

## Zastosowanie liniowego porządkowania obiektów do oceny aktywności ekonomicznej ludności w ujęciu województw

Agnieszka Sompolska-Rzechuła<sup>a</sup>

**Streszczenie.** Celem artykułu jest porównanie wyników zastosowania różnych metod porządkowania liniowego obiektów na przykładzie poziomu aktywności ekonomicznej ludności oraz wybór metody do ostatecznej oceny badanego zjawiska złożonego. Zaproponowano dwa podejścia. Pierwsze polega na wyborze spośród wyników uzyskanych za pomocą wszystkich rozpatrywanych metod wzorcowych i bezwzorcowych, drugie zaś zakłada wybór metody porządkowania osobno dla każdej z grup. Zagadnienie zilustrowano przykładem dotyczącym poziomu aktywności ekonomicznej ludności na podstawie danych uzyskanych z Badania Aktywności Ekonomicznej Ludności (BAEL) na koniec I kwartału 2019 r. w ujęciu województw. Zastosowano kilka wariantów metod bezwzorcowych, różniących się formułą normalizacji cech diagnostycznych, jak również metody wzorcowe: Hellwiga, TOPSIS oraz pozycyjną, bazującą na przestrzennej medianie Webera. W grupie bezwzorcowych metod porządkowania liniowego metodą dającą wyniki najbardziej zbliżone do wyników uzyskanych za pomocą wszystkich pozostałych metod okazała się metoda oparta na unitaryzacji zerowanej. W grupie metod wzorcowych wyniki takie uzyskano za pomocą metody Hellwiga.

**Słowa kluczowe:** porządkowanie liniowe, metody wzorcowe i bezwzorcowe, metody normalizacji cech, poziom aktywności ekonomicznej ludności, BAEL

**JEL:** C19, C38, J01

## Linear ordering of objects as applied to assessing economic activity of populations in voivodships

**Abstract.** The aim of the paper is to compare the results of adopting different methods of linear ordering of objects applied to evaluating the level of economic activity of the population, as well as to select the method for the final assessment of the studied complex phenomenon. Two approaches have been presented. The first involves choosing from the results obtained by adopting all the analysed pattern and patternless methods, while the other proposes the choice of a method separately for each group. The studied problem has been demonstrated on the example of the level of economic activity of the population, which was defined on the basis of the data for the end of the first quarter of 2019, presented for voivodships and drawn from the Labour Force Survey in Poland (LFS). The analysis involved using several variants of patternless methods that differed from one another according to which formula of the diagnostic feature normalisation they used, as well as the following standard methods: Hellwig's method, TOPSIS and the positional method based on Weber's spatial median. In the group of the patternless methods of linear ordering, the one which yielded results closest to the results obtained using all the other variants was the method based on zeroed unitarisation. In the group of the pattern methods, similar conditions were met by Hellwig's method.

**Keywords:** linear ordering, patternless and pattern method, methods of standardisation of variables, level of economic activity of the population, LFS

<sup>a</sup> Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, Wydział Ekonomiczny. ORCID: <https://orcid.org/0000-0002-7571-1152>.

## 1. Wprowadzenie

Metody porządkowania zbioru obiektów można podzielić na metody porządkowania liniowego i nieliniowego. Te pierwsze są wykorzystywane do oceny obiektów wielocechowych, ponieważ umożliwiają ich uszeregowanie według określonego kryterium ogólnego, od „najlepszego” do „najgorszego”. W badaniach społeczno-ekonomicznych bardzo często analizuje się zjawiska złożone, niepodlegające bezpośredniemu pomiarowi, wykorzystując zbiory cech diagnostycznych mierzonych na różnych skalach pomiarowych.

Liniowe porządkowanie obiektów otrzymuje się na podstawie cechy zwanej agregatową lub syntetyczną (albo metacehą), którą tworzy się poprzez agregację cech wyjściowych opisujących badane obiekty (Malina, 2004). Cecha syntetyczna ma charakter zmiennej ukrytej, ponieważ jej wartości nie są obserwowane bezpośrednio. Rozróżnia się dwie grupy metod wykorzystywanych do wyznaczenia wartości cechy syntetycznej: bezwzorcowe i wzorcowe.

Do polskich badań cechę syntetyczną wprowadził Hellwig (1968). Od tego czasu inspiruje ona badaczy i znajduje liczne zastosowania w ocenie zjawisk społeczno-gospodarczych<sup>1</sup>. Podczas badań tego rodzaju często pojawia się pytanie, które spośród metod porządkowania obiektów wybrać jako ostateczne do oceny analizowanego zjawiska.

Celem artykułu jest porównanie wyników zastosowania różnych metod porządkowania liniowego obiektów na przykładzie poziomu aktywności ekonomicznej ludności oraz wybór metody do ostatecznej oceny badanego zjawiska złożonego. Opisane w pracy badanie przeprowadzono, opierając się na procedurze przedstawionej w pracy Kukuły i Luty (2015). Do wyboru metody zaproponowano dwa podejścia<sup>2</sup>:

- wybór spośród wielu metod, zarówno wzorcowych, jak i bezwzorcowych;
- wybór metody porządkowania oddzielnie dla obu grup, tj. metod wzorcowych i bezwzorcowych.

Zaproponowanie drugiego podejścia jest uzasadnione tym, że metody bezwzorcowe i wzorcowe różnią się algorytmem liczenia oraz – jak pokazują badania – ostatecznymi wynikami porządkowań.

## 2. Metody liniowego porządkowania obiektów

Podstawowymi etapami w procedurze porządkowania liniowego są: selekcja cech, określenie ich charakteru (stymulanty, destymulanty, nominanty), wyznaczenie wag,

---

<sup>1</sup> Przykłady prac z zastosowaniem cechy syntetycznej: Bąk (2013, 2017), Kisielińska (2016), Kukuła (2000), Kukuła i Luty (2015, 2017), Lira, Wagner i Wysocki (2002), Malina i Zeliaś (1997), Młodak (2006), Nowak (1990), Sokołowski i Markowska (2017), Sompolska-Rzechuła (2018), Tarczyński i Łuniewska (2006), Waleśiak (1993), Wysocki (2010).

<sup>2</sup> W pracy Kukuły i Luty (2015) zaprezentowano podejście pierwsze, tj. wybór procedury spośród wielu metod, zarówno wzorcowych, jak i bezwzorcowych.

normalizacja cech diagnostycznych, wyznaczenie współrzędnych wzorca w przypadku agregacji wzorcowej i agregacja bezwzorcowa lub wzorcowa (Bąk, 2013).

Zmienne diagnostyczne powinny cechować się (Kukuła, 2000, s. 47 i 48):

- dużym znaczeniem w charakterystyce analizowanego zjawiska;
- dostępnością informacji o zmiennych;
- słabą korelacją między zmiennymi (a najlepiej jej brakiem);
- wysokim stopniem zmienności oraz względnej wartości informacyjnej.

W tworzeniu cechy syntetycznej decydujące jest kryterium merytoryczne. Jak twierdzi Kukuła (2000) oraz Kukuła i Luty (2018), skorelowanie między cechami nie przekreśla słuszności ich wyboru do badania, ponieważ liniowe porządkowanie obiektów opiera się na wartości zmiennej syntetycznej, którą otrzymuje się poprzez sumowanie ujednoczonych zmiennych diagnostycznych, uznanych z merytorycznego punktu widzenia za ważne<sup>3</sup>. Warunkiem uznania różnych cech za diagnostyczne jest ich zdolność do dyskryminacji obiektów. Aby ją określić, bada się, czy potencjalne cechy odznaczają się odpowiednią zmiennością. Można w tym celu wykorzystać standardowy współczynnik zmienności:

$$v_k = \frac{s_k}{\bar{x}_k} \cdot 100\% \quad (1)$$

gdzie  $\bar{x}_k$  i  $s_k$  oznaczają odpowiednio średnią arytmetyczną i odchylenie standardowe  $k$ -tej cechy.

Ze zbioru potencjalnych cech diagnostycznych eliminuje się te, dla których moduł współczynnika zmienności kształtuje się poniżej wybranej wartości progowej. Wartość tę często przyjmuje się na poziomie 10% (Młodak, 2006). Ponadto do oceny zdolności dyskryminacyjnej cech można wykorzystać miernik zróżnicowania zwany współczynnikiem względnej amplitudy wahań zmiennej (Kukuła, 2000). Stanowi on iloraz dwóch wartości zmiennej: maksymalnej i minimalnej oraz informuje o tym, ile razy zmienna obiektu najlepszego przewyższa wartość zmiennej obiektu najgorszego<sup>4</sup>.

Ustalony zbiór cech diagnostycznych tworzy podstawę dalszej analizy, w której należy określić charakter cech, czyli wyróżnić stymulanty, destymulanty i nominanty. Pojęcia stymulanty i destymulanty zostały wprowadzone przez Hellwiga (1968). Stymulanta oznacza cechę, której wyższa wartość świadczy o lepszej kondycji obiektu w danym kontekście. Tym samym maksymalna wartość stymulanty uznawana jest za najkorzystniejszą, a minimalna – za najmniej korzystną dla badanych obiektów.

<sup>3</sup> Uzasadnienie wskazanego podejścia można znaleźć w pracy Kukuły (2000, s. 47–52).

<sup>4</sup> Inne metody doboru zmiennych w liniowym porządkowaniu obiektów oparte na analizie współczynników korelacji zostały przedstawione w pracy Bąka (2017).

Destymulanta zaś to cecha, której niższe wartości oznaczają lepszą sytuację obiektu pod danym względem. Wobec tego wartość maksymalna destymulanty jest uznawana za najmniej korzystną, a najmniejsza – za najkorzystniejszą dla badanych obiektów (Walesiak, 1990). Nominanta z kolei charakteryzuje się tym, że poniżej wartości najkorzystniejszej z punktu widzenia oceny obiektów, tzn. optymalnego poziomu nasycenia lub wartości nominalnej (Młodak, 2006, s. 33), ma charakter stymulanty (a zatem większe wartości są korzystniejsze), a powyżej – destymulanty (co oznacza, że po przekroczeniu wartości optymalnej dalszy wzrost wartości cechy staje się niekorzystny), lub na odwrót. Nominanty są często pomijane w badaniach empirycznych ze względu na trudności związane z ustaleniem wartości nominalnych (Walesiak, 1990)<sup>5</sup>. Jeśli określenie charakteru cech stwarza trudności, stosuje się kryteria merytoryczne lub analizę korelacyjną.

Po rozpoznaniu charakteru cech należy je przekształcić. Najczęściej destymulanty zamienia się w stymulanty za pomocą przekształcenia (Walesiak, 1990):

- różnicowego:  $x'_{ij} = a - b \cdot x_{ij}$

lub

- ilorazowego:  $x'_{ij} = \frac{b}{x_{ij}}$

gdzie:

$a$  i  $b$  – stałe przyjmowane arbitralnie (najczęściej  $a \geq \max_{\substack{i \in \{1, \dots, n\} \\ j \in \{1, \dots, k\}}} x_{ij}$ ,  $b = 1$ ),

$i = 1, \dots, n$ ,

$j = 1, \dots, k$ ,

$n$  – liczba obiektów,

$k$  – liczba cech.

Kolejnym etapem budowy cechy syntetycznej jest normalizacja cech w celu wprowadzenia addytywności w zbiorach wartości cech o różnych mianach. Uzyskuje się to poprzez transformację ich wartości w taki sposób, aby uzyskać wartości pozbawione mian oraz ujednoczone co do zakresów liczbowych (Wysocki, 2010, s. 48).

W literaturze przedmiotu można znaleźć ogólną formułę przekształcenia normalizującego dla cech (Pawełek, 2008):

$$z_{ij} = \frac{x_{ij} - a}{b} \quad (i = 1, 2, \dots, n; j = 1, 2, \dots, k) \quad (2)$$

gdzie:

$z_{ij}$  – wartość znormalizowanej cechy  $X_j$ ,

$x_{ij}$  – wartość  $j$ -tej cechy dla  $i$ -tego obiektu,

$a, b$  – parametry normalizujące.

<sup>5</sup> Pojęcie nominanty zostało wprowadzone przez Borysa (1978).

Parametr  $b$  może występować jako odchylenie standardowe zmiennej, rozstęp zmiennej, średnia arytmetyczna, a także maksymalna lub minimalna wartość cechy. W niniejszej pracy wykorzystano następujące metody normalizacji cech (Kukuła i Luty, 2015):

- standaryzację:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (s_j \neq 0; \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, k) \quad (3)$$

gdzie:

$z_{ij}$  – wartość unormowana  $j$ -tej cechy dla  $i$ -tego obiektu,

$\bar{x}_j$  i  $s_j$  – średnia arytmetyczna i odchylenie standardowe  $j$ -tej cechy;

- standaryzację pozycyjną z medianą Webera<sup>6</sup>:

$$z_{ij} = \frac{x_{ij} - \theta_{0j}}{1,4826 \cdot \mathit{m\ddot{a}d}(X_j)} \quad (\mathit{m\ddot{a}d}(X_j) \neq 0) \quad (4)$$

gdzie:

$z_{ij}$  – wartość unormowana  $j$ -tej cechy dla  $i$ -tego obiektu,

$\theta_{0j}$  – mediana Webera,

$\mathit{m\ddot{a}d}(X_j)$  – medianowe odchylenie bezwzględne, w którym odległości cech badane są w odniesieniu do mediany Webera:

$$\mathit{m\ddot{a}d}(X_j) = \mathit{med}_{i=1, \dots, n} |x_{ij} - \theta_{0j}| \quad (j = 1, \dots, k);$$

- unitaryzację zerowaną:

$$z_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \quad (\max_i x_{ij} \neq \min_i x_{ij}) \quad (5)$$

gdzie:

$z_{ij}$  – wartość unormowana  $j$ -tej cechy dla  $i$ -tego obiektu,

$z_{ij} \in [0, 1]$ ;

<sup>6</sup> Mediana Webera stanowi wielowymiarowe uogólnienie klasycznego pojęcia mediany. Chodzi tu o wektor, który minimalizuje sumę odległości euklidesowych od danych punktów reprezentujących rozpatrywane obiekty, a więc znajduje się niejako pośrodku nich, ale jest jednocześnie uodporniony na występowanie obserwacji odstających (Młodak, 2006). Problem mediany Webera został przedstawiony także w pracach Młodaka (2009) oraz Sompolskiej-Rzechuły i Machowskiej-Szewczyk (2018).

- przekształcenie ilorazowe z punktem odniesienia przyjętym jako wartość minimalna cechy.

Stworzenie liniowego porządkowania obiektów wymaga następnie określenia wartości cechy syntetycznej. W przypadku metod bezwzorcowych wyznacza się w tym celu uśrednioną wartość unormowanych cech:

$$\mu_i = \frac{1}{k} \sum_{j=1}^k z_{ij} \quad (6)$$

gdzie  $\mu_i$  – wartość cechy syntetycznej dla  $i$ -tego obiektu.

Idea wzorcowych metod agregacji cech polega na wyznaczeniu odległości poszczególnych obiektów od obiektu modelowego. Wśród metod wzorcowych na uwagę zasługuje metoda zaproponowana przez Hellwiga (1968). Opiera się ona na znormalizowanych wartościach cech diagnostycznych  $X_1, X_2, \dots, X_k$ , które są traktowane jako jednakowo ważne. Odległości euklidesowe każdego obiektu od wzorca oblicza się według wzoru:

$$d_i = \sqrt{(z_{ik} - z_{0k})^2} \quad (i = 1, \dots, n) \quad (7)$$

gdzie:

$z_{0k}$  – znormalizowana wartość  $k$ -tej cechy dla obiektu wzorcowego,

$n$  – liczba obiektów.

Na podstawie wartości  $d_i$  konstruowany jest względny taksonomiczny miernik rozwoju, definiowany jako (Nowak, 1990):

$$\mu_i = 1 - \frac{d_i}{d_0} \quad (i = 1, \dots, n) \quad (8)$$

gdzie  $d_0 = \bar{d} + 2 \cdot s_d$  oznacza teoretycznie najmniej korzystną wartość odległości obiektu od wzorca (wartości większe od  $d_0$  są przyjmowane bardzo rzadko), przy czym  $\bar{d} = \frac{1}{n} \cdot \sum_{i=1}^n d_i$  oraz  $s_d = \sqrt{\frac{1}{n} (d_i - \bar{d})^2}$  to odpowiednio średnia arytmetyczna i odchylenie standardowe odległości obiektów od wzorca.

Syntetyczny miernik rozwoju Hellwiga przyjmuje zazwyczaj wartości z przedziału  $[0, 1]$ . Im mniejsza różnica wartości miernika  $\mu_i$  od 1, tym mniejsza różnica w poziomie rozwoju między danym obiektem a obiektem wzorcowym. Syntetyczny miernik rozwoju jest wypadkową wszystkich cech charakteryzujących badane obiek-

ty, zatem pozwala określić osiągnięty w określonym czasie przeciętny poziom wartości cech (Balicki, 2009).

Metodą porządkowania liniowego wykorzystującą zarówno wzorzec, jak i antywzorzec jest metoda TOPSIS (Technique for Order Preference by Similarity to an Ideal Solution; Hwang i Yoon, 1981). Polega ona na obliczeniu odległości euklidesowych każdego ocenianego obiektu od wzorca i od antywzorca rozwoju, co odróżnia ją od metody Hellwiga, która uwzględnia tylko odległości od wzorca (Wysocki, 2010). Ustala się współrzędne jednostek modelowych – wzorca i antywzorca rozwoju. Wartości wzorca ( $A^+$ ) i antywzorca rozwoju ( $A^-$ ) definiuje się jako (Wysocki, 2010):

$$A^+ = \left( \max_i(z_{i1}), \max_i(z_{i2}), \dots, \max_i(z_{ik}) \right) = (z_1^+, z_2^+, \dots, z_k^+) \quad (9)$$

$$A^- = \left( \min_i(z_{i1}), \min_i(z_{i2}), \dots, \min_i(z_{ik}) \right) = (z_1^-, z_2^-, \dots, z_k^-) \quad (10)$$

Jeśli jako formułę normującą przyjmie się unitaryzację zerowaną, jest to:

$$z^+ = \underbrace{(1, 1, \dots, 1)}_K \quad z^- = \underbrace{(0, 0, \dots, 0)}_K \quad (11)$$

Obliczenie odległości euklidesowych każdego obiektu od wzorca i antywzorca następuje według wzorów:

$$d_i^+ = \sqrt{\sum_{k=1}^K (z_{ik} - z_k^+)^2}, \quad d_i^- = \sqrt{\sum_{k=1}^K (z_{ik} - z_k^-)^2}, \quad i = 1, \dots, n \quad (12)$$

Natomiast wartość cechy syntetycznej wyznacza się następująco (Hwang i Yoon, 1981):

$$\mu_i = \frac{d_i^-}{d_i^+ + d_i^-} \quad (13)$$

przy czym  $0 \leq \mu_i \leq 1, i = 1, 2, \dots, n$ .

Im mniejsza jest odległość danego obiektu od wzorca rozwoju (a tym samym większa od antywzorca rozwoju), tym wartość cechy syntetycznej jest bliższa 1.

W budowie miary syntetycznej – szczególnie gdy cechy mają rozkłady asymetryczne – wykorzystuje się podejście pozycyjne, np. z medianą Webera. Mediana jako podstawowy parametr pozycyjny charakteryzuje się wysoką odpornością na występowanie obserwacji odstających i może być stosowana na etapie normalizacji cech oraz w przekształceniu normalizacyjnym. Wartość cechy syntetycznej oblicza się tak jak w przypadku metody Hellwiga, czyli:

$$\mu_i = 1 - \frac{d_i}{d_0^*} \quad (i = 1, \dots, n) \quad (14)$$

gdzie:

$$d_i = \text{med}_{j=1, \dots, k} |z_{ij} - \varphi_j|,$$

$\varphi_j = \max_{i=1, \dots, n} z_{ij}$  – wektor wzorca rozwoju, definiowany jako wektor, którego współrzędnymi są maksymalne wartości cech znormalizowanych,

$z_{ij}$  – wartość unormowana  $j$ -tej cechy dla  $i$ -tego obiektu,

$d_0^* = \text{med}(\mathbf{d}) + 2,5\text{mad}(\mathbf{d})$ , gdzie  $\mathbf{d} = (d_1, d_2, \dots, d_n)$  – wektor odległości od wzorca.

Wielkość  $d_0^*$  stanowi pozycyjny odpowiednik analogicznej wartości progowej  $d_0$  rozpatrywanej w podejściu opartym na typowych miarach odległości, takim jak metoda Hellwiga czy TOPSIS. Wartość 2,5, określana jako odpornościowa wartość progowa, wyznacza barierę korzystnych wartości pomiarów odległości obiektów od wzorca rozwoju (Młodak, 2006). Parametr ten, podobnie jak wartość 1,4826, został ustalony w wyniku badań empirycznych. Im większa wartość cechy syntetycznej, tym wyższy jest poziom rozwoju obiektu (Wysocki, 2010).

Jeżeli w badaniu rozpatruje się wyniki porządkowania liniowego obiektów otrzymane za pomocą kilku metod, to spośród uzyskanych porządkowań należy wybrać takie, które jest najbliższe wszystkim pozostałym rankingom. Przy wyborze metody porządkowania liniowego w badaniach należy wziąć pod uwagę postulaty, aby (Kukuła i Luty, 2017):

- dostosować wybór metody do wykorzystania jej własności, charakteru cech diagnostycznych oraz skali pomiaru zmiennych;
- wybrać metodę, która daje najbliższe wyniki końcowe względem pozostałych metod.

Do wyboru wyników najbliższych względem pozostałych można zaproponować następujące postępowanie (Kukuła i Luty, 2015):

- dokonanie porównania porządkowań uzyskanych za pomocą  $v$  metod liniowego porządkowania; liczba tych porównań wyniesie:  $\alpha = \frac{v \cdot (v-1)}{2}$ ;
- wyznaczenie miary podobieństwa porządkowań  $m_{pg}$  w postaci:



$$m_{pq} = 1 - \frac{2 \cdot \sum_{i=1}^n |c_{ip} - c_{iq}|}{n^2 - z} \quad p, q = 1, \dots, v \quad (15)$$

gdzie:

$c_{ip}$  – pozycja  $i$ -tego obiektu w porządkowaniu o numerze  $p$ ,

$c_{iq}$  – pozycja  $i$ -tego obiektu w porządkowaniu o numerze  $q$ ,

$z = \begin{cases} 0, n \in P \\ 1, n \notin P \end{cases}$ , a  $P$  jest zbiorem parzystych liczb naturalnych.

Miara  $m_{pq}$  przyjmuje wartości z przedziału  $[0, 1]$ . Wyniki porównań przedstawia się w formie macierzy  $\mathbf{M}$ :

$$\mathbf{M} = \begin{bmatrix} 1 & m_{12} & m_{13} & \dots & m_{1v} \\ m_{21} & 1 & m_{23} & \dots & m_{2v} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{v1} & m_{v2} & m_{v3} & \dots & 1 \end{bmatrix}$$

Macierz  $\mathbf{M}$  jest macierzą symetryczną o wymiarach  $v \times v$ ; na głównej przekątnej umieszczone są wyniki porównania rankingów o tych samych numerach, czyli  $m_{pq} = 1$ , gdy  $p = q$ . Ponadto  $m_{pq} = m_{qp}$ , gdy  $p \neq q$ .

Aby określić stopień podobieństwa rankingu uzyskanego w wyniku zastosowania  $p$ -tej metody porządkowania liniowego do pozostałych rankingów, należy obliczyć sumę elementów  $p$  wiersza (lub kolumny) macierzy  $\mathbf{M}$  pomniejszoną o 1. Sumę tę oznacza się symbolem  $u_p$ . Uzyskany wynik można uśrednić następująco:

$$\bar{u}_p = \frac{1}{v-1} \sum_{\substack{q=1 \\ p \neq q}}^v m_{pq} \quad p, q = 1, \dots, v \quad (16)$$

Należy wybrać tę metodę porządkowania liniowego, dla której:

$$\bar{u}_p = \max_p \bar{u}_p \quad (17)$$

### 3. Porównanie wyników liniowego porządkowania województw na przykładzie danych BAEL

Procedurę wyboru takiego liniowego porządkowania obiektów, które byłoby najbliższe wszystkim pozostałym rankingom, zaprezentowano na przykładzie danych dotyczących poziomu aktywności ekonomicznej ludności pochodzących z Badań Aktywności Ekonomicznej Ludności (BAEL) w I kwartale 2019 r. w ujęciu

województw (GUS, 2019). Badanie Aktywności Ekonomicznej Ludności prowadzone jest w Polsce kwartalnie od maja 1992 r., zgodnie z zaleceniami Eurostatu. Podstawę prawną badania stanowi coroczne rozporządzenie Rady Ministrów w sprawie programu badań statystycznych statystyki publicznej (GUS, 2019). Badanie jest realizowane metodą reprezentacyjną i ma na celu uzyskanie informacji o wielkości i strukturze zasobów pracy. W efekcie zostaje ustalona liczba osób zarówno aktywnych, jak i biernych zawodowo. W przekroju województw w publikacji GUS (2019, s. 100 i n.) można znaleźć informacje odnoszące się do aktywności ekonomicznej ludności z uwzględnieniem takich wskaźników, jak m.in.: odsetek aktywnych zawodowo (w tym odsetek pracujących i odsetek bezrobotnych) oraz odsetek biernych zawodowo, współczynnik aktywności zawodowej, wskaźnik zatrudnienia i stopa bezrobocia. Dla ludności w wieku 15 lat i więcej (GUS, 2019, s. 107 i n.) w ujęciu makroregionów oraz województw podane są informacje na temat liczby aktywnych zawodowo (w tym liczby pracujących i liczby bezrobotnych) oraz liczby biernych zawodowo. Ponadto zamieszczono informacje odnoszące się do współczynnika aktywności zawodowej, wskaźnika zatrudnienia i stopy bezrobocia<sup>7</sup>.

Poziom aktywności ekonomicznej ludności opisywany jest przez wiele cech dotyczących wykonywania pracy, pozostawania bezrobotnym lub biernym zawodowo (GUS, 2019). W badaniu omawianym w niniejszym artykule przeprowadzono liniowe porządkowania województw Polski ze względu na poziom aktywności ekonomicznej ludności w I kwartale 2019 r. na podstawie następujących cech:

- $X_1$  – liczba pracujących na 1000 ludności,
- $X_2$  – liczba bezrobotnych na 1000 pracujących,
- $X_3$  – liczba biernych zawodowo na 1000 pracujących,
- $X_4$  – współczynnik aktywności zawodowej w %,
- $X_5$  – wskaźnik zatrudnienia w %,
- $X_6$  – stopa bezrobocia ogółem w %,
- $X_7$  – stopa bezrobocia wśród kobiet w %,
- $X_8$  – stopa bezrobocia w miastach w %,
- $X_9$  – stopa bezrobocia na wsi w %.

Wybór cech do badania związany był z dostępnością danych w przekroju województw zawartych w publikacji GUS (2019). Cechy o numerach 1, 4, 5 uznano za stymulanty, a pozostałe – za destymulanty. Wyższa wartość cech uznanych za stymulanty świadczy o wyższym poziomie aktywności ekonomicznej ludności, natomiast wyższa wartość cech uznanych za destymulanty wskazuje na jej niższy poziom.

---

<sup>7</sup> Koncepcja BAEL oraz definicje odnoszące się do pojęć związanych z ludnością aktywną zawodowo znajdują się w publikacji GUS (2016).

W tabl. 1 zaprezentowano wartości wybranych parametrów opisowych aktywności ekonomicznej ludności.

**Tabl. 1.** Wybrane parametry opisowe cech informujących o aktywności ekonomicznej ludności

Parametry	Cechy								
	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$
Średnia .....	553,29	39,20	810,11	55,33	53,15	3,92	4,52	3,76	4,25
Wartość maksymalna .....	602,49	67,69	917,69	60,20	57,70	6,80	7,80	5,70	7,70
Wartość minimalna .....	521,46	18,35	659,79	52,10	49,60	1,80	2,10	2,00	2,30
Mediana .....	545,06	36,10	834,67	54,50	52,75	3,60	4,25	3,75	4,15
Odchylenie standardowe .....	22,18	12,64	70,82	2,20	2,36	1,28	1,44	0,96	1,61
Współczynnik zmienności .....	4,01	32,25	8,74	3,98	4,43	32,68	31,86	25,44	37,79
Iloraz skrajnych wartości .....	1,16	3,69	1,40	1,16	1,16	3,78	3,71	2,85	3,35
Współczynnik asymetrii .....	0,80	0,61	-0,63	0,79	0,58	0,60	0,55	-0,08	0,50

Źródło: obliczenia własne na podstawie GUS (2019).

Przy wyborze cech kierowano się dwoma kryteriami (Kukuła, 2014): przydatnością merytoryczną w ocenie badanego zjawiska oraz stopniem zmienności cech kwalifikowanych do zbioru zmiennych diagnostycznych.

Na podstawie informacji zawartych w tabl. 1 można stwierdzić, że cechy o numerach 1, 3, 4 i 5 charakteryzują się słabą zdolnością dyskryminacyjną obiektów, o czym świadczy niska wartość współczynników zmienności i ilorazów skrajnych wartości. Dla wymienionych cech współczynnik zmienności przyjmuje wartości niższe od 10%, a iloraz skrajnych wartości jest niższy od 2. Z tego powodu do dalszej analizy przyjęto cechy:  $X_2$ ,  $X_6$ ,  $X_7$ ,  $X_8$  i  $X_9$ . Wszystkie miały charakter destymulant, w związku z czym zostały przekształcone w stymulanty według przekształcenia ilorazowego.

W tabl. 2 przedstawiono wyniki porządkowań województw Polski uzyskanych na podstawie przyjętych cech diagnostycznych za pomocą wybranych metod bezwzorcowych na podstawie: standaryzacji (1), unitaryzacji zerowanej (2) i przekształcenia ilorazowego (3) oraz wzorcowych: Hellwiga (4), TOPSIS (5) i pozycyjnej (6).

**Tabl. 2.** Ranking województw ze względu na poziom aktywności ekonomicznej ludności w I kwartale 2019 r. na podstawie wybranych metod porządkowania liniowego

Województwa	Metody bezwzorcowe z formułą normującą			Metody wzorcowe		
	1	2	3	4	5	6
Dolnośląskie .....	3	3	3	3	3	3
Kujawsko-pomorskie .....	9	9	9	9	9	11
Lubelskie .....	16	16	16	16	16	16
Lubuskie .....	1	1	1	1	1	1
Łódzkie .....	12	12	12	12	12	15
Małopolskie .....	10	10	10	10	10	10

**Tabl. 2.** Ranking województw ze względu na poziom aktywności ekonomicznej ludności w I kwartale 2019 r. na podstawie wybranych metod porządkowania liniowego (dok.)

Województwa	Metody bezwzorcowe z formułą normującą			Metody wzorcowe		
	1	2	3	4	5	6
Mazowieckie .....	11	11	11	11	11	9
Opolskie .....	5	5	5	4	6	4
Podkarpackie .....	15	15	15	15	15	13
Podlaskie .....	6	4	4	5	4	2
Pomorskie .....	2	2	2	2	2	5
Śląskie .....	8	8	8	7	8	6
Świętokrzyskie .....	13	13	13	13	13	12
Warmińsko-mazurskie .....	14	14	14	14	14	14
Wielkopolskie .....	7	7	6	8	7	8
Zachodniopomorskie .....	4	6	7	6	5	7

Źródło: jak przy tabl. 1.

W tabl. 3 przedstawiono ocenę zgodności uporządkowań wybranymi metodami mierzoną przy użyciu współczynnika korelacji rang Kendalla (Gatnar i Walesiak, 2004).

**Tabl. 3.** Wartości współczynnika korelacji rang Kendalla według poszczególnych metod

Metody	1	2	3	4	5	6
1 .....	1,000	0,950	0,933	0,950	0,967	0,733
2 .....	0,950	1,000	0,983	0,967	0,983	0,783
3 .....	0,933	0,983	1,000	0,950	0,967	0,767
4 .....	0,950	0,967	0,950	1,000	0,950	0,783
5 .....	0,967	0,983	0,967	0,950	1,000	0,767
6 .....	0,733	0,783	0,767	0,783	0,767	1,000

Źródło: jak przy tabl. 1.

Na podstawie analizy wyników zawartych w tabl. 2 można stwierdzić, że w przypadku niektórych województw – np. opolskiego, podlaskiego, wielkopolskiego czy zachodniopomorskiego – porządkowania uzyskane za pomocą poszczególnych metod są różne, ale w przypadku kilku województw uzyskano takie same pozycje. Dotyczy to: dolnośląskiego, lubuskiego (pierwsza pozycja), małopolskiego, warmińsko-mazurskiego i lubelskiego (ostatnia pozycja).

Ocena zgodności uporządkowań za pomocą wybranych metod mierzona współczynnikiem korelacji rang Kendalla wskazuje na występowanie istotnych powiązań między pozycjami województw.

W kolejnym kroku analizy wyznaczono porządkowania najbardziej zbliżone do pozostałych porządkowań uzyskanych zarówno wybranymi metodami wzorcowymi, jak i bezwzorcowymi.

Dla każdej pary układów porządkowych przedstawionych w tabl. 2 wyznaczono wartość miary  $m_{pq}$ , zapisane w postaci macierzy  $\mathbf{M}$ , w której numer wiersza (kolumny) odpowiada metodzie o oznaczeniu przyjętym w tabl. 2:

$$\mathbf{M} = \begin{bmatrix} 1,000 & 0,969 & 0,990 & 0,953 & 0,969 & 0,813 \\ 0,969 & 1,000 & 0,984 & 0,969 & 0,984 & 0,844 \\ 0,990 & 0,984 & 1,000 & 0,953 & 0,969 & 0,844 \\ 0,953 & 0,969 & 0,953 & 1,000 & 0,953 & 0,859 \\ 0,969 & 0,984 & 0,969 & 0,953 & 1,000 & 0,828 \\ 0,813 & 0,844 & 0,844 & 0,859 & 0,828 & 1,000 \end{bmatrix}$$

Wektor wartości miary podobieństwa ma postać:

$$[\bar{u}_p] = [0,939; 0,950; 0,948; 0,938; 0,941; 0,838]$$

Liniowe porządkowanie województw na podstawie zmiennej syntetycznej uzyskanej według metody 2, czyli na podstawie unitaryzacji zerowanej, jest najbliższe wszystkim pozostałym rankingom.

Podobne postępowanie przeprowadzono oddzielnie dla wyników uzyskanych metodami bezwzorcowymi i oddzielnie – metodami wzorcowymi.

W przypadku porządkowań otrzymanych za pomocą trzech metod bezwzorcowych uzyskano następujący wektor wartości miary podobieństwa:

$$[\bar{u}_p] = [0,961; 0,977; 0,969]$$

Natomiast dla trzech metod wzorcowych wektor ten przyjął postać:

$$[\bar{u}_p] = [0,969; 0,891; 0,844]$$

Jak wynika z analizy, w grupie metod bezwzorcowych najbliższe wszystkim porządkowaniom jest porządkowanie uzyskane metodą opartą na unitaryzacji zerowanej, a w grupie metod wzorcowych – porządkowanie metodą Hellwiga.

W opartym na unitaryzacji zerowanej porządkowaniu liniowym województw Polski ze względu na poziom aktywności ekonomicznej ludności według BAEL w I kwartale 2019 r. pierwszą lokatę zajęło woj. lubuskie, a ostatnią – woj. lubelskie. Najwyższa pozycja woj. lubuskiego wynika z niskiej stopy bezrobocia ogółem, a także z niskiej stopy bezrobocia wśród kobiet. Natomiast w woj. lubelskim zanotowano najwyższe wartości wszystkich cech przyjętych w badaniu, czyli liczby bezrobotnych na 1000 pracujących oraz poziomu stopy bezrobocia: ogółem, wśród kobiet, w miastach i na wsi.

#### 4. Podsumowanie

W artykule przedstawiono procedurę wyboru metody liniowego porządkowania obiektów. Różne metody mogą dawać znacznie różniące się od siebie wyniki liniowego porządkowania obiektów, w związku z czym pojawia się pytanie, które wyniki należy uznać za ostateczne. Procedura wyboru porządkowania najbliższego wszystkim pozostałym – a tym samym eliminacja porządkowań odstających od reszty – została zaproponowana w dwóch ujęciach. Pierwsze z nich polegało na wyborze najlepszego porządkowania spośród wyników uzyskanych różnymi metodami wzorcowymi i bezwzorcowymi. W drugim ujęciu wybór optymalnej metody został dokonany oddzielnie dla obu grup metod. Zastosowanie drugiego podejścia jest uzasadnione odmiennym sposobem wyznaczania porządkowań w grupach metod wzorcowych i bezwzorcowych, co często prowadzi do różnych wyników.

Z przeprowadzonych analiz wynika, że w grupie metod bezwzorcowych najbliższe wszystkim porządkowaniom otrzymanym przy użyciu rozpatrywanych metod jest porządkowanie uzyskane metodą opartą na unitaryzacji zerowanej. W grupie metod wzorcowych jest to porządkowanie metodą Hellwiga. Biorąc pod uwagę wyniki porządkowań przeprowadzonych na podstawie zmiennej syntetycznej, jakie otrzymano przy użyciu wszystkich rozpatrywanych metod łącznie, należy stwierdzić, że najbardziej zbliżone do pozostałych porządkowań okazało się porządkowanie metodą bezwzorcową z unitaryzacją zerowaną. Ta metoda została przyjęta jako formuła normująca.

W porządkowaniu liniowym województw pod względem poziomu aktywności ekonomicznej ludności w I kwartale 2019 r. opartym na unitaryzacji zerowanej pierwszą lokatę zajęło woj. lubuskie, a ostatnią – woj. lubelskie.

#### Bibliografia

- Balicki, A. (2009). *Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne*. Gdańsk, Sopot: Wydawnictwo Uniwersytetu Gdańskiego.
- Bąk, A. (2013). Metody porządkowania liniowego w polskiej taksonomii – pakiet pllord. *Taksonomia*, 20(278), 54–62.
- Bąk, A. (2017). Statystyczne metody doboru zmiennych w porządkowaniu liniowym. *Taksonomia*, 28(468), 29–37. DOI: 10.15611/pn.2017.468.03.
- Borys, T. (1978). Metody normowania cech w statystycznych badaniach porównawczych. *Przegląd Statystyczny*, 25(2), 227–239.
- Gatnar, E., Walesiak, M. (2004). *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*. Wrocław: Wydawnictwo Akademii Ekonomicznej im. Oskara Langego.
- GUS. (2016). *Aktywność ekonomiczna ludności Polski w latach 2013–2015*. Pobrane z: <https://stat.gov.pl/obszary-tematyczne/rynek-pracy/pracujacy-bezrobotni-bierni-zawodowo-wg-bael/aktywnosc-ekonomiczna-ludnosci-polski-w-latach-2013-2015,5,4.html>.

- GUS. (2019). *Aktywność ekonomiczna ludności w I kwartale 2019*. Pobrane z: <https://stat.gov.pl/obszary-tematyczne/rynek-pracy/pracujacy-bezrobotni-bierni-zawodowo-wg-bael/aktywnosc-ekonomiczna-ludnosci-polski-i-kwartal-2019-roku,4,33.html>.
- Hellwig, Z. (1968). Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr. *Przegląd Statystyczny*, (4), 307–327.
- Hwang, C. L., Yoon, K. (1981). *Multiple Attribute Decision Making: Methods and Applications: A State-of-the-Art Survey*. Berlin, Heidelberg, New York: Springer-Verlag.
- Kisielińska, J. (2016). Ranking państw UE ze względu na potencjalne możliwości zaspokojenia zapotrzebowania na produkty rolnicze z wykorzystaniem metod porządkowania liniowego. *Zeszyty Naukowe Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie. Problemy Rolnictwa Światowego*, 16(3), 142–153.
- Kukuła, K. (2000). *Metoda unitaryzacji zerowanej*. Warszawa: Wydawnictwo Naukowe PWN.
- Kukuła, K. (2014). Regionalne zróżnicowanie stopnia zanieczyszczenia środowiska w Polsce a gospodarka odpadami. *Przedsiębiorczość i Zarządzanie*, 15(8), 183–198.
- Kukuła, K., Luty, L. (2015). Propozycja procedury wspomagającej wybór metody porządkowania liniowego. *Przegląd Statystyczny*, (2), 219–231.
- Kukuła, K., Luty, L. (2017). Jeszcze o procedurze wyboru metody porządkowania liniowego. *Przegląd Statystyczny*, (2), 163–176.
- Kukuła, K., Luty, L. (2018). O wyborze metody porządkowania liniowego do oceny gospodarki odpadami w Polsce w ujęciu przestrzennym. *Zeszyty Naukowe Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie*, 18(2), 183–192. DOI: 10.22630/PRS.2018.18.2.46.
- Lira, J., Wagner, W., Wysocki, F. (2002). *Mediana w zagadnieniach porządkowania obiektów wielocechowych*. W: J. Paradyś (red.), *Statystyka regionalna w służbie samorządu lokalnego i biznesu* (s. 87–99). Poznań: Internetowa Oficyna Wydawnicza Centrum Statystyki Regionalnej.
- Malina, A. (2004). *Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw*. Kraków: Wydawnictwo Akademii Ekonomicznej w Krakowie.
- Malina, A., Zeliaś, A. (1997). O budowie taksonomicznej miary jakości życia. *Taksonomia*, 4, 238–263.
- Młodak, A. (2006). *Analiza taksonomiczna w statystyce regionalnej*. Warszawa: Difin.
- Młodak, A. (2009). Historia problemu Webera. *Matematyka Stosowana*, 37(51/10), 3–21. DOI: 10.14708/ma.v37i51/10.267.
- Nowak, E. (1990). *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*. Warszawa: Państwowe Wydawnictwo Ekonomiczne.
- Pawelek, B. (2008). *Metody normalizacji zmiennych w badaniach porównawczych złożonych zjawisk ekonomicznych*. Kraków: Wydawnictwo Uniwersytetu Ekonomicznego.
- Sokołowski, A., Markowska, M. (2017). Iteracyjna metoda liniowego porządkowania obiektów wielocechowych. *Przegląd Statystyczny*, (2), 153–162.
- Sompolska-Rzechuła, A. (2018). *Pomiar i ocena jakości życia w ujęciu regionalnym*. Szczecin: Wydawnictwo Uczelniane Zachodniopomorskiego Uniwersytetu Technologicznego.

- Sompolska-Rzechuła, A., Machowska-Szewczyk, M. (2018). Warunki życia w powiatach ziemskich województwa zachodniopomorskiego w latach 2002–2015 w świetle zmian demograficznych. *Przegląd Statystyczny*, (1), 115–136.
- Tarczyński, W., Łuniewska, M. (2006). *Metody wielowymiarowej analizy porównawczej na rynku kapitałowym*. Warszawa: Wydawnictwo Naukowe PWN.
- Walesiak, M. (1990). Syntetyczne badania porównawcze w świetle teorii pomiaru. *Przegląd Statystyczny*, (1–2), 37–46.
- Walesiak, M. (1993). Zagadnienie oceny podobieństwa zbioru obiektów w czasie w syntetycznych badaniach porównawczych. *Przegląd Statystyczny*, (1), 95–102.
- Wysocki, F. (2010). *Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich*. Poznań: Wydawnictwo Uniwersytetu Przyrodniczego.