CHAPTER 2

CONFIRMATORY EVIDENCE FOR WEAK UNIQUENESS

In chapter 1 the sufficient condition $k_a+k_b+k_c \geq 2q+2$ for uniqueness (Kruskal, 1977) was explained. In practice, this sufficient condition is always fulfilled, because the probability that the PARAFAC algorithm yields (exactly) proportional columns is zero for empirical data. So in practice there is always uniqueness. In section 1.8 an example was given where the uniqueness of the PARAFAC components is weak. This diagnosis was based on the existence of alternative parameter matrices that do have rotational freedom and almost minimize the PARAFAC loss function. Now the question arises: How can it be determined to what extent, for any particular data set, the PARAFAC components are unique? In order to answer this question, we start with considering a necessary condition for uniqueness.

Suppose that a parameter matrix has k–rank 1, which means that there are at least two proportional columns in the matrix. Then there is in fact no uniqueness, hence having k–rank greater than 1 for the three parameter matrices is necessary for uniqueness. A proof follows in section 2.1. Accordingly, to verify the uniqueness of the PARAFAC solution in practice, one may reason as follows. If PARAFAC, constrained to have two proportional columns in A, B, or C, represents the data (almost) as well as unconstrained PARAFAC, then the PARAFAC components are *weakly* unique. On the other hand, the PARAFAC components are considered to be *strongly* unique if the discrepancy in fit between PARAFAC and PARAFAC with two proportional columns in one of the parameter matrices is large. In the present chapter such a constrained PARAFAC method will be described. First however, a necessary condition for uniqueness will be established.

2.1 A necessary condition for uniqueness

In chapter 1 it was explained that, if two columns in, for instance, the matrix C are proportional, then $k_C=1$ and therefore the above sufficient condition for uniqueness (Kruskal, 1977) is not satisfied. In fact, the condition that C does not have proportional columns is necessary for uniqueness, as will be proven now. No generality is lost by assuming that the first two columns in C are proportional. Let $A=(A_-|A_+)$, $B=(B_-|B_+)$, and $C=(C_-|C_+)$, where A_- is an $n\times2$ matrix, A_+ is an $n\times(q-2)$ matrix, B_- is an $m\times2$ matrix, B_+ is an $m\times(q-2)$ matrix, C_- is a $p\times2$ matrix with two proportional columns, and C_+ is a $p\times(q-2)$ matrix, then $AD_kB'=A_-D_{-k}B_-'+A_+D_{+k}B_+'$. From the fact that the first two columns in C are proportional it follows that $C_-=\mathbf{cd}'$ for a p–vector $\mathbf{c}$ and a 2–vector $\mathbf{d}$. Therefore, $D_{-k}=\text{Diag}(c_k\mathbf{1d}')=c_kD_-$, where c_k is a scalar and $D_-=\text{Diag}(\mathbf{1d}')$. Hence, $A_-D_{-k}B_-'=c_kA_-D_-B_-'=\overset{*}{A}\overset{*}{D}_k\overset{*}{B}{}'$, $k=1,\ldots,p$, where $\overset{*}{A}=A_-T$, $\overset{*}{B}=B_-D_-(T')^{-1}$, $\overset{*}{D}_k=c_kI$, for any non-singular matrix T. This shows that the first two columns in A and B are determined up to a non-singular transformation, hence the first two components are not unique.

Note that the non-uniqueness only occurs in the first two columns of A and B. Therefore, in case $q>2$, having proportional columns in C implies rotational freedom only for the components that correspond to the proportional columns in C and can therefore be called partial rotational freedom.

From the symmetry property and the above result, it follows that the condition that two columns in A or B are proportional is also sufficient for partial rotational freedom. For this reason, the PARAFAC components may be considered to be weakly unique if the discrepancy in fit between PARAFAC and PARAFAC constrained to have two proportional columns in A, B, or C is small.

2.2 An Algorithm for PARAFAC with two proportional columns in one of the parameter matrices

In order to examine the extent to which the PARAFAC components are uniquely determined it has been proposed to compare the discrepancy in fit between PARAFAC and PARAFAC with two proportional columns in, for instance, C. Therefore, the problem is to minimize the loss function

$$f(A,B,C)=\sum_{k=1}^{p}\|X_k-AD_kB'\|^2, \tag{2.1}$$

subject to the constraint that the first two columns in C are proportional.

In order to minimize $f(A,B,C)$ an ALS algorithm will be employed, in which first A is updated while B and C are fixed, next B is updated while A and C are fixed, and finally C is updated while A and B are fixed. From (2.1) it can be seen that to update A and to update B the PARAFAC algorithm can be used. This leaves us with updating C for fixed A and B, subject to the constraint that the first two columns in C are proportional. Imposing the identification constraint $\text{Diag}(C'C)=I_q$ implies that the first two columns of C are constrained to be equal. Specifically, using the fact that $\text{Vec}(\mathbf{ab}')=\mathbf{a}\otimes\mathbf{b}$ where Vec(.) denotes the vector containing all the elements of the matrix strung out rowwise into a column vector, and $\otimes$ denotes the (right) Kronecker product, we have

$$h(C)=\sum_{k=1}^{p}\|X_k-AD_kB'\|^2=\sum_{k=1}^{p}\|X_k-\sum_{r=1}^{q}\mathbf{a}_r\mathbf{b}_r'c_{kr}\|^2=\sum_{k=1}^{p}\|\text{Vec}(X_k)-\sum_{r=1}^{q}\mathbf{a}_r\otimes\mathbf{b}_rc_{kr}\|^2$$

$$=\sum_{k=1}^{p}\|\text{Vec}(X_k)-(\mathbf{a}_1\otimes\mathbf{b}_1|....|\mathbf{a}_q\otimes\mathbf{b}_q)\mathbf{c}_k\|^2, \tag{2.2}$$

where $\mathbf{c}_k'$ is row k in C. Let $\mathbf{d}_k$ denote the q−1 vector with $\mathbf{d}_k'=(c_{k1},c_{k3},...,c_{kq})$. We may write

$$(\mathbf{a}_1\otimes\mathbf{b}_1|...|\mathbf{a}_q\otimes\mathbf{b}_q)\mathbf{c}_k=(\mathbf{a}_1\otimes\mathbf{b}_1+\mathbf{a}_2\otimes\mathbf{b}_2|\mathbf{a}_3\otimes\mathbf{b}_3|...|\mathbf{a}_q\otimes\mathbf{b}_q)\mathbf{d}_k=U\mathbf{d}_k, \tag{2.3}$$

for k=1,...,p, where $U\equiv(\mathbf{a}_1\otimes\mathbf{b}_1+\mathbf{a}_2\otimes\mathbf{b}_2|\mathbf{a}_3\otimes\mathbf{b}_3|...|\mathbf{a}_q\otimes\mathbf{b}_q)$. From (2.2) and (2.3) it follows that the problem is to minimize

$$\mathrm{h}(\mathbf{d}_1,...,\mathbf{d}_p)=\sum_{k=1}^{p}\|\mathrm{Vec}(X_k)-U\mathbf{d}_k\|^2. \qquad (2.4)$$

From (2.4) it is clear that the minimizing $\mathbf{d}_k$ can be obtained as

$$\mathbf{d}_k=(U'U)^{-1}U'\mathrm{Vec}(X_k) \qquad (2.5)$$

for k=1,...,p. From (2.5) the solution for $\mathbf{c}_k'$ can be obtained as $\mathbf{c}_k'=(d_{k1},d_{k1},d_{k3},...,d_{kq})$, and D_k=Diag($\mathbf{1c}_k'$), for k=1,...,p. With the above procedures for updating the PARAFAC parameter matrices an ALS algorithm can be constructed which, after an arbitrary start, monotonically decreases f(A,B,C) until the function value stabilizes. A number of test trials revealed that the above algorithm is not sensitive to local minima. Because the danger of local minima cannot be ruled out completely, it is suggested to run more than one analysis, with different starting configurations for both A and B.

From the symmetry property it follows that analyzing the horizontal slices, or the lateral slices, by the above algorithm corresponds to PARAFAC subject to the constraint that the first two columns in the matrix A or B are proportional, respectively. So no additional algorithms need be developed to examine the strength of uniqueness of the PARAFAC components in other directions.

2.3 Examining uniqueness for empirical data

To illustrate how PARAFAC with proportional columns can be used to examine the uniqueness of the PARAFAC components, the results from analyzing three empirical data sets will be reported here. Prior to the analyses the data were centered and scaled to unit length over the occasions. The data were analyzed by PARAFAC and separately by PARAFAC with two proportional columns in A, B, or C, respectively, yielding four fit values per dimensionality. In three cases a degenerate PARAFAC solution was found. In these cases the data were analyzed by PFORTA in order to arrive at a non-degenerate solution. In these cases the uniqueness of the PFORTA components rather than the uniqueness of the PARAFAC components will be examined. In case of dimensionality 2, a non-degenerate solution is guaranteed for PARAFAC with two proportional columns in one of the parameter matrices, due to the freedom to orthonormalize one of the two other parameter matrices column-wise, see section 2.1. Unfortunately, the solutions for dimensionality 3, having one parameter matrix with two proportional columns, can be degenerate. Two such solutions were encountered. These two solutions will not be used for the examination of the degree of uniqueness of the PARAFAC or the PFORTA components. In Table 2.1 the percentages of the variance accounted for are reported for dimensionality 2 and 3.

Table 2.1 *The percentages of variance explained by PARAFAC with proportional columns in A, in B and in C, by unconstrained PARAFAC, and by PFORTA for three data sets.*

q=2	PARAFAC with prop. col. in			PARAFAC	PFORTA
Data set	*A*	*B*	*C*		
Affective Response	39.0	38.4	42.9	46.4	44.3
Tongue Shape	64.4	58.1	71.4	74.8	–
Triple Personality	41.0	57.4	42.8	60.9	–
q=3	PARAFAC with prop. col. in			PARAFAC	PFORTA
Data set	*A*	*B*	*C*		
Affective Response	48.1	53.8	50.5	54.0	51.7
Tongue Shape	82.2	79.2	84.8	86.6	–
Triple Personality	62.6	65.5	66.3	67.3	66.7

As a first data set the 32x6x8 Affective Response* data (see section 3.6) were analyzed. It was found that the two-dimensional PARAFAC solution is borderline degenerate (cosABC=–.78 after 2000 iterations, see section 3.6 for a more detailed description of these results), and that the three-dimensional solution is degenerate (cosABC=–.89 after 2000 iterations). For dimensionality 2, the discrepancy between PFORTA and PARAFAC with two proportional columns in C is 1.4% of variance explained, which indicates that the PFORTA components may be considered as weakly

*The author is obliged to Gudrun Eckblad who kindly made the data available.

unique. For dimensionality 3 of the PARAFAC solution having proportional columns in B, a value of $-.93$ of cosABC was encountered, hence this solution is degenerate. The discrepancy between PFORTA and PARAFAC with two proportional columns in C is 1.2%, which indicates that the PFORTA components for dimensionality 3 are also weakly unique.

The Tongue Shape data consist of heights of 13 tongue positions that were measured during the pronunciation of 10 vowels by five speakers, see Harshman, Ladefoged and Goldstein (1977) for a more detailed description of these data. The data were previously analyzed in two dimensions by Harshman, Ladefoged and Goldstein (1977), and Kruskal (1984), who called it a 'successful application'. For dimensionality 2, the discrepancy in fit between PARAFAC and PARAFAC with two proportional columns in C is 3.4% of variance explained, which indicates at least some degree of uniqueness. For dimensionality 3, the discrepancy in fit between PARAFAC and PARAFAC with two proportional columns in C is 1.8% of variance explained, which may be considered as some evidence for weak uniqueness of these PARAFAC components.

The Triple Personality data consist of scores from one person for 15 concepts on 10 variables (semantic differential scales) measured at six administrations, see Osgood and Luria (1954) or Kroonenberg (1983, pp. 227–242). For dimensionality 2, the discrepancy between PARAFAC and PARAFAC with two proportional columns in B is 3.5% of variance explained, which indicates at least some degree of uniqueness. For dimensionality 3, PARAFAC yielded a value of $-.94$ for cosABC, and PARAFAC with two proportional columns in B yielded a value of -1.00 for cosABC hence these

PARAFAC solutions are degenerate. For dimensionality 3 the discrepancy between PFORTA and PARAFAC with two proportional columns in C is 0.4% of variance explained, hence these PFORTA components are weakly unique.

2.4 Examining uniqueness by means of splithalf analysis

Harshman and Lundy (1984a, pp. 164–167) proposed to examine the uniqueness of the PARAFAC components by means of splithalf analysis. They reason that, if the data are split into two different halves and two separate PARAFAC analyses of these splithalves yield two sets of identical components, then the data contains enough systematic variation to determine components uniquely. To get an impression on how this approach works in practice, the results will be presented of a splithalf analysis of the so-called DAT* data. The DAT data consist of scores of 133 persons on nine variables that measure intelligence, at two occasions. The data were split into two sets by random assignment of the persons. The sets were separately centered within the occasions and scaled to unit length over the occasions. Each set was analyzed by an independent PARAFAC analysis with dimensionality 2. Next, Tucker's (1951) congruence coefficients for the corresponding columns of B and C across the two sets were computed as a measure for the proportionality of the components

*The author is obliged to Theo Nijsse who kindly made the data available.

across the two sets. In addition, for columns having unit length, say $\mathbf{b}_1$ and $\mathbf{b}_2$, the congruence coefficient equals $\mathbf{b}_1'\mathbf{b}_2$ and, because $\|\mathbf{b}_1 - \mathbf{b}_2\|^2 = 2(1 - \mathbf{b}_1'\mathbf{b}_2)$, it can also be seen as a measure of equality. Without loss of generality, the columns of B and C can be scaled to unit length. Hence, in case the congruence value, for the corresponding columns from PARAFAC analysis of each of the two splits, is close to 1, these columns are nearly equal. In case the congruence values encountered are greater than or equal to .85 (see Haven & Ten Berge, 1977), it will be said that the splithalf analysis indicates stable components. For the DAT data, five separate splithalf analyses were conducted. Two of the five splithalf analyses indicated stable components. The three lowest congruence values encountered for the unstable components were .31, .71, and .63. These results suggest that the PARAFAC components are not stable in the splithalf sense. In addition, it seems that it is hazardous to rely on only one splithalf analysis.

Now the question arises whether or not unstable components occur together with weak uniqueness. To shed a first light on this question, the same DAT data were analyzed by PARAFAC and by PARAFAC with two proportional columns in C. The data were preprocessed as before. The percentages of variance explained by PARAFAC and by PARAFAC with two proportional columns in C were 48.3 and 48.2, respectively. This discrepancy in fit can be regarded as negligible, and hence these PARAFAC components are weakly unique. This illustrates a case where the PARAFAC components are not stable in the splithalf sense and are weakly unique. For these data other components may exist, which, for instance, allow for an easier interpretation, and are stable in the splithalf sense.

Above, an example was presented where the PARAFAC components are unstable and the uniqueness is weak. To investigate the relation between stability and uniqueness more systematically, a simulation study was conducted.

2.5 A simulation study on PARAFAC uniqueness

It can be expected that, in case the PARAFAC components are stable in the splithalf sense, a characteristic of the population is responsible for this (see also Harshman & Lundy, 1984a, pp. 164–169). For instance, one might expect that stable sample components indicate unique PARAFAC components in the population. The question arises whether or not splithalf stability is sufficient to reveal such population characteristics. In order to answer this question a small simulation study was done. The data of this simulation study will also be used in chapter 4.

Three covariance matrices, Σ_1, Σ_2, and Σ_3, were constructed according to $\Sigma_i=\begin{pmatrix} BD_{1i} \\ BD_{2i} \\ BD_{3i} \end{pmatrix}(D_{1i}B' \,|\, D_{2i}B' \,|\, D_{3i}B')$, using one matrix $B=\begin{pmatrix} .85 & .15 \\ .75 & .25 \\ .15 & .85 \\ .25 & .75 \end{pmatrix}$, and three matrices $C_1=\begin{pmatrix} 1.50 & 0.50 \\ 1.00 & 1.00 \\ 0.50 & 1.50 \end{pmatrix}$, $C_2=\begin{pmatrix} 1.25 & 0.75 \\ 1.00 & 1.00 \\ 0.75 & 1.25 \end{pmatrix}$, and $C_3=\begin{pmatrix} 1.50 & 1.50 \\ 1.00 & 1.00 \\ 0.50 & 0.50 \end{pmatrix}$, where D_{ki} is row k of C_i, i=1,..,3. Such covariance matrices can be seen as the covariance matrices for data matrices of the form $(X_1 \,|\, X_2 \,|\, X_3)$, where the PARAFAC model holds for X_i, i=1,..,3, having orthonormal columns in A. By holding the matrix B constant and manipulating the matrices C_1, C_2, and

C_3, the PARAFAC components from Σ_1, Σ_2, and Σ_3, respectively, have different degrees of uniqueness. Accordingly, the covariance matrix Σ_1 is said to have 'strong' uniqueness, the covariance matrix Σ_2 is said to have 'medium' uniqueness and the covariance matrix Σ_3 is without uniqueness. These covariance matrices served as input for the SIMLIS program (Boomsma, 1983), which is a program for pseudo random sampling from a multivariate normal distribution. Samples of size n=40 and n=20 were drawn from these distributions. For each combination of degree of uniqueness and sample size, three samples were drawn, yielding 18 arrays of order n×4×3 in total.

Each of the 18 arrays was analyzed by the splithalf method followed by PARAFAC, using five splits. For the six samples drawn from the strongly unique population and the six samples drawn from the medium unique population all congruence values encountered were greater than .97. For the six samples drawn from the population without uniqueness one stable splithalf solution (out of 15 splithalf analyses) was encountered for n=20 and six stable splithalf solutions (out of 15) were encountered for n=40. The lowest congruence value encountered was .14 for n=20 and .25 for n=40. These results suggest that the PARAFAC components for samples drawn from populations having uniqueness are stable in the splithalf sense, even when the sample size is as low as 20. In addition, the PARAFAC components for samples from populations without uniqueness are not stable in the splithalf sense.

Next, to examine the degrees of uniqueness in the samples, the 18 arrays were analyzed by PARAFAC and by PARAFAC with two proportional columns in

C, with dimensionality 2. The discrepancies in percentage of fit between both methods for n=40 and n=20 were highly similar. For this reason, only the discrepancies for n=40 will be reported. The discrepancies encountered were, for strong uniqueness, 12.0, 13.6, and 15.9, for medium uniqueness, 3.6, 5.8, and 4.5 and, for non-uniqueness, 0.1, 0.1, and 0.1. It can be concluded that these discrepancies in the samples quite accurately reflect the extent of uniqueness in the population. It seems that both the proposed discrepancy in fit and an analysis of stability by splithalf can be used to make inferences about the degree of uniqueness of the components in the population.

2.6 Examining uniqueness in case the parameter matrices have disproportional columns

In case the PARAFAC solution of dimensionality 2 contains a matrix C with nearly proportional columns, it might be expected that the discrepancy between PARAFAC and PARAFAC constrained to have two proportional columns in C is small. On the other hand, it might be expected that in case clearly non-proportional columns in the parameter matrices are found, the uniqueness is strong. However, this expectation need not come true. This will be illustrated with the results of a PARAFAC analysis of a 42×4×2 empirical data set*. The data were centered within the occasions and

---

*The author is obliged to Fred Wolters who kindly made the data available.

scaled to unit length across the occasions. Tucker's (1951) congruence coefficients between the columns of the parameter matrices were computed as a measure for the proportionality of the columns in the parameter matrices. The values of the congruence coefficients are –.06, .17, and .66, respectively, for the columns of A, B, and C. Hence, the columns of the parameter matrices are clearly non–proportional. These parameter matrices have full rank, hence this PARAFAC solution is unique. However, the amount of explained variance by PARAFAC and by PARAFAC with proportional columns in C is 55.5 and 55.4, respectively. This demonstrates that these PARAFAC components are weakly unique. It can be concluded that clear non–proportionality of the columns in the PARAFAC parameter matrices is not sufficient for strong uniqueness of the PARAFAC components.

2.7 Discussion

It has been illustrated that splithalf analysis and PARAFAC with two proportional columns in one of the parameter matrices is useful to detect weak uniqueness of the PARAFAC and of the PFORTA components. In addition to this, one might want to use PARAFAC with two proportional columns in one of the parameter matrices for the purpose of representing a three–way array. This will be taken up in Chapter 3.

In case a splithalf analysis reveals that the PARAFAC components are not unique Harshman and Lundy (1984a, pp. 160–161) recommend to remove such non-unique components by appropriate preprocessing. Unfortunately, some preprocessing methods are rather difficult to understand from a substantive point of view. Instead of introducing different preprocessing methods in this study, alternative procedures for dealing with weakly unique PARAFAC solutions will be introduced. Specifically, in case the PARAFAC components are weakly unique, other components than the PARAFAC components exist that may be preferred, because, for instance, they have simple structure and allow therefore an easier interpretation. In the next three chapters, various constraints to impose on the PARAFAC parameter matrices will be considered in order to find (constrained) PARAFAC components that allow for an easier interpretation than the unconstrained PARAFAC components.