

ŠTEFAN POLÁČIK

**DELIMITÁCIA MNOHOZNAKOVÝCH HOMOGENÝCH REGIÓNOV
POMOCOU NUMERICKEJ TAXONÓMIE**

Štefan Poláčik: Delimitation of multisigned homogeneous regions by means of numerical taxonomy. Geogr. Čas., 29, 1977, 1; 3 figs, 7 tables, 40 refs.

The aim of the contribution is describing the technique to delimit multi-signed regions. Methodically the paper is divided into three parts: 1. arrangement of data; 2. regional grouping; 3. setting out the specific signs of delimited regions. To demonstrate the technique, the regions of the ČSSR and 19 indices from the employment rate in the branches of socialist sector of national economy have been chosen. The data have been transformed into the quadridimensional space of components. In the regional grouping only the differences between directly neighbouring territorial units have always been reckoned. In each cycle of grouping, to set out the most similar couples of territorial units, Wrocławian taxonomy — the dendrite method — is used. A means of setting out the specific signs of the individual regions was a standardization of indices.

Rozvoj výpočtovej techniky má bezprostredný vplyv i na geografiu. Prehľubuje sa najmä metodický potenciál geografie a súčasne sa vytvárajú podmienky na rýchle a exaktné vyhodnocovanie najrozmanitejších úloh, najmä klasifikačných a regionalizačných. Zvlášť treba zdôrazniť, že v súčasnosti je použitie výpočtovej techniky na aplikáciu novších klasifikačných techník v geografii nevyhnutnosťou. Stačí, ak prihliadneme na mnohodimenzný charakter javov a procesov, ktoré pozoruje geograf.

Do geografie, ale i do ostatných geovedných disciplín, začali v 60. rokoch tohto storočia prenikať rozličné zoskupovacie metódy na tvorbu typov, ktoré boli vypracované hlavne pre potreby antropológie a psychológie [7, 12, 14]. Významné miesto medzi klasifikačnými metódami má ešte i dnes Československého metóda rozdielov a podobností [7]. Československého myšlienka kvantitatívne porovnávať určitý súbor, charakterizovaný väčším počtom znakov — ukazovateľov, a získavať tak abstraktné relácie, resp. vzdialenosti je základom i modernej numerickej taxonómie. Dnes neprestajne pribúdajú metódy numerickej taxonómie a staršie metódy sa modifikujú. Tento trend sa dá vysvetliť tak, že numerická taxonómia je úzko spojená s vedami, v ktorých sa využíva, a tak môže pružne reagovať na ich špecifické požiadavky.

V literatúre sa možno najčastejšie stretnúť s grupovacími technikami, ktoré

vypracovali Sokal a Sneath [29, 30, 31]. Okrem toho sú rozšírené postupy Warda [37], Berryho [3] a wroclawská taxonómia [14]. Zvlášť významné miesto má v literatúre z oblasti numerickej taxonómie práca [30], ktorá koncentruje všetky významné poznatky a techniky. V príspevkoch [1, 4, 18, 19] sa analyzujú a porovnávajú viaceré metódy, a to nielen po stránke metodologickej, ale aj z aspektu ich prínosu a aplikovateľnosti v rozličných odvetviach geografie.

V geografickej literatúre sa numericko-taxonomické metódy používajú veľmi často. Vyčerpávajúci prehľad o ich používaní môžeme nájsť v práci [18]. V našej geografickej literatúre sa urobil prvý pokus v práci [28]. Na používanie a rozšírenie metód numerickej taxonómie malo priaznivý vplyv i vydávanie periodík Computer application in the natural and social sciences, resp. Computer contribution, kde boli publikované programy metód pre samočinné počítače (obyčajne vo FORTRAN-e), napr. [25, 26, 33]. Boli vypublikované i knihy, ktoré pozostávali iba z programov rôznych metód používaných v geografii [23, 34]. V súčasnosti sa od publikovania programov v zahraničnej literatúre upúšťa a presadzuje sa komerčná forma poskytovania programov.

REGIONÁLNE GRUPOVANIE, JEHO CIEĽ A PRÍPRAVA DÁT

Cieľom grupovacích postupov je vytvorenie takých skupín elementov, ktoré vykazujú maximálnu podobnosť z hľadiska zvolených znakov. V geografii môžu byť podrobované grupovaniu, napr. rozličné územné jednotky. Hlavné diferenciačné znaky, podľa ktorých možno metódy rozdeliť, sú: 1. spôsob výpočtu vzdialenosti alebo rozdielov medzi elementmi, 2. algoritmus grupovania.

Na určenie vzdialeností alebo rozdielov medzi elementmi bolo navrhnutých viacero koeficientov. Najčastejšie sa stretávame s využívaním euklidovskej metriky. Vychádza sa vždy z toho, že každý skúmaný element, ktorý charakterizuje M ukazovateľov, považujeme za bod v M-rozmernom priestore ukazovateľov. Súradnicová sústava takéhoto priestoru však najčastejšie nebýva ortogonálna, pretože ukazovatele bývajú úzko skorelované, čiže vyjadrujú určitú rovnakú vlastnosť. V takomto prípade je potrebné použiť komponentnú alebo faktorovú analýzu. Docielime tak ortogonalitu súradnicových osí a súčasne sa zmenší počet premenných. Hoci faktorová či komponentná analýza môžu geografovi slúžiť i ako samostatné metódy, v tomto prípade budú slúžiť ako nástroj na úpravu dát.

Každá numericko-taxonomická metóda má svoj špecifický algoritmus, ale rovnaký cieľ konštruovať typy. V geografii, najmä v posledných desaťročiach sa presadzujú typologické postupy ako nástroje na odhalenie priestorových zákonitostí, či regionalizáciu. Tradičný je mnohoparametrický typologický prístup, keď už v roku 1943 na základe veľkého počtu ukazovateľov z obyvateľstva a poľnohospodárstva urobil H a g o o d [15] typológiu územných jednotiek. Z novších prác treba spomenúť najmä B e r r y h o príspevky [2, 3, 4, 5] a z prác iných autorov [8, 13, 27, 28].

V geografii je potrebné zaoberať sa otázkou, či je možné vytvoriť regióny, t. j. kompaktné oblasti tak, aby elementy, z ktorých by sa tieto regióny skladali, mali z hľadiska mnohých vopred zvolených ukazovateľov podobný charakter. V antropológii, psychológii či v štatistike, kde numericko-taxonomické

Tabuľka 1
Štruktúra zamestnanosti krajov ČSSR z hľadiska 19 ukazovateľov

Ukazovateľ		Stredo- český	Juho- český	Západo- český	Severo- český	Výcho- dočeský	Juho- morav.	Severo- morav.	Západo- slov.	Stredo- slov.	Výcho- doslov.
priemysel	(1)	37,382	42,734	44,209	50,612	53,187	50,125	52,419	37,668	47,375	35,883
stavebníctvo	(2)	11,016	9,094	8,075	8,835	6,984	9,491	8,932	13,137	10,835	14,256
poľnohospodárstvo	(3)	3,712	8,664	8,135	5,782	4,938	4,267	3,556	5,622	3,715	5,369
lesníctvo	(4)	0,651	3,220	2,271	1,044	1,402	1,224	1,375	1,013	3,899	3,440
doprava	(5)	5,995	5,911	6,077	6,721	5,989	5,081	6,141	6,309	5,655	7,960
spoje	(6)	2,007	1,876	1,706	1,359	1,592	1,546	1,361	1,685	1,597	1,729
obchod a verejné stravovanie	(7)	10,653	8,743	8,759	7,798	7,824	8,229	7,795	8,536	7,572	8,110
materiálno-technické zásobovanie	(8)	1,786	0,852	0,978	0,903	0,943	0,989	0,749	0,978	0,905	0,786
nákup poľnohosp. výrobkov	(9)	0,497	1,177	0,835	0,593	0,817	0,656	0,432	0,776	0,585	0,715
Ostatné výrobné odvetvia	(10)	1,141	0,159	0,248	0,241	0,174	0,262	0,313	0,924	0,199	0,167
veda a výskum	(11)	4,604	0,280	0,370	0,614	0,857	2,011	0,725	2,646	1,026	0,896
komunálne služby	(12)	3,007	1,954	2,773	2,731	2,556	2,405	2,658	2,373	2,144	2,495
bytové hospodárstvo	(13)	1,873	0,844	1,099	1,168	0,633	0,798	1,045	0,678	0,482	0,723
zdravotníctvo a social. starostlivosť	(14)	4,071	4,512	5,889	4,216	4,365	4,202	4,308	4,791	4,568	5,490
školsťvo, kultúra, osвета	(15)	6,684	6,757	5,994	5,172	5,355	6,340	6,130	8,721	6,713	8,868
správa a súdnictvo	(16)	3,140	1,987	1,649	1,355	1,418	1,361	1,239	2,603	1,635	1,880
peňažníctvo a poisťov.	(17)	0,818	0,582	0,474	0,407	0,486	0,457	0,363	0,541	0,413	0,456
spoločenské organizácie	(18)	0,541	0,295	0,241	0,216	0,236	0,311	0,271	0,690	0,503	0,616
ostatné nevýrobné odvetvia	(19)	0,413	0,352	0,208	0,223	0,235	0,236	0,179	0,300	0,171	0,151

Tabuľka 2
Matica komponentných váh

Ukazovateľ	Komp. 1	Komp. 2	Komp. 3	Komp. 4	Disp.
priemysel	-0,71047	-0,67971	-0,00063	-0,15896	0,99205
stavebníctvo	0,43062	0,71333	-0,51315	-0,07202	0,96279
poľnohospodárstvo	-0,14381	0,37624	0,80157	0,31853	0,90623
lesníctvo	-0,37849	0,69490	0,16593	-0,10032	0,66374
doprava	0,00790	0,55061	-0,32517	0,59537	0,76346
spoje	0,75651	0,35660	0,45489	-0,03099	0,90737
Obchod a verejné stravovanie	0,92241	-0,11438	0,28296	0,15536	0,96813
materiálno-technické zásobovanie	0,88808	0,35963	0,05372	0,02691	0,92164
nákup poľnohosp. výrobkov	-0,13114	0,46259	0,82976	-0,10436	0,93059
ostatné výrobné odvetvia	0,91215	-0,13113	-0,19756	-0,07410	0,89375
veda a výskum	0,90521	-0,20668	-0,26171	-0,20415	0,97230
komunálne služby	0,38717	-0,52611	-0,29602	0,68017	0,97696
bytové hospodárstvo	0,67138	-0,49442	0,07372	0,44438	0,89812
zdravotníctvo a sociálna starostlivosť	-0,15529	0,65491	0,11521	0,57986	0,80254
školsstvo, kultúra a osвета	0,36945	0,83568	-0,30760	-0,11790	0,94339
správa a súdníctvo	0,95542	0,22985	0,06074	-0,06353	0,97340
peňažníctvo a poisťovníctvo	0,92719	-0,06011	0,32304	-0,05037	0,97019
spoločenské organizácie	0,58269	0,62403	-0,45639	-0,21776	0,98467
ostatné nevýrob. odvetvia	0,78742	-0,19521	0,48169	-0,22557	0,94106
Celková disperzia	8,22808	4,57450	2,87438	1,69547	17,3724
% celkovej disperzie	43,3057	24,0763	15,1283	8,9235	91,4340

metódy vznikali, nemalo zmysel zaoberať sa podobnými problémami. V týchto vedných disciplínach sa priestorové problémy neriešili — to bolo a je úlohou geografie. Geografi metódy numerickej taxonómie preberali a chápali ich často ako regionalizačné techniky. Pri regionalizácii je však potrebné, aby sa brala do úvahy i poloha územných jednotiek. Taký postup, pri ktorom sa berie do úvahy tak podobnosť územných jednotiek, ako aj ich poloha, sa nazýva regionálne grupovanie.

Prvý pokus, kde už možno hovoriť o regionálnom grupovaní, predstavuje práca [16]. Metóda je založená na faktorovej analýze. Hagood a Price použili pritom 104 ukazovateľov z poľnohospodárstva a obyvateľstva. Vymedzovanie regiónov bolo však komplikované. Jednoduchší postup si zvolila T. Czyszová [8]. Použila wroclawskú taxonómiu, kde sa na vyjadrenie podobnosti využíva dendrit. Vypočítala dve kritické hodnoty, ktoré jej umožnili rozdeliť dendrit, a teda i celé územie na regióny a subregióny. Kritické hodnoty boli síce odrazom rozdielov územných jednotiek (jedna predstavovala aritmetický priemer vzdialeností v dendrite, druhá bola súčtom aritmetického priemeru a smerodajnej odchýlky vzdialeností v dendrite), no pri ich použití, t. j. pri určení regiónov či subregiónov dochádza k veľmi veľkej generalizácii.

Cieľom tohto príspevku je podať korektnejší postup regionálneho grupovania. Na demonštrovanie postupu boli zvolené kraje ČSSR (Praha bola začle-

Tabuľka 3
Štandardizované komponentné skóre

Kraj	Komp. 1	Komp. 2	Komp. 3	Komp. 4
Stredočeský	2,65004	-0,96507	-0,07111	0,24629
Juhočeský	-0,09171	0,78332	2,31408	-0,81928
Západočeský	-0,30721	0,10998	1,12678	1,86439
Severočeský	-0,66535	-0,98163	-0,29510	0,90837
Východočeský	-0,67717	-0,80553	0,46855	-0,25757
Juhomoravský	-0,32024	-0,75140	-0,16915	-1,08526
Severomoravský	-0,79226	-0,92551	-1,07563	0,21606
Západoslovenský	0,90993	1,02283	-0,58794	-0,78228
Stredoslovenský	-0,64088	0,48017	-0,66878	-1,37457
Východoslovenský	-0,06511	2,03285	-1,04169	1,08384

nená do Stredočeského kraja a Bratislava do Západoslovenského kraja) a 19 ukazovateľov zo zamestnanosti v nasledovných odvetviach socialistického sektoru národného hospodárstva (údaje sú za rok 1974):

1. priemysel, 2. stavebníctvo, 3. poľnohospodárstvo, 4. lesníctvo, 5. doprava, 6. spoje, 7. obchod a verejné stravovanie, 8. materiálno-technické zásobovanie, 9. nákup poľnohospodárskych výrobkov, 10. ostatné výrobné odvetvia, 11. veda a výskum, 12. komunálne služby, 13. bytové hospodárstvo, 14. zdravotníctvo a sociálna starostlivosť, 15. školstvo, kultúra a osвета, 16. správa a súdnictvo, 17. peňažníctvo a poisťovníctvo, 18. spoločenské organizácie, 19. ostatné nevýrobné odvetvia.

Tabuľka 1 predstavuje štruktúru zamestnanosti jednotlivých krajov ČSSR a je základom všetkých výpočtov. Prvým krokom bol výpočet korelačnej matice, ktorá ukázala, že medzi niektorými ukazovateľmi existuje silná korelácia ($r_{1,2} = -0,79$, $r_{2,15} = 0,93$, $r_{17,19} = 0,9$ atď.), preto bolo potrebné použiť komponentnú analýzu. Pri výpočte hlavných komponentov a skóre boli použité postupy opísané a doporučované v prácach [17, 35]. Extrahované boli 4 komponenty, ktoré mali disperziu väčšiu ako 1. Spolu majú viac ako 91 % celkovej disperzie. Tabuľka 2 zobrazuje váhy (súradnice) ukazovateľov k jednotlivým komponentom. Každý komponent nahrádza určitú kombináciu základných ukazovateľov. Z tabuľky 2 možno zistiť, že komponenty nahrádzajú kombinácie ukazovateľov takto:

Komponent 1: spoje, obchod a verejné stravovanie, materiálno-technické zásobovanie, ostatné výrobné odvetvia, veda a výskum, bytové hospodárstvo, správa a súdnictvo, peňažníctvo a poisťovníctvo, ostatné nevýrobné odvetvia,

Komponent 2: stavebníctvo, lesníctvo, zdravotníctvo a sociálna starostlivosť, školstvo, kultúra a osвета, spoločenské organizácie.

Komponent 3: poľnohospodárstvo, nákup poľnohospodárskych výrobkov.

Komponent 4: doprava, komunálne služby.

Priemysel, t. j. ukazovateľ 1 nahrádzajú komponenty 1 a 2, hoci nadobúda záporné váhy. V tabuľke 3 sú štandardizované hodnoty komponentného skóre — sú to súradnice krajov ČSSR k všetkým komponentom a teda požadované vstupné hodnoty pre grupovanie.

Aby metóda mala regionalizačný charakter, potrebné je zaviesť do výpočtu také informácie, ktoré opíšu polohu územných jednotiek v skutočnom priestore. Pre automatizáciu výpočtu na samočinnom počítači je najvhodnejšia maticová forma opisu polohy územných jednotiek. Konštrukcia matice vhodnej na tento regionalizačný postup je veľmi jednoduchá. Do riadku i a stĺpca j sa vpisuje 1, resp. 0, ak oblasti i a j majú, resp. nemajú spoločné hranice. Stredočeský kraj napr. susedí so všetkými českými krajinami, preto v matici susedstva krajov ČSSR (tab. 4) sú v prvom riadku v druhom, treťom, štvrtom a piatom stĺpci jednotky. Matica susedstva územných jednotiek urýchľuje výpočty, pretože informuje o tom, ktoré koeficienty rozdielov medzi krajinami treba počítať. Za koeficient rozdielov medzi krajinami X a Y budeme považovať

hodnotu $V_{XY} = \sum_{i=1}^4 (X_i - Y_i)^2$, kde X_1, \dots, X_4 je komponentné skóre kraja X a Y_1, \dots, Y_4 je komponentné skóre kraja Y . Hodnota V_{XY} je druhá mocnina vzdialenosti bodov X a Y v štvorrozmernom priestore. Používanie takejto hodnoty je

Tabuľka 4
Matica susedstva krajov ČSSR

Kraj	ST. Č.	J. Č.	Z. Č.	SV. Č.	V. Č.	J. M.	S. M.	Z. S.	S. S.	V. S.
ST. Č.	X	1	1	1	1	0	0	0	0	0
J. Č.	1	X	1	0	1	1	0	0	0	0
Z. Č.	1	1	X	1	0	0	0	0	0	0
SV. Č.	1	0	1	X	1	0	0	0	0	0
V. Č.	1	1	0	1	X	1	1	0	0	0
J. M.	0	1	0	0	1	X	1	1	1	0
S. M.	0	0	0	0	1	1	X	0	1	0
Z. S.	0	0	0	0	0	1	0	X	1	0
S. S.	0	0	0	0	0	1	1	1	X	1
V. S.	0	0	0	0	0	0	0	0	1	X

ST. Č. — Stredočeský kraj

J. Č. — Juhočeský kraj

Z. Č. — Západočeský kraj

SV. Č. — Severočeský kraj

V. Č. — Východočeský kraj

J. M. — Juhomoravský kraj

S. M. — Severomoravský kraj

Z. S. — Západoslovenský kraj

S. S. — Stredoslovenský kraj

V. S. — Východoslovenský kraj

praktickejšie ako jej druhej odmocniny, t. j. skutočnej vzdialenosti. Odstraňuje sa tak relatívne zdĺhavá operácia odmocňovania, ale i drobné nepresnosti pri realizácii tejto operácie. V tabuľke 5 sú všetky koeficienty rozdielov medzi krajinami ČSSR, odpovedajúce matici susedstva územných jednotiek, t. j. tab. 4. Takúto maticu budeme nazývať redukovaná matica rozdielov medzi územnými jednotkami. Na takto vytvorenú maticu aplikujeme wroclawskú taxonómiu — metódu dendritu. Dendrit vytvárame tak, že v každom riadku matice (tab. 5) nájdeme najmenšiu hodnotu a príslušné kraje pospájame líniami (obr. 1). Dendritov môže byť i viac a každý z nich podáva informáciu o dvojiciach územ-

Tabuľka 5

Redukovaná matica rozdielov (vzdialeností) územných jednotiek

Kraj	ST. Č.	J. Č.	Z. Č.	SV. Č.	V. Č.	J. M.	S. M.	Z. S.	S. S.	V. S.
ST. Č.	X	17,39	13,95	11,48	11,64					
J. Č.	17,39	X	9,11		6,58	8,64				
Z. Č.	13,95	9,11	X	4,25						
SV. Č.	11,48		4,25	X	1,97					
V. Č.	11,64	6,58		1,97	X	1,22	2,63			
J. M.		8,64			1,22	X	2,76	4,92	1,95	
S. M.					2,63	2,76	X		4,69	
Z. S.						4,92		X	3,05	
S. S.						1,95	4,69	3,05	X	8,92
V. S.								8,92		X

ST. Č. — Stredočeský kraj

J. Č. — Juhočeský kraj

Z. Č. — Západočeský kraj

SV. Č. — Severočeský kraj

V. Č. — Východočeský kraj

J. M. — Juhomoravský kraj

S. M. — Severomoravský kraj

Z. S. — Západoslávsky kraj

S. S. — Stredoslávsky kraj

V. S. — Východoslávsky kraj

ných jednotiek, ktoré majú podobné vlastnosti vzhľadom na zvolené ukazovatele a súčasne susedia. Vo zvolenom príklade sú najpodobnejšie kraje Východočeský a Juhomoravský (koeficient rozdielov je 1,22), ktoré budeme v ďalších výpočtoch považovať za jednu územnú jednotku. Súčasne je však potrebné vyjadriť túto územnú jednotku súradnicami ku komponentom. Najvhodnejšie je nahradiť grupujúce sa územné jednotky ťažiskom bodov, ktoré tieto územné jednotky reprezentujú v priestore komponentov. Súradnice ťažiska vypočítame ako aritmetický priemer stĺpcových hodnôt komponentného skóre agregujúcich sa územných jednotiek. Napríklad súradnice novej územnej jednotky, ktorá nahradí Východočeský a Juhomoravský kraj, sa vypočítajú takto:

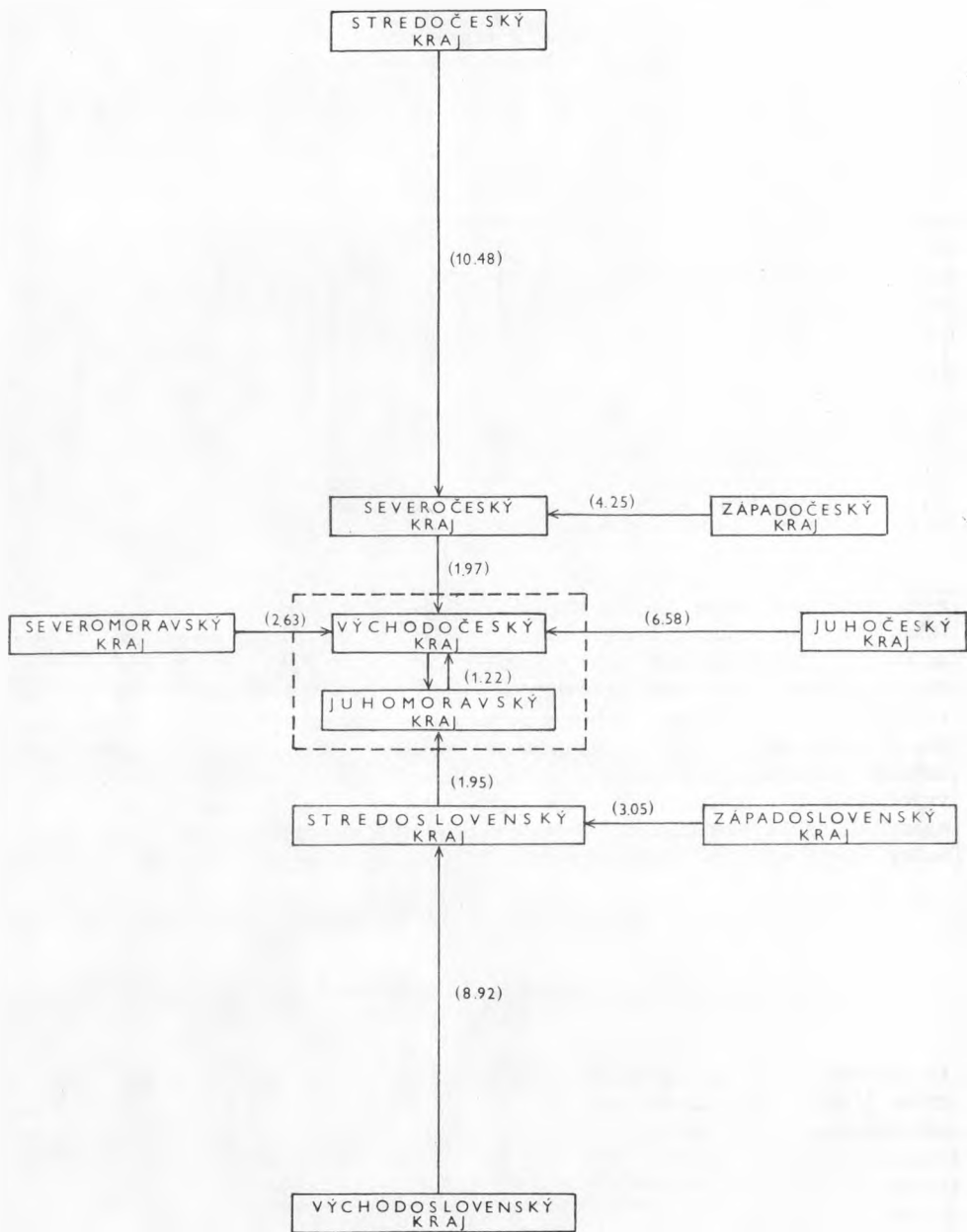
$$X_1 = \frac{[(-0,67717) + (-0,32024)]}{2} ;$$

$$X_2 = \frac{[(-0,80553) + (-0,7514)]}{2} ; \dots$$

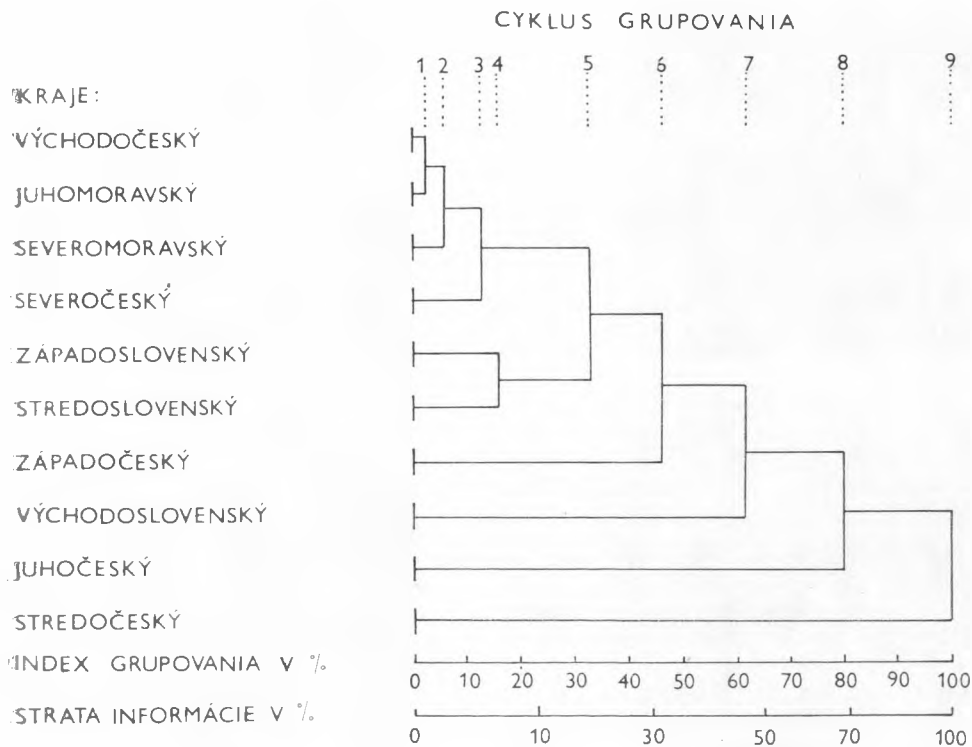
Po druhom cykle grupovacieho postupu sa ku krajom agregovaným v prvom cykle pridáva Severomoravský kraj. V tomto prípade nahrádzame až tri body ich ťažiskom, preto sa bude napr. prvá súradnica ťažiska X_1 rovnáť hodnote:

$$\frac{[(-0,67717) + (-0,32024) + (-0,79226)]}{3}$$

Ďalšie výpočty prebiehajú s agregovanou územnou jednotkou tak, že sa vypracuje nová matica susedstva územných jednotiek, vypočítava sa matica rozdielov atď. Výpočty končia splynutím všetkých územných jednotiek do jedného celku. Celý výpočet možno vyjadriť nasledovnou blokovou schémou:



Obr. 1. Dendrit prvého cyklu grupovania.



Obr. 2. Dendrogram regionálneho grupovania krajov ČSSR.

1. úprava dát
- ↓
- 2. vypracovanie matice susedstva územných jednotiek
- ↓
3. výpočet redukovanej matice rozdielov územných jednotiek
- ↓
4. konštrukcia dendritov
- ↓
5. identifikácia dvojíc najpodobnejších územných jednotiek
- ↓
6. nahradenie agregujúcich sa územných jednotiek ťažiskom bodov, ktoré ich reprezentujú v priestore komponentov

GRAFICKÉ ZNÁZORNENIE GRUPOVANIA A INDEX GRUPOVANIA

Vhodnou grafickou formou na zobrazenie grupovania, a teda i podobnosti medzi územnými jednotkami je dendrogram (obr. 2). Konštrukcia dendrogramu je jednoduchá. Obyčajne do vertikálneho smeru sa zapisujú grupované elementy a do vodorovného smeru etapy (cykly) grupovania. Dendrogram

Tabuľka 6

Hodnoty indexu grupovania a straty informácie

Cyklus grupovania	Index grupovania				Strata informácie			
	absolútne	relatívne	Kumulovane		absolútne	relatívne	Kumulovane	
			absolútne	relatívne			absolútne	relatívne
1	1	2,222	1	2,222	1,222	0,305	1,222	0,305
2	2	4,444	3	6,666	5,404	1,351	6,626	1,656
3	3	6,666	6	13,333	7,242	1,810	13,868	3,467
4	1	2,222	7	15,555	3,056	0,764	16,924	4,231
5	8	17,777	15	33,333	47,881	11,970	64,805	16,202
6	6	13,333	21	46,666	56,386	14,097	121,191	30,299
7	7	15,555	28	62,222	69,514	17,379	190,705	47,678
8	8	17,777	36	80,000	89,079	22,271	279,784	69,950
9	9	20,000	45	100,000	120,193	30,049	399,977	100,000

skonštruovaný uvedenou metódou má takú vlastnosť, že ak ho rozdelíme na ľubovoľnej úrovni, dostaneme vždy kompaktné oblasti, vnútorne homogénne. Pri prvom delení dendrogramu na dva regióny vystupuje samostatne Stredočeský kraj. Druhý región sa skladá z ostatných krajov ČSSR. Pri druhom delení na 3 regióny vystupujú samostatne už dva kraje: Stredočeský a Juhočeský. Ostatné kraje vytvárajú tretí región. Pri poslednom deviatom delení dostaneme tzv. triviálne delenie, kde každá územná jednotka stojí samostatne.

B. J. L. Berry zaviedol v práci [3] pojem strata informácie. Pri každom kroku grupovania totiž strácame informácie o základnom súbore, pretože sa používajú určité generalizované vyjadrenia agregovaných elementov. Strata informácie sa používa na určenie optimálneho rozdelenia, keďže ju možno považovať i za mieru homogenity. Výpočet koeficientov straty informácie je pomerne zložitý. Jednoduchšie je možné vypočítavať index grupovania, ktorý taktiež dobre poslúži na optimálne rozdelenie dendrogramu, resp. územia. Existujú dva spôsoby výpočtu indexu grupovania. Prvý spôsob je taký istý ako v práci [3], avšak s tým rozdielom, že všetky prvky matice vzdialeností (matice rozdielov) budú sa rovnať jednej. Druhý spôsob je jednoduchší. Za index grupovania považujeme pri istom kroku grupovania také číslo, ktoré je súčinom hodnôt vyjadrujúcich počty grupujúcich sa elementov. Napríklad pri deviatom cykle grupovania sa agreguje jeden kraj s ostatnými deviatimi, preto je index grupovania deviateho cyklu rovný súčinu hodnôt 1 a 9. Pri piatom cykle sa agregujú dva kraje (Západoslovenský a Stredoslovenský) so 4 kraji (Severočeským, Východočeským, Severomoravským a Juhomoravským). Index grupovania pri piatom cykle grupovania sa rovná súčinu hodnôt 2 a 4, t. j. 8. Na porovnanie sú v tabuľke 6 uvedené tak hodnoty indexu grupovania, ako aj koeficienty straty informácie. Z kumulovaných hodnôt indexu grupovania môžeme určiť optimálne delenie súborov. Najväčší skok badať v piatom kroku grupovania. Aby sme zamedzili zníženiu homogenity regiónov, musíme územie rozdeliť po štvrtom cykle grupovanie. Územie ČSSR sa tak rozpadne na 6 regiónov, homogénnych z aspektu 19 zvolených ukazovateľov. Sú to nasledovné regióny, znázornené i na priloženej mape:

R—I Severočeský, Východočeský, Severomoravský a Juhomoravský kraj,
 R—II Juhočeský kraj,
 R—III Západočeský kraj,



Obr. 3. Homogénne regióny ČSSR [vysvetlenie v texte].

Tabuľka 7
Špecifické znaky regiónov

Ukazovateľ	R — I	R — II	R — III	R — IV	R — V	R — VI
bytové hospodárstvo	-0,219	0,217	-0,393	2,072	-0,993	-0,683
komunálne služby	0,191	0,767	-1,655	1,459	-0,707	-0,055
zdravotníctvo a soc. starostlivosť	-0,843	1,632	-0,472	-1,147	-0,192	1,023
lesníctvo	-0,926	0,088	1,049	-1,551	0,068	1,272
priemysel	1,825	0,383	0,093	-0,961	-0,084	-1,256
poľnohospodárstvo	-0,734	1,218	1,501	-1,153	-0,567	-0,264
nákup poľnohosp. výrobkov	-0,684	0,376	1,962	-1,195	-0,279	-0,179
obchod a verejné stravovanie	-0,858	0,043	0,025	2,098	-0,647	-0,661
peňažníctvo a poisťov. spoje	-0,870	-0,491	0,316	2,082	-0,405	-0,631
ostatné nevýrobné odvetvia	-1,612	-0,182	0,803	1,563	-0,524	-0,047
materiálno-technické zásobovanie	-0,537	-0,621	0,976	1,655	-0,213	-1,259
ostatné výrobné odvetvia	-0,442	-0,181	-0,554	2,198	-0,273	-0,745
veda a výskum	-0,494	-0,518	-0,771	2,015	0,517	-0,748
správa a súdnictvo	-0,289	-0,787	-0,848	2,075	0,281	-0,431
spoločenské organizácie	-1,232	-0,674	-0,075	1,967	0,279	-0,264
stavebníctvo	-0,996	-1,142	-0,813	0,690	1,112	1,151
školsťvo, kultúra a osveta	-0,855	-1,140	-0,670	0,216	0,739	1,710
doprava	-1,092	-0,950	-0,230	-0,299	0,810	1,761
	-0,537	-0,322	-0,547	-0,433	-0,388	2,229

R — I Severočeský kraj, Východočeský kraj, Juhomoravský kraj, Severomoravský kraj
R — II Juhočeský kraj
R — III Západočeský kraj
R — IV Stredočeský kraj
R — V Západoslonský kraj, Stredoslonský kraj
R — VI Východoslonský kraj

R—IV Stredočeský kraj,
R—V Západoslonský a Stredoslonský kraj,
R—VI Východoslonský kraj.

CHARAKTERISTIKA VYMEDZENÝCH REGIÓNOV

Získané regióny sú odrazom ich podobnosti z hľadiska 19 ukazovateľov. Nemožno sa teda nazdávať, že sa dajú opísať pomocou jedného ukazovateľa. Každý región však má svoju špecifickú kombináciu ukazovateľov. Môžeme ju získať tak, že vypočítame štruktúru všetkých šiestich regiónov z hľadiska zvolených 19 ukazovateľov. Aby boli ukazovatele porovnateľné, štandardizujeme ich a dostaneme tabuľku, z ktorej ľahko zistíme špecifickú kombináciu uka-

zovateľov každého regiónu. Často je možné upraviť túto tabuľku do schodovitého tvaru tak, ako je to v tabuľke 7. Jednotlivé regióny majú nasledovné špecifické kombinácie znakov, zoradené podľa dôležitosti:

R—I 1. priemysel, 2. komunálne služby,

R—II 1. zdravotníctvo a sociálna starostlivosť, 2. poľnohospodárstvo, 3. komunálne služby, 4. priemysel, 5. nákup poľnohospodárskych výrobkov, 6. bytové hospodárstvo, 7. lesníctvo, 8. obchod a verejné stravovanie,

R—III 1. nákup poľnohospodárskych výrobkov, 2. poľnohospodárstvo, 3. lesníctvo, 4. ostatné nevýrobné odvetvia, 5. spoje, 6. peňažníctvo a poisťovníctvo, 7. priemysel, 8. obchod a verejné stravovanie,

R—IV. 1. materiálne-technické zásobovanie, 2. obchod a verejné stravovanie, 3. peňažníctvo a poisťovníctvo, 4. veda a výskum, 5. bytové hospodárstvo, 6. ostatné výrobné odvetvia, 7. správa a súdníctvo, 8. ostatné nevýrobné odvetvia, 9. spoje, 10. komunálne služby, 11. spoločenské organizácie, 12. stavebníctvo,

R—V 1. spoločenské organizácie, 2. školstvo, kultúra a osвета, 3. stavebníctvo, 4. ostatné výrobné odvetvia, 5. veda a výskum, 6. správa a súdníctvo, 7. lesníctvo,

R—VI 1. doprava, 2. školstvo, kultúra a osвета, 3. stavebníctvo, 4. lesníctvo, 5. spoločenské organizácie, 6. zdravotníctvo a sociálna starostlivosť.

ZÁVER

Hoci v posledných desaťročiach sa v geografickej literatúre mnoho písalo o regionalizácii, iba sporadicky sa v príspevkoch objavujú návrhy na delimitáciu regiónov, najmä komplexných — mnohoznakových. Práca má metodický charakter. Zamerali sme sa na výklad metódy, ktorou je možné dospieť k homogénnym mnohoznačným regiónom. Tento postup možno využiť všade tam, kde sa s úspechom uplatňuje faktorová, resp. komponentná analýza — najvhodnejší prostriedok na prípravu vhodných dát. Nevýhodou metódy (ako temer pri všetkých metódach numerickej taxonómie) je jej náročnosť na počet operácií. Najmä pri väčšom počte územných jednotiek je nevyhnutné použitie samočinného počítača. Výpočty pre tento článok boli pripravené na počítači WAWG 2200 inštalovanom na Geografickom ústave SAV.

LITERATÚRA

1. BEAVON, K. S. O., HALL, A. V.: A geotaxonomic approach to classification in urban and regional studies. *Geogr., Analysis* 4, 1972, 4, s. 407—415. — 2. BERRY, B. J. L.: A note concerning methods of classification. *Ann. Ass. Americ. Geogr.*, 48, 1958, s. 300—303. — 3. BERRY, B. J. L.: A method for deriving multi-factor uniform regions. *Przeegl. Geogr.*, 33, 1961, 2, s. 263—282. — 4. BERRY, B. J. L.: Grouping and regionalization: an approach to the problem using multivariate analysis. *Quant. Geogr.* [GARRISON, W. I., MARBLE, D. F.—ed], 13, 1967, Northwestern University, Ewanston. — 5. BERRY, B. J. L., RAY, D. M.: Multivariate socioeconomic regionalization: a pilot study in Central Canada. *Regional Studies* [RYMES, T., OSTRY, S.—ed], Toronto 1966. — 6. BLAŽEK, M.: Analyse géographique de la regionalisation administrative. *Economic Regionalization*. Academia, Praha 1967, s. 219—236. — 7. CZEKANOWSKI, J.: Metoda podobienstw w zastosowaniu do antropologii. *Prace Towarzystwa Naukowego Warszaw-*

skiego, 5, 1926. — 8. CZYZ, T.: Zastosowanie metody analizy czynnikowej do badania ekonomicznej struktury regionalnej Polski. IG PAN, Prace Geograficzne, 92, Wrocław 1971. — 9. DACEY, M. F.: Improved linkage analysis by a nearest neighbor method. Discussion paper, 9, 1958, Department of geography. Univ. of Washington. — 10. DIXON, W. J.: BMD Biomedical computer programs. Univ. of California, Los Angeles 1968, 600 strón.

11. DOMANSKI, R.: Procedura typologiczna w badaniach ekonomiczno-geograficznych. Przgl. Geogr., 36, 1964, 4, s. 627—655. — 12. EDWARDS, A. W. F., CAVALLI-SFORZA, L. L.: A method for cluster analysis. Biometrics, 21, 1963, s. 362—375. — 13. ERNST, J.: Niektóre metody okreslania regionow geograficzno-rolniczych. Annales UMCS, Sectio B, 21, 1966, 1, s. 1—26. — 14. FLOREK, K., LUKASZEWICS, J., PERKAL, J., STEINHAUS, H., ZUBRZICKI, S.: Taksonomia Wroclawska. Przgl. Antropolog., 17, 1951, s. 193—211. — 15. HAGOOD, M. J.: Statistical methods for delimitation of regions applied to data on agriculture and population. Soc. Forc., 21, 1943, s. 287—297. — 16. HAGOOD, M. J., PRICE, D. O.: Statistics for sociologist. New York 1957. — 17. HARMAN, H. H.: Modern factor analysis. Chicago 1960. — 18. CHOJNICKI, Z., CZYZ, T.: Metody taxonomii numerycznej w regionalizacji geograficznej. Panstwowe Wydawnictwo Naukowe, Warszawa 1973, 103 strón. — 19. JOHNSTON, R. J.: Grouping and regionalizing: some methodical and technical observations. Econom. Geogr., 46, 1970, 2, s. 293—305. — 20. LANKFORD, F. M.: Regionalization: theory and alternative algorithms. Geograph. Analys., 1, 1969, s. 196—212.

21. LANCE, G. N., WILLIAMS, W. T.: Computer program for hierarchical polythetic classification (Similarity analysis). Computer J., 9, 1966, s. 60—64. — 22. MATHER, P. M.: Cluster analysis. Computer Applications in the Natural and Social Sciences, 1. Nottingham 1969. — 23. MARBLE, D. F. (ed): Some computer programs for geographic ressearch. Department of geography, Northwestern Univ., Ewanston 1967. — 24. MARSZ, A.: Próba regionalizacji fizyczno-geograficznej wyspy Wolin. Prace fizyograficzne nad Polska zachodnia, 17, 1966, s. 59—108. — 25. Mc CAMMON, R. B., WENINGER, C.: The dendrograph. Computer Contrib., 48, 1970. — 26. PARKS, J. M.: FORTRAN IV program for Q—mode cluster analysis on distance function with printed dendrogram. Computer Contrib., 46, 1970. — 27. POCOCK, D. C. D., WISHART, D.: Methods of deriving multifactor uniform regions. Transactions of the Institute of British Geographers, 47, 1969, s. 73—98. — 28. POLÁČIK, Š.: Regionálne typy stredného a časti dolného Ponitria. Acta geograph. UC, 11, 1972, s. 171—199. — 29. SNEATH, P. H. A.: The application of computers to taxonomy. J. gener. Mikrobiol., 17, 1957, s. 201—206. — 30. SNEATH, P. H. A., SOKAL, R. R.: Numerical taxonomy. San Francisco 1973.

31. SOKAL, R. R., SNEATH, P. H. A.: Principles of numerical taxonomy. San Francisco 1963. — 32. SPENCE, N. A., TAYLOR, P. J.: Quantitative methods in regional taxonomy. Progr. in Geogr., 2, 1971, s. 3—64. — 33. TAYLOR, A. H.: Group average classifications. Computer Application in the Natural and Social Sciences, 12, 1971, s. 10—25. — 34. TOBLER, W. (ed): Selected computer programs. Department of Geography, University of Michigan, Ann Arbor 1970, 162 strón. — 35. ÜBERLA, K.: Faktorenanalyse. Springer Verlag, Berlin 1971. — 36. VIELROSE, E.: O porownywaniu dendrytow w taxonomii wroclawskej. Przgl. Statyst., 14, 1967, 1, s. 13—19. — 37. WARD, J. H.: Hierarchical grouping to optimize and objective function. J. Amer. Statist. Ass., 58, 1965, s. 236—244. — 38. WILLIAMS, G. N., LANCE, W. T.: A general theory of classificatory sorting strategies; 1. Hierarchical methods. The Computer J., 9, 1967, 4, s. 373—380. — 39. WILLIAMS, G. N., LANCE, W. T.: A general theory of classificatory sorting strategies; 2. Clustering systems. The Computer J., 10, 1967, 3, s. 271—276. — 40. WISHART, D.: FORTRAN II programs for 8 methods of cluster analysis [CLUSTAN 1]. Computer Contrib., 38, 1969.

DELIMITATION VIELMERKMALIGER HOMOGENER REGIONEN MITTELS DER
 NUMMERISCHEN TAXONOMIE

Hagoods erster Versuch, in Arbeit (15), komplexe Typen territorialer Einheiten zu schaffen, wurde lange nicht nachgeahmt. Erst Berrys Arbeiten (2, 3) haben diesen Trend belebt. Die Methoden der numerischen Taxonomie sind heute die meistverbreiteten Typisierungsmethoden. Sie ermöglichen die Schaffung von Typen aus Elementen, die durch eine grosse Menge von Kennziffern charakterisiert sind. In der Geographie können sie jedoch auch für Regionalisierungszwecke ausgenützt werden. Die für Regionalisierungszwecke bestadaptierte Methode ist die Wroclawer Taxonomie.

Für zehn Gebietseinheiten — Kreise der ČSSR (Tschechoslowakischen Sozialistischen Republik) haben wir 19 Kennziffern aus der Beschäftigung in Zweigen des sozialistischen Sektors der Volkswirtschaft ausgewählt. Es sind folgende Kennziffern: 1. Industrie; 2. Bauwesen; 3. Landwirtschaft; 4. Forstwirtschaft; 5. Verkehrswesen; 6. Kommunikationswesen; 7. Handel und öffentliche Verköstigung; 8. Material-technische Versorgung; 9. Einkauf landwirtschaftlicher Erzeugnisse; 10. übrige Produktionszweige; 11. Wissenschaft und Forschung; 12. Kommundaldienste; 13. Wohnungswirtschaft; 14. Gesundheitswesen und soziale Fürsorge; 15. Schulwesen, Kultur und Volksbildung; 16. Verwaltung und Gerichtswesen; 17. Finanzwesen und Versicherungswesen; 18. Gesellschaftliche Organisationen; 19. übrige Nichtproduktionszweige.

Aus dem methodischen Aspekt ist die Arbeit in drei Teile geteilt: 1. Datenverarbeitung; 2. regionale Gruppierung; 3. Festsetzung spezifischer Merkmale abgegrenzter Regionen.

1. Für jeden Kreis wurde die Beschäftigungsstruktur aus dem Aspekt von 19 Kennziffern ausgerechnet (Tab. 1). Manche Paare von Kennziffern waren in enger Korrelation. Aus diesem Grund haben wir uns entschlossen die Komponentenanalyse zu benutzen. In Tab. 2 sind die Komponentenwaagen von 19 Kennziffern zu vier Komponenten. Tab. 3 enthält standardisierte Werte des Komponentenscore der Kreise der ČSSR, d. h. Eintrittsdaten für regionale Gruppierung.

2. Der erste Schritt beim Verfahren der regionalen Gruppierung ist die Ausarbeitung einer Matrix der Nachbarschaft territorialer Einheiten (für die Kreise der ČSSR ist es Tab. 4). Die Werte 0 bzw. 1 in der Nachbarschaftsmatrix bedeuten, dass Kreispaare haben nicht bzw. haben gemeinsame Grenzen.

Der zweite Schritt ist die Berechnung der Entfernungen (Unterschiede) V^2_{XY} nach der Formel $V^2_{XY} = \sum_{j=1}^4 (X_j - Y_j)^2$, wo X_1, \dots, X_4 standardisierte Werte des Komponentenscore des Kreises X und Y_1, \dots, Y_4 standardisierte Werte des Komponentenscore des Kreises Y sind. Die Koeffizienten V^2_{XY} werden jedoch nur für benachbarte Kreise berechnet. In der reduzierten Matrix der Entfernungen (Tab. 5) müssen jene Räume ausgefüllt werden, welche in der Nachbarschaftsmatrix territorialer Einheiten gleich 1 sind.

Der dritte Schritt des Verfahrens ist die Schaffung von Dendriten. Im gewählten Beispiel haben wir einen Dendrit bekommen (Abb. 1). Es kann auch mehrere geben, vor allem bei grösserer Anzahl von territorialen Einheiten. Ein jeder Dendrit hilft Paare von territorialen Einheiten zu identifizieren, welche aneinander grenzen und relativ die ähnlichsten Eigenschaften haben. Die festgestellten Paare können sich im Gruppierungsprozess vereinen. Beim ersten Gruppierungszyklus haben der Ostböhmisches und Südmährische Kreis die ähnlichsten Eigenschaften. Aggregierte territoriale Einheiten ersetzen wir durch den Schwerpunkt der Punkte, welche diese territorialen Einheiten im Raum der Komponenten repräsentieren. Beim ersten Gruppierung

rungszyklus wird z. B. der Schwerpunkt von zwei Punkten ausgerechnet, beim zweiten Zyklus der Schwerpunkt von drei Punkten usw. Weiter verläuft das Gruppierungsverfahren durch den ersten Schritt, d. h. durch die Konstruktion der Nachbarschaftsmatrix territorialer Einheiten, wobei wir die Paare aggregierter Kreise für eine territoriale Einheit halten. Die Gruppierung endet, wenn alle territorialen Einheiten sich in eine Gesamtheit vereinigen. Die Gruppierung kann am besten mit Hilfe eines Dendrogramms dargestellt werden (Abb. 2).

Der Vorteil der zyklischen regionalen Gruppierung ist, dass die Begrenztheit der Informationen, welche die reduzierte Matrix der Entfernungen (Unterschiede) bieten kan, teilweise eliminiert wird.

Beim Informationsverlust, welchen Berry (3) eingeführt hat, ist ein guter Koeffizient der Gruppierungsindex, der bei der Gruppierung die Homogenität charakterisiert. Der Gruppierungsindex beim k - ten Zyklus ist das Produkt zweier Werte, welche die Anzahl der gruppierenden Elemente bei demselben Zyklus darstellen. Z. B. beim neunten Gruppierungszyklus aggregiert ein Kreis (Mittelböhmen) mit den übrigen neun Kreisen. Der Gruppierungsindex des Zyklus $k = 9$ gleicht dem Produkt $1 \times 9 = 9$. Für Vergleichung werden in Tab. 6 Werte des Gruppierungsindex, aber auch der Verlustkoeffizient der Information angeführt.

3. Aufgrund des Anstiegs des Gruppierungsindex wurden 6 homogene Regionen abgegrenzt. Für jede Region wurde die Struktur berechnet und alle Kennziffern wurden standardisiert. Durch Bearbeitung haben wir Tab. 7 gewonnen, aus der es möglich ist spezifische Kombinationen einzelner Regionen festzustellen. Das Bild der Territoriums nach der Aufteilung in Regionen stellt die beliegende Karte dar.

Für einzelne Regionen sind folgende Kombinationen der Kennziffern spezifisch: Region R—I (der Nordböhmisches Kreis, der Ostböhmisches Kreis, der Nordmährisches Kreis, der Südmährisches Kreis) Kennziffern: 1, 12.

Region R—II (der Südböhmisches Kreis): 14, 3, 12, 1, 9, 13, 4 und 7.

Region R—III (der Westböhmisches Kreis): 9, 3, 4, 19, 6, 17, 1 und 7.

Region R—IV (der Mittelböhmisches Kreis): 8, 7, 17, 11, 13, 10, 16, 19, 6, 12, 18, 2.

Region R—V (der Westslowakisches Kreis, der Mittelslowakisches Kreis): 18, 15, 2, 10, 11, 16, 4.

Region R—VI (der Ostslowakisches Kreis): 5, 15, 2, 4, 18 und 14.

Abb. 1. Dendrit des ersten Gruppierungszyklus.

Abb. 2. Dendrogramm regionaler Gruppierung der Kreise der ČSSR.

Abb. 3. Homogene Regionen der ČSSR.

Tab. 1. Beschäftigungsstruktur der Kreise der ČSSR aus dem Aspekt von 19 Kennziffern.

Tab. 2. Matrix der Komponentenwaagen.

Tab. 3. Standardisiertes Komponentenscore.

Tab. 4. Nachbarschaftsmatrix der Kreise der ČSSR.

Tab. 5. Reduzierte Matrix der Unterschiede (Entfernungen) territorialer Einheiten.

Tab. 6. Werte des Gruppierungsindex und des Informationsverlustes.

Tab. 7. Spezifische Merkmale der Regionen.

Übersetzt von A. Miš í k o v á

ВЫДЕЛЕНИЕ МНОГОПРИЗНАКОВЫХ ГОМОГЕННЫХ РЕГИОНОВ МЕТОДОМ НУМЕРИЧЕСКОЙ ТАКСОНОМЕТРИИ

Первая попытка Хагуда создать комплексные типы территориальных единиц в работе [15] долго не находила последователей. Только работы Берри [2, 3] продолжили этот тренд. Методы нумерической таксонометрии в настоящее время являются самыми распространенными методами типизации. Они позволяют выделять типы из элементов, характеризованных большим количеством показателей. В географии их можно также применить в целях районизации. Наиболее адаптируемым для районизации является вроцлавский метод.

Для десяти территориальных единиц — областей ЧССР мы избрали 19 показателей занятости в отраслях социалистического сектора народного хозяйства. Это следующие показатели: 1 — промышленность, 2 — строительство, 3 — сельское хозяйство, 4 — лесное хозяйство, 5 — транспорт, 6 — связь, 7 — торговля и общественное питание, 8 — материально-техническое снабжение, 9 — заготовка сельско-хозяйственных продуктов, 10 — остальные отрасли сферы производства, 11 — отрасль научных исследований, 12 — коммунальные услуги, 13 — жилищное хозяйство, 14 — здравоохранение и социальное обеспечение, 15 система образования, культуры и просвещения, 16 — сфера управлений и судов, 17 — денежное хозяйство и страхование, 18 — общественные организации, 19 — остальные непроизводительные отрасли.

Настоящая работа с методической точки зрения подразделяется на 3 части: 1 — обработка данных, 2 — региональное группирование, 3 — выявление специфических признаков выделяемых регионов.

1. Для каждой области вычислялась структура занятости по 19-ти показателям (табл. 1). Отдельные пары показателей были тесно скоррелированы. Поэтому мы решили применить метод комплексного анализа. В табл. 2 приведены компонентные веса 19-ти показателей в отношении к четырем компонентам. В табл. 3 содержатся стандартизованные значения компонентных отношений областей ЧССР, т. е. исходные данные для региональной группировки.

2. Первым шагом региональной группировки является разработка матрицы связей соседства территориальных единиц (для областей ЧССР это табл. 4). Значения 0 или 1 в этой матрице означают, что пары областей не имеют или имеют общие границы.

Вторым шагом является расчет расстояний (разниц) V^2_{XY} по формуле $V^2_{XY} = \sum_{j=1}^4 (X_j - Y_j)^2$, где X_1, \dots, X_4 — это стандартизованные значения компонентного отношения области X , далее Y_1, \dots, Y_4 — это стандартизованные значения компонентного отношения области Y . Коэффициенты V^2_{XY} вычисляются только для соседних областей. В редуцированной матрице расстояний (табл. 5) должны заполняться те места, которые в матрице связей соседства территориальных единиц равны единице.

Третий шаг это создание дендритов. В данном примере мы получили один дендрит (рис. 1). В случае большего количества территориальных единиц можно получить больше дендритов. Каждый дендрит позволяет идентифицировать пары соседних территориальных единиц и те единицы, которые обладают относительно схожими свойствами. Полученные единицы можно в процессе группировки объединять. В первом цикле группировки выявлены наиболее схожие свойства Восточно-чешской и Южно-моравской областей. Агрегированные территориальные единицы мы замещаем центрами тяжести точек, которым отвечают эти территориальные единицы в пространстве компонентов. В первом цикле группировки, например, вычисляется центр тяжести двух точек, во втором цикле центр тяжести трех точек и т. п. д. В дальнейшем процесс группировки продолжается первым шагом, т. е. созданием матрицы связей соседства принимаемой в качестве одной территориальной единицы. Группирование

заканчивается после того как все территориальные единицы объединены в одно целое. Лучше всего процесс группировки иллюстрируется с помощью дендрограммы (рис. 2).

Преимуществом циклического регионального группирования является элиминация ограниченности информации, которые может передать редуцированная матрица расстояний (разниц).

В случае потери информации, введенной Берии (3), к качеству хорошего коэффициента, характеризующего однородность группировки, можно принять индекс группировки. Индекс группировки в k -том цикле это произведение двух величин, представляющих собой число группирующихся элементов в том же самом цикле. Например, в девятом цикле группировки агрегируется одна область (Средне-чешская) с остальными девятью областями. Индекс группировки цикла $k = 9$ равен произведению $1 \times 9 = 9$. Для сравнения в табл. 6 приводятся значения индекса группировки, а также коэффициента потери информации.

3. На основании роста индекса группировки выделено шесть однородных регионов. Для каждого региона вычислялась структура и все показатели были стандартизованы. После исправлений мы получили табл. 7, по которой можно определить специфические комбинации отдельных регионов. Районизованная территория изображена на карте в приложении.

Для отдельных регионов характерна следующая комбинация показателей:

Регион Р — I (Северно-чешская, Восточно-чешская, Северо-моравская и Южно-моравская области) показатели 1, 12.

Регион Р — II (Южно-чешская область): 14, 3, 12, 1, 9, 13, 4 и 7.

Регион Р — III (Западно-чешская область): 9, 3, 4, 19, 6, 17, 1 и 7.

Регион Р — VI (Средне-чешская область): 8, 7, 17, 11, 13, 10, 16, 19, 6, 12, 18, 2.

Регион Р — V (Западно-словацкая и Средне-словацкая области): 18, 15, 2, 10, 11, 16, 4.

Регион Р — VI (Восточно-словацкая область): 5, 15, 2, 4, 18 и 14.

Рис. 1 Дендрит первого цикла группировки.

Рис. 2 Дендрограмма региональной группировки областей ЧССР.

Рис. 3 Гомогенные регионы ЧССР.

Табл. 1 Структура занятости областей ЧССР на основании 19 показателей.

Табл. 2 Матрица главных компонент.

Табл. 3 Стандартизованные компонентные отношения.

Табл. 4 Матрица связей соседства областей ЧССР.

Табл. 5 Редуцированная матрица разниц (расстояний) территориальных единиц.

Табл. 6 Значения индекса группировки и потери информации.

Табл. 7 Специфические признаки регионов.

Перевод: Л. Правдова