METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XII/2, 2011, str. 58–68

WIELOKRYTERIALNE METODY ANALIZY ZRÓŻNICOWANIA POLSKIEGO ROLNICTWA W 2009 ROKU

Agata Binderman
Katedra Ekonometrii i Statystyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: agata_binderman@sggw.pl

Streszczenie: W pracy przedstawiono wykorzystanie trzech metod wzorcowych do analizy regionalnego zróżnicowania poziomu rolnictwa w 2009 roku, po zastosowaniu ilorazowej normalizacji zmiennych. Wyniki porównano z rezultatami otrzymanymi przy pomocy bezwzorcowego, miernika rozwoju typu Pluty. Przy pomocy czterech mierników skonstruowano piąty, „wypadkowy" i na tej podstawie dokonano liniowego uporządkowania oraz pogrupowania polskich województw. Wyniki wskazują, że poziom rozwoju polskiego rolnictwa między województwami w roku 2009 był bardzo zróżnicowany. Z otrzymanych obliczeń można wnioskować, że wybór metody wzorcowej nie ma wpływu na grupowanie województw oraz, że ma niewielki wpływ na wartości ich mierników syntetycznych.

Słowa kluczowe: poziom rozwoju rolnictwa, metody wzorcowe, TOPSIS, metoda bezwzorcowa, miernik syntetyczny, klasyfikacja

WSTĘP

Rolnictwo polskie charakteryzuje się znaczącymi zasobami użytków rolnych oraz wysokim potencjałem produkcyjnym. Sytuacja ekonomiczna polskiej wsi jest jednak silnie zróżnicowana. Przekształcenia zachodzące w gospodarce wywołują wiele problemów, do których przede wszystkim należy zróżnicowanie przestrzenne rolnictwa. Temat ten był już przedmiotem wielu badań, w których zostało udokumentowane, że istnieje na wielu płaszczyznach, zarówno w przestrzeni geograficznej, jak i czasowej [Zegar 2002, Wilkin 2001, Hunek 2002, Krasowicz 2009, Borkowski, Szczesny 2002, Binderman 2007].

Niniejsza praca jest kontynuacją wcześniejszych szeregu prac autorki, które dotyczyły regionalnego zróżnicowania polskiego rolnictwa. Celem głównym przedstawionych badań było uporządkowanie i podział województw na grupy ze względu na poziom rozwoju rolnictwa, cel szczegółowy to wyznaczenie syntetycznych mierników zróżnicowania regionalnego rolnictwa.

Hipotezą badawczą w pracy było stwierdzenie, że poziom rozwoju polskiego rolnictwa pomiędzy województwami w roku 2009 był nadal bardzo zróżnicowany. Grupowanie województw pod względem poziomu rozwoju rolnictwa, przy ilorazowej normalizacji cech nie zależy w istotnym stopniu od zastosowanej metody (wzorcowej) konstruowania mierników syntetycznych.

METODY BADAŃ

Podstawowym celem analizy taksonomicznej jest dokonanie grupowania i porządkowanie obiektów będących elementami wielowymiarowej przestrzeni cech. Do klasyfikacji i grupowania stosowanych jest wiele metod [zob. Bartosiewicz 1976, Cieślak 1976, Gatnar, Walesiak 2009, Gatnar, Wywiał 2007, Hellwig 1968, Kukuła 2000, Malina 2004, Młodak 2006, Strahl D. 1990, Walesiak 2005, Zeliaś 2000]. Najczęściej stosowane sposoby do klasyfikacji obiektów (grupowania) to metody centroidalne, Warda, wiązań oraz jednorodnego rdzenia, k najbliższych sąsiedztw, taksonomii wrocławskiej. Do porządkowania obiektów stosowane są metody, które można podzielić na bezwzorcowe (oparte głównie na średnich matematycznych) i wzorcowe (Hellwiga, TOPSIS). Metody bezwzorcowe polegają na konstrukcji miernika syntetycznego agregatowego na podstawie tylko znormalizowanych wartości cech. Metody wzorcowe polegają na skonstruowaniu taksonomicznych mierników rozwoju (sztucznych punktów odniesienia), mierzeniu odległości od tych wzorców i na tej podstawie konstruowaniu miernika syntetycznego.

Badane zjawisko zostało opisane przez zmienne będące stymulantami. Każdy obiekt (obserwacja) opisany był za pomocą wektora, będącego elementem przestrzeni:

$$\Re^n := \{\mathbf{x} = (x_1, x_2, \ldots, x_n) : x_i \in \Re;\ i = 1, 2, \ldots, n\}, \Re = (-\infty, \infty), n \in \mathbb{N},$$

gdzie $n \geq 1$ jest ilością zmiennych zakwalifikowanych do oceny zjawiska.

Przyjęto, że $X = \Re^n$, $n \in N$, oznacza n-wymiarową przestrzeń wektorową. Rozważono problem polegający na klasyfikacji $m \in N$ obiektów $\mathbf{Q}_1$, $\mathbf{Q}_2$,...,$\mathbf{Q}_m$ za pomocą $n \in N$ zmiennych (cech). Wektor $\mathbf{x}_i=(x_{i1},x_{i2},...,x_{in}) \in X$, i=1,2,...,m, opisuje i-ty obiekt. Jeżeli $x_{ik} > x_{jk}$ ($x_{ik} \geq x_{jk}$) dla k=1,2,...,n, to można zapisać, że $\mathbf{x}_i > \mathbf{x}_j$, ($\mathbf{x}_i \geq \mathbf{x}_j$), gdzie $i, j \in [1,m]$.

Nietrudno zauważyć, że jeżeli $\mathbf{x}_i > \mathbf{x}_j$ i $\mathbf{x}_i \neq \mathbf{x}_j$, to w niektórych przypadkach naturalnym jest nazywać obiekt $\mathbf{x}_i$ lepszym (wyżej ocenianym) od obiektu $\mathbf{x}_j$. Oznacza to, że żadna ze składowych wektora $\mathbf{x}_i$ nie jest mniejsza od odpowiednich składowych wektora $\mathbf{x}_j$, a przynajmniej jedna z nich ma wartość większą, tj. istnieje takie $k \in [1,n]$, że $x_{ik} > x_{jk}$.

Przyjęto w pracy następujące oznaczenia:

$$x_{0,k} = \min_{1 \le i \le m} x_{ik}, \quad x_{m+1,k} = \max_{1 \le i \le m} x_{ik}, \quad k = 1, 2, \ldots, n;$$

oraz $\mathbf{x}_0 := (x_{0,1}, x_{0,2}, \ldots, x_{0,n})$, $\mathbf{x}_{m+1} := (x_{m+1,1}, x_{m+1,2}, \ldots, x_{m+1,n})$.

Obiekt $\mathbf{Q}_0$, opisany przez wektor $\mathbf{x}_0$ oraz $\mathbf{Q}_{m+}1$ przez wektor $\mathbf{x}_{m+1}$ (być może fikcyjne) są odpowiednio nie gorsze, nie lepsze od pozostałych obiektów, tj.:

$$\mathbf{x}_{m+1} \ge \mathbf{x}_i \text{ oraz } \mathbf{x}_i \ge \mathbf{x}_0 \text{ dla każdego i: } m \ge i \ge 1.$$

Fakt ten możemy również zapisać jako relacje: $\mathbf{Q}_{m+1} \succeq \mathbf{Q}_i$, $\mathbf{Q}_i \succeq \mathbf{Q}_0$, i=1,2,...,m [Panek 2000]. W przypadku, gdy obiekty $\mathbf{Q}_0$ i $\mathbf{Q}_{m+1}$ są różne od rozważanych obiektów $\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_m$, to spełniają one odpowiednio rolę obiektu *najlepszego* oraz *najgorszego* i są traktowane, jako wzorce. W związku z powyższym:

$$<\mathbf{x}_0, \mathbf{x}_{m+1}> := \{\mathbf{x} \in \Re^n_+ : \mathbf{x}_0 \le \mathbf{x} \le \mathbf{x}_{m+1}\}.$$

W pracy do mierzenia odległości między wektorami (obiektami) $\mathbf{x}, \mathbf{y} \in X$ wykorzystano prostą do interpretacji metrykę Euklidesa

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{j=1}^{n} \left(x_j - y_j\right)^2} \tag{1}$$

W dalszej części rozważań założono, że obiekty wzorcowe są różne, tj. $\mathbf{x}_0 \ne \mathbf{x}_{m+1}$. Przyjęcie takiego założenia powoduje, że $d(\mathbf{x}_0, \mathbf{x}_{m+1}) \ne 0$ oraz $d(\mathbf{x}_0, \mathbf{x}) + d(\mathbf{x}_{m+1}, \mathbf{x}) \ne 0$ dla $\mathbf{x} \in \langle \mathbf{x}_0, \mathbf{x}_{m+1} \rangle$.

W literaturze, do klasyfikacji obiektów wykorzystywać można następujące wzory, określające mierniki syntetyczne danego wektora $\mathbf{x} \in \langle \mathbf{x}_0, \mathbf{x}_{m+1} \rangle$:

$$m_1(\mathbf{x}) = \frac{d(\mathbf{x}_0, \mathbf{x})}{d(\mathbf{x}_0, \mathbf{x}_{m+1})} \tag{2}$$

$$m_2(\mathbf{x}) = 1 - \frac{d(\mathbf{x}_{m+1}, \mathbf{x})}{d(\mathbf{x}_0, \mathbf{x}_{m+1})} \tag{3}$$

$$m_3(\mathbf{x}) = \frac{m_1(\mathbf{x}) + m_2(\mathbf{x})}{2} = \frac{1}{2} + \frac{d(\mathbf{x}_0, \mathbf{x}) - d(\mathbf{x}_{m+1}, \mathbf{x})}{2d(\mathbf{x}_0, \mathbf{x}_{m+1})} \tag{4}$$

$$m_4(\mathbf{x}) = \sqrt{m_1(\mathbf{x}) m_2(\mathbf{x})} = \frac{1}{d(\mathbf{x}_0, \mathbf{x}_{m+1})} \sqrt{d(\mathbf{x}_0, \mathbf{x})\left(d(\mathbf{x}_0, \mathbf{x}_{m+1}) - d(\mathbf{x}, \mathbf{x}_{m+1})\right)} \tag{5}$$

$$m_5(\mathbf{x}) = \frac{m_1(\mathbf{x})}{1 + m_1(\mathbf{x}) - m_2(\mathbf{x})} = \frac{d(\mathbf{x}_0, \mathbf{x})}{d(\mathbf{x}_0, \mathbf{x}) + d(\mathbf{x}_{m+1}, \mathbf{x})} \tag{6}$$

Mierniki m_1 i m_2 wykorzystują jeden wzorzec, natomiast m_3, m_4 i m_5 dwa wzorce. Mierniki te można potraktować, jako narzędzia rozwiązywania wielokryterialnych problemów decyzyjnych. Mierniki m_1 i m_2 wykorzystują jedno kryterium natomiast m_3, m_4, m_5 wykorzystują dwa kryteria.

Podane wyżej mierniki są znormalizowane tj.

$$0 \le m_k(\mathbf{x}) \le 1 \quad \text{dla} \quad \mathbf{x} \in < \mathbf{x}_0, \mathbf{x}_{m+1} >,\ k = 1, 2, \ldots, 5.$$

Zauważmy ponadto, że:

$$m_k(\mathbf{x}_0) = 0,\ m_k(\mathbf{x}_{m+1}) = 1,\ m_k(\overline{\mathbf{x}}_a) = \frac{1}{2},\ \text{dla k} = 1,2,\ldots,5,$$

gdzie $\overline{\mathbf{x}}_a := (\mathbf{x}_0 + \mathbf{x}_{m+1}) / 2$ jest obiektem pośrednim, być może fikcyjnym.

Warto wspomnieć, że w pracy [Hellwig 1968] podany został miernik obiektów wykorzystujący tylko obiekt najlepszy i standaryzację cech. W pracach autorki [Binderman 2009, 2010] porównano wyniki otrzymane przy pomocy metody Z. Hellwiga z wynikami otrzymanymi poprzez wykorzystanie miernika m_3. Teoria i zastosowania miernika m_3 zostały podane również w serii prac autorki [Binderman 2006, 2007, 2008, 2009, 2010, 2010a, Binderman, Bojarski 2010] oraz w pracy [Binderman Z. 2010]. Miernik m_4 związany jest z metodą TOPSIS (Technique for order Preference by Similarity to Ideal Solution) [Hwang, Yoon 1981]. Metodę TOPSIS należy zakwalifikować do metod rankingowych służących do podejmowania decyzji wielokryterialnych. Metodami tego typu są również znane metody SAW (Simple Additive Weighting) [Chen, Hwang 1992, Neuman 1998, Zhang 2004] oraz AHP (Analytical Hierarchy Process) [Saaty 1980, 1995].

W pracach autorki [Binderman 2004, 2005, 2006] pokazano, że stosowanie metod opartych tylko na jednym wzorcu, w wielu przypadkach prowadzi do otrzymania błędnych wyników, które nie spełniają warunków poprawności [Jackson 1970]. Z powyższego powodu w niniejszej pracy, do analizy regionalnego zróżnicowania polskiego rolnictwa wykorzystano tylko te mierniki, które są oparte na dwóch wzorcach. Wyniki porównano z rezultatami otrzymanymi przy pomocy bezwzorcowego, izotonicznego miernika rozwoju typu M. Pluty [Malarska 2006, Malina 2004], miernik ten jest określony za pomocą wzoru :

$$p(\mathbf{x}) = \sum_{j=1}^{n} x_j. \tag{7}$$

Funkcję $p(\mathbf{x})$ można traktować, jako funkcję użyteczności [Panek 2000], funkcja ta nie jest funkcją znormalizowaną, gdyż jej wartości mogą należeć do przedziału [0, n]. Dla porównywalności liczbowej wyników otrzymanych za pomocą funkcji $p(\mathbf{x})$ z wynikami otrzymanymi za pomocą rozważanych wcześniej, znormalizowanych mierników dokonajmy normalizacji funkcji $p(\mathbf{x})$. Takiej normalizacji, która generuje relację preferencji identyczną z relacją indukowaną przez funkcję $p(\mathbf{x})$. W tym celu zauważmy, że jeżeli dana funkcja użyteczności u indukuje relacje preferencji obiektów zbioru $W:=(\mathbf{x}_0,\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_m,\mathbf{x}_{m+1})$ to funkcja złożona g(u(x)), gdzie g: $\Re \to \Re$ jest dowolną funkcją rosnącą, jest również funkcją użyteczności, generującą tą samą relację preferencji w zbiorze obiektów W, co funkcja u [Panek 2000].

Wykorzystując powyższą własność, funkcję p($\mathbf{x}$) możemy unormować na wiele sposobów [Binderman 2007]. Polega to na wybraniu takiej funkcji g, aby jej wartość dla obiektu najgorszego $\mathbf{x}_0$ wynosiła 0, wartość zaś dla obiektu najlepszego $\mathbf{x}_{m+1}$ wynosiła 1, tj., aby: $g(p(\mathbf{x}_0))=0$, $g(p(\mathbf{x}_{m+1}))=1$.

Funkcją o tej własności może być na przykład funkcja liniowa: $g(t):=(t-t_0)/(t_1-t_0)$, gdzie $t_1:= p(\mathbf{x}_{m+1})$, $t_0:= p(\mathbf{x}_0)$. Przy oczywistym założeniu, że $t_1> t_0$, gdyż dopuszczenie przypadku $t_1=t_0$ oznaczałoby, że wszystkie rozważane obiekty mają tą samą użyteczność, otrzymaną za pomocą funkcji u. W tym przypadku funkcja

$$m_6(\mathbf{x}):= g((p(\mathbf{x})) = \frac{p(\mathbf{x}) - p(\mathbf{x}_0)}{p(\mathbf{x}_{m+1}) - p(\mathbf{x}_0)}, \tag{8}$$

określona na zbiorze obiektów W, jest funkcją użyteczności mającą tą własność, że $m_6(\mathbf{x}_0) = 0$, $m_6(\mathbf{x}_{m+1}) = 1$.

Z uwagi na fakt, że miernik Pluty p($\mathbf{x}$) wykorzystuje ilorazową normalizację zmiennych, to w pracy również zastosowano tą metodę normalizacji, jest ona określona za pomocą wzoru:

$$z_{ij} = x_{ij}\left(\sum_{k=1}^{m} x_{kj}\right)^{-1}, \ i = 1, 2, ..., m; \ j = 1, 2, ..., n. \tag{9}$$

Takie podejście do problemu pozwala wychwycić różnice w poziomie cech diagnostycznych przy równoczesnej eliminacji wpływu jednostek miar poszczególnych cech na wynik badania.

Według L. Breimana [Breiman 1994], (zob. również Breiman 1996, 1998]), pojedynczy klasyfikator (miernik syntetyczny) może być daleki od optymalnego, natomiast kombinacje wielu mogą dać stabilny klasyfikator - bliski optymalnemu. Niestety w przypadku „słabych klasyfikatorów" w wyniku ich kombinacji można otrzymać klasyfikator „gorszy" od każdego z nich.

Z powyższego powodu, do analizy zróżnicowania poziomu rozwoju polskiego rolnictwa w roku 2009 wybrano klasyfikator, będący średnią geometryczną mierników m_3, m_4, m_5, m_6, który został obliczony według wzoru:

$$m(\mathbf{x}) = \sqrt[4]{m_3(\mathbf{x})\, m_4(\mathbf{x})\, m_5(\mathbf{x})\, m_6(\mathbf{x})}, \ \mathbf{x} \in [\mathbf{x}_0, \mathbf{x}_{m+1}], \tag{10}$$

gdzie mierniki m_3, m_4, m_5, m_6 są określone za pomocą wzorów (4), (5), (6), (8), odpowiednio. Warto zauważyć, że miernik ten, jako średnia geometryczna funkcji znormalizowanych jest również znormalizowany.

WYNIKI BADAŃ

Do badań empirycznych wykorzystano dane makroekonomiczne uzyskane z Głównego Urzędu Statystycznego dla poszczególnych 16 województw w 2009 r.

W pracy doktorskiej autorki [Binderman 2007] przedstawiono dyskusję związaną z wyborem zmiennych do analizy polskiego rolnictwa. Z wielu czynników, jakie wpływają na ocenę rozwoju rolnictwa, wyodrębniono 70 cech, a następnie w oparciu o dane z lat 1998–2005 dokonano ich selekcji pod względem merytorycznym oraz statystycznym. W niniejszej pracy został przyjęty ten sam wybór 10 następujących zmiennych. $\mathbf{X}_1$ – udział użytków rolnych w powierzchni ogółem, $\mathbf{X}_2$ – wskaźnik waloryzacji rolniczej przestrzeni produkcyjnej (w punktach), $\mathbf{X}_3$ – plony zbóż w tonach z 1 hektara, $\mathbf{X}_4$ – obsada bydła w sztukach dużych na 100 hektarów użytków rolnych, $\mathbf{X}_5$ – skup owoców z drzew w kg na 1 hektar powierzchni upraw, $\mathbf{X}_6$ – wartość skupu produktów rolnych ogółem w zł na 1 ha użytków rolnych, $\mathbf{X}_7$ – nakłady inwestycyjne w rolnictwie w zł na 1 hektar użytków rolnych, $\mathbf{X}_8$ – wskaźnik zatrudnienia na obszarach wiejskich w %, $\mathbf{X}_9$ – średnia powierzchnia gospodarstwa rolnego w hektarach, $\mathbf{X}_{10}$ – produkt krajowy brutto w zł na 1 mieszkańca.

Czynniki wybrane do badań były stymulantami, tj. większe wartości cech oznaczały wyższy poziom badanego zjawiska. Każde województwo było opisane przez wektor o dziesięciu współrzędnych i dodatnich wartościach, które to wektory utworzyły macierz danych $\mathbf{X} = \left[x_{ij} \right]_{16\times 10}$ o szesnastu wierszach i dziesięciu kolumnach. Województwa zostały uporządkowane alfabetycznie, tj. wektor $\mathbf{x}_1 = (x_{11}, x_{12}, ..., x_{110})$ określa województwo dolnośląskie, $\mathbf{x}_2 = (x_{21}, x_{22}, ..., x_{210})$ kujawsko-pomorskie, itd. Tabela 1 przedstawia dane wejściowe do analizy rozważanego w pracy problemu. Wektory wzorcowe i podstawowe dane statystyczne poddano wcześniej ilorazowej normalizacji cech, określonej za pomocą wzoru (9).

W tabeli 1 wektory $\mathbf{x}_{17}$, $\mathbf{x}_0$ określają hipotetyczne, wzorcowe województwa, odpowiednio najlepsze i najgorsze. Wektor $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, ..., \bar{x}_{10})$ jest określony przez średnie arytmetyczne cech, natomiast wektor $\mathbf{S} = (S_1, S_2, ..., S_{10})$ przez odchylenia standardowe cech, odpowiednio. Wektor $\mathbf{V}$ jest wektorem współczynników zmienności cech obliczonych według wzoru: $V = \bar{x}_i / S$, i=1,2,...,10. Zauważmy, że po ilorazowej normalizacji zmiennych, wektory określające wartości danej cechy dla województw mają charakter struktur tj.:

$$\sum_{i=1}^{16} x_{ij} = 1 \quad \text{dla} \quad j = 1, 2, ..., 10.$$

Przy użyciu podanych wcześniej funkcji obliczono syntetyczne mierniki województw, co następnie umożliwiło wyznaczenie rankingów województw, tj. uporządkować obiekty według wartości mierników syntetycznych. W tabeli 2 podane zostały otrzymane wartości mierników syntetycznych wskazujących na poziom rolnictwa polskich województw w roku 2009. Tabela 2 zawiera również wynikające z nich uporządkowania. Wykorzystane w tabeli mierniki m_3, m_4, m_5, m_6 są określone za pomocą wzorów (4) - (7).

Tabela 1. Dane wejściowe

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
x_1	0,063	0,069	0,074	0,021	0,032	0,052	0,055	0,059	0,065	0,073
x_2	0,069	0,066	0,068	0,087	0,060	0,088	0,057	0,047	0,083	0,059
x_3	0,074	0,069	0,054	0,045	**0,179**	0,073	0,054	0,060	0,044	0,047
x_4	**0,043**	0,058	0,066	0,024	0,016	0,025	0,057	0,051	0,080	0,059
x_5	**0,075**	0,057	0,056	0,085	0,087	0,069	0,062	0,069	0,043	0,064
x_6	0,064	0,064	0,057	0,056	0,051	0,026	0,056	0,079	0,016	0,059
x_7	0,072	0,056	0,050	0,089	0,132	0,138	0,067	0,061	0,049	**0,108**
x_8	0,067	**0,075**	**0,087**	0,042	0,005	0,039	0,079	0,083	0,071	0,058
x_9	0,057	0,065	0,056	0,039	0,056	0,018	0,048	0,047	0,019	0,047
x_{10}	0,063	**0,051**	0,048	**0,154**	0,028	0,072	0,066	0,074	0,078	0,050
x_{11}	0,053	0,061	0,064	0,046	0,090	0,049	0,058	0,058	0,095	0,065
x_{12}	0,055	0,060	0,062	0,054	0,034	0,028	0,087	**0,101**	0,020	0,074
x_{13}	0,068	0,064	0,051	0,058	0,074	0,022	0,052	0,051	0,029	0,055
x_{14}	0,058	0,061	0,063	0,087	0,066	0,069	0,055	0,037	0,112	0,051
x_{15}	0,068	0,060	0,071	0,092	0,067	**0,182**	**0,095**	0,083	0,071	0,071
x_{16}	0,052	0,063	0,073	0,020	0,021	0,051	0,051	0,042	**0,125**	0,062
x_{17}	0,075	0,075	0,087	0,154	0,179	0,182	0,095	0,101	0,125	0,108
x_0	0,043	0,051	0,048	0,020	0,005	0,018	0,048	0,037	0,016	0,047
$\bar{x}_i$	0,063	0,063	0,063	0,063	0,063	0,063	0,063	0,063	0,063	0,063
S	0,009	0,006	0,010	0,034	0,044	0,043	0,013	0,017	0,032	0,014
V	14%	9%	16%	54%	70%	69%	21%	27%	51%	23%

Źródło: opracowanie własne

Tabela 2. Wartości syntetycznych mierników i uporządkowanie województw w 2009 roku

Funkcje / Województwo	Wartości mierników				Uporządkowanie			
	m_3	m_4	m_5	m_6	m_3	m_4	m_5	m_6
Dolnośląskie	0,23	0,23	0,25	0,27	13	12	13	11
Kujawsko-pomorskie	0,41	0,41	0,42	0,41	5	5	6	5
Lubelskie	0,49	0,48	0,49	0,43	3	3	3	3
Lubuskie	0,17	0,16	0,20	0,17	15	15	15	15
Łódzkie	0,40	0,40	0,40	0,39	7	7	7	6
Małopolskie	0,21	0,21	0,23	0,23	14	14	14	13
Mazowieckie	0,61	0,61	0,60	0,58	1	1	1	2
Opolskie	0,24	0,23	0,27	0,32	10	11	10	9
Podkarpackie	0,16	0,15	0,18	0,14	16	16	16	16
Podlaskie	0,43	0,42	0,44	0,42	4	4	4	4
Pomorskie	0,37	0,37	0,38	0,36	8	8	8	8
Śląskie	0,24	0,23	0,26	0,28	12	13	11	10
Świętokrzyskie	0,24	0,24	0,26	0,23	11	10	12	14
Warmińsko-mazurskie	0,41	0,41	0,42	0,38	6	6	5	7
Wielkopolskie	0,60	0,60	0,59	0,62	2	2	2	1
Zachodnio-pomorskie	0,28	0,26	0,31	0,27	9	9	9	12

Źródło: obliczenia własne

Wykorzystując wcześniejsze obliczenia dokonano następnie podziału województw na 4 grupy, charakteryzujące się zbliżonym poziomem rozwoju rolnictwa. Podstawą podziału województw na klasy w danym roku, były wartości ich syntetycznych mierników. Wykorzystano do tego szeregi przedziałowe rozdzielcze, w których rozpiętości przedziałów klasowych były równe w przybliżeniu jednej czwartej rozstępu. W tabeli 3 przedstawiono klasyfikację województw sporządzoną na podstawie wyznaczonych średnich geometrycznych obliczonych za pomocą wzoru (10), ze wszystkich syntetycznych mierników dla poszczególnych województw.

Tabela 3. Podział województw na grupy

Klasa	Pozycja	Województwo	m(x)
I	1	Wielkopolskie	0,605
	2	Mazowieckie	0,600
II	3	Lubelskie	0,474
	4	Podlaskie	0,426
	5	Kujawsko-pomorskie	0,413
	6	Warmińsko-mazurskie	0,405
	7	Łódzkie	0,398
III	8	Pomorskie	0,368
	9	Zachodnio-pomorskie	0,278
IV	10	Opolskie	0,267
	11	Śląskie	0,253
	12	Dolnośląskie	0,244
	13	Świętokrzyskie	0,240
	14	Małopolskie	0,222
	15	Lubuskie	0,177
	16	Podkarpackie	0,157

Źródło: obliczenia własne

Wartości syntetycznych mierników dla poszczególnych województw świadczą o tym, że w roku 2009 regiony wykazywały silne zróżnicowanie ze względu na poziom rolnictwa. Z powyższej tabeli wynika, że wartość miernika syntetycznego województwa wielkopolskiego jest blisko czterokrotnie większa niż województwa podkarpackiego. Obserwując podział województw na cztery klasy, na podstawie średniej geometrycznej badanych mierników widać, że grupę pierwszą, charakteryzującą się najwyższym poziomem rolnictwa stanowią dwa województwa – wielkopolskie i mazowieckie. W ich przypadku wartości wszystkich mierników oscylowały wokół liczby 0,6 (na co wskazuje też średnia z mierników). Jest to wartość znacznie wyższa od wartości obliczonych dla pozostałych województw – najwyższa z nich nie przekracza 0,5. Można na tej podstawie wnioskować o wysokim poziomie rozwoju rolnictwa (przy tak wybranym sposobie mierzenia) w tych dwóch województwach w stosunku do pozostałych regionów. Na pozycję województwa wielkopolskiego miał wpływ przede wszystkim wysoki, bo najwyższy w kraju poziom dwóch cech: wartość skupu produktów rolnych ogółem (zł/ha UR) oraz nakłady inwestycyjne w rolnictwie (zł/ha UR). Natomiast na wysoką pozycję województwa mazowieckiego wpłynęła niewątpliwie najwyższa w Polsce wartość PKB (zł/mieszkańca) oraz dość wysoki poziom takich cech jak skup owoców z drzew (w kg na ha upraw) czy wartość skupu produktów rolnych ogółem (zł/ha UR).

Do drugiej klasy należy pięć województw: lubelskie, podlaskie, kujawsko-pomorskie, warmińsko-mazurskie i łódzkie. Można powiedzieć, że wykazują dość wysoki na tle kraju poziom rolnictwa. Wartości obliczonych dla nich mierników były powyżej średniej krajowej, tj. 0,35.

Grupa trzecia obejmuje dwa województwa o niskim, poniżej przeciętnego poziomie rolnictwa. Było to województwo pomorskie i zachodnio – pomorskie.

Najliczniejsza była ostatnia tj. czwarta grupa, co niestety wskazuje na to, że w 2009 roku prawie połowa polskich województw miała niski poziom rozwoju rolnictwa. Do tej grupy należało siedem następujących województw: opolskie, śląskie, dolnośląskie, świętokrzyskie, małopolskie, lubuskie i podkarpackie.

WNIOSKI

Na podstawie przeprowadzonych badań można wnioskować, że dla przyjętych cech, wybór metody wzorcowej nie miał większego wpływu zarówno na uporządkowanie liniowe, jak i na grupowanie województw. Wyniki otrzymane przy tym samym sposobie normalizacji zmiennych nie różnią się znacząco między sobą.

Opierając się na otrzymanych rezultatach można twierdzić, że najwyższym poziomem rolnictwa w 2009 roku charakteryzowało się województwo wielkopolskie. Wraz z województwem mazowieckim wyraźnie odbiegało swoim poziomem od pozostałych województw. Najniższym poziomem rolnictwa charakteryzowały się województwa podkarpackie i lubuskie. One również wyraźnie odbiegały swoim poziomem od pozostałych województw. Zastanawia niska pozycja województwa opolskiego.

Otrzymane wyniki pokazują duże zróżnicowanie polskiego rolnictwa w 2009 roku. Dla potwierdzenia tego zjawiska wskazana jest dalsza analiza występujących zależności – tym bardziej, że wyniki autorki [Binderman 2010] dotyczące polskiego rolnictwa w 2008 roku różnią się od prezentowanych w artykule. Wspomniane wyniki dla roku 2008 otrzymane zostały przy pomocy innego klasyfikatora i innej metody grupowania.

BIBLIOGRAFIA

Bartosiewicz S. (1976) Propozycja metody tworzenia zmiennych syntetycznych, Prace Naukowe AE we Wrocławiu, nr 84, Wrocław.

Binderman A. (2005) Klasyfikacja polskich województw według poziomu rozwoju rolnictwa, RNR, Seria G., T.92, Z.1, str. 42, wyd. „Wieś Jutra”, Warszawa.

Binderman A. (2006) Klasyfikacja danych na podstawie dwóch wzorców, Ekonomika i Organizacja Gospodarki Żywnościowej, SGGW Warszawa, z. 60, s. 25-34.

Binderman A. (2007) Wielowymiarowa analiza regionalnego zróżnicowania rolnictwa w Polsce, praca doktorska, SGGW, Warszawa.

Binderman A. (2008) Zastosowanie liniowej i nieliniowej funkcji użyteczności do badania poziomu rolnictwa w Polsce, MIBE IX, wyd. SGGW, 29-38.

Binderman A. (2009) Dynamika rozwoju rolnictwa w Polsce po akcesji do Unii Europejskiej, Roczniki Nauk Rolniczych, SERiA, Tom XI, Zeszyt 3.

Binderman A. (2010) Wpływ sposobu normalizacji zmiennych na ocenę regionalnego zróżnicowania rolnictwa, MIBE XI, wyd. SGGW.

Binderman A. (2010a) Porównanie poziomu rozwoju rolnictwa województw w Polsce, Roczniki Naukowe SERiA, T. XII, z. 2, s. 29-34.

Binderman Z. (2010) Zjawisko niedosytu w polu preferencji, indukowane przez miernik dwuwzorcowy, MIBE XI, wyd. SGGW.

Borkowski B, Szczesny W. (2002) Metody taksonomiczne w badaniach przestrzennego zróżnicowania rolnictwa, RNR, Seria G., T.89, Z.2, wyd. „Wieś Jutra”, Warszawa.

Breiman L. (1994) Bagging predictors, Technical Report 420, Departament of Statistics, University of California, CA, USA, September.

Breiman L. (1996) Arcing Classifiers, Technical Report 460, Departament of Statistics, University of California, CA, USA, February.

Breiman L. (1998) Bias-variance, regularization, instability and stabilization. C. M. Bishop, redaktor, Neural Networks and Machine Learning, Springer-Verlag, s. 27–56.

Chen S.J., Hwang C.L. (1992) Fuzzy Multiple Attribute Decision Making: Methods and Applications, Springer, New York.

Cieślak M. (1993) Ekonomiczne zastosowanie mierników syntetycznych ze zmiennym wzorcem, [w:] Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych, AE, Kraków.

Gatnar E., Walesiak M. (2009) Statystyczna analiza danych z wykorzystaniem programu R, PWN, Warszawa.

Gatnar E., Wywiał J. (2007) Wykorzystanie metod grupowania danych do wspomagania prac nad podziałem administracyjnym, Taksonomia 5, AE, Wrocław.

Hellwig Z. 1968; Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę kwalifikowanych kadr, „Przegląd Statystyczny”, z. 4.

Hunek E. (ed.) (2002) Rolnicza Polska wobec wyzwań współczesności, Wieś Jutra.

Hwang. C.L., Yoon K. (1981) Multiple Attribute Decision Making: Methods and Applications, Springer-Verlag, New York.

Jackson D. M. (1970) The stability of classifications of binary attribute data, Technical Report 70-65, Cornell University 1-13.

Krasowicz S. (2009) Regionalne zróżnicowanie zmian w rolnictwie polskim, Studia i Raporty IUNG-PIB, Wybrane elementy regionalnego zróżnicowania rolnictwa w Polsce, z.15, s.9-36.

Kukuła K. (2000) Metoda unitaryzacji zerowanej, PWN, Warszawa.

Malarska A. (2006) Materiały od wykładu: Klasyfikacja obiektów wielocechowych, KSEiS UŁ.

Malina A. (2004) Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw, AE, Seria Monografie nr 162, Kraków.

Młodak A. (2006) Analiza taksonomiczna w statystyce regionalnej, DIFIN Warszawa.

Neuman F. (1998): Data Fusion and data Quality, Procedings of the New techniques and Technologies for statistics seminar (NTTS"98).

Rocznik Statystyczny Województw 2010 r., GUS Warszawa.

Saaty T.L. (1980) The analytic hierarchy process, Mc-Graw Hill, New York.

Saaty T.L. (1995) Decision Making for Leaders, RWS Publications, New York.

Strahl D. (1990) Metody programowania rozwoju społeczno-gospodarczego, PWE, Warszawa.

Walesiak M. (2005) Uogólniona miara odległości w statystycznej analizie danych, wyd. AE, Wrocław.

Wilkin J. (2001) Polska Wieś 2000: Raport o stanie polskiej wsi, FAPA, Warszawa.

Zegar J. (2002) Zróżnicowanie regionalne rolnictwa, GUS, Warszawa .

Zeliaś A. (2000) Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym, AE Kraków.

Zhang W. (2004) Handover Decision Using Fuzzy MADM in Heterogeneous Networks, in Proc. IEEE WCNC'04 , Atlanta, GA, March.

ON MULTI-CRITERIA DECISION METHODS FOR STUDY OF A LEVEL OF DIFFERENTIATION OF POLISH AGRICULTURE IN 2009

Abstract: The paper presents the use of three standard methods for the analysis of regional differences in the level agriculture in 2009, after applying the quotient normalization of variables. The research compared with results obtained by using the measure of development of Pluta type - without model -. Of the four measures constructed the fifth, "resultant" synthetic indicator. Then was made the linear arrangement and division into group of Polish voivodships. The results showed that the level of development of Polish agriculture between voivodships in 2009 was very different. Classification of voivodships does not depend on the method and differentiation to a small extent.

Key words: cluster analysis, level of agricultural development, pattern measure, TOPSIS