Wiadomości Statystyczne. The Polish Statistician, 2024, *69*(5), 1–13 https://doi.org/10.59139/ws.2024.05.1
Studia metodologiczne / Methodological studies
Zgłoszony/submitted: 19.12.2022, poprawiony/revised: 30.01.2024, zaakceptowany/accepted: 20.03.2024

The optimal number of tetrads for measurement of dissimilarities in nonmetric multidimensional scaling[1]

Artur Zaborski[a]

Abstract. The direct determination of dissimilarities is the most popular and most frequently used method for raising input data in nonmetric multidimensional scaling, i.e. when variables are measured on an ordinal scale (e.g. in preference studies). Most methods for the direct measurement of similarities, including ranking, sorting, pairwise comparison, conditional ranking of similarities are, however, very laborious, especially when a large number of objects is tested. Thus, the research described in this article is based on the tetrad method, which is uncomplicated and less burdensome for the respondents.

In the proposed method, respondents are asked to evaluate four-element subsets (tetrads) from a set of *n* objects. The respondent is asked to indicate the pair with the most and the least similar elements in each tetrad. As the number of tetrads rapidly increases along with the number of objects, it becomes necessary to use the incomplete variant of the method, in which only some four-element subsets are presented to the respondents for evaluation.

The aim of the research presented in the article is to determine the size of the tetrad set that is sufficient to create a dissimilarity matrix used to perform nonmetric multidimensional scaling. The study was based on four distance matrices for 7, 9, 11 and 13 objects that were randomly selected voivodship capitals in Poland. The distances between the cities were expressed in kilometres. The Procrustes analysis and Spearman's rank correlation were used in the study. The findings show that the use of the tetrad method for the measurement of dissimilarities produces beneficial results already at the point when each pair of objects appears in the set of tetrads only once, which allows the number of opinions provided by the respondents to be significantly reduced.

Keywords: measurement of dissimilarities, direct indication of dissimilarities, tetrad method, nonmetric multidimensional scaling, Procrustes analysis

JEL: C38, C63, M31

Optymalna liczba tetrad do pomiaru niepodobieństw w niemetrycznym skalowaniu wielowymiarowym

Streszczenie. Bezpośrednie wyznaczanie niepodobieństw jest najpopularniejszym i najczęściej stosowanym sposobem uzyskiwania danych wejściowych w niemetrycznym skalowaniu wielowymiarowym, czyli gdy zmienne są mierzone na skali porządkowej (np. w badaniach preferencji).

[1] Artykuł został opracowany na podstawie referatu wygłoszonego na konferencji Multivariate Statistical Analysis MSA'2022, która odbyła się w dniach 7–9 listopada 2022 r. w Łodzi. / The article is based on a paper delivered at the Multivariate Statistical Analysis MSA'2022 Conference, held on 7–9 November 2022 in Łódź, Poland.

[a] Uniwersytet Ekonomiczny we Wrocławiu, Wydział Ekonomii i Finansów, Polska / Wroclaw University of Economics and Business, Faculty of Economics and Finance, Poland.
ORCID: https://orcid.org/0000-0003-1374-2268. E-mail: artur.zaborski@ue.wroc.pl.

Stosowanie większości metod bezpośredniego pomiaru podobieństwa, takich jak rangowanie, sortowanie, porównania parami czy warunkowe porządkowanie podobieństw, jest jednak bardzo pracochłonna, zwłaszcza przy dużej liczbie obiektów. Z tego powodu w badaniu omawianym w niniejszym artykule posłużono się metodą tetrad – nieskomplikowaną i nieuciążliwą dla respondentów.

Proponowana metoda polega na przedstawieniu respondentom do oceny czteroelementowe podzbiory (tetrady) z *n*-elementowego zbioru obiektów. Respondent proszony jest o wskazanie pary najbardziej i najmniej podobnych elementów w każdej tetradzie. Ponieważ liczba tetrad rośnie bardzo szybko wraz z liczbą obiektów, niezbędne jest zastosowanie niepełnego wariantu tej metody, w której respondentom przedstawia się do oceny jedynie część czteroelementowych podzbiorów.

Badanie omawiane w artykule ma na celu ustalenie wielkości zbioru tetrad, która jest wystarczająca do wyznaczenia macierzy niepodobieństw służącej do wykonania niemetrycznego skalowania wielowymiarowego. Badanie przeprowadzono na podstawie czterech macierzy odległości dla 7, 9, 11 i 13 obiektów, którymi były losowo wybrane miasta wojewódzkie w Polsce. Odległości między miastami wyrażono w kilometrach. Wykorzystano analizę Prokrustesa oraz współczynnik korelacji rang Spearmana. Wykazano, że zastosowanie metody tetrad do pomiaru niepodobieństw daje korzystne wyniki już wtedy, gdy każda para obiektów pojawia się w zbiorze tetrad tylko raz, co pozwala na znaczne ograniczenie liczby opinii wyrażanych przez respondentów.

Słowa kluczowe: pomiar niepodobieństw, bezpośrednie wyznaczanie niepodobieństw, metoda tetrad, niemetryczne skalowanie wielowymiarowe, analiza Prokrustesa

1. Introduction

Nonmetric multidimensional scaling is used when variables are measured on an ordinal scale. The development of nonmetric methods of classical scaling took place in the 1960s. The first algorithm was proposed by Shepard (1962). Shepard presented, for the first time, an iterative computer procedure of the proximity matrix analysis, in which the distances between points were measured on an ordinal scale. The following researchers also contributed extensively to the development of multidimensional scaling methodology in their studies: Carroll and Chang (1970), Guttman (1968), Kruskal (1964a, 1964b), de Leeuw and Heiser (1980), Takane et al. (1977), and others. The goal of nonmetric multidimensional scaling is to find such a mapping φ of a set of objects $O = (O_1,\ O_2, \dots ,\ O_n)$, e.g. people, products, companies, brands, advertisements, etc., with dissimilarities δ_{ij} $(i, j = 1, 2, \dots, n)$ into a set of points in an r-dimensional space (r is usually equal to 2 or 3) with d_{ij} distances between them, so that $\hat{d}_{ij} \approx d_{ij}$, where $\hat{d}_{ij}$ is a monotone regression of d_{ij} on δ_{ij} (see e.g. Borg & Groenen, 2005, p. 34; Kruskal, 1964a, 1964b). In nonmetric multidimensional scaling, dissimilarities are measured on an ordinal scale, so $\hat{d}_{ij}$ must satisfy the following condition:

$$\delta_{ij} \prec \delta_{i'j'} \Rightarrow \hat{d}_{ij} \leq \hat{d}_{i'j'}. \qquad (1)$$

The starting point in multidimensional scaling is the creation of a dissimilarity matrix. There are two ways of obtaining input dissimilarities in multidimensional scaling: when they are directly obtained from empirical subjective measurements of objects performed by subjects, they are called direct dissimilarities; by contrast, when they are not obtained from subjects, but calculated from a data matrix associated with these objects, they are labeled as derived dissimilarities. This article focuses only on direct dissimilarities.

So far, many more or less popular methods used for direct similarity measurement have been developed, e.g. sorting, pair comparisons, ranking, ranking of pairs, conditional ranking, and others. The differences in the application of various measurement methods may depend on the number of objects simultaneously presented to the respondents, the difficulty in assessing similarities and the total number of required ratings (see e.g. Bijmolt, 1996). When the number of objects is high, the number of direct assessments made by respondents becomes too large, and makes the task of dissimilarities measurement more difficult.

In view of the fact that the results of multidimensional scaling based on different collecting methods are similar (see e.g. Bijmolt, 1996; Humphreys, 1982; Zaborski, 2003), the choice of the method of measurement should be guided primarily by two criteria: the method should not be labour-intensive and expressing judgements should not be problematic for the respondents.

A method that largely meets these two criteria is the incomplete method of tetrads, proposed by Zaborski (2020). The idea is based on the method of triads (Burton & Nerlove, 1976) and on the theory of balanced incomplete block designs (see e.g. Burton & Nerlove, 1976; Morris, 2010; Rink, 1987). A preliminary comparative analysis of the method of triads and the method of tetrads for nine objects (Zaborski, 2022) showed that obtaining comparable results when using the method of triads requires over three times more respondents' assessments than in the case of the method of tetrads.

The purpose of the article is to determine the size of the tetrad set that is sufficient to create a dissimilarity matrix and to perform nonmetric multidimensional scaling.

2. Measurement of dissimilarities in nonmetric multidimensional scaling

There are three main approaches to collecting dissimilarities in nonmetric multidimensional scaling (Tsogo et al., 2000). The first approach is based on rankings and similarities ratings of pairs of objects. The second group of methods uses grouping and sorting tasks in order to calculate similarities. Finally, the third approach consists of pairwise comparisons of similarities.

The above-mentioned selection of methods tend to evoke subjective feelings of the respondents, i.e. fatigue, weariness resulting from making numerous assessments, or experiencing difficulties in expressing differences between objects. As a result, the collected data may be incomplete or the assessments made do not always fully reflect the opinions of the respondents.

Many studies, including those of Bijmolt (1996) and Humphreys (1982), have considered this problem. Although the research above did not confirm any significant influence of the measurement of dissimilarities on the multidimensional scaling results, other aspects differentiating the methods were identified. It is much more difficult to determine dissimilarities for an entire set of objects than for two or three of them, especially when some objects are only slightly different from each other. With a large number of objects, the respondents focus on extreme assessments, i.e. they notice the objects that are the most similar, the least similar, while ignoring the other objects in their assessments. As a result, the respondents' assessments are incomplete or often random. In turn, methods in which the respondents are presented with only two or three options require providing multiple answers, which results in fatigue and weariness. In such cases, incomplete analyses may offer a solution. A review of the research related to incomplete methods of measuring dissimilarities in multivariate scaling was presented by e.g. Tsogo et al. (2000).

3. The method of tetrads under incomplete block design

3.1. The methodological basis of the method of tetrads

In the complete method of tetrads (Zaborski, 2020, 2022), the subject is asked to consider all the possible groups of four objects $(O_i, O_j, O_k, O_l), i, j, k, l = 1, 2, \dots, n$, where $i \neq j \neq k \neq l \neq i \neq k$ and $j \neq l$, taken from the full set of n objects $O = (O_1, O_2, \dots, O_n)$. The respondent has to indicate the most similar pair and the least similar pair. On this basis the tetrad is formed, where the most similar objects are placed as the first and the second, and the least similar as the first and the fourth. If the object from the most similar pair (O_i, O_j) is not present in a pair of the least similar objects, then the most similar objects are placed as the second and the third. In this situation, the respondent should also be asked to indicate the second most similar pair of objects. For example, if (O_i, O_j) is the most similar pair, (O_i, O_k) is the second most similar pair and (O_k, O_l) is the least similar pair, the tetrad is $(O_k, O_i, O_j, O_l)^*$. The asterisk next to the tetrad indicates a case where the most similar objects are placed as the second and the third.

The number of ratings which a respondent must make for n objects in the method of tetrads is equal to the number of four-element combinations of the n-element set, and amounts to:

$$C_n^4 = \frac{n(n-1)(n-2)(n-3)}{24}, \tag{2}$$

where each pair appears in tetrads $(n-1)(n-2)/2$ times. Therefore, when the number of tetrads is considered too large to be practical, according to the theory of balanced incomplete block designs (see e.g. Rink, 1987), it can be reduced in such a way that all pairs of objects in tetrad sets are presented equally frequently, but less frequently than their potential maximum number. If λ denotes the number of tetrads in which each pair of objects occurs (e.g. using a $\lambda = 2$ design, each pair of objects appears together only twice in the questionnaire), then the reduced number of blocks L_λ equals:

$$L_\lambda = C_n^4 \cdot \frac{2\lambda}{(n-2)(n-3)} = \frac{\lambda n(n-1)}{12}, \tag{3}$$

and equation (3) must satisfy both of these defining relations (see e.g. Rink, 1987):

$$\begin{cases} nr = kL_\lambda \\ (n-1)\lambda = (k-1)r \end{cases}, \tag{4}$$

where:
k is the number of objects in one block (for tetrads $k = 4$),
r is the number of replications of each object in the reduced blocks,
$\lambda = 1, \dots, (n-1)(n-2)/2$.

The number of tetrads for different values of λ and n is shown in Table 1.

Table 1. Number of tetrads for different values of λ and n

n	λ						Full set of tetrads
	1	2	3	4	5	6	
6	.	.	.	.	.	15	15
7	.	7	.	14	.	21	35
8	.	.	14	.	.	28	70
9	.	.	18	.	.	36	126
10	.	15	.	30	.	45	210
11	.	.	.	.	.	55	330
12	.	.	33	.	.	66	495
13	13	26	39	52	65	78	715
14	.	.	.	.	.	91	1001
15	.	.	.	.	.	105	1365
16	20	40	60	80	100	120	1820
17	.	.	68	.	.	136	2380

Source: author's work.

As it is not possible to define a reduced number of blocks for all combinations of λ and n, not all the cells in Table 1 are filled.

3.2. Determining the dissimilarities based on tetrads

It is possible to include the evaluation of paired comparisons into a similarity matrix. The creation of a triangular similarity matrix is possible by assigning the number of points to the pair of objects in the tetrads. The number of points is equal to the number of pairs, for which it can be assumed that the similarity is smaller than the similarity of a given pair. The number of points assigned to pairs from the set of hypothetical tetrads marked with the consecutive letters of the alphabet are presented in Table 2.

Table 2. Number of points for pairs of objects A, B, C, D in example tetrads

Tetrad	Pairs in tetrads					
	AB	AC	AD	BC	BD	CD
ABCD	5	1	0	3	1	3
CABD*	5	4	1	1	3	0
CBAD*	5	1	3	4	1	0
DABC*	5	1	4	3	1	0
DBAC*	5	3	1	1	4	0

* The most similar objects are placed as the second and the third.
Note. The most similar pairs are marked in green, the second most similar pairs in light green, the least similar pairs in grey.
Source: author's work.

The value of element p_{ij} in the i-th row and the j-th column of the similarity matrix is equal to the sum of points awarded to a pair consisting of the i-th and the j-th objects in all blocks.

To discover the perceptual map by using nonmetric multidimensional scaling, the similarity matrix should be transformed into a matrix of dissimilarities. Dissimilarities δ_{ij} are determined in accordance with the formula:

$$\delta_{ij} = \begin{cases} 1 - \frac{p_{ij}}{\max r \cdot m_{ij}} \text{ for } m_{ij} \neq 0 \\ 0{,}5 \text{ for } m_{ij} = 0 \end{cases}, \tag{5}$$

where m_{ij} is the number of pairs (O_i, O_j) in blocks and $\max r$ is the maximum number of points that can be obtained by a pair of objects in a block. For tetrads, $\max r$ is equal to 5. The denominator in the second component for $m_{ij} \neq 0$ of equation (5) indicates the maximum possible number of points for pair (O_i, O_j), i.e. when in all blocks it was considered to be a pair of the most similar objects. Zaborski (2020, 2022) showed that owing to the use of formula (5), performing scaling on the basis of the incomplete tetrad method is possible, even when all pairs of objects cannot be presented equally frequently.

4. Results of the study

In order to make the study results independent of the respondents' subjective effects, the analysis was made on the basis of four distance matrices for 7, 9, 11, and 13 objects that were randomly selected voivodship capitals in Poland. The distances between cities were expressed in kilometres. Using the ALSCAL algorithm, multidimensional scaling was performed for each matrix. As a result, four configurations of n points representing cities ($n = 7, 9, 11, 13$) were obtained. Moreover, for each n, the distances of all pairs of objects (cities) were ranked from the smallest to the largest.

To determine the minimum number of tetrads necessary to recreate the known structure of the objects, for each set of n objects, 18 blocks of L_n tetrads were generated (three different blocks for each L_n value). For $n = 7$, the number of tetrads in a block was $L_7 = 4, 5, 6, 7, 9, 11$, for $n = 9$, $L_9 = 4, 6, 8, 10, 12, 14$, for $n = 11, L_{11} = 10, 11, 12, 13, 14, 16$ and for $n = 13, L_{13} = 7, 9, 11, 13, 15, 17$. As a result, 72 sets of tetrads were obtained. All L_n-element blocks, for which tetrads were created, were generated using the `ibd` function of the `ibd` package in the R programme (Mandal, 2019).

For each set of objects and for each block of tetrads, similarity matrices were calculated based on the ranking of distances for all pairs of objects. Next, they were transformed into a dissimilarity matrix according to formula (5) and multi-dimensional scaling was performed with the use of the MINISSA programme. MINISSA performs the basic model of nonmetric multidimensional scaling and it is available in the New MDSX multidimensional scaling package (Coxon & Davies, 1982).

The Procrustes statistic was used to test the quality of matching the resulting configurations of points to the configuration determined based on the distance matrices (Borg & Groenen, 2005; Cox & Cox, 2001):

$$R^2 = \frac{\left\{\mathrm{tr}\left(\mathbf{X}^{*\mathrm{T}}\mathbf{Y}\mathbf{Y}^{\mathrm{T}}\mathbf{X}^{*}\right)^{\frac{1}{2}}\right\}^2}{\mathrm{tr}(\mathbf{X}^{*\mathrm{T}}\mathbf{X}^{*})\mathrm{tr}(\mathbf{Y}^{\mathrm{T}}\mathbf{Y})}, \qquad (6)$$

where:

$\mathbf{X}^* = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{Y}\mathbf{Y}^{\mathrm{T}}\mathbf{X})^{\frac{1}{2}}(\mathbf{Y}^{\mathrm{T}}\mathbf{X})^{-1}$ – optimally rotated configuration $\mathbf{X}$, where:

$\mathbf{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n]^{\mathrm{T}}$ – the configuration of points determined on the basis of the incomplete blocks,

$\mathbf{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n]^{\mathrm{T}}$ – the configuration of points determined on the basis of the distance matrix. $R^2 \in (0, 1\rangle$, where 1 means a perfect match.

The quality of the matching configurations of points obtained on the basis of T_L^k (T_L^k – k-th set of L tetrads) to the configuration obtained on the basis of the distance matrix tested with the Procrustes statistics is presented in Table 3.

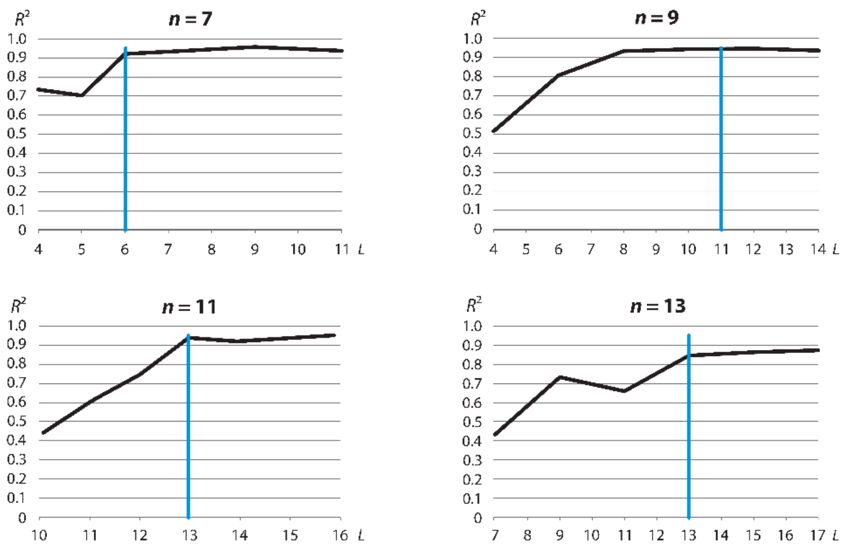
Table 3. Procrustes statistics for different sets of tetrads

T_L^k	R^2	T_L^k	R^2	T_L^k	R^2	T_L^k	R^2
***n* = 7**		***n* = 9**		***n* = 11**		***n* = 13**	
T_{11}^1	0.91709	T_{14}^1	0.93574	T_{16}^1	0.96095	T_{17}^1	0.92804
T_{11}^2	0.96246	T_{14}^2	0.92967	T_{16}^2	0.93460	T_{17}^2	0.91933
T_{11}^3	0.93234	T_{14}^3	0.94632	T_{16}^3	0.96023	T_{17}^3	0.88515
$\overline{T_{11}^k}$	0.93730	$\overline{T_{14}^k}$	0.93724	$\overline{T_{16}^k}$	0.95193	$\overline{T_{17}^k}$	0.91084
T_9^1	0.97065	T_{12}^1	0.96024	T_{14}^1	0.93327	T_{15}^1	0.92088
T_9^2	0.98040	T_{12}^2	0.94652	T_{14}^2	0.92061	T_{15}^2	0.91940
T_9^3	0.92021	T_{12}^3	0.93471	T_{14}^3	0.90692	T_{15}^3	0.91683
$\overline{T_9^k}$	0.95709	$\overline{T_{12}^k}$	0.94716	$\overline{T_{14}^k}$	0.92027	$\overline{T_{15}^k}$	0.91904
T_7^1	0.96827	T_{10}^1	0.96051	T_{13}^1	0.94660	T_{13}^1	0.90023
T_7^2	0.91936	T_{10}^2	0.93733	T_{13}^2	0.93274	T_{13}^2	0.91622
T_7^3	0.91009	T_{10}^3	0.93470	T_{13}^3	0.93682	T_{13}^3	0.89817
$\overline{T_7^k}$	0.93257	$\overline{T_{10}^k}$	0.94418	$\overline{T_{13}^k}$	0.93872	$\overline{T_{13}^k}$	0.90487
T_6^1	0.90302	T_8^1	0.94211	T_{12}^1	0.74459	T_{11}^1	0.73392
T_6^2	0.92113	T_8^2	0.95326	T_{12}^2	0.67872	T_{11}^2	0.69141
T_6^3	0.93950	T_8^3	0.90749	T_{12}^3	0.81894	T_{11}^3	0.66545
$\overline{T_6^k}$	0.92122	$\overline{T_8^k}$	0.93429	$\overline{T_{12}^k}$	0.74742	$\overline{T_{11}^k}$	0.69693
T_5^1	0.74699	T_6^1	0.83260	T_{11}^1	0.68085	T_9^1	0.82126
T_5^2	0.65601	T_6^2	0.89545	T_{11}^2	0.51807	T_9^2	0.67311
T_5^3	0.70576	T_6^3	0.69519	T_{11}^3	0.62874	T_9^3	0.89331
$\overline{T_5^k}$	0.70292	$\overline{T_6^k}$	0.80775	$\overline{T_{11}^k}$	0.60922	$\overline{T_9^k}$	0.79589
T_4^1	0.56308	T_4^1	0.82258	T_{10}^1	0.47152	T_7^1	0.41954
T_4^2	0.84216	T_4^2	0.48776	T_{10}^2	0.39246	T_7^2	0.34305
T_4^3	0.79820	T_4^3	0.23341	T_{10}^3	0.46272	T_7^3	0.31264
$\overline{T_4^k}$	0.73448	$\overline{T_4^k}$	0.51458	$\overline{T_{10}^k}$	0.44223	$\overline{T_7^k}$	0.35841

Note. $\overline{T_L^k}$ – mean of T_L^k, where $k = (1, 2, 3)$, L – number of tetrads in a block.
Source: author's calculations based on the ranking of the distances between all pairs of objects (cities).

Figure 1 shows how, depending on the number of tetrads, the mean values of the Procrustes statistics change for the different number of objects included in the study. The vertical line in the graphs of Figure 1 indicates the minimal number of tetrads generated so that each pair of objects appears in the set at least once.

Figure1. Relationship between the number of tetrads for different n and the value of R^2



Note. L – number of tetrads in a block. The vertical line indicates the minimal number of tetrads in a set, where each pair occurs at least once.
Source: author's calculations based on the ranking of the distances between all pairs of objects (cities).

Based on the test results presented in Table 3 and Figure 1, it can be concluded that the matching of the configurations obtained on the basis of tetrads to the configurations obtained on the basis of the distance matrix is very good if each pair of objects appears in the tetrad set at least once. The mean values of the Procrustes statistics in the study for these tetrad sets range from 0.905 for the set of 13 tetrads for 13 objects to 0.957 for the set of 9 tetrads for 7 objects, with a coefficient of variance equaling 0.00047. The value of the matching function clearly decreases only when each pair of objects does not appear at least once in the tetrad set.

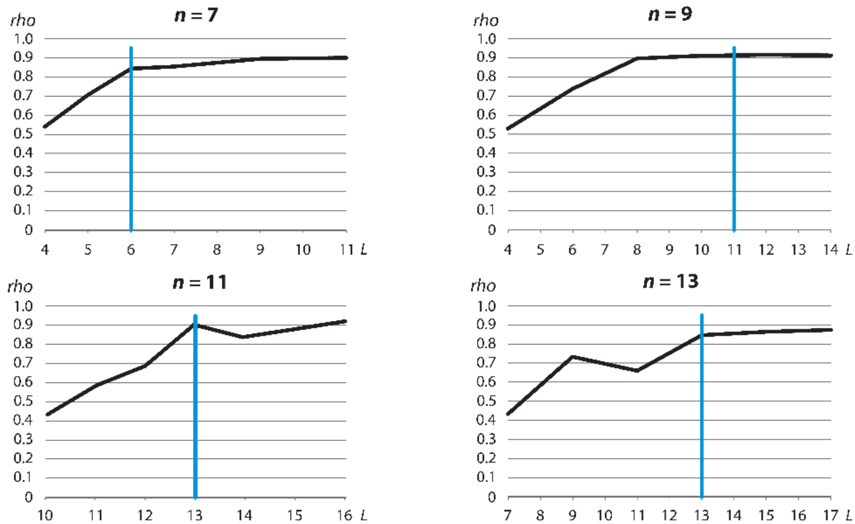
In the configuration of points obtained as a result of nonmetric multidimensional scaling, it is not the distances between the points that are important, but their rank order. Therefore, for all point configurations obtained by the tetrad method, the distances between points were ranked from the smallest to the largest. Then, the Spearman rank correlations for all distances obtained on the basis of tetrads to the distances for the input data were determined. The analysis results are presented in Table 4 and Figure 2. All the correlations presented in Table 4 are statistically significant.

Table 4. Spearman rank correlations for different sets of tetrads

T_L^k	Spearman *rho*	T_L^k	Spearman *rho*	T_L^k	Spearman *rho*	T_L^k	Spearman *rho*
***n* = 7**		***n* = 9**		***n* = 11**		***n* = 13**	
T_{11}^1	0.88182	T_{14}^1	0.91762	T_{16}^1	0.93088	T_{17}^1	0.89380
T_{11}^2	0.91299	T_{14}^2	0.90147	T_{16}^2	0.89387	T_{17}^2	0.90078
T_{11}^3	0.90650	T_{14}^3	0.91873	T_{16}^3	0.92193	T_{17}^3	0.82757
$\overline{T_{11}^k}$	0.90044	$\overline{T_{14}^k}$	0.91261	$\overline{T_{16}^k}$	0.91556	$\overline{T_{17}^k}$	0.87405
T_9^1	0.87403	T_{12}^1	0.92517	T_{14}^1	0.87684	T_{15}^1	0.86280
T_9^2	0.94026	T_{12}^2	0.91586	T_{14}^2	0.85729	T_{15}^2	0.86525
T_9^3	0.87143	T_{12}^3	0.90672	T_{14}^3	0.78023	T_{15}^3	0.86259
$\overline{T_9^k}$	0.89524	$\overline{T_{12}^k}$	0.91592	$\overline{T_{14}^k}$	0.83812	$\overline{T_{15}^k}$	0.86355
T_7^1	0.92468	T_{10}^1	0.92513	T_{13}^1	0.91320	T_{13}^1	0.82875
T_7^2	0.84156	T_{10}^2	0.90522	T_{13}^2	0.90433	T_{13}^2	0.86181
T_7^3	0.79741	T_{10}^3	0.90253	T_{13}^3	0.89279	T_{13}^3	0.84767
$\overline{T_7^k}$	0.85455	$\overline{T_{10}^k}$	0.91096	$\overline{T_{13}^k}$	0.90344	$\overline{T_{13}^k}$	0.84608
T_6^1	0.89351	T_8^1	0.91946	T_{12}^1	0.72965	T_{11}^1	0.67137
T_6^2	0.84156	T_8^2	0.89762	T_{12}^2	0.59322	T_{11}^2	0.63571
T_6^3	0.79740	T_8^3	0.87264	T_{12}^3	0.73889	T_{11}^3	0.67228
$\overline{T_6^k}$	0.84416	$\overline{T_8^k}$	0.89657	$\overline{T_{12}^k}$	0.68725	$\overline{T_{11}^k}$	0.65979
T_5^1	0.78182	T_6^1	0.78251	T_{11}^1	0.58146	T_9^1	0.70356
T_5^2	0.64416	T_6^2	0.77982	T_{11}^2	0.52460	T_9^2	0.65961
T_5^3	0.69221	T_6^3	0.65137	T_{11}^3	0.65202	T_9^3	0.83685
$\overline{T_5^k}$	0.70606	$\overline{T_6^k}$	0.73790	$\overline{T_{11}^k}$	0.58603	$\overline{T_9^k}$	0.73334
T_4^1	0.28442	T_4^1	0.69294	T_{10}^1	0.45203	T_7^1	0.50639
T_4^2	0.64416	T_4^2	0.57243	T_{10}^2	0.39271	T_7^2	0.36737
T_4^3	0.69221	T_4^3	0.32431	T_{10}^3	0.45733	T_7^3	0.42828
$\overline{T_4^k}$	0.54026	$\overline{T_4^k}$	0.52989	$\overline{T_{10}^k}$	0.43402	$\overline{T_7^k}$	0.43401

Note. As in Table 3.
Source: author's calculations based on the ranking of the distances between all pairs of objects (cities).

Figure 2. Relationship between the number of tetrads for different *n* and Spearman's rank correlation coefficient



Note. As in Figure 1.
Source: author's calculations based on the ranking of the distances between all pairs of objects (cities).

The conclusions of the correlation analysis are in line with those of Procrustes analysis. If each pair appears in the tetrad set at least once, the ranks are in strong agreement. For the analysed examples, when each pair appears in the tetrad set at least once, the mean value of Spearman's rank correlation coefficient ranges from 0.844 for the set of 6 tetrads for 7 objects to 0.916 for the set of 12 tetrads for 9 objects. The coefficient of variation for these results is 0.001571, which proves that the selection of the tetrad set for a given L_n is not significant.

5. Conclusions

The tetrad method can be classified as one of the group of methods in which dissimilarities are assessed on the basis of comparisons with pairs of objects. These methods require the respondents to make many assessments, causing fatigue and weariness. as a result of which the obtained results do not always fully reflect the respondents' attitudes.

The study confirms that, despite the above, the tetrad method can be an alternative to obtaining dissimilarity data when the variables are measured on an ordinal scale. It has been shown that satisfactory results in determining dissimilarities by means of the method of tetrads under incomplete block designs can be obtained already at the point when each pair of objects appears in the tetrad set at least once. This allows for a significant reduction in the number of multiple opinions provided by the respondents.

The results of the study also indicate that the choice of the incomplete set of tetrads has no significant effect on the results of nonmetric multidimensional scaling.

Despite the advantages, the main limitation of the incomplete tetrad method, as well as most methods of direct dissimilarity determination, is the possibility of using it for a large number of objects. For even a dozen or so objects, the need for respondents to make assessments may cause fatigue and boredom. In such a situation, the solution may be to aggregate the results of different respondents (or groups of respondents), provided that different subsets of the full set of objects are presented to individual respondents.

References

Bijmolt, T. H. A. (1996). *Multidimensional Scaling in Marketing: Towards Integrating Data Collection and Analysis*. Capelle a/d Ussel: Labyrint Publication.

Borg, I., & Groenen, P. J. F. (2005). *Modern Multidimensional Scaling. Theory and Applications* (2nd edition). Springer. https://doi.org/10.1007/0-387-28981-X.

Burton, M. L., & Nerlove, S. B. (1976). Balanced Design for Triads Tests: Two Examples from English. *Social Science Research, 5*(3), 247–267. https://doi.org/10.1016/0049-089X(76)90002-8.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*, *35*(3), 283–319. https://doi.org/10.1007/BF02310791.

Cox, T. F., & Cox, M. A. A. (2001). *Multidimensional Scaling* (2nd edition). Chapman & Hall/CRC. https://doi.org/10.1201/9780367801700.

Coxon, A. P. M, & Davies, P. M. (1982). *The User's Guide to Multidimensional Scaling with Special Reference to the MDS(X) Library of Computer Programs.* Heinemann Educational Books.

Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, *33*(4), 469–506. https://doi.org/10.1007/BF02290164.

Humphreys, M. A. (1982). Data collecting effects on nonmetric multidimensional scaling solutions. *Educational and Psychological Measurement*, *42*(4), 1005–1022. https://doi.org /10.1177/001316448204200408.

Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*(1), 1–27. https://doi.org/10.1007/BF02289565.

Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, *29*, 115–129. http://dx.doi.org/10.1007/BF02289694.

de Leeuw, J., & Heiser, W. (1980). Multidimensional scaling with restrictions on the configuration. In P. R. Krishnaiah (Ed.), *Multivariate Analysis V* (pp. 501–522). North-Holand.

Mandal, B. N. (2019). *ibd: Incomplete Block Designs*. www.r-project.org.

Morris, M. D. (2010). *Design of Experiments. An Introduction Based on Linear Models*. Chapman & Hall/CRC.

Rink, D. R. (1987). An improved preference data collection method: Balanced incomplete block designs. *Journal of the Academy of Marketing Science*, *15*(3), 54–61. https://doi.org/10.1007/BF02721954.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. Part I. *Psychometrika*, *27*, 125–140. https://doi.org/10.1007/BF02289630.

Takane, Y., Young, F. W., & de Leeuw, J. (1977). Nonmetric individual differences in multidimensional scaling: an alternating least squares method with optimal scaling features. *Psychometrika*, *42*, 7–67. https://doi.org/10.1007/BF02293745.

Tsogo, L., Masson, M. H., & Bardot, A. (2000). Multidimensional Scaling Methods for Many-Object Sets: A Review. *Multivariate Behavioral Research*, *35*(3), 307–319. https://doi.org/10.1207/S15327906MBR3503_02.

Zaborski, A. (2003). Wpływ alternatywnych metod pomiaru preferencji na wyniki skalowania wielowymiarowego. *Prace Naukowe Akademii Ekonomicznej w Katowicach: Analiza i prognozowanie zjawisk o charakterze niemetrycznym,* 55–69.

Zaborski, A. (2020). The Use of the Incomplete Tetrad Method for Measuring the Similarities in Nonmetric Multidimensional Scaling. *Folia Oeconomica Stetinensia, 20*(1), 519–530. https://doi.org/10.2478/foli-2020-0030.

Zaborski, A. (2022). Triads or tetrads? Comparison of two methods for measuring the similarity in preferences under incomplete block design. *Statistics in Transition new series*, *23*(3), 185–198. https://doi.org/10.2478/stattrans-2022-0037.