

Wykorzystanie taksonomii wrocławskiej do oceny zróżnicowania warunków przyrodniczych i ekonomicznych

1. Wstęp

Obecnie, wiele uwagi poświęca się badaniom przestrzennym dużych obszarów, takich jak kontynenty, państwa, czy regiony. Jest to spowodowane nasilającymi się procesami globalizacji i integracji europejskiej.

Do badań przestrzennych stosuje się wiele metod, które umożliwiają podział dużej zbiorowości na mniejsze grupy ułatwiające ich charakterystykę, a także umożliwiające określenie podobieństw i różnic.

Jedną z takich metod jest metoda **taksonomii wrocławskiej**, nazywana także metodą dendrytu wrocławskiego. Pozwala ona na tworzenie pewnych skupisk, w skład których wchodzi obiekty najbardziej do siebie podobne, a wszystkie te skupiska stanowią całą zbiorowość. Umożliwia ona także wybór reprezentanta z każdej grupy, mającego być przedmiotem szczegółowych i pogłębionych analiz. Jest to również metoda, która dla badacza przystępującego do poznania nowego problemu samorzutnie wyodrębnia grupy w nieznaną zbiorowości analizując kilka cech jednocześnie.

Niniejsza praca była również próbą przeprowadzenia podziału przestrzennego obszaru województwa kujawsko-pomorskiego. Grupowano 92 gminy wiejskie tego województwa, stanowiące podstawową jednostkę samorządu terytorialnego naszego kraju. Dane zostały zaczerpnięte z Rocznika Statystycznego z 1998 roku.

Metoda taksonomii wrocławskiej pozwoliła więc na określenie grup gmin na terenie województwa w sposób prosty, tani i praktycznie użyteczny, a także na wybór gmin „typowych”, reprezentujących daną grupę gmin.

2. Cele pracy

Podstawowe cele pracy można sformułować następująco:

- zapoznanie się z metodą dendrytu wrocławskiego,
- wyodrębnienie względnie jednorodnych grup gmin wiejskich pod względem warunków przyrodniczych i ekonomicznych,
- wyodrębnienie gmin „typowych”, charakteryzujących w sposób możliwie najlepsze warunki przyrodnicze i ekonomiczne w określonych grupach województwa kujawsko-pomorskiego.

3. Metoda badań

3.1. Ogólna charakterystyka zastosowanej metody

Istnienie tak dużej liczby metod pozwala na wybór metody, która w sposób najbardziej dokładny pozwala grupować duże zbiorowości. W pracy do grupowania gmin województwa kujawsko-pomorskiego zastosowano metodę *taksonomii wrocławskiej*.

Metoda ta służy do podziału pewnej zbiorowości obiektów na grupy jednorodne, w skład których wchodziłyby obiekty najbardziej do siebie podobne. Każdy taki obiekt scharakteryzowany jest pewną liczbą cech; zarówno cech, jak i obiektów może być dowolna ilość.

Grupowanie metodą taksonomii wrocławskiej polega na dendrytowym uporządkowaniu poszczególnych obiektów na podstawie największego wzajemnego podobieństwa. Jest ono interpretowane jako funkcja odległości między obiektami w przestrzeni wielowymiarowej. Wymiar przestrzeni określany jest przez liczbę cech diagnostycznych. Odległość pomiędzy poszczególnymi obiektami w przestrzeni obliczana jest według różnych wzorów na podstawie zestandaryzowanych cech. Standaryzacja przeprowadzona jest w celu doprowadzenia do porównywalności zmiennych opisujących różne właściwości badanych obiektów [2].

Standaryzację przeprowadza się według wzoru:

$$z_{ij} = \left| \frac{\bar{x}_{ij} - x_j}{\delta_j} \right| \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, k$$

gdzie:

- \bar{x}_{ij} – elementy macierzy,
- x_j – średnia cechy j ,
- δ_j – odchylenie standardowe cechy j ,
- n – liczba badanych obiektów,
- k – liczba cech diagnostycznych.

Zestandaryzowane wartości cech służą do obliczenia tzw. **odległości taksonomicznych**, które zestawione w odpowiedniej tablicy, zwanej macierzą odległości, obrazują stopień wzajemnego podobieństwa poszczególnych obiektów.

Odległość taksonomiczną (d) oblicza się jako pierwiastek z przeciętnej kwadratów różnic wartości zmiennych standaryzowanych:

$$d_{ij} = \sqrt{\sum_{j=1}^n (z_{ij} - z_{ji})^2} \quad i, j = 1, 2, \dots, n$$

gdzie:

- d_{ij} – odległość między obiektami „ i ” a obiektem „ y ”,
- $d_{ij} = d_{ji}$,
- $d_{ij} = 0$ dla $i = j$.

W wielu badaniach ekonomicznych istnieje potrzeba redukcji liczby zmiennych opisujących badany wycinek rzeczywistości. W tym celu przeprowadza się procedurę wyznaczenia **zmiennych diagnostycznych**, które powinny mieć następujące właściwości:

- nie powinny być skorelowane lub co najwyżej mogą być słabo skorelowane między sobą,
- cecha diagnostyczna powinna być w wysokim stopniu skorelowana z cechami nie diagnostycznymi.

Redukcja cech może być przeprowadzana w różny sposób. Najczęściej stosowana jest redukcja cech na podstawie współczynników korelacji i wskaźników zmienności.

Wskaźnik zmienności (W_z) oblicza się według wzoru:

$$W_z = \frac{\delta \cdot 100}{x}$$

gdzie:

- δ – odchylenie standardowe,
- x – średnia wartość cechy.

Technika ta polega na wyodrębnianiu grup silnie ze sobą skorelowanych, następnie z każdej grupy wybiera się cechę o najwyższym wskaźniku zmienności. To pozwala ocenić każdą cechę diagnostyczną z uwagi na przydatność merytoryczną na każdym etapie wyboru cech. Poprzez analizę związków korelacyjnych pomiędzy cechami i zastosowanie kryterium najbliższego sąsiedztwa wykreśla się dendryt najsilniejszych powiązań. Dendryt tworzy powiązania korelacyjne między badanymi cechami, co pozwala utworzyć skupienie cech najsilniej skorelowanych. Spośród każdego skupienia wyodrębnia się cechę diagnostyczną.

Kolejnym etapem grupowania jest wykreślenie spójnego dendrytu przedstawiającego najkrótsze odległości między poszczególnymi parami obiektów. Dendryt jest linią łamaną łączącą wszystkie punkty (obiekty), która może się rozgałęziać, lecz nie może zawierać łamanych zamkniętych.

Przy wykreślaniu dendrytu spójnego tworzą się naturalne skupienia obiektów bardzo do siebie podobnych. Tworzą one szereg niespójnych dendrytów, które określane są mianem dendrytów pierwszego rzędu. Poszczególne skupienia łączone są następnie ze sobą na podstawie najkrótszych połączeń międzyskupieniowych, otrzymuje się w ten sposób skupienia drugiego rzędu. Zakończenie tworzenia dendrytu następuje z chwilą połączenia wszystkich obiektów badanego zbioru, czyli po skonstruowaniu dendrytu spójnego.

Następnym etapem jest podział dendrytu na podstawie naturalnych skupień w celu wyodrębnienia grup obiektów podobnych. Podziału dokonuje się przez usunięcie niektórych połączeń. Najczęściej stosuje się podział na podstawie **odległości krytycznej** Hellwiga oraz analizy powiązań między dendrytami wyższych stopni.

Odległość krytyczną oblicza się według wzoru:

$$d_{kr} = d + 2\delta_d$$

gdzie:

- d_{kr} – odległość krytyczna,
- d – średnia arytmetyczna odległość w dendrycie,
- δ_d – odchylenie standardowe odległości w dendrycie.

Dzieląc dendryt przez usunięcie odległości krytycznych, oddziela się zwykle od dużej zbiorowości punktów pojedyncze lub niewielkie grupki obiektów nietypowych. Zaleca się więc często, aby podziału dendrytu dokonywać w miejscach powiązań wyższych rzędów, np.: trzeciego i wyższych.

4. Sposób przedstawienia danych

Przystępując do grupowania gmin wiejskich województwa kujawsko-pomorskiego zebrano podstawowe cechy charakteryzujące każdą gminę.

Pierwotnie zostały wybrane 32 cechy, przedstawione w tabeli 1.

Tabela 1. Wykaz wszystkich zmiennych przed redukcją

Numery zmiennych	Nazwy zmiennych
1	Powierzchnia gminy w km
2	Ogólna liczba ludności
3	Liczba mężczyzn
4	Liczba kobiet
5	Liczba mieszkańców przypadająca na km ²
6	Ludność w wieku produkcyjnym
7	Ludność w wieku poprodukcyjnym
8	Napływ ludności ogółem
9	Napływ z miasta
10	Napływ do wsi
11	Odptyw ogółem
12	Odptyw z miast
13	Odptyw ze wsi
14	Liczba pracujących
15	Liczba bezrobotnych
16	Podmioty gospodarcze
17	Sektor publiczny
18	Sektor prywatny
19	Ogólny dochód w gminie
20	Dochód przypadający na 1 mieszkańca gminy
21	Ogólna powierzchnia gruntów ornych
22	Powierzchnia użytków rolnych
23	Liczba indywidualnych gospodarstw
24	Przeciętna powierzchnia indywidualnego gospodarstwa
25	Przeciętna powierzchnia użytków rolnych
26	Ogólna liczba bydła w gminie
27	Ogólna liczba trzody chlewnej w gminie
28	Liczba poczty w gminie
29	Liczba telefonów w gminie
30	Długość dróg w km
31	Długość wodociągów w km
32	Długość sieci kanalizacyjnej w km

Źródło: opracowanie własne na podstawie Rocznika Statystycznego z 1998 r.

Wybrane zmienne zostały poddane obróbce statystycznej. Obliczono średnią, minimum, maksimum, odchylenie standardowe oraz współczynnik zmienności. Na podstawie współczynnika zmienności przeprowadzono redukcję zmiennych i pozostawiono 14 zmiennych. Wykaz tych cech zamieszczono w tabeli 2.

Tabela 2. Wykaz zmiennych po redukcji

Numery zmiennych	Nazwa zmiennych
2	Ogólna liczba ludności
5	Ludność przypadająca na km ²
6	Ludność w wieku produkcyjnym
9	Napływ z miasta
12	Odpływ z miast
14	Liczba pracujących
15	Liczba bezrobotnych
18	Sektor prywatny
19	Ogólny dochód w gminie
21	Ogólna powierzchnia gruntów ornych
23	Liczba indywidualnych gospodarstw w gminie
29	Liczba telefonów
30	Długość dróg w km
32	Długość sieci kanalizacyjnej w km

Źródło: opracowanie własne na podstawie Rocznika Statystycznego z 1998 r.

5. Wybór cech diagnostycznych

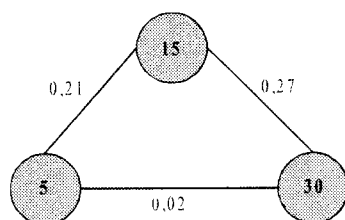
Wybranych po redukcji 14 zmiennych zostało ponownie poddanych obróbce statystycznej w celu wybrania cech diagnostycznych do grupowania taksonomicznego.

Zgodnie z założoną metodą badań, wyboru cech diagnostycznych dokonano na podstawie siły związków korelacyjnych i wskaźników zmienności cech. Ostatecznie dokonano wyboru trzech cech diagnostycznych. Są nimi:

- liczba bezrobotnych – cecha nr 15,
- liczba mieszkańców przypadająca na km² – cecha nr 5,
- długość dróg w km – cecha nr 30.

Podjęto także próbę wykonania dendrytów i grupowania przy użyciu innych cech diagnostycznych. Były to: w pierwszym dendrycie – liczba bezrobotnych, ogólna powierzchnia gruntów ornych i długość sieci kanalizacyjnej w km oraz w drugim dendrycie – liczba bezrobotnych, dochód na 1 mieszkańca i długość dróg. Próby te nie przyniosły jednak oczekiwanych rezultatów, dlatego zrezygnowano z nich i rozpoczęto grupowanie przy użyciu wyżej wymienionych cech diagnostycznych, które zakończyło się powodzeniem.

Stopień skorelowania cech diagnostycznych przedstawiono na rysunku 1.



Rys. 1. Stopień skorelowania cech diagnostycznych

Źródło: opracowanie własne

Jak widać z podanych wielkości, wzajemny stopień skorelowania tych cech jest niski. Odpowiadają więc formalnym wymogom stawianym cechom diagnostycznym.

6. Konstrukcja spójnego dendrytu oraz jego podział

Przy pomocy programu „Dendryt”, na podstawie cech diagnostycznych obliczono tablicę odległości i dokonano wyboru najmniejszych odległości pomiędzy wszystkimi gminami w województwie kujawsko-pomorskim. Wybrane wielkości posłużyły do wykreślenia spójnego dendrytu przedstawionego na rysunku 2.

Stosując zasadę Hellwiga, na podstawie odległości krytycznej wydzielono pięć gmin „nietypowych” (odległość krytyczna dla całej zbiorowości wyniosła 1,429). Oddzielone gminy to:

- gmina nr 83 – **Wąbrzeźno**, odległość krytyczna dla tej gminy wyniosła 1,585;
- gmina nr 49 – **Osiężyny**, odległość krytyczna – 1,481;
- gmina nr 73 – **Wielka Nieszawka**, odległość krytyczna – 2,515;
- gmina nr 69 – **Lubicz**, odległość krytyczna – 2,050;
- gmina nr 66 – **Wartubie**, odległość krytyczna – 1,712.

Stosując zaś zasadę podziału dendrytu podaną przez Groszka i Zagórskiego wyodrębniono 10 grup gmin i 5 gmin „nietypowych” dzieląc dendryt w miejscach połączeń III rzędu [1]. W celu sprawdzenia poprawności grupowania przeprowadzono obliczenia wartości średnich, odchyłeń standardowych oraz wskaźników zmienności wszystkich cech w badanych grupach. Wyliczeń dokonano w dwóch grupach:

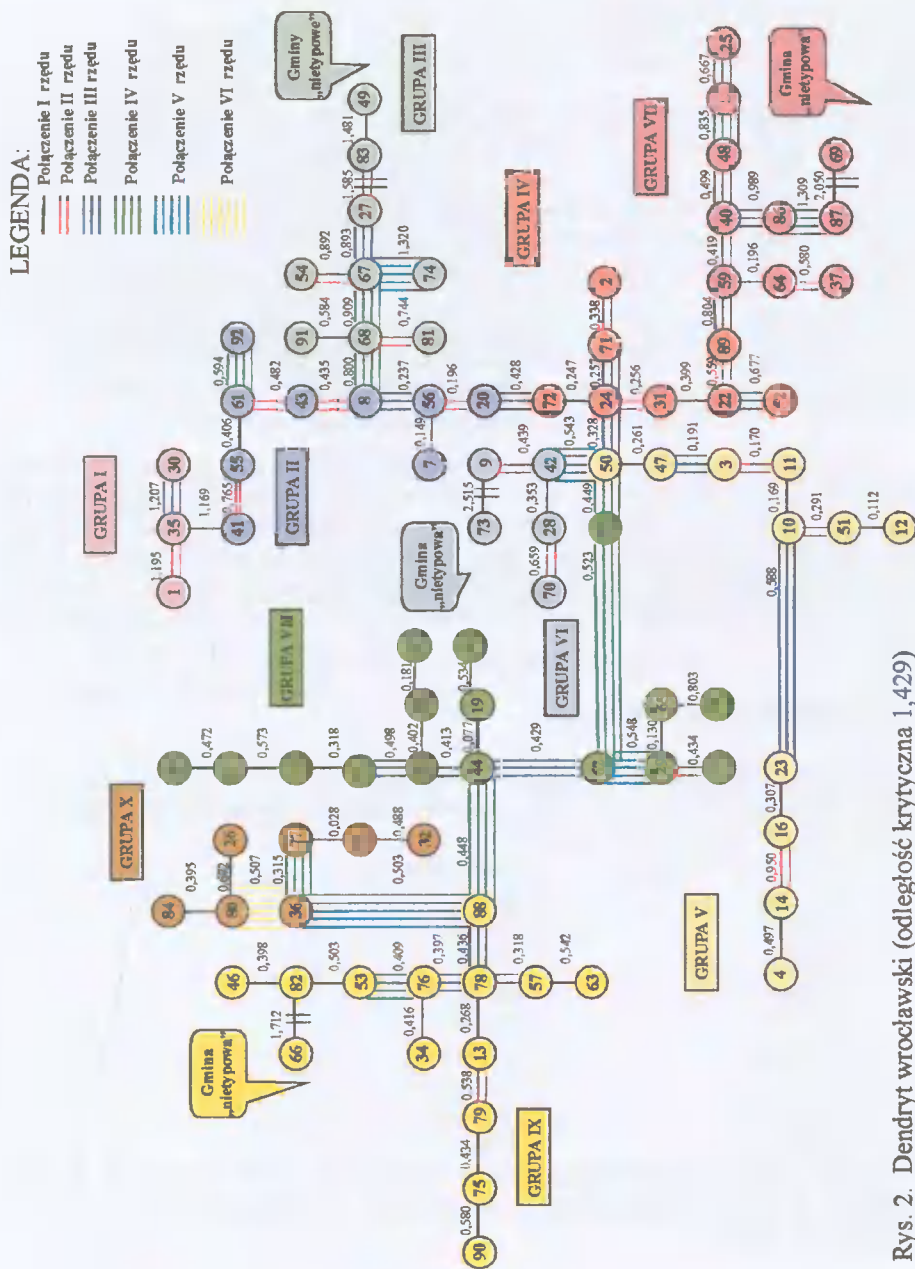
- bez gmin „nietypowych”,
- po dołączeniu tych gmin do grup, do których są najbardziej przystające.

Zostało to przedstawione w tabelach 3 i 4.

Tabela 3. Porównanie współczynników zmienności w grupach gmin (podział pierwszy – bez gmin „nietypowych”)

Grupa	Współczynnik zmienności			Numery gmin
	bezrobotni	ludność/km ²	drogi	
razem gminy	40,15	33,91	46,78	
I	6,23	13,16	16,80	1,35,30
II	17,91	9,77	8,48	41,55,61,92,43,8,56,7,20
III	20,62	13,60	13,41	68,81,67,91,54,74,2,7
IV	10,72	9,90	11,33	24,24,2,71,31,22,52,22
V	11,57	18,90	30,68	50,47,11,3,10,51,12,23,1,14,4
VI	13,71	16,90	1,74	42,28,9,70
VII	27,14	12,89	30,41	85,62,29,65,21,18,44,19,38,60,58,33,6,15,1,39
VIII	19,09	14,97	14,98	59,64,37,40,86,87,48,5,25
IX	14,60	17,71	26,87	88,78,57,63,13,79,75,90,76,34,53,82,46
X	26,79	7,39	7,61	36,77,45,32,80,26,84

Źródło: opracowanie własne



Rys. 2. Dendryt wrocławski (odległość krytyczna 1,429)

Tabela 4. Porównanie współczynników zmienności w grupach gmin (podział drugi – po dołączeniu gmin „nietypowych”)

Grupa	Współczynnik zmienności			Numery gmin
	bezrobotni	ludność/km ²	drogi	
razem gminy	40,15	33,91	46,78	
I	6,23	13,16	16,80	1,35,30
II	17,91	9,77	8,48	41,55,61,92,43,8,56,7,20
III	18,2	23,48	33,91	68,81,67,91,54,74,2,7,83,49
IV	10,72	9,90	11,33	24,24,2,71,31,22,52,22
V	11,57	18,90	30,68	50,47,11,3,10,51,12,23,16,14,4
VI	19,99	15,14	6,28	42,28,9,70,73
VII	27,14	12,89	30,41	85,62,29,65,21,18,44,19,38,60,58,33,6,15,17,39
VIII	18,66	19,14	26,74	59,64,37,40,86,87,48,5,25,69
IX	23,98	20,71	26,66	88,78,57,63,13,79,75,90,76,34,53,82,46,66
X	26,79	7,39	7,61	36,77,45,32,80,26,84

Źródło: opracowanie własne

Porównując oba zestawienia (tabela 3 i 4) można stwierdzić, że wyniki są zbliżone – gminy „nietypowe” dołączone do poszczególnych grup nie spowodowały znaczących zmian. Przyłączenie tych gmin nie doprowadziło do przekroczenia wartości współczynników zmienności dla poszczególnych grup w stosunku do ich wartości dla całej zbiorowości. Ostatecznie można więc przyjąć, że na podstawie wytypowanych cech diagnostycznych gminy województwa kujawsko-pomorskiego dzielą się na 10 podstawowych grup.

7. Wybór gmin typowych

Porównując średnie wartości cech można przyjąć, że dla poszczególnych grup najbardziej charakterystyczne (typowe) z punktu widzenia warunków przyrodniczych i ekonomicznych są następujące gminy:

- Aleksandrów Kujawski – I grupa gmin,
- Brodnica – II grupa,
- Chełmża – III grupa,
- Grua – IV grupa,
- Grażawy – V grupa,
- Brzozie – VI grupa,
- Dąbrowa – VII grupa,
- Kikół – VIII grupa,
- Lubiewo – IX grupa,
- Rojewo – X grupa.

Średnie wartości cech diagnostycznych charakteryzujących te gminy są zbliżone do średnich w grupie, a wytypowane gminy mieszczą się w środkowej części dendrytu w danej grupie gmin.

8. Wnioski

Taksonomia wrocławska służy do podziału pewnej zbiorowości obiektów na grupy jednorodne, w skład których wchodziłyby elementy najbardziej do siebie podobne. Za taką zbiorowość zostały uznane 92 gminy wiejskie województwa kujawsko-pomorskiego, które pogrupowano metodą dendrytową.

Przeprowadzone badania pozwalają więc na wysunięcie następujących wniosków:

1. Metoda dendrytowa jest skutecznym sposobem przestrzennego podziału danego obszaru na grupy jednorodne.
2. Ze względu na warunki przyrodnicze i ekonomiczne gminy wiejskie województwa kujawsko-pomorskiego można podzielić na 10 różniących się między sobą grup.
3. Dla każdej grupy gmin można ustalić gminę typową, tzn. najlepiej charakteryzującą daną grupę.
4. Czynnikiemami najsilniej delimitującymi gminy wiejskie województwa są: liczba bezrobotnych, długość dróg w km i liczba mieszkańców przypadająca na km².
5. Metoda ta pozwala także ocenić stopień zróżnicowania warunków przyrodniczych i ekonomicznych w wydzielonych grupach gmin wiejskich.

Literatura

1. Groszek T., Zagórski K., 1972. Zastosowanie taksonomii wrocławskiej, analizy czynnikowej do badań struktury społeczno-zawodowej ludności. *Wiadomości Statystyczne* 10.
2. Pluta W., 1977. Wielowymiarowa analiza porównawcza w badaniach ekonomicznych. PWE Warszawa.