

Zbyszko Chojnicki, Teresa Czyż

Podstawy metodologiczne zastosowania analizy czynnikowej

1. Wstęp

W ostatnich latach w światowej literaturze geograficznej analiza czynnikowa stała się powszechnie stosowaną metodą badawczą. Szybki rozwój zastosowań analizy czynnikowej jest symptomem wyraźnego postępu w dziedzinie zastosowania metod matematycznych w geografii.

Nowe doświadczenia badawcze w zakresie zastosowań analizy czynnikowej oraz krytyka jej wartości poznawczej uzasadniają przeprowadzenie dyskusji nad podstawami metodologicznymi metody czynnikowej, które warunkują dalsze, jeszcze efektywniejsze jej stosowanie.

Analiza czynnikowa jako metoda badania układów wielozmiennych w geografii ekonomicznej może wykazać się szerokim zakresem zastosowania (patrz T. Czyż, 1970). Ostatnio najliczniejszą grupę prac dotyczących zastosowań analizy czynnikowej tworzą prace z zakresu tzw. ekologii czynnikowej miast. Stanowią one kontynuację kierunku badawczego rozwiniętego w szkole chicagowskiej, zmierzającego do określenia podstawowych wymiarów strukturalnych zróżnicowania przestrzeni społeczno-ekonomicznej (B. J. L. Berry, F. E. Horton, 1970). Do nich należą prace dotyczące struktury przestrzennej miast: Indii (B. J. L. Berry, H. Spodek, 1971), Kanady (L. S. Bourne, G. M. Barber, 1971; K. E. Hayness, 1971), Szwecji (C. G. Janson, 1971), USA (R. Palm, D. Caruso, 1972; M. R. Greenberg, T. D. Boswell, 1972; R. Palm, 1973) oraz pojedynczych miast: New Castle (D. Parker, 1971), Christchurch (B. Donaldson, R. J. Johnston, 1973), Rio de Janeiro (F. B. Morris, G. F. Pyle, 1971), Londynu (H. S. Biel, 1972), Akry (R. R. Brand, 1972), Moskwy (J. W. Miedwiedkow, 1974).

Drugą sferę zastosowania analizy czynnikowej w geografii ekonomicznej stanowi badanie rozwoju ekonomicznego oraz regionalizacji i struktury regionalnej (J. H. P. Schilderick, 1970; G. A. Van der Knaap, 1971; S. D. Brunn, J. O. Wheeler, 1971; R. J. Saunders, 1971; Ch. Jaumotte, J. H. P. Paelinck, J. M. Leheureux, M. Pietquin, 1971; R. K. Semple, H. L. Gauthier, C. E. Youngmann, 1972; W. M. Żukowska, 1973; H. G. Kariel, P. E. Kariel, 1973; B. Ortmann, 1971; T. Czyż, 1971; A. Jelonek, 1971; J. A. Dawson, 1972; Z. Chojnicki, T. Czyż, 1972; W. R.

Black, 1973; J. Byfuglien, A. Norgard, 1973; D. Clark, 1973; M. S. Monmonier, 1973; C. R. Bryant, 1974).

Dalszy intensywny rozwój zastosowań nastąpił przez wprowadzenie analizy czynnikowej do badania przestrzeni percepcji środowiska (K. R. Cox, G. Zannaras, 1970; F. B. Stutz, 1973; F. E. Horton, J. Louvière, D. R. Reynolds, 1973; W. Zielinsky, 1974).

2. Analiza czynnikowa a analiza głównych składowych

Pojęcie „analiza czynnikowa” jest wieloznaczne i w ogólniejszym znaczeniu obejmuje się nią wiele metod (patrz D. N. Lawley, A. E. Maxwell, 1963). Spośród nich najbardziej znane i szeroko stosowane to analiza czynnikowa w węższym znaczeniu i analiza głównych składowych. Między tymi dwiema metodami istnieją zasadnicze różnice wyrażające się przede wszystkim w odmiennych założeniach metodologicznych.

W analizie czynnikowej zakłada się dekompozycję zmiennej obserwowalnej x w następujący sposób:

$$x_i = a_{i1} f_1 + a_{i2} f_2 + \dots + a_{im} f_m + a_i u_i \quad (i = 1, 2, \dots, n),$$

gdzie

- f_k – czynniki wspólne ($k = 1, 2, \dots, m$),
- a_i – ładunki czynnikowe ($i = 1, 2, \dots, n$),
- u_i – czynniki swoiste ($i = 1, 2, \dots, n$).

Założenia powyższego modelu są następujące: 1) liczba czynników wspólnych jest mniejsza niż liczba zmiennych, a więc $m < n$; 2) funkcja zmiennej x względem czynników f jest liniowa; 3) czynniki swoiste są nieskorelowane z czynnikami wspólnymi; 4) czynniki swoiste są wzajemnie nieskorelowane. Przy tych założeniach macierz kowariancji C między zmiennymi obserwowalnymi x może być rozłożona na: macierz kowariancji międzyczynnikowych ψ , macierz ładunków czynników wspólnych A (element a_{ij} tej macierzy jest kowariancją i -tej zmiennej z j -tym czynnikiem wspólnym) i macierz diagonalną wariancji czynników swoistych D , co zapisujemy wzorem¹:

$$C = A\psi A' + D.$$

W szczególnym przypadku, gdy czynniki wspólne są nieskorelowane macierz ψ jest macierzą jednostkową, a zatem:

$$C = AA' + D.$$

W analizie czynnikowej zakłada się, że kowariancję między zmiennymi generują czynniki wspólne, podczas gdy czynniki swoiste wyjaśniają wariację właściwą tylko jednej określonej zmiennej. Analiza czynnikowa jest w istocie rzeczą analizą macierzy kowariancji z zasobami wariancji wspólnej na głównej przekątnej. Celem analizy jest wyodrębnienie zbioru tylko czynników wspólnych, co pociąga za sobą redukcję liczb zmiennych.

¹ Niekiedy zamiast macierzy kowariancji używa się macierzy korelacji, tj. macierzy kowariancji dla zmiennych standaryzowanych (por. D. F. Morrison, 1967, s. 223).

Należy podkreślić, że u podstaw analizy czynnikowej leży hipoteza podziału wariancji całkowitej na dwie części: wariację wspólną i wariację specyficzną.

W przeciwieństwie do analizy czynnikowej – model głównych składowych nie uwzględnia struktury wariancji zmiennej. Analiza głównych składowych jest ortogonalną transformacją układu zmiennych (x_1, x_2, \dots, x_n) w zbiór nowych zmiennych (y_1, y_2, \dots, y_p) i ma postać modelu matematycznego wyrażonego równaniem:

$$x_i = b_{i1} y_1 + b_{i2} y_2 + \dots + b_{ik} y_k \quad \begin{matrix} (i = 1, 2, \dots, n), \\ (k = 1, 2, \dots, p), \end{matrix}$$

przy czym liczba zmiennych empirycznych równa się liczbie nowych zmiennych, tzw. składowych ($n = p$) i ogólna wariancja zmiennych x_i równa się wariancji składowych y_k .

W literaturze geograficznej z zakresu zastosowań metod czynnikowych często nie dostrzega się albo ignoruje różnice między analizą czynnikową i metodą głównych składowych. Zjawiskiem powszechnym jest mieszanie terminów tych dwóch metod, np. używa się zamiennie określenia czynnik i składowa. Ch. Jaumotte i in. (1971, s. 46) piszą: „analiza czynnikowa zwana także analizą głównych składowych”, W. R. Black (1973, s. 62) stwierdza: „Czwarty czynnik wyjaśnia 6,3% ogólnej zmienności”. W zastosowaniach nie precyzuje się często ściśle modelu i takie zabiegi, jak np. rotację ortogonalną ukośną, które możliwe są tylko w procedurze analizy czynnikowej nieprawidłowo stosuje się też w analizie głównych składowych (patrz P. M. Mather, 1972).

Należy zaznaczyć, że cele analizy czynnikowej i analizy głównych składowych mogą być zbieżne, jednak każda z nich jest bardziej użyteczna w innej sytuacji problemowej (por. M. G. Kendall, 1968, s. 37; D. Harvey, 1969, s. 343–345; L. J. King, 1969, s. 165 i dalsze). Przykładowo można tu przytoczyć stwierdzenie D. F. Morrisona (1967, s. 257), który pisze, że chociaż analiza głównych składowych może w pewnym stopniu realizować cel analizy czynnikowej, pozostaje wciąż ortogonalną transformacją, a nie metodą analizy struktury kowariancji. W analizie głównych składowych, przez podział ogólnej wariancji zmiennych na sukcesywnie coraz mniejsze części, wyodrębnia się nowe zmienne zwane składowymi. Jeżeli kilka pierwszych składowych ma znaczny udział w ogólnej wariancji, to można opisać wówczas system zmiennych bardziej „oszczędnie” za pomocą tylko tych kilku składowych. W ten sposób analiza głównych składowych może również prowadzić do redukcji danych (podobnie jak analiza czynnikowa), jednak jej podstawowym celem pozostaje ortogonalna transformacja zmiennych.

W związku z niniejszą dyskusją warto jeszcze zwrócić uwagę na fakt, że model analizy czynnikowej (w węższym znaczeniu) jest modelem statystycznym wykazującym wyraźne podobieństwo do równań regresji wielokrotnej (R. J. Rummel, 1970). Porównując równanie czynnikowe z równaniem regresji można stwierdzić wyraźną odpowiedniość terminów: czynniki odpowiadają zmiennym niezależnym w analizie regresji, ładunki czynnikowe – współczynnikom regresji, wariancja specyficzna – pozostałościom regresji, ponadto zasób zmienności wspólnej – współczynnikowi determinacji wielokrotnej. Jednak cel poznawczy i użyteczność modeli są zasadniczo różne. W analizie czynnikowej ładunki identyfikują liczbę i naturę nieznanymi wymiarów (czynników) przestrzeni zmiennych, natomiast w analizie regresji współczynniki określają udział każdego znanego wymiaru (zmiennej niezależnej) w dopasowaniu wektora do tej przestrzeni zmiennych.

W praktyce, w badaniach geograficznych spotyka się liniowy model czynnikowy, odpowiadający liniowej zależności między rozpatrywanymi zmiennymi i czynnikami. Tymczasem zastosowanie modelu czynnikowego w postaci wielomianu wyższego stopnia zapewniłoby w niektórych sytuacjach badawczych większą precyzję wyników. W gruncie rzeczy R. P. Mc Donald (1962) rozwinął już metody nieliniowej funkcji czynnikowej i opublikował odpowiednie programy obliczeń, wymagają one jednak ustalenia ich przydatności do badań empirycznych w geografii.

3. Decyzje arbitralne przy zastosowaniu analizy czynnikowej

Ogólnie biorąc, dotychczasowe zastosowania w badaniach geograficznych preferowały jednak metodę analizy czynnikowej. Ostatnio wysuwa się wiele krytycznych argumentów przeciwko stosowaniu metody analizy czynnikowej nawiązujących do arbitralności niektórych decyzji w procedurze badawczej tej metody. Równocześnie przypisuje się większą wartość wynikom pochodzącym z analizy głównych składowych. Wydaje się jednak, że w toku tej dyskusji należy pamiętać o tym, iż efektywne stosowanie każdej metody wymaga przede wszystkim właściwego dostosowania założeń metody do specyfiki problemu badawczego.

Do zagadnień stanowiących sedno krytyki analizy czynnikowej należą: 1. wybór zmiennych, 2. estymacja zasobu zmienności wspólnej, 3. rotacja czynników.

Ad 1. Planując analizę czynnikową, należy dokonać starannego doboru zmiennych. Oczywiście, że ta wskazówka odnosi się równie dobrze i do innych modeli matematycznych. W wyborze zmiennych do analizy czynnikowej kierujemy się wstępnie przyjętymi założeniami o analizowanym zjawisku. Wybiera się zmienne, z których każda jest istotna dla wybranego problemu badawczego. Zjawisko poddane badaniu musi być scharakteryzowane za pomocą wyjściowych zmiennych. Jeżeli pewne aspekty badanego fragmentu rzeczywistości zostaną pominięte w wyjściowym zbiorze zmiennych, nie pojawiają się one w strukturze czynnikowej.

Ad 2. W analizie czynnikowej stosuje się wiele sposobów estymacji zasobów zmienności wspólnej (Z. Piasecki, 1969). L. J. King (1969) estymuje zasób wariancji wspólnej jako współczynnik determinacji wielokrotnej zmiennej ze względu na pozostałe zmienne zbioru. W licznych zastosowaniach wprowadza się jedności na przekątną macierzy korelacji poddanej analizie czynnikowej, zakładając tym samym występowanie tylko wariancji wspólnej (K. E. Hayness, 1971). W tym przypadku analiza czynnikowa nie różni się od analizy głównych składowych.

Liczbę czynników wspólnych ustala się, wykorzystując kryterium Kaisera: wyodrębnia się tylko czynniki odpowiadające wartościom własnym większym od jedności.

Trzeba również zaznaczyć, że ponieważ w analizie czynnikowej tylko część całkowitej zmienności jest wyjaśniona przez czynniki wspólne, wartości czynnikowe są jedynie estymowanymi wartościami w porównaniu ze ścisłymi wartościami składowych.

Reasumując należy przyznać, że wybór metody estymacji zasobu zmienności wspólnej jest decyzją, która obciąża wyniki analizy czynnikowej pewną dozą subiektywizmu. Problem estymacji wariancji wspólnej nie występuje w analizie głównych składowych po pro-

stu dlatego, że zagadnienie wariancji wspólnej nie mieści się w jej założeniach modelowych.

Ad 3. Faktoryzacja macierzy kowariancji nie jest jednoznaczna nawet w przypadku ograniczenia się do struktur ortogonalnych. Problem wyboru jednej z dekompozycji nazywa się problemem rotacji. W rozwiązaniu problemu rotacji pierwsza decyzja dotyczy wyboru między strukturą ortogonalną i ukośną. W badaniach geograficznych operuje się najczęściej rotacjami ortogonalnymi, m. in. ze względu na ich łatwiejsze opracowanie statystyczne. Poprawność tego podejścia w wielu sytuacjach badawczych, gdy czynniki są skorelowane na podstawie przesłanek teoretycznych, kwestionowali R. J. Johnston (1971) i P. H. Rees (1971). Ostatnio spośród geografów rotację ukośną zastosowali: K. E. Hayness (1971) w procedurze testowania struktury ekologicznej Montrealu oraz F. E. Horton i in. (1973) w analizie czynników przestrzeni percepcji zmian systemu komunikacyjnego. Podstawę technik rotacji stanowi ogólna zasada tzw. prostej struktury. Jednak każdy program rotacji oparty jest na pewnym kryterium, które wymaga maksymalizacji lub minimalizacji pewnej funkcji, np. rotacja Varimax, Quartimax, Quartimin, Oblimin itd. Wybór jednego z programów rotacji jest subiektywny. Jednak z chwilą wyboru programu rotacja jest już jednoznacznie określona przez dane empiryczne (M. Nowakowska, 1973).

4. Założenia i interpretacja wyników analizy czynnikowej w geografii²

Jednym z podstawowych zagadnień jest problem roli poznawczej analizy czynnikowej w badaniach geograficznych.

Analizę czynnikową należy traktować jako metodę wyodrębniania czynników, tj. pewnych wielkości istotnych dla kształtowania się innych wielkości zarówno o charakterze klasyfikującym, jak i porządkującym. Metoda ta ma przede wszystkim charakter heurystyczny i odkrywczy oraz stanowi źródło hipotez faktualnych, jakkolwiek w pewnych przypadkach może też być traktowana jako metoda testowania hipotez.

Odkrywczy charakter metody czynnikowej opiera się na zasadach swoistej redukcji statystycznej, która przez określenie przestrzeni zmienności badanych cech pozwala odkryć pewne ich podstawowe właściwości, tj. czynniki. W ten sposób czynnik powiązany jest znaczeniowo z terminami obserwacyjnymi i stanowi ich uogólnienie. Analiza czynnikowa stanowi więc swoistą metodę budowy hipotezy na temat istotnych wielkości (czynników) w danym fragmencie rzeczywistości.

Istnieją dwa zasadnicze poglądy na zagadnienie interpretacji czynników: opisowy i teoretyczny (R. J. Rummel, 1970; P. H. Rees, 1971; M. Nowakowska, 1973). Interpretacja opisowa polega na ustaleniu nazwy danego podzbioru zmiennych zależnych, wyodrębnionych jako czynnik (tzn. o najwyższych ładunkach w danym czynniku), tak aby uogólnić znaczenie tych zmiennych. Czynniki tak pojmowane są terminami klasyfikacyjnymi i mają znaczenie opisowe. Zgodnie z tym poglądem analiza czynnikowa jest techniką opisu i może prowadzić do sformułowania twierdzeń opisowych.

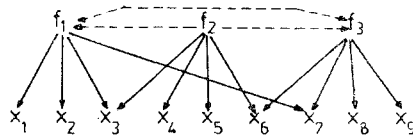
² W dalszym ciągu opracowania będziemy zajmowali się analizą czynnikową w ogólniejszym znaczeniu.

Interpretacji opisowej przeciwstawia się interpretację teoretyczną lub wyjaśniającą. Należy zwrócić uwagę, że wśród autorów prac dotyczących podstaw analizy czynnikowej istnieje w tym zakresie znaczne pomieszanie pojęć, uwidacznia się to między innymi w iunctim między interpretacją przyczynową i teoretyczną (por. R. J. Rummel, 1970), co nie jest właściwe, gdyż interpretacja przyczynowa stanowi tylko pewien przypadek interpretacji teoretycznej.

Interpretacja teoretyczna wiąże się przede wszystkim z traktowaniem czynników jako terminów (predykatów) teoretycznych, a więc takich, które denotują cechy lub relacje nieobserwowalne w przeciwieństwie do terminów obserwowalnych. W tym ujęciu czynnik traktuje się jako konstrukt pojęciowy zawierający nadwyżkę znaczeń (M. Nowakowska, 1973).

Zagadnienie, w jakim stopniu nadaje się czynnikowi charakter teoretyczny jest związane z wykorzystaniem pojęcia czynników do budowy teorii.

Wątpliwości budzi również sprawa interpretacji przyczynowej analizy czynnikowej. Wiąże się to z próbą potraktowania czynników jako przyczyn zmienności zjawisk opisywanych przez zmienne wyjściowe. Nie wdając się w szerszą analizę pojęcia przyczynowości i licznych trudności, jakie są z tym związane, należy zwrócić uwagę na to, że w badaniach geograficznych opierających się na obserwacji, a nie na eksperymencie, interpretację taką można oprzeć na statystycznej analizie związku przyczynowego wyrażonej w modelu Simona-Blałocka, tj. w postaci modelu równań rekurencyjnych. Analiza czynnikowa w tym ujęciu reprezentuje pewien specyficzny przypadek tego typu modelu przyczynowego (H. M. Blalock, 1964).



Ryc. 1. Struktura zależności
Fig. 1. Dependence structure

Analiza czynnikowa w węższym znaczeniu jest modelem przyczynowym, w którym zakłada się, że między zmiennymi obserwowalnymi x_i nie istnieje bezpośrednia zależność przyczynowa (ryc. 1). Przyczynami zjawisk obserwowalnych są tylko zmienne nieobserwowalne f_k występujące w postaci jednego lub większej liczby czynników. Jest rzeczą zrozumiałą, że dzięki zależności przyczynowej zmiennych f_k i x_i zachodzi również korelacja między zmiennymi obserwowalnymi. Tego rodzaju zależność ma jednak charakter zależności formalno-statystycznej, zwanej korelacją pozorną, gdyż między zmiennymi obserwowalnymi brak jest bezpośredniej więzi przyczynowej. Zakłada się, że przyczynami korelacji stwierdzonych w danym zbiorze zmiennych obserwowalnych mogą być jedynie czynniki. Czynniki mogą, lecz bynajmniej nie muszą być powiązane przyczynowo. Stosując liniowe równania jednoczesne, można zapisać każdą zmienną obserwowalną x_i ($i = 1, 2, \dots, n$) jako funkcję podstawowych czynników (zmiennych ukrytych) f_k ($k = 1, 2, \dots, m$) plus dodatkowy termin e_i , reprezentujący skutki działania ubocznych czynników, swoistych tylko dla określonej zmiennej x_i . W ten sposób otrzymuje się układ

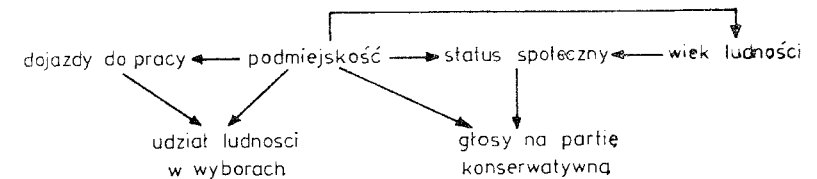
równań:

$$\begin{aligned} x_1 &= b_1 + a_{11}f_1 + a_{12}f_2 + \dots + a_{1m}f_m + e_1, \\ x_2 &= b_2 + a_{21}f_1 + a_{22}f_2 + \dots + a_{2m}f_m + e_2, \\ x_n &= b_n + a_{n1}f_1 + a_{n2}f_2 + \dots + a_{nm}f_m + e_n, \end{aligned}$$

dla których zakłada się, że: 1) wartości e_i są nieskorelowane, 2) niektóre współczynniki a_{ik} mogą być równe zero. W procedurze wyodrębniania kolejnych czynników sukcesywnie redukuje się wartości korelacji między zmiennymi x_i . Po wyodrębnieniu wszystkich m czynników, pozostałości korelacji powinny być równe zero lub bliskie zero (w przypadku błędu próby).

W niektórych sytuacjach empirycznych te założenia przyczynowe analizy czynnikowej wydają się niewłaściwe. Przypuśćmy, że zmiennymi obserwowalnymi są: poziom wykształcenia, wysokość dochodu, odsetek zatrudnionych w przemyśle, sprzedaż w handlu detalicznym. Proponuje się wyodrębnienie czynników, które wyjaśniają korelacje między tymi obserwowalnymi zmiennymi i mogą być stosowane w charakterystyce miasta. Tymczasem racjonalniejsze wydaje się traktowanie poziomu wykształcenia jako bezpośredniej przyczyny wysokości dochodu niż wyjaśnianie korelacji tych dwóch zmiennych obserwowalnych w kategoriach jakiegoś ukrytego czynnika, który trudno nawet nazwać. Konieczne byłoby więc rozpoznanie par obserwowalnych zmiennych, które są skorelowane przyczynowo i dokonanie modyfikacji równań tych zmiennych. W przeciwnym przypadku można popełnić błąd, określając niewłaściwe czynniki albo nadając zbyt „ubogą” interpretację wyodrębnionym czynnikom.

Należy zwrócić uwagę, że w świetle dotychczasowych zastosowań rola analizy czynnikowej w konstrukcji modeli przyczynowych sprowadza się zasadniczo tylko do wyodrębnienia podstawowych wymiarów zbioru zmiennych przyczynowych. Przykładem tego sposobu stosowania analizy czynnikowej są badania K. R. Coxa (1968) w dziedzinie geografii politycznej. K. R. Cox testuje przyczynowe zależności statystyczne między typami przedmieść i zachowaniem się głosujących w obszarze metropolitalnym Londynu. Funkcja analizy czynnikowej polega na redukcji 21 zmiennych społeczno-ekonomicznych, charakteryzujących środowiska przedmieść, do czterech istotnych wymiarów: 1) czynnika społecznego, 2) czynnika odległości od centrum Londynu, czyli podmiejskości, 3) czynnika dojazdów do pracy, 4) czynnika struktury wieku. Te cztery czynniki na podstawie założeń *a priori* traktuje się jako przyczyny dwóch zjawisk: 1) liczby głosujących na partię konserwatywną oraz 2) udziału ludności w wyborach, i rozwija się model przyczynowy Simona-Blałocka. Analiza korelacji cząstkowych pozwala określić działanie poszczególnych czynników i zmodyfikować łańcuch przyczynowy, co z kolei prowadzi do konstrukcji sukcesywnego modelu bardziej zgodnego z danymi empirycznymi (ryc. 2).

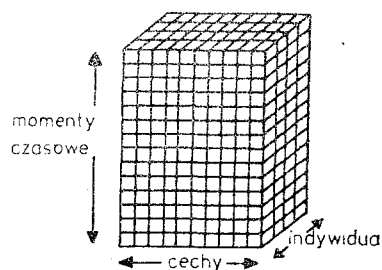


Ryc. 2. Zależności przyczynowe
Fig. 2. Causal dependencies

5. Badanie zmienności przestrzennej i czasowej w analizie czynnikowej

W geografii stosuje się wiele technik czynnikowych związanych z analizą trzech możliwych przekrojów prostopadłościanu danych i ich przestawień (R. J. Rummel, 1970).

Pierwszy sposób cięcia prostopadłościanu danych poziomo, równoległe do płaszczyzny „indywiduum – cecha” prowadzi do macierzy danych dla określonego momentu czasowego, w której cechy tworzą kolumny, a indywidua – wiersze (ryc. 3). Faktorowanie tej macierzy nosi nazwę techniki *R*. Operuje się macierzą korelacji między cechami i uzyskuje ładunki czynnikowe dla cech. Rozwiązania czynnikowe dokonuje się najczęściej przy zastosowaniu tej techniki, prowadzi ona do redukcji przestrzeni wieloczechowej. Jeżeli macierz typu indywiduum × cechy będzie transponowana, to indywidua tworzą ko-



Ryc. 3. Prostopadłościan danych
Fig. 3. Data cube

lumną, a cechy – wiersze. Ta technika zwana *Q* odnosi się do korelacji między jednostkami rozpatrywanymi w kategoriach cech. W tym układzie każdy czynnik reprezentuje „typ” (czyli wiązkę indywiduów). Ładunek czynnikowy określa więc stopień podobieństwa między jednostką a typem. Natomiast wartości czynnikowe identyfikują cechy, które różnicują jednostki w klasy. Ostatnio technika *Q* znajduje coraz szersze zastosowanie w zagadnieniach klasyfikacji przestrzennej. (Z. Chojnicki, T. Czyż, 1972; W. R. Black, 1973; W. Zelinsky, 1974). Algorytm i program rozwiązania techniką *Q* w języku Fortran IV, opracowany przez J. E. Klovana i J. Imbrie (1971) umożliwia operowanie w analizie czynnikowej macierzami o znacznych rozmiarach (1500 indywiduów × 50 zmiennych).

Drugi sposób cięcia prostopadłościanu danych jest pionowy, równoległy do płaszczyzny „moment czasowy – cecha” i daje w rezultacie macierz serii czasowych, obejmującą wartości cech dla jednego indywiduum w różnych momentach czasowych. Tę macierz można analizować dwiema technikami metody czynnikowej: techniką *O*, gdy badamy cechy jednego indywiduum w różnych momentach czasowych oraz techniką *P*, gdy badamy momenty czasowe w kategoriach cech jednej jednostki. Wydaje się, że tego rodzaju podejście w analizie czynnikowej stanowi raczej kierunek badawczy ekonomii.

Trzecie cięcie prostopadłościanu danych jest równoległe do płaszczyzny „moment czasowy – indywiduum” i daje macierz obejmującą zbiór jednostek opisanych w kategoriach tylko jednej cechy w różnych momentach czasowych. Istnieją dwa sposoby analizy tej

macierzy: techniki *T* i *S*. Technika *T* analizy macierzy typu „indywiduum × momenty czasowe” określa istotne momenty czasowe, potrzebne do opisu podobieństwa jednostek w kategoriach jednej cechy. Faktorowanie macierzy typu „momenty czasowe × indywiduum” zwane jest techniką *S* i pozwala określić grupy jednostek, które w kategoriach jednej cechy podobnie zmieniają się w czasie.

Interesującym przykładem zastosowania techniki *S* są badania D. Jeffreya, E. Casettiego i L. J. Kinga (1969) nad ekonomicznymi fluktuacjami w systemie miast. Ponieważ dotyczą one nowego pola zastosowań analizy czynnikowej w geografii ekonomicznej, warto im poświęcić trochę uwagi.

Dla próby obszarów metropolitalnych USA określono serie czasowe wielkości bezrobocia. Sformułowano hipotezę, że komponenty zmienności czasowej zjawiska w systemie obszarów metropolitalnych, takie jak trendy i oscylacje cykliczne, związane są z czynnikami zróżnicowanymi pod względem przestrzennego zasięgu oddziaływania. Seria czasowa zawiera trzy różne poziomy zmienności odpowiadające: 1) czynnikom działającym w całym systemie, 2) czynnikom wspólnym dla grupy elementów systemu, 3) czynnikowi specyficznemu dla każdego elementu. Do testowania tej hipotezy zastosowano analizę dwuczynnikową Holzingera dla macierzy korelacji obszarów metropolitalnych rozpatrywanych w kategoriach serii czasowych. Model struktury czynnikowej serii czasowej przybiera następującą postać:

$$Y_i = aX + b_m V_m + c_i Z_i,$$

gdzie

Y_i – i -ta seria czasowa,

X – czynnik ogólny występujący we wszystkich seriach czasowych zbioru,

V_m – czynnik wspólny dla serii czasowych dla m -tej grupy elementów,

Z_i – czynnik specyficzny dla elementu i .

a, b, c – ładunki czynnikowe.

Należy podkreślić, że proponowane podejście jest szczególnie wartościowe, gdyż pozwala realizować postulaty dynamicznego i systemowego ujmowania procesów w ujęciu przestrzennym.

Każda z wymienionych technik stwarza możliwość określenia modelu czynnikowego dla jednej warstwy prostopadłościanu danych. Sformułowanie systemu modeli przy zastosowaniu określonej techniki dla całego zbioru warstw może stanowić punkt wyjścia dla badań porównawczych (technika *M*). Analiza porównawcza struktur czynnikowych w przestrzeni i w czasie jest szeroko stosowana w geografii miast i w badaniach nad rozwojem ekonomicznym (B. J. L. Berry, H. Spodek, 1971; K. E. Hayness, 1971; Ch. Jaumotte i in. 1971). Należy jednak zauważyć, że istnieje pewne niebezpieczeństwo w stosowaniu tego rodzaju studiów porównawczych: często natura czynników zmienia się i serie rozwiązań czynnikowych nie są porównywalne. W związku z tym na uwagę zasługuje propozycja metodologiczna J. H. F. Schildernicka (1970) w sprawie wykorzystania rotacji w studiach porównawczych. Ponieważ w serii przestrzennej struktur czynnikowych dominujące cechy w odpowiadających kolejnych czynnikach mogą być często różne, co uniemożliwi analizę porównawczą, określony czynnik jest rotowany ze względu na tę samą zmienną.

W poszukiwaniu dalszych rozwiązań czynnikowych przyjmuje się przekształcenie zbioru danych o wymiarach $N \times n \times t$ (indywiduum \times cecha \times moment czasowy) w następujące macierze:

1) macierz typu $N \times nt$, przy założeniu, że określona cecha jednostki dla dwóch różnych momentów czasowych jest traktowana jako dwie różne zmienne; analiza czynnikowa tej macierzy wyodrębnia ładunki czynnikowe dla cechy określonego momentu czasowego;

2) macierz typu $Nt \times n$; faktorowanie tej macierzy daje ładunki czynnikowe dla każdej zmiennej i wartości czynnikowe dla każdego indywiduum w określonym momencie czasowym;

3) macierz typu $t \times nN$; analiza czynnikowa tej macierzy określa ładunki czynnikowe dla cechy określonej jednostki.

Zastosowanie tego rodzaju ujęcia prostopadłościanu danych w analizie czynnikowej przedstawia A. S. Banks (1974). Bada on rozwój krajów uprzemysłowionych i nieuprzemysłowionych w kategoriach ośmiu zmiennych społeczno-ekonomicznych w latach 1868–1963. Zakłada się, że układy zmian społeczno-ekonomicznych w badanym przekroju czasowym można zredukować do następujących podstawowych wymiarów: 1) urbanizacja, 2) budżet i handel, 3) komunikacja, 4) edukacja. Założenie to weryfikuje się przy zastosowaniu analizy czynnikowej macierzy typu $Nt \times n$. Studium porównawcze trendów czasowych wyodrębnionych czynników zmian pozwala scharakteryzować etapy rozwoju poszczególnych krajów.

Najbardziej interesującą propozycją metodologiczną jest próba rozciągnięcia modelu czynnikowego do analizy prostopadłościanu danych dokonana przez L. R. Tuckera (1963). Metodologia ta polega na jednoczesnej analizie czynnikowej łącznej macierzy kowariancji dla zbiorów zmiennych w przestrzeni i w czasie. L. R. Tucker przeprowadza ją w sposób konwencjonalny za pomocą trzech analiz czynnikowych dla dwuwymiarowych macierzy typu: $N_t X_n$, $n_t X_N$, $N_n X_t$, co pozwala wyodrębnić istotne cechy (m), istotne indywidua (p), istotne momenty czasowe (q). Punktem kulminacyjnym jest obliczenie macierzy rdzeniowej o wymiarach $m \times p \times q$, czyli 3-kierunkowej macierzy wartości czynnikowych w następujący sposób:

$${}_m A_n X_{Nt} = {}_m X_{Nt},$$

$${}_p B_N X_{mt} = {}_p X_{mt},$$

$${}_q C_t X_{pm} = {}_q G_{pm},$$

gdzie

${}_q G_{pm}$ – macierz rdzeniowa,

${}_n A_m$, ${}_N B_p$, ${}_t C_q$ – macierze ładunków czynnikowych.

Elementy macierzy rdzeniowej określają siłę zależności między czynnikami każdego z trzech wymiarów prostopadłościanu danych i reprezentują istotną jednostkę wyrażoną w kategoriach istotnych cech w istotnych momentach czasowych. Model czynnikowy przybiera więc postać:

$$X_{ijk} = \sum_m \sum_p \sum_q a_{im} b_{jp} c_{kq} g_{mpq} \quad (i = 1, 2, \dots, n),$$

$$(j = 1, 2, \dots, N),$$

$$(k = 1, 2, \dots, t).$$

Pierwszą próbę zastosowania tej wersji modelu czynnikowego w geografii podjął R. G. Cant (1971). Należy zaznaczyć, że rozszerzona koncepcja modelu czynnikowego może stanowić propozycję w zakresie budowy modelu czasoprzestrzennego. Badania empiryczne w dziedzinie analizy systemów regionalnych prowadzone w Zakładzie Geografii Ekonomicznej UAM mają na celu sprawdzenie efektywności modelu czasoprzestrzennego w ujęciu czynnikowym.

Bibliografia

- Banks A. S. 1974. Industrialization and development – a longitudinal analysis. *Economic Development and Cultural Change*, 22, 2, 320–337.
- Berry B. J. L. 1971. Introduction: the logic and limitations of comparative factorial ecology. *Economic Geography*, 47, 2, 209–219.
- Berry B. J. L., Horton F. E. 1970. *Geographic perspectives on urban systems*. Englewood Cliffs.
- Berry B. J. L., Spodek H. 1971. Comparative ecologies of large Indian cities. *Economic Geography*, 47, 2, 266–285.
- Biel H. S. 1972. Suburbia and voting behavior in the London Metropolitan Area. *Tijdschrift voor Economische en Sociale Geografie*, 63, 1, 39–43.
- Black W. R. 1973. Toward a factorial ecology of flows. *Economic Geography*, 49, 1, 59–67.
- Blaalock H. M. Jr. 1964. *Causal inferences in nonexperimental research*. Chapel Hill.
- Bourne L. S., Barber G. M. 1971. Ecological patterns of small urban centers in Canada. *Economic Geography* 47, 2, 258–265.
- Brand R. R. 1972. The spatial organization of residential areas in Accra, Ghana, with particular reference to aspects of modernization. *Economic Geography*, 48, 3, 284–298.
- Brunn S. D., Wheeler J. O. 1971. Spatial dimension of poverty in the United States. *Geografiska Annaler*, 53B, 1, 6–15.
- Bryant C. R. 1974. An approach to the problem of urbanization and structural change in agriculture: a case study from the Paris Region 1955–1968. *Geografiska Annaler*, 56B, 1, 1–27.
- Byfuglien J., Nordgard A. 1973. Region-building – a comparison of methods. *Norsk Geografisk Tidsskrift*, 27, 127–151.
- Cant R. G. 1971. Changes in the location of manufacturing in New Zealand 1957–1968: and application of three mode factor analysis. *New Zealand Geographer*, 7, 1, 38–55.
- Chojnicki Z., Czyż T. 1972. Zmiany struktury regionalnej Polski w świetle przepływów towarowych w latach 1958–1966. *Studia KPZK PAN*, 40, Warszawa.
- Clark D. 1973. The formal and functional structure of Wales. *Annals of the Association of American Geographers*, 63, 1, 71–84.
- Conway D., Hayness K. E. 1973. A geographical factor analysis model: some transatlantic reflections. *Area*, 5, 2, 106–109.
- Cox K. R. 1968. Suburbia and voting behavior in the London Metropolitan Area. *Annals of the Association of American Geographers*, 58, 1, 111–127.
- Cox K. R., Zannaras G. 1970. Designative perceptions of macrospace: concepts, a methodology and applications. Discussion Paper No 17, Department of Geography, The Ohio State University, Columbus.
- Czyż T. 1970. Zastosowanie metody czynnikowej w badaniach przestrzenno-ekonomicznych. *Przeгляд Geograficzny*, t. XLII, 3, 467–486.
- Czyż T. 1971. Zastosowanie metody analizy czynnikowej do badania ekonomicznej struktury regionalnej Polski. *Prace Geograficzne IG PAN*, nr 92. Wrocław.
- Dahms F. A. 1971. Urban residential structure: some neglected factors. *New Zealand Geographer*, 27, 2, 130–150.
- Davies W. K. D. 1972. Varimax and generality: second reply. *Area*, 4, 3, 207–208.

- Dawson J. A. 1972. Urban and rural retail structures in the East Midland Region of England. *Geographia Polonica*, 24, 157-172.
- Dieleman F. M. 1971. Factor analyse en multidimensionale groepering. *Tijdschrift voor Economische en Sociale Geografie*, 52, 4, 217-225.
- Donaldson B., Johnston R. J. 1973. Intraurban sectoral maps: further evidence from an extended methodology. *Geographical Analysis*, 5, 1, 45-54.
- Gauthier H. L. 1968. Transportation and the growth of the Sao Paulo economy. *Journal of Regional Science*, 8, 1, 77-94.
- Golant S., Burton I. 1970. A semantic differential experiment in the interpretation and grouping of environmental hazards. *Geographical Analysis*, 2, 2, 120-133.
- Greenberg M. R., Boswell T. D. 1972. Neighborhood deterioration as a factor in intraurban migration: a case study in New York City. *The Professional Geographer*, 24, 1, 11-15.
- Harvey D. 1969. *Explanation in geography*. London.
- Hayness K. E. 1971. Spatial change in urban structure: alternative approaches to ecological dynamics. *Economic Geography*, 47, 2, 324-335.
- Horton F. E., Louviere J., Reynolds D. R. 1973. Mass transit utilization: individual response data inputs. *Economic Geography*, 49, 2, 122-133.
- Janson C. G. 1971. A preliminary report on Swedish urban spatial structure. *Economic Geography*, 47, 2, 249-257.
- Jaumotte Ch., Paelinck J. H. P., Leheureux J. M., Pietquin M. 1971. The differential economic structures of the Belgian Provinces: a time varying factor analysis. *Regional and Urban Economics*, 1, 1, 41-75.
- Jeffrey D., Casetti E., King L. J. 1969. Economic fluctuations in a multiregional setting: a bifactor analytic approach. *Journal of Regional Science*, 9, 3, 397-404.
- Jelonk A. 1971. Regiony demograficzne Polski. *Zeszyty Naukowe Uniwersytetu Jagiellońskiego, Prace Geograficzne*, 30. Kraków.
- Johnston R. J. 1971. Some limitations of factorial ecologies and social area analysis. *Economic Geography*, 47, 2, 314-323.
- Jones K. J., Jones W. C. 1970. Toward a typology of American cities. *Journal of Regional Science*, 10, 2, 217-224.
- Kaiser H. F. 1970. A second generation Little Jiffy. *Psychometrika*, 35, 401-413.
- Kariel H. G., Kariel P. E. 1973. The continuum of technology: a non-grouping of countries. *Geografiska Annaler*, 55b, 1, 1-12.
- Kendall M. G. 1968. *A course in multivariate analysis*. London.
- King L. J. 1969. *Statistical analysis in geography*. Englewood Cliffs.
- Klován J. E., Imbrie J. 1971. An algorithm and Fortran-IV program for large-scale Q-mode factor analysis and calculation of factor scores. *Journal International Association for Mathematical Geology*, 3, 1, 61-77.
- Lawley D. N., Maxwell A. E. 1963. *Factor analysis as a statistical method*. London.
- Mather P. M. 1971. Comment on „Varimax and the destruction of generality”. *Area*, 3, 252-254.
- Mather P. M. 1972. Varimax and generality. *Area*, 4, 1, 27-30.
- Mc Donald R. P. 1962. A general approach to nonlinear factor analysis. *Psychometrika*, 27, 397-415.
- Mc Nown R. F., Singell J. D. 1974. A factor analysis of the socioeconomic structure of riot and crime prone cities. *The Annals of Regional Science*, 8, 1, 1-10.
- Meyer D. R. 1971. Factor analysis versus correlation analysis: are substantive interpretations congruent? *Economic Geography*, 47, 2, 337-343.
- Miedwiedkow J. W. 1974. Modeli ekologii człowieka w geografii. [W:] *Teorija i metodika geograficzeskich issledowanij ekologii czelowieka*. Moskwa, s. 49-74.
- Monmonier M. S. 1973. Maximum-difference barriers: an alternative numerical regionalization method. *Geographical Analysis*, 5, 3, 245-254.
- Morris F. B., Pyle G. F. 1971. The social environment of Rio de Janeiro in 1960. *Economic Geography*, 47, 2, 286-299.
- Morrison D. F. 1967. *Multivariate statistical methods*. Nowy Jork.
- Munton R. J. C., Norris M. 1969. The analysis of farm organization: an approach to the classification of agricultural land in Britain. *Geografiska Annaler*, 51B, 2, 95-103.
- Muraco W. A. 1972. Intraurban accessibility. *Economic Geography*, 48, 4, 388-405.
- Nowakowska M. 1973. Analiza czynnikowa: arbitralne decyzje w ramach modelu matematycznego. *Studia Socjologiczne*, 4, 51, 24-40.
- Ortmann B. 1971. Delimitation of economic planning regions in Western Germany. *Tijdschrift voor Economische en Sociale Geografie*, 52, 5, 308-317.
- Palm R. 1973. Factorial ecology and the community of outlook. *Annals of the Association of American Geographers*, 63, 3, 341-346.
- Palm R., Caruso D. 1972. Factor labelling in factorial ecology. *Annals of the Association of American Geographers*, 62, 1, 122-133.
- Parker D. 1971. A classical social area analysis: New Castle, N. S. W. and some comparisons. *The Australian Geographer*, 11, 6, 555-578.
- Piasecki Z. 1969. Analiza czynnikowa. *Metoda Hotellinga*. Program na maszynę cyfrową Gier. Centrum Obliczeniowe PAN, Warszawa.
- Pitts F. R. 1971. Factorial ecology of Seoul and Taegu, Korea: a preliminary report. *Economic Geography*, 47, 2, 300-302.
- Ray D. M. 1971. The location of United States manufacturing subsidiaries in Canada. *Economic Geography*, 47, 3, 389-400.
- Rees P. H. 1971. Factorial ecology: an extended definition, survey, and critique of the field. *Economic Geography*, 47, 2, 220-233.
- Rummel R. J. 1970. *Applied factor analysis*. Evanston.
- Saunders R. J. 1971. Population flows, spatial economic activity and urban areas in Appalachia. *The Annals of Regional Science*, 5, 3, 125-136.
- Schilderlinck J. H. P. 1970. Factor analysis applied to developed and developing countries. *Studies on Economics*, 1, Tilburg Institute of Economics, Groningen.
- Semple R. K., Gauthier H. L., Youngmann C. E. 1972. Growth poles in Sao Paulo, Brazil. *Annals of the Association of American Geographers*, 62, 4, 591-598.
- Soen D., Bieger J., Hönig M., Tamir M., Run M. 1973. Israeli urban settlement — an attempt at scaling and ranking according to population quality index. *Geoforum*, 13, 63-69.
- Stutz F. B. 1973. Distance and network effects on urban social travel fields. *Economic Geography*, 49, 2, 134-144.
- Tucker L. R. 1963. Implications of factor analysis of three-way matrices for measurement of change. [W:] *Problems in measuring change*, ed. C. W. Harris. Madison, s. 122-137.
- Van der Knaap G. A. 1971. Een indeling van Nederland naar „Economische Gezondheid”: een verkennend onderzoek. *Tijdschrift voor Economische en Sociale Geografie*, 62, 6, 332-350.
- Zelinsky W. 1974. Selfward bound? Personal preference patterns and the changing map of American society. *Economic Geography*, 50, 2, 144-179.
- Żukowskaja M. W. 1973. Primienienije metodow mnogomiernoj klasifikacji w zadacze tipologii stran. [W:] *Matiematiczeskije metody w ekonomikie i miezdunarodnych odnoszenijach*, red. W. Żukowskaja. Instytut Mirowoj Ekonomiki i Miezdunarodnych Otnoszenii. Akademia Nauk SSSR. Moskwa.