

УДК 332.122

КЛАССИФИКАЦИЯ РЕГИОНОВ РОССИИ ПО ДИНАМИКЕ ИХ РАЗВИТИЯ В ОКОЛОКРИЗИСНЫЙ ПЕРИОД*

ВИНЮКОВ ИГОРЬ АЛЕКСАНДРОВИЧ,*кандидат экономических наук, доцент кафедры «Математика-1», Финансовый университет, Москва, Россия***E-mail:** vinukov@mail.ru**БЛАГОВЕЩЕНСКИЙ ЮРИЙ НИКОЛАЕВИЧ,***доктор физико-математических наук, профессор, главный специалист Фонда ИНДЕМ, Москва, Россия***E-mail:** vkbun@ya.ru**КОННОВА ЛАРИСА ПЕТРОВНА,***кандидат педагогических наук, доцент кафедры «Математика-1», Финансовый университет, Москва, Россия***E-mail:** lpkonn@mail.ru**МАЕВСКИЙ ЕВГЕНИЙ ВАЛЕРЬЕВИЧ,***кандидат физико-математических наук, доцент кафедры «Математика-1», Финансовый университет, Москва, Россия***E-mail:** emaevskiy@mail.ru**ЯГОДОВСКИЙ ПЕТР ВЛАДИМИРОВИЧ,***кандидат физико-математических наук, доцент кафедры «Математика-1», Финансовый университет, Москва, Россия***E-mail:** korczak@list.ru

АННОТАЦИЯ

Статья посвящена разработке классификации субъектов Российской Федерации, основанной на данных государственной статистики за 2005–2011 гг. и нацеленной на оценку финансовой состоятельности регионов относительно обеспеченности бюджета, привлечения инвестиций, развития экспортно-импортных связей с зарубежными странами и их финансовой политики в отношениях с федеральным центром. Главным достижением является новый подход к формированию интегральных характеристик (индексов), которые одинаково хорошо работают в разные периоды развития экономики (диапазон 2005–2011 гг. содержит подъем, кризис и начальную стадию выхода из него). В итоге были сконструированы четыре индекса, отражающие отдельные грани финансового «благополучия» регионов, и на их базе сформированы достаточно простые критерии классификации. В результате все регионы разделены на шесть хорошо интерпретируемых классов. Показана необусловленность полученной классификации регрессиями по показателям благосостояния и оценены регрессии последних по финансовым индикаторам.

Ключевые слова: классификация субъектов РФ; финансовые показатели; главные компоненты; робастная нормировка; опорные регионы; бюджетная обеспеченность; рейтинг.

CLASSIFICATION OF RUSSIAN REGIONS ON THE DYNAMICS OF DEVELOPMENT IN THE YEARS 2005–2011**

IGOR A. VINYUKOV*PhD, Associate Professor of the Department «Mathematics 1» Financial University under the Government of the Russian Federation, Moscow, Russia***E-mail:** vinukov@mail.ru

* Статья подготовлена по результатам исследований, выполненных за счет бюджетных средств по государственному заданию Финансового университета в 2013 г.

** This article was prepared by the research, performed by the budget on the state task Finuniversiteta 2013.

YURY N. BLAGOVESCHENSKY

Doctor of physico-mathematical sciences, professor, chief specialist The INDEM FOUNDATION, Moscow, Russia

E-mail: vkbun@ya.ru

LARISA P. KONNOVA

PhD, Associate Professor of the Department «Mathematics 1» Financial University under the Government of the Russian Federation, Moscow, Russia

E-mail: lpkonn@mail.ru

EVGENY V. MAEVSKIY

PhD, Associate Professor of the Department «Mathematics 1» Financial University under the Government of the Russian Federation, Moscow, Russia

E-mail: emaevskiy@mail.ru

PETR V. YAGODOVSKY

PhD, Associate Professor of the Department «Mathematics 1» Financial University under the Government of the Russian Federation, Moscow, Russia

E-mail: korczak@list.ru

ABSTRACT

The article is devoted to the development of classification of the subordinate entities of the Russian Federation based on the data of the State statistical reporting for 2005–2011 and aimed at assessment of financial solvency of regions concerning budget backing, attracting of investments, promoting export-import relationship between foreign countries and their fiscal policy in the relations with the federal center. The main achievement is a new approach to formation of integrated characteristics (indexes) which work equally well during different periods of development of the economy (range 2005–2011 contains upswing, crisis and an initial stage of its surmounting). As a result 4 indexes reflecting separate sides of affluence of regions were designed, and common criteria of classification were created on their base. As a result all the regions were divided into 6 well interpreted classes. Shows non-conditionality the classification by regressions on indicators of well-being and evaluated regression of recent on financial indicators.

Keywords: classification of subjects of the Russian Federation, the financial indicators, principal components, a robust normalization, the reference regions, fiscal capacity, rating.

ВВЕДЕНИЕ

Для развивающихся стран, а Россия относится к их числу, быстрый экономический рост сопровождается, как правило, усилением неравенства между разными территориями в пределах одной страны [1]. Проведение федеральной региональной политики требует понимания тенденций развития территорий, анализа ситуации в краткосрочном и долгосрочном планах. Региональные власти также должны отслеживать тренды в развитии соседних территорий и подобных себе регионов, но на коротких интервалах наблюдений, 2–3 года, почти невозможно оценить характер развития регионов, сопоставить их между собой, выделить регионы, которые остро нуждаются в помощи для преодоления кризисных

явлений. Все это актуализирует проблемы создания классификаций субъектов Российской Федерации под самые разные целевые задачи: бюджетные потоки, развитие инфраструктуры и освоение территории, финансовые потоки, демографические проблемы и многое другое, требующее длительных наблюдений, мониторинга.

Наше исследование нацелено на создание классификации, которая базируется на данных государственной статистики (Росстат) за достаточно длинный период времени (2005–2011) и в целом характеризует регионы по сходству и различию таких латентных характеристик, как финансовая состоятельность регионов относительно обеспеченности бюджета, привлечения инвестиций, развития экспортно-импортных связей с зарубежными

странами и их финансовой политики в отношениях с федеральным центром.

Исследование опирается в определенной степени на работу Фонда ИНДЕМ [2] и на идеи, содержащиеся в работах [3, 4]. Мы существенно модернизировали экспертно-статистический метод опорных регионов из книги [2] в двух аспектах:

1) предложен новый подход к выделению устойчивых во времени групп лидеров и аутсайдеров;

2) применен новый метод сокращения размерности, в котором главные компоненты используются не как оси в новом пространстве, а лишь как инструмент для выбора разных подходящих расстояний в исходном пространстве, каждому из которых ставится в соответствие с помощью специальных процедур своя ось. Тем самым создается пространство меньшей размерности (по числу разных расстояний) (см. также [9]).

Новизна полученных результатов (по сравнению, например, с работами [5, 6, 7, 8]) состоит в том, что взятый нами период наблюдений еще никем не исследовался — это семь лет, 2005–2011 гг., которые содержат и период быстрого роста ВВП России, и жесткий кризисный 2009 г., и восстановительный период 2010 и 2011 гг. Далее была разработана и реализована на статистических данных новая технология классификации, отражающая разные аспекты финансовой состоятельности регионов.

Наконец, были разработаны новые приемы для исследования связи уже созданной классификации с показателями, характеризующими в той или иной мере благосостояние населения в регионах [3, 4].

Мы полагаем, что сформулированные в ходе исследования идеи конструирования трех-четырех индексов, в совокупности дающих адекватное представление о финансовой состоятельности регионов, могут дать хороший инструмент мониторинга и, тем самым, индикатор для федеральных чиновников, опираясь на который они смогут вовремя реагировать на сложные события финансовой жизни регионов.

КЛАССИФИКАЦИИ СУБЪЕКТОВ РОССИЙСКОЙ ФЕДЕРАЦИИ ПО ФИНАНСОВЫМ ПОКАЗАТЕЛЯМ МЕТОДОМ ОПОРНЫХ РЕГИОНОВ

В основу настоящего исследования положены статистические данные Росстата¹ за 2005–2011 гг. из сборников «Регионы России» и «Финансы России». Из имеющихся данных был скомпонован набор *финансовых индикаторов*, ставших рабочим материалом для построения классификации регионов РФ, и ряд *служебных индикаторов*, использованных для исследования связи между общими характеристиками благополучия регионов и их финансовыми позициями в разрабатываемой классификации.

Индикаторы, характеризующие финансовую деятельность регионов, формировались как относительные величины, по которым можно было бы сравнивать субъекты РФ независимо от их размеров (население) или экономического потенциала (ВРП — валовой региональный продукт). В итоге нами было сконструировано 11 финансовых индикаторов (табл. 1).

Из приведенного списка требуют пояснения всего два индикатора — w_5 и w_7 .

Первый из них, w_5 — задолженность по кредитам в расчете на одного человека, тыс. руб., является пересчетом суммы двух показателей: кредиты, предоставленные физическим лицам *в рублях*, млн руб., и кредиты, предоставленные физическим лицам *в валюте*, млн руб.

Чтобы учесть «продленное» действие инвестиций, за *уровень инвестиций* w_7 было взято взвешенное среднее за три года в пересчете на одного человека: если объем *прямых инвестиций* за текущий год и два предыдущих обозначить как $I(*)$, $I(*-2)$ и $I(*-1)$, то сначала вычислялось

$$I_{CP}(*) = (2 \times I(*-2) + 3 \times I(*-1) + 5 \times I(*)) / 10,$$

а затем

$$w_7 = I_{CP}(*)R(*)/P,$$

¹ Росстат — краткое название Федеральной службы государственной статистики РФ.

Список финансовых индикаторов

Код	Определение индикаторов через показатели
w_1	Сальдо (прибыль минус убыток) деятельности организаций в процентах от дохода консолидированного бюджета
w_2	Доход консолидированного бюджета на одного занятого в экономике, тыс. руб.
w_3	Расходы консолидированного бюджета на душу населения, тыс. руб.
w_4	Разность между доходом и расходами в процентах к доходу
w_5	Задолженность по кредитам в расчете на одного человека, тыс. руб.
w_6	Инвестиции в основной капитал на душу населения, тыс. руб.
w_7	Уровень иностранных инвестиций в расчете на душу населения, тыс. руб.
w_8	Экспорт в расчете на душу населения, тыс. руб.
w_9	Импорт в расчете на душу населения, тыс. руб.
w_{10}	Поступление налогов и других платежей в бюджет РФ в процентах от дохода
w_{11}	Задолженность по налогам и сборам в бюджет РФ в процентах от ВРП

где P — численность населения, тыс. человек;

$R^{(*)}$ — средний курс доллара в расчетном году.

Служебные индикаторы представлены в табл. 2.

Отметим, что 11 индикаторов ($y_1, y_2, y_3, y_{10} — y_{13}$ и $y_{17} — y_{20}$) были взяты из статистических таблиц Росстата без изменений, а остальные 10 являются расчетными: для возможностей сравнительного анализа регионов они были очевидным образом нормированы.

При проведении классификационного анализа выведены из рассмотрения 11 регионов, по разным причинам резко выделяющихся из общего ряда. Это две агломерации — Московская и Санкт-Петербургская; нефтяная провинция — Тюменская область с входящими в нее Ханты-Мансийским и Ямало-Ненецким автономными округами; два богатых полезными ископаемыми региона, добыча которых в значительной степени производится приезжими из других регионов вахтовым методом, — Ненецкий и Чукотский автономные округа; три бедных региона, во многом живущих за счет дотаций, — Республики Чеченская, Ингушская и Дагестан. В итоговой

классификации они подключаются к сформированным без них классам.

Таким образом, изначально классифицировалось 72 региона. С каждым из них связано 77 чисел — значения 11 финансовых индикаторов в каждом из рассматриваемых годов, иными словами, каждый регион задается точкой в 77-мерном пространстве.

Для обработки столь значительного объема данных обычно используется та или иная процедура снижения размерности, например за счет введения небольшого числа *индексов*, функций от индикаторов, чаще всего — линейных. Мы последуем этой идее и попробуем вместо 11 индикаторов подобрать существенно меньшее число характеристик (индексов), следя за тем, чтобы потери информации при такой операции были несущественными.

Для решения этой задачи на первом этапе были выделены две группы регионов — 5 *лидеров* и 5 *аутсайдеров*. Выделение было произведено путем анализа рангов всех 72 регионов — отдельно по каждому из 11 индикаторов, причем для w_{11} — *по убыванию*, а для всех остальных — *по возрастанию*. И это было сделано по каждому году из 7 лет, так что каждый отдельный регион получил по 77 рангов

Список дополнительных показателей

Код	Показатели
y_1	Площадь территории, тыс. кв. км
y_2	Численность населения на конец года, тыс. человек
y_3	Среднемесячная номинальная начисленная зарплата работников организаций, руб.
y_4	Добыча полезных ископаемых, тыс. руб. на душу населения
y_5	Обрабатывающие производства, тыс. руб. на душу населения
y_6	Производство и распределение электроэнергии, газа и воды, тыс. руб. на душу населения
y_7	Продукция сельского хозяйства, тыс. руб. на душу населения
y_8	ВРП на душу населения, тыс. руб.
y_9	Число малых предприятий на тысячу человек
y_{10}	Средние цены на вторичном рынке жилья, руб. за 1 кв. м
y_{11}	Удельный вес городского населения, доля от всего населения, %
y_{12}	Население моложе трудоспособного возраста, доля от всего населения, %
y_{13}	Население старше трудоспособного возраста, доля от всего населения, %
y_{14}	Общая безработица к числу занятых в экономике, %
y_{15}	Превышение общей безработицы над регистрируемой, раз
y_{16}	Среднегодовая численность занятых в экономике, % от всего населения
y_{17}	Среднедушевые доходы населения, руб.
y_{18}	Численность населения с доходами ниже прожиточного минимума, %
y_{19}	Потребительские расходы в среднем на душу населения, руб.
y_{20}	Число собственных легковых автомобилей на тысячу человек, шт.
y_{21}	Индекс Джини, %

(низкие ранги указывают на слабость региона, а высокие — на силу). Дополнительно были использованы данные рейтингового агентства «Эксперт РА» и ряда других источников. Результат этой работы приводится в табл. 3.

Отметим, что имеется несколько регионов (Кабардино-Балкарская, Карачаево-Черкесская Республики, Северная Осетия, Калмыкия и Адыгея), которые, несомненно, являются аутсайдерами, но их статистические данные по ряду позиций экстремальны, так что использовать их как основу для статистического анализа не представлялось возможным. Затем осреднением по лидерам и аутсайдерам были получены *опорные* (эталонные) *регион-лидер* и *регион-аутсайдер*.

Подчеркнем, что это *виртуальные* регионы, у которых значения индикаторов в каждый год равны среднему арифметическому значению *лидеров* для *региона-лидера* и *аутсайдеров* для *региона-аутсайдера*. Оказалось, что полученные таким образом опорные регионы подходят для дальнейшего анализа, так как хорошо разнесены и определяют направленность от «плохих» регионов к «хорошим».

На следующем шаге происходит построение индексов. Геометрически идея вычисления индекса проста, ее иллюстрирует *рис. 1*, где точки — это регионы. Опорные регион-лидер *B* и регион-аутсайдер *A* соединены координатной осью *AB* (пунктирная стрелка на

Таблица 3

Группы лидеров и аутсайдеров

Лидеры	Аутсайдеры
Иркутская обл.	Алтайский край
Липецкая обл.	Ивановская обл.
Татарстан	Курганская обл.
Самарская обл.	Пензенская обл.
Свердловская обл.	Марий Эл

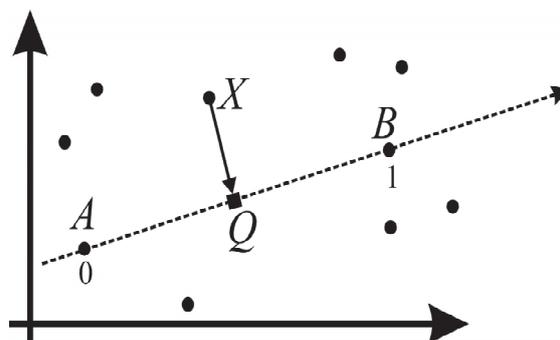


Рис. 1. Построение индекса по методу опорных регионов

рис. 1), шкала на которой такая, что аутсайдер A имеет координату 0, а лидер B — координату 1. Пусть точка Q — проекция точки X на ось AB ($XQ \perp AB$). Значение индекса для региона X равно координате точки Q на оси AB . Такой подход впервые был использован в книге [2].

В нашем случае каждый регион в каждый год задается точкой $X = (x_1, x_2, \dots, x_{11})$, где x_k — значение k -го индикатора в этом году, $1 \leq k \leq 11$. Пусть $B = (b_1, b_2, \dots, b_{11})$ и $A = (a_1, a_2, \dots, a_{11})$ — опорные регионы аутсайдер и лидер. Тогда значение индекса для региона X равно:

$$Pr_{L(A,B)}(X) = \frac{d^2(A, X) + d^2(A, B) - d^2(B, X)}{2d^2(A, B)}, \quad (1)$$

где $d(X, Y)$ — расстояние между точками-регионами X и Y в 11-мерном пространстве. Очевидно, что значение индекса существенно зависит от того, как именно определено расстояние $d(X, Y)$, так что, варьируя ими, можно конструировать различные индексы.

Припишем k -му индикатору вес $\mu_k \geq 0$ (сумма весов равна единице) и определим квадрат расстояния между $X = (x_1, x_2, \dots, x_{11})$ и $Y = (y_1, y_2, \dots, y_{11})$ равенством:

$$d^2(X, Y) = \sum_{1 \leq k \leq 11} (\mu_k \cdot (x_k - y_k))^2, \quad (2)$$

где $1 \leq k \leq 11$.

Теперь нужно для каждого индекса выбрать веса так, чтобы построенный набор индексов был как можно более информативным. При

этом следует помнить, что определяющие индексы веса будут использоваться на протяжении всех 7 лет, т.е. индексы должны оставаться «хорошими» в течение всего этого периода, несмотря на то что геометрическая конфигурация множества всех регионов существенно менялась.

Для подбора весов были использованы четыре главные компоненты, объясняющие примерно $\frac{3}{4}$ общей дисперсии: веса выбирались в соответствии с абсолютными значениями факторных нагрузок по результатам анализа 2010 и 2011 гг. как наиболее представительные для наших целей. Результаты выбора весов приведены в табл. 4.

Оказалось, что индексы, значения которых зависят от региона и года, построенные из чисто математических соображений, обладают достаточно экономически ясными интерпретациями, вытекающими из смысла индикаторов с ненулевыми весами:

- J_1 — общее состояние финансового благополучия;
- J_2 — бюджетная обеспеченность;
- J_3 — экспертно-инвестиционное благополучие;
- J_4 — использование импорта в развитии.

Каждый из них для данного региона и года вычисляется по формуле (1), в которой функция расстояния $d(X, Y)$ определяется формулой (2) с соответствующими весами из табл. 4, а в качестве координат отвечающей региону точки $X = (x_1, x_2, \dots, x_{11})$ берутся нормированные значения² индикаторов $w_1, \dots,$

² Мы использовали робастную нормировку, при которой значение индикатора конкретного региона в конкретный год заменяется разностью этого значения и медианы (всех регионов в текущий год), де-

Веса индикаторов для четырех индексов

№ индекса	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}
1	0,051	0,128	0,142	0	0,102	0,157	0,131	0,142	0,05	0,052	0,045
2	0,18	0,184	0,154	0	0	0	0	0	0,102	0,241	0,139
3	0,188	0	0	0,246	0,204	0,056	0,178	0,128	0	0	0
4	0,169	0	0	0	0,132	0	0	0	0,699	0	0

w_{12} для этого региона за этот год (нормируются и опорные регионы — лидер и аутсайдер). Каждый из этих индексов растет, если регион «перемещается» из района *аутсайде-ра* в направлении *лидера*, а потому мы можем перейти к рангам и снабдить каждый регион четырьмя порядковыми номерами $R(J_1), R(J_2), R(J_3), R(J_4)$, согласно тому месту, которое он занимает при упорядочивании по убыванию соответствующего индекса, а затем вычислить *взвешенный средний ранг*:

$$R_{CP} = \frac{4R(J_1) + 2R(J_2) + 3R(J_3) + R(J_4)}{10}.$$

Веса при рангах $R(J_1), \dots, R(J_4)$ выбирались в основном пропорциональными четырем наибольшим собственным значениям и немного корректировались³ под целые числа. Напомним, что на базе главных компонент, соответствующих этим собственным значениям, и были построены индексы.

Сформулированная в настоящей работе классификация основана на взвешенном среднем ранге R_{CP} . Обозначим через $R_n(t)$ значение этого ранга для региона n в год t . Тогда рост $R_n(t)$ по t указывает на ухудшение финансовых дел в регионе n , а убывание $R_n(t)$ — наоборот, на улучшение ситуации. Определим теперь для каждого региона две характеристики:

$$V_n = \max([R_n(2005) + R_n(2011)] / 2; R_n(2011)) —$$

ленной на *междециль* (половину расстояния между верхним и нижним децилями).

³ При этом вес индекса «Экспортно-инвестиционное благополучие» был усилен до числа 3, поскольку это направление деятельности региональных администраций является для субъектов РФ и России в целом одним из важнейших.

показатель хороших рейтингов и в начале периода, и в его конце с приоритетом 2011 г.;

$\Delta_n = R_n(2005) - R_n(2011)$ — *показатель динамики*: чем больше Δ_n , тем лучше оценивается регион в 2011 г. по сравнению с 2005 г.

Все регионы были разделены сначала на пять групп:

- 1) $V_n < 30$;
- 2) $30 < V_n < 60$ и $\Delta_n \geq 8$;
- 3) $30 < V_n < 60$ и $\Delta_n \leq -8$;
- 4) $30 < V_n < 60$ и $-8 < \Delta_n < +8$;
- 5) $V_n > 60$.

Первая группа — сильные регионы — состоит из 26 регионов, пять из которых явно выпадают из общего ряда. Это *Якутия* с алмазодобывающей промышленностью, *Сахалинская область* с проектом «Сахалин-2» и мощным заводом сжиженного газа, *Камчатский край* и *Магаданская область* с разнообразными и уникальными природными богатствами, *Калининградская область* с добычей янтаря и морским портом.

С точки зрения региональной политики, эти пять регионов можно рассматривать только индивидуально, не смешивая с другими субъектами РФ. Их объединяет лишь то, что каждый из них, будучи сильным, практически не имеет ничего общего с другими регионами России. Поэтому был выделен дополнительный класс «Нестандартно сильные регионы», в который вынесены эти пять субъектов.

По всем признакам к этому классу при-
мыкают и временно исключенные из анализа

Таблица 5

Состав класса «Стандартно сильные регионы»

Регионы		
Ленинградская обл.	Свердловская обл.	Новгородская обл.
Красноярский край	Пермский край	Самарская обл.
Республика Коми	Иркутская обл.	Республика Карелия
Белгородская обл.	Липецкая обл.	Томская обл.
Республика Татарстан	Вологодская обл.	Республика Хакасия
Мурманская обл.	Хабаровский край	Республика Башкортостан
Кемеровская обл.	Оренбургская обл.	Нижегородская обл.

регионы: Москва, Санкт-Петербург, Московская и Тюменская области, а также Ненецкий, Чукотский, Ханты-Мансийский и Ямало-Ненецкий автономные округа. Остальные регионы из первой группы отнесены нами к классу «Стандартно сильные регионы» (табл. 5).

Это достаточно развитые регионы, экономика которых имеет солидную базу: или полезные ископаемые, как Хакасия и Кемеровская область, или промышленное производство, как Липецкая и Самарская области, или инфраструктура и транспорт, как Мурманская область с морским портом и производством апатитового концентрата.

Вторая группа — это следующий класс, «Динамичная середина», восемь крепких середняков, позиции которых растут от года к году, даже кризис их почти не затронул, и они быстро вернулись к своим темпам роста (табл. 6).

Затем идет группа регионов-середнячков с хорошим потенциалом развития, но неустойчивыми позициями. Они тяжело перенесли кризисное время и полностью от его последствий не восстановились. В этот класс, названный «Тормозящая середина», вошли восемь субъектов РФ. Это Республики Алтай и Мордовия плюс шесть областей: Архангельская, Челябинская, Омская, Псковская, Астраханская и Орловская.

Следующая группа определяет класс «Среднеслабые регионы». Среди субъектов РФ из этого класса имеются регионы с разной

Таблица 6

Регионы, вошедшие в класс «Динамичная середина»

Регионы	
Калужская обл.	Курская обл.
Амурская обл.	Брянская обл.
Краснодарский край	Владимирская обл.
Приморский край	Ульяновская обл.

степенью «слабости», например, растущий потенциал наблюдается у Смоленской области, может «подняться» и Тульская область. Однако в этих регионах нет ни производственного, ни ресурсного капитала, который был бы исключительно важен для Федерации в целом, а потому представителям этого класса необходимы больше внимания центральных органов и новые идеи развития (табл. 7).

Последний класс — это «Слабые по существу регионы», для которых выполняется условие $V_n > 60$. В этот класс попадают все регионы-аутсайдеры и сюда следует отнести Чеченскую Республику, Дагестан и Ингушетию, хотя данные по ним неполные. Этот класс, возможно, следовало бы разделить на две или три группы с учетом специфики регионов (табл. 8).

Таблица 7

Состав класса «Среднеслабые регионы»

Регионы		
Тульская обл.	Смоленская обл.	Воронежская обл.
Новосибирская обл.	Рязанская обл.	Саратовская обл.
Волгоградская обл.	Удмуртская Республика	Кировская обл.
Ростовская обл.	Забайкальский край	Чувашская Республика
Республика Бурятия	Еврейская автономная обл.	Республика Тыва
Ярославская обл.	Тверская обл.	Ставропольский край

Таблица 8

Состав класса «Слабые по существу регионы»

Регионы	
Курганская обл.	Ивановская обл.
Алтайский край	Карачаево-Черкесия
Республика Марий Эл	Республика Адыгея
Тамбовская обл.	Республика Калмыкия
Костромская обл.	Пензенская обл.
Кабардино-Балкарская Республика	Республика Северная Осетия

СОПОСТАВЛЕНИЕ РЕЗУЛЬТАТОВ С ДАННЫМИ КЛАСТЕРНОГО АНАЛИЗА

Параллельно построению своей классификации регионов РФ авторы изучили, что может быть получено при обработке входных данных иными методами, в частности кластерным анализом.

Входные данные представляют собой значения для регионов РФ 11 финансовых индикаторов, среди которых есть сильно коррелированные. С помощью компонентного анализа можно получить из них меньшее число некоррелированных показателей, объясняющих большую часть суммарной вариации всего исходного набора показателей. Эти показатели являются собственными векторами корреляционной матрицы исходных индикаторов, соответствующими собственным значениям больше единицы. Оказалось, что в каждый год таких показателей ровно четыре. Доля дисперсии исходных данных, которая объясняется этими показателями, составляет от 73,6% в 2005 г. до 76,7% в 2011. Этот результат означает, что набор из 11 финансовых индикаторов в каждом из рассматриваемых

годов можно с незначительными искажениями спроектировать в некоторую (для каждого года свою) четырехмерную плоскость. В каждый год набор показателей свой: при переходе от одного года к другому репер собственных векторов с четырьмя собственными значениями больше единицы несколько поворачивается. Однако экономический смысл показателей год от года сохраняется. Так, показатель F_1 характеризует общее финансовое благополучие региона, показатель F_2 указывает на снижение поступлений налогов и рост дохода, расходов и задолженности по налогам, показатель F_3 указывает на прибыльность, рост поступлений налогов, увеличение экспорта, а F_4 — на отсутствие недоимок, профицитность и доступность кредитов.

Помимо годовых индикаторов, был также рассмотрен двухлетний набор 22 значений по 72 регионам за 2010 и 2011 г. Оказалось, что в этом случае существуют шесть значимых показателей (с дисперсиями больше единицы), которые вместе объясняют свыше 86% общей суммарной дисперсии финансовых индикаторов.

Сопряженность результатов кластерного анализа на шесть кластеров за 2011 г. с аналогичными результатами за 2010 г. (численность регионов в кластерах)

Распределение 72 регионов по кластерам двух кластеризаций и их пересечениям	По кластерам 2010 г. (объемы)					
	23	18	14	10	4	3
По кластерам 2011 г. с объемами: 24	9	8	4	-	2	1
24	11	5	6	2	-	-
18	3	3	4	7	1	-
3	-	2	-	-	1	-
2	-	-	-	-	-	2
1	-	-	-	1	-	-

Таким образом, двухлетний массив финансовых индикаторов можно с незначительными искажениями спроектировать в некоторую шестимерную плоскость. При этом четыре показателя с наибольшими дисперсиями по-прежнему объясняют $\frac{3}{4}$ общей дисперсии финансовых индикаторов, поэтому ниже именно они будут использованы для кластерного анализа.

Показатели F_1, F_2, F_3 и F_4 в этом случае имеют тот же экономический смысл, что и выше. Показатель F_5 растет с увеличением задолженности по налогам, а F_6 — при уменьшении разности между доходом и расходами.

Для каждого года кластерный анализ финансовых индикаторов был проведен для массивов значений четырех значимых показателей F_1, F_2, F_3 и F_4 с помощью метода k -средних с обычным евклидовым расстоянием между наблюдениями — регионами, так как обычно число получаемых типов регионов в исследованиях принадлежит интервалу 6–9. Поэтому и здесь рассматривались только два значения для числа кластеров 6 и 9.

В результате шесть кластеров в 2005 г. были образованы 34, 20, 12, 4, 1 и 1 регионами; в 2006 г. — 26, 18, 12, 10, 5 и 1; в 2007 г. — 21, 20, 14, 10, 6, 1; в 2008–40, 15, 10, 5, 1 и 1; в 2009–27, 23, 13, 5, 3 и 1; в 2010–23, 18, 14, 10, 5 и 2; в 2011 г. — 24, 24, 18, 3, 2 и 1 регионами. Видно, что размеры кластеров в разные годы абсолютно отличаются. Также хаотично происходят перестройки систем кластеров в разные годы. Для иллюстрации представим

сопряженность систем кластеров за 2010 и 2011 гг. (табл. 9).

Аналогичная картина имеет место и с девятью кластерами.

Кластерный анализ разделения 72 регионов на шесть кластеров по первым четырем показателям для двухлетних данных 2010 и 2011 гг. показал кластеры с числом регионов 32, 19, 15, 3, 2 и 1. Сопряженность этого результата и шестикластерного анализа по однолетним (за 2011 г.) данным приведена в табл. 10.

Примечательно, что три самых малых кластера полностью совпали.

Обратите внимание, что все регионы наибольшего кластера по данным 2010–2011 гг. (кроме одного) разошлись по двум крупнейшим кластерам по данным 2011 г. Были сделаны рисунки проекций наибольшего кластера по данным 2010–2011 гг. на координатные плоскости, задаваемые всевозможными парами компонент для 2010–2011 гг., но всякий раз регионы, которые должны разойтись по разным кластерам, оказывались перемешанными. Разделение видно на проекциях на координатные плоскости F_1F_4 и F_2F_4 по данным 2011 г. (рис. 2).

Объединение этих регионов в один кластер при учете двухлетней динамики 2010–2011 г. указывает на существование различий в классификациях регионов по статическим и динамическим оценкам их деятельности. Учет двухлетней динамики смазывает и затушевывает различия, видимые на годовых данных. Поэтому построение классификаций, учитывающих динамику, представляется,

Таблица 10

Сопряженность двухлетней (2010–2011 гг.) и однолетней (2011 г.) кластеризаций

Распределение 72 регионов по кластерам двух кластеризаций и их пересечениям	По кластерам 2010–2011 гг. (объемы)					
	32	19	15	3	2	1
По кластерам 2011 г. с объемами: 24	15	-	9	-	-	-
24	16	5	3	-	-	-
24	1	14	3	-	-	-
18	-	-	-	-	-	-
3	-	-	-	3	-	-
2	-	-	-	-	2	-
1	-	-	-	-	-	1



Рис. 2. Распределение наибольшего кластера 2010–2011 гг. по наибольшим кластерам 2011 г.

несомненно, интересной и заслуживающей внимания задачей.

Для сравнения результатов кластерного анализа по статическим данным за два года (2010–2011) и за последний, 2011 г., с классификацией регионов, построенной методом опорных регионов по семилетним данным за период 2005–2011 гг., были созданы таблицы сопряженности их кластеров (табл. 11, 12).

Итак, как было указано ранее [7], видно, что результаты кластерного анализа за разные годы резко различаются, причем различия

нарастают по мере увеличения временного разрыва между сравниваемыми кластеризациями. Поэтому методом кластерного анализа сформировать сколько-нибудь общую типологию, объединяющую ряд лет, невозможно. Вместе с тем кластерный анализ помещает в одну «корзину» (кластер) абсолютно разные регионы. Так, при построении шести кластеров по данным 2011 г. в один кластер были помещены такие непохожие регионы, как Хакасия и Пензенская область! Такого рода странности обусловлены, скорее всего, тем, что в кластерном анализе использовалось

Таблица 11

Сопряженность двухлетней (по данным 2010–2011 гг.) и полученной методом опорных регионов (по данным 2005–2011 гг.) кластеризаций

Распределение 72 регионов по кластерам двух кластеризаций и их пересечениям		Кластеры 2010–2011 гг. (объемы)					
		32	19	15	3	2	1
Из них по кластерам 2005–2011 гг. с							
названием и	объемом	2	16	3	-	-	-
Стандартно сильные регионы	21						
Среднеслабые	18	12	2	4	-	-	-
Слабые по существу регионы	12	8		4	-	-	-
Тормозящая середина	8	5	1	2	-	-	-
Динамичная середина	8	5	-	2	-	1	-
Нестандартно сильные регионы	5	-	-	-	3	1	1

Таблица 12

Сопряженность однолетней (по данным за 2011 г.) и полученной методом опорных регионов (по данным 2005–2011 гг.) кластеризаций

Распределение 72 регионов по кластерам двух кластеризаций и их пересечениям		Кластеры 2011 г. (объемы)					
		24	24	18	3	2	1
Из них по кластерам 2005–2011 гг. с							
названием и	объемом	-	4	17	-	-	-
Стандартно сильные регионы	21						
Среднеслабые	18	9	9	-	-	-	-
Слабые по существу регионы	12	10	2	-	-	-	-
Тормозящая середина	8	3	5	-	-	-	-
Динамичная середина	8	2	4	1	-	1	-
Нестандартно сильные регионы	5	-	-	-	3	1	1

евклидово расстояние, в котором не учитывается разная роль разных индикаторов. Однако определенное сопряжение кластеризации регионов по данным 2011 г. с классификацией методом опорных регионов случается, и это в известной мере подтверждает объективный характер последней.

РЕГРЕССИОННЫЕ СВЯЗИ ПОКАЗАТЕЛЕЙ БЛАГОСОСТОЯНИЯ С ФИНАНСОВЫМИ ИНДИКАТОРАМИ

Дополнительно было проведено исследование регрессионных связей показателей благосостояния в зависимости от совокупности финансовых индикаторов w_1, w_2, \dots, w_{12} . Для

исследования были взяты 16 показателей благосостояния $Y_3, Y_8, Y_9, Y_{10}, Y_{11}, Y_{12}, Y_{13}, Y_{14}, Y_{15}, Y_{16}, Y_{17}, Y_{18}, Y_{19}, Y_{20}, Y_{21}$ и u , где новый усредненный показатель u определяется формулой

$$u = \frac{Y_4 + Y_5 + Y_6 + Y_7}{Y_2}.$$

Цель состояла в том, чтобы для каждого из показателей благосостояния выяснить, существует ли линейная регрессия, связывающая значение этого показателя со значениями финансовых индикаторов. Если да, то регрессионная зависимость должна выполняться для всех регионов. Понятно, что в разные годы линейная регрессия будет разной, поэтому на втором этапе рассматривался вопрос об устойчивости регрессии на протяжении рассматривавшегося промежутка времени. Мы полагали, что линейная регрессия *хорошо* описывает показатель в данный год, если относительная погрешность не превышает 10%, и *посредственно* в противном случае. Для оценки устойчивости рассматривалась регрессия для каждого показателя, дополнительно рассматривалась усредненная регрессия (коэффициентами которой были средние значения коэффициентов регрессий для этого показателя за все годы) и вычислялась ее средняя относительная погрешность по всем регионам за все годы. Если она не превышала 10%, то регрессия для данного показателя считалась *устойчивой*, в противном случае — *неустойчивой*.

Из *табл. 13* видно, что регрессии для большинства показателей или хорошо описывают показатель каждый год, или посредственно. Случаев, когда год от года качество регрессий меняется, сравнительно немного. Это позволяет утверждать, что возможность или невозможность хорошо описать показатель благосостояния линейной регрессией по совокупности финансовых индикаторов носит не случайный характер, а является объективным свойством самого показателя, обусловленным его собственной природой.

НЕУДАЧА СОЗДАНИЯ НОВЫХ ПОКАЗАТЕЛЕЙ, ОПРЕДЕЛЯЮЩИХ ПОЛУЧЕННЫЕ КЛАССЫ РЕГИОНОВ

В завершение исследования была рассмотрена

задача о возможности построения пары показателей (критериев), которые определяли бы ту же классификацию регионов, что была построена выше методом опорных регионов. Существующие стандартные методы построения критериев классификаций, к сожалению, оказались неприемлемыми в силу специфики нашей ситуации.

Идея заключается в том, чтобы подобрать два показателя, каждый из которых представляет собой линейную функцию от показателей благосостояния (см. перечень в предыдущем разделе), соответствующим образом нормированных. Первый показатель выбирается так, чтобы разброс его значений в регионах одного класса был как можно меньше (в среднем по всем шести классам). Второй показатель должен быть таким, чтобы разброс значений по регионам из разных классов был как можно больше. Затем предлагается характеризовать каждый регион парой чисел — значениями двух показателей. Цель состоит в том, чтобы посмотреть, разделятся ли в этом двумерном пространстве регионы на те же шесть классов.

Чтобы иметь возможность сравнивать значения показателей, необходимо их нормировать. Была использована та же нормировка, что и выше: значение индикатора региона в данный год нормируется на разность этого значения и *медианы* значений индикатора в этот год, деленную на *междециль* (половину расстояния между верхним и нижним децилями). Пусть v_{ijk} — нормированное значение j -го показателя в i -й год на k -м регионе. Определим новый показатель ρ по формуле

$$\rho = r_1 v_{i,1,k} + r_2 v_{i,2,k} + \dots + r_{16} v_{i,16,k},$$

где r_1, \dots, r_{16} — некоторые числовые коэффициенты.

Пусть K_1, \dots, K_6 — кластеры регионов в классификации, проведенной методом опорных регионов. Для каждого класса регионов K_p рассмотрим сумму квадратов разностей $(v_{ik}r - v_{il}r)$ по всем парам регионов l и k ($k > l$) из класса K_p , разделим полученное выражение на $\frac{|K_p|(|K_p|-1)}{2}$ (это число пар регионов l и k в классе K_p) и просуммируем по всем классам. В итоге получится функционал:

Существование и устойчивость регрессионных связей показателей благосостояния с финансовыми индикаторами

№ п/п	Индикаторы	Существование регрессии	Устойчивость регрессии
1	y_3	Каждый год, кроме 2005 г., хорошая регрессия. В 2005 г. регрессия посредственная	Регрессия неустойчивая
2	u	Каждый год посредственная регрессия	Регрессия неустойчивая
3	y_8	Каждый год, кроме 2005 г., хорошая регрессия. В 2005 г. регрессия посредственная	Регрессия неустойчивая
4	y_9	Каждый год посредственная регрессия	Регрессия неустойчивая
5	y_{10}	Каждый год посредственная регрессия	Регрессия неустойчивая
6	y_{11}	Каждый год хорошая регрессия	Регрессия практически устойчивая
7	y_{12}	Каждый год хорошая регрессия	Регрессия неустойчивая
8	y_{13}	Каждый год хорошая регрессия	Регрессия неустойчивая
9	y_{14}	Каждый год посредственная регрессия	Регрессия неустойчивая
10	y_{15}	Каждый год посредственная регрессия	Регрессия неустойчивая
11	y_{16}	Каждый год хорошая регрессия	Регрессия устойчивая
12	y_{17}	В 2005–2007 гг. и в 2009 г. посредственная регрессия, в остальные годы хорошая регрессия	Регрессия неустойчивая
13	y_{18}	Каждый год посредственная регрессия	Регрессия неустойчивая
14	y_{19}	Каждый год посредственная регрессия	Регрессия неустойчивая
15	y_{20}	Каждый год, кроме 2008 г., посредственная регрессия. В 2005 г. регрессия хорошая	Регрессия неустойчивая
16	y_{21}	Каждый год хорошая регрессия	Регрессия устойчивая

$$m(r) = \sum_{p=1, \dots, 6} \frac{2}{|K_p|(|K_p|-1)} \sum_{\substack{k>l \\ k,l \in K_p}} (v_{ik}r - v_{il}r)^2,$$

который следует минимизировать.

Аналогично конструируется функционал:

$$M(r) = \sum_{p>q} \frac{1}{|K_p||K_q|} \sum_{\substack{k \in K_p \\ l \in K_q}} (v_{ik}r - v_{il}r)^2,$$

который следует максимизировать.

Здесь рассматриваются сумма квадратов разностей $(v_{ik}r - v_{il}r)$ по всем парам регионов l и k ($k > l$) из классов K_p и K_q , деленная на число таких пар $|K_p||K_q|$ и просуммированная по всем парам различных классов p и q . Оба этих функционала являются положительными полуопределенными квадратичными формами.

Минимум первого функционала достигается, если коэффициенты r_1, \dots, r_6 образуют собственный вектор матрицы квадратичной

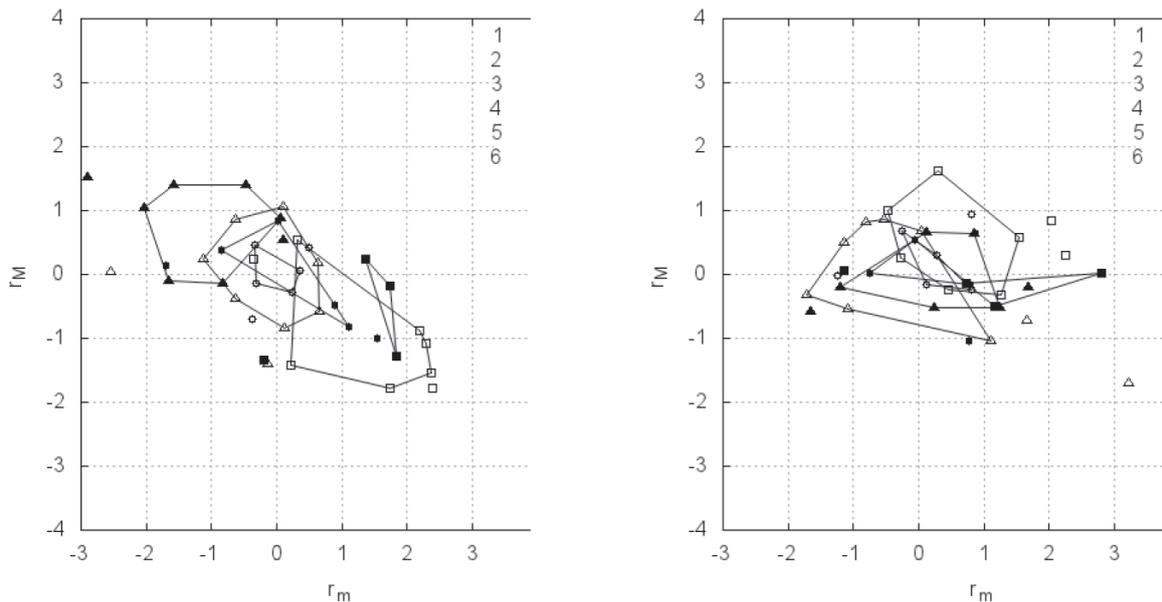


Рис. 3. Выпуклые оболочки классов в 2008 и 2011 г.

формы $m(r)$, соответствующий ее минимальному собственному значению. Максимум второго функционала достигается, когда коэффициенты r_1, \dots, r_{16} образуют собственный вектор матрицы квадратичной формы $M(r)$, соответствующий ее максимальному собственному значению. Таким образом, можно определить два набора коэффициентов вида r_1, \dots, r_{16} , первый — оптимальный для $m(r)$, а второй — для $M(r)$. Эти наборы задают пару показателей ρ_m и ρ_M , для каждого года свою. Значения показателей ρ_m и ρ_M в регионе определяют его координаты в текущем году. Для каждого года были построены диаграммы рассеяния регионов в координатах ρ_m и ρ_M . Затем в каждом классе был определен центр (как среднее арифметическое координат); два региона, наиболее удаленные от центра, не включались в наименьший выпуклый многоугольник (выпуклую оболочку) остальных регионов класса. Примеры результатов для 2008 и 2011 гг. приведены на рис. 3.

Оказалось, что регионы, входящие в определенный класс, лежат достаточно кучно (в основном, за исключением класса нестандартно сильных регионов). Но при этом многоугольники, соответствующие различным классам, существенно пересекаются. Поэтому получить заданную классификацию на основе проведенного анализа нефинансовых индикаторов не представляется возможным.

ЛИТЕРАТУРА

1. Scott A. J. Regional Push: Towards a Geography of Development and Growth in Low-and Middle-Income Countries // *Third World Quarterly*. Vol. 23. 2002. № 1. P. 137–161.
2. Региональная политика России: адаптация к разнообразию: Труды Фонда ИНДЕМ /Под общ. ред. Г.А. Сатарова). М.: Фонд ИНДЕМ, 2004.
3. Айвазян С.А. К методологии измерения синтетических категорий качества жизни населения // *Экономика и математические методы*. 2003. Т. 39. № 2. С. 33–53.
4. Российские регионы: экономический кризис и проблемы модернизации/ под ред. Л.М. Григорьева, Н.В. Зубаревич, Г.Р. Хасаева. М.: ТЭИС, 2011.
5. Винюков И.А., Крылова И.Б. Многомерная финансовая типология регионов России: Препринт. М.: Финансовый университет, 2011.
6. Иванова Е.И., Перекрест В.Т. Применение методов проблемной типологии субъектов РФ в задачах оценки эффективности информационного взаимодействия органов государственной власти и территориальных органов государственной статистики. СПб: Центр стратегического анализа общественных процессов, 2006.

7. Бутс Б., Дробышевский С., Кочеткова О., Мальгинов Г. и др. Типология российских регионов. М.: СЕРРА, 2002.
8. Бездудный М.А., Кучинский К.А., Пастухов Е.С. Типология регионов России по уровню развития банковской деятельности // Банковское дело. 2003. № 11. С. 33–42; № 12. С. 17–21.
9. Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Классификация и снижение размерности. М.: Финансы и статистика, 1989.

REFERENCES

1. Scott A.J. Regional Push: Towards a Geography of Development and Growth in Low-and Middle-Income Countries. *Third World Quarterly*, vol. 23, 2002, no. 1, pp. 137–161.
2. Regional'naiia politika Rossii: adaptatsiia k raznoobraziuu: Trudy Fonda INDEM (pod obshch. red. G.A. Satarova) [Regional Policy of Russia: Adaptation to a Variety of Proceedings of the INDEM Foundation (Ed. GA Satarov)]. Moscow: INDEM Foundation, 2004. (in Russ.)
3. Aivazian S.A. K metodologii izmereniia sinteticheskikh kategorii kachestva zhizni naseleniia [Methodologies for measuring population life quality categories]. *Ekonomika i matematicheskie metody — Economics and Mathematical Methods*, 2003, vol. 39, no. 2, pp. 33–53. (in Russ.)
4. Rossiiskie regiony: ekonomicheskii krizis i problemy modernizatsii/ pod red. L.M. Grigor'eva, N.V. Zubarevich, G.R. Khasaeva [Russian regions: the economic crisis and the problems of modernization. Ed. L.M. Grigorieva, N.V. Zubarevich, G.R. Khasaev], Moscow: TEIS, 2011. (in Russ.)
5. Vinyukov I.A., Krylova I.B. Mnogomernaia finansovaiia tipologia regionov Rossii: Preprint. Moscow, Finansovyi universitet, 2011.
6. Ivanova E.I., Perekrest V.T. Primenenie metodov problemnoi tipologizatsii sub"ektov RF v zadachakh otsenki effektivnosti informatsionnogo vzaimodeistviia organov gosudarstvennoi vlasti i territorial'nykh organov gosudarstvennoi statistiki [Application of methods of problem typology of subjects of the Russian Federation to the problems of assessing the effectiveness of information between state authorities and territorial state statistics bodies], SPb: Tsentr strategicheskogo analiza obshchestvennykh protsessov — St. Petersburg: Center for Strategic Analysis of social processes, 2006. (in Russ.)
7. Buts B., Drobyshevskii S., Kochetkova O., Mal'ginov G. i dr. Tipologiiia rossiiskikh regionov [The typology of Russian regions]. Moscow: СЕРРА, 2002. (in Russ.)
8. Bezduunyi M. A., Kuchinskii K. A., Pastukhov E. S. Tipologiiia regionov Rossii po urovniu razvitiia bankovskoi deiatel'nosti [The typology of Russian regions in terms of development of banking activities]. *Bankovskoe delo — Banking*, 2003, no. 11, pp. 33–42; no. 12, pp. 17–21. (in Russ.)
9. Aivazian S.A., Bukhshtaber V.M., Eniukov I.S., Meshalkin L.D. Prikladnaia statistika: Klassifikatsiia i snizhenie razmernosti [Applied Statistics: Classification and dimension reduction]. Moscow: Finansy i statistika — Finance and Statistics, 1989. (in Russ.)