Journal of Classification 19:161-178 (2002)
DOI: 10.1007/s00357-001-0037-7

Distance Models for Three-Way Tables and Three-Way Association

Mark de Rooij

Leiden University

Abstract: In the present paper we study distance models for the analysis of three-way contingency tables. Specifically, we will study three-way association under these models measured by the second order odds ratio. Two kinds of distance models will be studied: (a) Models for three-way tables where each way is treated on an equal footing; (b) Models for multiple two-way tables, where one of the three ways has a special importance. For the first kind of models, called triadic distance models, we will show that there exists a natural conjugacy between the Exponential$-p$ similarity function, the L_p-transform and the Minkowski$-p$ distance. For triadic distance models defined by the L_p-transform we will prove that they do not model three-way association. Moreover, triadic distance models defined by the L_p-transform are restricted multiple dyadic distances, where each dyadic distance is defined for a two-way margin of the three-way table. Distance models for three-way two-mode data, called three-way distance models, do succeed in modeling three-way association.

Keywords: Second order odds ratio; Triadic distance; Weighted Euclidean distance; Three-way two-mode data; INDSCAL.

Netherlands Organization for Scientific Research (NWO) is gratefully acknowledged for funding this project. This research was conducted while the author was supported by a grant of the Foundation for Behavioural and Educational Sciences of this organization (575-30-006).

Author's Address: Mark de Rooij, Department of Psychology, University of Leiden, PO Box 9555, 2300 RB Leiden, THE NETHERLANDS; e-mail: rooijm@fsw.leidenuniv.nl.

1. Introduction

The analysis of three-way tables has received much attention in the last few decades. Most models for two-way data analysis can readily be adapted to three-way data. For example, the log-linear model for three-way tables is well defined. Once we get into the field of scalar products and distances, however, things get a little more complicated. Joly and Le Calvé (1995, p. 192), for example, write

> "In the two-way case, matrix analysis and the theory of vector spaces provide well adapted tools. Therefore, it seems natural to generalize the notions of scalar products and distances to three-way tables (and to N-way tables). Indeed, the generalization of the scalar product defined as a trilinear map is immediate, but unfortunately it leads to many difficulties because of the theoretical complexity of the tensor product of order 3 (see Franc 1989; Denis and Dhorn, 1989). However, the generalization of the notion of distance turns out to be straightforward and very useful, at least in our opinion."

In the present paper we will show that generalizations of distance models, called triadic distance models, also have their limitations. Triadic distances are distances defined on triples of points. We will focus on one specific family of triadic distance models formed by the L_p-transform (see Section 3.1). These are triadic distances defined on dyadic distances. We will show that triadic distance models based on the L_p-transform do not model three-way association but only two-way marginal association.

Other distance models for three-way data, where the distances are not defined on triples of points, will also be discussed. More specifically, the INDSCAL-model (Carroll and Chang 1970), and the model proposed by Okada and Imaizumi (1997) for the analysis of three-way two-mode asymmetric proximity data are studied. As we will see, these models *do* model three-way association.

We will focus on models for the analysis of three-way contingency tables. More specifically, we study models for two kinds of three-way tables: (a) three-way tables where each way is treated equally; (b) multiple two-way contingency tables, i.e., two-way tables obtained for different cohorts, different samples, or different time-points. In this latter kind of tables, one of the three ways has a special importance, whereas the first kind of tables each way is equally important. For two-way contingency tables, De Rooij and Heiser (2000b) presented distance-association models. In the present paper these distance-association models will be generalized to three-way tables. We will discuss Minkowski distances and unfolding distances. It is important to note that the Minkowski distances and three-way generalizations thereof can only be applied to square tables, i.e., tables where each way corresponds to the same variable.

To study association in contingency tables we first need to define association. We follow Rudas (1998, p. 9): "...one qualitative definition of association is that it is the information in the joint distribution (i.e., in the contingency table) not contained in the marginal distributions". A measure of association congruent with this definition is the odds ratio. Therefore we will use the odds ratio, the conditional odds ratio, and the second order odds ratio (also called the ratio of odds ratios). An advantage of the odds ratio over other measures of association is its *variation independence* (Rudas 1998, p. 10); i.e., the odds ratio can vary independently from the marginals.

The odds ratio ($\theta_{ii'jj'}$) for a two-way contingency table is defined as

$$\theta_{ii'jj'} = \frac{\pi_{ij}\pi_{i'j'}}{\pi_{i'j}\pi_{ij'}}, \tag{1}$$

where π_{ij} is the expected probability of an observation in cell ij for $i = 1, \ldots, I$, $j = 1, \ldots, J$ under a specified model. The conditional odds ratio ($\theta_{ii'jj'|k}$, $k = 1, \ldots, K$) for a three-way table is equal to the odds ratio given a value on the third way. In mathematical terms,

$$\theta_{ii'jj'|k} = \frac{\pi_{ijk}\pi_{i'j'k}}{\pi_{i'jk}\pi_{ij'k}}. \tag{2}$$

The conditional odds ratio is a measure of two-way association for a slice of the three-way table. A measure of three-way association is the second order odds ratio, also called ratio of odds ratios ($\theta_{ii'jj'kk'}$); i.e.,

$$\theta_{ii'jj'kk'} = \frac{\pi_{ijk}\pi_{i'j'k}\pi_{i'jk'}\pi_{ij'k'}}{\pi_{i'j'k'}\pi_{i'jk}\pi_{ij'k}\pi_{ijk'}}, \tag{3}$$

which is equal to the ratio of conditional odds ratios for the slices k and k'; that is,

$$\theta_{ii'jj'kk'} = \frac{\theta_{ii'jj'|k}}{\theta_{ii'jj'|k'}}. \tag{4}$$

There is no three-way association when the second order odds ratio is equal to one, i.e., $\theta_{ii'jj'kk'} = 1$, or in other words, when the conditional odds ratio is independent of k. Below we will often use the logarithm of the second order odds ratio, which equals zero when there is no three-way association.

2. Models for Three-way Tables and Distance Restrictions

Basic models for the analysis of contingency tables are the log-linear models. We will present a brief discussion of these models. Before doing so, we will present the same models in a multiplicative framework. We end this section with a discussion of the transformation of association parameters of a multiplicative model into distances.

2.1. Multiplicative Models for the Cell Probabilities

The saturated multiplicative model for a three-way contingency table can be written as

$$\pi_{ijk} = \mu \alpha_i^R \alpha_j^C \alpha_k^P \eta_{ij}^{RC} \eta_{ik}^{RP} \eta_{jk}^{CP} \eta_{ijk}^{RCP}. \qquad (5)$$

The α parameters denote main effects of the variables; the η parameters denote association between variables. The term μ is a general constant. The saturated model is uninteresting in itself because it always fits perfectly and does not give any reduction of data. To obtain more interesting models it is possible to restrict sets of parameters. Usual restrictions are: (a) setting some association parameters equal to a prespecified value (1 for no association, another constant for uniform association); (b) restricting a set of association parameters to be of reduced rank. In the present paper we will constrain the association parameters to fulfill the metric axioms. First however, we will show the equivalent of model (5) in log-linear form.

2.2. The Log-linear Parameterization

A more familiar parameterization of model (5) is the log-linear model, which is obtained by taking the natural logarithm on both sides. We then obtain

$$\log(\pi_{ijk}) = \lambda + \lambda_i^R + \lambda_j^C + \lambda_k^P + \lambda_{ij}^{RC} + \lambda_{ik}^{RP} + \lambda_{jk}^{CP} + \lambda_{ijk}^{RCP}, \qquad (6)$$

where $\log(\mu) = \lambda$, $\log(\alpha_i^R) = \lambda_i^R$, etc.. Many authors have studied these models; we refer to Bishop, Fienberg, and Holland (1975), Fienberg (1980), and Agresti (1990).

Here we will introduce more notation. Let us rewrite (6) as

$$\log(\pi_{ijk}) = \lambda_{ijk}^I + \lambda_{ijk}^{II}, \qquad (7)$$

where $\lambda_{ijk}^I = \lambda + \lambda_i^R + \lambda_j^C + \lambda_k^P$, the terms that will not be transformed to distances but will be kept in the model as such; λ_{ijk}^{II} the set of (two-way and three-way) association terms that will be transformed to distances. Before every transformation we have to define the set λ_{ijk}^{II}.

2.3. Transformation of Association Parameters

The parameters for both the log-linear model and the multiplicative model are not always easily interpretable. Transformation of the parameters into a distance model may enhance interpretability: a small distance corresponds to a large association, i.e., a larger number than can be expected on basis of the marginal parameters (i.e., the set λ^{I}_{ijk}); a large distance corresponds to low association, i.e., a smaller number than can be expected from the marginal parameters. To attain this goal, a monotone decreasing function of the multiplicative association parameters should be used. A family of transformations is given by the exponential$-p$ similarity function; that is

$$\eta = \exp(-d^p), \tag{8}$$

for $p \geq 1$, d is a distance satisfying the metric axioms. We do not use subscripts here, because we will below apply this transformation to both two- and three-way association parameters. Two special cases deserve attention: The exponential decay function with $p = 1$ and the Gaussian transformation with $p = 2$. These transformation were proposed by, among others, Shepard (1957, 1987) and Nosofsky (1985). Taking the natural logarithm on both sides the transformation can be written as

$$\lambda = -d^p. \tag{9}$$

3. Triadic Distance Models

In the present section we will introduce triadic distance models. Such models have been proposed by Hayashi (1972), Cox, Cox, and Branco (1991), Pan and Harris (1991), Joly and Le Calvé (1995), Daws (1996), Heiser and Bennani (1997), and De Rooij and Heiser (2000a).

Joly and Le Calvé (1995) and Heiser and Bennani (1997) both gave an axiomatic framework for the study of triadic distances. Within these frameworks, a number of triadic distance models have been proposed. One family of such models is formed by the L_p-transform, which will be discussed in the next section. De Rooij and Heiser (2000a) presented triadic distance models for the analysis of three-way asymmetric proximities. The models are generalizations of the L_2-model, in the same sense that multidimensional unfolding models are generalizations of multidimensional scaling models (Heiser 1981, 1987). In the remainder of this paper our discussion will be limited to the family of triadic distance models formed by the L_p-transform and the generalizations proposed by De Rooij and Heiser (2000a).

3.1. The L_p-transform

Triadic distances via the L_p-transform are based on dyadic distances. The family of triadic distances formed by the L_p-transform is defined as

$$d_{ijk}^p = d_{ij}^p + d_{ik}^p + d_{jk}^p, \tag{10}$$

for $p \geq 1$, where d_{ijk} denotes a triadic distance, and d_{ij} a dyadic distance. Heiser and Bennani (1997) showed this family of triadic distances satisfies their axioms, if the dyadic distances satisfy the metric axioms of minimality, positivity, symmetry, and the triangle inequality. Two specific models received attention by Cox, Cox, and Branco (1991), Joly and Le Calvé (1995), and Heiser and Bennani (1997), that is, the perimeter model ($p = 1$) and the generalized Euclidean model ($p = 2$). We will focus on the generalized Euclidean model, but our conclusions generalize to other triadic distance models defined by the L_p-transform.

To enhance interpretation of the triadic distance model, and specifically the generalized Euclidean model, the reader is referred to De Rooij and Heiser (2000a), who provide an extensive discussion of the generalized Euclidean model with examples.

3.2. Natural Conjugacy between Exponential$-p$ Similarity Function, the L_p-transform and Minkowski$-p$ Distances

The most familiar distances are the Minkowski$-p$ distances, defined by

$$d_{ij}(\mathbf{X}) = (\sum_{m=1}^{M} |x_{im} - x_{jm}|^p)^{1/p}, \tag{11}$$

where $p \geq 1$ is the Minkowski parameter, $\mathbf{X}$ is the $I \times M$ matrix with coordinate values, and M is the dimensionality. The Minkowski distance with $p = 1$ is known as the city-block distance, the Minkowski distance with $p = 2$ is known as the Euclidean distance, and is probably the best known and most used distance. Another interesting distance is the dominance metric, for which $p = \infty$. The dominance distance equals the absolute difference between the two points on whichever dimension they are furthest apart.

De Rooij and Heiser (2000a) noted that for $p = 1, 2$ the triadic distance formed by the L_p-transform together with the Minkowski$-p$ distance is proportional to a natural measure of dispersion. For $p = 1$ the triadic distance is proportional to the sum of ranges over dimensions, which in one dimension is just the range; for $p = 2$ the triadic distance is proportional to the square root of the inertia of the three points considered, which in one dimension reduces to the standard deviation.

Here we generalize that observation by including the exponential$-p$ similarity function and other Minkowski metrics. Combining (8), (10), and (11) we have

$$\begin{aligned} -\lambda_{ijk}^{II} &= d_{ijk}^{p}(\mathbf{X}) \\ &= d_{ij}^{p}(\mathbf{X}) + d_{ik}^{p}(\mathbf{X}) + d_{jk}^{p}(\mathbf{X}) \\ &= \sum_m |x_{im} - x_{jm}|^p + \sum_m |x_{im} - x_{km}|^p + \sum_m |x_{jm} - x_{km}|^p. \end{aligned} \tag{12}$$

Using the exponential$-p$ similarity function, with the L_p-transform and the Minkowski$-p$ distance, we obtain additivity over dimensions, which is mathematically attractive. Special attention is needed here for the case $p = \infty$. The L_∞-transform selects the largest dyadic distance, and as noted above, the Minkowski$-\infty$ distance is equal to the distance between the two points on whichever dimension they are farthest apart. The triadic distance reduces to this distance, i.e.,

$$\begin{aligned} d_{ijk} &= \max(d_{ij}, d_{ik}, d_{jk}) \\ &= \max\left(\max_m |x_{im} - x_{jm}|, \max_m |x_{im} - x_{km}|, \max_m |x_{jm} - x_{km}|\right). \end{aligned} \tag{13}$$

In the remainder of this paper, we will use the conjunction Gaussian transformation, generalized Euclidean model, and Euclidean distances (all $p = 2$). However, *our conclusions generalize to other values of* $p \geq 1$.

3.3. Three-way Association

We will transform all two- and three-way association terms of model (6) to a triadic distance. That is, $\lambda_{ijk}^{II} = \lambda_{ij}^{RC} + \lambda_{ik}^{RP} + \lambda_{jk}^{CP} + \lambda_{ijk}^{RCP}$. Our model then becomes

$$\begin{aligned} \log(\pi_{ijk}) &= \lambda_{ijk}^{I} - d_{ijk}^2 \\ &= \lambda + \lambda_i^R + \lambda_j^C + \lambda_k^P - d_{ijk}^2, \end{aligned} \tag{14}$$

where d_{ijk} is defined as in (10) with $p = 2$.

Proposition 1: *The model as defined in (14) with d_{ijk} defined by the L_2-transform does not model three-way association.*

Proof: The log of the second order odds ratio under this model is equal to

$$\begin{aligned} \log(\theta_{ii'jj'kk'}) &= -\, d_{ijk}^2 - d_{i'j'k}^2 - d_{i'jk'}^2 - d_{ij'k'}^2 \\ &\quad + d_{i'j'k'}^2 + d_{i'jk}^2 + d_{ij'k}^2 + d_{ijk'}^2. \end{aligned} \tag{15}$$

Using the definition of the generalized Euclidean model to decompose the triadic distance we obtain

$$\begin{aligned}\log(\theta_{ii'jj'kk'}) &= -d^2_{ij} - d^2_{ik} - d^2_{jk} - d^2_{i'j'} - d^2_{i'k} - d^2_{j'k} \\ &\quad - d^2_{i'j} - d^2_{i'k'} - d^2_{jk'} - d^2_{ij'} - d^2_{ik'} - d^2_{j'k'} \\ &\quad + d^2_{i'j'} + d^2_{i'k'} + d^2_{j'k'} + d^2_{i'j} + d^2_{i'k} + d^2_{jk} \\ &\quad + d^2_{ij'} + d^2_{ik} + d^2_{j'k} + d^2_{ij} + d^2_{ik'} + d^2_{jk'}, \end{aligned} \quad (16)$$

from which all terms drop out; so we obtain

$$\log(\theta_{ii'jj'kk'}) = 0. \quad (17)$$

□

Conclusion: With triadic distance models we are not modeling any three-way association, but only two-way marginal association. Triadic distance models are useful, but they do not model three-way association. They do give us useful information on the two-way marginal association, as can be seen from the conditional odds-ratio. The log of the conditional odds ratio given k can be written as

$$\begin{aligned}\log(\theta_{ii'jj'|k}) &= -d^2_{ijk} - d^2_{i'j'k} + d^2_{i'jk} + d^2_{ij'k} \\ &= -d^2_{ij} - d^2_{i'j'} + d^2_{i'j} + d^2_{ij'}. \end{aligned} \quad (18)$$

The latter expression does not depend on k, indicating again that no three-way association is modeled.

3.4. Triadic versus Dyadic Models

Until now we did not distinguish between Minkowski distances and unfolding distances. Here we will make that distinction. Minkowski distances are based on one coordinate matrix $\mathbf{X}$ and have been defined earlier by (11). For the Euclidean metric the distance is

$$d_{ij}(\mathbf{X}) = \left[\sum_m (x_{im} - x_{jm})^2\right]^{1/2}. \quad (19)$$

We will call this expression the Euclidean case of the Minkowski distance or simply distance; it is symmetric because $d_{ij}(\mathbf{X}) = d_{ji}(\mathbf{X})$. A generalization of the Minkowski distance is the unfolding distance, based on *two* coordinate matrices $\mathbf{X}$ and $\mathbf{Y}$. The unfolding distance in the Euclidean metric is defined as

$$d_{ij}(\mathbf{X};\mathbf{Y}) = \left[\sum_m (x_{im} - y_{jm})^2\right]^{1/2}. \quad (20)$$

We will call this quantity the unfolding distance, and when we use it to represent association in a square table, the association is asymmetric because in general $d_{ij}(\mathbf{X};\mathbf{Y}) \neq d_{ji}(\mathbf{X};\mathbf{Y})$. The unfolding distance is equal to the corresponding Minkowski distance if and only if $\mathbf{X} = \mathbf{Y}$.

This distinction between the Minkowski and unfolding distance can be generalized to triadic distances, as was done by De Rooij and Heiser (2000a) for the generalized Euclidean model. The distance model is written

$$d_{ijk}(\mathbf{X}) = \left[d_{ij}^2(\mathbf{X}) + d_{ik}^2(\mathbf{X}) + d_{jk}^2(\mathbf{X})\right]^{1/2}. \tag{21}$$

The adaptation made by De Rooij and Heiser is to replace every dyadic distance in (21) by a dyadic unfolding distance (20). We then obtain a triadic unfolding distance

$$d_{ijk}(\mathbf{X};\mathbf{Y};\mathbf{Z}) = \left[d_{ij}^2(\mathbf{X};\mathbf{Y}) + d_{ik}^2(\mathbf{X};\mathbf{Z}) + d_{jk}^2(\mathbf{Y};\mathbf{Z})\right]^{1/2}, \tag{22}$$

which reduces to the triadic distance model if and only if $\mathbf{X} = \mathbf{Y} = \mathbf{Z}$.

Analogous to this distinction, we can define two triadic association models, both like (14). The first will be called Triadic Distance Association Model (TDAM) and is obtained by using the triadic distance defined by (21):

$$\log(\pi_{ijk}) \quad = \quad \lambda_{ijk}^I - d_{ijk}^2(\mathbf{X}). \tag{23}$$

The second is called Triadic Unfolding Association Model (TUAM) and is obtained by using the triadic unfolding distance (22):

$$\log(\pi_{ijk}) \quad = \quad \lambda_{ijk}^I - d_{ijk}^2(\mathbf{X};\mathbf{Y};\mathbf{Z}). \tag{24}$$

For both models (23) and (24), Proposition 1 holds. Because TDAM and TUAM do not model three-way association, a comparison to models based on dyadic distances is of interest.

Let us return to (7), and let $\lambda_{ijk}^{II} = \lambda_{ij}^{RC} + \lambda_{ik}^{RP} + \lambda_{jk}^{CP}$. Here we set $\lambda_{ijk}^{RCP} = 0$; that is, we do not model three-way association. Each two-way association will be transformed separately to a dyadic distance, $\lambda_{ij}^{RC} = -d_{ij}^2$, $\lambda_{ik}^{RP} = -d_{ik}^2$, and $\lambda_{jk}^{CP} = -d_{jk}^2$. Again, we can distinguish between Euclidean Minkowski and unfolding distances ((19) and (20), respectively). First consider the Dyadic Distance Association Model (DDAM), in which every two-way association is transformed separately to a Euclidean distance; that is

$$\log(\pi_{ijk}) \quad = \quad \lambda_{ijk}^I - d_{ij}^2(\mathbf{X}) - d_{ik}^2(\mathbf{Y}) - d_{jk}^2(\mathbf{Z}). \tag{25}$$

Each association term is based on a different coordinate matrix ($\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$), but all associations are symmetric. Secondly, consider the case in which every association term is transformed to an unfolding distance:

$$\log(\pi_{ijk}) \quad = \quad \lambda_{ijk}^I - d_{ij}^2(\mathbf{X}^A;\mathbf{X}^B) - d_{ik}^2(\mathbf{Y}^A;\mathbf{Y}^B) - d_{jk}^2(\mathbf{Z}^A;\mathbf{Z}^B), \tag{26}$$

where superscripts A and B are used to avoid a proliferation of matrix symbols, but still to distinguish between the two coordinate matrices of an unfolding distance. We will call this model the Dyadic Unfolding Association Model (DUAM). Of course it is possible to transform some association terms with a Euclidean Minkowski distance, and others by an Euclidean unfolding distance. We can thus make eight (including DDAM and DUAM) different dyadic association models, but we will not go into details here.

The triadic and dyadic unfolding association models can be used for any three-way contingency table. As an example, consider counts f_{ijk} of persons living in a specific region i (five categories: Aude, Gard, Herault, Lozere, and Pyrenees-orientale), having job j (nine categories: farmer, farm laborers, self-employed professionals, higher professionals, middle professionals, employees, workers, workers is services, and other categories) in year k (four categories: 1954, 1962, 1968, and 1975). This example is obtained from Bernard and Lavit (1985; see also Van der Heijden, 1987, p. 97). In general we feel that cross-classifications with many categories (more than three) are more suitable for analysis by distance models than, for example, with a $2 \times 2 \times 2$ cross-classification. The models based on Minkowski distances can only be used in the case where each way refers to the same variable, i.e., a three-way one-mode input matrix. For an example of such data, see De Rooij and Heiser (2000a) who analyzed a three-way transition table involving the political votes in Sweden, obtained over three consecutive elections (see Upton 1978, p. 128). In the case where we use the triadic unfolding models for three-way one-mode tables this practice leads to an asymmetric association pattern. Thus, the models based on unfolding distances are more generally applicable than the models based on Minkowski distances.

Comparing TDAM, TUAM, DDAM, and DUAM we find the following: If we place symmetry restrictions on DUAM, we obtain DDAM; if we place equality restrictions on DUAM (i.e., $\mathbf{X}^A = \mathbf{Y}^A$, $\mathbf{X}^B = \mathbf{Z}^A$, and $\mathbf{Y}^B = \mathbf{Z}^B$ in (26)), DUAM reduces to TUAM; If we place symmetry restrictions on TUAM, we obtain TDAM; if we place equality restrictions on DDAM (i.e., $\mathbf{X} = \mathbf{Y} = \mathbf{Z}$ in (25)), we obtain TDAM. This comparison is depicted in Figure 1. We can conclude that triadic association models are equal to dyadic association models with equality restrictions.

3.5. Conclusions, Part 1

In this section we considered triadic distance association models. More specifically, we covered triadic association models for contingency tables, i.e., multiplicative models in which association terms are transformed to triadic distances. Transformations from multiplicative parameters to distances were

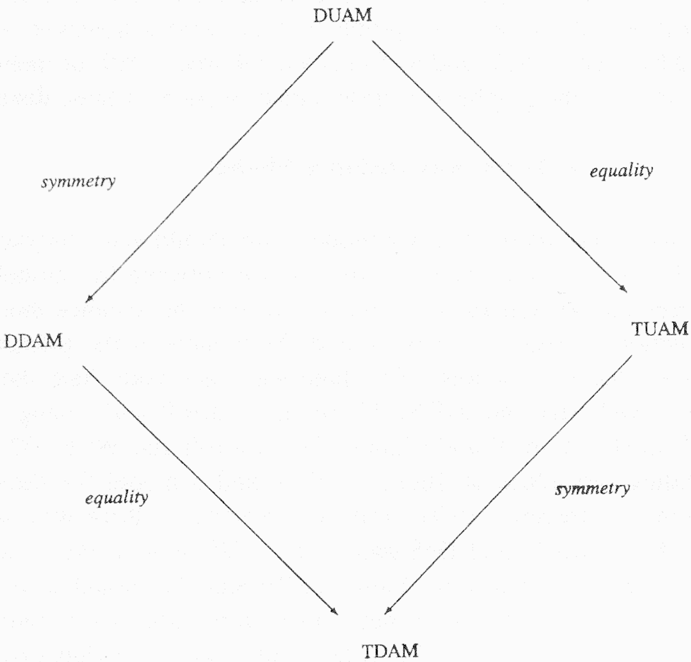

Figure 1: Relationships between different association models through symmetry and equality restrictions on the coordinate matrices.

performed using the exponential$-p$ similarity function. A natural conjugacy was found for the exponential$-p$ similarity function, the L_p-transform, and the Minkowski$-p$ metric.

Using one of the conjunctions ($p = 2$), we found that triadic association models do not model three-way association according to the second order odds ratio. Again, please recall that this conclusion can be generalized to other values of $p \geq 1$. Further expanding this result, we found that triadic association models are equal to dyadic association models with equality restrictions.

Let us return for the moment to the definition of the triadic unfolding distance (22):

$$\begin{aligned} d^2_{ijk}&(\mathbf{X};\mathbf{Y};\mathbf{Z}) \\ &= d^2_{ij}(\mathbf{X};\mathbf{Y}) + d^2_{ik}(\mathbf{X};\mathbf{Z}) + d^2_{jk}(\mathbf{Y};\mathbf{Z}) \\ &= 2\sum_m \left(x^2_{im} + y^2_{jm} + z^2_{km} - x_{im}y_{jm} - x_{im}z_{km} - y_{jm}z_{km}\right). \qquad (27) \end{aligned}$$

This distance formulation does not contain a trilinear term ($x_{im}y_{jm}z_{km}$). One way to model three-way association is to include such a trilinear a term in the definition of a triadic distance. Two problems then remain however: (a) the problems noted by Franc (1989) and by Denis and Dhorne (1989) as mentioned in the Introduction; (b) the graphical representation of such a triadic distance.

4. Three-way Distance Models

Having discussed triadic distance models, we should also analyze other distance models for three-way data. To distinguish between the models discussed in the previous section and here, we will denote the distance models in the previous section by triadic distances, and the models in the present section as three-way distance models. For three-way two-mode data there are many distance models like the INDSCAL model (Carroll and Chang 1970), the IDIOSCAL model (Carroll and Chang 1972; Carroll and Wish 1974), the three-mode scaling procedure of Tucker (1972), and a model for three-way city-block scaling by Heiser (1989). For an overview of three-way scaling models see Arabie, Carroll, and DeSarbo (1987). We would like to discuss the INDSCAL model (Carroll and Chang, 1970) and the model devised by Okada and Imaizumi (1997) for asymmetric proximity data. These two models lead to other generalizations of the distance-association models discussed by De Rooij and Heiser (2000b). These three-way distance models are different from the triadic distance models: no distance is specified between three points, but dyadic distances are specified, which are weighted for the third way. Triadic distance models are most useful for three-way one-mode data, or three-way three-mode data. The INDSCAL model and Okada and Imaizumi's (1997) model are most useful for three-way two-mode data, that is, for example, occupational mobility data at different points in time, or in different countries. To study three-way association, we will apply the Gaussian transformation to $\lambda_{ijk}^{II} = \lambda_{ij}^{RC} + \lambda_{ik}^{RP} + \lambda_{jk}^{CP} + \lambda_{ijk}^{RCP}$, that is

$$\log(\pi_{ijk}) \quad = \quad \lambda + \lambda_i^R + \lambda_j^C + \lambda_k^P - d_{ijk}^2. \tag{28}$$

We start our discussion with the INDSCAL model, and afterwards we discuss Okada and Imaizumi's model.

4.1. The INDSCAL Model

The INDSCAL model (Carroll and Chang, 1970) in squared form is written as

$$d_{ijk}^2 = \sum_m w_{km}(x_{im} - x_{jm})^2. \tag{29}$$

Consider now (14) with d^2_{ijk} defined by (29). Developing the second order odds ratio (3) we obtain

$$\begin{aligned}
\log(\theta_{ii'jj'kk'}) &= -(\sum_m w_{km}x^2_{im} + w_{km}x^2_{jm} - 2w_{km}x_{im}x_{jm}) \\
&\quad -(\sum_m w_{k'm}x^2_{i'm} + w_{k'm}x^2_{jm} - 2w_{k'm}x_{i'm}x_{jm}) \\
&\quad -(\sum_m w_{k'm}x^2_{im} + w_{k'm}x^2_{j'm} - 2w_{k'm}x_{im}x_{j'm}) \\
&\quad -(\sum_m w_{km}x^2_{i'm} + w_{km}x^2_{j'm} - 2w_{km}x_{i'm}x_{j'm}) \\
&\quad +(\sum_m w_{k'm}x^2_{im} + w_{k'm}x^2_{jm} - 2w_{k'm}x_{im}x_{jm}) \\
&\quad +(\sum_m w_{km}x^2_{i'm} + w_{km}x^2_{jm} - 2w_{km}x_{i'm}x_{jm}) \\
&\quad +(\sum_m w_{km}x^2_{im} + w_{km}x^2_{j'm} - 2w_{km}x_{im}x_{j'm}) \\
&\quad +(\sum_m w_{k'm}x^2_{i'm} + w_{k'm}x^2_{j'm} - 2w_{k'm}x_{i'm}x_{j'm}), \quad (30)
\end{aligned}$$

which can be rewritten as

$$\begin{aligned}
\log(\theta_{ii'jj'kk'}) &= 2\sum_m (w_{km} - w_{k'm}) \times (x_{im}x_{jm} + x_{i'm}x_{j'm} - x_{i'm}x_{jm} \\
&\quad -x_{im}x_{j'm}) \\
&= 2\sum_m (w_{km} - w_{k'm}) \times (x_{im} - x_{i'm})(x_{jm} - x_{j'm}). \quad (31)
\end{aligned}$$

The second order odds ratio equals zero if and only if, for all m, either $w_{km} = w_{k'm}$, or $x_{im} = x_{i'm}$, or $x_{jm} = x_{j'm}$.

The conditional odds ratio, given a source k is

$$\begin{aligned}
\log(\theta_{ii'jj'|k}) &= \sum_m \Big[w_{km}(x_{i'm} - x_{jm})^2 + w_{km}(x_{im} - x_{j'm})^2 \\
&\quad -w_{km}(x_{im} - x_{jm})^2 - w_{km}(x_{i'm} - x_{j'm})^2\Big] \\
&= 2\sum_m w_{km}(x_{im}x_{jm} + x_{i'm}x_{j'm} - x_{i'm}x_{jm} - x_{im}x_{j'm}) \\
&= 2\sum_m w_{km}(x_{im} - x_{i'm})(x_{jm} - x_{j'm}). \quad (32)
\end{aligned}$$

This form is the same as derived by De Rooij and Heiser (2000b) for one two-way table (i.e., only one source) when we set the weight for each source k on each dimension equal to one (i.e., equal dimension weights). If we define

$y_{im}^k = w_{km}^{1/2} x_{im}$ and collect the y_{im}^k in a matrix $\mathbf{Y}^k$, the conditional odds ratio for slice k is given by

$$d_{i'j}^2(\mathbf{Y}^k) + d_{ij'}^2(\mathbf{Y}^k) - d_{ij}^2(\mathbf{Y}^k) - d_{i'j'}^2(\mathbf{Y}^k). \tag{33}$$

So, if we make a plot of the coordinates for slice k, we can obtain the conditional log odds ratio by the sum and difference of some squared distances. The conditional odds ratio given some i is computationally more tedious, but after some mathematical reformulation we obtain

$$\begin{aligned}\log(\theta_{jj'kk'|i}) &= \sum_m (w_{k'm} - w_{km}) \\ &\quad \times \left[(x_{im} - x_{jm})^2 - (x_{im} - x_{j'm})^2\right].\end{aligned} \tag{34}$$

Both conditional odds ratios are dependent on the index of the given slice.

4.2. The Okada and Imaizumi Model

The model of Okada and Imaizumi (1997) is defined as

$$\begin{aligned} d_{ijk} = t_{ijk} &- \left(\frac{t_{ijk}^2}{\sum_m \left(\frac{x_{im}-x_{jm}}{u_{km}}\right)^2} \right)^{1/2} \times r_i \\ &+ \left(\frac{t_{jik}^2}{\sum_m \left(\frac{x_{jm}-x_{im}}{u_{km}}\right)^2} \right)^{1/2} \times r_j, \end{aligned} \tag{35}$$

where

$$t_{ijk} = w_k d_{ij} = w_k \left(\sum_m (x_{im} - x_{jm})^2 \right)^{1/2}. \tag{36}$$

The model uses two kinds of weights to model differences on the third mode k, symmetry weights w_k, and asymmetry weights u_{km}. For a detailed discussion on the model, fitting the model to data, and interpreting the model, we refer to Okada and Imaizumi (1997). Taking the square, the model can be written as

$$\begin{aligned} d_{ijk}^2 &= w_k^2 d_{ij}^2 + \left(\frac{w_k^2 d_{ij}^2}{a_{ijk}} \right) \times (r_i^2 + r_j^2) \\ &\quad + 2 w_k d_{ij} \left[\left(\frac{w_k^2 d_{ij}^2}{a_{ijk}} \right)^{1/2} \times (r_i - r_j) \right] \\ &\quad - 2 \left(\frac{w_k^2 d_{ij}^2}{a_{ijk}} \right) \times r_i \times r_j, \end{aligned} \tag{37}$$

where $a_{ijk} = \sum_m \left(\frac{x_{jm}-x_{im}}{u_{km}}\right)^2 = \sum_m \left(\frac{x_{im}-x_{jm}}{u_{km}}\right)^2$.

The second order odds ratio (3), under this model is given by

$$\begin{aligned}
\log(\theta_{ii'jj'kk'}) = \quad & w_k^2 \times \quad \left\{ d_{i'j}^2 \left(1 + \frac{(r_{i'} - r_j)^2}{a_{i'jk}} + 2\frac{r_{i'} - r_j}{\sqrt{a_{i'jk}}}\right)\right. \\
& + d_{ij'}^2 \left(1 + \frac{(r_i - r_{j'})^2}{a_{ij'k}} + 2\frac{r_i - r_{j'}}{\sqrt{a_{ij'k}}}\right) \\
& - d_{ij}^2 \left(1 + \frac{(r_i - r_j)^2}{a_{ijk}} + 2\frac{r_i - r_j}{\sqrt{a_{ijk}}}\right) \\
& \left. - d_{i'j'}^2 \left(1 + \frac{(r_{i'} - r_{j'})^2}{a_{i'j'k}} + 2\frac{r_{i'} - r_{j'}}{\sqrt{a_{i'j'k}}}\right)\right\} \\
& - w_{k'}^2 \times \quad \left\{ d_{i'j}^2 \left(1 + \frac{(r_{i'} - r_j)^2}{a_{i'jk'}} + 2\frac{r_{i'} - r_j}{\sqrt{a_{i'jk'}}}\right)\right. \\
& + d_{ij'}^2 \left(1 + \frac{(r_i - r_{j'})^2}{a_{ij'k'}} + 2\frac{r_i - r_{j'}}{\sqrt{a_{ij'k'}}}\right) \\
& - d_{ij}^2 \left(1 + \frac{(r_i - r_j)^2}{a_{ijk'}} + 2\frac{r_i - r_j}{\sqrt{a_{ijk'}}}\right) \\
& \left. - d_{i'j'}^2 \left(1 + \frac{(r_{i'} - r_{j'})^2}{a_{i'j'k'}} + 2\frac{r_{i'} - r_{j'}}{\sqrt{a_{i'j'k'}}}\right)\right\} . \quad (38)
\end{aligned}$$

This second order odds ratio equals zero if and only if either $w_k = w_{k'}$ and $u_{km} = u_{k'm}$ for all m, or all distances d_{ij}, $d_{i'j}$, $d_{ij'}$ and $d_{i'j'}$ are zero. This latter condition is very rare, however. If the asymmetry weights are not equal, the second order odds ratio equals zero if $w_k = w_{k'} = 0$. The model proposed by Okada and Imaizumi obviously is mathematically rather complex, as is necessary to give a nice representation of symmetry and asymmetry by distances in Euclidean space.

The conditional odds ratio, for a source k, is given by the first four lines of (38). Both expressions for the second order odds ratio and the conditional odds ratio are rather complex. In fact, the expressions are just the sum and differences (as in (15) and the first line of (18)) of the original distance definitions, with no reduction but just a reshuffling of the terms. Because the expressions cannot be simplified, we can again plot a configuration for each slice. The conditional odds ratio, given a slice k, is then given by the sum and differences of the squared distances in this plot, just as for the INDSCAL model (see Equation 33).

4.3. Conclusions, Part 2

The three-way distance-association models discussed in this section successfully model three-way association, in contrast to the triadic distance-association models in the Section 3. Therefore, the models have an advantage in situations where one is specifically interested in this three-way association. Both distance models discussed in this Section include a trilinear term. The model of Okada and Imaizumi is a generalization of the INDSCAL model, including parts for the skew-symmetric part of the data. Setting all radii (r_i) equal to each other, we obtain the INDSCAL model with equal dimension weights for every slice k.

5. Discussion

We studied three-way association measured by the second order odds ratio under distance models for three-way data. An interesting negative result was found: triadic distance association models do not model three-way association, but only two-way marginal association. Three-way distance association models *do* model three-way association. The three-way distance models both contain trilinear terms, but the triadic distance models do not. If a model includes trilinear terms, we are sure that three-way association is modeled. We feel, however, that trilinear terms are not the only way to include three-way information in the model. For example, the log-linear model does not have trilinear terms, but still is capable of modeling three-way association.

Triadic distance models and three-way distance models are intended for the analysis of different kinds of data. The triadic distance association models are better understood having three-way three-mode data, or three-way one-mode data; the three-way distance association models discussed in Section 4 are better understood having multiple two-way data matrices, i.e., three-way two-mode data. An advantage of the triadic distance model over the three-way distance model is the representation of each of the three two-way margins by distances, whereas the three-way distance models only give a distance representation of one of the three two-way margins by distances.

Because the triadic distance models do not model three-way association but only marginal two-way association, we should not use these models when we are especially interested in three-way association. The triadic distance models can give useful information about the structure in the data, however. Examples are given in the papers by Heiser and Bennani (1997) and De Rooij and Heiser (2000a). We do not want to argue here that triadic distances should not be used. They definitely can be useful, but one should recognize the properties of such models.

The study of triadic distances should be continued, until triadic distance models are found that (a) model three-way association; (b) have a clear graphical representation; and (c) overcome the mathematical difficulties in fitting them.

References

AGRESTI, A., (1990), *Categorical Data analysis*, New York: Wiley.

ARABIE, P., CARROLL, J.D., and DESARBO, W.S. (1987), *Three-Way Scaling and Clustering*, Newbury Park, CA: Sage.

BERNARD, M. CL., and LAVIT, CH. (1985), "Présentation des Donnees," *Statistique et Analyse des Données*, 10, 11-15.

BISHOP, Y. M. M., FIENBERG, S. E., and HOLLAND, P. W. (1975), *Discrete Multivariate Analysis: Theory and Practice*, Cambridge, MA: MIT press.

CARROLL, J. D., and CHANG, J. J. (1970), "Analysis of Individual Differences in Multidimensional Scaling via an N-Way Generalization of 'Eckart-Young' Decomposition," *Psychometrika*, 35, 283-319.

CARROLL, J. D., and CHANG, J. J. (1972), "*IDIOSCAL (Individual Differences in Orientation SCALing): A Generalization of INDSCAL Allowing IDIOsyncratic Reference Systems as well as an Analytic Approximation to INDSCAL*," Paper presented at meeting of the Psychometric Society, Princeton.

CARROLL, J. D., and WISH, M. (1974). Models and Methods for Three-Way Multidimensional Scaling, in *Contemporary Developments in Mathematical Psychology*, Eds., D. H. Krantz, R. C. Atkinson, R. D. Luce, and P. Suppes, San Francisco: Freeman, pp. 57-105.

COX, T. F., COX, M. A., and BRANCO, J. A. (1991), "Multidimensional Scaling for *N*-Tuples," *British Journal of Mathematical and Statistical Psychology* , 44, 195-206.

DAWS, J. T., (1996), "The Analysis of Free-Sorting Data: Beyond Pairwise Cooccurences," *Journal of Classification*, 13, 57-80.

DE ROOIJ, M., and HEISER, W. J. (2000a), "Triadic Distance Models for the Analysis of Asymmetric Three-Way Proximity Data," *British Journal of Mathematical and Statistical Psychology* , 53, 99-119.

DE ROOIJ, M., and HEISER, W. J. (2000b), "Graphical Representations and Odds-Ratios in a Distance Association Model for the Analysis of Cross-Classified Data," Technical report, PRM-2000-01, Leiden University.

DENIS, J. B., and DHORNE, T. (1989), "Orthogonal Tensor Decompositions of 3-Way Tables," in *Multiway Data Analysis*, Eds., R. Coppi and S. Bolasco, Amsterdam: Elsevier, pp. 31-37.

FIENBERG, S. E. (1980), *The analysis of Cross-Classified Categorical Data*, Cambridge, MA: MIT press.

FRANC, A. (1989), "Multiway Arrays: Some Algebraic Remarks," in *Multiway Data Analysis*, Eds., R. Coppi and S. Bolasco, Amsterdam: Elsevier, pp. 19-29.

HAYASHI, C. (1972), "Two Dimensional Quantifications Based on a Measure of Dissimilarity Among Three Elements," *Annals of the Institute of Statistical Mathematics*, 25, 251-257.

HEISER, W. J., (1981), *Unfolding Analysis of Proximity Data*, Unpublished Doctoral Dissertation, Department of Data Theory, Leiden University.

HEISER, W. J., (1987), "Joint Ordination of Species and Sites: The Unfolding Technique," in *Developments in Numerical Ecology*, Eds., P. Legendre and L. Legendre, Berlin: Springer-Verlag, pp. 189-221.

HEISER, W. J. (1989), "The City-Block model for Three-Way Multidimensional Scaling," in *Multiway Data Analysis*, Eds., R. Coppi and S. Bolasco, Amsterdam: Elsevier, pp. 395-404.

HEISER, W. J., and BENNANI, M. (1997), "Triadic Distance Models: Axiomatization and Least Squares Representation," *Journal of Mathematical Psychology* , 41, 189-206.

JOLY, S., and LE CALVÉ, G. (1995), "Three-Way distances," *Journal of Classification*, 12, 191-205.

NOSOFSKY, R., (1985), "Overall Similarity and the Identification of Separable Dimension Stimuli: A Choice Model Analysis," *Perception & Psychophysics*, 38, 415-432.

OKADA, A., and IMAIZUMI, T. (1997), "Asymmetric Multidimensional Scaling of Two-Mode Three-Way Proximities," *Journal of Classification*, 14, 195-224.

PAN, G., and HARRIS, D. P. (1991), "A New Multidimensional Scaling Technique Based upon Association of Triple Objects Pijk and its Application to the Analysis of Geochemical Data," *Mathematical Geology*, 23, 861-886.

RUDAS, T., (1998), *Odds Ratios in the Analysis of Contingency Tables*, Newbury Park, CA: Sage.

SHEPARD, R. N., (1957), "Stimulus and Response Generalization: A Stochastic Model Relating Generalization to Distance in Psychological Space," *Psychometrika*, 22, 325-345.

SHEPARD, R. N., (1987), "Towards a Universal Law of Generalization for Psychological Science," *Science*, 237, 1317-1323.

TUCKER, L. R (1972), "Relations between Multidimensional Scaling and Three-Mode Factor Analysis ," *Psychometrika*, 37, 3-27.

UPTON, G. J. G., (1978), *The Analysis of Cross-Tabulated Data*, Chichester, U.K.: Wiley.

VAN DER HEIJDEN, P. G. M., (1987), *Correspondence Analysis of Longitudinal Data*, Leiden: DSWO press.