APPENDIX A

Van IJzendoorn and Kroonenberg (1990) used the TUCKALS3 algorithm for Weighted PCA with $q_1=q_2=q$ and $q_3=1$. In the TUCKALS3 algorithm the order of updating the matrices in one iterative cycle implicitly is: G, A, G, B, G, C (see Kiers et al., 1992). It has to be proven that the TUCKALS3 algorithm produces the same series of function values as the Weighted PCA algorithm, if the same starting configurations are used.

Let A^i, B^i, $\mathbf{c}^i$, and G^i denote the i^{th} updates in TUCKALS3 for A, B, $\mathbf{c}$, and G, respectively, and let A^0, B^0, and $\mathbf{c}^0$ denote starting values for A, B and $\mathbf{c}$. According to Kiers et al. (1992) updates are found as follows. The first update of G is $G^1=A^{0\prime}\sum_{k=1}^{p}c_k^0X_kB^0$, the first update of A is

$$A^1=\mathrm{GS}\left(\sum_{k=1}^{p}c_k^0X_kB^0G^{1\prime}\right), \tag{A.1}$$

and the next update for G is $G^2=A^{1\prime}\sum_{k=1}^{p}c_k^0X_kB^0$. Next, by updating B it is found that

$$B^1=\mathrm{GS}\left(\sum_{k=1}^{p}c_k^0X_k'A^1G^2\right). \tag{A.2}$$

For convenience, let $Z^1\equiv\left(\sum_{k=1}^{p}c_k^0X_k'A^1\right)$ and $W^1\equiv\mathrm{GS}(Z^1)$. Then B^1 and Z^1 span the same column space, hence $B^1=W^1V^1$ for a certain rotation matrix V^1. Next, the third update of G satisfies $G^3=A^{1\prime}\sum_{k=1}^{p}c_k^0X_kB^1$. Using that $B^1=W^1V^1$, it is found that $G^3=Z^{1\prime}W^1V^1$. Hence, for the TUCKALS3 representation of X_k, it follows that

$$\hat{X}_k \equiv c_k^0 A^1 G^3 B^{1\prime} = c_k^0 A^1 Z^{1\prime} W^1 V^1 V^{1\prime} W^{1\prime} = c_k^0 A^1 Z^{1\prime} W^1 W^{1\prime}, \tag{A.3}$$

k=1,..,p. We may simplify this representation by noting that W^1 is a projector for the column space of Z^1, so that $Z^1 = W^1 W^{1\prime} Z^1$. By substituting Z^1 for $W^1 W^{1\prime} Z^1$ into (A.3) it is found that

$$\hat{X}_k = c_k^0 A^1 G^3 B^{1\prime} = c_k^0 A^1 Z^{1\prime}. \tag{A.4}$$

The first update of $\mathbf{c}$ satisfies $\mathbf{c}^1 = \mathrm{GS}(\mathbf{d}^1)$, with $d_k^1 = \mathrm{tr} A^{1\prime} X_k B^1 G^{3\prime}$, where d_k^1 denotes element k of $\mathbf{d}^1$. From (A.4) it follows that

$$d_k^1 = \mathrm{tr} A^{1\prime} X_k B^1 G^{3\prime} = \mathrm{tr} X_k' A^1 G^3 B^{1\prime} = \mathrm{tr} X_k' A^1 Z^1, \tag{A.5}$$

which completes the first full cycle of updatings in TUCKALS3.

In order to show that the Weighted PCA algorithm produces the same function value after one complete cycle, it is convenient to express the Weighted PCA loss function as

$$\mathrm{WPCA}(E,F,\mathbf{h}) = \sum_{k=1}^{p} \|X_k - h_k E F'\|^2, \tag{A.6}$$

where E plays the role of A, F the role of B and $\mathbf{h}$ the role of $\mathbf{c}$. Let E^i, F^i and $\mathbf{h}^i$ denote the i^{th} update of E, F and $\mathbf{h}$, respectively, and let starting configurations be given such that $F^0 = B^0 T^0$ and $\mathbf{h}^0 = \mathbf{c}^0$, where T^0 denotes an arbitrary non-singular matrix. Specifically, it will be proven that, after updating E and F, the representation $\hat{X}_k \equiv h_k^1 E^1 F^{1\prime}$ for X_k from the Weighted PCA algorithm equals the representation $\hat{X}_k = c_k^1 A^1 G^3 B^{1\prime}$ from one cycle of the TUCKALS3 algorithm, and hence TUCKALS$(A^1,B^1,\mathbf{c}^1,G^3)$=WPCA$(E^1,F^1,\mathbf{h}^1)$.

In the Weighted PCA algorithm (see section 3.3), the first update of E satisfies $E^1=\mathrm{GS}\left(\sum_{k=1}^{p} h_k^0 X_k F^0\right)$. By substitution of B^0T^0 for F^0, and $\mathbf{c}^0$ for $\mathbf{h}^0$ into the first update for E and by using the fact that $\left(\sum_{k=1}^{p} c_k^0 X_k B^0 T^0\right)$ and A^1 span the same column space, we find

$$E^1=\mathrm{GS}\left(\sum_{k=1}^{p} c_k^0 X_k B^0 T^0\right)=A^1S^1, \tag{A.7}$$

where S^1 denotes a certain rotation matrix. The first update of F satisfies $F^1=\sum_{k=1}^{p} h_k^0 X_k' E^1$, see section 3.3. Now substituting A^1S^1 for E^1, $\mathbf{c}^0$ for $\mathbf{h}^0$ and Z^1 for $\left(\sum_{k=1}^{p} c_k^0 X_k' A^1\right)$ into the first update of F, yields

$$F^1=\sum_{k=1}^{p} c_k^0 X_k' A^1 S^1=Z^1S^1=B^1T^1, \tag{A.8}$$

for a certain non-singular matrix T^1. Hence, it follows from (A.7), (A.8) and the substitution of $\mathbf{c}^0$ for $\mathbf{h}^0$ into $h_k^0E^1F^{1\prime}$ that $h_k^0E^1F^{1\prime}=c_k^0A^1S^1S^{1\prime}Z^{1\prime}=c_k^0A^1Z^{1\prime}$, which equals (A.4).

Now it will be shown that $\mathbf{c}^1=\mathbf{h}^1$. The first update of $\mathbf{h}$ satisfies $\mathbf{h}^1=\mathrm{GS}(\mathbf{g}^1)$, with $g_k^1=\mathrm{tr}E^{1\prime}X_kF^1=\mathrm{tr}X_k'E^1F^{1\prime}$, where g_k^1 denotes element k of $\mathbf{g}^1$. From substituting $A^1Z^{1\prime}$ for $E^1F^{1\prime}$ into $g_k^1=\mathrm{tr}X_k'E^1F^{1\prime}$, it follows that $g_k^1=\mathrm{tr}X_k'A^1Z^{1\prime}$, which equals (A.5). This proves that $\mathbf{c}^1=\mathbf{h}^1$, and therefore $h_k^1E^1F^{1\prime}=c_k^1A^1Z^{1\prime}=c_k^1A^1G^3B^{1\prime}$. Thus, after the first complete cycle, the same representations are found by the two algorithms and hence the TUCKALS3 algorithm and the Weighted PCA algorithm produce the same function values after one complete cycle. In passing, it has been shown that after the first iterative cycle it is true that $F^1=B^1T^1$ (A.8) and $\mathbf{h}^1=\mathbf{c}^1$. Hence, using exactly the same reasoning as above, it can be shown that, after the second iterative cycle, TUCKALS(A^2,B^2,$\mathbf{c}^2$,G^6)=WPCA(E^2,F^2,$\mathbf{h}^2$), and, more generally, that after the i^{th} iterative cycle TUCKALS(A^i,B^i,$\mathbf{c}^i$,G^{3i})=WPCA(E^i,F^i,$\mathbf{h}^i$), which completes the proof.

APPENDIX B

Generalized Perron–Frobenius Theorem: Let $\mathbf{x}_{jt}$ denote column j of the $n{\times}m$ frontal slice t of the $n{\times}m{\times}p$ three-way array X. Let it be given that

$$\mathbf{x}_{jt}'\mathbf{x}_{is} \geq 0, \text{ for } i=1,\ldots,m,\ j=1,\ldots,m,\ s=1,\ldots,p, \text{ and } t=1,\ldots,p. \tag{B.1}$$

Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ denote the parameter vectors that minimize the PARAFAC loss function for dimensionality 1. Then the elements of $X_s'\mathbf{a}$, $s=1,\ldots,p$, $\mathbf{b}$ and $\mathbf{c}$ can be taken to have no negative elements.

Proof: It will first be proven that $\mathbf{b}$ and $\mathbf{c}$ can be taken non–negative. It can always be arranged that $\mathbf{b}'\mathbf{b}=\mathbf{c}'\mathbf{c}=1$. Clearly, $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ minimize

$$\mathrm{f}(\mathbf{a},\mathbf{b},\mathbf{c}) = \sum_{s=1}^{p} \|X_s - \mathbf{a}c_s\mathbf{b}'\|^2. \tag{B.2}$$

From this and the optimality of $\mathbf{a}$, it follows that $\mathbf{a}=\sum_{t=1}^{p} c_tX_t\mathbf{b}$, see equation (1.4). Upon substituting $\sum_{t=1}^{p} c_tX_t\mathbf{b}$ for $\mathbf{a}$ in (B.2), it is found that

$$\mathrm{g}(\mathbf{b},\mathbf{c}) = \mathbf{b}'\left(\sum_{s=1}^{p} c_sX_s\right)'\left(\sum_{t=1}^{p} c_tX_t\right)\mathbf{b}, \tag{B.3}$$

is maximal with $\mathbf{b}'\mathbf{b}=\mathbf{c}'\mathbf{c}=1$. From (B.3) and (B.1) it follows that

$$\mathrm{g}(\mathbf{b},\mathbf{c}) = \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{s=1}^{p}\sum_{t=1}^{p} b_ib_jc_sc_t\mathbf{x}_{jt}'\mathbf{x}_{is} \leq \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{s=1}^{p}\sum_{t=1}^{p} |b_i|\,|b_j|\,|c_s|\,|c_t|\,\mathbf{x}_{jt}'\mathbf{x}_{is}, \tag{B.4}$$

where $|.|$ denotes the absolute value of (.). From (B.4) it can be seen that reflection of negative elements in **b** or **c** does not decrease the value of g(**b**,**c**); hence **b** and **c** can be chosen to have non-negative values only. Next, after substituting $\sum_{t=1}^{p} c_t X_t \mathbf{b}$ for **a** into $X_s'\mathbf{a}= \sum_{t=1}^{p} c_t X_s' X_t \mathbf{b}$ we find that $X_s'\mathbf{a}$, $s=1,\ldots,p$, is non-negative, because (B.1) is fulfilled and **b** and **c** can be taken non-negative, which completes the proof. It seems worthwhile noting that the above result generalizes the Perron-Frobenius theorem and that it can be generalized to the analysis of a so-called N-way array.

REFERENCES

Boomsma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and non-normality.* Doctoral dissertation, University of Groningen.

Braverman, E.M. (1970). Methods for the extremal grouping of parameters and the problem of determining essential factors. *Automation and Remote Control*, 1, 108–116.

Brogden, H.E. (1969). Pattern, structure, and the interpretation of factors. *Psychological Bulletin*, 72, 375–378.

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.

Carroll, J.D., De Soete, G, & Pruzansky, S (1989). Fitting of the latent class model via iteratively reweighted least squares CANDECOMP with non–negativity constraints (463–472). In R. Coppi and S. Bolasco (Eds.), *Analysis of multiway data matrices* (pp. 463–472). Amsterdam: North Holland.

Cattell, R.B. (1944). "Parallel proportional profiles" and other principles for determining the choice of factors by rotation. *Psychometrika*, 9, 267–283.

Cliff, N. (1966). Orthogonal rotation to congruence. *Psychometrika*, 31, 33–42.

Draper, N.R., & Smith, H. (1981). *Applied regression analysis.* New York: Wiley.

Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1, 211–218.

Eckblad, G. (1981). Assimilation resistance and affective response in problem solving. *Scandinavian Journal of Psychology*, 1981, 22, 1–16.

Escoufier, Y. (1988). Beyond correspondence analysis. In: H.H. Bock (Ed.) *Classification and related methods of data analysis* (pp. 505–514). Amsterdam: North Holland.

Gorsuch, R,L. (1983). *Factor analysis*. Philadelphia: Saunders.

Guttman, L. (1952). Multiple group methods for common–factor analysis: Their basis, computation, and interpretation. *Psychometrika*, 17, 209–222.

Harman, H.H. (1976). *Modern factor analysis* (3rd. ed.). Chicago: University of Chicago Press.

Harris, C.W., & Kaiser, H.F. (1964). Oblique factor analytic solutions by orthogonal transformations. *Psychometrika*, 29, 347–362.

Harshman, R.A. (1970). Foundations of the PARAFAC procedure: Models and conditions for an "exploratory" multi–modal 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., & DeSarbo, W.S. (1984). An application of PARAFAC to a small sample problem, demonstrating preprocessing, orthogonality constraints, and split–half diagnostic techniques. In H.G. Law, C.W. Snyder, J.A. Hattie, & R.P. McDonald (Eds.), *Research methods for multi–mode data analysis* (pp. 602–642). New York: Praeger.

Harshman, R.A., Ladefoged, P. & Goldstein, L. (1977). Factor analysis of tongue shapes. *Journal of the Acoustical Society of America*, 62, 693–707.

Harshman, R.A., & Lundy, M.E. (1984a). The PARAFAC model for three-way factor analysis and multidimensional scaling. In H.G. Law, C.W. Snyder, J.A. Hattie, & R.P. McDonald (Eds.), *Research methods for multi-mode data analysis* (pp. 122–215). New York: Praeger.

Harshman, R.A., & Lundy, M.E. (1984b). Data preprocessing and the extended PARAFAC model. In H.G. Law, C.W. Snyder, J.A. Hattie, & R.P. McDonald (Eds.), *Research methods for multi-mode data analysis* (pp. 216–284). New York: Praeger.

Haven, S., & Ten Berge, J.M.F. (1977). Tucker's coefficient of congruence as a measure of factorial invariance: An empirical study. *Heymans Bulletin* 290 EX. Department of Psychology, University of Groningen.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, 417–441.

Jöreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34, 183–202.

Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187–200.

Kiers, H.A.L. (1991). Hierarchical relations among three-way methods. *Psychometrika*, 56, 449–470.

Kiers, H.A.L., & Krijnen, W.P. (1991). An efficient algorithm for PARAFAC of three-way data with large numbers of observation units. *Psychometrika*, 56, 1, 147–152.

Kiers, H.A.L., Kroonenberg, P.M. & Ten Berge, J.M.F. (1992). An efficient algorithm for TUCKALS3 on data with large numbers of observation units. *Psychometrika*, 57, xxx–xxx.

Krijnen, W.P., & Ten Berge, J.M.F. (1992). A constrained PARAFAC method for positive manifold data. *Applied Psychological Measurement*, 16, xxx–xxx.

Kroonenberg, P.M. (1983). *Three-mode principal component analysis: Theory and applications*. Leiden: DSWO Press.

Kroonenberg, P.M. & De Leeuw, J. (1980). Principal components analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika*, 45, 69–97.

Kroonenberg, P.M., Ten Berge, J.M.F., Brouwer, P., & Kiers, H.A.L. (1989). Gram–Schmidt versus Bauer–Rutishauser in alternating least–squares algorithms for three-mode principal component analysis. *Computational Statistics Quarterly* 2, 81–87.

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. (1984). Multilinear methods. In H.G. Law, C.W. Snyder, J.A. Hattie, & R.P. McDonald (Eds.), *Research methods for multi-mode data analysis* (pp. 36–62). New York: Praeger.

Kruskal, J.B. (1989). Rank, decomposition, and uniqueness for 3-way and n-way arrays. In R. Coppi and S. Bolasco (Eds.), *Analysis of multiway data matrices* (pp. 7–18). Amsterdam: North Holland.

Kruskal, J.B., Harshman, R.A., & Lundy, M.E. (1989). How 3–MFA data can cause degenerate PARAFAC solutions, among other relationships. In R. Coppi and S. Bolasco (Eds.), *Analysis of multiway data matrices* (pp. 115–122). Amsterdam: North Holland.

Lawson, C.L., & Hanson, R.J. (1974). *Solving least squares problems.* New Jersey: Prentice–Hall.

Lundy, M.E., & Harshman, R.A. (1985). *Reference manual for the PARAFAC analysis package.* London, Ontario: Scientific Software Associates.

Mulaik, S.A. (1972). *The foundations of factor analysis.* New York: McGraw–Hill.

Nunnally, J.C. (1978). *Psychometric theory.* New York: McGraw–Hill.

Osgood, C.E. & Luria, Z. (1954). A blind analysis of a case of multiple personality. *Journal of Abnormal and Social Psychology*, 49, 579–791.

Ten Berge, J.M.F. (1986). *Three notes on three–way analysis.* Paper presented at the workshop on TUCKALS and PARAFAC, University of Leiden, July 2.

Ten Berge, J.M.F. (1989). Convergence of Parafac preprocessing procedures and the Deming–Stephan method of iterative proportional fitting. In R. Coppi and S. Bolasco (Eds.), *Analysis of multiway data matrices* (pp. 53–64). Amsterdam: North Holland.

Ten Berge, J.M.F., De Leeuw, J. & Kroonenberg, P.M. (1987). Some additional results on principal components analysis of three–mode data by means of alternating least squares algorithms. *Psychometrika*, 52, 183–191.

Ten Berge, J.M.F., Kiers, H.A.L., & De Leeuw, J. (1988). Explicit CANDECOMP/PARAFAC solutions for a contrived 2x2x2 array of rank three. *Psychometrika*, 53, 579–584.

Tenenhaus, M. (1988). Canonical analysis of two polyhedral cones and applications. *Psychometrika*, 53, 503–524.

Thurstone, L.L. (1947). *Multiple-factor analysis: A development and expansion of the vectors of mind.* Chicago: The University of Chicago Press.

Tucker, L.R. (1951). *A method for synthesis of factor analysis studies* (Personnel Research Section Report No. 984). Washington, D.C.: Department of the Army.

Van der Burg, E., De Leeuw, J. & Dijksterhuis, G.B. (1992). OVERALS: Nonlinear canonical correlation with k sets of variables. To appear in: R. Coppi & A. Di Ciaccio (Eds.), *Special Issue of Computational Statistics and Data Analysis devoted to Multiway Data analysis, software and applications.*

Van der Kloot, W.A. & Kroonenberg, P.M. (1985). External analysis with three-mode principal component models. *Psychometrika*, 50, 479–494.

Van Eldik, M.C.M., Neutel, R.J., Van der Meulen, B.F. & Spelberg, H.C.L. (1990). *Groningse Ontwikkelingsschalen. Constructiefase: Opzet en resultaten.* Internal Report, Department of Special Education, University of Groningen.

Van IJzendoorn, M.H. & Kroonenberg, P.M. (1990). Cross-cultural consistency of coding the strange situation. *Infant Behavior And Development*, 13, 469–485.

DE ANALYSE VAN DRIE-WEG GEGEVENS DOOR GERESTRICTEERDE PARAFAC METHODEN

Voor de exploratieve analyse van tweeweg-gegevens, bijvoorbeeld scores van personen op variabelen, is Principale Componenten Analyse (PCA) een nuttige methode. Met PCA kunnen dergelijke gegevens samengevat (gerepresenteerd) worden in een aantal componenten die een maximale hoeveelheid variantie verklaren. PCA komt neer op het ontbinden van de matrix met gegevens in twee matrixen: de matrix met coëfficiënten van de personen op de componenten en de matrix met coëfficiënten van de variabelen op de componenten (patroonmatrix). Indien drieweg-gegevens, bijvoorbeeld scores van personen op variabelen gemeten op een aantal tijdstippen (tijdstippen-matrix), beschikbaar zijn, lijkt PARAFAC (Harshman, 1970), een afkorting voor PARAllele FACtor-analyse, een nuttige methode voor de exploratieve analyse van dergelijke gegevens. In dit proefschrift worden gerestricteerde PARAFAC-methoden ontwikkeld om enerzijds de mate van uniciteit vast te stellen en anderzijds een fraaiere PARAFAC-representatie van de gegevens te vinden.

In hoofdstuk 1 wordt een overzicht gegeven van de belangrijkste eigenschappen van PARAFAC. Er wordt uitgelegd dat PARAFAC opgevat kan worden als een generalisatie van PCA en dat PARAFAC neerkomt op het ontbinden van het array met drieweg-gegevens in drie matrixen: de matrix met coëfficiënten van de personen op de componenten (componentenmatrix), de matrix met coëfficiënten van de variabelen op de componenten (patroonmatrix) en de matrix met coëfficiënten van de tijdstippen op de componenten (tijdstippen–matrix). De PARAFAC–representatie van de gegevens heeft als eigenschap dat de componenten, over de tijdstippen heen, variëren in belangrijkheid. Een van de opmerkelijke eigenschappen van PARAFAC is de uniciteit van de PARAFAC–componenten. Hiermee wordt bedoeld dat rotatie van de PARAFAC–componenten tot een slechtere fit leidt en om deze reden onwenselijk is. Er worden voldoende voorwaarden voor uniciteit besproken en er wordt uitgelegd dat deze bij de analyse van empirische drieweg–gegevens altijd vervuld zijn. Tevens wordt er met een aantal theoretische voorbeelden aangetoond dat de mate van uniciteit van de PARAFAC componenten gering kan zijn. In een dergelijk geval wordt van zwakke uniciteit gesproken. Als de uniciteit van de PARAFAC–componenten zwak is dan leidt rotatie van, op zijn minst, twee PARAFAC–componenten nauwelijks tot een slechtere fit. Afgezien van het geval van zwakke uniciteit, kan men zich in het algemeen afvragen of er PARAFAC–componenten bestaan die eenvoudiger te interpreteren zijn en daarmee een fraaiere representatie van de gegevens vormen. In dit proefschrift wordt, zowel voor het vastellen van zwakke uniciteit als voor het vinden van een fraaiere PARAFAC–representatie van de gegevens, een oplossing gezocht in het opleggen van bepaalde randvoorwaarden aan de PARAFAC–methode. Tevens wordt via split–half analyse onderzocht in welke mate de (gerestricteerde) PARAFAC–componenten stabiel zijn.

Hoofdstuk 2 heeft als doel het vaststellen van zwakke uniciteit. Er wordt aangetoond dat het hebben van twee proportionele kolommen in één van de drie PARAFAC–matrixen voldoende is voor rotatievrijheid, en dus voldoende is voor de afwezigheid van uniciteit. Er wordt een algoritme afgeleid dat het PARAFAC model fit aan de gegevens onder de restrictie dat één van de drie matrixen een proportioneel kolommenpaar heeft. Aan de hand van analyses van empirische drieweg-gegevens wordt geïllustreerd dat de fit van deze gerestricteerde variant van PARAFAC vrijwel even goed kan zijn als de fit van (ongerestricteerde) PARAFAC. In een dergelijk geval wordt geconcludeerd dat de uniciteit van de PARAFAC componenten zwak is.

In Hoofdstuk 3 wordt nagegaan of componenten met dezelfde relatieve belangrijkheid over de tijdstippen een fraaie representatie van de gegevens kunnen vormen. Hiervoor wordt het idee van proportionele kolommen in de tijdstippen-matrix verder uitgewerkt. Er wordt aangetoond dat PARAFAC onder de restrictie van proportionele kolommen in de tijdstippen-matrix neerkomt op PCA van een gewogen som van gegevens-matrixen per tijdstip (Gewogen PCA). Gewogen PCA past in een hiërarchie van methoden om drieweg-gegevens te analyseren. Er wordt een efficiënt algoritme voor Gewogen PCA afgeleid. Aan de hand van een analyse van empirische gegevens wordt geïllustreerd hoe rotatievrijheid bij Gewogen PCA benut kan worden om simple structure te benaderen.

Hoofdstuk 4 heeft als doel het analyseren van zogenaamde positive manifold gegevens. Hiermee worden (drieweg-) gegevens bedoeld waarbij de variabelen onderling positief correleren. Het hoofdstuk begint met een demonstratie van de rol van rotatie-vrijheid bij PCA van dergelijke gegevens. Na rotatie ontstaan er componenten die contrastvrije componenten genoemd kunnen worden, omdat de matrix met correlaties tussen de variabelen en de componenten per component geen teken-contrast bevat. Vervolgens wordt met een voorbeeld geïllustreerd dat contrastvrije componenten qua interpretatie fraaier zijn dan contrast-componenten. De definitie van contrast-componenten bij PCA wordt gegeneraliseerd naar PARAFAC. Van deze definitie wordt gebruik gemaakt om aan te tonen dat PARAFAC contrast–componenten kan leveren. Dit roept de vraag op of er contrastvrije PARAFAC–componenten gevonden kunnen worden waarmee de gegevens even goed gefit kunnen worden als met de PARAFAC–componenten. Er wordt een algoritme afgeleid om het PARAFAC model met contrast–vrije componenten aan de gegevens te fitten. Met deze gerestricteerde variant van PARAFAC wordt geïllustreerd, dat de fit van PARAFAC met contrastvrije componenten vrijwel even goed kan zijn als die van PARAFAC. In een dergelijk geval wordt de voorkeur gegeven aan de contrastvrije componenten vanwege de fraaiere interpretatie.

Hoofdstuk 5 heeft als doel het bepalen van PARAFAC–componenten die overeenkomen met niet–overlappende clusters van variabelen. In sociaal–wetenschappelijk onderzoek ontstaan regelmatig confirmatieve en exploratieve onderzoeksvragen met betrekking tot dergelijke clusters. De PARAFAC–componenten komen niet noodzakelijk overeen met niet–overlappende clusters van variabelen. Derhalve worden twee gerestricteerde PARAFAC–methoden geïntroduceerd. Bij de eerste methode wordt ervan uitgegaan dat de onderzoeker een hypothese heeft omtrent de manier waarop de variabelen in niet–overlappende clusters zijn ingedeeld. Deze hypothese wordt vertaald in restricties op de patroonmatrix van PARAFAC. Er wordt een algoritme afgeleid om deze gerestricteerde PARAFAC–variant aan de gegevens te fitten en er worden voorwaarden gegeven waaraan de componenten moeten voldoen om te kunnen spreken van een bevestiging van de hypothese (confirmatie). Bij de tweede methode wordt ervan uitgegaan dat de onderzoeker wil weten of de variabelen zinvol in niet–overlappende clusters zijn op te delen. Deze exploratieve vraag wordt beantwoord door de restrictie aan de patroonmatrix van PARAFAC op te leggen, dat elke variable slechts door middel van één component wordt samengevat. Er wordt een algoritme afgeleid om deze gerestricteerde PARAFAC-variant aan de gegevens te fitten. Beide methoden worden geïllustreerd aan de hand van analyses van empirische gegevens.

In hoofdstuk 6 wordt een schematisch overzicht gepresenteerd van de gerestricteerde PARAFAC-methoden die in dit onderzoek besproken zijn. Er wordt geconcludeerd dat, in geval van zwakke uniciteit, de voorgestelde gerestricteerde PARAFAC-methoden een nuttige aanvulling zijn op de (ongerestricteerde) PARAFAC-methode, mede omdat ze stabielere componenten plegen te leveren.

THE ANALYSIS OF THREE-WAY ARRAYS BY CONSTRAINED PARAFAC METHODS

In three-way data, scores are available from a number of cases on a number of variables at a number of occasions. Among the earliest methods for the exploratory analysis of three-way data is a model and technique called PARAFAC, which was initiated by Richard Harshman, and which can be seen as a generalization of Principal Components Analysis. Some properties of Principal Components Analysis are retained under this generalization, while others are lost. PARAFAC yields a decomposition of the three-way data array into three component matrices: one for the persons, one for the variables, and one for the occasions.
In this book – a companion volume to Kroonenberg's *Three-mode Principal Component Analysis* that also appeared in the DSWO Press M&T series – much attention is directed to the most salient property of the PARAFAC model, the uniqueness of its components. It turns out that the theoretical property of uniqueness does not exclude the existence of alternative representations that fit the data almost as well as the standard PARAFAC solution. In such cases, the uniqueness is called weak. A constrained PARAFAC variant is developed specifically to determine the degree of uniqueness of PARAFAC components. For data having weak uniqueness, three new constrained PARAFAC methods are introduced that allow for easier interpretations. The first one employs constant relative importances of the components across occasions, the second one determines contrast-free components, and the third one determines components that correspond to non-overlapping clusters of variables. The usefulness of the constrained PARAFAC methods is illustrated by various analyses of empirical three-way data, and by simulations.

DSWO PRESS

ISBN 90-9005493-6