k-rank of 3, the resulting pairwise combinations of k-ranks were (2,3) and (3,3). Although the KR product created from two matrices with k-rank of 3 yielded alternatives that were only permutation-scale transformations, the KR product with matrices having k-ranks of 2 and 3 were allowed to have alternatives that were not permutation-scale. Therefore, the resulting PARAFAC solution was non-unique.

4.6.2.2.3 The PARAFAC decomposition: k-ranks of 2, 3, and 4

The final decomposition involved a third matrix with k-rank = 4, resulting in pairwise combinations of (2,4), (3,4), and (2,3). The KR products with k-ranks of (2,4) and (3,4) yielded OC constraints that only permitted permutation-scale transformations (Table 4.2). However, the KR product with k-ranks of (2,3) did not have these restrictions, making this PARAFAC decomposition non-unique.

4.6.2.3 PARAFAC decompositions: A KR product with k-ranks of 3 and 3

In this case, the third matrix would be limited to having a k-rank of 2 or 3. A third matrix with k-rank of 4 would meet Kruskal's condition for uniqueness, $3 + 3 + 4 = 10 \geq 2R + 2$. The PARAFAC decompositions with KR products involving matrices with k-ranks and ranks of 3 were especially interesting because the tBSM had been unable to provide a means of investigating uniqueness when the third loading matrix also had k-rank = 3. After careful review, the appendix method employed by ten Berge and Sidiropoulos was to be inadequate, leaving the question of uniqueness when all three loading matrices had k-ranks of 3.

4.6.2.3.1 The PARAFAC decomposition: k-ranks of 3, 3, and 2

When the third matrix had k-rank of 2, the possible combinations were (2,3) and (3,3), and it was found that the KR product resulting from two matrices with k-ranks of 2 and 3 could have alternatives other than permutation-scale transformations (Table 4.2). Hence, a PARAFAC solution with k-rank = rank and loading matrices with k-ranks 2, 3, and 3 would be non-unique.

4.6.2.3.2 The PARAFAC decomposition: k-ranks of 3

In the case where all the matrices had rank and k-rank equal to 3, the only pairwise combination that resulted was when both loading matrices had rank and k-rank equal to 3. The evaluation of such KR products using the OCSA, demonstrated that the constraints imposed by the OC basis vectors would forced alternative KR products to be permutation-scale versions of the original.

Therefore, every pairwise permutation of the loading matrices resulted in a KR product with only permutation-scale alternatives; and, by definition, the PARAFAC solution was unique.

The only k-rank = rank scenario where the tBSM failed to provide definitive results were decompositions where the k-ranks of all matrices were 3, requiring the use of an alternative approach that incorrectly suggested non-uniqueness. The alternative method, the appendix method, was fully discussed in Chapter 3 and demonstrated the error in using numerical examples to evaluate uniqueness. The bigger issue at hand, however, is that ten Berge and Sidiropoulos used the findings in the appendix method to make a general statement about PARAFAC decompositions where k-rank = rank and R = 4.

Since the tBSM had shown all other k-rank = rank PARAFAC decompositions to be non-unique, the only missing piece to showing Kruskal's condition necessary and sufficient was the case where the k-ranks were 3, 3, and 3. Hence, the incorrect conclusion of non-uniqueness in this case caused ten Berge and Sidiropoulos to incorrectly conclude that Kruskal's condition was necessary and sufficient for R = 4 and k-rank = rank. The OCSA, however, was not impeded by the same problems that ten Berge and Sidiropoulos encountered when using column spaces (or solving for polynomials) and was able to finally show that such a PARAFAC solution would be unique. Hence, Kruskal's condition could not be necessary and sufficient for R = 4 and k-rank = rank.

4.6.2.4 PARAFAC decompositions: A KR Product with k-ranks of 3 and 4

Finally, in the case where the PARAFAC decomposition involved two loading matrices with k-ranks of 3 and 4, the only option for the third matrix was a k-rank of 2, otherwise Kruskal's sufficient condition would be invoked. In this final case, the pairwise combinations were (2,4), (3,4) and (2,3), a repeat of an earlier scenario which was identified as a non-unique PARAFAC solution.

4.6.3 PARAFAC Decomposition Uniqueness Conclusions (k-rank=rank)

ten Berge and Sidiropoulos concluded that when R = 4 and k-rank was equivalent to rank, Kruskal's condition for uniqueness would be necessary and sufficient. However, by the OCSA, it was possible to determine that PARAFAC solutions, when all k-ranks were 3, would be unique. Finding uniqueness when k-rank = rank invalidated ten Berge's and Sidiropoulos' conjecture of necessity. To see this, consider that since this decomposition was shown to be unique, uniqueness occurred outside of Kruskal's condition. Thus, PARAFAC decompositions were not unique only when the sum of the k-ranks was at least 2R + 2 and Kruskal's condition could no longer be considered necessary for R=4 and k-rank = rank. Even so, in addition to determining uniqueness, it would be possible to use the OCSA to identify additional conditions so that Kruskal's uniqueness condition could be necessary and sufficient for certain cases in R=4. These additional conditions will be described in detail in later sections.

4.6.4 Evaluating KR Products (k-rank < rank)

With the exception of an example to show that Kruskal's condition was not necessary for R = 4, the class of PARAFAC decompositions where a matrix had deficient k-rank (k-rank < rank) was not investigated for uniqueness. The one case that ten Berge and Sidiropoulos did investigate hinted that the reduced forms offered many permutations within a particular rank and k-rank, and the investigation was halted. However, by considering the pairwise combinations and the KR products, the task of investigating uniqueness for k-rank < rank was much less daunting. Further by analyzing the OC spaces of these KR products, it was possible to gain a better understanding of uniqueness and when it would occur.

When R = 4 and k-rank < rank, by the properties of k-rank, the loading matrices would have k-rank less than 4. However, a necessary condition for uniqueness was that k-rank was at least 2. Thus, the only case for R = 4 where k-rank < rank was when k-rank = 2 and rank = 3. In order for PARAFAC solutions to be in the class of solutions where k-rank < rank, at least one of the loading matrices would need to have k-rank of 2 and rank of 3.

The loading matrices with a rank of 3 and a k-rank of 2 have simplified forms with a 0 located in exactly one of the rows of the last column. For example,

$\begin{pmatrix} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$. From the structure of RREF, the last column provides the location of the coefficients for the linear combinations of the two linearly independent columns. The rationalization for this is attributed to deficient column rank. Since the k-rank of the loading matrix was 2, any two columns could be combined as a linear combination to form one of the others. Hence, when the RREF for this matrix is obtained, exactly one 0 is located in the last column to represent which column is not needed in the linear combination. Therefore, the matrix could have 3 different RREFs based on which row the 0 was located: $\begin{pmatrix} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & a_3 \end{pmatrix}$, and $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & a_3 \end{pmatrix}$. Accordingly, for each KR product that involved the loading matrix with k-rank < rank, every variation of the RREF would also have to be considered. ten Berge's and Sidiropoulos' realization of this was a primary reason for opting to investigate only cases where k-rank and rank were equivalent.

As in the case of k-rank = rank, each KR product will be described in terms of the following:

- The number of basis vectors in the OC space;
- The number of basis vectors that contain constraint elements; and,
- A description of the constraint elements.

4.6.4.1 KR Products: k-ranks and ranks were {2,3} and {2,2}

The loading matrix, with k-rank and rank of 2, would not have variations in the structure of its RREF. The loading matrix with k-rank of 2 and rank of 3 would have three variations, resulting in a total of three KR products. However, for this KR product, regardless of the position of the 0, the OC space had 2 basis vectors.

When the 0 was located in the first or second rows of the last column of the second loading matrix, only one of these vectors had constraint elements, ratios located in two positions. One ratio combined terms from both matrices and one combined terms from the matrix with rank and k-rank of 2.

On the other hand, when the zero was located in the last or third row of the loading matrix, the number of basis vectors with constraint elements was two, and not one. Instead of two ratios in one vector, however, each vector had one ratio. One vector contained the ratio which combined elements from the two loading matrices and the other vector contained the ratio which consisted only of elements from the matrix with k-rank and rank equal to 2. Thus, regardless of the position of the zero, two similar constraints resulted. In all three cases the constraints that resulted from using these basis vectors resulted in alternative KR products that were not permutation-scale versions of the original KR product.

Result 4.6 For the KR product composed of matrices with {k-rank, rank} of {2,3} and {2,2} the following was found:

- Two OC basis vectors resulted
- Depending on the position of the zero, one or both contained constraint elements
- One constraint element contained elements from both matrices and one contained elements from only one.

4.6.4.2 KR Products: k-ranks and ranks were {2,3} and {2,3}

Of all the cases when k-rank < rank for R = 4, this case proved to be the most intriguing. First noted by ten Berge and Sidiropoulos, this case was used as a counterexample to Kruskal's necessity for R = 4. Additionally, this case suggested that something "bigger" than k-rank was responsible for uniqueness.

In order to have k-rank = 2 and rank = 3, the last column of the RREF for each loading matrix had to contain exactly one zero element. The placement of this zero element resulted in nine different KR products and nine different sets of basis vectors for the OC to the column space of the KR products. Each set of basis vectors contained 5 vectors.

If the zeroes in the last column of the two loading matrices were located in the same row for each loading matrix, only one of the vectors would contain constraint elements. If, however, the zeroes were located in different rows for each of these two loading matrices, two of the five vectors would contain constraint elements. With the zeroes in the same row positions, the single vector with constraint elements contained one ratio, comprised of the elements from both loading matrices. When the zeroes were in different row positions, however, two constraint elements resulted, one ratio in each of the two vectors. The ratio in one of the vectors was comprised of elements from both matrices while the ratio in the other vector was comprised of elements from only one of the loading matrices.

Consider two examples, one with the zeroes located in the same row position and one with the zeroes located in different row positions.

Example 1 For the two matrices with zeroes in the same row position the RREFs are $\mathbf{M_1} = \begin{pmatrix} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ and $\mathbf{M_2} = \begin{pmatrix} 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & b_2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$. A set of basis vectors for $\mathcal{N}((\mathbf{M_1} \circ \mathbf{M_2})^{\mathbf{t}})$ could be

$$\left[\begin{bmatrix} 0 \\ -\frac{b_1 a_2}{b_2 a_1} \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right].$$

When these basis vectors are used in the inner product, the OC constraint that results is $\frac{h_2 g_1}{h_1 g_2} = \frac{b_2 a_1}{b_1 a_2}$. These constraints form a ratio where the elements cannot be separated into constraints on the individual loading matrices.

Example 2 For the two matrices with zeroes in different row positions, examples of the RREFs are $\mathbf{M_1} = \begin{pmatrix} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ and $\mathbf{M_2} = \begin{pmatrix} 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & b_3 \end{pmatrix}$. A set of basis vectors for $\mathcal{N}((\mathbf{M_1} \circ \mathbf{M_2})^{\mathbf{t}})$ could be

$$\left[\begin{bmatrix} 0 \\ 0 \\ -\frac{b_1 a_2}{b_3 a_1} \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -\frac{a_2}{a_1} \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right].$$

When these basis vectors are applied to arbitrary alternatives with similar matrix structures, the OC constraints that result are $\frac{g_1}{g_2} = \frac{a_1}{a_2}, \frac{h_1}{h_3} = \frac{b_1}{b_3}$. Two OC constraints result, each providing restrictions on a single loading matrix.

In the case where there were two basis vectors with constraint elements, the constraints limited alternatives to be permutation-scale transformations only. However, when only one vector contained a single ratio of elements from both loading matrices, alternative KR products were not constrained to be permutation-scale transformations.

Result 4.7 For the KR product composed of matrices with {k-rank, rank} of {2,3} and {2,3} the following was found:

- Five OC basis vectors resulted
- Same or different row position of the zero: one or two contained constraint elements, respectively
- Same or different row position of the zero: one ratio or two separable constraint elements resulted, respectively.

4.6.4.3 KR Products: k-ranks and ranks were {2,3} and {3,3}

Again in this case, the matrix with k-rank = 2 and rank = 3 would need exactly one of the elements in the last column of the RREF to be zero. The other matrix had equivalent k-rank and rank, so no zero was needed and no variations of the KR product form existed.

Regardless of the placement of this zero value, five basis vectors resulted and three of these contained constraint elements. Also, each of these three vectors contained exactly one ratio which consisted of elements from the loading matrices. When the zero was located in the first or second row of the matrix with k-rank = 2 and rank = 3, one ratio was comprised of elements from each of the loading matrices, one ratio was comprised of elements from the loading matrix with k-rank = rank = 3, and the last ratio was comprised of elements from the other loading matrix. However, when the zero was located in the last, third, row of the loading matrix with k-rank = 2 and rank = 3, two of the ratios were comprised of elements from both loading matrices and the other ratio was comprised only of elements from the loading matrix with k-rank = rank = 3. Additionally, one of the ratios comprised of elements from both matrices included the two elements which composed the ratio consisting of elements from only the loading matrix with equivalent rank and k-rank. Hence, after simplifying the constraints, the pattern of constraints that was found when the zero was located in the first or second position was the same as when the zero was located in the third row position of the last column.

Result 4.8 For the KR product composed of matrices with {k-rank, rank} of {2,3} and {3,3} the following was found:

- Five OC basis vectors resulted
- Three contained constraint elements
- One constraint element had elements from both loading matrices and two constrain elements had elements from only one loading matrix.

For all of these sets of basis vectors, the constraints were enough to restrict alternative KR products to be permutation-scale transformations.

4.6.4.4 KR Products: k-ranks and ranks were {2,3} and {4,4}

Finally, the last scenario in which k-rank < rank occurred when one of the matrices had k-rank of 2 and rank of 3 and one of the loading matrices had k-rank = rank = 4. Here, the full-column rank of one of the loading matrices would mean that the RREF would be the identity matrix. Thus, while the loading matrix with k-rank = 4 would help to increase the number of basis vectors for the OC space, any constraint elements would involve only the matrix with k-rank < rank.

In this case, there were eight basis vectors for the OC space. However, only one of these vectors contained a constraint element, a ratio comprised of the elements from the loading matrix with k-rank = 2 and rank = 3. The constraint from the OC to the column space of this KR product was enough to confine alternatives to be permutation-scale transformations.

Result 4.9 For the KR product composed of matrices with {k-rank, rank} of {2,3} and {4,4} the following was found:

- Eight OC basis vectors resulted
- Regardless of the position of the zero, one vector contained a constraint element
- The constraint element was composed only from the matrix with k-rank=2.

For R = 4 and k-rank < rank only one of the pairwise combinations resulted in a KR product where both of the loading matrices had deficient k-rank. Accordingly, this same KR product was the only instance where two separate descriptions of the null vectors and conclusions regarding alternatives occurred. For all other cases, where at least one matrix had k-rank = rank, the conclusions concerning alternative forms and the description of the constraints did not vary with the position of the zero.

Hence, the class of KR products created from loading matrices with deficient k-rank were completely described. However, at this point, these descriptions were limited to discussing the constraints and alternative KR products for pairwise combinations of loading matrices. The uniqueness, or non-uniqueness, of a PARAFAC solution could not be evaluated until the third matrix was considered.

4.6.5 Evaluating PARAFAC decompositions (k-rank < rank)

When k-rank < rank and R = 4, every PARAFAC decomposition must have at least one matrix with k-rank = 2 and rank = 3. With the exception of the case where two matrices had k-rank < rank and a third had full-column rank, this class of PARAFAC solutions had not been analyzed to determine uniqueness. The PARAFAC decompositions for each combination of matrices are presented below and assume that at least one of the matrices in the KR product have k-rank < rank. Thus, to distinguish between the different KR products when k-rank < rank, only one of the loading matrices in the KR product needs to be identified. Hence, in the underlined headings that follow, the k-rank and rank displayed refer to the loading matrix that is joined with a matrix with k-rank of 2 and rank of 3 to form the KR product. For each of these KR products, the various third matrices and the resulting PARAFAC decompositions were analyzed for uniqueness by considering whether every pairwise combination resulted in KR products with alternative forms that were not permutation-scale alternatives.

4.6.5.1 PARAFAC decompositions: k-rank and rank of 2 and 2

In the case where the second matrix had rank = k-rank = 2, the third matrix could have {k-rank, rank} combinations of {2,2}, {2,3}, {3,3}, or {4,4}. All resulting PARAFAC solutions were identified as non-unique. The solutions were found to be non-unique since a KR product composed of two matrices with k-rank and rank combinations $\left(\left\{\mathrm{k}(\mathbf{M_1}),\mathrm{r}(\mathbf{M_1})\right\},\left\{\mathrm{k}(\mathbf{M_2}),\mathrm{r}(\mathbf{M_2})\right\}\right)$ of ({2,2},{2,2}), ({2,2},{2,3}), and ({2,2}, {3,3}), respectively, had been found to have KR products with alternatives that were not restricted to be permutation-scale transformations (Table 4.2).

4.6.5.2 PARAFAC decompositions: k-rank and rank of 2 and 3

The next set of PARAFAC solutions would have two matrices with k-rank = 2 and rank = 3 and a third matrix with {k-rank, rank} combinations of {2,2}, {2,3}, {3,3}, or {4,4}. This set of PARAFAC decompositions was particularly interesting as the decision regarding whether the KR product was allowed permutation-scale transformations depended on the number of resulting constraints, or to use terminology from ten Berge and Sidiropoulos, the positions of the zeros. Additionally, it was with this KR product that ten Berge and Sidiropoulos were able to show that Kruskal was not necessary for R = 4.

With the exception of one PARAFAC decomposition, where the third matrix had k-rank = rank = 2, the uniqueness identification for all other PARAFAC solutions would depend, as suggested by ten Berge and Sidiropoulos, on the position of the zeroes. However, considering OC spaces could provide greater insight. The difference between the KR products, consisting of matrices with k-rank = 2 and rank = 3, with zeroes in the same position versus different positions was in the constraints requiring the column spaces to be equivalent. In the case where the zeroes were in the same position only one basis vector contained one constraint element, formed from elements from both matrices. On the other hand, when the zeroes were in different positions, two OC basis vectors contained constraint elements, one a ratio of elements from both matrices and the other a ratio of elements from only one matrix. Therefore, if only one of the OC basis vectors contained one constraint element, that pairwise combination of loading matrices would have a KR product with alternatives that were not limited to permutation-scale. However, if two of the OC basis vectors each contained one constraint element, the resulting KR product would be composed of alternatives that were restricted to be permutation-scale versions of the originals.

4.6.5.2.1 PARAFAC decomposition: k-ranks and ranks of {2,3}, {2,3}, {2,2}

PARAFAC solutions with a third matrix having k-rank = rank = 2 would be non-unique since KR products formed from matrices with matrix properties of {2,3} and {2,2} did not limit alternatives to permutation-scale transformations.

4.6.5.2.2 PARAFAC decomposition: k-ranks and ranks of {2,3}, {2,3}, {2,3}

In the case where the third matrix also had k-rank = 2 and rank = 3, the PARAFAC decomposition would be unique if each KR product had OC basis vectors with two constraint elements (in two vectors). Otherwise, if any two of the loading matrices formed a KR product with OC basis vectors with only one constraint (in one vector), the loading matrices would have alternatives that were not simply permutation-scale transformations, making the PARAFAC solution non-unique.

4.6.5.2.3 PARAFAC decomposition: k-ranks and ranks of {2,3}, {2,3}, {3,3}

If the third matrix had k-rank = rank = 3, the pairwise combination of the third matrix with any of the matrices with k-rank = 2 and rank = 3 would result in KR products where the alternatives were forced to be permutation-scale transformations. Consequently, the uniqueness or non-uniqueness of the PARAFAC solution would rest on the combination of the two loading matrices with k-rank = 2 and rank = 3. If the KR product created from this combination of matrices resulted in two constraint elements in two basis vectors of the OC space, the PARAFAC solution would be unique. Conversely, if there were only one constraint element in a single basis vectors of the OC space, the PARAFAC solution would be non-unique.

4.6.5.2.4 PARAFAC decomposition: k-ranks and ranks of {2,3}, {2,3}, {4,4}

The last option for the third matrix in this case was the example used by ten Berge and Sidiropoulos to show that Kruskal's condition was not necessary for R = 4. In fact, it was the only scenario where k-rank < rank had been evaluated for uniqueness. ten Berge and Sidiropoulos were the first to note the phenomenon that the row position of the zeroes for the two loading matrices with k-rank = 2 and rank = 3 had an impact on the conclusion of uniqueness for the PARAFAC solution when the third matrix had full-column rank. However, other than recognizing that the position of the zeroes had something to do with uniqueness or non-uniqueness, no further explanation for the uniqueness or non-uniqueness of the PARAFAC solution was given.

The utilization of OC spaces, however, did provide an answer to the question as to what was really different between the PARAFAC solutions and could be utilized to comment on uniqueness. Since the KR product of two loading matrices with k-rank of 2 and 4 and rank of 3 and 4, respectively, had alternatives limited to permutation-scale versions, the uniqueness of a PARAFAC solution would depend on the KR product formed from two matrices with k-rank = 2 and rank = 3. In fact, uniqueness would depend on whether or not the basis vectors for the OC to the column space of the KR product formed from loading matrices with k-rank = 2 and rank = 3 had one or two constraint elements in one or two basis vectors. If two constraint elements were present in two basis vectors, the solution would be unique; and if only one constraint element in only one basis vector was present, the solution would be non-unique.

4.6.5.3 PARAFAC decompositions: k-rank and rank of 3 and 3

One last additional scenario for PARAFAC decompositions, which has yet to be covered by the other cases, was when one of the loading matrices had k-rank = 2 and rank = 3, another loading matrix had k-rank = rank = 3, and the third matrix had k-rank = rank = 3 or k-rank = rank = 4.

4.6.5.3.1 PARAFAC decomposition: k-ranks and ranks of {2,3}, {3,3}, {4,4}

In this case, the PARAFAC solution would be unique. Every pairwise combination of the three loading matrices with k-rank and rank combinations $\left(\{k(\mathbf{M_1}),r(\mathbf{M_1})\},\{k(\mathbf{M_2}),r(\mathbf{M_2})\}\right)$ of ({2,3},{3,3}), ({2,3},{4,4}), and ({3,3},{4,4}) would result in KR products with constraints that limited alternative loading matrices to be permutation-scale versions of the original.

4.6.6 PARAFAC Decomposition Uniqueness Conclusions (k-rank < rank)

PARAFAC solutions where one of the matrices had deficient k-rank offered the richness necessary to decipher what underlying property was contributing to uniqueness or non-uniqueness. The case where two of the loading matrices had deficient k-rank was the first scenario where two KR products composed of loading matrices with the same matrix properties (same k-rank, especially) had two different conclusions concerning uniqueness. The investigation of the OC space basis vectors provided an explanation, more general than k-rank, as to when uniqueness would occur. The use of the OCSA made possible the identification of uniqueness for PARAFAC solutions when at least one of the matrices had deficient k-rank.

4.6.7 PARAFAC Decomposition Uniqueness (for all of R = 4)

It was now possible to identify, for the entire class of R = 4, the uniqueness of PARAFAC decompositions (Table 4.3). Kruskal's k-rank sum condition would be considered necessary in addition to sufficient if no PARAFAC decomposition could be found that was unique when the condition was not met. Since there were cases where uniqueness was determined without imposing Kruskal's condition, it was established that Kruskal's condition was not necessary for all of R = 4, even when k-rank and rank were equivalent. Moreover, it was now possible to formulate a condition based on the information gained from considering the OC to the column space of the KR products so that necessary and sufficient conditions would be available for R = 4.

Table 4.3 PARAFAC Decomposition Uniqueness for R = 4

k-rank (M1)	rank(M1)	k-rank (M2)	rank(M2)	k-rank (M3)	rank(M3)	Uniqueness?
2	2	2	2	2	2	No
2	2	2	3	2	2	No
2	2	3	3	2	2	No
2	2	4	4	2	2	No
2	2	3	3	4	4	No
2	3	2	3	2	2	No
2	3	3	3	2	2	No
2	3	4	4	2	2	No
3	3	3	3	2	2	No
3	3	4	4	2	2	No
2	3	2	3	2	3	No/Yes
2	3	3	3	2	3	No/Yes
2	3	4	4	2	3	No/Yes
3	3	3	3	2	3	Yes
3	3	4	4	2	3	Yes
3	3	3	3	3	3	Yes

4.6.8 Correcting the claims made by ten Berge and Sidiropoulos

The additional condition of k-rank = rank had been applied by ten Berge and Sidiropoulos so that Kruskal's condition would be necessary and sufficient for a particular class of decompositions where R = 4. However, the conclusion that PARAFAC solutions would be non-unique when all matrices had k-rank = rank was erroneously achieved when ten Berge and Sidiropoulos relied on their appendix method. The OCSA did not suffer from the same limitations as the tBSM or the appendix method and made clear that the case where all three loading matrices had k-rank = rank = 3 was unique.

Thus, it is not enough to require k-rank=rank. When k-rank=rank and rank is large, a unique decomposition is possible. In order to construct a class where all decompositions are non-unique when Kruskal is not met, the size of the ranks must be limited. Moreover, with the development and application of the OCSA, it is possible to determine uniqueness or non-uniqueness for all PARAFAC solutions in R = 4, regardless of the relationship of k-rank and rank. Hence, after correcting ten Berge and Sidiropoulos' claim about the class of decompositions where k-rank = rank, the next task would be to analyze the results for R = 4 and develop necessary and sufficient conditions for uniqueness when R = 4. Therefore, by the OCSA, it would be possible to offer necessary and sufficient conditions for R = 4 where none had existed.

4.6.8.1 Necessary and Sufficient Conditions for R = 4 and k-rank = rank

From Table 4.2 and Table 4.3, it was obvious that when k-rank = rank and the sum of the ranks for any pairwise matrix combination was less than R + 2, the PARAFAC solution would be non-unique. Therefore, when the pairwise sum of the ranks was at least R + 2 and k-rank=rank, the PARAFAC decomposition was unique. This is true when Kruskal's condition is not met as well as when it is met. Thus, a necessary and sufficient condition for uniqueness when R = 4 and k-rank=rank is that all of the pairwise sum of the ranks are at least R + 2.

Result 4.10 A PARAFAC solution where R = 4 and k-rank = rank is unique if and only if $r(\mathbf{M}_i) + r(\mathbf{M}_j) \geq R + 2$ for all $i \neq j$.

4.6.8.2 Necessary and Sufficient Conditions for all of R = 4

For R = 4, regardless of the equality of rank and k-rank, the obvious problem to the necessity of Kruskal's condition occurred for PARAFAC decompositions where two of the loading matrices had k-rank<rank (shaded lower section of Table 4.3). A variety of comments could be made regarding these solutions.

Each of these PARAFAC solutions had every pairwise combination of loading matrices where the sum of the ranks was greater than or equal to R + 2 = 6. Based solely on this observation, it would be possible to state that non-uniqueness was sure to occur for R = 4 when any pairwise combination of the loading matrices had ranks that summed to less than R + 2 = 6, as was stated for the class of PARAFAC solutions when k-rank = rank. However, this still ignored the PARAFAC solutions where uniqueness depended on the number of constraint elements / constraint vectors in the OC basis (or the position of the zero in the loading matrices). In order to incorporate these PARAFAC solutions, an additional condition when two of the loading matrices had k-rank < rank would have to be considered. Thus, for R = 4, uniqueness would not occur when any of the following were true:

- At least two of the loading matrices had ranks that summed to less than R + 2 = 6; or,
- When two of the loading matrices had k-rank<rank and only one ($\min(r(\mathbf{M}_i) - k(\mathbf{M}_i), r(\mathbf{M}_j) - k(\mathbf{M}_j))$) constraint element.

It should be noted that when Kruskal's condition was not met, no PARAFAC decomposition could have two matrices with k-rank<rank. Additioanlly, only one decomposition {(2,3), (4,4), (4,4)} had k-rank<rank when Kruskal's condition was met. In this case, all of the pairwise sum of the ranks were at least R + 2 =7. Therefore, when fewer than two of the loading matrices have k-rank<rank, the PARAFAC decomposition will be unique if and only if each of the pairwise sum of the ranks is at least R + 2. When at least two of the loading matrices have k-

rank<rank, then the decomposition is unique if and only if he pairwise sum of the ranks is at least R + 2 *and* there is more than one ($\min(r(\mathbf{M}_i)-k(\mathbf{M}_i), r(\mathbf{M}_j)-k(\mathbf{M}_j))$) OC constraint.

Result 4.11 A PARAFAC solution where R = 4 and fewer than two of the loading matrices have k-rank < rank is unique if and only if $r(\mathbf{M}_i)+r(\mathbf{M}_j)\geq R+2$ for all $i\neq j$.

Result 4.12 A PARAFAC solution where R = 4 and two of the loading matrices have k-rank < rank is unique if and only if $r(\mathbf{M}_i)+r(\mathbf{M}_j)\geq R+2$ for all $i\neq j$ and the number of basis vectors with constraint elements is greater than one or $\min(r(\mathbf{M_1})-k(\mathbf{M_1}), r(\mathbf{M_2})-k(\mathbf{M}_2))$.

Copyright © Heather Michele Clyburn Bush 2006

5. CONSTRAINTS FROM OC SPACES (R = 5 AND R = 6)

The investigation of the necessity of Kruskal's k-rank sum condition had been limited to the cases where R was 3 or 4. The approach of ten Berge and Sidiropoulos had been unable to answer the question of uniqueness for all decompositions where R = 4. The OCSA, on the other hand, had been able to succinctly evaluate all the possible PARAFAC decomposition forms for R = 4, adding understanding to why some solutions were unique while others were not. In an attempt to identify whether or not the same trends occurred for R larger than 4, cases where R was 5 and 6 were evaluated. Until this dissertation, general PARAFAC decompositions forms that did not meet Kruskal's k-rank sum condition had not been investigated for uniqueness. The results of applying the OCSA to PARAFAC decompositions with 5 and 6 columns are presented in this chapter.

5.1 What the Constraints Showed for R = 5 (k-rank = rank)

PARAFAC solution uniqueness when Kruskal's sufficient condition for uniqueness was not met was not investigated by ten Berge and Sidiropoulos for R = 5. Therefore, the OC to the column space of the various KR products would be the first assessment of uniqueness for R = 5.

For PARAFAC solutions with R = 5, the increased number of columns increased the number of variations when k-rank < rank. More importantly, however, the increased number of columns increased the complications of investigating alternative forms. With larger ranks and smaller k-ranks, the complexity of the problem grows to one that needs to be handled through computation. Unfortunately, the computational techniques needed to correctly quantify alternative solutions are unavailable at this time. Therefore, the subsequent discussions about PARAFAC solution uniqueness will be limited to the case where k-rank = rank. Even so, the discourse concerning uniqueness was the first thorough review of uniqueness when Kruskal was not met for R > 4. Additionally, the principles presented here for k-rank = rank will be applicable for evaluating the cases that could result with PARAFAC solutions with deficient k-rank and demonstrate that the answer to uniqueness does not lie simply with k-rank conditions.

For the cases where the KR product was formed by loading matrices where the k-rank and the ranks were equivalent, a similar pattern as R = 4 and k-rank = rank emerged (Table 5.1). The details and descriptions for each of these KR products were presented in the following pages where each KR product was described in terms of the following:

- The number of basis vectors in the OC space;
- The number of basis vectors that contain constraint elements; and,
- A description of the constraint elements.

Finally, for each of these KR products, a third matrix was considered and decisions on uniqueness were determined for the resulting PARAFAC solutions.

Table 5.1 Pairwise Combinations of Loading Matrices (k-rank=rank)

k-rank (M1)	k-rank (M2)	rank(M1)	rank(M2)	P/S Only?
2	2	2	2	No
2	3	2	3	No
2	4	2	4	No
2	5	2	5	Yes
3	3	3	3	No
3	4	3	4	Yes
3	5	3	5	Yes
4	4	4	4	Yes
4	5	4	5	Yes

5.1.1 Evaluating KR Products

Before proceeding to the specific cases where R = 5 and k-rank = rank, lessons from R = 4 were employed and the class where k-ranks and ranks were equivalent was further subdivided. The class of matrices where the sum of the ranks of the loading matrices was less than R+2 = 7 was considered separately from the cases where the KR products were composed of loading matrices with ranks which summed to 7 or more. The class of KR products with ranks summing to less than 7 would include KR products with loading matrices with k-ranks of

(2,2), (2,3), (2,4), and (3,3). The class of KR products where the ranks summing 7 or more would include the loading matrices with k-ranks of (2,5), (3,4), (3,5), (4,4), and (4,5).

5.1.1.1 KR Products: k-ranks of 2 and 2

The first case in this class for R = 5, where both loading matrices have k-rank and rank equal to 2, resulted in no constraints. Thus, like in the case where R = 4, KR products composed of two matrices with k-rank = rank = 2 could have alternatives that were not restricted to permutation scale alternatives.

Result 5.1 For the KR product composed of matrices with k-ranks of (2,3) the following was found:

- The OC resulted in the empty set, therefore,
- No vectors contained constraint elements
- No constraint elements were formed

5.1.1.2 KR Products: k-ranks of 2 and 3

In this case, there was only one basis vector for the OC to the column space of the KR product. The vector, however, contained three constraint elements from the two loading matrices: two ratios with elements from both loading matrices and one ratio with only elements from the loading matrix with k-rank = 2. Even so, these constraints were not sufficient to limit alternative loading matrices to be permutation-scale transformations.

Result 5.2 For the KR product composed of matrices with k-ranks of (2,3) the following was found:

- One OC basis vector resulted
- One vector contained constraint elements
- Two constraint elements formed from both loading matrices and one formed from the elements of only one

5.1.1.3 KR Products: k-ranks of 2 and 4

The next case considered was where the loading matrices had k-ranks of 2 and 4. All three of the basis vectors for the OC to the column space of the KR product contained constraint elements. In fact, two of the three vectors contained two sets of constraint elements or ratios. In each set, one ratio was comprised of elements from both loading matrices while the other contained only elements from the loading matrix with k-rank = 2. The third vector contained only one ratio which was formed from elements from both loading matrices and did not force permutation-scale alternatives.

Result 5.3 For the KR product composed of matrices with k-ranks of (2,3) the following was found:

- Three OC basis vectors resulted
- All three vectors contained constraint elements
- Three constraint elements were formed from elements from both loading matrices and two were formed from the elements of only one

5.1.1.4 KR Products: k-ranks of 3 and 3

Finally, the last case where both loading matrices had k-rank = rank = 3, resulted in four basis vectors for the OC to the column space of the KR product. Each of these four vectors contained two constraint elements which were ratios composed of elements from both loading matrices. As in the other k-rank cases for this subclass, the alternative loading matrices were not limited to permutation-scale versions of the originals.

Result 5.4 For the KR product composed of matrices with k-ranks of (3,3) the following was found:

- Four OC basis vectors resulted
- All four vectors contained constraint elements, two in each
- Eight constraint elements were formed from both loading matrices

5.1.1.5 KR Products: k-ranks of 2 and 5

In the first case, where k-ranks were 2 and 5, five basis vectors resulted. Of these five basis vectors, 3 contained constraint elements, ratios formed form elements of the loading matrix with k-rank of 2. Of course, since these KR products were evaluated using loading matrices in RREF, the matrix with k-rank of 5 would simply be the identity matrix and could not contribute any values to the KR product or the basis vectors for the OC space. These constraints were sufficient to require that all alternative forms be permutation-scale transformations.

Result 5.5 For the KR product composed of matrices with k-ranks of (2,5) the following was found:

- Five OC basis vectors resulted
- Three vectors with constraint elements resulted
- All three constraint elements were formed from elements from one of the loading matrices

5.1.1.6 KR Products: k-ranks of 3 and 4

In the case where the KR product was formed from two matrices with k-ranks of 3 and 4, both matrices would have elements that could contribute to the basis vectors of the OC space. In this case, all seven basis vectors contained constraint elements. Five of these vectors contained only one ratio of elements, while two contained two ratios of elements. For the five vectors with only one ratio, three of these vectors/ratios contained values from both matrices. The other two vectors with only one ratio, each had a ratio consisting of values from only one of the loading matrices; one had a ratio consisting of elements from the matrix with k-rank = 3 and the other had a ratio of values from the matrix with k-rank = 4. Both of the final two vectors with two ratios had one ratio with elements from the matrix with k-rank = 3 and one ratio with elements from both loading matrices. These constraints did impose that all alternative forms be permutation-scale transformations.

Result 5.6 For the KR product composed of matrices with k-ranks of (3,4) the following was found:

- Seven OC basis vectors resulted
- Seven constraint vectors resulted
- Five constraint elements were formed from elements from both loading matrices and four were formed from the elements of only one

5.1.1.7 KR Products: k-ranks of 3 and 5

When one loading matrix had k-rank of 3 and the other k-rank of 5, only the matrix with k-rank of 3 would contribute values to the constraints formed from the basis vectors of the OC space. Four out of ten of these vectors contained constraint elements from the matrix with k-rank = 3, which limited alternatives to be permutation-scale transformations.

Result 5.7 For the KR product composed of matrices with k-ranks of (3,5) the following was found:

- Ten OC basis vectors resulted
- Four vectors had constraint elements
- Four constraint elements were formed from elements of the loading matrix with k-rank = 3

5.1.1.8 KR Products: k-ranks of 4 and 4

When both of the matrices had k-ranks of 4, each of the eleven basis vectors containing exactly one ratio composed of elements from the loading matrices. Seven of these eleven ratios had elements from both loading matrices, while the remaining four basis vectors had ratios involving elements from only one of the loading matrices, two from one and two from the other. Again, as in all the previous cases, the constraints imposed by these ratios required that all alternative forms be permutation-scale transformations.

Result 5.8 For the KR product composed of matrices with k-ranks of (4,4) the following was found:

- Eleven OC basis vectors resulted
- Eleven vectors had constraint elements
- Seven constraint elements were formed from elements from both loading matrices and the other four were formed from elements from only one loading matrix

5.1.1.9 KR Products: k-ranks of 4 and 5

Finally, the last case in this subclass occurred when one of the loading matrices had k-rank = 4 and the other had k-rank = 5. Three of the fifteen basis vectors contained constraint elements from the loading matrix with k-rank of 4. Although each of these vectors contained only one ratio, the resulting constraints would force alternative KR products to be comprised of loading matrices that were simply permutation-scale versions of the originals.

Result 5.9 For the KR product composed of matrices with k-ranks of (4,5) the following was found:

- Fifteen OC basis vectors resulted
- Three vectors had constraint elements
- Three constraint elements were formed from elements of the loading matrix with k-rank = 4.

5.1.2 PARAFAC Decomposition Uniqueness (R = 5, k-rank = rank)

The pairwise combinations of loading matrices and the resulting KR products could now be used to evaluate when PARAFAC solutions with loading matrices with k-rank = rank were unique. The details of the process of identifying uniqueness were given for the case where R = 4. Each KR product that could be formed from the loading matrices of a PARAFAC solution were evaluated. When every KR product had alternatives that were restricted to permutation-scale transformations, uniqueness was determined. On the other hand, when any KR product of a PARAFAC solution was not restricted to permutation-scale transformations, non-uniqueness was established.

The PARAFAC solutions that were possible when all loading matrices had k-rank = rank and the sum of the k-ranks for the three matrices did not reach or exceed 2R+2 were listed, with conclusions concerning uniqueness, in Table 5.2.

Table 5.2 PARAFAC Decomposition Uniqueness for R = 5, k-rank = rank

k-rank (M1)	rank(M1)	k-rank (M2)	rank(M2)	k-rank (M3)	rank(M3)	Uniqueness?
2	2	2	2	2	2	No
2	2	3	3	2	2	No
2	2	4	4	2	2	No
2	2	5	5	2	2	No
3	3	3	3	2	2	No
3	3	4	4	2	2	No
3	3	5	5	2	2	No
4	4	4	4	2	2	No
4	4	5	5	2	2	No
2	2	5	5	3	3	No
3	3	3	3	3	3	No
3	3	4	4	3	3	No
3	3	5	5	3	3	No
4	4	4	4	3	3	Yes

Considering only the class of PARAFAC solutions where the k-rank = rank for all loading matrices, it was shown that imposing the additional condition that k-rank = rank would not cause Kruskal's condition to be necessary and sufficient. As in the case where R = 4, another condition would be needed.

5.1.3 Necessary and Sufficient Conditions for Uniqueness

From Table 5.2, it appeared that most PARAFAC decompositions when k-rank = rank and R = 5 were non-unique. However, the PARAFAC decomposition where the k-ranks were 4, 4, and 3 was found to be unique. This same pattern was seen in the case where R = 4 and k-rank = rank. From Table 5.1 and Table 5.2, it was obvious that when k-rank = rank and the sum of the ranks for any pairwise matrix combination was less than R + 2, the PARAFAC solution would be non-unique. Therefore, when the pairwise sum of the ranks was at least R + 2 and k-rank=rank, the PARAFAC decomposition was unique. This is true when Kruskal's condition is not met as well as when it is met. Thus, a necessary and sufficient condition for uniqueness when R = 5 and k-rank=rank is that all of the pairwise sum of the ranks are at least R + 2.

Result 5.10 A PARAFAC decomposition where k-rank = rank and R = 5 is unique if and only if $r(\mathbf{M}_i) + r(\mathbf{M}_j) \geq R + 2$ for all $i \neq j$.

It is important to note here that this is the same conclusion achieved for R = 3 and R = 4. The most recent approaches in this area have focused on evaluating individual cases for R and making conjectures that might extend to all R. To that point, it appears that the condition on the sum of the ranks will provide a necessary and sufficient conditions for uniqueness for all R when k-rank=rank.

5.2 What the Constraints Showed for R = 6 (k-rank = rank)

Prior to this work, the uniqueness of PARAFAC solutions have not been evaluated for R = 6. Below, the case where R = 6 and k-rank = rank was evaluated using the OC to the column space of the KR product. First, the KR products for each of the pairwise combinations were analyzed by investigating the constraints formed from the basis vectors of the OC space. After conclusions regarding alternatives were made, all three loading matrices of a PARAFAC solution could be assessed to identify whether or not the solution would be unique.

5.2.1 Evaluating KR products

For each of the KR products where k-rank = rank, the basis vectors for the OC space were described. Included in the description of the basis vectors would be the number of vectors with constraints and the number and type of constraint elements that resulted. Depending on the constraints that resulted, the alternatives could be restricted to permutation-scale transformations (Table 5.3).

Table 5.3 Pairwise Combinations of Loading Matrices (k-rank = rank)

k-rank (M1)	k-rank (M2)	rank(M1)	rank(M2)	P/S Only?
2	2	2	2	No
2	3	2	3	No
2	4	2	4	No
2	5	2	5	No
2	6	2	6	Yes
3	3	3	3	No
3	4	3	4	No
3	5	3	5	Yes
3	6	3	6	Yes
4	4	4	4	Yes
4	5	4	5	Yes
4	6	4	6	Yes
5	5	5	5	Yes
5	6	5	6	Yes

As in the case where R = 5, for ease in evaluating the differences in OC constraints, the KR products were divided into two classes:

- KR products where the sum of the ranks of the loading matrices was less than R+2 = 8, and
- KR products where the sum of the ranks of the loading matrices was at least R+2 = 8.

5.2.1.1 KR products: k-ranks of (2,2) or (2,3)

In both cases where the k-ranks of the loading matrices were 2 and 2 or 2 and 3, no orthogonal basis vectors resulted. Thus, no constraints were found for the alternative KR products with these matrices. Hence, for both types of KR products, alternatives were not restricted to be permutation-scale transformations.

Result 5.11 For the KR product composed of matrices with k-ranks of (2,2) and (2,3) the following was found:

- No OC basis vectors resulted
- None contained constraint elements
- No constraint elements were formed.

5.2.1.2 KR products: k-ranks of 2 and 4

In this case, the two basis vectors contained constraint elements from both matrices. Each vector contained three constraint elements; two were ratios formed from elements from both matrices, while one was a ratio containing only elements from the matrix with k-rank of 2.

Result 5.12 For the KR product composed of matrices with k-ranks of 2 and 4 the following was found:

- Two OC basis vectors resulted
- Both contained constraint elements
- Four constraint elements were formed from elements from both matrices and two were formed from elements of the matrix with k-rank of 2.

Alternative forms composed of loading matrices with these matrix properties were not limited to permutation-scale transformations.

5.2.1.3 KR product: k-ranks of 2 and 5

The next k-rank combination possible was when the k-ranks were 2 and 5. Again, in this case, alternatives were not restricted to be permutation-scale alternatives. Four basis vectors resulted for these KR products and all had constraint elements in the form of ratios. Three of the four vectors had two ratios: one with elements from both matrices and one with only elements from the matrix with k-rank of 2. The last vector had only one ratio but it was comprised of elements from both loading matrices.

Result 5.13 For the KR product composed of matrices with k-ranks of 2 and 5 the following was found:

- Four OC basis vectors resulted
- Four vectors had constraint elements
- Four constraint elements were formed from elements of both matrices and three elements were formed from elements of only one.

5.2.1.4 KR products: k-ranks of 3 and 3

In the next scenario, where both matrices had k-rank = 3, all three of the basis vectors contained constraint elements. In fact, each vector contained three ratios composed of elements from both loading matrices.

Result 5.14 For the KR product composed of matrices with k-ranks of 3 and 3 the following was found:

- Three OC basis vectors resulted
- Three vectors had constraint elements
- Nine constraint elements were formed from elements of both matrices.

The constraints imposed by these ratios were not sufficient to restrict alternative forms to be permutation-scale transformations.

5.2.1.5 KR product: k-ranks of 3 and 4

Likewise, in the case where one of the loading matrices had k-rank = 3 and the other had k-rank = 4, the constraints from considering the OC space did not require alternatives to be permutation-scale alternatives. In this case, all six of the basis vectors contained constraint elements. Two of these vectors had three ratios, two of which were created from elements of both matrices and one with only elements from the loading matrix with k-rank = 3. The remaining basis vectors had two ratios each, and both ratios contained elements from both loading matrices.

Result 5.15 For the KR product composed of matrices with k-ranks of 3 and 4 the following was found:

- Six OC basis vectors resulted
- Six vectors had constraint elements
- Twelve constraint elements were formed from elements of both matrices and two elements were formed from elements of only one.

The remaining cases where k-rank = rank involved matrices where the pairwise sum of the ranks was greater than or equal to R + 2 = 8. In each of these remaining cases, the conditions imposed by the elements obtained by evaluating the orthogonal complement to the column space of the KR product were enough to restrict all alternative forms to be permutation-scale versions of the original.

5.2.1.6 KR products: k-ranks of 2 and 6

For loading matrices with k-ranks of 2 and 6, four of the six basis vectors for the OC to the column space of the KR product had constraint elements. Each of these four vectors contained a single ratio comprised of elements from the loading matrix with k-rank of 2. Alternative forms where the loading matrices had k-rank of 2 and 6 were found to be restricted to permutation-scale versions of the originals.

Result 5.16 For the KR product composed of matrices with k-ranks of 2 and 6 the following was found:

- Six OC basis vectors resulted
- Four vectors had constraint elements
- Four constraint elements were formed from elements of the matrix with k-rank of 2.

5.2.1.7 KR products: k-ranks of 3 and 5

In the case where the k-ranks were 3 and 5, all nine basis vectors contained constraint elements. Four of these vectors each contained two ratios, one with elements from both and one with elements from the loading matrix with the smaller k-rank. The other five vectors contained only one ratio; three of these had elements from both matrices, while the other two had ratios comprised of elements from the matrix with k-rank of 3 and k-rank of 5, respectively.

Result 5.17 For the KR product composed of matrices with k-ranks of 3 and 5 the following was found:

- Nine OC basis vectors resulted
- Nine vectors had constraint elements
- Seven constraint elements were formed from elements of both matrices and six were formed from elements from only one matrix.

5.2.1.8 KR products: k-ranks of 3 and 6

The next case evaluated was when the loading matrices had k-ranks of 3 and 6. Twelve basis vectors resulted from considering the OC to the column space of the KR product, of these twelve, six contained constraint elements from the matrix with k-rank = 3.

Result 5.18 For the KR product composed of matrices with k-ranks of 3 and 6 the following was found:

- Twelve OC basis vectors resulted
- Six vectors had constraint elements
- Six constraint elements formed from the loading matrix with k-rank = 3.

5.2.1.9 KR products: k-ranks of 4

In the case where both of the loading matrices had k-rank = 4, ten basis vectors resulted. All ten of these vectors contained two constraint elements, ratios consisting of elements from both of the loading matrices.

Result 5.19 For the KR product composed of matrices with k-ranks of 4 and 4 the following was found:

- Ten OC basis vectors resulted
- Ten vectors had constraint elements
- Twenty constraint elements with elements from both loading matrices

5.2.1.10 KR products: k-ranks of 4 and 5

When one loading matrix had k-rank = 4 and the other k-rank = 5, fourteen basis vectors resulted, all with constraint elements. Three of these vectors had two ratios, one with elements from both loading matrices and one with elements from only the matrix with k-rank = 4. The remaining eleven vectors contained only one ratio each. Seven of these eleven had a ratio with elements from both loading matrices, two had elements from only the loading matrix with k-rank = 4, and two had elements from only the matrix with k-rank = 5.

Result 5.20 For the KR product composed of matrices with k-ranks of 4 and 5 the following was found:

- Fourteen OC basis vectors resulted
- Fourteen vectors had constraint elements
- Ten constraint elements were formed from elements of both loading matrices and seven had elements formed from only one loading matrix.

5.2.1.11 KR products: k-ranks of 4 and 6

With k-ranks of 4 and 6, six of the resulting eighteen basis vectors had one ratio containing elements from the loading matrix with k-rank = 4.

Result 5.21 For the KR product composed of matrices with k-ranks of 4 and 5 the following was found:

- Eighteen OC basis vectors resulted
- Six vectors had constraint elements
- Six constraint elements formed from the loading matrix with k-rank of 4.

5.2.1.12 KR products: k-ranks of 5

For the KR product composed of loading matrices where both k-ranks were 5, all nineteen basis vectors contained one ratio of elements from the loading matrices. Thirteen basis vectors contained a ratio consisting of elements from both loading matrices, three vectors had a ratio with elements from only one of the loading matrices, and three vectors had a ratio with elements from the other loading matrix.

Result 5.22 For the KR product composed of matrices with k-ranks of 5 and 5 the following was found:

- Nineteen OC basis vectors resulted
- Nineteen vectors had constraint elements
- Thirteen constraint elements were formed from elements of both loading matrices and six had elements from only one matrix.

5.2.1.13 KR products: k-ranks of 5 and 6

Finally, when the loading matrices had k-rank = 5 and k-rank = 6, only four of the twenty-four basis vectors had constraint elements from the matrix with k-rank = 5. Each of these four vectors contained only one ratio.

Result 5.23 For the KR product composed of matrices with k-ranks of 5 and 6 the following was found:

- Twenty-four OC basis vectors resulted
- Four vectors had constraint elements
- Four constraint elements were formed from elements of the loading matrix with k-rank of 5.

5.2.2 PARAFAC Decomposition Uniqueness (R = 6, k-rank = rank)

Utilizing the results of investigating the constraints obtained by the OCSA for each of the KR products where k-rank = rank, it was possible to ascertain which PARAFAC solutions would be unique and which would be non-unique when k-rank = rank. By considering all possible pairwise combinations of two loading matrices, it was possible to utilize Table 5.3 to determine whether or not all pairwise combinations would result in alternative KR products that were restricted to permutation-scale transformations only (Table 5.4).

Table 5.4 PARAFAC Decomposition Uniqueness for R = 6, k-rank = rank

k-rank (M1)	rank(M1)	k-rank (M2)	rank(M2)	k-rank (M3)	rank(M3)	Uniqueness?
2	2	2	2	2	2	N
2	2	3	3	2	2	N
2	2	4	4	2	2	N
2	2	5	5	2	2	N
2	2	6	6	2	2	N
3	3	3	3	2	2	N
3	3	4	4	2	2	N
3	3	5	5	2	2	N
3	3	6	6	2	2	N
4	4	4	4	2	2	N
4	4	5	5	2	2	N
4	4	6	6	2	2	N
5	5	5	5	2	2	N
5	5	6	6	2	2	N
3	3	3	3	3	3	N
3	3	4	4	3	3	N
3	3	5	5	3	3	N
3	3	6	6	3	3	N
4	4	4	4	3	3	N
4	4	5	5	3	3	N
4	4	6	6	3	3	N
5	5	5	5	3	3	Y
4	4	4	4	4	4	Y
4	4	5	5	4	4	Y

5.2.3 Necessary and Sufficient Conditions for Uniqueness

For k-rank = rank and R = 6, most PARAFAC solutions were non-unique. However, when the sum of each pairwise combination of ranks equaled or exceeded R+2 = 8, the PARAFAC solution was unique. Thus, k-rank = rank would not make Kruskal's sufficient condition necessary for R = 6. On the other hand, when the pairwise sum of the ranks was at least R + 2 and k-rank=rank, the PARAFAC decomposition was unique. This is true when Kruskal's condition is not met as well as when it is met. Thus, a necessary and sufficient condition for uniqueness when R = 6 and k-rank=rank is that all of the pairwise sum of the ranks are at least R + 2.

Result 5.24 A PARAFAC solution where k-rank = rank and R = 6 is unique if and only if $r(\mathbf{M}_i)+r(\mathbf{M}_j)\geq R+2$ for all $i \neq j$.

Thus, in all of the cases considered, R = 3, 4, 5, and 6, the same conditions for uniqueness were found when k-rank=rank. For many years, uniqueness was thought to be directly linked with the k-rank of the loading matrices. While this is still true, the reason necessary and sufficient conditions have been so elusive is that considering k-rank does not reveal the entire picture. Even though these results point to uniqueness conditions for the cases where k-rank = rank, these same methods could be applied to investigate cases with deficient k-rank. Additionally, with the advent of some computational tools it will be entirely possible to apply the OCSA to determine uniqueness for all the cases for a particular R. As will be discussed in the final chapter, it is expected that the application of this technique for k-rank < rank will answer much of the question about uniqueness as it will be possible to look at cases that appear similar but have differing uniqueness results and OC constraints.

Copyright © Heather Michele Clyburn Bush 2006

6. THEOREMS FOR RANK, K-RANK, AND OC CONSTRAINTS

From the results that were observed for R = 3, 4, 5, and 6 for k-rank = rank, it was shown that a condition on the pairwise sums of ranks was necessary and sufficient for uniqueness. As demonstrated in these cases, when the sum of the ranks of two loading matrices reached or exceeded R+2, it was possible to invoke Kruskal's condition, and so uniqueness, with particular third loading matrices. In other cases, non-uniqueness was found because the pairwise sum of the ranks was less than R+2. Thus, it would appear that understanding uniqueness and developing general necessary and sufficient conditions for uniqueness when k-rank=rank would depend on the differences between solutions where the sum of the ranks in the KR product achieves R+2 and where it does not. Specifically, what matrices are allowable when the sums do not achieve R+2 and when the sums were R+2 and greater.

The following theorems help in this search and could prove to be very useful as the approach in this dissertation is extended for general R, where k-rank = rank. Additionally, as computational advances are made, these theorems should also provide the basis for developing a set of conditions for general R without restricting k-rank to be equivalent to rank. By providing the theoretical foundation of how these KR products differ with respect to OC constraints, the first steps for obtaining conditions for uniqueness, possibly necessary and sufficient conditions for uniqueness, for any R have been made.

The first section of proofs provide general information about the types of loading matrices that can result when considering ranks that are restricted by the R+2 condition. In the second section, theorems are offered for the connections between the rank arguments and OC constraints for general R. Finally, uniqueness is addressed for general R, providing theoretical evidence that ten Berge and Sidiropoulos' conjecture regarding k-rank = rank decompositions is not correct.

6.1 Concepts of ranks, k-ranks, and OC constraints

In general, larger ranks seemed to supply enough constraints so that alternative forms were limited to permutation-scale transformations. The rationalization for this could also be found in the theory of OC spaces. As the ranks (or the number of rows in the RREFs) increased, the product of the two ranks would also increase. Since the number of basis vectors for the OC space was the difference between this product and the number of columns (R), as the rank increased so would the number of basis vectors for the OC to the column space of the KR product. When the k-rank was equivalent to the rank and rank < R, each basis vector would contain at least one constraint vector. Thus, for k-rank = rank < R, the number of constraints would increase as the number of basis vectors increased. Similarly, the increase in rank and dimension of the OC space, would increase the constraints in other KR product cases as well. The increased number of constraints on the alternative forms would limit the alternatives to permutation-scale transformations. Therefore, it appeared that conditions for uniqueness would need to incorporate ranks.

However, k-ranks must also be considered, as noted by Kruskal and others that have pursued uniqueness conditions. The k-rank of a matrix had a direct impact on the number of constraint elements that were contained in the OC basis vectors. The patterns found in considering the OC space basis vectors with constraints suggested that, although k-rank had served as an indicator for uniqueness, it was too general of a concept to provide necessary and sufficient conditions for uniqueness. Even so, the k-rank of a matrix would be related to the constraint elements that appeared in the OC basis vectors, and considering the basis vectors in addition to k-rank might provide the detail needed to offer necessary and sufficient conditions for uniqueness. Therefore, it would be necessary to link the relationship of k-rank, rank, OC basis vectors, and constraint elements together.

The results presented in the previous chapters described the OC space basis vectors, the resulting constraints, and their effects on uniqueness. Although these results were presented for specific R, the following theorems will prove certain properties for all R. From the results presented in Chapters 4 and 5, certain trends and properties became obvious in the OC constraints that resulted.

These trends and properties are best described by grouping the KR products into three types:

- KR products with k-rank = rank and the sum of the ranks totaling to less than R+2 , when the maximum rank is less than R
- KR products with one of the loading matrices with k-rank = rank = R and the other with k-rank = rank < R.
- KR products with k-rank = rank and the sum of the ranks are at least R+2 when the maximum rank is less than R.

Finally, using the results for KR products where k-rank = rank and the sum of the ranks is R+2 or greater, it will be possible to prove, for all R, that k-rank = rank will not make Kruskal's sufficient condition necessary. Future work in uniqueness will need to consider OC constraints and the resulting conditions on rank, and the work presented here will help to lay the groundwork for those who follow.

6.2 Theorems on the allowable ranks of loading matrices

For the sum of the ranks of two loading matrices that achieved R+2 or greater, Theorem 6.1 will show that when the largest rank of a loading matrix is not equivalent to the number of columns, the other loading matrix must have rank greater than 2. It has already been assumed that all loading matrices have a k-rank of at least 2, a necessary condition for uniqueness (Krijnen, 1993). Therefore, when the maximum rank does not reach R, the other loading matrices would have to be more than 2 in order to reach R+2.

Theorem 6.1 Let $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$ and $r(\mathbf{M_i}) + r(\mathbf{M_j}) \geq R + 2$, then $r(\mathbf{M_j}) > 2$.

Proof 6.1 Since $r(\mathbf{M_i}) + r(\mathbf{M_j}) \geq R + 2$, $r(\mathbf{M_j}) \geq R + 2 - r(\mathbf{M_i})$. Assuming, $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$, then $R + 2 - r(\mathbf{M_i}) > R + 2 - R = 2$ or $r(\mathbf{M_j}) > 2$.
□

For the sum of the ranks of two loading matrices that does not reach R+2, Theorem 6.2 demonstrates that both of the matrices in the sum have deficient rank or rank < R.

Theorem 6.2 Let $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k}))$ and $r(\mathbf{M_i}) + r(\mathbf{M_j}) < R + 2$, then $r(\mathbf{M_i}) < R$ and $r(\mathbf{M_j}) < R$.

Proof 6.2 Let $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k}))$. Since $r(\mathbf{M_i}) + r(\mathbf{M_j}) < R + 2$, $r(\mathbf{M_i}) < R + 2 - r(\mathbf{M_j})$. A necessary condition for uniqueness, which had already been assumed, was that $k(\mathbf{M_m}) \geq 2, \forall m$. By definition, the k-rank of a matrix was at least the rank of a matrix, or $k(\mathbf{M_m}) \leq r(\mathbf{M_m})$. Hence, if the lower bound on the k-rank of a matrix was 2, then $r(\mathbf{M_m}) \geq 2, \forall m$. Therefore, $r(\mathbf{M_i}) < R + 2 - r(\mathbf{M_j}) \leq R + 2 - 2 = R$. Hence, $r(\mathbf{M_i}) < R$. Since $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k}))$, $r(\mathbf{M_j}) \leq r(\mathbf{M_i}) < R$, so $r(\mathbf{M_j}) < R$. □

Consequently, the ranks that formed the sum, whether it was R+2 and greater or less than R+2, would be of certain sizes. Limiting the sum of the ranks to be less than R+2 would necessarily force the ranks of the loading matrices to be smaller, and allowing the sum to be R+2 or more would allow for the ranks to be larger. Additionally, for k-rank = rank and Kruskal's sufficient condition unsatisfied, the ranks of the loading matrices would be restricted further, as seen in Theorem 6.3.

Theorem 6.3 For $r(\mathbf{M_m}) = k(\mathbf{M_m}), \forall m, m = 1,2,3$, let $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) = R$ and $k(\mathbf{M_i}) + k(\mathbf{M_j}) + k(\mathbf{M_k}) < 2R + 2$, then $r(\mathbf{M_j}) < R$ and $r(\mathbf{M_k}) < R$.

Proof 6.3 Let $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) = R$ .When Kruskal's condition was not met, $k(\mathbf{M_i}) + k(\mathbf{M_j}) + k(\mathbf{M_k}) < 2R + 2$. A necessary condition for uniqueness, which had already been assumed, was that $k(\mathbf{M_m}) \geq 2, \forall m$. Therefore, $k(\mathbf{M_i}) + k(\mathbf{M_j}) < 2R + 2 - k(\mathbf{M_k}) < 2R + 2 - 2 = 2R$. Consequently, since $r(\mathbf{M_m}) = k(\mathbf{M_m}), \forall m$, $r(\mathbf{M_i}) + r(\mathbf{M_j}) < 2R$. It was assumed that $r(\mathbf{M_i}) = R$, therefore, $r(\mathbf{M_i}) + r(\mathbf{M_j}) = R + r(\mathbf{M_j}) < 2R$. Hence, $r(\mathbf{M_j}) < R$. Likewise, in a similar approach it could also be shown that $r(\mathbf{M_k}) < R$. □

6.3 OC Spaces and KR products

The relationship between the number of basis vectors in an OC space and the number of rows in a matrix was a relatively simple idea from elementary linear algebra, but had not heretofore been applied to KR products. The number of basis vectors for the OC space was based on the product of the ranks.

The loading matrices investigated in the OCSA (and the tBSM) were in reduced form. Therefore, the number of rows in each matrix was equivalent to the rank of the matrix. When the KR product was formed from these matrices, the number of rows that resulted in the KR product would be $r(\mathbf{M_i}) \cdot r(\mathbf{M}_j)$. Thus, the number of basis vectors in the OC space would be directly related to the product of the ranks of the loading matrices.

Regardless of k-rank, when KR products were created from any loading matrices with the same ranks, the same number of OC space basis vectors would result. Hence, when the KR product was formed from loading matrices having k-rank = rank, the number of basis vectors would be the same as KR products formed where k-rank < rank but the ranks were the same. Therefore, a KR product formed from loading matrices with k-ranks summing to R+2 or more would have the same number of OC basis vectors as a KR product formed from loading matrices with k-ranks summing to less than R+2 but with the same ranks summing to R+2 or more.

Hence, it was noted that any differences in permutation-scale restrictions between two sets of matrices with the same ranks would not be due to differences in the number of basis vectors, but in the constraint elements within those basis vectors. Thus, the ranks of the loading matrices, the size of their sum, and the constraint elements became of particular interest when trying to determine uniqueness for PARAFAC solutions. Therefore, theorems were created and proved to provide some general information regarding loading matrices and the sum of their ranks.

Theorem 6.4 The number of basis vectors in the OC space of $(\mathbf{M}_j \circ \mathbf{M}_i)$ is $\eta = r(\mathbf{M}_i)r(\mathbf{M}_j) - R$.

Proof 6.4 The dimension of a space is equivalent to the number of basis vectors. Also, the OC space of a matrix consisted of all **x** such that $\mathbf{Mx} = 0$. In the case of KR products, the OC space of interest was the OC to the column space of $(\mathbf{M}_j \circ \mathbf{M}_i)$. Hence, all vectors **x**, where $(\mathbf{M}_j \circ \mathbf{M}_i)^t \mathbf{x} = 0$, comprised the OC space to the column space of $(\mathbf{M}_j \circ \mathbf{M}_i)$. The size of $(\mathbf{M}_j \circ \mathbf{M}_i)^t$ was $r(\mathbf{M}_i)r(\mathbf{M}_j) \times R$ and, per the assumption that KR products were full column-rank, the $r\left((\mathbf{M}_j \circ \mathbf{M}_i)^t\right) = R$. Therefore, from linear algebra, the number of basis vectors in the OC to the column space of $(\mathbf{M}_j \circ \mathbf{M}_i)$ was $r(\mathbf{M}_i)r(\mathbf{M}_j) - R$. □

Although the above theorem was simply an application of a well-known principle in linear algebra to matrices in KR product form, the relationship of constraint elements and number of basis vectors had not been explored. The following theorems are presented to define, for all R, the number of OC space basis vectors that contain constraint elements. The theorems will define the number of basis vectors with constraints when the maximum rank of the loading matrices is less than and equal to R, Theorems 6.5 and 6.10, respectively. Additionally, theorems will be given that define the number of constraint vectors that result when the sum of the matrices are less than R+2, at least R+2 but maximum rank < R, and at least R+2 but maximum rank = R. These theorems are the first of their kind and will help to provide a path for determining conditions for uniqueness.

Theorem 6.5 For $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$, and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$, the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ was $\eta = r(\mathbf{M_i})r(\mathbf{M_j}) - R$.

Proof 6.5 In the case where $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$, and $r(\mathbf{M_i}) < R$, the matrices $\mathbf{M_i}$ and $\mathbf{M_j}$ had $R - r(\mathbf{M_i})$ and $R - r(\mathbf{M_j})$ columns with elements from $\mathbf{M_i}$ and $\mathbf{M_j}$, respectively. Additionally, because rank and k-rank were equivalent and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$, there were no elements equal to 0 in these columns. Hence, there would be at least $R - r(\mathbf{M_i})$ columns without any 0 entries. As a consequence, no rows of $(\mathbf{M_j} \circ \mathbf{M_i})$ would contain all 0 elements. Therefore, no columns of $(\mathbf{M_j} \circ \mathbf{M_i})^{\mathbf{t}}$ would be comprised of all 0 elements, and so none of the basis vectors of the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$, or the null space of $(\mathbf{M_j} \circ \mathbf{M_i})^{\mathbf{t}}$, would be unit null vectors. Hence, all basis vectors for the null space of $(\mathbf{M_j} \circ \mathbf{M_i})^{\mathbf{t}}$ would contain at least one constraint element or combination of constraint elements from $\mathbf{M_i}$ or $\mathbf{M_i}$. Theorem 6.4 established that the number of basis null vectors for the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ was $\eta = r(\mathbf{M_i})r(\mathbf{M_j}) - R$. Therefore, for $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$, and $r(\mathbf{M_i}) < R$, the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_i}$ is $r(\mathbf{M_i})r(\mathbf{M_j}) - R$. □

The pairwise sum of loading matrix ranks were evaluated in three classes: sums less than R+2; sums greater than or equal to R+2 where a loading matrix had full-column rank; and, sums greater than or equal to R +2 where the loading matrices were not full-column rank. For each of these sets of loading matrices the following results were found:

- A bound on the number of OC space basis vectors was obtained
- A bound on the number of basis vectors with constraint elements was obtained
- Necessary and sufficient conditions for when the bounds would be achieved were defined

Utilizing the elementary theorems presented above, it was possible to provide theoretical remarks on whether or not KR products would be limited to permutation-scale transformations, and ultimately whether PARAFAC solutions were unique.

Theorem 6.6 For $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$, if $r(\mathbf{M_i}) + r(\mathbf{M_j}) < R + 2$, then $\eta = r(\mathbf{M_i})r(\mathbf{M_j}) - R < 2R - 3$.

Proof 6.6 A necessary condition for uniqueness, which had already been assumed, was that $k(\mathbf{M_m}) \geq 2, \forall m$. Since $r(\mathbf{M_m}) = k(\mathbf{M_m}), \forall m$, $r(\mathbf{M_j}) \geq 2$. Additionally, under these conditions, it was found in Theorem 6.2 that $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$. Therefore, let $r(\mathbf{M_j}) = 2 + \rho$, where $0 \leq \rho \leq R - 3$. The bounds for ρ were established by requiring $2 \leq r(\mathbf{M_j}) < R$. Since there could only be positive integer number of columns, R, $2 \leq r(\mathbf{M_j}) \leq R - 1$. Hence if $r(\mathbf{M_j}) = 2 + \rho$, $0 \leq \rho \leq R - 3$. Also, it was assumed that $r(\mathbf{M_i}) + r(\mathbf{M_j}) < R + 2$. Therefore, $r(\mathbf{M_i}) < R - \rho$. From Theorem 6.4, $\eta = r(\mathbf{M_i})r(\mathbf{M_j}) - R$. By substitution,

$$\begin{aligned} r(\mathbf{M_i})r(\mathbf{M_j}) - R &< (R - \rho)r(\mathbf{M_j}) - R \\ &= (R - \rho)(2 + \rho) - R \end{aligned}.$$

Consider that $(R - \rho)(2 + \rho)$ is a second degree polynomial. For any R, $\eta < \max((R - \rho)(2 + \rho))$. The maximum of this polynomial in ρ would be achieved at $\rho = R - 2$. However, ρ was constrained to be less than or equal to $R - 3$. Hence, with this constraint, the maximum was achieved at $\rho = R - 3$, and $\eta < 3(R-1) - R = 2R - 3$. □

Theorem 6.7 For $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$, if $r(\mathbf{M_i}) + r(\mathbf{M_j}) < R + 2$, then the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ was less than $2R - 3$.

Proof 6.7 The maximum rank for a loading matrix must be less than R by Theorem 6.2. By Theorem 6.4, it was established that the number of basis vectors in the null space of $(\mathbf{M_j} \circ \mathbf{M_i})^t$, or OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$, with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ was $\eta = r(\mathbf{M_i})r(\mathbf{M_j}) - R$. From Theorem 6.6, it was found that under these conditions $\eta < 2R - 3$. Since the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ was η, the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ was less than $2R - 3$. □

Theorem 6.8 For $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$ where the maximum rank of a loading matrix is less than R, the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ is less than $2R - 3$ if and only if $r(\mathbf{M_i}) + r(\mathbf{M_j}) < R + 2$.

Proof 6.8 Theorem 6.6 showed that if $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$, and $r(\mathbf{M_i}) + r(\mathbf{M_j}) < R + 2$, then $\eta < 2R - 3$, which established sufficiency. It was assumed that $\eta < 2R - 3$, which implied that $r(\mathbf{M_i})r(\mathbf{M_j}) - R < 2R - 3$ or that

$r(\mathbf{M_i})r(\mathbf{M_j}) < 3(R-1)$. Hence, $r(\mathbf{M_i}) < \frac{3(R-1)}{r(\mathbf{M_j})}$. By substitution, this implied that $r(\mathbf{M_i}) + r(\mathbf{M_j}) < \frac{3(R-1)}{r(\mathbf{M_j})} + r(\mathbf{M_j})$. For any R, the rhs of this inequality was a function of $r(\mathbf{M_j})$. The lhs of the inequality, $r(\mathbf{M_i}) + r(\mathbf{M_j})$, would always be smaller than the $\max_{r(\mathbf{M_j})}\left(\frac{3(R-1)}{r(\mathbf{M_j})} + r(\mathbf{M_j})\right)$. Since this function must bound the sum of the ranks of the two loading matrices, certain restrictions must be applied. The first is that the rank of $r(\mathbf{M_j})$ must not be R or greater, which is assumed in the statement of the theorem. The second is that the $\max_{r(\mathbf{M_j})}\left(\frac{3(R-1)}{r(\mathbf{M_j})} + r(\mathbf{M_j})\right)$ cannot exceed $r(\mathbf{M_j})$+R since the sum of the ranks of the two matrices must be less than R. Therefore, taking into consideration these limits, the maximum is achieved at two values of $r(\mathbf{M_j})$, R-2 and 3. Thus, substituting in for $\max_{r(\mathbf{M_j})}\left(\frac{3(R-1)}{r(\mathbf{M_j})} + r(\mathbf{M_j})\right)$ yields $r(\mathbf{M_i}) + r(\mathbf{M_j}) < R + 2$, which satisfies necessity. □

Theorem 6.9 For $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$, if $r(\mathbf{M_j}) \geq 2$ and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) = R$, then $\eta \geq R$.

Proof 6.9 From Theorem 6.4 it was found that the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ was $\eta = r(\mathbf{M_i})r(\mathbf{M_j}) - R$. Substitution of the values for $r(\mathbf{M_i}) = R$ and $r(\mathbf{M_j}) \geq 2$ resulted in $\eta \geq 2R - R = R$. □

Theorem 6.10 For $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$, if $2 \leq r(\mathbf{M_j}) < R$ and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) = R$, then the number of unit basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ is equal to $\left[\left(r(\mathbf{M_j})\right)\left(r(\mathbf{M_j}) - 1\right)\right]$.

Proof 6.10 By Theorem 6.9, in this case, the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ was greater than or equal to R. Also, unit vectors would occur whenever a row of $(\mathbf{M_j} \circ \mathbf{M_i})$, and so a column of $(\mathbf{M_j} \circ \mathbf{M_i})^{\mathbf{t}}$, contained all 0 elements. Therefore, for each column of $(\mathbf{M_j} \circ \mathbf{M_i})^{\mathbf{t}}$ comprised of all 0 elements, a unit vector would result. The KR product, $(\mathbf{M_j} \circ \mathbf{M_i})$, was composed of R columns and $r(\mathbf{M_j})r(\mathbf{M_i})$ rows. It was assumed that $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) = R$ and $r(\mathbf{M_j}) = \min(r(\mathbf{M_i}), r(\mathbf{M_j})) < R$. Hence, the first $r(\mathbf{M_j})$ columns of $(\mathbf{M_j} \circ \mathbf{M_i})$ would be the columnwise Kronecker product of columns from identity matrices, which would result in a column of only one nonzero entry, a 1. Therefore, the remaining $R - r(\mathbf{M_j})$ columns of $(\mathbf{M_j} \circ \mathbf{M_i})$ would be created from the columnwise Kronecker product of columns from $\mathbf{M_i}$ and $\mathbf{M_j}$. Since $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m$, the elements in the last $R - r(\mathbf{M_j})$ columns of $\mathbf{M_j}$ would all be constraint elements. Also, because $r(\mathbf{M_i}) = R$, $\mathbf{M_i} = \mathbf{I_{R \times R}}$, an identity matrix, and had no constraint elements. Hence, while the first $r(\mathbf{M_j})$ columns of $(\mathbf{M_j} \circ \mathbf{M_i})$ would have no constraint elements, the last $R - r(\mathbf{M_j})$ columns would hold all the $r(\mathbf{M_j})$ constraint elements of $\mathbf{M_j}$. In fact, due to the nature of columnwise Kronecker products, a constraint element from $\mathbf{M_j}$ would occur in rows at R intervals from the last constraint element of $\mathbf{M_j}$, starting with the $r(\mathbf{M_j}) + 1$ row and column of $(\mathbf{M_j} \circ \mathbf{M_i})$. Let m_r^i be an element on the diagonal in the r^{th} row of $\mathbf{M_j}$. Since $\mathbf{M_i} = \mathbf{I_{R \times R}}$, $m_r^i \neq 0$. To create the last $R - r(\mathbf{M_j})$ columns of $(\mathbf{M_j} \circ \mathbf{M_i})$, the last $R - r(\mathbf{M_j})$ columns of $\mathbf{M_i}$ and $\mathbf{M_j}$ would be used. Therefore, a constraint element would occur in the last $R - r(\mathbf{M_j})$ columns of $(\mathbf{M_j} \circ \mathbf{M_i})$ whenever $m_r^i, r = r(\mathbf{M_j}) + 1, \ldots R$, was aligned with the constraint elements in the last $R - r(\mathbf{M_j})$ columns of $\mathbf{M_j}$. Since $\mathbf{M_i} = \mathbf{I_{R \times R}}$ and the last $R - r(\mathbf{M_j})$ columns of $\mathbf{M_j}$ were all nonzero, each of the last $R - r(\mathbf{M_j})$ columns of $(\mathbf{M_j} \circ \mathbf{M_i})$ would contain exactly

$r(\mathbf{M_j})$ constraint elements. Also, because $\mathbf{M_i} = \mathbf{I_{R\times R}}$, the rows of $(\mathbf{M_j} \circ \mathbf{M_i})$ with these constraint elements would not overlap with rows containing other constraint elements Therefore, there would be $r(\mathbf{M_j})(R - r(\mathbf{M_j}))$ rows of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from $\mathbf{M_j}$. In addition, these rows with constraint elements would not overlap with any row created from the columnwise Kronecker product of columns from $\mathbf{M_i}$ and the identity partition of $\mathbf{M_j}$. Hence, there were $r(\mathbf{M_j})$ rows containing only one nonzero, non-constraint element that resulted from combining columns of identity matrices and $r(\mathbf{M_j})(R - r(\mathbf{M_j}))$ rows with constraint elements. In total, there were $r(\mathbf{M_j}) + r(\mathbf{M_j})(R - r(\mathbf{M_j}))$ rows that contained at least one nonzero element. Hence, the KR product, $(\mathbf{M_j} \circ \mathbf{M_i})$, had $r(\mathbf{M_i})r(\mathbf{M_j})$ rows, $r(\mathbf{M_j}) + r(\mathbf{M_j})(R - r(\mathbf{M_j}))$ with at least one nonzero element. Therefore, $r(\mathbf{M_i})r(\mathbf{M_j}) - r(\mathbf{M_j}) + r(\mathbf{M_j})(R - r(\mathbf{M_j})) = r(\mathbf{M_j})(r(\mathbf{M_j}) - 1)$ rows had all zero elements. Hence, the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ would have $\left[\left(r(\mathbf{M_j})\right)\left(r(\mathbf{M_j}) - 1\right)\right]$ unit basis vectors. □

Theorem 6.11 For $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$, if $2 \le r(\mathbf{M_j}) < R$ and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) = R$, then the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from $\mathbf{M_i}$ or $\mathbf{M_j}$ was $\left[\left(R - r(\mathbf{M_j})\right)\left(r(\mathbf{M_j}) - 1\right)\right]$.

Proof 6.11 By Theorem 6.10**,** there are $\left[\left(r(\mathbf{M_j})\right)\left(r(\mathbf{M_j}) - 1\right)\right]$ unit basis vectors in the null space of $(\mathbf{M_j} \circ \mathbf{M_i})^{\mathbf{t}}$. From linear algebra, the total number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ would be $r(\mathbf{M_i})r(\mathbf{M_j}) - R$. If a basis vector was not a unit vector, then it would contain at least two nonzero elements. These elements would have resulted from the Khatri-Rao product of $\mathbf{M_i}$ or $\mathbf{M_j}$ and would be used to impose constraints on loading matrices from alternative

PARAFAC solutions. Therefore, if a basis vector was not a unit vector, then it was a vector with constraint elements from $\mathbf{M_i}$ or $\mathbf{M_j}$. Hence, the number of basis vectors in the null space of $(\mathbf{M_j} \circ \mathbf{M_i})^t$ with constraint elements was $\left(r(\mathbf{M_i})r(\mathbf{M_j})-R\right)-\left(r(\mathbf{M_j})\right)\left(r(\mathbf{M_j})-1\right)$. By substitution of $r(\mathbf{M_i})=R$, the number of basis with constraint elements was $\left(R-r(\mathbf{M_j})\right)\left(r(\mathbf{M_j})-1\right)$. Since Hence, $\left(R-r(\mathbf{M_j})\right)\left(r(\mathbf{M_j})-1\right)$ of the basis vectors have constraint elements from $\mathbf{M_i}$ or $\mathbf{M_j}$. □

Theorem 6.12 For $r(\mathbf{M_m})=k(\mathbf{M_m})$, $\forall m, m=1,2,3$, and $r(\mathbf{M_i})=\max(r(\mathbf{M_i}),r(\mathbf{M_j}),r(\mathbf{M_k}))<R$, if $r(\mathbf{M_i})+r(\mathbf{M_j})\geq R+2$, then $\eta \geq 2R-3$.

Proof 6.12 It was assumed that $r(\mathbf{M_i})<R$ and $r(\mathbf{M_i})+r(\mathbf{M_j})\geq R+2$. Therefore, $r(\mathbf{M_i})\geq R+2-r(\mathbf{M_j})$. Hence, by substitution, $r(\mathbf{M_i})r(\mathbf{M_j})\geq r(\mathbf{M_j})(R+2-r(\mathbf{M_j}))$. The rhs of the inequality was a function of $r(\mathbf{M_j})$, where $2\leq r(\mathbf{M_j})\leq R-1$. Therefore, $r(\mathbf{M_i})r(\mathbf{M_j})$ would always be greater than or equal to the $\min\left(r(\mathbf{M_j})(R+2-r(\mathbf{M_j})\right)$, which was achieved at R+2. However, $2\leq r(\mathbf{M_j})\leq R-1$. Thus, with the constraint of $2\leq r(\mathbf{M_j})\leq R-1$, the minimum was achieved at $R-1$ such that $r(\mathbf{M_j})(R+2-r(\mathbf{M_j})\geq 3(R-1)$. Hence, $\eta = r(\mathbf{M_i})r(\mathbf{M_j})-R\geq 2R-3$. □

Theorem 6.13 For $r(\mathbf{M_m})=k(\mathbf{M_m})$, $\forall m, m=1,2,3$, and $r(\mathbf{M_i})=\max(r(\mathbf{M_i}),r(\mathbf{M_j}),r(\mathbf{M_k}))<R$, if $r(\mathbf{M_i})+r(\mathbf{M_j})\geq R+2$, then the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ is greater than or equal to $2R-3$.

Proof 6.13 By Theorem 6.4, it was established that the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ was $\eta = r(\mathbf{M_i})r(\mathbf{M_j}) - R$. From Theorem 6.12, it was found that under these conditions $\eta \geq 2R - 3$. Since the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ was η, the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ was greater than or equal to $2R - 3$. □

Theorem 6.14 For $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$ and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$, the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from either matrix $\mathbf{M_i}$ or $\mathbf{M_j}$ is greater than or equal to $2R - 3$ if and only if $r(\mathbf{M_i}) + r(\mathbf{M_j}) \geq R + 2$.

Proof 6.14 From Theorem 6.12, it was found that if $r(\mathbf{M_i}) + r(\mathbf{M_j}) \geq R + 2$ then $\eta \geq 2R - 3$, which established sufficiency. Therefore, to establish necessity, it was assumed that $\eta \geq 2R - 3$ or that $r(\mathbf{M_i})r(\mathbf{M_j}) - R \geq 2R - 3$. Thus, $r(\mathbf{M_i}) \geq \frac{3(R-1)}{r(\mathbf{M_j})}$. By substitution, this implied that $r(\mathbf{M_i}) + r(\mathbf{M_j}) \geq \frac{3(R-1)}{r(\mathbf{M_j})} + r(\mathbf{M_j})$, which was a function of $r(\mathbf{M_j})$. The minimum of $\frac{3(R-1)}{r(\mathbf{M_j})} + r(\mathbf{M_j})$ was $R + 2$. Hence, $r(\mathbf{M_i}) + r(\mathbf{M_j}) \geq R + 2$, which supplied necessity. □

6.4 Theorems and Uniqueness

Using the theorems presented above, it was possible to form theoretical results regarding PARAFAC decomposition uniqueness and OC space basis vectors. First, for PARAFAC solutions where k-rank = rank, the number of OC space basis vectors with constraint elements was used to identify when KR products would have alternatives that were or were not limited to permutation-scale transformations.

The theoretical results, presented when (1) the pairwise sum of the ranks is at least R+2 and the maximum rank is less than R and (2) the pairwise sum of the ranks is less than R+2, are used, with the definition of uniqueness, to show that the additional condition of k-rank = rank was not able to limit the set of PARAFAC solutions so that Kruskal's condition would be necessary and sufficient.

Theorem 6.15 For $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$ and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$, if the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from $\mathbf{M_i}$ or $\mathbf{M_j}$ was greater than or equal to $2R-3$, then no alternative KR basis for the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ that was composed of non-permutation-scale versions of $\mathbf{M_i}$ and $\mathbf{M_j}$ existed.

Proof 6.15 Under the conditions of equivalent rank and k-rank and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$, in Theorem 6.5 it was shown that the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from $\mathbf{M_i}$ or $\mathbf{M_j}$ was equivalent to $\eta = r(\mathbf{M_i})r(\mathbf{M_j}) - R$; and, under those same conditions, in Theorem 6.12**,** η was found to be greater than or equal to $2R-3$. By Theorem 6.14 it was shown that $\eta \geq 2R-3$ if and only if $r(\mathbf{M_i}) + r(\mathbf{M_j}) \geq R+2$ when $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m$ and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$. Therefore, the ranks were equivalent to the k-ranks and $r(\mathbf{M_i}) + r(\mathbf{M_j}) \geq R+2$, which implied that $k(\mathbf{M_i}) + k(\mathbf{M_j}) \geq R+2$. By applying Kruskal's condition (Theorem 2.2)**,** no alternative KR basis for the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ that was composed of non-permutation-scale versions of $\mathbf{M_i}$ and $\mathbf{M_j}$ existed when $k(\mathbf{M_i}) + k(\mathbf{M_j}) \geq R+2$. Thus, when $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m$ and $r(\mathbf{M_i}) + r(\mathbf{M_j}) \geq R+2$, no alternative KR basis for the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ that was composed of non-permutation-scale versions of $\mathbf{M_i}$ and $\mathbf{M_j}$ existed.

Finally, this implied that when $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$ and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$, if the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from $\mathbf{M_i}$ or $\mathbf{M_j}$ was greater than or equal to $2R - 3$, then no alternative KR basis for the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ that was composed of non-permutation-scale versions of $\mathbf{M_i}$ and $\mathbf{M_j}$ existed. □

Theorem 6.16 For $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$ and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$, the PARAFAC solution would be unique if all of the basis vectors from the OCs to the column spaces of all of the resulting KR product permutations had $2R - 3$ or more vectors with constraint elements from either of the two matrices in the KR product.

Proof 6.16 From Theorem 6.15**,** it was established that, under the conditions of $r(\mathbf{M_m}) = k(\mathbf{M_m})$, $\forall m, m = 1,2,3$ and $r(\mathbf{M_i}) = \max(r(\mathbf{M_i}), r(\mathbf{M_j}), r(\mathbf{M_k})) < R$, if the number of basis vectors in the OC to the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ with constraint elements from $\mathbf{M_i}$ or was greater than or equal to $2R - 3$, then no alternative KR basis, composed of non-permutation-scale versions of $\mathbf{M_i}$ and $\mathbf{M_j}$,for the column space of $(\mathbf{M_j} \circ \mathbf{M_i})$ existed. Therefore, if every KR product permutation of $\mathbf{M_i}$, $\mathbf{M_j}$, and $\mathbf{M_k}$ resulted in a transpose such that the OC to the column space had $2R - 3$ or more basis vectors with constraint elements, no alternative KR basis, composed of non-permutation-scale versions of the loading matrices, for the column space of any of the KR products would exist. By the definition of uniqueness, a PARAFAC solution was unique if all KR product permutations of loading matrices, had no alternative KR basis, composed of non-permutation-scale versions of the loading matrices, for the column space of the KR product existed. Hence, a PARAFAC solution where all KR product permutations of the loading matrices had transposes such that the resulting OC spaces consisted of $2R-3$ or more basis vectors with constraint elements, would be unique. □

Theorem 6.17 For $r(\mathbf{M_m})=k(\mathbf{M_m})$, $\forall m, m=1,2,3$, Kruskal's condition is not necessary and sufficient.

Proof 6.17 Suppose that Kruskal's condition was necessary when $r(\mathbf{M_m})=k(\mathbf{M_m})$, $\forall m, m=1,2,3$. For $r(\mathbf{M_m})=k(\mathbf{M_m})$, $\forall m$, the necessity of the condition implied that PARAFAC uniqueness would occur only if Kruskal's condition was met or only when $k(\mathbf{M_i})+k(\mathbf{M_j})+k(\mathbf{M_k})\geq 2R+2$. The necessity of Kruskal's condition when rank was equal to k-rank could only be denied if uniqueness could be established for solutions with $r(\mathbf{M_m})=k(\mathbf{M_m})$, $\forall m$, and $k(\mathbf{M_i})+k(\mathbf{M_j})+k(\mathbf{M_k})<2R+2$. From Theorem 6.16, PARAFAC solutions are unique, when k-ranks and ranks are equivalent, if $r(\mathbf{M_i})=\max(r(\mathbf{M_i}),r(\mathbf{M_j}),r(\mathbf{M_k}))<R$ and all of the basis vectors from the Null spaces of all the transposes of the resulting KR product permutations had $2R-3$ or more vectors with constraint elements. Theorem 6.14 showed that requiring $2R-3$ or more basis vectors to have constraint elements was equivalent to requiring $r(\mathbf{M_i})+r(\mathbf{M_j})\geq R+2$. Since each KR product permutation had a transpose where the Null space had $2R-3$ or more basis vectors with constraint elements, $r(\mathbf{M_i})+r(\mathbf{M_j})\geq R+2$, $\forall i\neq j, i=1,2,3, j=1,2,3$. It was assumed that $r(\mathbf{M_1})+r(\mathbf{M_2})+r(\mathbf{M_3})<2R+2$ and $r(\mathbf{M_m})=k(\mathbf{M_m})$, $\forall m$. Thus, the only remaining loading was to show that there existed loading matrices such that each of the pairwise sums of ranks was at least R+2 and $r(\mathbf{M_1})+r(\mathbf{M_2})+r(\mathbf{M_3})<2R+2$.

Suppose that $r(\mathbf{M_2})+r(\mathbf{M_3})=R+2$ and $r(\mathbf{M_1})=\max(r(\mathbf{M_1}),r(\mathbf{M_2}),r(\mathbf{M_3}))<R$. Therefore, $r(\mathbf{M_2})=R+2-r(\mathbf{M_3})$. Adding $r(\mathbf{M_1})$ to each side of the equation preserved equality so that $r(\mathbf{M_1})+r(\mathbf{M_2})=r(\mathbf{M_1})+R+2-r(\mathbf{M_3})$. It was also assumed that $r(\mathbf{M_1})=\max(r(\mathbf{M_1}),r(\mathbf{M_2}),r(\mathbf{M_3}))<R$ so that $r(\mathbf{M_1})\geq r(\mathbf{M_3})$. Hence,

$r(\mathbf{M_1}) - r(\mathbf{M_3}) \geq 0$ and $r(\mathbf{M_1}) + r(\mathbf{M_2}) = r(\mathbf{M_1}) + R + 2 - r(\mathbf{M_3}) \geq R + 2$. Similarly, $r(\mathbf{M_1}) \geq r(\mathbf{M_2})$ and $r(\mathbf{M_1}) + r(\mathbf{M_3}) = r(\mathbf{M_1}) + R + 2 - r(\mathbf{M_2}) \geq R + 2$. Thus, $r(\mathbf{M_i}) + r(\mathbf{M_j}) \geq R + 2$, $\forall i \neq j,\ i = 1,2,3,\ j = 1,2,3$. Further, since it was assumed that $r(\mathbf{M_1}) = \max(r(\mathbf{M_1}), r(\mathbf{M_2}), r(\mathbf{M_3})) < R$, $R - r(\mathbf{M_1}) > 0$, which could be rewritten as $R - r(\mathbf{M_1}) > R + 2 - (R + 2)$. Since $r(\mathbf{M_2}) + r(\mathbf{M_3}) = R + 2$, $R - r(\mathbf{M_1}) > r(\mathbf{M_2}) + r(\mathbf{M_3}) - (R + 2)$. Rearranging the elements in the inequality resulted in $r(\mathbf{M_1}) + r(\mathbf{M_2}) + r(\mathbf{M_3}) < 2R + 2$.

Therefore, uniqueness had been established when rank and k-rank were equivalent and $k(\mathbf{M_1}) + k(\mathbf{M_2}) + k(\mathbf{M_3}) < 2R + 2$, or when Kruskal's condition was not met. Hence, uniqueness in the absence of Kruskal's condition implied that Kruskal's sufficient condition was not necessary. □

Copyright © Heather Michele Clyburn Bush 2006

7. DISCUSSION AND FUTURE RESEARCH

Necessary and sufficient conditions for the uniqueness of PARAFAC solutions had long been sought ever since Harshman introduced the idea of parallel proportional profiles. Kruskal's seminal work in the area provided sufficient conditions that quieted the pursuit of conditions for almost three decades. The search, however, was reopened when Bro introduced the idea of KR products (1998) and Liu and Sidiropoulos (2001) offered necessary conditions for uniqueness with KR products. However, it was the use of KR products and simplified forms, employed by ten Berge and Sidiropoulos, that provided the groundwork for truly investigating Kruskal's condition and the somewhat abstract concept of k-rank.

Although ten Berge and Sidiropoulos offered a method that proved adequate in refuting the idea that Kruskal's k-rank sum condition was necessary as well as sufficient, their approach was not entirely successful in providing additional conditions so that Kruskal's k-rank condition would be necessary and sufficient. Additionally, and more importantly, the investigation of column spaces failed to provide answers as to why some PARAFAC solutions were unique and others were not. Especially troublesome to ten Berge and Sidiropoulos was that k-rank appeared to be lacking in providing answers to the question.

OC spaces, however, were able to illuminate the differences that existed between PARAFAC solutions where the loading matrices had the same k-rank and rank but different uniqueness conclusions. In addition, exploring the constraints that arose from the OC spaces offered necessary and sufficient conditions for PARAFAC solution uniqueness.

The use of OC spaces laid a foundation so that the examination of PARAFAC decompositions could be reduced to evaluating the pairwise combinations of loading matrices and provided a simplified method for evaluating uniqueness. The limitation of the OCSA and the tBSM came in evaluating PARAFAC solutions with a large number of factors, or large R and k-rank < rank. However, whereas the tBSM and the direct use of column spaces did not offer a methodical approach to the problem, the OCSA provided a straightforward and systematic approach, needing only a more computational strategy for evaluating k-rank. The difficulty with large R and deficient k-rank is rooted in the many combinations of RREFs that result from having k-rank < rank. Although the number of these combinations was not insurmountably large, evaluating the many alternative possibilities for each of these combinations became increasingly complex.

The greatest factor contributing to the complexity was maintaining the matrix properties of alternative loading matrices. As experienced by ten Berge and Sidiropoulos, results could be misleading if the alternative matrices did not have the appropriate matrix properties. Identifying whether or not an alternative loading matrix had the appropriate k-rank became an overly difficult task. In order to utilize OC spaces for the evaluation of uniqueness for larger R and k-rank < rank, it would be necessary to computationally evaluate alternative solutions for k-rank.

To date, there are no mechanisms available to weed out alternative loading matrices with inappropriate k-rank. Future research in the uniqueness of PARAFAC decompositions will need to incorporate a computational tool to evaluate k-rank for alternative loading matrices in RREF. Additionally, the tool could be employed to produce the various RREF for loading matrices with k-rank<rank as well as selecting appropriate alternative loading matrices for evaluation. Thus, one of the remaining roadblocks to uniqueness conditions would be to develop a strategy for examining the many alternative possibilities for large R and k-rank < rank.

Many of the KR products composed of matrices with small ranks were found to have non-permutation-scale alternatives, while most of the cases with large ranks were limited to permutation-scale transformations. Necessary and sufficient conditions for all R appear to be rooted in the evaluation of the OC spaces. Although the mathematical principles needed to find and prove the necessary and sufficient conditions for uniqueness are well outside the scope of this dissertation, the investigation of specific cases suggests that OC spaces could finally provide an avenue for finding these conditions.

Copyright © Heather Michele Clyburn Bush 2006

8. BIBLIOGRAPHY

Appellof, C.J., and Davidson, E.R. 1981. Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents. *Analytical Chemistry* 53: 2053-2056.

Bro, R. 1997. PARAFAC: Tutorial & applications. *Chemom. Intell. Lab. Syst*. 38: 149-171.

Bro, R. 1998. Multi-way Analysis in the Food Industry. Models, algorithms and Applications, Ph.D. Thesis, University of Amsterdam, The Netherlands.

Bro, R. 1999. Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis. *Chemometrics and Intelligent Laboratory Systems* 46: 133-147.

Bro, R., and De Jong, S. 1997. A fast non-negativity-constrained least squares algorithm. *Journal of Chemometrics* 11: 393-401.

Bro, R., and Andersson, C.A. 1998. Improving the speed of multiway algorithms. Part II: Compression. *Chemometrics and Intelligent Laboratory Systems* 42: 105-113.

Bro, R., and Heimdal, H. 1996. Enzymatic browning of vegetables. Calibration and analysis of variance by multiway methods. *Chemometrics and Intelligent Laboratory Systems* 34: 85-102

Burdick, D. 1995. An introduction to tensor products with applications to multiway data analysis. *Chemom. Intell. Lab. Syst*. 28: 229-237.

Carroll, J. D., and 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, R. A. and Mackenzie, W. A. (1923) Studies in crop variation II. The manurial treatment of different potato varieties. *Journal of Agricultural Science* 13, 311-320.

Gantmacher, F.R., *The Theory of Matrices*, *Vols. I,II,* Chelsea Publishing Company, New York, 1959.

Harshman, R.A. 1970. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multinomial factor analysis. *UCLA Working Papers in Phonetics* 16: 1-84.

Harshman, R.A. 1972. Determination and proof of minimum uniqueness conditions for PARAFAC1. *UCLA Working Papers in Phonetics* 22: 111-117.

Harshman, R.A., and Lundy, M.E. 1984a. The PARAFAC model for three-way factor analysis and multidimensional scaling. In *Research Methods for Multimode Data Analysis* (H.G. Law, C.W. Snyder Jr., J.A. Hattie and R.P. Mcdonald, eds.) Praeger, New York.

Harshman, R.A., and Lundy, M.E. 1984b. Data preprocessing and the extended PARAFAC model. In *Research Methods for Multimode Data Analysis* (H.G. Law, C.W. Snyder Jr., J.A. Hattie and R.P. Mcdonald, eds.) Praeger, New York.

Ho, C.N., Christian, G.D., and Davidson, E.R. 1978. Application of the method of rank annihilation to quantitative analyses of multicomponent fluorescence data from the video fluorometer. *Analytical Chemistry* 50: 1108-1113.

Kiers, H.A.L. 2000. Towards a standardized notation and terminology in multiway analysis. *Journal of Chemnometrics* 14:105-122.

Krijnen, W.P. 1993 *The analysis of three-way arrays by constrained PARAFAC methods.* Leiden, Netherlands: DSWO Press.

Kroonenberg, P.M. 1983. *Three-Mode Principal Component Analysis*. DSWO Press, Leiden, The Netherlands.

Kroonenberg, P.M. 1989. Singular value decompositions of interactions in three-way contingency tables. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.), North-Holland.

Kruskal, J.B. 1977. Three-way arrays: Rank and uniqueness of trilinear decompositions with application to arithmetic complexity and statistics. *Linear Algebra Appl.* 18: 95-138.

Kruskal, J.B. 1984. Multilinear methods. In *Research Methods for Multimode Data Analysis*. (H.G. Law, C.W. Snyder, Jr., J.A. Hattie and R.P. McDonald, eds.), Praeger, New York.

Kruskal, J.B. 1989. Rank decomposition and uniqueness for 3-way and N-way arrays. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.), North-Holland.

Leurgans, S. and Ross, R. 1992. Multilinear Models: Applications in spectroscopy. *Statistical Science* 7(3): 289-319.

Leurgans, S., Ross R., Abel R. 1993 A decomposition for three-way arrays. *SIAM Journal of Matrix Analysis and Application* 14:1064-1083.

Liu, X. and Sidiropoulos, N.D. 2001. Cramer-Rao lower bounds for low-rank decomposition of multidimensional arrays. *IEEE Transactions on Signal Processing*, 49, 2074-2086.

McDonald, R.P., A simple comprehensive model for the analysis of covariance structures: Some remarks on applications, *British Journal of Mathematical and Statistical Psychology*, **33** (1980) 161.

Martin, A., Wiggs, C. L., Ungerleider, L. G., and Haxby, J. V. 1996. Neural correlates of category-specific knowledge. *Nature* 379: 649-652.

Mitchell, B. and Burdick, D. 1994. Slowly converging PARAFAC sequences: swamps and two factor degeneracies. *Journal of Chemometrics* 8: 155-168.

Rao, C.R, Mitra, S., Generalized inverse of matrices and its applications, Wiley, New York, 1971.

Rayens, W.S. and Mitchell, B. 1997. Two-factor degeneracies and a stabilization of PARAFAC. *Chemometrics and Intelligent Laboratory Systems* 38: 173-181.

Sanchez, E., and Kowalski, B.R. 1988. Tensorial calibration: II. Second-order calibration. *Journal of Chemometrics* 2: 265-280.

Sanchez, E., and Kowalski, B.R. 1990. Tensorial resolution: A direct trilinear decomposition. *Journal of Chemometrics* 4: 29-45.

Sidiropoulos, N.D. and Bro, R. 2000. On the uniqueness of multilinear decomposition of *N*-way arrays. *Journal of Chemometrics* 14: 229-239.

Smilde, A.K., Bro, R., and Geladi, P. 2004. *Multi-way Analysis: Applications in the Chemical Sciences*. John Wiley & Sons, Ltd. West Sussex, England.

Smilde, A.K. and Doornbos, D.A. 1992. Simple validatory tools for judging the predictive performance of PARAFAC and three-way PLS. *Journal of Chemometrics* 6: 11-28.

Smilde, A.K 1992. Three-way analysis. Problems and prospects. *Chemom. Intell. Lab. Syst.* 15: 143-157.

ten Berge, J.M. and Sidiropoulos, N.D. 2002. On uniqueness in CANDECOMP/PARAFAC. *Journal of Psychometrika* 67: 399-409.

Tucker, L.R. 1966. Some mathematical notes on three-mode factor analysis. *Psychometrika* 31: 279-311.

Wold, S., Geladi, P., Esbensen, K., and Ohman, J. 1987. Multi-way principal components- and PLS-analysis. *Journal of Chemometrics* 1: 41-56.

Copyright © Heather Michele Clyburn Bush 2006

9. VITA

DATE AND PLACE OF BIRTH	March 16, 1978	Murfreesboro, TN

EDUCATION

2006	Ph.D. Candidate in Statistics, University of Kentucky
2001	M.S. in Statistics University of Kentucky
1999	B.S. in Mathematics Mount Vernon Nazarene University Summa Cum Laude Valedictorian

PROFESSIONAL EXPERIENCE

03/04-present	Statistician REGISTRAT, Inc.
08/02-06/05	*Research Assistant* University of Kentucky, School of Public Health
06/00-08/02	*Teaching Assistant /Instructor* University of Kentucky, Statistics
08/99-05/00	*Teaching Assistant* University of Kentucky, Statistics

PROFESSIONAL SERVICE

Graduate Student Representative, Department of Statistics (08/00—06/05)
University of Kentucky New Teaching Assistant Orientation, Panel Member & Microteach Leader (2002)

PRESENTATIONS

"Introduction to Biostatistics". *Biomedical Research Undergraduate Training (BRUT) program*, Transylvania University, July 2004.

"Patterns of Cervical Cancer Occurrence Among Appalachian Subgroups," with Claudia M. Hopenhayn, Amy Christian, Bin Huang. 18th National Conference on Chronic Disease Prevention and Control. Washington, D.C. February 2004. Hopenhayn presenting.

"Issues of Existence and Uniqueness in $I \times J \times 2$ Arrays," Proposal Presentation at the University of Kentucky, Department of Statistics. October 2003.

"Does Arsenic Exposure Increase the Risk of Anemia During Pregnancy," with Claudia M. Hopenhayn and Gilbert A. Boissonneault. International Society for Environmental Epidemiology 15th Annual Conference. Perth, Australia. September 2003. Hopenhayn presenting.

PAPERS AND REPORTS

Hopenhayn, C., Bush, H.M., G. Boisonneault, Peralta C, Bingcang, A, I. Hertz-Picciotto "Does Arsenic Exposure Increase Maternal Anemia During Pregnancy?". Accepted Journal of Occupational and Environmental Medicine.

Hopenhayn, C., Bush, H.M., Christian, A., Shelton, B. "Comparative analysis of invasive cervical cancer incidence rates in three Appalachian states". Article in Press. Preventive Medicine.

Christian, A., Knight, E.A., Bush, H.M., and Anochie, L. (2005) "The Impact of Diabetes on the Commonwealth of Kentucky 2004". Kentucky Diabetes Prevention and Control Program, Department for Public Health

ABSTRACTS

Langberg, J.L, Smith, A., Heeke, S., Cooper, P., Bush, H.M., Brosius, L. (2004) "Tall Stature Increases the Risk of Atrial Fibrillation (AF) in Patients with Impaired Left Ventricular Function (LVD)". *AHA Scientific Sessions,* 04-SS-A-6965-AHA. Abstract

MEMBERSHIP IN PROFESSIONAL SOCIETIES

American Statistical Association
Institute of Mathematical Statistics

HONORS AND AWARDS

R.L. Anderson Award for Outstanding Teaching (2001)
Alpha Chi—National Honor Society—Undergraduate Inductee
Phi Delta Lambda—Nazarene Honor Society—Undergraduate Inductee
Outstanding Freshman Award—Undergraduate Award

CONFERENCES ATTENDED

CBI: Phase IV and Patient Registries Conferences, Philadelphia, PA, Sept. 2005.
Midwest Biopharmaceutical Statistics Workshop, Ball State University, May 2005.
Trilinear Methods in Chemistry and Psychology (by invitation), University of Kentucky, June 2003.
Midwest Biopharmaceutical Statistics Workshop, Ball State University, May 2003.
Statistical Methodology in Alzheimer's Disease Research II, University of Kentucky, May 2002.
Nonlinear Models Conference, University of Kentucky, December 1999.

Copyright © Heather Michele Clyburn Bush 2006