УДК 338.001.36, 338.1, 332.055, 339.9, 330.4

C. Математические и количественные методы
C. Mathematical and Quantitative Methods

МАТЕМАТИЧЕСКИЕ МОДЕЛИ РЕЙТИНГОВОГО АНАЛИЗА

Е. Г. ГОСПОДАРИК[1), 2)], ***М. М. КОВАЛЁВ***[1)]

[1)]*Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь*
[2)]*Финансовый университет при Правительстве Российской Федерации,*
пр. Ленинградский, 49, 125167, г. Москва, Россия

Рассматриваются современные математические модели рейтингов и рэнкингов, применяемые в сравнительном экономическом анализе. Изложен взгляд на теорию рейтингов как с позиций теории коллективного выбора, так и с позиции теории построения сводных индексов, которая описывает наиболее распространенные модели рейтингов, основанные на построении сводного рейтингового индекса. Представлены также модели рейтингов, базирующиеся на проведении факторного и кластерного анализа, а также графовые модели и модели, созданные на основе решения оптимизационных задач поиска порядка, наименее отклоняющегося от заданных.

Ключевые слова: рейтинг; рэнкинг; мажоритарная модель Кондорсе; кластер-анализ; балльная модель Борда; модель Симпсона; модель Копленда; модель Нэша; модель рейтинга на основе сводного рейтингового индекса; агрегация показателей; модель на основе линейных порядков; модель на основе факторного анализа; модель на основе кластер-анализа; оптимизационная модель разбиения на классы.

Образец цитирования:
Господарик ЕГ, Ковалёв ММ. Математические модели рейтингового анализа. *Журнал Белорусского государственного университета. Экономика*. 2023;2:4–19.
EDN: OYCNAG

For citation:
Gospodarik CG, Kovalev MM. Mathematical models of rating analyses. *Journal of the Belarusian State University. Economics.* 2023;2:4–19. Russian.
EDN: OYCNAG

Авторы:
Екатерина Геннадьевна Господарик – кандидат экономических наук, доцент; заведующий кафедрой аналитической экономики и эконометрики экономического факультета[1)], профессор департамента бизнес-аналитики[2)].
Михаил Михайлович Ковалёв – доктор физико-математических наук, профессор; профессор кафедры аналитической экономики и эконометрики экономического факультета.

Authors:
Catherine G. Gospodarik, PhD (economics), docent; head of the department of analytical economics and econometrics, faculty of economics[a], and professor at the department of business analysis[b].
gospodarik@bsu.by
Mikhail M. Kovalev, doctor of science (physics and mathematics), full professor; professor at the department of analytical economics and econometrics, faculty of economics.
kovalev@bsu.by

MATHEMATICAL MODELS OF RATING ANALYSES

C. G. GOSPODARIK[a, b], *M. M. KOVALEV*[a]

[a]*Belarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus*
[b]*Financial University under the Government of the Russian Federation, 49 Leningradskii Avenue, Moscow 125167, Russia*

Corresponding author: C. G. Gospodarik (gospodarik@bsu.by)

The article outlines modern mathematical models of ratings and rankings used in comparative economic analysis. A view on the theory of ratings is presented both from the point of the theory of collective choice and from the point of the theory of constructing composite indices, which give the most common rating models based on the construction of a composite rating index. Rating models based on factor and cluster analysis, as well as graph models and models based on optimisation problems of finding the order least deviating from the specified ones are also developed.

Keywords: rating; ranking; Condorcet majority model; cluster analysis; Bord score model; Simpson model; Copland model; Nash model; rating model based on a consolidated rating index; aggregation of indicators; model based on linear orders; model based on factor analysis; model based on cluster analysis; optimisation model partitions into classes.

Введение

В последнее время экономисты все чаще применяют рейтинговые исследования для сравнительного анализа различных производственных объектов, в первую очередь экономических процессов в странах и региональных объединениях в целях выявления скорости модернизации и конвергенции, т. е. выравнивания социально-экономического развития. Рейтинговый анализ широко применяется такими международными организациями, как ООН, Всемирный банк, МВФ, для сравнения экономических процессов в странах мира, а его результаты широко используются и обсуждаются учеными. Однако в поле зрения попадают только результаты и показатели, а модели рейтингования остаются в стороне, хотя они оказывают существенное влияние на результаты рейтингового анализа.

Теория рейтингов восходит к классическим работам Ж. Ш. Борда [1] и М. Кондорсе [2] времен Великой французской революции, посвященным моделям обработки и методам голосования. Эти труды предопределили развитие теории кооперативного (коллективного) принятия решений [3] и теории общественного выбора [4]. В основе этих теорий лежат математические модели агрегации индивидуальных предпочтений в единственный (рейтинговый) коллективный выбор, которые бурно развиваются и сегодня, несмотря на знаменитую теорему Эрроу [5] о несуществовании правила коллективного выбора, удовлетворяющего пяти естественным аксиомам: единогласию, отсутствию диктатора, транзитивности, охвату (полноте и универсальности), а также независимости от посторонних альтернатив. Иными словами, К. Д. Эрроу доказал, что справедливая демократия невозможна, а коллективное решение всегда должно совпадать с мнением одного (диктатор). В 1972 г. К. Д. Эрроу стал лауреатом Нобелевской премии по экономике за теорию благосостояния и экономического равновесия. Позднее А. Сен на основе теории общественного выбора разработал теорию справедливости и экономики благосостояния [6], за что в 1998 г. получил Нобелевскую премию по экономике. Заметим, что ему же принадлежит идея создания одного из первых и самых известных рейтингов – рейтинга ООН человеческого развития.

Необходимо отметить, что все эти теории возникли как поиск модели агрегации мнений людей (экспертов) по любому вопросу – от обработки голосов избирателей до выбора устраивавшего общество варианта экономической политики. Нас же будут интересовать модели агрегации не только предпочтений людей, но и численных показателей, когда одни из них лучше у одной страны (процесса, системы), а другие лучше у другой страны.

К числу самых простых и самых распространенных моделей рейтинга относится модель построения рейтингового индекса, который является сверткой отобранных для рейтинга показателей. База такой модели – построение композитного (или сводного) индекса, агрегирующего частные показатели. Методология построения композитных (сводных) индикаторов обобщена в методологических руководствах Статистической дирекции ОЭСР[1], на основе которых позднее этой организацией были даны рекомендации к измерению, например, конкурентоспособности[2], проведены экономический анализ глобализации[3] или обзор методов построения композитного индекса [7]. Обширный обзор литературы по теме построения композитных индексов имеется также в работе [8].

[1]Handbook on constructing composite indicators: methodology and user guide. Paris : OECD, 2008. 162 p.
[2]Methodologies to measure market competition. Paris : OECD, 2021. 55 p.
[3]Measuring globalisation : handbook on economic globalisation indicators. Paris : OECD, 2005. 230 p.

Опубликовано огромное количество работ, посвященных моделям рейтинга в конкретных областях мировой и национальной экономики. Использование рейтингового подхода к рассмотрению сложных процессов в экономике можно найти в [9–22].

Позиции российских ученых по вопросу построения сводных критериев эффективности различных систем подробно проанализированы в работе [23], где утверждается, что метод сводных показателей предложил выдающийся математик А. Н. Крылов еще в 1908 г., дав обзор моделей построения композитных индексов оценки качества жизни в [24].

Рейтинги и рэнкинги: уточнение понятий и исторические модели

Теория рейтингов (лат. *rating*) – это раздел математической экономики, изучающий методы измерения сравнительных преимуществ одних экономических объектов перед другими и отнесение к классу подобных ему объектов, разряду, категории (буквально слово *rating* переводится как «положение, класс, разряд, ранг»), с течением времени возможна миграция объекта в другой класс.

Рэнкинг (лат. *ranking*) – это упорядочивание объектов в линейный (от первого до последнего) список (шеренгу) по некоторому правилу.

С проблемой сравнения классификации (не только экономических систем) человечество сталкивается с незапамятных времен и практически с момента своего зарождения использует те или иные методики проведения, сравнения и отбора лучших (победителей) или упорядочения всех объектов, систем (соревнования, турниры, разведка, экзамены, экспертизы, выборы, советы, тендеры, конкурсы, консилиумы, комиссии и т. п.). Подходы и модели могут различаться, но суть их применения одна: некие экономические объекты, системы или процессы – государства, отрасли, регионы, города, предприятия, банки, гостиницы, тендерные предложения, университеты и т. д. – необходимо либо ранжировать (упорядочить от наилучшего к худшему), либо отнести к тому ли иному классу однородных равноценных объектов.

Есть три важные особенности подобных классов задач: 1) многокритериальность, при которой качество каждого объекта оценивается по многим, как правило несоизмеримым между собой, показателям; 2) противоречивость и субъективность информации[4]; 3) динамичность[5].

По исходной информации задачи рейтинга делятся на два класса: с заданной информацией о состоянии каждого объекта в отдельности и с заданной информацией о попарном сравнении всех объектов.

Идея рэнкинга, т. е. ранжирования от первого до последнего, принадлежит известному французскому ученому эпохи Великой французской революции Ж. Ш. Борда, позднее к нему примкнул и математик М. Кондорсе. На основе своих моделей они пытались построить теорию справедливых выборов – рейтинг политиков, программ, партий. Их идеи и модели легли в основу современной математической теории коллективного выбора, рейтингов. Это фактически раздел теории рейтингов. Понятие «рейтинг» употребляется чаще, заменяя понятие «рэнкинг». Следуя этой практике, мы тоже будем употреблять в обоих смыслах термин «рейтинг».

Необходимость использования рейтингов достаточно очевидна. К их достоинствам относятся следующие:

- устранение асимметрии экономической информации о рынках по Стиглицу – Акерлофу – Спенсу (лауреаты Нобелевской премии по экономике 2001 г. за это открытие);
- наглядность (достаточно легко сравнивать объекты по различным направлениям деятельности по единой сводной таблице или классу);
- компактность (агрегация многих показателей в небольшое число индексов улучшает способность людей воспринимать одновременно десятки показателей);
- сегментация (в отдельных индексах характеризуется только одна определенная характеристика экономического процесса (надежность, эффективность и т. д.));
- системность (дается целостное представление о всех сторонах объекта);
- робастность (устойчивость во времени, т. е. объекты могут менять класс или место, но сохранять некоторую стабильность).

Постановка задачи. В дальнейшем независимо от числа объектов и принципа отбора показателей, характеризующих их работу, исходные данные для рейтинга будем представлять в виде матрицы

---

[4]При построении рейтингов учитываются как объективные статистические данные (полные, неполные, вероятностные), так и субъективные, противоречивые оценки значений критериев, данные экспертами. Первые обычно называют жесткими данными (*harddate*), вторые – мягкими данными (*softdate*).

[5]Как правило, решение принимают в фиксированный момент времени, значения критериев были заданы в прошлые периоды, а интерес представляет качество объекта в будущем, т. е. по предыстории объекта необходимо классифицировать его будущее состояние, например, по репутации фирмы и некоторым ее финансовым показателям нужно составить представление о реальных возможностях реализации ее деловых предложений.

$\left[a_{ij}\right]_{n \times m}$, где n – число объектов; m – число показателей; a_{ij} – значение j-го показателя (*hard date*) для i-го объекта либо усредненное мнение экспертов (*soft date*). В динамических рейтингах такие матрицы $\left[a_{ij}^{t}\right]_{n \times m}$ задаются для каждого момента времени t.

Необходимо либо разбить объекты на классы родственных с возможно заданным числом классов (рейтинг), либо упорядочить их от первого до последнего (рэнкинг).

Приведем исторические модели рейтингов, которые агрегируют мнения экспертов.

Балльная модель Борда. В теории ранжирования очень популярна модель другого французского математика Ж. Ш. Борда. При построении данной модели упорядочиваем все n объектов от лучшего к худшему по каждому показателю j. За последнее место объект получает 0 баллов, за предпоследнее место он получает 1 балл и т. д. За первое место объект получает $n - 1$ баллов. Баллы, полученные за каждый показатель, суммируются. В итоговом рэнкинге на первое место ставится объект с наибольшей суммой баллов. Упорядочение, полученное этим методом, обобщает мнения всех экспертов или значения всех показателей, учитывая мнение большинства. Многие методики Всемирного банка основываются на модели Борда. Однако очевиден их недостаток: они только сравнивают показатели и не учитывают их значения, т. е. происходит потеря информации об объектах.

Мажоритарная модель Кондорсе. Буквальное применение модели Кондорсе к рэнкингу объектов в поставленной задаче формулируется следующим образом: наилучшим называется объект i (с необходимостью единственный), который лучше любого по правилу большинства (рейтинговых показателей, при которых $a_{ik} > a_{jk}$, больше, чем показателей k, при которых $a_{ik} < a_{jk}$). Затем определяется наилучший объект из оставшихся, и таким образом строится рэнкинг.

Соревнование моделей Борда и Кондорсе суммирует известная теорема Фишберна, гласящая, что существуют профили предпочтений по определению худших или лучших показателей, при которых победитель по Кондорсе не может быть выявлен ни при какой модели подсчета баллов.

Модель Копленда. Объект получает +1 балл, если согласно большинству экспертов или показателей он лучше какого-то другого объекта, получает –1 балл, если он хуже какого-то другого объекта. Объект получает 0 баллов, если он не лучше и не хуже других. Рейтинг суммирует баллы при сравнении объекта со всеми другими. Получается, что оценка Копленда равна разности между количеством объектов, которые лучше данного объекта, и количеством объектов, которые хуже его, по мнению экспертов или по значениям показателей. Модель Копленда рекомендуется использовать, когда необходимо упорядочить мнение большинства и узнать, по сравнению со сколькими объектами, по мнению большинства, каждый объект выигрывает или проигрывает. Если в полученном упорядочении сумма баллов всех объектов равна количеству экспертов, то рейтинг каждого объекта указывает на то, сколько экспертов поставили его на первое место, и на то, что нет больших расхождений во мнениях экспертов. Если же сумма баллов всех объектов больше количества экспертов, это свидетельствует о том, что у экспертов значительное расхождение во мнениях. В этом случае оценки могут оказаться необъективными.

Модель Симпсона. Для каждых двух объектов i и i' находим число u_{ij} экспертов или показателей, указывающих, что объект i лучше, чем объект j. Оценкой Симпсона $u_{ij\min}$ объекта i будем называть наименьшее из чисел $(u_{i1}, \ldots, u_{in})$. Объекты размещаем в порядке убывания оценки Симпсона. В результате упорядочения происходит усреднение мнений экспертов. Эту модель рекомендуется использовать при отсутствии больших расхождений экспертов во мнениях об упорядочении объектов. Тогда модель выделит самые лучшие объекты по мнению самых необъективных экспертов. С помощью этой модели можно обнаружить необъективность экспертов.

Модель Нэша. Объекты упорядочиваются по произведению всех своих показателей (аналог среднего геометрического без корня). Оценка 0 не учитывается. Эту модель не рекомендуется использовать при большом разбросе оценок.

В перечисленных исторических моделях все эксперты (показатели) равноправны, т. е. вес каждого из них равняется 1. Но при обработке мнений экспертов или значений показателей можно пользоваться методом упорядочения с весами экспертов или показателей, но это вопрос отдельной теории утилитарных коллективных функций.

Модель рейтинга на основе сводного рейтингового индекса (агрегация показателей)

Построение модели состоит из четырех этапов: отбора показателей, перехода к безразмерным показателям (шкалирование), агрегации показателей в сводный рейтинговый индекс, упорядочения в линейный список (рэнкинг). Рассмотрим подробнее эти этапы.

Этап 1: формулировка цели рейтинга и отбор адекватной цели рейтинговых показателей. Каждая модель рейтинга, имеющая свою цель, опирается на соответствующий набор показателей. Можно использовать немного показателей, это хорошо известно специалистам по распознаванию образов, однако определить, какие показатели существенны, трудно, поэтому на практике применяют достаточно много показателей.

Этап 2: шкалирование показателей. Этот этап представляет собой превращение показателей в безразмерные величины (баллы) путем сравнения либо с лучшим, либо со средним, либо с эталонным значением или другим способом. Шкалирование проводят по одной из формул построения композитных (в русскоязычной практике сводных) индикаторов (индексов) согласно рекомендациям ОЭСР[6]:

- ранжирование (фактически это модель Борда)

$$\bar{a}_{ij} = n - \mathrm{Rank}_{ij}; \tag{1}$$

- сравнение с лучшим значением

$$\bar{a}_{ij} = \frac{a_{ij}}{\max a_{ij}}; \tag{2}$$

- решкалирование (*re-scaling*)[7]

$$\bar{a}_{ij} = \frac{a_{ij} - \min a_{ij}}{\max a_{ij} - \min a_{ij}}; \tag{3}$$

- сравнение со средним значением

$$\bar{a}_{ij} = \frac{a_{ij}}{a_{cpj}} \text{ или } \bar{a}_{ij} = \frac{a_{ij} - a_{cpj}}{a_{cpj}}; \tag{4}$$

- стандартизация

$$\bar{a}_{ij} = \frac{a_{ij} - a_{cpi}}{\sigma_j}; \tag{5}$$

- квантильное шкалирование

$$\bar{a}_{ij} = \begin{cases} 25, \text{ если } a_{ij} \in 25\ \% \text{ (1-й худший квантиль)}, \\ 50, \text{ если } a_{ij} \in \{25\ \%, 50\ \%\} \text{ (2-й худший квантиль)}, \\ 75, \text{ если } a_{ij} \in \{50\ \%, 75\ \%\} \text{ (3-й худший квантиль)}, \\ 100, \text{ если } a_{ij} \in \{75\ \%, 100\ \%\} \text{ (4-й худший квантиль)}. \end{cases} \tag{6}$$

В приведенных формулах a_{cpj} – среднее значение показателя j; $\sigma_j = \sum_{i=1}^{n} (a_{ij} - a_{cpj})^2$ – дисперсия.

Выбор подходящей формулы шкалирования из (1) – (6) является непростой задачей, при этом нужно учитывать как свойства данных, так и цель построения рейтинга. Для того чтобы оценить влияние используемой формулы шкалирования на получаемые результаты, необходимо провести тесты на устойчивость рейтинга. Формула ранжирования Борда (1) применяется при большом разбросе показателей, при этом теряется информация об их значении, а используется только их отношение друг с другом. В двух наиболее часто употребляемых формулах (2) и (3) производится отображение показателей на интервал [0, 1] или [0, 100] в случае умножения на 100. Однако значительный выброс показателя у объекта дает ему балл 1, существенно уменьшая количество баллов у остальных и стирая их различия. Чтобы этого избежать, часто берут логарифмические значения показателей. Для динамических рейтингов важна робастность, т. е. непрерывность рейтингового индекса для каждого объекта во времени, достижение которой невозможно при использовании формул (2) и (3), так как лучший и худший показатели для следующего года изменятся. Чтобы избежать такой ситуации, берут пороговые значения для худшего и лучшего показателей на длительную перспективу. Так, в рейтинге стран по индексу человеческого развития за пороги для ожидаемой продолжительности жизни взяты 85 и 20 лет, для валового национального дохода – 100 и 75 000 долл. США.

[6]Handbook on constructing composite indicators…

[7]В решкалировании балл от 0 до 1 иногда умножают на 100, так делают, например, при построении глобального инновационного индекса GII.

Этап 3: построение рейтингового (композитного) критерия как некоторой функции свертки шкалированных показателей. Наибольшее распространение получили аддитивные или мультипликативные функции $F(i)$ в следующем виде:

$$F(i) = \sum_{j=1}^{m} f\left(\overline{a}_{ij}\right) \text{ или } F(i) = \prod_{j=1}^{m} f\left(\overline{a}_{ij}\right),$$

где f – некоторая возрастающая функция, которая выбирается из одного из следующих классов:

$$F(i) = \sum_{j=1}^{n} c_j a_{ij} \text{ – линейная функция (взвешенное среднее) или}$$

$$F(i) = (\operatorname{sign} q) \prod_{i=1}^{m} \overline{a}_{ij}^{\,q} \text{ – степенная функция.}$$

В модели Нэша степенная функция имеет вид $F(i) = \prod_{j=1}^{m} \overline{a}_{ij}$. Очень часто берется среднее геометрическое: $F(i) = \sqrt[m]{\prod_{j=1}^{m} \overline{a}_{ij}}$.

При большом разбросе показателей используют логарифмическую функцию

$$F(i) = \sum_{j=1}^{n} \log a_{ij} \text{ или } F(i) = \prod_{j=1}^{m} \log a_{ij}.$$

Если хотят увеличить разброс, используют экспоненциальную функцию

$$F(i) = (\operatorname{sign} p) \prod_{j=1}^{m} e^{p a_{ij}}.$$

Можно дать простейшие рекомендации по выбору рейтинговой функции. Чаще всего из-за простоты отдается предпочтение линейной функции с обоснованием весов, которые выбираются как усреднение мнений экспертов (метод Дельфи) или математически с помощью метода главных компонент. Однако в большинстве рейтингов берут равные веса, считая показатели равнозначными.

Логарифмическую функцию рекомендуется использовать, когда оценки объектов имеют большой разброс, при применении такой функции вперед выйдут объекты с равномерными значениями показателей.

При использовании экспоненциальной функции с положительной степенью ($p > 0$) более высокие оценки сильнее влияют на результат, чем более низкие. Такую функцию хорошо использовать, когда необходимо дать преимущество резко вырвавшимся вперед по отдельным показателям объектам. Выбор значения коэффициента p зависит от величины показателей. Если показатели выражаются большими числами, то коэффициент p лучше брать ближе к 0, потому что при суммировании не будут учитываться низкие оценки. Если показатели выражаются небольшими числами, то коэффициент p рекомендуется брать равным 1 и больше, чтобы слагаемые не были слишком маленькими и рейтинги отличались друг от друга на большую величину. Для диапазона оценок от 0 до 100 удобны коэффициенты, равные 0,1.

В противоположность предыдущей функции более низкие показатели экспоненциальной функции с отрицательной степенью ($p < 0$) сильнее влияют на результат благодаря отрицательному показателю степени. Эту функцию хорошо использовать, когда нужно, чтобы более низкие показатели были тоже учтены в конечном результате. Для диапазона от 0 до 100 предлагается принимать p равным –0,1.

Степенную функцию с положительной степенью ($0 < q < 1$) рекомендуется применять при необходимости сгладить высокие оценки, поскольку большие числа при возведении в степень q уменьшаются в большее число раз, чем меньшие оценки, а при $q > 1$ большие оценки влияют на рейтинг сильнее, чем меньшие.

При использовании степенной функции с отрицательной степенью большие оценки при возведении в степень q уменьшатся в большее число раз, чем меньшие оценки. При подсчете рейтинга большие оценки сглаживаются сильнее, чем меньшие. Для диапазона оценок от 0 до 100 часто используется значение $q = -0{,}3$. Функцию рекомендуется использовать при желании сгладить высокие оценки.

Этап 4: линейное упорядочение объектов. Этот процесс осуществляется по убыванию значений рейтингового сводного индекса $F(i)$.

Главный недостаток почти всех рейтингов, построенных на основе композитного индекса, – использование единственного результирующего критерия для упорядочивания объектов. Действительное состояние объекта характеризуется множеством экономических показателей. Свести все эти показатели к одному числу (баллу), которое определяет место в рейтинге с помощью математических методов, большой проблемы не представляет. Однако проблема сравнения объектов является принципиально многокритериальной, и поэтому любой линейный список (это хорошо знают математики) характеризуется субъективизмом и в весах отдельных показателей, и в схеме отбора самих показателей, и в функции свертки показателей. Поэтому к рейтингам, построенным в виде списков, в которых все объекты ранжированы от первого до последнего места, даже если это сделано только внутри групп однотипных объектов (больших, средних, малых), следует относиться как к мнению экспертов, составлявших модель, а не как к объективной реальности. Безусловно, могут быть подобраны более адекватные реальности и менее субъективные модели. Существует простой критерий: если одни и те же объекты в разных рейтингах, составленных разными институтами по разным моделям, попадают в первые или последние десятки, то так оно и есть. Если при этом будет выявлена корреляция между действительными результатами и тем, о чем сигнализировал некоторый рейтинг, это придаст данному рэнкингу вес и авторитет.

Пример 1: модель индекса человеческого развития. Данный индекс ООН рассчитывает с 1990 г. (*human development index*, HDI), берет среднее геометрическое трех шкалированных показателей: продолжительности жизни, времени обучения (текущее и ожидаемое) и валового национального дохода:

$$\mathrm{HDI} = \left(\overline{\mathrm{HEALTH}} \cdot \overline{\mathrm{EDU}} \cdot \overline{\mathrm{GNI}}\right)^{\frac{1}{3}}, \tag{7}$$

где HEALTH – здоровье; EDU – образование; GNI – валовый национальный доход.

Расчет индекса HDI по формулам (3) и (7) приведем на примере Беларуси:

$$\overline{\mathrm{HEALTH}}_{\text{Беларусь}} = \frac{72,4 - 20}{85 - 20} = 0,806,$$

$$\overline{\mathrm{EDU}}_{\text{Беларусь}} = \frac{1}{2}\left(\overline{\mathrm{EDU}}_{\text{Беларусь}}^{\text{expected}} + \overline{\mathrm{EDU}}_{\text{Беларусь}}^{\text{mean}}\right) = \frac{1}{2}\left(\left(\frac{15,2 - 0}{18 - 0}\right) + \left(\frac{12,1 - 0}{15 - 0}\right)\right) = 0,826,$$

$$\overline{\mathrm{GNI}}_{\text{Беларусь}} = \frac{\ln(18,849) - \ln(100)}{\ln(75\,000) - \ln(100)} = 0,791.$$

Валовый национальный доход берется по паритету покупательной способности в долларах США 2017 г.

В результате HDI для Беларуси составляет 0,808, что соответствует 60-му месту в мире (для сравнения: HDI мирового лидера – Швейцарии – равен 0,962).

ООН вычисляет и другие композитные рейтинговые индексы: индекс неравенства в человеческом развитии, индекс гендерного развития, индекс гендерного неравенства, многомерный индекс бедности – с помощью аналогичных моделей (среднее геометрическое и шкалирование по формуле (3), но с определением экспертных весов по методу главных компонент).

Пример 2: модель ООН рейтинга стран по достижению целей устойчивого развития. Новая идея в этой модели рейтинга ООН заключается в том, что большое число показателей (93) агрегируется не сразу в сводный индекс, а сначала в 17 субиндексов, каждый из которых отвечает за одну из 17 целей развития ООН, и только затем вычисляется сводный индекс как среднее арифметическое своих субиндексов (ООН считает, что все 17 целей тысячелетия равнозначны). Шкалирование всех показателей осуществляется по формуле (3) с умножением на 100, а их свертка в субиндексы, как правило, представляет собой среднее арифметическое или среднее геометрическое. Беларусь в рейтинге 2023 г. занимает 34-е место в мире (77,5 балла), США – 39-е место, Финляндия – 1-е место.

Заметим, что такая иерархическая агрегация показателей (сначала строятся субиндексы показателей на нижнем уровне, затем сводный индекс из субиндексов на верхнем уровне) присуща большинству рейтингов, причем количество таких уровней бывает и три, и четыре.

Пример 3: рейтинг глобального инновационного развития. В век экономики знаний популярность завоевал рейтинг Всемирной организации интеллектуальной собственности (WIPO), созданный на основе глобального инновационного индекса (*Global innovation index*, GII), ранжирующий к октябрю каждого года 132 страны мира. Новая идея в модели этого рейтинга – разбиение субиндексов, а следовательно и показателей, на две группы: input (потенциал) и output (результат), что позволяет находить рейтинг по индексу output/input, определяющему отношение результатов к потенциалу, который показывает, насколько эффективно страны используют свой инновационный потенциал. Например, Беларусь в 2022 г. оказалась на 86-м месте по потенциалу и на 63-м месте по результатам (в основном за счет ИКТ-экспорта, сложности экспортной продукции, внедрения ИСО-стандартов, производительности труда), что говорит об эффективном использовании потенциала. США по потенциалу занимает 2-е место, а по результатам – 5-е место в отличие от Китая, который пока по потенциалу находится на 21-м месте, а по результатам – на 8-м месте в мире.

Пример 4: рейтинг скорости развития процесса. Практически во всех рейтингах сравнивают текущее состояние того или иного экономического процесса в странах. При этом очень важно для прогноза будущего сравнивать скорость развития процесса, т. е. скорость догоняющей или обгоняющей мирового лидера (США) модернизации. Для оценки скорости процесса в сводном индексе необходимо агрегировать в качестве показателей индексы роста, что не требует применения процедур шкалирования. В подобных рейтингах в лидеры будут выходить страны, находящиеся на низком уровне и поэтому стремительно догоняющие мировых лидеров.

Рейтинговая модель линейных порядков

Модель линейных порядков удобно применять с помощью показателей, когда для всех объектов имеются данные об их попарном сравнении экспертами, т. е. для любых объектов i и j дано число C_{ij}, показывающее, насколько i-й объект лучше j-го, например, C_{ij} может показывать число показателей, по которым i-й объект лучше j-го. По заданной матрице $\left\| C_{ij} \right\|_{n \times n}$ необходимо найти перестановку объектов $\pi^* = \left(\pi_1^*, \ldots, \pi_n^* \right)$, на которой достигается следующее равенство:

$$\sum \sum C_{\pi^*(i)\pi^*(j)} = \max_{\pi} \sum \sum C_{\pi(i)\pi(j)}. \tag{8}$$

Формула (8) – это так называемая задача линейных порядков. Чтобы ее свести к задаче дискретной оптимизации, нужно ввести следующие переменные: $x_{ji} = 1$, если в искомом порядке π объект i предшествует объекту j, и $x_{ij} = 0$ в противном случае.

Тогда задача линейных порядков (8) эквивалентна следующей задаче дискретной оптимизации:

$$\max \sum \sum C_{ij} x_{ij},$$

$$x_{ij} = 0 \text{ или } 1, \; x_{ij} + x_{ji} = 1, \; i, j = 1, \ldots, n, \tag{9}$$

$$x_{ij} + x_{jk} + x_{ki} \le 2, \; i, j, k = 1, \ldots, n. \tag{10}$$

Последняя задача решается стандартными методами дискретной оптимизации, например построения фасетных неравенств, отсекающих нецелочисленные вершины многогранника (9) и (10) [23]. Модель применялась для ранжирования белорусских банков [12] и дала хорошие результаты.

Модель свертки рейтинговых индексов

Во многих ситуациях анализа экономических процессов необходимо из нескольких рейтинговых индексов получить один, например, для расчета роста совокупной факторной производительности [25]. Простейшие модели для агрегации индексов $I_1, \ldots, I_s$, переведенных в одинаковую меру от 0 до 1, имеют следующий вид:

$$I = \frac{\sum_{k=1}^{s} I_k}{s},$$

$$I = \sum_{k=1}^{s} \frac{I_k^* - I_k}{I_k^*},$$

$$I = \frac{100 \left(\sum_{k=1}^{s} \frac{I_k}{I_k^*} \right)}{s},$$

где I_k^* – значение у лучшей страны k-го индекса.

Модернизируем предложенную в [25] более сложную модель агрегации индексов следующим образом:

$$I = \frac{\sum_{k=1}^{s} \left(\frac{\ln I_k^1}{\ln I^{*1}} e \left(\frac{I_k^1}{I^{*1}} - \frac{I_k^0}{I^{*0}} \right) \right)}{s},$$

где I_k^1, I_k^0 – индексы страны в двух k-х рейтингах: последнем и предыдущем.

Рейтинговая модель факторного анализа

Ранжирование на основе композитного индекса сводит показатели в один фактор. Ранжирование на основе моделей с двумя, тремя и более факторами эффективнее. Например, модель на основе двухфакторного анализа представляет объекты точками на плоскости факторов, что особенно удобно при составлении динамических рейтингов и отслеживании траектории движения объекта, так как в двухфакторной динамической модели во времени объекты-точки превращаются в траектории, указывающие, в каком направлении они движутся (рис. 1). Идеей построения моделей факторного анализа выступают выделение из многочисленных показателей некоррелированных главных факторов (главных компонент) и исключение остальных. Впрочем, существует мнение, что метод главных компонент (*principal component analysis*, PCA) слишком упрощает идеи факторного анализа.

В рейтинговой факторной модели предполагается, что известные показатели a_{ij} являются линейными функциями некоторых неизвестных нам факторов, определенных на множестве рейтингуемых объектов. Некоторые из них являются общими для нескольких показателей, а некоторые другие выступают характерными, возможно, только для одного показателя. Характерные факторы полагаются ортогональными друг другу и, следовательно, не вносящими вклад в ковариацию. Отбрасывая характерные факторы (приравнивая их к нулю), получаем следующее выражение, аппроксимирующее каждый вектор-показатель:

$$A^j = x_{j1}F_1 + \ldots + x_{jp}F_p + \ldots + x_{jp}F_r, \;\; j = 1, \ldots, m,$$

где x_{jp} – вес фактора p в показателе j; $A^j = (a_{1j}, \ldots, a_{nj})$ – вектор значений показателя j на множестве объектов.

Необходимо определить факторы и факторную матрицу $\left[x_{jp}\right]_{m \times r}$ так, чтобы они наиболее точно аппроксимировали значения вектор-показателя, при этом желательно, чтобы факторы несли в себе некий содержательный смысл. Например, в анализе экономических объектов факторами могут быть масштабность, эффективность, надежность.

Метод главных компонент применяется в три этапа.

Этап 1: построение корреляционной матрицы. Значения каждого из показателей $\left[a_{ij}\right]_{n \times m}$ считаются нормированными, т. е. среднее арифметическое каждого показателя равно 0, дисперсия равна 1. Если это условие не выполняется, то проводим нормировку

$$\overline{a}_j = \frac{1}{n}\sum_{i=1}^{n} a_{ij}.$$

Сделаем замену значений показателей на их нормированные значения:

$$a_{ij} = \frac{\left(a_{ij} - \overline{a}_j\right)}{\sigma_j},$$

где $\overline{a}_j$ – среднее арифметическое показателя j; $\sigma_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(a_{ij} - \overline{a}_j\right)^2}$ – дисперсия.

Построим ковариационную матрицу $C = A^T \times \frac{A}{n}$, затем вычислим корреляционную матрицу $R = \left[r_{ij}\right]_{m \times m}$:

$$r_{ij} = \frac{c_{ij}}{\sqrt{c_{ii}c_{jj}}}$$

– и нормируем у нее диагональ:

$$r_{ii} = \frac{1}{n}\sum_{k=1}^{n} r_{ik}, \;\; i \neq k, i = 1, \ldots, n,$$

которая есть общность показателя i.

Этап 2: выделение главных факторов и факторной матрицы $\left[x_{jp}\right]_{m \times r}$. Сначала ищем коэффициенты x_{i1} при первом факторе так, чтобы достигался максимум

$$\max\left(x_{11}^2 + x_{21}^2 + \ldots + x_{m1}^2\right) \tag{11}$$

при условии, что

$$\sum_{p=1}^{r} x_{jp}x_{kp} = r_{jk}, \; j, \; k = 1, \ldots, m. \tag{12}$$

Идея задачи (11), (12) – сделать среднеквадратичное отклонение максимальным. С помощью множителей Лагранжа легко доказывается, что искомый максимум является ничем иным, как максимальным собственным решением следующего уравнения:

$$\det\left(R-\lambda E\right)=0.$$

При помощи метода вращений Якоби находим все собственные значения редуцированной матрицы R и соответствующие им собственные векторы q_{11}, …, q_{n1}. В итоге получим искомые коэффициенты при всех факторах F_p:

$$x_p=\frac{q_{j1}\sqrt{\lambda_1}}{\sqrt{q_{11}^2+q_{21}^2+\ldots+q_{n1}^2}}.$$

Этап 3: вращение факторов. Идея этого этапа заключается в том, чтобы каждый следующий фактор сделать ортогональным предыдущим. Полученную факторную матрицу $\left[x_{jp}\right]_{m\times r}$ размерности $m \times r$ вращаем путем умножения на матрицу поворота T. Для двухфакторного анализа матрица поворота T имеет следующий вид:

$$T=\begin{bmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{bmatrix},$$

где α – угол поворота. Оптимальное значение угла поворота α получаем путем перебора всех значений от 0 до $\frac{\pi}{2}$ с шагом $\frac{\pi}{720}$.

Окончательный поворот будет произведен на угол, при котором для превалирующих в данном факторе показателей сумма квадратов отклонений (дисперсия) от абсолютного максимума (M) должна быть минимальна:

$$\sum_{p=1}^{r}\sum_{j=T_p}\left(M_j-a_j\right)^2\to\min,$$

здесь T_p – множество показателей, которые должны нести в p-м факторе основную нагрузку.

Умножив справа факторную матрицу $X=\left[x_{jp}\right]_{m\times r}$ на T, получим окончательную матрицу, показывающую расположение объектов в полученных факторах F_1, …, F_r.

Экономисты не обязаны знать все тонкости рейтинговой модели факторного анализа, они могут воспользоваться программами в *SPSS* или *STATISTICA*. Здесь модельные нюансы приведены для лучшего понимания роли факторов, иногда их называют главными компонентами.

В двух- или трехфакторных моделях объекты удобно разбивать на классы, указывая интервалы значений для классов по каждому фактору. Например, объект попадает в класс A, если фактор масштабности находится в интервале $[a, b]$, если меньше a, то это класс A^-, если больше, то это класс A^+ (рис. 1). Таким образом, в 1999 г. объект (страна) попадает в класс (A, B^-), хотя в 1994 г. был в классе (A^-, B^+).

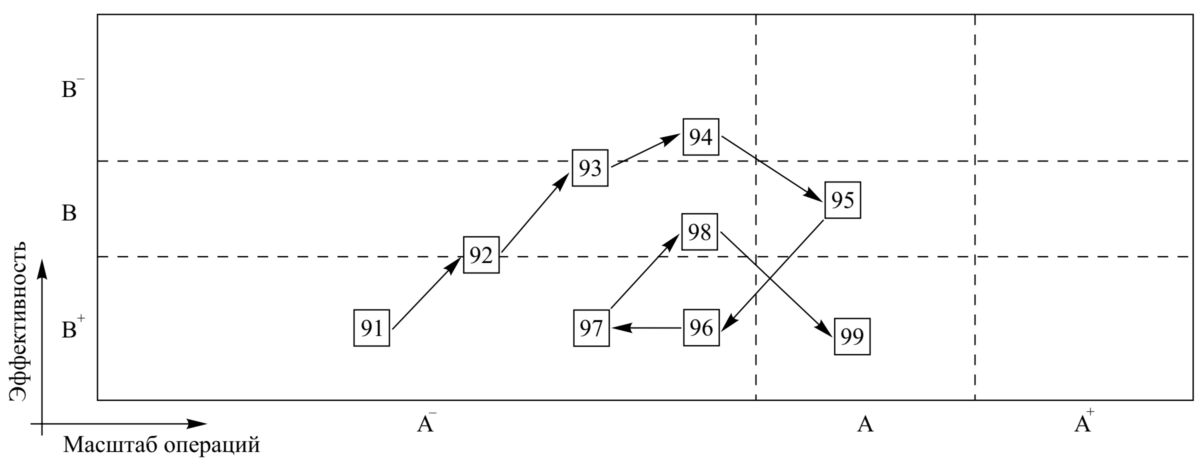

Рис. 1. Динамика развития объекта в двухфакторном пространстве *масштабность – эффективность*

Fig. 1. The dynamics of the development of an object in a two-factor space *scale – efficiency*

Под масштабностью понимается доля страны в мировом ВВП по ППС, в мировом населении и мировом экспорте, а под эффективностью – ВВП на жителя, продолжительность жизни, энергоемкость ВВП, коэффициент Джини.

Рейтинговая модель кластер-анализа

Рейтинговые модели, основанные на разделении однотипных объектов на кластеры (группы), являются основным инструментом работы международных рейтинговых агентств, которые определяют место государств, банков, предприятий в международных рейтингах путем отнесения к определенному кластеру.

В рейтинговой модели кластер-анализа сначала отбираются рейтинговые показатели и каждый объект представляется в m-мерном метрическом пространстве, определяющемся некоторой метрикой (евклидовой, Минковского, Чебышева и др.). Далее с помощью одного из методов кластер-анализа объекты группируются в несколько однотипных кластеров. Точнее, ведется поиск наилучшей группировки (кластеризация) объектов в пространстве отобранных рейтинговых показателей. Количество таких множеств (кластеров) и их содержательный смысл могут быть либо заданы заранее, либо получиться автоматически. Рейтинговые модели кластер-анализа бывают двух классов: разделяющие и агломеративные (объединяющие). Их различие состоит в том, что в моделях первого класса все объекты объединяются сначала в один кластер, а потом пошагово делятся, а в моделях второго класса кластеров сначала столько же, сколько и объектов, и они пошагово объединяются.

Объединяющие модели кластер-анализа. Каждый объект i рассматривается как m-мерный вектор $A_i = (a_{i1}, \ldots, a_{im})$ его показателей. Все объекты дают n точек в m-мерном метрическом пространстве, и их необходимо разбить на заданное число кластеров максимально близко расположенных точек.

При построении этой модели все объекты рассматриваются как совокупность одноэлементных кластеров. Из них выбираются два объекта, которые в смысле некоторой меры близости наиболее близки друг к другу, и объединяются в один кластер. Новая совокупность кластеров содержит на один кластер меньше. С помощью повторения процесса получающаяся совокупность кластеров сводится к требуемому количеству кластеров, каждый из которых содержит близкие объекты.

В качестве меры близости кластеров может применяться, например, расстояние между их центрами. Изначально матрица расстояний $D = \left\{d_{ij}^2\right\}$, где d_{ij} есть дисперсия между вектор-показателями объектов i и j, содержит n строк и n столбцов. После объединения двух кластеров в новый кластер матрицу расстояний необходимо пересчитать. Теперь эта матрица содержит на одну строку и на один столбец меньше. Пересчитывать необходимо лишь одну строку и один столбец для объединяемых кластеров, поскольку остальные элементы не изменятся. Вычисления можно свести к минимуму, выразив эти новые элементы через элементы старой матрицы расстояний. При этом вариативные модели могут использоваться из-за разных мер близости двух кластеров.

В модели максимального локального расстояния два кластера объединяются, если максимальное расстояние между точками одного кластера и точками другого минимально.

Модель минимального локального расстояния аналогична предыдущей, но с точностью до наоборот. Эта модель склонна создавать крупные кластеры.

В модели Уорда в качестве функции расстояния применяется внутригрупповая сумма квадратов отклонений, которая является суммой квадратов расстояний между каждым объектом и средним объектом в кластере. На каждом шаге объединяются такие два кластера, которые приводят к минимальному увеличению внутригрупповой суммы квадратов расстояний.

В центроидной модели близость кластеров определяется как расстояние между их центрами.

В модели групповых средних два кластера объединяются, если их среднее сходство максимально среди прочих. Среднее сходство между кластерами определяется как среднее сходство между всеми парами объектов из соответствующих кластеров.

Приведем рекурсивную схему, интегрирующую различные модели кластер-анализа за счет разных способов перерасчета расстояний.

Квадрат расстояния (дисперсия) между объектами i и j определяется по следующей формуле:

$$d_{ij}^2 = \left(A_i - A_j\right)^{\mathrm{T}} \left(A_i - A_j\right).$$

Сначала необходимо найти два кластера p и q, расстояние (дисперсия) между которыми минимально:

$$d_{pq}^2 = \min\left\{d_{ij}^2\right\},\ i = 1, \ldots, j-1,\ j = 2, \ldots, n,\ n_i > 0,\ n_j > 0.$$

Затем кластер p объединяется с кластером q. Мера близости при этом увеличивается на величину, рассчитываемую по формуле

$$W_{pq} = \frac{1}{2} d_{pq}^2.$$

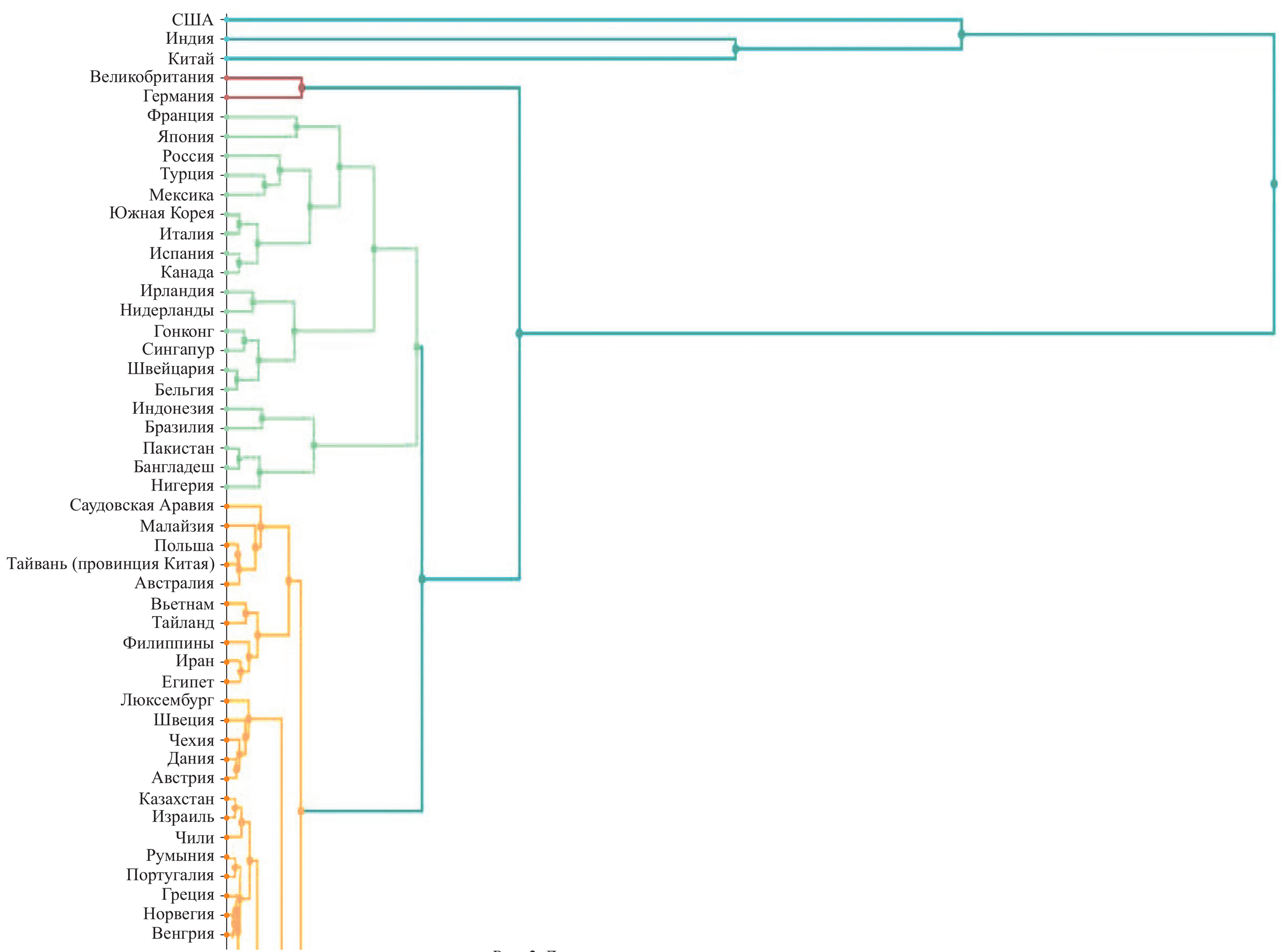

Рис. 2. Дендрограмма кластер-анализа стран мира
Fig. 2. Dendrogram of cluster analysis of countries

Далее строка и столбец, соответствующие новому кластеру h, пересчитываются по следующей формуле:

$$d_{hk}^2 = \alpha_i d_{hi}^2 + \alpha_j d_{hi}^2 + \beta d_{ij}^2 + \gamma \left| d_{hi}^2 - d_{hj}^2 \right|,$$

где значения параметров α, β, γ соответствуют каждой модели и имеют следующие значения: $\alpha_i = \alpha_j = \frac{1}{2}$, $\beta = 0$, $\gamma = -\frac{1}{2}$ (для модели минимальных локальных расстояний); $\alpha_i = \alpha_j = \frac{1}{2}$, $\beta = 0$, $\gamma = \frac{1}{2}$ (для модели максимальных локальных расстояний); $\alpha_i = \alpha_j = \frac{1}{2}$, $\beta = \frac{1}{4}$, $\gamma = 0$ (для модели медиан); $\alpha_i = \frac{n_i}{n_k}$, $\alpha_j = \frac{n_j}{n_k}$, $\beta = \gamma = 0$ (для модели средних групп); $\alpha_i = \frac{n_i}{n_k}$, $\alpha_j = \frac{n_j}{n_k}$, $\beta = -a_i a_j$, $\gamma = 0$ (для центроидной модели); $\alpha_i = \frac{(n_h + n_i)}{(n_h + n_k)}$, $\alpha_j = \frac{(n_h + n_j)}{(n_h + n_k)}$, $\beta = -\frac{n_h}{(n_h + n_k)}$, $\gamma = 0$ (для модели Уорда).

Наконец, возвращаемся к вычислению квадратов кластеров и повторяем процедуру до тех пор, пока не получим требуемое количество кластеров, например три: объекты с высокими показателями (класс А), объекты со средними показателями (класс В) и объекты с низкими показателями (класс С).

Пример 5: кластеризация государств по модели Уорда. Развитие каждого государства характеризуется тремя показателями: долей в мировом ВВП по ППС, долей в населении и долей в торговле. Фрагмент дендрограммы, построенной по матрице мер близости всех стран мира, представлен на рис. 2.

Модель выделила пять кластеров государств (и их объединений): сверхзначимые, или супербольшие (Китай, США, ЕС, Индия), значимые, или большие (Великобритания, Германия, Франция, Япония, Россия, Турция, Мексика, Корея, Италия, Испания, Канада, Ирландия, Нидерланды), среднезначимые, или средние (Польша и др.), малозначимые, или малые (Беларусь, Израиль и др.), микрозначимые (Литва, Латвия, Эстония и др.).

Рейтинговая модель оптимизационного разбиения на классы

Число классов заранее не определяется. С каждым показателем k сопоставляется некоторое бинарное отношение R_k, которое характеризует похожесть объектов i и j по показателю k. Обычно $R_k = \left\{ (i, j) : \left| a_{ik} - a_{jk} \right| \le e_k \right\}$, где e_k есть заданная вариация значений показателя k (если объекты i и j имеют различия в показателе k не более e_k, то считается, что они схожи по этому показателю). Заметим, что R_k может быть мнением эксперта по вопросу отнесения объектов к одному классу. Искомое разбиение объектов на непересекающиеся классы однотипных объектов представляет собой ничто иное, как отношение эквивалентности R на всем множестве объектов. Искомые классы определены отношением эквивалентности R. Для определения наилучшей группировки необходимо сравнить искомое отношение эквивалентности R с каждым из бинарных отношений R_k, т. е. подсчитать количество рассогласований между R и R_k:

$$\left| R \Delta R_k \right| = \left\{ (i, j) / (i, j) \in R \text{ и } (i, j) \notin R_k \text{ или } (i, j) \notin R \text{ и } (i, j) \in R_k \right\}.$$

Также нужно просуммировать количество рассогласований по отношениям R_k для всех показателей k и найти оптимальное отношение эквивалентности R^*, на котором достигается $\min \sum_{k=1}^{m} \left| R \Delta R_k \right|$, где $\left| R, R_k \right|$ – выбранная метрика, например число рассогласований.

Сформулированная задача с нефиксированным числом классов сводится к задаче булевой линейной оптимизации, для чего каждое бинарное отношение R_k, R на множестве объектов представляется с помощью булевой матрицы следующим образом: для каждого k, $1 \le k \le m$, и для каждой пары объектов (i, j) пусть

$$r_{ij}^{(k)} = \begin{cases} 1, \text{ если } (i, j) \in R_k, \\ 0, \text{ если } (i, j) \notin R_k, \end{cases} \qquad r_{ij} = \begin{cases} 1, \text{ если } (i, j) \in R, \\ 0, \text{ если } (i, j) \notin R. \end{cases}$$

Нелинейную целевую функцию $\sum_{k=1}^{m} \left| R \Delta R_k \right|$ можно линеаризовать, а требования к искомому отношению эквивалентности R (рефлексивность, симметричность и транзитивность) выразить с помощью линейных неравенств, исключив часть переменных r_{ij}, в силу того что $r_{ij} = r_{ji}$ (порядок i и j несуществен, так что можно считать $i < j$), и определить новые веса $w_{ij} = c_{ij} + c_{ji}$. В результате возникает следующая задача линейной булевой оптимизации поиска матрицы $\left[x_{ij}\right]_{n \times n}$, определяющей оптимальное отношение эквивалентности R:

$$\min \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} x_{ij},$$

при условиях, что

$$x_{ij} + x_{jk} - x_{ik} \le 1,\ 1 \le i < j < k \le n,$$
$$x_{ij} - x_{jk} + x_{ik} \le 1,\ 1 \le i < j < k \le n,$$
$$-x_{ij} + x_{jk} + x_{ik} \le 1,\ 1 \le i < j < k \le n,$$
$$x_{ij} \in \{0, 1\},\ 1 \le i < j \le n.$$

Результатом решения последней задачи является разбиение множества объектов на классы, которые объединяют близкие друг другу объекты. Эксперты анализируют полученный результат и определяют названия классов.

Этой модели можно дать графовую интерпретацию, если все объекты представить вершинами полного n-вершинного графа, ребра которого взвешены числами w_{ij}. В таком случае задача кластеризации объектов сводится к задаче разбиения на клики полного графа со взвешенными ребрами, а $\left[x_{ij}\right]$ является их матрицей инциденций.

Когда число классов фиксировано, возникает задача разбиения взвешенного полного графа на заданное число однородных клик, это значит, что сумма ребер с обоими концами в клике должна быть минимальна по всем кликам.

Рейтинговая модель оболочечного анализа данных

Модели оболочечного анализа данных (*data envelopment analysis*, DEA), предложенные специалистами по теории принятия решений А. Чарнсом, В. Купером и Е. Родом еще в 1978 г. [26], широко применяются в экономическом анализе эффективности работы объектов, в том числе и для их рейтингования, например, в [12] она использовалась для составления рейтинга белорусских банков.

Основная идея рейтинговых моделей оболочечного типа состоит в разбиении на две группы (*output* (выход – результаты), например объемы выпуска, и *input* (вход – ресурсы), например затраты и последующее определение эффективности работы каждого объекта) всех показателей, характеризующих объекты как отношение свертки выходных показателей к свертке входных с неизвестными агрегирующими коэффициентами для выхода ($u_1, \ldots, u$) и для входа ($v_1, \ldots, v$), которые определяются как решение задач максимизации эффективности для каждого i-го объекта следующего вида:

$$\max \theta_i = \frac{\sum_{j=1}^{n_1} u_j a_{ij}}{\sum_{j=n_1+1}^{m} v_j a_{ij}}. \tag{13}$$

Далее рейтинг объектов строится в порядке убывания эффективностей θ_i по преобразованию входов в выходы.

Задача (13) легко линеаризуется и решается стандартными программами линейного программирования.

Заключение

Проведенное исследование позволяет сделать следующие выводы.

1. Популярность рейтингов и их широкое применение для сравнительного анализа процессов экономического, социального, инновационного, промышленного, экономического развития и конкуренции стран, регионов, интеграционных объединений, организаций требуют понимания математической модели рейтинга, которое позволит более полно выявлять глубинные причины низких или высоких позиций в рейтингах и указывать механизмы улучшения положения в них.

2. Для многих важных экономических процессов рейтинги составляют разные исследовательские центры, и, разумеется, результаты этих рейтингов не совпадают, хотя, как правило, коррелируют. Чтобы прояснить ситуацию, полезно составлять консенсус-рейтинг как агрегацию мест или лучших рейтинговых композитных индексов. Для их агрегации можно применять все модели, описанные в статье. Иногда,

когда рейтингов с разным количеством стран много, как, например, при измерении степени глобализации или продвижения к цифровой экономике, применяют специальные методики консенсус-рейтингов, как, например, в [27; 28], модели свертки рейтингов глобализации и цифровизации и связь между ними. Проблема консенсус-рейтингов является особо важной для инвестиционных и кредитных рейтингов и поэтому широко изучается отдельно [29]. В России, например, создан агрегатор независимых оценок качества предметного образования в вузах России, где по правилу Борда агрегируются места, затем вузы кластеризуются в лиги в зависимости от суммы баллов Борда (Национальный агрегированный рейтинг). Хотя мы считаем, что эта проблема полностью аналогична проблеме рейтинга, где показатели – это места или композитные индексы в отдельных рейтингах.

3. Большинство рейтингов стран учитывают достигнутые показатели процесса, однако политиков больше интересует скорость в догоняющем развитии. Существует очень мало рейтингов, которые агрегировали бы индексы роста показателей, т. е. таких рейтингов, где все показатели измерялись бы одной шкалой, соответственно, отпала бы необходимость их предварительного шкалирования.

Библиографические ссылки

1. Borda JC. *Mémoire sur les elections au scrutiny. Histoire de l'Academie Royale des Sciences*. Paris: [s. n.]; 1781.
2. Condorcet M. *Essai sur l'applecation de l'analyse à la probabilité des decisions rendued à la pluralité de voix*. Paris: Imprimerie royale; 1784. 304 p.
3. Мулен Э. *Кооперативное принятие решений: аксиомы и модели*. Москва: Мир; 1991. 464 с.
4. Нуреев РМ. *Теория общественного выбора*. Москва: Издательский дом ВШЭ; 2005. 531 с.
5. Эрроу КД. *Коллективный выбор и индивидуальные ценности*. Москва: Издательский дом ВШЭ; 2004. 204 с.
6. Сен А. *Идея справедливости*. Москва: Издательство Института экономической политики имени Е. Т. Гайдара; 2016. 520 с.
7. Mazziotta M, Pareto A. Methods for constructing composite indices: one for all or all for one. *Rivista Italiana di Economia Demagrafica e Statistica*. 2013;LXVII(2):67–80.
8. Greco S, Isshizako A, Tasion M, Torrisi G. On the methodological framework of composite indices: a review of the issues of weighting, aggregation and robustness. *Social Indicators Research*. 2019;141:61–94.
9. Горбач АВ, Ковалёв ММ. Как определяются международные рейтинги государств. *Вестник Ассоциации белорусских банков*. 2000;33:13–19.
10. Ковалёв ММ, Новик ВВ. Факторы, влияющие на экономический рост, конкурентоспособность и страновые рейтинги. В: Коваль ММ, редактор. *Вопросы внешней политики Республики Беларусь*. Минск: БГУ; 2001. с. 68–117.
11. Ковалёв ММ, Шибеко ИТ. Методики расчета банковских рейтингов. *Банковский вестник*. 1999;8:30–39.
12. Ковалёв ММ, Господарик ЕГ. *Банковская аналитика*. Минск: Издательский центр БГУ; 2020. 458 с.
13. Ковалёв М, Петрик О. Рейтинг информационных сайтов белорусских банков. *Вестник Ассоциации белорусских банков*. 2003;45:25–31.
14. Ковалёв ММ, Шашко АА. Инвестиционный рейтинг основных городов Беларуси. *Белорусский банковский бюллетень*. 2003;19:36–44.
15. Ковалёв ММ, Шашко АА. Малое предпринимательство в Беларуси: рейтинг развития МП в регионах. *Вестник Ассоциации белорусских банков*. 2004;28:16–34.
16. Ковалёв ММ, Шашко АА. Развитие инновационного потенциала Республики Беларусь: инновационный рейтинг областей и г. Минска. *Вестник Ассоциации белорусских банков*. 2004;38–39:24–32.
17. Girlich E, Kovalev MM, Kozulin AV. *Ranking: theorie and praxis*. Magdeburg: Otto-von-Guericke-Universität; 2001. 24 S.
18. Ковалёв ММ, Гедранович АБ. Международные рейтинги университетов. *Вышэйшая школа*. 2007;1:12–15.
19. Ковалёв ММ, Листопад НИ, Мимнюкович ЕА. Сайт – зеркало вуза и инструмент маркетинговой политики. *Вышэйшая школа*. 2011;3:58–60.
20. Ковалёв ММ, Гедранович АБ, Минюкович ЕА. Белорусский опыт рейтингования вузов. В: *Система рейтинга вузов: национальная и мировая практика. Материалы Международного научно-практического семинара; 19–20 марта 2010 г.; Алма-Ата, Казахстан*. Алма-Ата: КазНТУ имени К. И. Сатпаева; 2010. с. 41–52.
21. Ковалёв ММ, Листопад НИ, Минюкович ЕА. Вебометрический рейтинг университетов. *Информатизация образования*. 2009;2:63–73.
22. Ковалёв ММ, Курбацкий АН, Листопад НИ. Экспертная система анализа тендерных предложений компьютерного оборудования и софтвера. *Информатизация образования*. 1997;1:68–92.
23. Федотов ЮВ, Ховамов НВ. *Методы построения сводных оценок эффективности деятельности сложных производственных систем*. Санкт-Петербург: НИИ менеджмента Санкт-Петербургского университета; 2006. 33 с.
24. Жгун ТВ. Алгоритмы построения интегрального индикатора качества сложной системы для ряда последовательных наблюдений. *Вестник ЮУрГУ. Серия: Вычислительная математика и информатика*. 2017;6(1):5–25. DOI: 10.14529/cmse170101.
25. Головенчик ГГ, Ковалёв ММ. *Цифровая экономика*. Минск: Издательский центр БГУ; 2019. 395 с.
26. Charnes A, Cooper W, Rhodes E. Measuring the effectiveness of decision-making units. *European Journal of Operational Research*. 1978;2(6):429–444.
27. Bolotashvili G, Kovalev M, Girlich E. New facets of the linear ordering polytope. *SIAM Journal on Discrete Mathematics*. 1999;12(3):326–336.
28. Головенчик ГГ. *Цифровизация белорусской экономики в современных условиях глобализации*. Минск: Издательский центр БГУ; 2019. 257 с.
29. Буздалин АВ, Заночкин АЮ, Курбангалеев МЗ, Смирнов СН. Агрегация кредитных рейтингов как задача построения консенсуса в системе экспертных оценок. *Глобальные рынки и финансовый инжиниринг*. 2017;4(3):181–207. DOI: 10.18334/grfi.4.3.38830.

References

1. Borda JC. *Mémoire sur les elections au scrutiny. Histoire de l'Academie Royale des Sciences*. Paris: [s. n.]; 1781.

2. Condorcet M. *Essai sur l'applecation de l'analyse à la probabilité des decisions rendued à la pluralité de voix*. Paris: Imprimerie royale; 1784. 304 p.

3. Moulin H. *Kooperativnoe prinyatie reshenii: aksiomy i modeli* [Cooperative decision making: axioms and models]. Moscow: Mir; 1991. 464 p. Russian.

4. Nureev RM. *Teoriya obshchestvennogo vybora* [The theory of collective choise]. Moscow: Publishing House of the Higher School of Economics; 2005. 531 p. Russian.

5. Arrow KD. *Kollektivnyi vybor i individual'nye tsennosti* [Collective choise end individual values]. Moscow: Publishing House of the Higher School of Economics; 2004. 204 p. Russian.

6. Sen A. *Ideya spravedlivosti* [The idea of fairness]. Moscow: Publishing House of the Gaidar Institute for Economic Policy; 2016. 520 p. Russian.

7. Mazziotta M, Pareto A. Methods for constructing composite indices: one for all or all for one. *Rivista italiana di Economia Demagrafica e Statistica*. 2013;LXVII(2):67–80.

8. Greco S, Isshizako A, Tasion M, Torrisi G. On the methodological framework of composite indices: a review of the issues of weighting, aggregation and robustness. *Social Indicators Research*. 2019;141:61–94.

9. Gorbach AV, Kovalev MM. [How are international ratings of states determined]. *Vestnik Assotsiatsii belorusskikh bankov*. 2000;33:13–19. Russian.

10. Kovalev MM, Novik VV. [Factors influencing economic growth, competitiveness and country rankings]. In: Kovalev MM, editor. *Voprosy vneshnei politiki Respubliki Belarus'* [Foreign policy issues of the Republic of Belarus]. Minsk: Belarusian State University; 2001. p. 68–117. Russian.

11. Kovalev MM, Shibeko IT. [Methods of calculating of bank ratings]. *Bankovskii vestnik*. 1999;8:30–39. Russian.

12. Kovalev MM, Gospodarik EG. *Bankovskaya analitika* [Bank analytics]. Minsk: Publishing House of the Belarusian State University; 2020. 458 p. Russian.

13. Kovalev M, Petrik O. [Information sites rating of Belarusian banks]. *Vestnik Assotsiatsii belorusskikh bankov*. 2003;45:25–31. Russian.

14. Kovalev MM, Shashko AA. [Investment rating of the main cities of Belarus]. *Belorusskii bankovskii byulleten'*. 2003;19:36–44. Russian.

15. Kovalev MM, Shashko AA. [Small business in Belarus: ranking of small business development in the regions]. *Vestnik Assotsiatsii belorusskikh bankov*. 2004;28:16–34. Russian.

16. Kovalev MM, Shashko AA. [Development of the innovative potential of the Republic of Belarus: innovation rating of regions and Minsk]. *Vestnik Assotsiatsii belorusskikh bankov*. 2004;38–39:24–32. Russian.

17. Girlich E, Kovalev MM, Kozulin AV. *Ranking: theorie and praxis*. Magdeburg: Otto-von-Guericke-Universität; 2001. 24 S.

18. Kovalev MM, Gedranovich AB. [International ratings of universities]. *Vyshjejshaja shkola*. 2007;1:12–15. Russian.

19. Kovalev MM, Listopad NI, Mimnyukovich EA. [Site is a mirror of university and marketing politics instrument]. *Vyshjejshaja shkola*. 2011;3:58–60. Russian.

20. Kovalev MM, Gedranovich AB, Minyukovich EA. [Belarusian experience of universities ranking]. In: *Sistema reitinga vuzov: natsional'naya i mirovaya praktika. Materialy Mezhdunarodnogo nauchno-prakticheskogo seminara; 19–20 marta 2010 g.; Alma-Ata, Kazakhstan* [The system of universities rating: national and world practice. Materials of the International scientific and practice seminar; 2010 March 19–20; Alma-Ata, Kazakhstan]. Alma-Ata: Satbayev University; 2010. p. 41–52. Russian.

21. Kovalev MM, Listopad NI, Minyukovich EA. [Webometric universities rating]. *Informatizatsiya obrazovaniya*. 2009;2:63–73. Russian.

22. Kovalev MM, Kurbatskii AN, Listopad NI. [Expert system for analysing tender proposals for computer equipment and software]. *Informatizatsiya obrazovaniya*. 1997;1:68–92. Russian.

23. Fedotov YuV, Khovamov NV. *Metody postroeniya svodnykh otsenok effektivnosti deyatel'nosti slozhnykh proizvodstvennykh system* [Methods for constructing summary assessments of the efficiency of complex production systems]. Saint Petersburg: NII menedzhmenta Sankt-Peterburgskogo universiteta; 2006. 33 p. Russian.

24. Zhgun TV. An algorithm for constructing integral quality indicator of complex systems for a sequence of observations. *Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering*. 2017;6(1):5–25. Russian. DOI: 10.14529/cmse170101.

25. Colovenchik GG, Kovalev MM. *Tsifrovaya ekonomika* [Digital economy]. Minsk: Publishing House of the Belarusian State University; 2019. 395 p. Russian.

26. Charnes A, Cooper W, Rhodes E. Measuring the effectiveness of decision-making units. *European Journal of Operational Research*. 1978;2(6):429–444.

27. Bolotashvili G, Kovalev M, Girlich E. New facets of the linear ordering polytope. *SIAM Journal on Discrete Mathematics*. 1999;12(3):326–336.

28. Golovenchik GG. *Tsifrovizatsiya belorusskoi ekonomiki v sovremennykh usloviyakh globalizatsii* [Digitalisation of Belarusian economics in the contemporary conditions of globalisation]. Minsk: Publishing House of the Belarusian State University; 2019. 257 p. Russian.

29. Buzdalin AV, Zanochkin AYu, Kurbangaleev MZ, Smirnov SN. Aggregation of credit ratings as a task of building consensus in the system of expert assessments. *Global Markets and Financial Engineering*. 2017;4(3):181–207. Russian. DOI: 10.18334/grfi.4.3.38830.

Статья поступила в редколлегию 15.10.2023.
Received by editorial board 15.10.2023.