

**СТАТИСТИЧЕСКИЙ АНАЛИЗ ДЕЯТЕЛЬНОСТИ БАНКОВ С ПРИМЕНЕНИЕМ
УСОВЕРШЕНСТВОВАННОЙ КЛАСТЕРИЗАЦИИ****Оксана Станиславовна ДОЦЕНКО**

кандидат экономических наук, доцент кафедры бухгалтерского учета, анализа и аудита,
Севастопольский государственный университет, Севастополь, Российская Федерация
plakiddin@mail.ru
ORCID: отсутствует
SPIN-код: 8275-6226

История статьи:

Получена 07.05.2018
Получена в доработанном
виде 18.05.2018
Одобрена 29.05.2018
Доступна онлайн 27.07.2018

УДК 330.101.52**JEL:** C1, C6, G2**Аннотация**

Предмет. Исследование направлено на расширение способов анализа деятельности банков, характеризующихся множеством показателей. Акцент сделан на возможности прогнозной оценки принципов деятельности отдельно взятого банка и построении однородных классификаций банков с точки зрения принципов их работы в современных условиях. Предметом исследования являются процедуры статистического и математического анализа банковской деятельности.

Цели. Совершенствование способов многомерной группировки банков и прогнозирования их деятельности с применением математических моделей: кластеризация на основе меры сходства между банками-скалярами – взвешенного евклидова расстояния и векторного видоизменения кластерного анализа на основе меры сходства – угла наклона между банками-векторами, характеризующимися взвешенными показателями.

Методология. Комплекс статистических и математических приемов: статистическое наблюдение, метод группировки банков по отдельным показателям и по комплексу показателей с применением факторного, кластерного, дискриминантного и векторного анализа, метод обобщающих показателей, методы анализа временных рядов и взаимосвязей между показателями.

Результаты. Разработаны обобщенные классификации банков по трем признакам – размеру активов, взвешенному евклидову расстоянию и углу наклона между взвешенными характеристиками банков, что позволяет избежать противоречий в группировке объектов и уточнить принципы их работы по кластерам с учетом размеров активов и динамики по углам наклона. Проведен прогноз политики отдельно взятого банка с использованием векторного представления его показателей, что позволило оценить устойчивость этого банка на протяжении исследуемого промежутка времени. Приемы такого анализа могут быть использованы как руководством самих банков, так и клиентами и контрагентами банковских учреждений при выборе устойчивого и надежного банка.

Ключевые слова: банки, кластерный анализ, взвешенное евклидово расстояние, вектор, угол наклона

Выводы. Представленная методика прогрессивнее известных классических процедур. В результате ее применения группировка банков по комплексу показателей получается более однородной и интерпретация принципов банковской деятельности точнее. Она достаточно проста и оперирует открытыми данными, используя которые, можно с большой степенью вероятности выбрать наиболее надежный банк.

© Издательский дом ФИНАНСЫ и КРЕДИТ, 2018

Для цитирования: Доценко О.С. Статистический анализ деятельности банков с применением усовершенствованной кластеризации // *Экономический анализ: теория и практика*. – 2018. – Т. 17, № 7. – С. 1352 – 1366.
<https://doi.org/10.24891/ea.17.7.1352>

Банки, представляющие основу банковской системы, выступают центром аккумуляции банковских ресурсов и их перераспределения, предоставляют клиентам сотни видов банковских продуктов и услуг. Их роль возрастает еще больше в условиях развития и становления рыночных отношений, когда они работают в целях получения максимальных

доходов наряду с обслуживанием клиентов по линии мобилизации и направления финансовых ресурсов [1, с. 4]. Основой эффективного управления банками, исходной базой принятия управленческих решений, выбора наиболее устойчивого банка является статистический анализ деятельности кредитного учреждения. С помощью такого анализа разрабатывают стратегию и тактику развития банков, реально и всесторонне оценивают достигнутые результаты деятельности. Наибольшую актуальность статистический анализ деятельности банков приобретает в современных условиях, в постоянно меняющейся общеэкономической и социально-политической ситуации, что влияет на их надежность и политику.

Значительный вклад в исследование разных аспектов банковской деятельности внесли такие отечественные и зарубежные ученые, как Э.Дж. Долан, К.Д. Кемпбелл, М.З. Бор, Н.А. Головач, В.Н. Кочетков¹, П. Роуз, В.В. Иванов, Дж.Ф. Синки, Р.С. Портер и др. [2–8]. Этими авторами проведено разделение всех приемов статистико-экономического анализа банковской деятельности на три группы – логические, элиминирования и математические. Логические способы позволяют выявить только общий характер и направление связи и развития явления. Способы элиминирования применяются для определения влияния отдельных факторов на изменение результативного показателя посредством устранения воздействий всех факторов, кроме одного. Математические способы предназначены для решения оптимизационных задач различных аспектов деятельности банков. Особое место здесь отводится многомерному факторному анализу, который необходим для комплексной оценки банковской деятельности. Такая задача встает всегда, когда нужно, например, дать обобщенную рейтинговую оценку результатам деятельности нескольких банков. Поскольку местоположение банков по разным показателям будет неодинаковым (например, по активам банк будет занимать первое место,

по капиталу – второе, а по финансовому результату – пятое и т.д.), многомерный анализ проводится не только по абсолютным и относительным показателям, но и по евклидову расстоянию, которое позволяет учитывать степень близости банков к банку-эталону.

Несмотря на то что сами банки эффективно внедряют прогрессивные принципы управления, информационные технологии и программно-технические средства практически во всех областях своей деятельности, выбранная тема исследования недостаточно изучена. В особенности это касается способов анализа банковской системы в целом, когда банки представляют собой большую совокупность экономических объектов, каждый из которых характеризуется множеством показателей. Установлено, что в большинстве своем подходы к такому анализу деятельности банков основаны на применении элементарных статистических методов к отдельным параметрам без учета их совокупностей. Так, российские ученые М.В. Ершов и В.М. Зубов, прогнозируя стратегию государства и банковской системы, предлагают использовать показатели операционной эффективности и рентабельности, которые сравнивают с показателями других стран и по отраслям в России [9]. Российский экономист М.В. Ключников предлагает сравнивать у банков различные показатели ликвидности и делать выводы о финансовом положении конкретного банка в сравнении с другими банковскими учреждениями [10]. Журнал «Бизнес» периодически производит ранжирование коммерческих банков по одному из показателей (по чистым активам или по капиталу) с последующей группировкой по ранжированному показателю. Другие часто публикуемые методики основываются на расчетах темпов роста и прироста банковских активов, кредитов, депозитов, прибыли и рентабельности. Анализ динамики показателей деятельности коммерческих банков производится как по отдельным банкам (преимущественно первой двадцатки лидеров по активам), так и по общей совокупности

¹ Кочетков В.Н., Омельченко А.В. Економічний аналіз діяльності комерційних банків і підприємств. Київ: Скарби, 2001. 334 с.

банков. Кроме того, для исследования банковской системы экономистами применяются методы индексного, коэффициентного анализа и анализа структурных изменений [11–14].

Очевидно, что простота в расчетах и экономическом обосновании полученных данных – существенный фактор при анализе больших объемов информации. Однако погрешности в этих расчетах, приводящие при определенных условиях к недостоверным прогнозам, неверная оценка политики банков и в результате – ошибочное представление об их надежности и устойчивости – проблема как для руководства самого банка, так и для его клиентов.

Мониторинг существующих методик дает возможность понять, что даже самые авангардные исследования в этой области с использованием сложных экономико-математических методов нуждаются в дальнейшей доработке. В особенности это касается применения технологий многомерной классификации банков по определенным показателям с использованием кластерного анализа [15]. Первые исследования в направлении многомерных классификаций проводились еще в 1960-х гг. американскими классиками математической статистики Т. Андерсоном, М.Д. Кендаллом [16, 17]. В настоящее время методы группировок на основе множества показателей рассматривают в своих работах, например, В.В. Майба [18], Н.П. Тихомиров² и др.

Сегодня экономистами используется упрощенная мера сходства между объектами – евклидово расстояние, в формуле которого не учитывается вес того или иного показателя, характеризующего банк (или другой экономический объект). Одним из последствий такого упрощенного анализа, по нашему мнению, является попадание в один кластер, казалось бы, совершенно различных по значению показателей банков, что трудно поддается объяснению. В результате возникает задача создания более совершенной методики

анализа деятельности банков на основе открытой и доступной информации, чтобы расширить способы статистической оценки банковской деятельности. Это позволит учитывать роль каждого показателя банковской деятельности в общей политике банка, будет способствовать более точному и обоснованному распределению банков по кластерам.

Основой для исследования функционирования банковской системы явился комплексный подход к статистическому анализу деятельности банков. Нами разработана методика статистического исследования банковской системы любого государства на базе открытой и доступной информации (на примере деятельности группы банков), основанная на двух принципах:

- на углубленном статистическом анализе (с использованием усовершенствованных методов многомерного статистического анализа – кластерного анализа и его векторного видоизменения), который предусматривает сравнение комплекса показателей деятельности банковских учреждений в динамике;
- на прогнозной оценке политики банков с точки зрения способности сохранять свои существенные свойства на заданном уровне в течение фиксированного промежутка времени при определенных условиях.

Основным преимуществом этой методики является возможность сопоставить деятельность различных банков по небольшому количеству обобщенных параметров, что позволит руководству банков-конкурентов в любой момент времени оценить деятельность соперников, а клиентам выбрать наиболее надежный банк для выгодного вложения денег.

Для построения данной методики использовались следующие статистические приемы: статистическое наблюдение и упорядочение, суть которых сводилась к выбору совокупности банков с определенным набором показателей их деятельности, метод группировки банков по отдельным

² Тихомиров Н.П., Тихомирова Т.М., Урмаев О.С. Методы эконометрики и многомерного статистического анализа. М.: Экономика, 2011. 647 с.

показателям и по комплексу показателей с применением факторного, кластерного и дискриминантного анализа, метод обобщающих статистических показателей, который используется при нормировании данных; методы анализа взаимосвязей между показателями; методы анализа временных рядов для определения дальнейших принципов работы банков. Для расчетов применялись стандартные пакеты прикладных статистических программ.

Предлагаемая методика является универсальной. Ее алгоритм может быть использован для установления большей прозрачности работы любой банковской системы. Достаточно оперировать большими объемами информации – совокупностью объектов (банков), характеризуемых множеством показателей. Исследования проводились по случайной выборке из 70 российских банков с 2011 по 2017 г.³ В целях демонстрации самой методики и акценте на ее универсальности и применимости к любому государству (и не только к банкам, но и к любой экономической совокупности объектов, характеризуемых множеством показателей) названия банков были заменены их порядковыми номерами.

В проведенном исследовании каждый из банков характеризуется десятью показателями-скалярами:

- активами x_1 ;
- обязательствами x_2 ;
- кредитно-инвестиционным портфелем x_3 ;
- кредитами юридическим лицам x_4 ;
- кредитами физическим лицам x_5 ;
- собственным капиталом x_6 ;
- уставным капиталом x_7 ;
- депозитами физических лиц x_8 ;
- депозитами юридических лиц x_9 ;
- финансовым результатом x_{10} .

Эти показатели, по нашему мнению, полно и объективно характеризуют политику банков. Пример представлен в *табл. 1*.

Во избежание большого разброса и потери данных был применен метод нормирования по следующему алгоритму:

- нахождение математического ожидания для каждого показателя по всем банкам $mean(x)$;
- нахождение среднеквадратического отклонения $\sigma = std(x)$;
- центрирование значений $x_{центр} = x - mean(x)$;
- нормирование значений $x_{норм} = x_{центр} / std(x)$.

Пример представлен в *табл. 2*.

Это позволило рассматривать относительные (безразмерные) показатели деятельности банков, с помощью которых производились дальнейшие расчеты. Таким же образом были составлены таблицы нормированных значений для показателей деятельности банков по состоянию на 2011 г. и на 2017 г.

Для определения общих закономерностей и принципов деятельности банковских учреждений предложена их многомерная группировка с помощью метода кластеризации (классификации) банков на основании меры сходства между ними – взвешенного евклидова расстояния с использованием следующей формулы⁴:

$$d = \sqrt{\sum_{n=1}^{10} w_n \left(x_n^{(i)} - x_n^{(j)} \right)^2},$$

где w_n – весовые коэффициенты при показателях деятельности коммерческих банков;

$x_n^{(i)}$ и $x_n^{(j)}$ – нормированные значения n -го показателя для i -го и j -го банков.

Необходимость такой усовершенствованной кластеризации генеральной совокупности была вызвана тем, что на сегодняшний день группирование действующих банков

³ Banki.ru. URL: <http://www.banki.ru>

⁴ Елисеева И.И., Юзбашев М.М. Общая теория статистики. М.: Финансы и статистика, 2002. 480 с.

осуществляется либо только по одному показателю (как правило, по размерам активов), что делает совокупность неоднородной с точки зрения других показателей, либо по множеству равноправных показателей без учета роли (веса) этих показателей (как правило, на основании меры сходства – невзвешенное евклидово расстояние), что делает классификацию менее совершенной:

$$d = \sqrt{\sum_{n=1}^{10} (x_n^{(i)} - x_n^{(j)})^2}.$$

Для определения весов всех показателей деятельности банков применялись следующие процедуры:

- факторный анализ, где была проанализирована взаимосвязь между десятью показателями деятельности банков и установлено, что параметр «кредитно-инвестиционный портфель» наиболее тесно связан с остальными параметрами. В результате этого было составлено уравнение множественной регрессии:

$$x_3 = a_0 + a_1x_1 + a_2x_2 + a_3x_4 + \dots + a_9x_{10},$$

где x_3 – нормированное значение кредитно-инвестиционного портфеля;

x_1, \dots, x_{10} – нормированные значения остальных показателей деятельности банков;

- множественная регрессия, где были найдены стандартизированные коэффициенты регрессии β , на основании которых определены веса каждого показателя⁵:

$$w_n = \frac{\beta_i^2}{\sum_1^n \beta_i^2}.$$

В ходе исследования (с помощью кластерного анализа) совокупность была разбита на четыре группы (четыре кластера – A, B, C, D) банков на основании следующих мер сходства: взвешенное евклидово расстояние по десяти

показателям, невзвешенное евклидово расстояние по десяти показателям, невзвешенное евклидово расстояние по одному показателю – активы. Вероятность попадания каждого банка в определенный кластер определялась с помощью дискриминантного анализа. Целью таких разбиений было дальнейшее сопоставление полученных результатов (в результате кластеризации) друг с другом и с результатами обычного группирования банков по активам на четыре группы (банки-гиганты, крупные банки, средние банки и мелкие банки), которое часто осуществляется экономистами. Наглядное расхождение в результатах представлено для кластера A на *рис. 1*.

Анализ банков остальных трех кластеров B, C и D , сформированных на основе различных мер сходства, показал невысокий процент соответствий группировок по кластерам B и (особенно) C и высокое соответствие (но не абсолютное) по самому многочисленному кластеру D . Очевидно, что уточнение весов группировочных признаков дает изменения в кластеризации, которая является более обоснованной, чем кластеризация на основании меры сходства между объектами – невзвешенное евклидово расстояние (об этом свидетельствует развернутый авторский анализ: в кластеры по d_3 крупные банки попадают к крупным, а мелкие к мелким. Необоснованное распределение банков здесь встречается, но реже, чем в кластеризации по d_2). Что касается кластеризации по d_1 и группировки по активам, то такие группировки не являются однородными и были представлены с точки зрения демонстрации расхождений в результатах исследований.

Таким образом, предложенные ранее способы классификации объектов не дают пользователям достаточной информации относительно деления совокупности на сравнительные группы. Даже усовершенствованная кластеризация (по d_3) имеет существенный недостаток, связанный, по нашему мнению, с ограниченным набором показателей (которые являются скалярными величинами) и специфичностью формулы,

⁵ Елисеева И.И., Юзбашев М.М. Общая теория статистики. М.: Финансы и статистика, 2002. 480 с.

определяющей заданную меру сходства, что не позволяет в достаточной степени обосновывать отнесение некоторых банков к определенной группе. Это может привести к ошибочной интерпретации их политик, например, как показали дальнейшие исследования структуры показателей деятельности банков по группам, в один кластер могут попасть как очень крупные, так и средние банки (рис. 2).

На рис. 2 кластеры представлены графиками средних значений показателей деятельности банков по состоянию на 2017 г. В кластере С вместе с банком-гигантом № 2 находятся хоть и крупные, но все же значительно уступающие ему по своим показателям другие банки. Именно поэтому кластер разделен на два графика, где на одном представлен только банк № 2 с крупнейшими показателями, а на другом – все остальные банки, попавшие в этот кластер, у которых по всей совокупности показатели значительно меньше, чем у банка № 2. В то же время в разные кластеры могут попасть банки, принципы действия которых очень похожи (кластеры А, В и оставшиеся банки из кластера С).

В связи с этим нами предложено векторное видоизменение кластерного анализа – группирование банков на основании дополнительной меры сходства – угла наклона φ между банками-кортежами (векторной интерпретацией взвешенных показателей деятельности банковских учреждений), которое более точно определяет направление (близость) политик исследуемых банков.

В основе этой теории лежит геометрическое представление объектов (банков), определенных не точками-скалярами в n -мерном пространстве (где мера сходства между ними – евклидово расстояние), а векторами, имеющими направления и угол наклона между собой.

В аналитической геометрии показано, что объект с безразмерными характеристиками вида (x_1, x_2, \dots, x_n) – далее n -мерный банк-кортеж – может рассматриваться как вектор $\vec{x}^{(1)}$ в n -мерном пространстве. Начинается этот вектор в начале координат – точка $(0, 0, \dots, 0)$,

а заканчивается в точке с координатами (x_1, x_2, \dots, x_n) . Этих векторов может быть несколько. Допустим, мы рассматриваем два банка-кортежа с одинаковым набором характеристик (показателей). При такой интерпретации разность между векторами представляет собой вектор $\vec{d}_{1,2}$:

$$\vec{d}_{1,2} = \vec{x}^{(2)} - \vec{x}^{(1)}.$$

Здесь появляется новое качество: если при скалярном представлении банков-кортежей разница между ними характеризуется только скаляром d – евклидовым расстоянием между кортежами, то при векторном представлении банков-кортежей появляются новые характеристики, а именно углы наклонов векторов \vec{d} , $\vec{x}^{(2)}$, $\vec{x}^{(1)}$ к осям координат и друг относительно друга. Для дальнейших исследований нам особенно важен угол φ между банками-кортежами $\vec{x}^{(2)}$, $\vec{x}^{(1)}$, который позволяет прояснить в определенной степени политику этих банков.

Например, один банк-кортеж представлен в виде

$$\vec{x}^{(1)} = (x_1, x_2, \dots, x_{10}),$$

а другой банк-кортеж представлен в виде

$$\vec{x}^{(2)} = (y_1, y_2, \dots, y_{10}).$$

Если $y_i = kx_i$, где k – коэффициент, то при $k > 0$ векторы $\vec{x}^{(2)}$, $\vec{x}^{(1)}$ ориентированы в одном направлении

$$\vec{x}^{(2)} \uparrow\uparrow \vec{x}^{(1)}, \varphi = 0^\circ,$$

то есть политики банков похожи; при $k < 0$ векторы противоположно направлены

$$\vec{x}^{(2)} \uparrow\downarrow \vec{x}^{(1)}, \varphi = 180^\circ,$$

то есть политики банков противоположны.

Если $x_1y_1 + x_2y_2 + \dots + x_{10}y_{10} = 0$, то векторы перпендикулярны

$$\vec{x}^{(2)} \perp \vec{x}^{(1)}, \varphi = \pm 90^\circ,$$

то есть политики банков индифферентны, никак не связаны друг с другом.

Все остальные соотношения дают другие углы, в той или иной мере близкие к указанным.

Определим степень статистической связи между кортежами. Для этого применим операцию скалярного произведения между векторами и введем понятие начального смешанного момента для двух банков-кортежей. В результате получаем следующую формулу:

$$\cos \varphi = \frac{M_{11}(\vec{x}^{(1)}, \vec{x}^{(2)})}{\sqrt{M_2(\vec{x}^{(1)}) M_2(\vec{x}^{(2)})}},$$

где M_{11} и M_2 – ковариационные начальные моменты второго порядка.

Правая часть полученной формулы – это коэффициент ковариации. Он показывает степень статистической связи между кортежами (или векторами). Если коэффициент ковариации равен нулю, то связь между кортежами отсутствует; если равен единице, то связь функциональная; если близок к единице, то связь сильная; если близок к нулю, то связь слабая; если отрицателен, то связь обратная. Следовательно, угол φ также дает возможность оценить степень статистической связи. Близкая к этому интерпретация статистической связи между объектами, характеризующимися одинаковыми показателями, впервые была дана академиком А.Н. Колмогоровым в монографии [19]. Однако в области исследования банковской деятельности такой способ анализа нам не встречался.

Таким образом, совокупность банков может быть разбита на кластеры, в которых рассматривается не только степень отдаленности банков от банка-эталона, но и учитываются направления банков-векторов друг относительно друга. Сила связи между такими банками-кортежами $\vec{x}^{(2)}$, $\vec{x}^{(1)}$ в зависимости от коэффициента ковариации $\cos \varphi$ представлена в табл. 3.

При расчетах в авторской методике за банк-эталон брался базисный банк № 1 с показателями на 2011 г. и с ним сравнивались все остальные 70 банков. Эти расчеты проводились по состоянию на 2011, 2014, 2017 гг. Поэтому φ в предыдущей формуле – это угол наклона между показателями деятельности i -го банка и базисного банка № 1 на 2011 г.

Для удобства вычислений применялась преобразованная формула [20]

$$\begin{aligned} \cos \varphi &= \cos(\vec{x}^{(1)}, \vec{x}^{(2)}) = \\ &= \frac{x_1^{(1)} x_1^{(2)} + x_2^{(1)} x_2^{(2)} + \dots + x_{10}^{(1)} x_{10}^{(2)}}{\sqrt{((x_1^{(1)})^2 + (x_2^{(1)})^2 + \dots + (x_{10}^{(1)})^2) \times \\ &\quad \times ((x_1^{(2)})^2 + (x_2^{(2)})^2 + \dots + (x_{10}^{(2)})^2)}}, \end{aligned}$$

где $\vec{x}^{(1)}$ – базисный банк № 1;

$\vec{x}^{(2)}$ – поочередно другие банки, сравниваемые с базисным банком № 1.

Для 70 банков φ был рассчитан с учетом весовых коэффициентов w_n для показателей деятельности.

Дальнейшие исследования показали, что классификация банков по d_3 имеет больший процент соответствия с классификацией по φ , чем d_2 с φ . На основании этого нами впервые были созданы обобщенные классификации банков на 2011, 2014, 2017 гг. по трем признакам – размеру активов, взвешенному евклидову расстоянию и углу наклона между взвешенными характеристиками банков, что позволяет избежать противоречий в группировке объектов и уточнить принципы их работы по кластерам с учетом размеров активов (группировка по активам приведена с учетом критериев, предлагаемых экономистами) (табл. 4). Здесь ПБ, ПИ, ПП – это соответственно политики близкие, индифферентные и противоположные, интерпретированные с помощью данных, представленных в табл. 3.

Данные, представленные в табл. 4, можно объяснить следующим образом. Несмотря на

то что банки-гиганты по активам № 1, № 2 и № 3 попадают в одну группу лидеров, по комплексу признаков они находятся в разных кластерах. В то же время эти банки придерживаются политики базисного банка № 1 на 2011 г., о чем свидетельствует близость (ПБ) их принципов работы к базисному банку (при проведении кластеризации по d_3 , как и по φ в совокупности участвовали 70 банков и базисный № 1 с показателями на 2011 г., чтобы была возможность получить более корректные результаты при сравнении этих кластеризаций). О том, что показатели этих банков значительно отличаются от показателей базисного банка № 1 на 2011 г., говорят и расчеты за 2014 и 2017 гг., что свидетельствует о динамичности банковской системы и мгновенной реакции на внешние и внутренние факторы.

Особый интерес в *табл. 4* представляет кластер *D*, где именно благодаря векторной модификации усовершенствованного кластерного анализа по d_3 можно получить ответ на вопрос – почему, например, в один и тот же кластер попали и средние, и мелкие банки: они похожи в принципах своей работы (их политики противоположны относительно политики базисного банка № 1 на 2011 г.). Аналогичны рассуждения относительно попадания в этот кластер 12 крупных и 9 средних банков – их принципы работы близки базисному банку № 1 на 2011 г. В то же время векторная кластеризация помогает еще более точно определить позиции средних банков: несмотря на то что по активам и обобщающим характеристикам они попали в один кластер, принципы работы относительно базисного банка № 1 на 2011 г. у них различны.

Итак, по результатам анализа данных, представленных в *табл. 4*, клиентам банков и банкам-конкурентам можно порекомендовать обратить внимание на наиболее успешные и

надежные банки: это все банки-гиганты, так как их политики близки к надежному устойчивому базисному банку № 1 и крупные и средние банки с политиками ПБ.

С помощью векторной кластеризации была определена динамика политик 70 банков по кластерам, которая позволяет сделать вывод о достаточной устойчивости банковской системы для исследуемого промежутка времени (*табл. 5*).

В результате дальнейших исследований был сделан прогноз деятельности базисного банка № 1 с точки зрения изменения углов φ с течением времени. По нашему мнению, подобный анализ каждого из банков прояснит ситуацию относительно способности банковской системы любого государства сохранять свои существенные свойства на заданном уровне в течение фиксированного промежутка времени при определенных условиях. Для анализа были взяты данные балансов за 24 месяца с 2015 по 2017 г. Использование метода анализа временного ряда – ARIMA & autocorrelation functions – модель авторегрессии и проинтегрированного скользящего среднего системы STATISTICA позволило дать вероятностную оценку долгосрочного прогноза политики банка (*рис. 3*).

Было установлено, что политика банка № 1 на 01.11.2020 будет достаточно близкой к политике базисного банка № 1 на 2011 г. (с вероятностью 0,9 гарантировалось, что угол между кортежами будет находиться в интервале от $-29,7^\circ$ до $57,8^\circ$).

Таким образом, предложенная методика проведения статистического анализа банковской деятельности позволит более точно оценить принципы деятельности банков и выбрать наиболее надежный и устойчивый к внешним изменениям банк.

Таблица 1**Показатели деятельности банков за 2017 г. (фрагмент), ден. ед.****Table 1****Indicators of bank activities in 2017 (a fragment), monetary unit**

Банк	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
№ 1	19 149	17 308	15 598	7 785	3 608	2 189	244	8 244	6 978	15
№ 2	20 453	18 395	16 236	7 807	5 813	2 103	184	8 703	4 234	382
№ 3	8 227	7 843	4 668	378	1 126	370	49	5 137	2 103	29
...
№ 70	88	55	68	26	3	33	4	18	19	0,9

Источник: авторская разработка*Source:* Authoring**Таблица 2****Нормированные показатели банков за 2017 г. (фрагмент)****Table 2****Normalized indicators of banks for 2017: A fragment**

Банк	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
№ 1	4,5	4,3	4,9	4,5	3,8	5,1	5,6	4,9	4,3	-0,17
№ 2	5,11	4,9	5,3	4,8	6,1	5,1	4,2	5,2	2,9	8,1
№ 3	2,4	3,6	2,1	1,1	0,8	0,9	1	3,4	2,4	-0,1
...
№ 70	-0,3	-0,3	-0,3	-0,3	-0,2	-0,3	-0,2	-0,2	-0,3	-0,2

Источник: авторская разработка*Source:* Authoring**Таблица 3****Геометрическая оценка принципов деятельности банков $\vec{x}^{(1)}$ и $\vec{x}^{(2)}$ (для любого государства)****Table 3****Geometrical evaluation of the principles of bank activities $\vec{x}^{(1)}$ and $\vec{x}^{(2)}$ (for any State)**

φ	$\cos \varphi$	Сила связи	Политики банков
$\varphi < 60^\circ$	$0,5 < \cos \varphi \leq 1$	Сильная, прямая	Близкие
$120^\circ < \varphi < 240^\circ$	$-1 \leq \cos \varphi < -0,5$	Сильная, обратная	Противоположные
$60^\circ \leq \varphi \leq 120^\circ$	$-0,5 \leq \cos \varphi \leq 0,5$	Слабая	Индиifferentные
$240^\circ \leq \varphi \leq 300^\circ$			

Источник: авторская разработка*Source:* Authoring

Таблица 4**Обобщенная классификация банков на 2017 г. (фрагмент)****Table 4****Generalized classification of banks in 2017: A fragment**

Активы, ден. ед.	Кластер А			Кластер В			Кластер С			Кластер D		
	ПБ	ПИ	ПП	ПБ	ПИ	ПП	ПБ	ПИ	ПП	ПБ	ПИ	ПП
Свыше 2 500 (банки- гиганты)	№ 1 (базисный) 1 банк	–	–	№ 1 1 банк	–	–	№ 3 – 28 ⁷ № 2 – 9 ⁸ и т.д. 7 банков	–	–	–	–	–
От 1 300 до 2 500 (крупные банки)	–	–	–	–	–	–	–	–	–	12 банков	–	–
От 400 до 1 300 (средние банки)	–	–	–	–	–	–	–	–	–	9 банков	2 банка	8 банков
Менее 400 (мелкие банки)	–	–	–	–	–	–	–	–	–	–	–	31 банк

Источник: авторская разработка*Source:* Authoring**Таблица 5****Динамика политики банков по кластерам А, В, С, D****Table 5****Dynamics of banks' policy by cluster A, B, C, D**

Политика	2011 г.				2014 г.				2017 г.			
	А	В	С	Д	А	В	С	Д	А	В	С	Д
ПБ	2	1	6	17	1	1	7	21	2	9	2	17
ПИ	–	–	–	5	–	–	–	2	–	–	–	1
ПП	–	–	–	40	–	–	–	39	–	–	–	40

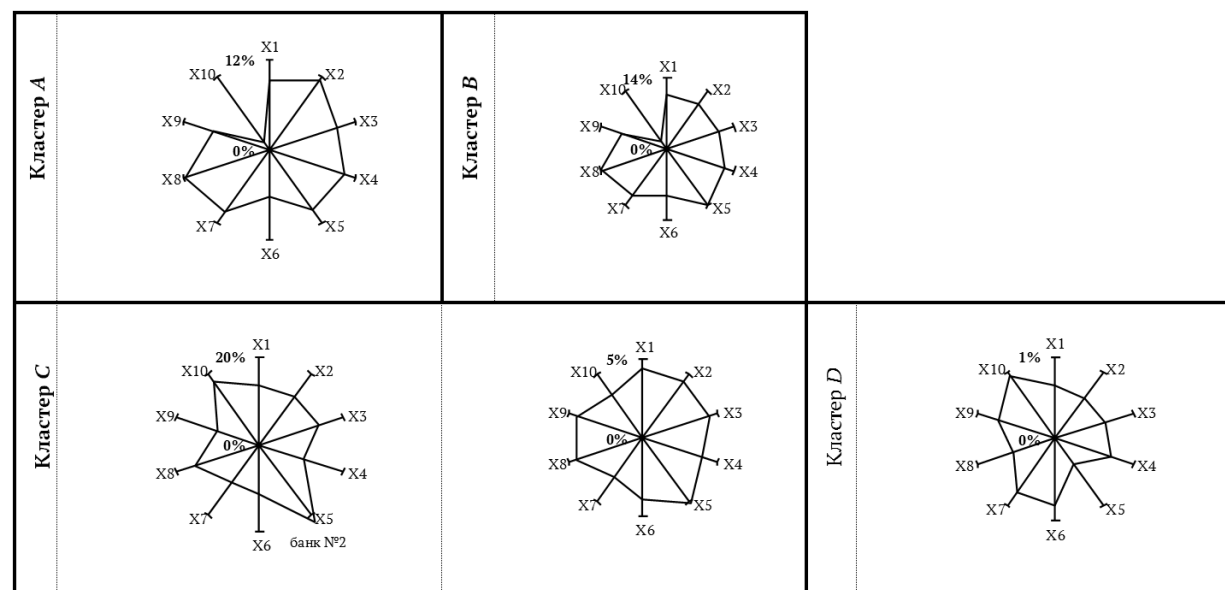
Источник: авторская разработка*Source:* Authoring

Рисунок 1**Соответствия классификаций банков для кластера А на 2017 г., %****Figure 1****Conformity of banks' classifications with the A cluster in 2017, percentage**

	По активам	$d_1 = \sqrt{\sum_{n=1} (x_n^{(i)} - x_n^{(j)})^2}$ (1 показатель)	$d_2 = \sqrt{\sum_{n=1}^{10} (x_n^{(i)} - x_n^{(j)})^2}$ (10 показателей)	$d_3 = \sqrt{\sum_{n=1}^{10} w_n (x_n^{(i)} - x_n^{(j)})^2}$ (10 показателей)
По активам	–	0,25	6,7	0,13
d_1	0,25	–	50	0
d_2	6,7	50	–	0
d_3	0,13	0	0	–

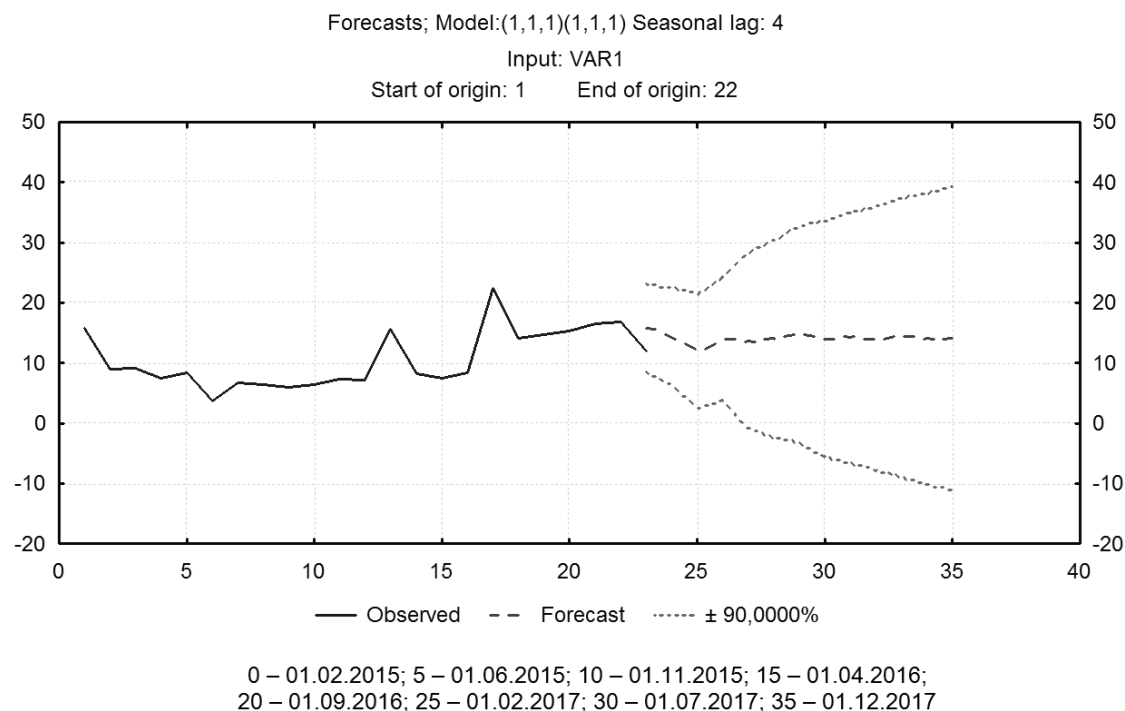
Источник: авторская разработка

Source: Authoring

Рисунок 2**Структура показателей деятельности банков по кластерам (средние значения) на 2017 г.****Figure 2****Structure of indicators of bank activities by cluster (average values) for 2017**

Источник: авторская разработка

Source: Authoring

Рисунок 3**График углов наклона показателей деятельности банка № 1 с построенным прогнозом их изменений****Figure 3****A graph of slope angles of Bank No. 1 performance indicators with the forecast of their changes**

Источник: авторская разработка

Source: Authoring

Список литературы

1. Петров А.Ю., Петрова В.И. Комплексный анализ финансовой деятельности банка. М.: Финансы и статистика, 2007. 560 с.
2. Долан Э.Дж., Кэмпбелл К.Д., Кэмпбелл Р.Дж. Деньги, банковское дело и денежно-кредитная политика. М.: Профико, 1993. 446 с.
3. Бор М.З., Пятенко В.В. Менеджмент банков: организация, стратегия, планирование. М.: ИКЦ «ДИС», 1997. 288 с.
4. Головач Н.А. Методологія та організація статистичної інформації в аналізі банківської діяльності. In: Прикладна статистика: проблеми теорії та практики. Київ, Інформаційно-аналітичне агентство, 2010, 368 р.
5. Rose P.S. Commercial Bank Management. Boston, McGraw-Hill/Irwin, 2002, 803 p.
6. Иванов В.В. Надежность вашего банка. М.: ФБК-пресс, 1997. 176 с.
7. Синки Дж.Ф. Финансовый менеджмент в коммерческом банке и в индустрии финансовых услуг. М.: Альпина Бизнес Букс, Альпина Паблишер, 2007. 1024 с.
8. Porter R.S. Introduction to Banking Regulation, Supervision and Bank Analysis. EDI, World Bank, 1993, 104 p.

9. Ершов М.В., Zubov В.М. Эффективность банковской системы: актуальные аспекты // Деньги и кредит. 2005. № 10. С. 3–9.
10. Ключников М.В. Анализ показателей, характеризующих финансовую деятельность коммерческих банков // Финансы и кредит. 2003. № 20. С. 40–45.
URL: <https://cyberleninka.ru/article/v/analiz-pokazateley-harakterizuyuschih-finansovuyu-deyatelnost-kommercheskih-bankov>
11. Саблина Е.А. Статистическая оценка состояния банковской системы России после кризиса // Вопросы статистики. 2011. № 7. С. 68–76.
12. Замковой С.В. Моделирование динамики банковской системы и финансовых рынков // Банковское дело. 2002. № 7. С. 9–12.
13. Demirguc-Kunt A., Detragiache E. The Determinants of Banking Crises in Developing and Developed Countries. *IMF Staff Papers*, 1998, vol. 45, no. 1, pp. 81–109.
14. Kaminsky G., Reinhart C. The Twin Crises: The Causes of Banking and Balance-of-Payments Problems. *The American Economic Review*, 1999, vol. 89, iss. 3, pp. 473–500.
URL: <https://doi.org/10.1257/aer.89.3.473>
15. Chincarini L., Asherie N. An Analytical Model for the Formation of Economic Clusters. *Regional Science and Urban Economics*, 2008, vol. 38, iss. 3, pp. 252–270.
URL: <https://doi.org/10.1016/j.regsciurbeco.2008.01.010>
16. Андерсон Т. Введение в многомерный статистический анализ. М.: Физматгиз, 1963. 500 с.
17. Кендалл М., Стюарт А. Многомерный статистический анализ и временные ряды. М.: Наука, 1976. 736 с.
18. Майба В.В. Кластерна структура банківської системи України. In: Прикладна статистика: проблеми теорії та практики, Київ, Інформаційно-аналітичне агентство, 2008, pp. 463–474.
19. Колмогоров А.Н. Основные понятия теории вероятностей. М.: Наука, 1974. 119 с.
20. Харкевич А.А. Борьба с помехами. М.: Наука, 1965. 274 с.

Информация о конфликте интересов

Я, автор данной статьи, со всей ответственностью заявляю о частичном и полном отсутствии фактического или потенциального конфликта интересов с какой бы то ни было третьей стороной, который может возникнуть вследствие публикации данной статьи. Настоящее заявление относится к проведению научной работы, сбору и обработке данных, написанию и подготовке статьи, принятию решения о публикации рукописи.

STATISTICAL ANALYSIS OF BANK ACTIVITIES USING THE ADVANCED CLUSTERING

Oksana S. DOTSENKO

Sevastopol State University, Sevastopol, Russian Federation
plakiddin@mail.ru
ORCID: not available

Article history:

Received 7 May 2018
Received in revised form
18 May 2018
Accepted 29 May 2018
Available online
27 July 2018

JEL classification: C1, C6,
G2

Keywords: bank, cluster
analysis, weighted Euclidean
distance, vector, slope angle

Abstract

Subject The article addresses procedures for statistical and mathematical analysis of bank activities.

Objectives The study aims to improve methods of multidimensional grouping of banks and forecasting their activities using mathematical models, i.e. clustering based on similarity between scalar banks – weighted Euclidean distance, and vector modification of cluster analysis based on similarity – the angle of slope between banks-vectors characterized by weighted indicators.

Methods I apply statistical and mathematical techniques, like statistical observation, banks' grouping by individual indicators and a suite of metrics using the factor, cluster, discriminant and vector analysis, methods of analyzing time series and relationships between indicators.

Results I developed generalized classifications of banks based on three attributes, namely, asset size, weighted Euclidean distance and the slope angle between weighted characteristics of banks. It helps avoid contradictions in grouping the objects and clarify principles of their work by clusters, considering the size of assets and dynamics in slope angles. The forecast of the particular bank's policy using the vector representation of its indicators enabled to assess the bank's stability over the analyzed period of time. The analysis techniques may be useful for bank CEOs, customers and counterparties when choosing a stable and reliable bank.

Conclusions Under the presented methodology, the bank grouping on the basis of a set of indicators is more homogeneous and the principles of bank activities have better interpretations.

© Publishing house FINANCE and CREDIT, 2018

Please cite this article as: Dotsenko O.S. Statistical Analysis of Bank Activities Using the Advanced Clustering. *Economic Analysis: Theory and Practice*, 2018, vol. 17, iss. 7, pp. 1352–1366.
<https://doi.org/10.24891/ea.17.7.1352>

References

1. Petrov A.Yu., Petrova V.I. *Kompleksnyi analiz finansovoi deyatel'nosti banka* [Comprehensive analysis of bank's financial activities]. Moscow, Finansy i statistika Publ., 2007, 560 p.
2. Dolan E.G., Campbell C.D., Campbell R.G. *Den'gi, bankovskoe delo i denezhno-kreditnaya politika* [Money, Banking and Monetary Policy]. Moscow, Profiko Publ., 1993, 446 p.
3. Bor M.Z., Pyatenko V.V. *Menedzhment bankov: organizatsiya, strategiya, planirovanie* [Management of banks: Organization, strategy, planning]. Moscow, IKTs DIS Publ., 1997, 288 p.
4. Golovach N.A. Методологія та організація статистичної інформації в аналізі банківської діяльності. In: Прикладна статистика: проблеми теорії та практики. Київ, Інформаційно-аналітичне агентство, 2010, 368 p.
5. Rose P.S. *Commercial Bank Management*. Boston, McGraw-Hill/Irwin, 2002, 803 p.

6. Ivanov V.V. *Nadezhnost' vashego banka* [The reliability of your bank]. Moscow, FBK-press Publ., 1997, 176 p.
7. Sinkey J.F. *Finansovyi menedzhment v kommercheskom banke i v industrii finansovykh uslug* [Commerical Bank Financial Management in the Financial Services Industry]. Moscow, Al'pina Biznes Buks, Al'pina Publisher Publ., 2007, 1024 p.
8. Porter R.S. *Introduction to Banking Regulation, Supervision and Bank Analysis*. EDI, World Bank, 1993, 104 p.
9. Ershov M.V., Zubov V.M. [Efficiency of the banking system: Challenging issues]. *Den'gi i kredit = Money and Credit*, 2005, no. 10, pp. 3–9. (In Russ.)
10. Klyuchnikov M.V. [Analysis of indicators that characterize the financial activities of commercial banks]. *Finansy i kredit = Finance and Credit*, 2003, no. 20, pp. 40–45.
URL: <https://cyberleninka.ru/article/v/analiz-pokazateley-harakterizuyuschih-finansovuyu-deyatelnost-kommercheskih-bankov> (In Russ.)
11. Sablina E.A. [Statistical assessment of the Russian banking system condition after the crisis]. *Voprosy Statistiki*, 2011, no. 7, pp. 68–76. (In Russ.)
12. Zamkovo S.V. [Modeling the dynamics of the banking system and financial markets]. *Bankovskoe delo = Banking*, 2002, no. 7, pp. 9–12. (In Russ.)
13. Demirguc-Kunt A., Detragiache E. The Determinants of Banking Crises in Developing and Developed Countries. *IMF Staff Papers*, 1998, vol. 45, no. 1, pp. 81–109.
14. Kaminsky G., Reinhart C. The Twin Crises: The Causes of Banking and Balance-of-Payments Problems. *The American Economic Review*, 1999, vol. 89, iss. 3, pp. 473–500.
URL: <https://doi.org/10.1257/aer.89.3.473>
15. Chincarini L., Asherie N. An analytical model for the formation of economic clusters. *Regional Science and Urban Economics*, 2008, vol. 38, iss. 3, pp. 252–270.
URL: <https://doi.org/10.1016/j.regsciurbeco.2008.01.010>
16. Anderson T. *Vvedenie v mnogomernyi statisticheskii analiz* [An Introduction to Multivariate Statistical Analysis]. Moscow, Fizmatgiz Publ., 1963, 500 p.
17. Kendall M., Stuart A. *Mnogomernyi statisticheskii analiz i vremennye ryady* [The Advanced Theory of Statistics: Design and Analysis, and Time-Series]. Moscow, Nauka Publ., 1976, 736 p.
18. Maiba V.V. Кластерна структура банківської системи України. In: Прикладна статистика: проблеми теорії та практики. Київ, Інформаційно-аналітичне агентство, 2008, pp. 463–474.
19. Kolmogorov A.N. *Osnovnye ponyatiya teorii veroyatnostei* [Basic concepts of the probability theory]. Moscow, Nauka Publ., 1974, 119 p.
20. Kharkevich A.A. *Bor'ba s pomekhami* [Interference control]. Moscow, Nauka Publ., 1965, 274 p.

Conflict-of-interest notification

I, the author of this article, bindingly and explicitly declare of the partial and total lack of actual or potential conflict of interest with any other third party whatsoever, which may arise as a result of the publication of this article. This statement relates to the study, data collection and interpretation, writing and preparation of the article, and the decision to submit the manuscript for publication.