

Michał Bernard Pietrzak

DWUETAPOWA PROCEDURA BUDOWY PRZESTRZENNEJ MACIERZY WAG Z UWZGLĘDNIENIEM ODLEGŁOŚCI EKONOMICZNEJ

Słowa kluczowe: *ekonometria przestrzenna, dane przestrzenne, przestrzenna macierz wag, odległość ekonomiczna*

Abstrakt: W artykule przedstawiono istotny problem ekonometrii przestrzennej, jakim jest konstrukcja przestrzennej macierzy wag. Macierz ta wyraża potencjalne interakcje przestrzenne między badanymi obszarami oraz stanowi podstawę analiz przestrzennych. Celem artykułu jest przeprowadzenie rozważań dotyczących różnorodnych sposobów definiowania przestrzennej macierzy wag. Rozpatrzono kolejne sposoby ustalania sąsiedztwa między obszarami w postaci sąsiedztwa w sensie wspólnej granicy, sąsiedztwa określonego na podstawie odległości fizycznej oraz sąsiedztwa ustalonego na podstawie kryterium liczby k najbliższych sąsiadów. Rozważona została również kwestia wyznaczania siły oddziaływania dla zdefiniowanego wcześniej układu sąsiedztwa, gdzie pod uwagę zostały wzięte takie podejścia jak standaryzacja macierzy wierszami do jedności, czy wyznaczenie siły oddziaływania na podstawie własności fizycznych obszarów. W wyniku przeprowadzonych rozważań zaproponowana została dwuetapowa procedura budowy macierzy przestrzennej wag (sąsiedztwa). Zaproponowana została również definicja odległości ekonomicznej, która stanowi poprawne narzędzie dla ustalania siły oddziaływania między sąsiadującymi obszarami. Wprowadzona odległość ekonomiczna jest ważną alternatywą określania siły oddziaływania dla ekonomicznych procesów przestrzennych wobec odległości opartej na charakterystykach fizycznych badanych obszarów oraz wobec propozycji standaryzacji wierszami do jedności.

TWO-STAGE PROCEDURE OF BUILDING A SPATIAL WEIGHT MATRIX WITH THE CONSIDERATION OF ECONOMIC DISTANCE

Keywords: *spatial econometrics, spatial data, matrix, economic distance*

Abstract: The article discusses an essential problem of spatial econometrics which is the construction of a spatial weight matrix. This matrix expresses potential spatial interactions between the researched areas and forms the basis for spatial analyses. The objective of the paper was to consider various ways of defining the a spatial weight matrix. Consideration was given to the methods of identifying adjacency between areas in the form of neighbourhood in the sense of a common border, the adjacency identified on the basis of the physical distance and on the criterion of the k number of the nearest neighbours. Also, the issue of identifying the force of impact for the previously defined neighbourhood system was taken into account. The approaches that were considered include the standardisation of matrices by rows to unity, or the identification of the force of impact based on the physical properties of areas. The outcome of the considerations is the proposed two-stage procedure of building a spatial weight matrix. A definition of economic distance was proposed, the definition which constitutes an adequate instrument for measuring the force of interaction between neighbouring areas. The introduced economic distance is an important alternative to the identification of the force of impact of economic spatial processes in relation to the distance based on the physical properties of the researched areas and in relation to the proposal of the standardisation by rows to unity.

WPROWADZENIE

Ekonometria przestrzenna rozwija się dynamicznie od połowy lat siedemdziesiątych XX wieku, a do ważnych prac pionierskich w tej dziedzinie zaliczyć należy Cliff, Ord (1973, 1981), Bivand (1981), Klaassen, Paelinck, Wagenaar (1982), Klaassen, Pealinck (1983), Anselin (1988), Zeliaś (1991). Wśród najnowszych prac propagujących ekonometrię przestrzenną wyróżnić można Arbia (2006), Kopczewska (2005), Szulc (2007), Bivand, Pebesma, Gómez-Rubio (2008), LeSage, Pace (2009), Suchecki (2010). Przedmiot analiza prowadzonych w ramach ekonometrii przestrzennej stanowią przestrzenne procesy ekonomiczne. Procesy te określane są najczęściej dla obszarów nieregularnych, co jest wynikiem przyjętej nomenklatu-

ry definiowania obszarów, w ramach której prowadzony jest pomiar podstawowych charakterystyk obszaru¹. Wykonywane analizy przestrzenne dotyczące Polski, czy Unii Europejskiej prawie zawsze oparte są na klasyfikacji NUTS, co wynika z dostępności danych statystycznych². W związku z tym prowadzone w artykule rozważania ograniczone zostaną do obszarów nieregularnych³.

Należy podkreślić bardzo istotną kwestię, iż wszystkie badania dla danych w ramach klasyfikacji NUTS są badaniami dla danych przestrzennych⁴. Dane przestrzenne charakteryzują się dwoma podstawowymi własnościami, heterogenicznością przestrzenną oraz istnieniem zależności przestrzennych (por. Anselin 1988). Każda analiza ekonomiczna, nieuwzględniająca choć jednej z wymienionych własności danych przestrzennych prowadzi do błędów poznawczych, co podważa wiarygodność jej wyników. Poruszona kwestia wskazuje na konieczność rozwijania ekonometrii przestrzennej, gdzie w pracy Anselin (1988) określona została jako dziedzina, która w ramach swoich metod i technik oraz przy uwzględnieniu dwóch kluczowych własności danych przestrzennych w postaci przestrzennej heterogeniczności oraz zależności przestrzennych, zapewnia narzędzia pozwalające na poprawną specyfikację modeli, ich estymację, testowanie hipotez oraz predykcję.

Treść artykułu dotyczy ważnego problemu ekonometrii przestrzennej jakim jest wyrażenie potencjalnych interakcji przestrzennych w zależności od położenia obszarów. Podstawową rolę odgrywają tutaj wagi przestrzenne, które dla wszystkich obszarów określone są za pomocą przestrzennej macierzy wag (sąsiedztwa). Za pomocą wag przestrzennych wyrażona jest

¹ W Unii Europejskiej, jak i w Polsce pomiar najważniejszych charakterystyk społeczno-gospodarczych wykonywany jest dla obszarów określonych według klasyfikacji NUTS (Nomenclature of Units for Territorial Statistics). Nomenklatura NUTS obowiązuje w Polsce od 01.05.2004 roku, a celem jej wprowadzenia było przyjęcie wspólnego podziału administracyjnego dla wszystkich krajów Unii Europejskiej, w ramach którego gromadzone oraz publikowane są dane statystyczne. Poziom NUTS 0 określa kraje Unii Europejskiej. Wyższe poziomy klasyfikacji oznaczają w przypadku Polski NUTS 1 regiony, NUTS 2 województwa, NUTS 3 podregiony, NUTS 4 powiaty oraz NUTS 5 gminy.

² Prowadzone analizy mogą dotyczyć dowolnych obszarów regularnych, jak i nieregularnych. Ze względu jednak na koszt oraz czas pozyskania danych, najczęściej korzysta się z gotowych danych statystycznych opracowanych przez wiarygodne instytucje, które z kolei prowadzą badania zgodnie z klasyfikacją NUTS.

³ W przypadku rozważań dla obszarów nieregularnych istotną sprawą jest ich umiejscowienie w przestrzeni za pomocą pary współrzędnych geograficznych. W tym celu wyznaczany jest centroid, który jako wewnętrzny punkt reprezentuje centrum obszaru. Odległość między obszarami liczona jest jako odległość euklidesowa między centroidami obszarów.

⁴ Dane przestrzenne traktowane są w analizach najczęściej jako dane przekrojowe, co może prowadzić do błędnych wyników prowadzonych analiz.

siła oddziaływania dwóch obszarów uznanych za sąsiadujące albo brak tego oddziaływania. Cel artykułu stanowi przeprowadzenie rozważań dotyczących różnorodnych sposobów definiowania przestrzennej macierzy wag. W wyniku realizacji założonego celu zaproponowana zostanie dwuetapowa procedura budowy przestrzennej macierzy wag. W wyniku rozważań zaproponowana zostanie także definicja odległości ekonomicznej⁵. Wprowadzona odległość ekonomiczna stanowi ważną alternatywę określania siły oddziaływania dla ekonomicznych procesów przestrzennych wobec odległości opartej na charakterystykach fizycznych badanych obszarów takich jak powierzchnia, długość granic, odległość euklidesowa mierzona względem centroidów obszarów oraz wobec propozycji standaryzacji wierszami do jedności.

DWUETAPOWA PROCEDURA BUDOWY MACIERZY SĄSIEDZTWA

W ramach prowadzonych analiz przestrzennych jedną z najważniejszych kwestii jest ustalenie przestrzennej macierzy wag. Jest to związane z faktem, że stosowane miary zależności przestrzennych oraz modele przestrzenne oparte są na macierzy sąsiedztwa, która wyraża potencjalne interakcje przestrzenne.

Istnieją dwa podejścia konstrukcji macierzy wag. Pierwsze podejście polega na przyjęciu macierzy sąsiedztwa jako egzogenicznej względem modelu, gdzie wagi ustalane są *a priori*. Takie podejście rodzi problem, w którym struktura przestrzenna, jaką badacz zamierza odkryć, jest z góry zakładana poprzez odpowiednią postać macierzy wag, zanim analiza zostanie wykonana. Oznacza to, że wynik w znacznym stopniu jest uzależniony od przyjętej macierzy sąsiedztwa.

W podejściu drugim zakłada się przyjęcie macierzy sąsiedztwa jako endogenicznej i estymację parametrów macierzy wraz z parametrami modelu. Założenie endogeniczności macierzy sąsiedztwa wydaje się jeszcze bardziej problematyczne. Każdą macierz sąsiedztwa można potraktować jako sumę dwóch informacji, o tym które obszary są ustalone jako sąsiedzi (dodatnie wagi) oraz o tym jaka jest pomiędzy nimi potencjalna siła oddziaływania (wartość wag). Egzogeniczność macierzy wag wymusza ustalenie całej informacji *a priori*. W przypadku endogeniczności macierzy wag, estymowane parametry odpowiadają jedynie za siłę oddziaływania, natomiast informacja dotycząca sąsiadów musi być ustalona *a priori*. Rodzi to

⁵ Definicja ta nie jest uniwersalna. Stanowi jedno z możliwych podejść odnośnie pomiaru odległości ekonomicznej.

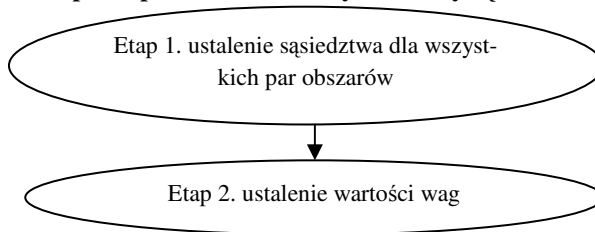
ten sam problem, jaki pojawia się w przypadku egzogenicznej macierzy sąsiedztwa. Dodatkowa problematyczność przyjęcia endogenicznej macierzy sąsiedztwa wynika również z faktu, że kryterium estymacji parametrów jest jak najlepsze dopasowanie do danych empirycznych. Otrzymane zostają takie oceny parametrów, dla których wartość zależności przestrzennych jest największa. W przypadku, gdy w analizowanym procesie przestrzennym na jego wewnętrzną strukturę⁶ składa się również heterogeniczność przestrzenna oraz własność ta nie zostanie założona w modelu, to wartość zależności przestrzennych będzie znacznie zawyżona. Dzieje się tak dlatego, iż w przypadku istnienia heterogeniczności danych możliwa jest błędna identyfikacja zależności przestrzennych. Z kolei w przypadku istnienia zależności przestrzennych błędna identyfikacja na przykład w postaci trendu przestrzennego. Obydwa przypadki błędnej specyfikacji heterogeniczności oraz zależności przestrzennych dla modeli przestrzennych rozważone zostały np. w pracach Pietrzaka (2009, 2010). Oznacza to, że źle wyspecyfikowany model ekonometryczny może prowadzić do większych błędów poznawczych w przypadku wykorzystania endogenicznej macierzy sąsiedztwa w porównaniu z podejściem egzogenicznym.

Zdaniem autora zalecane powinno być wykorzystanie egzogenicznej macierzy wag. Jest to najprostsze rozwiązanie i w mniejszym stopniu naraża badacza na błędną identyfikację wewnętrznej struktury procesu. Często stosowanym podejściem rozwiązania problemu egzogeniczności macierzy sąsiedztwa jest ustalenie kryterium, na podstawie którego, z potencjalnego zbioru macierzy, wybrana zostanie właściwa. Kryterium to stanowią wartości miar autokorelacji, wartości testów statystycznych lub wartości funkcji wiarygodności dla przyjętych modeli (por. Cliff, Ord 1981; LeSage, Pace 2009). Właściwą macierzą wag ma być ta, gdzie dla stosowanych miar autokorelacji przestrzennej, testu statystycznego, czy funkcji wiarygodności otrzymane zostały najlepsze wartości ze względu na przyjęte kryterium. Niestety takie rozwiązanie naraża badacza na błędną identyfikację wewnętrznej struktury procesu, w wyniku czego może prowadzić do błędnego wyboru macierzy sąsiedztwa oraz zawyżonych ocen zależności przestrzennych. Dlatego wybór konkretnej macierzy wag powinien wynikać ze specyfiki analizowanego zjawiska ekonomicznego oraz jego determinant, a nie z najlepszego dopasowania macierzy sąsiedztwa do danych empirycznych. Zaproponowana poniżej dwuetałowa procedura budowy macierzy sąsiedztwa powinna zapewnić dobór odpowiedniej macierzy w ramach

⁶ Wewnętrzną strukturę danych tworzą niejednorodność systematyczna oraz jednorodność systematyczna. Wyrazem niejednorodności systematycznej jest heterogeniczność przestrzenna. Natomiast zależności przestrzenne stanowią składnik jednorodności systematycznej.

prorowadzonych analiz ekonomicznych. W pierwszym etapie procedury ustalane jest sąsiedztwo dla wszystkich par sąsiadów. W drugim etapie natomiast określana jest potencjalna siła oddziaływania między obszarami. Procedura została przedstawiona na schemacie 1.

Schemat 1. Dwuetapowa procedura budowy macierzy sąsiedztwa



Źródło opracowanie własne.

Zgodnie z proponowaną procedurą w etapie pierwszym wynik ustalania sąsiedztwa dla wszystkich par obszarów stanowi kwadratowa macierz sąsiedztwa. Zawiera ona elementy w_{ij} , gdzie sąsiedztwo określone jest za pomocą wartości równej jeden, w przeciwnym wypadku element w_{ij} jest równy zero. Każdy element macierzy sąsiedztwa określany jest zgodnie z poniższym wzorem

$$w_{ij} = \begin{cases} 1, & \text{gdy warunek } g^* \text{ zachodzi} \\ 0, & \text{gdy warunek } g^* \text{ nie zachodzi} \end{cases} \quad (1)$$

Warunek g^* jest definiowany najczęściej na trzy sposoby⁷. Pierwszy sposób ustalenia sąsiadów polega na tym, że warunek g^* jest spełniony wówczas, gdy obszary posiadają wspólną granicę. Warunek ten prowadzi do utworzenia binarnej macierzy sąsiedztwa pierwszego rzędu⁸, której pierwsze wykorzystanie w analizach przestrzennych znaleźć można w pracach Moran (1948), Geary (1954).

⁷ Sposób ustalenia sąsiedztwa może być dowolny. Ważne jest, żeby zdefiniowane sąsiedztwo korespondowało z podjętym problemem badawczym.

⁸ Należy zauważyć, że istnieje możliwość tworzenia macierzy sąsiedztwa wyższych rzędów. Sąsiadami drugiego rzędu będą wszystkie obszary, które zostały uznane za sąsiadów obszarów-sąsiadów pierwszego rzędu.

W drugim przypadku warunek g^* zachodzi, jeżeli odległość między obszarami jest mniejsza od z góry założonej odległości granicznej, co zapisane może zostać jako

$$d_{ij} \leq d^*, \quad (2)$$

gdzie d_{ij} jest odległością euklidesową między obszarami, a d^* oznacza wybraną odległość graniczną⁹. Warunek określony we wzorze (2) prowadzi do utworzenia macierzy sąsiedztwa na podstawie odległości, której zastosowanie znaleźć można w pracy Cliff, Ord (1981).

Ostatni sposób określenia warunku g^* polega na przyjęciu za sąsiadów z góry ustalonej liczby k najbliższych obszarów w sensie odległości. Ostatni sposób prowadzi do utworzenia tak zwanej macierzy sąsiedztwa dla najbliższych sąsiadów. Przykład wykorzystania tej macierzy w badaniach przestrzennych znaleźć można np. w pracy LeSage, Pace (2009). Wszystkie trzy sposoby ustalenia sąsiedztwa prowadzą do utworzenia binarnej macierzy sąsiedztwa. W macierzy tego typu zakłada się, że siła oddziaływania między obszarem a jego wszystkimi sąsiadami jest taka sama.

Wybór odpowiedniego sposobu ustalania sąsiedztwa zależy jest od rozpatrywanego problemu. Jednak w analizach ekonomicznych, gdzie badane są obszary nieregularne, najbardziej odpowiedni, według autora, jest wybór sąsiadów w sensie sąsiedztwa wspólnej granicy. Ze względu na różne kształty obszarów, środek ciężkości nie musi być zdefiniowany dokładnie na środku obszaru. Może to powodować duże nieścisłości w definiowaniu sąsiedztwa na podstawie kryterium odległości. Bardzo problematyczny jest również wybór odpowiedniej wartości odległości granicznej d^* . Tak samo jest w przypadku ustalenia liczby k najbliższych sąsiadów. W literaturze, co zostało wskazane we wcześniejszej części artykułu, sugerowany jest taki dobór odległości lub liczby k najbliższych sąsiadów, który optymalizuje przyjęte kryterium. Ponieważ za tym krokiem stoi niebezpieczeństwo w postaci błędnej identyfikacji wewnętrznej struktury procesu ekonomicznego¹⁰, kryterium to należy odradzać w ramach czynionych analiz. Nawet przy błędnej specyfikacji struktury wewnętrznej procesu, przy zmianie odległości granicznej lub zmianie liczby k najbliższych sąsiadów, badacz

⁹ Opis dotyczący wzorów obowiązywać będzie dla wszystkich wzorów w tekście. W razie potrzeby opis zostanie rozszerzony.

¹⁰ Dla każdej macierzy sąsiedztwa proponowany jest pomiar siły zależności przestrzennych za pomocą miar statystyki przestrzennej (na przykład statystyki I Morana). W takim podejściu nie uwzględnia się struktury wewnętrznej badanego zjawiska ekonomicznego, co oznacza, że prawdopodobieństwo błędnej identyfikacji jest bardzo duże.

zawsze otrzyma optymalną wartość kryterium. Jednak wynik może być fałszywy na przykład poprzez zakłócenia wywołane istnieniem heterogeniczności przestrzennej. Reasumując, w trakcie analizy obszarów nieregularnych, najbardziej intuicyjne jest określenie sąsiedztwa poprzez przyjęcie kryterium wspólnej granicy.

Proponowana w artykule procedura składa się z dwóch etapów, gdzie po ustaleniu sąsiadów na etapie pierwszym, następuje przejście do drugiego etapu i ustalana jest siła oddziaływania poprzez nadanie nowych wartości w binarnej macierzy sąsiedztwa. Etap drugi uzasadniony jest faktem, że w przypadku wielu procesów ekonomicznych siła oddziaływania jest zmienna i wynika z podobieństwa oraz różnic między obszarami. Oznacza to, że im bardziej podobne do są do siebie dwa sąsiadujące obszary pod względem badanego procesu lub jego determinant, to tym większa jest siła wspólnego oddziaływania.

Najczęściej stosowaną metodą przekształcenia binarnej macierzy sąsiedztwa pierwszego rzędu jest jej standaryzacja wierszami do jedności, w wyniku czego uzyskiwana jest standaryzowana macierz sąsiedztwa pierwszego rzędu (zob. Anselin 1988). Przyjęcie takiej macierzy oznacza, że siła oddziaływania wszystkich sąsiadów jest taka sama. Oprócz prostoty utworzenia takiej macierzy, argumentem przemawiającym za wykorzystaniem standaryzowanej macierzy sąsiedztwa pierwszego rzędu jest wykorzystanie opóźnienia przestrzennego zmiennej objaśnianej *WY*. Każdy element wektora *WY* oznacza opóźnienie przestrzenne dla wybranego obszaru przestrzennego, które jest nieważoną średnią arytmetyczną z wartości procesu obszarów sąsiadujących. Opóźnienie przestrzenne można również określić jako średni poziom zjawiska wśród sąsiadów wybranego obszaru. Pozwala to na łatwą interpretację modelu ekonometrycznego, a także możliwość porównania różnych modeli ekonometrycznych.

Alternatywę dla standaryzacji wierszami do jedności stanowi podejście zakładające istnienie oddziaływania między obszarami o różnym nasileniu. Koncepcja zmiennej siły oddziaływania wprowadzona została w pracach Cliff, Ord (1973, 1981)¹¹, gdzie binarna macierz sąsiedztwa została uogólniona poprzez uwzględnienie fizycznych własności obszarów w celu lep-

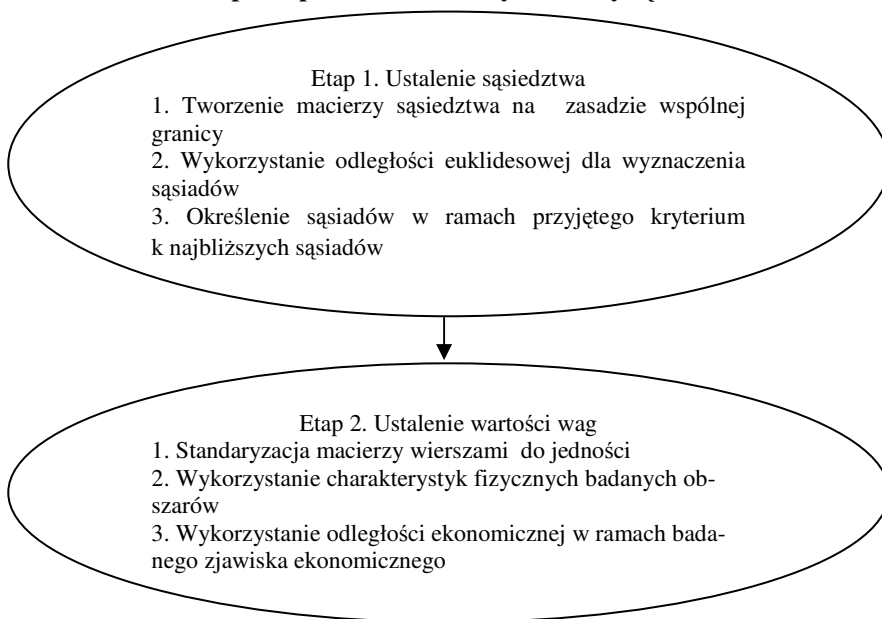
¹¹ W pracach tych przyjęte zostały kryteria oparte na właściwościach fizycznych obszarów. Wynika to z faktu, że kryteria te autorzy uważali za istotne dla badanych przez nich zjawisk, które związane były bezpośrednio z odległością euklidesową. Należy podkreślić, że autorzy widzieli znacznie szerszy problem tworzenia macierzy sąsiedztwa. Twierdzili, że w związku z charakterem badanego zjawiska przyjmowane mogą być różne hipotezy odnośnie do interakcji między obszarami. Wskazywali na konieczność uzasadnienia procedury tworzenia przestrzennej macierzy wag, tak by uzyskane zależności nie były iluzoryczne. Autorzy sugerowali również wykorzystanie innych kryteriów, na przykład dostępności transportu pomiędzy obszarami.

szego opisu potencjalnych interakcji przestrzennych. W pracach tych wykorzystano kryteria odwrotności odległości oraz długości wspólnej granicy, które pozwoliły na przebudowę binarnych macierzy sąsiedztwa pierwszego rzędu. Zgodnie z przyjętymi kryteriami wartości wyjściowej macierzy zostały zmienione według wzoru

$$w_{ij} = \begin{cases} d_{ij}^{-a} \beta_{ij}^b, & w_{ij}^* = 1 \\ 0, & w_{ij}^* = 0 \end{cases} \quad (3)$$

gdzie w_{ij}^* oznacza wagi binarnej macierzy sąsiedztwa, d_{ij} to odległość euklidesowa między obszarami, β to długość wspólnej granicy, natomiast a , b to ustalone a priori parametry, większe od zera.

Schemat 2. Dwuetapowa procedura budowy macierzy sąsiedztwa



Źródło opracowanie własne.

Dwa opisane powyżej sposoby przekształcania binarnej macierzy sąsiedztwa pierwszego rzędu i jednocześnie określania siły interakcji przestrzennych między obszarami są jednak krytykowane ze względu na nieadekwatność w stosunku do analizowanego procesu ekonomicznego. Naturalnym założeniem jest zróżnicowanie siły oddziaływania w zależności od

wybranych obszarów, co wynika z podobieństwa oraz różnic obszarów w zakresie kształtowania się badanego zjawiska oraz jego determinant. Standaryzowana macierz sąsiedztwa zakłada tą samą siłę oddziaływania dla wszystkich obszarów, bez względu na zmienność analizowanego procesu przestrzennego. Dodatkowo pojawia się problem, że obszar posiadający mniej sąsiadów posiada silniejsze interakcje z sąsiadującymi obszarami, niż obszar o większej liczbie sąsiadów. Z kolei wykorzystanie kryteriów fizycznych, związanych z odległością geograficzną czy długością wspólnej granicy może prowadzić do znacznego zniekształcenia siły oddziaływania między obszarami, ponieważ analizowane zjawisko ma charakter ekonomiczny i nie zależy od determinant fizycznych. Argumenty te wskazują na potrzebę wykorzystania odległości ekonomicznej, która różnicowałaby siłę oddziaływania przestrzennego i byłaby w bezpośredniej relacji do zjawiska ekonomicznego, które jest badane.

Po przeprowadzeniu powyższych rozważań proponowaną dwuetapową procedurę tworzenia macierzy sąsiedztwa można uszczegółowić w postaci schematu 2. Zarówno dla etapu pierwszego, jak i drugiego podane zostały w punktach alternatywne sposoby ustalania sąsiedztwa oraz siły oddziaływania między obszarami. W każdym z etapów badacz musi podjąć decyzję odnośnie wykorzystania dokładnie jednej zasady ustalenia sąsiedztwa i jednej zasady ustalenia wartości wag. Wybór innych zasad prowadzi do utworzenia alternatywnej przestrzennej macierzy wag.

OKREŚLENIE ODLEGŁOŚCI EKONOMICZNEJ

Informacja dotycząca podobieństwa obszarów ze względu na badane zjawisko jest kluczowa dla wykonywanych analiz ekonomicznych i jej pominięcie znacznie zubaża prowadzone analizy. Można przyjąć, że uwzględnienie podobieństwa obszarów poprzez wykorzystanie odległości ekonomicznej powinno pozwolić na ustalenie realnej siły oddziaływania między sąsiadującymi obszarami. Jest to bardzo ważny fakt, ponieważ sąsiedzi w sensie wspólnej granicy mogą posiadać różne odległości ekonomiczne w ramach rozważanego zjawiska ekonomicznego, a co za tym idzie różnić się siłą interakcji przestrzennych. W związku z postawionym problemem należy dokładniej wyjaśnić definicję ekonomicznej odległości. Określona ona zostanie jako miara podobieństwa sąsiadujących obszarów pod względem rozpatrywanego procesu ekonomicznego oraz jego determinant. Oznacza to, że dowolnie wybrany obszar powinien charakteryzować się zmienną odległością ekonomiczną w zależności od wybranego sąsiada. W związku z sugerowanym wykorzystaniem analizowanego procesu eko-

nomicznego Y lub n procesów ekonomicznych determinujących zmienność badanego zjawiska¹², odległość ekonomiczną¹³ można zapisać jako

$$d_{ij} = \begin{cases} |y_i - y_j|^{\delta_1} + \frac{1}{k_1 + 1} \sum_{z=0}^{k_1} |x_{1i,t-z} - x_{1j,t-z}|^{\delta_2} + \dots \\ \dots + \frac{1}{k_2 + 1} \sum_{z=0}^{k_2} |x_{2i,t-z} - x_{2j,t-z}|^{\delta_3} + \frac{1}{k_n + 1} \sum_{z=0}^{k_n} |x_{ni,t-z} - x_{nj,t-z}|^{\delta_{n+1}}, & i \neq j \\ 0, & i = j \end{cases} \quad (4)$$

gdzie Y jest procesem objaśnianym, X_i są to procesy objaśniające¹⁴, stałe k_1, k_2, \dots, k_{n+1} oznaczają wielkość opóźnienia po czasie dla przyjętych procesów, dla których liczona jest średnia arytmetyczna, $\delta_1, \delta_2, \dots, \delta_{n+1}$ są to stałe normalizujące, przyjmujące wartości większe od zera

Wykorzystanie ustalonych zmiennych do wyznaczenia odległości ekonomicznej zależy od badacza. Może on ograniczyć zakres wykorzystanych zmiennych tylko do zmiennej objaśnianej Y ¹⁵, do jednej lub więcej zmiennych objaśniających X lub jedynie do średnich z wartości zmiennej objaśnianej lub zmiennych objaśniających. W związku z określoną odległością ekonomiczną, binarną macierz sąsiedztwa można przekształcić zgodnie ze wzorem

$$w_{ij} = \begin{cases} \frac{1}{d_{ij}}, & i \neq j \\ 0, & i = j \end{cases} \quad (5)$$

¹² Zmienność analizowanego zjawiska ekonomicznego oraz zmienność jego determinant najlepiej wyrażają podobieństwo ekonomiczne między regionami, na podstawie którego wyznacza się następnie siłę interakcji przestrzennych. Możliwe jest też wzięcie pod uwagę wartości średnich dla determinant, szczególnie takich które charakteryzują się dużą zmiennością w czasie (na przykład inwestycje).

¹³ Proponowana odległość ekonomiczna nie jest rozumiana w sensie metryki. Metryka mogłaby służyć dla ustalenia sąsiedztwa. Przy już ustalonym sąsiedztwie odległość ekonomiczna określa podobieństwo między sąsiadującymi obszarami.

¹⁴ Wszystkie wykorzystane procesy muszą zostać poddane operacji normalizacji.

¹⁵ Pojawia się tutaj zagrożenie wystąpienia tych samych wartości zjawiska dla różnych obszarów, co spowoduje, że odległość ekonomiczna będzie wynosić zero. Sugerowane jest wtedy przyjęcie wartości równej najmniejszej możliwej jednostce, jaką może przyjąć zjawisko. W przypadku bezrobocia, które jest publikowane w wartościach zaokrąglonych do jednego miejsca po przecinku, sugerowaną wartością byłoby 0,1.

gdzie na podstawie uzyskanej macierzy, w wyniku standaryzacji wierszy do jedności, uzyskiwana jest macierz sąsiedztwa oparta na odległości ekonomicznej.

Zaproponowana odległość ekonomiczna może zostać wykorzystana w przestrzennej analizie stopy bezrobocia. Jako determinanty stopy bezrobocia można przyjąć inwestycje oraz liczbę podmiotów gospodarczych przypadające na mieszkańców. W wielu sytuacjach nie znamy wszystkich powodów odnośnie wyboru miejsca inwestycji albo utworzenia podmiotu gospodarczego w danym obszarze lub nie znamy ich wcale. Jeszcze trudniej wytłumaczyć fakt, że u wybranych sąsiadów taka inwestycja lub nowo utworzone podmioty gospodarcze powodują pozytywny wpływ na rozwój ekonomiczny, a dla innych sąsiadów zmiany te są neutralne. Wyróżnione determinanty przekładają się bezpośrednio na sytuację ekonomiczną regionu oraz pośrednio na regiony sąsiadujące. Informacją, którą dysponujemy są dane statystyczne w postaci stopy bezrobocia, wielkości inwestycji oraz liczby podmiotów, świadczące o istniejącej sytuacji w regionie. Dane te warto wykorzystać do ustalenia podobieństwa między obszarami, poprzez ich wykorzystanie w obliczeniu odległości ekonomicznej. W ten sposób w utworzonej macierzy sąsiedztwa uwzględnia się zastaną sytuację dotyczącą rozwoju regionu oraz bezrobocia. Największą siłę oddziaływania przestrzennego będą posiadać sąsiadujące obszary o największym podobieństwie, bliskości w sensie rozwoju. Są to obszary, gdzie analizowana stopa bezrobocia oraz jej determinanty są na podobnym poziomie. Jest bardzo prawdopodobne, że uchwycone mechanizmy oddziaływać będą również w kolejnym okresie. W przypadku stopy bezrobocia istnieją silne mechanizmy (zależności przestrzenne) przeciwdziałające zmianom w wytworzonym układzie regionów. W rezultacie tworzą się skupiska obszarów charakteryzujących się niską stopą bezrobocia, przyciągających inwestycje oraz przedsiębiorców, jak i skupiska obszarów o niskim rozwoju gospodarczym. Oznacza to, że jeżeli zmiana nastąpi w dowolnie określonym regionie, to uwzględniając opisane powyżej zależności przestrzenne, należy się spodziewać najsilniejszego oddziaływania tej zmiany w regionach o największym stopniu podobieństwa. Wynika z tego również, że zmiana może okazać się neutralna wobec obszaru o niskim stopniu podobieństwa wyrażonego odległością ekonomiczną, mimo iż jest obszarem bezpośrednio sąsiadującym.

ZAKOŃCZENIE

Kluczowym problemem w przestrzennych analizach procesów ekonomicznych jest ustalenie odpowiedniej macierzy wag przestrzennych. Konstrukcję każdej takiej macierzy można opisać za pomocą proponowanej w artykule dwuetapowej procedury, gdzie poszczególne etapy powinny posiadać uzasadnienie w odniesieniu do badanego zjawiska ekonomicznego. W etapie pierwszym ustalane jest sąsiedztwo między obszarami. Najczęściej stosowana jest jedna spośród trzech metod, sąsiedztwo w sensie wspólnej granicy, sąsiedztwo określone za pomocą odległości oraz sąsiedztwo wyznaczonego na podstawie liczby k najbliższych sąsiadów. Po ustaleniu sąsiedztwa dla wszystkich obszarów następuje przejście do etapu drugiego, gdzie ustalana jest potencjalna siła oddziaływania między obszarami. W artykule czynione jest założenie, że siła ta powinna być największa dla sąsiadujących obszarów o największym stopniu podobieństwa w sensie analizowanego zjawiska ekonomicznego. Standaryzacja macierzy wierszami do jedności oraz wykorzystanie odległości opartej na charakterystykach fizycznych obszarów nie koresponduje bezpośrednio z badanym zjawiskiem i może prowadzić do ustalenia niewłaściwej siły oddziaływania. Dobrym rozwiązaniem wydaje się wykorzystanie odległości ekonomicznej, gdzie wzięcie pod uwagę wartości samego zjawiska oraz jego determinant może przyczynić się do poprawnego ustalenia potencjalnej siły oddziaływania między sąsiadującymi obszarami. Niewątpliwie poruszone w artykule sprawy wymagają rozpatrzenia empirycznego zastosowania.

Przedstawiona dwuetapowa procedura może być wyjątkowo pomocna dla osób początkujących w analizach przestrzennych, ponieważ jej zastosowanie wymaga przyjęcia określonych założeń, co do sąsiedztwa, jak i potencjalnych interakcji i jednocześnie może być pomocna w lepszym zrozumieniu uzyskanej konstrukcji. Procedura może też służyć do jasnego uzasadnienia zastosowanej macierzy sąsiedztwa w zaawansowanych badaniach naukowych.

LITERATURA

- Anselin L. (1988), *Spatial Econometrics: Method and Models*, Kluwer Academic Publishers, Netherlands.
- Anselin L., Florax R. J. G. M., Rey, S. J. (2004), *Advances in Spatial Econometrics. Methodology, Tools and Applications*, Springer-Verlag, Berlin.
- Arbia G. (2006), *Spatial Econometrics*, Springer-Verlag GmbH.

- Bivand R. (1981), *Modelowanie geograficznych układów czasoprzestrzennych*, PWN, Warszawa-Poznań
- Bivand, R., Pebesma, E. J., Gómez-Rubio, V. (2008), *Applied Spatial Data Analyses with R*, Springer, New York.
- Clif A., Ord J. (1973), *Apatial Autocorrelation*, Pion, London.
- Clif A., Ord J. (1981), *Apatial Processes, Models and Applications*, Pion, London.
- Cressie N. A. C. (1993), *Statistics for Spatial Data*, John Wiley & Sons, New York.
- Geary R. (1954), *The contiguity ratio and statistical mapping*, "TEhe Incorporated Statistician, 5, s. 115-145.
- Getis A., Mur, J. Zoller, H. (2004), *Spatial Econometrics and Spatial Statistics*, Palgrave Macmillan.
- Haining R. P. (2005), *Spatial Data Analysis. Theory and Practice*, Cambridge University Press, 3rd ed., Cambridge.
- Klaassen J. H. P., Paelinck L. H., Wagenaar S. (1982), *Systemy przestrzenne*, PWN, Warszawa
- Klaassen J. H. P., Paelinck L. H. (1983), *Ekonometria przestrzenna*, PWN, Warszawa
- Kopczewska K. (2006), *Ekonometria i statystyka przestrzenna*, CeDeWu, Warszawa.
- LeSage J. P., Pace R. K. (red) (2004), *Advances in Econometrics: Spatial and Spatiotemporal Econometrics*, Elsevier, Amsterdam.
- LeSage J.P., Pace R. K. (2009), *Introduction to Spatial Econometrics*, CRC Press.
- Moran P. (1948), *The interpretation of statistical maps*, „Journal of the Toyal Statistical Society”, Series B,10, s. 243-251.
- Pietrzak M. (2009), *Analiza danych przestrzennych a jakość informacji*, „Modelowanie preferencji a ryzyko '09”, T. Trzaskalik (red.), AE Katowice, 2009, s. 323-339.
- Pietrzak M. (2010), *Problem identyfikacji struktury danych przestrzennych*, „Acta Universitatis Nicolai Copernici, Ekonomia XLI”, Wydawnictwo UMK, w druku
- Schabenberger, O., Gotway, C. A. (2005), *Statistical Methods for Spatial Data Analysis*, Texts in Statistical Science, Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton, London.
- Suchecky B. (2010), *Ekonometria przestrzenna. Metody i modele analizy danych przestrzennych*, Wydawnictwo C.H.Beck, Warszawa
- Szulc E. (2007), *Ekonometryczna analiza wielowymiarowych procesów gospodarczych*, Wydawnictwo UMK, Toruń 2007.
- Zeliaś A. (red) (1991), *Ekonometria przestrzenna*, PWE, Warszawa.