Journal of Classification 20:115-136 (2003)
DOI: 1007/s00357-003-0008-2

A Comparison of the Multidimensional Scaling of Triadic and Dyadic Distances

John C. Gower

The Open University, U.K.

Mark de Rooij

Leiden University, The Netherlands

Abstract: We examine the use of triadic distances as a basis for multidimensional scaling (MDS). The MDS of triadic distances (MDS3) and a conventional MDS of dyadic distances (MDS2) both give Euclidean representations. Our analysis suggests that MDS2 and MDS3 can be expected to give very similar results, and this is strongly supported by numerical examples. We have concentrated on the perimeter and generalized Euclidean models of triadic distances, both of which are linear transformations of dyadic distances and so might be suspected of explaining our findings; however an MDS3 of the nonlinear variance definition of triadic distance also closely approximated the MDS2 representation. An appendix gives some matrix results that we have found useful and also gives matrix representations and alternative derivations of some known properties of triadic distances.

Keywords: Multidimensional scaling; Triadic distance; Perimeter distance; Generalized Euclidean distance.

Authors' Addresses: John C. Gower, Department of Statistics, Walton Hall, The Open University, Milton Keynes, MK7 6AA, U.K.; Mark De Rooij, Department of Psychology, Leiden University, P.O.Box 9555, 2300 RB Leiden, The Netherlands.

1. Introduction

There has been recent interest in extending the idea of defining distance or dissimilarity between two objects to defining triadic distances between three objects labelled *i, j* and *k*. (Cox, Cox and Branco, 1991; Daws, 1996; De Rooij and Heiser, 2000; De Rooij, 2001; Heiser and Bennani, 1997; Hayashi, 1972; Joly and Le Calvé, 1995; Pan and Harris, 1991). Triadic distances δ_{ijk} are usually defined as functions of the pair-wise, or dyadic, distances δ_{ij}. Thus, $\delta_{ijk} = f(\delta_{ij}, \delta_{jk}, \delta_{ik})$. There have been many proposals for defining the function f(.) but here we confine our attention to two of the more popular:

$$\delta_{ijk} = \delta_{jk} + \delta_{ik} + \delta_{ij} \quad \text{(perimeter model)} \tag{1}$$

and

$$\delta^2_{ijk} = \delta^2_{jk} + \delta^2_{ik} + \delta^2_{ij} \quad \text{(generalized Euclidean model)}. \tag{2}$$

These are both symmetric functions, in the sense that it is immaterial in which order the suffices are presented, so they may be regarded as analogous to symmetric dissimilarity coefficients.

One way of analyzing a complete set of $M = n(n-1)(n-2)/6$ observed symmetric triads d_{ijk} among n objects, is to find a set of points in r dimensions that generate $m = n(n-1)/2$ pair-wise Euclidean distances δ_{ij} that form triadic distances $\delta_{ijk} = f(\delta_{ij}, \delta_{ik}, \delta_{jk})$ that minimize some criterion of goodness of fit. An obvious choice is a generalization of metric *stress* or *sstress*, commonly used in multidimensional scaling. Thus we may seek to minimize:

$$\text{Stress:} \quad \sum_{i<j<k}^{n} (d_{ijk} - \delta_{ijk})^2 \tag{3}$$

$$\text{or Sstress} \quad \sum_{i<j<k}^{n} (d^2_{ijk} - \delta^2_{ijk})^2 . \tag{4}$$

Here we investigate how such a form of multidimensional scaling (MDS3) may differ from the usual multidimensional scaling of dissimilarities (MDS2).

2. Analysis

It is convenient to vectorise the two-way array $\{\delta_{ij}\}$ and the three-way array $\{\delta_{ijk}\}$ by writing $\delta_2 = (\delta_{21}, \delta_{31}, \delta_{32}, \delta_{41}, \delta_{42}, \delta_{43}, \delta_{51}, \delta_{52}, \dots)'$ and $\delta_3 =$

$(\delta_{321}, \delta_{421}, \delta_{431}, \delta_{432}, \delta_{521}, \delta_{531}, \delta_{532}, \delta_{541}, \ldots)'$. Then (1) may be written as the linear transformation:

$$\delta_3 = \mathbf{C}\delta_2 \tag{5}$$

where $\mathbf{C}$ is a matrix with M rows and m columns. A similar result holds for (2), with squared values in the vectors δ_2 and δ_3. Note we do not include the diagonal plane values, that is, triadic distances for which a subscript repeats (e.g. δ_{112}). However, these might be included without materially changing the following results.

Thus, $\mathbf{C}$ is a simple indicator matrix with a unit wherever both column-labels are a subset of the row-labels. Thus, $\mathbf{C}$ has three units in every row, which pick out the *jk, ik, ij* pairs for each of the triadic distances given as the row-names. Table 1 shows the case $n = 6$ but the general form is evident.

Table 1. The matrix **C** for n = 6.

	d_{21}	d_{31}	d_{32}	d_{41}	d_{42}	d_{43}	d_{51}	d_{52}	d_{53}	d_{54}	d_{61}	d_{62}	d_{63}	d_{64}	d_{65}
d_{321}	1	1	1												
d_{421}	1			1	1										
d_{431}		1		1		1									
d_{432}			1		1	1									
d_{521}	1						1	1							
d_{531}		1					1		1						
d_{532}			1					1	1						
d_{541}				1			1			1					
d_{542}					1			1		1					
d_{543}						1			1	1					
d_{621}	1										1	1			
d_{631}		1									1		1		
d_{632}			1									1	1		
d_{641}				1							1			1	
d_{642}					1							1		1	
d_{643}						1							1	1	
d_{651}							1				1				1
d_{652}								1				1			1
d_{653}									1				1		1
d_{654}										1				1	1

2.1 Unrestricted MDS3

With δ_3 given by (5), the stress criterion may be written[1] $\|\mathbf{d}_3 - \mathbf{C}\delta_2\|$. Generally, δ_3 will be constrained in some way (often to be Euclidean distances in a specified number of dimensions) but initially we assume that there is no constraint. There are two possibilities: (i) $\mathbf{d}_3$ are observed triads, not usually satisfying the conditions for triadic distances (see the Appendix for necessary and sufficient conditions) (ii) $\mathbf{d}_3 = \mathbf{C}\mathbf{d}_2$, derived from observed dyadic distances $\mathbf{d}_2$.

First we examine case (i), observing that minimizing $\|\mathbf{d}_3 - \mathbf{C}\delta_2\|$ has the unconstrained ordinary least-squares estimate[2] of δ_2 given by:

$$\hat{\mathbf{d}}_2 = (\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\mathbf{d}_3 \tag{6}$$

for the dyadic distances, whence perimeter triadic distances may be estimated as:

$$\hat{\mathbf{d}}_3 = \mathbf{C}\hat{\mathbf{d}}_2 = \mathbf{C}(\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\,\mathbf{d}_3 \tag{7}$$

with a residual sum-of-squares:

$$S_U = \|\mathbf{d}_3 - \hat{\mathbf{d}}_3\| = \mathbf{d}_3'[\mathbf{I} - \mathbf{C}(\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}']\mathbf{d}_3. \tag{8}$$

The matrix $\mathbf{C}'\mathbf{C}$ is of order *m*, so in numerical work can be quite large and its inverse could be inaccurate. Fortunately, the inverse has a simple algebraic form (22) which, with other results given in the appendix, allow the unrestricted estimates (6), (7) and (8) to be calculated accurately and efficiently.

2.2 MDS3 with Euclidean Constraints

The restricted residual sum-of-squares is given by:

---

[1] In this paper $\|\mathbf{X}\|$ always denotes the squared L_2 –norm *trace*($\mathbf{X}'\mathbf{X}$), so we dispense with the superfix of the common notation $\|.\|^2$.

[2] Roman letters ($\mathbf{d}$) normally denote data and Greek letters parameters (δ) with estimates $\hat{\delta}$. Here $\hat{\mathbf{d}}_2$ is an estimate but uses a roman letter because, as we shall see below $\hat{\mathbf{d}}_2$ and $\hat{\mathbf{d}}_3$ may be treated as data and, indeed, in model (ii) may be synonyms for $\mathbf{d}_2$ and $\mathbf{d}_3$. $\hat{\delta}_2$ and $\hat{\delta}_3$ are reserved for constrained estimates.

$$S_R = \|\mathbf{d}_3 - \mathbf{C}\hat{\delta}_2\| = \|(\mathbf{d}_3 - \hat{\mathbf{d}}_3) + (\hat{\mathbf{d}}_3 - \mathbf{C}\hat{\delta}_2)\|, \qquad (9)$$

where $\hat{\delta}_2$ are dyadic distances chosen to minimize S_R, subject to desired constraints. Using (7) and (8), (9) may be expanded as:

$$S_R = \|\mathbf{d}_3 - \hat{\mathbf{d}}_3\| + \|\hat{\mathbf{d}}_3 - \mathbf{C}\hat{\delta}_2\| + 2\ \mathbf{d}_3'(\mathbf{I} - \mathbf{C}(\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}')\mathbf{C}(\hat{\mathbf{d}}_2 - \hat{\delta}_2). \qquad (10)$$

The final term of (10) vanishes, whence:

$$S_R = S_U + S_S \qquad (11)$$

where

$$S_S = \|\hat{\mathbf{d}}_3 - \mathbf{C}\hat{\delta}_2\|. \qquad (12)$$

The importance of (11) is that it allows us to examine the additional contribution to the residual sum-of-squares induced by the constraint on δ_2; this result may be used as part of an analysis of variance, discussed below. Equation (12) means that, without loss of generality, we may estimate the constrained dyadic distances from the unconstrained estimates of the triadic perimeter distances $\hat{\mathbf{d}}_3 = \mathbf{C}\hat{\mathbf{d}}_2$ rather than from $\mathbf{d}_3$ itself. Thus, we have reduced problem (i) to the form of problem (ii), discussed further below.

Suppose now that $\hat{\delta}_2$ minimizes $\|\mathbf{d}_3 - \mathbf{C}\delta_2\|$ under some constraint and consider setting $\delta_2 = \lambda\hat{\delta}_2$ for some scalar λ. This remains an admissible solution *provided $\lambda\hat{\delta}_2$ satisfies the constraints imposed*, a condition that is certainly satisfied when $\hat{\delta}_2$ are Euclidean (or any other Minkowski) distances, squared distances, metrics and ultrametrics. The residual sum-of-squares is:

$$\|\mathbf{d}_3 - \lambda\mathbf{C}\hat{\delta}_2\| = \mathbf{d}_3'\mathbf{d}_3 + \lambda^2\hat{\delta}_2'\mathbf{C}'\mathbf{C}\hat{\delta}_2 - 2\lambda\ \mathbf{d}_3'\mathbf{C}\hat{\delta}_2 \qquad (13)$$

which is minimized when

$$\lambda = \frac{\mathbf{d}_3'\mathbf{C}\hat{\delta}_2}{\hat{\delta}_2'\mathbf{C}'\mathbf{C}\hat{\delta}_2}.$$

We know that (13) is minimized when $\lambda = 1$; so $\hat{\delta}_2'\mathbf{C}'\mathbf{C}\hat{\delta}_2 = \mathbf{d}_3'\mathbf{C}\hat{\delta}_2$, which on substitution into (13) gives:

$$S_R = \|\mathbf{d}_3 - \mathbf{C}\hat{\delta}_2\| = \mathbf{d}_3'\mathbf{d}_3 - \hat{\delta}_2'\mathbf{C}'\mathbf{C}\hat{\delta}_2. \qquad (14)$$

The first term on the right-hand-side of (14) is the total sum-of-squares and the second term is the sum-of-squares of the fitted triadic distances. The result is valid for the wide variety of constraints mentioned above. Indeed, it may be generalized to show that if $\hat{\gamma}$, $\hat{\delta}$ minimize $||\mathbf{d}_3 - \gamma - \delta||$ under constraints that remain satisfied by $\mu\hat{\gamma} + \lambda\hat{\delta}$, for arbitrary scalars λ and μ, then $||\mathbf{d}_3 - \hat{\gamma} - \hat{\delta}|| = ||\mathbf{d}_3|| - ||\hat{\gamma} + \hat{\delta}||$. Potentially, this generalization allows our methodology to be extended to a wider class of problems.

Combining (8), (11) and (14) gives:

$$||\mathbf{d}_3|| = ||\mathbf{C}\hat{\delta}_2|| + ||\mathbf{d}_3 - \hat{\mathbf{d}}_3|| + ||\hat{\mathbf{d}}_3 - \mathbf{C}\hat{\delta}_2||$$

or, in words:
Total ss = Fitted ss + unconstrained residual ss + increase in residual ss induced by the constraint on δ_2. This is an orthogonal analysis of variance which may be represented in the tabular form:

	Sum-of-squares	r-dimensions	2-dimensions
Restricted Fit	$\|\|\mathbf{C}\hat{\delta}_2\|\|$	$\frac{1}{2}r(2n-r-1)$	$2n-3$
S_S	$\|\|\hat{\mathbf{d}}_3 - \mathbf{C}\hat{\delta}_2\|\|$	$m - \frac{1}{2}r(2n-r-1)$	$m-2n+3$
S_U	$\|\|\mathbf{d}_3 - \hat{\mathbf{d}}_3\|\|$	$M-m$	$M-m$
S_R	$\|\|\mathbf{d}_3 - \mathbf{C}\hat{\delta}_2\|\|$	$M - \frac{1}{2}r(2n-r-1)$	$M-2n+3$
Total	$\|\|\mathbf{d}_3\|\|$	M	M

The final two columns give the number of parameters estimated that are to be associated with each term when a distance model is fitted, the first such column being for an r-dimensional fit and the second for the most important case of $r = 2$. For the unrestricted fit these come from the usual least-squares "regression" fit, so may be regarded as degrees of freedom ($M - m$ d.f. for S_U and M for the Total sum-of-squares and m d.f. for the unrestricted fit obtained by summing the first two rows of the analysis of variance table). The remaining terms merely indicate the number of parameters fitted and should not be treated as d.f.; nevertheless, they may be useful in judging the relative contributions of the various models. S_S gives the increase in the residual sum-of-squares due to imposing the constraint, which when added to S_U gives S_R (see equation (11)).

In deriving the above results, we have only assumed (5) without specifying whether the elements of δ_2 and δ_3 are squared or not. Thus the results hold both for minimizing stress associated with the perimeter model (1), and minimizing sstress associated with the generalized Euclidean model (2). Note that the latter is not the same as specifying d_{ijk} through the generalized Euclidean model (which requires taking square-roots) and fitting δ_2 by minimizing stress, which is an approach we adopt in one of our examples.

2.3 Equivalence of MDS3 and a Weighted MDS2

Thus, even in case (i) it is easy to derive a best fitting set of perimeter triadic distances (7) and their generating dyadic distances (6) which may be treated as in case (ii). The unconstrained least-squares solution certainly need not yield Euclidean distances; $\hat{\mathbf{d}}_2$ may not even be positive. However, if the unconstrained fit is poor, or is inconsistent with distance models, there seems little point in moving on to consider constrained MDS solutions. Hence, in the following, we assume that $\mathbf{d}_3$ is a triadic distance. This assumption is exact for case (ii) and is valid for case (i) providing we substitute $\hat{\mathbf{d}}_2$ for, $\mathbf{d}_2$ as we have shown to be permissible and without loss of generality, provided $\hat{\mathbf{d}}_2$ is a good approximation to being Euclidean. When it is not Euclidean, we show in the example of section 3.2 how our assumptions may be validated by adding a constant to $\hat{\mathbf{d}}_2$. Thus we consider minimizing:

$$\|\mathbf{d}_3 - \delta_3\| = \|\mathbf{C}\mathbf{d}_2 - \mathbf{C}\delta_2\| = (\mathbf{d}_2 - \delta_2)'\mathbf{C}'\mathbf{C}(\mathbf{d}_2 - \delta_2) \qquad (15)$$

where $\mathbf{d}_2 = \hat{\mathbf{d}}_2$ for case (i) and is an observed set of dyadic values $\mathbf{d}_2$ in case (ii).

In (15), δ_2 may be constrained as deemed fit. Equation (15) shows that minimizing stress for triadic distances is the same as minimizing dyadic stress with weights $\mathbf{C}'\mathbf{C}$. A similar result applies to sstress in association with the generalized Euclidean model. We have two estimates of the dyadic distances: (i) $\hat{\delta}_{2T}$ obtained from minimizing triadic stress (15) and (ii) $\hat{\delta}_{2D}$ obtained from minimizing conventional stress, which is (15) with $\mathbf{C}'\mathbf{C}$ replaced by a unit matrix. We are concerned with how $\hat{\delta}_{2T}$ compares with $\hat{\delta}_{2D}$. Therefore, to understand the effect of the weighting in (15), we need to examine $\mathbf{C}'\mathbf{C}$ in detail. Table 2 shows $\mathbf{C}'\mathbf{C}$ for the case $n = 6$. Writing w_{ijkl} for the element of $\mathbf{C}'\mathbf{C}$ corresponding to the row labelled d_{ij} and column labelled d_{kl}, we see that:

Table 2. The matrix **C′C** for n = 6.

	d_{21}	d_{31}	d_{32}	d_{41}	d_{42}	d_{43}	d_{51}	d_{52}	d_{53}	d_{54}	d_{61}	d_{62}	d_{63}	d_{64}	d_{65}
d_{21}	4	1	1	1	1		1	1			1	1			
d_{31}	1	4	1	1		1	1		1		1		1		
d_{32}	1	1	4		1	1		1	1			1	1		
d_{41}	1	1		4	1	1	1			1	1			1	
d_{42}	1		1	1	4	1		1		1		1		1	
d_{43}		1	1	1	1	4			1	1			1	1	
d_{51}	1	1		1			4	1	1	1	1				1
d_{52}	1		1		1		1	4	1	1		1			1
d_{53}		1	1			1	1	1	4	1			1		1
d_{54}				1	1	1	1	1	1	4				1	1
d_{61}	1	1		1			1				4	1	1	1	1
d_{62}	1		1		1			1			1	4	1	1	1
d_{63}		1	1			1			1		1	1	4	1	1
d_{64}				1	1	1				1	1	1	1	4	1
d_{65}				1			1	1	1	1	1	1	1	1	4

Table 3. The matrix $\mathbf{P}_6$

	1	2	3	4	5	6
d_{21}	1	1				
d_{31}	1		1			
d_{32}		1	1			
d_{41}	1			1		
d_{42}		1		1		
d_{43}			1	1		
d_{51}	1				1	
d_{52}		1			1	
d_{53}			1		1	
d_{54}				1	1	
d_{61}	1					1
d_{62}		1				1
d_{63}			1			1
d_{64}				1		1
d_{65}					1	1

$$\left.\begin{aligned} w_{ijkl} &= 4 \quad && if \quad && i, j = k, l \\ w_{ijkl} &= 1 \quad && if \quad && i \in (k,l) \quad or \quad j \in (k,l) \\ w_{ijkl} &= 0 \quad && otherwise && \end{aligned}\right\}$$

This pattern is repeated for general values of n but with the first line replaced by $w_{ijkl} = n\text{-}2$.

We shall define a matrix $\mathbf{P}_s$ with $s(s\text{-}1)/2$ rows and s columns. The special case $s = n$ will be written without the suffix; thus $\mathbf{P}$ with m rows and n columns is synonymous with $\mathbf{P}_n$. Any row of $\mathbf{P}_s$ labelled d_{ij} has a unit in columns i and j and is zero elsewhere. Table 3 shows $\mathbf{P}_6$.

Thus, $\mathbf{C}$ of Table 1 may be written in the form:

<table>
<tr><th></th><th>d_{21}</th><th>d_{31}</th><th>d_{32}</th><th>d_{41}</th><th>...</th><th>d_{43}</th><th>d_{51}</th><th>...</th><th>d_{54}</th><th>d_{61}</th><th>...</th><th>d_{65}</th></tr>
<tr><td>d_{321}</td><td>$\mathbf{I}_1$</td><td colspan="2">$\mathbf{P}_2$</td><td colspan="9"></td></tr>
<tr><td>d_{421}</td><td colspan="3" rowspan="3">$\mathbf{I}_3$</td><td colspan="3" rowspan="3">$\mathbf{P}_3$</td><td colspan="6" rowspan="3"></td></tr>
<tr><td>d_{431}</td></tr>
<tr><td>d_{432}</td></tr>
<tr><td>d_{521}</td><td colspan="6" rowspan="3">$\mathbf{I}_6$</td><td colspan="3" rowspan="3">$\mathbf{P}_4$</td><td colspan="3" rowspan="3"></td></tr>
<tr><td>⋮</td></tr>
<tr><td>d_{543}</td></tr>
<tr><td>d_{621}</td><td colspan="9" rowspan="3">$\mathbf{I}_{10}$</td><td colspan="3" rowspan="3">$\mathbf{P}_5$</td></tr>
<tr><td>⋮</td></tr>
<tr><td>d_{654}</td></tr>
</table>

It may also be verified that the off-diagonal values of $\mathbf{C}'\mathbf{C}$ are given by $\mathbf{PP}'$ which has diagonal values $2\mathbf{I}$. Thus, we have:

$$\mathbf{C}'\mathbf{C} = (n - 4)\mathbf{I} + \mathbf{PP}' \tag{16}$$

Assuming the approximate independence of the residuals, (16) implies that the two parts contribute to stress in the ratio $trace[(n\text{-}4)\mathbf{I}]:trace(\mathbf{PP}') = (n\text{-}4)m:2m = (n\text{-}4):2$. For quite modest values of n, the term $(n\text{-}4)\mathbf{I}$ dominates, suggesting that minimizing triadic stress (or sstress) might very closely approximate minimizing dyadic stress. That is, we expect the approximation $\hat{\delta}_{2T} \sim \hat{\delta}_{2D}$ to be a good one.

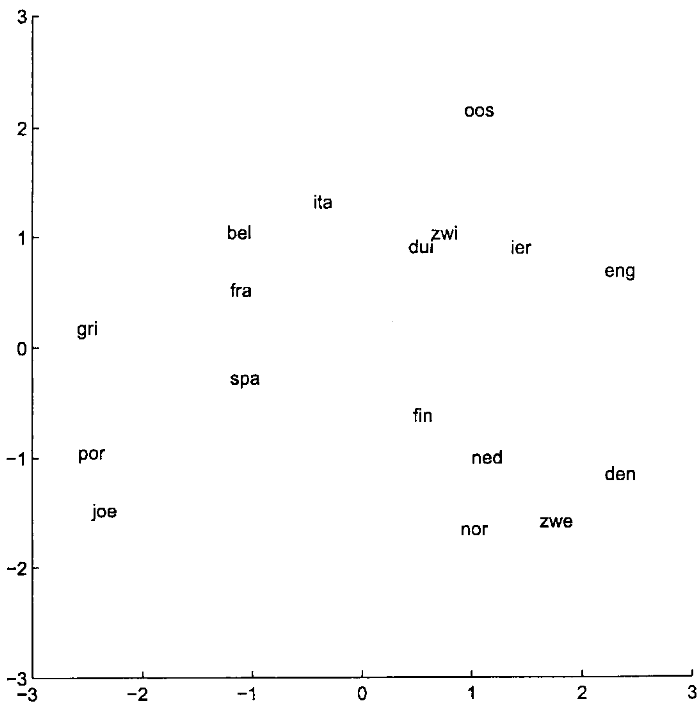

Figure 1. MDS2 solution to country data.

3. Numerical Results

The expected close relationship between MDS2 and MDS3 is investigated in the following examples. In the first, the data exemplify case (ii) and, in the second, case (i). We also briefly mention case (ii) data where $\mathbf{d}_3$ are not perimeter or generalized Euclidean distances.

3.1 Example 1

Basic dyadic distances between 17 countries were calculated from four variables – power-distance, individualism/collectivism, masculinity/ femininity and avoidance of insecurity (Hofstede, 1980). Figures 1, 2 and 3 show the multidimensional scalings of the 17 countries. Figure 1 shows a conventional MDS2 obtained by minimizing the dyadic stress criterion. Our goal is to compare the solution with the representation obtained from triadic distance models. Therefore, we also formed the triadic distances from the four-dimensional dyadic distances according to the perimeter model (1) and the generalized Euclidean model (2). These triadic distances were then used to obtain fitted triadic distances in two dimensional space by minimizing (3) in the potentially more general form which allows given weights w_{ijk}:

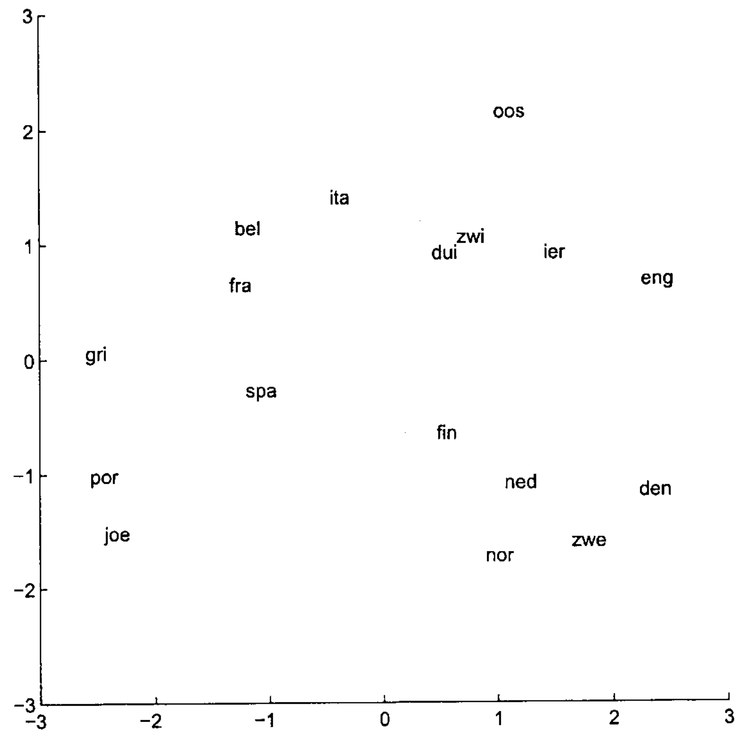

Figure 2. MDS3 perimeter mode solution to country data.

$$\sum_{i<j<k}^{n} w_{ijk}(d_{ijk} - \delta_{ijk})^2$$

This general form of the triadic stress criterion may be minimized by the iterative majorization algorithms (MDS3) proposed by Heiser and Bennani (1997). We minimized (3) by choosing the weights to be zero on the super diagonal and on the diagonal plane and unity elsewhere.

Figures 2 and 3 show the MDS3 (minimizing stress) of the triadic distances calculated from the dyadic distances according to the perimeter model (1) and the generalized Euclidean model (2). We are not concerned here with the substantive interpretation of these figures but only with the fact, as expected from our above analysis, that the three solutions are virtually identical to each other and to the MDS2 configuration of Figure 1. The stress values were 208.59 (perimeter model) and 69.37 (generalized Euclidean model). The stress values for the two solutions cannot be directly compared as there is a difference of scale of about three. Indeed, 208.59 is nearly the same as 3×69.37 but we think that this degree of agreement is accidental. The MDS2 solution had a stress of 14.68 which translates into triadic stresses for the two models of 235.48 and 76.17. That the dyadic solutions were slightly less good than the triadic is purely a consequence of minimizing triadic stress; if we had compared at the level of dyadic stress the converse would have been found, for dyadic stress would then have been minimized and the translated triadic stresses would have been greater. We did not minimize triadic sstress but would expect similar findings.

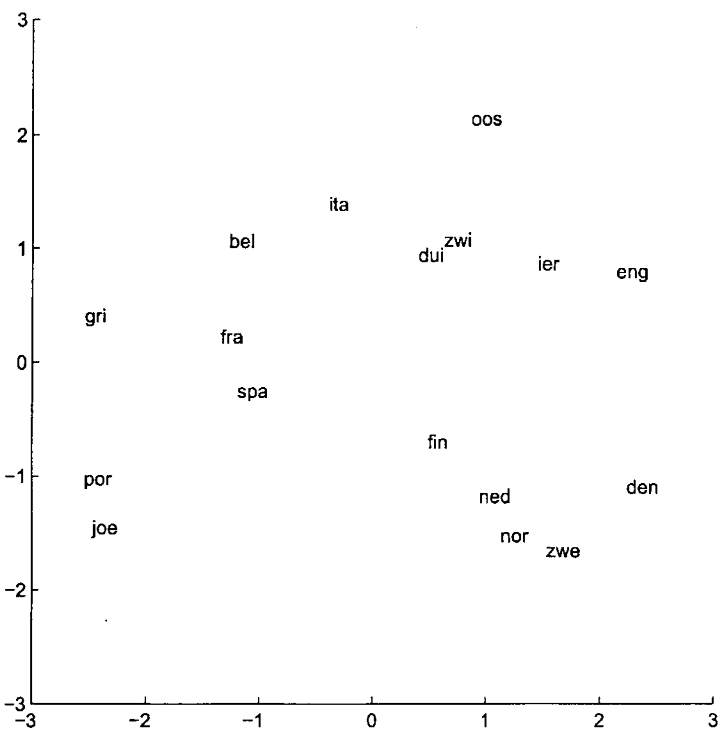

Figure 3. MDS3 generalized Euclidean model solution to country data.

3.2 Example 2

The previous example used data of type (ii) where triadic distances are derived from observed dyadic distances. In this next example we are concerned with observed data on the unproductivity of triadic teams (Hayashi, 1972) and can use the analysis of variance approach discussed in section 3. In this example n = 6, m = 15 and M = 20. The total sum-of-squares is 853. For the perimeter model, the residual sum-of squares for the unrestricted model is 8.92 giving an excellent fit. However, the fitted values $\hat{\mathbf{d}}_3$ contain one negative value and the corresponding fitted dyadic values $\hat{\mathbf{d}}_2$

have three negative values. Thus, even though the fit is good, there seems little point in going on to fit a Euclidean model, at least not in the context of metric multidimensional scaling. To make progress we consider adding a constant k to all the elements of $\hat{\mathbf{d}}_2$ in the hope that by choosing a suitable value of k, the new constrained estimates of δ_2 will give better Euclidean fits. (For a more extended discussion of negative dissimilarities see Heiser, 1991). Thus, we set:

$$\delta_2 = \gamma_2 - k\mathbf{1} \tag{17}$$

where γ_2 is constrained to be Euclidean. Equation (17) represents a change of model that has an effect on the analysis of variance. Equation (10) remains valid but now:

$$S_S = \|\hat{\mathbf{d}}_3 - \mathbf{C}(\gamma_2 - k\mathbf{1})\| = \|\hat{\mathbf{d}}_3 - (\mathbf{C}\gamma_2 - 3k\mathbf{1})\|. \tag{18}$$

For given k, $\hat{\gamma}_2$ is obtained by minimizing (18) rearranged as $\|(\hat{\mathbf{d}}_3 + 3k\mathbf{1}) - \mathbf{C}\gamma_2\|$ which is (11) with each element of $\hat{\mathbf{d}}_3$ increased by $3k$. With this small change, the previous algorithm remains valid. For given γ_2, k is obtained by minimizing (18) rearranged as $\|(\hat{\mathbf{d}}_3 - \mathbf{C}\gamma_2) + 3k\mathbf{1}\|$, giving an estimate:

$$\hat{k} = -\tfrac{1}{3M}\mathbf{1}'(\hat{\mathbf{d}}_3 - \mathbf{C}\gamma_2).$$

Based on these updating steps, it is fairly straightforward to estimate optimal values $\hat{k}$ and $\hat{\gamma}_2$ by using an alternating least-squares algorithm. We decided to examine three values of k: that for which $\hat{\mathbf{d}}_2 + k\mathbf{1}$ becomes Euclidean, that for which it becomes metric, and the optimal value. The Cailliez (1983) solution to the additive constant problem yields $k = 5.5698$ as the smallest constant that ensures that $\hat{\mathbf{d}}_2 + k\mathbf{1}$ is Euclidean embeddable. Note that with this value of k, (17) would give an exact fit ($S_S = 0$) in n-2 = 4 dimensions. Here we consider two-dimensional solutions based on this value of k. The "worst" triangle given by the unconstrained fit $\hat{\mathbf{d}}_2$ is associated with the points 4, 5 and 6 with $\hat{d}_{45} = -0.7500$, $\hat{d}_{46} = 0.9167$ and $\hat{d}_{56} = -0.5833$. This triangle is closed, thus satisfying the metric inequality, by adding 2.25 to each side and this is the smallest quantity that ensures that the metric inequality is satisfied for all triangles. The iterative procedure gave an estimated optimal value $\hat{k} = 2.0043$ and $S_S = 3.70$, associated with a two-dimensional solution $\hat{\gamma}_2$. Like other algorithms of its type, local minima were found. Out of 20 random starts, ten found the value tabulated, whereas ten found $\hat{k} = 1.9176$ with $S_S = 4.28$.

Thus, we set k = 5.5698, k = 2.25 and $\hat{k}$ = 2.0043 and repeated our analysis, adding $3k$ to each element of $\hat{\mathbf{d}}_3$. The analyses of variance are shown below.

	$k = 5.5698$	$k = 2.25$	$\hat{k} = 2.00$	2-dimensions
Restricted Fit	755.47	836.76	840.38	9 (10)
S_S	88.62	7.33	3.70	6 (5)
S_U	8.92	8.92	8.92	5
S_R	97.53	16.24	12.62	11 (10)
Total	853	853	853	20

The line labelled Restricted Fit in the analysis of variance is obtained by subtracting S_R from the total sum-of-squares. For the optimal value of $\hat{k}$, this line can be interpreted as the fitted sum of squares $||\mathbf{C}(\hat{\gamma}_2 - \hat{k}\,\mathbf{1})||$ but suboptimal values lead to non-orthogonal terms; however, the residual lines remain fully valid. The residual sums-of-squares are all small when compared with the total. The choice of k = 5.5698 has given a value of S_S about ten times that of S_U, each associated with a similar number of parameters. This suggests that the two-dimensional Euclidean fit has failed to account for all the structure in $\mathbf{d}_3$. The fit is much better for k = 2.25. The optimal solution improves a little on the solution for k = 2.25 but the MDS3 diagrams are indistinguishable. Because an extra parameter has been fitted, the number of parameters associated with each line of the analysis of variance for k = 2.00 needs adjustment; adjusted values are shown in parentheses. The direct MDS2 solution that minimizes $||\hat{\mathbf{d}}_2 - (\gamma_2 - k\mathbf{1})||$ yields a value of $\hat{k}$ = 1.7867 and a stress of 1.6447. Figure 4 shows the MDS2 and MDS3 two-dimensional configurations. The differences are slight, showing that the degree of approximation can be excellent even with n as low as 6.

Another example of close agreement is shown in Figure 12.1 (i) and (ii) of Cox and Cox (2001).

3.3 Remarks on Other Triadic Distances

Analysis has shown that the MDS3 of two triadic distances (1) and (2) might be expected to give similar multidimensional scalings to the corresponding MDS2. This theoretical result was borne out by examples whose MDS3 gave virtually the same representation as when the simple dyadic distances were used. This does not imply that the same necessarily

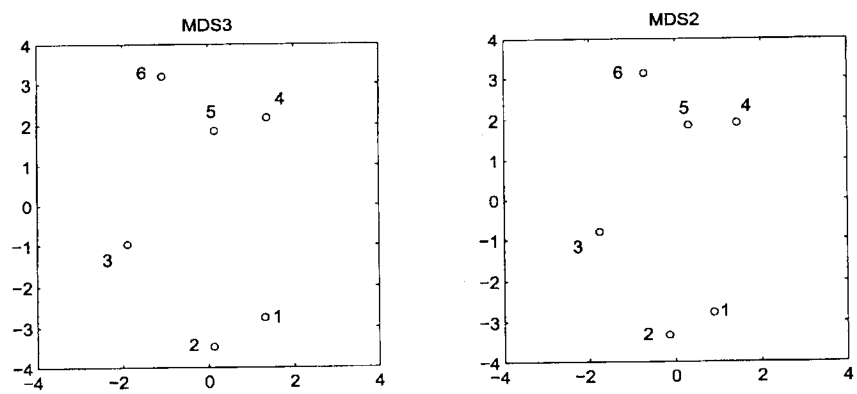

Figure 4. MDS solutions for Hayashi team data.

applies to other definitions of triadic distances. Indeed, with some triadic distances we might expect major differences in the configurations obtained by MDS3 and MDS2. For example, if we define $\delta_{ijk} = variance(\delta_{jk}, \delta_{ik}, \delta_{ij})$, it can be seen that adding any constant onto the dyadic distances leaves δ_{ijk} unaltered[3]. Thus the triangles with sides (2,3,4), (3,4,5) and (4,5,6) have the same triadic variances, as will any whose sides differ by a constant from (1,2,3). The three specified triangles share a side of length 4. Figure 5 shows the three triangles with the shared side superimposed. The remaining vertices are in very different places. The whole figure may be reflected about AB or the perpendicular through the center of AB without affecting triadic variance, showing that even triangles with a shared side and the same variance may be located in a host of different positions C_i. Metric information seems to be lost, leading one to expect that the MDS2 and MDS3 of triadic variance may differ greatly when analyzed by any form of metric MDS. To see if this were indeed so, we reanalyzed the data of section 3.1, after converting to triadic variance. Figure 6 shows the result, with a stress of 125.09. Contrary to expectation although not quite so good as with models (1) and (2), the approximation remains quite satisfactory. Presumably, this is because the 680 triadic variances, fitted by 136 dyadic distances impose strong constraints on admissible configurations.

[3] It is not difficult to show by counterexample that this definition of a triadic distance does not satisfy the axioms of Heiser and Bennani (1997) nor that of Joly and Le Calvé (1995), but nevertheless has given an approximation that compares well with configurations that derive from definitions that do satisfy the axioms and, indeed, with the known initial dyadic distances.

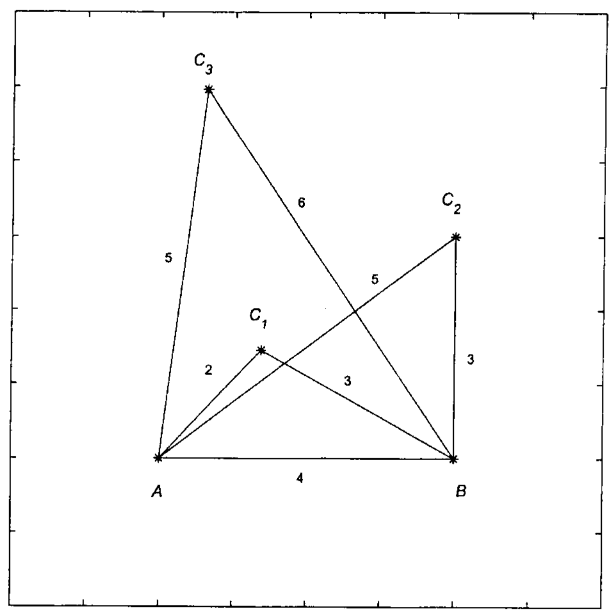

Figure 5. Three points C with equal variance to A and B.

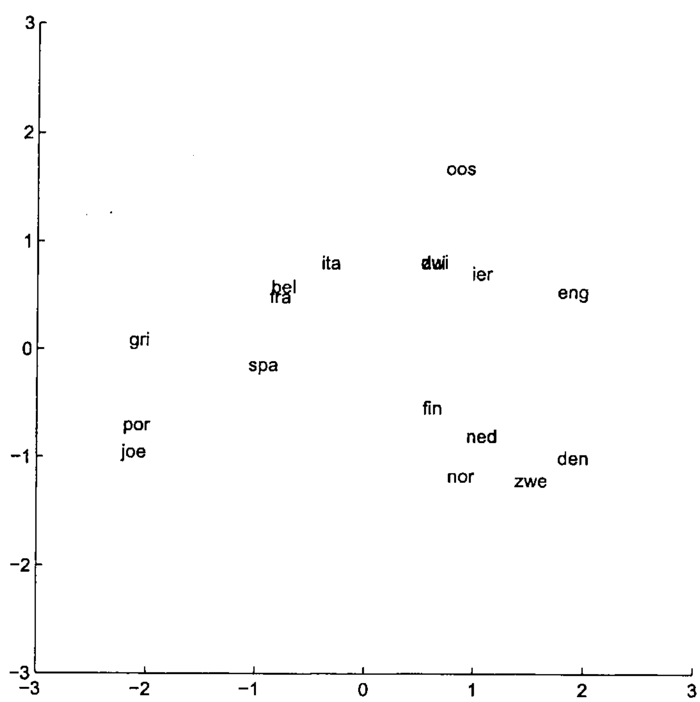

Figure 6. MDS3 variance solution to country data.

4. Conclusion

We have shown that minimizing triadic stress (3) by fitting the perimeter model (1), without loss of generality, may be achieved in two stages: (a) an unconstrained fit followed by (b) a constrained fit to the triadic distances $\mathbf{C}\hat{\mathbf{d}}_2$ derived from the first stage. The procedure may be summarized in an analysis of variance that can give guidance on the type of constraint that may be justified by the data. Analysis shows that the MDS3 of $\mathbf{d}_3$ (or equivalently of $\mathbf{C}\hat{\mathbf{d}}_2$) yields a configuration that generates distances $\hat{\delta}_{2T}$ that are closely approximated by $\hat{\delta}_{2D}$ generated by a conventional MDS2 of $\hat{\mathbf{d}}_2$. These results apply equally to whether $\mathbf{d}_3$ represents observed triads (i.e. model (i)) or is derived from observed dyadic distances $\mathbf{d}_2$ (i.e. model (ii)). With model (i) the comparison of $\hat{\delta}_{2T}$ with $\hat{\delta}_{2D}$ is internal and may be judged only insofar as we get a good fit. With model (ii), an MDS2 of $\mathbf{d}_2$ gives an external "true" solution which also may be compared with $\hat{\delta}_{2T}$ and $\hat{\delta}_{2D}$. Equivalent results apply to minimizing triadic sstress (4) by fitting the generalized Euclidean model (2).

The clear indication is that an MDS3 of triadic distances gains little or nothing over the conventional MDS2 of the dyadic distances $\hat{\mathbf{d}}_2$. Perhaps this should be no surprise when the fitted triadic distances are linear transformations of dyadic distances, as with (1) and (2) and in our first examples. However, it seems that non-linear transformations (e.g. triadic variance) also have little effect. Then $\mathbf{d}_3 = \mathbf{C}\mathbf{d}_2$ for no matrix $\mathbf{C}$, and fitting (1) can be regarded as fitting the wrong model. Nevertheless, the MDS2 of $\mathbf{d}_2$, agreed well with $\hat{\delta}_{2T}$ and $\hat{\delta}_{2D}$. As noted by de Rooij (2001, Chapter 5; 2002) the problem seems to be that definitions of triadic distance in terms of dyadic distances do not model true three-way interactions. It is clear that the properties of different triadic distance coefficients need close examination; this is the subject of current work.

Appendix

The equations (5) $\mathbf{d}_3 = \mathbf{C}\delta_2$ are overdetermined when $n > 5$. However, if the elements of $\mathbf{d}_3$ are indeed triadic perimeter distances, these equations will be consistent and will have a unique solution for δ_2. For sets of $n = 5$ points; the matrix $\mathbf{C}$ is square of order 10 (see the first ten rows and columns of Table1) and may be inverted to give ten elements of δ_2. By choosing different quintuples, each of the elements of δ_2 will be found repeatedly (e.g. 12345, 12678 both lead to evaluations of δ_{12}). When the equations are inconsistent, the values of the common terms will disagree. When all solutions are consistent, there will be no disagreement, and the whole of δ_2 may be built up by considering enough overlapping quintuples. Then, a δ_2 will have been found that generates the triadic distances $\mathbf{d}_3$. A more symmetric approach is to adopt the solution:

$$\delta_2 = (\mathbf{C}'\mathbf{C})^{-1}\,\mathbf{C}'\mathbf{d}_3 \tag{19}$$

When $\mathbf{d}_3$ is a true triadic distance satisfying (1), (19) must recover its generating dyadic values δ_2, and when $\mathbf{d}_3$ is not triadic $\mathbf{C}\delta_2$ gives the least-squares estimate $\hat{\mathbf{d}}_3$ with its generating dyadic distances $\hat{\mathbf{d}}_2$. Therefore when the elements of $\mathbf{d}_3$ are triadic distances, we must have:

$$\mathbf{d}_3 = \mathbf{C}\delta_2 = \mathbf{C}(\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\mathbf{d}_3.$$

It follows that:

$$(\mathbf{I} - \mathbf{C}(\mathbf{C}'\mathbf{C})^{-1}\,\mathbf{C}')\mathbf{d}_3 = 0 \tag{20}$$

is a necessary condition both for the validity of the perimeter model and when δ_2 is defined to contain square-distances, for the generalized Euclidean model. Also (20) is a sufficient condition, because it implies that $\mathbf{d}_3 = \mathbf{C}\delta_2$ where δ_2 is given by (19). Equation (20) implies that all M elements of the vector should be checked for zero but this can be avoided by calculating their sum-of-squares:

$$\mathbf{d}_3'(\mathbf{I} - \mathbf{C}(\mathbf{C}'\mathbf{C})^{-1}\,\mathbf{C}')\mathbf{d}_3 = 0, \tag{21}$$

i.e. we have only to check (8) that $S_U = 0$. Noting that:

$$(n-4)(\mathbf{C}'\mathbf{C})^{-1} = \mathbf{I} - \frac{1}{2(n-3)}\mathbf{PP}' + \frac{2}{3(n-2)(n-3)}\mathbf{11} \tag{22}$$

(20) may be rewritten:

$$(n-4)\mathbf{d}_3 = \mathbf{C}(\mathbf{I} - \frac{1}{2(n-3)}\mathbf{PP}' + \frac{2}{3(n-2)(n-3)}\mathbf{11}')\mathbf{C}'\mathbf{d}_3. \tag{23}$$

This result requires $n \geq 5$; for $n = 3, 4$ $\mathbf{C}'\mathbf{C}$ is of deficient rank.

The necessary and sufficient condition (23) is well known in its expanded form (see e.g. Proposition 2 of Chepoi and Fichet, 1998). To obtain the expanded form from our matrix result (23) we proceed as follows. From the definitions of $\mathbf{C}$ (Table 1) and $\mathbf{P}$ (Table 2) we immediately have the following typical elements for the given vectors:

$$\left.\begin{array}{lrl} \mathbf{C1} & =31 & (M\times 1) \\ \mathbf{C'd}_3 & =\mathrm{d}_{*ij} & (m\times 1) \\ \mathbf{CC'd}_3 & =\{\mathrm{d}_{*ij}+\mathrm{d}_{*jk}+\mathrm{d}_{*ik}\} & (M\times 1) \\ \mathbf{P'C'd}_3 & =\mathrm{d}_{**i} & (n\times 1) \\ \mathbf{PP'C'd}_3 & =\{\mathrm{d}_{**i}+\mathrm{d}_{**j}\} & (m\times 1) \\ \mathbf{CPP'C'd}_3 & =2\{\mathrm{d}_{**i}+\mathrm{d}_{**j}+\mathrm{d}_{**k}\} & (M\times 1) \end{array}\right\}$$

where the * notation means *sum over all non-repeated suffices*. When inserted into (23) these give:

$$\mathrm{d}_{ijk}=\frac{(\mathrm{d}_{*ij}+\mathrm{d}_{*jk}+\mathrm{d}_{*ik})}{\binom{n-4}{1}}-\frac{(\mathrm{d}_{**i}+\mathrm{d}_{**j}+\mathrm{d}_{**k})}{\binom{n-3}{2}}+\frac{\mathrm{d}_{***}}{\binom{n-2}{3}}. \qquad (24)$$

Thus, the M conditions (24) are the same as the single matrix conditions (20), (21) or (23). Sometimes it is convenient to replace the summations in (24) by summing over all suffices *excluding* i, j and k. Using a dot notation for this kind of summation, (24) becomes:

$$\mathrm{d}_{ijk}=\frac{(\mathrm{d}_{\cdot ij}+\mathrm{d}_{\cdot jk}+\mathrm{d}_{\cdot ik})}{\binom{n-3}{1}}-\frac{(\mathrm{d}_{\cdot\cdot i}+\mathrm{d}_{\cdot\cdot j}+\mathrm{d}_{\cdot\cdot k})}{\binom{n-3}{2}}+\frac{\mathrm{d}_{\cdot\cdot\cdot}}{\binom{n-3}{3}} \qquad (25)$$

but, because we have used division by $n-5$ in deriving this result, (25) requires $n > 5$ whereas (24) is also valid for $n = 5$. When $n = 6$, (25) gives:

$$\mathrm{d}_{ijk}=\frac{(\mathrm{d}_{\cdot ij}+\mathrm{d}_{\cdot jk}+\mathrm{d}_{\cdot ik})}{3}-\frac{(\mathrm{d}_{\cdot\cdot i}+\mathrm{d}_{\cdot\cdot j}+\mathrm{d}_{\cdot\cdot k})}{3}+\mathrm{d}_{\cdot\cdot\cdot} \qquad (26)$$

the sums chosen being over the three suffices remaining after allowing for i, j and k.

A remarkable result (Chepoi and Fichet, 1998) is that it is necessary and sufficient for (26) to hold for all $\binom{n}{6}$ 6-tuples for the elements of $\mathbf{d}_3$ to be triadic perimeter distances. Necessity is obvious, Chepoi and Fichet give an elegant proof of sufficiency; here we note that sufficiency follows directly from (25) and (26). We outline the proof.

Consider all $\binom{n-3}{3}$ 6-tuples that include *i*, *j* and *k*, thus excluding all those that do not involve d_{ijk}. Then for given *p, q, r* ≠ *i j, k*:

$$\left.\begin{array}{lll} d_{ijk} & \text{occurs} \binom{n-3}{3} & \text{6-tuples} \\ d_{pij} & \text{occurs} \binom{n-4}{2} & \text{6-tuples} \\ d_{pqi} & \text{occurs} \binom{n-5}{1} & \text{6-tuples} \\ d_{pqr} & \text{occurs} \binom{n-6}{0} & \text{6-tuples} \end{array}\right\}$$

Thus, the last line derives from the observation that *i, j* and *k* use three of the suffices, leaving only one 6-tuple (*i, j, k, p, q, r*) to deliver d_{pqr}. Similarly, (*i, j, k, p, q*) leaves *n – 5* possible suffices to select for the remaining member of the 6-tuple. Summing over all 6-tuples that include *i, j* and *k* and over all *p, q,* and *r* (26) gives:

$$\binom{n-3}{3} d_{ijk} \doteq \binom{n-4}{2}\frac{(d_{.ij}+d_{.jk}+d_{.ik})}{3} - \binom{n-5}{1}\frac{(d_{..i}+d_{..j}+d_{..k})}{3} + \binom{n-6}{0} d_{...}$$

summation occurring over the full set of permissible suffices, excluding *i, j* and *k*. A little rearrangement leads to (25), thus proving sufficiency.

Turning now to the expression of dyadic distances in terms of triadic distances, we may write (19) in expanded form as:

$$\delta_{ij} = \frac{d_{*ij}}{n-4} - \frac{(d_{**i}+d_{**j})}{(n-3)(n-4)} + \frac{2d_{***}}{(n-2)(n-3)(n-4)}. \qquad (27)$$

or, equivalently:

$$\delta_{ij} = \frac{d_{.ij}}{n-2} - \frac{(d_{..i}+d_{..j})}{(n-2)(n-3)} + \frac{2d_{...}}{(n-2)(n-3)(n-4)} \qquad (28)$$

now with summation over all suffices excluding i and j.

Although (19) or (27) allow dyadic terms to be calculated from given or fitted triadic distances, there is no guarantee that the elements of δ_2, so found, are distances, or even that they are positive. Chepoi and Fichet (1998) and Heiser and Bennani (1997) give various definitions of triadic metrics and ultrametrics. Here we note the simple alternative of defining the elements of δ_3 to inherit the metric properties of the elements of δ_2. Thus, if the elements of δ_2 satisfy the metric inequality, the ultrametric inequality or are Euclidean, we *define* the elements of δ_3 to be triadic metrics, triadic ultrametrics or triadic Euclidean distances. This form of definition ties in with our MDS problems where we fit a Euclidean δ_2 to $\mathbf{d}_3$.

For interest and completeness, we have found the spectral decomposition of $\mathbf{C}'\mathbf{C}$. The result is:

$$\mathbf{C}'\mathbf{C} = (n-4)\left[\mathbf{I} - \tfrac{1}{n-2}\mathbf{PP}' + \tfrac{2}{(n-1)(n-2)}\mathbf{11}'\right] + 2(n-3)\left[\tfrac{1}{n-2}\left(\mathbf{PP}' - \tfrac{4}{n}\mathbf{11}'\right)\right] + 3(n-2)\left[\tfrac{1}{m}\mathbf{11}'\right] \quad (29)$$

where $n\text{-}4$, $2(n\text{-}3)$ and $3(n\text{-}2)$ are the only non-zero eigenvalues. The matrices in square brackets [ ] are idempotent, are mutually orthogonal and sum to a unit matrix. Their ranks give the multiplicities of the corresponding eigenvalues as $\lambda_1 = n\text{-}4$ with multiplicity $\mu_1 = \frac{1}{2}n(n\text{-}3)$, $\lambda_2 = 2(n\text{-}3)$ with multiplicity $\mu_2 = (n\text{-}1)$ with $\lambda_3 = 3(n\text{-}2)$ an unrepeated eigenvalue, so that $\mu_3 = 1$. Independent eigenvectors may be taken to be any μ_i columns from the corresponding idempotent matrices. We derived this result from the observation that:

$$(\mathbf{C}'\mathbf{C})^r = \alpha_r\mathbf{I} + \beta_r\,\mathbf{11}' + \gamma_r\mathbf{PP}'$$

for all r, where α_r, β_r, γ_r are scalars. The modified Leverier-Faddeev algorithm (Gower, 1975) shows that (29) must have the same form for suitable coefficients α, β, γ which, after considerable detailed but elementary algebra, are found to be as given in (29).

References

CAILLIEZ, F. (1983) "The analytical solution to the additive constant problem", *Psychometrika 48*, 305 - 308.

CHEPOI, V. and FICHET, B. (1998), "A note on three-way dissimilarities and their relationship with two-way dissimilarities", In: *Data, Science, Classification and Related methods.* Paper presented at the IFCS 1998 Rome, Italy.

COX T.F. and COX. M.A.A. (2001), *Multidimensional Scaling*. – 2nd edition. Monographs on Statistics and Applied Probability 88, Boca Raton, Chapman and Hall/ CRC.

COX, T.F., COX, M.A.A., and BRANCO, J.A. (1991), "Multidimensional scaling or *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. (2000), "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. (2001), *Distance models for transition frequency data*. Unpublished doctoral thesis, Leiden University.

DE ROOIJ, M. (2002), "Distance models for three-way tables and three-way association", *Journal of Classification, 19,* 161-178

GOWER, J.C. (1985), "A modified Leverrier-Faddeev algorithm", *Linear Algebra and its Applications, 31,* 61-70.

HAYASHI, C. (1972), "Two dimensional quantifications based on the measure of dissimilarity among three elements", *Annals of the Institute of Statistical Mathematics, 24,* 251-257.

HEISER, W.J. (1991), "A generalized majorization method for least squares multidimensional scaling of pseudodistances that may be negative", *Psychometrika, 56,* 7-27.

HEISER, W.J. and BENNANI, M. (1997), "Triadic distance models: axiomatization and least squares representation", *Journal of Mathematical Psychology, 41,* 189-206.

HOFSTEDE, G. (1980), *Culture's consequences: International differences in work-related values.* Beverley Hills, CA, Sage Publications.

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

PAN, G. and HARRIS, D.P. (1991), "A new multidimensional scaling technique based upon association of triple objects P_{ijk} and its application to the analysis of geochemical data", *Mathematical Geology, 23,* 861-886.