A C T A   U N I V E R S I T A T I S   L O D Z I E N S I S
FOLIA OECONOMICA 3(314)2015
http://dx.doi.org/10.18778/0208-6018.314.04

*Artur Zaborski**

SWITCH PREFERENCE ANALYSIS BY THE DRIFT VECTORS METHOD

Abstract. The matrix of switch preference data (e.g. one's preference for brand *j*, given that one has already defined his/her first choice for brand *i*) is not symmetric. The averaging of appropriate off-diagonal proximity entries for such data leads to the loss of information, hence the necessity to use appropriate methods for asymmetric data. Among the chosen methods of asymmetric multidimensional scaling, particular attention was paid to the drift vectors method. This method enables to present simultaneously on the perceptual map both the configuration of points representing the analyzed objects and the vectors indicating the direction and the strength of changes in the respondents' preferences.
Keywords: preference analysis, multidimensional scaling, asymmetric data, drift vectors.

1. INTRODUCTION

In multidimensional scaling the data matrix that underlies the analysis of perceptual judgments is usually square and symmetric. It is a matrix which represents a set of empirically obtained similarities or proximities between pairs of objects. In preference research proximities may by asymmetric, and therefore they cannot always be fully represented by the distances between points in multidimensional space. If the proximities deviate from being symmetric due to an error only, the most common procedure is to average off-diagonal matrix elements in accordance with the formula:

$$p_{ij}^{*} = \frac{p_{ij} + p_{ji}}{2}$$

However, there is a broad class of data in which the matrix of dissimilarities is not intended to satisfy the condition of symmetry (see: Holyoak, Gordon (1983), Tversky, Gati (1982), Harshman *et al*. (1982), Chino (1978)). Examples include:

* Ph.D., Chair of Econometrics and Computer Science, Wroclaw Univeristy of Economics.

- trade flows between countries,
- the number of people declaring that the object *j* is their top priority, even though in the earlier period it was the object *i*,
- the frequency with which brand *j* is incorrectly perceived to be brand *i* in testing of a new package design,
- the likelihood that the consumer is purchasing brand *j* on condition that on the previous purchases it was brand *i*.

For such data matrix symmetrisation by averaging the corresponding amounts of dissimilarities leads to the loss of valuable information. Thus, it is necessary to apply appropriate methods for asymmetric data.

Tha aim of the article is the analysis of the changes that have occurred in the preferences of the University of the Third Age members. First, it is shown that asymmetric proximity matrix can be decomposed into a symmetric and a skew-symmetric part. Within the scope of multidimensional scaling methods for asymmetric data, for the analysis the drift vectors method was used. This method allows the presentation on a perceptual map both symmetric and skew-symmetric components of the proximity matrix.

2. THE DIVISION OF METHODS FOR ASYMMETRIC DATA

Each asymmetric proximity square matrix can be decomposed by presentation as the sum of the matrices:

$$\mathbf{P} = \mathbf{A} + \mathbf{B},$$

where:

$\mathbf{A} = \dfrac{(\mathbf{P} + \mathbf{P}^{\mathbf{T}})}{2}$ – symmetric matrix,

$\mathbf{B} = \dfrac{(\mathbf{P} - \mathbf{P}^{\mathbf{T}})}{2}$ – skew-symmetric matrix.

As for such decomposition $\sum_{i,j} p_{ij}^2 = \sum_{i,j} a_{ij}^2 + \sum_{i,j} b_{ij}^2$, the analysis of the data contained in the asymmetric proximity matrix may consist of two parts: the analysis of the symmetrical components and the analysis of the skew-symmetric components. Hence, the division of multidimensional scaling methods for asymmetric data into those, in which the analysis will be made only on the basis of an asymmetric matrix (e.g. unfolding, slide vector model, hill-climbing model, dominance point model),and those, which are based on a skew-symmetric matrix or on the symmetric and the skew-symetric matrix (e.g. signed-distance model, radius-distance model, drift vectors model).

Several algorithms and models of multidimensional scaling for asymmetric data have been introduced in the last three decades. Some of them deal with two-mode three-way(?) proximities (see e.g.: Zielman, Heiser (1993), DeSarbo et al. (1992), Okada, Imaizumi (1997)), and some deal with one-mode two-way proximities (see e.g.: Borg, Groenen (2005), Okada, Imaizumi (2007)). The method presented in the next section belongs to the second group of methods.

3. THE DRIFT VECTORS METHOD

In the drift vectors method the symmetric part and the skew-symmetric part of the preference data are simultaneously displayed on the perceptual map. The symmetric part of the data is depicted by non-metric multidimensional scaling representation of the symmetrized data. The skew-symmetric values are embedded into the configuration of points representing objects by drawing arrows (drift vectors) from each point to any other point. These vectors correspond in length and direction to the values in appropriate rows of the skew-symmetric matrix and show the power and the direction of preference changes. The length and the direction angle of the drift vectors in two-dimensional space are computed as follows (see: Borg, Groenen (2005):

1. Compute matrix **A** for the asymmetric proximity matrix **P**, where: $a_{ij} = \dfrac{p_{ij} + p_{ji}}{2}$.
2. Perform non-metric multidimensional scaling for the matrix **A**.
3. Construct vectors $\mathbf{y}_{ij} = \mathbf{x}_j - \mathbf{x}_i$ for all points $\mathbf{x}_i$ and $\mathbf{x}_j$ ($i \neq j$) obtained by multidimensional scaling.
4. Norm $\mathbf{y}_{ij}$ to unit length to get $\mathbf{y}^*_{ij} = \mathbf{y}_{ij} / (\mathbf{y}^T_{ij}\mathbf{y}_{ij})^{1/2}$.
5. Multiply $\mathbf{y}^*_{ij}$ by elements b_{ij} of the skew-symmetric component of the proximity matrix to obtain $\mathbf{z}_{ij} = b_{ij} \cdot \mathbf{y}^*_{ij}$.
6. Average the n vectors $\mathbf{z}_{ij}$ to obtain the drift vector for point i: $\mathbf{w}_i = n^{-1}\sum_j \mathbf{z}_{ij}$.
7. Compute $\mathbf{w}_i$'s length as: $\|\mathbf{w}_i\| = \sqrt{w_{i1}^2 + w_{i2}^2}$.

8. Compute the direction angle relative to the Y-axis:

$$\alpha_i = \arccos\left(\mathbf{w}_i \mathbf{u}^T / \sqrt{\mathbf{w}_i \cdot \mathbf{w}_i^T}\right), \text{ where } \mathbf{w}_i = (w_{i1}, w_{i2}),\ \mathbf{u} = (0,1).$$

4. EMPIRICAL EXAMPLE

The aim of the study was the analysis of changes that have occurred in the preferences of the University of the Third Age members in Boleslawiec (Lower Silesia province). The study compared the students' choices regarding the willingness to participate in certain forms of classes, which are: English, German, computer skills, gymnastics, yoga, swimming, weight training, nordic walking, painting and handicraft. The study involved 109 people who consistently participated in activities whithin a definite period of time.

Table 1 shows changes which occurred in the choices made by the University members in the above activities in each year from 2009 to 2013. Elements of the i-th row and j-th column show the number of members who at the beginning of the period chose i-th class, and the j-th class at the end of the period. The presented matrix should be considered as a similarity matrix because the high value of an element means that students willingly switch activities, and in accordance with their preference, they treat them as similar.

Table 1. Switching data among classes in subsequent years

Classes		Engl.	Germ.	Comp.	Gym.	Yoga	Swim.	Weight	Nord.	Paint.
1	2	3	4	5	6	7	8	9	10	11
12009 – 2010	English	22	0	1	1	3	0	0	0	0
	German	1	15	2	2	4	1	1	1	0
	Computer skills	2	0	5	4	5	4	4	0	1
	Gymnastics	0	0	0	50	4	2	2	1	1
	Yoga	0	0	0	0	9	3	1	0	1
	Swimming	0	0	0	2	0	22	2	2	0
	Weight training	0	0	0	0	0	0	5	0	0
	Nordic walking	0	0	0	0	0	0	0	8	0
	Painting and handicraft	0	0	2	0	0	0	0	1	7
2010 – 2011	English	14	0	0	3	0	2	1	0	0
	German	0	9	1	1	2	4	0	1	1
	Computer skills	0	0	11	1	0	2	1	1	0
	Gymnastics	0	0	0	56	2	2	0	1	0
	Yoga	0	0	0	1	18	4	0	2	0
	Swimming	1	0	0	1	2	25	1	2	0
	Weight training	0	0	0	0	1	1	8	0	0
	Nordic walking	0	0	0	0	0	0	0	7	0
	Painting and handicraft	0	0	0	0	0	0	0	0	8

Table 1 (cont.)

1	2	3	4	5	6	7	8	9	10	11
2011 – 2012	English	19	1	0	0	1	0	0	0	0
	German	0	8	0	1	1	0	0	0	0
	Computer skills	0	0	7	2	2	2	0	0	0
	Gymnastics	0	0	1	55	4	5	3	0	0
	Yoga	1	0	0	1	17	1	0	2	0
	Swimming	0	0	0	0	0	41	1	0	0
	Weight training	0	0	0	1	0	0	9	0	0
	Nordic walking	0	0	0	0	1	1	0	12	0
	Painting and handicraft	0	0	0	0	0	0	0	0	12
2012 – 2013	English	20	0	0	1	0	1	0	1	0
	German	0	8	0	0	0	1	0	0	0
	Computer skills	0	0	10	1	0	0	2	0	0
	Gymnastics	0	0	1	48	3	2	3	2	0
	Yoga	0	0	0	0	15	1	0	2	0
	Swimming	0	0	0	3	1	47	2	1	0
	Weight training	0	0	0	0	0	3	11	0	0
	Nordic walking	0	0	0	1	0	0	0	11	0
	Painting and handicraft	0	0	0	0	0	0	0	0	9

Source: own elaboration.

After the averaging of corresponding non-diagonal elements of the proximity matrix, non-metric multidimensional scaling was performed using PROXSCAL algorithm for similarity data for every period, and on the obtained configurations of points drift vectors were plotted (see: Figure 1).

The points distribution on perceptual maps shows that from 2009 to 2012 it is difficult to notice the similarity in preferences with respect to classes included in the study. However, on the basis of the drift vectors it can be concluded that the preferences of the University's members turn in the direction of movement activities. Instead of foreign languages and computer skills, activities which improve physical fitness are increasingly preferred. It is particularly evident in the period 2009–2010, when there is a clearly visible trend of switching preferences towards swimming and weight training. This may be explained by the fact that it was the first year when students could choose these classes. In the period 2012–2013 there are shown two groups of classes considered as similar in terms of preferences. The first one includes all physical activities, and the second comprises English, German, computer skills, painting and handicraft. However, there is still a visible trend of changes in the preferences to classes belonging to the first group.

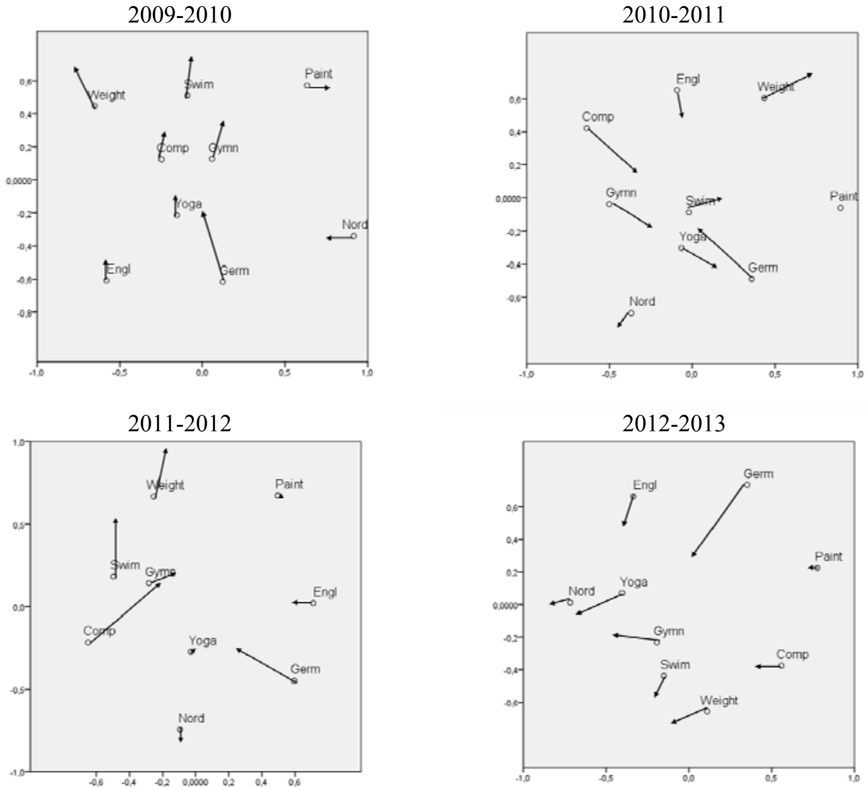

Figure 1. Classes configurations with drift vectors

Source: own elaboration.

5. FINAL REMARKS

As researchers become aware of the importance of the asymmetry in proximity relationships in preference research, several models and algorithms of multidimensional scaling for asymmetric data have been introduced. In the paper the algorithm and the use of the drift vectors method for switch preferences analysis was presented. The undoubted advantage of the drift vectors method is that unlike many other asymmetric multidimensional scaling methods it has a simple geometrical interpretation. In the model each object (form of classes) is represented as a point in multidimensional space, and drift vectors, which are also embedded in the same multidimensional space, show the power and the direction of the preference changes.

The directions of the drift vectors obtained in the study indicate that the preferences of University's members are moving towards physical activities. In place of classes which develop foreign language and computer skills, classes which improve physical fitness are increasingly preferred.

REFERENCES

Borg, I., Groenen, P. (2005), *Modern multidimensional scaling. Theory and applications. Second Edition*, Springer-Verlag, New York.

Chino N. (1978), A graphical technique for representing the asymmetric relationship between N objects, *Behaviometrika*, no 5, 23–40.

DeSarbo W. S., Johnson M.D., Manrai A.K., Manrai L.A., Edward E.A. (1992), TSCALE: A New Multidimensional Scaling Procedure Based on Tversky's Contrast Model, *Psychometrika*, 57, 43–69.

Harshman R.A., Green P.E., Wind Y., Lundy M.E. (1982), A model for the analysis of asymmetric data in marketing research, *Marketing Science*, vol. I, no 2, 205–242.

Holyoak K.J., Gordon P.C. (1983), Social reference points, *Journal of Personality and Social Psychology*, no 44, 881–887.

Okada A., Imaizumi T. (1987), Nonmetric multidimensional scaling of asymmetric proximities, *Behaviometrika*, no 21, 81–96.

Okada A., Imaizumi T. (2007), Multidimensional scaling of asymmetric proximities with a dominance point, *Advances in Data Analysis Studies in Classification, Data Analysis, and Knowledge Organization*, (red.) R. Decker, H.J. Lenz, Springer-Verlag, Berlin, Heidelberg, 307–318.

Tversky A., Gati I. (1982), Features of similarity, *Psychological Review*, no 89, 123–154.

Zielman B., Heiser W.J. (1996), Analysis of Asymmetry by a Slide-Vector, *Psychometrika*, 58, 101–114.

Artur Zaborski

ANALIZA ZMIAN PREFERENCJI Z WYKORZYSTANIEM METODY WEKTORÓW DRYFU

Streszczenie. Punktem wyjścia w skalowaniu wielowymiarowym jest symetryczna macierz niepodobieństw. Jednak macierz danych o zmianach preferencji (np. prawdopodobieństwo, że konsument dokonuje zakupu marki *j* pod warunkiem, że przy wcześniejszych zakupach była to marka *i*, liczba osób deklarujących, że marka *j* jest przez nich najbardziej preferowana, mimo że we wcześniejszym okresie była to marka *i* i in.) nie jest symetryczna. Dla takich danych uśrednienie odpowiednich wartości niepodobieństw prowadzi do utraty cennych informacji dotyczących analizowanego zjawiska, stąd konieczność stosowania metod właściwych dla danych niesymetrycznych. Spośród wybranych metod niesymetrycznego skalowania wielowymiarowego szczególną uwagę zwrócono na metodę wektorów dryfu. Metoda ta umożliwia równoczesną prezentację na mapie percepcyjnej konfiguracji punktów reprezentujących analizowane obiekty, jak również wektorów wskazujących kierunek i siłę zmian zachodzących w preferencjach respondentów.

Słowa kluczowe: badania preferencji, skalowanie wielowymiarowe, dane niesymetryczne, wektory dryfu.