

INSTYTUT GEOGRAFII
POLSKIEJ AKADEMII NAUK

PRACE GEOGRAFICZNE Nr 92

TERESA CZYŻ

ZASTOSOWANIE
METODY ANALIZY CZYNNIKOWEJ
DO BADANIA
EKONOMICZNEJ STRUKTURY
REGIONALNEJ POLSKI

WROCŁAW • WARSZAWA • KRAKÓW • GDAŃSK
ZAKŁAD NARODOWY IMIENIA OSSOLIŃSKICH
WYDAWNICTWO POLSKIEJ AKADEMII NAUK

PRACE GEOGRAFICZNE IG PAN

51. Kostrowicki A. S., Regionalizacja zoogeograficzna Palearktyki w oparciu o faunę motyli tzw. większych (Macrolepidoptera). 1965, s. 100+21 ilustr., zł 30,—
52. Gerlach T., Współczesny rozwój stoków w dorzeczu górnego Grajcarka (Beskid Wysoki — Karpaty Zachodnie). 1966, s. 111, 20 ilustr., zł 33,—
53. Klimek K., Deglacjacja północnej części Wyżyny Śląsko-Krakowskiej w okresie zlodowacenia środkowopolskiego. 1966, s. 136+26 ilustr., zł 32,—
54. Kosmowska-Suffczyńska D., Rozwój rzeźby w trzeciorzędzie okolic Ostrowca Świętokrzyskiego i Ćmielowa. 1966, s. 114+22 ilustr. +7 fot. +2 mapy, zł 33,—
55. Ziemońska Z., Obieg wody w obszarze górskim na przykładzie górnej części dorzecza Czarnego Dunajca. 1966, s. 111+16 ilustr. +2 wkładki, zł 34,—
56. Ratajski L., Mapy przemysłu, ich właściwości metodyczne i kartometryczne. 1966, s. 115+22 ilustr., zł 28,—
57. Więckowski K., Osady denne Jeziora Mikołajskiego. 1966, s. 112+12 ilustr. +7 fot., zł 24,—
58. Szostak M., Pochodzenie Jeziora Śniardwy i jego zasoby wodne. 1967, s. 70+11 ilustr., zł 20,—
59. Rościszewski M., Siemek Z., Rolnictwo krajów słabo rozwiniętych (Egipt, Syria, Turcja). 1967, s. 109+9 ilustr., zł 24,—
60. Ziętara T., Rola gwałtownych ulew i powodzi w modelowaniu rzeźby Beskidów. 1968, s. 116+12 ilustr. + 11 fot., zł 33,—
61. Urbaniak U., Wydmy Kotliny Płockiej. 1967, s. 79+43 ilustr. +8 fot., zł 21,—
62. Jewtuchowicz S., Geneza Pradoliny Warszawsko-berlińskiej między Nerem a Moszczenicą. 1967, s. 102+42 ilustr. +19 fot., zł 30,—
63. Dziewoński K., Baza ekonomiczna i struktura funkcjonalna miast. Studium rozwoju pojęć, metod i ich zastosowań. 1967, s. 135, zł 32,—
64. Rychłowski B., Regionalizacja ekonomiczna — zagadnienia podstawowe. 1967, s. 139, zł 33,—
65. Bączyk J., Masy wodne południowego Bałtyku i wpływ ich ruchów na polską strefę przybrzeżną. 1968, s. 120+32 ilustr., zł 31,—
66. Szulc H., Typy wsi Śląska Opolskiego na początku XIX wieku i ich geneza. 1968, s. 105+14 ilustr. +2 wkładki, zł 30,—
67. Szewczyk J., Włóka. Pojęcie i termin na tle innych średniowiecznych jednostek pomiaru ziemi. 1968, s. 113+6 ilustr., zł 30,—
68. Wojciechowski K., Zagadnienie metody bilansu wodnego Thornthwaite'a i Mathera w zastosowaniu do Polski. 1968, s. 79+23 ilustr., zł 18,—
69. Praca zbiorowa. Problemy regionalizacji fizycznogeograficznej. Materiały z sympozjum, zorganizowanego przez PTG w dniach 16—24 września 1966. 1968, s. 114+4 ilustr. +1 wkładka, zł 28,—
70. Pulina M., Zjawiska krasowe we wschodniej Syberii. 1968, s. 94+34 ilustr. +4 fot., zł 19,—
71. Szupryczyński J., Niektóre zagadnienia czwartorzędu na obszarze Spitsbergenu. 1968, s. 127+15 ilustr. +35 fot. +1 wkładka, zł 34,—
72. Kosiński L., Migracje ludności w Polsce w latach 1950—1960. 1968, s. 106+41 ilustr., zł 28,—
73. Korolec H., Procesy brzegowe i zmiany linii brzegowej Jeziora Mikołajskiego. 1968, s. 67+16 ilustr. +6 fot. +1 wkładka, zł 24,—
74. Praca zbiorowa. Ostatnie zlodowacenie skandynawskie w Polsce. 1968, s. 216+12 ilustr. +11 fot., zł 67,—
75. Praca zbiorowa. Procesy i formy wydymowe w Polsce. Zbiór prac pod redakcją R. Galona. 1969, s. 386+69 ilustr. +68 fot., zł 98,—
76. Iwanicka-Lyra E., Delimitacja aglomeracji wielkomiejskich w Polsce. 1969, s. 117+12 ilustr., zł 28,—

INSTYTUT GEOGRAFII
POLSKIEJ AKADEMII NAUK



PRACE GEOGRAFICZNE NR 92

ГЕОГРАФИЧЕСКИЕ ТРУДЫ
№ 92

ТЕРЕСА ЧИЖ

ПРИМЕНЕНИЕ ФАКТОРНОГО АНАЛИЗА
ДЛЯ ИЗУЧЕНИЯ
ЭКОНОМИЧЕСКОЙ РАЙОННОЙ СТРУКТУРЫ ПОЛЬШИ

*

GEOGRAPHICAL STUDIES
№ 92

TERESA CZYŻ

THE APPLICATION OF FACTOR ANALYSIS IN THE
STUDY OF POLAND'S ECONOMIC REGIONAL
STRUCTURE

INSTYTUT GEOGRAFII
POLSKIEJ AKADEMII NAUK

PRACE GEOGRAFICZNE NR 92

TERESA CZYŻ

ZASTOSOWANIE
METODY ANALIZY CZYNNIKOWEJ
DO BADANIA
EKONOMICZNEJ STRUKTURY
REGIONALNEJ POLSKI

WROCLAW • WARSZAWA • KRAKÓW • GDAŃSK
ZAKŁAD NARODOWY IMIENIA OSSOLIŃSKICH
WYDAWNICTWO POLSKIEJ AKADEMII NAUK

1971

Komitet redakcyjny

REDAKTOR NACZELNY: M. KIEŁCZEWSKA-ZALESKA
ZASTĘPCA REDAKTORA NACZELNEGO: K. DZIEWOŃSKI
CZŁONKOWIE: R. GALON, L. STARKEL
SEKRETARZ: I. STAŃCZAK

Praca wykonana pod kierunkiem
doc. dra hab. ZBYSZKO CHOJNICKIEGO

Redaktor wydawnictwa Zofia Pawłowska

Redaktor techniczny Wanda Szajkiewicz

Printed in Poland

Zakład Narodowy im. Ossolińskich — Wydawnictwo. Wrocław 1971.
Nakład: 870+130 egz. Objętość: ark. wyd. 14,30, ark. druk. 7,13+15 wkł., ark.
A1. 9. Papier powlek. kl. V, 80 g, 70×100. Oddano do składania
17 IV 1971. Podpisano do druku 30 IX 1971. Druk ukończono w listop. 1971.
Wrocławska Drukarnia Naukowa. — Nr zam. 630/71 — K-8 — Cena zł 38.—

SPIS TREŚCI

Wstęp	7
1. Cel pracy	7
2. Założenia badawcze	9
I. Zagadnienie analizy czynnikowej	17
1. Formalizacja zagadnienia analizy czynnikowej	18
2. Podstawy matematyczno-statystyczne analizy czynnikowej	19
3. Metody wyodrębniania czynników	24
4. Rotacja czynników	31
5. Interpretacja czynników	33
6. Alternatywne techniki analizy czynnikowej	34
II. Zastosowania metody czynnikowej w badaniach przestrzenno-ekonomicznych	35
1. Badania empiryczne	35
2. Analiza czynnikowa jako metoda badawcza w geografii ekonomicznej	48
III. Studium ekonomicznej struktury regionalnej Polski	53
1. Koncepcja ukrytej struktury przestrzeni społeczno-ekonomicznej w świetle analizy czynnikowej	53
2. Procedura regionalizacji ekonomiczno-geograficznej	54
2.1. Macierz obserwacji	54
A. Interpretacja empiryczna macierzy	54
B. Sformułowanie matematyczne	61
C. Rozkłady zmiennych	62
2.2. Analiza przestrzeni wieloczechowej	64
A. Wstępna analiza korelacji zmiennych metodą dendrytową	64
B. Analiza korelacji metodą analizy czynnikowej	66
B 1. Struktura czynnikowa przestrzeni wieloczechowej	66
B 2. Identyfikacja podstawowych czynników przestrzeni wieloczechowej	69
2.3. Zastosowanie czynników do rozwiązania taksonomicznego zagadnienia regionalizacji	73
A. Klasyfikacja jednoczynnikowa	74
B. Klasyfikacja dwuczynnikowa	78
C. Klasyfikacja wieloczynnikowa	81
C 1. Typologia przestrzenna	81
C 2. Klasyfikacja regionalna	82

3. Wstępna hipoteza o strukturze regionalnej Polski	94
4. Zróżnicowanie strefowe a układ wojewódzki	97
Uwagi końcowe	100
Literatura	104
Применение факторного анализа для изучения экономической районной структуры Польши (Резюме)	111
The application of factor analysis in the study of Poland's economic regional structure (Summary)	113
Mapy 1—11 na końcu książki pod opaską	

WSTĘP

1. CEL PRACY

Celem pracy jest poznanie ekonomicznej struktury regionalnej Polski z punktu widzenia jednolitości.

U podstaw niniejszego studium leży koncepcja ukrytej struktury przestrzeni społeczno-ekonomicznej. Jako metodę zidentyfikowania tej struktury wprowadza się analizę czynnikową.

Praca składa się z dwóch zasadniczych części, metodologicznej i empirycznej.

Pierwsza część zawiera: 1) omówienie podstaw matematyczno-statystycznych metody analizy czynnikowej w badaniach układów wielocechowych; 2) przegląd dotychczasowych zastosowań metody czynnikowej w badaniach przestrzenno-ekonomicznych i na tym tle przeprowadzoną charakterystykę analizy czynnikowej jako metody badawczej w geografii ekonomicznej, opartą o różne koncepcje jej zastosowania; 3) określenie, na podstawie powyższych rozważań, roli analizy czynnikowej jako metody identyfikacji i redukcji wymiarów przestrzeni społeczno-ekonomicznej.

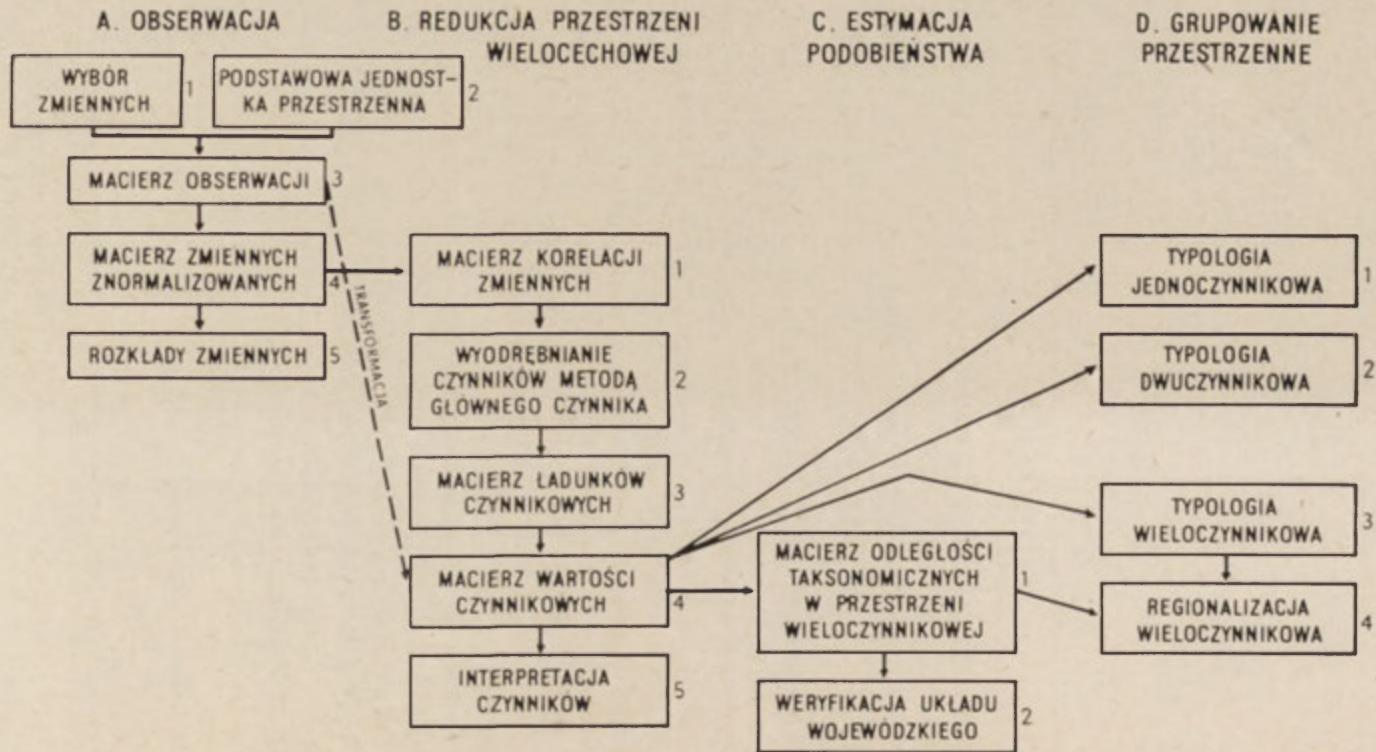
Tak obszerne ujęcie tych zagadnień podyktowane jest małą znajomością metody analizy czynnikowej w polskiej geografii ekonomicznej.

Druga część pracy ma charakter empiryczno-poznawczy. Stanowi właściwe studium ekonomicznej struktury regionalnej Polski, które sprowadza się do wyznaczenia układu regionalnego i analizy zróżnicowania tego układu.

Metoda rozwiązania czynnikowego zagadnienia regionalizacji ma postać algorytmu o charakterze uniwersalnym, który uwzględnia cztery zasadnicze czynności (ryc. 1):

- A. Obserwację
- B. Redukcję przestrzeni wielocechowej
- C. Estymację podobieństwa
- D. Grupowanie przestrzenne.

W oparciu o materiał statystyczny, obejmujący 33 cechy społeczno-ekonomiczne dla 324 podstawowych jednostek przestrzennych na obszarze Polski w 1965 r., wyodrębnia się metodą głównych czynników H. Hotellinga istotne wymiary rzeczywistości społeczno-gospodarczej kraju. Następnie dokonuje się oszacowania



Ryc. 1. Algorytm typologii przestrzennej i regionalizacji ekonomiczno-geograficznej przy zastosowaniu metody analizy czynnikowej

Fig. 1. Spatial typology and economic regionalization algorithm based on the application of factor analysis

A. Observation: 1 — selection of variables, 2 — identification of areal unit, 3 — data matrix, 4 — standard score matrix, 5 — distributions of variables; B. Reduction of a multivariate space: 1 — correlation matrix, 2 — extraction of factors by principal factor method, 3 — factor matrix, 4 — factor score matrix, 5 — interpreting factors; C. Estimation of similarity: 1 — distance matrix, 2 — verification of the administrative pattern of voivodships; D. spatial grouping: 1 — one-factor typology, 2 — two-factor typology, 3 — multifactor typology, 4 — multifactor regionalization

podobieństwa oraz grupowania i regionalizacji jednostek w przestrzeni czynników wspólnych przy zastosowaniu metod taksonomii numerycznej w postaci odległości taksonomicznej, wskaźnika przyrodniczego J. Perkala i dendrytowej metody najbliższego sąsiada.

Badanie dyferencjacji struktury przestrzeni społeczno-ekonomicznej w ujęciu czynnikowym na podstawie cech odzwierciedlających właściwości rzeczywistości społeczno-gospodarczej ma na celu: 1) wykrycie istotnych układów zjawisk w przestrzeni Polski w języku przestrzeni cząstkowych oraz ustalenie ich hierarchii; 2) określenie stopnia jednorodności tych przestrzeni i ustalenie, czy mają regularną strukturę regionalną; 3) zbadanie pionowych i poziomych relacji przestrzeni cząstkowych; 4) próbę integracji tych układów prostych w układ złożony struktury regionów jednolitych Polski; 5) zbadanie stosunku istniejącej struktury regionalnej do podziału wojewódzkiego kraju.

2. ZAŁOŻENIA BADAWCZE

Do podstawowych zadań poznawczych geografii ekonomicznej należy analiza ogólnej przestrzeni społeczno-ekonomicznej. Badanie prawidłowości w zakresie występowania zjawisk społeczno-ekonomicznych w tej przestrzeni jest ściśle związane z koncepcją regionalną. Regionalizacja stanowi tylko początkowe stadium w porządkowaniu i formułowaniu generalizacji przestrzennych, jednak ze względu na dotychczasowy słaby rozwój (szczególnie od strony metodologicznej) pozostaje nadal istotnym zagadnieniem geografii ekonomicznej.

Zgodnie z nowoczesną koncepcją regionalizację traktuje się jako samą czynność podziału przestrzeni na regiony ekonomiczne w oparciu o właściwe i efektywne metody naukowe. Prowadzi ona do wykrycia elementów składowych przestrzeni społeczno-ekonomicznej, które tworzą specyficzną jej strukturę, nazywaną strukturą regionalną. Studium z zakresu ekonomicznej struktury regionalnej trzeba więc rozpocząć od postępowania regionalizacyjnego.

Regionalizacja implikuje uprzednie wyjaśnienie zakresu terminu „region”. K. Dziewoński (1967, s. 57) wyróżnił, dokonując integracji poglądów teoretycznych, trzy podstawowe znaczenia regionu, wzajemnie ze sobą powiązane: region-narzędzie badania, region-narzędzie działania, region-przedmiot poznania.

Region w sensie poznawczym traktuje się jako przedmiot badania. W tym znaczeniu region ekonomiczny stanowiąc formę generalizacji cech przestrzeni jest narzędziem analizy geograficzno-ekonomicznej. Ta definicja regionu nie ma jednak charakteru operacyjnego i nie może być w tej postaci podstawą regionalizacji. Bardziej przydatne praktycznie jest określenie D. Whittleseya (1954), które stanowi próbę zawężenia zakresu pojęcia regionu-narzędzia analitycznego. W jego ujęciu region to zwarty obszar, jednorodny w zakresie określonych kryteriów odnoszących się do współwystępowania pewnego zespołu cech, między którymi istnieją związki przyczynowe. Jednorodność można rozpatrywać ze względu na: 1) podo-

bieństwo cech podstawowych jednostek przestrzennych sąsiadujących ze sobą lub 2) stopień ścisłości powiązań. Stąd wyróżnia się dwa podstawowe typy regionów: a) jednolite (regiony formalne), odpowiadające układom rozmieszczenia zjawisk, b) węzłowe (nodalne, funkcjonalne, organizacji przestrzennej), odpowiadające układom powiązań.

W pierwszym okresie rozwoju regionalizacji ekonomicznej kryteriów wydziałania regionów poszukiwano na podstawie jednorodności przestrzennej występowania określonych zjawisk społeczno-ekonomicznych¹. Przykładem tego rodzaju ujęć były próby regionalizacji ekonomicznej w międzywojennej literaturze polskiej².

Z czasem koncepcji jednolitości, ujętej w kategoriach strefowości, została przeciwstawiona koncepcja powiązań przestrzennych. Przeciwstawienie to nie mogło mieć jednak charakteru bezwzględny. Każda z tych koncepcji służy bowiem oświetleniu różnych aspektów rzeczywistości społeczno-gospodarczej.

Współcześnie teoria pola zakłada integrację tych dwóch koncepcji regionalizacji i operuje twierdzeniem, że struktura przestrzeni społeczno-ekonomicznej jest złożona; obejmuje układy powiązań i układy rozmieszczeń wzajemnie powiązane i w zasadzie izomorficzne³.

U podstaw teorii regionalizacji leży koncepcja pojmowania regionalizacji jako wnej formy klasyfikacji, rozwinięta w ostatnich latach przez G. de Jonga (1962), pe. Bunge (1962), D. Grigga (1965, 1967), A. Wróbla (1967), Z. Wysockiego (1964, 1968), Z. Chojnickiego (1970). Co prawda przyjmuje się, że regionalizacja ze względu na aspekt przestrzenny jest metodą badawczą właściwą geografii; jednak równocześnie dowodzi się, że stanowi procedurę analogiczną do stadium klasyfikacji w innych naukach⁴.

Między regionalizacją i klasyfikacją występują istotne podobieństwa, jednak procedury i pojęcia tych dwóch dziedzin nie pozostają w jednoznacznej odpowiedniości. Regionalizacja ze względu na przedmiot badania i cel naukowy poznania jest z natury geograficzna i tylko pod względem formalnym utożsamia się z klasyfikacją. Klasyfikacja jako czynność bardziej precyzyjna, lepiej znana i zorganizowana teoretycznie może wystąpić w roli modelu teoretycznego regionalizacji.

Akceptacja tego poglądu upoważnia do rozważania regionalizacji w terminach klasyfikacji.

Region rozumiany *ex definitione* jako pewien zwarty obszar, jednolity w zakresie określonych kryteriów, jest klasą obszarów. W tym ujęciu regionalizacja jako proces

¹ A. Wróbel (1956).

² J. Piekałkiewicz, S. Rutkowski (1927), J. Czyżewski (1930), J. Ernst (1932) i inne.

³ B. J. L. Berry (1966).

⁴ Należy zwrócić uwagę, że w geografii problemy klasyfikacji występują w dwóch formach. Pierwsza forma jest grupowaniem obserwacji, które wykazują podobieństwo w zakresie cech, np. klasyfikacja miast na podstawie struktury zatrudnienia albo klasyfikacja przemysłu na podstawie preferencji lokalizacyjnych. Najbardziej charakterystyczna dla badań geograficznych jest druga forma klasyfikacji — regionalizacja.

wyodrębniania klas odbywa się poprzez czynności klasyfikacji lub podziału logicznego. Jako odpowiednik klasyfikacji przyjmuje się regionalizację syntetyczną, a jako odpowiednik podziału logicznego — regionalizację analityczną.

Należy zaznaczyć, że w praktyce badawczej szczególnie w odniesieniu do względnie małych obszarów znalazła zastosowanie i uzasadnienie przede wszystkim regionalizacja syntetyczna.

Regionalizacja syntetyczna jest procedurą grupowania, mającą na celu wyodrębnienie ciągłych zbiorów jednolitych podstawowych jednostek przestrzennych.

Punkt wyjściowy w regionalizacji stanowi podstawowa jednostka przestrzenna, traktowana jako „indywiduum” klasyfikowanej przestrzeni. Jednostka przestrzenna w przeciwieństwie do indywiduum ze schematu klasyfikacji nie tworzy wyraźnie wyodrębnionej całości, gdyż powierzchnia ziemi jest ciągła. Określenie obszaru jednostkowego zawiera zawsze element arbitralności. Z reguły podstawowa jednostka przestrzenna jest definiowana w kontekście konkretnej procedury regionalizacyjnej. W regionalizacji ekonomiczno-geograficznej stosuje się najczęściej najmniejsze jednostki podziału administracyjnego, dla których posiada się dane statystyczne.

Jednostki te są nieregularne; zróżnicowane są bowiem pod względem wielkości i kształtu. Poza tym zakłada się, że elementarna jednostka jest przestrzennie niepodzielna. Nie uwzględniając zmienności zjawisk w obszarze wewnętrznym jednostki, traktuje się ją właściwie jako punkt. Dane statystyczne odnoszą się bowiem do całości obszaru jednostkowego i nie podlegają dezagregacji na podstawie jego dalszego podziału przestrzennego. W związku z tym powstają zniekształcenia pomiaru przestrzennego, co w następstwie obniża wartość ustalanego układu regionalnego. Oczywiście, że gęsta siatka analizowanych jednostek przestrzennych prowadzi do bardziej precyzyjnego uchwycenia granic regionalnych. Zniekształceniu pomiaru przestrzennego zapobiega w pewnym stopniu standaryzacja nieregularnych obszarów⁵.

Dane liczbowe przypisane do jednostki przestrzennej-punktu dotyczą wartości cech zjawisk społeczno-ekonomicznych występujących na jej obszarze.

W tym miejscu należy zaznaczyć, że w regionalizacji — w odróżnieniu od klasyfikacji — oprócz właściwości podstawowych jednostek przestrzennych bierze się równocześnie pod uwagę ich lokalizację przestrzenną. Konieczność uwzględnienia położenia, unikalnej cechy jednostki przestrzennej, decyduje o specyfice zagadnienia regionalizacji w porównaniu z klasyfikacją.

Procedurę grupowania według wzajemnych różnic i podobieństw między podstawowymi jednostkami przestrzennymi w zakresie badanych zjawisk identyfikuje się z opisem klasyfikującym szeregującym (z relacją porządkującą), czyli z klasyfikacją typologiczną⁶. Ta operacja w języku analizy przestrzennej, zwana typolo-

⁵ Patrz T. J. Coppock (1960), A. Robinson (1956), O. D. Duncan, P. P. Cuzzort, B. Duncan (1961), E. Thomas, D. L. Anderson (1965).

⁶ Z. Chojnicki (1970).

gią przestrzenną, grupuje w klasę (typ) jednolite jednostki przestrzenne, występujące bez ciągłości przestrzennej, a więc przemieszane z jednostkami zaliczanymi do innych klas. Należy zwrócić uwagę, że jedynie w przypadku wyjątkowym, kiedy klasy występują jako zgrupowania zwarte przestrzennie, typologia przestrzenna jest równocześnie regionalizacją. A więc typologia przestrzenna może pośrednio prowadzić do regionalizacji, ale nie zawsze każda typologia jest regionalizacją.

Pojęcie regionu, oparte na założeniu porządku przestrzennego, wymaga jednolitości i przylegania jednostek przestrzennych.

Regionalizacja jest więc specyficzną typologią przestrzenną, grupowaniem z pewnymi ograniczeniami, wynikającymi z przymusu ciągłości przestrzennej klas (typów). W tym ujęciu regionalizację rozpatruje się jako procedurę dwustopniową. Obejmuje ona właściwą typologię przestrzenną oraz analizę rozmieszczenia geograficznego typów, która prowadzi do delimitacji zwartych jednostek wyższego rzędu, czyli regionów, przez przyjęcie szerszych przedziałów klasowych podobieństwa.

Regionalizacja zrozumiana jako procedura jednostopniowa może prowadzić do regionów ogólnych lub regionów specyficznych⁷.

W pierwszym przypadku regionalizacja sprowadza się do agregacji jednostek podobnych i przyległych przestrzennie. Przy takim postępowaniu wynikowe jednostki regionalne mogą być również klasami jednoelementowymi. W obrazie przestrzennym takiego układu regionalnego pewne klasy będą się powtarzać, a więc poszczególne jednostki regionalne nie cechuje unikalność. W tym ujęciu regionalizacja jest formą klasyfikacji.

W drugim przypadku chodzi o uzyskanie jednostek przestrzennych wyższego rzędu unikalnych i specyficznych. Charakter regionu specyficznego jest określony nie tylko przez warunki wewnętrzne, lecz również przez położenie geograficzne. Zakłada się, że region jest „indywidualnością” o sobie tylko właściwej kombinacji cech, niepowtarzalnej w przestrzeni. Według D. Grigga (1965) regiony specyficzne nie tworzą systemów regionalnych, a ich przydatność w formułowaniu indukcyjnych generalizacji wydaje się dyskusyjna.

Interesującą konkluzją przedstawionych wyżej ustaleń na temat analogii między regionalizacją i klasyfikacją jest to, że w konstrukcji i oszacowaniu wartości układów regionalnych mogą być przydatne zasady klasyfikacji.

Sformułowania zasad klasyfikacji w odniesieniu do regionalizacji dokonał D. Grigg (1965). Jednak zdaniem jego zasady te dotyczą raczej strategii niż taktyki regionalizacji, a ponadto nie pozwalają eliminować elementu subiektywizmu.

Regionalizacja nie ma dotąd opracowanej szczegółowej metodologii opartej na zastosowaniu metod matematyczno-statystycznych. Ponieważ w postępowaniu badawczym w zakresie regionalizacji istotną rolę odgrywa typologia przestrzenna, zachodzi możliwość wstępnego rozwiązania problemu regionalizacji za pomocą znanych metod taksonomii numerycznej, rozwiniętych w ramach metodologii nauk

⁷ D. Grigg (1965, 1967).

przyrodniczych⁸. Jednak w zasadniczym etapie regionalizacji jest dużo dowolności. Konieczność przestrzegania ciągłości przestrzennej prowadzi do wielkich trudności i wymaga nowych, ścisłych i efektywnych metod regionalizacji⁹.

Dotychczas problem uwzględnienia sąsiedztwa pod względem metodologicznym pozostaje nie rozwiązany. Formuluje się jedynie pewne ogólne zasady postępowania w tym zakresie.

Z kolei należy zastanowić się, w jakim sensie układ regionalny może być rozpatrywany jako model, a regionalizacja jako procedura budowy modeli.

Region jest abstrakcją rzeczywistości. Aby przedstawić rzeczywistość opisowo lub za pomocą symboli, trzeba upraszczać, nie można reprodukować wszystkich cech przestrzeni społeczno-ekonomicznej. Dzieje się tak dlatego, że rzeczywistość jest wysoce złożona ze względu na wielość uwarunkowań i powiązań przyczynowo-skutkowych zjawisk społeczno-ekonomicznych.

Stwierdzenie prowadzi do refleksji, że również układy regionalne są modelami. Układ regionalny jest oparty na wyselekcjonowanych właściwościach „regionalizowanych” indywiduali. Trzeba przy tym podkreślić, że miarą użyteczności danej generalizacji przestrzennej w zakresie opisu i wyjaśniania jest istotność właściwości, wybranych do przedstawienia rzeczywistości.

Tak rozumiany model układu regionalnego formuluje się dla celów dalszej analizy i weryfikacji.

Delimitacja obszarów jednolitych może być punktem wyjścia do analizy relacji między regionem — przedmiotem poznania (obszarem jednolitym) a regionem — narzędziem działania (województwem).

Problem transformacji regionalizacji w sensie poznawczym w regionalizację w sensie podziału administracyjnego jest jednak złożony¹⁰. Podziału administracyjnego nie wyprowadza się z samych właściwości przestrzeni, jest on oparty przede wszystkim na zasadach dobrego administrowania. Konstrukcja jego wymaga uwzględnienia dodatkowych przesłanek w postaci elementów arbitralnych, odnoszących się do wielkości, ilości, kształtu jednostek administracyjnych oraz pewnych wielkości zagregowanych w jednostkach, np. ludności. Podział administracyjny może być najwyżej korygowany na podstawie opisu generalizującego właściwości przestrzeni społeczno-ekonomicznej.

Z tego punktu widzenia wyniki badań struktury regionalnej stanowią podstawę do ewentualnej korekty podziału administracyjnego lub materiał wyjściowy do opracowania podziału kraju dla celów badawczych na regiony statystyczne¹¹.

Przegląd dotychczasowych prac dotyczących analizy struktury regionalnej Polski opartej na kryteriach jednolitości nasuwa szereg uwag natury merytorycznej

⁸ R. Sokal, P. Sneath (1964).

⁹ L. J. King (1969), s. 194—215.

¹⁰ A. Wróbel (1963, 1965).

¹¹ Por. A. Kukliński (1961), s. 7—44, oraz T. Lijewski (1963), s. 655—661.

i metodologicznej. Krytyczna ocena tych opracowań ze względu na różnorodność problemów musi mieć charakter wielopłaszczyznowy.

Przede wszystkim należy zwrócić uwagę na fakt, że większość podziałów regionalnych Polski jest związana z koncepcją regionalizacji cząstkowej. Nie negując wartości poznawczej regionalizacji cząstkowej, warto zwrócić uwagę na celowość podziału przestrzeni na regiony ogólnoeconomiczne. Regionalizacja ogólnoeconomiczna realizuje postulat całkowitego i strukturalnego ujmowania rzeczywistości społeczno-gospodarczej w jej przestrzennym aspekcie. Trzeba więc stwierdzić, że dopiero regionalizacja kompleksowa prowadzi do odwzorowania przestrzennego zróżnicowania całej rzeczywistości gospodarczej i jako taka stanowi narzędzie syntezy geograficznej¹².

Regionalizacja ma charakter generalizacji wyższego rzędu w stosunku do podstawowej jednostki przestrzennej, najczęściej jednostki podziału administracyjnego.

Na podstawie stopnia dokładności przestrzennej dotychczasowe podziały regionalne Polski można podzielić na podziały w układzie wojewódzkim i układzie powiatowym.

Województwa stanowią wyjściowe jednostki przestrzenne dla układów regionalnych K. Secomskiego (1956), K. Dziewońskiego (1957), S. Berezowskiego (1959, 1969), B. Rychłowskiego (1960), W. Kawalca (1965). Regionalizacja w układzie wojewódzkim prowadzi jednak do znacznej utraty szczegółowości w analizie przestrzennej zjawisk. Daje w rezultacie wielkie jednostki regionalne, o granicach pokrywających się z granicami województw, które nie są tworami jednorodnymi.

Układ odniesienia złożony z powiatów zapewnia realizację programu badawczego przy niższym poziomie generalizacji. Powiat ze względu na znacznie mniejszą powierzchnię w porównaniu z województwem można traktować jako jednostkę względnie jednolitą.

Układ powiatowy uwzględniono w klasycznej z punktu widzenia interesującej nas problematyki pracy J. Piekalkiewicza i S. Rutkowskiego (1927). Również W. Kawalec (1965) wydzielił mikroregiony, a A. Kukliński (1961) i T. Lijewski (1963) regiony statystyczne odpowiadające jednostkom ponadpowiatowym.

Aktualnie przeprowadzenie podziału regionalnego oderwanego od granic wojewódzkich, celem lepszego odwzorowania rzeczywistości społeczno-gospodarczej, umożliwi rozwój informacji statystycznej na szczeblu powiatu oraz postęp w zakresie elektronowych technik obliczeniowych.

Kryteria wyróżniania regionów ogólnoeconomicznych powinny mieć charakter kompleksowy i dotyczyć całości gospodarki na badanym obszarze. Ze względu na to, że ustalenie a priori znaczenia poznawczego pojedynczych cech zjawisk społeczno-economicznych jest zadaniem trudnym, istnieje konieczność uwzględnienia

¹² Pojęciami „regionalizacja całkowita” (kompleksowa) i „regionalizacja cząstkowa” operuje J. Kostrowicki (1967), s. 27. B. Rychłowski (1965), s. 29—56, oraz (1967) jako ich synonimy traktuje regionalizację integralną i rodzajową.

w regionalizacji ogólnoeconomicznej dużego zbioru cech. Dla tego ujęcia reprezentatywna jest w literaturze polskiej praca J. Piekałkiewicza i S. Rutkowskiego (1927).

Z metodologicznego punktu widzenia dotychczasowe próby regionalizacji wielo cechowej Polski, oparte głównie na opisie kartograficznym charakteryzują się prymitywizmem. Z reguły identyfikacji układu regionów ogólnych dokonywano drogą demarkacji granic regionów prostych, jednocechowych i konstrukcji ujęcia syntetycznego na podstawie wizualnego porównania tych elementarnych układów. Takie rozwiązanie nie pozwalało uniknąć elementów subiektywnych i intuicyjnych.

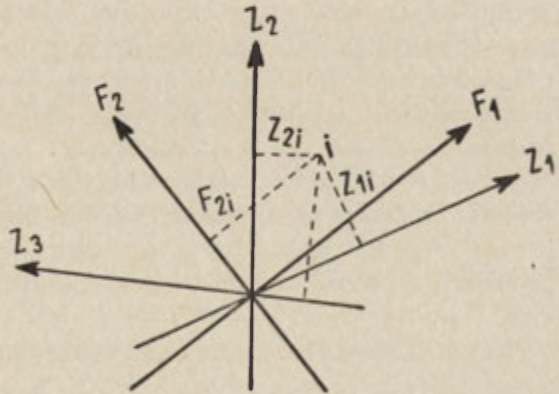
Uzyskanie w pełni wartościowej regionalizacji wielo cechowej Polski związane jest z zastosowaniem ścisłych i efektywnych metod badawczych w postaci technik statystycznych i algorytmów. Mimo słabego przenikania metod matematyczno-statystycznych do badań regionalizacyjnych istnieje już kilka prób wprowadzenia metod analizy statystycznej do wstępnej fazy regionalizacji Polski — typologii przestrzennej. Do nich należą prace Z. Wysockiego (1966), B. Kostrubca (1968), S. E. Browna i Ch. E. Trotta (1968).

I. ZAGADNIENIE ANALIZY CZYNNIKOWEJ

W analizie wielocechowej stwierdzenie podobieństwa w przebiegu rozkładu wartości poszczególnych zmiennych sugeruje, że niektóre zmienne „zachodzą” na siebie (pokrywają się), a więc różnicują obiekty w ten sam sposób. Istnienie korelacji między zmiennymi pozwala z kolei na sformułowanie hipotezy, że u podłoża zmienności zjawisk ukrywa się jakaś bardziej istotna struktura, której elementami są metacechy.

Analiza czynnikowa obejmuje grupę metod matematyczno-statystycznych pozwalających na sprowadzenie pierwotnego (wyjściowego) zbioru zmiennych, które charakteryzują obiekty poddane obserwacji, do znacznie mniejszej liczby hipotetycznych zmiennych, zwanych czynnikami. Te nowe zmienne zawierają podstawową informację o oryginalnych zmiennych. Analiza czynnikowa wyodrębnia więc czynniki (wymiary, składniki) leżące u podstaw korelacji stwierdzonych w danym zbiorze zmiennych. Te czynniki mogą być traktowane jako przyczyny zaobserwowanej zmienności; badacz interpretuje je jako mające ważne znaczenie dla pomiaru, opisu i wyjaśnienia zmienności.

Jeżeli rozpatruje się zbiór N obiektów (elementów) określonych za pomocą n cech (zmiennych niezależnych i zależnych) o wartościach rzeczywistych, to elementy tego zbioru można zinterpretować jako punkty przestrzeni euklidesowej n -wymiarowej. Jednak zbiór punktów, pomimo że jest wyrażony za pomocą n zmiennych, może w rzeczywistości leżeć, przynajmniej w przybliżeniu, w przestrzeni o mniejszej liczbie m wymiarów. Tak np. punkt położony w układzie trzech współrzędnych prostokątnych może



Ryc. 2. Interpretacja geometryczna zagadnienia czynnikowego

Fig. 2. The geometrical interpretation of the factorial problem

znaleźć się na jednej z płaszczyzn utworzonej przez dwie osie współrzędne (ryc. 2). Trzeba zatem określić m nowych zmiennych ($m < n$) i dokonać w ten sposób redukcji przestrzeni cechowej.

Jeżeli uwzględni się tylko m czynników i następnie rzutuje się zbiór punktów przestrzeni n -wymiarowej na podprzestrzeń rozpiętą na m czynnikach, to uzyska się możliwość scharakteryzowania każdego elementu zbioru za pomocą m nowych współrzędnych.

1. FORMALIZACJA ZAGADNIENIA ANALIZY CZYNNIKOWEJ

Model analizy czynnikowej konstruuje się jako założenie wstępne, które jest sformułowane w postaci układu równań

$$(1) \quad \begin{cases} z_1 = a_{11}F_1 + a_{12}F_2 + \dots + a_{1m}F_m + a_1U_1 \\ \dots \\ z_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jm}F_m + a_jU_j \\ \dots \\ z_n = a_{n1}F_1 + a_{n2}F_2 + \dots + a_{nm}F_m + a_nU_n \end{cases}$$

Zmienne empiryczne, wejściowe z_j ($j = 1, 2, \dots, n$) są wyrażone jako liniowe funkcje hipotetycznych (ukrytych) czynników F_m i U_j . F_1, F_2, \dots, F_m to zmienne niezależne zwane czynnikami wspólnymi, $a U_1, U_2, \dots, U_n$ są tzw. czynnikami swoistymi. Parametry równania a_{jm}, a_j ($j = 1, 2, \dots, n$) nazywają się ładunkami czynnikowymi i określają wagę danego czynnika w opisie zmiennych empirycznych.

Analiza czynnikowa jest liniowym modelem matematycznym. Rozwiązanie czynnikowe polega więc na konkretyzacji modelu, czyli nadaniu parametrom równań modelu liczbowych wartości. Jeżeli całkowicie skonkretyzuje się relację matematyczną, nadając parametrom określone wartości, wtedy można powiedzieć, że posiada ona pewną strukturę czynnikową.

Jednak w modelu czynnikowym nie można bliżej określić empirycznego odpowiednika czynnika. Model ten nie jest więc bezpośrednio sprawdzalny. Można go jedynie sprawdzić pośrednio na podstawie określenia stopnia korelacji między czynnikiem a obserwowalnymi cechami zjawisk. Model czynnikowy należy więc do klasy modeli pomiaru.

Należy zwrócić uwagę również na dwie możliwe interpretacje modelu czynnikowego, deterministyczną i probabilistyczną. W niniejszym opracowaniu stosuje się ujęcie deterministyczne, ponieważ w dalszych badaniach empirycznych operuje się danymi wyczerpującymi.

Analizowana populacja jest skończona. Liczba jednostek należących do tej populacji jest relatywnie mała; jest ona równa (w przybliżeniu) liczbie powiatów Polski. Gdy populacja jest mała i nie można z niej pobrać próbki o dowolnej liczebności, to nie ma sensu stosowanie metod probabilistycznych.

Badaniem statystycznym zostaje objęty cały zbiór, a więc wszystkie problemy są natury deterministycznej.

2. PODSTAWY MATEMATYCZNO-STATYSTYCZNE ANALIZY CZYNNIKOWEJ

W analizie czynnikowej przyjmuje się dwa założenia o zmiennych i czynnikach:

1. Zmienne i czynniki są znormalizowane.

2. Wszystkie czynniki, wspólne i specyficzne, są nie skorelowane, tzn. że ich współczynniki korelacji są równe zero, a więc w interpretacji geometrycznej czynniki te reprezentują osie do siebie prostopadłe.

Dany jest zbiór złożony z N obiektów. Dla każdego elementu tego zbioru zmierzono n cech (własności), czyli każdemu elementowi zbioru przyporządkowano układ n liczb rzeczywistych. Dana jest więc macierz obserwacji X składająca się z N wierszy i n kolumn. Macierz X przekształca się w macierz Z również typu $N \times n$, zawierającą znormalizowane wartości poszczególnych zmiennych.

Normalizacja zmiennych polega na wyrażeniu wartości danej zmiennej w liczbie odchyłeń standardowych, jaka je dzieli od średniej arytmetycznej. Normalizacji dokonuje się na podstawie wzoru

$$(2) \quad z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j},$$

gdzie x_{ij} = wartość zmiennej j dla indywiduum i , \bar{x}_j = średnia N wartości zmiennej j , s_j = odchylenie standardowe zmiennej j .

Rozkład wartości zmiennej sprowadzonej do skali mierzonej w jednostkach odchylenia standardowego ma średnią arytmetyczną równą zero, odchylenie standardowe (i wariancję) zaś równe jedności.

Normalizacja zmiennych pozwala porównywać cechy wyrażone w różnorodnych jednostkach mianowanych. Liczby niemianowane otrzymane w wyniku normalizacji pozostają w tym samym stosunku do siebie, w jakim pozostawały odpowiadające im liczby mianowane przed normalizacją.

Wariancję zmiennej j , wyrażoną w terminach wag czynników, zapisuje się w postaci

$$(3) \quad s_j^2 = \frac{\sum_{i=1}^N z_{ij}^2}{N} = a_{j1}^2 \frac{\sum_{i=1}^N F_{i1}^2}{N} + a_{j2}^2 \frac{\sum_{i=1}^N F_{i2}^2}{N} + \dots + a_{jm}^2 \frac{\sum_{i=1}^N F_{im}^2}{N} + a_j^2 \frac{1}{N} \sum_{i=1}^N U_{ij} +$$

$$+ 2 \left(a_{j1} a_{j2} \frac{\sum_{i=1}^N F_{i1} F_{i2}}{N} + a_{j,m-1} a_{jm} \frac{\sum_{i=1}^N F_{i,m-1} F_{im}}{N} + \right.$$

$$\left. + a_{j1} a_j \frac{\sum_{i=1}^N F_{i1} U_{ij}}{N} + \dots + a_{jm} a_j \frac{\sum_{i=1}^N F_{im} U_{ij}}{N} \right).$$

Ponieważ zakłada się, że zmienne i czynniki są unormowane, powyższe równanie przybiera postać

$$(4) \quad s_{z_j}^2 = 1 = a_{j1}^2 + a_{j2}^2 + \dots + a_{jm}^2 + a_j^2 + 2(a_{j1}a_{j2}r_{F_1, F_2} + \dots + a_{jm}a_jr_{F_m, U_j}).$$

Po uwzględnieniu założenia o ortogonalności czynników powyższe wyrażenie można zredukować do postaci

$$(5) \quad s_{z_j}^2 = 1 = a_{j1}^2 + a_{j2}^2 + \dots + a_{jm}^2 + a_j^2.$$

A więc wariancja zmiennej j równa się sumie kwadratów wag czynnikowych należących do tej zmiennej. Składniki tej sumy reprezentują porcje jednostkowej wariancji zmiennej z_j przypisane poszczególnym czynnikom. Na przykład a_{j1}^2 jest częścią wariancji zmiennej j , którą można wyjaśnić za pomocą czynnika F_1 ; a_{j2}^2 jest udziałem czynnika F_2 w wariancji zmiennej z_j . Całkowity udział czynnika F_i w wariancji wszystkich zmiennych określa wzór $\sum_{j=1}^n a_{ji}^2$. Natomiast wyrażenie

$$(6) \quad h_j^2 = a_{j1}^2 + a_{j2}^2 + \dots + a_{jm}^2$$

nazywa się zasobem zmienności wspólnej i jest tą częścią wariancji zmiennej j , która może być wyjaśniana za pomocą wspólnych czynników. a_j^2 nazywa się swoistością i jest tą częścią wariancji zmiennej j , która jest właściwa tylko tej jednej, określonej zmiennej. Stąd

$$(7) \quad s_j^2 = 1 = h_j^2 + a_j^2.$$

Należy zwrócić uwagę, że współczynnik korelacji między zmienną j a czynnikiem F_1 można zapisać symbolicznie:

$$(8) \quad r_{z_j F_1} = \frac{\sum z_j F_1}{N} = a_{j1} \frac{\sum F_1^2}{N} + a_{j2} \frac{\sum F_1 F_2}{N} + \dots + a_j \frac{\sum F_1 U_j}{N} = a_{j1} \frac{\sum F_1^2}{N} + a_{j2} r_{F_1 F_2} + \dots + a_j r_{F_1 U_j}.$$

Jeżeli czynniki są znormalizowane, to wtedy suma kwadratów ładunków czynników jest równa jedności i jeżeli czynniki są wzajemnie niezależne, to wszystkie sumy iloczynów stają się równe zero. Powyższe wyrażenie upraszcza się więc do postaci $r_{z_j F_1} = a_{j1}$ co oznacza, że waga czynnika jest współczynnikiem korelacji między zmienną a czynnikiem.

Według klasycznej koncepcji modelu czynnikowego istnieją dwie postaci tego modelu: 1) analiza głównych komponentów, 2) analiza czynnikowa.

W analizie głównych komponentów model czynnikowy można przedstawić jako

$$(9) \quad z_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jm}F_m \quad (j = 1, 2, \dots, n; n = m).$$

Ta koncepcja pojmowania postaci modelu czynnikowego opiera się więc na założeniu istnienia tylko czynników wspólnych. Dokonuje się analizy całkowitej wariancji zmiennej (równiej jedności) wyłącznie w kategoriach czynników wspólnych. Metoda czynnikowa w tym przypadku jest zasadniczo transformacją ortogonalną zbioru zmiennych z_1, z_2, \dots, z_n w zbiór nowych zmiennych F_1, F_2, \dots, F_n o tej samej liczebności. Ogólna zmienność czynników F_1, F_2, \dots, F_n równa się oryginalnej wariancji układu zmiennych z_1, z_2, \dots, z_n .

W analizie czynnikowej, która leży u podstaw metodologicznych niniejszej pracy, model ma postać

$$(10) \quad z_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jm}F_m + a_j U_j,$$

a więc zakłada istnienie czynników wspólnych i czynników specyficznych.

W analizie czynnikowej dokonuje się rozbicia całkowitej zmienności układu obserwacji wraz z różnymi interpretacjami jej składników, związanych z określonymi źródłami zmienności. Wariancja zmiennej obejmuje wariancję wspólną, wyjaśnioną przez czynniki wspólne, oraz wariancję związaną z czynnikiem specyficznym.

Jednak w samym rozwiązaniu czynnikowym interesuje nas przede wszystkim wariancja wspólna (mniejsza od jedności), a więc nie wyodrębnia się z reguły czynników specyficznych.

Należy jednak zaznaczyć, że rozwiązanie matematyczne analizy czynnikowej może być dokonane w ten sam sposób jak rozwiązanie głównych komponentów.

Podstawowym problemem analizy czynnikowej jest wyznaczenie układu czynników wspólnych F_t ($t = 1, 2, \dots, m$), co jest równoważne z określeniem dla każdego czynnika F_t odpowiadającego mu wektora a_t ($a_{1t}, a_{2t}, \dots, a_{nt}$), przy czym $t = 1, 2, \dots, m$, natomiast a_{jt} są wagami czynnika F_t w kolejnych zmiennych z_1, z_2, \dots, z_n . Bada się po prostu udziały czynników wspólnych w zmienności wspólnej określonej zmiennej z_j .

W rozwiązaniu czynnikowym wychodzi się od zmiennych

$$(11) \quad z'_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jm}F_m \quad (j = 1, 2, \dots, n),$$

które z założoną dokładnością przybliżą zmienne z_j . Wariancja tych zmiennych

$$(12) \quad s_j^2 = h_j^2$$

jest mniejsza od wariancji całkowitej s_j^2 o udział czynnika U_j i równa się zmienności wspólnej zmiennej z_j . Dla zmiennych z' można obliczyć współczynniki korelacji, które wynoszą:

$$(13) \quad r'_{jk} = \frac{1}{N} \sum_{i=1}^N z'_{ij} z'_{ik} = a_{j1} a_{k1} \frac{1}{N} \sum_{i=1}^N F_{i1}^2 + a_{jm} a_{km} \frac{1}{N} \sum_{i=1}^N F_{im}^2 + \\ + a_{j1} a_{k2} \frac{1}{N} \sum_{i=1}^N F_{i1} F_{i2} + \dots + a_{jm} a_{k,m-1} \frac{1}{N} \sum_{i=1}^N F_{im} F_{i,m-1}.$$

Po uproszczeniach wynikających z założeń otrzymuje się

$$r'_{jk} = a_{j1}a_{k1} + a_{j2}a_{k2} + \dots + a_{jm}a_{km} \quad (k, j = 1, 2, \dots, n).$$

Dla $j \neq k$ współczynniki korelacji r'_{jk} są równe współczynnikom r_{jk} , ponieważ z założenia czynnik U_j nie wpływa na korelację między zmiennymi; natomiast $r'_{jj} = h_j^2$.

Tak więc analiza czynnikowa jest w istocie rzeczą analizą macierzy korelacji z zasobami zmienności wspólnej na głównej przekątnej zamiast jedności. Ta macierz nosi nazwę zredukowanej macierzy korelacji \mathbf{R}'

$$(14) \quad \mathbf{R}' = \begin{bmatrix} h_1^2 & r_{12} & \dots & r_{1n} \\ r_{21} & h_2^2 & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & h_n^2 \end{bmatrix}.$$

Macierz \mathbf{R}' jest macierzą symetryczną stopnia n , zawierającą $\frac{1}{2}(n^2 - n)$ współczynników korelacji. Najczęściej używany jest współczynnik korelacji K. Pearsona według momentu iloczynowego

$$(15) \quad r_{jk} = \frac{\sum_{i=1}^N (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^N (x_{ij} - \bar{x}_j)^2 \sum_{i=1}^N (x_{ik} - \bar{x}_k)^2}}$$

lub

$$(15a) \quad r_{jk} = \frac{\sum_{i=1}^N z_{ij}z_{ik}}{N}.$$

W przypadku zmiennych jakościowych stopniowalnych, ujętych w formie rang, można się również posługiwać współczynnikiem korelacji rang Ch. Spearmana, wyrażonym wzorem

$$(16) \quad R_{jk} = 1 - \frac{6 \sum_{i=1}^N (x_{ij} - x_{ik})^2}{N(N^2 - 1)},$$

gdzie x_{ij} , x_{ik} są rangami.

Dodatni współczynnik korelacji oznacza, że zachodzi zgodna współzmiennność obu zmiennych; ujemny — wskazuje, że zachodzi współzmiennność jakby przeciwna.

bieżna. Zerowa lub bliska zeru wartość współczynnika oznacza, że między zmiennymi nie ma związku korelacyjnego.

Wartości h_j^2 nie są znane z góry, gdyż nie można ich uzyskać na drodze eksperymentalnej. Istnieją jednak sposoby szacowania h_j^2 . L. L. Thurstone (1935, s. 85–91) podaje kilka metod estymacji zasobu zmienności wspólnej. Najczęściej h_j^2 określa się na podstawie wzoru

$$(17) \quad h_j^2 = \frac{r_{jk} r_{jl}}{r_{kl}} \quad (k, j = 1, 2, \dots, n),$$

gdzie r_{jk} i r_{jl} są najwyższymi współczynnikami korelacji zmiennej z_j . Ponieważ $r_{jk} = a_{j1} a_{k1}$, $r_{jl} = a_{j1} a_{l1}$, $r_{kl} = a_{k1} a_{l1}$, a więc

$$h_j^2 = \frac{a_{j1} a_{k1} a_{j1} a_{l1}}{a_{k1} a_{l1}} = a_{j1}^2.$$

Współczynniki a_{jk} tworzą macierz czynnikową \mathbf{A} .

$$(18) \quad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}.$$

Równanie

$$(19) \quad r_{jk} = a_{j1} a_{k1} + a_{j2} a_{k2} + \dots + a_{jm} a_{km}$$

można zapisać w postaci macierzowej

$$(19a) \quad \begin{bmatrix} h_1^2 & r_{12} & \dots & r_{1n} \\ r_{21} & h_2^2 & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & h_n^2 & \dots & h_n^2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{bmatrix},$$

a więc

$$(20) \quad \mathbf{R}' = \mathbf{A} \cdot \mathbf{A}^T.$$

Jest to równanie analizy czynnikowej, które przyjmuje się jako podstawę dla określenia ładunków czynnikowych.

Po rozwiązaniu tego równania ze względu na m czynników oblicza się macierz wartości czynnikowych

$$(21) \quad \mathbf{F} = \mathbf{Z} \cdot \mathbf{A},$$

która zawiera wartości czynnikowe dla N jednostek obserwacji. Każdy element f_{ip} tej macierzy jest tzw. wartością czynnikową, czyli wartością daną dla jednostki i ze względu na czynnik p .

Transformacja macierzy Z w macierz F oznacza przeprowadzenie zbioru punktów z przestrzeni n -wymiarowej do równoważnego zbioru punktów w spłaszczonej przestrzeni $m < n$ -wymiarowej, zwanej przestrzenią czynników wspólnych.

Teoria efektywnej liczby czynników oparta jest na podstawowych twierdzeniach algebry macierzy.

Rząd macierzy obserwacji $[x_{ij}]$ określa maksymalną liczbę wektorów liniowo niezależnych tego układu¹³. Jeśli zatem rząd macierzy $[x_{ij}]$ jest równy m , to w układzie wektorów kolumnowych jest m wektorów liniowo niezależnych, a każdy zbiór $m+1$ wektorów ($m+1 \leq n$) tego układu jest liniowo zależny. Oznacza to, że elementy macierzy $[x_{ij}]$ są liniowo zależne od m zmiennych. Stąd n zmiennych można przedstawić w postaci kombinacji liniowej m czynników.

Ponieważ rząd iloczynu macierzy przez jej przestawienie jest równy rzędowi macierzy (w przypadku zmiennych znormalizowanych), rząd macierzy korelacji równa się więc rzędowi macierzy obserwacji¹⁴. Stąd, jeżeli macierz korelacji jest rzędu $m < n$, zmienność układu określa liniowa przestrzeń m -wymiarowa. Można wtedy znaleźć liniową transformację do m nowych zmiennych, które całkowicie wyjaśniają zmienność.

M. S. Bartlett (1950) opracował test istotności dla liczby wspólnych czynników. Jednak na podstawie tego testu otrzymuje się dużo więcej czynników, niż można je zinterpretować w praktyce.

H. H. Harman (1960) twierdzi, że analizę czynnikową należy kontynuować tylko do momentu przeanalizowania 95% zasobu zmienności wspólnej, gdyż pozostałe czynniki mają małe znaczenie.

Z reguły w praktyce badawczej, wykorzystując formułę o aproksymacji rzędu macierzy korelacji, wyodrębnia się tylko czynniki odpowiadające wartościom własnym większym od jedności¹⁵.

3. METODY WYODRĘBNIANIA CZYNNIKÓW

Koncepcja analizy czynnikowej powstała i rozwinęła się w psychologii anglosaskiej. Za twórcę tej metody uważa się Ch. Spearmana, który wprowadził pojęcie pojedynczego czynnika ogólnego dla wyjaśnienia wyników testów inteligencji¹⁶. Ch. Spearman badał układ czterech testów i stwierdził, że dodatnie korelacje między testami uzdolnień wywołuje pewien wspólny czynnik, jeśli współczynniki korelacji każdego dwóch kolumn i wierszy tabeli korelacji są proporcjonalne. Z tabeli korelacji, w oparciu o kryterium proporcjonalności, wyprowadził równania różnicy

¹³ Rzędem m macierzy nazywamy najwyższy stopień wyznacznika różnego od zera dającego się wyjąć z tej macierzy.

¹⁴ Dowód tego twierdzenia podaje K. J. Holzinger, H. H. Harman (1941), s. 58, 317—319.

¹⁵ P. O. Pedersen (1967 b), L. King (1969).

¹⁶ Ch. Spearman (1904), s. 201—229, Patrz: L. L. Thurstone (1935), s. 137—141; B. Fruchter (1954), s. 8, oraz T. Czyż (1967).

czwórkowej, które stosował jako dowód występowania tylko jednego wspólnego czynnika w układzie korelacji.

Dopiero L. L. Thurstone stworzył podstawy teoretyczne analizy wieloczynnikowej, która jest generalizacją teorii Ch. Spearmana. L. L. Thurstone (1931, s. 406—427; 1935, s. 72—77) dążył do wyodrębnienia wszystkich czynników, które mogą rzeczywiście tkwić w korelacjach danego układu zmiennych. Stwierdził, że różnica czwórkowa Ch. Spearmana jest rozwinięciem podwyznacznika drugiego stopnia w ujęciu rachunku macierzowego i wprowadził pojęcie rzędu macierzy korelacji. Według L. L. Thurstone'a rząd macierzy korelacji wyznacza więc liczbę czynników potrzebną do wytłumaczenia korelacji danego układu zmiennych. W ten sposób L. L. Thurstone określa analizę czynnikową jako liniowy model matematyczny i rozszerza możliwości zastosowania metody czynnikowej, gdyż poddaje analizie macierze korelacji wyższego rzędu.

Wyodrębnienia ładunków czynnikowych z macierzy korelacji L. L. Thurstone dokonał metodą centroidalną¹⁷.

Obok metody centroidalnej istnieją metody blisko z nią spokrewnione, oparte na tej samej zasadniczej koncepcji. L.L. Thurstone podaje jeszcze cztery metody: metodę przekątną oraz warianty metody centroidalnej: metodę grup, metodę grupowania, metodę grup wielokrotnych¹⁸.

Sformułowanie modelu wieloczynnikowego spowodowało dalszy silny rozwój badań nad założeniami analizy czynnikowej z udziałem nie tylko psychologów, ale również statystyków.

Do najwybitniejszych należą prace: H. Hotellinga (1933), D. N. Lawley'a (1943), K. J. Holzingera i H. H. Harmana (1941), M. S. Bartletta (1950), R. B. Cattella (1952), B. Fruchtera (1954), T. W. Andersona (1958), M. G. Kendallalla (1957), K. G. Jöreskoga (1963), D. N. Lawley'a i A. E. Maxwella (1963), J. Okonia (1964) oraz praca H. H. Harmana (1960) — kompendium nowoczesnych numerycznych metod czynnikowych.

Spośród metod wyodrębniania ładunków czynnikowych największe uznanie matematyków ze względu na jednoznaczność wyników zdobyła metoda głównych składników, której podstawy sformułował K. Pearson (1901), a alternatywną procedurę rozwinęli H. Hotelling (1933) i T. L. Kelley (1935). Modyfikację tej metody stanowi metoda głównego czynnika H. H. Harmana (1960), w której analizie się tylko zasób zmienności wspólnej każdej zmiennej.

W metodzie głównego czynnika układ odniesienia tworzą wektory reprezentujące zmienne, a więc N punktów reprezentujących obiekty rozpatruje się w przestrzeni

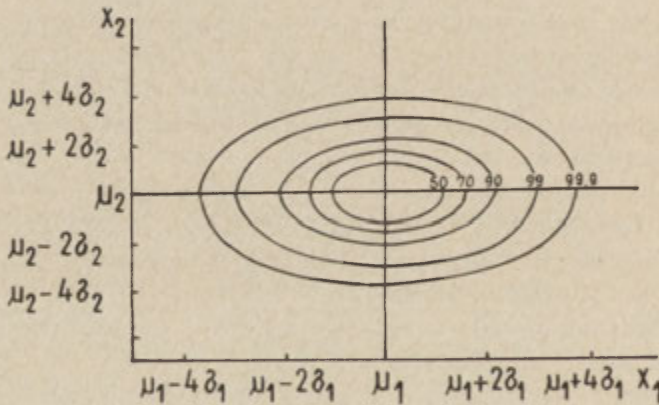
¹⁷ Podstawy metody centroidalnej pierwszy opracował C. Burt (1917). Metodę tę rozwinął L. L. Thurstone (1931), s. 406—427. Za K. G. Joreskog (1963), s. 120. Patrz L. L. Thurstone (1935), s. 92—119, K. J. Holzinger, H. H. Harman (1941), s. 180—198, T. Czyż (1967), s. 139—141.

¹⁸ Patrz R. B. Cattell (1952), s. 167—187, B. Fruchter (1954), s. 52—59, 87—99.

n -wymiarowej. Te punkty przestrzeni n -wymiarowej są zawarte w przestrzeni wspólnego czynnika, tylko m -wymiarowej, gdy m jest rzędem zredukowanej macierzy korelacji. Czynniki wyznaczają główne osie elipsoidy uformowanej przez „chmurę” punktów. Wzdłuż tych osi wariancje zmiennych są ekstremalne.

W interpretacji geometrycznej rozwiązanie głównego czynnika jest równoważne z wyznaczeniem hiperpłaszczyzny m -wymiarowej przechodzącej przez środek ciężkości zbioru, od której średniokwadratowa odległość wszystkich punktów jest ekstremum. Hiperpłaszczyznę wyznaczają osie, wzdłuż których wariancje zmiennych są maksymalne. Transformacja głównego czynnika jest stąd rotacją oryginalnego układu współrzędnych x_1, x_2, \dots, x_n do układu określonego przez główne osie elipsoidy.

Przypuśćmy, że obiekty rozpatruje się w przestrzeni 2-wymiarowej¹⁹. Jeżeli dwie zmienne są nie skorelowane, to linie konturowe normalnej powierzchni korelacji stanowią elipsy, których osie główne są równoległe do odpowiednich osi współrzędnych x_1 i x_2 . Środkiem tej rodziny elips jest środek ciężkości, tzn. punkt o współrzędnych (\bar{x}_1, \bar{x}_2) . Obszar ograniczony elipsą może być uważany za charakterystykę rozproszenia (ryc. 3).



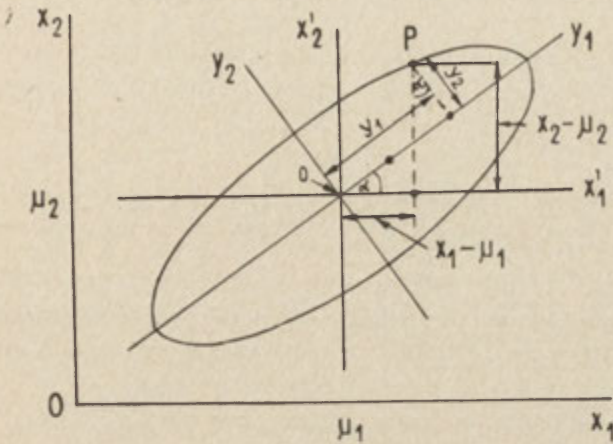
Ryc. 3. Osie główne i linie konturowe powierzchni korelacji w przypadku ortogonalności zmiennych. Źródło: L. J. King (1969)

Fig. 3. Principal axes and contour lines of correlation surface for independence of variables. Source: L. J. King (1969)

Załóżmy teraz, że zmienne x_1 i x_2 są skorelowane i każda z nich ma rozkład normalny. Linie konturowe stanowią, tak samo jak w przypadku niezależności, szereg elips współśrodkowych i podobnych. Osie główne y_1 i y_2 pokrywające się z prostymi regresji ortogonalnej nie są jednak już równoległe do osi x_1 i x_2 . Powierzchnia normalna została obrócona z położenia, dla którego korelacja jest

¹⁹ G. U. Yule, M. G. Kendall (1966), s. 251—263.

równa zero, o kąt α . Punkt P w układzie współrzędnych ortogonalnych y_1 i y_2 ma nowe współrzędne czynnikowe (ryc. 4).



Ryc. 4. Osie główne i linie konturowe powierzchni korelacji w przypadku zależności zmiennych. Źródło: L. J. King (1969)

Fig. 4. Principal axes and contour lines of correlation surface for dependence of variables. Source: L. J. King (1969)

Powyższy przypadek dotyczy sytuacji dwuwymiarowej, ale może być z łatwością uogólniony na przypadek wielowymiarowy.

Z algebraicznego punktu widzenia wybór nowych osi jest równoznaczny z wyodrębnieniem czynników w określonym porządku stosownie do ich udziału w całkowitym zasobie zmienności wspólnej. Analizę rozpoczyna się od czynnika F_1 , którego udział w zmienności wspólnej zmiennych jest największy. Potem oblicza się macierz pierwszych pozostałości korelacji. Następnie wyznacza się drugi czynnik ortogonalny F_2 z maksymalnym udziałem w pozostałej zmienności wspólnej. Ten proces kontynuuje się aż do momentu przeanalizowania całej zmienności wspólnej (zmiennych).

Model głównego czynnika przedstawia wzór

$$22) \quad z_n = a_{n1}F_1 + a_{n2}F_2 + a_{n3}F_3 + \dots + a_{nm}F_m,$$

gdzie a_{jk} jest ładunkiem czynnika F_k ($k = 1, 2, \dots, m$) w zmiennej z_j ($j = 1, 2, \dots, n$)²⁰.

Pierwsze stadium metody głównego czynnika obejmuje wyodrębnienie współczynników a_{j1} ($j = 1, 2, \dots, n$) pierwszego wspólnego czynnika F_1 .

²⁰ Procedurę matematyczną przedstawia się wg H. Harmana (1960), s. 154—185.

Suma udziałów pierwszego czynnika F_1 w ogólnym zasobie zmienności wspólnej wynosi

$$(23) \quad A_1 = a_{11}^2 + a_{21}^2 + \dots + a_{n1}^2.$$

Rozwiązanie problemu sprowadza się do znalezienia takich wartości współczynników a_{j1} , dla których A_1 przyjmuje wartość maximum, przy czym spełniony jest warunek wyrażony wzorem

$$(24) \quad r_{jk} = r'_{jk} = \sum_{p=1}^m a_{jp} a_{kp} \quad (j, k = 1, 2, \dots, n),$$

gdzie r_{jj} jest zasobem zmienności wspólnej h_j^2 zmiennej s_j ($j = 1, 2, \dots, n$).

Po wyznaczeniu maksimum funkcji n zmiennych, w której zmienne są związane $n(n+1)/2$ dodatkowymi równaniami (24), stosuje się metodę współczynników nieoznaczonych (dowolnych) Lagrange'a.

Tworzy się funkcję pomocniczą Lagrange'a w postaci

$$(25) \quad 2T = A_1 - \sum_{jk=1}^n \mu_{jk} r_{jk} = A_1 - \sum_{j,k=1}^n \sum_{p=1}^m \mu_{jk} a_{jp} a_{kp},$$

gdzie $\mu_{jk} = \mu_{kj}$ są nieoznaczonymi współczynnikami Lagrange'a. Warunkiem koniecznym istnienia ekstremum funkcji jest, aby pochodne cząstkowe funkcji T względem n zmiennych były równe zero, czyli

$$(26) \quad \frac{\partial T}{\partial c_{j1}} = a_{j1} - \sum_{k=1}^n \mu_{jk} a_{k1} = 0 \quad (j = 1, 2, \dots, n)$$

$$(27) \quad \frac{\partial T}{\partial c_{jp}} = - \sum_{k=1}^n \mu_{jk} a_{kp} = 0 \quad (p \neq 1).$$

Zestawiając łącznie równania (26) i (27) otrzymuje się

$$(28) \quad \frac{\partial T}{\partial a_{jp}} = \delta_{jp} a_{j1} - \sum_{k=1}^n \mu_{jk} a_{kp} \quad \begin{matrix} (j = 1, 2, \dots, n) \\ (p = 1, 2, \dots, m), \end{matrix}$$

gdzie δ_{1p} jest symbolem Kroneckera; $\delta_{jp} = 1$, jeśli $p = 1$; a $\delta_{jp} = 0$, jeśli $p \neq 1$.

Po pomnożeniu wyrażenia (28) obustronnie przez a_{j1} i określeniu sumy w odniesieniu do wszystkich j otrzymuje się

$$(29) \quad \delta_{jp} \sum_{j=1}^n a_{j1}^2 - \sum_{j=1}^n \sum_{k=1}^n \mu_{jk} a_{j1} a_{kp} = 0.$$

Stwierdzając na podstawie (26), że wyrażenie $\sum_{j=1}^n \mu_{jk} a_{j1}$ jest równe a_{k1} , i przyjmując

$\sum_{j=1}^n a_{j1}^2 = \lambda_1$, równanie (29) można zapisać w postaci

$$(30) \quad \delta_{1p} \lambda_1 - \sum_{k=1}^n a_{k1} a_{kp} = 0.$$

Z kolei mnożąc wyrażenie (30) przez a_{jp} i sumując, dla p otrzymuje się równanie

$$(31) \quad a_{j1} \lambda_1 - \sum_{i=1}^n a_{k1} \left(\sum_{p=1}^m a_{jp} a_{kp} \right) = 0,$$

a po uwzględnieniu warunku (24)

$$(32) \quad \sum_{k=1}^n r_{jk} a_{k1} - \lambda_1 a_{j1} = 0.$$

Powyższe wyrażenie reprezentuje układ n równań, jedno dla każdej wartości j , które można zapisać w postaci

$$(33) \quad \begin{bmatrix} h_1^2 - \lambda & r_{12} & \dots & r_{1n} \\ r_{21} & h_2^2 - \lambda & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & h_n^2 - \lambda \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{n1} \end{bmatrix} = 0,$$

gdzie λ jest wartością λ_1 z wyrażenia (32).

Te równania stanowią podstawę do obliczenia n niewiadomych, tj. współczynników a_{j1} .

Warunkiem koniecznym i dostatecznym na to, aby układ n jednorodnych równań miał rozwiązanie niezerowe, jest, by wyznacznik główny tego układu był równy zeru, czyli jeżeli

$$(34) \quad \begin{vmatrix} (h_1^2 - \lambda) & r_{12} & \dots & r_{1n} \\ r_{21} & (h_2^2 - \lambda) & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & (h_n^2 - \lambda) \end{vmatrix} = 0.$$

Rozwijając ten wyznacznik, otrzymuje się wielomian stopnia n względem zmiennej λ . Wielomian ten przybiera postać wielomianu charakterystycznego, a równanie (34) jest równaniem charakterystycznym.

Wszystkie pierwiastki tego równania są rzeczywiste. Podstawiając zamiast λ w (33) pierwiastek równania charakterystycznego, otrzymuje się układ $(n-1)$ jednorodnych liniowych równań. Ten układ równań ma szereg rozwiązań, z których wszystkie są proporcjonalne do jednego szczególnego rozwiązania. Z analizy

maksymalizacji A_1 wynika, że czynnikiem proporcjonalności jest $\lambda_1 = \sum_{j=1}^n a_{j1}^2 = A_1$.

Stąd A_1 podlegające maksymalizacji równa się jednemu z pierwiastków charakterystycznych równania, mianowicie największemu pierwiastkowi λ_1 .

Największy pierwiastek λ_1 równania (34) podstawia się do układu równań (33) i otrzymuje się pewne rozwiązanie: $a_{11}, a_{21}, \dots, a_{n1}$. Następnie, aby spełnić zależność (23), otrzymany wektor normalizuje się, czyli dzieli się każdą z jego składowych przez pierwiastek kwadratowy z sumy ich kwadratów i mnoży się przez $\sqrt{\lambda_1}$. W rezultacie otrzymuje się

$$(35) \quad a_{j1} = \frac{a_{j1} \sqrt{\lambda_1}}{\sqrt{(a_{11}^2 + a_{21}^2 + \dots + a_{n1}^2)}} \quad (j = 1, 2, \dots, n),$$

a więc szukane współczynniki F_1 w modelu głównego czynnika (22).

Pierwiastki λ równania charakterystycznego (34) są tzw. wartościami własnymi macierzy \mathbf{R}' , natomiast rozwiązaniem układu równań (33), odpowiadającym danej wartości własnej, jest tzw. wektor własny tej macierzy.

Po określeniu współczynników a_{j1} pierwszego czynnika F_1 kolejnym problemem jest wyodrębnienie czynnika, który wyjaśnia maksimum pozostałej zmienności wspólnej. Po wyznaczeniu pierwszego czynnika tzw. pierwsze pozostałości korelacji wynoszą

$$36) \quad {}_1r_{jk} = r_{jk} - a_{j1}a_{k1} = a_{j2}a_{k2} + a_{j3}a_{k3} + \dots + a_{jm}a_{km}.$$

Uogólniając, macierz pierwszych pozostałości korelacji wyraża się wzorem

$$(36a) \quad \mathbf{R}_1 = \mathbf{R}' - \mathbf{Q}_1,$$

gdzie

$$(37) \quad \mathbf{Q}_1 = [a_{j1}][a_{j1}]^T \quad (j = 1, 2, \dots, n).$$

Aby określić współczynniki drugiego czynnika, należy zmaksymalizować wyrażenie

$$(38) \quad A_2 = a_{12}^2 + a_{22}^2 + \dots + a_{n2}^2,$$

a więc sumę udziałów F_2 w pozostałej zmienności wspólnej. Uwzględniając warunek (36), ekstremum funkcji (38) wyznacza się jak w przypadku pierwszego czynnika. Określa się maksymalną wartość własną macierzy \mathbf{R}_1 i odpowiadający jej wektor własny, którego składowe po unormowaniu są szukanymi współczynnikami a_{j2} ($j = 1, 2, \dots, n$) w układzie (22).

W ten sposób w procedurze wyodrębniania m czynników wyznacza się maksymalne wartości własne kolejno tworzonych macierzy pierwszych, drugich, ..., $(m-1)$ -szych pozostałości korelacji. L. R. Tucker (1944, s. 149–153) w rozwią-

zaniu głównego czynnika proponuje operowanie wyłącznie oryginalną macierzą korelacji, bez obliczania macierzy pozostałości, co upraszcza procedurę obliczeniową.

Bezpośrednie rozwiązanie równania charakterystycznego (34) i układów liniowych równań jednorodnych jest bardzo skomplikowane algebraicznie.

H. Hotelling w algorytmie wyznaczania głównego czynnika wprowadził iteracyjną metodę (kolejnych przybliżeń) obliczania maksymalnego pierwiastka równania charakterystycznego (34), czyli maksymalnej wartości własnej kolejno dla zredukowanej macierzy korelacji \mathbf{R}' , macierzy pierwszych pozostałości \mathbf{R}_1 , aż do macierzy $(m-1)$ -szych pozostałości korelacji²¹. Dla obliczonych wartości własnych określa się wektory własne, a więc współczynniki kolejnych czynników wspólnych F_1, F_2, \dots, F_m w układzie (22).

Metoda iteracji polega na tym, że dokonuje się arbitralnego wyboru n liczb, które traktuje się jako zerowe przybliżenie rozwiązania układu i wielokrotnie przekształca się je przy zastosowaniu obserwowanych korelacji w pierwsze, drugie, ..., k -te przybliżenie rozwiązania. Proces iteracyjny jest zbieżny (tzn. liczba przybliżeń potrzebnych do określenia zadaną dokładnością pierwiastków układu jest niewielka), gdy różnica między wartościami pierwiastków otrzymanych z dwóch kolejnych przybliżeń jest dowolnie mała. W tym przypadku obliczanie dalszych przybliżeń można przerwać, a ostatnie przybliżenie przyjąć za rozwiązanie równania²².

W procedurze obliczania wartości własnych i odpowiadających im wektorów własnych oprócz schematu iteracji H. Hotellinga stosuje się w praktyce metodę Jacobi i jej nowoczesne modyfikacje. Metody Jacobi znalazły przede wszystkim zastosowanie w przypadku realizacji analizy czynnikowej na maszynach cyfrowych. Szczegółowy opis matematycznych założeń tych metod zawierają prace H. H. Harmana (1960) oraz K. Styś (1967).

4. ROTACJA CZYNNIKÓW

Rozwiązanie czynnikowe wyznacza m -wymiarową przestrzeń wspólnego czynnika. Kontynuując analogię do geometrii, o zmiennych można mówić jako o punktach przestrzeni, która jest określona przez m osi odniesienia, reprezentujących czynniki.

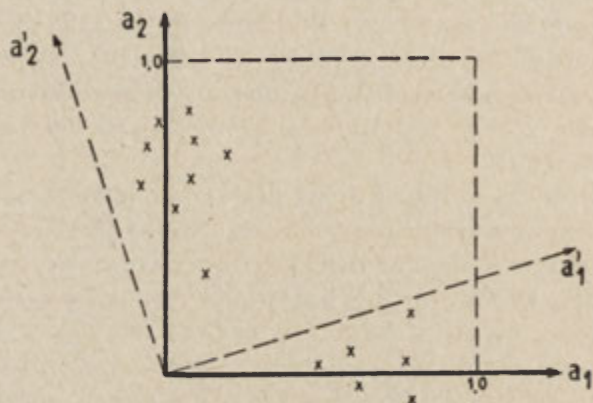
Układ wektorów zmiennych tej przestrzeni jest elementem niezmiennym (kąty zawarte między wektorami są zdeterminowane przez macierz korelacji). Jednak rzuty tej stałej konfiguracji na różnie ułożone układy odniesienia (czyli zbiory wag czynnikowych) mogą być wzajemnie przekształcone i są w tym sensie równo-

²¹ H. Hotelling (1933), s. 27—35, K. J. Holzinger, H. H. Harman (1941), s. 184—186, 363—366.

²² Prosty przykład zastosowania metody iteracji w rozwiązaniu głównego czynnika podaje T. Czyż (1967), s. 154—157.

ważne, jednak pod warunkiem, że nie przesunie się początku układu współrzędnych, lecz tylko dokona się obrotu układu odniesienia wokół tego punktu (ryc. 5).

Ta operacja obrotu osi współrzędnych w procedurze analizy czynnikowej nosi nazwę rotacji. Obracając układ odniesienia można znaleźć nieskończenie wiele rozwiązań. Stwierdzono, że istnieje jakaś jedna określona pozycja układu odniesienia,



Ryc. 5. Rotacja:

a'_1 — a'_2 osie przed rotacją
 a_1 — a_2 osie układu prostej struktury

Fig. 5. Rotation:

a'_1 — a'_2 axes before rotation
 a_1 — a_2 axes in simple structure

która daje zbiór ładunków czynnikowych o szczególnym znaczeniu, a wszelkie inne pozycje są jej matematycznymi przekształceniami. Ustalenie najwłaściwszej pozycji układu odniesienia jest jednym z najtrudniejszych problemów analizy czynnikowej.

Według L. L. Thurstone'a (1935, s. 150—163) w przypadku rotacji zasadą postępowania jest dążenie do tzw. prostej struktury. „Prostota” takiej struktury ładunków czynnikowych polega na tym, że każda zmienna ma stosunkowo najprostszą zawartość czynnikową, tj. dominujący ładunek jakiegoś jednego czynnika, i odwrotnie — miarą danego czynnika są tylko niektóre spośród analizowanych zmiennych. Określenie prostej struktury ma więc duże znaczenie dla interpretacji uzyskanych czynników.

W interpretacji geometrycznej rozwiązanie prostej struktury charakteryzuje się skupianiem się punktów reprezentujących zmienne wokół osi czynnikowych (ryc. 5).

Rozwiązania prostej struktury można dokonać graficznie lub analitycznie za pomocą metody Varimax i Quartimax.

Należy podkreślić, że brak jest dotychczas zgodności poglądów na temat konieczności zastosowania rotacji w procedurze analizy czynnikowej. Większość autorów

stoi na stanowisku, że żadna z metod wyodrębniania czynników nie daje takich wymiarów, które by bez rotacji mogły być uznane za ostateczne²³. Jednak R. J. Wherry (1959), B. J. L. Berry (1967a) pomijają to przekształcenie i w ten sposób upraszczają procedurę matematyczną.

5. INTERPRETACJA CZYNNIKÓW

Struktura czynnikowa zawiera układ skonkretyzowanych zależności liniowych między zmiennymi i czynnikami wspólnymi. Czynniki są zmiennymi ukrytymi, bezpośrednio nieobserwowalnymi²⁴.

Model czynnikowy należy więc do klasy struktur ukrytych. W tej sytuacji trzeba podkreślić, że wartość poznawcza modelu zależy będzie od właściwego rozpoznania zmiennych modelu oraz ich pojęciowej identyfikacji.

Interpretacja czynników jest najważniejszą częścią procedury analizy czynnikowej i posiada decydujące znaczenie dla adekwatnego poznania rzeczywistości. Operacja ta jednak jest bliżej nie zidentyfikowana metodologicznie. Próby sformułowania bardziej szczegółowych dyrektyw postępowania badawczego w tym zakresie nie dały dotąd rezultatów.

Interpretacja czynników jest w dużym stopniu wyznaczona przez konstrukcję teoretyczną i pojęciową, którą badacz uważa za najwłaściwszą. Wprowadza ona więc element subiektywizmu do analizy czynnikowej. Zasady interpretacji nasuwają wiele wątpliwości, oparte są na założeniach arbitralnych i nie prowadzą do konkretyzacji czynników w sposób dostatecznie ścisły. Podstawą interpretacji jest macierz korelacji zmiennych pierwotnych oraz wagi czynników, które mają postać współczynnika korelacji między zmienną a czynnikiem.

Najczęściej dokonuje się analizy względnego udziału cech wyjściowych wchodzących w skład czynnika. Najistotniejsza dla opisu czynnika będzie cecha, w której udział tego czynnika jest największy. Taka sytuacja zaistnieje jedynie w szczególnym przypadku prostej struktury czynnikowej.

Jeżeli czynnik ma charakter czynnika ogólnego z dodatnimi ładunkami we wszystkich cechach, można go interpretować jako „średnią” zmiennych wyjściowych.

Natomiast czynnik bipolarny interpretowany w kategoriach cech wyjściowych wyraża równocześnie właściwość, która może mieć zarówno ujemne, jak i dodatnie natężenie²⁵.

²³ R. B. Cattell (1952), s. 66.

²⁴ Często określa się je również mianem zmiennych fikcyjnych lub teoretycznych w odróżnieniu od zmiennych „prawdziwych” lub empirycznych.

²⁵ Czynnikiem bipolarnym nazywa się czynnik dodatnio skorelowany z pewnymi cechami oraz ujemnie skorelowany z pozostałymi zmiennymi zbioru, np. stabilność i mobilność ludności.

6. ALTERNATYWNE TECHNIKI ANALIZY CZYNNIKOWEJ

Podstawą metody czynnikowej są dwie techniki *R* i *Q*. Jeżeli operuje się macierzą korelacji między zmiennymi, to uzyskuje się ładunki czynnikowe dla zmiennych. Jest to metoda standardowa, stosowana od kilkudziesięciu lat i znana jako technika *R*.

Technika *Q* odnosi się do korelacji między jednostkami rozpatrywanymi w kategoriach wielu cech. Te korelacje określają stopień podobieństwa jednostek i wykazują grupy lub typy jednostek. W tym układzie każdy czynnik reprezentuje „typ” taksonu. Ładunek czynnikowy określa więc stopień podobieństwa między jednostką a typem. Im wyższy ładunek czynnikowy, tym bardziej typowa jest jednostka. Najbardziej typowym członem taksonu jest ta jednostka, która ma najwyższą wagę czynnika. Natomiast wartości czynnikowe identyfikują cechy, które różnicują jednostki w klasy.

Należy zwrócić uwagę, że podczas gdy technika *R* prowadzi jedynie do redukcji przestrzeni wielocechowej i określenia zbioru diagnostycznych, to zastosowanie techniki *Q* daje bezpośrednio klasyfikację zbioru.

Dotychczas większość prac z zakresu analizy czynnikowej jest oparta na technice *R*. Jednak wprowadzona przez W. Stephensona (1936) technika *Q* znajduje coraz szersze, efektywne zastosowanie.

Koncepcję operowania równocześnie techniką *R* i *Q* zawiera procedura w ujęciu D. R. Saundersa o postaci tzw. bezpośredniej analizy czynnikowej²⁶.

²⁶ D. R. Saunders (1950) za B. J. L. Berry, H. G. Barnum (1962), s. 45.

II. ZASTOSOWANIA METODY CZYNNIKOWEJ W BADANIACH PRZESTRZENNO-EKONOMICZNYCH

1. BADANIA EMPIRYCZNE

Rozwój analizy czynnikowej był początkowo związany z jej zastosowaniem w naukach behavioralnych. Jednak zakres zastosowań tej metody szybko się rozszerzył na nowe dziedziny wiedzy, operujące analizą wielocechową.

Sformułowania koncepcji analizy czynnikowej w odniesieniu do przestrzennego kształtowania się zjawisk społeczno-ekonomicznych dokonał pierwszy M. G. Kendall (1939). Należy zaznaczyć, że praca M. G. Kendalla nie wywarła wówczas istotnego wpływu na rozwój analizy czynnikowej w badaniach geograficznych. W naukach społecznych pod względem liczby praktycznych zastosowań analizy czynnikowej geografii ekonomiczną początkowo znacznie wyprzedzała socjologia. Jednak wraz z rozwojem metod matematyczno-statystycznych analiza czynnikowa stopniowo przeniknęła do wszystkich gałęzi geografii. Szczególnie bogaty dorobek w zakresie zastosowań metod czynnikowych ma geografia anglosaska. W polskiej literaturze geograficznej pierwsi zwrócili uwagę na szczególne znaczenie metody czynnikowej dla dokonywania klasyfikacji przestrzennej zjawisk Z. Chojnicki i A. Wróbel (1961, s. 621).

Na tym miejscu pragnę przedstawić choćby w zarysie szereg najbardziej reprezentatywnych zastosowań analizy czynnikowej w badaniach przestrzenno-ekonomicznych. Wśród nich znalazły się ze względu na pokrewną problematykę również niektóre prace socjologiczne interesujące z geograficznego punktu widzenia.

Prezentowane zastosowania dotyczą przede wszystkim trzech grup zagadnień:

- 1) badań wybranych elementów struktury przestrzenno-ekonomicznej ze szczególnym uwzględnieniem rozmieszczenia produkcji rolnej, układów produkcji przemysłowej, układów działalności usługowej, sieci transportowej i sieci osadniczej;
- 2) analizy rozwoju i wzrostu ekonomicznego w makro- i mikroskali;
- 3) badania ekonomicznej struktury regionalnej.

Użyteczność analizy czynnikowej w badaniach z zakresu geografii rolnictwa zasugerował M. G. Kendall (1939). W formie eksperymentu zastosował on analizę czynnikową w procedurze regionalizacji upraw w Wielkiej Brytanii. M. G. Kendall badał 48 „counties” w zakresie 10 cech dotyczących plonów głównych ziemiopło-

dów, uwzględniając kilka przekrojów czasowych. Metodą analizy czynnikowej wyodrębnił czynnik zinterpretowany jako produktywność, który przykładowo w 1925 r. wyjaśniał 47,6% ogólnej zmienności. Na podstawie wartości tej metacechy dokonał uporządkowania „counties”.

Za klasyczne w tej dziedzinie uważa się również dwie prace M. J. Hagood (1943) oraz M. J. Hagood i D. O. Price (1957).

Pierwsza z nich zawiera obok właściwej analizy czynnikowej dalsze etapy proceduralne związane z regionalizacją. Wielostopniowa analiza zbioru 104 cech dla 48 stanów stanowi punkt wyjścia w delimitacji rolniczo-ludnościowych regionów na obszarze USA. W oparciu o wynikowy czynnik rolniczo-demograficzny i dodatkowo wyprowadzony współczynnik podobieństwa stanów o postaci współczynnika korelacji dokonuje się regionalizacji. Jednak ta procedura na etapie łączenia stanów zawiera elementy arbitralne.

W drugiej pracy analizę czynnikową zastosowano do konstrukcji wskaźnika poziomu życia ludności rolniczej na podstawie 4 mierników wartości konsumpcji rodziny farmerskiej w 196 „counties” USA.

Trzeba zaznaczyć, że wyodrębnienie w powyższych dwóch opracowaniach tylko jednego czynnika, wyjaśniającego 10% zmienności, poważnie obniża wartość wyników analizy.

Z dalszych zastosowań zbliżoną koncepcję wykorzystania modelu czynnikowego w tej dziedzinie reprezentują prace J. D. Henshall (1966), W. M. Żukowskiej (1964), W. M. Żukowskiej i L. Karpowa (1967), B. J. L. Berry'ego (1967a).

J. D. Henshall przy zastosowaniu tej techniki statystyczno-matematycznej dokonała testowania hipotezy opartej na obserwacji stwierdzającej, że istotną rolę w wyjaśnianiu zmienności wewnętrznej struktury chłopskiej gospodarki rolnej na Barbados odgrywa liczba i struktura ludności farm. Analizie poddano 32 cechy społeczno-ludnościowe, organizacyjno-techniczne i produkcyjne rolnictwa, charakteryzujące próbę złożoną ze 116 farm, wybraną według schematu losowania warstwowego w ujęciu przestrzennym. 12 czynników, wyjaśniających 99% ogólnej zmienności, traktuje się jako podstawowe wymiary struktury rolnej. Wśród nich czynnik demograficzny zajmuje trzecie miejsce za czynnikiem wpływu miasta i czynnikiem rozproszenia ziemi, wyjaśniając 12% ogólnej zmienności. Na podstawie tego stwierdzenia przyjmuje się hipotezę wstępną.

W studium W. M. Żukowskiej analiza czynnikowa leży u podstawy określenia *ex post* istotnych właściwości rolnictwa ekstensywnego Kanady. Dane wyjściowe tworzą wartość 11 wskaźników ludnościowych i rolniczych dla 33 okręgów prowincji Saskatchewan i Alberta. Trzy czynniki wspólne tego zbioru formułuje się następująco: 1) specjalizacja towarowa oparta na produkcji pszenicy, 2) poziom rozwoju stosunków kapitalistycznych w rolnictwie, 3) wzrost liczby ludności. Ponadto W. M. Żukowska przeprowadza badania rozkładu przestrzennego wartości czynnikowych i na tej podstawie dokonuje charakterystyki stref glebowo-klimatycznych Kanady.

Wersję powyższego studium stanowi praca W. M. Żukowskiej i L. Karpowa.

Schemat badawczy B. J. L. Berry'ego przedstawia się następująco: 47 cech rolniczych i społecznych dla 555 „municipalities” i „census townships” prowincji Ontario w Kanadzie redukuje się do 3 podstawowych czynników, wyjaśniających 60% ogólnej zmienności, które interpretuje się jako: „farm poverty”, „rural nonfarm poverty”, „disadvantage social”. Jednak analiza czynnikowa nie jest celem samym w sobie, studium „rural poverty” zamyka próba delimitacji jednolitych regionów rolniczych. Regionalizację w 4 regiony i 17 subregionów, nie pozbawioną elementów subiektywnych, B. J. L. Berry ujmuje jako pewną modyfikację typologii przestrzennej w 3 wymiarach.

Rozszerzoną koncepcję pojmowania modelu czynnikowego zawiera praca J. C. Henshall i L. J. King (1966). W badaniu strukturalnym cech gospodarki rolnej Barbados, w oparciu o reprezentatywną próbę złożoną ze 150 farm, początkowo zastosowano technikę *R*, która dała 4 czynniki wspólne dla zbioru 48 cech dotyczących rodzajów upraw i hodowli zwierzęcej. Z kolei zastosowanie techniki *Q* umożliwiło wyodrębnienie 30 czynników (przy czym pierwszy wyjaśniał aż 40% ogólnej zmienności), które interpretuje się jako podstawowe typy farm w zakresie produkcji roślinnej i zwierzęcej.

W literaturze polskiej na możliwość zastosowania modelu czynnikowego do wyodrębnienia cech diagnostycznych dla analizy rozmieszczenia produkcji rolnej zwrócili uwagę Z. Hładyniuk (1966) i A. Zeliaś (1968).

Z. Hładyniuk podjął próbę podziału wielostopniowego województwa lubelskiego na rejony do ekonomicznej optymalizacji kierunków produkcji pasz dla bydła mlecznego. W oparciu o metodę analizy czynnikowej zredukował 12 pierwotnych zmiennych, opisujących 418 podstawowych obszarów rolniczych (gromad, osad, miast) do 5 czynników. Grupowania dokonał przy pomocy metody skupisk, której jednak bliżej nie wyjaśnia.

W ujęciu A. Zeliaśa sformułowane zależności między 12 cechami rolniczymi dla 17 powiatów województwa krakowskiego prowadzą do identyfikacji 3 czynników, wyrażonych w kategoriach cech wyjściowych. Czynniki pierwsze utożsamia się z kierunkiem zbożowym, czynnik drugi — z kierunkiem ekstensywnym hodowli, a czynnik trzeci określa się jako ekstensywny kierunek gospodarki. Po uwzględnieniu dodatkowego kryterium, dotyczącego najwyższych ładunków czynnikowych, za cechy diagnostyczne w rejonizacji produkcji rolnej traktuje się obszary: 1) uprawy koniczyny, 2) uprawy ziemniaków, 3) pogłowia bydła, 4) pogłowia owiec.

W zakresie geografii przemysłu warto wymienić dwa ciekawe zastosowania modelu czynnikowego — do określenia podstawowych mierników rozmieszczenia przemysłu oraz badania zróżnicowania przestrzennego procesu uprzemysłowienia.

Pierwsze podejście prezentuje praca J. L. Morrisona, M. W. Scriptera i R. H. T. Smitha (1968). Autorzy ci opierają się na obszernym materiale statystycznym, macierz obserwacji obejmuje 11 zmiennych dotyczących zatrudnienia

i produkcji przemysłowej dla 2474 „counties” w 1958 r. Analiza czynnikowa wykazuje, że istnieją 2 czynniki, czyli 2 podstawowe źródła zmienności tego układu. Zmienne w układzie dwuczynnikowym tworzą 3 „wiązki”. Średnie wartości tak pogrupowanych 11 zmiennych określa się jako mierniki kolejno: środków produkcji, działalności przemysłowej, nowych nakładów kapitału. Na podstawie powyższych sumarycznych charakterystyk dokonuje się analizy struktury przestrzennej przemysłu USA.

W tym ujęciu autorzy przedstawiają odmienny sposób wykorzystania modelu czynnikowego. Analiza czynnikowa nie jest w tym przypadku narzędziem redukcji czy selekcji zmiennych, lecz dostarcza jedynie kryteriów klasyfikacji zmiennych w podzbiory.

B. Kostrubiec (1969) zgodnie z nowoczesną tendencją do ujęć dynamicznych bada za pomocą analizy czynnikowej relatywny rozwój uprzemysłowienia powiatów Polski w latach 1946–1956. Operuje zbiorem 22 cech dotyczących zatrudnienia w poszczególnych gałęziach przemysłu. Odrębnie dla 1946 i 1956 r. dokonuje klasyfikacji powiatów na bazie 3 czynników. Czynniki te, zasadnicze składowe wzrostu, zmieniają się w czasie. W rozpatrywanych przekrojach czasowych odpowiednie udziały cech wyjściowych, wchodzących w ich skład, znacznie odbiegają od siebie. Dlatego problem klasyfikacji dynamicznej powiatów w zakresie uprzemysłowienia rozwiązuje się obierając ruchomy układ współrzędnych.

Wzorcowe zastosowanie analizy czynnikowej w studiach empirycznych nad transportem stanowi praca W. L. Garrisona i D. F. Marble (1964). Analizie czynnikowej poddano macierz 104 połączeń lotniczych między 59 miastami Wenezueli w 1957 r. w zapisie zero-jedynkowym. Techniką Q liniowej transformacji czynnikowej zredukowano macierz stopnia 59×59 do macierzy $59 \times r$ ($r < 59$). Poprzez interpretację 4 wspólnych czynników wyjaśniających 38% ogólnej obserwowanej zmienności zidentyfikowano pewne strukturalne zależności występujące wewnątrz systemu transportowego. Pierwszy czynnik nazwano czynnikiem ogólnej liczby połączeń, drugi — czynnikiem bezpośrednich połączeń z Caracas, trzeci — czynnikiem regionalizacji systemu na obszar zachodni i wschodni, czwarty — czynnikiem subregionalizacji obszaru zachodniego. Fakt możliwości opisanie tylko 38% wariacji przez czynniki wspólne W. L. Garrison i D. F. Marble tłumaczą „efektem sąsiada”. Zjawisko to polega na istnieniu połączeń przede wszystkim z najbliższymi sąsiadami i może być wyjaśnione przez czynniki specyficzne.

Próbę przydatności zastosowania modelu czynnikowego w odniesieniu do badania rozmieszczenia działalności usługowej przeprowadził B. J. L. Berry (1961b). Dane statystyczne dotyczyły liczby zakładów 6 kategorii usług dla 9 „census divisions” 1954 r. Analiza głównych składników wykazała istnienie 3 istotnych zespołów cech zmienności przestrzennej, na które przypada 93,7% całkowitej zmienności. Dalsze etapy proceduralne obejmowały konstrukcję „drzewa połączeń” opartą o pojęcie odległości taksonomicznej w przestrzeni dwuczynnikowej i delimitację 4 typów regionalnych połączonej z pomiarem straty informacji. Natomiast w studium

struktury handlowej Chicago B. J. L. Berry (1963) transformował macierz zawierającą wartości 10 zmiennych dla 63 centrów handlowych metodą osi głównych do rozmiarów 63×2 . Pierwszy czynnik obejmował wielkość, funkcje i transakcje centrów, a drugi — zmienność opartą na obszarze, ludności i dochodach obszarów centrów handlowych.

W przypadku wykorzystania modelu czynnikowego do analizy sieci osadniczej można wydzielić 3 sfery zastosowań: 1) badanie wewnętrznego zróżnicowania przestrzennego wielkich miast; 2) studia porównawcze z zakresu struktury społeczno-ekonomicznej miast; 3) badanie hierarchii ośrodków centralnych.

Z prac nad zróżnicowaniem przestrzennym miasta z metodologicznego punktu widzenia interesujące są studia empiryczne miast kapitalistycznych wprowadzające problem segregacji społecznej dzielnic. Należy zaznaczyć, że badanie powiązania struktury społecznej ludności miasta ze sposobem jej rozmieszczenia i mieszkania zbliża geografie ekonomiczną do socjologii miasta, a zwłaszcza ekologicznego kierunku w socjologii. Zresztą pierwszą analizę czynnikową w ramach badań obszarów społecznych miast przeprowadzili właśnie socjologowie²⁷.

C. F. Schmid i K. Tagashira (1964) zidentyfikowali drogą wielostopniowej analizy czynnikowej podstawowe wymiary demograficzne miasta Seattle. F. L. Sweetser (1965b) dokonał analizy rozkładu przestrzennego ludności miasta Bostonu w ujęciu 7-czynnikowym w układzie pierścieniowych stref miasta.

Kolejną próbą zastosowania analizy czynnikowej do wydzielenia homogenicznych, pod względem społecznym i etnicznym, dzielnic miasta były badania C. W. Carey'a (1966) na obszarze Manhattanu. 33 cechy dotyczące stanu i struktury ludności oraz zasobów mieszkaniowych dla 269 „tracts of census” sprowadzono do 6 czynników, wyjaśniających 80% ogólnej zmienności. Jako czynniki wiodące można traktować ogólny czynnik ludnościowy oraz czynnik subpopulacji Puer Rican, na które przypada 60% zmienności.

Prace z zakresu tej problematyki były kontynuowane przez P. O. Pedersena (1967a, b). Jako novum w badaniach obszaru metropolitalnego Kopenhagi wprowadził on ujęcie dynamiczne. Zbiór danych tworzyły wartości 14 cech demograficznych dla 67 stref obszaru metropolitalnego w 1950 r. i 76 stref w 1960 r. Wyodrębniono 3 czynniki w postaci: 1) urbanizacji, 2) stanu społeczno-ekonomicznego, 3) wzrostu, zmiany, i dokonano weryfikacji ich interpretacji za pomocą analizy regresji. Dalsze etapy postępowania badawczego obejmowały analizę zmian w strukturze czynnikowej i geograficznym rozkładzie czynników w okresie 1950—1960.

Studia porównawcze w zakresie struktury społeczno-ekonomicznej miast z uwzględnieniem metod czynnikowych prowadzili D. O. Price, F. L. Sweetser, J. K. Had-den i E. F. Borgatta, C. A. Moser i W. Scott. D. O. Price (1942) w oparciu o analizę czynnikową wyznaczył podstawowe wymiary ośrodków wielkomiejskich. Za pomocą 4 czynników wyjaśnił korelacje przy zmienności 15 cech dla 95 miast.

²⁷ M. D. Ardsol i inni (1958), W. Bell (1955).

Czynniki te zidentyfikowano, ale prowizorycznie jako: 1) stopień dojrzałości miasta, 2) rząd ośrodka w zakresie usług, 3) poziom życia w mieście, 4) obroty handlowe na głowę mieszkańca.

F. L. Sweetser (1965c) w studium porównawczym struktury ludności Helsinek i Bostonu w oparciu o 20 zmiennych wyodrębnił 3 podstawowe wymiary. Czynniki zinterpretowano jako: 1) status społeczno-ekonomiczny, 2) przyrost naturalny, 3) urbanizację.

W tym samym kierunku rozwinęli badania J. K. Hadden i E. F. Borgatta (1965). Przeprowadzili oni 8 analiz czynnikowych 65 zmiennych dotyczących ludności i bazy mieszkaniowej dla zbioru 644 miast USA i podzbiorów odpowiadających albo klasom wielkości miast, albo typom miast (centralne, satelityczne, niezależne). Zmienne zredukowano do 16 czynników. Jednak czynniki różnych macierzy nie mają jednakowej wagi i interpretacji. Do najważniejszych należą: status społeczno-ekonomiczny, ludność kolorowa, rozkład wieku, centra oświaty, mobilność mieszkańców oraz gęstość zaludnienia.

W ujęciu C. A. Mosera i W. Scotta (1961) analiza głównych komponentów stanowi podstawę do ustalenia typologii miast. Celem pracy było według autorów zbadanie możliwości określenia stosunkowo jednolitych miast wyznaczonych na podstawie zespołu 57 zmiennych społeczno-ekonomicznych w ich wzajemnych powiązaniach. Dla 157 miast (powyżej 50 tys.) Anglii i Walii cechy zgrupowano w 8 podstawowych klas. Różnice między miastami wyrażone 57 zmiennymi zostały uznane za pochodne 4 głównych czynników, wyrażających 60% wszystkich różnic. Pierwszy czynnik nazwano czynnikiem zróżnicowania społecznego, drugi i trzeci czynnikami etapu rozwoju (drugi — w latach 1931—1950, a trzeci po 1950 r.), czwarty zaś — przeludnienia mieszkaniowego. W końcowej klasyfikacji uzyskano podział na 3 główne typy miast: 1) miasta wypoczynkowe, administracyjne i handlowe, 2) przemysłowe, 3) przedmieścia i miasta typu podmiejskiego oraz 14 grup miast.

Zdaniem K. Dziewońskiego to analityczne ujęcie ilościowe, potwierdzające zresztą intuicyjne rozpoznanie, prowadziłyby do najbardziej precyzyjnej typologii, gdyby dotyczyło jedynie całkowitych jednostek osadniczych²⁸. Prawdopodobnie znalazłoby to odzwierciedlenie w znakomitym poprawieniu procentu różnic wyrażonych przez 4 główne czynniki.

Podobne studia prowadzili Q. Ahmad (1965) dla miast Indii i L. J. King (1966), R. K. Semple (1969) dla miast Kanady.

A. L. Mabogunje (1965) zastosował analizę czynnikową w badaniach nad oddziaływaniem kolonialnego rozwoju ekonomicznego na system miejski Nigerii. Analizy dokonano na podstawie 11 zmiennych demograficznych danych dla 329 miast. Trzy czynniki wyjaśniające około 80% zmienności zinterpretowano następująco: 1) udział młodzieży albo dzieci, 2) ludność starsza, 3) udział kobiet. Następnie

²⁸ Rec. K. Dziewoński (1964).

układy miast badano w przestrzeni trzyczynnikowej w świetle pewnych hipotez dotyczących regionalnego rozwoju ekonomicznego.

W literaturze polskiej pierwszymi zastosowaniami analizy czynnikowej do klasyfikacji miast są prace S. Lewińskiego (1968) oraz J. Bobińskiego i K. Zagórskiego (1969).

W ujęciu S. Lewińskiego analiza czynnikowa jako metoda redukcji zbioru 30 cech dla 250 miast polskich do 4 czynników stanowi punkt wyjścia w grupowaniu dendrytowym miast.

Studium J. Bobińskiego i K. Zagórskiego jest próbą określenia poziomu rozwoju 735 miast Polski przy zastosowaniu transformacji czynnikowej 34 wskaźników do 4 podstawowych wymiarów: skali względnego uprzemysłowienia, budownictwa mieszkaniowego, infrastruktury miejskiej, zaludnienia.

Odrębną grupę tworzą prace poświęcone analizie ośrodków centralnych. Na szczególną uwagę zasługuje koncepcja zastosowania bezpośredniej analizy czynnikowej do określenia hierarchii ośrodków centralnych, sformułowana przez B. J. L. Berry, H. G. Barnum, R. J. Tennanta²⁹. Dla zbioru danych dotyczących występowania lub niewystępowania 104 funkcji usługowych w 47 ośrodkach centralnych 9 „counties” południowo-zachodniej części stanu Iowa (USA) wyodrębniono techniką *R* i *Q* 8 par komponentów. Pierwsza para komponentów, wyjaśniająca 60% ogólnej zmienności, obejmuje czynnik wielkości, który porządkuje ośrodki według liczby wykonywanych funkcji, oraz czynnik powszechności, określający porządek występowania funkcji w ośrodku. Drugą parę komponentów identyfikuje się jako czynnik bipolarny „a city — village factor”, który wskazuje na istnienie dwóch klas ośrodków: osiedli i dużych miast oraz dwóch klas odpowiadających im funkcji: funkcji poziomu osiedla i funkcji poziomu dużego miasta. Trzecią klasę ośrodków i funkcji: „towns” i „town-level functions”, wyznacza trzecia para komponentów. W ten sposób wydzielając na podstawie głównych komponentów trzy klasy ośrodków: „villages”, „towns”, „cities”, różniących się stopniem rozwoju funkcji, ustalono hierarchię sieci ośrodków centralnych.

S. Illeris i P. O. Pedersen (1968) przyjęli jako kryterium wyznaczenia systemu ośrodków centralnych cechę o charakterze wektora w postaci liczby rozmów telefonicznych. W tym ujęciu badanie stopnia centralności ośrodków wiąże się z delimitacją ich stref wpływu i prowadzi do wydzielenia regionów węzłowych. Analiza czynnikowa materiału statystycznego, pochodzącego z badania reprezentacyjnego, dla 62 okręgów telefonicznych Danii (okręg składa się z miasta i jego zaplecza) wykazała istnienie 10 czynników, wyjaśniających 79% całkowitej zmienności. Wartości czynnikowe identyfikują 7 silnych ośrodków centralnych, a wagi czynnikowe — ich strefy wpływu.

J. O. A bi odun (1967) zastosowała analizę komponentów jako metodę weryfikacji

²⁹ B. J. L. Berry, H. G. Barnum (1962), B. J. L. Berry, H. G. Barnum, R. J. Tennant (1962).

modelu rozmieszczenia ośrodków centralnych Christallera. W studium hierarchii miast południowo-zachodniej Nigerii dokonuje się podziału 28 funkcji na 3 klasy i przypisuje się im odpowiednie wagi 10,5,1. Analiza komponentów macierzy korelacji funkcji ośrodków określiła 10 głównych czynników. Dwa pierwsze, wyjaśniające 68,01% ogólnej zmienności, traktuje się jako istotne. W układzie dwuczynnikowym bada się funkcjonalne odległości ośrodków i dokonuje się hierarchicznego grupowania 74 ośrodków w 5 kategorii. Rozkład ośrodków poszczególnych rzędów w systemie regionalnym Ijebu Province jest w zasadzie zgodny z modelem Christallera dla $k = 3$.

W literaturze ekonomiczno-geograficznej z punktu widzenia zastosowań analiza czynnikowa ściśle związana jest z pomiarem stopnia rozwoju społeczno-ekonomicznego. W studiach kwantytatywnych makro- i mikrorozwoju, opartych na wielkich zbiorach danych statystycznych, stanowi ona najbardziej efektywne narzędzie postępowania badawczego.

Do badań nad rozwojem gospodarczym świata metodę tę wprowadził B. J. L. Berry (1960, 1961a). Wykazał on, że za pomocą analizy czynnikowej 43 zmienne stosowane do charakterystyki ekonomicznego rozwoju 95 krajów można zredukować do 4 głównych komponentów, wyjaśniających 92% zmienności. Czynniki światowego rozwoju gospodarczego są: 1) skala technologiczna, 2) skala demograficzna, 3) różnice dochodów i stosunków zagranicznych, 4) wielkie i małe kraje. W końcowej regionalizacji B. J. L. Berry uwzględnił tylko skalę technologiczno-demograficzną.

M. Megee (1965a, 1966) przeprowadziła podobne badania dla 153 państw świata w 1963 r. Operowanie wartościami absolutnymi 56 zmiennych zwiększyło precyzję analizy. Wprowadzono 4 główne wymiary rozwoju ekonomicznego, na które przypada 76% całkowitej zmienności: 1) rozwój przemysłu, 2) infrastruktura, 3) handel wewnętrzny i zagraniczny, 4) wymiar społeczny. W oparciu o dwuwymiarową koncepcję rozwoju ekonomicznego określono światowy system regionalny, złożony z 5 regionów i grupy państw izolowanych.

Badania stopnia i charakteru rozwoju społeczno-ekonomicznego w skali międzynarodowej dotyczy również opracowanie L. F. Schnore (1961).

Większość zastosowań metody odnosi się jednak do badania poziomu ekonomicznego w mikroskali. Tego rodzaju badania M. Megee (1963, 1965 b, c) prowadziła dla stanów USA i Meksyku. Szczególnie interesujące pod względem metodologicznym jest opracowanie dotyczące 48 stanów USA, w którym wprowadzono analizę zmian natury i względnego znaczenia czynników wzrostu w czasie. Aby można było porównać bezpośrednio rezultaty dwóch analiz czynnikowych dla dwóch różnych okresów (1940 i 1960 r.), uwzględniono takie same zmienne w postaci 75 cech ekonomicznych. Wyodrębniono 4 czynniki, różne dla dwóch badanych okresów. Stąd konkluduje się, że strukturę i stopień rozwoju ekonomicznego w 1960 r. określają inne metacechy niż w 1940 r. Podczas gdy w 1940 r. najistotniejszym czynnikiem stopnia rozwoju była produkcja energii, to w 1960 r. stały się nim produkcja

przemysłowa i budownictwo. Natomiast dochód z działalności usługowej jest drugim pod względem wagi czynnikiem w obydwu latach.

Badania rozwoju regionów średniej skali kontynuowali również B. R. Olsen i G. Garb (1965) w USA oraz R. E. Roberts i G. W. Mc Bee (1968) w Meksyku.

Wielu autorów w studiach z tego zakresu schodzi poniżej szczebla stanowego, co nadaje badaniom wyższy stopień szczegółowości przestrzennej. Najczęściej podstawową jednostką przestrzenną jest „county”.

J. H. Thompson, S. C. Sufrin, P. R. Gould i M. A. Buck (1962) dokonali próby sformułowania wymiarów „ekonomicznego zdrowia” dla 50 „counties” stanu Nowy Jork. Centralnym zagadnieniem pracy jest wyróżnienie i przestudiowanie wspólnych składników w 9 cechach, dotyczących warunków ekonomicznych i tendencji rozwojowych. Wyprowadzone komponenty otrzymały subiektywną interpretację ekonomiczną: pierwszy to miernik ogólnego rozwoju, drugi reprezentuje różnice między zurbanizowanymi i wiejskimi obszarami, a trzeci jest miarą wzrostu ekonomiczno-demograficznego. Na powyższym studium Nowego Jorku wzorowali się W. H. Bell i D. W. Stevenson (1964), prowadząc badania „ekonomicznego zdrowia” Ontario.

W ujęciu D. Thompsona (1970) badania poziomu ekonomicznego sprawdzają się przede wszystkim do identyfikacji głównych wymiarów tzw. „środowiska społeczno-ekonomicznego” poszczególnych „counties”. Analiza czynnikowa dużej masy statystycznej (72 zmienne dla 300 „counties” wybranych losowo) pozwoliła wyróżnić 11 czynników. Na pierwszy plan wysunęły się czynniki: poziom życia, zmiana, i urbanizacja, które wykazują podobieństwo do czynników „ekonomicznego zdrowia”.

Analiza głównych wymiarów poziomu ekonomicznego przy zastosowaniu metod czynnikowych stanowi często początek postępowania badawczego w zakresie typologii przestrzennej i regionalizacji ogólnoekonomicznej. Czynniki stanowią kryteria delimitacji ekonomicznych układów przestrzennych.

To podejście realizuje szereg prac dotyczących USA, m. in. J. P. Cole’a i C. A. M. Kinga (1968), B. G. Jonesa i W. W. Goldsmitha (1968), B. H. Stevensa i C. A. Bracketta (1968), F. E. Hortona i H. Mc Conella (1969).

Najbardziej systematyczne i wyczerpujące studium stanowi praca B. G. Jonesa i W. W. Goldsmitha. Badanie różnic i podobieństw między „minor civil divisions” w kategoriach tylko 5 czynników, zamiast 44 zmiennych, prowadzi do uogólnień typologicznych i subregionalizacji 3 „counties” stanu New York. Czynniki sformułowano ogólnie jako różne stadia rozwoju urbanizacji i rolnictwa.

Nieco odmienny charakter posiada praca F. E. Hortona i H. Mc Conella. Realizuje ona rozszerzoną koncepcję analizy struktury ekonomicznej w aspekcie zmian czasoprzestrzennych. W tym empirycznym studium analiza czynnikowa jest użyteczna we wstępnej penetracji danych, obejmujących wartości 9 cech — wskaźników relatywnego znaczenia poszczególnych działów gospodarki 248 wiejskich „counties” środkowo-zachodniej części USA, w dwóch okresach czasowych. Główne wymiary

struktury ekonomicznej identyczne dla 1953 i 1964 r., interpretuje się jako czynniki: 1) bipolarny, działalności usługowej i rolnictwa, 2) przemysłu, 3) górnictwa. Dokonuje się ekstrapolacji wyników na zbiór złożony z 899 „counties”. Klasyfikacja „counties” w oparciu o kryterium minimalizacji odległości w przestrzeni trzyczynnikowej i analizę dyskryminacji prowadzi do 4 kategorii „counties”. Rozkłady przestrzenne kategorii wykazują pewne zmiany w czasie. W 1964 r. w porównaniu z 1953 r. 662 „counties” pozostaje w tej samej kategorii, a 237 przechodzi do innej kategorii. Jako wiodące określa się dwa procesy zmian ekonomicznych: 1) rolnictwo – przemysł, 2) działalność usługowo-przemysłowa.

W dziedzinie typologii przestrzennej Polski na szczególną uwagę zasługuje praca S. E. Browna i Ch. E. Trotta (1968). Istota zastosowania analizy czynnikowej leży w redukcji 60 cech ekonomicznych (dla siatki województw Polski) do 4 podstawowych miar czynników: 1) rolnictwa, 2) przemysłu ciężkiego, 3) stopnia uspołecznienia rolnictwa, 4) wzrostu przemysłu. Analiza macierzy obserwacji z dwóch różnych okresów czasowych (lata 1958 i 1964) ma na celu uchwycenie ewentualnych zmian w strukturze i przestrzennym rozkładzie czynników. Postępowanie badawcze zamyka próba klasyfikacji dynamicznej województw w 4 wymiarach oparta na algorytmie J. H. Warda.

Należy zaznaczyć, że wysoki stopień generalizacji danych (układ wojewódzki) oraz krótki czasokres nakłada znaczne ograniczenia przestrzenno-czasowe na wyniki analizy. Pod tym względem większą precyzją odznacza się, dokonana przez Główny Urząd Statystyczny, próba określenia poziomu rozwoju ekonomicznego powiatów Polski, która prowadzi do uogólnień typologicznych³⁰. Po wykonaniu analizy czynnikowej macierzy korelacji 12 wskaźników poziomu ekonomicznego dla 317 powiatów okazało się, że pierwszy składnik stanowi najbardziej ogólną charakterystykę wszystkich zależności. Pierwszy składnik tłumaczy 38,41% ogólnej wariancji. Jednocześnie wysokie ładunki na tym składniku osiągnęła większość wskaźników. Składnik ten można więc traktować jako wystarczająco dobrą estymację zjawiska ogólnego rozwoju ekonomicznego. Analiza rozmieszczenia powiatów na skali wymiaru ogólnego rozwoju ekonomicznego prowadzi do klasyfikacji powiatów w 5 grup rozwoju. W świetle tych wyników charakteryzuje się układ wojewódzki. Po rotacji można również zinterpretować wymiar II i III jako wymiar uprzemysłowienia i wymiar rozwoju rolnictwa; wymiar I pozostaje odzwierciedleniem zjawiska ogólnego rozwoju ekonomicznego.

W ujęciu S. Rokity (1966) analiza ogranicza się do badania poziomu warunków bytu ludności w 17 powiatach województwa krakowskiego i charakteryzuje się pewnym prymitywizmem analitycznym (zastosowanie metody centroidalnej, dość dowolna interpretacja czynników). Wyniki wykazują, że na terenie województwa krakowskiego na poziom warunków bytowych ludności, określony 10 badanymi

³⁰ Określenie poziomu ekonomicznego powiatów (1968).

cechami wpływają trzy czynniki: rekreacja, industrializacja i urbanizacja. Przykładowo przedstawia się rozkład przestrzenny czynnika rekreacji.

W zakresie typologii przestrzennej i regionalizacji przeprowadzono również badania empiryczne w ujęciu czynnikowym dla Szwajcarii, Jugosławii, Indii i Kanady.

W pracy D. Steinera (1965) zasadniczym elementem badawczym jest identyfikacja i estymacja zależności, jakie zachodzą na poziomie kantonów Szwajcarii między 7 cechami, co prowadzi do określenia 2 istotnych czynników: ogólnego rozwoju gospodarczego oraz zmian w gospodarce rolnej. Badanie różnic i podobieństw w układzie dwuczynnikowym stanowi podstawę klasyfikacji kantonów w 8 grup.

Praca J. C. Fishera (1966) to studium wewnętrznej struktury Jugosławii oparte na klasyfikacji typologicznej miast i komun. Analiza czynnikowa struktury miast jako wskaźnika zróżnicowań regionalnych objęła 26 cech dla 55 miast. Cztery główne czynniki różnicujące miasta: 1) czynnik względnego rozwoju, 2) zatrudnienia w budownictwie i transporcie, 3) tradycyjnego zatrudnienia i 4) ludności czynnej zawodowo, stworzyły podstawę pomiaru odległości wieloczynnikowej między miastami, prowadząc w ten sposób do klasyfikacji typologicznej. Ustalone tą drogą typy miast miały charakter wybitnie regionalny. Transformacja danych dla 581 komun (63 cechy) metodą czynnikową pozwoliła na ustalenie następujących czynników: urbanizacji, względnego rozwoju, specjalizacji zawodowej, oraz charakteru rolniczego. W tym przypadku klasyfikacja zatraciła w dużej mierze charakter regionalny.

B. J. L. Berry i V. L. S. P. Rao (1968) stosując metodę czynnikową do analizy struktury regionalnej Andhra Pradesh, jednego z najbardziej heterogenicznych stanów Indii w zakresie ogólnego poziomu i orientacji rozwoju społeczno-ekonomicznego. Analizy dokonuje się w układzie 189 „talkus”. Podejście czynnikowe umożliwiło określenie różnic regionalnych, pierwotnie wyrażonych za pomocą 60 zmiennych, w kategoriach tylko 13 czynników, wyjaśniających 95% zmienności wspólnej. Najbardziej istotną różnicą społeczno-ekonomiczną jest wiejsko-miejska dychotomia (czynnik pierwszy i drugi). Z praktycznego punktu widzenia znajomość podstawowych układów przestrzennych staje się przydatna do planowania regionalnego, dotyczącego w przypadku Andhra Pradesh przede wszystkim rozwoju izolowanych społeczności wiejskich.

W studiach z zakresu regionalizacji Kanady B. J. L. Berry i D. M. Ray (1966) poddali analizie czynnikowej dane dla 120 „counties” Ontario i Quebec w zakresie 88 zmiennych demograficznych, rolniczych, przemysłowych i otrzymali 3 czynniki: 1) różnice między francuską i angielską Kanadą, 2) zmienność na skali rolniczo-miejskiej, 3) kontrasty między północną tarczą i południowymi nizinami.

Osobnym a ważnym rozdziałem badań, w którym model czynnikowy znajduje zastosowanie, to badanie ekonomicznej struktury regionalnej w świetle powiązań przestrzenno-ekonomicznych. Tego typu zastosowanie analizy czynnikowej wprowadza B. J. L. Berry (1968b).

Punktem wyjścia w procedurze badawczej jest macierz przestrzennego oddziały-

wania. Opierając się na założeniu, że niezmiernie wiele oddziaływań reprezentuje jedynie pewną liczbę podstawowych układów zachowania, drogą transformacji czynnikowej dochodzi się do macierzy zachowania.

Analizy układów przestrzennego zachowania B. J. L. Berry dokonuje w ramach analizy przestrzeni ekonomicznej w postaci ogólnej teorii pola. Zajmuje się korelacją układów zachowania i układów strukturalnych. Jako podstawowe założenie badawcze przyjmuje, że oddziaływania i cechy są wzajemnie powiązane i w zasadzie izomorficzne. Określenie relacji między układami zachowań i układami strukturalnymi za pomocą analizy kanonicznej prowadzi do integracji regionalizacji oddziaływań i cech, tj. regionalizacji węzłowej i jednolitej.

Modele czynnikowe dotyczące struktury przestrzennej i przestrzennego oddziaływania znalazły przykładowe zastosowanie w badaniach B. J. L. Berry nad strukturą regionalną Indii (1966, 1967a, 1968b).

Analiza czynnikowa macierzy cech dotyczyła 325 „districts” i 92 cech społeczno-ekonomicznych. Pierwszy czynnik wyznaczył regiony przemysłowo-miejskie, drugi — obszary intensywnego rolnictwa opartego na irygacji, trzeci — część wschodnią i zachodnią kraju na bazie różnic w produkcji rolnej. Łącznie zidentyfikowano 9 czynników.

Analiza macierzy oddziaływań objęła 36 „trade blocks” i 63 rodzaje przepływów towarowych. Każdy wiersz macierzy reprezentuje ukierunkowane powiązanie między parą miejsc. ($36^2 - 36$) diad charakteryzuje się w kategoriach 63 typów przepływów towarowych, tworzących kolumny macierzy. Określono strukturę 12-czynnikową. Pierwsze trzy czynniki zinterpretowano jako: 1) specjalizację produkcyjną Bombaju, Madrasu i Kalkuty oraz ich zapleczy, 2) przepływy produktów zachodu i północnego zachodu w kierunku południa, 3) przepływy produktów wschodu z silnym podkreśleniem roli Madrasu jako ośrodka redystrybucji.

Analiza kanoniczna wykazała 3 istotne układy współzależności między typami przestrzennego zachowania i wymiarami struktury przestrzennej. Podstawowymi czynnikami regionalizacji są: 1) wymiana wyspecjalizowanych produktów z regionów przemysłowych, miejskich w zamian za wyspecjalizowane produkty rolne innych regionów, 2) przepływy towarowe w 4 największych regionach, łączące się z wewnątrzregionalną wymianą produktów wyspecjalizowanych, 3) przepływy wyspecjalizowanych produktów rolnych między okręgami rolnymi i miejskimi.

W studium B. J. L. Berry’ego metodę czynnikową zastosowano również bezpośrednio do 63 macierzy przepływów towarowych stopnia 36, sformułowanych dla poszczególnych rodzajów przewożonych ładunków, w ujęciu ilościowym i wartościowym. Analiza czynnikowa dokonana techniką R i Q prowadzi do funkcjonalnej regionalizacji Indii z punktu widzenia nadawców (źródeł masy towarowej) i z punktu widzenia odbiorców (miejsc przeznaczenia).

O analogicznych badaniach prowadzonych w Polsce sygnalizują Z. Chojnicki i T. Czyż (1968, 1971).

Powyższe zastosowania metody czynnikowej dotyczyły w każdym przypadku

analizy zbioru wartości cech o znanym rozkładzie przestrzennym, w postaci macierzy geograficznej. Istotną sprawą pozostaje stwierdzenie różności tych macierzy pod względem stosowanych skal pomiaru zjawisk. W badaniach geograficznych opartych na analizie czynnikowej stosowano najczęściej macierz danych w skali ilorazowej, najbardziej precyzyjnej i najmocniejszej ze względu na możliwość wykonywania wszystkich operacji analizy statystycznej. Należy jednak zaznaczyć, że zachodziła niekiedy konieczność zastosowania również innych skal ze względu na jakościowy charakter cech lub możliwość uproszczenia operacji rachunkowych.

Przykładowo: B. J. L. Berry, H. G. Barnum, R. J. Tennant (1962), W. L. Garrison i D. F. Marble (1964), J. D. Henshall i L. J. King (1966) dokonali analizy czynnikowej macierzy systemu dwójkowego, tj. danych w skali nominalnej. Dane uzyskane ze skal porządkowych stosowali w studiach empirycznych analizy czynnikowej B. J. L. Berry (1960), J. H. Thompson i inni (1962), W. H. Bell i D. W. Stevenson (1964).

Trzeba jednak stwierdzić, że wyniki analizy czynnikowej zbioru cech dychotomicznych lub rangowanych są mało precyzyjne. Istnieją również pewne problemy związane z interpretacją wartości czynnikowych, ponieważ są one w tym przypadku wartościami wyprowadzonymi z liniowej kombinacji rang.

Z przedstawionego powyżej przeglądu wynika, że stosuje się różne techniki analizy czynnikowej macierzy geograficznej. Macierz geograficzna składa się z N wierszy odpowiadających liczbie jednostek przestrzennych i n kolumn reprezentujących zmienne. Macierz ta jest macierzą prostokątną, ponieważ z reguły $N > n$. Analiza czynnikowa tej macierzy przebiega od sformułowania macierzy korelacji stopnia n do określenia macierzy ładunków czynnikowych o wymiarach $n \times n$ w metodzie głównych komponentów i macierzy $n \times m$ w analizie czynnikowej, gdyż liczba czynników m może być równa lub mniejsza od n . Macierz wartości czynnikowych ma N wierszy oraz odpowiednio n lub m kolumn.

Macierz geograficzna kwadratowa jest stopnia N i zawiera powiązania między parami miejsc, np. migracje, przepływy towarowe, rozmowy telefoniczne. Ta macierz z wierszami reprezentującymi źródła i kolumnami odpowiadającymi miejscom przeznaczenia nosi nazwę macierzy transakcji. Rozwiązania czynnikowego dokonuje się przy zastosowaniu techniki R i Q . Te dwie techniki zastosował B. J. L. Berry w analizie struktury przepływów towarowych, traktując je jako odrębne podejścia, oraz w badaniu hierarchii ośrodków centralnych, ujmując je łącznie w postaci bezpośredniej analizy czynnikowej³¹.

Inną techniką jest wprowadzona przez M. Megee (1965c) technika M , która dobrze uzupełnia techniki R i Q , dla wykazywania zmian zachodzących w czasie. Technika M wiąże się z operowaniem macierzami geograficznymi dla t okresów jednocześnie. R. B. Cattell (1952) wymienia jeszcze technikę P , która znajduje zastosowanie do badania trendu zjawisk w ujęciu przestrzennym.

³¹ B. J. L. Berry (1966), B. J. L. Berry, H. G. Barnum (1962).

2. ANALIZA CZYNNIKOWA JAKO METODA BADAWCZA W GEOGRAFII EKONOMICZNEJ

Powyższy przegląd studiów geograficznoekonomicznych umożliwia wskazanie kręgu zagadnień, przy rozpatrywaniu których model czynnikowy znalazł już zastosowanie. Dalsze możliwości rozszerzenia zastosowań związane są z dopasowywaniem ogólnej postaci liniowego modelu czynnikowego do konkretnych sytuacji będących przedmiotem badań empirycznych oraz opracowaniem modelu czynnikowego w postaci wielomianu wyższego stopnia.

Na podstawie dotychczasowych doświadczeń trudno zdobyć się na ogólną ocenę modelu czynnikowego, należącego do najnowszych i stąd mało wypróbowanych narzędzi poznawania zjawisk przestrzenno-ekonomicznych. W związku z tym, że współcześnie na terenie geografii ekonomicznej obserwuje się ruch naukowy podejmujący problematykę zastosowań analizy czynnikowej, należy zdać sobie sprawę z pewnych przesłanek natury metodologicznej, które te zastosowania uzasadniają i warunkują.

Analiza czynnikowa jako model podziału zmiennych na podzbiory

Analiza czynnikowa zastosowana do macierzy współzależnych zmiennych umożliwia rozwikłanie zależności w odrębne układy. Każdy układ ujawnia się jako czynnik określający wyraźną więźkę maksymalnie powiązanych ze sobą zmiennych. W tym ujęciu analiza czynnikowa jest więc metodą badania heterogenicznych agregacji cech.

Metoda czynnikowa jako metoda ortogonalizacji zmiennych

Analiza czynnikowa stanowi etap wstępny w schemacie postępowania w zakresie wielu innych technik statystycznych, np. analizy dyskryminacji, które zakładają konieczność transformacji oryginalnych zmiennych³². Rola analizy czynnikowej jest ściśle związana z ortogonalizacją danych. Analiza czynnikowa sprowadza (bez straty istotnej zmienności) zbiór wyjściowych skorelowanych zmiennych do mniejszego zbioru statystycznie niezależnych wartości czynnikowych, a więc równocześnie działa jako metoda redukcji.

Metoda czynnika jako metoda redukcji masy informacji

Potrzebę stosowania metod redukcji w geografii uzasadniają dwie przesłanki: 1) konieczność wykorzystania wielkich zbiorów danych statystycznych oraz 2) postulat ekonomiczności, żądający, by możliwie mała ilość zmiennych wyjaśniła możliwie

³² Z. Casetti (1964).

wiele zjawisk. Z reguły duże wymiary macierzy geograficznych utrudniają opis i analizę zjawisk.

Analiza czynnikowa jest podstawową metodą statystyczno-matematyczną łączenia zmiennych. Sprowadza macierz oryginalnych zmiennych do układu wartości wspólnych czynników zbioru w postaci macierzy o mniejszych wymiarach. Czynniki wspólne skupiają informację dotąd rozproszoną w oryginalnych zmiennych. Redukcja w ujęciu czynnikowym pociąga za sobą stosunkowo małą stratę informacji.

Wprowadzając w postaci czynników niewielką liczbę podstawowych „wymiarów”, metoda czynnikowa ułatwia operacje matematyczne w zredukowanej przestrzeni własności. Przede wszystkim ulegają znacznemu uproszczeniu zagadnienia taksonomiczne. Należy zaznaczyć, że redukcję w ujęciu czynnikowym przeprowadza się nie tylko ze względów praktycznych, lecz również dla celów teoretycznych³³.

Analiza czynnikowa jako model pomiaru

W dążeniu do poznania rzeczywistości ze względu na złożoność zjawisk społeczno-ekonomicznych i ich zależności istnieje konieczność ograniczenia tej różnorodności i ustalenia pewnej ilości podstawowych „wymiarów”. Należy zaznaczyć, że w przeciwieństwie do innych uproszczonych sposobów redukcji przestrzeni wielocechowej (np. metody różnic przeciętnych J. Czekanowskiego, wskaźników przyrodniczych J. Perkala), które prowadzą do łączenia cech w jeden indeks, metoda czynnikowa pozwala na stopniowanie agregacji zmiennych i zatrzymanie redukcji na kilku podstawowych „wymiarach”, wartych zbadania, ale bezpośrednio nieobserwowalnych.

Analizę czynnikową można traktować jako metodę ustalania ukrytej struktury. Podstawowym postulatem analizy czynnikowej jest założenie, że istnieje zbiór czynników ukrytych, których zachodzenie jest niezbędne do wyjaśnienia zależności między obserwowalnymi cechami zjawisk. Stosując model ukrytej struktury dąży się do przeprowadzenia podziału zbioru cech na podzbiory, które posiadają ukryty wspólny czynnik. Związki między zmiennymi pochodzą wyłącznie z występowania ukrytego czynnika. Poszczególne cechy obserwowalne są wskaźnikami dla wszystkich możliwych wartości ukrytego czynnika.

Analizy zjawisk społeczno-ekonomicznych dokonuje się w kategoriach obserwowalnych cech tych zjawisk, ponieważ często nie zna się treści empirycznej kategorii opisowych. Po prostu nie potrafimy powiedzieć, jakie obserwowalne fakty i procesy należy dostrzec, aby można stwierdzić, czy zachodzi, czy też nie zachodzi określone zjawisko. Podstawę sądów o zjawiskach stanowią cechy zwane wskaźnikami³⁴. Tak więc liczba zatrudnionych w przemyśle jest wskaźnikiem zjawiska industrial-

³³ Patrz: Analiza czynnikowa jako model pomiaru.

³⁴ Definicja wskaźnika w tym znaczeniu jest następująca: Wskaźnikiem zjawiska *Z* nazywa się taką cechę *W*, której zaobserwowanie pozwoli nam orzec, że zaszło zjawisko *Z*. Według S. Nowaka (1965).

zacji, a posiadanie telewizora — wskaźnikiem poziomu konsumpcji. Wskaźnik jest definiensem zjawiska lub o istnieniu zjawiska inferujemy z faktu zajścia wskaźnika.

Należy zaznaczyć, że analiza czynnikowa jest właśnie metodą, która zmierza do wykrycia charakteru tych nieobserwowalnych zjawisk, traktowanych jako ukryte czynniki i prowadzi do sformułowania jednoznacznej definicji zjawisk wskaźnikowanych w kategoriach cech. W tym ujęciu czynniki są pewnymi strukturami, reprezentującymi zbiór rzeczywistych zmiennych w sferze zjawisk społeczno-ekonomicznych. Geograf często rozważa te matematyczne czynniki jako odpowiedniki „element-complex” R. Hartshorne’a (1959, s. 74, 123—127).

Analiza czynnikowa stwarza więc nowe poważne możliwości poznawcze w zakresie identyfikacji ukrytych zjawisk. Jest modelem pomiaru, gdy chodzi o wykrycie czynników, które nie są bezpośrednio obserwowalne. Przykładowo — przy założeniu, że struktura regionalna jest liniową funkcją ukrytych prostych układów, można wykryć, jakie czynniki w liniowym modelu obrazują prostą strukturę. Analiza czynnikowa w roli modelu pomiaru wprowadza więc nową kategorię ukrytej struktury przestrzennej³⁵.

Układy czynnikowe mogą być traktowane jako kategorie opisowe lub klasyfikacyjne. Wychodząc z nowoczesnej koncepcji przyczynowości jako matematycznej zależności między zjawiskami, za R. B. Cattellem (1952) definiuje się również czynniki jako przyczyny układów, które one reprezentują. Wykryte czynniki zwykle nie są przyczynami prostymi, jednorodnymi, ale o złożonej strukturze.

Analiza czynnikowa w roli modelu pomiaru wielkości podstawowych w badaniach społeczno-ekonomicznych operuje skalowaniem. Badacz często chce rozwinąć skalę, na której indywidua mają przypisaną określoną wartość i mogą być porównywane. Skala może się odnosić do takich zjawisk, jak urbanizacja bądź poziom rozwoju ekonomicznego. Podstawowym problemem w rozwinięciu skali jest ważenie łączonych cech. Analiza czynnikowa dostarcza rozwiązania przez podział cech według niezależnych źródeł zmienności. Każdy czynnik reprezentuje skalę opartą na empirycznych zależnościach między cechami. Analiza czynnikowa operuje dla cech podlegających łączeniu wagami, wyprowadzonymi z modelu matematycznego. Wynikowe wartości czynnikowe są skalami rozwiniętymi przez połączenie tych cech, z uwzględnieniem określonych wag.

Charakter czynnika musi być inferowany z treści cech obserwowalnych, które mogą składać się na daną skalę. Do tej skali powinny być włączone cechy logicznie ze sobą powiązane, o treściowym pokrewieństwie. W budowie skali pomagają więc dokładna znajomość przedmiotu. Badacz musi wiedzieć, co chce mierzyć.

Należy jednak zaznaczyć, że można formułować właściwie tylko pewne hipotezy co do natury ujawnionego czynnika. Konieczna jest więc empiryczna kontrola czynników.

³⁵ Patrz P. R. Gould (1967).

Analiza czynnikowa jako metoda eksperymentu

Analiza czynnikowa odgrywa zasadniczą rolę w zakresie tych nauk, w których kontrolowany eksperyment jest trudny lub niemożliwy do przeprowadzenia³⁶. Do tej kategorii nauk należy geografia, nauka obserwacyjna o wysokiej kompleksowości badanych zjawisk. W analizie przestrzenno-ekonomicznej poszczególne sytuacje mogą być badane w naturalnych rzeczywistych warunkach, a nie w kontrolowanych warunkach laboratoryjnych, jak np. w fizyce. Utrudnia to wykrywanie przyczynowych zależności o charakterze prawidłowości bezwyjątkowych w chaosie występujących relacji³⁷.

W tym przypadku celowe jest zastosowanie analizy czynnikowej, która spełnia funkcję „laboratorium”³⁸. Kompleks współzależnych zjawisk zostaje poddany systematycznemu badaniu; podlega: 1) „rozplątaniu” przez wyodrębnienie różnych źródeł zmienności, 2) redukcji do stosunkowo prostej kombinacji liniowej czynników systematycznych (funkcjonalnych całości) na drodze izolacji nieistotnych zależności oraz 3) rozpoznaniu.

Analiza czynnikowa jako metoda klasyfikacji

Analiza czynnikowa służy do rozwiązywania zagadnień taksonomicznych. Może stanowić metodę grupowania współzależnych zmiennych w nowe kategorie, takie jak urbanizacja, industrializacja (przy zastosowaniu techniki *R*) albo metodę grupowania jednostek obserwacji w typy na zasadzie największego podobieństwa lub zachowań (technika *Q*).

Analiza czynnikowa jako narzędzie testowania hipotez

Analiza czynnikowa znajduje zastosowanie do weryfikacji sformułowanych *ex ante* hipotetycznych „wymiarów”. Wymiar oznacza zwykle wiązkę wysoko skorelowanych cech albo zachowań. Analiza czynnikowa jest regułą postępowania w procedurze sprawdzania empirycznego istnienia tych postulowanych wymiarów. Za pomocą testów statystycznych istotności można określić granicę, od której począwszy zaobserwowane różnice między wynikami analizy czynnikowej a hipotezą będziemy uważać za istotne, tj. świadczące przeciwko sprawdzanej hipotezie. W praktyce analizę czynnikową stosuje się również do sprawdzania hipotez dotyczących układów przestrzennych zjawisk³⁹.

³⁶ R. B. Cattell (1952), s. 3—4, 357—359.

³⁷ Z. Chojnicki (1969).

³⁸ R. J. Rummel (1967).

³⁹ Patrz J. C. Abiodun (1967).

Analiza czynnikowa a teoria geografii

Analiza czynnikowa odgrywa zasadniczą rolę w konstruowaniu teorii geografii ekonomicznej, zarówno jako technika indukcyjna, jak i technika dedukcyjna, w zależności od każdorazowego podejścia i interpretacji wyników. Znana jest przede wszystkim jako procedura uogólniająca, oparta na metodach indukcji. Rzadziej stosowana jest jako metoda systematyzacji twierdzeń⁴⁰.

⁴⁰ Patrz R. Domański (1967).

III. STUDIUM EKONOMICZNEJ STRUKTURY REGIONALNEJ POLSKI

1. KONCEPCJA UKRYTEJ STRUKTURY PRZESTRZENI SPOŁECZNO-EKONOMICZNEJ W ŚWIETLE ANALIZY CZYNNIKOWEJ

Według K. Dziewońskiego (1967) ogólną przestrzeń społeczno-ekonomiczną, należącą do kategorii metaukładów, można traktować jako zbiór cząstkowych przestrzeni społeczno-ekonomicznych. Zbiór ten mimo mnogości jest zbiorem skończonym. Przestrzenie społeczno-ekonomiczne, ujmujące różne zjawiska społeczno-ekonomiczne (demograficzne, przemysłowe, rolnicze, transportowe, usługowe) „...dotyczą jednej i tej samej rzeczywistości, stanowiąc jej wycinkowe odbicia czy modele, są one współzależne, ale oczywiście nieidentyczne” (s. 37). Ogólna przestrzeń społeczno-ekonomiczna jest wypadkową wszystkich przestrzeni cząstkowych, które tworzą jej ukrytą strukturę.

W ramach takiej koncepcji przestrzeni analiza czynnikowa jest metodą ustalania ukrytej struktury ogólnej przestrzeni społeczno-ekonomicznej. Za pomocą tej metody dokonuje się agregacji danych, dotyczących całości gospodarki narodowej w podstawowe układy czynnikowe, reprezentujące określone przestrzenie. Czynniki identyfikują więc podstawowe przestrzenie elementarne ogólnej przestrzeni społeczno-ekonomicznej kraju. Przestrzenie te nie wyczerpują całej zbiorowości przestrzeni cząstkowych, jednak mają podstawowe znaczenie, gdyż czynniki je determinujące są najbardziej istotne. Analiza czynnikowa prowadzi zatem równocześnie do redukcji ogólnej przestrzeni społeczno-ekonomicznej.

Należy zwrócić uwagę na pozorną sprzeczność między założeniem o ortogonalności teoretycznych czynników a korelacją empirycznych przestrzeni cząstkowych. Niezależność czynników ma jednak charakter względny: korelacje między czynnikami są bardzo słabe w porównaniu z silnymi korelacjami, występującymi między zmiennymi w obszarze wewnętrznym układu każdego czynnika.

W ramach cząstkowych przestrzeni społeczno-ekonomicznych występują podprzestrzenie — regiony cząstkowe, które tworzą ich strukturę. Tak więc podział przestrzeni pojedynczego czynnika jest związany z koncepcją regionalizacji cząstkowej.

Regiony cząstkowe mają odbicie w ogólnej przestrzeni społeczno-ekonomicznej,

która jest sumą przestrzeni cząstkowych. Integracja regionów cząstkowych prowadzi do wydzielenia podprzestrzeni ogólnej przestrzeni społeczno-ekonomicznej, tzn. sumarycznych, ogólnych regionów ekonomicznych. Ta czynność jest odpowiednikiem regionalizacji całkowitej (kompleksowej).

W interpretacji czynnikowej tego rodzaju podejście oznacza klasyfikację zbioru rozpatrywanego w przestrzeni wieloczynnikowej na podzbiory maksymalnie jednorodnie i ciągłe.

2. PROCEDURA REGIONALIZACJI EKONOMICZNO-GEOGRAFICZNEJ

2.1. MACIERZ OBSERWACJI

A. Interpretacja empiryczna macierzy

Punkt wyjścia w postępowaniu badawczym stanowi macierz geograficzna, czyli uporządkowany zbiór zdarzeń geograficzno-ekonomicznych. Zdarzeniami są zjawiska związane z całością działalności gospodarczej na obszarze Polski. Zdarzenia te, określone co do charakteru zjawiska, jego wielkości i lokalizacji, traktuje się jako „cechy” obszaru. Np. zbiór obiektów materialnych zagregowanych na danym obszarze (ludność) rozpatruje się jako jego cechę. Operując pojęciem „cecha obszaru”, dokonuje się pewnego skoku myślowego. W określeniu tym tkwi pewien element konwencji, nieodłączny od każdej generalizacji. Konwencja czy arbitralność uwidaczniają się jeszcze wyraźniej, jeśli weźmie się pod uwagę nieciągły (z reguły) charakter zjawisk społeczno-ekonomicznych⁴¹. W rzeczywistości cechy stanowią więc podstawę sądów o zjawiskach, które występują na jakimś obszarze w sposób mniej lub bardziej ciągły.

W niniejszym studium cechy odnoszą się do rozmieszczenia (stanów) zjawisk i mają charakter cech skalarnych. Cechom przyporządkowane są liczby rzeczywiste, które stanowią miary tych cech. Stosowanie cech mierzalnych w operacji badawczej zapewnia uchwycenie natężenia zjawisk, a więc umożliwia porównanie stanów badanego zjawiska w różnych punktach przestrzeni społeczno-ekonomicznej Polski. Cechy odzwierciedlają treść gospodarczą obszaru; dotyczą ludności, produkcji przemysłowej i rolniczej, transportu, usług (tab. 1). Dane wykorzystane w niniejszej pracy są wzięte z dostępnych źródeł informacji statystycznej (tab. 2)⁴².

Należy zaznaczyć, że z punktu widzenia celów niniejszego opracowania wyjściowy materiał statystyczny nie reprezentuje, niestety, wszystkich istotnych własności przestrzeni społeczno-ekonomicznej Polski. Jednak wystąpiły duże trudności w doborze cech, wynikające z poważnych braków w materiałach statystycznych w przekroju powiatowym. Również z identycznych przyczyn układ wartości cech nie

⁴¹ A. Wróbel (1967), s. 78—79.

⁴² Ze względu na ograniczoność miejsca nie zamieszcza się olbrzymiego zbioru wyjściowych danych statystycznych.

Tabela 1. Cechy społeczno-ekonomiczne

Nr cechy	Nazwa cechy
1	Gęstość zaludnienia w 1965 r.
2	Przyrost naturalny na 1000 ludności w 1965 r.
3	Ludność miejska w % ogółu ludności w 1965 r.
4	Ludność czynna zawodowo w % ogółu ludności w 1960 r.
5	Ludność czynna zawodowo poza rolnictwem w % ludności czynnej zawodowo w 1960 r.
6	Ludność utrzymująca się ze źródeł pozarolniczych w % ogółu ludności w 1960 r.
7	Ludność rolnicza na 100 ha użytków rolnych w 1960 r.
8	Nakłady inwestycyjne w cenach bieżących na 1 mieszkańca w 1965 r. (w zł)
9	Nakłady inwestycyjne w cenach bieżących na 1 km ² w 1965 r. (w tys. zł)
10	Wartość środków trwałych na 1 km ² (w tys. zł) w 1961 r.
11	Zatrudnienie w przemyśle na 1000 ludności w 1965 r.
12	Zatrudnienie w przemyśle na 100 km ² w 1965 r.
13	Produkcja globalna przemysłu na 1 mieszkańca w 1965 r. (w zł)
14	Użytki rolne w % ogólnej powierzchni w 1965 r.
15	Udział uprawy zbóż w ogólnej powierzchni zasiewów w 1965 r. (w %)
16	Udział uprawy ziemniaków w ogólnej powierzchni zasiewów w 1965 r. (w %)
17	Udział uprawy buraków cukrowych w ogólnej powierzchni zasiewów w 1965 r. (w %)
18	Bydło na 100 ha użytków rolnych w 1965 r. (w szt.)
19	Trzoda na 100 ha użytków rolnych w 1965 r. (w szt.)
20	Zbiory pszenicy na 1 mieszkańca w 1965 r. (w q)
21	Indywidualne gospodarstwa rolne zelektryfikowane w % ogółu gospodarstw indywidualnych w 1965 r.
22	Powierzchnia leśna na 1 mieszkańca w 1965 r. (w ha)
23	Lesistość w 1965 r.
24	Długość linii kolejowych przypadająca na 100 km ² w 1965 r. (w km)
25	Długość dróg z komunikacją autobusową przypadająca na 100 km ² w 1965 r. (w km)
26	Liczba pociągokilometrów w ciągu doby przypadająca na 100 km ² w 1965 r.
27	Liczba autobusokilometrów w ciągu doby przypadająca na 100 km ² w 1965 r.
28	Liczba pociągokilometrów w ciągu doby przypadająca na 10 000 mieszkańców w 1965 r.
29	Liczba autobusokilometrów w ciągu doby przypadająca na 10 000 mieszkańców w 1965 r.
30	Liczba ludności przypadająca na 1 punkt sprzedaży w uspołecznionym handlu detalicznym w 1965 r.
31	Sprzedaż w handlu detalicznym uspołecznionym na 1 mieszkańca w 1965 r. (w zł)
32	Abonenci telewizji na 1000 ludności w 1965 r.
33	Lekarze medycyny na 10 000 ludności w 1965 r.

odpowiada w pełni postulatowi jedności czasu, wymaganej w badaniach statystycznych. Większość danych co prawda pochodzi z 1965 r. (65% cech), jednak dane dotyczące struktury ludności (zawarte tylko w spisach) są przyjęte według stanu z 1960 r., a wartość środków trwałych na podstawie 1961 r. Jednakże ta nieznaczna niejednorodność nie może zaważyć w sposób istotny na wynikach analizy przestrzennej,

Tabela 2. Źródła danych statystycznych

Lp.	Źródła	Dane statystyczne dotyczące powiatów Polski
1	<p>Przekroje terenowe 1945—1965, Seria „Statystyka Regionalna”, Nr 7, GUS, Warszawa 1967, s. 1014—1048</p> <p>s. 1042—1048</p> <p>s. 1057—1069</p> <p>s. 1070—1073</p> <p>s. 1075—1087</p> <p>s. 1094—1099</p> <p>s. 1127—1141</p> <p>s. 1174—1188</p> <p>s. 1189—1202</p>	<p>Powierzchnia w 1965 r. Ludność w 1960 i 1965 r. Gęstość zaludnienia w 1965 r. Przyrost naturalny na 1000 ludności w 1965 r. Ludność miejska w % ogółu ludności w 1965 r.</p> <p>Ludność czynna zawodowo w 1960 r. Ludność czynna zawodowo poza rolnictwem w 1960 r. Ludność utrzymująca się ze źródeł pozarolniczych w % ogółu ludności w 1960 r. Ludność czynna zawodowo w rolnictwie w 1960 r.</p> <p>Zatrudnienie w przemyśle na 1000 ludności w 1965 r. Zatrudnienie w przemyśle na 100 km² w 1965 r.</p> <p>Produkcja globalna przemysłu w 1965 r.</p> <p>Powierzchnia zasiewów zbóż w % ogólnej powierzchni zasiewów w 1965 r. Powierzchnia uprawy ziemniaków w % ogólnej powierzchni zasiewów w 1965 r. Powierzchnia uprawy buraków cukrowych w % ogólnej powierzchni zasiewów w 1965 r.</p> <p>Bydło w 1965 r. Trzoda chlewna w 1965 r. Zelektryfikowane indywidualne gospodarstwa rolne w % ogólnej liczby gospodarstw w 1965 r.</p> <p>Liczba ludności przypadająca na 1 punkt sprzedaży w uspołecznionym handlu detalicznym w 1965 r. Sprzedaż w handlu detalicznym uspołecznionym na 1 mieszkańca w 1965 r.</p> <p>Abonenci telewizji na 1000 ludności w 1965 r. Lekarze na 10 000 ludności w 1965 r.</p>
2	<p>Rozwój gospodarczy powiatów w latach 1950—1965, „Statystyka Regionalna”, z. 5, GUS, Warszawa 1967, s. 29—51 s. 124—149</p>	<p>Ludność rolnicza na 100 ha użytków rolnych w 1960 r. Powierzchnia użytków rolnych w 1960 i 1965 r.</p>

Lp.	Źródła	Dane statystyczne dotyczące powiatów Polski
3	Lijewski T., Lenk J., Piotrowska H., Rozwój komunikacji kolejowej i autobusowej w Polsce w okresie 1946—1965, Dokumentacja Geograficzna IG PAN, Warszawa 1967, s. 131—142	Długość linii kolejowej na 100 km ² w 1965 r. Długość dróg z komunikacją autobusową na 100 km ² w 1965 r. Liczba pociągokilometrów w ciągu doby przypadająca na 100 km ² w 1965 r. Liczba autobusokilometrów w ciągu doby przypadająca na 100 km ² w 1965 r. Liczba pociągokilometrów w ciągu doby przypadająca na 10 000 mieszkańców w 1965 r. Liczba autobusokilometrów w ciągu doby przypadająca na 10 000 mieszkańców w 1965 r.
4	Rocznik Statystyczny Leśnictwa i Przemysłu Drzewnego 1966, Warszawa 1967, s. 118—129	Powierzchnia leśna na 1 mieszkańca w 1965 r. Lesistość w 1965 r.
5	Nakłady inwestycyjne według powiatów 1961—1965 Seria „Statystyka Regionalna”, z. 4, GUS, Warszawa 1966, s. 116—150 s. 151—162 s. 170—180	Nakłady inwestycyjne w 1965 r. Nakłady inwestycyjne na 1 mieszkańca w 1965 r. Wartość środków trwałych na 1 km ² w 1961 r.
6	Materiały Instytutu Ekonomiki Rolnej	Zbiory pszenicy w 1965 r.

ponieważ wielkości cech ludnościowych w okresie 1960—1965 nie wykazywały nagłych zmian.

W kontekście proponowanej procedury regionalizacyjnej zdarzenia geograficzno-ekonomiczne mają określoną lokalizację w zmodyfikowanym układzie administracyjnym Polski z 1965 r. Za podstawową jednostkę przestrzenną ugrupowania

zjawisk przyjęto powiat. Ze względu na wymiary geograficzne Polski jest to dla zagadnień struktury gospodarczej skala generalizacji zbyt daleko posunięta. Powiatów nie można traktować jako obszary dostatecznie jednorodne. Jednak ze względu na niedostępność informacji statystycznych dla jednostek mniejszych niż powiat, tj. gromad, stało się to koniecznością. Należy również wziąć pod uwagę fakt, że uwzględnienie zbioru złożonego z powiatów zapewnia realizację programu badawczego; z dalszym wzrostem ilości jednostek przestrzennych zwiększają się nieproporcjonalnie trudności techniczne analizy, nawet w przypadku zastosowania maszyn cyfrowych.

Modyfikacja faktycznego układu powiatowego polegała na: 1) włączeniu 52 miast o statusie powiatu, położonych wewnątrz powiatów wiejskich do właściwych powiatów; 2) zaliczeniu 9 miast o statusie powiatu, położonych na styku kilku powiatów, do tego powiatu, z którego zostały w swoim czasie wydzielone; 3) potraktowanie 5 miast wydzielonych na prawach województw, zespołu 10 miast Górnośląskiego Okręgu Przemysłowego oraz Trójmiasta jako odrębnych jednostek przestrzennych.

W ten sposób zbiór 391 powiatów i 5 miast wydzielonych wg podziału administracyjnego z 1965 r. sprowadzono do 324 jednostek przestrzennych rangi powiatu (mapa 1).

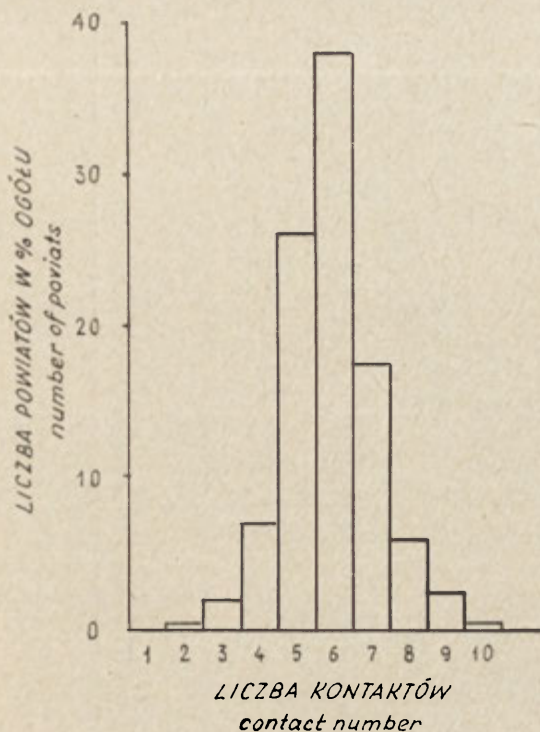
Włączenie miast stanowiących powiaty do powiatów wiejskich i traktowanie tych dwóch jednostek podziału administracyjnego kraju jako jednej jednostki przestrzennej uznano za uzasadnione, ponieważ miasto jest integralnie związane z obszarem powiatu wiejskiego, w pewnym sensie leży na jego obszarze.

Przyjęty układ odniesienia zjawisk odpowiada właściwie podziałowi administracyjnemu, ściśle związanemu z układem przestrzennej organizacji działalności gospodarczej.

Siatka podstawowych jednostek przestrzennych jest układem nieregularnym; powiaty zróżnicowane są pod względem wielkości powierzchni i kształtu. Rozpiętość między najmniejszą i największą jednostką waha się w granicach 1:10, jeśli chodzi o powierzchnię (m. Poznań 220 km² i Białystok 2308 km²), oraz 1:75 (Ustrzyki Dolne 17,3 tys. i Górnośląski Okręg Przemysłowy 1309,2 tys.) w zakresie ludności.

W związku ze zmienną wielkością jednostek powstaje zagadnienie zniekształcenia pomiaru przestrzennego. Celem ograniczenia stopnia tego zniekształcenia zastosowano standaryzację nieregularnych obszarów jednostkowych przez wprowadzenie zamiast wartości bezwzględnych cech odpowiednich wskaźników struktury i natężenia. Wskaźniki te mają postać mierników per capita (produkcja globalna przemysłu na głowę mieszkańca, zbiory pszenicy na 1 mieszkańca itp.) lub koncentracji przestrzennej (gęstość zaludnienia, ludność rolnicza przypadająca na 100 ha użytków rolnych itp.).

W zakresie form geometrycznych jednostek przyjęty układ terytorialny charakteryzuje się wyraźną nieregularnością. Jednak badania rozkładu liczby kontaktów dla poszczególnych jednostek wykazały, że w tym schemacie przestrzennym wystę-

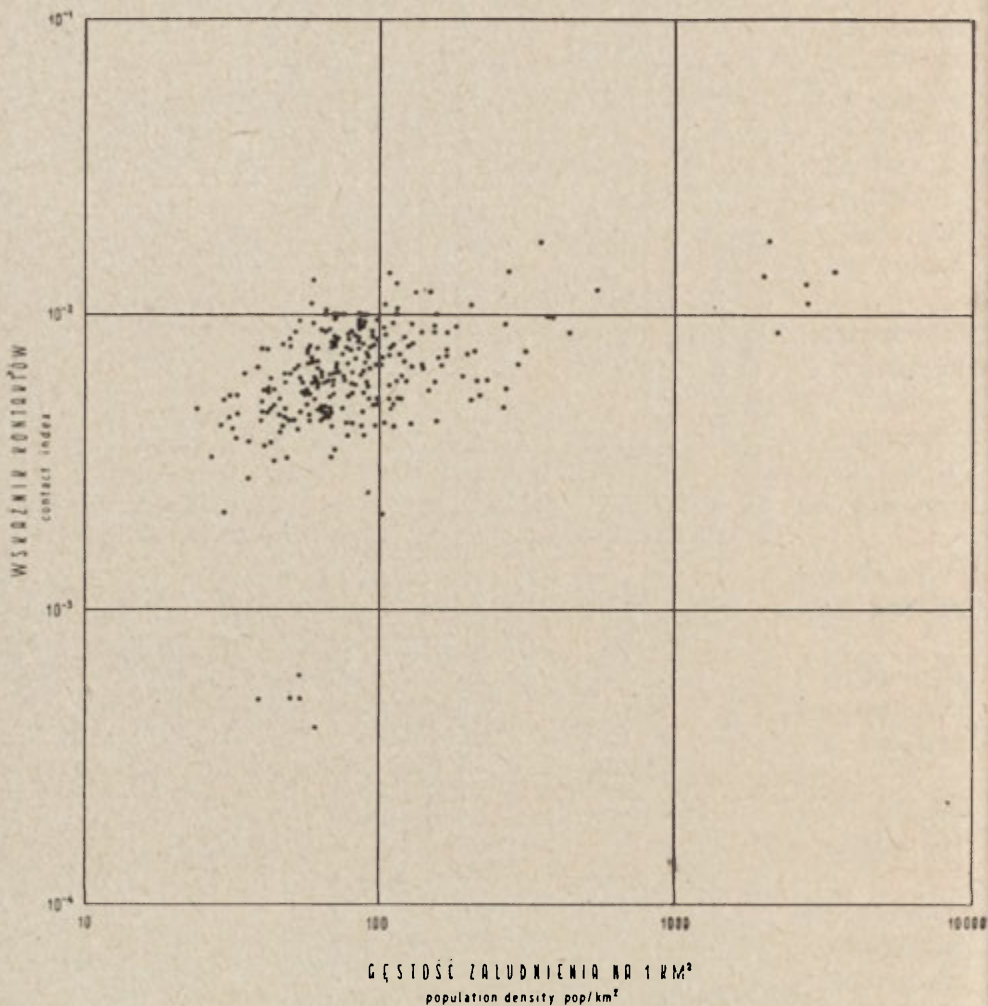


Ryc. 6. Rozkład liczby kontaktów według podstawowych jednostek przestrzennych (powiatów)

Fig. 6. Histogram of contact numbers by areal units (poviats)

Tabela 3. Podstawowe jednostki przestrzenne wg liczby kontaktów

Liczba kontaktów	Liczba podstawowych jednostek przestrzennych	
	w liczbach bezwzględnych	w %
2	1	0,4
3	5	1,9
4	18	7,0
5	67	26,2
6	98	38,3
7	45	17,6
8	15	5,9
9	6	2,3
10	1	0,4
Razem	256	100,0



Ryc. 7. Zależność między wskaźnikiem kontaktów a gęstością zaludnienia według podstawowych jednostek przestrzennych (powiatów)

Fig. 7. Relationship of contact index and population density by areal units (poviats)

pują pewne prawidłowości. Liczba kontaktów dla 256 jednostek przestrzennych zmienia się od 2 do 10, przy czym 38,3% jednostek ma 6 kontaktów (tyle co heksagon), a średnia liczba kontaktów wynosi 5,9 (tab. 3 i ryc. 6)⁴³. Tak więc można stwierdzić, że ukrytą formą tego rzeczywistego układu przestrzennego jest układ heksagonalny, który w świetle badań teoretycznych, traktuje się jako optymalne rozwiązanie organizacji przestrzeni.

⁴³ Ze zbioru jednostek wyłączono powiaty nadmorskie i graniczne.

Równocześnie badania zależności między gęstością zaludnienia a wskaźnikiem kontaktów ujawniły, że jednostki o większej gęstości zaludnienia wykazują tendencję do wzrostu liczby kontaktów (ryc. 7)⁴⁴. Konfiguracja układu związana jest więc z przestrzenią ludności, co wydaje się prostą konsekwencją charakteru rozpatrywanego układu.

Należy zaznaczyć, że zastosowanie powiatu jako jednostki podstawowej o charakterze quasi-naturalnym w procedurze delimitacji regionów ogólnoeconomicznych wymaga traktowania rzeczywistego podziału przestrzeni, czyli regionalizacji w sensie poznawczym, jako pewnej funkcji wstępnie przyjętego układu odniesienia.

B. Sformułowanie matematyczne

Celem wyspecyfikowania zbioru złożonego z 324 podstawowych jednostek przestrzennych w przestrzeni cechowej 33-wymiarowej można zastosować jeden z dwóch równoważnych sposobów:

- 1) wymienić składowe $P_1, P_2, \dots, P_{324}(x_{1,1}, x_{1,2}, \dots, x_{1,33}), (x_{2,1}, x_{2,2}, \dots, x_{2,33}), \dots, (x_{324,1}, x_{324,2}, \dots, x_{324,33})$;
- 2) podać macierz

$$\begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,33} \\ x_{2,1} & x_{2,2} & \dots & x_{2,33} \\ \dots & \dots & \dots & \dots \\ x_{324,1} & x_{324,2} & \dots & x_{324,33} \end{bmatrix}$$

W tej macierzy każdy wiersz przedstawia jedną jednostkę przestrzenną opisaną jako punkt (wektor) przestrzeni 33-wymiarowej. Każda kolumna macierzy to współrzędne punktów (wartości cech) odłożone na jakiejś osi 33-wymiarowego prostokątnego układu współrzędnych.

Wprowadzenie miar poszczególnych składowych wektorów jest równoznaczne z określeniem 33-wymiarowej przestrzeni wektorowej.

Wartości zmiennych pierwotnych podlegają jednak przekształceniu. Przeprowadza się standaryzację poszczególnych zmiennych za pomocą wzoru

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

Wartości parametrów \bar{x}_j i s_j (średnia arytmetyczna i odchylenie standardowe) zmiennych pierwotnych podaje tab. 4. Wyniki standaryzacji zawiera macierz zmiennych standaryzowanych, która stanowi ostateczną postać macierzy obserwacji i jest przedmiotem dalszej analizy statystycznej⁴⁵.

⁴⁴ Wskaźnik kontaktu P. Haggetta jest stosunkiem liczby kontaktów do powierzchni powiatu. P. Haggett (1965).

⁴⁵ Macierzy zmiennych standaryzowanych nie zamieszczono w pracy ze względu na ograniczoną ilość miejsca.

Tabela 4. Średnie arytmetyczne i odchylenie standardowe zmiennych pierwotnych

Nr zmiennej	x_j	s_j
1	145,52	365,53
2	11,89	5,40
3	35,47	20,46
4	47,65	6,41
5	42,43	20,71
6	48,63	18,49
7	60,13	21,97
8	3716,04	3202,14
9	686,87	2215,22
10	5701,80	22916,71
11	80,18	64,24
12	2209,13	8771,28
13	15442,31	13249,41
14	61,03	13,30
15	51,94	5,62
16	18,37	4,72
17	3,09	3,30
18	53,29	29,85
19	69,94	22,40
20	145,74	127,58
21	78,82	15,97
22	0,40	0,40
23	24,63	11,95
24	8,88	5,65
25	25,15	7,42
26	177,07	300,96
27	559,03	492,03
28	156,36	87,88
29	575,39	203,73
30	173,10	23,82
31	7591,60	2314,40
32	45,03	31,87
33	7,76	5,36

C. Rozkłady zmiennych

Analiza wartości zmiennych prowadzi do stwierdzenia, że 98,89% wartości zbioru zmiennych jest zawartych w przedziale $(-3, +3)^{46}$. Przedziały zmienności poszczególnych cech nie są wcale identyczne, a krzywe rozkładu mają przebieg bardzo nieregularny (tab. 5). Rozstęp zawarty ściśle w granicach odchyień od średniej mają zmienne 2,4,5,6,14,18,21,23, a w pobliżu tych granic 3,7,13,16,19,25,28,32

⁴⁶ Patrz przyp. 45.

Tabela 5. Przedziały zmienności cech znormalizowanych

Nr cechy	Wartość minimalna	Wartość maksymalna
1	-0,332	9,100
2	-1,812	2,058
3	-1,480	3,155
4	-2,330	2,549
5	-1,614	2,736
6	-1,965	2,687
7	-2,008	3,362
8	-1,154	9,410
9	-0,287	8,413
10	-0,247	9,534
11	-1,186	4,091
12	-0,250	11,174
13	-1,104	3,654
14	-2,927	2,074
15	-4,914	1,807
16	-3,000	3,035
17	-0,936	4,010
18	-0,780	1,193
19	-2,595	3,110
20	-1,140	4,584
21	-2,330	1,326
22	-0,998	4,192
23	-2,061	2,968
24	-3,068	4,906
25	-1,986	3,051
26	-0,588	10,569
27	-0,945	7,583
28	-1,779	3,205
29	-2,608	4,195
30	-3,068	4,906
31	-1,680	5,078
32	-1,350	3,451
33	-1,075	6,972

(tab. 1). Są to przede wszystkim cechy demograficzne oraz cechy dotyczące produkcji rolnej i transportu (w ujęciu per capita). Stosunkowo mały zakres zmienności tych cech jest wynikiem powszechności w skali Polski reprezentowanych przez nie zjawisk ludnościowych i rolniczych.

Natomiast charakterystyczną właściwością rozkładów pozostałych zmiennych jest duże rozproszenie wartości wokół średniej i wyraźna dodatnia asymetria. Wyjątkowo wysokie skrajne dodatnie wartości wykazują zmienne 1,8,9,10,12,26,27,33 (o postaci wskaźników koncentracji przestrzennej, z wyjątkiem cechy 8 i 33), ponie-

waż największe różnice występują w kraju w stanie uprzemysłowienia, intensywnej gospodarki rolnej i natężenia ruchu transportowego. Szczegółowe badanie rozkładu przestrzennego wartości skrajnych tych zmiennych wykazało, że ich występowanie związane jest przede wszystkim z jednostkami przestrzennymi o charakterze miejscowym (Górnośląski Okręg Przemysłowy, Trójmiasto, m. Poznań, m. Warszawa, m. Łódź, m. Wrocław).

2.2. ANALIZA PRZESTRZENI WIELOCECHOWEJ

A. Wstępna analiza korelacji zmiennych metodą dendrytową

Zmienne zbioru wykazują zależności korelacyjne. Macierz korelacji o postaci

$$R = \begin{bmatrix} 1 & r_{1,2} & \dots & r_{1,33} \\ r_{2,1} & 1 & \dots & r_{2,33} \\ \dots & \dots & \dots & \dots \\ r_{33,1} & r_{33,2} & \dots & 1 \end{bmatrix}$$

jest stopnia 33 i zawiera $\frac{1}{2}(33^2 - 33) = 528$ współczynników korelacji (tab. 6). Współczynniki korelacji obliczono ze wzoru K. Pearsona:

$$r_{jk} = \frac{\sum_{i=1}^N (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^N (x_{ij} - \bar{x}_j)^2 \sum_{i=1}^N (x_{ik} - \bar{x}_k)^2}} \quad \begin{matrix} (i = 1, 2, \dots, 324) \\ j, k = 1, 2, \dots, 33) \end{matrix}$$

Wartość krytyczna współczynnika korelacji ustalona przy 322 stopniach swobody dla poziomu istotności $\alpha = 0,01$ wynosi 0,14413⁴⁷. Jeżeli wartość r jest zawarta między $-0,14413$ a $+0,14413$, to współczynnik korelacji jest nieistotny.

Zmienne charakteryzują się dodatnimi i ujemnymi zależnościami z innymi zmiennymi. Jednak dominują korelacje dodatnie. Korelacje istotne w zbiorze stanowią 66,6% zależności, co jest dowodem ścisłych zależności zjawisk społeczno-ekonomicznych. Stosunek między liczebnością współczynników pozytywnych i negatywnych kształtuje się jak 1:0,6. Przewagę istotnych zależności ujemnych nad dodatnimi wykazują zmienne związane z działalnością rolniczą: 2,4,7,14,15,16,20,22,23,29,30. W przypadku zmiennych 14,15,16,19,20,28,29,30 ujemne współczynniki są maksymalnymi korelacjami tych zmiennych. Natomiast wysokie pozytywne zależności z innymi zmiennymi ($0,65 \leq r_{\max} \leq 0,98$) charakteryzują wskaźniki 1,3,5,6,9,10, 11,12,13,21,24,25,26,27,31,32,33, dotyczące urbanizacji, uprzemysłowienia, transportu, usług.

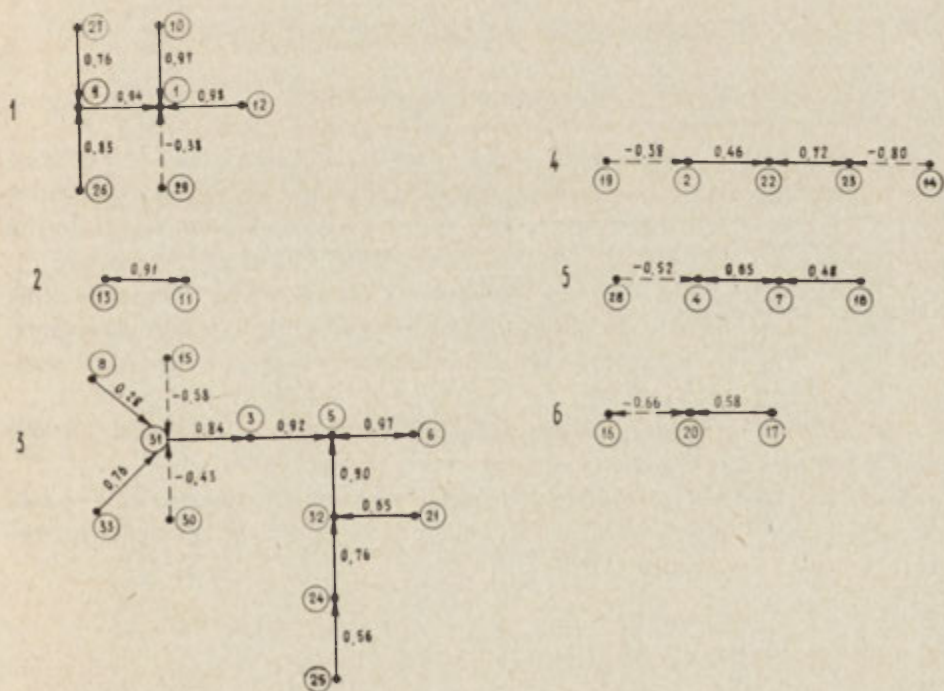
Zależności występujące w zbiorze charakteryzują się wysoką złożonością. W celu wyodrębnienia podstawowych układów zmiennych zależnych zastosowano metodę dendrytową grupowania zmiennych na podstawie kryterium maksymalnej korelacji (tab. 6).

⁴⁷ J. E. Freund (1968), s. 351, 353.

Pierwsze stadium grupowania obejmuje ustalenie podstawowych par wzajemnie zależnych zmiennych. Każda zmienna zawarta w „parze” jest silniej skorelowana z drugą zmienną aniżeli z jakąkolwiek inną zmienną zbioru. W rezultacie wydzielono 6 podstawowych par zmiennych:

1	0,98	12
5	0,97	6
22	0,72	23
4	0,65	7
13	0,91	11
16	-0,66	20

W drugim stadium grupowania uwzględnia się również maksymalne korelacje o charakterze zależności jednokierunkowej. Takie postępowanie prowadzi do wyodrębnienia 6 układów powiązań, które kształtują morfologię całego zbioru zmiennych (ryc. 8).



Ryc. 8. Układy zmiennych
Fig. 8. Clusters of variables

Należy zwrócić uwagę, że sformułowany model układów zmiennych, nie uwzględniając wszystkich istotnych zależności, stanowi znacznie uproszczony obraz niezwykle skomplikowanych powiązań występujących faktycznie w zbiorze. Jednak pomimo upraszczających założeń model ten może pełnić funkcję hipotetyczną.

Proponowany podział cech na podzbiory traktuje się jako podstawę budowy prostej hipotezy empirycznej o występowaniu 6 układów zbieżności zjawisk w przestrzeni. Układy te interpretuje się w kategoriach cech jako: układy wskaźników uprzemysłowienia (1. i 2.), układ wskaźników urbanizacji (3.), układ wskaźników leśnictwa (4.) oraz układy wskaźników rolnictwa (5. i 6.).

Zakłada się, że układy rozmieszczeń tych zjawisk tworzą ukrytą strukturę przestrzenną rzeczywistości społeczno-gospodarczej kraju.

Podstawowa para każdego układu reprezentuje punkt zbieżności cech i odgrywa zasadniczą rolę w formowaniu układu, stanowiąc jego jądro. Analiza podstawowych par pozwala więc rozpoznać cechy diagnostyczne, które w roli kryteriów klasyfikacyjnych należy uwzględnić w analizie przestrzennej.

W dalszym postępowaniu badawczym ta hipoteza o układach zbieżności zjawisk podlega sprawdzeniu. Jako metodę weryfikacji wprowadza się analizę czynnikową.

B. Analiza korelacji metodą analizy czynnikowej

B. 1. Struktura czynnikowa przestrzeni wielocechowej

W ujęciu analizy czynnikowej zagadnienie wyodrębnienia układów maksymalnie powiązanych ze sobą zmiennych sprowadza się do wyznaczenia podstawowych wymiarów przestrzeni 33-cechowej, sprzężonych z określonymi zespołami zmiennych.

W niniejszym studium dokonano rozwiązania czynnikowego tego zagadnienia przy zastosowaniu metody głównego czynnika wraz z algorytmem iteracyjnym Hotellinga. Program na analizę czynnikową napisany w języku Gier-Algol realizował metodę głównego czynnika na maszynie cyfrowej Gier⁴⁸.

Wprowadzenie schematu metodologicznego opartego na właściwej analizie czynnikowej wymaga wyznaczenia zredukowanej macierzy korelacji.

Elementami głównej przekątnej tej macierzy są zasoby zmienności wspólnej poszczególnych zmiennych $h_1^2, h_2^2, \dots, h_{33}^2$ obliczone na podstawie zmodyfikowanego przez Z. Piaseckiego wzoru (17)⁴⁹:

$$h_j^2 = \frac{1}{4} \left(\frac{r_{jk} r_{jl}}{r_{kl}} + 3 \right) \quad (j = 1, 2, \dots, 33).$$

⁴⁸ Program i obliczenia wykonał mgr Z. Piasecki z Centrum Obliczeniowego PAN w Warszawie.

⁴⁹ Dokładne objaśnienie tej metody estymacji h_j^2 znajduje się na s. 23.

Wprowadzając do macierzy korelacji R wartości liczbowe h_j^2 zamiast 1, eliminuje się z układu czynniki swoiste i zakłada się analizę czynników wspólnych (tab. 7). Zasób zmienności wspólnej przybiera wartości mniejsze lub równe jedności. Należy jednak zwrócić uwagę, że w przypadku zmiennej 5 i 7 zasób y zmienności wspólne

Tabela 7. Szacowane zasoby zmienności wspólnej h_j^2

Nr	h_j^2	Nr cechy	h_j^2
1	0,9986683	18	0,5654865
2	0,8049619	19	0,7730830
3	0,9604756	20	0,8821039
4	0,9080238	21	0,8636629
5	1,0047250	22	0,9891510
6	0,9800217	23	0,8842774
7	1,3304120	24	0,9037152
8	0,7823057	25	0,8543148
9	0,9766616	26	0,9287587
10	0,9889747	27	0,8983838
11	0,9971272	28	0,8021598
12	0,9898941	29	0,7697129
13	0,9578688	30	0,8168269
14	0,8267933	31	0,9314887
15	0,8959972	32	0,9521775
16	0,8937810	33	0,9153072
17	0,9070719		

wynoszą odpowiednio 1,004 i 1,330, a więc mają wartość nawet powyżej jedności. Te błędy wynikają z faktu, że określenie zasobu zmienności wspólnej ma charakter szacunkowy. Zastosowanie ścisłej metody w oparciu o ustalenie rzędu zredukowanej macierzy korelacji jest jednak niemożliwe, gdyż w przypadku dużego zbioru zmiennych sposób ten jest bardzo skomplikowany pod względem rachunkowym.

Ze zredukowanej macierzy korelacji wyodrębniono 4 czynniki wspólne odpowiadające wartościom własnym powyżej jedności, wyjaśniające 72,58% zasobu zmienności wspólnej. Strukturę czynnikową przedstawia tab. 8.

Tabela 8. Struktura czynnikowa

Czynnik	Wartość własna	% wykorzystania zmienności	
		wspólnej	skumulowanej
F_1	11,4750	38,33	38,33
F_2	4,7574	15,90	54,23
F_3	3,5644	11,90	66,13
F_4	1,9308	6,45	72,58

Charakter hierarchiczny tej struktury wyjaśnia sama procedura wyodrębniania czynników. Analizę rozpoczyna się od wyznaczenia czynnika F_1 , którego udział w ogólnej zmienności wspólnej jest maksymalny i wynosi 38,33%. Następnie wyznacza się pierwsze pozostałości korelacji i pozostałą, nie wyjaśnioną dotąd, zmienność wspólną. Z kolei wyodrębnia się drugi czynnik F_2 , nie skorelowany z pierwszym,

Tabela 9. Macierz czynnikowa

Cecha	F_1	F_2	F_3	F_4
1	0,7885	0,4675	-0,1579	-0,2324
2	-0,1530	-0,6501	-0,0471	-0,0197
3	0,8825	-0,2012	-0,0903	0,0505
4	-0,4681	0,6551	-0,0298	0,2506
5	0,8921	-0,2851	-0,0708	0,1822
6	0,8570	-0,3281	-0,0542	0,2165
7	-0,0678	0,7837	0,1561	0,5658
8	0,2631	-0,0485	-0,0388	0,2044
9	0,7816	0,4346	-0,1712	-0,2052
10	0,7896	0,3987	-0,1647	-0,2569
11	0,7754	-0,0468	0,0740	0,4259
12	0,7668	0,4220	-0,1682	-0,2129
13	0,7461	-0,0162	0,1239	0,3588
14	-0,3682	0,3192	0,6970	-0,2065
15	-0,6064	0,1167	-0,2220	0,3299
16	-0,0061	0,3974	-0,5177	0,2603
17	0,0591	-0,1270	0,7255	-0,2845
18	0,0995	0,0535	0,4253	0,4289
19	0,2076	0,3504	0,1486	-0,2142
20	-0,1435	-0,2506	0,7208	-0,2187
21	0,5356	-0,4860	0,1756	0,2437
22	-0,2450	-0,5018	-0,6424	-0,0782
23	-0,0512	-0,4888	-0,6563	0,3631
24	0,7858	-0,1584	0,2190	-0,1010
25	0,4858	-0,2506	0,5745	0,0643
26	0,7803	0,2945	-0,0935	-0,1921
27	0,7467	0,3883	0,0901	0,1368
28	0,1527	-0,5610	0,0758	-0,1608
29	-0,2966	-0,2825	0,1173	0,1178
30	-0,2099	0,5386	0,0303	0,0782
31	0,8617	-0,1589	-0,0091	-0,0861
32	0,8945	-0,2464	0,1217	0,0775
33	0,7743	0,1075	-0,1797	-0,0304
% wyko- rzystania zmienności wspólnej	38,33	15,90	11,90	6,45

którego udział jest ponownie maksymalny, ale już w nowo obliczonej zmienności wspólnej. Po wyznaczeniu tego czynnika F_2 procent wyjaśnionej zmienności wspólnej wzrasta do 54,23%. Proces ten jest kontynuowany aż do określenia czynnika F_4 , co jest równoznaczne z wyjaśnieniem 72,58% zasobu ogólnej zmienności wspólnej. Wyznaczone w ten sposób cztery czynniki mają tę właściwość, że są uporządkowane według ich wkładu do zmienności wspólnej.

Ładunki czynników w poszczególnych zmiennych zbioru zawiera macierz czynnikowa 33×4 (tab. 9).

Ładunki czynnikowe są współczynnikami kombinacji liniowej układu zmiennych, która w tym wariancie rozwiązania przybiera postać

$$z_j = a_{j1}F_1 + a_{j2}F_2 + a_{j3}F_3 + a_{j4}F_4 \quad (j = 1, 2, \dots, 33).$$

B. 2. Identyfikacja podstawowych czynników przestrzeni wielocechowej

Na podstawie wielkości ładunków czynnikowych, które w interpretacji statystycznej mają postać współczynników korelacji między czynnikiem a zmiennymi, można zidentyfikować układy zmiennych sprzężonych z poszczególnymi czynnikami. Czynniki są bipolarne, a więc cechy mogą być skorelowane z czynnikiem dodatnio lub ujemnie. Przyjmując arbitralnie za istotne — ładunki czynnikowe powyżej 0,5, wyróżniono grupy cech przypisanych czterem czynnikom (tab. 10 i ryc. 9). Ponieważ czynniki określają zespoły zmiennych maksymalnie ze sobą powiązanych, traktuje się je jako przyczyny korelacji stwierdzonych w danym podzbiorze zmiennych.

Czynnik F_1 , wyjaśniający 38,33% zasobu zmienności wspólnej, jest czynnikiem wiodącym. Struktura czynnikowa F_1 , ujęta w kategoriach cech skorelowanych dodatnio z tym czynnikiem, składa się z następujących elementów: 1) ludności miejskiej i ludności nierolniczej, 2) działalności przemysłowej, 3) transportu, 4) działalności usługowej. Czynnikiem ten obejmuje więc zespół cech podstawowych, stanowiących wskaźniki zjawiska urbanizacji i zjawiska uprzemysłowienia. Kryterium ludności miejskiej i kryterium odsetka zatrudnionych poza rolnictwem, o charakterze korygującym, są podstawowymi miernikami dla określenia stopnia urbanizacji w warunkach Polski. Według S. Leszczyckiego i A. Wróbla (1967) szybko postępujące w Polsce procesy urbanizacji wyrażają się nie tylko w narastaniu układów osadniczych typu miejskiego, lecz także w przenikaniu cech życia miejskiego na obszary wiejskie. Statystycznie znajduje to wyraz w zmianach struktury zawodowej ludności wiejskiej (powiększanie się udziału zajęć nierolniczych). Najważniejszą przyczyną dominacji liczbowej ludności nierolniczej na obszarach wiejskich są dojazdy do pracy do zakładów przemysłowych w miastach, które świadczą o opóźnionym procesie rozwoju miasta w stosunku do wzrostu uprzemysłowienia. Ponadto ludność wiejska zatrudniona w przemyśle nie zawsze rozstaje się z pracą na roli. Powstaje coraz liczniejsza grupa chłopo-robotników. Drugą przyczyną jest proces uprzemysłowienia powiatów wiejskich.

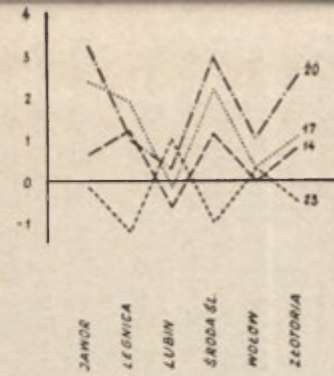
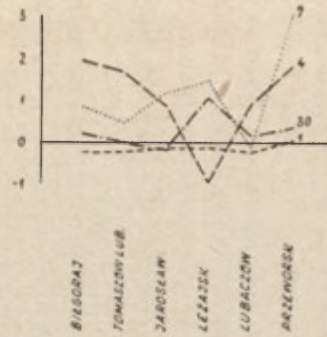
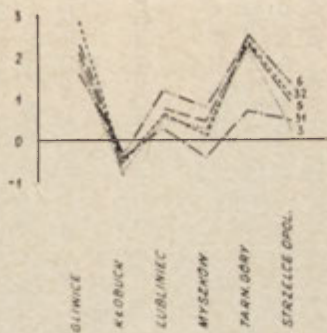
Tabela 10. Interpretacja czynników

Czynnik	Ładunki	Zmienne
F_1	0,7885	Gęstość zaludnienia
	0,8825	Ludność miejska
	0,8921	Ludność czynna zawodowo poza rolnictwem
	0,8570	Ludność utrzymująca się ze źródeł pozarolniczych
	0,7816	Nakłady inwestycyjne na 1 mieszkańca
	0,7896	Wartość środków trwałych na 1 mieszkańca
	0,7754	Zatrudnienie w przemyśle na 1000 mieszkańców
	0,7668	Zatrudnienie w przemyśle na 100 km ²
	0,7461	Produkcja globalna przemysłu na 1 mieszkańca
	0,5356	Indywidualne gospodarstwa rolne zelektryfikowane
	0,7858	Długość linii kolejowych na 100 km ²
	0,7803	Liczba pociągokilometrów na dobę na 100 km ²
	0,7467	Liczba autobusokilometrów na dobę na 100 km ²
	0,8617	Sprzedaż w handlu detalicznym na 1 mieszkańca
0,8945	Abonenci telewizyjni na 1000 mieszkańców	
0,7743	Lekarze na 10 000 mieszkańców	
-0,6064	Udział uprawy zbóż	
F_2	0,6551	Ludność czynna zawodowo
	0,7837	Ludność rolnicza na 100 ha użytków rolnych
	0,5386	Liczba ludności przypadająca na 1 punkt sprzedaży detalicznej
	-0,6501	Przyrost naturalny
	-0,5018	Powierzchnia leśna na 1 mieszkańca
	-0,5610	Liczba pociągokilometrów na dobę na 10 000 mieszkańców
F_3	0,6970	Użytki rolne
	0,7255	Udział uprawy buraków cukrowych
	0,7208	Zbiory pszenicy na 1 mieszkańca
	0,5745	Długość dróg z komunikacją autobusową na 100 km ²
	-0,6563	Lesistość
	-0,6424	Powierzchnia leśna na 1 mieszkańca
F_4	0,5658	Ludność rolnicza na 100 ha użytków rolnych

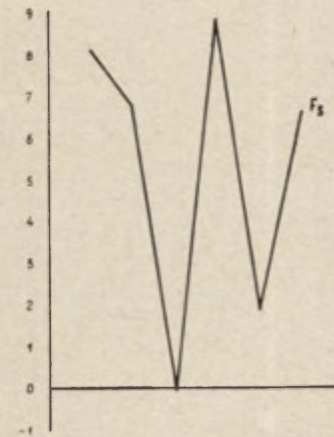
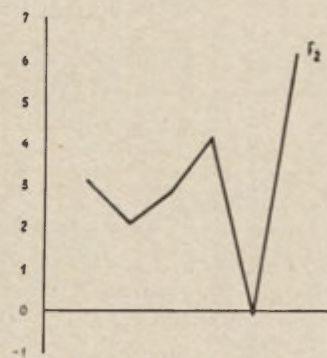
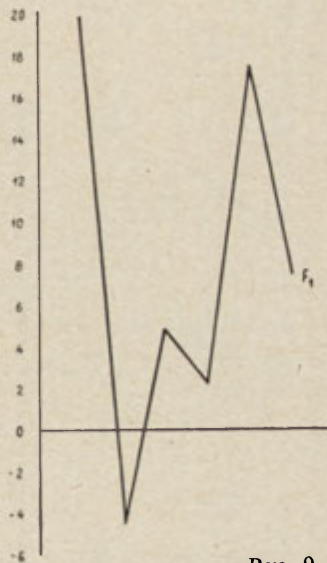
Czynnik F_1 zawiera w sobie również „czyste” wskaźniki industrializacji: zatrudnienie w przemyśle, produkcja globalna przemysłu, wartość środków trwałych, nakłady inwestycyjne, stopień zagospodarowania transportowego.

Należy zwrócić uwagę, że agregacja cech obserwowalnych urbanizacji i uprzemysłowienia w kategorię czynnika F_1 ma pełne uzasadnienie ze względu na istnienie

ZMIENNE
Variables



WARTOŚĆ
CZYNNIKA
Factor score



Ryc. 9. Zależność między zmiennymi a czynnikiem
Fig. 9. Collapsing of variables into basic pattern of factor
<http://rcin.org.pl>

różnych sprzężeń zwrotnych między tymi procesami. Czynniki F_1 i F_2 identyfikują się jako czynniki poziomu urbanizacji i przemysłowienia. Urbanizacja i industrializacja stanowią elementy jednego szerszego procesu modernizacji. W postaci czynnika F_1 dokonuje się identyfikacji syndromu modernizacji, co może mieć pewną wartość dla określenia stopnia rozwoju gospodarczego kraju. Wyprowadzona kategoria czynnika F_1 wyraża równocześnie często spotykane określenie gospodarki miejskiej.

Czynnik F_2 realizuje 15,9% zasobu zmienności wspólnej i ujawnia drugi układ powiązania zmiennych. Czynniki te wykazują zgodne związki z cechami: ludność rolnicza na 100 ha użytków rolnych, ludność czynna zawodowo, liczba ludności przypadająca na 1 punkt sprzedaży detalicznej. Jest rzeczą uderzającą, że kształtowanie się rolniczej gęstości zaludnienia pozostaje w ścisłej zależności dodatniej z liczbą ludności czynnej zawodowo. Wyjaśnia się to nieścistością współczynnika aktywności ludności, spowodowaną stosowaniem specyficznej zasady przy wydzielaniu ludności aktywnej. W statystyce polskiej do ludności pracującej włącza się wszystkie kobiety mieszkające na wsi i to powoduje zawyżenie liczby czynnych zawodowo na obszarach rolniczych. Zależność między zaludnieniem rolniczym a liczbą ludności przypadającą na 1 punkt sprzedaży detalicznej ma postać więzi przyczynowo-skutkowej o charakterze raczej pośrednim.

Można więc stwierdzić, że cecha „rolnicza gęstość zaludnienia” jest wiodącą zmienną tego układu. W konsekwencji utożsamia się z nią czynnik F_2 . Czynniki F_2 określa się jako wymiar społeczny rolnictwa.

Czynnik F_3 pod względem wartości wyjaśnianej zmienności wspólnej (11,9%) można traktować jako w przybliżeniu równorzędny czynnikowi F_2 . Interpretacja czynnika F_3 oparta jest na stwierdzeniu jego dodatniej korelacji z pewnymi symptomami intensywności i produktywności rolnictwa, jak udział użytków rolnych, udział upraw intensywnych w stosunku do ogółu upraw, gęstość pogłowia zwierząt, zbiory pszenicy na 1 mieszkańca. Czynniki F_3 identyfikują się jako czynniki poziomu intensywności gospodarki rolnej.

Czynniki F_2 i F_3 są więc czynnikami rolniczymi — reprezentują przestrzenie zjawisk demograficznych i produkcyjnych w rolnictwie.

Interpretacja czynnika F_4 jest trudna. Wskaźnikami tego czynnika są wszystkie zmienne, brak jest jednak istotnych ładunków tego czynnika. Wyjątek stanowi cecha „zaludnienie rolnicze” z wagą 0,56, która równocześnie ma wyższy ładunek czynnika F_2 . Czynniki F_4 pozostają więc w postaci metacechy teoretycznej.

Analiza porównawcza układów zmiennych wyodrębnionych metodą dendrytową i układów określonych przez poszczególne czynniki pozwala ustalić stopień zgodności tych układów. Zgodność nie jest zupełna. Jednak zasadnicze rozbieżności istnieją jedynie w zakresie liczby podstawowych układów. Układ pierwszego czynnika wchłania trzy układy dendrytowe (1, 2, 3) i stanowi agregację cech wchodzących w ich skład. Układ drugiego czynnika pokrywa się z układem dendrytowym 5.

Pozostałe układy mniej istotne nie znajdują swoich odpowiedników w obu schematach.

Metoda czynnikowa uwzględnia jednak pełne „drzewko korelacji” zmiennych i daje lepszą prezentację morfologii przestrzeni. Z tego względu w dalszym postępowaniu badawczym operuje się schematem układów czynnikowych jako bardziej precyzyjnym.

2.3. ZASTOSOWANIE CZYNNIKÓW DO ROZWIĄZANIA TAKSONOMICZNEGO ZAGADNIENIA REGIONALIZACJI

Podstawowa informacja o zbiorze, wyrażona w kategoriach czynników, zawarta jest w macierzy wartości czynnikowych:

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} & f_{14} \\ f_{21} & f_{22} & f_{23} & f_{24} \\ \dots & \dots & \dots & \dots \\ f_{33,1} & f_{33,2} & f_{33,3} & f_{33,4} \end{bmatrix}$$

Macierz tę otrzymuje się w wyniku przeprowadzenia wyjściowego zbioru 324 jednostek przestrzennych z przestrzeni euklidesowej 33-cechowej w przestrzeń 4-wymiarową czynników wspólnych. Transformacji tej dokonuje się na podstawie równania

$$F_{324 \times 4} = Z_{324 \times 33} \cdot A_{33 \times 4}$$

Macierz wartości czynnikowych stopnia 324×4 zawiera tzw. wartości czynnikowe, czyli wartości dane dla jednostki i ze względu na czynnik p ($i = 1, 2, \dots, 324$, $p = 1, 2, 3, 4$). Wartości czynnikowe wyrażone są liczbami niemianowanymi w postaci

Tabela 11. Obszar zmienności czynników

Czynnik	Wartość maksymalna	Wartość minimalna
F_1	65,86	-15,71
F_2	19,33	-10,67
F_3	10,41	-11,55
F_4	5,63	- 6,64

znormalizowanej (a więc ze średnią arytmetyczną równą zeru). Zakresy zmienności poszczególnych czynników nie są jednakowe (tab. 11), co pozostaje w ścisłym związku ze zróżnicowanym stopniem zmienności przestrzennej zjawisk ujętych przez te czynniki.

A. Klasyfikacja jednoczynnikowa⁵⁰

Klasyfikację jednoczynnikową jednostek przestrzennych przeprowadza się w oparciu o funkcję liniowo porządkującą w postaci wartości czynnikowej. Na osi liczbowej odcina się zbiór 324 jednostek przestrzennych na podstawie wartości określonego czynnika i dzieli na klasy o maksymalnej jednorodności. Klasy podstawowych jednostek przestrzennych są skupieniami typologicznie homogenicznymi lub krótko typami. Utworzono 9 klas, których granice zawiera tab. 12.

Tabela 12. Przedziały klasowe wartości czynnikowych

Klasa (typ)	Zakres wartości
I	—20 do —10
II	—10 do — 3
III	— 3 do — 1
IV	— 1 do 0
V	0 do + 1
VI	+ 1 do + 3
VII	+ 3 do +10
VIII	+10 do +20
IX	+20 do +70

Prezentację morfologii poszczególnych przestrzeni czynnikowych trzeba rozpocząć od czynnika wiodącego F_1 .

Mapa rozkładu przestrzennego wartości czynnika F_1 wykazuje, że jednostki przestrzenne zaliczane do jakiegoś typu występują najczęściej w sposób nieciągły, tj. przemieszane są z jednostkami zaliczonymi do innych typów. Mozaika typów przestrzennych jest dowodem nieciągłości przestrzeni uprzemysłowienia i urbanizacji (mapa 2).

Typy występują w 105 przypadkach jako pojedyncze izolowane jednostki. Jest to najczęstsza w zbiorze forma występowania przestrzennego typu, niezależna od kategorii typu (patrz tab. 13). Jako enklawy rozproszone po całym kraju występują zarówno jednostki zaliczane do typu o najwyższej wartości czynnikowej (wysoki stopień urbanizacji i uprzemysłowienia) — Warszawa, m. Łódź, Trójmiasto, m. Kraków, m. Wrocław, m. Poznań, Szczecin, Wałbrzych, Bielsko-Biała, jak i jednostki nieuprzemysłowione, o charakterze wiejskim: Sejny, Kolno, Proszowice, Żuromin, Włoszczowa, Limanowa, Pajęczno, Poddębice. Pozostałych 219 jednostek przestrzennych graniczy co najmniej z jedną jednostką tego samego typu i tworzy grupy dwuelementowe, rzadziej wieloelementowe. Występowanie klas dwuelementowych jest charakterystyczne przede wszystkim dla typów dodatnich (urbanizacja i uprzemysłowienie powyżej poziomu przeciętnego). Tego rodzaju tworzy przestrzenne

⁵⁰ Uwzględniono tylko 3 czynniki.

Tabela 13. Grupy jednostek przestrzennych wg typów

Liczba elementów w grupie	Typy									Ogółem	Ilość grup przestrzennych
	I	II	III	IV	V	VI	VII	VIII	IX		
	Liczba jednostek										
1	8	9	20	12	9	17	11	10	9	105	105
2	4	—	10	—	6	8	8	4	2	42	21
3	3	3	—	3	3	—	6	—	—	18	5
4	12	8	—	—	—	—	—	4	—	24	6
5	—	5	5	—	—	5	5	—	—	20	4
6	—	12	—	—	—	—	—	—	—	12	2
8	8	—	8	—	—	—	16	—	—	32	4
10	—	—	—	—	—	—	—	10	—	10	1
11	11	22	—	—	—	—	—	—	—	33	3
28	—	28	—	—	—	—	—	—	—	28	1
	46	87	43	15	18	30	46	28	11	324	152

obserwuje się w zachodniej i środkowej części Polski. Dobrą ich reprezentację stanowią jednostki: GOP-Będzin, Toruń-Bydgoszcz, Malbork-Elbląg, Kołobrzeg-Koszalin, Ostrów Wlk.-Kalisz, Nisko-Tarnobrzeg, Płock-Włocławek. Tendencję do tworzenia zbiorów wieloelementowych wykazują przede wszystkim typy ujemne, i to na obszarze wschodniej Polski. Zbiory te tworzą obszary o rozwiniętym kształcie, mało zwarte, wydłużone, z licznymi enklawami w postaci jednostek należących do różnych typów. Dotyczy to szczególnie zbioru 28-elementowego typu II, położonego na pograniczu województw: olsztyńskiego, białostockiego i warszawskiego (ryc. 10).

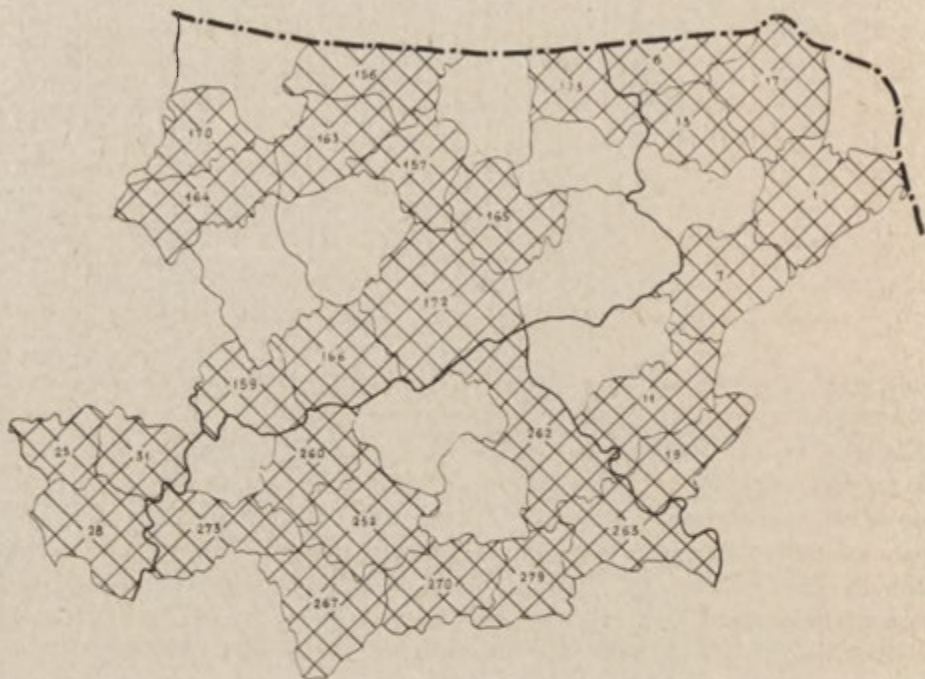
Poza tym na uwagę zasługują dobrze wykształcone dwa przestrzennie ciągle zgrupowania jednostek: 1) górnośląskie, złożone z 10 powiatów typu bardzo silnego stopnia urbanizacji i uprzemysłowienia — względnie zwarte; 2) opolskie — obejmujące 7 jednostek wysoko uprzemysłowionych i zurbanizowanych (typ VII) — o kształcie wydłużonym.

W poszukiwaniu pewnych prawidłowości przestrzennych w rozmieszczeniu typów uprzemysłowienia i urbanizacji użytecznym środkiem może okazać się zwiększenie stopnia generalizacji przez przyjęcie szerszych przedziałów klasowych podobieństwa. Redukcja 9 typów do 2 kategorii podstawowych: 1) urbanizacji i uprzemysłowienia, 2) niedorozwoju w zakresie urbanizacji i nieuprzemysłowienia, opisanych przez dodatnie i ujemne wartości czynnika F_1 , nie gwarantuje jednak ciągłości przestrzennej typów.

Podstawowym efektem tej generalizacji staje się wyraźny podział Polski w zakresie rozmieszczenia zjawisk urbanizacji i uprzemysłowienia na część zachodnią i wschodnią.

Dobrze wykształcona zachodnia zurbanizowana strefa przemysłowa nie jest jednak jednolita. Obejmuje kilka wyraźnie wyodrębnionych skupień jednostek typu

dodatniego, oddzielonych pasmami lub enklawami słabo uprzemysłowionymi. W obrębie tej strefy można wyróżnić: 1) obszar górnośląski łączący się z obszarem dolnośląskim, który z kolei znajduje swoje północne przedłużenie w obszarze zielonogórskim i gorzowskim, a na wschodzie przechodzi w obszar wielkopolski; 2) obszar nadmorski, mało spoisty; 3) obszar dolnej Wisły.



Ryc. 10. Rozwinięty kształt obszaru jednolitego

Fig. 10. Unfolding shape of uniform area

Oczywiście, że formy występujące w tej strefie zróżnicowane są wewnętrznie pod względem stopnia urbanizacji i uprzemysłowienia. Jądra ich stanowią jednostki: GOP-Będzin, Bielsko-Biała, m. Wrocław, m. Poznań, Szczecin, Trójmiasto, charakteryzujące się maksymalnym nasileniem zjawiska uprzemysłowienia.

Polska środkowa i wschodnia ma w odróżnieniu od Polski zachodniej charakter wybitnie rolniczy; obszary zagospodarowania typu przemysłowego są niewielkie i występują wyspowo. Są to obszary: łódzki, warszawski, płocko-włocławski (z Kutnem), giżycko-ełcki, tarnobrzesczo-niski oraz pojedyncze jednostki: m. Kraków, Tarnów, Rzeszów, Kielce, Radom, Lublin, Białystok, Olsztyn.

Reasumując, w świetle powyższych wyników analizy rozkładu przestrzennego czynnika F_1 , należy stwierdzić, że układy przestrzenne tworzone przez ten czynnik wykazują formowanie się ugrupowań regionalnych typu semiurbanistycznego,

a więc reprezentowanych zarówno przez tzw. obszary metropolitalne i miejskie, jak i tereny uprzemysławiane na aktywnie rozwijających się terenach rolniczych.

Mapa rozkładu czynnika F_2 wykazuje, że zróżnicowanie przestrzenne gęstości ludności rolniczej w Polsce jest znaczne (mapa 3). Wstępne porównanie mapy rozkładu czynnika F_2 z mapą rozmieszczenia uspołecznionych form gospodarowania rolniczego prowadzi do wniosku, że rozkład ludności rolniczej jest w znacznym stopniu negatywem rozmieszczenia uspołecznionych gospodarstw rolnych⁵¹. Na tej podstawie można więc wzbogacić dotychczasową treść czynnika F_2 i sformułować hipotezę, że przestrzeń czynnika F_2 reprezentuje społeczno-własnościowe stosunki w rolnictwie.

W oparciu o dwa przedziały wartości tego czynnika (ujemne i dodatnie) wyróżniamy dwa typy: typ pierwszy — normalnego nasycenia ludnością lub względnych niedoborów ludności rolniczej i dominacji gospodarstw uspołecznionych, oraz typ drugi, przeciwstawny — zagęszczenia ludności rolniczej. Analiza rozprzestrzenienia tych typów prowadzi do podziału Polski na dwie duże jednostki regionalne o wyraźnych granicach, cechujące się zasadniczym przeciwieństwem właściwości struktur społecznych rolnictwa.

Region Polski zachodniej i północnej charakteryzuje się gęstością ludności rolniczej poniżej przeciętnej krajowej, daleko posuniętym uspołecznieniem gospodarki rolnej, przewagą gospodarstw średnich i większych w sektorze prywatnym. Region ten zawierając 165 podstawowych jednostek przestrzennych obejmuje województwa: szczecińskie, koszalińskie, gdańskie, olsztyńskie, północno-zachodni kraniec białostockiego (powiaty: Gołdap, Olecko, Ełk, Augustów), zielonogórskie, poznańskie, bez powiatów wschodnich (Słupca, Konin, Koło, Turek, Pleszew, Kalisz, Ostrzeszów), bydgoskie bez części południowo-wschodniej (tj. powiatów: Radziejów, Włocławek, Aleksandrów Kujawski, Lipno, Golub-Dobrzyń, Rypin), wrocławskie, opolskie, powiaty województwa katowickiego: Lubliniec, Tarnowskie Góry, Gliwice, Zabrze. W kategoriach rozpatrywanej przestrzeni jest regionem wysoce jednolitym, zwartym, zawiera bowiem tylko dwie niewielkie enklawy typu drugiego — m. Poznań i Trójmiasto.

Region Polski środkowej i wschodniej (na obszarze województw: łódzkiego, warszawskiego, białostockiego, kieleckiego, lubelskiego, katowickiego, krakowskiego, rzeszowskiego) o nadwyżkach ludności rolniczej i wyraźnej przewadze drobnych gospodarstw indywidualnych charakteryzuje się tylko względną jednolitością. 15 jednostek, stanowiących 10⁰/₀ liczebności podzbioru, reprezentuje odmienny typ. Tworzą one kraniec południowo-wschodni regionu (gdzie gospodarstwa uspołecznione powstały na ziemiach poukraińskich), kraniec południowo-zachodni (silnie uprzemysłowiony) oraz kilka enklaw (o gospodarce rolnej typu podmiejskiego).

Badanie rozkładu geograficznego czynnika F_3 pozwala stwierdzić wysoki stopień

⁵¹ Por. mapę „Rozmieszczenie PGR i spółdzielni produkcyjnych” (opracowanie S. Hauzera), [w:] Zarys geografii ekonomicznej Polski (1967), s. 152.

niejednorodności przestrzeni poziomu intensywności rolnictwa. Układ przestrzenny typów rolnictwa, opartych na 8 przedziałach wartości czynnika F_3 , ma charakter mozaikowy (mapa 4). Poszczególne typy rolnictwa przeplatają się wzajemnie w przestrzeni tworząc obraz dość skomplikowany. Dopiero generalizacja układów typów rolnictwa przez łączenie pokrewnych typów „dodatnich” i „ujemnych” w zwarte jednostki terytorialne doprowadziła do wyróżnienia 2 stref wysoce intensywnego rolnictwa. Pierwsza z nich stanowi oś rozpatrywanej przestrzeni i biegnie z południowego zachodu w kierunku północnym. Strefa ta ma formę zniekształconą i można w niej wyróżnić kilka członów w postaci obszarów śląskiego, wielkopolskiego (Wielkopolska południowa i wschodnia) i wschodniopomorskiego z odgałęzieniem kujawskim i łódzko-warszawskim. Druga strefa południowej Polski, o przebiegu raczej równoleżnikowym, jest znacznie słabiej wykształcona i składa się z obszarów miechowsko-sandomierskiego i lubelskiego. Obszar międzystrefowy i obszary leżące po zewnętrznej stronie wyróżnionych stref, charakteryzujące się — ogólnie biorąc — niskim poziomem gospodarki rolnej (poniżej przeciętnej krajowej), zawierają tylko 3 enklawy wysoce intensywnego rolnictwa: obszary pyrzycki, przeworski i kętrzyński.

B. Klasyfikacja dwuczynnikowa

Klasyfikacja dwuczynnikowa oparta jest na analizie podstawowych jednostek przestrzennych w przestrzeni dwuwymiarowej. Liczbę tego rodzaju klasyfikacji określa liczba kombinacji z 3 czynników po dwa. Można więc dokonać klasyfikacji w układach: $F_1 F_2, F_2 F_3, F_1 F_3$.

Wstępna analiza tych układów dwuczynnikowych odbywa się w kategoriach zmiennych wyjściowych. Na wykresach dwuczynnikowych, na podstawie ładunków czynnikowych, wyznaczono położenie każdej zmiennej względnej każdej pary czynników, reprezentowanej przez ortogonalny układ odniesienia (ryc. 11).

Rozmieszczenia zmiennych nie wykazują żadnych prawidłowości. Brak jest wyraźnego trendu w układzie punktów, zmienne są rozproszone w czterech ćwiartkach. Taka postać diagramu rozproszenia prowadzi do wniosku, że dla układu dwuczynnikowego (dla każdej pary czynników) nie można ustalić w zasadzie zespołu wspólnych istotnych cech.

Z kolei w układach dwuczynnikowych rozpatruje się zbiór podstawowych jednostek przestrzennych. Klasyfikacji jednostek dokonuje się stosując podział zbioru wg kwadrantów. Typy jednostek określa się w kategoriach różnych kombinacji zgodności lub niezgodności znaków wartości czynnikowych.

Jednostki poszczególnych kwadrantów reprezentują następujące typy:

- I ćwiartka — typ zgodny dodatni (+ +)
- II ćwiartka — typ niezgodny (+ -)
- III ćwiartka — typ zgodny ujemny (- -)
- IV ćwiartka — typ niezgodny (- +).



Ryc. 11. Diagram rozproszenia zmiennych w układzie czynnikowym

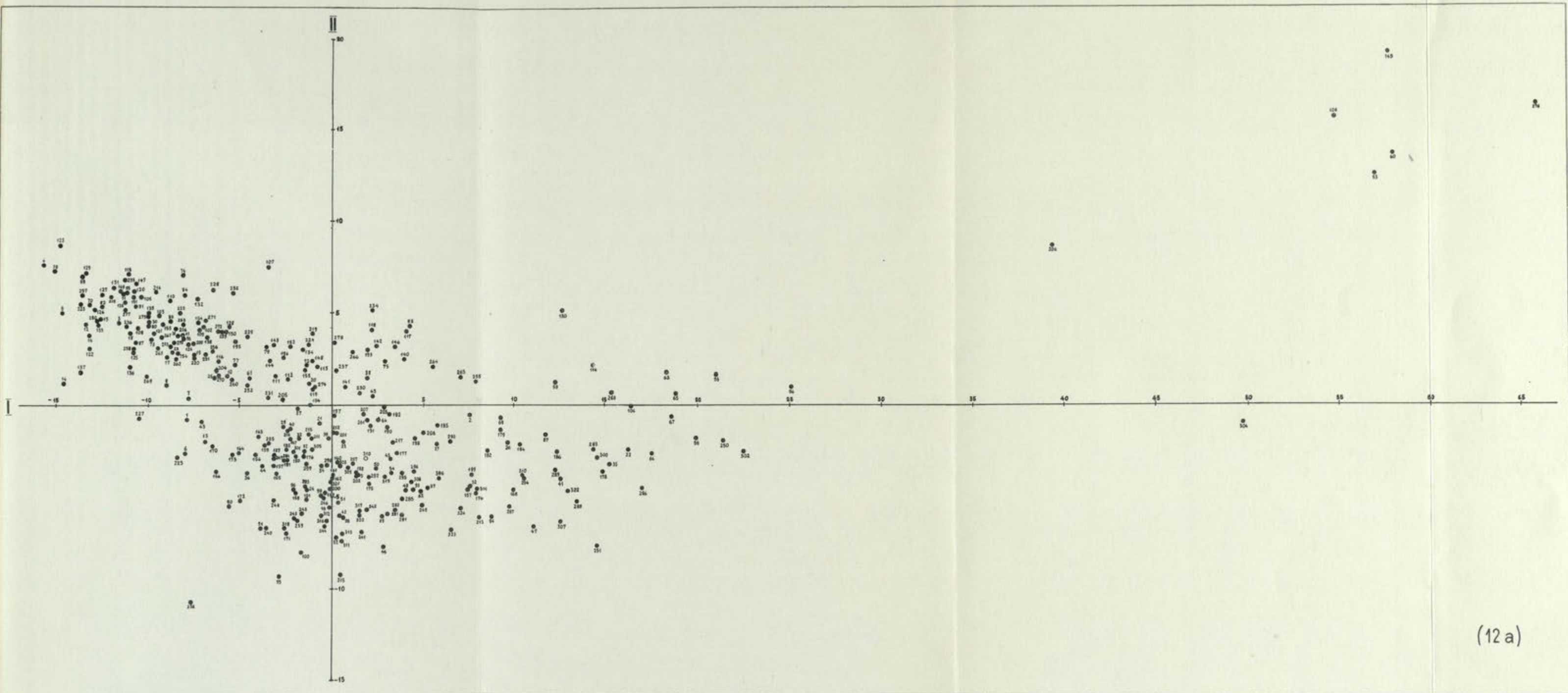
Fig. 11. A scatter diagram of variables in the factorial system

Uporządkowanie zbioru w oparciu o 4 typy oznacza wysoki stopień generalizacji. Skupiska typologiczne są tylko względnie jednorodne: odległości taksonomiczne między jednostkami określonego kwadrantu, traktowane jako miary podobieństwa, odznaczają się dużym zróżnicowaniem (szczególnie w układzie $F_1 F_2$ i $F_1 F_3$). Jednak w przypadku dużego zbioru operowanie tylko 4 głównymi typami może uprościć całe postępowanie typologiczne, które przy zwiększeniu liczby typów choćby o dwa komplikuje się niepomierne.

Rozkład jednostek poszczególnych typów w układach dwuczynnikowych przedstawiają ryc. 12 a, b, c. Najwyższy współczynnik zgodności Michenera (0,56) charakteryzuje układ $F_1 F_3$, gdyż wysoka intensywność rolnictwa towarzyszy uprzemysłowieniu⁵². Natomiast niską wartość tego współczynnika dla układu $F_1 F_2$ (0,31) tłumaczy się występowaniem kombinacji: „wysokie uprzemysłowienie — zagęszczenie ludności rolniczej”, która jest w zasadzie sprzeczna z rzeczywistością (ten typ reprezentują wyjątkowo jednostki miejskie).

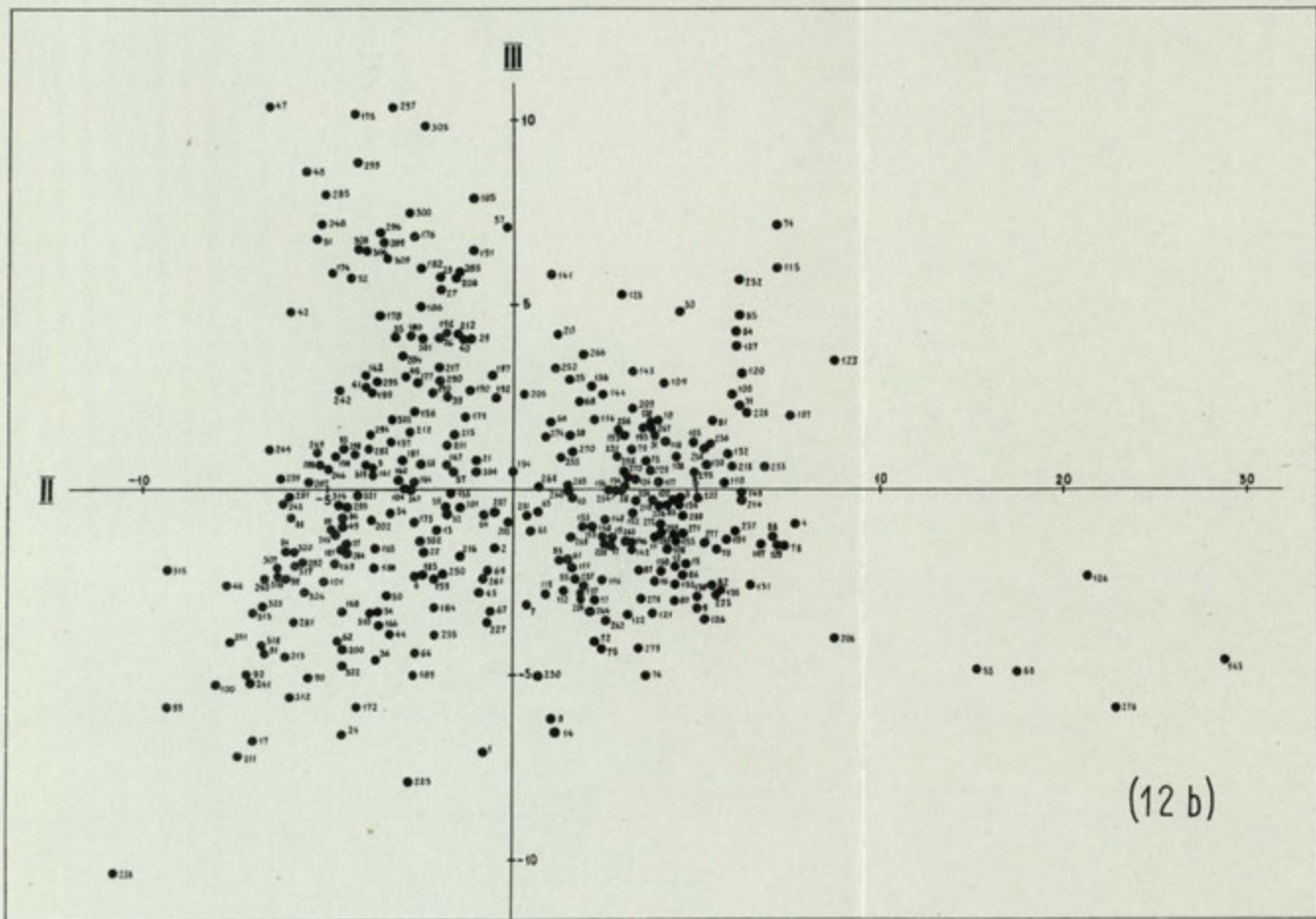
Geograficzny obraz układów dwuczynnikowych charakteryzuje występowanie typów z reguły bez ciągłości przestrzennej. Przykładowo w układzie $F_2 F_3$ dokonuje się syntezy przestrzeni poziomej intensywności gospodarki rolnej z podłożem odzwierciedlającym stosunki społeczno-własnościowe w rolnictwie i otrzymuje się wypadkową konfigurację zintegrowanej przestrzeni. Pierwotna morfologia przestrzeni czynnika F_2 o charakterze wybitnie regionalnym ulega zaburzeniu (mapa 5). W regionie Polski zachodniej i północnej nastąpiło zróżnicowanie w obszarze wewnętrznym w kategoriach typów: 1) niezgodnego (gęstość zaludnienia rolniczego poniżej przeciętnej krajowej, duża intensywność gospodarki rolnej) oraz 2) zgodnego, ujemnego. Typ pierwszy obejmuje południową i zachodnią część województwa opolskiego, prawie całe wrocławskie, południową i północno-wschodnią część poznańskiego i bydgoskiego, wschodnią część gdańskiego, zachodnią i północną część olsztyńskiego, obszary gryficki, pyrzycki, szprotawski, szamotulski — a więc reprezentowany przez obszary o wysokiej intensywności gospodarki rolnej, częściowo wywołanej sąsiedztwem dużych aglomeracji ludności i przemysłu, a częściowo związanej z występowaniem dobrych gleb, bez większych nadwyżek siły roboczej w rolnictwie. Drugi typ obejmuje województwa: szczecińskie, koszalińskie, zachodnią część gdańskiego, północną część bydgoskiego, południową część olsztyńskiego, zielonogórskie, zachodnią część poznańskiego i pogranicze poznańsko-koszalińskie, wschodnią część opolskiego, enklawy w województwie wrocławskim — a więc obszary szczególnie w północnej części Polski charakteryzujące się słabym zaludnieniem (lasy), stosunkowo małym arealem użytków rolnych i ekstensywnym charakterem gospodarki rolnej. W regionie Polski środkowej i wschodniej typ zgodny dodatni, z intensywnym rolnictwem i nadwyżkami siły roboczej, tworzy kilka wyraźnie wykształconych plam: Kujawy, zachodnia część województwa

⁵² Współczynnik zgodności C. D. Michenera $S = m/m + u$, gdzie m = liczba zgodności, u = liczba niezgodności. R. R. Sokal, P. H. Sneath (1964).

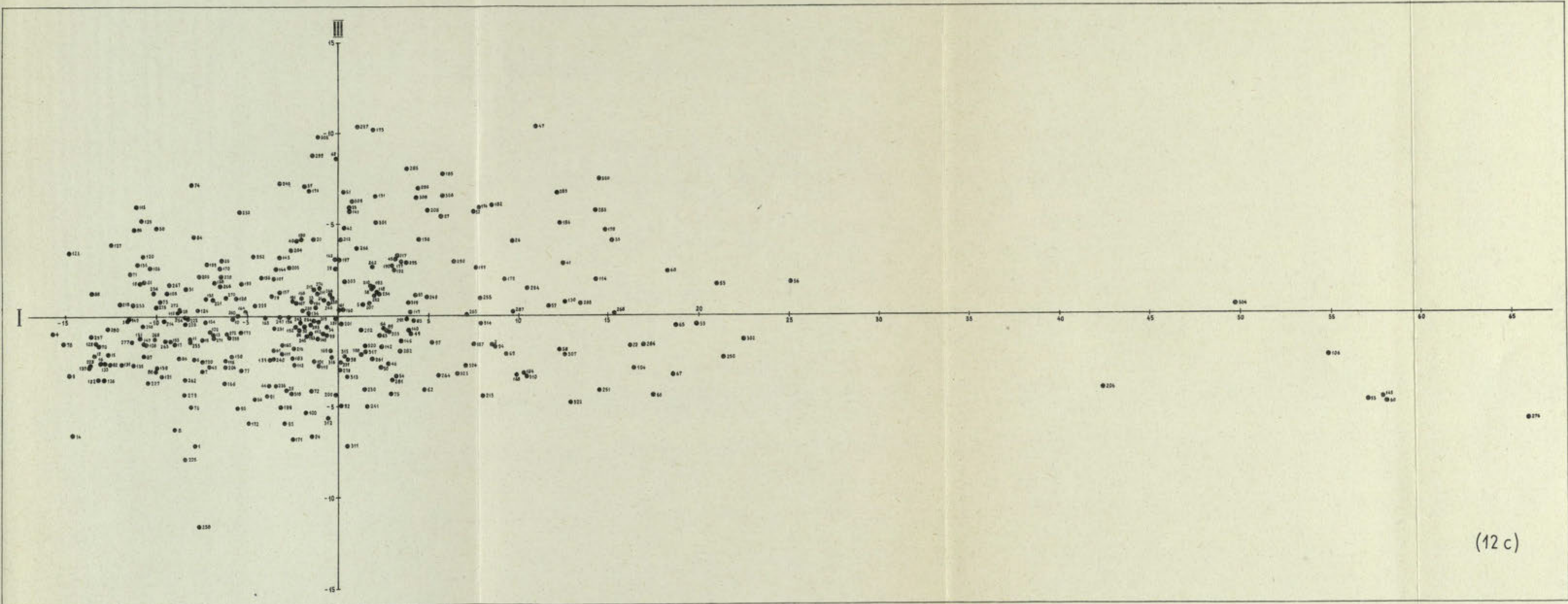


(12 a)

Ryc. 12 (a,b,c) Typologia dwuczynnikowa
 Fig. 12 (a,b,c) Typology in two factor system. Each dot stands for a powiat



(12 b)



(12 c)

warszawskiego, północna część łódzkiego; północna część krakowskiego, południowa część kieleckiego, województwo lubelskie, obszar przeworski — z bardzo dobrymi glebami na tle obszarów o niskiej intensywności gospodarki rolnej, dużym rozdrobnieniu gospodarstw, a więc poważnych utajonych nadwyżkach siły roboczej

C. Klasyfikacja wieloczynnikowa

Klasyfikacja wieloczynnikowa w języku analizy przestrzennej oznacza klasyfikację ogólnej przestrzeni społeczno-ekonomicznej.

W niniejszej pracy proponuje się dwie drogi rozwiązania klasyfikacji przestrzeni. Pierwsza z nich jest procedurą typologiczną, połączoną z analizą rozmieszczenia geograficznego typów. Druga przybiera postać klasyfikacji regionalnej, realizującej postulat jednolitości i przylegania jednostek przestrzennych, a więc właściwej regionalizacji, która prowadzi do uzyskania obrazu struktury regionalnej kraju.

C 1. Typologia przestrzenna

W klasyfikacji typologicznej, wieloczynnikowej podstawę określenia podobieństwa między jednostkami stanowią wskaźniki przyrodnicze J. Perkala (1953).

Wskaźnik wielkości J. Perkala wyrażony jest wzorem

$$(39) \quad W_i = \frac{1}{3} \sum_{j=1}^3 f_{ij} \quad (i = 1, 2, \dots, 324).$$

Jako średnia arytmetyczna wartości czynnikowych dla jednostki przestrzennej jest funkcją porządkującą, zależną od zespołu wszystkich (trzech) czynników.

Należy zwrócić uwagę, że składowe wskaźniki W_j są czynnikami o hipotetycznej względnej ortogonalności. Porównanie rozkładu wartości czynników F_1 i F_2 pozwala wnioskować o ujemnej korelacji między zjawiskiem uprzemysłowienia a gęstością zaludnienia rolniczego (współczynnik zgodności = $-0,31$)⁵³. Czynniki nie tworzą

Tabela 14. Rozkład liczebności wg wartości wskaźnika Perkala

Klasy (typy)	Przedziały wartości W_i	Liczba jednostek
I	poniżej -3	49
II	-3 do 0	157
III	0 do $+3$	81
IV	powyżej $+3$	37

⁵³ Współczynnik zgodności G. T. Fechnera $S_F = m - u/m + u$

więc zgodnego zespołu cech, stąd interpretacja wskaźnika W_i jest trudna. Można go jedynie traktować formalnie jako miarę podobieństwa.

Podział zbioru na klasy (typy) w oparciu o wartości W_i przedstawia tab. 14.

Geograficzny obraz takiego układu typologicznego daje mapa 6. Jednostki typu IV tworzą zwarty obszar górnośląski, poza tym obszarem stanowią liczne enklawy jako jednostki miejskie. Typ I występuje bez ciągłości przestrzennej tworząc zespoły, zawierające od 2 do 11 elementów na wschodzie i północnym zachodzie Polski, oraz 9 enklaw (Sucha, Kolbuszowa, Białobrzegi, Końskie, Poddębice, Goleniów, Czarnków, Grudziądz, Wyszaków). Dominują wyraźnie typy II i III, które przeplatając się wzajemnie budują szeroką strefę Polski środkowej i południowej z odgałęzieniem północno-wschodnim.

C. 2. Klasyfikacja regionalna

Punkt wyjścia do grupowania regionalnego stanowi estymacja wieloczynnikowego podobieństwa jednostek, oparta na koncepcji odległości taksonomicznej R. Bachi⁵⁴. W ujęciu geometrii analitycznej ta miara podobieństwa jednostek jest odległością między punktami w m -wymiarowej przestrzeni euklidesowej (m oznacza liczbę czynników, w rozpatrywanym przypadku $m = 4$) i wyraża się wzorem

$$(40) \quad d_{ij} = \left[\sum_{p=1}^4 (f_{ip} - f_{jp})^2 \right]^{1/2} \quad (i, j = 1, 2, \dots, 324).$$

Im mniejsza jest wartość odległości, tym większe podobieństwo jednostek.

Macierz odległości \mathbf{D} stopnia 324 zawiera odległości między każdym punktem (jednostką) a wszystkimi innymi punktami w przestrzeni czteroczynnikowej.

Ze wzoru (40) wynika, że $d_{ii} = 0$, $d_{ij} = d_{ji}$, a więc wyprowadzona macierz

$$(41) \quad \mathbf{D} = \begin{bmatrix} 0 & d_{12} & \dots & d_{1,324} \\ d_{21} & 0 & \dots & d_{2,324} \\ \dots & \dots & \dots & \dots \\ d_{324,1} & d_{324,2} & \dots & 0 \end{bmatrix}$$

jest macierzą symetryczną zawierającą 324 (324 - 1) odległości taksonomicznych.

Na podstawie macierzy odległości dokonuje się grupowania jednostek przestrzennych w klasy o maksymalnej jednolitości.

Jednak przy delimitacji regionów należy dodatkowo uwzględnić taką metacechę, jak położenie geograficzne jednostek zaliczanych do określonych klas, co pociąga za sobą duże trudności w zakresie doboru techniki grupowania.

Metody grupowania dendrytowego, szeroko stosowane w porządkowaniu przestrzennym, takie jak taksonomia wrocławska, metoda grupowania hierarchicznego

⁵⁴ R. Bachi (1963). Patrz R. Stone (1960), B. J. L. Berry (1958).

J. H. Warda, drzewo połączeń B. J. L. Berry (adaptacja metody średniego połączenia R. R. Sokala i C. D. Michenera), nie realizują postulatu przyległości przestrzennej agregowanych jednostek podobnych⁵⁵. Są mało przydatne w grupowaniu regionalnym, gdyż nie gwarantują (zabezpieczają) ciągłości przestrzennej regionów.

Próbie rozwiązania problemu sąsiedztwa w regionalizacji jednocechowej w kategoriach statystycznych, jednak nie uwieńczoną powodzeniem, podjęła M. J. Hago od (1943)⁵⁶.

L. J. King (1969, s. 194–215) proponuje modyfikację algorytmu grupowania dendrytowego B. J. L. Berry'ego przez wprowadzenie przymusu ciągłości.

Odległość taksonomiczna d_{AB} w macierzy odległości jest kodowana dodatnio, jeśli jednostki przestrzenne A i B są przyległe, i ujemnie, jeśli są nieprzyległe. Procesowi grupowania podlegają tylko elementy dodatnie.

W niniejszej pracy metoda ta nie mogła być zastosowana ze względu na znaczne wymiary macierzy odległości i związane z tym trudności natury technicznej. Elektroniczne maszyny cyfrowe w polskich ośrodkach obliczeniowych nie dysponują pamięcią operacyjną o pojemności 104 652 liczb (tyle elementów zawiera wyjściowa macierz odległości).

Autorka wprowadza jako metodę grupowania regionalnego dendrytową metodę najbliższego sąsiada (w sensie odległości taksonomicznej i odległości fizycznej), która zawiera pewne idee taksonomii wrocławskiej.

Omówienie układu operacji związanych z grupowaniem regionalnym w proponowanym ujęciu metodologicznym poprzedza próba weryfikacji założenia porządku przestrzennego w rzeczywistości gospodarczej Polski, która stanowi podstawę koncepcji regionu jednolitego. Rozwiązanie tak sformułowanego problemu jest równoznaczne z udzieleniem odpowiedzi na pytanie, czy jednostki przestrzenne położone blisko siebie są bardziej do siebie podobne niż jednostki przestrzenne oddalone od siebie⁵⁷.

W dochodzeniu tym należy oprzeć się na analizie odległości taksonomicznej określonej jednostki do sąsiadów (czyli jednostek przyległych przestrzennie).

Średnia odległość jednostki do sąsiada (przy średniej dla całego zbioru jednostek równej 8,95) wykazuje duże zróżnicowanie przestrzenne (mapa 7). Wysokie wartości średniej odległości do sąsiada są charakterystyczne dla jednostek miejskich i powiatów z nimi graniczących (maksymalna wartość 62,39 dla Warszawy). Niskie wartości spotyka się najczęściej na obszarach słabo rozwiniętych gospodarczo, odizolowanych od centrów działalności przemysłowej (minimalna wartość dla Parczewa 2,67).

⁵⁵ K. Florek, J. Łukaszewicz, J. Perkal, H. Steinhaus, S. Zubrzycki (1951), B. J. L. Berry (1961 b), R. R. Sokal, P. H. A. Sneath (1964), J. H. Ward (1963).

⁵⁶ Uwagi krytyczne na marginesie tej pracy sformułował C. L. Gregory (1949).

⁵⁷ O. D. Duncan, R. P. Cuzzort, B. Duncan (1963), s. 52, R. C. Geary (1968).

W oparciu o próbę złożoną z 16 powiatów województwa zielonogórskiego badano kształtowanie się wartości współczynnika przyległości przestrzennej (tab. 15).

Tabela 15. Współczynniki przyległości dla powiatów województwa zielonogórskiego

Nr jednostki	Nazwa powiatu	Średnia odległość do sąsiada	Średnia odległość do pozostałych powiatów Polski	Współczynnik przyległości
309	Głogów	7,11	12,49	1,76
310	Gorzów Wlk.	12,13	15,87	1,31
311	Krosno Odrz.	5,44	14,68	2,70
312	Lubsko	7,96	13,06	1,64
313	Międzyrzecz	4,57	12,51	2,74
314	Nowa Sól	6,90	14,16	2,05
315	Słubice	7,51	13,75	1,83
316	Strzelce Kraj.	5,51	11,88	2,15
317	Sulechów	5,16	11,75	2,28
318	Sulęcín	6,19	12,74	2,06
319	Szprotawa	4,96	11,63	2,34
320	Świebodzin	5,30	11,77	2,22
321	Wschowa	5,85	10,66	1,82
322	Zielona Góra	10,96	18,29	1,67
323	Żagań	5,11	14,49	2,84
324	Żary	5,16	14,22	2,75

Współczynnik przyległości przestrzennej jako stosunek średniej odległości określonego powiatu do pozostałych powiatów Polski i średniej odległości tego powiatu do sąsiadów wyraża się wzorem

$$(42) \quad c_i = \frac{(N-1)^{-1} \sum_{j=1}^{N-1} d_{ij}}{k_i^{-1} \sum_{j=1}^* d_{ij}},$$

gdzie

k_i = liczba przylegających jednostek dla i -tej jednostki,

$\sum_{j=1}^*$ = sumowanie dotyczy jedynie przyległych jednostek przestrzennych.

Współczynnik c_i jest zawsze większy od jedności, co świadczy o tym, że występuje wyższy stopień podobieństwa w przypadku przyległych jednostek.

Procedura właściwego grupowania regionalnego przy zastosowaniu metody najbliższego sąsiada jest wieloetapowa i przedstawia się następująco.

1) Punktem wyjścia jest wybór z macierzy **D** odległości między przyległymi jednostkami przestrzennymi i odwzorowanie ich na mapie (mapa 8).

2) Zakłada się, że dwie jednostki (lub podzbiory) przyległe są istotnie różne, jeśli

odległość między parą jednostek (lub podzbiorów) jest większa niż pewna wartość krytyczna K . Liczbę K szacuje się w sposób następujący. Oblicza się średnią arytmetyczną odległości $d_i = \min_j d_{ij}$ (symbol * oznacza „połączenia” z sąsiadami), czyli średnią odległość do najbliższego sąsiada dla wszystkich powiatów

$$(43) \quad \bar{d} = \frac{1}{N} \sum_{i=1}^N d_i$$

oraz odchylenie standardowe

$$(44) \quad s_d = \left[\frac{1}{N} \sum_{i=1}^N (d_i - \bar{d})^2 \right]^{1/2}$$

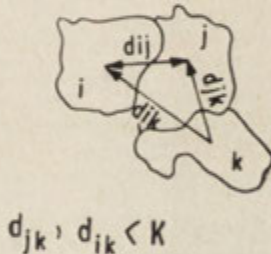
i przyjmuje się, że w pierwszym stadium grupowania

$$(45) \quad K = \bar{d} = 4,71,$$

a w drugim stadium

$$(46) \quad K = \bar{d} + s_d = 4,71 + 7,20 = 11,91.$$

3) W zbiorze odległości „sąsiadów” określa się minimalną wartość $d_{ij} < 4,71$, łączy się jednostki i i j jako najbardziej podobne w grupę 2-elementową, która przyjmuje postać jądra regionu. Następnie spośród jednostek przyległych do grupy wybiera się, na podstawie odległości taksonomicznej, najbliższego sąsiada k , pod



Ryc. 13. Graf grupowania w metodzie najbliższego sąsiada

Fig. 13. A graph of grouping by the nearest neighbour method

warunkiem, że d_{jk} i $d_{ik} < 4,71$ jednostkę k włącza się do grupy (ryc. 13). W przypadku niespełnienia się tego warunku grupowanie przerywa się i określa się na podstawie zbioru odległości jądro nowej grupy. Procedurę kontynuuje się. Ta zasada grupowania sprawia, że w grupie jednostek odległości między przyległymi i nieprzyległymi przestrzennie elementami nie przekraczają wartości krytycznej.

W wyniku takiego postępowania dokonuje się delimitacji maksymalnie jednolitych obszarów, które są układami 7-, 6-, 5-, 4-, 3-, a najczęściej 1-elementowymi (mapa 9). Zwiększając stopień generalizacji przez przyjęcie wartości krytycznej $d = 11,91$ włącza się przy zastosowaniu metody najbliższego sąsiada obszary 1-elementowe do układów wieloelementowych (mapa 10).

W rezultacie otrzymuje się siatkę 95 mikroregionów Polski. Wśród nich jako

mikroregiony 1-elementowe występują miasta o statusie województw, GOP, Trójmiasto, Białystok, Lublin, Ustrzyki Dolne, Pszczyna (tab. 16). 84% układu stanowią mikroregiony 1–5-elementowe. Dominują jednostki regionalne obejmujące 2 powiaty.

Tabela 16. Mikroregiony według liczby podstawowych jednostek przestrzennych

Liczba podstawowych jednostek przestrzennych	Liczba mikroregionów
1	11
2	30
3	18
4	11
5	10
6	8
7	2
8	3
9	1
10	1
Ogółem	95

Mikroregiony są bardzo zróżnicowane pod względem wielkości powierzchni (220–9779 km²) i liczby ludności (17,3–1425,5 tys. mieszkańców, tab. 17).

W trzecim stadium grupowania dokonuje się łączenia mikroregionów w jednostki wyższego rzędu. Kolejność operacji jest następująca. Jako czynność wstępną traktuje się wyprowadzenie charakterystyk czynnikowych dla mikroregionów według 4 czynników w postaci średnich arytmetycznych z wartości czynnikowych przypisanych jednostkom powiatowym wchodzącym w skład mikroregionów. Na ich podstawie oblicza się odległości taksonomiczne między mikroregionami w przestrzeni 4-czynnikowej.

Średnia odległość

$$(47) \quad \bar{e} = \frac{\sum_{i=1}^{95} e_i}{95}$$

dla par mikroregionów będących najbliższymi sąsiadami wynosi 9,12 i stanowi wartość krytyczną podobieństwa na tym szczeblu grupowania.

Rozwijając grupowanie mikroregionów przy zastosowaniu grafowej metody najbliższego sąsiada (już poprzednio omówionej) dokonuje się delimitacji 24 regionów ogólnoeconomicznych (mapa 11).

Tabela 17. Mikroregiony ogólnoeconomiczne Polski

Mikroregiony	Powierzchnia	Ludność	Podstawowe jednostki przestrzenne	
			nr	nazwa
1	1498	376,7	250	Szczecin
			251	Wolin
2	3483	160,4	93	Kołobrzeg
			242	Gryfice
			244	Kamień Pomorski
3	9343	446,5	241	Goleniów
			249	Stargard Szczeciński
			247	Nowogard
			245	Łobez
			99	Świdwin
			89	Białogard
			94	Koszalin
			98	Szczecinek
4	5591	217,2	101	Złotów
			248	Pyrzyce
			246	Myślibórz
			243	Gryfino
			239	Chojna
			240	Choszczno
5	3504	104,6	100	Wałcz
			92	Drawsko Pomorskie
6	4098	136,8	96	Sławno
			95	Miastko
			91	Człuchów
7	4661	324,5	97	Słupsk
			46	Lębork
			49	Puck
			54	Wejherowo
8	246	532,5	53	Trójmiasto
9	6370	278,3	90	Bytów
			43	Kartuzy
			44	Kościerzyna
			24	Chojnice
			36	Tuchola
			50	Starogard Gdański
10	2680	168,0	34	Świecie
			42	Gdańsk
11	1396	86,8	48	Nowy Dwór Gdański
			52	Tczew
12	2070	246,9	47	Malbork
			41	Elbląg
			51	Sztum
13	5687	444,0	45	Kwidzyn
			26	Grudziądz

Mikroregiony	Powierz- chnia	Ludność	Podstawowe jednostki przestrzenne				
			nr	nazwa			
14	4854	143,6	161	Ława			
			169	Ostróda			
			168	Olsztyn			
			166	Nidzica			
			172	Szczytno			
15	2606	158,0	171	Pisz			
			21	Brodnica			
			167	Nowe Miasto Lubawskie			
16	1238	97,8	159	Działdowo			
			23	Chełmno			
			37	Wąbrzeźno			
17	5667	371,0	25	Golub-Dobrzyń			
			31	Rypin			
			28	Lipno			
			280	Żuromin			
			273	Sierpc			
			267	Płońsk			
			18	3229	216,5	260	Mława
			19	3140	137,3	252	Ciechanów
						270	Pułtusk
						170	Pasłęk
164	Morąg						
20	4480	206,4	163	Lidzbark Warmiński			
			158	Braniewo			
			156	Bartoszyce			
			162	Kętrzyn			
			157	Biskupiec			
21	3928	158,2	173	Węgorzewo			
			6	Gołdap			
			13	Olecko			
			17	Suwałki			
22	3037	142,6	165	Mrągowo			
			160	Giżycko			
			5	Ełk			
23	2493	117,7	269	Przasnysz			
			258	Maków Mazowiecki			
24	3771	195,0	9	Kolno			
			262	Ostrołęka			
			279	Wyszków			
25	6918	324,5	14	Sejny			
			1	Augustów			
			7	Grajewo			
			11	Łomża			
			19	Zambrów			
			263	Ostrów Mazowiecka			

Mikroregiony	Powierzchnia	Ludność	Podstawowe jednostki przestrzenne				
			nr	nazwa			
26	3494	147,1	4	Dąbrowa Białostocka			
			12	Mońki			
			16	Sokołka			
27	2308	234,3	2	Białystok			
28	7326	340,5	18	Wysokie Mazowieckie			
			10	Łapy			
			3	Bielsk Podlaski			
			8	Hajnówka			
			15	Siemiatycze			
			257	Łosice			
			277	Węgrów			
			275	Sokołów Podlaski			
			29	2353	140,9	259	Mińsk Mazowiecki
						272	Siedlce
253	Garwolin						
256	Grójec						
271	Ryki						
77	Kozienice						
134	Puławy						
121	Biała Podlaska						
30	7709	633,2				135	Radzyń Podlaski
						131	Łuków
			129	Lubartów			
			133	Parczew			
			137	Włodawa			
			31	8861	489,8	75	Kielce
						72	Ilża
83	Radom						
32	4151	679,5				78	Lipsko
			88	Zwolenie			
			79	Opatów			
33	1397	91,4	84	Sandomierz			
			132	Opole Lubelskie			
			128	Kraśnik			
34	4125	418,6	130	Lublin			
			310	Gorzów Wielkopolski			
35	1201	318,1	316	Strzelce Krajeńskie			
			315	Słubice			
			318	Sulęcín			
			313	Międzyrzecz			
			320	Świebodzin			
			317	Sulechów			
			200	Międzychód			
			189	Czarnków			
			36	9157	443,5	311	Krosno Odrzańskie
						37	2085

Mikroregiony	Powierzchnia	Ludność	Podstawowe jednostki przestrzenne	
			nr	nazwa
38	2718	171,4	312	Lubsko
			213	Trzcianka
			188	Chodzież
39	5132	376,0	202	Oborniki
			210	Szamotuły
			201	Nowy Tomyśl
			207	Poznań
			216	Wolsztyn
			321	Wschowa
40	220	438,2	206	m. Poznań
41	3280	280,6	190	Gniezno
			212	Środa
			217	Września
			192	Jarocin
			197	Kościąn
42	1804	135,8	211	Śrem
			199	Leszno
43	2795	256,9	191	Gostyń
			208	Rawicz
			198	Krotoszyn
			203	Ostrów Wielkopolski
			193	Kalisz
44	2532	302,8	32	Sępólno Krajeńskie
			39	Wyrzysk
			33	Szubin
			215	Wągrowiec
			40	Żnin
45	3696	219,4	29	Mogilno
			209	Słupca
46	1753	108,2	205	Pleszew
			196	Konin
			30	Radziejów
			195	Koło
			143	Łęczycza
			141	Kutno
			254	Gostynin
			144	Łowicz
			153	Skierniewice
			214	Turek
47	9779	850,8	149	Poddębice
			22	Bydgoszcz
48	2016	139,8	35	Toruń
			27	Inowrocław
			20	Aleksandrów Kujawski
49	2669	511,1	38	Włocławek
50	5341	570,1		

Mikroregiony	Powierzchnia	Ludność	Podstawowe jednostki przestrzenne	
			nr	nazwa
			266	Płock
			274	Sochaczew
51	1388	294,6	255	Grodzisk Mazowiecki
			268	Pruszków
52	446	1252,6	276	m. Warszawa
53	1955	256,9	261	Nowy Dwór Mazowiecki
			278	Wołomin
54	1226	218,5	265	Piaseczno
			264	Otwock
55	3791	350,2	324	Żary
			322	Zielona Góra
			314	Nowa Sol
			323	Żagań
			307	Zgorzelec
56	2620	153,9	319	Szprotawa
			291	Lubin
			281	Bolesławiec
57	1276	123,1	290	Lubań
			292	Lwówek Śląski
58	2038	234,3	308	Złotoryja
			289	Legnica
			285	Jawor
59	5033	318,8	309	Głogów
			284	Góra
			303	Wołów
			293	Milicz
			301	Trzebnica
			295	Oleśnica
60	1568	123,0	299	Środa Śląska
			305	Wrocław
61	1163	111,1	296	Oława
			174	Brzeg
62	225	474,2	304	m. Wrocław
63	1784	121,0	297	Strzelin
			176	Grodzków
			181	Niemodlin
64	3079	215,9	298	Syców
			180	Namysłów
			177	Kluczbork
			194	Kępno
			204	Ostrzeszów
65	214	744,1	145	m. Łódź
66	5923	739,1	152	Sieradz
			142	Łask
			146	Łódź

Mikroregiony	Powierz- chnia	Ludność	Podstawowe jednostki przestrzenne	
			nr	nazwa
67	1677	133,0	140 *	Brzeziny
			148	Piotrków Trybunalski
			155	Wieruszów
68	1651	115,8	154	Wieluń
			147	Pajęczno
69	1710	123,1	139	Bełchatów
			183	Olesno
70	4065	270,6	61	Kłobuck
			151	Rawa Mazowiecka
			80	Opoczno
71	3490	290,8	82	Przysucha
			70	Białobrzegi
			150	Radomsko
			76	Końskie
72	850	131,4	86	Szydłowiec
			294	Nowa Ruda
73	3523	347,0	288	Kłodzko
			175	Głubczyce
			185	Prudnik
			182	Nysa
			306	Ząbkowice
			282	Bystrzyca Kłodzka
74	1649	446,6	283	Dzierżoniów
			302	Wałbrzych
			300	Świdnica
			286	Jelenia Góra
75	1175	194,5	287	Kamienna Góra
			184	Opole
76	5801	864,0	179	Krapkowice
			187	Strzelce Opolskie
			62	Lubliniec
			58	Częstochowa
			63	Myszków
			69	Zawiercie
			178	Koźle
			186	Racibórz
77	1163	193,2	66	Tarnowskie Góry
			59	Gliwice
78	1307	432,0	60	GOP
			68	Wodzisław
79	472	1309,2	65	Rybnik
			67	Tychy
80	4204	1425,5	55	Będzin
			104	Chrzanów
			114	Oświęcim

Mikroregiony	Powierzchnia	Ludność	Podstawowe jednostki przestrzenne	
			nr	nazwa
			56	Bielsko-Biała
			57	Cieszyn
81	502	66,7	64	Pszczyna
82	4521	611,5	113	Olkusz
			107	Kraków
			102	Bochnia
			110	Myślenice
			108	Limanowa
83	230	520,1	106	m. Kraków
84	7785	655,3	87	Włoszczowa
			73	Jędrzejów
			109	Miechów
			81	Pińczów
			71	Busko-Zdrój
			85	Staszów
			105	Dąbrowa Tarnowska
			103	Brzesko
85	970	109,6	115	Proszowice
			74	Kazimierza Wielka
86	5951	702,9	119	Żywiec
			118	Wadowice
			116	Sucha
			112	Nowy Targ
			111	Nowy Sącz
87	4716	463,6	220	Gorlice
			222	Jasło
			236	Strzyżów
			218	Brzozów
			233	Ropczyce
			223	Kolbuszowa
88	1884	181,2	237	Tarnobrzeg
			230	Nisko
89	2428	353,8	229	Mielec
			219	Dębica
			117	Tarnów
90	2557	206,1	120	Bełżyce
			123	Bychawa
			127	Krasnystaw
91	6762	483,9	138	Zamość
			125	Hrubieszów
			136	Tomaszów Lubelski
			124	Chelm
92	3779	189,5	126	Janów Lubelski
			122	Biłgoraj
			227	Lubaczów

Mikroregiony	Powierzchnia	Ludność	Podstawowe jednostki przestrzenne	
			nr	nazwa
93	3769	485,8	226	Leżajsk
			221	Jarosław
			232	Przeworsk
			228	Łańcut
			234	Rzeszów
94	4323	321,5	224	Krosno
			235	Sanok
			231	Przemyśl
			225	Lesko
95	1187	17,3	238	Ustrzyki Dolne

3. WSTĘPNA HIPOTEZA O STRUKTURZE REGIONALNEJ POLSKI

W przypadku układu regionalnego opartego na kryteriach strefowości gospodarczej pojęcie struktury staje się korelatem pojęcia całości, zbudowanej w pewien sposób z określonych elementów. Badanie struktury regionalnej układu ogranicza się więc do analizy zróżnicowania i ustalenia charakterystycznych właściwości poszczególnych części składowych układu oraz analizy ich rozmieszczenia.

Uzyskany podział Polski na 24 jednolite regiony ogólnoeconomiczne składa się z następujących jednostek⁵⁸:

- I — region szczecińsko-koszaliński,
- II — region słupsko-chojnicko-gdański,
- III — region grudziądzko-elbląsko-olsztyński,
- IV — region lidzbarsko-szczytnicko-ełcki,
- V — region ostrołęcko-białostocko-suwalski,
- VI — region gorzowski,
- VII — region poznański,
- VIII — region bydgosko-konińsko-płocki,
- IX — region brodnicko-ciechanowski,
- X — region bielsko-białopodlaski,
- XI — region zielonogórsko-bolesławiecki,
- XII — region głogowsko-kluczborski,
- XIII — region łódzki,
- XIV — region wieluńsko-konecki,
- XV — region warszawski,

⁵⁸ Regionom nadano nazwy od największego lub centralnie położonego miasta w danych regionach albo kilku miast.

- XVI — region kielecko-lubelsko-siedlecki,
- XVII — region chełmsko-zamojski,
- XVIII — region wałbrzysko-klódzki,
- XIX — region wrocławsko-nyski,
- XX — region opolsko-częstochowski,
- XXI — region krakowski,
- XXII — region rzeszowski,
- XXIII — region katowicki,
- XXIV — region nowosądecko-tarnowsko-tarnobrzeski.

Regiony mają bardzo różną wielkość powierzchni (tab. 18). Do największych regionów należą szczecińsko-koszaliński (26 137 km²), bydgosko-konińsko-płocki (25 254 km²) i ostrołęcko-białostocko-suwalski (22 922 km²). Najmniejsze regiony występują przede wszystkim na południu Polski (najmniejszą powierzchnię zajmuje region wałbrzysko-klódzki — 3674 km²).

Tabela 18. Regiony ogólnoeconomiczne Polski

Region	Powierzchnia w km ²	Ludność w tys. osób
I	26137	1476,8
II	15375	1272,1
III	13071	1043,5
IV	15511	629,9
V	22922	1176,8
VI	11242	527,7
VII	13231	1487,5
VIII	25254	2399,6
IX	11502	745,5
X	18540	971,2
XI	9725	861,5
XII	9822	657,8
XIII	8669	1786,0
XIV	10883	810,2
XV	5015	2022,6
XVI	18583	2140,8
XVII	9319	690,0
XVIII	3674	772,5
XIX	7100	1065,2
XX	8127	1168,3
XXI	13506	1896,5
XXII	17774	1477,7
XXIII	6485	3233,4
XXIV	10263	1237,9
Polska	311730	31551,0

Pod względem zaludnienia dolną granicą wielkości regionów jest liczba 527,7 tys. osób dla regionu gorzowskiego, największym zaludnieniem odznaczają się regiony: katowicki — 3233,4 tys. osób, bydgosko-konińsko-płocki — 2399,6 tys., warszawski — 2022,6 tys., kielecko-lubelsko-siedlecki — 2140,8 tys.

Układ regionalny charakteryzuje się również dużą różnorodnością kształtów: od zwartych (np. region poznański) do znacznie zniekształconych (np. region nowosądecko-tarnowsko-tarnobrzeski), często z enklawami.

Regiony są jednostkami o bardzo różnych stopniach jednolitości. Żaden region tego układu nie jest jednolity w sensie absolutnym. Stopień jednolitości regionu w obszarze wewnętrznym determinuje wartość krytyczna podobieństwa. W dodatku względną jednolitość regionu zaburza włączenie do poszczególnych obszarów jednolitych, na podstawie arbitralnych założeń, specyficznych mikroregionów, występujących w ich obszarze wewnętrznym lub partiach brzeżnych. Te mikroregiony są diametralnie różne od otaczającego obszaru. Reprezentują one obszary metropolitalne, odznaczające się znacznie wyższym poziomem rozwoju gospodarczego w porównaniu z sąsiadami (z reguły są wtedy mikroregionami jednoelementowymi) lub obszary zacofane gospodarczo (np. mikroregion zwoleński).

Należy jeszcze stwierdzić, że występowanie ognisk w obszarze jednorodnym nadaje wyprowadzonemu układowi regionalnemu specyficzny charakter o pewnych symptomach węzłowości.

Regiony charakteryzują się silnym zróżnicowaniem profilów ekonomicznych. W układzie regionów znajdują odzwierciedlenie podstawowe elementy strukturalne rzeczywistości społeczno-gospodarczej Polski. Dążąc do uchwycenia podobieństwa poszczególnych regionów w zakresie charakteru gospodarczego wydzielono następujące typy regionów:

- 1) przemysłowe: katowicki, warszawski;
- 2) przemysłowo-rolnicze: łódzki, wałbrzysko-łódzki, wrocławsko-nyski, opolsko-częstochowski, poznański, zielonogórsko-bolesławiecki, bydgosko-konińsko-płocki;
- 3) rolniczo-przemysłowe: szczecińsko-koszaliński, słupsko-chojnicko-gdański, grudziądzko-elbląsko-olsztyński, gorzowski, kielecko-lubelsko-siedlecki, nowosądecko-tarnowsko-tarnobrzeski;
- 4) rolnicze: lidzbarsko-szczytnicko-ęłcki, ostrołęcko-białostocko-suwalski, brodnicko-ciechanowski, bielsko-białopodlaski, głogowsko-kluczborski, wieluńsko-konecki, chełmsko-zamojski, krakowski, rzeszowski.

Bliższe określenie stopnia zróżnicowania regionalnego wchodzi już w zakres badania struktury wewnętrznej regionu.

Powyższa próba określenia złożoności struktury regionalnej oparta na kryterium strefowości może mieć jedynie charakter wstępnej hipotezy. Jako sposób weryfikacji tego układu można zaproponować badania oparte na kryteriach węzłowości, które pozwolą ostatecznie ustalić relacje układu strefowego i układu węzłowego. Synteza regionów jednolitych i węzłowych zagwarantuje jeszcze bliższe rzeczywistości odwzorowanie przestrzennych układów gospodarczych.

4. ZRÓŻNICOWANIE STREFOWE A UKŁAD WOJEWÓDZKI

Porównanie układu strefowego makroregionalnego z układem wojewódzkim prowadzi do stwierdzenia wyraźnej niezgodności tych układów nie tylko w przebiegu granic, lecz również w ilości regionów. Charakter heterogeniczny województw potwierdziły dodatkowe badania nad jednolitością województw, przeprowadzone w oparciu o wartość odległości taksonomicznej (czynnikowej).

Wprowadzono trzy podstawowe miary:

I — miarę zwartości województwa wyrażoną wzorem

$$(48) \quad Z_g = \frac{\sum d_{jk}}{p},$$

gdzie

j, k = przyległe przestrzennie powiaty województwa g ,

d_{jk} = odległość taksonomiczna między powiatami przyległymi przestrzennie,

p = liczba par powiatów przyległych przestrzennie;

II — miarę odległości międzywojewódzkich wyrażoną wzorem

$$(49) \quad M_{gh} = \frac{\sum d_{uv}}{q},$$

gdzie

u, v = powiaty graniczne dwóch sąsiadujących województw g i h ,

q = liczba par powiatów granicznych sąsiednich województw g i h ,

d_{uv} = odległość taksonomiczna między powiatami granicznymi;

III — statystykę H o postaci

$$H = \frac{\text{wariancja odległości międzywojewódzkich}}{\text{wariancja odległości wewnątrzwojewódzkich}}$$

wyrażoną wzorem

$$(50) \quad H_g = \frac{p}{q} \cdot \frac{\sum (d_{uv} - \bar{M}_{gh})^2}{\sum (d_{jk} - Z_g)^2}$$

Wartości tych parametrów statystycznych dla poszczególnych województw zawierają tab. 19 i 20.

Najbardziej „zwarte” w kategoriach odległości taksonomicznych są województwa: koszalińskie (5,58), olsztyńskie (5,91), zielonogórskie (7,02), poznańskie (7,05), a najbardziej zróżnicowane: gdańskie (16,33), katowickie (16,08), łódzkie (11,37), wrocławskie (11,19), warszawskie (11,01).

W strefie granicznej największe odległości występują w przypadku województwa

Tabela 19. Macierz odległości międzywojewódzkich i miar zwartości województw

	białostockie	bydgoskie	gdańskie	katowickie	kieleckie	koszalińskie	krakowskie	lubelskie	łódzkie	olsztyńskie	opolskie	poznańskie	rzeszowskie	szczecińskie	warszawskie	wrocławskie	zielonogórskie	Średnia odległość w strefie granicznej
białostockie	8,50																	5,58
bydgoskie		9,97																6,27
gdańskie		6,19	16,33															7,54
katowickie				16,08														12,79
kieleckie				19,12	9,11													7,82
koszalińskie		6,40	5,42			5,58												5,82
krakowskie				15,46	6,01		9,45											10,26
lubelskie	3,34				7,13			8,21										7,04
łódzkie		4,95		13,68	6,10				11,37									7,68
olsztyńskie	8,14	7,38	10,51							5,91								8,57
opolskie				7,86					10,95		9,25							7,38
poznańskie		5,93				7,62			5,73		5,37	7,05						6,17
rzeszowskie					11,26		8,67	9,41					8,22					9,57
szczecińskie						4,77						10,43		8,79				6,36
warszawskie	3,29	6,33			5,04			5,28	7,55	7,99					11,01			5,83
wrocławskie											t5,45	6,45				11,19		5,68
zielonogórskie												6,00		8,13		5,21	7,02	6,06

Tabela 20. Analiza wariancji odległości taksonomicznych dla układu wojewódzkiego

Województwo	Wariancja odległości wewnątrz-wojewódzkich	Wariancja odległości między-wojewódzkich	H
Białostockie	31,3771	21,5396	0,6825
Bydgoskie	32,3060	11,8967	0,3682
Gdańskie	334,4934	13,7964	0,0412
Katowickie	177,9978	84,4612	0,4745
Kieleckie	26,4593	24,3540	0,9204
Koszalińskie	9,1255	7,0016	0,7675
Krakowskie	92,0631	80,4337	0,8737
Lubelskie	41,1686	12,0648	0,2930
Łódzkie	213,2913	25,2719	0,1185
Olsztyńskie	16,6972	20,5961	1,2335
Opolskie	11,5455	12,3218	1,1545
Poznańskie	49,1986	7,8251	0,1590
Rzeszowskie	16,3916	12,6603	0,7724
Szczecińskie	32,1203	21,5824	0,6716
Warszawskie	235,5992	9,7308	0,0413
Wrocławskie	130,3181	5,0285	0,0386
Zielonogórskie	14,2221	9,1273	0,6418

katowickiego (12,79), szczególnie na granicy z województwami kieleckim i krakowskim, oraz województwa krakowskiego (10,26), najmniejsze zaś w przypadku województw: białostockiego, wrocławskiego, koszalińskiego i warszawskiego.

Jednak tylko w przypadkach województw: koszalińskiego, krakowskiego, olsztyńskiego, rzeszowskiego średnia odległość taksonomiczna w obrębie województwa jest mniejsza niż w strefie granicznej, a dla zielonogórskiego, poznańskiego, lubelskiego wartości te zbliżone są do siebie.

Statystyka H tylko w przypadku województw olsztyńskiego i opolskiego kształtuje się powyżej 1, co oznacza, że tylko 2 województwa Polski są słabiej zróżnicowane w obszarze wewnętrznym niż zewnętrznym.

W świetle powyższych wyników staje się również wątpliwa teza, że zagadnienie strefowości powinno występować jako zagadnienie podziałów pośrednich pomiędzy województwem a powiatem. Punktem wyjścia krytyki koncepcji wyczerpującego podziału województw na podobszary jednolite stał się wyprowadzony układ mikroregionalny. Granice mikroregionów ogólnoeconomicznych z reguły naruszają granice województw.

Przeprowadzone badania udowodniły, że kryteriów jednorodności układu wojewódzkiego należy poszukiwać w powiązaniach społeczno-economicznych, gdyż o podziale na regiony — narzędzia działania w większym stopniu rozstrzygają układy węzłowe niż układy strefowe.

UWAGI KOŃCOWE

W konkluzji pracy należy zwrócić uwagę, że nosi ona symptomy swoistego eksperymentowania, jest pewną propozycją w dziedzinie rozwiązania taksonomicznych problemów regionalizacji ekonomiczno-geograficznej.

Wyprowadzonego podziału regionalnego Polski zarówno w ujęciu mikro-, jak i makroskopowym nie można więc absolutyzować.

Takie podejście do rezultatów pracy znajduje uzasadnienie we względności każdego poznania naukowego i konieczności wielokrotnego stosowania ujęć regionalnych. Zagadnienie to podnosi K. Dziewoński (1969, s. 260) stwierdzając: „... w diagnozach regionalnych trzeba coraz konsekwentniej stosować wielorakie metody analityczne. Każda metoda analityczna zawiera w sobie elementy ograniczające, każda stanowi generalizację, każda opiera się na pewnych założeniach upraszczających rzeczywistość i skutkiem tego każda zawiera *implicite* określoną ilość błędów i odchyień od rzeczywistości”.

Pozostaje więc dokonanie autooceny pracy i przedstawienie pewnych uwag co do dyskusyjnych problemów wiążących się z interpretacją założeń badawczych.

Z punktu widzenia poznawczego w niniejszym studium nie doprowadza się do końca analizy nad podstawami zróżnicowania regionalnego zjawisk społeczno-ekonomicznych w Polsce, uwypuklając jedynie zagadnienie strefowości gospodarczej.

Przede wszystkim dokonuje się świadomej koncentracji wysiłku na wypracowaniu i sprawdzeniu właściwego modelu metodologicznego nowoczesnej regionalizacji, opartego na zastosowaniu metod matematyczno-statystycznych.

Główną metodą zastosowaną w pracy jest metoda czynnikowa. Metoda ta najlepiej realizuje postulat analizy strukturalnej złożonej rzeczywistości społeczno-gospodarczej. Badanie ogólnej przestrzeni społeczno-ekonomicznej w kategoriach istotnych przestrzeni czynnikowych jest badaniem jej ukrytej struktury.

Sam model czynnikowy nie rozstrzyga oczywiście, jakimi wskaźnikami opisu rzeczywistości należy się posługiwać i jakie wielkości empiryczne przypisać czynnikom. W konkretnych przypadkach można jednak te kwestie rozstrzygnąć na podstawie odpowiednio dobranego materiału statystycznego.

Analiza czynnikowa oparta jest na liniowych przekształceniach zmiennych na tzw. czynniki. Aby więc zbliżyć wyniki analizy do rzeczywistości, konieczny jest

prawidłowy dobór cech wyjściowych. Upřednio zwrócono już uwagę na fakt, że z punktu widzenia celów niniejszego opracowania wyjściowy materiał statystyczny nie reprezentuje, niestety, wszystkich istotnych własności przestrzeni społeczno-ekonomicznej. Wynika to z trudności zebrania odpowiedniego materiału statystycznego. Szczególnie dyskusyjna jest sprawa doboru wyjściowych zmiennych charakteryzujących gospodarke rolną; szersze uwzględnienie wielkości dotyczących poziomu gospodarki rolnej (wielkość produkcji czystej przypadającej na 1 ha użytków rolnych, wielkość produkcji wytworzonej przez 1 zatrudnionego w rolnictwie) wzbogaciłoby chyba w pewnym stopniu uzyskany obraz przestrzennego zróżnicowania obszaru Polski w dziedzinie rolnictwa.

Ta sytuacja utrudnia uzyskanie pełnego zbioru reprezentatywnych czynników, znamienych pod względem merytorycznym, ogranicza więc możliwości adekwatnego poznania rzeczywistości społeczno-gospodarczej Polski.

Należy jednak zaznaczyć, że w ramach już przyjętego zbioru cech przypisanie tym cechom wag czynnikowych oparte jest na analizie matematycznej. Na występowaniu tych właśnie wag polega wartość praktyczna modelu czynnikowego. Na podstawie wag, czyli stopnia korelacji między czynnikiem a innymi zjawiskami społeczno-ekonomicznymi, odbywa się identyfikacja czynników, która przy zachowaniu największej ostrożności może stać się źródłem szeregu odkrywczych hipotez.

W ramach niniejszej pracy interpretacja czynników, zawierająca zawsze pewien element subiektywny, była szczególnie niebezpieczna ze względu na symptomatyczny charakter zmiennych objaśniających i ze względu na pominięcie etapu rotacji wektorów czynnikowych w procedurze matematycznej.

Z punktu widzenia metodologicznego pewne zastrzeżenia może budzić wprowadzenie założenia o braku korelacji między czynnikami. Badanie empiryczne prowadzi do stwierdzenia, że ortogonalność czynników ma charakter względny. Wskazane byłoby więc przeprowadzenie rotacji ukośnej, uwzględniającej postulat korelacji czynników. Należy jednak podkreślić, że brak jest dotąd zgodności poglądów na temat celowości zastosowania rotacji ukośnej w procedurze analizy czynnikowej. Według R. B. Cattella (1952) czynniki ukośne ułatwiają stwierdzenie tożsamości czynników. Natomiast argumenty zwolenników osi prostopadłych są następujące. Z wyjątkiem przypadków wysokiej korelacji osi układu odniesienia wzory czynników uzyskane każdą z dwóch możliwych dróg nie różnią się w sposób istotny i mogą być podobnie interpretowane. W tych najczęstszych przypadkach, w których korelacje osi są nieznaczne, stosowanie rotacji ukośnej komplikuje bardzo obliczenia i nie prowadzi do różnic, które miałyby większe praktyczne znaczenie. Koncepcja czynników ukośnych przyjęła się przede wszystkim w psychologii. W geografii ekonomicznej próbę zastosowania tej koncepcji podjął R. K. Semple (1969) w pracy dotyczącej analizy podstawowych wymiarów społeczno-ekonomicznych miast południowego Ontario.

Również większą precyzję wyników być może zapewniłoby zastosowanie modelu o postaci wielomianu wyższego stopnia zamiast modelu czynnikowego liniowego.

Zależność zmiennych od czynników może być nieliniowa. Aktualnie nie jest jednak możliwe przeprowadzenie studiów empirycznych w tej dziedzinie ze względu na brak wypracowanych podstaw matematycznych modelu czynnikowego wyższego stopnia.

Uwagi krytyczne w sprawie zastosowania analizy czynnikowej nie oznaczają, że należy zrezygnować z tej metody. Jest to jedna z zasadniczych dróg ujmowania struktury rzeczywistości i we wstępnej fazie każdej regionalizacji odgrywa zasadniczą rolę. Metoda czynnikowa wyodrębnia z przestrzeni elementy strukturalne i zależności reprezentacyjne, pozwala więc prowadzić badania ograniczone do elementów pierwszorzędnych, wiodących. Innymi słowami przygotowuje podmiot działania regionalizacyjnego.

Wydaje się, że pewne deformacje uzyskanych obrazów regionalnych Polski w stosunku do rzeczywistości wynikają przede wszystkim z zasad grupowania regionalnego.

Kryteria regionalizacji mają co prawda rodowód empiryczno-statystyczny, wypracowany na podstawie konkretnego materiału obserwacyjnego, nie podlegały jednak weryfikacji statystycznej.

Konstrukcja przedziałów klasowych w klasyfikacji przestrzeni cząstkowych oparta jest na analizie rozsiewu znormalizowanych wartości czynnikowych. Ustalono 9 przedziałów o maksymalnej jednorodności. Jednak ze względu na skomplikowany charakter zróżnicowania przestrzeni w dalszym postępowaniu badawczym ten schemat podziału na klasy okazał się mało przydatny i został zastąpiony dwiema kategoriami dychotomicznymi, niewątpliwie o bardziej ograniczonej wartości.

Wydzielenie podprzestrzeni w zintegrowanej przestrzeni ogólnoeconomicznej oparte jest na ustaleniu przedziałów podobieństwa i zasad rozgraniczenia regionów.

Odległość taksonomiczna nie jest szczególnie dogodnym miernikiem podobieństwa. Przybiera wartości w przedziale od 0 do $+\infty$. Ponieważ nie da się w sposób ogólny określić górnej granicy wartości, jaką ona może przybierać (maksymalne niepodobieństwo), tym samym brak jest obiektywnej skali porównawczej przy badaniu stopnia podobieństwa różnych par jednostek. Na wielkie trudności natrafia się również przy określaniu wartości podobieństwa, zapewniającej dokładne rozróżnienia międzygrupowe. Klasyfikacja bowiem opiera się na wykrywaniu podobieństw przy jednoczesnym wyznaczaniu punktów, w których zaczyna się różnica, a więc należy przerwać proces grupowania.

Rzeczywisty rozkład odległości taksonomicznych wykluczał możliwość odgraniczenia w sposób prosty (np. na podstawie gradientów) grup wieloelementowych. W tej sytuacji podjęto próbę statystycznego oszacowania wartości krytycznej podobieństwa w postaci przeciętnej odległości taksonomicznej do najbliższego sąsiada powiększonej o wartość odchylenia standardowego w zależności od poziomu grupowania. Nie rozwiązany jednak został problem pomiaru efektywności poszczególnych szczebli grupowania. Również należy zdać sobie z tego sprawę, że jest wiele sposobów statystycznego określania wartości krytycznej podobieństwa, np. metoda

pomiaru straty informacji J. H. Warda (1963). Dlatego istnieje pytanie, czy dokonano tego najlepiej.

W procedurze grupowania zgodnie z założeniem, że żaden region nie powinien składać się tylko z jednej podstawowej jednostki przestrzennej, konieczne było również wprowadzenie tzw. zewnętrznych reguł przy włączaniu diametralnie różnych jednostek-enklaw lub jednostek przejściowych do alternatywnych regionów. Takie postępowanie w konsekwencji prowadzi do zaburzenia względnej jednolitości wydzielonych regionów.

W pracy przedstawiono jedynie pewne propozycje metodologiczne w zakresie każdorazowych kryteriów włączenia lub wyłączenia określonych punktów statystycznych w obręb pewnych podprzestrzeni, czyli regionów. Dalsze badania empiryczne z dziedziny regionalizacji powinny dokonać sprawdzenia proponowanych metod taksonomicznych.

LITERATURA

- Abiodun J. O. 1967. Urban hierarchy in developing country. *Econ. Geogr.*, V. 45, No 4, s. 347—367.
- Adcock C. I. 1954. *Factorial analysis for non-mathematicians*, Melbourne.
- Adelman I., Morris C. T. 1966. A quantitative study of social and political determinants of fertility. *Economic Development and Cultural Change*, V. 14, No 2, s. 129—195.
- Ahmad Q. 1965. Indian cities: characteristics and correlates, Research Paper, No 102, Dep. Geogr. Univ. of Chicago, Chicago.
- Anderson T. W. 1958. *An introduction to multivariate statistical analysis*, New York.
- Bachi R. 1963. Standard distance measures and related methods for spatial analysis, *Reg. Sci. Ass. Papers X*, European Congress, Zurich 1962, s. 83—132.
- Banks C. 1954. The factorial analysis of crop productivity. *J. Roy. Stat. Soc.*, B. 16, s. 100—111.
- Bartlett M. S. 1950. Tests of significance in factor analysis. *Brit. J. Psych. Statist.*, V. 3, s. 77—85.
- Behrens F. 1965. *Faktorenanalyse der Arbeitsproduktivität und Kybernetik*, Verlag Die Wirtschaft, Berlin.
- Bell W. 1955. Economic, family and ethnic status: an empirical test. *Amer. Sociol. Rev.*, V. 20, s. 45—52.
- Bell W., Stevenson. D. W. 1964. An index of economic health for Ontario countries and districts. *Ontario Econ. Rev.*, V. 2, s. 7—14.
- Berezowski S. 1959. Problem podziału Polski na regiony gospodarcze. *Gosp. plan.*, nr 3, s. 56—63.
- 1960. Regionalizacja gospodarcza, [w:] *Geografia ekonomiczna Polski. Praca zbiorowa pod red. S. Berezowskiego*. Warszawa, s. 405—475.
- Berry B. J. L. 1958. A note concerning methods of classification. *Ann. Ass. Am. Geogr.*, V. 48, s. 300—303.
- 1960. An inductive approach to the regionalization of economic development, [w:] Norton Ginsburg (ed.), *Essays on geography and economic development*. Research Paper, No 62, Dep. Geogr. Univ. of Chicago, Chicago, s. 78—107.
- 1961 a. Basic patterns of economic development, [w:] Norton Ginsburg (ed.), *Atlas of economic development, Part VIII*, Chicago, s. 110—119.
- 1961 b. A method for deriving multifactor uniform regions. *Przegl. geogr.*, t. 33, z. 2, s. 263—282.
- 1963. Commercial structure and commercial blight. Research Paper, No 85, Dep. Geogr. Univ. of Chicago, Chicago.
- 1966. *Essays on commodity flows and the spatial structure of the Indian economy*. Research Paper, No 111, Dep. Geogr. Univ. of Chicago, Chicago.
- 1967 a. The mathematics of economic regionalization, General Report and Resolution of the Meeting Economic Regionalization, Proceedings of the 4th General Meeting of the Commission on Methods of Economic Regionalization of the International Geographical Union, September 7—12, 1965 in Brno, Prague, s. 77—100.
- 1967 b. Grouping and regionalization: an approach to the problem using multivariate analysis, [w:] W. L. Garrison, D. F. Marble (ed.), *Quantitative Geogr. Stud. in Geogr.*, No 13, Northwestern Univ. Evanston.

- Berry B. J. L. 1968 a. Numerical regionalization of political economic space. *Geogr. Polon.*, 15, s. 27—35.
- 1968 b. Interdependency of spatial structure and spatial behavior: a general field theory formulation. *The Reg. Sci. Ass. Papers*, V. XXI, s. 205—227.
- Berry B. J. L. Barnum H. G. 1962. Aggregate relations and elemental components of central place systems. *J. Reg. Sci. Ass.*, V. 4, No 1, s. 35—68.
- Berry B. J. L. Barnum H. G., Tennant R. J. 1962. Retail location and consumer behaviour. *Reg. Sci. Ass. Pap. and Proc.*, V. IX, s. 65—106.
- Berry B. J. L. Rao V. L. S. P. 1968. Urban-rural duality in the regional structure of Andhra Pradesh. A challenge to regional planning and development. *Geogr. Ztschr.*, Heft 21, s. 1—40.
- Berry B. J. L. Ray D. M. 1966. Multivariate socio-economic regionalization: a pilot study in Central Canada, [w:] T. Rymes, S. Ostry (ed.) *Regional Statistical Studies*, Toronto.
- Bobiński J., Zagórski K. 1969. Zastosowanie analizy czynnikowej do określania poziomu rozwoju miast, [w:] *Mierniki rozwoju regionów*, *Bibl. Wiad. Statyst.*, t. 9, GUS, Warszawa, s. 230—254.
- Brown S. E., Trott Ch. E. 1968. Grouping tendencies in an economic regionalization of Poland. *Ann. Ass. Am. Geogr.*, V. 58, No 2, s. 327—342.
- Bunge W. 1962. *Theoretical geography*, Lund.
- Burt C. 1917. *The distributions and relations of educational abilities*, London.
- Cattell R. B. 1952. *Factor analysis: an introduction and manual for the psychologist and social scientist*, New York.
- Carey G. W. 1966. The regional interpretation of Manhattan population and housing patterns through factor analysis. *Geogr. Rev.*, V. 56, s. 551—560.
- Casetti E. 1964. Multiple discriminant functions, Technical Report No 11, Computer Applications in the Earth Sciences Project, Dep. Geogr. Northwestern University.
- Chojnicki Z. 1969. Metody matematyczne w badaniach geograficznych. *Czas. geogr.*, t. 11, z. 2, s. 175—188.
- 1970. Podstawy teoretyczne zastosowania metod matematycznych w badaniach przestrzennych rolnictwa, *Biul. KPZK PAN*, z. 61, Warszawa, s. 7—41
- Chojnicki Z., Czyż T. 1968. Wyznaczanie układów powiązań na podstawie międzyregionalnych przepływów towarowych przy zastosowaniu metody analizy czynnikowej, *Spraw. PTPN za I płr.* 1968, nr 1, s. 135—137.
- — 1971. Structural changes of economic regions in Poland. A study by factor analysis of commodity flows, *Geogr. Polon.* (w druku).
- Chojnicki Z., Wróbel A. 1961. Metody matematyczno-statystyczne w geografii ekonomicznej. *Przegl. geogr.*, t. 33, z. 4, s. 615—629.
- Chorley R. J., Haggett P. 1967. *Models in geography*, London.
- Cole J. P., King C. A. M. 1968. *Quantitative geography. Techniques and theories in geography*, London.
- Cooley W. W., Lohnes P. R. 1962. *Multivariate procedures for the behavioral sciences*, London.
- Coppock T. J. 1960. The parish as a geographical statistical unit. *Tijdschr. Econ. en Soc. Geogr.*, 51, s. 317—326.
- Czyż T. 1967. Wyznaczanie regionów jednolitych metodą analizy czynników wielokrotnych. *Przegl. geogr.*, t. 39, z. 1, s. 135—160.
- 1968. The application of multifactor analysis in economic regionalization. *Geogr. Polon.*, 15, s. 115—133.
- Czyżewski J. 1930. Regiony drzewno-gospodarcze Polski, [w:] *Pamiętnik Zjazdu Stowarzyszenia Etnografów*, t. 2, s. 76.
- Domański R. 1964. Procedura typologiczna w badaniach ekonomiczno-geograficznych. *Przegl. geogr.* t. 36, z. 4, s. 627—655.

- 1965. Uwagi o analizie czynnikowej w związku z artykułem M. Megee. *Biul. KPZK PAN*, z. 34, Warszawa, s. 211—214.
- 1967. Konstruowanie teorii w geografii ekonomicznej. *Przegl. geogr.*, t. 39, z. 1, s. 85—102.
- Duncan O. D., Cuzzort R. P., Duncan B. 1961. *Statistical geography. Problems in analyzing areal data*, Glencoe, Ill.
- Dziewoński K. 1957. Niektóre problemy badania regionów gospodarczych w Polsce. *Przegl. geogr.*, t. 29, z. 4, s. 719—739.
- 1964. (rec.) C. A. Moser, W. Scott. *British towns. A statistical study of their social and economic differences* Centre for Urban Studies, Report No 2, London 1961. *Przegl. geogr.*, t. 36, s. 363—366.
- 1967. Teoria regionu ekonomicznego. *Przegl. geogr.*, t. 39, z. 1, s. 33—50.
- 1968. Economic regionalization. A report of progress. *Geogr. Polon.*, 15, s. 9—24.
- 1969. W diagnozach regionalnych należy stosować wielorakie metody analityczne, [w:] *Mierniki rozwoju regionów*, *Bibl. Wiad. Statyst.*, t. 9, GUS, Warszawa, s. 258—260.
- Dziewoński K., Wróbel A. 1964. Regional structure and economic regions of Poland. *Methods of economic regionalization. Geogr. Polon.*, 5, s. 47—58.
- Emrick H. W. 1965. A method for intra-regional comparisons. *Profess. Geogr.* 17 (3), s. 18—19.
- Ernst J. 1932. Regiony geograficzno-rolnicze Polski. *Czas. geogr.*, t. 10, s. 143—168.
- Fisher J. C. 1966. *Yugoslavia. A multinational state. Regional difference and administrative response*, San Francisco.
- Florek K., Łukaszewicz J., Perkal J., Steinhaus H., Zubrzycki S. 1951. *Taksonomia wrocławska. Przegl. antrop.*, t. 8.
- Freund J. E. 1968. *Podstawy nowoczesnej statystyki*, Warszawa.
- Fruchter B. 1954. *Introduction to factor analysis*. New York.
- Garrison W. L., Marble D. F. 1964. Factor analytic study of the connectivity of a transportation network. *Reg. Sci. Ass. Papers*, V. XII, s. 231—238.
- Geary R. C. 1968. The contiguity ratio and statistical mapping, [w:] B. J. L. Berry, D. F. Marble, *Spatial analysis. A reader in statistical geography*. Prentice Hall, Inc. Englewood Cliffs, New Jersey.
- Gould P. R. 1967. On the geographical interpretation of eigenvalues. *Inst. Brit. Geogr. Transactions*, No 42, s. 53—86.
- Gregory C. L. 1949. Advanced techniques in the delineation of regions. *Rural Sociology*, V. XIV, s. 59—63.
- Grigg D. 1965. The logic of regional systems. *Ann. Ass. Amer. Geogr.*, V. 55, s. 465—491.
- 1967. Regions, model and classes, [w:] R. J. Chorley, P. Haggett, *Models in geography*, London, s. 461—507.
- Hadden J. K., Borgatta E. F. 1965. *American cities. Their social characteristics*, Chicago.
- Haggett P. 1965. *Locational analysis in human geography*, London.
- Hagood M. J. 1943. Statistical methods for delineation of regions applied to data on agriculture and population. *Social Forces*, V. 21, s. 287—297.
- Hagood M. J., Price D. O. 1957. *Statistics for sociologists*, New York.
- Harman H. H. 1960. *Modern factor analysis*, Chicago.
- Hartshorne R. 1959. *Perspective on the nature of geography*, Chicago.
- Hładyniuk Z. 1966. Zastosowanie współczesnych metod obliczeniowych do oznaczania granic rejonów rolniczych (na przykładzie województwa lubelskiego). *Zagad. Ekon. Rolnej* 6 (78), s. 156—157.
- Holzinger K. J., Harman H. H. 1941. *Factor analysis: a synthesis of factorial methods*, Chicago.
- Horton F. E., Mc Connell H. 1969. Spatial and temporal changes in the Middlewestern rural economy. *Geogr. Polon.* (w druku).

- Hotelling H. 1933. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, 24, s. 417—441, 498—520.
- Henshall J. D. 1966. The demographic factor in the structure of agriculture in Barbados. *Inst. Brit. Geogr. Transactions*, V. 38, s. 183—195.
- Henshall J. D., King L. J. 1966. Some structural characteristic of peasant agriculture in Barbados. *Econ. Geogr.*, V. 42, No 1, s. 74—84.
- Illeris S., Pedersen P. O. 1968. Central places and functional regions in Denmark. Factor analysis of telephone traffic. *Sa rtryk at Geografisk Tidsskrift*, 67 bind, s. 1—18.
- Isard W. 1960. *Methods of regional analysis: an introduction to regional science*, New York, s. 293—305.
- Johnston R. J. 1970. Grouping and regionalizing: some methodological and technical observations. *Econ. Geogr.*, V. 46, No 2, s. 293—305.
- Jonassen Ch. T., Peres S. H. 1960. Interrelationships of dimensions of community systems, a factor analysis of eighty-two variables, Columbus.
- Jones B. G., Goldsmith W. W. 1968. A factor analysis. Approach to subregional definition in Chenango, Delaware and Otsego counties. *Geogr. Polon.*, 15, s. 59—114.
- de Jong G. 1962. Chorological differentiation as the fundamental principle of geography, Groningen.
- Joreskog K. G. 1963. *Statistical estimation in factor analysis: a new technique and its foundation*, Stockholm.
- Kawalec W. 1965. *Okręgi przemysłowe i regiony ekonomiczne w Polsce*, Warszawa.
- Kendall M. G. 1939. The geographical distribution of crop productivity in England. *J. Roy. Stat. So. A*, 102, s. 21—62.
- 1957. *A course in multivariate analysis*, London.
- Kendal M. G., D. N. Lawley. 1956. The principles of factor analysis. *J. Roy. Stat. So. A*, 119.
- Kendal M. G., B. B. Smith. 1950. Factor analysis. *J. Roy. Stat. So. B*, V. XII, No 1, s. 60—85.
- Kelley T. L. 1935. *Essential traits of mental life*, Cambridge.
- King L. J. 1966. Cross-sectional analysis of Canadian urban dimensions: 1951 and 1961. *Can. Geogr.*, V. 10, s. 205—224.
- 1969. *Statistical analysis in geography*, Englewood Cliffs, N. J.
- Kostrowicki J. 1967. Zagadnienie specjalizacji i integracji w geografii ekonomicznej. *Przegl. geogr.*, t. 39, z. 1, s. 13—32.
- Kostrubiec B. 1968. Klasyfikacja dynamiczna i wielocephowa Polski, *Biul. KPZK PAN*, z. 35, Warszawa, s. 28—49.
- 1969. Wieloczynnikowa analiza zróżnicowania przestrzennego procesu uprzemysłowienia Polski w latach 1946—1956. *Biul. KPZK PAN*, z. 52, Warszawa, s. 63—80.
- Kukliński A. 1961. Problemy badań nad lokalizacją poszczególnych gałęzi przemysłu w Polsce. *Biul. KPZK PAN*, nr 7, s. 7—44.
- Labbe S. 1965. L'analyse factorielle. *Metra*, V. IV, No 3, s. 421—439.
- Lawley D. N. 1943. The application of the maximum likelihood method to factor analysis. *Bri. J. Psych.* 33, s. 172—175.
- Lawley D. N., Maxwell A. E. 1963. *Factor analysis as a statistical method*, London.
- Leszczycki S., Wróbel A. 1967. Statystyczne wskaźniki procesu urbanizacji, *Bibl. Wiad. Stat.*, t. 4, GUS, Warszawa, s. 129—137.
- Lewiński S. 1968 a. Taxonomic methods in regional studies. *Geogr. Polon.*, 15, s. 189—198.
- 1968 b. Metody numeryczne i ich zastosowanie dla charakterystyki miast. *IUA*, Warszawa (mpis).
- Liberg M. 1961. Multiple factor analysis as a method of agricultural research, *Netherl. J. Agric. Sci.*, V. 9, z. 4.

- Lijewski T. 1963. Próba podziału Polski na regiony statystyczne. *Przeł. geogr.*, t. 35, z. 4, s. 655—661.
- Lundqvist J. 1969. Faktoranalys inom kulturgeografien. *Choros*, No 11, Kulturgeografiska Institutionem Goteborgs Univ.
- Mabogunje A. L. 1965. Urbanization in Nigeria: a constraint on economic development. *Econ. Devel. and Cultural Change*, V. 13, s. 413—438.
- Megee M. 1963. Social and economic factors in the differential growth of Mexican States. Paper presented to the First Latin American Conference on Regional Science held at Cendes, Caracas 1963, s. 143—178.
- 1965 a. On economic growth and the factor analysis method. *The South. Econ. J.* January, s. 215—228.
- 1965 b. Economic factors and economic regionalization in the United States. *Geogr. Annaler*, B, 47, s. 125—137.
- 1965 c. Nowe dziedziny zastosowania analizy czynnikowej: sprawdzenie hipotez dotyczących rozwoju gospodarczego. *Biul. KPZK PAN*, z. 34, Warszawa, s. 187—209.
- 1966. Problems in regionalization and measurement. *Peace Research Society, Papers*, IV, Cracow Conference 1965, s. 7—35.
- Morrison J. L., Scriptor M. W., Smith R. H. T. 1968. Basic measures of manufacturing in the United States 1958. *Econ. Geogr.*, V. 44, No 4, s. 296—311.
- Moser C. A., Scott W. 1961. British towns. A statistical study of their social and economic difference, Edinburgh.
- Nowak S. 1965. *Studia z metodologii nauk społecznych*, Warszawa.
- Okóń J. 1964. Analiza czynnikowa w psychologii. Warszawa. Określenie poziomu rozwoju ekonomicznego powiatów. 1968. *Seria Statystyka Regionalna*, nr 14, GUS, Warszawa.
- Olsen B. R., Garb G. 1965. An application of factor analysis to regional economic growth. *J. Reg. Sci.*, V. 6, No 1, s. 51—56.
- Pearson K. 1901. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 6, s. 559—572.
- Pedersen P. O. 1967 a. An empirical model of urban population structure. A factor analytical study of the population structure in Copenhagen. *Proceedings of the First Scandinavian-Polish Regional Science Seminar, Committee for Space Economy and Regional Planning Polish Academy of Sciences, Studies*, V. XVII, Warszawa, s. 193—214.
- 1967 b. Scalogramtechnik og faktoranalyse, Regionale Analysemetoder, Forelesninger pa RSA—seminar i Oslo juni 1966, Oslo, s. 87—119.
- Perkal J. 1953. O wskaźnikach antropologicznych. *Przeł. antrop.*, z. 19, s. 209—221.
- Piasecki Z. 1969. Analiza czynnikowa metodą Hotellinga. Program na maszynę cyfrową Gier. Centrum Obliczeniowe PAN, Warszawa.
- Piekałkiewicz J., Rutkowski S. 1927. Okręgi gospodarcze Polski. *Kwart. statyst.*, t. 4, z. 3.
- Price D. O. 1942. Factor analysis in the study of metropolitan centers. *Social Forces*, V. 20, s. 449—455.
- Roberts R. E., Mc Bee G. W. 1968. Modernization and economic development in Mexico: a factor analytic approach. *Econ. Devel. and Cultural Change*, V. 16, s. 603—612.
- Robinson A. 1956. The necessity of weighting values in correlation analysis of areal data. *Ann. Ass. Am. Geogr.*, V. 46, s. 233—236.
- Rodoman B. B. 1968. Mathematical aspects of the formalization of regional geographic characteristics. *Geogr. Polon.*, 15, s. 37—57.
- Rokita S. 1966. Analiza czynnikowa w badaniach regionalnych. *Przeł. statyst.*, R. XIII, z. 3, s. 245—260.
- Rummel R. J. 1967. Understanding factor analysis. *Journal of Conflict Resolution*, December, s. 444—480.

- Rychłowski B. 1960. Podział Polski na regiony ekonomiczne. *Gosp. plan.*, nr 1.
 — 1965. Typy i rodzaje regionalizacji ekonomicznej. *Przeł. geogr.*, t. 37, z. 1, s. 29—56.
 — 1967. Regionalizacja ekonomiczna. *Prace geogr. IG PAN*, nr 64.
- Saunders D. R. 1950. *Practical methods in direct factor analysis of psychological score matrixes*. Unpublished Ph D thesis, University of Illinois.
- Schmid C. F., Tagashira K. 1964. Ecological and demographic indices: a methodological analysis. *Demography*, V. 1, No 1, s. 194—211.
- Schnore L. F. 1961. The statistical measurement of urbanization and economic development. *Land Econ.*, V. 37, 3, s. 229—245.
- Secomski K. 1956. *Wstęp do teorii rozmieszczenia sił wytwórczych*, Warszawa.
- Semple R. K. 1969. An oblique „simple structure” factor analysis of viability measurement for southern Ontario towns, *Dep. Geogr. Univ. of Toronto, Discussion Paper Series, No 2*, s. 1—14.
- Shevky E., Williams M. 1949. *The social areas of Los Angeles: analysis and typology*. Univ. of California Press.
- Soja E. 1968. *The geography of modernization in Kenya*. Syracuse.
- Sokal R. R., Sneath P. H. 1964. *Principles of numerical taxonomy*, San Francisco.
- Solomon H. 1960. *Mathematical thinking in the measurement of behavior*. Glencoe, Ill.
- Spearman Ch. 1904. General intelligence objectively determined and measured. *Am. J. Psych.*, V. XV, s. 201—294.
- Spence N. A. 1968. A multivariate uniform regionalization of British counties on the basis of employment data for 1961. *Regional Studies*, 2, s. 87—104.
- Steiner D. 1965. Die Faktorenanalyse — Ein modernes statisches Hilfsmittel des Geographen für die objektive Raumgliederung und Typenbildung., *Geogr., Helv.*, No 1, s. 20—34.
- Stephenson W. 1936. The inverted factor technique, *Brit. J. Psychol.*, 26, s. 344—361.
- Stevens B. H., Brackett C.A. 1968. Regionalization of Pennsylvania countries for development planning. *Geogr. Polon.*, 15, s. 153—187.
- Stone R. 1960. A comparison of the economic structure of regions based on the concept of distance. *J. Reg. Sci.*, V. 2, No 2, s. 1—20.
- Styś K. 1967. *Program na analizę czynnikową metodą kierunków głównych*, Warszawa.
- Sweetser F. L. 1965 a. Factorial ecology: Helsinki 1960, *Demography*, V. 2.
 — 1965 b. Factorial ecology: zonal differentiation in metropolitan Boston 1960. Paper presented at Population Ass. of America Annual Meeting, April 1965, Chicago.
 — 1965 c. Factor structure as ecological structure in Helsinki and Boston. *Acta Sociologica*, V. 8, s. 205—255.
- Thomas E., Anderson D. L. 1965. Additional comment on weighting values in correlation analysis of areal data. *Ann. Ass. Am. Geogr.*, V. 55, s. 492—505.
- Thompson D. 1970. Some comments on the relevance of multivariate analysis to geography: a methodological review. *Geogr. Polon.*, 18, s. 157—175.
- Thompson J. H., Sufirin S. C., Gould P. R., Buck M. A. 1962. Toward a geography of economic health: the case of New York State. *Ann. Ass. Am. Geogr.*, V. 52, No 1, s. 1—20.
- Thurstone L. L. 1931. Multiple factor analysis. *Psych. Rev.*, V. XXXVIII, No 5, s. 406—427.
 — 1935. *The vectors of mind*, Chicago.
 — 1947. *Multiple-factor analysis. A development and expansion of the vectors of mind*, Chicago.
- Tucker L. R. 1944. The determination of successive principal components without computation of tables of residual correlation coefficients. *Psychometrika*, 9, s. 149—153.
- Van Ardsol M. D., Camilleri S. F., Schmid C. F. 1958. An investigation into the generality of the Shevky social area indices. *Am. Sociol. Rev.*, V. 23, s. 277—284.
- Ward J. H. 1963. Hierarchical grouping to optimize and objective function. *J. Am. Stat. Ass.*, V. 58, s. 236—244.

- Wherry R. J. 1959. Hierarchical factor solutions without rotation. *Psychometrika*, V. XXIV, No 1, s. 45—51.
- Whittlesey D. 1954. The regional concept and the regional method, [w:] *Amer. Geogr., Inventory and Prospect*, Syracuse.
- Wróbel A. 1956. Kryteria i metody delimitacji regionów gospodarczych. *Dokum. geogr. IG PAN*, nr 3, Warszawa.
- 1963. Region ekonomiczny — narzędzie analizy i planowania. *Gosp. plan.*, nr 6, s. 37—43.
- 1965. Pojęcie regionu ekonomicznego a teoria geografii. *Prace geogr. IG PAN*, nr 48.
- 1967. Pojęcie regionu a metoda regionalna. *Przegl. geogr.*, t. 39, z. 1, s. 73—84.
- 1969. Agregatowe i wskaźnikowe metody pomiaru i oceny stopnia rozwoju gospodarczego regionów, [w:] *Mierniki rozwoju regionów, Bibl. Wiad. Statyst.*, t. 9, GUS, Warszawa, s. 41—56.
- Wysocki Z. 1964. Niektóre problemy z metody badania obrazu regionalnego Polski w okresie 20-lecia powojennego. *Czas. geogr.*, t. 35, z. 3—4, s. 429—443.
- 1965. Zagadnienie taksonomii geograficznej. *Przegl. geogr.*, t. 38, z. 2, s. 313—339.
- 1966. Próba typologii i systematyki geograficznej struktur gospodarstwa narodowego na przykładzie Polski, *Prace Wrocł. Tow. Nauk. B*, nr 126, Wrocław.
- 1968. O problemie klasyfikacji i porządkowania w geografii ekonomicznej. *Przegl. geogr.*, t. 40, z. 3, s. 585—621.
- Zarys geografii ekonomicznej Polski. 1967. Praca zbiorowa pod red. S. Leszczyckiego i L. Kosińskiego, Warszawa.
- Zeliaś A. 1968. Analiza czynnikowa w badaniach nad rejonizacją produkcji rolniczej. *Zagadnienia Ekonomiki Rolnej*, 5 (89), s. 83—97.
- Žukovskaja W. M. 1964. Opyt primienienija mnogofaktornogo analiza dla ekonomiko-geograficeskoj charakteristiki sielskogo chozajstwa stepnych provincij Kanady, [w:] *Koliczestwiennyje metody issledowanija w ekonomiczeskoj gieografii*, pod red. I. M. Majergojza, Moskowskij Finał Wśiesojuz. *Gieogr. Obszcz. Moskwa*, s. 122—165.
- Žukovskaja W. M., Karpow L. 1967. Application of multifactor analysis in the identification and classification of agricultural regions. *Reg. Sci. Ass. Papers, XX, Hague Congress 1967*, s. 55—62.
- Yule G. U., Kendall M. G. 1966. *Wstęp do teorii statystyki*, Warszawa.

ПРИМЕНЕНИЕ ФАКТОРНОГО АНАЛИЗА ДЛЯ ИЗУЧЕНИЯ ЭКОНОМИЧЕСКОЙ РАЙОННОЙ СТРУКТУРЫ ПОЛЬШИ

Резюме

Работа является синтетическим анализом экономической районной структуры Польши с точки зрения однородности. В основу работы легла концепция скрытой структуры общественно-экономического пространства. Как метод идентификации этой скрытой структуры введен метод факторного анализа.

В первой части работы — методологической, представлены теоретические основы и сфера применения факторного анализа, вторая часть — эмпирически-познавательная, выполняет главную цель работы — изучение экономической районной структуры Польши.

Факторный анализ охватывает группу статистико-математических методов, позволяющих свести первичное множество переменных, характеризующих исследуемые объекты, к значительно меньшему числу гипотетических переменных, называемых факторами. Эти новые переменные содержат основную информацию об оригинальных переменных. Итак, факторный анализ выделяет факторы (величины), которые легли в основу корреляций выявленных в данном множестве переменных. К этим факторам можно подойти как к причинам замеченной изменчивости.

Факторный анализ — это линейная математическая модель. Модель факторного анализа строится как вступительную предпосылку, представляемую уравнением:

$$z_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jm}F_m + a_jU_j.$$

Эмпирические переменные z_j ($j = 1, 2, \dots, n$) выражены как линейные гипотетические переменные (скрытые факторы) F_m и U_j . F_1, F_2, \dots, F_m это независимые величины, называемые общими факторами, а U_1, U_2, \dots, U_n это т. наз. простые факторы. Параметры уравнения a_{jm} , a_j называются факторной нагрузкой и определяют вес данного фактора в описании эмпирических переменных.

Факторное решение состоит в конкретизации модели, т. е. в числовом выражении параметров модельных уравнений. Конкретизация линейных зависимостей переменных и общих факторов, проводимая с помощью разных математических методов выделения факторов ведет к определению т. наз. факторной структуры, являющейся основным элементом в факторной процедуре.

Факторы являются скрытыми переменными, непосредственно ненаблюдаемыми. Итак, факторная модель принадлежит к классу скрытых структур. Идентификация факторов имеет решающее значение для адекватного познания действительности. Эта операция однако методологически не идентифицирована.

Обзор применений факторного анализа в экономико-географических исследованиях в мировой литературе позволяет определить факторный анализ как 1) модель деления множества переменных на подмножества, 2) метод ортогонализации переменных, 3) метод ре-

дукции массы информации, 4) модель измерения, 5) метод эксперимента, 6) метод классификации, 7) орудие проверки гипотез, 8) обобщающую процедуру.

В изучении экономической структуры районов Польши факторный анализ является методом выявления скрытой структуры общей социально-экономического пространства страны.

Общее социально-экономическое пространство понимается по К. Дзевоньскому как множество частичных пространств, составляющих его скрытую структуру.

Согласно этой концепции строится алгоритм факторного решения вопроса районирования, учитывающий 4 этапа: 1) наблюдение, 2) редукция многообразного пространства, 3) оценка сходств, 4) территориальное группирование.

Опираясь на статистический материал, охватывающий 33 социально-экономические признака для 324 основных территориальных единиц в Польше в 1965 г., выделяются с помощью метода главного фактора со схемой итерации Хотеллинга 4 фактора, выясняющие 72,5% общей части дисперсии. Программа факторного анализа, записанная на языке ГИЭР-АЛГОЛ IV, выполняла метод на ЭВМ ГИЭР. Три основных фактора толкуются как структурные системы социально-экономической действительности страны в виде: 1) системы индустриализации и урбанизации, 2) системы общественных отношений и землевладения в сельском хозяйстве, 3) системы интенсивности сельского хозяйства.

Анализ территориального распределения этих факторов ведет к выводу, что структура этих частичных пространств, в основном, типологического, а не районного характера. Постепенная интеграция этих систем с помощью методов оценки сходства (показатель Перкаля, таксономное расстояние) и метод группирования по теории графов (м. пр. метод ближайшего соседа) в системы двух-, трех-, четырехфакторные позволяет провести делимитацию 95 микрорайонов и при дальнейшем обобщении — 24 однородных районов Польши.

Исследование соотношения выделенной районной системы с воеводской системой подтверждает, что воеводская система неоднородна.

В заключении автор утверждает, что работа является в некоторой степени своего рода экспериментом, предложением в области таксономного решения вопросов экономико-географического районирования. Проба определения сложности районной структуры Польши, опирающейся на критерий зонирования, рассматривается как вступительная гипотеза, которую следует проверить с помощью дальнейших эмпирических исследований.

Перевела Ханна Деренговска

THE APPLICATION OF FACTOR ANALYSIS IN THE STUDY OF POLAND'S ECONOMIC REGIONAL STRUCTURE

Summary

The present study is a synthetic presentation of the economic regional structure of Poland from the point of view of its uniformity. The study is based on the concept of latent structure of socio-economic space. To identify this latent structure, the factor analysis has been applied.

The first part of the study, dealing with methodological problems discusses the theoretical assumptions and the scope of applicability of factor analysis. The second part, of empirical-cognitive nature, fulfills the actual task of the study, that is presenting the economic regional structure of Poland.

Factor analysis comprises several mathematical-statistical techniques allowing of reducing an initial set of variables that characterize the objects under observation to a considerably smaller number of hypothetical variables called factors. These new variables contain the fundamental body of information on the initial variables. Thus, factor analysis identifies the factors (dimensions) underlying the correlations stated in a given set of variables. These factors may be considered to be the causes of the variation observed.

Factor analysis is a linear mathematical model. The model of factor analysis is built as a preliminary assumption in the form of the equation

$$z_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jm}F_m + a_j U_j.$$

The empirical variables z_j ($j = 1, 2, \dots, n$) are expressed as linear functions of hypothetical variables (latent factors) F_m and U_j . F_1, F_2, \dots, F_m are independent variables called the common factors, and U_1, U_2, \dots, U_n are what is called the specific factors. The parameters of the equation, a_{jm} and a_j , are called the factor loadings and determine the „weight” of a factor in the description of empirical variables.

To give a factorial solution to this model, the parameters of the model equations are substituted by numerical values. This concretization of the linear dependencies between the variables and the common factors is made by definite mathematical methods of extracting factors, and it leads to the determination of the „factorial structure”, which constitutes a fundamental stage in the procedure of factor analysis.

The factors are latent, that is directly unobservable, variables. The factorial model belongs therefore to the class of latent structures. The identification of the factors is of decisive importance to an adequate description of reality. This operation however methodologically is unidentified. A survey of the applications of factor analysis in the world literature on economic-geographical investigations allows of appraising factor analysis as 1) a model of division of a set of variables into subsets, 2) a method of orthogonalization of variables, 3) a method of reducing of information mass, 4) a model of measurements, 5) an experimental method, 6) a method of classification, 7) an instrument for testing hypotheses, 8) a generalizing procedure.

In the study of the economic regional structure of Poland, factor analysis is a method of disclosing the latent structure of the overall socio-economic space of the country.

Following K. Dziewoński, the overall socio-economic space of the country is treated as a set of partial elementary spaces which constitute its latent structure. These fundamental elementary spaces of the overall space identify the essential factors.

Accordingly, the algorithm of the factorial solution of the regionalization is constructed. This is completed in four stages: 1) observation, 2) reduction of a multivariate space, 3) estimation of similarity, 4) spatial grouping.

Using the statistical materials covering 33 socio-economic characteristics of 324 areal units on the territory of Poland in 1965, 4 factors explaining 72.5% of the common variation are identified by the principal factor method with Hotelling's iteration system. A program for factor analysis written in the Gier-Algol IV language was adjusted to a Gier computer. Three fundamental factors were interpreted as the structural patterns of the country's actual socio-economic state: 1) the pattern of industrialization and urbanization, 2) the pattern of social forms of ownership in agriculture, 3) the pattern of intensity of agriculture economy.

The analysis of the spatial distribution of these factors shows that the structure of the elementary spaces is principally typological rather than regional in its character. A gradual integration of these patterns by methods of similarity estimation (Perkal's index, taxonomic distance) and by graph grouping (among others, the nearest neighbor method) into two-, three-, and four-factor systems makes possible a final delimitation of 95 uniform microregions and, after further extension of the level of generality, of 24 uniform regions of Poland.

An examination of the relationship between the regional pattern and the administrative pattern of voivodships confirms the view that the latter is not uniform.

In conclusion, the author remarks that this study has some specific experimental features, that it is a proposition of the taxonomic solution of the problems of economic-geographical regionalization. This attempt to define the complex nature of Poland's regional structure based on the criterion of zoning must be viewed as a preliminary hypothesis, which ought to be tested in further empirical investigations.

Translated by Zygmunt Nierzada

WAŻNIEJSZE DOSTRZEŻONE BŁĘDY DRUKU

Stronica	Wiersz	Jest	Powinno być
10	19 od g.	wnej	pewnej
10	20 od g.	pe	W.
10	21 od g.	W964	W 1964
10	22 od g.	le	ze
10	23 od g.	ednak	jednak
28	8 od g.	<i>s_j</i>	<i>z_j</i>
34	8 od d.	zbioru diagnostycznych	zbioru cech diagnostycznych
67	4 od g.	wspólne	wspólnej
73	14 od g.	$f_{33,1}$ $f_{33,2}$ $f_{33,3}$ $f_{33,4}$	$f_{324,1}$ $f_{324,2}$ $f_{324,3}$ $f_{324,4}$
84	2 od d.	e dnostkami	jednostkami
98			
(tab. 19)	2 od d.	t5,45	5,45
107	22 od g.	endall	Kendall
107	25 od g.	Kendal	Kendall
107	26 od g.	Kendal	Kendall

T. Czyż, Zastosowanie...

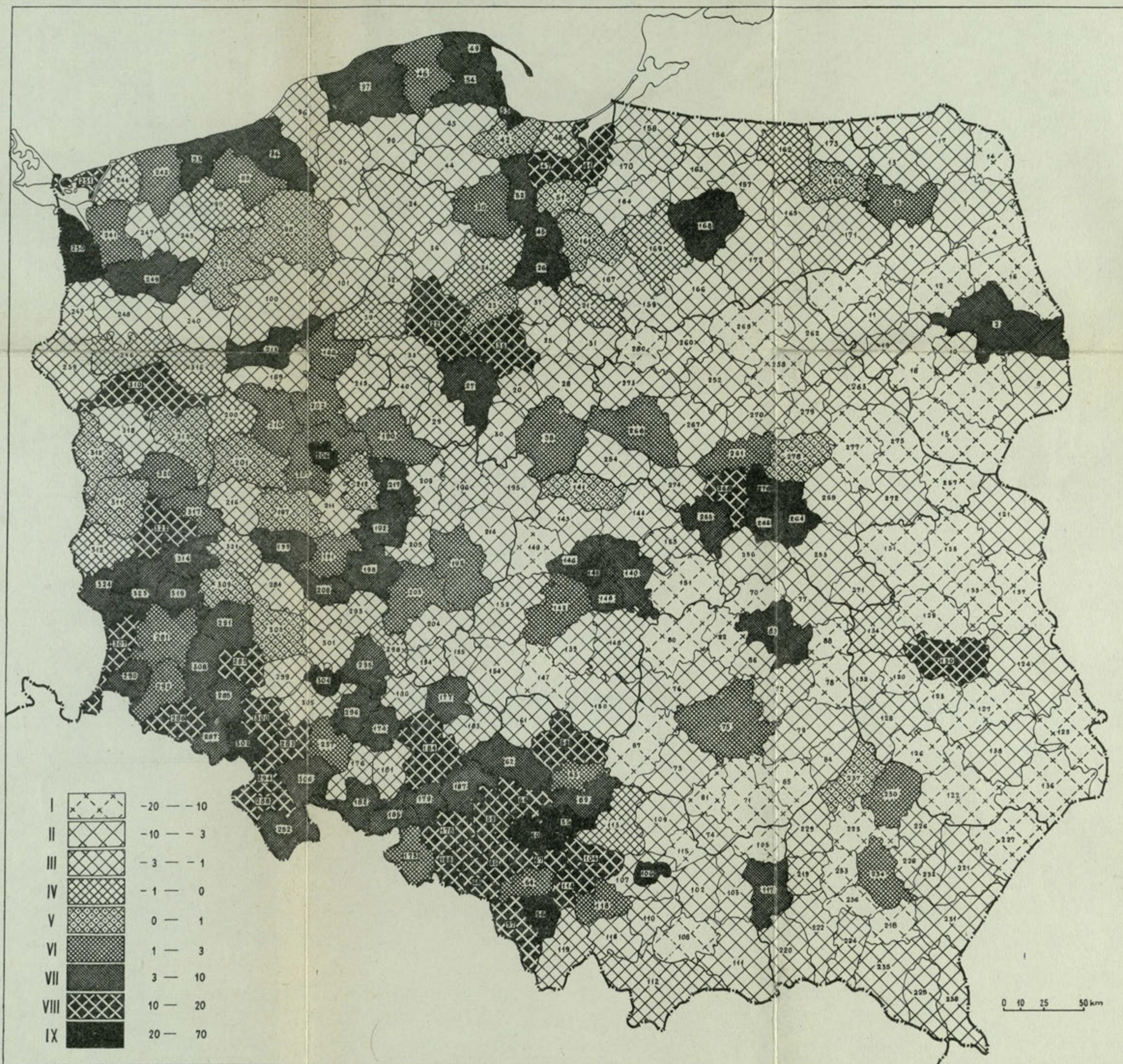
77. Praca zbiorowa. Z zagadnień ludnościowych krajów gospodarczo słabo rozwiniętych. 1969, s. 146+6 ilustr., zł 32,—
78. Korcelli P., Rozwój struktury przestrzennej obszarów metropolitalnych Kalifornii. 1969, s. 124+34 ilustr., zł 28,—
79. Koter M., Geniza układu przestrzennego Łodzi przemysłowej. 1969, s. 130+13 ilustr. +2 wkładki, zł 34,—
80. Kaszowski L., Kotarba A., Wpływ katastrofalnych wezbrań na przebieg procesów fluwialnych (na przykładzie potoku Kobylanka na Wyżynie Krakowskiej)
Nowak W. A., Rzeźba podczwartorzędowa i ewolucja układu sieci dolinnej w północno-środkowej części Wyżyny Małopolskiej. 1970, s. 124+71 ilustr. +1 zał. +12 fot., zł 30,—
81. Stola W., Próba typologii rolnictwa Poniżnia. 1970, s. 146+30 ilustr. +9 fot., zł 39,—
82. Praca zbiorowa. Studia z geografii średnich miast w Polsce. Problematyka Tarnowa (w druku)
83. Wiśniewski E., Struktura i tekstura sandru ostródzkiego oraz teras górnej Drwęcy (w druku)
84. Skoczek J., Wpływ podłoża atmosfery na przebieg dobowy bilansu cieplnego powierzchni czynnej. 1970, s. 96+49 ilustr. +10 fot., zł 21,—
85. Jewtuchowicz S., Rozwój rzeźby okolic Łęczycy po zlodowaceniu środkowopolskim. 1970, s. 78+26 ilustr. +5 fot., zł 18,—
86. Olechnowicz-Bobrowska B., Częstość dni z opadem w Polsce. 1970, s. 75+26 ilustr., zł 18,—
87. Baza ekonomiczna i struktura funkcjonalna miast
Dziwioński K., Studium rozwoju pojęć, metod i ich zastosowań. Jerczyński M., Metody pośrednie identyfikacji i pomiaru. 1971, s. 182+2 ilustr., zł 44,—
88. Rościszewski M., Kierunki ewolucji rolnictwa w krajach Maghrebu. 1970, s. 127+8 ilustr., zł 30,—
89. Adrjanowska E., Przestrzenne powiązania produkcyjne stoczni gdańskich. 1971, s. 105+16 ilustr., zł 22,—
90. Różycka W., Metody oceny warunków fizjograficznych dla potrzeb planowania przestrzennego miast. 1971, s. 204+2 wkładki, zł 40,—
91. Ciołkosz A., Miszański J., Wykorzystanie zdjęć lotniczych w geografii stosowanej (w druku)

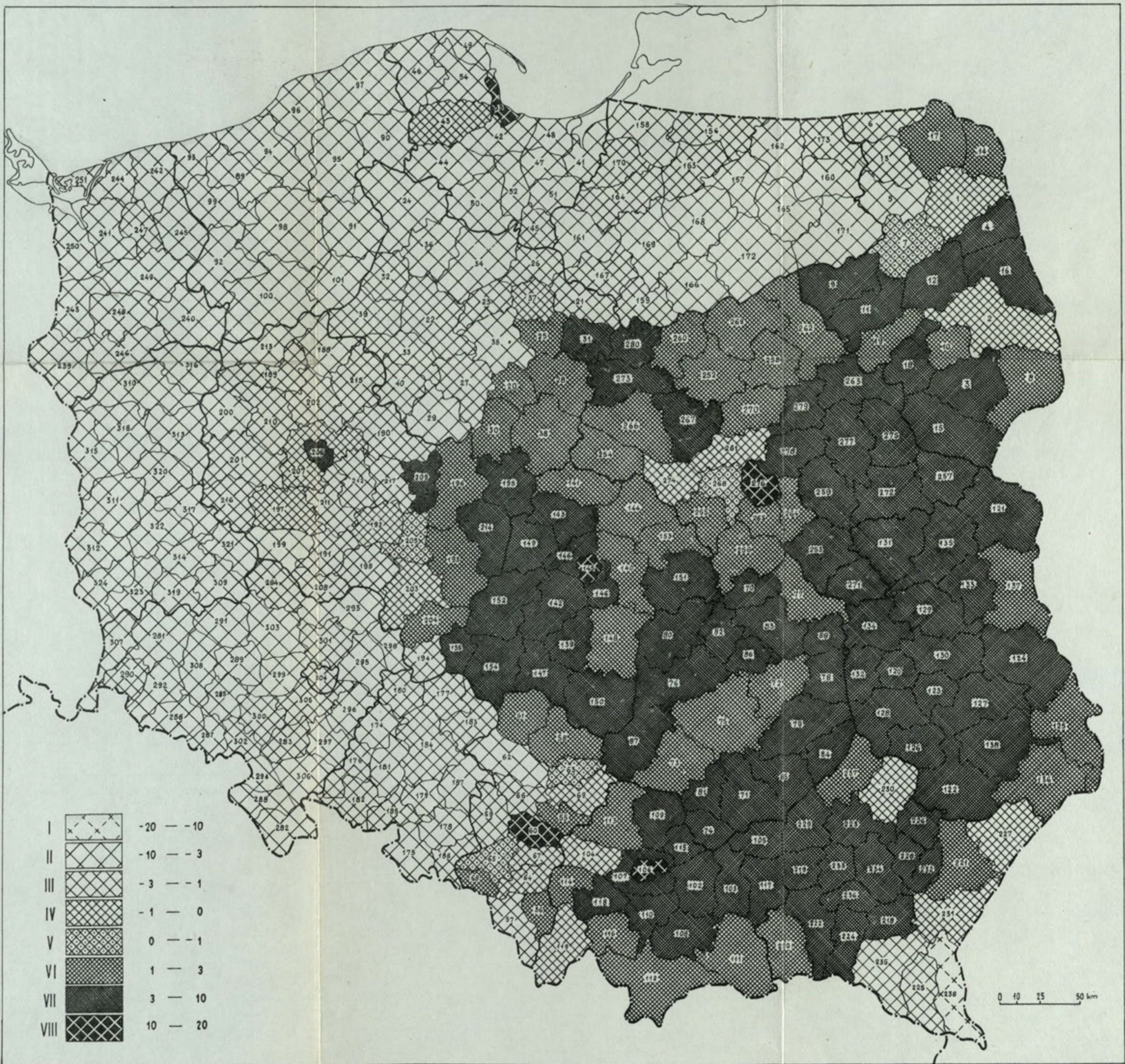
Varia

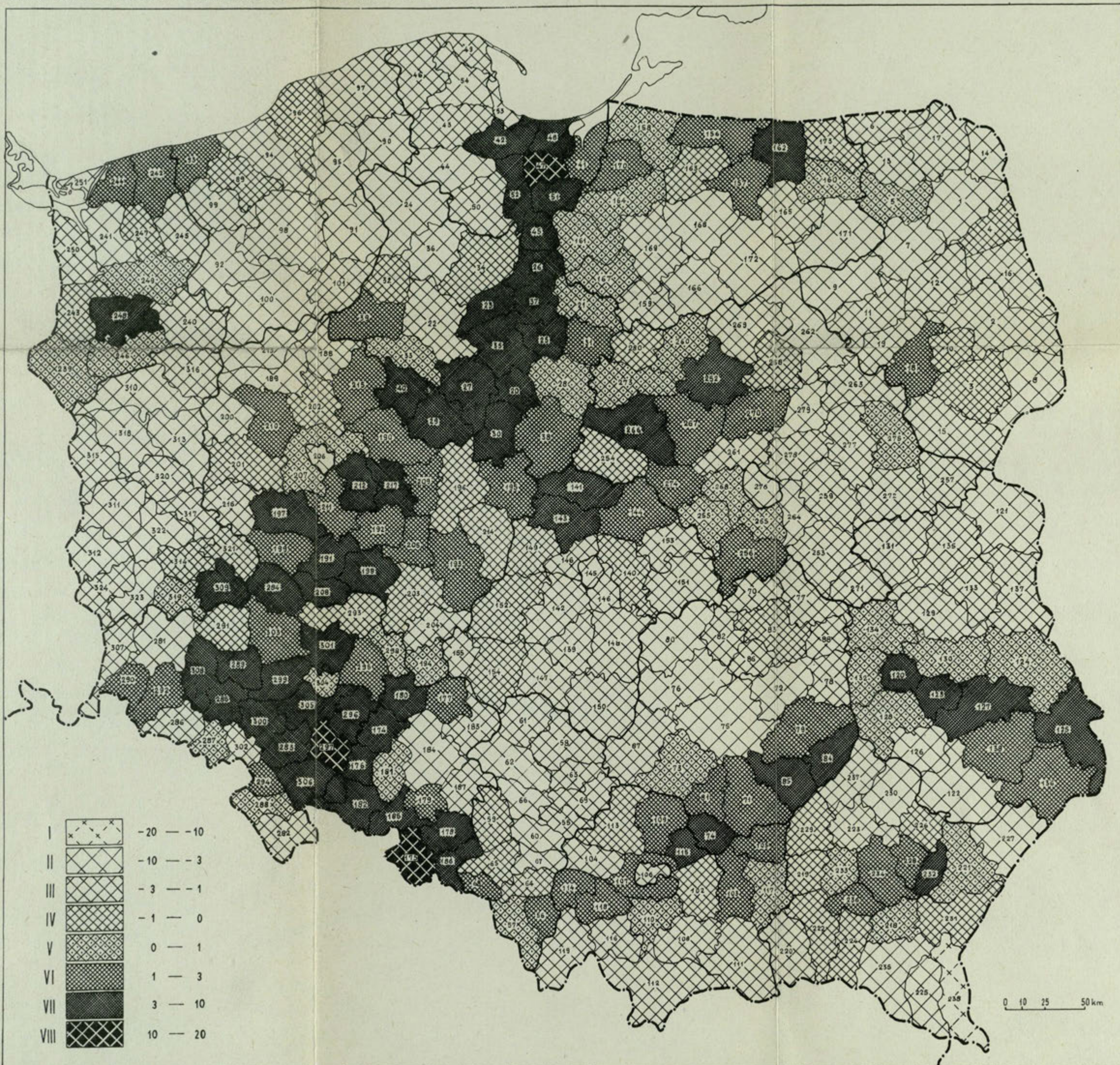
Centralny katalog zbiorów kartograficznych w Polsce

- Zeszyt 1. Katalog atlasów i dzieł geograficznych 1482—1800. 1961, s. 248, zł 72,—
- Zeszyt 2. (uzupełniający) Katalog atlasów i dzieł geograficznych 1482—1800. 1963, s. 124, zł 28,—
- Zeszyt 3. Katalog atlasów 1801—1919. 1965, s. 343, zł 76,—
- Zeszyt 4. Katalog atlasów 1920—1945. 1968, s. 160, zł 48,—

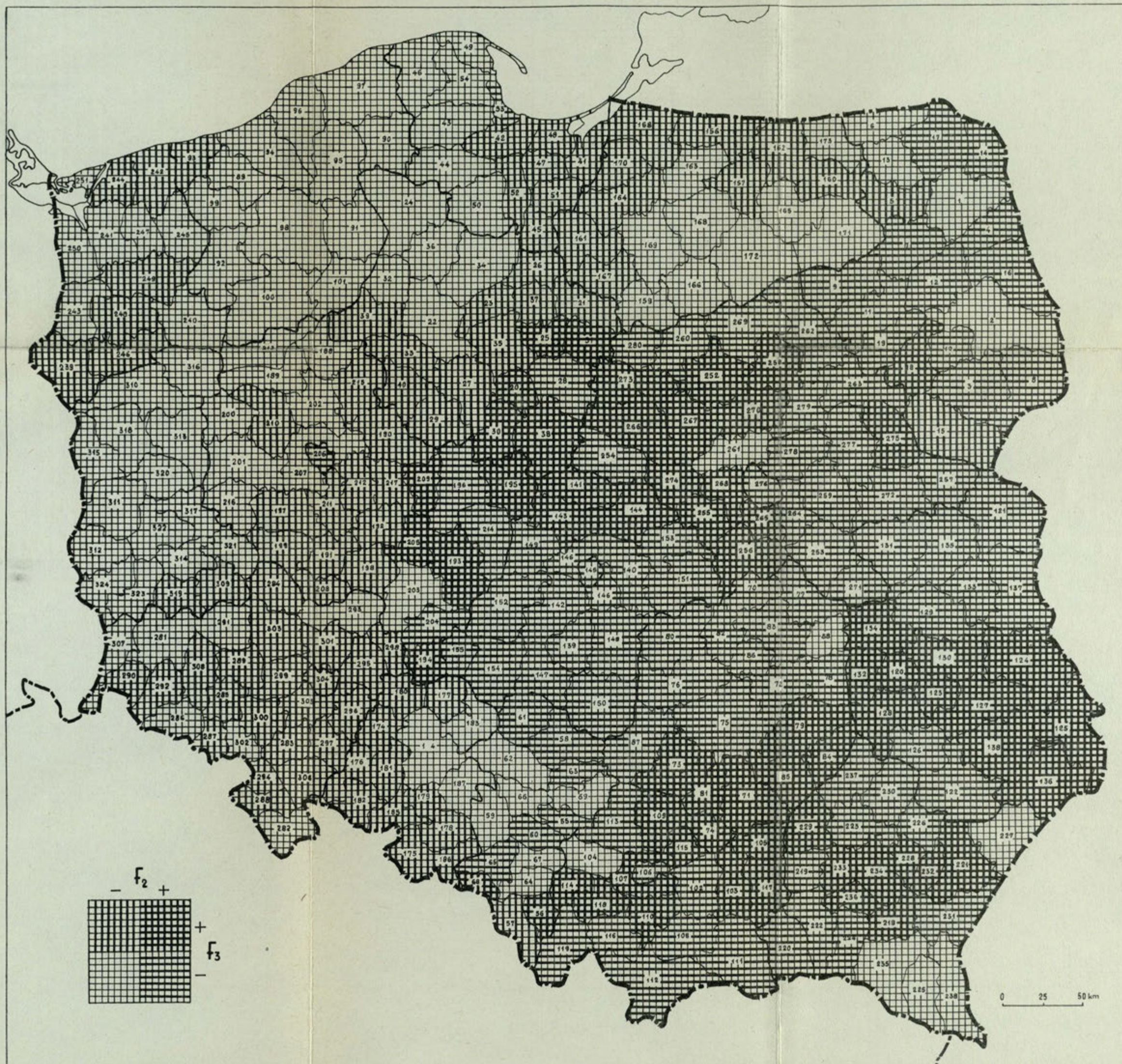




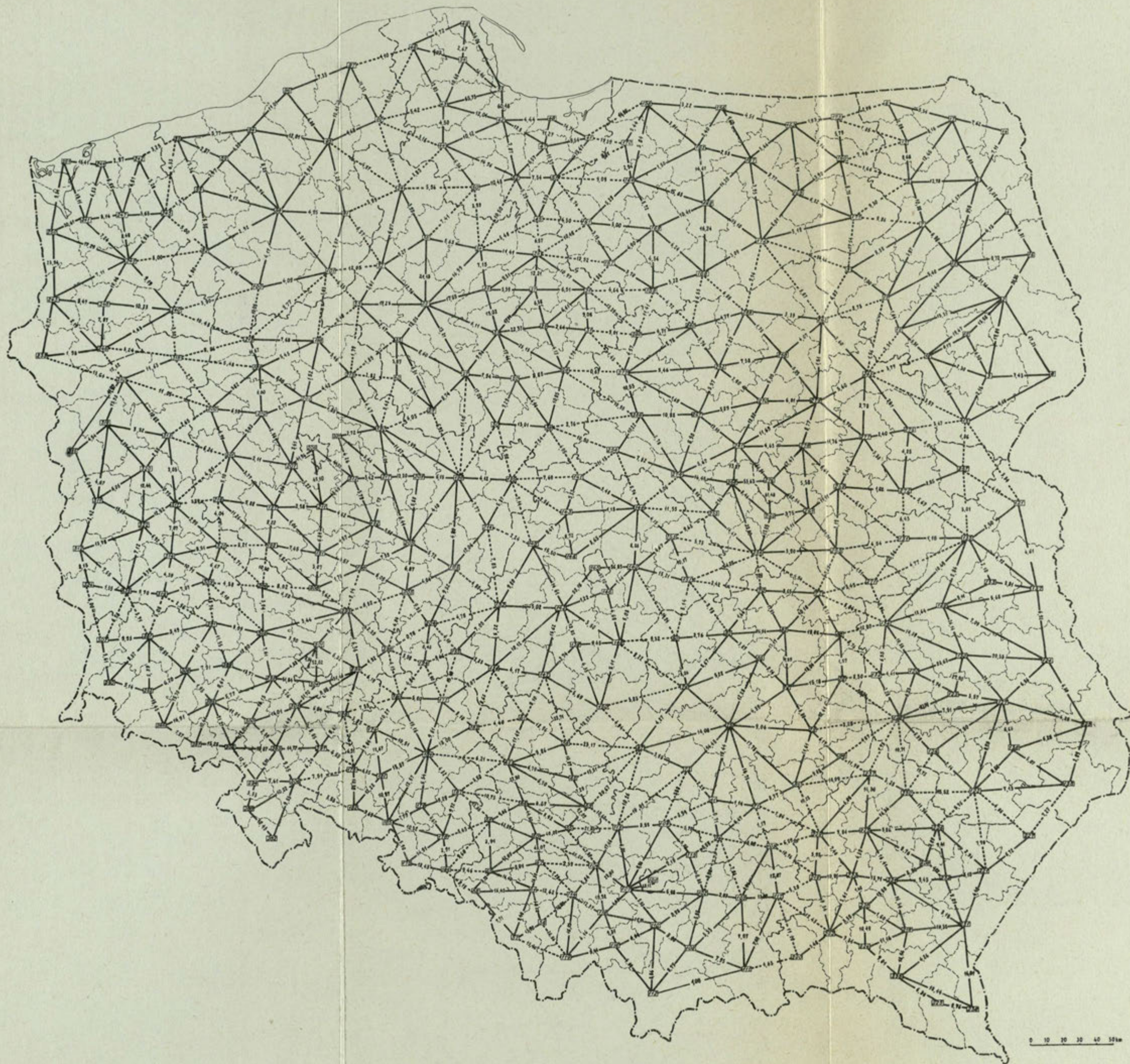


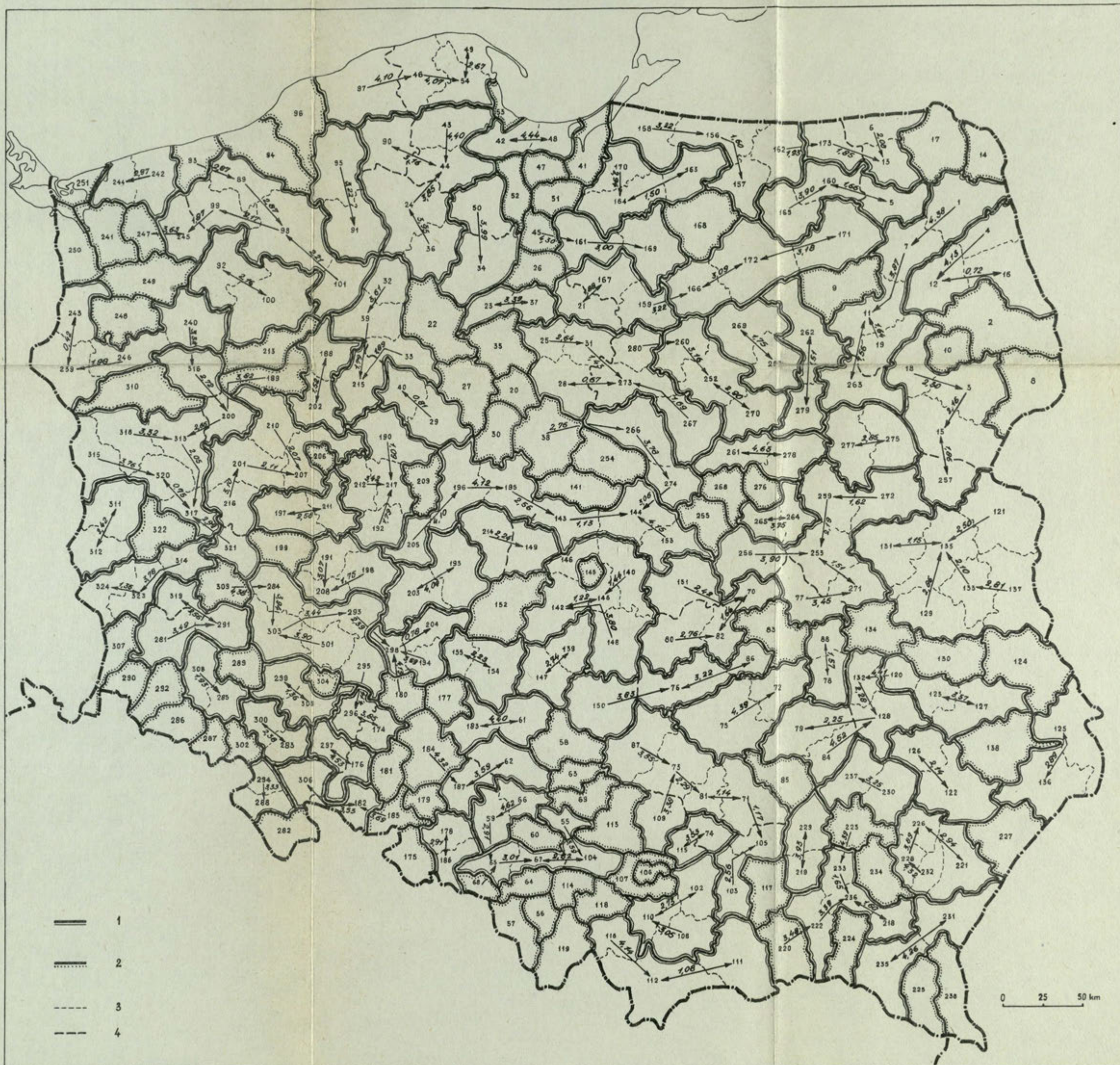


Mapa 4. Rozkład przestrzenny wartości czynnika poziomu intensywności gospodarki rolnej (F_3)
 Map 4. The spatial distribution of the score of the factor "level of intensity of agriculture economy" (F_3)

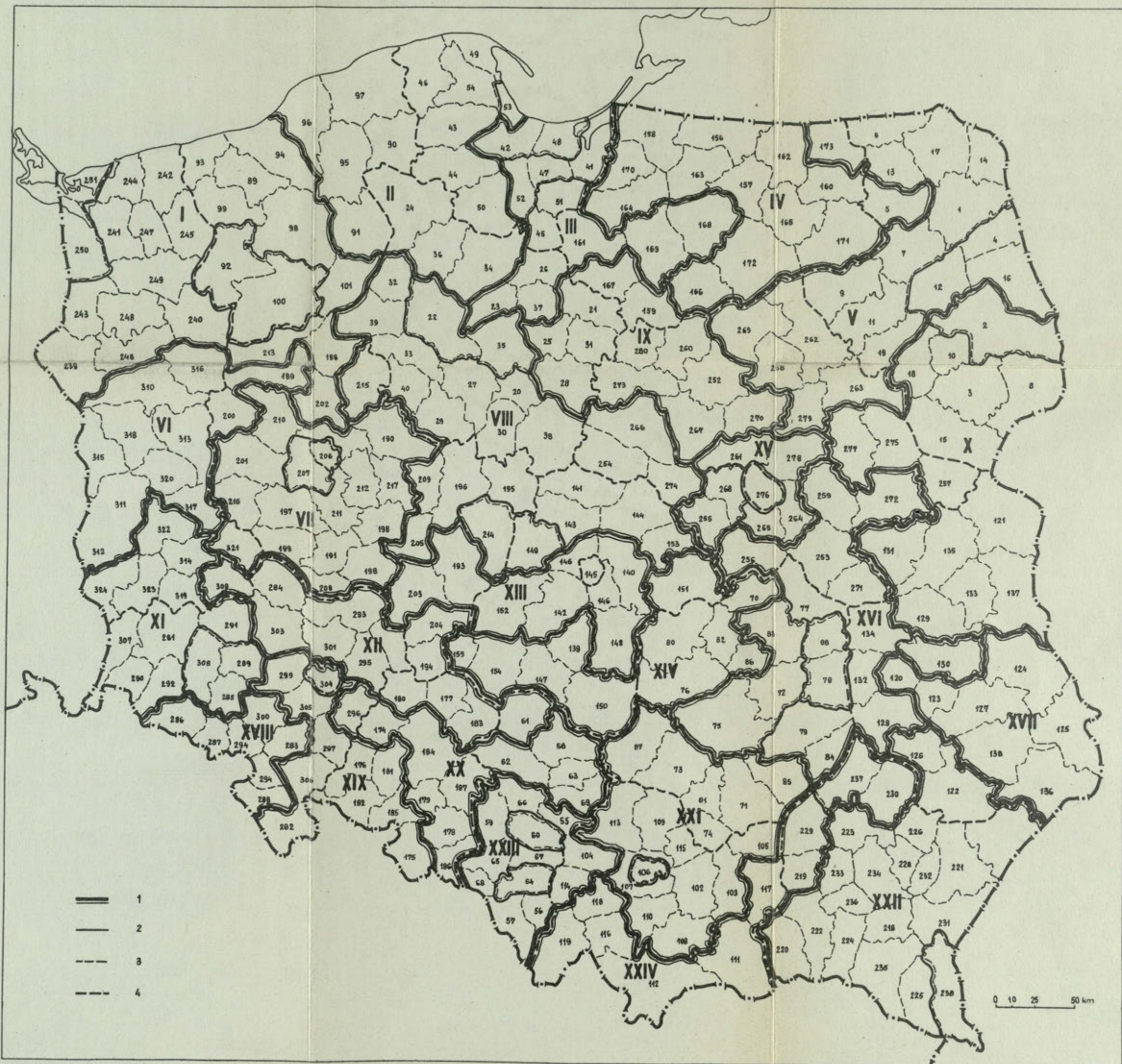












Mapa 11. Regiony ogólnoeconomiczne Polski 1965 r.

1 — granice regionu; 2 — granice mikroregionów włączonych na zasadzie arbitralnej do regionu; 3 — granice powiatu; 4 — granice województwa

Map 11. The economic regions of Poland, 1965

1 — region boundaries; 2 — microregion boundaries including to region on based of arbitrary decision; 3 — powiat boundaries; 4 — voivodship boundaries

Cena zł 38.—