

Iwona Konarzewska

Uniwersytet Łódzki
e-mail: i_konarzewska@uni.lodz.pl

RANKINGI WIELOKRYTERIOWE W WARUNKACH ZALEŻNOŚCI LINIOWEJ KRYTERIÓW – PRZYKŁAD BADANIA ŁADU ŚRODOWISKOWEGO W POLSCE W ROKU 2014

MULTI-CRITERIA RANKINGS UNDER LINEAR DEPENDENCE – ENVIRONMENTAL ORDER IN POLAND IN 2014

DOI: 10.15611/pn.2017.468.10

JEL Classification: G11, C40

Streszczenie: W pracy kontynuowane są rozważania na temat konstrukcji wskaźników syntetycznych przy prowadzeniu badań z zakresu wielowymiarowej analizy porównawczej czy też wielokryteriowej analizy dyskretnej (MCDA). Zakłada się najczęściej, że brane pod uwagę kryteria są niezależne. W pracy poddano empirycznej analizie porównawczej, zaproponowaną przez Autora, metodę MCPC tworzenia zmiennej syntetycznej w oparciu o wykorzystanie analizy głównych składowych macierzy kowariancji wartości kryteriów z metodami SAW i TOPSIS. Przeprowadzone badanie dotyczyło wskaźników ładu środowiskowego dla województw Polski w roku 2014 – silnie współzależnych. W pracy pokazano wyniki rankingów oraz zależność zmiennej syntetycznej z metody MCPC od liczby składowych głównych. Badania pokazały, że uzyskiwane za pomocą MCPC rankingi były odporne na sposób ważenia kryteriów. Metodę można traktować jako alternatywę dla innych metod wielokryteriowych, odporną na występowanie zależności liniowych wśród zestawu kryteriów.

Słowa kluczowe: rankingi wielokryteriowe, SAW, TOPSIS, MCPC, zależność liniowa kryteriów, metoda głównych składowych.

Summary: We continue considerations on synthetic indicators construction for multi-criteria rankings (MCDA) or multivariate comparative analysis methods. The methods usually require the criteria to be independent. We present comparative empirical analysis of the proposed method, based on principal components of the criteria covariance matrix (MCPC), with SAW and TOPSIS. The research deals with environmental order indicators for voivodships in Poland in the year 2014 – strongly collinear. We present rankings and an analysis of the relation between the synthetic variables constructed by MCPC as a function of the number of principal components involved. MCPC rankings occurred relatively strong on changing criteria weights. MCPC can be considered as rational alternative multi-criteria method when linear interdependence among criteria is present.

Keywords: multi-criteria rankings, SAW, TOPSIS, MCPC, criteria linear dependence, principal components method.

1. Wstęp

Praca jest wynikiem kontynuacji badań podjętych nad własnościami metod wielokryteriowej analizy dyskretnej i wielowymiarowej analizy porównawczej w warunkach występowania zależności liniowych wśród kryteriów¹.

Celem niniejszej pracy jest próba analizy własności, zaproponowanej przez Autora, metody tworzenia rankingów, MCPC (*Multi-Criteria Principal Components*), bazującej na ortogonalizacji macierzy wartości kryteriów. Podjęto próbę odpowiedzi na następujące pytania teoretyczne:

- Czy MCPC prowadzi do rankingów zasadniczo różniących się od wyników metod SAW i TOPSIS?
- Jak silna jest zależność wyników od liczby uwzględnionych składowych głównych?
- Czy wyniki metody MCPC są porównywalne z wynikami procedury uwzględniającej preselekcję kryteriów w celu spełnienia warunku ich niezależności?

Analizę własności metody MCPC przeprowadzono w oparciu o empiryczne wartości wskaźników zrównoważonego rozwoju dla województw Polski – kryteriów ładu środowiskowego wyznaczonych dla roku 2014². Wśród wziętych efektywnie pod uwagę 12 wskaźników wykryto trzy silne zależności o charakterze liniowym. Wyniki porównywanych metod rangowania: MCPC, SAW i TOPSIS, przedstawiono w przypadku pełnego zestawu kryteriów.

2. Rankingi wielokryteriowe – wybrane metody

W pracy przyjęto założenia takie jak w pracy [Konarzewska 2016], tzn.: A oznacza macierz o wymiarach $N \times K$ unormowanych wartości K kryteriów dla N obiektów; $N \geq K$. Wszystkie kryteria są mierzone na skali przedziałowej bądź ilorazowej.

Sposoby ustalania wag dla kryteriów oparte zostały na doświadczeniach metod wielowymiarowej analizy porównawczej: do pomiaru zmienności wykorzystano klasyczny współczynnik zmienności, do pomiaru informatywności zaś miarę VIF (*Variance Inflation Factor*), związaną ze współczynnikiem korelacji wielorakiej wybranego kryterium ze zbiorem pozostałych [Konarzewska 2016]. Do analiz porównawczych wybrano metody klasyczne: SAW³ i TOPSIS⁴ oraz zaproponowaną przez

¹ Temat ten podjęto w pracy [Konarzewska 2016].

² <http://swaid.stat.gov.pl/> (25.08.2016).

³ *Simple Additive Weighting* [Churchman, Ackoff 1954].

⁴ *Technique for Order Preference by Similarity to Ideal Solution* [Chen, Hwang 1992]

Autora metodę MCPC. Metoda MCPC stanowi modyfikację metody SAW i ma na celu uodpornienie wyników rankingów na występujące zależności między kryteriami. Metoda wykorzystuje rozkład SVD (*Singular Value Decomposition*) macierzy wartości kryteriów według wartości osobliwych⁵. Przedstawiono przykład zastosowania metod do przeprowadzenia rankingu województw Polski pod względem wartości wskaźników ładu środowiskowego w roku 2014.

Przyjęto następujące oznaczenia: f_{ik} – wartość k -tej funkcji kryterium dla i -tego obiektu, $k = 1, \dots, K$, $i = 1, \dots, N$, w_k – waga dla kryterium k -tego; $\sum_{k=1}^K w_k = 1$ oraz $\forall k \ w_k \geq 0$.

2.1. Metoda addytywnego ważenia SAW

Konstruuje się zmienną syntetyczną według wzoru:

$$Q_i = \sum_{k=1}^K w_k \bar{f}_{ik}, \quad (1)$$

gdzie symbol \bar{f}_{ik} oznacza znormalizowaną wartość k -tego kryterium dla i -tego obiektu porównań.

Największa wartość zmiennej Q_i odpowiada najlepszemu z porównywanych obiektów.

2.2. Metoda TOPSIS

W metodzie TOPSIS elementy ważonej znormalizowanej macierzy wartości kryteriów oblicza się według wzoru:

$$\omega_{ik} = \frac{w_k f_{ik}}{\sqrt{\sum_{i=1}^N f_{ik}^2}}. \quad (2)$$

Ranking końcowy uzyskuje się na podstawie malejących wartości wskaźnika C_i^* , zdefiniowanego jako

$$C_i^* = \frac{D_i^-}{D_i^* + D_i^-}, \quad (3)$$

gdzie D_i^* oraz D_i^- są odległościami euklidesowymi wariantu i -tego od rozwiązań: „idealnego” i „anty-idealnego”.

⁵ Algorytm numeryczny dekompozycji SVD macierzy podają np. [Golub, Reinsch 1971]. Do przeprowadzenia obliczeń wykorzystano: Matrix and Linear Algebra for Excel v.2.3.2, <http://digilander.libero.it/foxes>, autor: Leonardo Volpi.

2.3. Wybrane informacje o metodzie głównych składowych

Celem analizy głównych składowych jest wyłonienie kombinacji liniowych obserwowalnych kryteriów, ortogonalnych względem siebie, które dekomponują łączną wariancję układu kryteriów.

Dokonyjemy dekompozycji macierzy kowariancji wartości kryteriów według wzoru:

$$\Sigma = \mathbf{V}\Lambda\mathbf{V}^T, \quad (4)$$

gdzie Σ oznacza macierz kowariancji wartości kryteriów o wymiarach $K \times K$, Λ jest diagonalną macierzą wartości własnych macierzy kowariancji uporządkowanych w ciąg nierosnący, a \mathbf{V} macierzą ortogonalną związanych z nimi wektorów własnych o wymiarach $K \times K$. Pierwsza główna składowa charakteryzuje się największą wariancją, równą największej wartości własnej. Suma wartości własnych macierzy kowariancji to łączna wariancja układu kryteriów. Udział wartości własnej, odpowiadającej wybranej składowej, w sumie wartości własnych jest miernikiem stopnia objaśnienia wariancji układu przez główną składową. W przypadku niezależności kryteriów wszystkie wartości własne macierzy kowariancji są sobie równe (w przypadku dekompozycji macierzy korelacji, są równe jedności).

2.4. Metoda MCPC

Kroki metody MCPC:

1. Obliczenie wartości własnych macierzy kowariancji kryteriów po normalizacji.
2. Podjęcie decyzji o liczbie głównych składowych $K^* < K$, istotnych dla objaśnienia łącznej wariancji kryteriów (np. objaśnienie co najmniej 75% wariancji łącznej).
3. Dekompozycja macierzy \mathbf{A} według wartości osobliwych (SVD)

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T, \quad (5)$$

gdzie \mathbf{U} (\mathbf{U}^*) – macierz normalizowanych wektorów własnych związanych z niezerowymi wartościami własnymi macierzy $\mathbf{A}\mathbf{A}^T$ o wymiarach $N \times K$ ($N \times K^*$), \mathbf{D} (\mathbf{D}^*) – diagonalna macierz o wymiarach $K \times K$ ($K^* \times K^*$) pierwiastków z wartości własnych macierzy $\mathbf{A}^T\mathbf{A}$, \mathbf{V} (\mathbf{V}^*) – macierz o wymiarach $K \times K$ ($K \times K^*$) normalizowanych wektorów własnych macierzy $\mathbf{A}^T\mathbf{A}$.

4. Wyznaczenie macierzy \mathbf{Y}^* o wymiarach $N \times K^*$, której kolumnami są wektory głównych składowych:

$$\mathbf{Y}^* = \mathbf{A}\mathbf{V}^* = \mathbf{U}^*\mathbf{D}^*, \quad (6)$$

* oznacza ograniczenie się do K^* głównych składowych.

5. Wartości zmiennej syntetycznej S_i , $i = 1, \dots, N$, uzyskiwane są za pomocą wzoru:

$$\mathbf{S} = \mathbf{Y}^* (\mathbf{V}^{*T} \mathbf{W}) \mathbf{1} = \mathbf{Y}^* \mathbf{W}^* \mathbf{1} = \mathbf{Y}^* \mathbf{w}^*, \quad (7)$$

gdzie \mathbf{W} oznacza diagonalną macierz wag dla kryteriów, o wymiarach $K \times K$, przyjętych na zasadach jak w metodzie SAW, $\mathbf{1}$ oznacza kolumnowy wektor jedynek, \mathbf{w}^* to wektor „wag” dla głównych składowych (uwaga: elementy tego wektora mogą być ujemne).

6. Ranking obiektów uzyskiwany jest w porządku malejących wartości S_i .

3. Badanie wielokryteriowe ładu środowiskowego w Polsce w roku 2014 – rankingi województw

Zebrano dane o wartościach szesnastu wskaźników zrównoważonego rozwoju charakteryzujących ład środowiskowy w Polsce w roku 2014 według województw⁶. Wskaźniki przedstawione są w tabeli 1. Wszystkie wskaźniki przekształcono do postaci stymulant oraz poddano odpowiedniemu przesunięciu (w przypadku występowania wartości ujemnych) i unitaryzacji.

W pierwszym etapie badania do sporządzenia rankingu województw wykorzystano 12 wskaźników. Stopień uwarunkowania macierzy kowariancji, mierzony ilorazem największej do najmniejszej wartości własnej, wynosił 225 297. Wykryto występowanie trzech silnych związków o charakterze liniowym:

- między wskaźnikami S4 i S6⁷,
- między wskaźnikami S15 i S16⁸,
- wiążących wskaźniki S2, S3, S4, S6, S12⁹.

W tabeli 2 przedstawiono wagi¹⁰ przyporządkowane wskaźnikom: proporcjonalne do współczynników zmienności (oznaczone jako w^a), proporcjonalne do łącznej siły zmienności i pojemności informacyjnej mierzonej za pomocą VIF (*Variance Inflation Factors*) (oznaczone jako w^{VIF}). Poza tym, dla porównania, w badaniu przyjęto dodatkowo wagi równe. Zwracają uwagę zerowe lub bardzo małe wagi w^{VIF} dla wskaźników związanych zależnościami liniowymi.

⁶ W badaniu pominięto cztery z nich, oznaczone w tabeli 1 „*”, ze względu na słabą zmienność.

⁷ Wartość własna macierzy korelacji wynosi $1,7416 \cdot 10^{-5}$; współczynnik korelacji 0,99995.

⁸ Wartość własna macierzy korelacji równa się 0,01483; współczynnik korelacji – 0,94798.

⁹ Wartość własna macierzy korelacji równa jest 0,03374.

¹⁰ Wagi w tabeli 2 zostały, dla większej czytelności, wyrażone w procentach; sumują się zatem do 100%.

Tabela 1. Wskaźniki ładu środowiskowego

Wskaźnik	Definicja	Jednostka pomiaru	Pożądany kierunek zmian
S1*	Odpady komunalne zebrane selektywnie w relacji do ogółu odpadów komunalnych zebranych w ciągu roku	%	MAX
S2	Ilość zmieszanych odpadów komunalnych z gospodarstw domowych zebranych w ciągu roku przypadająca na 1 mieszkańca	kg	MIN
S3	Udział ścieków komunalnych i przemysłowych oczyszczanych w ogóle ścieków wymagających oczyszczenia	%	MAX
S4	Emisja zanieczyszczeń powietrza z zakładów szczególnie uciążliwych – gazowych	t/rok	MIN
S5*	Emisja zanieczyszczeń powietrza z zakładów szczególnie uciążliwych	t/rok	MIN
S6	Emisja dwutlenku węgla z zakładów szczególnie uciążliwych – pyłowych	t/rok	MIN
S7	Udział energii odnawialnej w produkcji energii elektrycznej ogółem	%	MAX
S8	Nakłady na środki trwałe służące ochronie środowiska związane z oszczędzaniem energii elektrycznej na 1 mieszkańca	zł	MAX
S9	Udział obszarów prawnie chronionych w powierzchni ogółem	%	MAX
S10*	Udział powierzchni obszarów Natura 2000 w powierzchni ogółem – udział obszarów specjalnej ochrony ptaków	%	MAX
S11	Udział powierzchni obszarów Natura 2000 w powierzchni ogółem – udział specjalnych obszarów ochrony siedlisk w powierzchni ogółem	%	MAX
S12	Nakłady na środki trwałe służące ochronie środowiska i gospodarce wodnej według kierunków inwestowania: ochrona różnorodności biologicznej i krajobrazu na 1 mieszkańca	zł	MAX
S13	Zasoby eksploatacyjne wód podziemnych – przyrost lub ubytek w stosunku do roku poprzedniego	hm ³	MAX
S14*	Udział powierzchni odnowień i zalesień w powierzchni lasów ogółem	%	MAX
S15	Udział powierzchni użytków rolnych w powierzchni ogółem	%	MAX
S16	Lesistość	%	MAX

Uwaga: W badaniu pominięto wskaźniki oznaczone „*”.

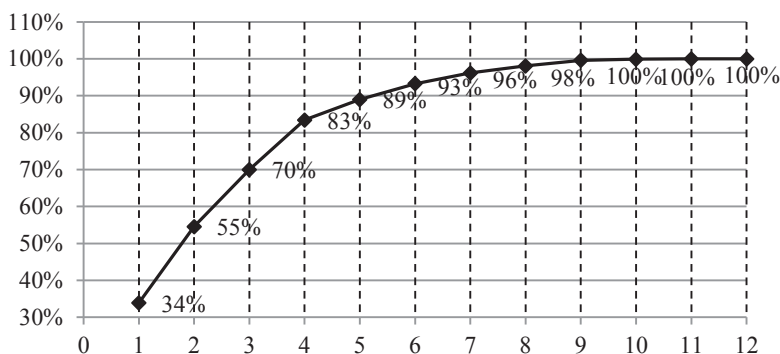
Źródło: opracowanie własne na podstawie: <http://swaid.stat.gov.pl/> (25.08.2016).

Tabela 2. Wagi dla kryteriów ładu środowiskowego (w %)

Wskaźnik/ wagi	S2	S3	S4	S6	S7	S8	S9	S11	S12	S13	S15	S16
w^a	8,26	3,87	4,61	4,58	11,72	9,63	9,29	8,72	16,26	9,55	4,94	8,57
w^{VIF}	5,70	1,58	0,00	0,00	17,46	13,99	11,36	19,04	12,47	16,90	0,54	0,96

Źródło: opracowanie własne.

Wykres osypiska na rys. 1 uzasadnia dokonanie wyboru w metodzie MCPC czterech głównych składowych – objaśniają one ponad 83% łącznej zmienności rozpatrywanych kryteriów ładu środowiskowego. W tabeli 3 przedstawiamy wyniki rankingów.



Rys. 1. Procent objaśnienia zmienności układu 12 kryteriów ładu środowiskowego przez główne składowe – kumulacyjnie

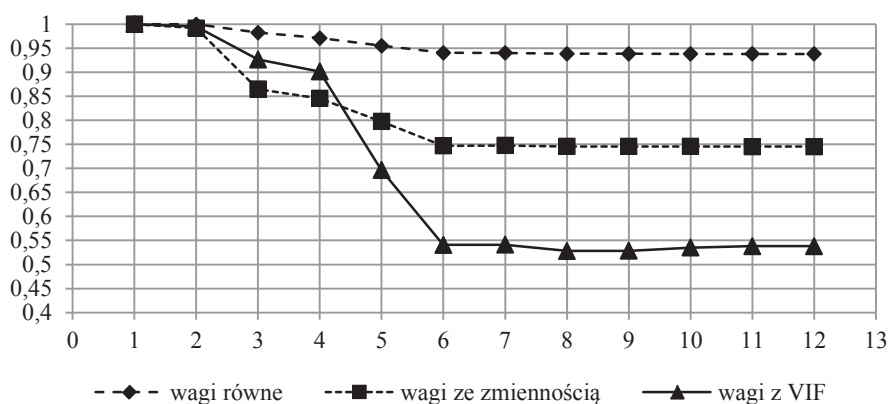
Źródło: opracowanie własne.

Wszystkie zastosowane metody, bez względu na sposób doboru wag, pozycjonują województwo podkarpackie na czele rankingu. W przypadku równych wag dla wszystkich kryteriów, wartości współczynników korelacji uzyskanych dla zmiennych syntetycznych w metodzie SAW oraz MCPC i TOPSIS były podobne, na poziomie 0,97 (pomiędzy MCPC i TOPSIS – 0,91). Wagi uwzględniające siłę zmienności wskaźników spowodowały obniżenie współczynników do poziomów: SAW i MCPC – 0,92, SAW i TOPSIS – 0,94, MCPC i TOPSIS – 0,79. Wzięcie pod uwagę wag w^{VIF} jeszcze mocniej zróżnicowało wyniki rankingów; współczynniki korelacji dla zmiennych syntetycznych wynosiły: SAW i MCPC – 0,65, SAW i TOPSIS – 0,86, MCPC i TOPSIS – 0,78. Zaobserwowano, że rankingi uzyskiwane za pomocą MCPC były najbardziej „odporne” na sposób ważenia kryteriów. Zbadano również, jak wpływa na wyniki metody MCPC włączanie do obliczeń kolejnych składowych głównych macierzy kowariancji wskaźników. Wyniki przedstawione są na rys. 2.

Tabela 3. Rankingi województw – 12 kryteriów ładu środowiskowego

Metoda rangowania	Wagi równe			Wagi w^a			Wagi w^{vif}		
	SAW	MCPC-4	TOPSIS	SAW	MCPC-4	TOPSIS	SAW	MCPC-4	TOPSIS
DOLNOŚLĄSKIE	12	13	13	14	14	14	14	14	14
KUJAWSKO-POMORSKIE	6	7	4	7	10	3	5	9	2
LUBELSKIE	9	6	10	12	7	13	15	7	15
LUBUSKIE	3	3	2	2	2	2	4	2	3
ŁÓDZKIE	13	14	14	9	12	8	9	13	4
MAŁOPOLSKIE	5	4	6	4	3	5	8	3	5
MAZOWIECKIE	14	15	12	13	15	11	11	15	10
OPOLSKIE	15	12	15	15	13	16	16	12	16
PODKARPACKIE	1	1	1	1	1	1	1	1	1
PODLASKIE	2	2	3	3	4	4	2	4	6
POMORSKIE	10	10	9	10	9	10	12	10	12
ŚLĄSKIE	16	16	16	16	16	15	13	16	13
ŚWIĘTOKRZYSKIE	7	8	8	5	5	9	3	5	7
WARMIŃSKO-MAZURSKIE	4	5	5	6	6	6	6	6	11
WIELKOPOLSKIE	11	11	11	11	11	12	10	11	9
ZACHODNIOPOMORSKIE	8	9	7	8	8	7	7	8	8

Źródło: opracowanie własne.



Rys. 2. Współczynniki korelacji liniowej zmiennych syntetycznych w metodzie MCPC jako funkcja liczby uwzględnionych składowych głównych – 12 kryteriów

Źródło: opracowanie własne.

Decyzja o liczbie składowych głównych ma istotny wpływ na wyniki rankingów, zwłaszcza przy metodzie ważenia uwzględniającej pojemność informacyjną kryteriów – w badaniu dokonano wyboru czterech składowych głównych, ale widać, że rankingi różniłyby się zasadniczo już przy wyborze liczby składowych od sześciu wzwyż¹¹.

Przeprowadzono również analizy, w których z zestawu wskaźników eliminowano sukcesywnie wskaźniki: S6 oraz S15, wprowadzające występowanie silnych związków liniowych. Stopień uwarunkowania macierzy kowariancji uległ zmniejszeniu do wartości 279,02, a następnie 108,33. Zaobserwowano, że wraz z poprawą uwarunkowania wartości współczynników korelacji zmiennych syntetycznych w metodzie MCPC przy uwzględnianiu kolejnych składowych głównych rosły.

4. Podsumowanie

Przeprowadzone badanie pokazało, że wskaźniki ładu środowiskowego prowadzą do rankingów wielokryteriowych, w których kryteria mogą być silnie liniowo współzależne. Stwierdzono, że sposób ważenia kryteriów wywiera istotny wpływ na wartości zmiennych syntetycznych i uzyskiwane wyniki rankingów. Zbadano, czy proponowana metoda MCPC, z uwzględnieniem uzasadnionej liczby głównych składowych macierzy kowariancji, daje wyniki podobne do metody SAW, zastosowanej do ograniczonej liczby kryteriów (po usunięciu wskaźników liniowo związanych). W przypadku wag proporcjonalnych do zmienności uzyskano współczynnik korelacji dla odpowiednich zmiennych syntetycznych 0,88, a w przypadku wag uwzględniających dodatkowo VIF – 0,63. Zaproponowana przez Autora metoda MCPC okazała się rozsądną alternatywą dla metody SAW i TOPSIS, odporną na występowanie związków liniowych wśród kryteriów.

Literatura

- Chen S. J., Hwang C.L., 1992, *Fuzzy Multiple Attribute Decision Making: Methods and Applications*, Springer, Berlin.
- Churchman C.W., Ackoff R.L., 1954, *An approximate measure of value*, Journal of Operations Research Society of America, vol. 2(1).
- Golub G.H., Reinsch C., 1971, *Singular Value Decomposition and Least Squares Solutions*, [w:] *Handbook for Automatic Computation*, vol. II (chief ed. Bauer F.L). Wilkinson J. H., Reinsch C., *Linear Algebra*, Springer Verlag, Berlin.
- Konarzewska I., 2016, *Rankingi wielokryteriowe a współzależność liniowa kryteriów*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 426, Taksonomia 26, s. 69-79.

¹¹ Zauważmy, że metoda SAW uwzględnia wszystkie 12 składowych głównych.