И.Б. Шумакова

МНОГОМЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ НАУЧНОГО КАДРОВОГО ПОТЕНЦИАЛА САНКТ-ПЕТЕРБУРГА

I.B. Shumakova

A MULTIVARIATE STATISTICAL ANALYSIS OF THE SCIENTIFIC PERSONNEL POTENTIAL OF SAINT PETERSBURG

Рассматриваются закономерности межгодовых изменений показателей научного кадрового потенциала вузов, учреждений Академии наук и отраслевых институтов за период 1999–2004 гг. с помощью методов факторного и кластерного анализа. Выявлены квазиоднородные группы и репрезентативные показатели, наилучшим образом характеризующие современное состояние кадрового потенциала научных организаций города.

Ключевые слова: высшее образование, научный потенциал, мониторинг, факторный анализ, кластерный анализ, классификация.

Regularities of interannual changes in parameters of the scientific personnel potential of higher educational institutions, establishments of the Academy of Sciences and sectoral institutes are considered with the help of factor and cluster analyses for the period 1999–2004. Quasihomogeneous groups and representative parameters describing the modern state of the personnel potential of the city's scientific organizations in the best way are revealed.

Keywords: higher education, scientific potential, monitoring, factor analysis, cluster analysis, classification.

В течение ряда лет под руководством Комитета по науке и высшей школе Правительства СПб осуществляется мониторинг научного и научно-технического потенциала научных организаций города с целью получения объективной информации о текущем состоянии его научно-технической сферы [Научный…, 1999-2003]. Эта информация может быть использована для принятия управленческих решений и разработке мероприятий, направленных на ее развитие. Отметим, что проведение мониторинга является весьма сложной и трудоемкой задачей, ибо в городе работают 50 государственных и 60 негосударственных высших учебных заведений (вузов), 44 научных учреждения Российской академии наук, 19 научных учреждений Российской медицинской академии, Российской сельскохозяйственной академии и Российской академии образования, а также более 500 отраслевых научных учреждений, включая 12 государственных научных центров.

Мониторинг проводится по нескольким направлениям:

– кадровый потенциал научных организаций;

– научное оборудование;

– финансирование;

– результативность научных исследований и разработок.

Каждое из этих направлений, в свою очередь, было разделено на следующие группы:

- государственные вузы;
- негосударственные вузы;
- учреждения Академии наук;
- отраслевые институты.

При этом кадровый потенциал состоял из 70 показателей, научное оборудование – из 8 показателей, финансирование – из 64 показателей, результативность научных исследований и разработок – из 120 показателей. В сумме получается 262 показателя, по которым осуществлялось их осреднение. Понятно, что прогноз такого количества показателей вряд ли возможен, да и, по-видимому, не нужен. К тому же, не все показатели являются равноценными. Однако «взвесить» их, то есть каждому придать тот или иной вес, чрезвычайно сложно. Ибо данная задача в значительной степени носит неформальный характер. Действительно, в зависимости от поставленной цели каждый из показателей может представлять определенный интерес.

Поэтому разделим задачу выделения наиболее «значимых» показателей на две части: формальную, когда их определение осуществляется тем или иным статистическим способом, и экспертную, когда приоритетные показатели выделяются непосредственно группой экспертов. В рамках данной работы ограничимся анализом кадрового потенциала научно-педагогических организаций города.

Исходной информацией послужили результаты мониторинга за 1999–2004 гг., приведенные в работе [научный, 1999-2003] по тем организациям, которые приняли участие в заполнении анкет в виде формы 3-НК. В табл. 1 приводятся оценки наиболее важных показателей кадрового потенциала научных организаций города за указанные годы. Общая численность научных работников в городе на 2004 г. составила: докторов наук – 5224 чел., в том числе по вузам 3231 чел., или 62 %, кандидатов наук – 17 038 чел., в том числе по вузам 10 502 чел., или 62 %, аспирантов – 12 685 чел., в том числе по вузам 11 344 чел., или 89 %. Следовательно, основной научный потенциал города сосредоточен в вузах, причем эта тенденция в последние годы усиливается.

Из табл. 1 следует, что некоторые показатели имеют прямо противоположные тенденции своих изменений. Так, общее число докторов наук в городе за рассматриваемое шестилетие заметно уменьшилось, что обусловлено аналогичным уменьшением их численности в вузах. Однако одновременно отмечался рост числа докторов наук в отраслевых институтах города. Общая численность кандидатов наук в городе уменьшалась еще более быстрыми темпами. Это также обусловлено главным образом уменьшением численности кандидатов наук в вузах. Если уменьшение числа докторов в значительной степени связано с естественными причинами, т.е. их смертностью, то уменьшение числа кандидатов наук означает в основном их переход в другие сферы деятельности, прежде всего в коммерческие структуры. Отметим также, что в течение 1999–2001 гг. отмечалось заметное уменьшение численности докторов и кандидатов наук, но начиная с 2002 г. происходит их стабилизация.

Таблица 1

Числовые данные кадрового потенциала научных организаций города за 1999–2004 гг., чел.

	1999 г.	2000 г.	2001 г.	2002 г.	2003 г.	2004 г.
Доктора наук: всего	5778	5633	5383	5248	5239	5224
Доктора наук: вузы	3833	3499	3062	2812	2923	3231
Доктора наук: РАН	1136	1059	1124	1321	850	878
Доктора наук: отрасл. ин-ты	809	1075	1197	1115	1466	1115
Кандидаты наук: всего	22433	20844	18256	17693	18506	17038
Кандидаты наук вузы	14572	12836	10125	9630	10744	10502
Кандидаты наук: РАН	2845	2822	2807	3087	2398	2075
Кандидаты наук: отрасл. ин-ты	5016	5186	5324	4976	5364	4461
Аспиранты: всего	12292	13746	13204	12745	14309	12685
Аспиранты: вузы	12256	12987	11554	11184	12610	11344
Аспиранты: РАН	746	769	810	777	319	435
Аспиранты: отрасл. ин-ты	787	805	848	784	1380	906
Доктора наук: госвузы	3567	3270	2795	2670	2658	3073
Кандидаты наук: госвузы	13567	12285	9530	9197	10077	10018
Аспиранты: госвузы	11944	11928	10864	11074	12100	11209
Докторов наук: негосвузы	266	229	267	142	265	158
Кандидаты наук: негосвузы	1005	551	595	433	667	484
Аспиранты: негосвузы	312	370	690	110	510	135
Диссоветы: всего	457	449	404	389	408	366
Диссоветы: вузы	346	329	278	281	294	267
Диссоветы: РАН	40	44	44	49	30	30
Диссоветы: отрасл. ин-ты	71	76	82	59	84	69

Естественно, что для отдельных групп организаций наблюдаются разнонаправленные изменения численности научных работников. К сожалению, выявление их закономерностей стандартными статистическими методами затруднено по двум причинам: во-первых, из-за короткой продолжительности временных рядов, а во-вторых, вследствие того, что выборка в определенной степени не является однородной. Это связано с тем, что количество организаций, принимавших участие в заполнении формы 3-НК, не оставалось строго постоянным год от года. Однако указанные изменения касались в основном малочисленных учреждений, поэтому они не могли существенно сказаться на оценках показателей кадрового потенциала.

Тем не менее, данное обстоятельство следует учитывать и при обработке данных табл. 1 желательно исключить мелкомасштабные шумы, характеризующие, в основном, случайные колебания характеристик, что в результате может позволить выделить некоторые латентные (скрытые) от взгляда исследователя закономерности. Очевидно, наиболее адекватным аппаратом решения данной задачи является метод факторного анализа (ФА). Основная формула факторного анализа в матричном виде записывается следующим образом:

$$\mathbf{X} = \mathbf{F} \cdot \mathbf{A}' + \mathbf{E}, \qquad (1)$$

где X – матрица исходных данных размером m×n (m – число столбцов, n – число строк); **F** – матрица значений общих факторов размером k×n (k ≤ m/2); **A′** – матрица коэффициентов связи общих факторов и исходных переменных размером k×n, называемых факторными нагрузками; **E** – матрица остатков или характерных факторов (k×n).

В настоящее время известно большое число методов факторного анализа [Айвазян, 1998; Афифи, 1982; Болч, 1979; Дубров, 1998; Иберла, 1980]. Наиболее широкое распространение получил *метод главных факторов*, включенный в большинство современных пакетов прикладных статистических программ. Основой его служит метод главных компонент (МГК), с помощью которого решается задача расчета неких вторичных признаков (или функций) в m-мерном пространстве, более эффективно описывающих структурные закономерности, содержащиеся в матрице исходных данных по сравнению с заданными переменными. Вторичные признаки являются линейными комбинациями от исходных переменных и имеют существенно различную дисперсию. Из геометрической интерпретации МГК следует, что переход из исходного пространства признаков в новую систему координат осуществляется простым ортогональным поворотом координатных осей на угол α против часовой стрелки, причем такое вращение происходит без искажения геометрической структуры облака наблюдений в пространстве главных компонент [Айвазян, 1998].

После этого происходит переход из *m*-мерного пространства в *k*-мерное пространство общих факторов, вследствие чего удаляются те оси, вдоль которых наблюдаемая изменчивость не выходит за рамки принятых ошибок.

Затем проводится вторичное вращение осей уже в пространстве *k* общих факторов таким образом, чтобы как можно большее число факторных нагрузок оказалось близким к нулю, а остальные факторные нагрузки, наоборот, были бы максимально приближены к ± 1. Такая процедура получила название *принципа простой структуры*. В результате этого происходит перераспределение дисперсии наблюдений и одновременно искажение геометрической структуры исходных данных. В тех случаях, когда достигается лучшая интерпретируемость полученных результатов, использование ФА можно считать оправданным.

Таким образом, общая структурная схема вычисления общих факторов может быть представлена следующим образом (рис. 1):

X – матрица исходных данных размером m×n;

Z – стандартизованная матрица того же размера;

R – корреляционная матрица размером m×m;

R^h – редуцированная корреляционная матрица размером m×m;

λ – вектор-столбец собственных чисел матрицы R^h длиной m;

A_0 – матрица первоначальных факторных нагрузок размером m×m;

λ^k – вектор-столбец собственных чисел матрицы R^h длиной k;

A – матрица окончательных факторных нагрузок размером k×m;
F – матрица главных факторов размером k×n.

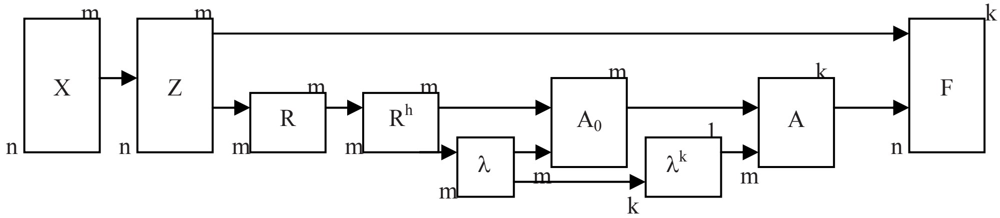

Рис. 1. Структурная схема расчета общих факторов [6]

Переход от обычной корреляционной матрицы к редуцированной состоит в том, что на главной диагонали последней вместо единиц стоят общности. Напомним, *общность* представляет собой долю дисперсии переменных, которая поддается объяснению через общие факторы. Итак, величина общности может быть записана в виде $h^2 = 1 - u^2$, где u^2 – дисперсия характерности. Для определения общности используются несколько способов: метод наибольшей корреляции, метод Барта, метод триад, метод малого центроида, метод множественного коэффициента корреляции [Иберла, 1980; Многомерный …, 1999]. Наиболее часто по умолчанию в пакетах прикладных программ используется последний способ, в соответствии с которым для каждой переменной рассчитывается величина *R* со всеми другими переменными и подставляется на главную диагональ вместо единицы.

Принципиальным моментом ФА является то, что осуществляется процедура вторичного вращения общих факторов с целью улучшения их интерпретируемости. В соответствии с принципом простой структуры ее главная задача, состоящая в оценке достаточности числа поворотов, обычно решается на основе специальных критериев, которые базируются на представлении величины дисперсии факторных нагрузок как меры сложности структуры факторов. С этой целью используются различные критерии, из которых наиболее широкое распространение получил критерий *варимакс Кайзера*. Рассчитанные в результате такого ортогонального вращения значения a_{ij} принимаются в качестве окончательных факторных нагрузок.

Если же при использовании ФА не достигается содержательной интерпретации факторных нагрузок, то применение данного метода теряет смысл. Очевидно, в этом случае следует ограничиться разложением исходной матрицы на главные компоненты, не требующим обязательной интерпретации результатов.

Итак, исходная матрица показателей кадрового потенциала имела размер 6×22 (табл. 1). Нетрудно видеть, что оценки некоторых показателей отличаются друг от друга на несколько порядков. Кроме того, ряды показателей для отдельных лет (n = 22) имеют хорошо выраженное многомодальное распределение. При этих условиях расчет корреляционной матрицы размером 6×6 непосредственно по их исходным значениям теряет смысл. Поэтому предварительно был осуществлен переход к стандартизованным переменным во времени, т.е. по строкам, по формуле вида

$$z = (x - \bar{x})/\sigma,$$

которая обладает тем важным свойством, что при любом распределении случайной величины X среднее значение стандартизованной переменной равно нулю, а дисперсия ее равна единице. Хотя новые ряды стандартизованных показателей по-прежнему носят резко выраженный ассиметричный характер, однако для них уже достаточно корректно может быть осуществлен расчет ранговых коэффициентов корреляции, которые не требуют соответствия исходных данных нормальному распределению. Нами использовались коэффициенты Спирмена, на основе которых вычислялась корреляционная матрица и затем проведен факторный анализ. В табл. 2 приводятся оценки собственных чисел и скорости их сходимости, отражающие вклад в дисперсию исходной матрицы.

Таблица 2

Оценки собственных чисел λ_j и скорости их сходимости, %

Общий фактор	Факторный анализ до вращения (МГК)		Факторный анализ после вращения	
	λ_j	Скорость сходимости, %	λ_j	Скорость сходимости, %
1	2,36	39	2,17	36
2	1,56	26	1,64	27
3	0,62	10	,73	12

Как видно из табл. 2, первые три общих фактора в сумме описывают 75 % дисперсии исходного поля, однако величина третьего фактора меньше единицы. Поэтому, исходя из рекомендации Кайзера, его можно исключить из анализа. Таким образом, ограничимся первыми двумя факторами, которые описывают две трети дисперсии исходной матрицы и, следовательно, наиболее крупномасштабные ее закономерности. В результате фильтруются мелкомасштабные особенности, обусловленные, в том числе неоднородностью выборки.

Факторные нагрузки матрицы показателей кадрового потенциала даны в табл. 3. Смысл их заключается в том, что они показывают степень однородности межгодовых колебаний показателей относительно той доли дисперсии, которую они описывают. Из табл. 3 видно, что если период с 1999 по 2001 г. полностью описываются первым фактором, то период с 2002 по 2004 г. уже вторым фактором. Очевидно, главной тенденцией рассматриваемой совокупности показателей является противоположный характер изменений показателей кадрового состава 1999–2000 гг. и 2001 г. Другими словами, если эти показатели за период с 1999 по 2000 гг. были выше среднего, то в 2001 г. они стали меньше нормы. Второй фактор, как известно, ортогонален (некоррелирован) с первым. Он отражает противоположный характер изменений показателей кадрового состава между 2002 и 2003–2004 гг. Правда, данная тенденция выражена несколько слабее. Итак, имеем 4 типа межгодовых изменений показателей. Близкие по характеру колебаний 1999–2000 и 2003–2004 гг. и резко отличающиеся от них 2001 и 2002 гг. При этом два последних года практически не связаны друг с другом.

Таблица 3

Распределение первых двух факторных нагрузок матрицы показателей кадрового потенциала

Факторная нагрузка	1999 г.	2000 г.	2001 г.	2002 г.	2003 г.	2004 г.
1	-0,857	-0,668	0,682	0,219	0,551	-0,123
2	-0,084	-0,072	-0.342	-0,768	0,691	0,650

Подтверждение сказанному следует из корреляционной матрицы 6×6, приведенной ниже (табл. 4). Действительно, значимая обратная корреляция отмечается для показателей 1999 и 2000 гг. с показателями 2001 г., а также между показателями 2002 и 2003 гг. Если теперь обратиться к исходной матрице (табл. 1), то мы видим, что действительно, в 2001 г. произошло существенное уменьшение количества докторов и кандидатов наук в вузах и в целом по городу, а также уменьшение числа диссоветов в вузах и в целом по городу. Именно этими показателями, очевидно, обусловлена отрицательная корреляция. Положительная корреляция отмечается между 2003 и 2004 гг. и отрицательная этих лет с 2002 г., что также соответствует результатам табл. 3. Однако оценки коэффициентов корреляция в табл. 4 не настолько высоки, чтобы по ним делать однозначные выводы, как это оказалось возможным по оценкам факторных нагрузок в табл. 3 после их второго вращения.

Таблица 4

Распределение выборочных коэффициентов корреляции Спирмена между показателями кадрового состава научных организаций города (n = 22) за отдельные годы периода с 1999 по 2004 г.

	1999	2000	2001	2002	2003	2004
1999	1,000	0,390	-0,572	-0,293	-0,518	0,121
2000	0,390	1,000	-0,598	-0,091	-0,207	-0,076
2001	-0,572	-0,598	1,000	0,137	0,054	-0,455
2002	-0,293	-0,091	0,137	1,000	-0,461	-0,408
2003	-0,518	-0,207	0,054	-0,461	1,000	0,386
2004	0,121	-0,076	-0,455	-0,408	0,386	1,000

Принимая во внимание большое число показателей большой интерес представляет задача выделения квазиоднородных групп показателей и минимального числа «центральных» показателей, которые бы наилучшим образом характеризовали всю совокупность исходных данных в целом. Очевидно, данная задача может быть решена в рамках *кластерного анализа*, позволяющего осуществить разбиение множества объектов (характеристик, показателей и т.д.) на ряд однородных подмножеств по степени их сходства. При этом однородность может трактоваться как максимальная идентичность («похожесть») межгодовых колебаний показателей внутри группы. Для каждой такой группы вначале находится центр «тяжести», наилучшим образом характеризующий колебания группы в целом, а затем показатель, наиболее близкий к центру «тяжести». Этот показатель мы будем считать основным, прогноз которого необходим в первую очередь, ибо он наилучшим образом характеризует всю группу.

В кластерном анализе (КА) понятие однородности двух или нескольких объектов означает близость их физических состояний, которая трактуется как геометрическая близость в многомерном пространстве признаков [Айвазян, 1998]. Данное положение можно рассматривать как основной постулат КА. Менее важным является знание числа выделяемых классов, которое в общем случае до начала классификации может быть и неизвестно. В качестве метрики близости в классическом КА обычно используется какая-либо разновидность расстояния Минковского, наиболее часто – евклидова метрика, имеющая четкий геометрический смысл: геометрическое расстояние в многомерном пространстве признаков.

В данной работе использован алгоритм иерархической агломеративной классификации, суть которой состоит в последовательном объединении объектов (признаков, показателей и т.д.) до тех пор, пока все они не соединятся в единый кластер. Результаты такого объединения представляются в виде дендрограммы. Из иерархических процедур наибольшего предпочтения заслуживает алгоритм Уорда [Малинин, 2003]. Данный алгоритм отличается от других процедур тем, что использует методы дисперсионного анализа для оценки расстояний между кластерами. Он минимизирует сумму квадратов расстояний для двух кластеров, формирование которых происходит на каждом шаге.

Данный метод приводит к образованию кластеров приблизительно равных размеров с минимальной внутриклассовой дисперсией [Многомерный …, 1999]. Этим он близок к эвристическим алгоритмам, минимизирующим внутриклассовую дисперсию (расстояния) и максимизирующим межклассовую дисперсию (расстояния). В целом алгоритм Уорда весьма эффективен, но он стремится создавать кластеры малого размера, что практически не сказывается на качестве классификации при относительно небольших размерах исходной выборки.

Принципиальный недостаток большинства иерархических процедур, в том числе и алгоритма Уорда, заключается в невозможности автоматического выбора оптимального числа классов. Считается, что исследователь должен сам выполнять эту операцию, исходя из каких-либо априорных физических представлений, либо задавая тем или иным способом пороговое (критическое) расстояние. В последнем случае при выполнении численных расчетов довольно часто возникает ситуация неопределенности, когда возможно несколько вариантов определения оптимального числа классов.

На рис. 2 представлена рассчитанная методом Уорда дендрограмма исходных показателей из табл. 1, визуальный анализ которой показывает, что исследуемая совокупность является довольно однородной. От нее отделяются лишь два показателя (кандидаты наук и аспиранты), характеризующие учреждения

Академии наук. Естественно, что подобный результат вряд ли можно признать заслуживающим внимания. Очевидно, содержащиеся в исходных данных закономерности скрыты от наблюдателя и для их нахождения необходимы преобразования, позволяющие выявить латентные факторы, описывающие исходную совокупность.

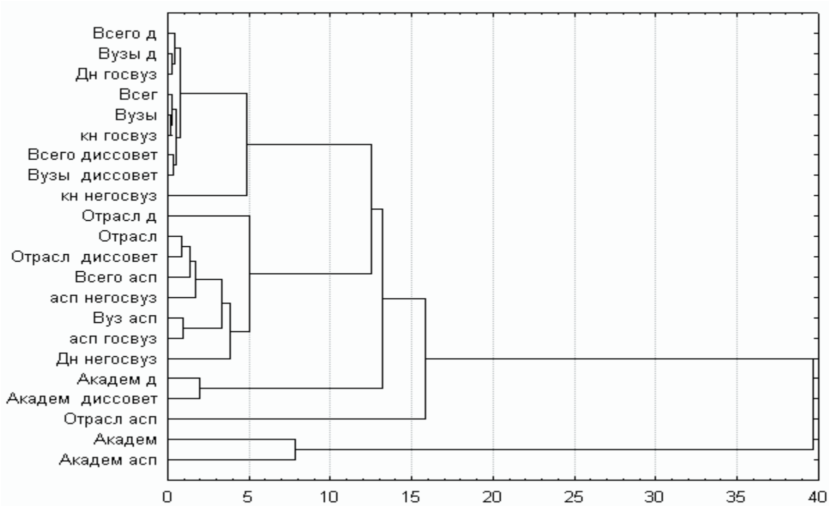

Рис. 2. Дендрограмма показателей кадрового потенциала города, построенная методом Уорда в исходном пространстве признаков

Обратимся теперь к рис. 3, на котором представлена дендрограмма тех же показателей кадрового потенциала в пространстве первых двух факторов. Нетрудно видеть, что она принципиально отличается от дендрограммы, построенной по исходным данным. Даже визуально легко выделить три класса. В данном случае использование критериев качества классификации представляется уже излишним. Дополнительно на рис. 4 приводится распределение показателей кадрового потенциала непосредственно в пространстве первых двух факторов, которое подтверждает, что рассматриваемая совокупность разделяется на три компактных групп точек, отделенные друг от друга значительными (межгрупповыми) расстояниями, превосходящими расстояния внутри групп точек.

В первый класс (нижняя группа на рис. 3) вошли все показатели отраслевой науки, представленной более чем 500 научными учреждениями, а также дополнительно показатели негосударственных вузов, включающих данные 60 учебных заведений. Данный факт представляется весьма любопытным и означает, что показатели кадрового состава негосударственных вузов и отраслевых учреждений в течение шести рассматриваемых лет изменяются близким образом. Однако причины этого требуют дополнительного осмысления. Второй класс (центральная группа на рис. 3) составляют показатели академической науки, представляющей 44 учреждения РАН и 19 учреждений других академий. Наконец, третья группа состоит из показателей 49 государственных вузов, а также показателей, характеризующих итоговые суммы для всего города. Это означает, что именно *государственные вузы*, несмотря на то, что их число значительно уступает научным учреждениям других типов, в конечном счете, дают *решающий вклад в итоговые показатели научного кадрового потенциала города*. Действительно, в вузах работает почти две трети докторов и кандидатов наук, сосредоточено 73 % диссертационных советов от суммарного количества в городе.

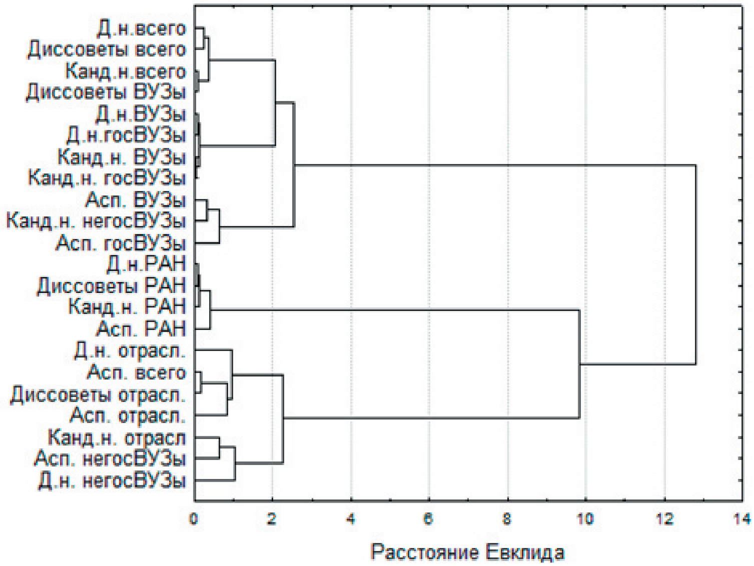

Рис. 3. Дендрограмма показателей кадрового потенциала города, построенная методом Уорда в пространстве первых двух факторов

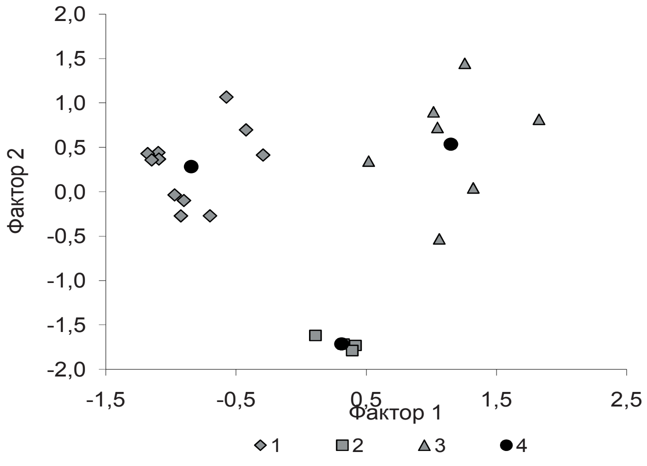

Рис. 4. Распределение показателей кадрового потенциала научных организаций города в пространстве первых двух факторов

Итак, совместное применение факторного анализа и иерархической классификации позволяет однозначно выделить три однородные структуры показателей кадрового потенциала: государственные вузы, академические учреждения, отраслевые организации и негосударственные вузы. На первый взгляд данные результаты, исключая последнюю группировку, достаточно очевидные и тривиальные. Однако это не совсем так. Во-первых, они стали «достаточно очевидными» только в факторном пространстве. Во-вторых, соответствие полученных результатов реально существующим научным структурам повышает степень доверия к использованному нами подходу анализа данных. В-третьих, может быть сформулирован целый ряд новых результатов, непосредственно вытекающих из методологии ФА и КА.

Один их них касается уже рассмотренного выше анализа межгодовой изменчивости показателей кадрового потенциала, невозможной не только в рамках традиционного одномерного статистического анализа, но даже и в рамках теории малых выборок. Кроме того, достаточно очевидным является то, что 1-й фактор описывает преимущественно первую и третью группировки показателей, причем связь между ними противофазная, а 2-й фактор описывает главным образом вторую группировку.

Одновременно с этим представляется возможным получение дополнительных результатов, направленных на оптимизацию системы мониторинга кадрового потенциала. Естественно, для этого необходимо выполнить ранжирование показателей, определить те из них, которые наиболее точно отражают все выделенные группировки в факторном пространстве. Очевидно, с этой целью, прежде всего, следует рассчитать центры тяжести каждой группировки. Хотя это может быть осуществлено несколькими способами, но воспользуемся самым простым – арифметическим осреднением факторов для каждой однородной группировки. После этого определяем расстояние от каждого показателя до центра группы при помощи эвклидовой метрики по формуле

$$D(x_j, x_{j+1}) = \sqrt{\sum_{i=1}^{m}(x_{i,j} - x_{i,j+1})^2} \to \min, \qquad (2)$$

где m – число признаков ($m = 2$).

Очевидно, тот показатель, для которого отмечается минимальное значение эвклидовой метрики до центра тяжести, можно принять в качестве центрального. Для первой группы это отраслевые диссоветы, для второй – академические диссоветы, для третьей группы очень близкие значения евклидовой метрики имеют количество кандидатов наук всех вузов и число диссоветов в вузах. Последний показатель примем в качестве центрального. Итак, *именно число диссертационных советов следует считать, очевидно, самым репрезентативным показателем кадрового потенциала города за период с 1999 по 2004 г., который является интегральным параметром, наилучшим образом, отражающим современное состояние научных организаций города.*

В заключение подчеркнем, что данные результаты следует рассматривать как ориентировочные и относительные, ибо получены по довольно короткой выборке. Вполне возможно, что с увеличением объема выборки взаимосвязи между показателями могут измениться.

Литература

1. *Айвазян С.А., Мхитарян В.С.* Прикладная статистика и основы эконометрики. – М.: ЮНИТИ, 1998. – 1022 с.
2. *Афифи А., Эйзен С.* Статистический анализ: подход с использованием ЭВМ. – М.: Мир, 1982. – 488 с.
3. *Болч Б., Хуань К.* Многомерные статистические методы экономики. – М.: Статистика, 1979. – 317 с.
4. *Дубров А.М., Мхитарян В.С., Трошин Л.И.* Многомерные статистические методы. – М.: Финансы и статистика, 1998. – 350 с.
5. *Иберла К.* Факторный анализ. – М.: Статистика, 1980. – 398 с.
6. *Малинин В.Н., Гордеева С.М.* Физико-статистический метод прогноза океанологических характеристик. – Мурманск: изд. ПИНРО, 2003. – 164 с.
7. Многомерный статистический анализ в экономике / Под ред. В.Н. Тамашевича. – М.: ЮНИТИ-ДАНА, 1999. – 598 с.
8. Научный и научно-технический потенциал Санкт-Петербурга. 1999–2003 гг. – СПб., с.