MULTIWAY DATA ANALYSIS
R. Coppi and S. Bolasco (Editors)
Elsevier Science Publishers B.V. (North-Holland), 1989

RANK, DECOMPOSITION, AND UNIQUENESS FOR 3-WAY AND *N*-WAY ARRAYS

J. B. Kruskal
AT&T Bell Laboratories
Murray Hill, NJ 07974-2070

Decomposition of a matrix as a sum or linear combination of outer product matrices underlies both the bilinear methods (factor analysis, principal components analysis, and correspondence analysis) and the fundamental concept of matrix rank. The decomposition of a 3-way array in the corresponding manner underlies PARAFAC, and can be used to define rank of 3-way arrays. In fact, decomposition and rank can be generalized to N-way arrays. The generalizations are so natural and mathematically appealing that they have been independently introduced several times in the mathematical literature. The many differences between 3-way arrays and 2-way arrays with respect to decomposition and rank are discussed.

One difference concerns *rotational uniqueness* of decompositions. For 2-way arrays, this holds only in trivial cases, but for 3-way arrays, it holds for many decompositions of interest, including most PARAFAC solutions. This is why PARAFAC solutions generally do not face the rotation problem, even though factor analysis solutions do. Many people consider the rotational uniqueness of PARAFAC solutions to be a major advantage of this model.

This paper also introduces the dimensionality vector of N-way arrays, which is closely connected to the number of factors used in 3-mode factor analysis. There is a final section explaining how various concepts are related to data analysis models.

1. INTRODUCTION

Rank is a fundamental property of matrices. For those interested in three-way arrays, it is natural to ask how rank generalizes to three-way and N-way arrays. This paper discusses a generalization that was introduced in (Kruskal, 1976) and further discussed in (Kruskal, 1977). This generalization, which is based on the decomposition of the array into outer product arrays, is so natural and mathematically appealing that it was independently introduced at least three times before 1976 in the mathematical literature, most recently in computational complexity theory of bilinear operations (see papers in the 1970's and 1980's by authors such as V. Strassen, V. Ya. Pan, and J. Ja' Ja'). It also discusses the uniqueness of the associated decompositions of 3-way arrays, based on theorems from (Kruskal, 1977) and extensive experimental calculations by Harshman and Lundy (unpublished). After that it introduces another concept that is much simpler than rank, the dimensionality vector of an array.

Rank, uniqueness theorems, and dimensionality vectors all turn out to be useful to the data analysis of three-way arrays. The final section explains the relevance and importance of these concepts to several models used in this area.

2. RANK AND N-ADIC DECOMPOSITIONS

One of the most fundamental concepts of matrix theory is the rank of a matrix. Although many parts of matrix theory have long ago been generalized to 3-way and N-way arrays in connection with tensor theory and Grassman algebras, the concept of rank was not included in these developments. For anyone interested in 3-way arrays, it is natural to ask how the concept of rank can be generalized to 3-way and N-way arrays.

An N-array means an N-way array. The value of N is the *order* of the array. Clearly

vector = 1-array = array of order 1
matrix = 2-array = array of order 2.

The *outer product* of N vectors a, b, c, ... with elements a_i, b_j, c_k, ... is the array $a \otimes b \otimes c \cdots$ whose elements are $a_i b_j c_k \cdots$. Note that N can be 2, 3, 4, ... or even 1. Define an *outer product array* to be any array which can be expressed as an outer product. Every 1-array is an outer product array; a 2-array is an outer product array if and only if it is a matrix of rank ≤ 1. The word *dyad* is widely used for an outer product matrix. In addition, *triad* is used here for an outer product 3-array, and *N-ad* for an outer product N-array.

An *N-adic decomposition* of an array X is a sum or a linear combination of N-ads which is equal to X:

$$X = \sum_{r=1}^{R} a_r \otimes b_r \otimes c_r \cdots \tag{1}$$

The *rank or the number of factors of the decomposition* is the number R of terms in it. A sum can be converted to a linear combination by inserting coefficients of 1, and a linear combination can be converted to a sum by absorbing the coefficients in one of the factors, so the ambiguity in the definition makes no difference. A rank R dyadic decomposition of a matrix X is easily seen to be the same thing as a factorization of X into a product ab' of two matrices, where a and b each have R columns.

The *rank of an* N-array X is the smallest number R such that X has an N-adic decomposition of rank R.

For order 2, this definition reduces directly to one of the many classical definitions of matrix rank, namely, the rank of a matrix X is equal to the smallest number R such that X has a dyadic decomposition of rank R. Every non-zero 1-array has rank 1. For any N, an N-array has rank 0 if and only if the array is 0, and has rank 1 if and only if the array is a non-zero N-ad.

The rank of a 3-array is the smallest number of triads needed to decompose it. To illustrate the meaning of this definition, consider the simplest case: $2 \times 2 \times 2$ arrays. The maximum possible rank is 3, though proving this fact would require several pages. The four ranks are illustrated.

Rank is the *same* concept for arrays of all orders, not a series of separate concepts for matrices, 3-arrays, 4-arrays, etc. The meaning of this statement is illustrated by an example. Consider a 2-array X that is $I \times J$. Define a 3-array Y that is $1 \times I \times J$ whose values are just those of X, and define a 4-array Z that is $1 \times 1 \times I \times J$ whose values are also just those of X. In formal mathematical terms X, Y, and Z are all different, since they have orders 2, 3, and 4 respectively, so the ranks of X, Y, and Z could all be different without damage to the meaning of rank. In fact, however, arrays of different formal order that have identical elements can be proved to have the same rank. Thus rank is a single concept whose value depends on the entries of the array but not

rank 0				rank 1			
		0	0			10	20
0	0	0	0	1	2	30	60
0	0			3	6		
rank 2				**rank 3**			
		0	1			0	1
1	0	1	0	−1	0	1	0
0	1			0	1		

on its order.

3. PROPERTIES OF RANK AND *N*-ADIC DECOMPOSITION

Arrays of all orders behave just like matrices with respect to many simple properties of rank. For example, regardless of order, $\text{rank}(X)=\text{rank}(kX)$ for any scalar $k \neq 0$, and $\text{rank}(X+Y) \leq \text{rank}(X) + \text{rank}(Y)$. Also, 3-arrays behave like matrices with respect to some intermediate properties of rank. For example, a famous lower-bound theorem of Frobenius (see, e.g. (Mirsky, 1955),)

$$\text{rank}(X) \geq \text{rank}(UX) + \text{rank}(XW) - \text{rank}(UXW) \quad \text{for all size-compatible matrices } X,\ U,\ V \tag{2}$$

can be generalized to 3-arrays (see Kruskal, 1977) and probably to arrays of all orders. However, with respect to some deeper properties of rank, arrays of order 3 behave very differently than arrays of order 2.

Difference 1: There are straight-forward algorithms to compute the rank of a matrix, but there is no known algorithm for computing the rank of a 3-array.

Difference 2: It is fairly easy to determine the rank of a matrix, but it is usually extremely difficult to determine the rank of a 3-array, even one that looks regular. There is a much-studied $9 \times 9 \times 9$ array whose rank has been bounded between 18 and 23 but is still unknown.

Difference 3: The rank of a matrix of real numbers does not change if it is considered to be a complex matrix that just happens to contain entries with no imaginary part. However, the rank of a 3-array of real numbers does change if it is considered to be a complex array. For example, the last $2 \times 2 \times 2$ array shown above has rank 2, not 3, when considered as a complex array. In general, the rank of a 3-array depends on the base field being used, i.e., the field from which the elements are drawn when forming triads. In this paper, the base field is always the field of real numbers.

Difference 4: Let $R_{\max}(I, J)$ and $R_{\max}(I, J, K)$ be the maximum possible rank of $I \times J$ and $I \times J \times K$ arrays respectively. It is well-known that $R_{\max}(I, J) = \min(I, J)$, but $R_{\max}(I, J, K)$ is unknown and difficult to determine. The following weak inequalities are known:

$$\max(I, J, K) \le R_{\max}(I, J, K) \le \min(IJ, IK, JK).$$

One significant special case has been determined (Ja' Ja', 1979; Kruskal, future), namely,

$$R_{\max}(2, J, K) = m + \min(m, n) \quad \text{where}$$
$$m = \min(J, K) \quad \text{and}$$
$$n = \lfloor \max(J, K)/2 \rfloor,$$

where $\lfloor x \rfloor$ means the greatest integer $\le x$. For $J = K$, this reduces to $\lfloor 3J/2 \rfloor$. The proofs by Ja' Ja' and by Kruskal, which are entirely different, are both long and complicated. It has also been proved (Kruskal, future) that $R_{\max}(3, 3, 3) = 5$. Even this proof is long.

Difference 5: This difference comes as a real surprise. In the IJ-dimensional space of $I \times J$ matrices, almost all matrices have maximum rank. Here "almost all" is used in its standard mathematical meaning, i.e, the complementary set (of matrices not having maximum rank) has volume 0. On the other hand, consider the smallest nondegenerate 3-arrays, those that are $2 \times 2 \times 2$. They form an 8-dimensional space. As noted above, the maximum rank is 3, and of course the set A_3 of arrays of rank 3 has non-zero volume, but the set A_2 of arrays of rank 2 *also* has non-zero volume! In fact, by a Monte Carlo calculation, A_2 fills about 79% of the space and A_3 fills only about 21%. The set of arrays of rank ≤ 1 has zero volume. (The percentages actually refer to arrays on any fixed sphere around the origin in 8-space. The extrapolation to all arrays is justified because rank is constant over rays from the origin. The calculation was performed by filling 1000 $2 \times 2 \times 2$ arrays with random normal numbers, and calculating a certain polynomial that is defined only for $2 \times 2 \times 2$ arrays: If it is positive, the array has rank 2; if negative, rank 3, and if 0, the rank can be 0, 1, 2, or 3. The 0 case, which is expected to occur with probability 0, did not occur.)

The fact that A_3 and A_2 both have non-zero volume has an important practical consequence, which is described in Kruskal, Harshman, and Lundy (this volume) under the discussion of degenerate solutions to the PARAFAC model.

Difference 6: The dyadic decomposition of a rank R matrix is never unique, though decompositions of this kind sometimes achieve uniqueness by means of added constraints. In contrast, the triadic decomposition of a rank R 3-array is frequently unique. A precise definition of uniqueness and an often applicable theorem guaranteeing it for 3-arrays are presented below, in Sections 5 and 6.

The assertion that dyadic decompositions are never unique might seem to contradicted by the uniqueness properties of the Singular Value Decomposition (SVD) and of the Eigenvector Decomposition (ED), since each of these is associated with a dyadic decomposition that shares the uniqueness properties. In each of these cases, however, the uniqueness can be ascribed to added constraints, as we now explain. Suppose that $X = P D Q'$ is the SVD of X, so P and Q each have orthonormal columns and D is square and diagonal with positive entries on the main diagonal. Then X is equal to a linear combination of the outer products of the columns of P with the corresponding columns of Q, using coefficients from D. The well-known uniqueness property of the SVD when the diagonal entries of D are unequal yields uniqueness for this dyadic decomposition. The orthogonality of the columns of P and of the columns of Q are added constraints that suffice to provide this uniqueness.

Similarly, if $X = E \Lambda E'$ is the Eigenvalue Decomposition of a square matrix D that happens to have a complete set of eigenvalues, then X is equal to a linear combination of the outer products of each column of E with itself, using coefficients from Λ. The well-known uniqueness property of the ED when the diagonal entries of Λ are unequal yields uniqueness for this dyadic decomposition. The equality of the two factors of each dyad is an added constraint that suffices to provide this uniqueness.

4. GENERALIZING ROW AND COLUMN RANK

Not only do the concepts of row rank and column rank generalize to 3-arrays and N-arrays, but the generalized concepts are of great practical importance when dealing with 3-arrays, for a reason mentioned below. Recall that the row rank is a property of the set of rows of a matrix, likewise for the column rank, and that row rank = column rank = rank. All these properties generalize in a natural way.

A 2-array has sets of 1-arrays, namely, rows and columns. Similarly, an N-array has sets of $(N-1)$-arrays, which are called slabs. A 2-array has 2 such sets; an N-array has N such sets. Specifically, for each ν from 1 to N, define the set of slabs *in the* ν*-th direction*. Each of these slabs is formed by fixing the ν-th subscript of X and letting all the other subscripts vary. Each slab is an $(N-1)$-array. For a 2-array, slabs in direction 1 are rows and slabs in direction 2 are columns.

Now we need a property that applies to a set of slabs, i.e., arrays. The first property that comes to mind, based on the matrix case, is dimensionality. This is used below in defining the dimensionality vector of an array. Here however another property is used, one which is just the same as dimensionality when applied to the matrix case, namely, rank of a *set* of arrays:

> The *rank* of a set of N-arrays $\{Y_p\}$ is the smallest number R such that every array Y_p can be written as a linear combination of a single collection of R N-ads. [Here it is necessary to use linear combination and not sum, since different arrays Y_p may need different coefficients.]

Now the analogues of row rank and column rank of X are $\mathrm{rank}_\nu(X)$ for each ν from 1 to N, where

> $\mathrm{rank}_\nu(X)$ is the rank of the set of slabs of X in the ν-th direction.

For a 2-array X, $\mathrm{rank}_1(X)$ is row rank, and $\mathrm{rank}_2(X)$ is column rank. It is not hard to prove the following lemma.

> *Lemma 1*: $\mathrm{rank}_\nu(X) = \mathrm{rank}(X)$ for all ν.

Working with simple 3-arrays X for which it is possible to determine the rank, I have always found it easier in practice to use the definition of $\mathrm{rank}_\nu(X)$ for this purpose than to work directly with the original definition of rank. Furthermore, there are 3 different approaches available, according to whether one looks at $\mathrm{rank}_1(X)$, $\mathrm{rank}_2(X)$, or $\mathrm{rank}_3(X)$. It is often much easier to deal with one of these instead of another.

5. ROTATIONAL UNIQUENESS: BASIC CONCEPTS

Great attention has been paid in data analysis to decomposing data matrices into outer products,

$$X \cong \hat{X} \equiv \sum_{r=1}^{R} a_r \otimes b_r \,, \tag{3}$$

where each a_r and b_r is a vector. This decomposition underlies factor analysis, principal components analysis, correspondence analysis, and some other methods. Attention has also been paid to decomposing data 3-arrays into outer products in the same way,

$$Y \cong \hat{Y} \equiv \sum_{r=1}^{R} a_r \otimes b_r \otimes c_r \ , \tag{4}$$

where each a_r, b_r, and c_r is a vector. This decomposition underlies PARAFAC and CANDECOMP. In both equations, each outer product corresponds to one factor, i.e., to one source of influence. Assuming there are no superfluous terms, R is the rank of the fitted array in both cases.

There are some obvious *elementary changes* that can be made in any N-adic decomposition, like the two above, without changing its value. First, the factors or outer products can be permuted or relabeled, e.g., factor 1 is renamed factor 4, etc. Second, multipliers that cancel each other can be inserted, e.g., if $\alpha_r \beta_r \gamma_r = 1$ for all r, then

$$\hat{Y} \equiv \sum (\alpha_r a_r) \otimes (\beta_r b_r) \otimes (\gamma_r c_r) \ . \tag{5}$$

Two N-adic decompositions are *equivalent* if they have the same rank and one can be obtained from another by elementary changes.

A rank R N-adic decomposition of an N-array X is *rotationally unique*, often shortened to just *unique*, if all rank R decompositions of X are equivalent to it.

(If equivalence were not restricted to decompositions of the same rank, it would be necessary to include other operations among the elementary changes, such as adjoining a zero N-ad to the decomposition, and breaking one of the existing N-ads into several N-ads by expressing one factor of it as a sum of several vectors.)

By and large, decompositions of rank 1 (and 0) are unique regardless of rank, as the following lemma states precisely.

Lemma 4i: For any N, a rank 1 N-adic decomposition of an array is unique if the array contains no zero slabs in any direction.

This elementary fact is called Lemma 4i because it is essentially the same as Theorem 4i from Kruskal (1977).

It happens to be a mathematical fact that both dyadic and triadic decompositions are usually unique if their ranks are small enough, and are literally never unique if their ranks are large enough. However, the dividing lines between small and big rank are very different for dyadic and triadic decompositions. For dyadic decompositions, the dividing line is always 1. A dyadic decomposition is unique under mild conditions if it has rank ≤ 1 by Lemma 4i, while it is easily proved that a dyadic decomposition is never unique if it has rank ≥ 2. For triadic decompositions, the dividing line increases with the size of the array For $2 \times 2 \times 2$ arrays, it can be proved that a rank 2 decomposition of a rank 2 array is always unique, while a rank 3 decomposition is never unique. For $R \times R \times R$ arrays, Theorem 4a shows that many decompositions of rank $\lfloor 3R/2-1 \rfloor$ are unique, while Conjecture 4a suggests that none of higher rank are unique.

6. ROTATIONAL UNIQUENESS: SOME THEOREMS

A series of results that prove rotational uniqueness and that subsume almost all previous uniqueness results occur in (Kruskal, 1977), though a fairly simple independent result meant for use with INDSCAL occurs as Corollary 5 in (de Leeuw & Pruzansky, 1978). I explain Theorem 4a from (Kruskal, 1977) here and discuss how it applies to practical computational solutions from PARAFAC. Harshman hopes to apply these results to PARAFAC-2, and they might be applicable to other models as well. Earlier uniqueness results were achieved by Jennrich (unpublished) as recounted by Harshman in Section V, pp. 61-62 from (Harshman, 1970), Harshman (1972), Carroll (unpublished), and Kruskal (1976).

Some new terminology is needed. Suppose X has a triadic decomposition

$$X = \sum_{r=1}^{R} a_r \otimes b_r \otimes c_r \qquad (6)$$

Let a, b, c be the matrices whose values are a_{ir}, b_{jr}, c_{kr}. Define a *trio with R columns or factors* to be a triple of matrices, like (a, b, c), each of which has R columns. The matrices are referred to individually as the *loading matrices* of the trio for modes 1, 2, and 3. Thus a triadic decomposition of rank R is effectively the same thing as a trio with R columns. The two phrases are just different descriptions of the same thing, and are used here interchangeably. The r-th triad of the decomposition is the outer product of the r-th columns of the three loading matrices.

Suppose the matrices in trio (a, b, c) have sizes $I \times R$, $J \times R$, and $K \times R$. Define the *triple product* $[a, b, c]$ of the trio to be the 3-array of size $I \times J \times K$ whose ijk element is $\sum_r a_{ir} b_{jr} c_{kr}$. Then a trio is a triadic decomposition of X if and only if X is the triple product of the trio.

Equivalence of trios is defined implicitly through the meaning of trios as triadic decompositions, but it is helpful to see the meaning of equivalence directly in trio form. Permutation or relabeling of factors means multiplying the three loading matrices by a fixed permutation matrix p: if p is a permutation matrix, then $[ap, bp, cp] = [a, b, c]$. Insertion of multipliers that cancel each other means multiplying the three loading matrices by diagonal matrices α, β, γ whose product is the identity matrix: if $\alpha\beta\gamma=$ the identity matrix, then $[a\alpha, b\beta, c\gamma] = [a, b, c]$. Two trios (a, b, c) and $(\bar{a}, \bar{b}, \bar{c})$ are equivalent if and only if they have the same number of columns and $\bar{a} = ap\alpha$, $\bar{b} = bp\beta$, $\bar{c} = cp\gamma$, for p, α, β, γ, having the properties just mentioned. Theorems 4i, 4a, 4b, 4c, and 4d from (Kruskal, 1977) all yield rotational uniqueness as their conclusion. It is also possible to prove rotational uniqueness for one loading matrix without necessarily proving it for the whole decomposition. Two matrices a and $\bar{a}$ of the same size are called rotationally equivalent if $\bar{a} = ap\alpha$ for some permutation matrix p and some nonsingular diagonal matrix α. Theorems 3a, 3b, 3c, and 3d all yield rotational equivalence for a single loading matrix as their chief conclusion.

Suppose a matrix a has column rank k. Then a has some set of k independent columns. However, some other set of k columns might not be independent. Consider the property that *every* set of k columns is independent. Call this property *universal k-column independence.* Universal k-column independence implies $\text{rank}(a) \geq k$, but is much stronger. Every matrix a is universally 0-column independent. If a has no 0 columns, then a is also universally 1-column independent. If there are no 0-columns and if no pair of columns is proportional to each other, then a is universally 2-independent. Every matrix a has a largest integer I_1 for which it is universally I_1-independent. Harshman and Lundy (1984, p. 162) introduced the term *k-rank* for this value, since it is a rank-like number.

Theorem 4a. Assumption 1: $[a, b, c] = [\bar{a}, \bar{b}, \bar{c}]$.

Assumption 2: The two trios have the same number of columns.

Let R be the number of columns in the two trios. Let I_1, J_1, K_1 be the k-ranks of a, b, c.

Assumption 3: $I_1 + J_1 + K_1 \geq 2R + 2$.

Conclusion: (a, b, c) and $(\bar{a}, \bar{b}, \bar{c})$ are equivalent.

In practice, it is the corollary below that is most useful.

Corollary 4a: Suppose a trio (a, b, c) has R columns. Let I_1, J_1, K_1 be the k-ranks of a, b, c. If $I_1 + J_1 + K_1 \geq 2R + 2$, then (a, b, c) is rotationally unique.

The conditions of this theorem turn out to be met in almost all practical direct applications of PARAFAC, with the exception of the case when $R = 1$, which is filled by Lemma 4i above (or by Theorem 4i from Kruskal, 1977).

Theorem 4a is slightly misstated in (Kruskal, 1977), because the letters I_0, J_0, K_0 are unfortunately used instead of I_1, J_1, K_1, and the context gives an additional meaning to I_0, J_0, K_0. Thus the theorem appears to assume additional information about I_1, etc., but this additional information is not needed for the result. Theorem 4i is also slightly misstated in that paper, in that the hypothesis stated there is weaker than the hypothesis stated here for Lemma 4i and is inadequate; the hypothesis there is merely that the array is non-zero.

Theorem 4a does not state that its upper bound for R is sharp, i.e., it does not state that solutions fail to be unique if R is larger than the largest value allowed by Theorem 4a. However, a substantial amount of (unpublished) computational experience with synthetic data sets by Harshman and Lundy strongly suggests the following:

Conjecture 4a: If the rank of a decomposition is too large to be guaranteed unique by Theorem 4a or by Lemma 4i, then the decomposition is not unique.

7. ROTATIONAL UNIQUENESS: ERROR-FREE DATA

Now we show how Theorem 4a and Lemma 4i can be applied to PARAFAC solutions. PARAFAC consists of two steps:

(i) preprocessing the data 3-array X to yield a new 3-array which is called Y;

(ii) finding the trio (a, b, c) which yields the least-squares fit to $Y \cong \hat{Y} \equiv [a, b, c]$.

Only the second step is of concern here, so the data arrays are referred to as Y.

In discussing the relevance of the two results to PARAFAC solutions, we start with an unusual example from (Harshman, 1970), where there is a test of PARAFAC with simulated data. This example is unusual because it contains more factors than the size of the array in any direction. Harshman generated a trio $(\bar{a}, \bar{b}, \bar{c})$ of 8×10 loading matrices using random numbers, and took their triple product to form an $8 \times 8 \times 8$ data set Y which contains 10 factors. No random errors were added to Y. He applied PARAFAC and obtained a decomposition (a, b, c) equivalent except for roundoff error to $(\bar{a}, \bar{b}, \bar{c})$. The fitted array $\hat{Y}$ agreed with Y to within roundoff error. To apply the theorem, note that by constuction each of 8×10 matrices $\bar{a}$, $\bar{b}$, $\bar{c}$ has both rank and k-rank equal to min(8, 10) = 8. Then Theorem 4a yields that $(\bar{a}, \bar{b}, \bar{c})$ is rotationally unique if $8 + 8 + 8 \geq 2R + 2$, i.e., if $R \leq 11$. Thus Harshman's use of $R = 10$ falls within the theorem. In view of Conjecture 4a, it is fortunate that Harshman did not start with matrices that were 8×12 instead of 8×10! Of course, Harshman's work was done years before Theorem 4a was discovered.

In the same paper, and in subsequent unpublished work with Lundy, Harshman has used simulated data with and without error in a more systematic way to study rotational uniqueness. We present here an overview of the empirical results obtained for error-free data. In these studies, error-free $I \times J \times K$ data Y are always generated by filling an R-column trio $(\bar{a}, \bar{b}, \bar{c})$ with random numbers, where $R \leq \min(I, J, K)$, and taking the triple product Y as the data. PARAFAC is always used to obtain several S-factor solutions $Y \cong \hat{Y} \equiv [a, b, c]$. The results fall into three different cases.

Case (i): Solutions with $S = R$ factors.
In this case Y is almost always equal to $\hat{Y}$, and almost all solution trios (a, b, c) are equivalent to the generating trio $(\bar{a}, \bar{b}, \bar{c})$ and hence to each other. Failure of $Y = \hat{Y}$ can occur due to a local minimum solution, and in this case (a, b, c) is of course not equivalent to $(\bar{a}, \bar{b}, \bar{c})$. No other failures of equivalence are observed.

Case (ii): Solutions with $S = R + 1$ factors.
In this case Y is almost always equal to $\hat{Y}$, and the solution trios (a, b, c) are never equivalent to each other. Local minimum solutions can occur here also.

Case (iii): Solutions with $S < R$ factors.
In this case it is never true that $Y = \hat{Y}$. For most data sets Y, all solution trios (a, b, c) yield approximately equal values of $\hat{Y}$, and these solution trios are approximately equivalent. Sometimes there are two (or occasionally even more) different values of $\hat{Y}$. The $\hat{Y}$ corresponding to the smallest sum of squared residual errors is in all likelihood the global minimum solution, while the others are merely local minimum solutions. When such local minima occur, solution trios (a, b, c) yielding approximately equal values of $\hat{Y}$ are approximately equivalent, but solution trios yielding different values of $\hat{Y}$ cannot be equivalent of course.

In case (iii), the solution trios (a, b, c) are frequently equivalent to subtrios of $(\bar{a}, \bar{b}, \bar{c})$, and the factors of $(\bar{a}, \bar{b}, \bar{c})$ which appear in (a, b, c) are those which contribute the greatest variance to the triple product. When (a, b, c) is not equivalent to a subtrio of $(\bar{a}, \bar{b}, \bar{c})$, it is generally the case that one factor from (a, b, c) appears to be a compromise or combination in some sense between two factors of $(\bar{a}, \bar{b}, \bar{c})$, and that all the other factors from (a, b, c) agree with factors from $(\bar{a}, \bar{b}, \bar{c})$.

How do these results match the theorem? First consider Case (i), solutions having R factors. In all the investigations under discussion at this point, the loading matrices $\bar{a}$, $\bar{b}$, $\bar{c}$ used to create the data have at least as many rows as columns (i.e., $I \geq R$, etc.). Then the method of construction guarantees that both the rank and the k-rank of $\bar{a}$, $\bar{b}$, $\bar{c}$ are all R because, e.g., $\min(I, R) = R$. The numerical condition of Corollary 4a now becomes $R + R + R \geq 2R + 2$, i.e., $R \geq 2$. Thus by Corollary 4a if $R \geq 2$, or by Lemma 4i if $R = 1$, the trio $(\bar{a}, \bar{b}, \bar{c})$ is rotationally unique, so all decompositions of Y must be equivalent to it and hence to each other, as observed.

Next consider Case (ii), solutions having $R + 1$ factors. The following will be proved:

Assertion: No rank $R + 1$ decomposition (a, b, c) of Y can satisfy the conditions of Corollary 4a or of Lemma 4i.

Thus from Conjecture 4a, such a decomposition would not be expected to be unique, in accordance with the Case (ii) results stated above. By the way, the chief evidence for Conjecture 4a is just the nonuniqueness observed in Case (ii) situations. To prove the assertion above, define a new trio $(\tilde{a}, \tilde{b}, \tilde{c})$ of rank $R + 1$ which consists of the triads of $(\bar{a}, \bar{b}, \bar{c})$ together with the zero triad. (The zero triad means the outer product of three zero vectors). Obviously $[\tilde{a}, \tilde{b}, \tilde{c}] = Y$. This trio does not satisfy the conditions of Corollary 4a because the k-ranks of $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ are all 0, and the trio obviously does not satisfy the conditions of Lemma 4i. Now suppose some other rank $R+1$ decomposition (a, b, c) of Y satisfies the conditions of Corollary 4a or Lemma 4i. Then it would be unique, and so it would have to be equivalent to $(\tilde{a}, \tilde{b}, \tilde{c})$. But it is easy to check that if two trios are equivalent, then their loading matrices have the same k-ranks. Thus either both trios satisfy the conditions of Corollary 4a or neither satisfies the conditions of Corollary 4a, and likewise for Lemma 4i. Thus if (a, b, c) satisfied the conditions of either Corollary 4a or Lemma 4i, then it would not satisfy those conditions. This contradiction proves the assertion.

Harshman (1970) noted a phenomenon of "partial uniqueness" in a Case (ii) example in which $R=4$. Though the 5-factor decompositions were not unique, and having 5 factors could not be equivalent to the 4-factor generating trio $(\bar{a}, \bar{b}, \bar{c})$, he observed that each 5-factor solution trio shared two factors with $(\bar{a}, \bar{b}, \bar{c})$. Different solutions shared different factors. Recent computational experiments by Kruskal, carried out after the Multiway '88 meeting, have greatly clarified and partially explained this phenomenon, but there is not enough space to include this material here.

Last consider Case (iii), solutions with $S < R$ factors for some fixed S. In Cases (i) and (ii), $Y = \hat{Y}$ always holds, so the trio decomposing $\hat{Y}$ is also a decomposition of Y. Here however $Y \neq \hat{Y}$, and it is important to remember that the trio decomposes $\hat{Y}$. As in Case (i), Corollary 4a or Lemma 4i can be applied to show that each decomposition is rotationally unique, so that trios corresponding to approximately equal $\hat{Y}$ values should be approximately equivalent, as the Case (iii) results show they are.

8. DIMENSIONALITY VECTOR OF A MATRIX

Here is another generalization of row rank and column rank of a matrix. It is a far simpler concept than that introduced above, but useful nevertheless. Let $\dim_v$ be the dimensionality of the space generated by all slabs of X in direction v. Then the *dimensionality vector* of X is the N-tuple $(\dim_1, \cdots, \dim_N)$.

Before continuing, note that there is another useful characterization of $\dim_v$ based on fibers of X in direction v. The *fibers of X in direction* v are all vectors which can be formed by letting the v-th subscript of X vary and fixing all the other subscripts. Fibers are complementary in concept, and orthogonal in the array, to the corresponding slabs. Then a legitimate alternative definition of $\dim_v$ is the dimensionality of the space generated by the fibers of X in direction v. To see why this second definition agrees with the first one, consider a matrix Y whose rows are all the fibers in direction v of X. Then each column of Y consists of the elements of a slab in direction v. One definition of $\dim_v$ is the row rank of Y and the other definition is the column rank of Y.

The dimensionality vector of the zero N-array is (0, ... , 0). The dimensionality vector of a non-zero vector is always the 1-tuple (1), and the dimensionality vector of a matrix is the pair (row rank, column rank), and hence equal to (R,R) where R is the ordinary rank of the matrix. The 3-arrays shown earlier with ranks 0, 1, 2, 3 have dimensionality vectors respectively (0,0,0), (1,1,1), (2,2,2), and (2,2,2). Unlike rank, as discussed in the last paragraph of Section 2, the dimensionality vector does depend on the formal order of an array. Recall the three arrays X, Y, and Z discussed there. If the dimensionality vector of X is say, (I_0, J_0), then the dimensionality vectors of Y and Z are $(1, I_0, J_0)$, and $(1, 1, I_0, J_0)$.

What triples of numbers (R, S, T) can serve as the dimensionality vectors of 3-arrays? It is not hard to prove that dimensionality vector (R, S, T) satisfies the inequalities $R \le ST$, $S \le TR$, $T \le RS$. Furthermore, a triple (R, S, T) is the dimensionality vector of some array if and only if it satisfies these inequalities. Let us get some concrete idea of what triples are possible, writing them always with $R \le S \le T$ for simplicity. If $R=0$, then $S=T=0$, so the triple is (0,0,0) and the corresponding array is the zero array. If $R=1$, then $S=T$, so the triples are (1,1,1), (1,2,2), (1,3,3), etc. If $R=2$, then $S \le T \le 2S$, so the possible triples are (2,2,2), ... , (2,2,4), (2,3,3), ... , (2,3,6), etc.

What relationships are there between rank and dimensionality vector for 3-arrays? Here are two weak inequalities, where the dimensionality vector of X is (R, S, T):

$$\max(R, S, T) \le \mathrm{rank}(X) \le \min(RS, RT, ST) \; . \tag{7}$$

Applying results mentioned above in Section 3, if dimensionality vector is (3,3,3), then rank ≤ 5, and if dimensionality vector is $(2,S,T)$, then the theorem of Ja' Ja' gives a sharp upper bound on rank. It is not hard to construct a 3-array with dimensionality vector (3,4,5) and rank 6, and it can be proved that there is no smaller way to achieve all 4 values distinct.

9. CONNECTIONS WITH DATA ANALYSIS MODELS

Rank relates strongly to the PARAFAC and the CANDECOMP methods. (These two very similar methods were independently invented, but both rest on the same model for 3-way arrays.) Perhaps it may relate also to other methods. If X is a 3-array of data, then rank(X) is the smallest number of factors or dimensions with which PARAFAC and CANDECOMP can fit the data exactly. Each triad in the decomposition is one factor or dimension of the PARAFAC or CANDECOMP solution. The original references to PARAFAC and CANDECOMP are (Carroll & Chang, 1970) and (Harshman, 1970) respectively, though much has been written about them since.

Information about maximum possible ranks of arrays, regrettably still fragmentary, can be helpful to data analysts when interpreting the graph of rank versus dimensionality in order to decide how many dimensions to use.

The dimensionality vector relates strongly to 3-mode factor analysis (3-MFA), and may perhaps relate also other methods. If X is a 3-array of data, then the dimensionality vector of X is the size of the smallest core G for which 3-MFA can fit X exactly. Some basic initial references to 3-MFA are (Tucker, 1963), (Tucker, 1964), (Tucker, 1966), though a great deal has been written about it since.

It is helpful to understand the relationship between decomposing a matrix into dyads and a 3-array into triads as explained in this paper. The corresponding strong parallelism between factor analysis and PARAFAC makes the process of interpreting their results quite similar, with one exception. Direct PARAFAC solutions of real data are usually rotationally unique, while factor analysis solutions are not, and this difference substantially modifies the similarity. In other words, factor analysis solutions face the rotation problem, while PARAFAC solutions generally do not. The rotational uniqueness of PARAFAC-CANDECOMP solutions is seen by many people as a major advantage.

The rotational uniqueness property of PARAFAC is seen to be not merely the uniqueness of a particular model but rather to reflect the uniqueness of a fundamental decomposition (of 3-arrays into triads). This uniqueness, which was first expected intuitively and then demonstrated empirically, has now been fully justified mathematically and is much better understood. The number of dimensions up to which rotational uniqueness may be expected to hold has been determined and explained mathematically, and found to agree with the results of Monte Carlo computations. The k-rank conditions of Theorem 4 clarify the limitations in another direction on when rotational uniqueness may be expected to hold. The theorem on rotational uniqueness is aiding an investigation by Harshman and Lundy into models involving linearly dependent dimensions, by clearly showing that they can exist, and giving some of the conditions for this to happen.

Some unusual behavior of PARAFAC, namely, two-factor degeneracies, is now understood, and by analogy other degeneracies are partly understood.

REFERENCES

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

de Leeuw, J., & Pruzansky, S. (1978). A New Computational Method to Fit the Weighted Euclidean Distance Model. *Psychometrika, 43*(4), 479-490.

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

Harshman, R. A. (1972). *Determination and Proof of Minimum Uniqueness Conditions for PARAFAC1* . University of California at Los Angeles. UCLA Working Papers in Phonetics, 22:111-117.

Ja' Ja', J. (1979). Optimal Evaluation of Pairs of Bilinear Forms. *SIAM Journal on Computing, 8*(3), 443-461.

Kruskal, J. B. (1976). More Factors than Subjects, Tests, and Treatments: An Indeterminacy Theorem for Canonical Decomposition and Individual Differences Scaling. *Psychometrika, 41*(3), 281-293.

Kruskal, J. B. (1977). Three-Way Arrays: Rank and Uniqueness of Trilinear Decompositions, with Applications to Arithmetic Complexity and Statistics. *Linear Algebra and Its Applications, 18*, 95-138.

Kruskal, J. B. (future). The Maximum Rank of Arrays.

Mirsky, L. (1955). *An Introduction to Linear Algebra*. Oxford University Press.

Tucker, L. R. (1963). Implications of factor analysis of three-way matrices for measurement of change. In C. W. Harris (Ed.), *Problems in Measuring Change*. Madison, Wisconsin: University of Wisconsin Press.

Tucker, L. R. (1964). The extension of factor analysis to three-dimensional matrices. In N. Fredericksen & H. Gulliksen (Eds.), *Contributions to Mathematical Psychology*. New York: Holt, Rinhart, and Winston.

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