

SHORT REVIEW**A GENERAL APPROACH TO ANALYSIS OF THREE-MODE DATA**

W. N. WICKREMASINGHE

*Department of Statistics, Faculty of Science, University of Colombo, P. O. Box 1490, Colombo.**(Received: 25 January 2005 ; accepted: 27 April 2005)*

Abstract: Three-mode data pertain to measurements related to three entities or "modes". For example, measurements of a number of objects on a number of variables at several different occasions, would make up a three-mode data set. Such data frequently occur in fields such as Agriculture, Biology, Chemistry, Education, Medicine, and Psychology. The primary objective of any analysis of such data should be to examine the interaction between these three factors simultaneously and interpret results based on the presence or absence of this interaction. In the past, several approaches were employed to analyze such data but none looked at the interaction between the three modes simultaneously. It was not until after 1980's that appropriate general methods were available. This paper reviews exploratory as well as confirmatory approaches to the analysis of three-mode data. The most general approach that studies three-way interaction simultaneously, called Tucker3 analysis, is mainly focused on. Two illustrations are also provided and potential for future research is highlighted.

Key words: Three-mode analysis, Three-mode principal components analysis, Three-way data analysis, Tucker3 analysis

INTRODUCTION

Three-mode data pertain to measurements related to three entities or "modes". For example, measurements of a number of objects on a number of variables at several different occasions or conditions, would make up a "three-mode" data set. Such data frequently occur in fields such as Agriculture, Chemistry, Education, Medicine, and Psychology. Examples also exist in other fields, but may not be common. The objective of any analysis of such data should be to examine the interaction between these three modes (factors) simultaneously and then to interpret results based on the absence or the presence of three-factor interaction. When the number of modes becomes higher, the analysis too becomes complicated. On the other hand, four or higher order interactions are not so common in data analysis. Thus in the past, the focus has been mainly on three-way data. The objective of this paper is to review the methods of analysis of three-mode data that are three-way generalizations of the Principal Components Analysis (PCA). Focus will be mainly

on the most general approach called Tucker3 analysis that helps one to simultaneously study the interaction between all three modes (factors).

PROBLEMS WITH THE USUAL ANOVA APPROACH

Three-way data can be classified into two categories: one is data that already exist, for example, mortality rates from 17 cancer sites in 43 different countries over a 9-year period;¹ the other is data from designed experiments such as three-factor factorial. If the experiment is replicated, then the analysis is straight forward.^{2,3} However, problems arise when it is non-replicated, *i.e.* when it has one observation per cell. The usual three-way analysis of variance (ANOVA) can not be used directly in this case, as the experimental error is mixed up with the three-factor interaction (if any). Common practice in the past has been to ignore the three-factor interaction and to use it as error.⁴ This approach fails in the presence of real three-factor interaction. In a typical three-mode analysis, a three-factor factorial model without three-factor interaction term is fitted to data and the residual from this fitting is decomposed using an appropriate model with the view to separate three-factor interaction from random error.¹

The analysis of three-mode data is two-fold: one is exploratory, where no attempt is made to come to conclusions based on statistical tests of hypotheses; the other is confirmatory, which uses probabilistic ideas and statistical theory.

EXPLORATORY METHODS*Common tools*

For exploratory analysis, two methods are commonly used.⁵ Both methods are three-way

generalizations of the principal components analysis (PCA). One is CANDECAMP/PARAFAC (CP) analysis where CANDECAMP⁶ stands for CANonical DECOMposition, a standard multivariate technique⁷ and PARAFAC^{8, 9, 10} stands for PARALLEL FACTOR analysis. CANDECAMP and PARAFAC are supposed to employ the same model and same fitting approach,⁵ and thus considered as one method, CP. The second is the most general method. It was proposed by Tucker^{11, 12} and commonly known as Tucker's three-mode analysis, or simply, Tucker3 analysis. The focus of this review is mainly on Tucker3 analysis.

Let $\{y_{ijk}; i = 1, 2, \dots, \ell; j = 1, 2, \dots, m; k = 1, 2, \dots, n\}$ denote a three-way data array, where i, j , and k represent the three modes respectively. The array $\{y_{ijk}\}$ can be thought of as original observations or residuals from fitting a three-factor-interaction-free model to three-way data. In the CP approach, y_{ijk} is modeled as

$$y_{ijk} = \sum_{r=1}^u g_{ir} h_{jr} e_{kr} + \varepsilon_{ijk} \quad (1)$$

where g_{ir}, h_{jr} , and e_{kr} denote elements of "components matrices" G, H , and E of order $\ell \times u, m \times u$, and $n \times u$, respectively; ε_{ijk} denotes a random error term. Model (1) is fitted to data by minimizing $\sum_{ijk} \varepsilon_{ijk}^2$ over G, H and E by means of alternating least squares algorithm.^{6, 8} One can easily see that when $\ell = 1$, or $m = 1$, or $n = 1$ i.e. for ordinary two-way data, model (1) reduces to the usual PCA model. Kiers⁵ give details of the CP model and its relationship to PCA; Harshman and Lundy⁹, and Harshman¹³ discuss uniqueness properties of CP solution.

Tucker's model^{11, 12} for three-way data is given by

$$y_{ijk} = \sum_{p=1}^s \sum_{q=1}^t \sum_{r=1}^u g_{ip} h_{jq} e_{kr} c_{pqr} + \varepsilon_{ijk} \quad (2)$$

which is an extended version of the CP model. In CP model (1) same number of components, u , is used in each mode, where as in model (2) different numbers s, t , and u can be used in different modes. Moreover, in CP each component in mode G is related to precisely one component in mode H and one in mode E , where as in Tucker's each component is related to every component of both

other modes.⁵ Further, the relations between the s components of mode G , the t components of mode H , and the u components of mode E are captured by the $s \times t \times u$ "core" array denoted by C . In matrix notation, (2) can be expressed as

$$Y_g = GC_g(E' \otimes H') + E_g \quad (3)$$

where Y_g is an $\ell \times mn$ matrix version of the $\ell \times m \times n$ data array; G, H , and E (with elements g_{ip}, h_{jq} , and e_{kr} respectively) are the component matrices of order $\ell \times s, m \times t$, and $n \times u$; C_g is the $s \times t \times u$ matrix version of the $s \times t \times u$ core array; E_g is an $\ell \times mn$ matrix of random errors, and \otimes stands for the matrix direct product. Two more equivalent expressions of (3) exist and they are

$$Y_h = HC_h(G' \otimes E') + E_h \quad (4)$$

and

$$Y_e = EC_e(H' \otimes G') + E_e \quad (5)$$

with corresponding definitions. However, one needs to use only (3), (4), or (5) when analyzing three-way data, as the final result is invariant to the version of the model expression.⁴

Tucker's solution and non-uniqueness

In contrast to the fact that CP solution is uniquely determined, a solution obtained by Tucker3 analysis is not unique.⁵ Any arbitrary non-singular transformations of all component matrices simultaneously, do not affect the model representation and thus gives another solution. For example, consider model (3) above.

Let $(\hat{G}, \hat{H}, \hat{E}, \hat{C}_g)$ be a solution such that $\hat{Y}_g = \hat{G}\hat{C}_g(\hat{E}' \otimes \hat{H}')$ gives a good approximation to Y_g . Consider non-singular matrices $S_{s \times s}, T_{t \times t}$, and $U_{u \times u}$. Then, $(\tilde{G}, \tilde{H}, \tilde{E}, \tilde{C}_g)$ is also a solution, where $\tilde{G} = \hat{G}S$; $\tilde{H} = \hat{H}T$; $\tilde{E} = \hat{E}U$, and $\tilde{C}_g = S^{-1}\hat{C}_g((U')^{-1} \otimes (T')^{-1})$.⁵ This is true, as $(\hat{G}S)S^{-1}\hat{C}_g((U')^{-1} \otimes (T')^{-1})(\hat{E}U)' \otimes (\hat{H}T)'$
 $= \hat{G}\hat{C}_g((U')^{-1} \otimes (T')^{-1})(U\hat{E}' \otimes T\hat{H}')$

$$\begin{aligned}
 &= \hat{G}\hat{C}_g((U')^{-1}U'\hat{E}' \otimes (T')^{-1}T'\hat{H}') \\
 &= \hat{G}\hat{C}_g(\hat{E}' \otimes \hat{H}') = \hat{Y}_g.
 \end{aligned}$$

Number of components

One key step in Tucker3 analysis is the determination of the number of components in each mode that give an adequate fit. Timmerman and Kiers¹⁴ provided a method called DIFFIT. However, since DIFFIT is a systematic approach based on a large number of models with all sensible combinations of number of components, it is found to be time-consuming. An approximate but fast method that compared well with DIFFIT was later proposed by Kiers and der Kinderen¹⁵ for choosing the number of components. Experience has shown us that $s = t = u = 1$ gives adequate fits for most moderate sized data sets where as in many other cases it is usually $s \leq 2, t \leq 2$, and $u \leq 2$.

General analysis

Once the number of components have been chosen, one can use the software called "3WAY PACK" specifically developed for this purpose by Kroonenberg¹⁶ for fitting Tucker3 model. The steps of analysis are detailed below:

Consider models (2) and (3) above. In (2), the coefficients c_{pqr} function as weights for the combination of component of G, H , and E . If a particular c_{pqr} , say c_{221} , is zero, then the 2nd, 2nd, and 1st combination of components of G, H and E respectively, does not contribute towards the estimation of the data based on the model. For example, assume $s = 2, t = 2$ and $u = 1$. Then, if $c_{221} = 0$, $(g_2 \otimes h_2 \otimes e_1)c_{221}$ does not contribute to the estimation of y , where

$$G = (g_1 \ g_2) = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ \dots & \dots \\ \dots & \dots \\ g_{s1} & g_{s2} \end{bmatrix}; H = (h_1 \ h_2) = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ \dots & \dots \\ \dots & \dots \\ h_{m1} & h_{m2} \end{bmatrix}; \text{ and } E = e_1 = \begin{bmatrix} e_{11} \\ e_{21} \\ \dots \\ \dots \\ e_{n1} \end{bmatrix}.$$

In general, non-singular transformations of the component matrices and corresponding transformation of the core, commonly known as orthogonal rotations, are carried out to obtain a simple structure of the core, C , to enhance interpretability.¹⁷ However, rotation of the core to get a simple interpretable structure, may

sometimes destroy the simplicity of the component matrices.¹⁸ Hence, a method called joint orthomax rotation¹⁸ that sees a compromise between the simplicity of both the core and the component matrices, has been proposed.

In Tucker3 analysis, usually the unrotated solution of model (3) is such that the component matrix G consists of unit normalized eigenvectors of $\hat{Y}_g\hat{Y}_g'$; H consists of unit normalized eigenvectors of $\hat{Y}_h\hat{Y}_h'$; and E consists of unit normalized eigenvectors of $\hat{Y}_e\hat{Y}_e'$. This unrotated solution is sometimes called the principal axes solution.^{19,5} Interpretation of the results is mainly based on the core array and the component matrices. This can be informal, or formal, depending on the type of analysis used. These will be illustrated later in this paper.

Graphical displays

Kiers⁵ discusses graphical display of three-mode data (i) based on individual modes, and (ii) based on combinations of two modes, commonly known as joint biplots²⁰ or joint plots. In case (i), entities of each mode is considered as points in a high-dimensional space with low-dimensional approximations given by Tucker3 or CP solutions. In Tucker3 for example, ℓ rows of \hat{Y}_g give coordinates for ℓ points in an s -dimensional subspace of \mathfrak{R}^{mn} . This approximated configuration is displayed by finding a basis of the s -dimensional subspace⁵ as follows:

1. Re-write model (3) as $\hat{Y}_g = \hat{G}\hat{F}'$, say, where $\hat{F} = (\hat{E} \otimes \hat{H})\hat{C}'$.
2. Use the method of approximate biplot^{21, 22} as for a two-way matrix.

An ordinary matrix of rank 2 can be exactly represented by a biplot. When the rank is higher than 2, if the matrix can be satisfactorily approximated by a rank 2 matrix, then the biplot of the latter gives an adequate representation of the former. A systematic way to find a rank 2 approximation is through singular value decomposition (SVD).²¹

For the principal axes solution mentioned in the previous section with $s = 2$, \hat{F} is columnwise orthogonal and when the two columns of \hat{F} are divided by the square roots of the corresponding eigenvalues of $\hat{Y}_g\hat{Y}_g'$ it becomes orthonormal. So, \hat{G} is rescaled by multiplying the two columns by

the square roots of the associated eigenvalues. This rescaled \hat{G} is said to contain "principal coordinates".⁵ One can plot these principal coordinates in each mode separately, to get an adequate representation of the data matrix.

The above procedure is for Tucker3 model. The same procedure works for u -dimensional CP model by simply replacing \hat{C}'_g by I'_g , where I_g is the matrix version of I which is a $u \times u \times u$ array which has unit elements in positions (i, i, i) , $i = 1, 2, \dots, u$, and zeros elsewhere. Kiers⁵ discusses ways to display data in CP by means of biplot approach while keeping the "unique axes" feature.

In case (ii), combinations of any two modes are plotted, called joint plots, at each component of the third mode. For example, to plot all combinations of G and E mode entities, it suffices to use co-ordinates of H . Joint plots are discussed by Kiers,⁵ Murakami and Kroonenberg,²⁰ and See and Smith.²³ These will be further discussed in the illustration.

A list of references on several other aspects of the three-mode analysis based on exploratory approach, could be found in Kroonenberg's.²⁴ For example, computational,^{25, 26} clustering in the absence or presence of missing values^{27, 28}, and mathematical,^{12, 29} are some of these aspects. However, majority of Kroonenberg's²⁴ bibliography deals with direct applications.

CONFIRMATORY METHODS

Attempts to test statistical hypotheses in three-way data with one observation per cell are relatively few. Most of these are in the context of designed experiments; three-factor factorial, in particular. The most common model used is Tucker3 for the reasons discussed earlier.

General model

Consider the general model for nonadditive three-way data,⁴

$$x_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \tau_{ij} + \rho_{ik} + \phi_{jk} + \pi_{ijk} + \varepsilon_{ijk}^* \quad (7)$$

for $i = 1, 2, \dots, \ell$; $j = 1, 2, \dots, m$; and $k = 1, 2, \dots, n$, where it is assumed that all parameters except the random error term are fixed and subject to usual sum-to-

zero restrictions. Subscripts i, j , and k represent the three modes, as earlier; and ε_{ijk}^* 's are assumed to be i.i.d. $N(0, \sigma^2)$. Since model (7) is saturated, conventionally it has been assumed that $\pi_{ijk} = 0$. Let the residuals from the model with $\pi_{ijk} = 0$ be denoted by y_{ijk} . When x_{ijk} 's and y_{ijk} 's are presented as ℓmn -vectors x and y respectively, we get

$$y = E_0 x \quad (8)$$

where $E_0 = (I_\ell - \frac{1}{\ell} J_\ell) \otimes (I_m - \frac{1}{m} J_m) \otimes (I_n - \frac{1}{n} J_n)$ under the assumptions of model (7). It can be easily seen that E_0 is symmetric and idempotent.

Simulation of residuals

Now consider Tucker3 model, i.e. model (2) above. Once the number of components, and are determined, one can re-write model (2) as

$$y = \hat{y} + \varepsilon \quad (9)$$

for the residuals using ℓmn -vectors for corresponding parts. This gives $V(y) = V(\varepsilon)$,

but from (8), we get $V(y) = E_0 V(x) E_0' = E_0 (\sigma^2 I) E_0' = \sigma^2 E_0 E_0' = \sigma^2 E_0$, since E_0 is symmetric and idempotent. Therefore, the ℓmn -vector of errors, ε , in (9) is i.i.d. $N(0, \sigma^2 E_0)$. By taking $\sigma^2 = 1$ without loss of generality, this result can be used to simulate the residual array y for three-mode data with known levels ℓ, m , and n .³⁰

Testing for three-way interaction

Marasinghe and Boik³¹ discuss a special case of model (2) with $s = t = u = 1$ and review a one-degree-of-freedom score test and a likelihood ratio test for the three-way interaction π_{ijk} and propose a three-degree-of-freedom score test for π_{ijk} as follows; Under the assumption of $s = t = u = 1$, model (2) reduces to

$$y_{ijk} = cg_i h_j e_k + \varepsilon_{ijk}, \quad (10)$$

where π_{ijk} is modeled as $cg_i h_j e_k$ with corresponding vectors g, h , and e summing to zero and having unit norm. Maximum likelihood estimates of parameters of model (10) are derived³¹ using alternating least squares approach.³² Marasinghe and Boik³¹ proposed a likelihood

ratio test to test for π_{ijk} by testing $H_0 : c = 0$. The test is to reject H_0 for large values of $\hat{c}^2 / y'y$. Boik³³ provided critical values for this test. Boik and Marasinghe³⁴ also proposed an approximation to the above test for which \hat{c}^2 is easier to compute than in the usual case with alternating least squares algorithm. Marasinghe and Boik³¹ also proposed an estimate of σ^2 in the presence of three-way interaction. However, no mention has been made on how to use it for tests of hypotheses.

Finding sub-areas free of three-way interaction

Suppose the test proposed by Marasinghe and Boik³¹ leads to rejection of $H_0 : c = 0$. Then, a question one would raise is whether one can find a sub-area free of three-way interaction in the assumed form. Wickremasinghe and Johnson⁴ proposed a likelihood ratio test for testing $H_0 : Sg = 0, Th = 0, \text{ and } Ue = 0$, where $S_{q_0 \times \ell}, T_{r_0 \times m}$, and $U_{s_0 \times n}$ are contrast matrices of rank $q_0 (< \ell - 1), r_0 (< m - 1)$, and $s_0 (< n - 1)$, respectively. If not rejected, the test would help identify a sub-area free of three-way

interaction. The test is to reject H_0 for small values of $\Lambda = \frac{y'y - \hat{c}^2}{y'y - \hat{c}^{***2}}$, where \hat{c}^2 is the maximum likelihood estimate of c^2 as given by Marasinghe and Boik³¹ and \hat{c}^{***2} is the square of the largest three-mode singular value of $Y^* = (I - S^-S)Y((I - T^-T) \otimes (I - U^-U))$, where $Y_{\ell \times m \times n}$ is the matrix version of the lmn -vector y ; ¹⁹ S^-, T^- , and U^- denote the Moore-Penrose g-inverses of the corresponding matrices, respectively. Wickremasinghe³⁵ provides critical points based on an approximation for the above test for limited cases. Joint plot^{5, 20, 23} may be considered as an "exploratory" alternative of the likelihood ratio test approach above.⁴ However, this needs to be examined.

ILLUSTRATION

1. Exploratory

Kroonenberg *et al.*¹ discuss an interesting application of three-mode analysis in the field of Medicine: Usually in large comparative studies

Table 1: Time component of the three-mode analysis

Year	1968/69	1970/71	1972/73	1974/75	1976/77	1978/79	1980/81	1982/83	1984/85
Component 1	-0.52	-0.42	-0.27	-0.11	0.02	0.17	0.26	0.37	0.49

Table 2: Sizes of interaction of stomach cancer and lung cancer with country

Country	Stomach	Lung
Guatemala	7.7	-2.8
USA Whites	4.3	7.9
Canada	3.1	7.4
USA Non-white	2.1	7.0
Scotland	1.8	8.0
Denmark	0.5	7.2
Greece	2.7	-4.0
Cuba	3.1	-4.0
Spain	-1.9	-4.2
Bulgaria	-1.4	-5.1
Hungary	-4.1	-1.0
Costa Rica	-4.4	-1.6
Austria	-4.5	-0.5
Japan	-8.2	-0.4
Chile	-13.0	-1.2

of cancer mortality rates, comparative conclusions are drawn by visual inspection, and results are presented by cancer site or by country. This makes comparisons between cancer sites, or between countries, difficult. They showed how countries, cancer sites, and trends can be analyzed simultaneously, using exploratory three-way techniques. A three-mode data set consisting of 17 tumour sites at each of 43 countries for a period of 9 years was analyzed. Kroonenberg *et al.*¹ approximated the three-mode data set with a $2 \times 2 \times 1$ Tucker3 model which explained 51% of the total variation in data. Even though this application is presented here for illustrative purposes, it's doubtful whether the explanation (*i.e.* 51%) given by the "core" is adequate. In cases like this, we see the importance of methodology to test the significance of the magnitude of the three-way interaction. An attempt has been made by Kroonenberg *et al.*¹ to present results initially in an ANOVA, as for a non-replicated three-factor experiment. However, the F-ratios given are not exact, but pseudo values, for reasons given at the beginning of this paper. More work based on statistical approach is needed for conclusions in this context. Only summary results from the three-way approach are given here. Table 1 shows the coefficients of the single time component given by the Tucker3 analysis.

Note that there is an increasing time trend in Table 1. The coefficients of the time components contain information on the changes over years in the relationships between cancer sites and countries. In particular, an increasing linear trend indicates that the specific relationships are linearly increasing or decreasing over time; a positive value in the joint plot for sites and countries correspond to an increase, and a negative value, a decrease, over time.¹ In making a joint plot, Kroonenberg *et al.*¹ included only countries and sites with at least one inner product larger than 4.0. Table 2 shows sizes of interaction of stomach cancer and lung cancer with selected countries.

When entries of Table 2 are multiplied by corresponding elements of Table 1, one gets an approximation of the three-way interaction.^{1,23} By observing the interactions in Table 2 (or the corresponding joint plot which is not given here) one can say that mortality due to stomach cancer in Chile and Japan were more rapidly decreasing

than in other countries. A similar but weaker trend was observed for Austria, Hungary, and Costa Rica. This was in addition to the general decreasing trend for stomach cancer. Further, Guatemala's mortality rate due to stomach cancer was increasing. The three-way interaction also showed that mortality due to lung cancer was increasing in Scotland, USA, Canada, and Denmark.¹

2. Confirmatory

Wickremasinghe and Johnson⁴ analyzed data on hardness of gold fillings by 3 dentists, using each of 3 methods, for each of 4 types. They used Tucker3 model with $s = t = u = 1$. This model with single multiplicative term explained 92% of the residual variation. Marasinghe and Boik's³¹ method further confirmed that three-factor interaction in the assumed form is significant. The estimates of the parameters of the Tucker3 model using alternating least squares algorithm were,

$$\hat{c}' = -5.5269; \hat{g}' = (-.0706 \quad .8363 \quad -.4172 \quad -.3485);$$

$$\hat{h}' = (-.2758 \quad .8035 \quad -.5276);$$

It was found that the hypothesis of the form $H_0 : Sg = 0, Th = 0,$ and $Ue = 0,$ for

$$S = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}, T = [1 \quad 0 \quad -1]$$

and $U = [0 \quad 1 \quad -1]$ was not rejected at 5%

level. This means that the sub-area containing all combinations of (Type1, Type3, Type4) \times (Method1, Method3) \times (Dentist2, Dentist3) is free of three-way interaction. Thus, there is only one observation, *i.e.* the value corresponding to the combination (Type2, Method2, Dentist1), that causes three-way interaction in the data. After removing this observation, one could go ahead and analyze the usual way. Details could be found in Wickremasinghe and Johnson.⁴

FUTURE WORK

This review mainly focuses on Tucker's three-mode PCA, as this is the most general approach to analysis of three-mode data with one observation per cell. A majority of the work listed in Kroonenberg's²⁴ bibliography, is concerned with direct applications of Tucker3 analysis in an exploratory context. Very little has been done on

confirmatory type analyses. For example, most current methods deal with single multiplicative term explaining three-way interaction in non-replicated three-factor experiments. This is a special case of Tucker3 model with $s = t = u = 1$. Statistical methods are needed for cases beyond this. However, experience has shown that usually a good approximation to three-mode data can be achieved with a core of $2 \times 2 \times 2$ or less, in most cases. Exact statistical methods are needed to identify the correct model. Sometimes, three-way interaction, when present, is caused by just a few observations. When these are identified and removed, the resulting sub-area is free of three-way interaction. Methods based on exact use of alternating least squares approach, are needed for this process. Further, it would be worthwhile to investigate the relationship, if any, between the exploratory joint plot approach and the likelihood ratio approach, for determining the correct model for a sub-area of the three-mode data.

References

1. Kroonenberg P. M., Murakami T., & Coebergh J. W. W. (2002). Added value of three-way methods for the analysis of mortality trends illustrated with worldwide female cancer mortality. *Statistical Methods in Medical Research* **11**: 275-292.
2. Cochran W. G. & Cox G. M. (1957). *Experimental Designs*. Second Edition, New York, Wiley.
3. Kuehl R. O. (2000). *Design of Experiments: Statistical Principles of Research Design and Analysis*. Second Edition, Duxbury Press.
4. Wickremasinghe W. N. & Johnson D. E. (2002). Testing subhypotheses and estimating σ^2 in the nonreplicated three-way multiplicative interaction model. *Communications in Statistics - Simulations* **31**(4): 605-618.
5. Kiers H. A. L. (2000). Some procedures for displaying results from three-way methods. *Journal of Chemometrics* **14**: 151-170.
6. Carrol 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.
7. Johnson D. E. (1998). *Applied Multivariate Methods for Data Analysts*. Duxbury Press.
8. Harshman R. A. (1970). Foundations of the PARAFAC procedure: models and conditions for an 'explanatory' multimode factor analysis. *UCLA Working Papers Phonet* **16**: 1-84.
9. Harshman R. A. & Lundy M. E. (1984). The PARAFAC model for three-way factor analysis and multidimensional scaling. In: *Research Methods for Multimode Data Analysis*, (Eds. H. G. Law, C. W. Snyder, J. A. Hattie & R. P.R. McDonald). Praeger: New York. 122-215.
10. Harshman R. A. & Lundy M. E. (1994). PARAFAC: Parallel factor analysis. *Computational Statistics and Data Analysis* **18**: 39-72.
11. Tucker L. R. (1963). Implications of factor analysis of three-way matrices for measurement of change. In: *Problems in Measuring Change*, (Ed. C. W. Harris) University of Wisconsin Press.
12. Tucker L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika* **31**: 279-311.
13. Harshman R. A. (1972). Determination and proof of minimum uniqueness conditions for PARAFAC1. *UCLA Working Papers Phonet* **22**: 111-117.
14. Timmerman M. E. & Kiers H. A. L. (2000). Three-mode principal components analysis: choosing the numbers of components and sensitivity to local optima. *British Journal of Mathematical and Statistical Psychology* **53**: 1-16.

15. Kiers H. A. L. & der Kinderen A. (2003). A fast method for choosing the numbers of components in Tucker3 analysis. *British Journal of Mathematical and Statistical Psychology* **56**: 119-125.
16. Kroonenberg P. M. (2003). *Applications Guide: 3Way Pack*. Department of Education, Leiden University, The Netherlands.
17. Kiers H. A. L. (1997). Three-mode orthomax rotation. *Psychometrika* **62**: 579-598.
18. Kiers H. A. L. (1998). Joint orthomax rotation of the core and component matrices resulting from three-mode principal components analysis. *Journal of Classification* **15**: 245-263.
19. Kroonenberg P. M. (1983). *Three Mode Principal Component Analysis: Theory and Applications*. DSWO Press, Leiden.
20. Murakami T. & Kroonenberg P. M. (2003). Three-mode models and individual differences in semantic differential data. *Multivariate Behavioral Research* **38**(2): 247-283.
21. Gabriel K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika* **58**(3): 453-467.
22. Bradu D. & Gabriel K. R. (1978). The biplot as a diagnostic tool for models of two-way tables. *Technometrics* **20**(1): 47-67.
23. See K. & Smith E. P. (1996). A graphical tool for three-factor experiments with one observation per cell. *Communications in Statistics -Simulation* **25** (3): 709-732.
24. Kroonenberg P. M. (2004). <http://www.leidenuniv.nl/fsw/three-mode/>.
25. Anderson C. A. & Bro R. (1998). Improving the speed of multi-way algorithms: Part I. Tucker3. *Chemometrics and Intelligent Laboratory Systems* **42**: 93-103.
26. Alsberg B. K. & Kvalheim O. M. (1994). Compression of three-mode data arrays by B-splines prior to three-mode principal component analysis. *Chemometrics and Intelligent Laboratory Systems* **23**: 29-38.
27. Basford K. E. (1985). The mixture method of clustering applied to three-way data. *Journal of Classification* **2**: 109-125.
28. Hunt L. A. & Basford K. E. (2001). Fitting a mixture model to three-mode three-way data with missing information. *Journal of Classification* **18**: 209-226.
29. Ten Berge J. M. F. & Kiers H. A. L. (1999). Simplicity of core arrays in three-way principal component analysis and the typical rank of $p \times q \times 2$ arrays. *Linear Algebra and Its Applications* **294**: 169-179.
30. Wickremasinghe W. N. & Johnson D. E. (2004). Comparison of three-mode principal components of several non-replicated three-way experiments using a graphical approach. (unpublished manuscript).
31. Marasinghe M. G. & Boik R. J. (1993). A three-degree of freedom test of additivity in three-way classifications. *Computational Statistics and Data Analysis* **16**: 47-61.
32. Kroonenberg P. M. & De Leeuw J. (1980). Principal components analysis of three-mode data by means of alternating least squares algorithm. *Psychometrika*. **65**: 69-97.
33. Boik R. J. (1990). A likelihood ratio test for three-mode singular values: upper percentiles and an application to three-way ANOVA. *Computational Statistics and Data Analysis* **10**: 1-9.
34. Boik R. J. & Marasinghe M. G. (1989). Analysis of nonadditive multi-way classifications. *Journal of American Statistical Association* **84**: 1059-1064.
35. Wickremasinghe W. N. (1995). Testing subhypotheses in the nonreplicated three-way multiplicative interaction model. *Ph. D. Thesis*, Kansas State University, Manhattan, Kansas.