

Joanna Trzęsiok, Michał Trzęsiok

Akademia Ekonomiczna w Katowicach

METRYKI I ICH ZASTOSOWANIE W METODACH WIELOKRYTERIALNYCH WSPOMAGAJĄCYCH PODEJMOWANIE DECYZJI

1. Wstęp

Jednym z podstawowych elementów procesu uczenia się, pozyskiwania wiedzy z obserwowania otaczającej nas rzeczywistości jest odnajdywanie analogii. W przypadku badania struktur złożonych nieuniknionym jest dokonanie pomiaru podobieństwa. Pozwala to na formalne postawienie problemu i wykorzystanie narzędzi matematyczno-statystycznych do skonstruowania modelu oraz wnioskowania na jego podstawie.

Pomiar podobieństwa jest głównym elementem wielu metod wielowymiarowej analizy statystycznej, ze szczególnym uwzględnieniem metod klasyfikacji oraz porządkowania liniowego. W artykule rozważane będzie zadanie porządkowania liniowego w kontekście wielokryteriowych metod wspomagających podejmowanie decyzji.

Wśród miar podobieństwa wyróżnia się miary odległości oraz bliskości. Duży stopień podobieństwa obiektów oznacza małe wartości miar odległości lub równoważnie duże wartości funkcji bliskości. Ze względu na tę wzajemną odpowiedniość w dalszej części rozpatrywane będą jedynie miary odległości. Pomiar odległości odbywa się przez obliczenie wartości pewnej ustalonej funkcji – *metryki*.

Materiał dotyczący przestrzeni metrycznych sprawia na ogół trudności studentom pierwszego roku. Szczególnie problematyczne jest dla nich zrozumienie potrzeby wprowadzania różnych miar odległości. Przedstawienie prostego przykładu ilustrującego zastosowanie metryk w metodach wspomagających podejmowanie decyzji umożliwia podkreślenie znaczenia wyboru metryki. Pozwala również na porównanie własności tychże metryk.

2. Podejście dystansowe w wielokryterialnej analizie decyzyjnej

2.1. Sformułowanie problemu

Rozważmy sytuację, w której zadaniem decydenta jest wybór jednego spośród pewnej ustalonej (skończonej) liczby możliwych wariantów decyzyjnych. Postulujemy, aby dokonany wybór nie był przypadkowy, lecz by maksymalizował poziom użyteczności decydenta. Sytuacja taka często występuje w zagadnieniach ekonomicznych, np. przy wyborze najlepszej oferty po ogłoszeniu przetargu lub w procesie rekrutacji przy wyborze najlepszego kandydata.

Najczęściej wybór dokonywany jest nie na podstawie jednej cechy charakteryzującej badane obiekty, lecz jednocześnie ze względu na wiele cech, nazywanych kryteriami wyboru. Rozważany więc będzie problem wielokryterialny w skończonym zbiorze wariantów decyzyjnych. Zagadnienie trywializuje się, gdy w zbiorze wariantów decyzyjnych występuje wariant dominujący – najlepszy ze względu na każde kryterium. Najczęściej jednak mamy do czynienia z sytuacją, kiedy poszczególne kryteria wskazują różne warianty i trzeba ustalić rozwiązanie kompromisowe na podstawie sprzecznych wskazań kryteriów cząstkowych.

Kryteria cząstkowe mogą być reprezentowane przez zmienne mierzone na różnych skalach pomiarów. Wyróżniamy cztery rodzaje skal (zob. M. Walesiak (2002)):

- nominalną,
- porządkową (rangową),
- przedziałową,
- ilorazową.

Skale pomiaru badanych cech mają kluczowe znaczenie dla doboru miar odległości oraz innych narzędzi wykorzystywanych w dalszej analizie.

Rozważać będziemy przypadek, gdy wszystkie kryteria wyboru są mierzone na skali ilorazowej.

W praktyce dla decydenta niektóre kryteria mogą mieć większe znaczenie od innych. W takim przypadku możliwe jest uwzględnienie tych preferencji poprzez wprowadzenie systemu wag. Niech:

$W = \{ W_j : j = 1, 2, \dots, N \}$ oznacza zbiór wariantów decyzyjnych,

$K_1, K_2, \dots, K_M$ – kryteria wyboru,

k_{ij} – wartość i -tego kryterium dla j -tego wariantu decyzyjnego,

w_i – wagi przyporządkowane przez decydenta i -temu kryterium.

2.2. Rozwiązanie problemu

Jedno z podejść w metodach wielokryterialnej analizy decyzyjnej to tzw. podejście dystansowe, w którym wybór optymalnego rozwiązania spośród skończonej liczby wariantów decyzyjnych (przypadek dyskretny), odbywa się przez wska-

zanie wariantu „najbliższego” pewnemu teoretycznemu rozwiązaniu, które przyjmujemy za wzorcowe. Aby móc wskazać rozwiązanie „najbliższe”, musimy mieć możliwość mierzenia odległości między rozwiązaniami.

Z formalnego punktu widzenia, aby rozwiązać problem decyzyjny, wystarczy skonstruować przestrzeń metryczną z wyróżnionym elementem reprezentującym rozwiązanie wzorcowe oraz elementami, będącymi odpowiednikami wariantów decyzyjnych. Następnie należy obliczyć wartości metryki między wzorcem a każdym punktem reprezentującym rozważane warianty. Ten wariant, którego odległość od wzorca jest najmniejsza, jest wskazany jako optymalny wybór.

Poszczególne warianty W_j reprezentowane są przez wektory:

$$(k_{1j}, k_{2j}, \dots, k_{mj}) \in \mathbf{R}^M. \quad (1)$$

W metodach porządkowania liniowego zachodzi potrzeba pozbawienia wartości zmiennych mian oraz ujednolicenia rzędów wielkości. W tym celu stosuje się transformacje normalizacyjne, na przykład:

- standaryzację: $z_{ij} = \frac{k_{ij} - \bar{k}_i}{s_i}, \quad (2)$

- unitaryzację: $z_{ij} = \frac{k_{ij} - \min_j \{k_{ij}\}}{\max_j \{k_{ij}\} - \min_j \{k_{ij}\}}, \quad (3)$

- normalizację: $z_{ij} = \frac{k_{ij}}{\max_j \{k_{ij}\}}, \quad (4)$

gdzie $\bar{k}_i$, s_i oznaczają odpowiednio średnią arytmetyczną oraz odchylenie standardowe dla i -tej zmiennej.

W kolejnym kroku konstruowany jest wektor wzorcowy. Można uczynić to arbitralnie, wskazując pożądaną wartość zmiennych $K_1, K_2, \dots, K_M$. Najczęściej jednak wykorzystuje się do tego celu rzeczywiste wartości zmiennych z rozpatrywanych wariantów. Konieczna jest wtedy identyfikacja charakteru kryteriów pod względem preferencji wśród zmiennych.

Jeżeli wraz ze wzrostem wartości kryterium K_i użyteczność wariantu maleje, to powiemy, że kryterium to ma charakter *destymulacyjny*. Natomiast jeśli wraz ze wzrostem wartości kryterium K_i użyteczność wariantu rośnie – o K_i powiemy, że ma charakter *stymulacyjny* (zob. Z. Hellwig (1968)). Mogą występować również kryteria z wyróżnioną pożądaną wartością lub przedziałem takich wartości, powyżej i poniżej którego wartości kryteriów oznaczają mniejszą użyteczność decyden-ta. Kryteria takie nazywamy *nominantami* (zob. T. Borys (1980)).

Po zidentyfikowaniu charakteru kryteriów można przejść do ustalenia współrzędnych wektora wzorcowego. Oznaczmy wektor wzorcowy przez:

$$\hat{\mathbf{k}} = (\hat{k}_1, \hat{k}_2, \dots, \hat{k}_M) \in \mathbf{R}^M. \quad (5)$$

Przyjmujemy:

- a) $\hat{k}_i = \min_j z_{ij}$, jeśli kryterium K_i ma charakter destymulacyjny,
- b) $\hat{k}_i = \max_j z_{ij}$, jeśli kryterium K_i ma charakter stymulacyjny,
- c) $\hat{k}_i$ jako nominalny poziom zmiennej K_i (znormalizowanej), jeśli kryterium K_i ma charakter nominacyjny.

Wektor wzorcowy jest obiektem teoretycznym i zazwyczaj nie należy do zbioru wariantów decyzyjnych poza szczególnym, wspomnianym już przypadkiem, kiedy jeden z wariantów zdominuje pozostałe – zostanie on wtedy wyznaczony jako wzorzec.

Jak już wspomniano, w celu wyłonienia rozwiązania optymalnego poszukujemy wariantu najbliższego wektorowi wzorcowemu. Zatem w kolejnym etapie dokonujemy wyboru miary odległości (metryki).

Metryką w przestrzeni liniowej $\mathbf{R}^M$ nazywamy dowolną funkcję

$$d : \mathbf{R}^M \times \mathbf{R}^M \rightarrow [0, +\infty),$$

spełniającą warunki:

- a) $\bigwedge_{\mathbf{x}, \mathbf{y} \in \mathbf{R}^M} d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$ (warunek identyczności),
- b) $\bigwedge_{\mathbf{x}, \mathbf{y} \in \mathbf{R}^M} d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (warunek symetrii),
- c) $\bigwedge_{\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{R}^M} d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ (warunek trójkąta).

W wielokryterialnych metodach wspomagających podejmowanie decyzji najczęściej wykorzystywane metryki to:

- metryka miejska: $d_1(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^M w_i \cdot |x_i - y_i|$, (6)

- metryka euklidesowa: $d_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^M w_i \cdot (x_i - y_i)^2}$, (7)

- metryka maksimum: $d_\infty(\mathbf{x}, \mathbf{y}) = \max_{1 \leq i \leq M} \{w_i \cdot |x_i - y_i|\}$, (8)

gdzie $\mathbf{x} = (x_1, x_2, \dots, x_M) \in \mathbf{R}^M$, $\mathbf{y} = (y_1, y_2, \dots, y_M) \in \mathbf{R}^M$,

natomiast $(w_1, w_2, \dots, w_M) \in \mathbf{R}_+^M$ jest wektorem wag ustalonym przez decydenta.

Stosując ustaloną postać metryki, obliczamy odległości między wektorem wzorcowym a każdym z rozważanych wariantów decyzyjnych:

$$d(\hat{\mathbf{k}}, \mathbf{z}_j) = d((\hat{k}_1, \hat{k}_2, \dots, \hat{k}_M), (z_{1j}, z_{2j}, \dots, z_{Mj})), \quad j = 1, 2, \dots, N. \quad (9)$$

Wariant, którego odległość od wzorca jest najmniejsza, jest wskazany jako rozwiązanie optymalne.

3. Przykład ilustrujący podejście dystansowe w wielokryterialnej analizie decyzyjnej

Rozważmy sytuację, w której decydent chce dokonać modernizacji budynku. Stoi przed wyborem jednej z siedmiu ofert firm budowlanych. W przeglądanych ofertach decydent wyróżnił cztery zmienne, które stanowić będą kryteria wyboru:

K_1 – planowany koszt materiałów [zł],

K_2 – planowany koszt pracy [zł],

K_3 – czas realizacji usługi [liczba dni],

K_4 – okres gwarancji [liczba miesięcy].

Oznaczmy przez k_{ij} wartość i -tego kryterium dla j -tej oferty (variantu W_j). Wartości kryteriów dla poszczególnych wariantów (ofert) przedstawia tab. 1.

Tabela 1. Wartości kryteriów dla poszczególnych ofert

	W_1	W_2	W_3	W_4	W_5	W_6	W_7
K_1	26 000	30 000	25 000	28 000	30 000	27 000	25 000
K_2	23 000	22 000	24 000	21 000	19 000	12 000	17 000
K_3	18	21	25	35	32	45	38
K_4	9	18	18	15	12	12	12

Źródło: dane umowne.

Trzy pierwsze kryteria mają charakter destymulant, czwarte zaś – stymulanty.

Zauważmy, że w zbiorze wariantów decyzyjnych nie występuje wariant dominujący. Poszczególne kryteria, rozważane osobno, wskazują różne warianty jako rozwiązania optymalne. Uzasadnione wydaje się wykorzystanie podejścia dystansowego opisanego w podrozdziale 2 do rozwiązania postawionego problemu decyzyjnego.

W rozważanym zagadnieniu zakładamy ponadto, że decydent ustalił wagi w_i dla poszczególnych kryteriów równe: 0,2, 0,4, 0,3 oraz 0,1 odpowiednio.

Dokonujemy normalizacji wartości kryteriów według formuły (4). Wartości znormalizowanych kryteriów oznaczonych przez z_i przedstawia tab. 2.

Tabela 2. Znormalizowane wartości kryteriów

	W_1	W_2	W_3	W_4	W_5	W_6	W_7
z_1	0,867	1,000	0,833	0,933	1,000	0,900	0,833
z_2	0,958	0,917	1,000	0,875	0,792	0,500	0,708
z_3	0,400	0,467	0,556	0,778	0,711	1,000	0,844
z_4	0,500	1,000	1,000	0,833	0,667	0,667	0,667

Źródło: obliczenia własne na podstawie tab. 1.

W kolejnym kroku ustalone zostaną wartości poszczególnych kryteriów, które przyjmiemy za wzorcowe. Wzorzec tworzą minimalne wartości trzech pierwszych kryteriów o charakterze destymulant oraz maksymalna wartość czwartego kryterium – stymulanty.

Wobec tego wektor wzorcowy $\hat{k}$ przybiera postać:

$$\hat{k} = \left(\min_j z_{1j}; \min_j z_{2j}; \min_j z_{3j}; \max_j z_{4j} \right) = (0,833; 0,5; 0,4; 1).$$

Wektorowi wzorcowemu odpowiadają wartości teoretycznej (nieistniejącej) oferty o najmniejszym koszcie materiałów równym 25 000 zł, o najmniejszym koszcie pracy – 12 000 zł, o najkrótszym – 18-dniowym czasie realizacji, oraz o najdłuższym – 18-miesięcznym okresie gwarancyjnym.

W następnym etapie, zgodnie z procedurą opisaną w podrozdziale 2, następuje obliczenie odległości wariantów decyzyjnych od wzorca dla jednej ustalonej metryki. Rozważany przykład zaś stanowić może podstawę do wstępnej analizy własności metody w zależności od wyboru metryki. Dlatego zbadane zostaną odległości poszczególnych wariantów od wzorca dla trzech standardowo wykorzystywanych miar odległości danych wzorami (6)-(8). Obliczone wartości przedstawia tab. 3.

Otrzymane wyniki pozwalają na liniowe uporządkowanie wariantów decyzyjnych od najbliższych wektorowi wzorcowemu (najlepszych) do najdalszych (tab. 4).

Tabela 3. Odległości poszczególnych wariantów decyzyjnych od wzorca dla różnych metryk

	W_1	W_2	W_3	W_4	W_5	W_6	W_7
$d_1(\hat{k}, W_j)$	0,240	0,220	0,247	0,300	0,277	0,227	0,250
$d_2(\hat{k}, W_j)$	0,331	0,276	0,328	0,322	0,282	0,346	0,296
$d_\infty(\hat{k}, W_j)$	0,183	0,167	0,200	0,150	0,117	0,180	0,133

Źródło: obliczenia własne na podstawie danych z tab. 2.

Celem metod wspomagających podejmowanie decyzji jest wskazanie rozwiązania optymalnego. Informacja, którą uzyskuje decydent po zastosowaniu podejścia dystansowego, jest pełniejsza, gdyż zawiera również ranking wariantów decyzyjnych.

Tabela 4. Liniowe uporządkowanie wariantów decyzyjnych (zaczynając od wariantów najbliższych wzorcowi) dla różnych metryk

	1.	2.	3.	4.	5.	6.	7.
d_1	W_2	W_6	W_1	W_3	W_7	W_5	W_4
d_2	W_2	W_5	W_7	W_4	W_3	W_1	W_6
d_∞	W_5	W_7	W_4	W_2	W_6	W_1	W_3

Źródło: obliczenia własne na podstawie danych z tab. 3.

Otrzymane liniowe uporządkowania ofert w sposób istotny zależą od wyboru metryki. Dlatego szczególnie ważna dla badacza i decydenta jest znajomość własności poszczególnych metryk przy stosowaniu narzędzi wykorzystujących miary odległości. W przypadku metryk miejskiej i euklidesowej na wartość pomiaru wpływ mają wszystkie kryteria. Możliwa jest sytuacja, w której jako optymalne wskazane zostanie rozwiązanie, w którym jedno, szczególnie istotne dla decydenta kryterium będzie bardzo dalekie od wartości wzorcowej, ale wszystkie pozostałe będą bliskie. Zjawisko zastępowania wpływu niekorzystnych wartości jednych zmiennych pożądanymi wartościami innych nazywamy substytucyjnością kryteriów. Jednym ze sposobów ograniczania negatywnego zjawiska substytucyjności jest stosowanie wag. Jest to jednak sposób niedoskonały. Często lepszym rozwiązaniem jest wybór metryki maksimum. Wpływ na jej wartość mają nie wszystkie kryteria, lecz jedno, wskazujące najmniej korzystny element oferty. Ponadto wartości metryki maksimum są interpretowalne.

4. Uwagi końcowe

Należy zwrócić uwagę na problem wyboru skali w której mierzymy. W literaturze spotkać można wiele propozycji miar odległości, gdy wszystkie zmienne mierzone są na skali nominalnej (m.in. Sokala i Michenera, Jaccarda, Czekanowskiego), na skali porządkowej (Kendalla), na skali przedziałowej (miejska, euklidesowa, maksimum, Jeffreysa-Matusita, Mahalanobisa) (zob. M. Walesiak (2002)). W pracy (M. Walesiak (2002)) przedstawiono również uogólnioną miarę odległości, którą można stosować, gdy kryteria opisujące warianty decyzyjne mierzone są na skalach różnych rodzajów, co bardzo rozszerzyło zakres zastosowań prezentowanego podejścia dystansowego. Zaznaczyć jednak należy, że zarówno uogólniona miara odległości, jak i niektóre z wymienionych powyżej miar nie są metrykami, gdyż nie spełniają warunku trójkąta.

Podkreślić trzeba także, że pionierskie prace metodologiczne dotyczące metod wykorzystujących pomiar odległości m.in. w zagadnieniach taksonomicznych i w badaniach nad miernikami rozwoju gospodarczego, opublikowane zostały w latach sześćdziesiątych XX wieku przez wrocławskie środowisko naukowe ze szczególnym wkładem prof. Hellwiga.

Literatura

- T. Borys (1980). *Elementy teorii jakości*. Państwowe Wydawnictwo Naukowe PWN. Warszawa.
- Z. Hellwig (1968). *Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr*. Przegląd Statystyczny z. 4. Str. 307-327.
- S. Krawczyk (2001). *Metody ilościowe w planowaniu*. Tom I. Wydawnictwo C.H. Beck. Warszawa.

-
- T. Trzaskalik (1990). *Wielokryterialne dyskretne programowanie dynamiczne. Teoria i zastosowania w praktyce gospodarczej*. Prace Naukowe AE w Katowicach. Katowice.
- M. Walesiak (2002). *Uogólniona miara odległości w statystycznej analizie wielowymiarowej*. Wydawnictwo Akademii Ekonomicznej we Wrocławiu.

MATRICES AND THEIR APPLICATIONS TO THE MULTICRITERIA OPTIMIZATION SUPPORTING THE DECISION MAKING PROCESS

Summary

One of the most common approaches in the multicriteria optimization is called "the distance approach". It can be used in a discrete case. The optimal solution is the one which is the closest, in the sense of a given distance function, to the pattern. The pattern is a theoretical (ideal) solution set by the decision maker. The essential role is played by the distance function which is simply the metric in the decision space.

In teaching mathematics the presentation of an example illustrating this method allows to stress the importance of choosing the metrics and comparing their properties. The paper focuses on the issue of controlling the criteria substitution in particular.