

УДК 519.681

КЛАСТЕРНО-РЕГРЕССИОННЫЙ АНАЛИЗ ФИНАНСОВЫХ ПОКАЗАТЕЛЕЙ БАНКОВ УКРАИНЫ НА ОСНОВЕ МГУА

Л.В. Сарычева¹, А.П. Сарычев²

¹Национальный горный университет,

²Институт технической механики НАН Украины и ГКА Украины

sarycheval@ntmu.org.ua, sarychev@prognoz.dp.ua,

Для аналізу діяльності банків України запропоновано кластерно-регресійний підхід, що засновано на принципах методу групового урахування аргументів. Підхід може бути використано для прогнозування стану банків.

Ключові слова: показники фінансової діяльності банків, кластеризація, регресія.

For the analysis of activity of banks of Ukraine it is offered the cluster-regression approach which is based on principles of the Group Method of Data Handling. The approach can be used for forecasting a state of banks.

Key words: parameters of financial activity of banks, clasterization, regression.

Для анализа деятельности банков Украины предложен кластерно-регрессионный подход, который основан на принципах метода группового учёта аргументов. Подход может быть использован для прогнозирования состояния банков.

Ключевые слова: показатели финансовой деятельности банков, кластеризация, регрессия.

1. Постановка задачи

К банкам предъявляются требования по обеспечению экономической стабильности и финансовой устойчивости. Поэтому моделирование показателей банковской финансовой деятельности, анализ их взаимосвязей, установление закономерностей с целью прогнозирования состояния банков является актуальной задачей.

Банки характеризуются различными показателями: 1) форма собственности (унитарные и коллективные, акционерные и кооперативные); 2) происхождение капитала (с отечественным капиталом, с участием иностранного капитала, со 100% иностранным капиталом); 3) территориальный признак (международные, национальные, региональные); 4) размер активов (крупнейшие, крупные, средние и малые); 5) финансовое положение (стабильные, проблемные, кризисные и банкроты); 6) наличие филиалов (много филиалов, мало филиалов, без филиалов); 7) диапазон операций, сектор рынка функционирования и др.

Национальный Банк Украины (НБУ) группирует банки в четыре группы по размеру активов. От того, в какой группе находятся те или иные банки, зависит степень банковского надзора, уровень развития и функционирования банковской системы.

В силу разнообразия украинских банков по многим приведенным характеристикам их финансовая деятельность не поддается общим закономерностям. Установление закономерностей для каждой из четырех групп, выделенных по классификации НБУ, является трудной задачей в силу неоднородности этих групп. Выделение же однородных групп возможно только по совокупности показателей финансовой деятельности, причем в фиксированном временном диапазоне, с учетом динамики изменения показателей.

В данной работе предлагается кластерно-регрессионный анализ финансовых показателей банков Украины, идея которого состоит в эмергентном сочетании итеративных процедур объективной кластеризации банков и построения регрессионных моделей для каждого кластера.

Цель работы – по ежемесячным банковским данным статистической отчетности о финансовой деятельности за годовой промежуток времени создать модель объективной кластеризации [1, 2] и регрессионные модели зависимости финансового результата банков для отдельных кластеров.

2. Исходные данные

Данные для расчетов получены на официальном сайте Ассоциации Украинских Банков и НБУ.

Объекты анализа – банки Украины.

Показатели деятельности банков представляют собой:

– ежеквартальные характеристики [3] (100 показателей V_1, V_2, \dots, V_{100} для 175-ти банков на 01.10.2012 г. и 01.01.2013 г., состав банков отличен);

– ежемесячные характеристики банковской деятельности (из шести основных групп) за 2012 год [4] (22 показателя, число банков в разные месяцы различно – 140 банков на 01.02.2012 г., 138 банков на 01.05.2012 г. и т. д.).

Показатели объектов могут быть разделены на такие группы:

А – Активы и обязательства: X_1 – «Активи», X_2 – «Балансовий капітал», X_3 – «Зобов'язання» (млн.грн);

КИП – Кредитно-инвестиционный портфель: X_4 – «КІП», X_5 – «Міжбанківські кредити», X_6 – «Резерв під заборгованість банків», X_7 – «Кредити юридичним особам», X_8 – «Кредити фізичним особам», X_9 – «Резерв під кредити та заборгованість клієнтів», X_{10} – «Цінні папери» (млн.грн);

К – капитал: X_{11} – «Капітал» (млн.грн), X_{12} – «Капітал» (млн.євро), X_{13} – «Статутний капітал» (млн.грн);

ДФ – Депозиты физических лиц: X_{14} – «Сума депозитів фізичних осіб», X_{15} – «Кошти до запитання», X_{16} – «Строкові депозити» (млн.грн);

ДЮ – Депозиты юридических лиц: X_{17} – «Сума депозитів юридичних осіб», X_{18} – «Кошти до запитання», X_{19} – «Строкові депозити» (млн.грн);

Ф – финансовый результат: X_{20} – «Фінансовий результат» (млн.грн), X_{21} – «Прибутковість статутного капіталу» (%), X_{22} – «Прибутковість активів (%)».

3. Разведочный анализ данных

Проводится анализ главных компонент, корреляционный, факторный, дисперсионный анализ показателей; попарно показатели анализируются с помощью диаграмм рассеяния. Выбираются показатели для дальнейшего анализа. Анализ первых трех главных компонент (рис. 1) показал наличие таких кластеров: Приватбанк (пять точек в правом верхнем углу на рис. 1а, 1б, 1в; одна точка в левом верхнем на рис. 1г, 1д) и Ощадбанк образуют отдельные далеко отстоящие кластеры, а основная масса малых банков группируется в один кластер. Предварительная оценка числа кластеров – от 5-ти до 11-ти, существуют как малочисленные, так и крупные кластеры.

Анализ коэффициентов парной корреляции (рис. 2а) показал, что показатели группы Φ (финансовый результат) X_{20} , X_{21} , X_{22} слабо коррелируют с показателями X_1, X_2, \dots, X_{19} .

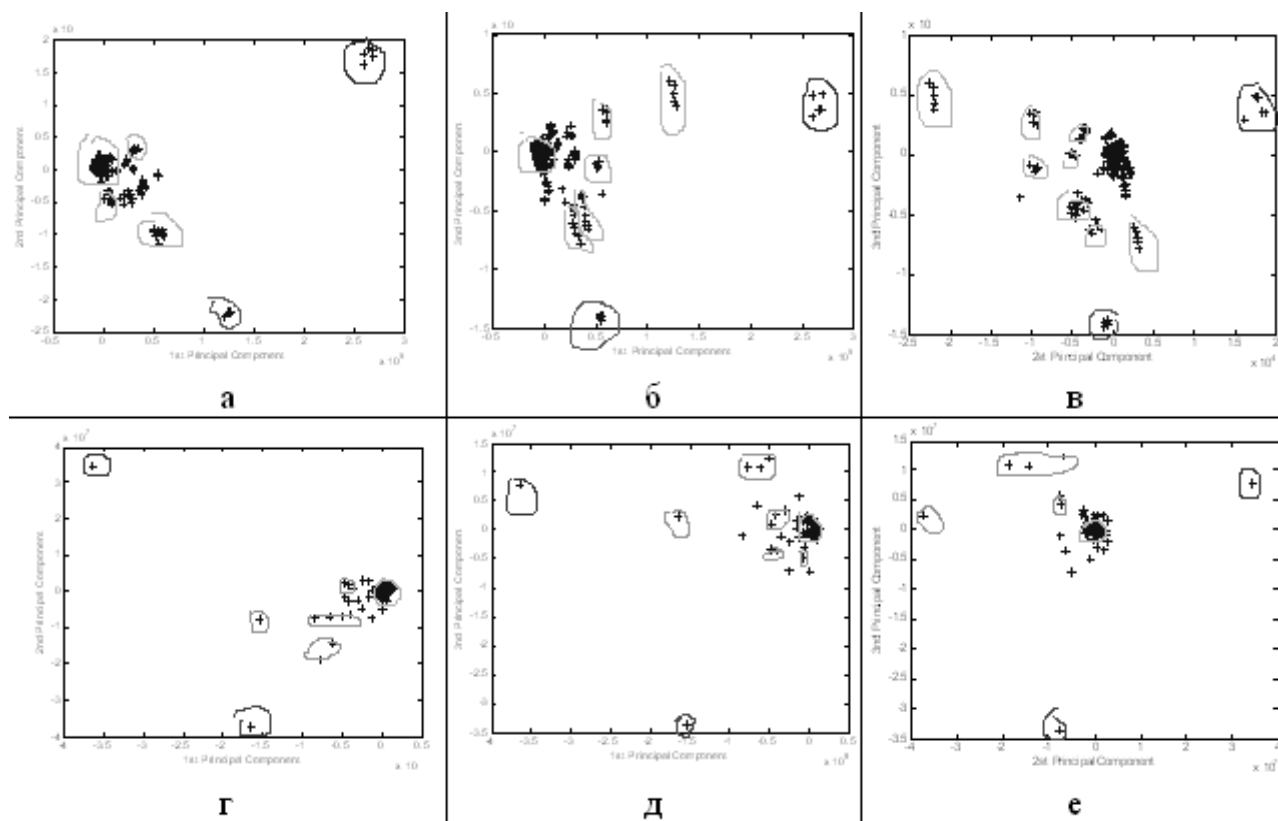


Рис. 1. Банки в пространстве главных компонент (а, г – 1-2-я; б, д – 1-3-я; в, е – 2-3-я компоненты):
а, б, в – по показателям X_1, X_2, \dots, X_{22} за 5 месяцев (период 01.01.2012-01.06.2012);
г, д, е – по показателям V_1, V_2, \dots, V_{100} (на 01.10.2012)

Корреляционный (рис. 2б) и дисперсионный анализ показателей V_1, V_2, \dots, V_{100} свидетельствуют о том, что более 60% показателей можно не использовать при кластеризации.

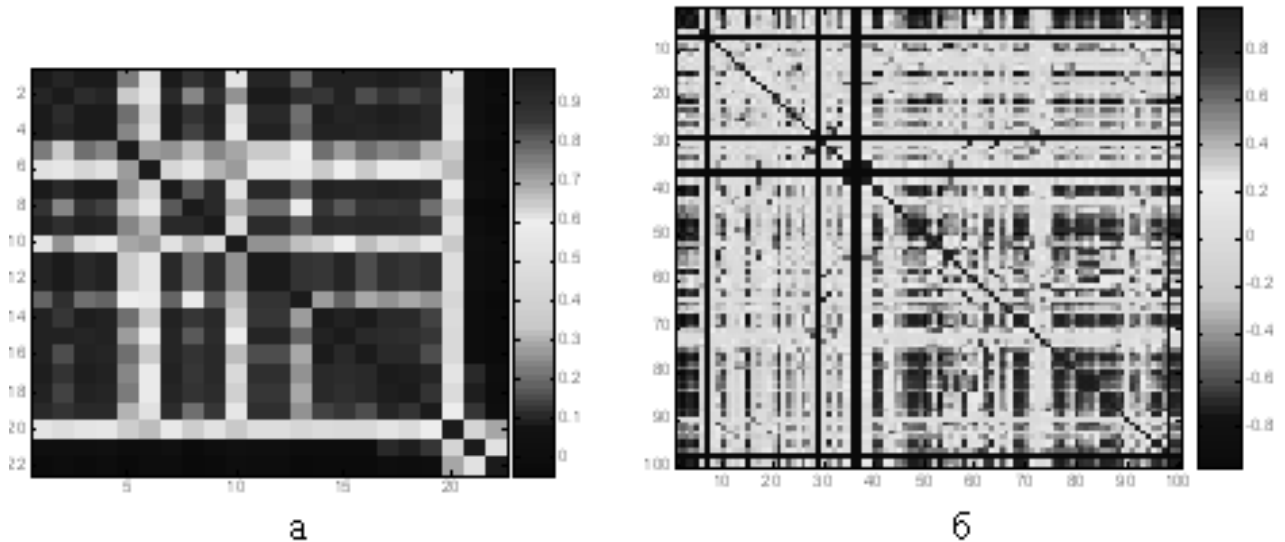


Рис.2. Коэффициенты парной корреляции между показателями X_1, X_2, \dots, X_{22} (а) и показателями V_1, V_2, \dots, V_{100} (б)

4. Методика кластерно-регрессионного анализа финансовых показателей банков Украины

Исходное пространство признаков X_1, X_2, \dots, X_{22} используем для объективной кластеризации [1, 2], выделив в нем функцию цели – финансовый результат банка $Y=X_{20}$.

Задача состоит в том, чтобы разделить оба пространства на некоторые подпространства (кластеры) и установить соответствие кластеров пространства признаков и кластеров пространства функции цели [5]:

$$X \Leftrightarrow Y.$$

Объективная кластеризация (ОК) основана на переборе вариантов ансамблей признаков и числа кластеров по критерию непротиворечивости при заданной таблице данных. В результате перебора находится число признаков и число кластеров для всех непротиворечивых кластеризаций, а все имеющиеся объекты распределяются по кластерам.

Новые измерения, не участвующие в кластеризации, также относятся к тому или иному кластеру по правилу «ближнего соседа» или по минимуму расстояния до центра кластера.

Алгоритм объективной кластеризации изложен в работе [1].

Пусть $x_{ij}(t_s)$ – измерения признаков, характеризующих заданное множество объектов в момент времени t_s ($i=1, 2, \dots, n$ – номер объекта, $j=1, 2, \dots, m$ – номер признака, $s=1, 2, \dots, L$ – номер момента времени). Исходные данные представляют собой блочную матрицу $[X(t_1) \mid X(t_2) \mid \dots \mid X(t_L)]$.

Предположим (для ежемесячных данных), что объекты (банки), близкие по своим свойствам (входящие в один кластер) в момент времени t , могут быть близки и в момент времени $(t+1)$. Пусть $K(t_s)=\{K_1(t_s), K_2(t_s), \dots, K_k(t_s)\}$ –

кластеризация объектов в момент времени $t=t_s$, $s=1,2,\dots,L$, где k – число кластеров, $1 < k < n$:

$$\bigcup_{i=1}^k K_i(t_s) = X(t_s), K_i(t_s) \neq \emptyset, K_i(t_s) \cap K_j(t_s) = \emptyset, \text{ при } i \neq j; i, j = 1, 2, \dots, k.$$

Каждому i -му объекту (отождествленному с точкой $(x_{i1}(t_s), x_{i2}(t_s), \dots, x_{im}(t_s))$, $i=1,2,\dots,n$, евклидова пространства R^m) кластеризация $K(t_s)$ ставит в соответствие номер кластера, к которому он принадлежит в момент времени t_s . По результатам $K(t_s)$ и $K(t_{s+1})$ определяется кластеризация $K(t_s, t_{s+1})$ (рис. 3) [2]. Номер кластера является не более чем меткой, то есть можно поменять нумерацию кластеров, сохранив при этом состав входящих в них элементов.

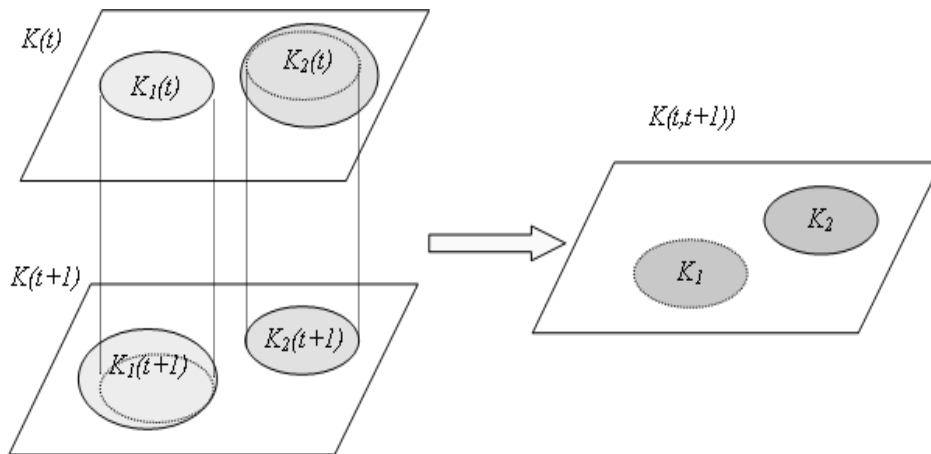


Рис. 3. Преемственность кластеризаций

Выделенные ядра кластеров расширяют путем доклассификации оставшихся объектов. Кластеризация оценивается с двух точек зрения: удачно ли она разделила представленные данные (точность) и сохраняется ли она и для других множеств данных, полученных на том же объекте наблюдений (непротиворечивость).

Перебор вариантов ансамблей признаков и числа кластеров приводит к множественному решению: на плоскости «ансамбль признаков – число кластеров» находится несколько вариантов, дающих непротиворечивые кластеризации. Расчет критерия непротиворечивости ведется только при равном числе кластеров на двух подвыборках А и В.

В качестве А и В для ежемесячных данных X используются выборки по соседним месяцам. Для ежеквартальных данных V выборка разбивается на подвыборки А и В по диполям [5].

Для сокращения объема вычислений сравнение вариантов кластеризации проводится, начиная с $k=4$ кластеров (НБУ выделяет 4 группы банков) до $k=11$ кластеров.

Однозначный выбор единственной кластеризации определяется критерием точности прогноза величины Y (финансового результата для объектов заданного кластера).

Для такого прогноза строится регрессионная модель по принципам метода группового учета аргументов [6, 7]. Класс синтезируемых моделей имеет вид

$$\hat{y} = \sum_{q=1}^s \theta_q \cdot \prod_{j=1}^m x_j^{\alpha_{qj}},$$

где \hat{y} – выходная переменная; s – число членов в модели; θ_q , $q = 1, 2, \dots, s$ – коэффициенты; x_j , $j = 1, 2, \dots, m$ – входные переменные; m – число входных переменных; α_{qj} – показатель степени, в которой переменная x_j входит в q -й член.

Поиск модели ведётся как в классе линейных, так и в классе нелинейных по входным переменным моделей; коэффициенты в моделях оцениваются по методу наименьших квадратов.

5. Модели кластеризации и регрессии для анализа банков Украины

Экспериментальные исследования проведены для ежемесячных (выборка **X**) и ежеквартальных данных (выборка **V**).

В таблице 1 представлены результаты трех объективных кластеризаций: выборки **V** с числом кластеров $k=11$ и выборки **X** с числом кластеров $k=4$ и $k=11$. Из-за громоздкости представлены не все группы (I, II, III, IV) - 175 банков, а только лучшие по финансовым результатам за рассматриваемый период - группа I (15 банков) и группа II (20 банков).

Объективная кластеризация (табл. 1, правая колонка) для числа кластеров $k=4$ не совпадает с классификацией НБУ по финансовым результатам банков (группа I, II, III, IV), но близка к ней.

В отдельный кластер выделяются Приватбанк и Ощадбанк. Но построить приемлемые прогнозные модели в случае $k=4$ удастся только для 1-го и 2-го кластера. Кластеризация для $k=5$ почти не отличается от случая $k=4$. Только Приватбанк и Ощадбанк образуют здесь отдельные кластеры.

Хорошие модели в смысле соответствия кластеров пространства признаков и кластеров пространства функции цели $X \Leftrightarrow Y$ дает результат при $k=11$ (табл. 1, левая колонка для выборки **V**, четвертая колонка для выборки **X**). Здесь в отдельный кластер (кроме Приватбанка и Ощадбанка) выделяется Родовид банк.

В качестве примера полученных моделей приведем наименее сложную модель для кластера {Дельта банк, Перший Український міжнародний банк, Альфа банк, Сбербанк Росії}: «Чистий процентний дохід» (V_3) как функция от трех показателей «Усього активів» (V_{68}), «Статутний капітал» (V_{70}), «Усього зобов'язань» (V_{99}):

$$V_3 = 0,479138 V_{68} + 0,00702 V_{70} - 0,50654 V_{99},$$

для которой множественный коэффициент корреляции составляет $R=0,99$.

Таблица 1

Выборка V № кластера k =11	Банк	«Процентні доходи»	Выборка X	
			№ кластера k =11	№ кластера k =4
Група I				
1	ПРИВАТБАНК	18153983	1	1
2	УКРЕКСІМБАНК	7535962	нет данных	нет данных
2	ОЩАДБАНК	9735742	2	1
3	РАЙФФАЙЗЕН БАНК АВАЛЬ	6335415	нет данных	нет данных
3	ПРОМІНВЕСТБАНК	4222253	3	2
3	УКРСОЦБАНК	4559268	3	2
3	ВТБ БАНК	4080458	3	2
4	ДЕЛЬТА БАНК	3218491	4	2
4	ПУМБ	3104469	4	2
4	"НАДРА"	1686251	нет данных	нет данных
4	СБЕРБАНК РОСІЇ	2447669	4	2
4	УКРСИББАНК	3124883	нет данных	нет данных
4	АЛЬФА-БАНК	3823231	4	2
4	БАНК ФІНАНСИ ТА КРЕДИТ	2045512	4	2
5	УКРГАЗБАНК	2165829	5	3
Група II				
4	ОТП БАНК	2659907	4	2
5	БРОКБІЗНЕСБАНК	1321166	4	2
5	ВІЕЙБІ БАНК	1243756	6	3
5	КРЕДІ АГРІКОЛЬ БАНК	1520061	6	3
5	КРЕДИТПРОМБАНК	1231124	5	3
5	ІНГ БАНК УКРАЇНА	884773	5	3
6	"ФІНАНСОВА ІНІЦІАТИВА"	1419327	5	3
6	"ПІВДЕННИЙ"	1037894	5	3
6	ІМЕКСБАНК	1281608	6	3
6	РОДОВІД БАНК	255596	11	3
6	"ХРЕЩАТИК"	828884	6	3
6	БАНК ФОРУМ	766221	нет данных	нет данных
6	БАНК КРЕДИТ ДНІПРО	805696	6	
6	ЕРСТЕ БАНК	897579	нет данных	нет данных
6	УНІВЕРСАЛ БАНК	805244	6	3
7	УНКРЕДИТ БАНК	500393	7	3
7	БТА БАНК	308969	7	3
7	БАНК КИЇВСЬКА РУСЬ	623075	7	3
7	БАНК 3/4	240205	8	4
7	ПРАВЕКС-БАНК	569091	7	3

На основе регрессионных моделей, построенных по принципам МГУА для прогнозирования финансового результата для объектов заданного кластера, выделены объекты с наибольшими (по модулю) остатками регрессии.

Они образуют множество банков, показатели которых не подчиняются «общей тенденции» соответствующего кластера: {Укрсоцбанк, Райффайзен банк Аваль, Брокбизнесбанк, Родовід банк, Сітібанк, КІБ КРЕДІ АГРІКОЛЬ, Платинум банк, Экспобанк, Банк Меркурий, Акцент-банк, Інтербанк}.

Выводы

В работе предложен новый подход к анализу показателей финансовой деятельности банков Украины, основанный на совместном применении итерационных процедур объективной кластеризации и регрессионного анализа, реализованных по принципам метода группового учета аргументов. Результаты экспериментов на реальных статистических данных финансовой деятельности банков за 2012 год показали адекватность моделей.

Литература

1. Сарычева Л. В. Объективный кластерный анализ данных на основе МГУА / Л. В. Сарычева // Проблемы управления и информатики. – 2008. – № 2. – С. 86–104.
2. Sarycheva L. The spatial-temporary approach in problems of clusterization / L. Sarycheva // II International Workshop on Inductive Modelling “IWIM–2007” : 22–26 September 2007, Prague : Proc. – Prague : Czech Technical University, 2007. – P. 180–187. Эл. ресурс. URL: <http://www.iwim2007.org/Proceedings>.
3. Основні показники діяльності банків України [Електронний ресурс]. – Режим доступу: http://www.bank.gov.ua/Bank_Supervision/dynamics.htm
4. Офіційний сайт Української асоціації банків [Електронний ресурс]. – Режим доступу: <http://aub.org.ua/>
5. Ивахненко А. Г. Объективная компьютерная кластеризация. Ч. 1. Теоретические вопросы / А. Г. Ивахненко, С. А. Петухова, Н. А. Ивахненко // Автоматика. – 1986. – №3. – С. 3–11.
6. Сарычева Л. В. Прогнозирование в задачах эколого-социально-экономического мониторинга регионов / Л. В. Сарычева // Моделювання та інформатизація соціально-економічного розвитку України. Зб. наук. пр. – К. : ДНДІМЕ, 2003. – №3. – С. 147–165.
7. Сарычев А. П. Усредненный критерий регулярности метода группового учета аргументов в задаче поиска наилучшей регрессии / А. П. Сарычев // Автоматика. – 1990. №5. – С. 28–33.