

Paweł CABAŁA
Uniwersytet Ekonomiczny w Krakowie
cabalap@uek.krakow.pl

ANALIZA SPÓJNOŚCI OCEN W PROCESIE PODEJMOWANIA DECYZJI STRATEGICZNYCH

Streszczenie. W artykule omówiono zagadnienie pomiaru spójności ocen w procesie podejmowania decyzji strategicznych. Decyzje strategiczne wymagają często formułowania ocen na podstawie subiektywnych przekonań decydentów, a wybór ostatecznego rozwiązania jest efektem analizy wielu czynników jednocześnie. W artykule przedstawiono propozycję podziału systemów oceny, następnie wyjaśniono pojęcie spójności ocen. Spójność ocen powiązano z zasadą przechodniości preferencji. Do pomiaru odstępstw od tej zasady zaproponowano wskaźnik konsystencji Kendalla oraz wskaźnik spójności Saaty'ego.

Słowa kluczowe: systemy oceniania, spójność preferencji, decyzje strategiczne, pomiar spójności ocen, metoda porównań parami

PREFERENCE CONSISTENCY ANALYSIS IN THE PROCESS OF MAKING STRATEGIC DECISIONS

Summary. The article discusses the problem of consistency measurement in the process of strategic decision making. Strategic decisions often require the formulation of uncertain assumptions which are based on subjective evaluations of decision-makers and the strategic choice is the result of analysis of many factors simultaneously. At the beginning of the article a proposal of the classification of assessment systems is characterized, then the concept of consistency and the essence of pairwise comparisons are presented. Consistency is associated with the principle of preference transitivity. Kendall's coefficient of consistence and Saaty's consistency ratio are proposed as the tools of measurement.

Keywords: evaluating systems, consistency of preferences, strategic decisions, measuring the consistency of assessments, method of pairwise comparisons

1. Uwagi wstępne

Problematyka podejmowania decyzji strategicznych wiąże się ściśle z zagadnieniem oceny. Ocena to stosunek podmiotu oceniającego do przedmiotu oceny (obiektu) wyrażony w formie sądu wartościującego. W decyzjach strategicznych brane są często pod uwagę czynniki o charakterze jakościowym, które są oceniane ze względu na rozmaite kryteria. Na przykład w tworzeniu scenariuszy rozwoju przedsiębiorstwa badana jest lista kilkudziesięciu czynników determinujących zmiany (*driving forces*). Lista takich czynników jest następnie oceniana pod kątem siły ich wpływu na system oraz pod kątem stopnia ich przewidywalności (niepewności). Do konstrukcji scenariuszy przyjmuje się ostatecznie 2 lub 3 zmienne, których siła oddziaływania i niepewność są największe. Wybór tych czynników (obiektów) wymaga zatem ich oceny (uszeregowania) z uwagi na stopień spełnienia kryteriów (wpływ i niepewność). Jakościowy charakter ma także ocena tzw. krytycznych czynników sukcesu w analizie strategicznej czy analiza szans i zagrożeń w otoczeniu przedsiębiorstwa.

W teorii podejmowania decyzji pojawia się również potrzeba identyfikacji rozkładu prawdopodobieństwa stanów natury, który można określić apriorycznie, na przykład porównując parami wyróżnione stany. Korzysta się tu najczęściej z subiektywnych opinii osób posiadających dostateczną wiedzę i doświadczenie (ekspertów), co wiąże się z możliwością pomyłek i błędów, które obniżają jakość ocen.

W celu podniesienia jakości ocen stosowane są różne rozwiązania. Jednym z nich jest włączenie w proces większej liczby kompetentnych osób. Uzyskanie zadowolającej zgodności ocen niezależnych ekspertów uznaje się za bardziej wiarygodne źródło informacji. Innym sposobem weryfikacji wiarygodności informacji jest identyfikacja błędów logicznych, które pojawiają się podczas oceny zjawisk jakościowych.

Błędy logiczne ujawniają się m.in. w stosowaniu metody porównań parami. Porównując parami elementy (obiekty, zjawiska, czynniki), jesteśmy szczególnie podatni na naruszenie zasady przechodniości preferencji. Przykładowo, jeżeli dla trzech elementów A , B oraz C uznajemy, że $A \succ B$ i $B \succ C$, to stwierdzenie $A \prec C$ jest niedopuszczalne. Tego typu błędy występują często, gdy ocena ma charakter jakościowy, a liczba porównywanych obiektów jest duża.

Celem artykułu jest prezentacja metod wykorzystywanych w identyfikacji takich błędów. Zaprezentowane zostały dwie metody pomiaru spójności preferencji (ocen), które można wykorzystać w porównaniach parami danego zbioru elementów – współczynnik konsystencji Kendalla oraz wskaźnik spójności Saaty'ego. Stosowanie tych mierników pozwala podnieść wiarygodność ocen, zwiększając tym samym efektywność decyzji strategicznych.

Wyniki wielu badań eksperymentalnych wskazują, że ludzie (w tym również eksperci) popełniają podstawowe błędy logiczne w procesie podejmowania decyzji. Systematyczne badania w tym zakresie rozpoczął w 1969 r. A. Tversky. Badania te były kontynuowane

w latach 70. i 80. XX w. przez licznych autorów i są obecnie przedmiotem debat w związku z tzw. ekonomią behawioralną¹.

2. Systemy oceniania

Ocena to postawa kogoś (podmiot oceniający, ekspert) wobec czegoś (przedmiot oceny, obiekt), ze względu na coś (kryterium oceny). Wynika z tego, że oceniać może jedna lub więcej osób, jeden lub więcej obiektów, stosując jedno lub więcej kryteriów. A zatem istnieje 8 systemów oceny, które pokazano w tabeli 1.

Tabela 1

Systemy oceny

System oceny	Podmiot oceniający (<i>kto?</i>)	Przedmiot oceny (<i>co?</i>)	Kryterium oceny (<i>ze względu na co?</i>)
I	Ocena indywidualna	Wiele obiektów	Jedno kryterium
II	Ocena grupowa	Wiele obiektów	Jedno kryterium
III	Ocena indywidualna	Wiele obiektów	Wiele kryteriów
IV	Ocena grupowa	Wiele obiektów	Wiele kryteriów
V	Ocena indywidualna	Jeden obiekt	Jedno kryterium
VI	Ocena grupowa	Jeden obiekt	Jedno kryterium
VII	Ocena indywidualna	Jeden obiekt	Wiele kryteriów
VIII	Ocena grupowa	Jeden obiekt	Wiele kryteriów

Źródło: Opracowanie własne.

Wybór określonego systemu oceny zależy od istoty problemu badawczego oraz dostępnych środków (koszty, czas). Ocena grupowa jest wynikiem niezależnych opinii ekspertów (np. badanie delfickie) lub bezpośrednich uzgodnień między ekspertami (np. burza mózgów). Ocenie może podlegać jeden obiekt (np. system zapewnienia jakości) lub wiele obiektów (np. czynniki ryzyka). W ocenie może być zastosowane jedno kryterium (np. kolor) lub większa liczba kryteriów (np. kolor, kształt, waga).

W przypadku oceny wielu obiektów wykorzystuje się jedną z trzech ogólnych metod: bezpośrednią (ocena każdego obiektu oddzielnie w przyjętej z góry skali), szeregowania (ocena obiektów ze względu na ich pozycję względem pozostałych obiektów) oraz porównań parami (ocena wszystkich par obiektów).

Ocena wielu obiektów polega na określeniu znaczenia (istotności, pozycji, rangi) każdego z tych obiektów z uwagi na dane kryterium. Ocena taka umożliwi mocne lub słabe uporządkowanie obiektów. Uporządkowanie mocne występuje wtedy, gdy dla wszystkich

¹ Michel-Kerjan E. , Slovic P. (red.): The Irrational Economist. Making Decisions in a Dangerous World, PublicAffairs, New York 2010.

obiektów zachodzi wyłącznie relacja preferencji. Słaby porządek występuje wtedy, gdy zachodzi relacja indyferencji przynajmniej między dwoma obiektami.

Wieloobiektowe systemy ocen (systemy I i II w tablicy 1) mogą być albo asymetryczne (mocno porządkujące), albo niesymetryczne (słabo porządkujące). Mówiąc bardziej obrazowo: w systemach asymetrycznych, jeżeli relacje zachodzą w jednym kierunku, to nie mogą zachodzić w drugim kierunku, natomiast w systemach niesymetrycznych, jeżeli relacja zachodzi w jednym kierunku, to może, ale nie musi, zachodzić w drugim kierunku. Asymetrycznym systemem oceniania jest szeregowanie obiektów w kolejności ich znaczenia. W systemach niesymetrycznych obiekty mogą zajmować taką samą pozycję.

3. Pojęcie spójności ocen

Jeżeli ekspert ocenia obiekty (czynniki, zjawiska, osoby, zdarzenia) z danego punktu widzenia (ze względu na określone kryterium różnicujące), to aktywność taką można nazwać porządkowaniem. Zasadą każdego ustanowionego porządku jest relacja preferencji lub indyferencji (\succ , \sqcup). Relacja preferencji jest asymetryczna (jeżeli $X \succ Y$, to nieprawdą jest, że $Y \succ X$) oraz przechodnia (jeżeli $X \succ Y$ i $Y \succ Z$, to $X \succ Z$). Przechodnia jest również relacja indyferencji, lecz w odróżnieniu od relacji preferencji jest symetryczna (jeżeli $X \sim Y$, to $Y \sim X$). Relacje preferencji oraz indyferencji powinny być ponadto zupełne, co oznacza, że między dowolnymi dwoma obiektami, z przyjętego do rozważań zbioru obiektów, zachodzi albo relacja preferencji, albo relacja indyferencji.

W logice formalnej oraz teorii mnogości spójnymi nazywane są takie relacje w zbiorze Z , które dla dowolnych x i y (elementów tego zbioru) oznaczają, że albo x jest równe y , albo x jest w relacji do y , albo y jest w relacji do x . W dalszej kolejności spójność ocen (preferencji) jest rozumiana jako własność opisująca stopień, w jakim oceny badanych obiektów spełniają zasadę przechodniości preferencji lub indyferencji.

Analiza spójności preferencji dotyczy przypadków, w których obiekty są porównywane parami w celu ustalenia ich relatywnego znaczenia ze względu na określone kryterium oceny. Przykładowo porównywać parami można dostawców z uwagi na ryzyko utraty ciągłości dostaw, pracowników pod kątem ich wkładu w osiągnięcia zespołu czy klientów z perspektywy możliwości składania kolejnych zamówień.

W metodzie porównań parami zakłada się najczęściej, że stosunek odwrotności preferencji będzie przez oceniającego zachowany. Przykładowo, jeżeli oceniamy, że obiekt A jest lepszy od obiektu B ($A \succ B$), to uznajemy automatycznie, że obiekt B jest gorszy od obiektu A ($A \prec B$). Dzięki temu liczba koniecznych porównań redukuje się do $n(n-1)/2$, gdzie n jest liczbą ocenianych obiektów.

Pomiar spójności preferencji jest uzależniony od przyjętej skali ocen i sprowadza się do wskazania jednoznacznego sposobu ustalenia liczby odstępstw od reguły przechodniości indyferencji i preferencji. Najczęściej spotykanymi metodami pomiaru są: współczynnik konsystencji (*coefficient of consistence*) zaproponowany przez M. Kendalla oraz indeks i wskaźnik spójności (*consistency index, consistency ratio*) opracowane przez T. Saaty'ego. Należy zwrócić uwagę, że wyrażenie *coefficient of consistence* bywa tłumaczone jako współczynnik zgodności². Podobnie wyrażenie *consistency ratio* T. Saaty'ego jako wskaźnik zgodności³. Warto jednak wyraźnie oddzielić pojęcie spójności od pojęcia zgodności ocen, ponieważ w istocie dotyczą one różnych zjawisk. Spójność odzwierciedla bardziej konsekwentne przestrzeganie zasady przechodniości preferencji podmiotu oceniającego, podczas gdy zgodność (koherencja) jest wyrazem zbieżności opinii niezależnych ekspertów (miarą zgodności jest np. współczynnik konkordancji).

Poniżej przedstawimy dwie metody porównań parami. Efektem zastosowania pierwszej metody jest uporządkowanie słabe badanych obiektów, efektem zastosowania drugiej uporządkowanie mocne. W obu przypadkach istnieje potrzeba weryfikacji spójności preferencji.

4. Współczynnik konsystencji Kendalla

Najprostszym sposobem porównań parami jest przypisanie każdej parze obiektów jednej z dwóch liczb (not):

1 – gdy obiekt X jest uznany za ważniejszy od obiektu Y ($X \succ Y$),

0 – gdy obiekt X uznajemy za mniej ważny od obiektu Y ($X \prec Y$).

Efektem tak przeprowadzonej oceny jest mocny porządek rangowy.

Wyniki porównań parami przedstawia się w macierzy. Na przekątnej – ze względu na przyjęty system ocen, który nie uwzględnia indyferencji – znajdują się puste pola. Z uznania stosunku odwrotności preferencji ($X \prec Y \leftrightarrow Y \succ X$) wynika, że ekspert ocenia jedynie pary poniżej (lub powyżej) przekątnej. Jeżeli obiekt w wierszu X uzyskał notę 1 w porównaniu z obiektem w kolumnie Y , to automatycznie obiekt w wierszu Y w zestawieniu z obiektem w kolumnie X otrzymuje notę 0.

Dla dowolnej triady obiektów X, Y, Z wyniki porównań sprowadzalne do sekwencji: $X \succ Y \succ Z \succ X$ lub $X \prec Y \prec Z \prec X$ naruszają zasadę przechodniości preferencji. W każdej niespójnej triadzie spójniki preferencji są zwrócone w jednym kierunku.

Gdy liczba obiektów jest nieparzysta (n), to maksymalna liczba niespójnych triad wynosi $(n^3 - n)/24$. Dla parzystej liczby obiektów istnieje natomiast $(n^3 - 4n)/24$ wszystkich

² Kendall G.M., Buckaland W.R.: Słownik terminów statystycznych, PWE, Warszawa 1986 (tłum. M. Kanton).

³ Trzaskalik T.: Metody wielokryterialne na polskim rynku finansowym, PWE, Warszawa 2006, s. 69.

możliwych niespójnych triad. Minimalna liczba niespójnych triad w obu przypadkach jest równa zero. Na tej podstawie oblicza się współczynnik spójności (*coefficient of consistence*), zwany także współczynnikiem konsystencji⁴.

W celu uniknięcia konieczności zliczania wszystkich niespójnych triad do obliczania konsystencji porównań parami warto wykorzystać następujące formuły⁵:

$$K = \frac{12 \sum_{i=1}^n (R_i - \bar{R})^2}{n^3 - n}, \text{ gdy } n \text{ jest nieparzyste, oraz} \quad (1)$$

$$K = \frac{12 \sum_{i=1}^n (R_i - \bar{R})^2 - 3n}{n^3 - 4n}, \text{ gdy } n \text{ jest parzyste} \quad (2)$$

gdzie:

K – współczynnik konsystencji,

n – liczba porównywanych obiektów (czynników);

R_i – suma ocen dla i -obektu czynnika;

\bar{R} – średnia z sum ocen wszystkich obiektów ($i=1, \dots, n$).

Współczynnik K przyjmuje wartości od 0 do 1, gdzie 0 oznacza całkowity brak spójności porównań, a 1 pełną spójność ocen.

5. Wskaźnik spójności Saaty'ego

Zaproponowany przez T. Saaty'ego wskaźnik spójności opracowany został dla potrzeb metody AHP (Analytic Hierarchy Process), która stanowi rozwiniętą koncepcję definiowania i oceny złożonych systemów. W metodzie AHP elementy systemu porównywane są parami przy użyciu następujących not (a_{ij}):

- 2,3,4,5,6,7,8,9 – gdy obiekt X jest uznany za ważniejszy od obiektu Y ($X \succ Y$);
- 1 – gdy obiekty X i Y uznawane są za równie ważne;
- 1/2, 1/3, 1/4, 1/5, 1/6, 1/7, 1/8, 1/9 – gdy obiekt X jest mniej ważny od obiektu Y ($X \prec Y$).

Wynikiem porównań jest macierz kwadratowa \mathbf{A} , której wyrazy (noty) a_{ij} są liczbami odzwierciedlającymi preferencje decydenta. Z przyjętego systemu not wynika, że na przekątnej macierzy \mathbf{A} znajdują się liczby 1. Zakłada się także stosunek odwrotności preferencji, tj. $a_{ij} = 1/a_{ji}$.

⁴ Kendall M.: Rank correlation methods, Charles Griffin & Company, London 1962, s. 144-161.

⁵ Ferguson G.A., Takane Y.: Analiza statystyczna w psychologii i pedagogice, Wydawnictwo Naukowe PWN, Warszawa 2009, s. 446-448.

Macierz porównań parami \mathbf{A} jest następnie normalizowana. W wyniku normalizacji powstaje macierz \mathbf{B} , której wyrazy obliczymy według poniższej formuły:

$$b_{ij} = \frac{a_{ij}}{\sum_{i=1}^n a_{ij}} \quad (3)$$

Wartość b_{ij} to po prostu stosunek elementu a_{ij} do sumy elementów j -tej kolumny macierzy \mathbf{A} . Średnie arytmetyczne wierszy znormalizowanej macierzy \mathbf{B} są wagami badanych elementów:

$$w_i = \frac{1}{n} \sum_{j=1}^n b_{ij} \quad (4)$$

Macierz porównań parami \mathbf{A} jest w pełni spójna, gdy dla każdego i, j oraz k prawdziwa jest następująca równość: $a_{ik} \cdot a_{kj} = a_{ij}$, gdzie a_{ij} jest oceną porównania elementu i z elementem j .

W macierzy w pełni spójnej $\lambda_{\max} = n$, tj. maksymalna wartość własna tej macierzy jest równa liczbie porównywanych ze sobą elementów. Takie przypadki występują, gdy stosujemy dokładnymi wynikami pomiarów na skali ilorazowej. W metodzie AHP przyjęta jest słabsza skala pomiaru, mająca dodatkowo ograniczoną rozpiętość. Dlatego za spójne przyjmuje się macierze porównań, dla których maksymalna wartość własna jest zbliżona do n .

Różnica $\lambda_{\max} - n$ jest wielkością, na podstawie której konstruuje się wskaźniki spójności porównań. Do oceny spójności T. Saaty proponuje na początku obliczenie indeksu spójności (*consistency index*) według następującej formuły:

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (5)$$

Indeks spójności CI (*consistency index*) rośnie wraz ze wzrostem niespójności oszacowań.

Dla dużej liczby macierzy, o wyrazach będących losowo wygenerowanymi ocenami w przyjętej skali, oblicza się średnią wartość indeksów CI . Średnią wartość indeksów spójności oznaczamy symbolem r . Różni się ona w zależności od wymiarów (stopnia) macierzy \mathbf{A} , czyli liczby porównywanych ze sobą elementów (n). Wyznaczone metodą stymulacji wartości krytyczne indeksów CI pokazano w tabeli 2.

Tabela 2

Średnie wartości losowe indeksów spójności dla macierzy o wymiarach od 2 do 10

Wymiary macierzy (n):	2	3	4	5	6	7	8	9	10
Średnie losowe CI (r):	0	0,52	0,89	1,11	1,25	1,35	1,40	1,45	1,49

Zródło: Saaty T.: Decision making for Leaders: The Analytic Hierarchy Process for decisions in a complex world, University of Pittsburgh, RWS Publications, Pittsburgh 2001, s. 83.

Oceny porównań parami eksperta powinny różnić się znacznie od średniej wartości losowych porównań parami (r). Wyrazem tej różnicy jest wskaźnik spójności CR (*consistency ratio*):

$$CR = \frac{CI}{r} = \frac{\lambda_{\max} - n}{r(n-1)} 100\% \quad (6)$$

Oceny uznawane są za spójne, gdy wskaźnik $CR < 10\%$.

6. Podsumowanie

Wysoka spójność preferencji nie gwarantuje, że uzyskane oceny adekwatnie odzwierciedlają badany stan rzeczy. Pomiar spójności odgrywa jednak ważną rolę podczas weryfikacji wystawionych ocen. W przypadku niskiej spójności proces oceny należy powtórzyć lub zastosować inną metodę pomiaru.

Ogólnie w analizie preferencji najpierw należy badać spójność ocen (o ile jest to możliwe ze względu na przyjętą metodę oceniania), a następnie zgodność (o ile oceny są wystawiane przez co najmniej dwie osoby). Analiza spójności pozwala wyeliminować pomyłki oraz stwierdzić, czy przedmiot badań jest w pełni rozumiany przez ekspertów. Analiza zgodności pozwala natomiast ustalić, czy eksperci stosują właściwie przyjętą skalę oceny oraz czy obiekty są oceniane podobnie ze względu na przyjęte kryteria.

Badanie spójności ocen jest ściśle powiązane ze sposobem oceniania. Identyfikacja odstępstw od reguły przechodniości preferencji jest możliwa w przypadku stosowania metody porównań parami. Obok sposobu oceny wpływ na jakość ocen (w tym ich spójność) ma wiele czynników, m.in. jednoznaczne zdefiniowanie obiektów oceny, dodatkowe informacje pozostające do dyspozycji oceniającego czy miejsce i czas oceny. Z uwagi na naturę problemów strategicznych rozstrzygającym czynnikiem wydaje się być precyzja w definiowaniu zmiennych. Zaprezentowane współczynniki pomiaru spójności mogą być wykorzystywane nie tylko do weryfikacji jakości dokonywanych ocen, lecz także stanowić ważny instrument badania efektywności procesu podejmowania decyzji strategicznych. W szczególności interesujące wydaje się powiązanie miar spójności z miarami zgodności ocen.

Bibliografia

1. Cabała P.: *Podjęcie decyzji w warunkach niepełnej informacji*. Wybrane zagadnienia, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków 2014

2. Cabała P.: Planowanie scenariuszowe w zarządzaniu bezpieczeństwem strategicznym przedsiębiorstwa, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków 2012
3. Ferguson G.A., Takane Y.: Analiza statystyczna w psychologii i pedagogice, Wydawnictwo Naukowe PWN, Warszawa 2009
4. Grzegorzczak A.: Zarys logiki matematycznej, PWN, Warszawa 1969
5. Hurley W.J.: The analytic hierarchy process: a note on an approach to sensitivity which preserves rank order, „Computers & Operations Research”, Vol. 28, 2001
6. Iida Y.: The Number of Circular Triads in a Pairwise Comparison Matrix and a Consistency Test in the AHP, „Journal of Operations Research Society of Japan”, Vol. 52, No. 2, 2009
7. Kendall G.M., Buckland W.R.: Słownik terminów statystycznych, PWE, Warszawa 1986
8. Kendall M.: Rank correlation methods, Charles Griffin & Company, London 1962
9. Michel-Kerjan E., Slovic P. (red.): The Irrational Economist. Making Decisions in a Dangerous World, PublicAffairs, New York 2010
10. Saaty T.: Decision making for Leaders: The Analytic Hierarchy Process for decisions in a complex world, University of Pittsburgh, RWS Publications, Pittsburgh 2001
11. Sharplin A.: Strategic Management, McGraw-Hill, New York 1985
12. Trzaskalik T.: Metody wielokryterialne na polskim rynku finansowym, PWE, Warszawa 2006

Abstract

The article discusses the problem of consistency measurement in the process of strategic decision making. Strategic decisions often require the formulation of uncertain assumptions which are based on subjective evaluations of decision-makers and the strategic choice is the result of analysis of many factors simultaneously. At the beginning of the article a proposal of the classification of assessment systems is characterized, then the concept of consistency and the essence of pairwise comparisons are presented. Consistency is associated with the principle of preference transitivity. Kendall's coefficient of consistence and Saaty's consistency ratio are proposed as the tools of measurement.