

Artur Zaborski

Uniwersytet Ekonomiczny we Wrocławiu

ANALIZA PREFERENCJI SŁUCHACZY UNIwersYTETU TRZECIEGO WIEKU Z WYKORZYSTANIEM WYBRANYCH METOD NIESYMETRYCZNEGO SKALOWANIA WIELOWYMIAROWEGO

Wprowadzenie

Od ukazania się w 1962 roku pionierskich publikacji Shepada¹ dotyczących procedury skalowania wielowymiarowego metoda ta znalazła szerokie zastosowanie w wielu dziedzinach badawczych, a zwłaszcza w badaniach marketingowych.

Skalowanie wielowymiarowe jest zbiorem technik mających na celu prezentację w przestrzeni r -wymiarowej (najczęściej jest to przestrzeń dwu- lub trójwymiarowa) zbioru obiektów $O = \{O_1, \dots, O_n\}$ traktowanych jako punkty w przestrzeni wielowymiarowej. Przykładami obiektów mogą być np.: konsument, produkt, marka, przedsiębiorstwo, rynek testowy i inne.

Punkt wyjścia w skalowaniu wielowymiarowym stanowi zazwyczaj symetryczna macierz zawierająca $\frac{n(n-1)}{2}$ elementów δ_{ij} , obrazujących niepodobieństwa pomiędzy ocenianymi obiektami O_i oraz O_j , gdzie $i, j = 1, 2, \dots, n$ (n – liczba obiektów). Dla danego zbioru obiektów oraz niepodobieństw δ_{ij} poszukuje się takiego odwzorowania zbioru obiektów w zbiór punktów w przestrzeni r -wymiarowej, aby dopasowanie konfiguracji punktów w przestrzeni wielowymiarowej i r -wymiarowej było możliwie najlepsze.

¹ R.N. Shepard: Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function. Part I and II. „Psychometrika” 1962, Vol. 27, No. 1, s. 125-140; Vol. 27, No. 2, s. 219-246.

W badaniach preferencji macierz niepodobieństw często nie jest symetryczna. Jeżeli niesymetryczność wynika np. z błędu pomiaru lub z tendencyjności ocen respondentów, najbardziej popularnym zabiegiem jest uśrednienie nie-diagonalnych elementów macierzy zgodnie z formułą:

$$\delta_{ij}^* = \delta_{ji}^* = \frac{\delta_{ij} + \delta_{ji}}{2}. \quad (1)$$

Jednak istnieje szeroka gama danych, dla których macierz niepodobieństw z założenia nie spełnia warunku symetryczności. Są nimi np.²:

- wielkość wymiany handlowej między krajami,
- liczba osób deklarujących, że obiekt j jest przez nich najbardziej preferowany, mimo że we wcześniejszym okresie był to obiekt i ,
- częstotliwość, z jaką marka j jest błędnie postrzegana jako marka i w testach projektów nowych opakowań,
- prawdopodobieństwo, że konsument dokonuje zakupu marki j pod warunkiem, że przy wcześniejszych zakupach była to marka i .

Dla takich danych symetryzacja macierzy poprzez uśrednienie odpowiednich wielkości niepodobieństw prowadzi do utraty cennych informacji dotyczących analizowanego zjawiska. Tym samym konieczne jest zastosowanie metod właściwych dla danych niesymetrycznych.

Do analizy zmian preferencji słuchaczy Uniwersytetu Trzeciego Wieku w pracy zastosowano dwie metody niesymetrycznego skalowania wielowymiarowego: grawitacyjną analizę *unfolding*, w której wiersze i kolumny macierzy niepodobieństw traktowane są jako odrębne punkty, oraz metodę wektorów dryfu pozwalającą na wskazanie kierunku i siły zmian w preferencjach.

1. Cel i przedmiot badań

Celem badania była analiza zmian zachodzących w preferencjach słuchaczy Uniwersytetu Trzeciego Wieku w Bolesławcu (woj. dolnośląskie). W badaniu porównano wybory, jakich dokonywali słuchacze odnośnie do chęci uczestnictwa w określonych formach zajęć w roku akademickim 2009/2010 oraz 2013/2014.

² Zob. m.in.: K.J. Holyoak, P.C. Gordon: Social reference points. „Journal of Personality and Social Psychology” 1983, No. 44, s. 881-887; A. Tversky, I. Gati: Features of similarity. „Psychological Review” 1982, No. 89, s. 123-154; R.A. Harshman, P.E. Green, Y. Wind, M.E. Lundy: A model for the analysis of asymmetric data in marketing research. „Marketing Science” 1982, Vol. I, No. 2, s. 205-242; N. Chino: A graphical technique for representing the asymmetric relationship between N objects. „Behaviormetrika” 1978, No. 5, s. 23-40.

W badaniu wzięto pod uwagę zajęcia, na które słuchacze uczęszczają zgodnie z dokonaniem przez siebie wyborem. Są to: j. angielski, j. niemiecki, obsługa komputera, gimnastyka rehabilitacyjna, gimnastyka chińska i joga, zajęcia na basenie oraz marsze z kijkami. Nie uwzględniono cyklicznych otwartych wykładów tematycznych oraz specjalistycznych zajęć rozwijających zainteresowania i umiejętności słuchaczy, takich jak: warsztaty malarskie, rękodzieło czy warsztaty muzyczne. Spośród wszystkich słuchaczy Uniwersytetu badaniem objęto 109 osób, które niezmiennie uczestniczyły w zajęciach w analizowanym okresie.

Tabela 1 wskazuje, jakie zmiany zaszły w dokonywanych przez słuchaczy wyborach zajęć w 2013 roku w stosunku do 2009 roku. Element i -tego wiersza oraz j -tej kolumny przedstawia liczbę słuchaczy, którzy w 2009 roku wybrali i -te zajęcia, zaś w 2013 roku j -te zajęcia. Przedstawioną macierz należy traktować jako macierz podobieństw, ponieważ duża wartość elementu macierzy oznacza, że słuchacze chętnie dokonują zamiany zajęć, a tym samym, zgodnie ze swoimi preferencjami, traktują je jako podobne.

Tabela 1

Macierz podobieństw między wybranymi zajęciami

	Rodzaje zajęć	a	b	c	d	e	f	g	m_i
a	j. angielski	17	0	0	3	4	4	3	31
b	j. niemiecki	0	8	0	2	7	10	3	30
c	obsługa komputera	1	3	6	5	4	12	1	32
d	gimnastyka rehabilitacyjna	0	0	0	45	4	6	2	57
e	gimnastyka chińska i joga	0	0	0	0	6	3	2	11
f	basen	0	0	0	1	0	29	1	31
g	nordic walking	0	0	0	1	0	0	9	10
	m_j	18	11	6	57	25	64	21	

Dla każdej niesymetrycznej macierzy kwadratowej można dokonać jej dekompozycji poprzez przedstawienie w postaci sumy macierzy:

$$\Delta = \mathbf{M} + \mathbf{N}, \quad (2)$$

gdzie \mathbf{M} jest macierzą symetryczną, a \mathbf{N} jest macierzą skośnosymetryczną. Ponieważ dla tak przeprowadzonej dekompozycji zachodzi zależność:

$$\sum_{ij} \delta_{ij}^2 = \sum_{ij} m_{ij}^2 + \sum_{ij} n_{ij}^2, \quad (3)$$

dlatego analiza danych zawartych w macierzy niesymetrycznej może składać się z dwóch elementów: analizy części symetrycznej oraz analizy części skośnosymetrycznej. Dla danych zawartych w tabeli 1 macierze \mathbf{M} i \mathbf{N} są równe:

$$\mathbf{M} = \begin{bmatrix} 17 & 0 & 0,5 & 1,5 & 2 & 2 & 1,5 \\ 0 & 8 & 1,5 & 1 & 3,5 & 5 & 1,5 \\ 0,5 & 1,5 & 6 & 2,5 & 2 & 6 & 0,5 \\ 1,5 & 1 & 2,5 & 4,5 & 2 & 3,5 & 1,5 \\ 2 & 3,5 & 2 & 2 & 6 & 1,5 & 1 \\ 2 & 5 & 3,5 & 3,5 & 1,5 & 2,9 & 0,5 \\ 1,5 & 1,5 & 1,5 & 1,5 & 1 & 0,5 & 9 \end{bmatrix}; \mathbf{N} = \begin{bmatrix} 0 & 0 & -0,5 & 1,5 & 2 & 2 & 1,5 \\ 0 & 0 & -1,5 & 1 & 3,5 & 5 & 1,5 \\ 0,5 & 1,5 & 0 & 2,5 & 2 & 6 & 0,5 \\ -1,5 & -1 & -2,5 & 0 & 2 & 2,5 & 0,5 \\ -2 & -3,5 & -2 & -2 & 0 & 1,5 & 1 \\ -2 & -5 & -6 & -2,5 & -1,5 & 0 & 0,5 \\ -1,5 & -1,5 & -0,5 & -0,5 & -1 & -0,5 & 0 \end{bmatrix}$$

Dla przedstawionego przykładu suma kwadratów podobieństw z wyłączeniem elementów diagonalnych wynosi 480, z czego 51,7% wyjaśnione jest przez symetryczną macierz \mathbf{M} , zaś 48,3% przez skośnosymetryczną macierz \mathbf{N} . Wynika z tego, że blisko połowa podobieństw jest wyjaśniana przez niesymetryczną część macierzy.

2. Grawitacyjna analiza *unfolding*

W klasycznej analizie *unfolding* przyjmuje się założenie, że preferencje wszystkich respondentów determinowane są przez takie same wymiary. Nie uwzględnia się jednak pewnych czynników, które mają wpływ na dokonywane wybory. Mogą to być np.: popularność zajęć, dostępność, umiejętności oraz sprawność fizyczna respondentów i in. Te dodatkowe czynniki w modelu grawitacyjnym określa się jako masy. W zależności od tego, czy są to czynniki wynikające z własności badanych obiektów czy respondentów, nazywa się je masą obiektu bądź masą respondenta.

Model grawitacyjny określony jest równaniem³:

$$f_{ij} = \varphi_i \left(\frac{M_j^\beta S_i^\alpha}{d_{ij}^2} \right), \quad (4)$$

gdzie:

f_{ij} – ocena preferencji i -tego respondenta względem j -tego obiektu,

φ_i – pewna monotonicznie niemalejąca funkcja dla i -tego respondenta,

$$d_{ij} = \sqrt{\sum_{a=1}^r (x_{ja} - y_{ia})^2} \text{ – odległość między } \mathbf{x}_j \text{ a } \mathbf{y}_i,$$

³ W.S. DeSarbo, J. Kim, S.C. Choi, M. Spaluding: A Gravity-Based Multidimensional Scaling Model for Deriving Spatial Structures Underlying Consumer Preference/Choice Judgments. „Journal of Consumer Research” 2002, Vol. 29, s. 91-100.

M_j – masa j -tego obiektu (np. udział w rynku, wartość marki),

S_i – masa i -tego respondenta (np. wielkość dochodu, lojalność względem marki),

β i χ – parametry masy.

W modelu grawitacyjnym przyjmujemy założenie, że oceny preferencji są wprost proporcjonalne do iloczynu mas obiektu i respondenta, a odwrotnie proporcjonalne do odległości między punktem respondenta \mathbf{y}_i a punktem obiektu \mathbf{x}_j na mapie percepcyjnej.

Ideę modelu grawitacyjnego wykorzystano do analizy zmian preferencji w czasie. W tym celu dane podobieństw p_{ij} zawarte w tabeli 1 przekształcono na niepodobieństwa δ_{ij} za pomocą wzoru:

$$\delta_{ij} = \left(\frac{m_i \cdot m_j}{p_{ij}} \right)^{1/2}, \quad (5)$$

gdzie:

m_i i m_j – masy (w analizowanym przykładzie są to sumy elementów odpowiednich wierszy i kolumn macierzy podobieństw).

Na podstawie otrzymanej macierzy niepodobieństw przeprowadzono analizę *unfolding*, w wyniku której otrzymano wspólną konfigurację punktów $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ obrazujących oceny preferencji słuchaczy względem wybranych zajęć w 2009 roku oraz punktów $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_n)$ obrazujących oceny preferencji w 2013 roku (zob. rys. 1). Jeżeli oceny p_{ij} były równe zeru, wtedy niepodobieństwa δ_{ij} potraktowano jako brakujące dane.

Do wyznaczania w kolejnych cyklach iteracyjnych konfiguracji punktów w analizie *unfolding* wykorzystano transformację Guttmana. W wyniku jej zastosowania współrzędne punktów po K -tym cyklu iteracyjnym wynoszą⁴:

$$\begin{aligned} \mathbf{X}^K &= \mathbf{V}^+ [\mathbf{B}_{11} \mathbf{X}^* + \mathbf{B}_{12} \mathbf{Y}^*] \\ \mathbf{Y}^K &= \mathbf{V}^+ [\mathbf{B}_{12}^T \mathbf{X}^* + \mathbf{B}_{22} \mathbf{Y}^*] \end{aligned} \quad (6)$$

gdzie:

$$[\mathbf{V}^+]_{n \times n} = n^{-1} (\mathbf{I} - (2n)^{-1} \mathbf{1}\mathbf{1}^T),$$

$\mathbf{1}$ – kolumnowy wektor jedynek,

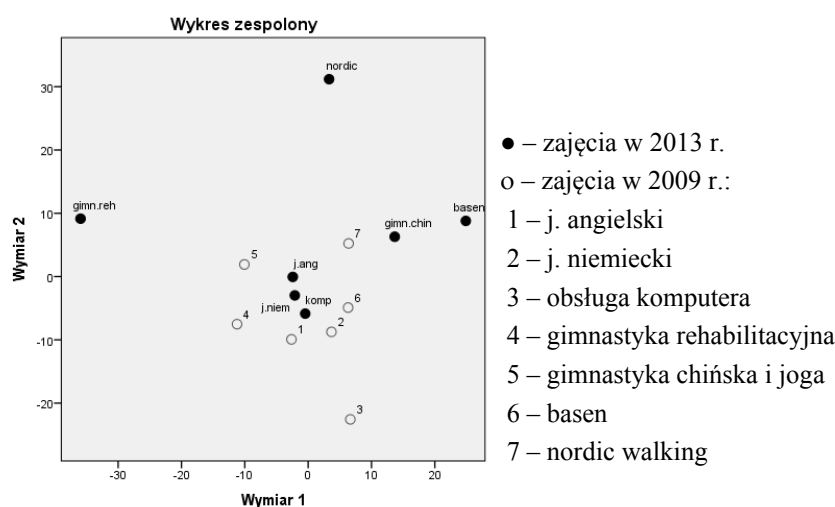
⁴ Ror. A. Zaborski: Analiza unfolding z wykorzystaniem modelu grawitacji. W: Taksonomia 20. Klasyfikacja i analiza danych – teoria i zastosowania. Red. K. Jajuga, M. Walesiak. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu 2013, nr 278, s. 217-218.

$$B_{12} - \text{macierz o elementach } b_{ij} = \begin{cases} -\delta_{ij} & \text{dla } d_{ij}(X^*, Y^*) \neq 0 \\ d_{ij}(X^*, Y^*) & \\ 0 & \text{dla } d_{ij}(X^*, Y^*) = 0 \end{cases},$$

$$B_{11} - \text{diagonalna macierz o elementach } b_{ii} = -\sum_j b_{ij},$$

$$B_{22} - \text{diagonalna macierz o elementach } b_{jj} = -\sum_i b_{ij}.$$

X^*, Y^* – macierze X i Y wyznaczone w $K-1$ cyklu iteracyjnym.



Rys. 1. Konfiguracja punktów reprezentujących zajęcia w 2009 r. i 2013 r.

Rozkład punktów na mapie percepcyjnej wskazuje, że w 2009 roku trudno jest zauważyć podobieństwa w preferencjach względem ujętych w badaniu zajęć. Ponieważ był to początek funkcjonowania Uniwersytetu, prawdopodobnie słuchacze nie byli w pełni świadomi swoich oczekiwań i dokonywane wybory były przypadkowe. W 2013 roku widoczne są grupy zajęć ocenianych pod względem preferencji jako podobne. Do pierwszej grupy należą: nauka obsługi komputera, j. niemiecki i j. angielski, do drugiej: zajęcia na basenie oraz gimnastyka chińska i joga. Zajęciami, które są bardzo popularne, jednak różnią się pod względem podobieństwa preferencji, są marsze z kijkami oraz gimnastyka rehabilitacyjna.

3. Metoda wektorów dryfu

Metoda wektorów dryfu umożliwia równoczesną prezentację symetrycznej i skośnosymetrycznej części danych. Symetryczną część danych obrazuje konfiguracja punktów reprezentujących obiekty otrzymana w wyniku niemetrycznego skalowania wielowymiarowego, zaś skośnosymetryczna część przedstawiona jest za pomocą wektorów dryfu wskazujących kierunek i siłę zmian w preferencjach. Algorytm wyznaczania konfiguracji punktów reprezentujących obiekty (w przeprowadzonym badaniu są to poszczególne rodzaje zajęć) oraz wektorów dryfu jest następujący⁵:

1. Na podstawie symetrycznej macierzy **M** przeprowadzić niemetryczne skalowanie wielowymiarowe.

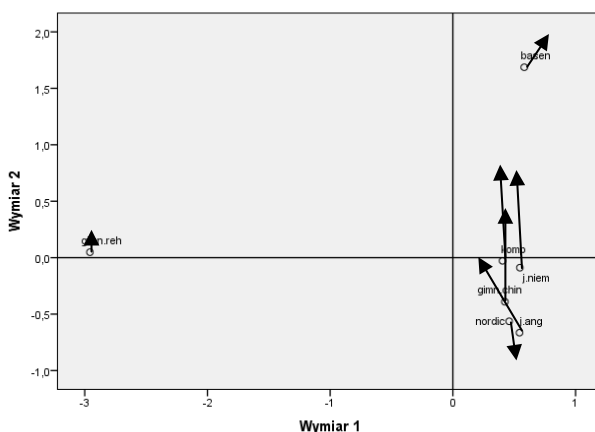
2. Dla wszystkich punktów \mathbf{z}_i i \mathbf{z}_j (przy czym $i \neq j$), otrzymanych w wyniku skalowania wielowymiarowego, wyznaczyć wektory $\mathbf{a}_{ij} = \mathbf{z}_j - \mathbf{z}_i$.

3. Unormować wektory tak, aby ich długość była równa 1, tzn.: $\mathbf{b}_{ij} = \mathbf{a}_{ij} / (\mathbf{a}_{ij}^T \mathbf{a}_{ij})^{1/2}$.

4. Wyznaczyć wektory $\mathbf{c}_{ij} = n_{ij} \cdot \mathbf{b}_{ij}$ (n_{ij} – elementy skośnosymetrycznej macierzy **N**).

5. Wyznaczyć wektor dryfu \mathbf{d}_{ij} jako: $\mathbf{d}_{ij} = n^{-1} \sum_j \mathbf{c}_{ij}$.

Konfigurację punktów reprezentujących analizowane formy zajęć oraz wektorów wskazujących kierunki i siłę zmian w preferencjach przedstawia rys. 2.



Rys. 2. Wspólna konfiguracja punktów reprezentujących zajęcia oraz wektorów dryfu

⁵ I. Borg, P. Groenen: Modern multidimensional scaling. Theory and applications. Second Edition. Springer-Verlag, New York 2005, s. 503.

Na podstawie kierunków wektorów dryfu można wyciągnąć wniosek, że preferencje słuchaczy Uniwersytetu kierują się w stronę zajęć ruchowych. W miejsce zajęć rozwijających umiejętności językowe słuchaczy oraz obsługę komputera coraz bardziej preferowane są zajęcia poprawiające sprawność fizyczną, a w szczególności są to zajęcia na basenie.

Podsumowanie

W pracy zaprezentowano zastosowanie dwóch metod niesymetrycznego skalowania wielowymiarowego do analizy zmian zachodzących w preferencjach słuchaczy Uniwersytetu Trzeciego Wieku względem wybranych zajęć w okresie od 2009 do 2013 roku. Grawitacyjną analizę *unfolding* wykorzystano do porównania podobieństwa zajęć ze względu na preferencje słuchaczy w badanym okresie, zaś metodę wektorów dryfu do wskazania kierunku i siły zmian zachodzących w preferencjach. Analiza danych wykazała, że w badanym okresie preferencje słuchaczy Uniwersytetu kierują się w stronę zajęć ruchowych. W miejscu nauki języków obcych i nauki obsługi komputera coraz częściej wybierają oni zajęcia z gimnastyki, a przede wszystkim zajęcia na basenie.

Literatura

- Borg I., Groenen P.: Modern multidimensional scaling. Theory and applications. Second Edition. Springer-Verlag, New York 2005.
- Chino N.: A graphical technique for representing the asymmetric relationship between N objects. „Behaviormetrika” 1978, No. 5.
- DeSarbo W.S., Kim J., Choi S.C., Spaluding M.: A Gravity-Based Multidimensional Scaling Model for Deriving Spatial Structures Underlying Consumer Preference/Choice Judgments. „Journal of Consumer Research” 2002, Vol. 29.
- Harshman R.A., Green P.E., Wind Y., Lundy M.E.: A model for the analysis of asymmetric data in marketing research. „Marketing Science” 1982, Vol. 1, No. 2.
- Holyoak K.J., Gordon P.C.: Social reference points. „Journal of Personality and Social Psychology” 1983, No. 44.
- Shepard R.N.: Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function. Part I and II. „Psychometrika” 1962, Vol. 27, No. 1, 2.
- Tversky A., Gati I.: Features of similarity. „Psychological Review” 1982, No. 89.
- Zaborski A.: Analiza unfolding z wykorzystaniem modelu grawitacji. W: Taksonomia 20. Klasyfikacja i analiza danych – teoria i zastosowania. Red. K. Jajuga, M. Waleśiak. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu 2013, nr 278.

**THE PREFERENCE ANALYSIS OF UNIVERSITY
OF THIRD AGE MEMBERS USING SELECTED ASYMMETRIC
MULTIDIMENSIONAL SCALING METHODS**

Summary

This paper has a methodological and empirical character. Its aims at analysis the changes that have occurred in the preferences of the University of the Third Age members from 2009 to 2013. The study used two methods of asymmetric multidimensional scaling: the gravity unfolding analysis and the drift vectors method. Gravity unfolding analysis was used for the graphical presentation of the respondents' preference similarities in a definite period, and the drift vectors method was used to indicate the direction and the strength of this changes.