

УДК 004.81
DOI 10.17513/snt.39627

АЛГОРИТМ АНАЛИЗА ДАННЫХ ОБ ИНВЕСТИЦИОННЫХ ПРОЦЕССАХ В РЕГИОНАХ РФ С ИСПОЛЬЗОВАНИЕМ МЕТОДА ГЛАВНЫХ КОМПОНЕНТ

Ильясов Б.Г., Макарова Е.А., Закиева Е.Ш., Бойцов А.Н.

ФГБОУ ВО «Уфимский университет науки и технологий», Уфа, e-mail: ea-makarova@mail.ru

В статье рассматриваются вопросы разработки алгоритма анализа данных о состоянии инвестиционных процессов в регионах РФ. Предложенный алгоритм основан на применении метода главных компонент. Используются статистические данные о социально-экономическом положении регионов РФ. Сформированы три выборки данных, характеризующих: инвестиционный процесс в регионе с учетом состояния основных фондов, структуру инвестиций в основной капитал по видам основных фондов, а также динамику инвестиционного процесса и его результативность в строительстве. Алгоритм содержит пять шагов: формирование выборок данных; анализ сформированных выборок с помощью метода главных компонент; формирование интегральной выборки и проведение компонентного анализа; корректировка интегральной выборки и заключительное построение главных компонент с целью выявления итоговой кластерной структуры регионов; расчет рейтинга регионов по уровню инвестиционного развития. Построены диаграммы рассеяния для всех анализируемых выборок, выявлены характеристики регионов по выделенным главным компонентам. Построена итоговая диаграмма рассеяния, сформированы кластеры регионов, определены их характеристики. Выполнен расчет рейтинга регионов по уровню инвестиционного развития как базы экономического развития регионов в целом. Предложенный алгоритм целесообразно применять при разработке информационных систем поддержки принятия решений в области инвестиционной политики.

Ключевые слова: инвестиционный процесс, регион, анализ данных, метод главных компонент, кластер, диаграмма рассеяния

ALGORITHM FOR DATA ANALYSIS ON INVESTMENT PROCESSES IN THE REGIONS OF THE RUSSIAN FEDERATION USING THE PRINCIPAL COMPONENT METHOD

Pyasov B.G., Makarova E.A., Zakieva E.Sh., Boytsov A.N.

Ufa University of Science and Technology, Ufa, e-mail: ea-makarova@mail.ru

The article discusses the development of an algorithm for analyzing data on the state of investment processes in the regions of the Russian Federation. The proposed algorithm is based on the application of the principal component method. Statistical data on the socio-economic situation of the regions of the Russian Federation are used. Three data samples were formed that characterize: the investment process in the region, taking into account the state of fixed assets, the structure of investments in fixed assets by types of fixed assets, as well as the dynamics of the investment process and its effectiveness in construction. The algorithm contains four steps: the formation of data samples; analysis of the generated samples using the principal component method; the formation of an integral sample and component analysis; correction of the integral sample and the final construction of the principal components in order to identify the final cluster structure of regions; calculation of the rating of regions by the level of investment development. Scattering diagrams are constructed for all analyzed samples, the characteristics of the regions for the selected main components are revealed. The final scattering diagram is constructed, clusters of regions are formed, their characteristics are determined. The calculation of the rating of regions according to the level of investment development as the basis for the economic development of the regions as a whole is carried out. The proposed algorithm should be used in the development of information systems to support decision-making in the field of investment policy.

Keywords: investment process, region, data analysis, principal component method, cluster, scattering diagram

В современных условиях развития экономики РФ одной из наиболее актуальных задач является совершенствование процедур поддержки принятия решений при управлении инвестиционными процессами на региональном уровне. Существующие проблемы регионального экономического развития имеют тенденцию к стремительному воздействию на ход инвестиционных процессов [1–3]. Среди таких проблем можно выделить: замедление темпов экономического роста регионов, сокращение промышленного производства, неполное исполнение бюджетных средств, выделяемых на реализацию национальных

проектов, нестабильность инвестиций в основной капитал [4–6].

Решение проблем управления инвестиционными процессами на региональном уровне требует разработки систем поддержки принятия решений, которые основаны на использовании результатов анализа широкой группы показателей, характеризующих различные стороны деятельности регионов и оказывающих влияние на экономический рост и благосостояние населения. Среди этих показателей должны быть выделены не только показатели, которые отражают особенности непосредственно инвестиционной деятельности в регионах,

но и такие показатели, на основе которых можно оценить уровень как социального развития, так и экономического развития в целом, которые отражают результативность инвестиционных вложений с учетом инвестиционного лага. Возникает необходимость в разработке алгоритма анализа данных об инвестиционных процессах на региональном уровне, который позволяет интегрировать показатели и строить обобщенные характеристики состояния инвестиционных процессов регионов, а также формировать кластерную структуру регионов.

Цель исследования заключается в разработке алгоритма анализа данных об инвестиционных процессах в регионах и составлении рейтинга инвестиционного развития регионов на основе применения метода главных компонент, что необходимо для поддержки принятия решений при реализации различных сценариев инвестиционной политики.

Материалы и методы исследования

Предлагаемый алгоритм анализа данных об инвестиционных процессах и составлении рейтинга инвестиционного развития регионов включает пять шагов.

На первом шаге формируются выборки на основе данных, представленных в статистических сборниках «Регионы России» [7]. Сформированы три выборки А1, А2 и А3, включающие в себя в целом 20 показателей.

В выборку А1 включены признаки, характеризующие инвестиционный процесс в регионе с учетом состояния основных фондов. Это признаки: инвестиции в основной капитал; инвестиции в основной капитал на душу населения; иностранные инвестиции в регион (сальдо); стоимость основных фондов; степень износа основных фондов; валовый региональный продукт (ВРП); ВРП на душу населения; число предприятий и организаций.

В выборке А2 присутствуют признаки, характеризующие структуру инвестиций в основной капитал по видам основных фондов: инвестиции в жилые здания и помещения; инвестиции в здания (кроме жилых), инвестиции на улучшение земель; инвестиции в машины, оборудования, транспортные средства; инвестиции в объекты интеллектуальной собственности, а также сальдированный финансовый результат организаций.

Выборка А3 включает признаки, характеризующие динамику инвестиционного процесса и его результативность в строительстве: индекс физического объема инвестиций, индекс физического объема валового регионального продукта, индекс промышленного производства, индекс производства продукции сельского хозяйства, количество введенных в эксплуатацию жилых и нежилых зданий, объем введенных в эксплуатацию жилых и нежилых зданий, площадь введенных в эксплуатацию жилых и нежилых зданий.

На втором шаге проводится анализ сформированных выборок методом главных компонент. Цель анализа данных на основе выборки А1 состоит в формировании характеристик регионов, различающихся по уровню инвестиционного развития, при этом учитывается текущее состояние основных фондов и уровень достигнутого экономического развития в целом. Построены главные компоненты (ГК), весовые коэффициенты которых представлены в табл. 1, а также сформированы названия главных компонент. Для первой главной компоненты ГК11 сформировано название – «Уровень экономического развития регионов с учетом инвестиций», для второй главной компоненты ГК12 – «Уровень инвестиционного развития с учетом численности населения», для третьей главной компоненты ГК13 – «Степень износа основных фондов».

Таблица 1

Весовые коэффициенты признаков в главных компонентах (выборка А1)

Название признака	ГК11	ГК12	ГК13
Валовой региональный продукт	0,450162	-0,0677153	0,0665183
Валовой региональный продукт на душу населения	0,134584	0,680818	-0,057786
Инвестиции в основной капитал	0,449461	-0,00780874	0,046134
Инвестиции в основной капитал на душу населения	0,112073	0,689901	-0,105016
Иностранные инвестиции в регион (сальдо)	0,419041	-0,092768	-0,0147691
Стоимость основных фондов на конец года	0,443649	-0,0226858	0,088154
Степень износа основных фондов на конец года	-0,095933	0,131661	0,982901
Число предприятий и организаций	0,427834	-0,171527	0,0706795

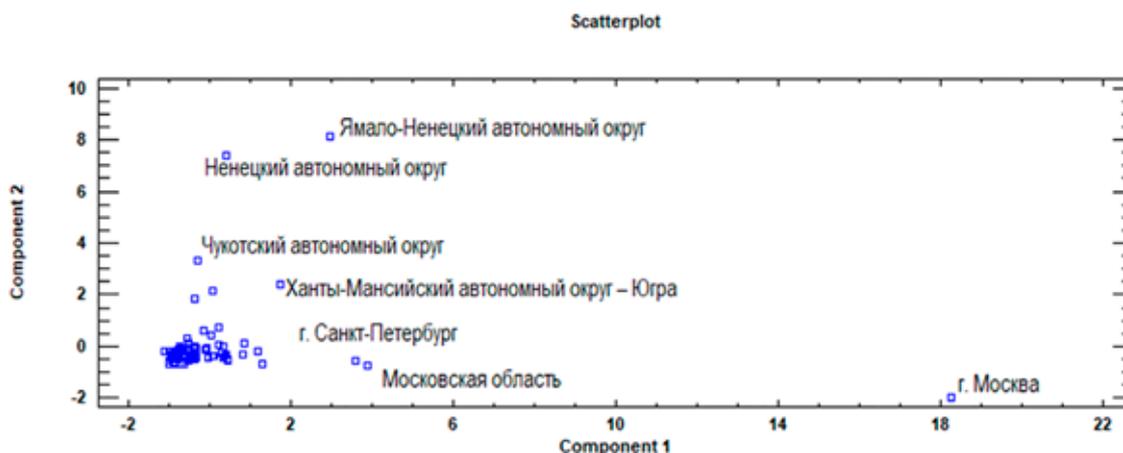


Рис. 1. Распределение регионов в пространстве первых двух ГК (выборка А1)

Таблица 2

Весовые коэффициенты признаков в главных компонентах (выборка А2)

Название признаков	ГК21	ГК22	ГК 23
Сальдированный финансовый результат организаций	-0,182566	0,660114	-0,159523
Инвестиции в жилые здания и помещения	-0,427643	-0,197754	-0,76539
Инвестиции в здания (кроме жилых), расходы на улучшение земель	0,699397	0,0292984	-0,0271111
Инвестиции в машины, оборудование, транспортные средства	-0,534129	-0,156849	0,618574
Инвестиции в объекты интеллектуальной собственности	-0,096653	0,706882	0,0732248

Распределение регионов на диаграмме рассеяния в пространстве двух главных компонент, ГК11 и ГК12, представлено на рис. 1. Показано, что большинство регионов находятся в области довольно малых значений обеих главных компонент, однако это обусловлено значительным отдалением г. Москвы по направлению ГК11 от всех остальных регионов. Необходимо выделить Московскую область и г. Санкт-Петербург, отличающиеся достаточно выраженным преимуществом экономического и инвестиционного развития (по ГК11) относительно других регионов. Нахождение региона в области больших значений главной компоненты ГК12 характеризует высокий уровень инвестиционных вложений на душу населения, а также ВРП – тоже на душу населения. Вследствие учета численности населения лидерами по ГК12 становятся Ямало-Ненецкий автономный округ и Ненецкий автономный округ. Довольно большой уровень инвестиционного и экономического развития с учетом численности населения характерен для Чукотского и Ханты-Мансийского автономных округов. Наибольший износ

основных фондов (наибольшее значение ГК13) присутствует в Удмуртской и Чувашской республиках, Республике Марий Эл, Оренбургской и Рязанской областях.

Анализ данных на основе выборки А2 также проводится с помощью метода главных компонент и имеет целью формирование характеристик регионов, различающихся по структуре инвестиционных вложений в основной капитал. Весовые коэффициенты признаков в главных компонентах приведены в табл. 2.

На основе полученных значений весовых коэффициентов признаков в главных компонентах сформированы названия для первых двух главных компонент. Первая главная компонента ГК21 имеет название «Инвестиции в землю с учетом инвестиций в оборудование и жилые здания», название второй главной компоненты ГК22 – «Инвестиции в интеллектуальную собственность с учетом финансового результата организаций», третьей главной компоненты ГК23 – «Инвестиции в жилые здания». На рис. 2 представлена диаграмма рассеяния, полученная при анализе выборки А2.

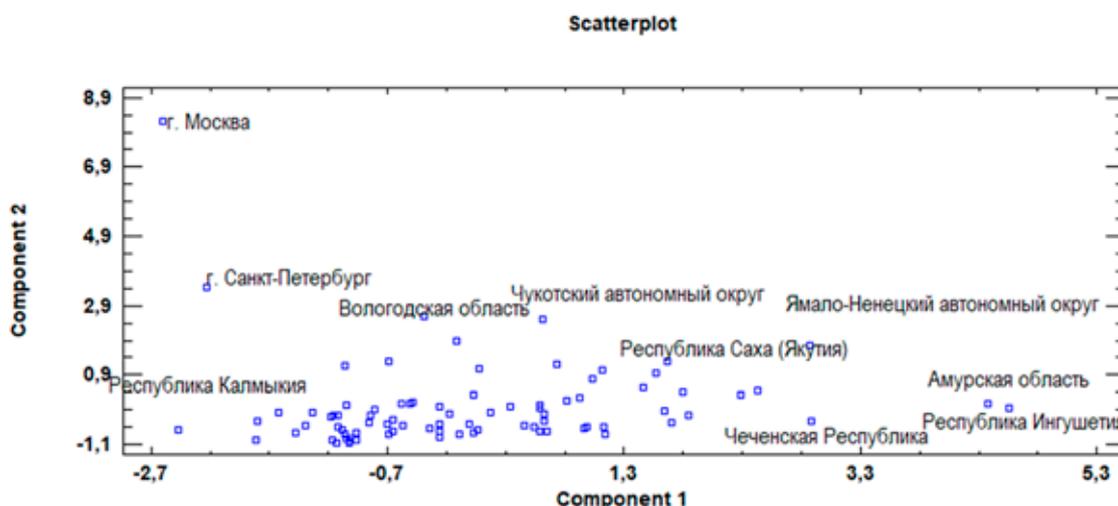


Рис. 2. Распределение регионов в пространстве первых двух ГК (выборка А2)

Таблица 3

Весовые коэффициенты признаков в главных компонентах (выборка А3)

Название признака	ГК31	ГК32	ГК33
Количество введенных в эксплуатацию жилых и нежилых зданий	0,531884	-0,0423312	0,0255631
Объем введенных в эксплуатацию жилых и нежилых зданий	0,595673	0,045948	0,0523732
Площадь введенных в эксплуатацию жилых и нежилых зданий	0,597546	0,0364365	0,0369021
Индекс физического объема валового регионального продукта	-0,0665812	0,698388	0,710349
Индекс физического объема инвестиций в основной капитал	0,0279085	0,712056	-0,700462

Анализ расположения регионов на диаграмме позволяет выявить характерные особенности структуры инвестиционных вложений в основной капитал в регионах РФ. Состав ГК21 позволяет выделить, с одной стороны, регионы с преимущественным вложением инвестиций в оборудование и жилые здания, а также малыми земельными инвестициями (это область малых значений ГК21, где расположены Москва и Санкт-Петербург); и, с другой стороны, регионы, характеризующиеся значительными инвестиционными вложениями в землю и небольшими инвестициями в оборудование (это область больших значений ГК21, где расположены, в частности, Республика Ингушетия и Амурская область). Подключение к анализу распределения регионов вдоль ГК22 позволяет выделить регионы с наибольшими вложениями инвестиций в интеллектуальную собственность (г. Москва, г. Санкт-Петербург, Чукотский автономный округ, Ямало-Ненецкий автономный округ, Вологодская область и др.); а также регионы с малыми инвестиционными вложениями в области интеллектуальной собственности

(в частности, Республика Калмыкия, Чеченская Республика, Республика Ингушетия).

Компонентный анализ данных на основе выборки А3 позволил построить три главные компоненты, кумулятивная дисперсия которых составила 73,6%. Исключение признаков, вносящих наименьший вклад в формирование компонент (индекс промышленного производства и индекс производства в сельском хозяйстве), позволило повысить дисперсию до 92,9%. Цель анализа скорректированной выборки А3 состоит в формировании характеристик регионов, различающихся по результативности инвестиционного процесса в строительстве с учетом динамики инвестиционного обеспечения экономического развития.

Сформированы названия для главных компонент, состав которых определяется степенью влияния признаков (табл. 3). Название первой главной компоненты ГК31 – «Ввод в эксплуатацию жилых и нежилых зданий», второй главной компоненты ГК32 – «Динамика инвестиционного и экономического развития (одинаковые знаки)», третьей главной компоненты ГК33 – «Дина-

мика экономического развития (противоположные знаки)».

Диаграмма рассеяния демонстрирует сосредоточение большинства регионов в области невысоких значений ГК31 и средних значений ГК32 (рис. 3). Для этих регионов характерен небольшой объем завершеного строительства и средний темп инвестиционного развития. Лидерами по развитию инвестиций в строительство являются Московская область и Москва. Довольно высокий уровень развития инвестиционных процессов в строительстве в Краснодарском крае, Республике Татарстан и Республике Башкортостан. Высокие темпы сбалансиро-

ванного инвестиционного и экономического развития имеют экономики Магаданской и Курской областей, очень низкие – в Ненецком автономном округе и Республике Калмыкия.

На третьем шаге составлена интегральная выборка Agen на основе ранее разработанного метода [8, 9], проведен компонентный анализ. В выборку Agen вошли 9 признаков, которые соответствуют главным компонентам, полученным в результате анализа выборок A1, A2 и A3. Весовые коэффициенты признаков для первых двух главных компонент ГК1gen и ГК2gen представлены в табл. 4.

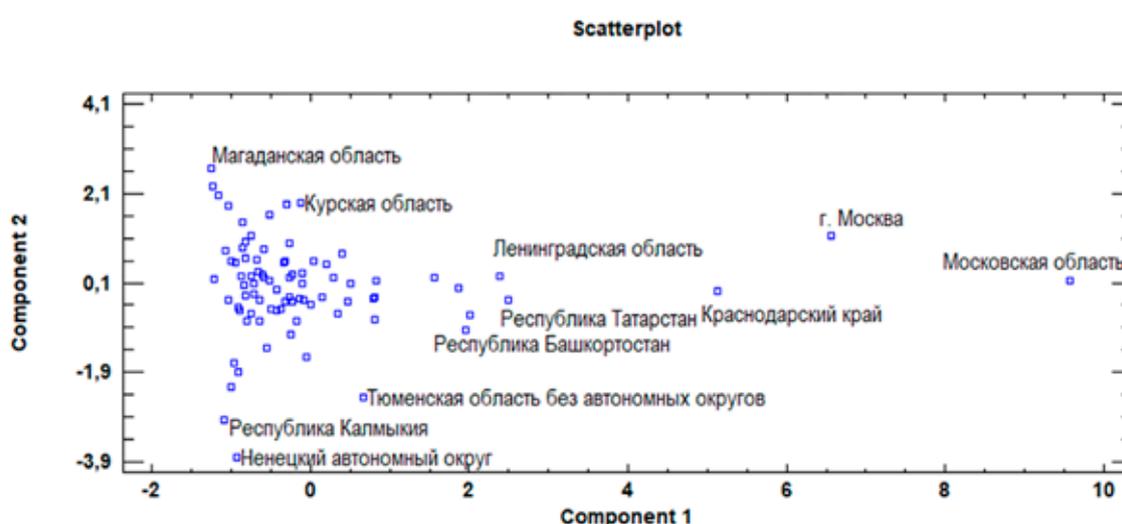


Рис. 3. Распределение регионов в пространстве первых двух КГ (выборка A3)

Таблица 4

Весовые коэффициенты признаков в главных компонентах (выборка Agen)

Название признаков	Номер ГК	ГК1gen	ГК2gen
Уровень экономического развития регионов с учетом инвестиций	ГК11	0,367859	0,134849
Уровень инвестиционного развития с учетом численности населения	ГК12	-0,144414	0,598891
Степень износа основных фондов	ГК13	0,372455	0,0607167
Инвестиции в землю с учетом инвестиций в оборудование и жилые здания	ГК21	-0,352001	-0,296169
Инвестиции в интеллектуальную собственность с учетом финансового результата организаций	ГК22	0,352	0,296187
Инвестиции в жилые здания	ГК23	-0,352001	-0,296179
Ввод в эксплуатацию жилых и нежилых зданий	ГК31	0,308667	-0,433768
Динамика инвестиционного и экономического развития (одинаковые знаки)	ГК32	0,367159	-0,100264
Динамика инвестиционного и экономического развития (противоположные знаки)	ГК33	0,321413	-0,3976

Анализ состава признаков, определяющих ГК1gen и ГК2gen, а также знаков коэффициентов при них позволяет заключить, что, во-первых, компонента ГК1gen характеризует в целом уровень экономического и инвестиционного развития регионов с учетом инвестиционных вложений в оборудование и интеллектуальную собственность; и, во-вторых, компонента ГК2gen характеризует уровень инвестиционного развития с учетом численности населения, а также инвестиции в строительство (с противоположными знаками). Особенность компоненты ГК2gen позволяет выявить регионы, в которых инвестиционные расходы на душу населения высокие, что присуще регионам с небольшой численностью населения и высокой долей трудовых ресурсов (рис. 4). К ним относятся северные регионы с развитой нефтедобывающей промышленностью, например Ямало-Ненецкий автономный округ (АО), Ханты-Мансийский автономный округ. Перечисленные особенности и определили названия компонент: для ГК1gen – «Уровень экономического и инвестиционного развития регионов с учетом вложений в оборудование и интеллектуальную собственность», для ГК2gen – «Уровень инвестиционного развития с учетом численности населения и вложений в строительство».

На диаграмме присутствует облако с высокой плотностью расположения регионов, структуру которого выявить на этом шаге не удалось. Выделены регионы, которые находятся далеко от этого облака и являются регионами-лидерами по уровню инвестиционного развития: г. Москва, Московская область, г. Санкт-Петербург, Ямало-Ненецкий автономный округ и Ненецкий автономный округ. Именно эти регионы вносят наибольший вклад в поляризацию данных. Требуется исключение их

из выборки Agen для проведения дальнейшего анализа.

На четвертом шаге формируется скорректированная выборка Agen_cor, проводится компонентный анализ с целью выявления кластерной структуры множества регионов, из которого на третьем шаге исключены регионы-лидеры.

На пятом шаге формируется рейтинг регионов по уровню инвестиционного развития.

Результаты исследования и их обсуждение

Анализ результатов проведения компонентного анализа скорректированной выборки Agen_cor позволил заключить, что состав признаков, определяющих содержание построенных первых двух главных компонент ГК1gen_cor и ГК2gen_cor, не изменился, поэтому названия компонент сохранены. Изменились только знаки признаков, определяющих ГК2gen_cor, на противоположные. Вследствие этого, изменилось направление оси ординат, определяемой ГК2gen_cor, и произошли соответствующие изменения в расположении регионов вдоль оси ГК2gen_cor (рис. 5).

На итоговой диаграмме рассеяния построены 9 квадрантов, полученных путем выделения интервалов с низкими, средними и высокими значениями ГК1gen_cor и ГК2gen_cor, в рамках которых выделены 10 кластеров. Наиболее высокий уровень экономического и инвестиционного развития (область высоких значений ГК1gen_cor) характерен для первых пяти одиночных кластеров (Краснодарский край, Свердловская область, Республика Татарстан, Красноярский край, Ханты-Мансийский автономный округ – Югра). Они представляют собой самые привлекательные регионы для инвестиционных вложений.

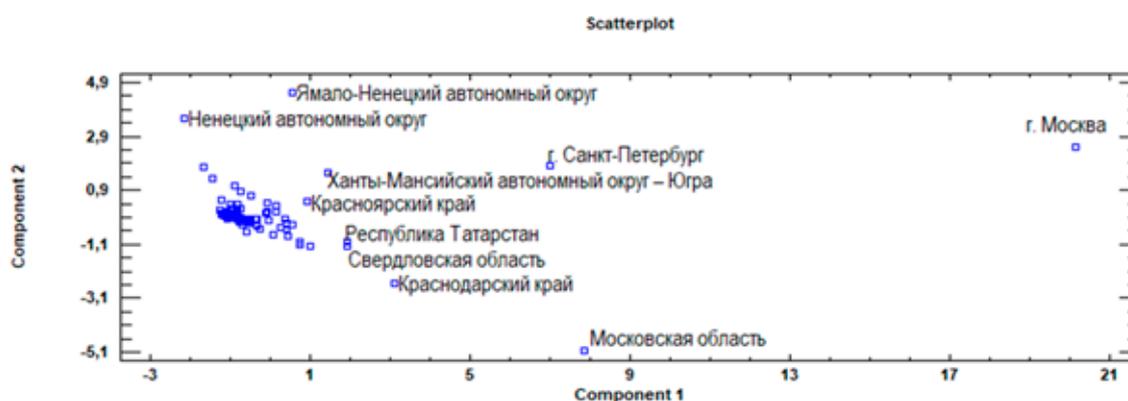


Рис. 4. Распределение регионов в пространстве первых двух ГК (выборка Agen)

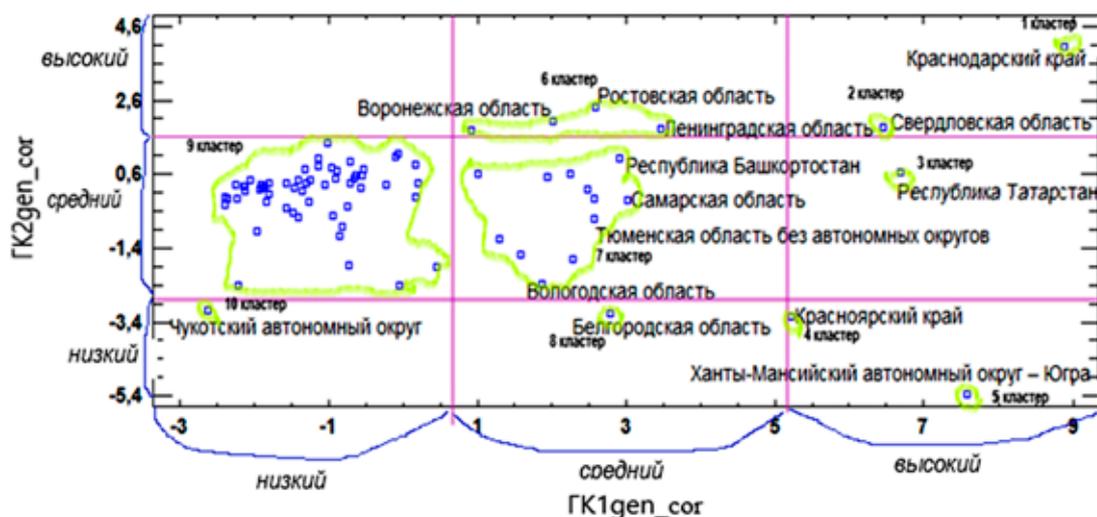


Рис. 5. Распределение регионов в пространстве первых двух ГК (выборка Aген_cor)

Таблица 5

Состав и характеристики кластеров регионов

Номер кластера	Состав	Уровень экономического и инвестиционного развития регионов с учетом вложений в оборудование и интеллектуальную собственность, GK1gen_cor	Уровень инвестиционного развития с учетом численности населения и вложений в строительство, GK2gen_cor
1	Краснодарский край	Очень высокий	Очень высокий
2	Свердловская область	Высокий	Довольно высокий
3	Республика Татарстан	Высокий	Средний
4	Красноярский край	Довольно высокий	Довольно низкий
5	Ханты-Мансийский автономный округ – Югра	Высокий	Очень низкий
6	Ростовская область, Воронежская область, Ленинградская область	Средний	Довольно высокий
7	Республика Башкортостан, Самарская область, Пермский край, Нижегородская, Челябинская, Кемеровская, Иркутская области и др.	Средний	Средний
8	Белгородская область	Средний	Довольно низкий
9	Рязанская, Томская области и др.	Низкий	Средний
10	Чукотский АО	Низкий	Довольно низкий

При анализе распределения регионов вдоль компоненты GK2gen_cor необходимо учитывать отрицательный знак коэффициента при признаке «Уровень инвестиционного развития с учетом численности населения». Наиболее высокий уровень инвестиционного развития с учетом численности населения (область малых значений GK2gen_cor) присущ северо-восточным регионам с малой численностью населения, в частности

Ханты-Мансийскому автономному округу (ХМАО), Чукотскому АО, Красноярскому краю; вложения в строительную отрасль здесь малы. И наоборот, большие инвестиции в строительство и недостаточно высокий уровень инвестиционных вложений с учетом высокой численности населения характерен для южных регионов, таких, например, как Краснодарский край, Ростовская область, Воронежская область.

Таблица 6

Рейтинг ТОП-10 регионов по уровню инвестиционного развития

Место в рейтинге	Регион	Место в рейтинге	Регион
1	г. Москва	6	Краснодарский край
2	Московская область	7	Свердловская область
3	г. Санкт-Петербург	8	Республика Татарстан
4	ХМАО	9	Красноярский край
5	Ямало-Ненецкий АО	10	Республика Башкортостан

Характеристики всех выделенных кластеров регионов представлены в табл. 5. Построен рейтинг регионов РФ по уровню экономического и инвестиционного развития, для этого используются значения компоненты ГК1gen. Полученный рейтинг регионов соответствует представленным выше диаграммам рассеяния (рис. 4 и 5). ТОП-10 регионов приведен в табл. 6.

Результаты построения рейтинга могут быть использованы для поддержки принятия решений на макро- и мезоуровнях социально-экономических систем.

Заключение

Таким образом, предложен алгоритм анализа данных об инвестиционных процессах и составления рейтингов регионов в области инвестиционного развития. Данный алгоритм основан на применении метода главных компонент и позволяет построить кластеры регионов, различающихся по уровню инвестиционного развития как базы экономического развития регионов в целом. Предложенный алгоритм целесообразно применять при разработке информационных систем поддержки принятия решений в области инвестиционной политики.

Список литературы

1. Агеева Е.В., Баранов А.О. Прогноз развития инвестиционного комплекса России в 2022–2023 гг. // ЭКО. 2022. № 5. С. 111–130.
2. Зубаревич Н.В. Регионы России в новых экономических условиях // Журнал Новой экономической ассоциации. 2022. № 3 (55). С. 226–234.
3. Гумба Х.М., Уваров С.С., Гамисония С.С. Устойчивое развитие экономики на мезоуровне: факторы и рейтинги // Недвижимость: экономика, управление. 2018. № 2. С. 37–42.
4. Батищева Г.А., Журавлева М.И., Новожилов А.С., Рязанов И.О. Анализ факторов развития национальной и региональной экономики // Вестник РГЭУ (РИНХ). 2020. № 2. С. 26–33.
5. Прогноз социально-экономического развития Российской Федерации на 2023 год и на плановый период 2024 и 2025 годов. [Электронный ресурс]. URL: <https://www.garant.ru/products/ipo/prime/doc/405257655/> (дата обращения: 29.04.2023).
6. Постановление Совета Федерации Федерального Собрания Российской Федерации от 26.04.2022 № 155 «О мерах Правительства Российской Федерации по обеспечению устойчивости экономики» [Электронный ресурс]. URL: <http://www.council.gov.ru/activity/documents/135388/> (дата обращения: 29.04.2023).
7. Регионы России. Социально-экономические показатели. Статистический сборник. М.: Росстат, 2022. 1122 с.
8. Ильясов Б.Г., Макарова Е.А., Закиева Е.Ш., Гиздадуллина Э.С. Оценка данных о доходах населения в региональном разрезе методом главных компонент // Экономика региона. 2019. Т. 15. Вып. 2. С. 601–617.
9. Ильясов Б.Г., Макарова Е.А., Закиева Е.Ш., Солнцев О.В. Технология формирования кластерной структуры сектора малого и среднего предпринимательства на основе метода главных компонент // Современные наукоемкие технологии. 2021. № 4. С. 51–57.