

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

Nr 385

Taksonomia 25

**Klasyfikacja i analiza danych –
teoria i zastosowania**

Redaktorzy naukowci

Krzysztof Jajuga

Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2015

Redaktor Wydawnictwa: Aleksandra Śliwka

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Beata Mazur

Projekt okładki: Beata Dębska

Tytuł dofinansowany ze środków Narodowego Banku Polskiego
oraz ze środków Sekcji Klasyfikacji i Analizy Danych PTS

Informacje o naborze artykułów i zasadach recenzowania
znajdują się na stronie internetowej Wydawnictwa
www.pracnaukowe.ue.wroc.pl
www.wydawnictwo.ue.wroc.pl

Publikacja udostępniona na licencji Creative Commons
Uznanie autorstwa-Użycie niekomercyjne-Bez utworów zależnych 3.0 Polska
(CC BY-NC-ND 3.0 PL)



© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2015

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)
e-ISSN 2392-0041 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)
ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Zamówienia na opublikowane prace należy składać na adres:
Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
tel./fax 71 36 80 602; e-mail:econbook@ue.wroc.pl
www.ksiegarnia.ue.wroc.pl

Druk i oprawa: TOTEM

Spis treści

Wstęp.....	9
Tomasz Bartłomowicz: Segmentacja konsumentów na podstawie preferencji wyrażonych uzyskanych metodą Maximum Difference Scaling	11
Barbara Batóg, Jacek Batóg, Andrzej Niemiec, Wanda Skoczylas, Piotr Waśniewski: Zastosowanie metod klasyfikacyjnych w identyfikacji kluczowych indyktorów osiągnięć w zarządzaniu wynikami przedsiębiorstw	20
Iwona Bąk: Wykorzystanie statystycznej analizy danych w badaniach turystyki transgranicznej na obszarach chronionych.....	28
Beata Bieszk-Stolorz: Ocena stopnia deprecjacji kapitału ludzkiego z wykorzystaniem nieliniowych modeli regresji.....	37
Mariola Chrzanowska, Nina Drejerska: Małe i średnie przedsiębiorstwa w strefie podmiejskiej Warszawy – określenie znaczenia lokalizacji z wykorzystaniem drzew klasyfikacyjnych.....	45
Adam Depta: Próba modelowania strukturalnego jakości życia osób jękaających się jako konstrukt ukrytego na podstawie kwestionariusza SF-36v2	53
Katarzyna Dębkowska: Wielowymiarowa analiza kondycji finansowej przedsiębiorstw sektora e-usług	63
Krzysztof Dmytrów, Mariusz Doszyń: Taksonomiczna procedura wspomagania kompletacji produktów w magazynie	71
Mariusz Doszyń, Sebastian Gnat: Propozycja procedury taksonomiczno-ekonometrycznej w indywidualnej wycenie nieruchomości.....	81
Marta Dziechciarz-Duda, Anna Król: Zastosowanie analizy <i>unfolding</i> i regresji hedonicznej do oceny preferencji konsumentów	90
Katarzyna Frodyma: Współzależność między poziomem rozwoju gospodarczego a udziałem energii ze źródeł odnawialnych w końcowym zużyciu w krajach Unii Europejskiej.....	99
Hanna Gruchociak: Porównanie struktury lokalnych rynków pracy wyznaczonych przy wykorzystaniu różnych metod w Polsce w latach 2006 i 2011 .	111
Alicja Grześkowiak, Agnieszka Stanimir: Postrzeganie środowiska pracy przez starszą i młodszą generację pracowników	120
Marta Hozer-Koćmiel, Christian Lis: Klasyfikacja krajów nadbałtyckich ze względu na czas prac wykonywanych w gospodarstwie domowym	129
Tadeusz Kufel, Magdalena Osińska, Marcin Błażejowski, Paweł Kufel: Zegar cyklu koniunkturalnego państw UE i USA w latach 1995-2013 w świetle badań synchronizacji.....	138
Aleksandra Łuczak: Wykorzystanie rozszerzonej interwałowej metody TOPSIS do porządkowania liniowego obiektów	147

Aleksandra Łuczak, Feliks Wysocki: Zintegrowane podejście do ustalania współczynników wagowych dla cech w zagadnieniach porządkowania linowego obiektów	156
Małgorzata Markowska, Danuta Strahl: Wykorzystanie klasyfikacji dynamicznej do identyfikacji wrażliwości na kryzys ekonomiczny unijnych regionów szczebla NUTS 2.....	166
Aleksandra Matuszewska-Janica, Marta Hozer-Koćmiel: Struktura zatrudnienia oraz wynagrodzenia kobiet i mężczyzn a przedmiotowa struktura gospodarcza w państwach UE.....	178
Anna M. Olszewska: Zastosowanie analizy korespondencji do badania związku pomiędzy zarządzaniem jakością a innowacyjnością przedsiębiorstw	187
Małgorzata Podogrodzka: Metoda aglomeracyjna w ocenie przestrzennego zróżnicowania starości demograficznej w Polsce	195
Ewa Roszkowska, Tomasz Wachowicz: Ocena ofert negocjacyjnych spoza dopuszczalnej przestrzeni negocjacyjnej.....	201
Ewa Roszkowska, Tomasz Wachowicz: Zastosowanie metody <i>unfolding</i> do wspomagania procesu negocjacji	210
Małgorzata Rószkiewicz: Próba diagnozy uwarunkowań poziomu wskaźnika braku odpowiedzi w środowisku polskich gospodarstw domowych.....	219
Marcin Salamaga: Próba identyfikacji muzycznych profili melomanów z wykorzystaniem drzew klasyfikacyjnych i regresyjnych	229
Agnieszka Sompolska-Rzechuła: Określenie czynników wpływających na prawdopodobieństwo poprawy poziomu rozwoju społecznego z wykorzystaniem modelu logitowego	239
Iwona Staniec: Wykorzystanie analizy czynnikowej w identyfikacji konstruktywów ukrytych determinujących ryzyko współpracy.....	248
Agnieszka Stanimir: Skłonność do zagranicznej mobilności młodszych i starszych osób	257
Mirosława Sztemberg-Lewandowska: Problemy decyzyjne w funkcjonalnej analizie głównych składowych.....	267
Tomasz Szubert: Demograficzno-społeczne determinanty określające subiektywny status jednostki w polskim społeczeństwie	276
Piotr Tarka: Własności 5- i 7-stopniowej skali Likerta w kontekście normalizacji zmiennych metodą Kaufmana i Rousseeuwa	286
Joanna Trzęsiok: Nielklasyczne metody regresji a problem odporności	296
Katarzyna Wawrzyniak: Ocena podobieństwa wyników uporządkowania województw uzyskanych różnymi metodami porządkowania	305
Katarzyna Wójcik, Janusz Tuchowski: Wykorzystanie metody opartej na wzorcach w automatycznej analizie opinii konsumenckich.....	314
Anna Zamojska: Zastosowanie analizy falkowej w ocenie efektywności funduszy inwestycyjnych	325

Summaries

Tomasz Bartłomowicz: Segmentation of consumers based on revealed preferences obtained with the Maximum Difference Scaling method	19
Barbara Batóg, Jacek Batóg, Andrzej Niemiec, Wanda Skoczylas, Piotr Waśniewski: Application of classification methods to identify the key performance indicators of performance management	27
Iwona Bąk: The application of statistical data analysis in the studies of cross-border tourism in protected areas.....	36
Beata Bieszk-Stolorz: Evaluating human capital depreciation by means of non-linear regression models.....	44
Mariola Chrzanowska, Nina Drejerska: Small and medium enterprises in the Warsaw suburban zone – determination of a localization’s role using classification trees	52
Adam Depta: An attempt of structural modelling of the quality of life of stuttering people as a latent construct, based on SF-36v2 questionnaire ...	62
Katarzyna Dębowska: Multidimensional analysis of financial condition of e-business services	70
Krzysztof Dmytrów, Mariusz Doszyń: Taxonomic procedure of supporting order-picking of products in a warehouse	80
Mariusz Doszyń, Sebastian Gnat: Taxonomic and econometric methods in individual real estate evaluation.....	89
Marta Dziechciarz-Duda, Anna Król: The application of unfolding analysis and hedonic regression in the investigation of consumers’ preferences	98
Katarzyna Frodyma: Interdependence between the level of economic development and the share of renewable energy in gross final energy consumption in the European Union.....	110
Hanna Gruchociak: Comparison of local labour markets structure designated using different methods in Poland in 2006 and 2011 years.....	119
Alicja Grześkowiak, Agnieszka Stanimir: Perception of working environment by older and younger generation of workers.....	128
Marta Hozer-Koćmiel, Christian Lis: Classification of the Baltic Sea Region countries due to the time of household work.....	137
Tadeusz Kufel, Magdalena Osińska, Marcin Błażejowski, Paweł Kufel: Business cycle clock for the EU and the USA in 1995-2013 in the light of synchronization research.....	146
Aleksandra Łuczak: The use of the extended interval TOPSIS methods for linear ordering of objects.....	155
Aleksandra Łuczak, Feliks Wysocki: Integrated approach for determining the weighting coefficients for features in issues of linear ordering of objects.....	165

Małgorzata Markowska, Danuta Strahl: The application of dynamic classification for the identification of vulnerability to economic crisis in the EU NUTS 2 regions	177
Aleksandra Matuszewska-Janica, Marta Hozer-Koćmiel: The structure of male and female employment and remuneration vs. the basic economy structure in the EU countries	186
Anna M. Olszewska: The application of the correspondence analysis for the study of the relations between quality management and innovation in the enterprises.....	194
Małgorzata Podogrodzka: Agglomeration method in the age and ageing in Poland by voivodships.....	200
Ewa Roszkowska, Tomasz Wachowicz: Scoring the negotiation offers from the outside of the feasible negotiation space	209
Ewa Roszkowska, Tomasz Wachowicz: Application of the unfolding analysis to negotiation support.....	218
Małgorzata Rószkiewicz: An attempt to diagnose the determinants of non-response rate in Polish households surveys	228
Marcin Salamaga: Attempt to identify music lovers profiles using classification and regression trees	238
Agnieszka Sompolska-Rzechuła: The definition of factors influencing the probability of improving the level of human development using the logit model.....	247
Iwona Staniec: The use of factor analysis to identify hidden constructs – determinants of the cooperation risk	256
Agnieszka Stanimir: Willingness to mobility abroad among younger and older persons	266
Mirosława Sztemberg-Lewandowska: Decision problems in functional principal components analysis.....	275
Tomasz Szubert: Socio-demographic factors determining subjective social status of an individual in Polish society	285
Piotr Tarka: Normalization methods of variables and measurement on 5 and 7 point Likert scale	295
Joanna Trzęsiok: Non-classical regression methods vs. robustness	304
Katarzyna Wawrzyniak: The evaluation of the similarity of the voivodships' orderings obtained by means of different methods.....	313
Katarzyna Wójcik, Janusz Tuchowski: Using pattern-based opinion mining.....	324
Anna Zamojska: Mutual funds performance measurement – wavelets analysis approach.....	333

Aleksandra Łuczak, Feliks Wysocki

Uniwersytet Przyrodniczy w Poznaniu

e-mails: luczak@up.poznan.pl; wysocki@up.poznan.pl

ZINTEGROWANE PODEJŚCIE DO USTALANIA WSPÓLCZYNNIKÓW WAGOWYCH DLA CECH W ZAGADNIENIACH PORZĄDKOWANIA LINIOWEGO OBIEKTÓW

Streszczenie: W pracy podano propozycję zintegrowanego podejścia do ustalania współczynników wagowych dla cech i ich wykorzystania w procedurze tworzenia miernika syntetycznego. Zaproponowane podejście zintegrowane polega na wykorzystaniu zarówno procedur statystycznych (metoda CCSD – CC – *Correlation Coefficient*, SD – *Standard Deviation*), jak i opinii ekspertów oraz metod analitycznego procesu hierarchicznego w wersji klasycznej (AHP) lub rozmytej (FAHP). Proponowane podejście zostało zastosowane do badania poziomu rozwoju społeczno-gospodarczego województw w Polsce przy wykorzystaniu metody TOPSIS. Badania wykazały, że zastosowanie różnych systemów wag i sposobów ich włączania do procesu konstrukcji miernika syntetycznego wpływa tylko w niewielkim stopniu na zmiany w uporządkowaniach liniowych województw. W związku z tym w procesie konstrukcji mierników syntetycznych największą wagę należy przykładać do właściwego w sensie merytorycznym doboru cech.

Słowa kluczowe: porządkowanie liniowe obiektów, współczynniki wagowe cech, metoda CCSD, metoda FAHP, metoda TOPSIS.

DOI: 10.15611/pn.2015.385.17

1. Wstęp

W literaturze dotyczącej ustalania współczynników wagowych cech wskazuje się na trzy sposoby tworzenia wag. „Wagi ustala się albo metodą ekspertów (metoda *a priori*), albo z użyciem algorytmów obliczeniowych opierających się na informacjach zawartych w danych pierwotnych (surowych). Można też wykorzystać metodę opartą na obu tych ujęciach” [Walesiak 2011, s. 47]. Jednak poglądy dotyczące systemu współczynników wagowych dla cech nie są w pełni wykrystalizowane, a najczęściej w badaniach empirycznych zakłada się jednakowe wagi dla wszystkich cech (zob. np. [Grabiński 1988; Walesiak 2011]).

W pracy przedstawiono propozycję zintegrowanego podejścia do tworzenia systemu wag dla cech polegającą na wykorzystaniu zarówno procedur statystycznych, jak i opinii ekspertów oraz metod analitycznego procesu hierarchicznego w wersji klasycznej lub rozmytej. W proponowanym podejściu zastosowano procedurę statystyczną opartą na metodzie CCSD (CC – *Correlation Coefficient*, SD – *Standard Deviation*) [Wang, Luo 2010; Łuczak, Wysocki 2014]. Metoda ta uwzględnia stopień skorelowania wartości danej cechy z wartościami miernika syntetycznego obliczonymi z wyłączeniem tej cechy oraz jej odchylenie standardowe. Natomiast w podejściu merytorycznym wykorzystano metodę rozmytego analitycznego procesu hierarchicznego, która pozwala agregować oceny ważności cech dokonane przez ekspertów [Chang 1996]. Proponowane zintegrowane podejście zostało zastosowane do oceny poziomu społeczno-gospodarczego województw w Polsce w 2011 roku. Do konstrukcji syntetycznego miernika poziomu rozwoju zastosowano metodę TOPSIS.

2. Metodyka badań

Ustalenie systemu wag dla cech prostych wykorzystywanych w procedurze tworzenia miernika syntetycznego może być przeprowadzone na podstawie zintegrowanego podejścia łączącego analizę statystyczną i merytoryczną według wzoru:

$$w_{k(z)} = \alpha \cdot w_{k(s)} + (1 - \alpha) \cdot w_{k(m)},$$

gdzie: $w_{k(z)}$ – waga zintegrowana k -tej cechy ($k = 1, \dots, K$), $w_{k(s)}$ – waga k -tej cechy uzyskana na podstawie analizy statystycznej, $w_{k(m)}$ – waga k -tej cechy uzyskana na podstawie analizy merytorycznej, α – parametr określający udział wagi statystycznej w tworzeniu wagi zintegrowanej, $\alpha \in \langle 0,1 \rangle$, przy czym zintegrowane współczynniki wagowe spełniają warunki $w_{k(z)} \in \langle 0,1 \rangle$, $\sum_{k=1}^K w_{k(z)} = 1$.

W podejściu statystycznym wykorzystuje się procedury statystyczne, bazując na informacjach o cechach tkwiących tylko w samej macierzy danych, a w szczególności analizę zmienności cech i korelacji między cechami albo tylko jedną z tych analiz [Wysocki 2010]. Natomiast w podejściu opartym na opiniach ekspertów można wykorzystać np. analityczny proces hierarchiczny w wersji klasycznej lub rozmytej.

Do ustalenia systemu wag dla cech metodami statystycznymi wykorzystano metodę CCSD według formuły [Wang, Luo 2010; Łuczak, Wysocki 2014]:

$$w_{k(s)} = s_k \sqrt{1 - R_k} / \sum_{k=1}^K s_k \sqrt{1 - R_k}, \quad (k = 1, \dots, K),$$

gdzie: s_k jest odchyleniem standardowym z wartości znormalizowanych k -tej cechy, R_k ($-1 \leq R_k \leq 1$) jest współczynnikiem korelacji pomiędzy k -tą cechą a wartościami skorygowanego miernika utworzonego z pozostałych cech, ($j = 1, \dots, K, j \neq k$). Otrzymane współczynniki wagowe spełniają warunki

$$w_{k(s)} \in \langle 0, 1 \rangle, \sum_{k=1}^K w_{k(s)} = 1.$$

Zakłada się, że jeśli R_k ma wartość bliską jeden, to usunięcie k -tej cechy ze zbioru wyjściowego miałoby niewielki wpływ na wartości miernika syntetycznego, co oznacza, że k -tej cesze może być przypisana odpowiednio mała waga. Jeśli R_k ma małą wartość, w skrajnym przypadku bliską -1 , to dodanie k -tej cechy do zbioru cech tworzących miernik będzie miało wpływ na jego wartości. Wtedy k -ta cecha powinna uzyskać relatywnie większą wagę [Wang, Luo 2010; Łuczak, Wysocki 2014].

W podejściu merytorycznym wykorzystano metodę rozmytą analitycznego procesu hierarchicznego (FAHP) [Chang 1996; Łuczak, Wysocki 2011], która pozwala zagregować oceny ważności cech dokonane przez ekspertów, a jednocześnie umożliwia eliminację cech o najmniejszym znaczeniu merytorycznym. W tym przypadku ustala się wagi cech na podstawie rozmytych opinii ekspertów, które są bardziej realistyczne aniżeli opinie dokładne [Łuczak, Wysocki 2011]. Istotą postępowania w FAHP jest [Łuczak, Wysocki 2011]:

1. Utworzenie struktury hierarchicznej wielokryterialnego problemu oceny obiektów poprzez określenie jej elementów składowych ze sobą powiązanych: główne kryterium oceny, kryteria podrzędne, cechy proste oraz oceniane obiekty.

2. Porównanie ważności cech parami między sobą w odniesieniu do kryteriów podrzędnych oraz kryteriów podrzędnych w odniesieniu do kryterium głównego przy wykorzystaniu dziewięciostopniowej rozmytej skali preferencji według Saaty'ego [Wang, Cheng, Kun-Cheng 2009].

3. Utworzenie macierzy porównań parami i obliczenie na tej podstawie tzw. priorytetów globalnych, które reprezentują ważności rozpatrywanych cech w odniesieniu do kryterium głównego. Priorytety globalne przyjmuje się jako współ-

czynniki wagowe cech. Wagi te spełniają warunki: $w_{k(m)} \in \langle 0, 1 \rangle, \sum_{k=1}^K w_{k(m)} = 1$.

Kolejnym etapem w procesie konstrukcji miernika syntetycznego jest przeprowadzenie ważenia cech. Waznieniu mogą podlegać [Walesiak 2011]:

podejście I – wartości cech znormalizowanych (wagi liniowe):

$$z_{ik(z)} = w_{k(z)} \cdot z_{ik},$$

gdzie: z_{ik} ($i = 1, \dots, N; k = 1, \dots, K$) przedstawia znormalizowaną wartość k -tej cechy w i -tym obiekcie;

podejście II – cząstkowe odległości (np. euklidesowe) wyznaczone dla i -tego obiektu ze względu na k -tą cechę od obiektu modelowego z_k . Wówczas całkowita odległość obiektu i -tego od wzorca jest definiowana jako średnia ważona odległości cząstkowych (d_{ik}) (por. [Gordon 1999]):

$$d_i(\mathbf{w}) = \sqrt{\sum_{k=1}^K w_{k(z)} (z_{ik} - z_k)^2}.$$

Następnie ustalone zostają współrzędne obiektów modelowych – wzorca, np. według zasady [Wysocki 2010]:

$$A^+ = (\max_i(z_{i1(\bullet)}), \max_i(z_{i2(\bullet)}), \dots, \max_i(z_{iK(\bullet)})) = (z_1^+, z_2^+, \dots, z_K^+)$$

i antywzorca rozwoju:

$$A^- = (\min_i(z_{i1(\bullet)}), \min_i(z_{i2(\bullet)}), \dots, \min_i(z_{iK(\bullet)})) = (z_1^-, z_2^-, \dots, z_K^-),$$

gdzie (\bullet) oznacza ustalanie współrzędnych na podstawie ważonych wartości znormalizowanych (podejście I) albo wartości znormalizowanych cech (podejście II).

Po ustaleniu wzorca i antywzorca rozwoju oblicza się odległości euklidesowe każdego ocenianego obiektu od wzorca z^+ i antywzorca rozwoju z^- :

$$\text{podejście I: } d_i^+ = \sqrt{\sum_{k=1}^K (z_{ik(z)} - z_k^+)^2} \quad \text{oraz} \quad d_i^- = \sqrt{\sum_{k=1}^K (z_{ik(z)} - z_k^-)^2}$$

jest równoważne

$$d_i^+ = \sqrt{\sum_{k=1}^K w_{k(z)}^2 (z_{ik} - z_k^+)^2} \quad \text{oraz} \quad d_i^- = \sqrt{\sum_{k=1}^K w_{k(z)}^2 (z_{ik} - z_k^-)^2},$$

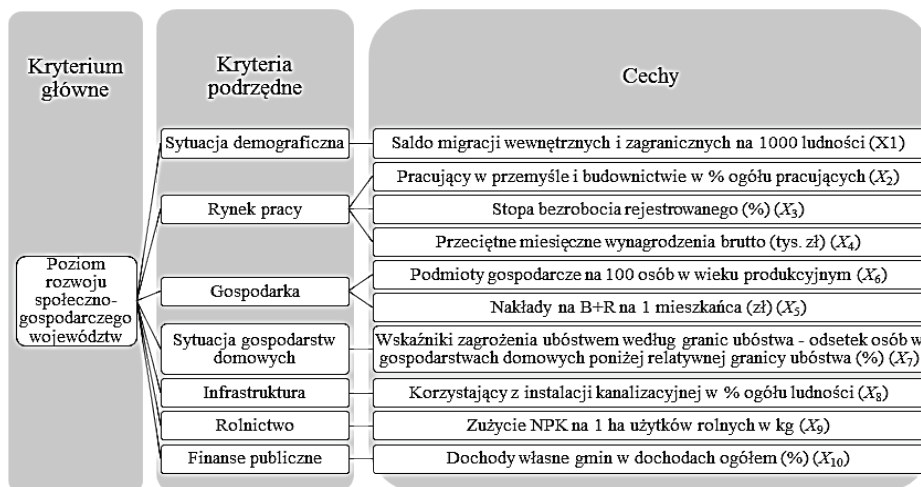
$$\text{podejście II: } d_i^+ = \sqrt{\sum_{k=1}^K w_{k(z)} (z_{ik} - z_k^+)^2} \quad \text{oraz} \quad d_i^- = \sqrt{\sum_{k=1}^K w_{k(z)} (z_{ik} - z_k^-)^2}.$$

Do konstrukcji miernika syntetycznego zastosowano metodę TOPSIS za pomocą miernika [Hwang, Yoon 1981; Wysocki 2010]: $S_i = d_i^- / (d_i^- + d_i^+)$, ($i = 1, 2, \dots, N$), przy czym $0 \leq S_i \leq 1$.

Im mniejsza jest odległość danego obiektu od obiektu modelowego – wzorca rozwoju, a tym samym większa od drugiego bieguna – antywzorca rozwoju, tym wartość miernika syntetycznego jest bliższa 1.

3. Wyniki badań

Podstawę badań stanowiły dane statystyczne z Głównego Urzędu Statystycznego [Bank Danych Lokalnych 2011]. W badaniach dokonano wyboru siedmiu kryteriów i 10 cech¹ opisujących poziom rozwoju społeczno-gospodarczego województw (rys. 1).



Rys. 1. Struktura hierarchiczna problemu oceny poziomu rozwoju społeczno-gospodarczego województw

Źródło: opracowanie własne.

Przyjęto, że dwie cechy mają charakter destymulant (X_3 , X_7), a pozostałe – stymulant. Cechy zostały znormalizowane za pomocą metody unitaryzacji zerowanej.

Porównując otrzymane wartości wag metodami CCSD, AHP i FAHP, można zauważyć istotne różnice między nimi (tab. 1). Współczynnik korelacji między systemami wag uzyskanymi metodami AHP i FAHP wskazuje na ich dużą zgodność (0,809). Natomiast współczynnik korelacji pomiędzy systemem wag uzyskanym metodami FAHP i CCSD wskazuje na średni ujemny związek (-0,602), a metodami AHP i CCSD – słaby ujemny (-0,278). W przypadku metody CCSD, poza trzema cechami (X_4 , X_6 , X_{10}), pozostałe miały zbliżone współczynniki wagowe. W sensie statystycznym największy współczynnik wagowy został przyporządkowany cesze *zużycie NPK na 1 ha użytków rolnych* (0,132). Taki wynik budzi jednak wątpliwości merytoryczne. Potwierdzają to wartości wag uzyskane metodami AHP i FAHP, które

¹ Ze względu na bardzo silne skorelowanie wartości cechy dotyczącej *PKB na 1 mieszkańca w zł* z wartościami pozostałych cech została ona usunięta ze zbioru cech opisujących poziom rozwoju społeczno-gospodarczego.

Tabela 1. Wartości współczynników wagowych dla cech uzyskane metodami CCSD, AHP i FAHP

Kryteria	Cechy	Wagi uzyskane metodami		
		CCSD	AHP	FAHP
Sytuacja demograficzna	Saldo migracji wewnętrznych i zagranicznych na 1000 ludności (X_1)	0,100	0,121	0,149
Rynek pracy	Pracujący w przemyśle i budownictwie w % ogółu pracujących (X_2)	0,130	0,080	0,073
	Stopa bezrobocia rejestrowanego (%) (X_3)	0,117	0,080	0,073
	Przeciętne miesięczne wynagrodzenia brutto (zł) (X_4)	0,084	0,080	0,073
Gospodarka	Nakłady na B+R na 1 mieszkańca (zł) (X_5)	0,103	0,168	0,137
	Podmioty gospodarcze na 100 osób w wieku produkcyjnym (X_6)	0,074	0,168	0,137
Sytuacja gospodarstw domowych	Wskaźniki zagrożenia ubóstwem według granic ubóstwa – odsetek osób w gospodarstwach domowych poniżej relatywnej granicy ubóstwa (%) (X_7)	0,108	0,157	0,180
Infrastruktura	Korzystający z instalacji kanalizacyjnej w % ogółu ludności (X_8)	0,119	0,047	0,015
Rolnictwo	Zużycie NPK na 1 ha użytków rolnych w kg (X_9)	0,132	0,021	0,000
Finanse publiczne	Dochody własne gmin w dochodach ogółem (%) (X_{10})	0,034	0,078	0,163
	Σ	1,000	1,000	1,000
Współczynniki korelacji między systemami wag	CCSD	1		
	AHP	-0,278	1	
	FAHP	-0,602	0,809	1

Źródło: obliczenia własne na podstawie danych statystycznych pochodzących z *Bank Danych Lokalnych* [2011], <http://www.stat.gov.pl/bdl>.

dla tej cechy wynoszą odpowiednio tylko 0,021 i 0, co w drugim przypadku prowadzi do jej eliminacji ze zbioru wyjściowego. Natomiast największe współczynniki wagowe uzyskane metodą AHP mają dwie cechy związane z gospodarką: *nakłady na B+R na 1 mieszkańca (zł)* (0,168) oraz *podmioty gospodarcze na 100 osób w wieku produkcyjnym* (0,168). Z kolei w metodzie FAHP największy współczynnik wagowy został przyporządkowany *wskaźnikowi zagrożenia ubóstwem* (0,180). Następnie obliczono wartości syntetycznego miernika rozwoju metodą TOPSIS (tab. 2).

Przedstawione wyniki obliczeń (tab. 2) wykazały, że:

1. Uzyskane klasyfikacje generalnie trafnie odzwierciedlają różnice w poziomie rozwoju społeczno-gospodarczego województw – najwyższym poziomem rozwoju charakteryzuje się województwo mazowieckie, a najniższym województwa Polski wschodniej (świętokrzyskie, lubelskie, podlaskie, warmińsko-mazurskie, podlaskie),

2. Zastosowanie wag merytorycznych doprowadziło do większego zróżnicowania wartości i zakresu zmienności mierników syntetycznych w podejściu zarówno I, jak i II do ważenia cech. Rozstęp między wartością maksymalną i minimalną miernika dla $\alpha = 0$ (wagi merytoryczne) wynosi 0,688 dla podejścia I, a dla podejścia II – 0,613, podczas gdy dla $\alpha = 1$ (wagi statystyczne) odpowiednio 0,407 i 0,438.

3. Wartości mierników syntetycznych uzyskanych w podejściu I wykazują przeważnie większe zróżnicowanie i większy obszar zmienności (maksymalny rozstęp dla $\alpha = 0$ wynosi 0,688, a minimalny dla $\alpha = 1 - 0,407$, natomiast w podejściu II odpowiednio 0,613 i 0,438. Różnice te wynikają prawdopodobnie z „umieszczenia” wag cech w procedurze konstrukcji miernika syntetycznego. Przypomnijmy, że w podejściu I ważeniu podlegają wartości znormalizowane cech za pomocą przekształcenia liniowego (wagi występują w pierwszej potęgze), natomiast w podejściu II ważone są cząstkowe odległości euklidesowe, wówczas wagi występują jako pierwiastki kwadratowe z ich wartości wyjściowych, co „osłabia” siłę ich wpływu na kształtowanie wartości mierników syntetycznych.

Tabela 2. Wartości syntetycznych mierników poziomu rozwoju społeczno-gospodarczego województw w zależności od podejścia do zagadnienia ważenia cech i tworzenia systemu wag zintegrowanych

Województwa ^{a)}	Podejście I					Podejście II					
	wagi merytoryczne		α			wagi merytoryczne		α			wagi statystyczne
	0	0,25	0,5	0,75	1	0	0,25	0,5	0,75	1	
Mazowieckie	0,835	0,787	0,728	0,668	0,612	0,786	0,743	0,707	0,675	0,644	
Dolnośląskie	0,630	0,619	0,613	0,613	0,621	0,616	0,616	0,615	0,615	0,615	
Śląskie	0,562	0,565	0,571	0,581	0,592	0,567	0,571	0,575	0,579	0,582	
Pomorskie	0,535	0,536	0,545	0,561	0,582	0,544	0,553	0,561	0,569	0,577	
Małopolskie	0,545	0,537	0,514	0,478	0,437	0,528	0,505	0,484	0,464	0,446	
Wielkopolskie	0,497	0,503	0,519	0,542	0,570	0,513	0,524	0,535	0,545	0,556	
Łódzkie	0,507	0,499	0,490	0,483	0,480	0,485	0,482	0,478	0,474	0,470	
Zachodniopomorskie	0,485	0,473	0,468	0,472	0,485	0,471	0,475	0,478	0,481	0,484	
Opolskie	0,481	0,482	0,493	0,513	0,541	0,465	0,478	0,491	0,505	0,518	
Lubuskie	0,470	0,466	0,461	0,457	0,454	0,452	0,450	0,448	0,446	0,444	
Kujawsko-pomorskie	0,359	0,365	0,386	0,419	0,458	0,353	0,375	0,396	0,415	0,434	
Podkarpackie	0,218	0,234	0,252	0,269	0,281	0,229	0,239	0,249	0,259	0,269	
Warmińsko-mazurskie	0,173	0,202	0,243	0,289	0,332	0,206	0,237	0,264	0,288	0,311	
Podlaskie	0,182	0,198	0,224	0,258	0,291	0,204	0,225	0,244	0,262	0,279	
Lubelskie	0,154	0,176	0,203	0,230	0,252	0,182	0,199	0,214	0,228	0,241	
Świętokrzyskie	0,147	0,163	0,182	0,199	0,214	0,173	0,182	0,190	0,198	0,206	
Rozstęp	0,688	0,624	0,546	0,469	0,407	0,613	0,561	0,517	0,477	0,438	
Współczynnik zmienności (%)	45,57	41,73	36,91	32,41	29,43	41,35	37,96	35,21	32,92	30,95	

^{a)} Uporządkowanie liniowe województw według wartości syntetycznego miernika rozwoju uzyskanego przy $\alpha = 0$, podejście II.

Źródło: obliczenia własne na podstawie danych statystycznych Głównego Urzędu Statystycznego z *Bank Danych Lokalnych* [2011], <http://www.stat.gov.pl/bdl>.

Tabela 3. Rankingi województw według różnych systemów wag zintegrowanych dla cech

Województwa	Rankingi województw według systemów wag zintegrowanych (α)									
	0		0,25		0,5		0,75		1	
	I ^{a)}	II	I	II	I	II	I	II	I	II
Dolnośląskie	2	2	2	2	2	2	2	2	1	2
Kujawsko-pomorskie	11	11	11	11	11	11	11	11	9	11
Lubelskie	15	15	15	15	15	15	15	15	15	15
Lubuskie	10	10	10	10	10	10	10	10	10	10
Łódzkie	6	7	7	7	8	8	7	8	8	8
Małopolskie	4	5	4	6	6	7	8	9	11	9
Mazowieckie	1	1	1	1	1	1	1	1	2	1
Opolskie	9	9	8	8	7	6	6	6	6	6
Podkarpackie	12	12	12	12	12	13	13	14	14	14
Podlaskie	13	14	14	14	14	14	14	13	13	13
Pomorskie	5	4	5	4	4	4	4	4	4	4
Śląskie	3	3	3	3	3	3	3	3	3	3
Świętokrzyskie	16	16	16	16	16	16	16	16	16	16
Warmińsko-mazurskie	14	13	13	13	13	12	12	12	12	12
Wielkopolskie	7	6	6	5	5	5	5	5	5	5
Zachodniopomorskie	8	8	9	9	9	9	9	7	7	7

^{a)} Podejście I – ważenie znormalizowanych wartości cech, podejście II – ważenie cząstkowych odległości euklidesowych wyznaczonych dla k -tej cechy.

Źródło: obliczenia własne na podstawie danych statystycznych z *Bank Danych Lokalnych* [2011].

Tabela 3 przedstawia zestawienie rang województw otrzymane w wyniku zastosowania różnych systemów wag zintegrowanych i dwóch podejść do ważenia cech. Można łatwo zauważyć, że występuje pełna lub prawie pełna zgodność uzyskanych rang w przypadku większości województw, niezależnie od zastosowanego systemu wag. Znaczne różnice w rankingach pojawiają się jedynie w przypadku województwa małopolskiego. Warto więc zwrócić uwagę na miejsca rankingowe uzyskane przez województwo małopolskie przy zastosowaniu różnych zestawów współczynników wagowych cech (tab. 3). W systemie zakładającym podejście I i wagi merytoryczne ($\alpha = 0$) województwo to uplasowało się na wysokim 4. miejscu, a przy $\alpha = 1$ (wagi ustalone na podstawie analizy statystycznej) dopiero na 11. miejscu. W przypadku zastosowania podejścia II różnice były nieco mniejsze, bowiem województwo to zajmowało odpowiednio pozycje 5. i 9. Jak wydaje się, przyczyna znacznie gorszej lokaty uzyskanej przez województwo małopolskie ze względu na poziom rozwoju społeczno-gospodarczego przy zastosowaniu wag ustalonych tylko metodami statystycznymi ($\alpha = 1$) leży w bardzo niskim poziomie nawożenia mineralnego (x_9) w tym województwie i największej wartości współczynnika wagowego przyporządkowanego tej cesze ($w_{9(s)}=0,132$). Zużycie NPK na 1 ha użytków rolnych w województwie małopolskim w 2011 roku wynosiło zaledwie 55,8 kg, podczas gdy w Polsce było ponad dwukrotnie (126,6 kg/ha UR),

a w województwie opolskim nawet ponad trzykrotnie (181,9 kg/ha UR) wyższe. Natomiast zastosowanie systemu wag merytorycznych za pomocą metody FAHP doprowadziło do wyeliminowania tej cechy ze zbioru wyjściowego cech opisujących poziom rozwoju społeczno-gospodarczego województw ($w_{g(m)}=0$) i w konsekwencji do znacznej poprawy pozycji rankingowej województwa małopolskiego. Z tych powodów przypisanie wysokiej wagi tej cesze spowodowało obniżenie rangi województwa małopolskiego.

4. Zakończenie

Przeprowadzone badania i analizy wykazały, że:

1. Oparcie rozwiązania problemu ustalania wag dla cech tylko na procedurach statystycznych może prowadzić do błędnych rankingów, przynajmniej w przypadku niektórych obiektów. Dlatego rozsądnym rozwiązaniem problemu ustalania systemu wag dla cech wydaje się łączenie obydwu analiz –statystycznej i merytorycznej opierającej się na opiniach ekspertów.

2. Wyznaczenie wag merytorycznych metodą FAHP ma jeszcze jedną zaletę, pozwala bowiem eliminować cechy o niewielkim znaczeniu merytorycznym ze zbioru tworzącego miernik syntetyczny. W badaniach poziomu rozwoju społeczno-gospodarczego cecha *zużycie NPK na 1 ha UR w kg (x_9)* jako mało ważna w sensie merytorycznym została wyeliminowana, podczas gdy na podstawie analizy statystycznej uzyskała ona najwyższą wagę spośród wszystkich branych pod uwagę cech.

3. Zintegrowane wagi cech mogą być włączane do procedury budowy miernika syntetycznego albo na etapie normalizacji cech (podejście I) albo obliczania odległości euklidesowych każdego obiektu wielocechowego od obiektu modelowego (podejście II). Badania potwierdziły przypuszczenie, że wartości mierników syntetycznych uzyskanych w podejściu I wykazują przeważnie większe zróżnicowanie i większy obszar zmienności aniżeli w podejściu II.

4. Zastosowanie różnych systemów wag i sposobów ich włączania do procesu konstrukcji miernika syntetycznego wpływa tylko w niewielkim stopniu na zmiany w uporządkowaniach liniowych województw. Potwierdza to tezę wcześniej wysuniętą przez prof. S. Bartosiewicz [2011], że w zagadnieniach konstrukcji miernika syntetycznego ranking obiektów zależy przede wszystkim od wybranego zbioru cech. Inne zabiegi związane np. z normalizacją i ważeniem cech tylko w niewielkim stopniu wpływają na zmiany w rankingu obiektów. Wynika stąd, że w procesie konstrukcji mierników syntetycznych największą wagę należy przykładać do właściwego w sensie merytorycznym doboru cech. Pomocna w rozwiązaniu tego problemu może być metoda FAHP, której zastosowanie umożliwia eliminację ze zbioru wyjściowego cech o niewielkim znaczeniu merytorycznym.

Literatura

- Bank Danych Lokalnych*, 2011, <http://www.stat.gov.pl/bdl>.
- Bartosiewicz B., 2011, *Opowieść o skutkach subiektywizmu w analizie wielowymiarowej*, [w:] Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 176, Taksonomia 18, *Klasyfikacja i analiza danych. Teoria i zastosowania*, Wydawnictwo UE we Wrocławiu, Wrocław, s. 17-20.
- Chang D.-Y., 1996, *Application of the extent analysis method on fuzzy AHP*, *European Journal of Operational Research*, vol. 95, no. 3, s. 649-655.
- Gordon A.D., 1999, *Classification*, Chapman and Hall, 2nd Edition, London, s. 30.
- Grabiński T., 1988, *Metody statystycznej analizy porównawczej*, [w:] Zeliaś A. (red.), *Metody statystyki międzynarodowej*, PWE, s. 235-259.
- Hwang C.L., Yoon K., 1981, *Multiple Attribute Decision Making: Methods and Applications*, Springer, Berlin.
- Łuczak A., Wysocki F., 2011, *Porządkowanie liniowe obiektów z wykorzystaniem rozmytych metod AHP i TOPSIS*, *Przegląd Statystyczny*, tom 53, nr 1-2, s. 3-23.
- Łuczak A., Wysocki F., 2014, *Ustalanie systemu wag dla cech w zagadnieniach porządkowania liniowego obiektów*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 327. Taksonomia 22, *Klasyfikacja i analiza danych. Teoria i zastosowania*, Wydawnictwo UE we Wrocławiu, Wrocław, s. 49-59.
- Walesiak M., 2011, *Uogólniona miara odległości GDM w statystycznej analizie wielowymiarowej z wykorzystaniem programu R*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław.
- Wang J.-W., Cheng C.-H., Kun-Cheng H., 2009, *Fuzzy hierarchical TOPSIS for supplier selection*, *Applied Soft Computing*, vol. 9, s. 377-386.
- Wang Y.-M., Luo Y., 2010, *Integration of correlations with standard deviations for determining attribute weights in multiple attribute decision making*, *Mathematical and Computer Modelling*, vol. 51, s. 1-12.
- Wysocki F., 2010, *Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich*, Wyd. Uniwersytetu Przyrodniczego w Poznaniu, Poznań.

INTEGRATED APPROACH FOR DETERMINING THE WEIGHTING COEFFICIENTS FOR FEATURES IN ISSUES OF LINEAR ORDERING OF OBJECTS

Summary: This paper presents a proposal for an integrated approach to determine the weighting factors for features and their use in the procedure of creating a synthetic measure. It consists of two procedures: statistical (CCSD method – CC – correlation coefficient, SD – standard deviation) and substantive (expert opinions and methods of analytical hierarchy process in the classic version (AHP) or fuzzy (FAHP)) and is applied to study the level of socio-economic development of voivodships in Poland. The studies have shown that different systems of weights and methods of their incorporation into the construction process of synthetic measure influence changes in linear orderings of voivodships to some extent only. Therefore, for the construction of synthetic measures greater weight should be given to the proper selection of features in the substantive sense.

Keywords: linear ordering of objects, the weight factors of features, CCSD method, FAHP, TOPSIS.