

УДК 332.1

ББК 65.9(2Рос-16)

Исследование социально-экономических различий субъектов Сибирского федерального округа методом кластерного анализа

В.И. Псарев, А.Ю. Юдинцев, Г.Н. Трошкина

Алтайский государственный университет (Барнаул, Россия)

The Research of Socio-Economic Differences of the Siberian Federal District by a Cluster Analysis Method

V.I. Psarev, A.Yu. Yudintsev, G.N. Troshkina

Altai State University (Barnaul, Russia)

Рассматривается проблема определения социально-экономической дифференциации субъектов Сибирского федерального округа (СФО). На основе оперативных данных территориального органа Федеральной службы государственной статистики по Алтайскому краю, характеризующих экономическое и социальное положение Алтайского, Красноярского, Забайкальского краев, Новосибирской, Омской, Томской, Иркутской, Кемеровской областей, республик Алтай, Тыва, Хакасия, Бурятия по состоянию на конец 2014 г., проведено исследование степени различий социально-экономического развития с использованием методов кластерного анализа. Используется следующее подпространство показателей социально-экономического развития на конец 2014 г.: доходы, расходы на душу населения, среднемесячная начисленная заработная плата работников организаций, объемы производства продукции сельского хозяйства, выполненных работ по виду деятельности «Строительство», оборот розничной торговли, численность работающих, безработных из экономически активного населения. В качестве меры схожести использовалось евклидово расстояние в пространстве нормализованных переменных. Нормализованные переменные вычислялись как отклонения исходных значений от средней величины выборки, деленных на дисперсию. Последовательным применением метода *k*-средних построена кластерная структура СФО для разбиений от двух до шести кластеров. Определены расположение, состав, статистические характеристики кластеров различных уровней, степени схожести и различия субъектов СФО.

Ключевые слова: социально-экономическая дифференциация, сглаживание регионального неравенства, кластерный анализ.

DOI 10.14258/izvasu(2015)2.1-25

The article deals with the problem of determining socio-economic differences within the constituent entities of the Siberian Federal District. The authors have conducted comprehensive research of the degree of socio-economic development difference using cluster analysis methods based on the operational data of the Altai Federal Statistics Service. The data characterize economic and social situation in the Altai Territory, Krasnoyarsk Territory, Zabaykalsk Territory, Novosibirsk Oblast, Omsk Oblast, Tomsk Oblast, Irkutsk Oblast, Kemerovo Oblast, the Republics of Altai, Tuva, Khakassia and Buryatia by the end of 2014. The following indicators of socio-economic development are used: per capita income, per capita consumption, the average monthly gross payroll, agricultural output, the amount of construction work performed, the retail trade turnover, total employment rate and unemployment rate in the economically active population. As a measure of similarity, the authors use the Euclidean distance in the normalized variables space. The normalized variables have been calculated as a ratio of the difference between the initial values and the mean value and the variance for each variables set. The use of *k*-means method led to building the cluster structure of the Siberian Federal District with 2- to 6- cluster division. The authors define the location, composition, statistical characteristics of different levels of clusters, the degree of similarity and dissimilarity between the constituent entities of the Siberian Federal District.

Key words: socio-economic differences, regional disparities smoothing, cluster analysis.

Исследование и мониторинг степени социально-экономических различий субъектов РФ, оценка уровня их социально-экономического развития, как и проблема регионального развития в целом, несомненно, являются актуальными и сложными задачами региональной экономики, имеющим глубинные корни. Так, в статье [1] рассмотрены основные задачи и ключевые принципы территориальной политики региона, показано наличие противоречия между экономическими и социальными целями развития территориальных систем. Вопросам выравнивания межрегиональных различий в инвестиционной сфере посвящена работа [2]. Проблема пространственного развития Российской Федерации, политика сглаживания пространственной поляризации и политика поляризованного развития, их достоинства и недостатки, инструменты, способствующие сокращению межрегиональной дифференциации, исследованы в [3]. Методика оценки социальных различий в Российской Федерации на основе заданной иерархической структуры с выраженной демографической и социальной спецификой изложена в [4]. Опыт применения методов многомерного статистического анализа для оценки уровней социально-экономического развития регионов России показан в [5].

Оперативные данные, характеризующие экономическое и социальное положение Алтайского края в сравнении с Красноярским, Забайкальским краями,

Новосибирской, Омской, Томской, Иркутской, Кемеровской областями, республиками — Алтай, Тыва, Хакасия, Бурятия, по состоянию на конец 2014 г. приведены в бюллетене территориального органа Федеральной службы государственной статистики по Алтайскому краю «Социально-экономический мониторинг субъектов РФ Сибирского федерального округа за 2014 год» [6].

Для выявления степени различий социально-экономического развития регионов, классификации субъектов по степени близости уровня развития, выявления универсальной характеристики степени дифференциации используем методы кластерного анализа. Социально-экономическое положение субъекта можно представить точкой в многомерном пространстве либо подпространстве социально-экономических параметров. Рассмотрим следующее подпространство социально-экономических параметров: X_1 — доходы на душу населения; X_2 — расходы на душу населения; X_3 — начисленная заработная плата работников организаций, среднемесячные значения за январь — ноябрь 2014 г., руб.; X_4 — объем производства продукции сельского хозяйства в хозяйствах всех категорий за 2014 г. (на душу населения), руб.; X_5 — объем выполненных работ по виду деятельности «Строительство»; X_6 — оборот розничной торговли за 2014 г., млн руб., X_7 — численность работающих; X_8 — численность безработных;

Таблица 1

Исходные значения параметров социально-экономического развития субъектов СФО на начало 2015 г. [5]

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
РА	15856,6	9695,6	21624,4	43609	7466,1	20368,4	92,3	9,5	213,5
РБ	21334,9	15549,5	26951,4	15404	20035,0	138997,1	421,6	37,3	978,6
РТ	13322,1	6357,0	26819,0	17350	5035,8	18785,7	102,0	17,9	313,6
РХ	17976,9	12975,3	28138,0	22463	13339,6	64080,5	237,4	12,2	535,6
АК	17209,1	12868,3	19023,0	45260	32971,6	306526,3	1091,3	112,5	2384,7
ЗК	19548,5	13721,9	28360,8	17146	25536,6	135835,4	481,0	53,5	1087,5
КК	23261,6	18669,4	33328,5	26620	135815,7	490769,2	1456,8	67,3	2859,8
ИО	19420,1	13565,6	30687,7	22565	84477,6	287935,8	1133,0	120,0	2415,7
КО	19229,4	13493,3	26226,0	18022	121652,0	339052,4	1312,4	97,1	2725,3
НО	22585,7	19550,2	26660,7	25959	54767,9	462588,5	1353,1	88,8	2746,7
ОО	23046,8	16850,0	25668,6	41893	59352,8	314250,0	968,0	73,7	1978,5
ТО	19915,1	13745,0	31726,6	22258	37866,4	120356,8	492,6	41,3	1074,3
Ср. зн.	19392,2	13920,1	27101,2	26545,8	49859,8	224962,2	761,8	60,9	1609,5
Дисп.	2996,1	3604,5	3978,0	10862,1	43594,7	163832,8	509,0	38,5	1010,2

Здесь введены краткие обозначения субъектов Сибирского федерального округа: РА — Республика Алтай, РБ — Республика Бурятия, РТ — Республика Тыва, РХ — Республика Хакасия, АК — Алтайский край, ЗК — Забайкальский край, КК — Красноярский край, ИО — Иркутская область, КО — Кемеровская область, НО — Новосибирская область, ОО — Омская область, ТО — Томская область. Для выполнения процедуры нормализации данных для каждого параметра определены среднее значение — «Ср. зн.» и дисперсия — «Дисп.».

X_9 — численность экономически активного населения (в возрасте 15–72 лет), тыс. чел. (в среднем за IV квартал 2014 г.). Значения величин X_1 — X_9 приведены в таблице 1.

Параметры социально-экономического развития имеют разные размерности и изменяются в широком диапазоне значений. Для выполнения статистического анализа проведем процедуру нормализации исходных переменных — перейдем к новым, нормализованным переменным V_1, V_2, \dots, V_9 , которые представляют собой отклонения исходных значений от средних величин, деленных на дисперсию. Таким образом, все нормализованные переменные будут безразмерными, иметь нулевое среднее значение и единичную дисперсию. Значения нормализованных переменных приведены в таблице 2.

Для каждого нормализованного значения V_i легко можно получить исходное значение параметра X_i : $X_i = S + V_i \cdot D$, где S и D — среднее значение и дисперсия выборки соответственно.

Выберем в качестве меры длины в пространстве нормализованных переменных V_1, V_2, \dots, V_9 евклидово расстояние. Единицей измерения расстояния является дисперсия. В таблице 3 приведены взаимные расстояния объектов мониторинга.

Из анализа таблицы перекрестных расстояний можно увидеть, что объекты расположены на достаточно больших расстояниях друг от друга и для выявления кластеров необходимо использовать дополнительные предположения. Проведем кластерный анализ при помощи метода k -средних с целью выяснения степени различия (схожести) объектов и кла-

Таблица 2

Нормализованные значения подпространства социально-экономического развития субъектов СФО

	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9
РА	-1,1801	-1,1720	-1,3768	1,5709	-0,9725	-1,2488	-1,3152	-1,3356	-1,3818
РБ	0,6484	0,4520	-0,0377	-1,0257	-0,6841	-0,5247	-0,6683	-0,6136	-0,6245
РТ	-2,0260	-2,0982	-0,0709	-0,8466	-1,0282	-1,2585	-1,2962	-1,1174	-1,2827
РХ	-0,4724	-0,2621	0,2606	-0,3759	-0,8377	-0,9820	-1,0302	-1,2655	-1,0630
АК	-0,7287	-0,2918	-2,0307	1,7229	-0,3874	0,4978	0,6473	1,3395	0,7674
ЗК	0,0522	-0,0550	0,3166	-0,8653	-0,5579	-0,5440	-0,5516	-0,1928	-0,5167
КК	1,2915	1,3176	1,5654	0,0068	1,9717	1,6224	1,3653	0,1656	1,2376
ИО	0,0093	-0,0983	0,9016	-0,3665	0,7941	0,3844	0,7292	1,5343	0,7980
КО	-0,0543	-0,1184	-0,2200	-0,7847	1,6468	0,6964	1,0817	0,9395	1,1044
НО	1,0659	1,5620	-0,1107	-0,0540	0,1126	1,4504	1,1616	0,7240	1,1257
ОО	1,2198	0,8128	-0,3601	1,4129	0,2178	0,5450	0,4051	0,3318	0,3653
ТО	0,1745	-0,0486	1,1628	-0,3947	-0,2751	-0,6385	-0,5288	-0,5097	-0,5298

Таблица 3

Взаимные расстояния между субъектами СФО

	РА	РБ	РТ	РХ	АК	ЗК	КК	ИО	КО	НО	ОО	ТО
РА	0,00	4,08	3,03	2,84	4,52	3,86	7,52	5,87	6,08	6,47	4,92	4,04
РБ	4,08	0,00	3,93	1,81	4,75	0,99	4,97	3,69	4,04	3,88	3,39	1,59
РТ	3,03	3,93	0,00	2,52	5,74	3,37	7,65	5,55	5,86	6,89	6,06	3,64
РХ	2,84	1,81	2,52	0,00	5,01	1,58	5,93	4,43	4,86	5,13	4,21	1,68
АК	4,52	4,75	5,74	5,01	0,00	4,41	5,61	3,87	3,83	3,93	3,08	4,83
ЗК	3,86	0,99	3,37	1,58	4,41	0,00	4,87	3,10	3,65	3,92	3,39	1,07
КК	7,52	4,97	7,65	5,93	5,61	4,87	0,00	3,10	3,06	2,60	3,45	4,56
ИО	5,87	3,69	5,55	4,43	3,87	3,10	3,10	0,00	1,68	2,75	3,02	3,13
КО	6,08	4,04	5,86	4,86	3,83	3,65	3,06	1,68	0,00	2,76	3,28	3,87
НО	6,47	3,88	6,89	5,13	3,93	3,92	2,60	2,75	2,76	0,00	2,22	4,09
ОО	4,92	3,39	6,06	4,21	3,08	3,39	3,45	3,02	3,28	2,22	0,00	3,38
ТО	4,04	1,59	3,64	1,68	4,83	1,07	4,56	3,13	3,87	4,09	3,38	0,00

стерной структуры СФО. Методика кластерного анализа описана в [7].

На первом этапе методом k -средних сформируем два кластера: $A1$ и $A2$, найдем их состав и характеристики. Состав кластеров $A1$ и $A2$ приведен в таблице 4.

Кластеры первого уровня расположены в точках, значения координат центров которых приведены в таблице 5.

Расстояние между центрами кластеров — 1,8683, кластер $A1$ имеет размер 1,811; $A2$ — 2,0791. Для кластера $A1$ пограничное положение занимают объекты: РА, РТ, а в составе кластера $A2$ имеется объект,

расположенный еще дальше от центра, — расстояние от центра до АК составляет 1,0395. Можно ожидать, что при выполнении разбиения на три кластера Алтайский край выделится из кластера $A2$. Действительно, в результате выполнения разбиения на три кластера второго уровня — $B1$, $B2$, $B3$ получаем, что кластер $A1$ не претерпел изменений и, соответственно, преобразовался в кластер $B1$ (РА, РБ, РТ, РХ, ЗК, ТО), а кластер $A2$ разбился на два кластера: $B2$ (АК, ОО) и $B3$ (КК, ИО, КО, НО).

Состав кластеров второго уровня, характеристика объектов, входящих в кластер, и положение центров приведены в таблицах 6 и 7.

Таблица 4

Состав кластеров $A1$ и $A2$ субъектов СФО и удаленность от центра кластера

Состав кластера $A1$	Удаленность от центра кластера	Состав кластера $A2$	Удаленность от центра кластера
РА	0,905538	АК	1,039546
РБ	0,578177	КК	0,875199
РТ	0,802908	ИО	0,550830
РХ	0,203928	КО	0,569021
ЗК	0,432309	НО	0,518540
ТО	0,531402	ОО	0,585251

Таблица 5

Расположения центров и стандартные отклонения координат для кластеров верхнего уровня

№	Центр кластера $A1$		Стандартное отклонение	Центр кластера $A2$		Стандартное отклонение
	V	X		V	X	
1	-0,46724	17992,3	0,986333	0,467238	20792,1	0,83857
2	-0,53065	12007,4	0,934356	0,530645	15832,8	0,806956
3	0,042436	27270,0	0,826002	-0,04244	26932,4	1,230203
4	-0,32289	23038,5	0,965076	0,322891	30053,1	1,009162
5	-0,72593	18213,3	0,282078	0,725926	81506,3	0,925045
6	-0,86608	83070,7	0,342323	0,866075	366853,7	0,531609
7	-0,89838	304,5	0,363068	0,898378	1219,1	0,362191
8	-0,83911	28,6	0,465359	0,839106	93,2	0,541853
9	-0,89974	700,6	0,391166	0,899737	2518,4	0,322687

Таблица 6

Состав кластеров $B1$, $B2$ и $B3$ субъектов СФО и удаленность объектов от центра кластера

Состав кластера $B1$	Удаленность от центра кластера	Состав кластера $B2$	Удаленность от центра кластера	Состав кластера $B3$	Удаленность от центра кластера
РА	0,905538	АК	0,513915	КК	0,641449
РБ	0,578177	ОО	0,513915	ИО	0,501930
РТ	0,802908			КО	0,495839

Состав кластера B1	Удаленность от центра кластера	Состав кластера B2	Удаленность от центра кластера	Состав кластера B3	Удаленность от центра кластера
PX	0,203928			HO	0,552568
ЗК	0,432309				
ТО	0,531402				

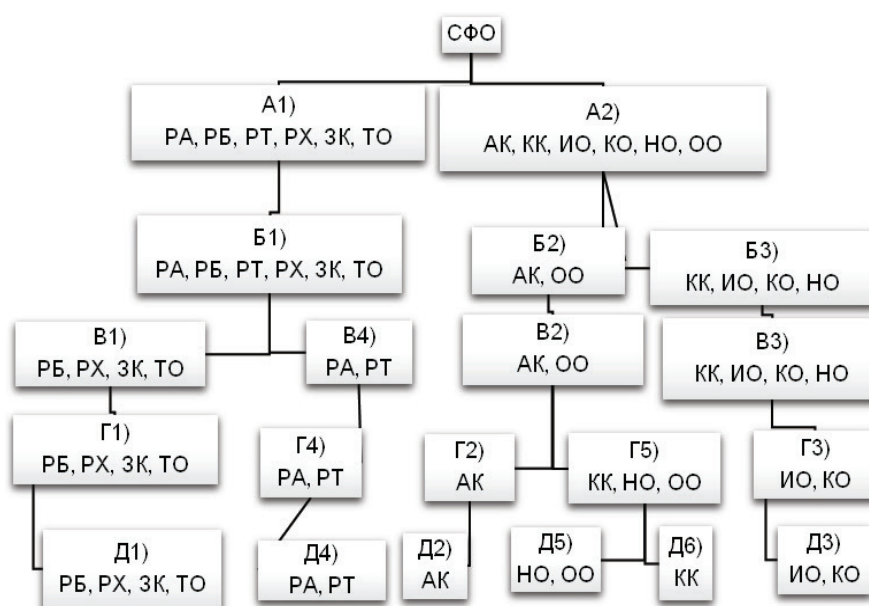
Таблица 7

Положения центра кластеров и стандартные отклонения координат для кластеров второго уровня: B1, B2 и B3

№	Центр кластера B1		Стандартное отклонение	Центр кластера B2		Стандартное отклонение	Центр кластера B3		Стандартное отклонение
	V	X		V	X		V	X	
1	-0,46724	17992,3	0,986	0,24556	20127,9	1,377	0,578077	21124,2	0,700
2	-0,53065	12007,4	0,934	0,26052	14859,1	0,781	0,665706	16319,6	0,89
3	0,042436	27270,0	0,826	-1,19545	22345,7	1,181	0,534069	29225,7	0,853
4	-0,32289	23038,5	0,965	1,56788	43576,3	0,219	-0,2996	23291,5	0,362
5	-0,72593	18213,3	0,282	-0,08482	46162,1	0,427	1,131297	99178,4	0,841
6	-0,86608	83070,7	0,342	0,52142	310387,9	0,033	1,038402	395086,5	0,593
7	-0,89838	304,5	0,363	0,52621	1029,6	0,171	1,084464	1313,8	0,265
8	-0,83911	28,6	0,465	0,83564	93,1	0,712	0,840837	93,3	0,56
9	-0,89974	700,6	0,391	0,56632	2181,6	0,284	1,066446	2686,8	0,188

Последовательно выполняя повторные разбиения и увеличивая требуемое количество кластеров, можно проследить развитие кластерной структуры исследуемого составного объекта. На рисунке приведена кластерная структура СФО при последовательных разбиениях до шести кластеров.

Рассмотрим динамику изменения кластерной структуры в процессе увеличения количества разбиений. Кластер B1 расщепляется на два кластера: B1 (РБ, РХ, ЗК, ТО) и B4 (РА, РТ), которые в дальнейшем сохраняют свой состав вплоть до шестой итерации. Это свидетельствует о сильной схожести субъектов РБ, РХ,



Кластерная структура СФО

ЗК, ТО и РА, РТ в данном подпространстве социально-экономических параметров. Действительно, максимальное расстояние от середины кластера В1–Г1–Д1 составляет 0,3828 для РХ. Для остальных субъектов, входящих в этот кластер, расстояния до центра меньше: РБ — 0,3102, ЗК — 0,1916, ТО — 0,2974. Расстояние между элементами кластера Д4 (РА, РТ) также невелико — 0,5051. Еще более схожи субъекты,

составляющие кластерную пару Д5 (НО, ОО), они расположены на расстоянии 0,3702, а наиболее маленький кластер — это пара Д3 (ИО, КО) с размером 0,2806. Субъекты АК и КК представляют собой кластеры, состоящие из одного элемента. В таблице 8 приведены координаты центров кластеров шестого уровня, матрица взаимных расстояний между центрами кластеров шестого уровня отражена в таблице 9.

Таблица 8

Координаты центров кластеров Д1, Д2, ..., Д6

	Д1	Д2	Д3	Д4	Д5	Д6
В1	0,100670	-0,72866	-0,022524	-1,60306	1,142831	1,291475
В2	0,021592	-0,29180	-0,108377	-1,63512	1,187403	1,317615
В3	0,425590	-2,03075	0,340785	-0,72387	-0,235441	1,565447
В4	-0,665418	1,72286	-0,575624	0,36216	0,679450	0,006837
В5	-0,588727	-0,38739	1,220448	-1,00032	0,165171	1,971707
В6	-0,672300	0,49785	0,540380	-1,25363	0,997707	1,622429
В7	-0,694726	0,64732	0,905450	-1,30568	0,783357	1,365337
В8	-0,645399	1,33950	1,236907	-1,22652	0,527877	0,165570
В9	-0,683475	0,76735	0,951238	-1,33226	0,745492	1,237605

Таблица 9

Матрица взаимных расстояний между центрами кластеров шестого уровня.
Под главной диагональю отображены квадраты расстояний

	Д1	Д2	Д3	Д4	Д5	Д6
Д1	0,000000	2,419893	1,507439	1,036221	1,516820	2,806952
Д2	1,555600	0,000000	1,253324	1,645832	1,248361	1,870631
Д3	1,227778	1,570820	0,000000	1,860829	0,758563	0,975066
Д4	1,073755	2,708761	3,462686	0,000000	3,781297	6,138366
Д5	1,231593	1,117301	0,870955	1,944556	0,000000	0,948721
Д6	1,675396	3,499259	0,987454	2,477573	0,900072	0,000000

Расстояния между кластерами Д3–Д5 и Д3–Д6 менее одной дисперсии, что свидетельствует о достаточной условности разделения объектов, входящих в эти кластеры. Кластеры Д1, Д2 и Д4 определены более достоверно, поскольку удалены от других кластеров на расстояниях больше дисперсии.

Применение кластерного анализа для всего набора статистических показателей, описывающих социально-экономическое положение субъектов СФО, также возможно, но в факторном пространстве по-

сле предварительно проведенного факторного анализа. Необходимо отметить сложность и неоднозначность проблемы интерпретации перемещения объекта в пространстве социально-экономических параметров как отражающее «положительный» или «отрицательный» эффект в развитии соответствующего субъекта. Для формирования подобных заключений необходимо введение для статистических показателей весовых коэффициентов, определяющих положительное направление в пространстве параметров.

Библиографический список

1. Троцковский А.Я., Щетинин М.П. Концептуальные основы регулирования территориального развития на мезоуровне // Известия Алт. гос. ун-та. — 2010. — №2/2.
2. Шваков Е.Е. Особенности формирования и содержание инвестиционной политики в различных субъектах Российской Федерации // Региональная экономика: теория и практика. — 2014. — №4.
3. Вдовкина Е.Г., Мамченко О.П. Об ускоренном внедрении и эффективном использовании информационных технологий как инструмента сглаживания пространственной поляризации // Известия Алт. гос. ун-та. — 2012. — №2/1.
4. Зубаревич Н.В. Социальная дифференциация регионов и городов России [Электронный ресурс]. — URL: <http://gtmarket.ru/laboratory/expertize/5278>.
5. Латышева М.А. Статистическое исследование дифференциации российских регионов по уровню социально-экономического развития // Вестник Волгогр. ун-та. Серия 3 : Экономика. Экология. — 2010. — №1.
6. Социально-экономический мониторинг субъектов РФ Сибирского федерального округа за 2014 год : стат. бюл. / Территориальный орган Федеральной службы государственной статистики по Алтайскому краю. — Барнаул, 2015.
7. Дюран Б. Кластерный анализ. — М., 2012.