



### Information about the authors

**V. Chernov** – D.Sc. (Economics), Professor of Department of Computer Science and Control in Technical and Economic Systems of Vladimir State University (87 Gorkogo Str., 600000, Vladimir, Russia, e-mail: vladimir.chernov44@mail.ru).

**O. Dorokhov** – Ph.D. (Engineering), Associate Professor of Department of Information Systems of Simon Kuznets Kharkiv National University of Economics (9a Nauky Ave., 61166, Kharkiv, Ukraine, e-mail: aleks.dorokhov@meta.ua).

*Стаття надійшла до ред.  
01.09.2016 р.*

JEL Classification: C15; R12

УДК 621.382:519.254

## ПОВЫШЕНИЕ УСТОЙЧИВОСТИ РЕШЕНИЯ ЗАДАЧ КЛАССИФИКАЦИИ МЕТОДАМИ КЛАСТЕРНОГО АНАЛИЗА С КОРРЕКТНЫМ НОРМИРОВАНИЕМ ДАННЫХ

**Ершов И. А.  
Стукач О.В.**

**Аннотация.** На примере решения одной задачи классификации исследована проблема нормирования данных в кластерном анализе. Показано, что некорректный выбор метрики и нормирования переменных и наблюдений приводит к неверной интерпретации результатов. Предложен вариант использования сочетания метода *k*-средних и построения дендрограмм в программном пакете Statistica для повышения устойчивости результата. Этот подход позволяет добиться устойчивости результата и его правильной интерпретации за счёт полного сохранения статистических характеристик кластеризуемых данных. Приведены уточнённые данные по кластеризации культур биоэнергетического потенциала в областях Украины

**Ключевые слова:** кластеризация в экономике, классификация областей, расстояние, урожайность, собираемость.

## ПІДВИЩЕННЯ СТІЙКОСТІ РІШЕННЯ ЗАДАЧ КЛАСИФІКАЦІЇ МЕТОДАМИ КЛАСТЕРНОГО АНАЛІЗУ З КОРЕКТНИМ НОРМУВАННЯМ ДАНИХ

**Ершов І. А.  
Стукач О. В.**

**Анотація.** На прикладі рішення однієї задачі класифікації досліджено проблему нормування даних в кластерному аналізі. Показано, що некоректний вибір метрики і нормування змінних і спостережень призводить до невірної інтерпретації результатів. Запропоновано варіант використання поєднання методу *k*-середніх і побудови дендрограм у програмному пакеті Statistica для підвищення стійкості результату. Цей підхід дозволяє домогтися стійкості результату та його правильної інтерпретації за рахунок повного збереження статистичних характеристик кластеризованих даних. Наведено уточнені дані по кластеризації культур біоенергетичного потенціалу в областях України.

**Ключові слова:** кластеризація в економіці, класифікація областей, відстань, врожайність, збируваність.

# INCREASING THE STABILITY OF SOLVING CLASSIFICATION TASKS BY MEANS OF CLUSTER ANALYSIS WITH CORRECT DATA NORMALIZATION

I. Ershov  
O. Stukach

**Abstract.** *Using the example of solving an individual classification task, the problem of data normalization in cluster analysis is studied. It is shown that an incorrect choice of metrics and normalization of variables and cases leads to incorrect interpretation of results. To increase the stability of final outcomes, an approach implying the combination of the k-means method and drawing dendrograms in the software package «Statistica» is proposed. The approach allows achieving stability of the result and its correct interpretation by means of full representation of statistical characteristics for the clustered data. The specified data on the bio-energy potential of clustered plants in regions of Ukraine are shown.*

**Keywords:** *clustering in economics, classification of regions, distance, yield, crop retrieval.*

**Введение.** В последнее время в связи с развитием новых направлений в анализе больших объемов данных возрос интерес к вопросам формирования корректного информационного базиса в исследовании экономического роста [1] и, в частности, к кластерному анализу. Данный интерес также подтверждается созданием инновационных территориальных кластеров, позволяющих создать сильный импульс в развитии страны [2]. В статье [3] представлен пример анализа статистических данных провинций и городов Китая, которые показали, что именно статистический анализ способен решить проблему низкого уровня конкурентоспособности высокотехнологичных отраслей промышленности. Достигается данная цель путем использования региональных преимуществ для разработки ведущих высокотехнологичных отраслей промышленности, стратегии специализации и кластеризации.

Достаточно неплохой обзор методов и задач кластерного анализа приведен в работах [4; 5]. Несмотря на множество алгоритмов и методов кластеризации, многие исследователи их используют без обоснования, а получаемые результаты носят случайный характер [5]. Один из характерных примеров такого рода рассмотрен в статье [6]. Проблема кластерного анализа, а в данном случае метода  $k$ -средних, в том, что он собирает группы относительно центра кластеризации, используя какой-либо метод нахождения наименьшего расстояния. А это означает, что будут создаваться некие области в пространстве переменных, внутри которых будут элементы данной группы. Несложно догадаться, что схожие данные не всегда располагаются в этих областях, но алгоритмы кластерного анализа этого не учитывают. Для оценки правильности результата должна использоваться априорная информация о природе исследуемых переменных, что зачастую кардинально изменяет результат. Это связано с отсутствием корректного определения расстояния между наблюдениями для каждой конкретной задачи. В качестве расстояния используют евклидово, а такой выбор зачастую объясняется лишь его наибольшей распространенностью.

В [6] предлагается учитывать дисперсию каждого фактора, которую он объясняет в исходном материале. Для этого требуется «каждый фактор умножить на корень квадратный из соответствующего ему собственного числа». Данное преобразование поможет избежать ошибки, связанной с неправильной группировкой, но недостатком данного подхода является очевидная необходимость дополнительных расчетов собственных значений, что в условиях больших данных усложняет получение результата и все равно не доказывает его правильность.

Логичным развитием этих рассуждений может служить работа [7], где предлагается сконструировать для каждого объекта локальную метрику, которая будет обеспечивать максимальную сферу действия. В этом случае результат должен гарантировать корректный выбор метрики и его правильная интерпретация. Однако в данном случае можно столкнуться с несовпадением метрик сравниваемых объектов. Это хорошо показано при проведении статистической обработки макроэкономических данных в статье [8]. Ведь близость значений показателей в различных метриках не означает их фактическую близость, и об этом необходимо помнить [9].

В данном исследовании предлагается сравнительный и уточняющий подход к оценке результатов решения задач классификации, основанный не на выборе метрик и обеспечении точной близости объектов в формируемых кластерах, а на сопоставлении результатов, получаемых с помощью разных способов классификации, выбора числа кластеров и оптимального формирования устойчивого решения по критерию повторяемости результатов анализа.

**Кластерный анализ и нормирование данных.** Данный подход заключается в том, что переменные одного типа группируются и рассматриваются не по отдельности, а как группа. Это уменьшает количество рассматриваемых элементов, что крайне упрощает анализ данных, особенно при больших выборках с несколькими тысячами наблюдений.

В этом подходе принципиально важна нормировка данных [8; 10], и она может не проводиться в единственном случае – одинаковости масштаба и единиц измерения переменных. Примером таких данных могут служить работы [11; 12].

Существует множество различных способов нормирования, среди которых наиболее распространены следующие:

$$z_1 = \frac{x - \bar{x}}{\sigma}, \quad (1)$$

$$z_2 = \frac{x}{x}, \quad (2)$$

$$z_3 = \frac{x}{x_{\max}}, \quad (3)$$

$$z_4 = \frac{x - \bar{x}}{x_{\max} - x_{\min}}, \quad (4)$$

где  $x$  – случайная величина;

$z$  – ее нормированное значение;

$\bar{x}$  – среднее арифметическое значение;

$\sigma$  – среднее квадратическое отклонение;

$x_{\max}$ ,  $x_{\min}$  – максимальное и минимальное значения случайной величины соответственно.

При выборе формулы нормировки важно руководствоваться практическим смыслом приведения наблюдений к сопоставимым значениям с сохранением статистических характеристик разброса. Более того, каждая переменная может иметь свои единицы измерения, и без нормирования кластерный анализ в общем случае приведет к неправильным результатам.

Формула (1) приводит переменные к нормированным с нулевым средним значением и единичной дисперсией. Среднее квадратическое отклонение как основной показатель рассеивания значений делает эту формулу наиболее часто используемой. В статье [13] отмечается, что при использовании данной формулы возникают сложности в интерпретации результатов, связанные с отрицательными значениями элементов. Но при формировании выводов нет никаких препятствий вернуться к ненормированным данным.

Формулы (2) и (3) не учитывают статистический разброс наблюдений, только лишь изменяя масштаб значений наблюдений. Как показывает практика, нормировка по этим формулам в кластерном анализе большого объема разнородных данных бесполезна. В формуле (4) используется размах вместо среднее квадратического отклонения. Результат кластеризации практически совпадает с тем, который получается при использовании соотношения (1).

**Пример кластерного анализа с оценкой устойчивости решения.** Для анализа был использован полный набор данных из статьи [14]. Переменные имеют сильно различающиеся средние значения, единицы измерения и смысл (собираемость, урожайность, цена реализации), поэтому их необходимо нормировать. Исходя из выводов, сделанных ранее, проведем нормирование по формуле (1). Это не было сделано авторами статьи [14], поэтому результат нельзя считать достоверным.

Вначале проведем иерархическую кластеризацию данных методом одиночной связи и построим дендрограмму, в качестве метрики используя евклидово расстояние. Это позволяет получить компактные и хорошо сгруппированные данные. Результаты получены в программном пакете Statistica 8.0 (рис. 1).

После построения дендрограммы требуется разделить данные на группы (кластеры). Это делается выбором порогового расстояния. Пороговое расстояние – значение, меньше которого вариация значений наблюдений не существенна для поставленной задачи.

На рис. 1 обозначены два пороговых расстояния (2,5 и 3,5), поэтому мы можем разделить все данные на три либо на четыре кластера. Малое число кластеров, как правило, не имеет практического смысла, а их большое число сложно интерпретировать. В нашем случае правильнее выделить четыре кластера, так как первые три существенно различаются. Рассматриваемые данные могут создать иллюзию, что в некоторых кластерах обязательно должно быть по одному элементу, но это не так, потому что количество элементов в кластерах зависит исключительно от выборки.

После определения количества кластеров и их состава найдем средние значения нормированных переменных, входящих в соответствующий кластер, и построим график (рис. 2).

По этому графику можно сделать следующие выводы.

1. Все анализируемые переменные хорошо кластеризуются в смысле разности расстояний между ними в различных кластерах, что свидетельствует о значимости результата.

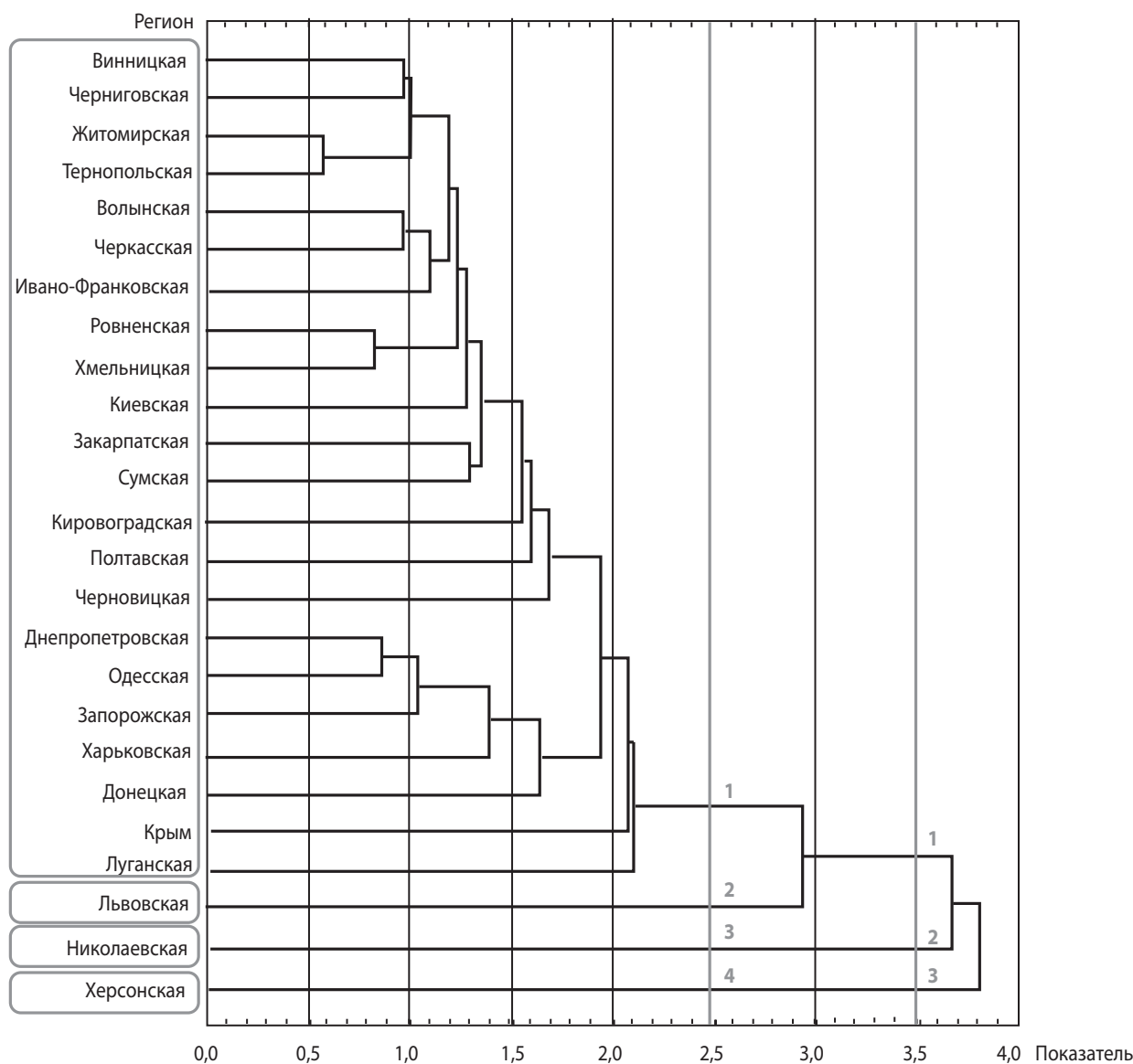


Рис. 1. Дендрограмма данных с выделенными группами

2. Самая лучшая область для выращивания рапса и кукурузы – Львовская. Она единственная формирует кластер 2, на что указывает также дендрограмма (рис. 1).

3. В Львовской области хороший сбыт рапса, поскольку там он имеет наибольшую цену реализации.

4. В кластерах 3 (Николаевская область) и 4 (Херсонская область) самая высокая цена реализации кукурузы.

5. Все остальные области Украины формируют единый кластер (1). Экономические показатели в этом кластере близки в смысле расстояния между ними. Поэтому для улучшения экономических показателей вопрос о том, какие конкретно культуры выращивать в областях кластера 1 и какие продавать, не играет роли. И наоборот, области кластеров 2–4 имеют важное значение, и случайный выбор выращиваемых культур никак не будет способствовать экономическому росту.

**Оценка результата.** Для того чтобы быть уверенным в достоверности результатов, требуется провести проверку другим методом кластеризации. В качестве него использован метод  $k$ -средних. В этом методе субъективен выбор количества кластеров, что может повлиять на интерпретацию результатов.

В нашем случае, чтобы определить количество кластеров, можно разделить данные на две части (кукуруза и рапс), а затем построить 3D-графики, чтобы на основании их сказать, какое количество кластеров имеет смысл выбрать (рис. 3, 4).

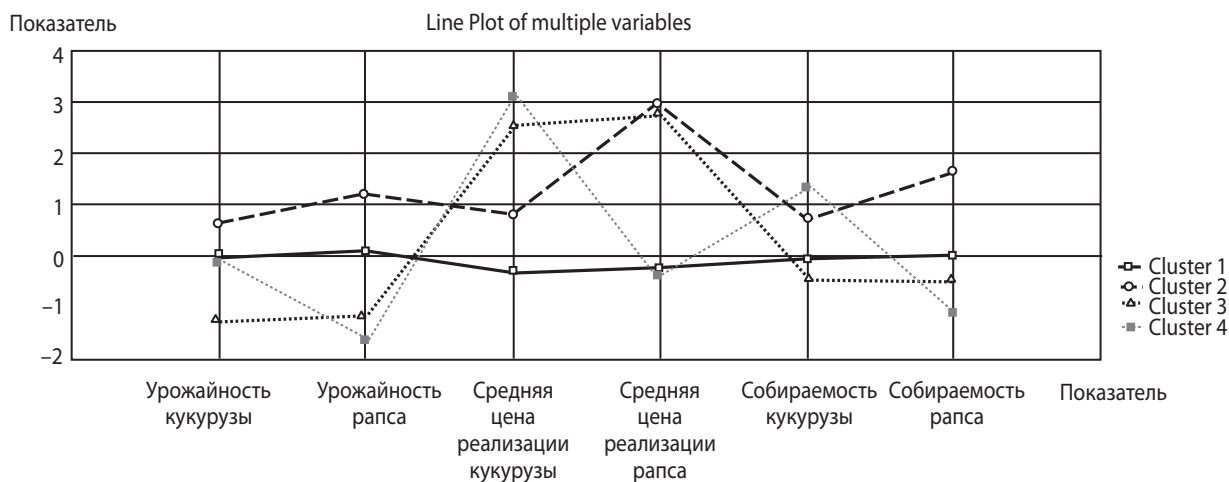


Рис. 2. График средних значений нормированных переменных

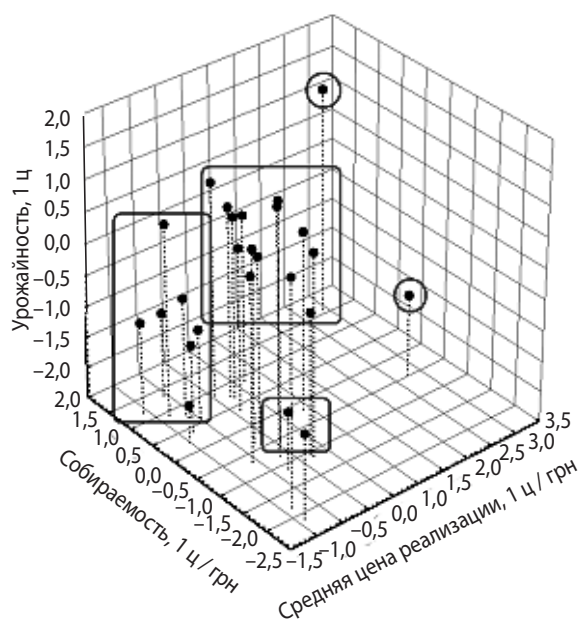


Рис. 3. Разделение областей на кластеры (рапс)

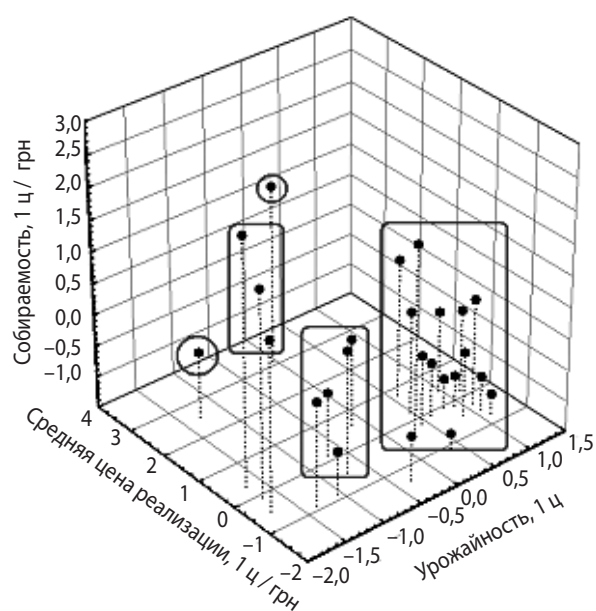


Рис. 4. Разделение областей на кластеры (кукуруза)

Было получено пять групп нормированных значений наблюдений, поэтому данные можно кластеризовать на пять кластеров. Это совпадает с предположением, следующим из рис. 1, если выбрать пороговое расстояние 1,75. В этом случае самый большой кластер будет поделен на два, что наглядно проявляется на рис. 3, 4.

На рис. 5 и рис. 6 приведены результаты кластерного анализа методом  $k$ -средних для нормированных данных при выбранном числе кластеров для кукурузы и рапса.

Наиболее хорошо показал себя кластер 4, который имеет как высокую урожайность, так и высокую собираемость (рис. 5). В кластере под номером 3 явное превосходство в средней цене реализации, но вызвано это низкой урожайностью, так что выращивать биокультуры в областях этого кластера не имеет смысла. Однако имеет смысл подумать о продажах в этих регионах, потому что, возможно, доходы с продаж окупят расходы на перевозку этого товара.

В итоге можно сказать, что приоритетные области для выращивания кукурузы находятся в кластере 4, а именно: Ивано-Франковская, Львовская, Полтавская, Черновицкая. Наиболее приоритетные области для реализации кукурузы находятся в кластере 3: Николаевская, Херсонская.

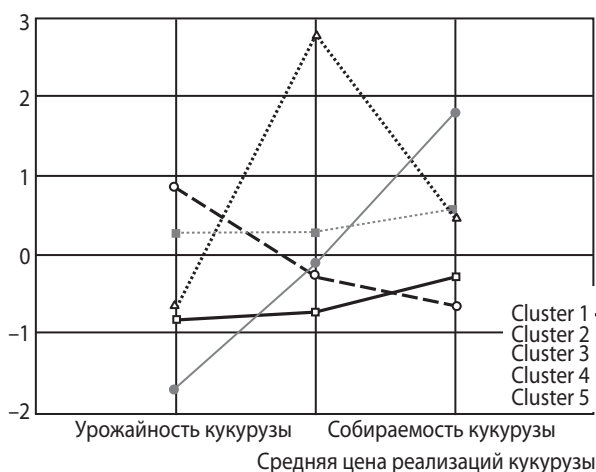


Рис. 5. Средние значения кластеров (кукуруза)

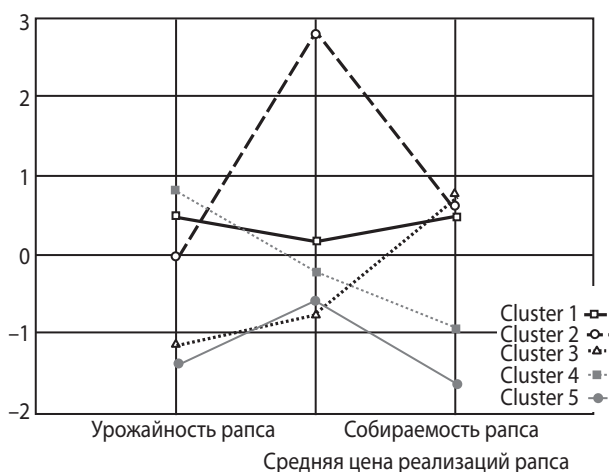


Рис. 6. Средние значения кластеров (рапс)

Наиболее приоритетным по выращиванию и реализации рапса (рис. 6) является кластер 3 (Львовская область). На дендрограмме это также отдельный кластер.

Результаты, полученные методом построения дендрограммы, можно считать значимыми, так как они почти полностью совпадают с результатами, полученными методом  $k$ -средних.

**Применение разных видов нормировки.** Как говорилось ранее, применение различных видов нормирования приводит к отличающимся результатам. Используем для нормирования данных формулы (2) и (3), а затем проведем аналогичные операции для данных из [14]. Дендрограммы показаны на рис. 7 и рис. 8.

Статистический разброс значений при таком нормировании не учитывается, что наглядно можно видеть из дендрограмм. Из-за того, что данные методы не учитывают среднее квадратическое отклонение показателей, область с наибольшей ценой реализации выталкивает все остальные области на задний план. В данном случае это область Херсонская. В итоге возникает неопределенность в делении на кластеры. Очевидно, что пороговое расстояние разделит набор наблюдений на две группы: Херсонская область и все остальные. Однако из предыдущих результатов нам известно, что данная область выделяется только по средней цене реализации. Соответственно, последующий анализ приведет к неправильным результатам. Нужно отметить, что если бы мы использовали данную нормировку сразу или не использовали ее вообще, то получили бы неверный результат.

В случае использования формулы (4) мы получаем аналогичные с формулой (1) результаты (рис. 9).

Деление наблюдений на кластеры не изменилось, но для формирования пяти кластеров пороговое расстояние нужно уменьшить до 0,59.

**Результаты и вывод.** Таким образом, в работе рассмотрена проблема нормировки данных в кластерном анализе с использованием методов иерархической кластеризации и  $k$ -средних. На примере кластеризации набора данных из работы [14] показано, как значительно изменяется результат при использовании разных методик нормирования и интерпретации визуального анализа. Устойчивости результата можно добиться только с использованием сочетания различных методов кластерного анализа и корректного выбора метрики или нормирования. Предложен один из способов обоснования результата кластеризации, основанный на корректном применении нормирования данных и визуализации в виде графиков средних значений или центров кластеров. Предложенный в статье подход может быть развит при определении расчета расстояния для показателей биоэнергетического потенциала.

**Литература:** 1. Серова И. А. Вопросы корректности оценки экономического роста. *Экономика развития*. 2016. № 4 (80). С. 63–70. 2. Glinitsky V., Serga L., Chemezova E., Zaykov K. Clusterization Economy as a Way to Build Sustainable Development of the Region // 13th Global Conference on Sustainable Manufacturing, Binh Du'ong New City, 16–18 Sept. 2015. ISSN: 22128271. DOI: 10.1016/j.procir.2016.01.050. 3. Zhaopeng Chu, Bing Li. A study on the selection of China's regional leading high-tech industries by grey incidence analysis // The 2nd International Conference on Information Science and Engineering, Hangzhou, 4–6 Dec. 2010. ISBN: 978-142448096-8. DOI: 10.1109/ICISE.2010.5691813. 4. Кузнецов Д. Ю., Трошина Т. Л. Кластерный анализ и его применение. *Ярославский педагогический вестник*. 2006. Вып. 4. С. 103–107. 5. Бериков В. С., Лбов Г. С. Современные тенденции в кластерном анализе // Информационно-телекоммуникационные системы: Всеросс. конкурс. отбор обзор-

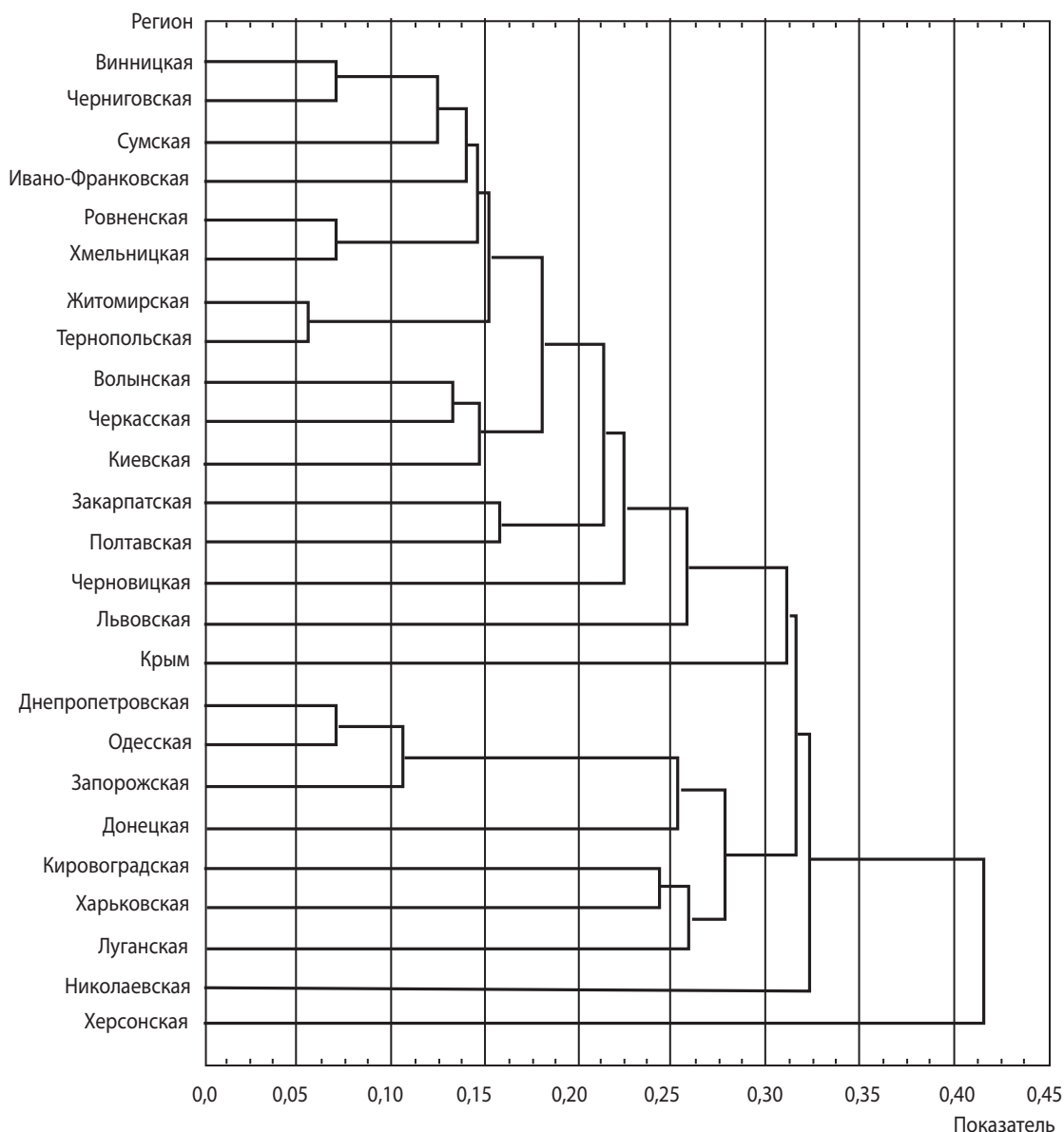


Рис. 7. Дендрограмма при нормировке по формуле (2)

аналит. статей по приоритет. направл. 2008. URL: <http://www.ict.edu.ru/ft/005638/62315e1-st02.pdf>. 6. Галицкая Е. Г. Галицкий Е. Б. Кластеры на факторах: как избежать распространенных ошибок? *Социология: методология, методы, математическое моделирование*. 2006. № 22. С. 145–161. 7. Карпов Л. Е., Юдин В. Н. Методы добычи данных при построении локальной метрики в системах вывода по прецедентам // Препринты Института системного программирования РАН – 2006 г. Препринт 18. С. 1–42. 8. Егорова Ю. В., Золотухина А. В. Проблемы и возможности статистической обработки макроэкономических данных (на примере оценки научно-технического потенциала региона). *Экономика и предпринимательство*. 2016. № 3–1 (68–1). С. 475–479. 9. Луценко Е. В., Коржаков В. Е. Некоторые проблемы классического кластерного анализа. *Вестник Адыгейского государственного университета*. 2011. Сер. 4. № 2. URL: <http://cyberleninka.ru/article/n/nekotorye-problemy-klassicheskogo-klasternogo-analiza>. 10. Вуколов Э. А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL: учеб. пособие. М.: ФОРУМ: ИНФРА-М, 2004. 464 с. 11. Черных А.В., Малеев Ю.В., Шевцов А.Н. Изучение особенностей топографии околощитовидных желез с использованием методов многомерного статистического анализа. *Український журнал клінічної та лабораторної медицини*. 2013. Т. 8. № 3. С. 185–190. 12. Дьячкова О. В. Моделирование формирования доверия потребителей на основании кластерного анализа данных. *Технологический аудит и резервы производства*. 2015. № 6/5 (26). С. 20–24. 13. Саранча М. А. Методологические проблемы интегральной оценки туристско-рекреационного потенциала,

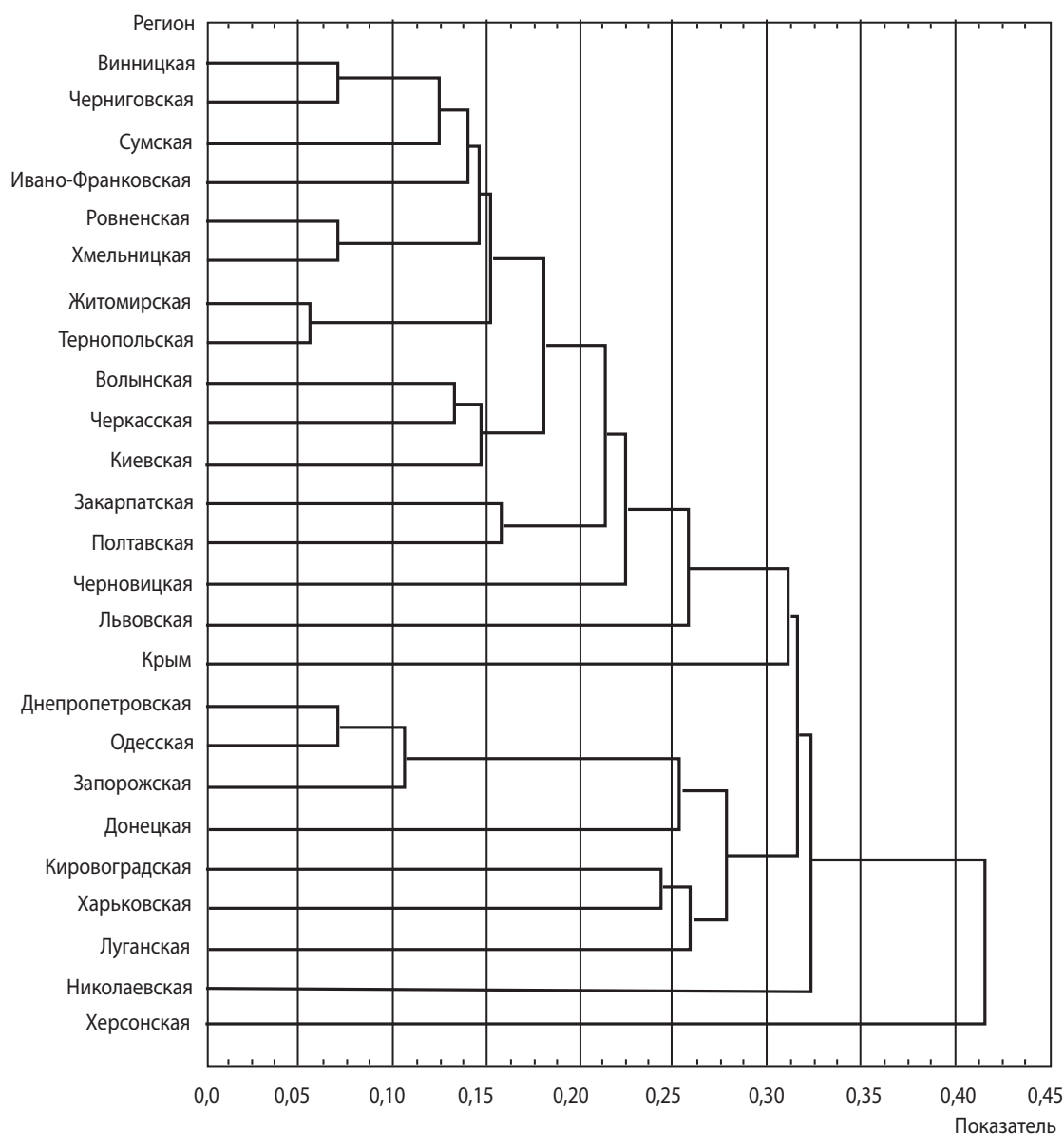


Рис. 8. Дендрограма при нормировке по формуле (3)

территории. *Вестник Удмуртского университета*. 2011. Вып. 1. С. 118–127. 14. Федорченко Б. С. Кластерний аналіз біоенергетичного потенціалу областей України. *Облік і фінанси*. № 2 (64). 2014. С. 173–178.

**References:** 1. Serova I. A. *Voprosy korrektnosti otsenki ekonomicheskogo rosta* [The Issues of Correctness of the Assessment of Economic Growth]. *Ekonomika razvitiya*. 2016. No. 4 (80). P. 63–70. 2. Glinskiy V., Serga L., Chemezova E., Zaykov K. Clusterization Economy as a Way to Build Sustainable Development of the Region // 13th Global Conference on Sustainable Manufacturing, Binh Du'ong New City, 16–18 Sept. 2015. ISSN: 22128271. DOI: 10.1016/j.procir.2016.01.050. 3. Zhaopeng Chu, Bing Li. A study on the selection of China's regional leading high-tech industries by grey incidence analysis // The 2nd International Conference on Information Science and Engineering, Hangzhou, 4–6 Dec. 2010. ISBN: 978-142448096-8. DOI: 10.1109/ICISE.2010.5691813. 4. Kuznetsov D. Yu., Troshina T. L. *Klasternyy analiz i yego primeneniye* [Cluster Analysis and its Application]. *Yaroslavskiy pedagogicheskiy vestnik*. 2006. Issue 4. P. 103–107. 5. Berikov V. S., Lbov G. S. *Sovremennyye tendentsii v klasternom analize* // Informatsionno-telekommunikatsionnyye sistemy: Vseross. konkurs. otbor obzor-analit. statey po prioritet. napravl. 2008. URL: <http://www.ict.edu.ru/ft/005638/62315e1-st02.pdf>. 6. Galitskaya E. G., Galitskiy E. B. *Klastery na faktorakh: kak izbezhat rasprostranennykh oshibok?* [Clusters on Factors: How to Avoid Common Mistakes?] *Sotsiologiya: metodologiya, metody*,

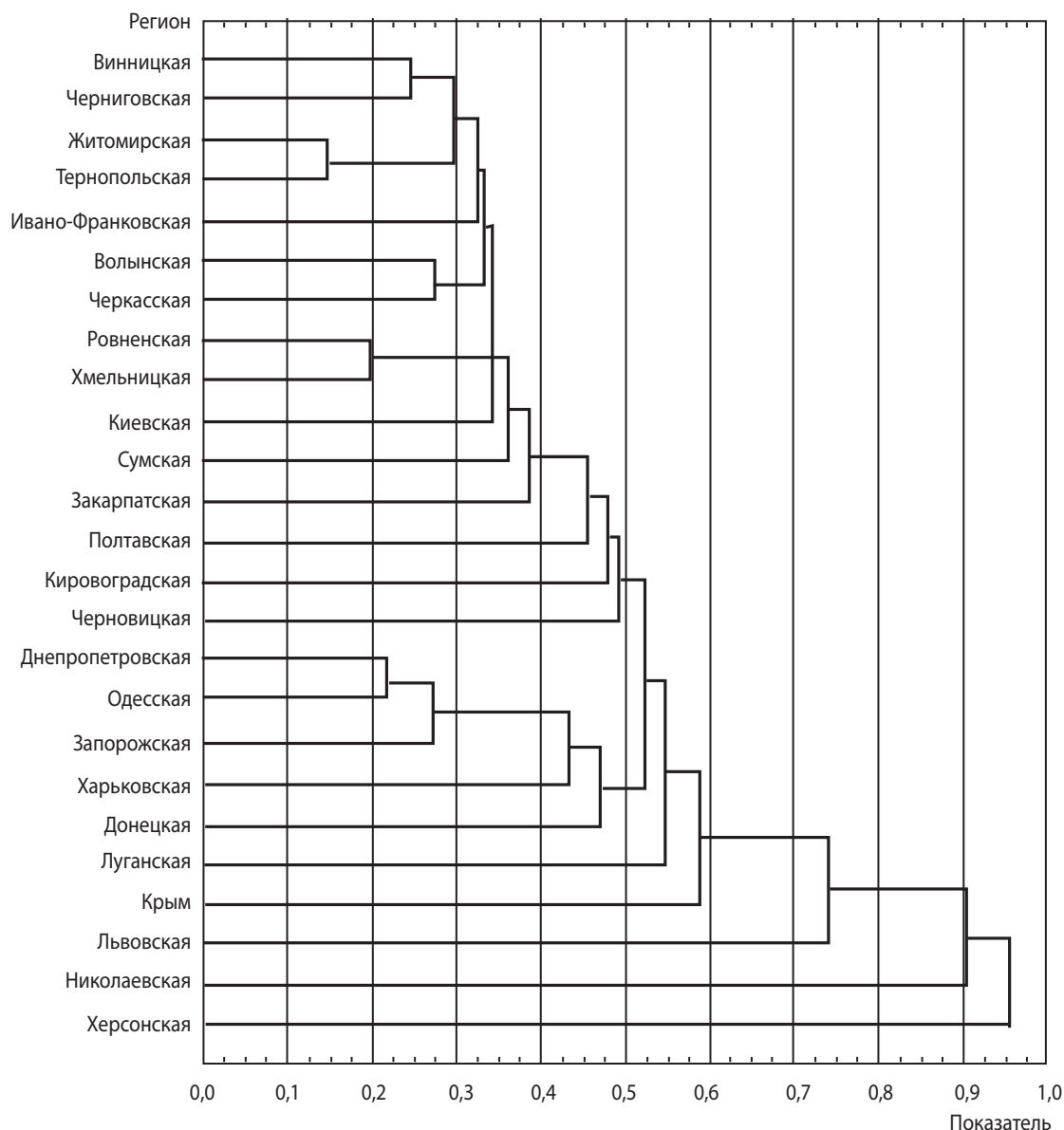


Рис. 9. Дендрограмма при нормировке по формуле (4)

*matematicheskoye modelirovaniye*. 2006. No. 22. P. 145–161. **7.** Karpov L. E., Yudin V. N. Metody dobychi dannykh pri postroyenii lokalnoy metriki v sistemakh vyvoda po pretседentam // Preprinty Instituta sistemnogo programmirovaniya RAN – 2006 g. Preprint 18. P. 1–42. **8.** Yegorova Yu. V., Zolotukhina A. V. *Problemy i vozmozhnosti statisticheskoy obrabotki makroekonomicheskikh dannykh (na primere otsenki nauchno-tekhnicheskogo potentsiala regiona)* [Challenges and Opportunities for Statistical Treatment of Macroeconomic Data (for Example, the Assessment of Scientific and Technical Potential of the Region)]. *Ekonomika i predprinimatelstvo*. 2016. No. 3–1 (68–1). P. 475–479. **9.** Lutsenko E. V., Korzhakov V. E. *Nekotoryye problemy klassicheskogo klasternogo analiza* [Some Problems of the Classical Cluster Analysis]. *Vestnik Adygeyskogo gosudarstvennogo universiteta*. 2011. Ser. 4. No. 2. URL: <http://cyberleninka.ru/article/n/nekotoryye-problemy-klassicheskogo-klasternogo-analiza>. **10.** Vukolov Ye. A. *Osnovy statisticheskogo analiza. Praktikum po statisticheskim metodam i issledovaniyu operatsiy s ispolzovaniyem paketov STATISTICA i EXCEL: ucheb. posobiye*. M.: FORUM; INFRA-M, 2004. 464 p. **11.** Chernykh A. V., Maleev Yu. V., Shevtsov A. N. *Izucheniye osobennostey topografii okoloshchitovidnykh zhelez s ispolzovaniyem metodov mnogomernogo statisticheskogo analiza* [The Study of the Topography of the Parathyroid Glands with the Use of Multivariate Statistical Analysis]. *Ukrainskii zhurnal klinichnoi ta laboratornoi medytyny*. 2013. Vol. 8. No. 3. P. 185–190. **12.** Dachkova O. V. *Modelirovaniye formirovaniya doveriya potrebitel'ey na osnovanii klasternogo analiza dannykh* [Modeling the Formation of Consumer Trust on the Basis of Cluster Analysis of the Data]. *Tekhnologicheskii audit i rezervy proizvodstva*. 2015. No. 6/5 (26). P. 20–24. **13.** Sarancha M. A. *Metodologicheskiye prob-*

*lemy integralnoy otsenki turistsko-rekreatsionnogo potentsiala territorii* [Methodological Problems of Integrated Assessment of Tourist-Recreational Potential of the Territory]. *Vestnik Udmurtskogo universiteta*. 2011. Issue 1. P. 118–127. **14.** Fedorchenko B. S. *Klasternyi analiz bioenerhetychnoho potentsialu oblastei Ukrainy* [Cluster Analysis of Bioenergy Potential of Regions of Ukraine]. *Oblik i finansy*. No. 2 (64). 2014. P. 173–178.

#### Информация об авторах

**Стукач Олег Владимирович** – доктор технических наук, профессор кафедры систем управления и мехатроники Национального исследовательского Томского политехнического университета (просп. Ленина, 30, г. Томск, 634050, Россия; e-mail: tomsk@ieee.org).

**Ершов Иван Анатольевич** – магистрант кафедры систем управления и мехатроники Национального исследовательского Томского политехнического университета (просп. Ленина, 30, г. Томск, 634050, Россия; e-mail: tomsk@ieee.org).

#### Інформація про авторів

**Стукач Олег Володимирович** – доктор технічних наук, професор кафедри систем управління і мехатроніки Національного дослідницького Томського політехнічного університету (просп. Леніна, 30, м. Томськ, 634050, Росія; e-mail: tomsk@ieee.org).

**Ершов Іван Анатолійович** – магістрант кафедри систем управління і мехатроніки Національного дослідницького Томського політехнічного університету (просп. Леніна, 30, м. Томськ, 634050, Росія; e-mail: tomsk@ieee.org).

#### Information about the authors

**O. Stukach** — D.Sc (Engineering), Professor of Department of Control Systems and Mechatronics of National Research Tomsk Polytechnic University (30 Lenina Ave., Tomsk, 634050, Russia; e-mail: tomsk@ieee.org).

**I. Ershov** — Master Student of Department of Control Systems and Mechatronics of National Research Tomsk Polytechnic University (30 Lenina Ave., Tomsk, 634050, Russia; e-mail: tomsk@ieee.org).

*Стаття надійшла до ред.  
30.08.2016 р.*

JEL Classification: C32; L60

УДК 338.2

## ІМІТАЦІЙНЕ МОДЕЛЮВАННЯ КОНКУРЕНТОСПРОМОЖНОСТІ МЕТАЛУРГІЙНОЇ ГАЛУЗІ УКРАЇНИ

**Ялдин І. В.**

**Анотація.** Доведено доцільність застосування імітаційної моделі для моделювання конкурентоспроможності металургійної галузі. Імітаційна модель конкурентоспроможності металургійної галузі складається з блоків: продукція, виробництво, наукоємність, інвестиції, фінанси, концентрація, інновації. Сценарне моделювання розвитку металургійної галузі включає розгляд сценаріїв: базового, песимістичного, оптимістичного, збільшення експорту високотехнологічної продукції. Побудовано імітаційну модель, що дозволила дослідити динаміку складових конкурентоспроможності за різних умов і виявити можливості для її підвищення. Пріоритетним сценарієм для зростання конкурентоспроможності металургійної галузі України було обрано сценарій збільшення експорту високотехнологічної продукції.

**Ключові слова:** імітаційне моделювання, конкурентоспроможність, металургійна галузь, сценарії.