

*Budownictwo i Architektura 5 (2009) 61-76*

# Podobieństwo w wycenie nieruchomości

Jacek Zyga

*Katedra Geotechniki, Wydział Budownictwa i Architektury,  
Politechnik Lubelska, e-mail: j.zyga@pollub.pl*

**Streszczenie:** Podobieństwo, mimo iż jest pojęciem powszechnie stosowanym, w praktyce wielu dziedzin (technicznych i nie technicznych) jest pojęciem o niesprecyzowanej definicji, bazującym na potocznym jego pojmowaniu. W praktyce oznacza to brak racjonalnych, opartych na dowodach, ocen porównywania dowolnych obiektów. Artykuł niniejszy podejmuje próbę zdefiniowania podobieństwa i zastosowania wybranych metod jego pomiaru w analizach danych stosowanym w wycenie nieruchomości.

**Słowa kluczowe:** podobieństwo, odległość, odstęp, wycena nieruchomości.

## 1. Definicja podobieństwa obiektów

Podobieństwo jest pojęciem trudnym do zdefiniowania a tym bardziej do zmierzenia.

Założmy, że mamy trzy figury geometryczne (Rys. 1).



Rys. 1. Przykłady figur podobnych.

Fig. 1 Examples of similar figures.

Kwadraty  $A$ ,  $B$  są tej samej wielkości a kwadraty  $A$  i  $C$  mają ten sam kolor. Które zatem z par kwadratów są bardziej lub mniej podobne?

W odniesieniu do figur i brył geometrycznych istotnym czynnikiem podobieństwa jest zachowanie proporcji odpowiednich elementów liniowych przy zachowaniu identyczności odpowiednich kątów obu przyrównywanych kształtów [11]. W przypadku obiektów charakteryzowanych innymi cechami podobieństwo staje się pojęciem mniej konkretnym. Praca [14] przez określenie „podobny” oznacza na przykład tyle co: „mający pewne cechy zgodne z cechami innego obiektu; przypominający kogoś, coś (...); prawie jednakowy” a także „tego samego typu, rodzaju”.

Podobieństwo jest wielkością, która odzwierciedla siłę i ilość relacji zachodzących pomiędzy dwoma przedmiotami i ich cechami. Matematycznie może być ono opisane skalą od -1 do +1 albo od 0 do 1. Podobieństwo  $S_{AB}$  między obiektem  $A$  oraz obiektem  $B$  możemy mierzyć kilkoma sposobami, zależnie od obranej skali pomiaru i posiadanych danych.

W praktyce pomiarowi często podlega „niepodobieństwo” czyli „odmienność” rozumiana jako przeciwieństwo „podobieństwa”:

$$d_{AB} = 1 - S_{AB} \text{ przy } S_{AB} \in <0,1>, \quad (1a)$$

lub

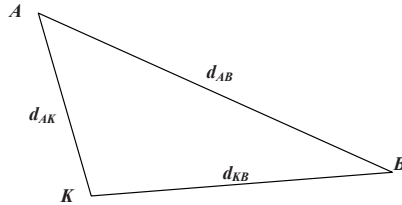
$$2 d_{AB} = 1 - S_{AB} \text{ przy } S_{AB} \in <-1,1>. \quad (1b)$$

Miarą odmienności jest z kolei rozbieżność ocen porównywanych przedmiotów, sformułowanych według zdefiniowanych wcześniej skal.

Do wyznaczenia rozbieżności ocen porównywanych przedmiotów niezbędne jest określenie:

- cech poddanych ocenie,
- zespołów rang jakimi formułowane mają być oceny,
- oraz uszeregowanie tych rang.

Zabieg ten tworzy wielowymiarową przestrzeń (metryczną lub niemetryczną), w której każdemu z porównywanych przedmiotów przypisana jest konkretna pozycja. Rozbieżność ocen porównywanych przedmiotów może być wyznaczona jako „odległość” pomiędzy tymi pozycjami (Rys. 2).



Rys. 2. Graficzne przedstawienie pojęcia „odległość”.

Fig. 2. Graphical representation of the concept of “the distance”.

Odległość  $d$  (nazywana też odstępem) jest wielkością spełniającą co najmniej trzy poniższe warunki:

- jest większa lub równa zero ( $d_{AB} \geq 0$ ),
- jest równa zero wtedy i tylko wtedy gdy obiekt  $A$  porównywany jest sam z sobą ( $d_{AA} = 0$ ),
- jest symetryczna ( $d_{AB} = d_{BA}$ ),

Gdy spełnia również warunek ( $d_{AB} \leq d_{AK} + d_{KB}$ ) jest nazywana metryką.

Ostatecznie podobieństwo lub odmiennosc dwóch wybranych obiektów, opisanych zbiorami pewnych cech, mogą być przedstawione za pomocą zagregowanego współczynnika, na który składają się oceny rozbieżności (odległości między ocenami analogicznych cech) obu porównywanych obiektów.

## 2. Wybrane metody pomiaru podobieństw i odmienności

Metodyka określania podobieństwa lub odmienności pomiędzy dwoma obiektami jest zróżnicowana ze względu na rodzaj cech poddawanych porównaniom bezpośrednim. Oddzielne sposoby obliczania współczynników podobieństwa  $S_{AB}$  (oraz odmienności (odstępu)  $d_{AB}$ ) przewiduje się dla:

- obiektów o cechach o charakterze binarnym,
- obiektów o cechach skategoryzowanych (bez ich wartościowania),

- obiektów o cechach o charakterze porządkowym,
- obiektów o cechach o charakterze jakościowym.

W porównywaniu obiektów opisanych cechami o charakterze binarnym (jest/brak, 1/0, tak/nie) stosowane są cztery podstawowe parametry porównawcze:  $p$  – liczba przypadków ocen pozytywnych dla obu obiektów,  $q$  – liczba przypadków gdy dla obiektu  $A$  ocena była pozytywna a dla obiektu  $B$  ocena była negatywna,  $r$  – liczba przypadków gdy dla obiektu  $B$  ocena była pozytywna a dla obiektu  $A$  ocena była negatywna,  $s$  – liczba przypadków ocen negatywnych dla obu obiektów.

Tabela 1. Zestawienie ocen obiektów A i B.

Table 1. Evaluation statement of objects A and B.

		Obiekt B	
		Ocena pozytywna	Ocena negatywna
Obiekt A	Ocena pozytywna	$p$	$q$
	Ocena negatywna	$r$	$s$

Na bazie powyższych parametrów różni autorzy kreują szeroki wachlarz współczynników podobieństwa/odmienności, z których najbardziej znane to:

- współczynnik prostego dopasowania (*simple matching coefficient*):

$$S_{AB} = (p+s)/t, \text{ gdzie: } t = p + q + r + s, \quad (2a)$$

$$d_{AB} = 1 - S_{AB} = (q+r)/t, \quad (2b)$$

- współczynnik Jaccarda (*Jaccard's coefficient*):

$$S_{AB} = p/(p+q+r), \quad (3a)$$

- odległość Jaccarda (*Jaccard's distance*):

$$d_{AB} = (q+r)/(p+q+r). \quad (3b)$$

Należałoby w powyższym zestawieniu ująć także odległość Hamminga a także stanowiącą jej uogólnienie odległość Leveshteina, które oceniają liczbę różnic (odmienności) z ciągach znaków o tej samej (lub w ogólności różnej) długości. Z uwagi na typowo informatyczne zastosowanie tych metod oceny podobieństwa w niniejszym opracowaniu poprzestaje się na przywołaniu powyższych metod.

Ocena podobieństwa lub odmienności obiektów opisanych zbiorami różnych cech (bez ich wartościowania) wymaga albo

- wstępnego przetworzenia opisów na zespół cech binarnych i analizowania ich według reguł opisanych powyżej,
- albo identyfikacji grupy wspólnych cech z obu zbiorów (iloczyn zbiorów cech  $A$  i  $B$ ) i oceny jego wielkości w proporcji do zbioru wszystkich cech (suma zbiorów cech  $A$  i  $B$ ), matematyczny opis tej proporcji ma postać indeksu Jaccarda:

$$s_{AB} = J(A, B) = \frac{|A \cap B|}{|A \cup B|}. \quad (4a)$$

Dopełnieniem współczynnika podobieństwa  $S_{AB}$  jest odległość Jaccarda (*Jaccard's distance*) obliczany według formuły:

$$d_{AB} = J_{\delta}(A, B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}, \quad (4b)$$

- pewną analogią indeksu Jaccarda jest współczynnik Sørensen

$$s_{AB} = QS = \frac{2 \cdot C}{A + B}, \quad (5)$$

gdzie  $A$  i  $B$  to liczby elementów w poszczególnych zbiorach a  $C$  to liczba elementów występujących w obu zbiorach równocześnie

Dla oceny podobieństw (lub raczej odmienności) obiektów o cechach o charakterze porządkowym możliwe są dwa typowe toki postępowania:

- W przypadku opisanego cech porównywanych obiektów wielkościami posiadającymi interpretację liczbową można wykorzystać ich wielkości nieznormalizowane. Otrzymane jednak charakterystyki podobieństwa będą wtedy wielkościami o charakterze względnym, których interpretacja wymaga uwzględnienia skali zastosowanych ocen analizowanych cech.
- W celu zobiektywizowania charakterystyk podobieństwa można jednak dokonać normalizacji ocen poszczególnych cech obiektów. Możliwe jest to jednak wyłącznie dla cech o jednakowych mianach lub cech opisanych rangami należącymi do jednej skali rang. Istotą normalizacji jest założenie, że każda z ocen może zostać przedstawiona jako zmienna ilościowa, co pozwala na przetworzenie każdego z szeregow ocen, rozciągniętych na zróżnicowanych skalach, do zbioru liczb z przedziału  $\langle 0, 1 \rangle$ . W przypadku zdefiniowania ocen w postaci rang o ustalonym interwale zastosowanie znajduje tzw. „znormalizowana transformacja rangowa” (*Normalized Rank Transformation*), w której w pierwszej kolejności dokonuje się zamiany rang na szereg liczbowy (najczęściej liczb naturalnych) by w drugiej kolejności zamieć go na szereg ocen z zakresu  $\langle 0, 1 \rangle$ .

Zidentyfikowane oceny cech porównywanych obiektów (znormalizowanych lub nie) stanowią współrzędne w  $n$ -wymiarowym układzie ujednoczonych współrzędnych, co pozwala na wyznaczenie jako miary ich odmienności tzw. odległości (odstępu) dwóch punktów zdefiniowanych wskazanymi współrzędnymi. Wykorzystać tu można kilka typów odległości opisanych w literaturze pod następującymi nazwami:

- Odległość Euklidesowa (*Euclidean distance*),
- Odległość Spearmana (*Spearman Distance*),
- Odległość Minkowskiego (*Minkowski Distance*),
- Odległość Czebyszewa (*Chebyshev /Maximum Distance*),
- Odległość Kendalla (*Kendall Distance*),
- Odległość Cayleya (*Cayley Distance*),
- Odległość Ulama (*Ulam Distance*).

Najprostsza w interpretacji jest odległość euklidesowa mająca swoje podstawowe zastosowanie z geometrii (dla  $n = 2$ ):

$$d_{AB}^{eu} = \sqrt{\sum_{i=1}^n (x_{Bi} - x_{Ai})^2}, \quad (6)$$

gdzie:  $x_{Bi}$  – oznacza ocenę  $i$  – tej cechy odpowiednio dla obiektu  $B$ ,  $x_{Ai}$  – oznacza ocenę  $i$  – tej cechy odpowiednio dla obiektu  $A$ ,  $n$  – oznacza liczbę cech, a jednocześnie liczbę wymiarów przestrzeni.

Analogiczną oceną odmienności jest odległość Spearmana, która jest kwadratem odległości euklidesowej

$$d_{BA}^{Sp} = (d_{AB}^{eu})^2 = \sum_{i=1}^n (x_{Bi} - x_{Ai})^2. \quad (7)$$

Podobną oceną jest formuła (*Spearman Footrule Distance*) nazywana także metryką miejską (metryką Manhattan)

$$d_{AB}^{SpF} = \sum_{i=1}^n |x_{Bi} - x_{Ai}|. \quad (8)$$

Powyższe miary odległości stanowią szczególne przypadki tzw. odległości Minkowskiego, formułowanej jako uogólniona miara odległości między dwoma punktami w przestrzeni  $n$ -wymiarowej [18]. Stosując wcześniejszą metodę zapisu formuła odległości Minkowskiego przyjmuje kształt:

$$d_{AB}^m = L_m(x_B, x_A) = \left( \sum_{i=1}^n |x_{Bi} - x_{Ai}|^m \right)^{\frac{1}{m}}, \quad (9)$$

gdzie:  $x_{Bi}$  – oznacza ocenę  $a$  – tej cechy odpowiednio dla obiektu  $B$ ,  $x_{Ai}$  – oznacza ocenę  $a$  – tej cechy odpowiednio dla obiektu  $A$ ,  $n$  – oznacza liczbę cech,  $m$  – oznacza liczbę całkowitą  $m \in \langle 1, \infty \rangle$ .

Kolejnym specjalnym przypadkiem odległości Minkowskiego jest odległość Czebyszewa mająca postać:

$$d_{AB}^{ch} = \max_i |x_{Bi} - x_{Ai}| = \lim_{m \rightarrow \infty} \left( \sum_{i=1}^n |x_{Bi} - x_{Ai}|^m \right)^{\frac{1}{m}}. \quad (10)$$

W praktyce określa największą z możliwych odległości pomiędzy punktami wskazanymi przez współrzędne  $x$  w przestrzeni  $R^n$ . W szachach (przykład zagadnienia odległości na płaszczyźnie) jest to odległość między polami szachownicy wyrażona w ruchach, które musi wykonać figura króla. Stąd pochodzi jej angielska nazwa „*chessboard distance*” [19].

Omówione wyżej odległości, można zastosować do oceny podobieństw lub raczej odmienności obiektów o cechach wyrażonych w sposób jakościowy. Podobne zastosowanie mogą mieć również takie mierniki (opisywane poniżej) jak:

- Odległości Bray’a Curtisa,
- Odległości Canberry,
- Współczynniki korelacji.

Do identyfikacji odmienności w szeregach cech lub znaków stosowane są z kolei tzw. odległości Kendalla, Ulama oraz Cayleya. Sprowadzają się do wyznaczenia minimalnej liczby kombinacji umożliwiających zamianę jednego uszeregowania na drugie. Wymienione metody wyznaczają odległości pomiędzy obiektami (ciągami ocen) w postaci liczb całkowitych i zawierają się w przedziale  $\langle 0, \infty \rangle$ .

Cechą wspólną badania podobieństw i odmienności obiektów o cechach wyrażonych w sposób jakościowy jest zakładana różnorodność kategorii i jednostek w jakich dokonywane są oceny porównywanych cech. Aktualna pozostaje zatem uwaga o konieczności starannej interpretacji wyników „odległości” porównywanych obiektów wyznaczanych na podstawie odległości Minkowskiego i jej przypadków szczególnych (Odległości Spearmana, Czebyszewa, euklidesowej). Drogą zobiektywizowania dokonywanych ocen odmienności może być, tak jak w przypadku obiektów o cechach o charakterze porządkowym, normalizacja ocen w zakresie poszczególnych cech ale również konieczna jest modyfikacja wzorów rozwiązujących. Jednak nawet po normalizacji, w przypadkach oceny obiektów o skrajnych ocenach cech porównawczych, wyznaczone odległości będą przybierać wartości większe niż 1 ( $d > 1$ ). Są one bowiem funkcją rzędu przestrzeni w jakiej są wyznaczone. Innymi słowy zależą od liczby cech jakimi opisane są obiekty. Dlatego w celu sprowadzenia oszacowań odległości do przedziału  $\langle 0, 1 \rangle$  należy zasugerować znormalizowanie przy użyciu czynnika maksymalnej odległości możliwej do wyznaczenia w danej przestrzeni. Dla szczególnych przypadków odległości sugeruje się następujące postaci wzorów:

- odległość euklidesowa (6) po znormalizowaniu przyjmie postać:

$$d_{AB}^{eu} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{Bi} - x_{Ai})^2}, \quad (11)$$

- odległość Spearmana po znormalizowaniu przyjmie postać:

$$d_{BA}^{Sp} = \frac{1}{n} \sum_{i=1}^n (x_{Bi} - x_{Ai})^2. \quad (12)$$

Problem nieznormalizowanych ocen występuje także w przypadku wyznaczenia odległości Bray’a Curtisa. Jakkolwiek odległość ta

$$d_{AB}^{bc} = \frac{\sum_{i=1}^n |x_{Bi} - x_{Ai}|}{\sum_{i=1}^n (x_{Bi} + x_{Ai})} = \frac{\sum_{i=1}^n |x_{Bi} - x_{Ai}|}{\sum_{i=1}^n (x_{Bi}) + \sum_{i=1}^n (x_{Ai})} \quad (13)$$

ma postać ilorazu (dzielenie różnic ocen poszczególnych cech obiektów przez ich sumę) to wynik maksymalny odległości w przestrzeni współrzędnych nieznormalizowanych nie osiąga wartości 1. Normalizacja ocen powoduje jednak, że wyniki przyjmują zakres zmienności  $\langle 0, 1 \rangle$ . Należy jednak pamiętać, że poprzez normalizację definiowane są współrzędne o wartości „0” przez co w niektórych wypadkach określenie odległości może być matematycznie niemożliwe.

Odległość Canberra stanowi wariant powyższej metody. Wyznacza ją suma ułamkowych różnic odpowiednich ocen w parach obiektów:

$$d_{AB}^{can} = \sum_i^n \frac{|x_{Bi} - x_{Ai}|}{|x_{Bi}| + |x_{Ai}|}. \quad (14)$$

Odległość ta posiada praktyczne cechy podobne do cech odległości Bray'a Curtisa. W jej przypadku jednak normalizacja ocen  $x$  nie dokonuje transformacji wyników rachunku odległości na przedział  $\langle 0, 1 \rangle$ . Podobnie jak w przypadku odległości Spearmana proponuje się modyfikację wzoru (14) do postaci

$$d_{AB}^{can} = \frac{1}{n} \sum_i^n \frac{|x_{Bi} - x_{Ai}|}{|x_{Bi}| + |x_{Ai}|}, \quad (15)$$

zapewniającej znormalizowane oszacowania odległości porównywanych obiektów. Tak jak w przypadku formuły (13) zastosowanie wzorów (14) i (15) z wykorzystaniem ocen znormalizowanych może prowadzić jednak do nieokreśloności wyników. Przy zestawieniu ocen minimalnych mianownik powyższych wzorów jest równy zero.

### 3. Pojęcie podobieństwa w odniesieniu do nieruchomości

W procedurach wyceny nieruchomości, zarówno ze względów merytorycznych jak też prawnych, kwestia podobieństwa pełni ważną rolę. Jest na przykład zasadniczą przesłanką umożliwiającą grupowanie obiektów, co pozwala z kolei ma określanie przedmiotowe i terytorialne poszczególnych segmentów rynku. Poprzez ocenę podobieństwa każda nieruchomość jest przyporządkowywana do danego segmentu a także poddawana porównaniom. Nawet w procedurach wyceny w podejściu kosztowym kwestia podobieństwa odgrywa kluczową rolę. Określenie kosztu wykonania wskazanego zakresu robót budowlanych, całego budynku czy jego tzw. scalonego elementu wymaga doboru właściwych wskaźników cenowych. Właściwość tych wskaźników jest natomiast wprost proporcjonalna do podobieństwa wycenianego elementu (lub obiektu czy konkretnego zakresu robót) do analogicznego elementu cennikowego.

Mimo trudnej do przecenienia roli „podobieństwa” szeroko rozumiane prawo opiera się na jego potocznej definicji. Jedyną wyrażoną wprost definicją z tego zakresu jest definicja zawarta w artykule 4 Ustawy z dnia 21 sierpnia 1997 r. o gospodarce nieruchomościami [16] określająca pojęcie podobieństwa na potrzeby gospodarki nieruchomościami i ich wyceny. Według treści powołanego przepisu przez nieruchomość podobną „*należy (...) rozumieć nieruchomość, która jest porównywalna z nieruchomością stanowiącą przedmiot wyceny, ze względu na położenie, stan prawny, przeznaczenie, sposób korzystania oraz inne cechy wpływające na jej wartość*”.

Wobec definicji słownikowej przymiotnika „porównywalny” jak i czasownika „porównywać” („porównywalny” to taki, który może być z czymś porównywany; „porównywać” to „zestawiając jakieś zjawiska, rzeczy, osoby zwracać uwagę na cechy wspólne i cechy różne”) wyłania się jednak tautologia tej definicji. Określona bowiem w ustawowej definicji „*nieruchomość podobna*” jest „*nieruchomością, która może być zestawiona z nieruchomością wycenianą (przyrównana do niej) ze zwróceniem uwagi na ich cechy wspólne (podobieństwa) i cechy różne (odmienności)*”. Zestawieniu takiemu może podlegać zatem nieruchomość przedmiotowa (wyceniana)



i dowolna inna. Nie przesądza to jednak o ich wzajemnym podobieństwie. W definicji ustawowej brakuje bowiem kryteriów co począć z wynikami dokonanego zestawienia.

Pozytywną odpowiedzią na dostrzeżone braki może być definicja podobieństwa oparta wprost na definicji pojęcia „podobny” ze „Słownika współczesnego języka polskiego” (przytoczonej wyżej). Transponując tę definicję można powiedzieć, że *„nieruchomość podobna” to nieruchomość mająca pewne cechy zgodne z cechami nieruchomości wycenianej; przypominająca tę nieruchomość pod pewnymi względami, tego samego typu lub rodzaju, prawie z nią jednakowa*”. Jest to oczywiście definicja, która też wymaga uściśleń. Na potrzeby różnych obszarów wyceny nieruchomości może i powinna być ona doprecyzowywana. Odnosi się ona jednak do zasadniczych cech pojęcia podobieństwa i umożliwia kreowanie kryteriów jego oceny.

#### 4. Jak oceniać podobieństwo nieruchomości?

W obszarze rozważań dotyczących wyceny nieruchomości można wyróżnić kilka zagadnień, w których występuje ocena podobieństwa:

- dokonujący wyceny mimowolnie posługuje się nią w procesie budowy bazy danych reprezentującej konkretny rynek lokalny (zarówno w podejściu porównawczym jak i dochodowym), w odniesieniu do samych nieruchomości, ich cech fizycznych, ekonomicznych etc.,
- używana jest w procedurach podejścia porównawczego (zwłaszcza przy porównywaniu parami),
- w podejściu dochodowym używana jest także przy ocenie rentowności i poszczególnych składowych ryzyka związanego z inwestowaniem,
- w podejściu kosztowym towarzyszy wyborowi elementów cennikowych.

Do oceny podobieństwa w poszczególnych zakresach zastosowań wystarczą najprostsze z opisywanych powyżej metod:

Do oceny przynależności poszczególnych nieruchomości do określonego (zawężonego) rynku lokalnego przydatne są procedury przystosowane do porównywania obiektów opisanych cechami o charakterze binarnym. Taka jest bowiem idea klasyfikowania tych obiektów (albo spełniają określony warunek albo nie). Przy zastosowaniu dowolnego opisu cechy wzorca danego rynku stosownym narzędziem oceny wzajemnych podobieństw będzie współczynnik prostego dopasowania (2a). Wygodniejszy wydaje się jednak taki sposób definiowania cech nieruchomości na rynku lokalnym (cech wzorcowych), który by wymuszał stosowanie wyłącznie pozytywnych wartości w ocenach cech, które mają być pozytywnie zweryfikowane. Wtedy oceny podobieństwa można dokonywać poprzez współczynnik Jaccarda (3a).

Zastosowanie któregośkolwiek współczynnika wymaga jednak:

- identyfikacji cech poszczególnych nieruchomości,
- wyboru wiodących cech rynkowych (czynników cenotwórczych, szczególnych dla wybranego rynku lokalnego),
- sformułowania warunków, których spełnienia wymaga pozytywna ocena cechy,
- binarnej oceny cech poszczególnych nieruchomości.



Przy zastosowaniu współczynnika Jaccarda proporcja liczby cech wspólnych (cech nieruchomości porównywanej i zbioru wiodącego ocenionych jednocześnie pozytywnie) do liczby wszystkich cech opisujących rynek stanowi informację jak bardzo poszczególne nieruchomości przystaje (jest podobna) do opisanego wybranymi cechami rynku lokalnego. Z uwagi na zdefiniowanie rynku lokalnego wyłącznie pozytywnymi ocenami czynnik  $r$  przyjmie wartość 0 (zero) zatem współczynnik podobieństwa wyniesie

$$S_{AB} = p/(p+q) \quad (16)$$

Zastosowanie współczynnika prostego dopasowania (2a) jakkolwiek wymagające zwiększonej uwagi w definiowaniu ocen binarnych prowadzi do takich samych wniosków jak zastosowanie współczynnika Jaccarda.

Oba powyższe mierniki podobieństwa są bardzo rygorystyczne. Dzięki temu, że rozróżniają wyłącznie zgodność lub brak zgodności pewnych cech z zadaniem wzorcem, potrafią wyeliminować przypadki niepełnego dopasowania do zadanego układu cech uznanych wstępnie za istotne. Za pozytywny wynik testu traktowana będzie bowiem wyłącznie wartość  $S_{AB} = 1$ .

#### Przykład 1:

Poszukiwany jest rynek lokali mieszkalnych o następujących cechach:

1. Nieruchomości lokalowych (stanowiących przedmiot odrębnej własności),
2. Położonych w dzielnicy A,
3. Położonych w niskiej zabudowie wielorodzinnej (do 5-io kondygnacji),
4. Nie remontowanych (standard przeciętny lub pogorszony).

Ocenie poddane zostają:

lokal I:

1. Stanowiący nieruchomość lokalową,
2. Położony w dzielnicy A,
3. Położony w budynku 5-io kondygnacyjnym,
4. Po remoncie (wymiana stolarki),
5. Okna lokalu wychodzą na park,,
6. Położony w pobliżu przystanku autobusowego, z dużym parkingiem.

lokal II:

1. Stanowiący nieruchomość lokalową,
2. Położony w dzielnicy A,
3. Położony w budynku 11-to kondygnacyjnym,
4. Nie remontowany,
5. Okna lokalu wychodzą na ulicę,
6. położony w pobliżu przystanku autobusowego, z dużym parkingiem.

lokal III:

1. Stanowiący nieruchomość lokalową,
2. Położony w dzielnicy B,
3. Położony w budynku 5-io kondygnacyjnym,
4. Po remoncie (wymiana stolarki),
5. Okna lokalu wychodzą na park,
6. Położony daleko od przystanku autobusowego, brak parkingu.

Cechy nr 5 i 6 jako nie istotne w ocenie przynależności do zdefiniowanego rynku nie są brane pod uwagę. Liczba wszystkich ocenianych cech wynosi zatem 4.

Tabela 2. Oceny binarnej zgodności cech.  
Table 2. Evaluation of binary features accordance.

	Cechy rynku	lokal I	lokal II	lokal III
Nieruchomości lokalowa (stanowiąca przedmiot odrębnej własności)	1	1	1	1
lokal położony w dzielnicy A	1	1	1	0
lokal położony w niskiej zabudowie wielorodzinnej (do 5-iu kondygnacji)	1	1	0	1
Lokal nie remontowany (standard przeciętny lub pogorszony)	1	0	1	0
$p$		3	3	2
$q$		1	1	2
$p+g$		4	4	4
Skorygowany współczynnik Jaccarda		0,75	0,75	0,5

Wnioski:

- Żaden z ocenianych lokali nie należy w pełni do wskazanego rynku lokalnego.
- Lokale nr I i II są bardziej podobne do wzorca rynku lokalnego niż lokal nr III.

W tej sytuacji można rozszerzyć rynek poprzez ograniczenie liczby cech wiodących (na przykład pominięcie cechy nr 4 – „lokal nie remontowany”).

Tabela 3. Wtórne oceny binarnej zgodności cech.  
Table 3. Secondary evaluation of binary features accordance.

	Cechy rynku	lokal I	lokal II	lokal III
Nieruchomości lokalowa (stanowiąca przedmiot odrębnej własności)	1	1	1	1
lokal położony w dzielnicy A	1	1	1	0
lokal położony w niskiej zabudowie wielorodzinnej (do 5-iu kondygnacji)	1	1	0	1
$p$		3	2	2
$q$		0	1	1
$p+g$		3	3	3
Współczynnik podobieństwa (przynależności do zdefiniowanego rynku lokalnego)		1	0,67	0,67

Wnioski ponowne:

- Lokal nr I należy do wskazanego (opisanego trzema cechami wiodącymi) rynku lokalnego.
- Lokale nr III i II są tak samo podobne (lub raczej nie podobne) do wzorca rynku lokalnego.

Przykład 2:

Poszukiwany jest ten sam rynek lokali mieszkalnych co w przykładzie 1, opisany cechami analogicznymi do obiektów porównywanych, inaczej zdefiniowano jednak cechę 3 i 4:

- Nieruchomość lokalowa (lokal stanowiący przedmiot odrębnej własności),
- Lokal położony w dzielnicy A,
- Lokal NIE położony w zabudowie wysokiej wielorodzinnej (powyżej 5-iu kondygnacji),
- Lokal NIE wyremontowany (standard przeciętny lub pogorszony),

(obecność czynnika negującego w ocenach cech 3 i 4) zasygnalizowano poniżej ocenami „0”).

Tabela 4. Oceny binarnej zgodności poszerzonego zbioru cech.

Table 4. Evaluation of binary accordance of extended set of features.

	Cechy rynku	lokal I	lokal II	lokal III
Nieruchomości lokalowa (stanowiąca przedmiot odrębnej własności),	1	1	1	1
lokal położony w dzielnicy A	1	1	1	0
lokal położony w zabudowie wysokiej wielorodzinnej (powyżej 5-iu kondygnacji),	0	0	1	0
Lokal remontowany	0	1	0	1
$p$		2	2	1
$q$		0	0	1
$r$		1	1	1
$s$		1	1	1
$p+s$		3	3	2
$p+g+r+s$		4	4	4
Prosty współczynnik dopasowania		0,75	0,75	0,5

Wnioski analogiczne jak w przykładzie 1.

Dla rozszerzonego rynku przy ograniczeniu liczby cech wiodących (pominięcie cechy nr 4).

Tabela 5. Wtórne oceny binarnej zgodności.

Table 5. Secondary evaluation of binary accordance of extended set of features.

	Cechy rynku	lokal I	lokal II	lokal III
Nieruchomości lokalowa (stanowiąca przedmiot odrębnej własności),	1	1	1	1
lokal położony w dzielnicy A	1	1	1	0
lokal położony w zabudowie wysokiej wielorodzinnej (powyżej 5-iu kondygnacji),	0	0	1	0
$p$		2	2	1
$q$		0	0	1
$r$		0	1	0
$s$		1	0	1
$p+s$		3	2	2
$p+g+r+s$		3	3	3
Prosty współczynnik dopasowania		1	0,67	0,67

Wnioski analogiczne jak w przykładzie 1

Współczynniki dopasowania (prosty oraz Jaccarda) mogą być też użyte w analizach podobieństwa odrębnych rynków (porównanie spełnienia pewnych warunków przez dwie populacje) lub nieruchomości w parach. Otrzymane w tych analizach miary dopasowania (podobieństwa) nie muszą być już oceniane tak rygorystycznie. Stanowią bowiem czystą informację. Informacja ta podlega oczywiście ocenie, jednak kryteria jej akceptowalności mogą być przedmiotem arbitralnych założeń.

Zakwalifikowanie pewnych nieruchomości do próby reprezentatywnej nie oznacza, że dysponujemy materiałem jednorodnym pod względem jakościowym. Kolejnym etapem oceny podobieństwa (nazwijmy je wewnętrznym) w grupie obiektów zgromadzonych w próbie rynkowej może być badanie natężenia cech przyjętych za najważniejsze przy jednoczesnej kontroli założeń poczynionych w tym zakresie. W ogólności może bowiem okazać się, że cechy wybrane jako wiodące przy głębszej analizie nie będą cechami różnicującymi grupę. Nie oznacza to oczywiście odrzucenia cechy jako tej identyfikującej rynek ale może mieć wpływ na uznanie, iż jej waga (w rozumieniu stosowanym w procedurach podejścia porównawczego) jest zaniedbywalna.

Narzędziem przydatnym na tym etapie mogą być odległości  $d$  wyznaczone według zmodyfikowanych wzorów (11), (12), (15) lub wzoru (13), łatwo modyfikowalne na współczynniki podobieństwa  $S$  według formuły (1a).

Oszacowanie tych odległości lub współczynników podobieństwa metodą „każdy z każdym” umożliwi zestawienie wyników w tablicy kwadratowej (o jednakowej liczbie kolumn i wierszy), która może stanowić numeryczną podstawę wnioskowania, które z przyjętych do analizy nieruchomości są do siebie podobne i czy to podobieństwo jest istotne. Ilustruje to poniższy przykład.

### Przykład 3:

Do analizy przyjęto pięć obiektów (lokale A, B, C, D, E) wyłonione z szerokiego rynku jako spełniające warunki analogiczne jak w przykładach 1 i 2. Spełnienie tych warunków oznacza, że w każdy z lokali posiada pozytywną ocenę cech 1-4

i dysponuje różnymi ocenami w innych zakresach cech. Dla uproszczenia przyjęto, że są to cechy 5 i 6 z przykładu 1.

Tabela 6. Opis cech porównywanych lokali.  
Table 6. Description of features of compared flats.

		Zakres ocen cechy
Cecha 1	lokalowa jest przedmiotem odrębnej własności	0, 10
Cecha 2	lokal położony w dzielnicy A (oceny opisują jakość położenia względem np. centrum dzielnicy)	20,30,40,50
Cecha 3	lokal położony w niskiej zabudowie wielorodzinnej (oceny opisują wiek i stan budynków)	30, 40
Cecha 4	lokal nieremontowany (oceny opisują stan techniczny lokalu)	20, 30, 40
Cecha 5	okna lokalu wychodzą na park	0, 10
Cecha 6	bliskość przystanku komunikacji miejskiej	20, 30, 40

Tabela 7. Oceny cech lokali z próby rynkowej.  
Table 7. Feature evaluations of flats from data set.

	lokal A	lokal B	lokal C	lokal D	lokal E
Cecha 1	10	10	10	0	0
Cecha 2	50	40	30	20	20
Cecha 3	40	40	30	30	30
Cecha 4	40	40	30	30	20
Cecha 5	10	10	0	0	0
Cecha 6	40	40	30	30	20

Dla odróżnienia od poniższych ocen szeregujących oceny pierwotne (opisane powyżej) opisano rangami wyrażonymi krotnościami liczby 10.

Tabela 8. Oceny porządkowe lokali z próby rynkowej.  
Table 8. Ordinal feature evaluations of flats from data set.

	lokal A	lokal B	lokal C	lokal D	lokal E
Cecha 1	2	2	2	1	1
Cecha 2	4	3	2	1	1
Cecha 3	2	2	1	1	1
Cecha 4	3	3	2	2	1
Cecha 5	2	2	1	1	1
Cecha 6	3	3	2	2	1

Tabela 9. Znormalizowane oceny porządkowe lokali z próby rynkowej.  
Table 9. Normalized ordinal feature evaluations of flats from data set.

	lokal A	lokal B	lokal C	lokal D	lokal E
Cecha 1	1	1	1	0	0
Cecha 2	1	0,666667	0,333333	0	0
Cecha 3	1	1	0	0	0
Cecha 4	1	1	0,5	0,5	0
Cecha 5	1	1	0	0	0
Cecha 6	1	1	0,5	0,5	0

Tabela 10. Oszacowanie odległości euklidesowej według formuły (11) w oparciu o oceny znormalizowane.

Table 10. Estimation of Euclidean distance (eq. 11) and normalized evaluations.

	lokal A	lokal B	lokal C	lokal D	lokal E
lokal A	0,00	0,14	0,70	0,87	1,00
lokal B	0,14	0,00	0,66	0,81	0,95
lokal C	0,70	0,66	0,00	0,43	0,52
lokal D	0,87	0,81	0,43	0,00	0,29
lokal E	1,00	0,95	0,52	0,29	0,00

Tabela 11. Oszacowanie odległości Spearmana według formuły (12) w oparciu o oceny znormalizowane.

Table 11. Estimation of Spearman distance (eq. 12) and normalized evaluations.

	lokal A	lokal B	lokal C	lokal D	lokal E
lokal A	0,00	0,02	0,49	0,75	1,00
lokal B	0,02	0,00	0,44	0,66	0,91
lokal C	0,49	0,44	0,00	0,19	0,27
lokal D	0,75	0,66	0,19	0,00	0,08
lokal E	1,00	0,91	0,27	0,08	0,00

Tabela 12. Oszacowanie odległości Bray'a Curtisa według formuły (13) w oparciu o oceny znormalizowane.

Table 12. Estimation of Bray-Curtis distance (eq. 13) and normalized evaluations.

	lokal A	Lokal B	lokal C	lokal D	lokal E
lokal A	0,00	0,03	0,44	0,71	1,00
lokal B	0,03	0,00	0,42	0,70	1,00
lokal C	0,44	0,42	0,00	0,40	1,00
lokal D	0,71	0,70	0,40	0,00	1,00
lokal E	1,00	1,00	1,00	1,00	---*

\*/określenie odległości „pozycji” najgorszego lokalu względem siebie samego jest niemożliwe na podstawie współrzędnych znormalizowanych.

Tabela 13. Oszacowanie odległości Canberra według formuły (15) w oparciu o oceny znormalizowane.

Table 13. Estimation of Canberra distance (eq. 15) and normalized evaluations.

	lokal A	Lokal B	lokal C	lokal D	lokal E
lokal A	0,00	0,03	0,53	0,78	1,00
lokal B	0,03	0,00	0,50	0,78	1,00
lokal C	0,53	0,50	---*	---*	---*
lokal D	0,78	0,78	---*	---*	---*
lokal E	1,00	1,00	---*	---*	---*

\*/określenie odległości jest numerycznie niemożliwe, bo w niektórych przypadkach odległości Canberra nie mogą być liczone ze współrzędnych znormalizowanych (przy zestawieniu ocen minimalnych mianowników wzorów 14 i 15 jest równy zero).

Tabela 14. Oszacowanie odległości Bray'a Curtisa według formuły (13) w oparciu o oceny nie normalizowane.

Table 14. Estimation of Bray-Curtis distance (eq. 13) and non-normalized evaluations.

	lokal A	Lokal B	lokal C	lokal D	lokal E
lokal A	0,00	0,03	0,23	0,33	0,45
lokal B	0,03	0,00	0,20	0,30	0,43
lokal C	0,23	0,20	0,00	0,11	0,25
lokal D	0,33	0,30	0,11	0,00	0,14
lokal E	0,45	0,43	0,25	0,14	0,00

Tabela 15. Oszacowanie odległości Canberra według formuły (15) w oparciu o oceny nie normalizowane.

Table 15. Estimation of Canberra distance (eq. 15) and non-normalized evaluations.

	lokal A	Lokal B	lokal C	lokal D	lokal E
lokal A	0,00	0,02	0,23	0,33	0,43
lokal B	0,02	0,00	0,21	0,32	0,42
lokal C	0,23	0,21	0,00	0,11	0,22
lokal D	0,33	0,32	0,11	0,00	0,11
lokal E	0,43	0,42	0,22	0,11	0,00

 Zacieniowanie komórki wskazuje uznanie podobieństwa między obiektami.  
Shading of cells points the recognition of objects similarity.

Przyjmując za kryterium podobieństwa (dla oszacowań opartych na ocenach znormalizowanych) odległość  $d < 0,5$  można uznać, że w większości oszacowań lokale A i B (predefiniowane jako lepsze) a także C, D i E (średnie i gorsze) uznać można za podobne. Wykazują to rachunki wykonane według skorygowanych formuł odległości euklidesowej (11) i Spearmana (12). Oceny podobieństwa według odległości Bray'a Curtisa (13) oraz Canberra (15) w przypadku obiektów o ocenach wysokich, są zbieżne. W oszacowaniach Bray'a Curtisa i Canberra wyraźnie zaznacza się jednak tendencja do uwypuklenia odmienności względem obiektów ocenianych niskimi rangami.

Pewnej uwagi wymaga interpretacja wyników oszacowania odległości Bray'a Curtisa (13) oraz Canberra (15) opartych na ocenach nienormalizowanych. Zgodnie z oczekiwaniami nawet między skrajnie ocenianymi obiektami odległości nie przekroczyły one bowiem wielkości 0,45. Oceniając jednak otrzymane poszczególne wielkości odległości w proporcji do maksymalnych oszacowań z poszczególnych wierszy odpowiednich tablic, dochodzi się do wniosków bardzo zbliżonych do powyższych, wywiedzionych z oszacowań według skorygowanych formuł odległości euklidesowej (11) i Spearmana (12). Oceny oparte na formułach Bray'a Curtisa czy Canberra grupują analizowane obiekty w małe grupy A-B, B-C itd. (identyfikowane jako komórki z tablic wyników, układające się w pobliżu diagonalnej poszczególnych tablic) odrzucając tezę podobieństwa dla pozostałych par.

Przeprowadzone przykładowe obliczenia dowodzą praktycznej użyteczności omówionych wyżej formuł rachunkowych dla potrzeb analizy danych rynkowych w procedurach wyceny nieruchomości. Wybrane formuły stosowane osobno lub równoległe mogą stanowić istotną pomoc z ocenie spójności i doborze danych rynkowych zwłaszcza w przypadkach dużych zbiorów informacji. Wyniki wykonanych przy ich pomocy rachunków mogą z kolei stanowić matematyczny dowód w dyskusji nad poprawnością doboru danych i ich odpowiednością względem przedmiotu wyceny.



## Literatura

- [1] Gregory S., *Metody statystyki w geografii*, PWN, Warszawa, 1976.
- [2] Jażdżewska I., *Statystyka dla geografów*, Wyd. Uniwersytetu Łódzkiego, Łódź, 2003.
- [3] Jokiel B., Kostrubiec B., *Statystyka z elementami matematyki dla geografów*, Państwowe Wydawnictwo Naukowe, Warszawa, 1981.
- [4] Kardi T., *Similarity Measurement*, <http://people.revoledu.com/kardi/tutorial/Similarity/>
- [5] Kendall, M., *Słownik terminów statystycznych*, PWE, Warszawa, 1986.
- [6] Koronacki J., Mielniczuk J., *Statystyka dla studentów kierunków technicznych i przyrodniczych*, WNT, Warszawa, 2006.
- [7] Kowalczyk T., Pleszczyńska E., Ruland F., (red.), *Grade Models and Methods for Data Analysis with Applications for the Analysis of Data Populations*, Berlin Heidelberg New York, Springer Verlag, 2004, seria: Studies in Fuzziness and Soft Computing vol. 151.
- [8] Kryszwicki W., Bartos J., Dyczka W., Królikowska K., Wasilewski M., *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach, Część 2. Statystyka matematyczna*, PWN, Warszawa, 2006.
- [9] Luszniwicz A., Słaby T., *Statystyka z pakietem komputerowym STATISTICA PL. Teoria i zastosowania*, Wydawnictwo C.H. Beck, Warszawa, 2001.
- [10] Maksimowicz-Ajchel A., *Wstęp do statystyki*, Wyd. UW, Warszawa, 2007.
- [11] *Mały słownik matematyczny*, Wiedza powszechna, Warszawa 1970.
- [12] Piechota J., *Statystyka nieparametryczna. Modele i zadania*, WPHU Opta, Warszawa, 1996.
- [13] Sierpiński W., *Wstęp do teorii mnogości i topologii*, PZWS, Warszawa, 1965.
- [14] *Słownik współczesnego języka polskiego*, Wydawnictwo WILGA, Warszawa, 1996.
- [15] Yule G.U., Kendall M. G.: *Wstęp do teorii statystyki*, PWN, Warszawa, 1966.
- [16] Ustawa z dnia 21 sierpnia 1997 r. o gospodarce nieruchomościami, (Tekst jednolity: Dz. U. z 2004 r. Nr 261, poz. 2603).
- [17] Rozporządzenie Rady Ministrów z dnia 21 września 2004 r. w sprawie wyceny nieruchomości i sporządzania operatu szacunkowego (Dz. U. nr 207 poz. 2109.).
- [18] [http://pl.wikipedia.org/wiki/Odleg%C5%82o%C5%9B%C4%87\\_Minkowskiego](http://pl.wikipedia.org/wiki/Odleg%C5%82o%C5%9B%C4%87_Minkowskiego)
- [19] [http://pl.wikipedia.org/wiki/Odleg%C5%82o%C5%9B%C4%87\\_Czebyszewa](http://pl.wikipedia.org/wiki/Odleg%C5%82o%C5%9B%C4%87_Czebyszewa)

## Similarity in real estates valuation

Jacek Zyga

*Lublin University of Technology, Faculty of Civil Engineering and Architecture,  
Department of Geotechnics, e-mail: j.zyga@pollub.pl*

**Abstract:** Similarity, is although a notion universally practical, in effect of many spheres (technical and not technical) is a notion about the not explicit definition, basing itself on current to his apprehension. In effect it means the lack rational, leaning on proofs, marks of comparing of facultative objects. The present article makes an attempt of defining of the resemblance and the use of chosen methods of his measurement in data analyses to practical in the pricing immovables.

**Key words:** Similarity, distances, real estates evaluation.