

Т.Б. Вітряк
Vitriak, T.B.

Особливості кластерного моделювання регіональних ринків праці України

Cluster modeling features of regional labor markets in Ukraine

У статті охарактеризовані методи кластеризації та запропоновано методу розподілення регіонів України за макроекономічними показниками. Проведено кластерне ранжування регіональних ринків праці на основі таксонометричного показника. Обґрунтовано стійкість запропонованої моделі.

Ключові слова: кластер, кластерне моделювання, ієрархічна кластеризація, агломеративний метод, метод *k*-середніх, дендрограма, таксонометричний показник.

This article describes clustering methods and it offers the method of classification of Ukraine's regions by macroeconomic indicators. A cluster rank list of regional labor markets was made on the basis of taxonomic indicators. The stability of the proposed model was substantiated.

Keywords: cluster, cluster modeling, hierarchical clustering, agglomerative method, *k*-means method, dendrogram, taxonomic indicator.

Постановка проблеми. У сучасних умовах розвитку економіки особливої актуальності набувають проблеми ефективної зайнятості населення України, створення ринку робочої сили і запобігання масового безробіття. Ринок праці посідає центральне місце серед інших ринків, перебуває під впливом багатьох факторів, більшість яких залежить від загальної економічної ситуації в країні та макроекономічних показників. Ці характеристики мають суттєві територіальні розбіжності, що зумовлює необхідність поглибленого вивчення регіональних особливостей ринку праці України та розподілення його на певні кластери. Із урахуванням типологічних особливостей ринок праці здатний не тільки діяти в межах певних завдань, а й самостійно створювати умови для їх вирішення.

Аналіз останніх досліджень і публікацій. Сучасна концепція кластерів запропонована М. Портером і продовжена в наукових дослідженнях зарубіжних учених (І. Дахмена, Р. Коуза, А. Маршала, В. Прайса та ін.). Ефективна методика вивчення конкурентоспроможності регіонів та виявлення регіональних економічних кластерів запропонована С. Бобровим, Н. Жуковим і В. Яровим [3]. Стратегія формування виробничих регіональних кластерів та аналіз їх ефективності розглядається у роботах М. Войнаренка [6].

Вітряк Тетяна Борисівна, кандидат економічних наук, науковий співробітник проблемної науково-дослідної лабораторії соціально-економічних проблем ринку праці, Інститут підготовки кадрів державної служби зайнятості України (Київ).

Vitriak, Tetiana Borysivna, PhD (Econ.), Researcher, Research Laboratory of Socio-Economic Problems of the Labour Market, Ukrainian State Employment Service Training Institute (Kyiv).

© Вітряк Т.Б., 2016

© Vitriak, T.B., 2016

Однак незважаючи на існуючий практичний досвід і наявність наукових розробок, проблеми кластеризації ринків праці України недостатньо вивчені та потребують подальшого дослідження, особливо на регіональному рівні.

Мета статті полягає у вивченні можливостей кластерного моделювання регіональних ринків праці України та їх розподіл за таксонометричним показником.

Виклад основного матеріалу. Методика кластерного аналізу базується на поняттях подібності об'єктів. За допомогою підбору найбільш "подібних" одиниць виконується розподіл сукупності на кластери (групи). Кластерний аналіз – це багатовимірна статистична процедура, яка виконує збір даних, що містять інформацію про вибірку об'єктів, і потім упорядковує об'єкти у порівняно однорідні групи. На відміну від інших методів цей вид аналізу дає можливість класифікувати об'єкти не за однією ознакою, а за кількома одночасно. Для цього вводяться відповідні показники, що характеризують певну міру близькості за всіма класифікаційними параметрами [1].

Мета кластерного аналізу полягає в пошуку наявних структур, що виражається в утворенні груп схожих між собою об'єктів – кластерів. Водночас його дія полягає й у привнесенні структури в досліджувані об'єкти. Це означає, що методи кластеризації необхідні для виявлення структури в даних, яку нелегко знайти при візуальному обстеженні або за допомогою експертів.

Методи кластерного аналізу можна розділити на ієрархічні та неієрархічні [2].

Для вирішення завдань дослідження пропонується використовувати наступні методи багатовимірної класифікації: ієрархічні агломеративні методи і метод *k*-середніх.

Пропонований алгоритм групування адміністративних регіонів за рівнем зайнятості населення та макроеко-

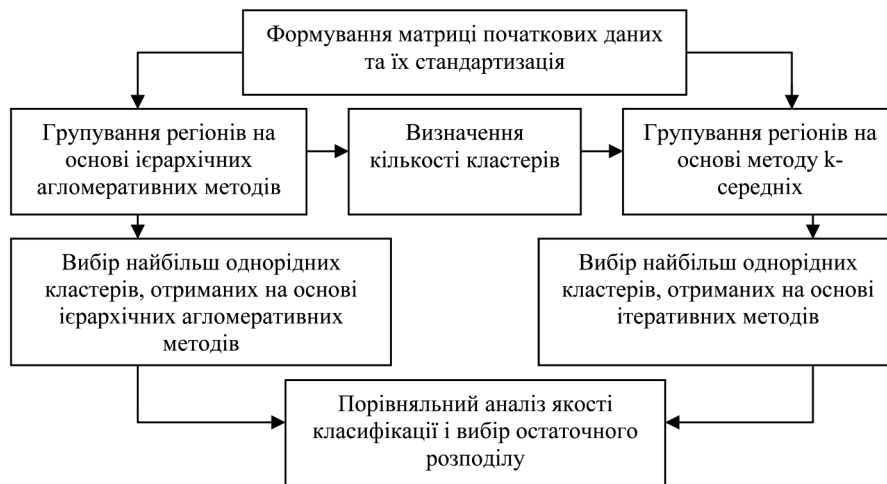


Рис.1. Результати діагностики за методикою Ш. Шварца

Джерело: розроблено автором на основі [3].

номічними показниками України наведено на рис. 1. Реалізація запропонованого алгоритму дозволить сформувати стійкі кластери регіонів з аналогічними тенденціями зайнятості населення.

Головним підсумком ієрархічного кластерного аналізу є дендрограма. При її інтерпретації дослідники стикаються з проблемою того ж роду, що і тлумачення результатів факторного аналізу – відсутністю однозначних критеріїв виділення кластерів. Рекомендується використовувати два способи – візуальний аналіз дендрограми і порівняння результатів кластеризації, виконаної різними методами.

Результати ієрархічного кластерного аналізу можна перевіряти ітеративним кластерним аналізом за методом k-середніх. Його характерна особливість полягає в тому, що кластери формуються, виходячи з поставлених умов

розподілу (параметрів), які в процесі роботи алгоритму можуть бути змінені користувачем для досягнення бажаної якості розподілу. Ітеративні методи відносяться до швидкодіючих, що дозволяє використовувати їх для обробки великих масивів вихідної інформації [4].

Метод k-середніх належить до групи ітеративних методів еталонного типу. На відміну від ієрархічних процедур, метод k-середніх не вимагає обчислення і зберігання матриці відстаней або подібностей між об'єктами. Алгоритм цього методу передбачає використання тільки вихідних значень змінних.

Для проведення і апробації запропонованого методу кластерного аналізу використовувалися середні дані Державної служби статистики України за 2010–2014 рр. (табл. 1). 3 об'єктивних причин статистичні дані

Таб.1. Середні макроекономічні показники України за регіонами

Ум. позн.	Регіони	Рівень зайнятості, %	Продуктивність праці, тис. грн/осіб	Капітальні інвестиції, млн грн	Економічно активне населення, тис. осіб	Середньомісячна заробітна плата штатних працівників, грн
	Україна	58,8	71,5	29 752	20 676,8	2 929
C1	Вінницька	58,1	46,6	19 740	764,3	2 350
C2	Волинська	58,2	43,0	17 943	475,0	2 265
C3	Дніпропетровська	60,9	92,1	42 216	1 635,9	3 055
C4	Житомирська	58,8	42,5	18 226	607,1	2 310
C5	Закарпатська	57,4	36,7	15 510	580,6	2 313
C6	Запорізька	60,0	62,6	28 452	878,3	2 859
C7	Івано-Франківська	53,6	52,5	20 716	591,2	2 447
C8	Київська	58,6	82,7	36 012	805,8	3 011
C9	Кіровоградська	57,4	50,8	21 505	459,3	2 351
C10	Львівська	57,8	51,8	19 211	1 179,2	2 503
C11	Миколаївська	59,3	54,1	24 290	575,8	2 766
C12	Одеська	58,2	61,4	26 801	1 111,9	2 642
C13	Полтавська	58,0	82,9	35 986	707,1	2 720
C14	Рівненська	58,3	41,7	17 508	540,6	2 525
C15	Сумська	58,4	46,7	20 614	557,3	2 425
C16	Тернопільська	54,7	38,9	15 598	482,3	2 120
C17	Харківська	60,3	62,2	28 679	1 362,7	2 668
C18	Херсонська	58,4	39,6	17 400	522,2	2 211
C19	Хмельницька	58,2	43,5	18 529	617,5	2 361
C20	Черкаська	58,7	52,0	22 668	614,3	2 402
C21	Чернівецька	57,4	32,4	13 698	417,7	2 230
C22	Чернігівська	59,2	46,9	17 478	522,0	2 237

Джерело: розроблено автором на основі даних Державної служби статистики.

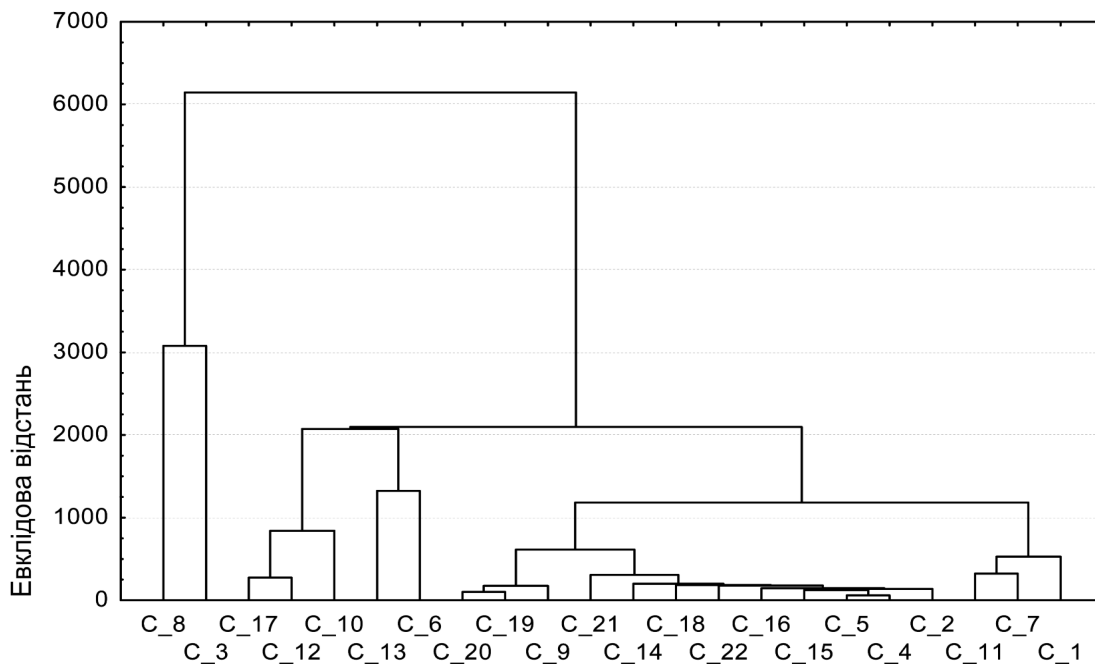


Рис.2. Результати групування на основі ієрархічних агломеративних методів

Джерело: розроблено автором.

Донецької та Луганської областей за останні два роки є неповними, тому вони були виключені з кластерної моделі для запобігання перекручування результатів.

За допомогою програми STATISTICA отримано результати графічного аналізу у вигляді дендрограми (рис. 2), яка дозволяє зробити висновок, що досліджувана сукупність регіонів може бути розділена на три групи об'єктів за макропоказниками, близькими за характером розвитку.

Для початку процедури класифікації повинні бути задані k випадково обраних об'єктів, які використовуються як еталони, тобто центри кластерів. Можливі дві модифікації методу k -середніх. Перша передбачає перерахунок центру тяжіння кластера після кожної зміни його складу, а друга – лише після того, як буде завершено перегляд усіх даних. В обох випадках ітеративний алгоритм цього методу мінімізує дисперсію всередині кожного кластера, хоча в реальному вигляді такий критерій оптимізації не використовується [5].

Обчислювальні процедури більшості ітеративних методів класифікації зводяться до виконання наступних кроків [6]:

- вибір кількості кластерів, на які повинна бути розподілена сукупність, завдання початкового розподілу об'єктів та визначення центрів ваги кластерів;
- відповідно до обраних методами подібності визначається новий склад кожного кластера;
- після повного перегляду всіх об'єктів і розподілу їх за кластерами здійснюється перерахунок центрів їх тяжіння.

Для пошуку оптимальної конфігурації необхідно порівняти значення потенційних сукупних результатів усіх учасників кластера при всіх можливих варіантах, визначити еталонну конфігурацію, найвигіднішу для всіх учасників, і обрати оптимальний склад кластера, орієнтуючись на еталон. Це завдання можливо вирішити за допомогою таксонометричного методу.

В основу таксонометричного методу (методу евклідових відстаней) закладено вибір еталона, в даному випадку – “еталонної конфігурації” та порівняння опти-

мальних параметрів (координат) її вектора з відповідними параметрами векторів інших можливих конфігурацій, тобто знаходження евклідових відстаней, за якими і проводиться ранжування можливих структур кластера: найменша відстань відповідає вищому місцю.

Принцип даної методики полягає у представленні всіх даних за бажаними перевагами від участі в кластері у вигляді матриці, де окремий рядок є вектором переваг варіанта конфігурації, координатами якого є ті ж значення сукупних результатів.

Оскільки показники моделі зайнятості населення мають різну природу і неспіврозмірні значення, проводиться нормування елементів матриці. Для цього слід замінити вихідні дані на матрицю K . Елементи матриці K розраховуються за формулою:

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

де $j = 1, 2, 3, 4$ – номер показника, $i = 1, 2, \dots$;

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij},$$

де n – номер спостереження;

σ_j – середньоквадратичне відхилення j -того показника, що розраховується за формулою:

$$\sigma_j = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2} = \sqrt{\overline{(x_{ij}^2)} - (\bar{x}_j)^2}.$$

Використовуючи програму STATISTICA нами отримані нормовані дані матриці K (табл. 2).

Для формування точки – еталона серед елементів матриці обираються максимальні значення. Таким чином, еталоном вважатиметься точка E_0 з координатами $K_{01}, K_{02}, K_{03}, K_{04}, K_{05}$, отриманими таким способом:

$$K_{0j} = \max k_{ij}, j = \overline{1,5}.$$

Таб.2. Матриця нормованих значень

Адміністративні райони	Рівень зайнятості, %	Продуктивність праці, тис. грн/осіб	Капітальні інвестиції, млн грн	Економічно активне населення, тис. осіб	Середньомісячна заробітна плата штатних працівників, грн	Евклідова відстань	Таксонометричний показник
	X_1	X_2	X_3	X_4	X_5	C_{i0}	d_i
Вінницька	-0,114	-0,409	-0,319	-0,147	-0,896	5,8539	0,3631
Волинська	-0,054	-0,527	-0,497	-0,827	-1,057	6,8145	0,2586
Дніпропетровська	1,565	1,072	0,816	1,902	0,442	0,1996	0,9783
Житомирська	0,306	-0,543	-0,506	-0,517	-0,972	6,4538	0,2979
Закарпатська	-0,534	-0,732	-0,502	-0,579	-0,966	6,9159	0,2476
Запорізька	1,025	0,112	-0,169	0,121	0,070	3,9771	0,5673
Івано-Франківська	-2,812	-0,217	-0,355	-0,554	-0,712	7,2862	0,2073
Київська	0,186	0,766	0,596	-0,049	0,359	3,0921	0,6636
Кіровоградська	-0,534	-0,273	-0,441	-0,864	-0,894	6,5655	0,2857
Львівська	-0,294	-0,240	0,076	0,828	-0,606	4,4089	0,5203
Миколаївська	0,606	-0,165	-0,358	-0,590	-0,106	5,2791	0,4257
Одеська	-0,054	0,073	0,144	0,670	-0,342	3,7339	0,5938
Полтавська	-0,174	0,773	-0,072	-0,281	-0,194	4,4305	0,5180
Рівненська	0,006	-0,569	-0,512	-0,673	-0,564	6,3590	0,3082
Сумська	0,066	-0,406	-0,499	-0,634	-0,754	6,2759	0,3172
Тернопільська	-2,152	-0,660	-0,495	-0,810	-1,333	8,0302	0,1264
Харківська	1,205	0,099	0,136	1,260	-0,292	2,9402	0,6801
Херсонська	0,066	-0,637	-0,530	-0,716	-1,160	6,9479	0,2441
Хмельницька	-0,054	-0,510	-0,447	-0,492	-0,875	6,3273	0,3116
Черкаська	0,246	-0,233	-0,454	-0,500	-0,797	5,9492	0,3528
Чернівецька	-0,534	-0,872	-0,551	-0,962	-1,124	7,5215	0,1817
Чернігівська	0,546	-0,400	-0,516	-0,717	-1,111	6,5638	0,2859
Еталон	1,7447	2,1673	2,8494	2,9029	2,1792		
Середня відстань $C_0 = 5,542$							
Середнє квадратичне відхилення $\sigma_0 = 1,825$							

Джерело: розроблено автором.

Оскільки матриця K складається із нормованих значень, кожен стовпець матриці являє собою вектор, координати якого у сумі дорівнюють нулю.

Порівняння можливих варіантів конфігурації кластера з показниками еталонної конфігурації здійснювалось за допомогою визначення евклідової відстані за формулою:

$$C_{i0} = \sqrt{\sum_{j=1}^N (k_{ij} - k_{0j})^2},$$

де k_{0j} , k_{ij} – нормовані координати відповідно еталонної і досліджуваної конфігурації.

Середня відстань за всією сукупністю визначалася за формулою:

$$C_0 = \frac{1}{m} \sum_{i=1}^m C_{i0}.$$

Середнє квадратичне відхилення за всією сукупністю визначалось:

$$\sigma_0 = \sqrt{\frac{1}{m} \sum_{i=1}^m (C_{i0} - C_0)^2}.$$

Таксонометричний показник досліджуваної конфігурації складає:

$$d_i = 1 - C_{i0}/C^*,$$

де $C^* = C_0 + 2\sigma_0$.

Чим ближче значення таксонометричного показника до одиниці, тим досконаліша конфігурація, тобто тим з більшою ймовірністю наявні ресурси використовуються ефективно.

За таксонометричним показником побудуємо рейтинг можливих конфігурацій та складемо кластери (табл. 3).

У результаті кластерного ранжування за рівнем зайнятості населення та факторних макроекономічних показників найближчим до еталона виявилися адміністративні регіони кластеру 1, до якого входять Дніпропетровська, Харківська, Київська, Одеська, Запорізька, Львівська і Полтавська області, їх середній таксонометричний показник є найвищим серед усієї сукупності адміністративних регіонів України і складає 0,646. У кластер 2 увійшли дев'ять адміністративних регіонів, їх середній таксонометричний показник складає 0,327, де лідерами є

Таб.3. Кластерне ранжування на основі таксонометричного показника

Кластери	Адміністративні регіони	d_i	Середній таксонометричний показник по кластеру
1	Дніпропетровська	0,9783	0,646
	Харківська	0,6801	
	Київська	0,6636	
	Одеська	0,5938	
	Запорізька	0,5673	
	Львівська	0,5203	
	Полтавська	0,5180	
2	Миколаївська	0,4257	0,327
	Вінницька	0,3631	
	Черкаська	0,3528	
	Сумська	0,3172	
	Хмельницька	0,3116	
	Рівненська	0,3082	
	Житомирська	0,2979	
	Чернігівська	0,2859	
	Кіровоградська	0,2857	
3	Волинська	0,2586	0,211
	Закарпатська	0,2476	
	Херсонська	0,2441	
	Івано-Франківська	0,2073	
	Чернівецька	0,1817	
	Тернопільська	0,1264	

Джерело: розроблено автором.

Миколаївська, Вінницька та Черкаська області. Найбільш віддалений від еталона є кластер 3, до якого увійшли шість адміністративних регіонів із середнім таксонометричним показником 0,211.

Взаємозв'язок між ознаками можна визначити передусім на основі бальних оцінок, що вимірюються методами рангової кореляції. Рангами називають числа натурального ряду, котрі згідно зі значеннями ознаки надаються елементам сукупності і певним чином упорядковують її. Ранжування проводиться за кожною ознакою окремо: перший ранг надається найменшому значенню ознаки, останній – найбільшому або навпаки. Кількість рангів дорівнює обсягу сукупності. Очевидно, зі збільшенням обсягу сукупності ступінь “розпізнаваності” елементів зменшується. Проведемо рангове ранжування сукупності з використанням рівня зайнятості населення за регіонами та таксонометричним показником (табл. 4).

Окрім рангів для кожного елемента із сукупності ознак x_i у в таблиці розраховані D_i – різниця рангів $R_{x_i} - R_{y_i}$ та D^2 – квадрат різниці рангів пари відповідних елементів. Для розрахунку коефіцієнта рангової кореляції Спірмена використовують формулу:

$$\rho = 1 - \frac{\sum_{i=1}^m D_i^2}{m(m^2 - 1)}$$

Коефіцієнт кореляції рангів свідчить про наявність прямого зв'язку між рівнем зайнятості населення по регіонах та таксонометричним показником кластерного ранжування:

$$\rho = 1 - \frac{6 \times 804}{22(22^2 - 1)} = 1 - \frac{4824}{10626} = 0,546$$

Достовірність коефіцієнта кореляції рангів перевіряємо за таблицею Фішера. Табличне значення коефіцієнта кореляції при $\alpha=0,05$ і $k=n-m-1=22-2-1=19$ становить $\rho_{табл}=0,352$. Оскільки $\rho_{факт} > \rho_{0,05}$ ($0,546 > 0,352$), це свідчить про те, що розподіл адміністративних регіонів України на три кластери є доцільним і підтверджує їх однорідність, а також передбачає помірний взаємозв'язок.

Значущість коефіцієнта рангової кореляції Спірмена перевіряється за формулою:

$$\rho_{xy} = \rho - \frac{\sqrt{m-2}}{\sqrt{1-\rho^2}}, \text{ тобто}$$

$$\rho_{xy} = 0,546 - \frac{\sqrt{22-2}}{\sqrt{1-0,546^2}} = 2,917$$

Табличне значення коефіцієнта кореляції Спірмена при $\alpha=0,05$ і $k=n-m=22-2=20$ становить $\rho_{табл}=2,086$. Оскільки $\rho_{факт} > \rho_{табл}$ ($2,917 > 2,086$), можна зробити висновок, що із вірогідністю 95% коефіцієнт рангової кореляції є достовірним, а факторні ознаки моделі – не випадковими.

Для наочного подання розташування кластерів до еталонної конфігурації нами побудована пелюсткова діаграма (рис. 3).

Результати проведеного аналізу дозволяють розглянути пропозиції для кожного кластера та сформулювати стратегії подальшого розвитку ринку праці України.

Таб. 4. Рангове розподілення сукупності з використанням рівня зайнятості населення по регіонах та таксонометричних показниках

	Рівень зайнятості, %	R_y	d_i	R_x	$R_x - R_y$	$(R_x - R_y)^2$
Вінницька	58,1	15	0,3631	9	-6	36
Волинська	58,2	12	0,2586	17	5	25
Дніпропетровська	60,9	1	0,9783	1	0	0
Житомирська	58,8	6	0,2979	14	8	64
Закарпатська	57,4	18	0,2476	18	0	0
Запорізька	60,0	3	0,5673	5	2	4
Івано-Франківська	53,6	22	0,2073	20	-2	4
Київська	58,6	8	0,6636	3	-5	25
Кіровоградська	57,4	18	0,2857	16	-2	4
Львівська	57,8	17	0,5203	6	-11	121
Миколаївська	59,3	4	0,4257	8	4	16
Одеська	58,2	12	0,5938	4	-8	64
Полтавська	58,0	16	0,5180	7	-9	81
Рівненська	58,3	11	0,3082	13	2	4
Сумська	58,4	9	0,3172	11	2	4
Тернопільська	54,7	7	0,1264	22	15	225
Харківська	60,3	2	0,6801	2	0	0
Херсонська	58,4	9	0,2441	6	-3	9
Хмельницька	58,2	12	0,3116	12	0	0
Черкаська	58,7	7	0,3528	10	3	9
Чернівецька	57,4	18	0,1817	21	3	9
Чернігівська	59,2	5	0,2859	15	10	100
Разом						804

Джерело: розроблено автором.

Висновки. Аналіз сучасного стану регіональних ринків праці дає можливість дійти висновку, що ефективний механізм їх регулювання повинен включати досить різнопланові та диференційовані за силою впливу підсистеми, не обмежуючись лише сферою трудових відносин. Для його формування слід перш за все визначитися з пріоритетами регулювання попиту на робочу силу та її пропозиції. Розробка ефективної стратегії управління процесами в різних кластерах регіональних ринків праці вимагає комплексного підходу до вирішення економічних, соціальних, демографічних проблем, які безпосередньо впливають та визначають його стан.

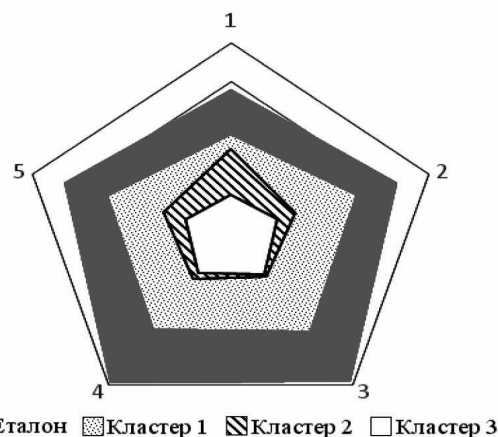


Рис. 3. Відношення показників кластерів до еталона

Джерело: розроблено автором.

Список використаних джерел

1. Наконечний С.І. До питання математичного моделювання техніко-економічних процесів АПК / С.І. Наконечний, С.С. Савіна, Т.С. Наконечний // Економіка АПК. – 2009. – №1(171). – С. 16–21.
2. Катренко А.В. Системний аналіз об'єктів та процесів комп'ютеризації / А.В. Катренко. – Львів : Новий світ, 2003. – 424 с.
3. Бобров С.Я. Актуальные вопросы формирования кластеров как инструмента повышения конкурентоспособности региона / С.Я. Бобров, Н.В. Жуков, В.В. Яровий // Фундаментальные исследования. Серия : Экономические науки. – 2007. – №12. – С. 508–509.
4. Плюта В. Сравнительный многомерный анализ в экономических исследованиях. Методы таксономии и факторного анализа / В. Плюта. – М. : Статистика, 2008. – 152 с.
5. Федосеев В.В. Экономико-математические методы и прикладные модели / В.В. Федосеев, А.Н. Гармаш, Д.М. Дайитбегов и др. – М. : ЮНИТИ, 2002. – 456 с.
6. Войнаренко М. Концепція кластерів – шлях до відродження виробництва на регіональному рівні / М. Войнаренко // Економіст. – 2000. – №1. – С. 29–31.

Стаття надійшла 10.11.2015