

Экспериментальные траектории процессов конвергенции и дивергенции неравномерности доходов населения регионов России

Д. Б. Берг¹  , Д. М. Балунгу¹ , А. Г. Шеломенцев² , К. С. Гончарова³ 

¹ Уральский федеральный университет
имени первого Президента России Б. Н. Ельцина,
г. Екатеринбург, Россия

² Институт социально-экономических исследований,
г. Уфа, Россия

³ Институт экономики Уральского отделения РАН,
г. Екатеринбург, Россия

 bergd@mail.ru

Аннотация. Исследование посвящено проблеме дифференциации доходов населения регионов России. Целью работы является разработка методики анализа процессов дифференциации доходов населения регионов России на базе теории динамических систем и машинного обучения, а также ее апробация на фактическом аналитическом материале. Гипотеза исследования заключается в предположении одновременного сосуществования процессов конвергенции и дивергенции дифференциации доходов населения регионов России, зависящих от внешних и внутренних факторов. Указанные процессы являются объектом исследования. Информационной базой исследования являются данные Росстата о значениях индекса Джини 80 регионов за период с 1995 по 2018 г. Для построения экспериментальных траекторий помимо индекса Джини использованы две независимые динамические переменные – его первая и вторая производные по времени, что позволило построить три различных пространства состояний (от одномерного до трехмерного). Методом кластеризации «*k*-средних» всё наблюдавшееся множество состояний было разделено на пять кластеров, количество которых было предварительно определено тестом «на осыпь» («метод локтя»). В результате расчетов было доказано преобладание конвергентных процессов над дивергентными в течение исследованного периода. Было обнаружено, что индивидуальные траектории движения отдельных регионов в пространстве состояний существенно отличаются: траектории некоторых регионов могут быть локализованы в пределах только одного кластера, тогда как отдельные части траекторий других могут принадлежать одновременно нескольким кластерам. Подавляющее большинство траекторий расположены в пределах 2–3 кластеров. Теоретическая значимость полученных результатов заключается в углублении представлений о региональной специфике динамики изменений дифференциации доходов населения субъектов Федерации. Практическая значимость результатов исследования заключается в расширении инструментальной поддержки принятия решений при реализации государственной политики в сфере регулирования дифференциации доходов населения на региональном уровне.

Ключевые слова: машинное обучение; кластеризация; теория динамических систем; пространство состояний; экспериментальные траектории; территориальные диспропорции; дифференциация уровня жизни; конвергенция/дивергенция доходов населения.

1. Введение

Неоднородность распределения доходов и межрегиональное неравенство всегда были и остаются острыми вопросами политической и экономической повестки дня стран с большой территорией, к которым относится и Россия. Обеспечение стабильного социально-экономического развития регионов России является необходимым условием устойчивого макроэкономического роста страны. Субъекты Федерации, вследствие географической протяженности последней, находятся под влиянием своих специфических природно-климатических и социально-экономических условий. Кроме того, вместе со всей страной они оказываются погружены в условия динамично меняющейся мировой экономической конъюнктуры.

Совокупность влияния разномастных (как региональных, так и макроэкономических) факторов внешней среды на такую комплексную динамическую систему, как регион, ставит перед теорией и практикой управления сложную задачу обеспечения его стабильного развития. Накопленные к настоящему времени территориальные диспропорции регионального развития значительно усложняют решение данной задачи. Одним из приоритетов социально-экономической политики в таких условиях является обеспечение необходимого уровня жизни населения и выравнивания диспропорций в его дифференциации, с одной стороны, и обеспечения темпов экономического роста — с другой.

Актуальность решения обозначенной проблемы обуславливается ее высокой значимостью для населения, поскольку высокая дифференциация доходов населения регионов приводит к росту социальной напряженности, снижению доверия населения к органам

власти, снижению эффективности национальной экономики, а также влиянием дифференциации доходов населения на неравенство индивидов в доступе к качественным услугам социальной сферы — образованию, здравоохранению, культуре, являющихся основой социального воспроизводства общества.

Принятие любых управленческих решений в такой крайне чувствительной сфере должно базироваться на аналитических подходах, комплексно и адекватно учитывающих изменения в процессах социально-экономического развития территорий. Необходимо не просто провести сравнительный анализ состояния дел на территориях, а постараться выявить существующие (если таковые имеются) долгосрочные тенденции. Такое возможно только при комплексной сравнительной идентификации ситуаций и проблем.

Уровень неравномерности распределения доходов населения формируется в результате влияния многих микро- и макроэкономических факторов, значимость которых варьируется от региона к региону. В настоящее время установлено существование значимой статистической связи между индексом Джини (характеристикой, наиболее часто используемой для измерения неравномерности доходов) и практически любым параметром социально-экономической сферы региона. Однако в результате такой детализации из поля зрения исследователя могут пропасть общие, интегральные тенденции, а именно они представляют наибольший интерес.

Поэтому настоящее исследование сфокусировано на сравнительном анализе временных рядов региональных значений индекса Джини. Исследование проводится в контексте известного конвергентного подхода к анализу уровня доходов населения, классическое описание которого представлено Sala-i-Martin [1] (исследование конвергенции на базе

неоклассической модели роста проведено им для больших временных рядов 110 стран, включая страны ОЭСР, штаты США и префектуры Японии), в последствии обобщенного им в соавторстве с Варго.

Varro и Sala-i-Martin [2] на эмпирических данных об экономике регионов разных стран как развитых, так и бедных, исследуют условия совпадения результатов действия различных механизмов конвергенции. При этом конвергенция территориальных различий дифференциации доходов населения интерпретируется как адаптационное свойство социально-экономических систем, характеризующееся сходством в поведении групп населения сравниваемых территорий, интеграции их мотивации и сближения ценностей. Дивергенция рассматривается как усложнение и увеличение разнообразия поведения групп населения, ранее имевших общие черты, в том числе за счет появления новых характеристик и моделей поведения. Вариации значения индекса Джини со временем рассматриваются авторами как индикатор сложных адаптационных изменений комплексной социально-экономической системы региона к динамичным изменениям его внутренней и внешней среды.

Наиболее перспективным подходом для решения задачи интегрального описания поведения сложной системы во времени является подход теории динамических систем. Согласно этому подходу, для описания поведения системы используется пространство ее состояний, так называемое «фазовое пространство», базовыми векторами которого являются три независимых переменных: 1) значение некоторого интегрального параметра системы (описывает положение системы); 2) его первая производная по времени (описывает скорость изменения данного положения); 3) вторая производная

по времени (отражает знак и величину силы, меняющей скорость изменения положения). Интегральным параметром, характеризующим неоднородность распределения доходов населения, является индекс Джини.

В рамках такого подхода траектория движения каждого региона в пространстве состояний динамической системы оказывается представлена последовательным множеством точек, каждая из которых соответствует состоянию региона за конкретный период (год). Общее множество точек пространства состояний, формирующих траектории движения всех регионов за исследуемый период, может быть разделено на кластеры, в каждом из которых значения координат точек будут наиболее близки друг другу. Локализация участков траекторий разных регионов в пределах одного и того же кластера означает сближение значений их параметров, характеризующих поведение исследуемых систем (регионов) в части неоднородности распределения в них доходов населения, т.е. конвергенцию. И наоборот, переход траектории некоторого региона из одного кластера в другой в определенный момент времени будет экспериментальным подтверждением процесса дивергенции параметров неоднородности распределения доходов населения данного региона по отношению к регионам, сохранившим свое положение в пределах прежнего кластера. Задача кластеризации (кластерного анализа, таксономии, классификации, распознавания образов и т.п.) в настоящее время выполняется методами так называемого «машинного обучения».

Целью исследования является разработка методики анализа процессов конвергенции и дивергенции дифференциации доходов населения регионов России на базе методов теории динамических систем и машинного обучения и ее апро-

бация на фактическом аналитическом материале.

Гипотеза исследования заключается в предположении одновременного существования процессов конвергенции и дивергенции дифференциации доходов населения регионов России, зависящих от внешних и внутренних факторов.

Объектом исследования являются процессы конвергенции и дивергенции дифференциации доходов населения страны в разрезе регионов.

В ходе настоящего исследования будут построены три различных фазовых пространства, координатами которых являются динамические параметры, отражающие неоднородность распределения доходов населения в регионах страны. В этих пространствах точками будут отмечены состояния регионов за исследуемый период, образующих, в свою очередь, траектории движения регионов в пространстве состояний. Наблюдающееся множество состояний регионов (точек фазового пространства) будет разделено на кластеры методами машинного обучения. Сравнительный анализ принадлежности участков траекторий регионов тому или иному кластеру позволит выявить особенности процессов конвергенции и дивергенции неоднородности распределения доходов населения страны в разрезе ее регионов.

Структура статьи. Текст статьи состоит из шести разделов и списка литературы. Во введении обосновывается важность разработки новых аналитических подходов для управления таким чувствительным социальным аспектом экономики региона, как неравномерность распределения доходов. По результатам обзора состояния изученности предметной области делается вывод о возможности использования подходов теории динамических систем и методов машинного обучения для развития аналитического инстру-

ментария исследований в данной области. В методологическом разделе описывается общий подход к проведению исследования, его информационная база, а также используемые математические методы. В следующем разделе приводится описание полученных результатов — экспериментальных траекторий регионов и их отдельных состояний — по каждому из трех исследуемых фазовых пространств. Сравнительный анализ траекторий, выявляющий особенности исследуемых процессов конвергенции и дивергенции, проводится в ходе обсуждения полученных результатов. В заключительном разделе полученные результаты обобщаются, приводится их теоретическая и практическая значимость, а также возможные направления дальнейших исследований.

2. Обзор состояния изученности предметной области

Проблемы региональной дифференциации и неравенства доходов населения вошли в повестку дня мирового сообщества в 60-х гг. С этого момента начинается ее активное изучение. Теоретико-методологический подход к исследованию этой проблемы традиционно связывается с работами С. Кузнеця середины XX в. Одна из первых его работ Kuznets [3] выполнена в традициях классической теории экономического роста и посвящена исследованию долгосрочной зависимости изменения неравенства доходов от экономического роста страны и определению ведущих факторов этого влияния.

Как правило, исследование явления дифференциации доходов населения проводится с применением методов статистической эконометрики, и в первую очередь корреляционно-регрессионного анализа. Построенная в результате такого анализа статистическая модель по-

зволяет лучше понять механизм влияния тех или иных социально-экономических факторов на неравномерность распределения доходов населения.

Малкина [4] исследует влияние уровня доходов и заработной платы на показатель внутрирегиональной дифференциации доходов населения.

Юревич [5] оценивает влияние индекса Джини на темпы прироста ВВП, а также объема ВВП на величину инвестиций в основной капитал.

Малкина [6] выполнила анализ многофакторных взаимосвязей нормального и избыточного неравенства с реальными доходами на душу населения, а также установила связи типов неравенства с показателями качества человеческого капитала, уровня жизни и благосостояния населения в регионах РФ.

Зубаревич и Сафронов [7] выявили зависимость динамики уровня жизни российского населения в постсоветский период от таких параметров, как доходы, расходы и потребление домохозяйств.

Овчарова и др. [8] использовали регрессионный анализ таких детерминант неравенства, как уровень образования, регион проживания, детская нагрузка, присутствие в домохозяйстве пожилых и др. для оценки вклада этих признаков в дисперсию подушевых расходов за 20 лет с 1994 по 2014 г. Вклад этих признаков оказался различным на разных этапах развития российской экономики.

Любимов [9] использовал регрессионный анализ для анализа панельных данных и выявил отрицательную корреляцию между экономическим ростом и неравенством доходов.

Григорьев и др. [10] использовали взаимодополняющие регрессионные модели для анализа временных рядов данных за период 2006–2012 гг. и обнаружили наличие положительного влияния прироста неравенства в распределении

дохода на темп экономического роста в регионах России.

Niea & Xingb [11] изучали зависимости неравенства доходов от уровня образования и семейного положения. Они показали, что уровень образования больше всего способствует увеличению неравенства доходов в городских районах в период с 1990 по 2009 г. и уровня развития финансового сектора.

Altunbas & Thorntonab [12] посредством квантильного регрессионного анализа показали, что в группе из 121 страны влияние уровня развития финансового сектора на неравенство доходов отличается для стран с низким, средним и высоким уровнем дохода.

Liu et al. [13] исследовали взаимосвязь неравномерности доходов населения и эмиссии диоксида углерода с привлечением моделей авторегрессии и квантильной регрессии.

Одновременно с построением регрессионных моделей для изучения явления региональной дифференциации населения по уровню доходов нередко используется кластерный анализ, который позволяет, во-первых, типологизировать объекты исследования так, чтобы они находились на наименьшем расстоянии друг от друга, во-вторых, обеспечить выявление естественного расслоения объектов исследования на четко выраженные группы (кластеры), лежащие друг от друга на некотором расстоянии.

Ибрагимова и Франц [14] изучали психосоциальное самочувствие населения, в ходе которого была проведена кластеризация регионов России по социально-экономическим показателям. Они также изучали взаимосвязи между уровнем социально-экономического развития региона проживания и объективным материальным положением жителей, а также его субъективным восприятием индивидами.

Афони́на и др. [15] разработали методику оценки качества жизни на основании совокупности методов корреляционного, регрессионного и кластерного анализа. Эта методика позволила им оценить благосостояние населения регионов Центрального федерального округа РФ с учетом неравенства распределения доходов в период с 2008 по 2016 г.

Войнова [16] провела сравнительное исследование классических методов декомпозиции доходного неравенства по источникам доходов и по группам домохозяйств, различающихся по социально-демографическим характеристикам (тип населенного пункта проживания, уровень образования и пр.). Автор рассмотрела теоретические ограничения и возможные области применения каждого из них, выявила ключевые отличия от остальных методов, а также присущие им достоинства и недостатки.

В исследованиях территориальных особенностей кластерный анализ широко применяют в задачах разделения географических единиц на однотипные группы на основании различных интегральных индексов.

Repiská et al. [17] сравнивают три метода кластерного анализа на выборке из 15 стран Европейского союза, используя эти индексы в качестве агрегированных показателей. Результаты их исследования подчеркивают потенциальную конвергенцию между странами внутри одной группы.

Fahmiyah & Ningrum [18] классифицировали провинции Индонезии по индексу человеческого развития с использованием метода *k-means* («*k*-средних») и обнаружили существование четырех различных групп, качественно отличающихся друг от друга.

Larasati et al. [19] для тех же 34 провинций Индонезии использовали метод усеченного кластера (TCLUST) и выделили только три кластера.

Sihombing et al. [20] произвели группировки округов/городов на территории одной из провинций Индонезии на основе показателей индекса человеческого развития с использованием алгоритма кластеризации *k-means*. Эти результаты были использованы для формирования программных целей правительства по увеличению человеческого потенциала.

Olimpia [21] анализировала различия между государствами — членами Европейского союза (ЕС) с точки зрения конкурентоспособности, человеческого капитала и инклюзивного развития. Автор утверждает, что могут быть большие различия в накоплении человеческого капитала при низких различиях в инклюзивности или конкурентоспособности, как это наблюдается в таких странах, как Болгария, Венгрия, Хорватия, Словацкая Республика и Италия.

Kharlamova et al. [22] в исследовании взаимосвязи между технологиями и неравенством доходов в ЕС одновременно используют регрессионный и кластерный анализ для выявления влияния изменений производительности труда на неравенство доходов. Оказалось, что влияние изменений производительности труда на неравенство доходов существенно в одних европейских странах, но незначительно в других.

Panzer & Postiglione [23] для анализа 110 провинций Италии с точки зрения количественного измерения регионального неравенства и его территориального распределения предложили новый класс мер на основе индекса Джини, используемого в расчетах с учетом веса населения соответствующего региона в населении страны.

Espoir [24] исследовал конвергенцию неравенства доходов между странами и определил различные «клубы» (кластеры) или устойчивые состояния неравенства. Согласно результатам этого исследования, во всем мире не суще-

ствуется равномерной конвергенции неравенства доходов. Вместо этого страны объединяются в пять отдельных клубов (кластеров), причем большинство стран попадают в три клуба (кластера) с высокой степенью конвергенции. Анализ показывает, что неравенство доходов между клубами (кластерами) со временем увеличивается, а неравенство доходов внутри клубов уменьшается.

Ogundari [25] исследовал конвергенцию неравенства доходов в африканских странах. Было выявлено пять групп стран со схожими характеристиками, которые сходятся к разным равновесным уровням неравенства доходов. Причем различия в неравенстве доходов со временем уменьшились внутри каждого такого кластера.

В целом методы машинного обучения в последние годы приобрели значительную популярность, поскольку представляют собой лучшую альтернативу традиционным эконометрическим методам в части анализа нелинейных зависимостей и выявления сложных закономерностей в данных. Однако в подавляющем большинстве исследований как территориальных различий, так и неоднородности распределения доходов населения для проведения кластерного анализа используются имеющиеся статистические данные «как есть», а «динамикой показателей» исследователи, как правило, называют их изменение во времени. Известно, что динамика — наука о силах, действующих на систему, а статистические данные отражают лишь результат влияния этих сил на социально-экономическую систему региона. «Истинно» динамические переменные в таких исследованиях не используются.

В экономике теория динамических систем используется для объяснения поведения и анализа динамики экономических систем на базе соответствующих математических моделей и концепций.

Эти модели позволяют анализировать эволюцию системы в целом на основании значений, описывающих состояние системы макропеременных.

Расчеты эволюции динамической системы традиционно основаны на описывающих систему дифференциальных уравнениях (в том случае, когда известны или предполагаются известными) действующие на систему силы и правила взаимодействия отдельных подсистем внутри системы.

Kydland & Prescott [26] представили модель реального делового цикла (RBC), в которой подчеркивается роль экзогенных шоков и колебаний совокупного предложения на рынке. В этой модели используются динамические методы стохастического общего равновесия (DSGE) для изучения циклических колебаний экономической активности.

Dinopoulos & Segerstrom [27] использовали модели нелинейной динамики для анализа закономерностей экономического роста и развития. Эти модели позволили продемонстрировать важность эндогенных механизмов, которые объясняют различия в темпах роста в разных странах.

Kenett & Havlin [28] провели обзор применения теории сетей в экономике. Они подчеркивают важность использование моделей динамических систем для исследования формирования и эволюции экономических сетей, таких как цепочки поставок, социальные сети и финансовые сети.

Если не известны внутренние механизмы, обеспечивающие эволюцию экономических систем и невозможно использовать соответствующие дифференциальные уравнения, то можно использовать экспериментальные данные о поведении системы, полученные в ходе компьютерных экспериментов (например, методами имитационного моделирования).

Теория динамических систем оказывается полезной в объяснении появления эффекта коллективного поведения агентов в моделях, задающих только правила их индивидуального поведения и взаимодействия друг с другом, таких как агент-ориентированное моделирование. Об этом на примерах финансовых рынков, рынков труда и работы промышленных предприятий подробно написано в обзоре Tesfatsion [29].

Экспериментальные траектории эволюции динамических систем могут быть построены также по результатам наблюдений за реальными экономическими системами. Однако такие работы практически отсутствуют. Редким исключением являются исследования конкурентных стратегий поведения агентов, в которых построение траекторий и идентификация типов стратегий проводится в соответствующем двумерном фазовом пространстве, образованном значением некоторого макропараметра и его первой логарифмической производной по времени. Berg & Pörkov [30] в качестве такой величины использованы активы банков (валюта нетто-баланса), а Berg et al. [31] использовали годовые объемы транзакций различных российских платежных систем.

Таким образом, включение в датасет машинного обучения для кластеризации регионов по неоднородности распределения доходов населения, помимо традиционно используемого индекса Джини дополнительных независимых динамических переменных, полученных на основе теории динамических систем, позволит выявить новые закономерности протекания процессов конвергенции и дивергенции неоднородности распределения доходов населения страны и определить их региональные особенности.

3. Методология исследования

Методологический подход настоящего исследования базируется на использовании подходов теории динамических систем и машинного обучения. Состояние региона (определяется за год) характеризуется значением индекса Джини G , а также двумя другими независимыми переменными — его первой и второй производной по времени G' и G'' соответственно.

Трехмерный вектор состояния региона (G , G' , G'') может быть спроецирован на одно- и двухмерные пространства. Множество состояний всех регионов за весь период исследования может быть разделено на некоторое количество подмножеств (классов, кластеров), в каждом из которых вектора состояний будут наиболее близки друг к другу. Все состояния считаются независимыми («эргодическая гипотеза»).

Они классифицируются методами машинного обучения, после чего строится траектория движения каждого региона в данном пространстве. Переход региона из одного кластера в другой означает, что имеет место быть процесс дивергенции с оставшимися регионами первого кластера и процесс конвергенции — с регионами, входящими во второй кластер.

Настоящее исследование включает выполнение следующих основных этапов:

- 1) подготовка данных для расчета траекторий;
- 2) определение оптимального количества кластеров в пространстве состояний;
- 3) классификация состояний регионов в одно-, двух- и трехмерном пространствах (одномерное — значение индекса Джини, двухмерное — значение индекса и его первая производная, трехмерное — значение индекса, первая и вторая его производные).

Для проведения необходимых расчетов нами использовались следующие пакеты (библиотеки) прикладных программ: Pandas (анализ данных), NumPy (выполнение математических операций над многомерными массивами данных), Sklearn (машинное обучение), Matplotlib.pyplot и Matplotlib.lines (визуализация данных) и ряд других.

3.1. Подготовка данных

Подготовка данных для расчета траекторий включала следующие шаги:

- формирование временных рядов значений индекса Джини для каждого региона на основании данных государственной статистики;
- расчет временных рядов первой производной индекса Джини по времени (G') — скорости изменения G ;
- расчет временных рядов второй производной индекса Джини по времени (G'') — темпов изменения скорости G' .

Согласно теории динамических систем, переменные G , G' и G'' являются независимыми.

Временные ряды значений индекса Джини G составлены по статистическим данным Росстата (издание «Регионы России. Социально-экономические показатели»). Использован интервал 1995–2018 гг.

При расчете значений временного ряда G' количество точек в ряду уменьшается на одну (поскольку рассчитывается разность значений между соседними точками), поэтому ряд ограничивается 2017 г., а при расчете временного ряда G'' — еще на одну, его последнее значение соответствует 2016 г. Таким образом результаты всех последующих расчетов соответствуют наиболее короткому временному интервалу из использованных в расчете.

Всего в работе рассчитываются траектории 80 регионов. Ряд регио-

нов исключен из анализа ввиду отсутствия части значений в исследуемый период ввиду административных изменений (Пермский край, Крым и др.). Значения индекса Джини для всех регионов на всем исследуемом интервале времени находятся в диапазоне значений от 0,231 до 0,623.

На рис. 1а приведены изменения усредненного по исследуемым регионам значения индекса Джини. Видно, что наблюдается устойчивая тенденция к его росту. Небольшое двухлетнее падение в конце 90-х гг. связано с последствиями кризиса 1998 г., в дальнейшем восстанавливается прежний тренд к росту. Следующее плавное снижение имеет место с 2015 г., что также может быть следствием введенных против России санкций и существенной девальвации рубля в тот период.

Однако среднее значение индекса дает мало информации о его региональной неоднородности. Поэтому было рассчитано отношение усредненного по всем регионам размаха (отклонениям от среднего) к самому усредненному значению (рис. 1б). Оказалось, что на протяжении всего исследуемого периода происходит относительно плавное выравнивание региональных различий в значении индекса Джини — от двухкратного в 1995 г. к 0,2 в 2018 г. Это означает, что пучок траекторий регионов по данному индексу становится в целом более узким, т. е. в масштабах страны процесс конвергенции преобладает над процессом дивергенции. Мощным толчком к этому оказался кризис 1998 г. (рис. 1б).

Таким образом, данные на рис. 1 показывают, что в условиях снижения территориальной неоднородности распределения доходов населения между регионами разница между богатыми и бедными в стране увеличивается.

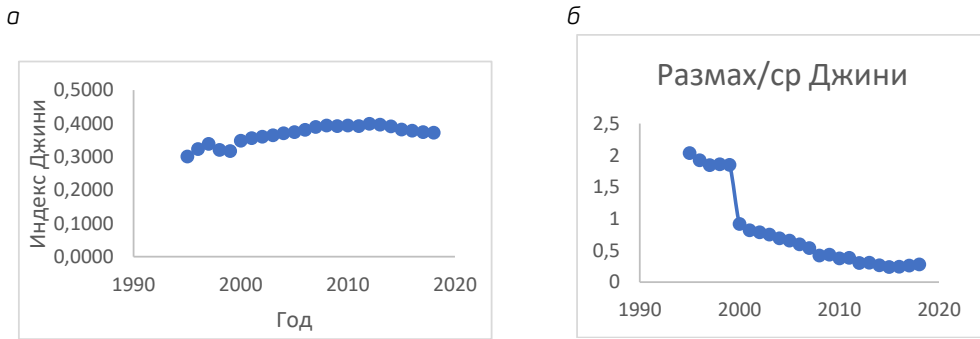


Рис. 1. Изменение усредненного по регионам значения индекса Джини (а) и отношения усредненного по регионам размаха относительно среднего к самому усредненному значению (б)

Figure 1. Changes of the regionally averaged value of the Gini index (a) and the ratio of the regionally averaged range relative to the average to the averaged value itself (b)

Источник: составлено авторами.

3.2. Определение оптимального количества кластеров

При кластеризации данных раздельными методами необходимо заранее задать число кластеров, на которые будет разбито исходное множество данных. Это может быть сделано экспертным методом или путем сравнения результатов классификации при различном количестве заданных кластеров. В данной работе используется второй вариант.

Сравнение результатов классификации проводится «методом локтя» (Elbow Method), предложенным Thorndike в 1953 г. [32]. Суть данного метода заключается в том, чтобы провести классификацию исходных данных для различного количества кластеров k , например, от 2 до 8. Для каждого значения k вычисляется сумма квадратов ошибок и отображается график зависимости суммы квадратов ошибок от количества кластеров.

Если представить такой график в виде руки, то точка, где изгиб графика больше всего напоминает изгиб локтя, даст нам наилучшее значение k . То есть необходимо определить наименьшее количество кластеров, при ко-

тором сумма квадратов ошибок остается небольшой.

Сумма квадратов ошибок определяется по формуле:

$$SSE = \sum_{k=1}^K W(C_k), \quad (1)$$

где $W(C_k)$ — это внутрикластерная вариация. Она находится по формуле:

$$W(C_k) = \frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^p (x_{ij} - x_{i'j})^2, \quad (2)$$

где $|C_k|$ обозначает число точек (наблюдений) в k -м кластере.

Пример использования «метода локтя» (другое название — тест «на осыпь») для определения оптимального количества кластеров при кластеризации методом « k -средних» (метод подробно описан в п. 3.3) в одномерном пространстве (G) приведен на рис. 2. При переходе к классификации от четырех до пяти кластеров, вариация уменьшилась незначительно, поэтому целесообразно проводить расчеты для пяти кластеров. Для других исследованных пространств результат расчета количества кластеров k оказался аналогичным.

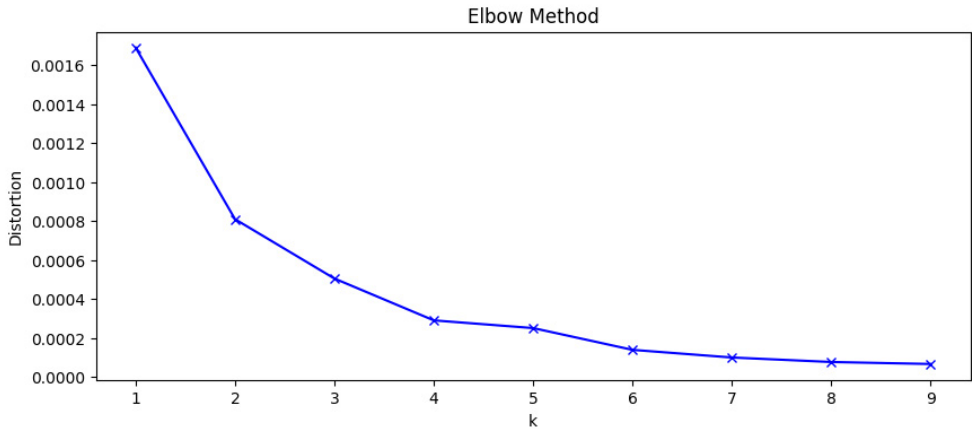


Рис. 2. Пример определения числа кластеров «методом локтя» (Elbow Method), k – заданное количество кластеров, Distortion – вариация внутрикластерных расстояний

Figure 2. An example of determining the number of clusters using the «Elbow Method», k is the specified number of clusters, Distortion is the variation of intra-cluster distances

Источник: составлено авторами.

3.3. Метод кластеризации (k -средних)

Для решения задачи кластеризации был использован метод « k -средних», поскольку он обладает такими сильными сторонами, как высокая эффективность при простоте его реализации, достаточный уровень качества выполненной кластеризации и возможность распараллеливания вычислительных процедур.

Алгоритм « k -средних» группирует N элементов исходного множества данных в k классов (кластеров), каждый из которых имеет один элемент (среднее значение) в качестве центра кластера, представляющего все элементы множества из этого кластера.

Алгоритм « k -средних» является одним из алгоритмов классификации с разбиением, при котором необходимо определение начальных значений центров кластеров, что подробно описано Eltibi & Ashour [33].

Выходные данные « k -средних» также зависят от выбранных значений центров при кластеризации, поэтому в данном методе начальное значение центральной точки кластера становится

основой для определения кластера (см. Likas et al. [34]). В качестве меры расчета расстояния используют евклидовое расстояние:

$$d = \sqrt{(x_1 - x_2)^2 + (\mu_1 - \mu_2)^2}.$$

Математическая постановка задачи классификации. Пусть дан набор из n наблюдений $X = \{x_1, x_2, \dots, x_n\}$, $x_i \in R^d$, $i = 1, \dots, n$. При этом k — заданное число кластеров, $k \in N, k \leq n$.

Требуется разделить множество X на k кластеров S_1, S_2, \dots, S_k :

$$S_i \cap S_j = \emptyset, i \neq j$$

$$\bigcup_{i=1}^k S_i = X.$$

Алгоритм k -средних разбивает набор X на k -наборов S_1, S_2, \dots, S_k таким образом, чтобы минимизировать сумму квадратов расстояний от каждой точки кластера до его центра. Введем обозначение $S = \{S_1, S_2, \dots, S_k\}$. Тогда действие алгоритма равносильно поиску:

$$\arg \min \sum_{i=1}^k \sum_{x \in S_i} \rho(x, \mu_i)^2,$$

где μ_i — центры кластеров, $i = 1, \dots, k$; $\rho(x, \mu_i)$ — функция расстояния между x и μ_i .

Классификация элементов исходного множества с использованием данного метода проводится в несколько этапов.

1. Определение числа кластеров k .
2. Инициализация k центров кластеров путем рандомизации элементов исходного множества.
3. Назначение каждого элемента соответствующему центру кластера путем расчета расстояния между данным элементом и данным центром с последующим выбором центра, ближайшего к этому элементу.
4. Вычисление новых центров кластеров путем усреднения всех элементов, содержащихся в каждом кластере.
5. Если ни один из N элементов не изменяет принадлежность к кластеру, то кластеризация завершена. В противном случае повторяются шаги 3–5.

стеру, то кластеризация завершена. В противном случае повторяются шаги 3–5.

4. Результаты расчетов траекторий движения регионов в пространстве состояний с использованием методов машинного обучения

4.1. Одномерное пространство (G)

Одномерное пространство образовано значениями индекса Джини (G) в пределах интервала значений от 0,231 до 0,623. Результаты классификации по пяти кластерам приведены в табл. 1. Нумерация кластеров в работе начинается с 0 в соответствии с результатами расчетов в использованном пакете прикладных программ машинного обучения.

Таблица 1. Результаты классификации регионов по значению индекса Джини

Table 1. Results of classification of regions according to the value of the Gini index

Параметр	Кластер 1	Кластер 2	Кластер 3	Кластер 4	Кластер 5
Центр кластера, G	0,306	0,353	0,386	0,422	0,565
Интервал значений G	0,231–0,33	0,331–0,37	0,371–0,404	0,405–0,486	0,501–0,623
Количество регионов (уникально)	63	75	76	37	2
Количество точек	246	675	653	304	18

Источник: составлено авторами.

Каждый кластер характеризуется интервалом значений индекса (границы кластера) и координатами центра кластера (точка, от которой сумма расстояний до каждого элемента кластера минимальна). Количество точек (наблюдавшихся состояний неравномерности доходов на-

селения в регионе — значений индекса Джини), сформировавших данный кластер, указано в нижнем ряду табл. 1.

Видно, что имеем два крупных кластера 2 и 3 (более чем по 650 точек каждый), со смежным диапазоном значений индекса Джини. Также имеем два

небольших кластера 1 и 4 (246 и 304 точек), находящихся слева и справа от совокупности двух больших кластеров. Самый маленький кластер, 5-й, образован из 18 точек с аномально большими значениями индекса Джини (более 0,5) всего двух регионов (г. Москва и Республики Тыва).

За весь анализируемый период в каждом из трех кластеров (1, 2 и 3) «побывали» почти все регионы страны (63, 75 и 76 соответственно), и только в 4-м кластере — около половины. Последовательность значений индекса Джини формирует траекторию движения соответствующего кластера в пространстве.

Таким образом становится очевидно, что даже при наблюдении траекто-

рий движения регионов в однопараметрическом пространстве одновременно наблюдаются процессы конвергенции и дивергенции. При переходе некоторого региона из одного кластера в другой имеет место быть процесс дивергенции относительно входящих в исходный кластер других регионов и конвергенции — по отношению к регионам, входящим в конечный кластер.

4.2. Двумерное пространство (G , G')

Двумерное пространство образовано значениями индекса Джини (G) и его первой производной по времени G' . Результаты классификации по пяти кластерам приведены в табл. 2.

Таблица 2. Результаты классификации регионов в двумерном пространстве по значениям G и G'

Table 2. Results of classification of regions in two-dimensional space according to the values of G and G'

Параметр	Кластер 1	Кластер 2	Кластер 3	Кластер 4	Кластер 5
Значение G центра кластера	0,302	0,350	0,383	0,420	0,565
Интервал значений G	0,231–0,344	0,308–0,369	0,367–0,403	0,403–0,486	0,501–0,623
Количество регионов (уникально)	63	74	77	38	2
Количество точек	205	599	662	331	18

Источник: составлено авторами.

За счет введения второй независимой переменной (G') интервалы значений самого индекса Джини, соответствующих разным кластерам, стали пересекаться (см. кластеры 1 и 2, 2 и 3 табл. 2).

Как видно из сравнения данных таблиц 1 и 2, значения индекса Джини центров кластеров сместились, но незначи-

тельно (в третьем знаке после запятой). Аномальный кластер, 5-й, из 18 точек остался без изменений. Количество уникальных регионов, попадавших в кластеры 2, 3 и 4, изменилось всего на 1. Чуть более значительно изменилось количество таких «попаданий» — максимальные изменения коснулись 1-го и 2-го кластеров, у них стало на 41 и 76 точек

меньше соответственно. Кластеры 3-й и 4-й, наоборот, немного «подросли» — на 9 и 27 точек соответственно. При общем сокращении количества точек на 80 (ввиду использования более короткого временного ряда G^t), увеличение кластера 4 выглядит весьма существенным.

В двухмерном пространстве траектории движения регионов можно построить особенно наглядно (рис. 3). Примеры траекторий разных типов демонстрируют такие регионы, как г. Москва, Московская область и Ямало-

Ненецкий автономный округ (ЯНАО). Траектория региона «Московская область» проходит через все четыре основных кластера (1–4). Траектория ЯНАО сохраняет свою принадлежность единственному кластеру — 4 — на протяжении всего анализируемого периода. Траектория г. Москва принадлежит двум кластерам — 4 и 5, регион переходит из кластера 5 с аномально высоким значением индекса Джини в кластер 4 — наглядно виден процесс конвергенции г. Москвы с таким регионом, как ЯНАО.

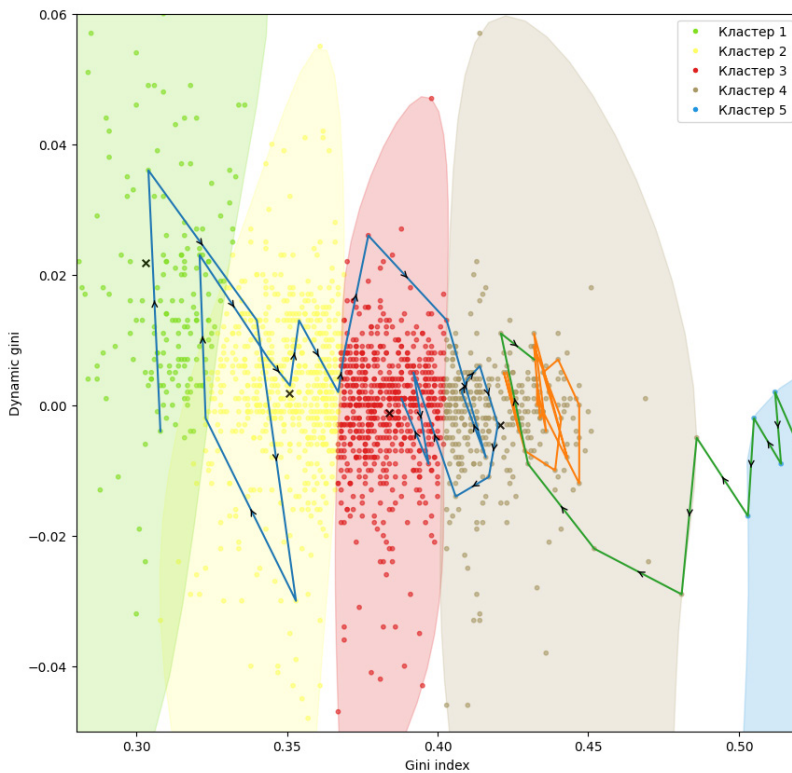


Рис. 3. Траектории движения регионов в двухпараметрическом пространстве (G, G^t). Центры кластеров отмечены значком «x». Цвета кластеров: 1 – зеленый, 2 – желтый, 3 – красный, 4 – коричневый, 5 – голубой. Примеры отдельных траекторий показаны линиями, соединяющими точки. Направление движения по каждой траектории указано стрелкой. Траектории регионов показаны цветными линиями: 1 – г. Москва (зеленая), 2 – ЯНАО (красная); 3 – Московская область (синяя)

Fig. 3. Trajectories of regions in two-parameter space (G, G^t). Cluster centers are marked with an “x”. Cluster colors: 1 – green, 2 – yellow, 3 – red, 4 – brown, 5 – blue. Examples of individual trajectories are shown with lines connecting the points. The direction of movement along each trajectory is indicated by an arrow. The trajectories of the regions are shown by colored lines: 1 – Moscow (green), 2 – Yamal-Nenets Autonomous Okrug (red); 3 – Moscow region (blue)

Источник: составлено авторами.

4.3. Трехмерное пространство (G, G', G'')

Трехмерное пространство образовано значениями индекса Джини (G), его первой G' и второй G'' производными по времени. Результаты классификации по ряти кластерам приведены в табл. 3.

Количество точек в трехмерном пространстве по сравнению с двухмерным сократилось на 80 ввиду более короткого временного ряда G'' . Интервал значений G в этой таблице не показан ввиду его низкой информативности.

Как видно из сравнения данных табл. 2 и 3, значения G центров кластеров при переходе от двухмерного пространства к трехмерному сместились более значительно, чем при переходе от одномерного к двухмерному, максимальный сдвиг составил 0,016 для кластера 2. Кластер 5 с аномально высоким значением G не изменился, в нем по-прежнему 18 точек траекторий тех же двух регионов.

Заметно изменилось количество точек в кластерах 1–4, причем произошло их перераспределение из 1 и 2 кластеров в пользу 3 и 4. Также изменилось количество «попадавших» в кластеры

1–4 регионов, наиболее значительно — для 2 и 4 кластеров. Распределение точек по кластерам трехмерного пространства наглядно представлено на рис. 4. Траектории движения регионов на этом рисунке не показаны ввиду недостаточной наглядности при визуализации пространства на плоскости.

5. Обсуждение: сравнительный анализ результатов расчетов траекторий (процессы конвергенции и дивергенции)

Сравнение результатов расчетов траекторий движения регионов в трех разных пространствах состояний, характеризующих неравномерность доходов населения и построенных на основании трех независимых переменных (абсолютное значение индекса Джини, его первая и вторая производные по времени) показывает, что использование динамических переменных (первой и второй производной индекса) существенно расширяет возможности изучения исследуемых процессов и позволяет извлечь из имеющихся данных новую информацию.

Таблица 3. Результаты классификации регионов в трехмерном пространстве по значениям G, G' и G''

Table 3. Results of classification of regions in three-dimensional space according to the values of G, G' and G''

Параметр	Кластер 1	Кластер 2	Кластер 3	Кластер 4	Кластер 5
Значение G центра кластера	0,305	0,334	0,373	0,414	0,565
Количество регионов (уникально)	68	64	76	51	2
Количество точек	138	384	750	445	18

Источник: составлено авторами.

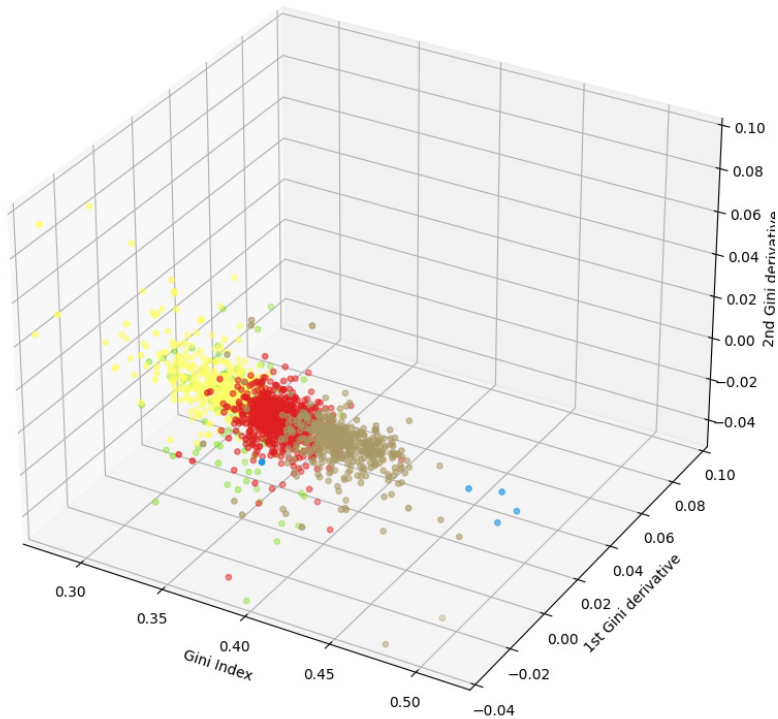


Рис. 4. Распределение регионов по кластерам в трехмерном пространстве состояний (G, G', G''). Цвета кластеров: 1 – зеленый, 2 – желтый, 3 – красный, 4 – коричневый, 5 – голубой

Fig. 4. Distribution of regions by clusters in three-dimensional space (G, G', G''). Cluster colors: 1 – green, 2 – yellow, 3 – red, 4 – brown, 5 – blue

Источник: составлено авторами

В табл. 4 приведено количество точек в каждом кластере по годам для всех трех изученных пространств, что интегрально отражает траектории движения регионов в этих пространствах, видно, что первоначально распределенные по всем пяти кластерам, со временем множество состояний регионов становится более компактными — они концентрируются в трех кластерах, причем в один из них попадает только 10 % регионов, остальные распределяются между оставшимися двумя.

Наблюдающееся уменьшение разнообразия состояний регионов по неравномерности доходов населения происходит постепенно (табл. 4), что подтверждает предварительный вывод (см. описание данных в п. 3) о преобладании процес-

сов конвергенции над процессами дивергенции в отношении неравномерности распределения доходов населения между регионами.

Несмотря на общую тенденцию к конвергенции в течение последних 20 лет, активно наблюдался и процесс дивергенции, особенно при рассмотрении небольших групп и даже пар отдельных регионов. Подробное рассмотрение перераспределения количества регионов между кластерами от года к году показывает (табл. 4), что целая группа регионов (от единиц до десятков) может менять свое положение за один год.

При этом между остающимися в кластере регионами и перешедшей в другой кластер группой расстояние в соответствующем пространстве увеличивается,

а значит, имеет место процесс дивергенции. Так, например, во втором из исследованных пространств (G, G') на 2012 г. в кластерах 2, 3 и 4 было 2, 46 и 32 региона соответственно, а через 5 лет, в 2017 г. их соотношение существенно поменялось — в кластере 2 количество регионов существенно увеличилось — до 37, а в четвертом, наоборот, уменьшилось до девяти (в третьем кластере осталось сравнимое с прежним количество — 34).

Сохранение своего положения регионом в рамках одного и того же кластера (независимо от типа пространства состояний) однозначно свидетельствует о его стабильности (в нашем случае — о стабильности степени неравномерности доходов населения). Поскольку за исследуемый 20-летний период макроэкономическая ситуация в стране менялась очень сильно, то сохранение регионом своего положения означает высокую степень его адаптации к таким внешним воздействиям, которая, в свою очередь, является следствием принятия своевременных управленческих решений с учетом региональных особенностей (природных и человеческих ресурсов, промышленного потенциала и др.).

Сгруппируем результаты расчетов так, чтобы было наглядно видно, скольким кластерам принадлежит траектория каждого региона в исследованных пространствах (табл. 5).

Оказывается, практически отсутствуют регионы, траектория которых локализована в пределах только одного кластера. Нет и таких, чья траектория касается всех пяти кластеров. Траектории только 10 % регионов относятся к двум кластерам, остальные — к трем или четырем. Это означает, что в пределах страны за последние 20 лет наблюдались интенсивные процессы дивергенции и конвергенции региональной неоднородности распределения доходов населения.

Все полученные и обсужденные выше результаты относятся к разделению пространств состояний регионов на пять кластеров. Несмотря на то, что такое разделение объективно обосновано результатами предварительно проведенного теста «на осыпь», полученные распределения точек (отдельных состояний регионов) в исследованных пространствах (рис. 3, 4) не позволяют визуально подтвердить целесообразность проведения границы между некоторыми кластерами, например между 2 и 3 на рис. 4. Поэтому остается открытым вопрос устойчивости полученных результатов к априорно выбранному количеству кластеров (в частности, к классификации на 4 и 6 кластеров), что является одним из методических ограничений данной работы.

Отдельного обсуждения требует вопрос о возможности присвоения собственных названий выделенным в данной работе кластерам. Традиционно собственные имена дают выделенным классам (или кластерам) на основании общих особенностей входящих в каждый их них элементов. Кластеризация (метод «классификации без учителя») позволяет достоверно отделить элементы одного кластера от элементов другого, но, как правило, не позволяет ответить на вопрос о том, в чем заключается принципиальная разница элементов одного кластера от элементов другого (в отличие от методов «классификации с учителем», где такие отличия заданы априорно). Это еще одно методическое ограничение данной работы. Поиск таких особенностей и их обоснование представляют сами по себе интересную исследовательскую задачу, однако она находится за пределами темы настоящей статьи и может рассматриваться как одно из возможных направлений продолжения исследований.

Таблица 4. Траектории движения регионов по кластерам
 Table 4. Trajectories of movement of regions by clusters

Кластер	Количество регионов в соответствующем кластере по годам																								
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	
	G																								
1	49	28	12	32	35	27	18	16	14	8	5	2	0	0	0	0	0	0	0	0	0	0	0	0	0
2	14	33	42	30	30	39	45	45	36	35	36	32	24	16	20	14	18	4	8	13	27	31	40	43	
3	5	7	14	6	5	8	11	10	23	28	28	33	36	40	40	44	40	46	47	46	36	36	33	31	
4	5	6	6	6	4	5	5	8	6	8	10	12	19	23	19	21	21	30	25	21	17	13	7	6	
5	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	
	G, G'																								
1	49	33	5	22	30	25	14	11	9	3	3	1	0	0	0	0	0	0	0	0	0	0	0	0	
2	12	27	46	40	35	41	45	45	40	38	36	28	20	13	11	11	13	2	4	6	21	28	37		
3	7	8	15	5	3	8	13	15	23	28	29	38	38	41	47	43	42	46	49	52	41	37	34		
4	5	6	8	7	6	5	7	8	7	10	11	12	21	25	21	25	24	32	27	22	18	15	9		
5	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0		

Окончание табл. 4

Кластер	Количество регионов в соответствующем кластере по годам																								
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	
1	40	44	2	10	19	10	5	1	4	0	1	0	1	0	1	0	0	0	0	0	0	0	0		
2	19	6	39	46	39	42	37	37	28	29	23	13	3	5	1	1	0	0	0	2	5	9			
3	8	18	21	9	9	18	27	31	39	36	40	44	45	41	44	42	47	39	44	44	53	51			
4	6	6	12	9	7	9	10	10	8	14	15	22	30	33	33	36	32	41	36	34	22	20			
5	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0			

G, G', G''

Источник: составлено авторами.

Таблица 5. Особенности траекторий отдельных регионов
Table 5. Features of the trajectories of individual regions

	1D	2D (Q, Q')	3D
1	1: Ямало-Ненецкий автономный округ	1: Ямало-Ненецкий автономный округ	3: Красноярский край, Ханты-Мансийский автономный округ, Ямало-Ненецкий автономный округ
2	8: Алтайский край, город Москва, Костромская область, Красноярский край, Ненецкий автономный округ, Республика Калмыкия, Тверская область, Ханты-Мансийский автономный округ	7: Алтайский край, город Москва, Красноярский край, Ненецкий автономный округ, Республика Калмыкия, Тверская область, Ханты-Мансийский автономный округ	7: Город Москва, Иркутская область, Кемеровская область — Кузбасс, Ненецкий автономный округ, Республика Бурятия, Республика Калмыкия, Чукотский автономный округ
3	48: Амурская область, Брянская область, Владимирская область, Волгоградская область, Вологодская область, Еврейская автономная область, Ивановская область, Иркутская область, Кабардино-Балкарская Республика, Калининградская область, Калужская область, Камчатский край, Кемеровская область — Кузбасс, Кировская область, Краснодарский край, Курская область, Ленинградская область, Липецкая область, Мурманская область, Оренбургская область, Пензенская область, Пермский край, Приморский край, Псковская область, Республика Ингушетия, Республика Бурятия, Республика Алтай, Республика Ингушетия, Республика Карелия, Республика Марий Эл, Республика Мордовия, Республика Северная Осетия — Алания, Республика Татарстан	49: Амурская область, Брянская область, Владимирская область, Волгоградская область, Вологодская автономная область, Еврейская автономная область, Ивановская область, Иркутская область, Кабардино-Балкарская Республика, Калининградская область, Калужская область, Камчатский край, Кемеровская область — Кузбасс, Кировская область, Краснодарский край, Курская область, Ленинградская область, Липецкая область, Мурманская область, Оренбургская область, Пензенская область, Пермский край, Приморский край, Псковская область, Республика Ингушетия, Республика Бурятия, Республика Карелия, Республика Марий Эл, Республика Мордовия, Республика Северная Осетия-Алания, Республика Татарстан (Татарстан), Республика Тыва, Республика Хакасия, Рязанская область, Саратовская область, Смоленская область, Ставропольский край, Тверская область, Тульская область,	36: Алтайский край, Владимирская область, Волгоградская область, Вологодская область, Еврейская автономная область, Ивановская область, Кабардино-Балкарская Республика, Калининградская область, Камчатский край, Кировская область, Костромская область, Краснодарский край, Курганская область, Ленинградская область, Мурманская область, Пермский край, Приморский край, Псковская область, Республика Алтай, Республика Ингушетия, Республика Карелия, Республика Коми, Республика Марий Эл, Республика Мордовия, Республика Северная Осетия — Алания, Республика Татарстан (Татарстан), Республика Тыва, Республика Хакасия, Рязанская область, Саратовская область, Смоленская область, Ставропольский край, Тверская область, Тульская область,

Окончание табл. 5

	1D	2D (Q, Q')	3D
3	(Татарстан), Республика Хакасия, Ростовская область, Рязанская область, Саратовская область, Смоленская область, Ставропольский край, Тамбовская область, Томская область, Тульская область, Удмуртская Республика, Ульяновская область, Хабаровский край, Челябинская область, Чувашская Республика — Чувашия, Чукотский автономный округ, Ярославская область	Республика Тыва, Республика Хакасия, Ростовская область, Рязанская область, Самарская область, Саратовская область, Смоленская область, Ставропольский край, Тамбовская область, Тульская область, Удмуртская Республика, Ульяновская область, Челябинская область, Чувашская Республика — Чувашия, Чукотский автономный округ, Ярославская область	Удмуртская Республика, Чувашская Республика — Чувашия
4	23: Астраханская область, Белгородская область, Воронежская область, город Санкт-Петербург, Забайкальский край, Карачаево-Черкесская Республика, Курганская область, Магаданская область, Московская область, Нижегородская область, Новосибирская область, Омская область, Орловская область, Республика Адыгея (Адыгея), Республика Башкортостан, Республика Дагестан, Республика Коми, Республика Саха (Якутия), Республика Саха (Якутия), Республика Тыва, Самарская область, Свердловская область, Свердловская область	23: Астраханская область, Белгородская область, Воронежская область, город Санкт-Петербург, Забайкальский край, Карачаево-Черкесская Республика, Курганская область, Магаданская область, Московская область, Нижегородская область, Новосибирская область, Омская область, Орловская область, Республика Адыгея (Адыгея), Республика Башкортостан, Республика Дагестан, Республика Коми, Республика Саха (Якутия), Сахалинская область, Свердловская область, Томская область, Хабаровский край	34: Амурская область, Астраханская область, Белгородская область, Брянская область, Воронежская область, город Санкт-Петербург, Забайкальский край, Калужская область, Карачаево-Черкесская Республика, Курская область, Липецкая область, Магаданская область, Московская область, Нижегородская область, Новосибирская область, Омская область, Оренбургская область, Орловская область, Пензенская область, Республика Адыгея (Адыгея), Республика Башкортостан, Республика Дагестан, Республика Саха (Якутия), Ростовская область, Самарская область, Сахалинская область, Свердловская область, Тамбовская область, Томская область, Ульяновская область, Хабаровский край, Челябинская область, Ярославская область

Источник: составлено авторами.

Таким образом сравнительный анализ полученных в результате расчетов экспериментальных траекторий доказывает гипотезу об одновременном сосуществовании процессов конвергенции и дивергенции региональной неоднородности распределения доходов населения.

Также исследование наглядно подтверждает высказанное в начале статьи на основании анализа данных описательной статистики предположение о преобладании процессов конвергенции над процессами дивергенции: в течение исследованного 20-летнего периода 90 % регионов в итоге локализируются в границах всего двух близко расположенных кластеров. Определить тип конвергенции в данной работе не представляется возможным, поскольку традиционно используемые для этого логарифмические функции не могут быть применены к анализу временных рядов неаддитивных величин, к которым относится и индекс Джини.

Данные результаты находятся в хорошем согласии с выводами, представленными Коломак [35] (выводы сделаны автором на основании данных о среднедушевых доходах населения на более коротком — 15-летнем интервале), о том, что в стране в целом процессы конвергенции преобладают над процессами дивергенции.

Также наши результаты в точности соответствует основным требованиям к такого типа работам, изложенным в статье Найдена и Белоусовой [36], о формировании методической базы межтерриториальных сравнений. В частности, соблюдение требования достоверности обеспечивается использованием данных официальной статистики, требования сопоставимости использованием безразмерных величин, требование наличия достаточно длинного временного ряда для характеристики процесса развития исследуемого явления исполь-

зованием 20-го временного ряда, превышающего длину временных рядов, используемых другими исследователями, требование ориентации на методы автоматизированной обработки исходных данных применением методов машинного обучения и т. д.

6. Заключение

Результатом данного исследования является разработка методики анализа процессов конвергенции и дивергенции региональной неоднородности доходов населения страны путем построения траекторий движения отдельных регионов в пространстве состояний различной размерности с использованием методов машинного обучения и ее апробация на данных Росстата.

Для построения пространства состояний, в соответствии с теорией динамических систем, использовалось до трех независимых переменных, характеризующих величину неоднородности доходов населения (индекс Джини G), скорость ее изменения (первая производная индекса Джини по времени G') и темпы изменения этой скорости (вторая производная индекса Джини по времени G''). Проанализирован характер данных траекторий в трех пространствах различной размерности: (G) , (G, G') , (G, G', G'') .

Классификация всех наблюдавшихся состояний 80 регионов за период с 1995 по 2018 г. выполнена методом « k -средних», количество кластеров для классификации определялось тестом «на осыпь» («методом локтя») — для всех исследованных пространств их количество оказалось одинаковым и равно пяти.

В результате расчетов траекторий была подтверждена гипотеза об одновременном сосуществовании процессов конвергенции и дивергенции региональной дифференциации доходов населения субъектов Федерации. Было

доказано преобладание конвергентных процессов над дивергентными: в течение исследованного 20-летнего периода 90 % регионов в итоге локализируются в масштабе всего двух близко расположенных кластеров.

Оказалось, что индивидуальные траектории движения отдельных регионов в пространстве состояний существенно отличаются: траектории некоторых регионов локализируются в пределах только одного кластера (например, Ямало-Ненецкого АО в пространствах (G) , (G, G') и (G, G', G'')), тогда как отдельные части траекторий других могут принадлежать четырем различным кластерам (таких регионов оказалось 20–25 % от всего количества исследованных). Однако подавляющее большинство траекторий расположены в пределах 2–3 кластеров. Полученные результаты полностью соответствуют заявленной цели исследования.

Теоретическая значимость полученных результатов заключается в углублении представлений о региональной специфике динамики изменений дифференциации доходов населения субъектов Федерации. Практическая значимость результатов исследования заключается в расширении инструментальной поддержки принятия решений при реализа-

ции государственной политики в сфере регулирования дифференциации доходов населения на региональном уровне.

Дальнейшие направления проведенных в настоящей работе исследований могут быть сфокусированы как на продолжении развития используемых эконометрических методов, так и на выявлении социально-экономических механизмов формирования той или иной траектории региона в пространстве исследуемых состояний. Так, развитие методики может заключаться, например, в исследовании траекторий процессов конвергенции и дивергенции пар регионов. В этом случае увеличение расстояния между траекториями в паре со временем будет означать дивергенцию, а уменьшение — конвергенцию. Привлечение независимых данных об изменении внешних и внутренних условий, а также проводимой социально-экономической политики в регионе для объяснения объективно наблюдающихся процессов конвергенции и дивергенции (переход из одного класса в другой, сближение или расхождение траекторий в паре сравнения и др.) позволят уточнить региональные особенности формирования неоднородности доходов населения.

Список использованных источников

1. *Sala-i-Martin X.* The Classical Approach to Convergence Analysis // *The Economic Journal*. 1996. Vol. 106, Issue 437. Pp. 1019–1036. <https://doi.org/10.2307/2235375>
2. *Barro R. J., Sala-i-Martin X.* *Economic Growth*. Second Edition. Cambridge, MIT Press. 2004. 654 p. URL: <http://piketty.pse.ens.fr/files/BarroSalaIMartin2004.pdf>
3. *Kuznets S.* Economic Growth and Income Inequality // *American Economic Review*. 1955. Vol. 45, No. 1. Pp. 1–28. URL: <https://www.aeaweb.org/aer/top20/45.1.1-28.pdf>
4. *Малкина М. Ю.* Оценка факторов внутрирегиональной дифференциации доходов населения РФ // *Пространственная экономика*. 2015. № 3. С. 97–119. <http://dx.doi.org/10.14530/se.2015.3.097-119>
5. *Юревич М. А.* Социальное неравенство, инвестиции и экономический рост // *Вопросы регулирования экономики*. 2019. Т. 10, № 4. С. 35–46. <https://doi.org/10.17835/2078-5429.2019.10.4.035-046>
6. *Малкина М. Ю.* Взаимосвязь типов неравенства с показателями уровня жизни и благосостояния населения регионов России // *Terra Economicus*. 2017. Т. 15, № 4. С. 46–63. <http://dx.doi.org/10.23683/2073-6606-2017-15-4-46-63>

7. Зубаревич Н. В., Сафронов С. Г. Люди и деньги: доходы, потребление и финансовое поведение населения российских регионов в 2000–2017 гг. // Известия РАН. Серия географическая. 2019. № 5. С. 3–17. <https://doi.org/10.31857/S2587-5566201953-17>
8. Овчарова Л. Н., Попова Д. О., Рудберг А. М. Декомпозиция факторов неравенства доходов в современной России // Журнал Новой экономической ассоциации. 2016. № 3 (31). С. 170–185. <https://doi.org/10.31737/2221-2264-2016-31-3-8>
9. Любимов И. Л. Неравенство и экономический рост: вызовы для российской экономики // Российское предпринимательство. 2017. Т. 17, № 1. С. 11–22. <http://dx.doi.org/10.18334/rp.17.1.2194>
10. Григорьев Р. А., Крамин М. В., Крамин Т. В., Тимирязова А. В. Неравенство распределения дохода и экономический рост в регионах России в посткризисный период // Экономика региона. 2015. № 3. С. 102–113. <https://doi.org/10.17059/2015-3-9>
11. Niea H., Xingb C. Education expansion, assortative marriage, and income inequality in China // China Economic Review. 2019. Vol. 55. Pp. 37–51. <https://doi.org/10.1016/j.chieco.2019.03.007>
12. Altunbas Y., Thorntonab J. The impact of financial development on income inequality: A quantile regression approach // Economics Letters. 2019. Vol. 175. Pp. 51–56. <https://doi.org/10.1016/j.econlet.2018.12.030>
13. Liu C., Jiang Y., Xie R. Does income inequality facilitate carbon emission reduction in the US? // Journal of Cleaner Production. 2019. Vol. 217. Pp. 380–387. <https://doi.org/10.1016/j.jclepro.2019.01.242>
14. Ибрагимова З. Ф., Франц М. В. Неравенство доходов, его субъективное восприятие и влияние на психосоциальное самочувствие населения // Статистика и экономика. 2018. Т. 15, № 4. С. 52–60. <http://dx.doi.org/10.21686/2500-3925-2018-4-52-60>
15. Афонина Т. Н., Ломанов А. О., Шуметов В. Г. Оценка качества жизни населения регионов Центральной России с учетом дифференциации денежных доходов // Региональная экономика: теория и практика. 2018. Т. 16, № 10. С. 1976–1994. <https://doi.org/10.24891/re.16.10.1976>
16. Войнова В. И. Подходы к декомпозиции доходного неравенства: отличия и условия применения // Учет и статистика. 2021. № 3 (63). С. 88–98. <https://doi.org/10.54220/1994-0874.2021.63.3.010>
17. Repiská R., Grisáková N., Štetka P. Hierarchical Clustering of EU Countries Based on HDI and EPI Index // Proceedings from the EDAMBA 2021 Conference. 2022. Pp. 420–430. <https://doi.org/10.53465/EDAMBA.2021.9788022549301.420-430>
18. Fahmiyah I., Ningrum R. Human Development Clustering in Indonesia: Using K-Means Method and Based on Human Development Index Categories // Journal of Advanced Technology and Multidiscipline. 2023. Vol. 2, No. 1. Pp. 27–33. <https://doi.org/10.20473/jatm.v2i1.45070>
19. Larasati S., Nisa K., Herawati N. Robust Principal Component Trimmed Clustering of Indonesian Provinces Based on Human Development Index Indicators // Journal of Physics: Conference Series. 2021. Vol. 1751. 012021. <http://dx.doi.org/10.1088/1742-6596/1751/1/012021>
20. Sihombing S., Sembiring R., Irawan E. Application of the K-Means Clustering Algorithm in Grouping Regencies/Cities in North Sumatra Province Based on Human Development Index Indicators // Jurnal Mandiri IT. 2022. Vol. 11, No. 1. Pp. 20–25. <https://doi.org/10.35335/mandiri.v11i1.124>
21. Olimpia N. Disparities Regarding Competitiveness, Human Capital and Inclusive Development in the EU: A Cluster Analysis // Annals — Economy Series. 2019. Vol. 1. Pp. 61–71. URL: <https://econpapers.repec.org/RePEc:cbu:jrnlec:y:2019:v:1:p:61-71>
22. Kharlamova G., Stavitskiy A., Zarotiadis G. The impact of technological changes on income inequality: the EU states case study. Journal of International Studies. 2018. Vol. 11, No. 2. Pp. 76–94. <https://doi.org/10.14254/2071-8330.2018/11-2/6>
23. Panzera D., Postiglione P. Measuring the Spatial Dimension of Regional Inequality: An Approach Based on the Gini Correlation Measure // Social Indicators Research. 2019. Vol. 148. Pp. 379–394. <https://doi.org/10.1007/s11205-019-02208-7>

24. *Espoir D. K.* Convergence or divergence patterns in income distribution across countries: A new evidence from a club clustering algorithm // *Cogent Economics & Finance*. 2022. Vol. 10, Issue 1. 2025667. <https://doi.org/10.1080/23322039.2022.2025667>
25. *Ogundari K.* Club Convergence in Income Inequality in Africa // *Social Indicators Research*. 2023. Vol. 167. Pp. 319–337. <https://doi.org/10.1007/s11205-023-03108-7>
26. *Kydland F. E., Prescott E. C.* Time to Build and Aggregate Fluctuations // *Journal of Econometrica*. 1982. Vol. 50, No. 6. Pp. 1345–1370. <https://doi.org/10.2307/1913386>
27. *Dinopoulos E., Segerstrom P.* Intellectual property rights, multinational firms and economic growth // *Journal of Development Economics*. 2010. Vol. 92, Issue 1. Pp. 13–27. <https://doi.org/10.1016/j.jdeveco.2009.01.007>
28. *Kenett D. Y., Havlin S.* Network science: a useful tool in economics and finance // *Mind & Society*. 2015. Vol. 14. Pp. 155–167. <https://doi.org/10.1007/s11299-015-0167-y>
29. *Tesfatsion L.* Agent-based computational economics: modeling economies as complex adaptive systems // *Journal of Information Sciences*. 2003. Vol. 149, Issue 4. Pp. 262–268. [https://doi.org/10.1016/S0020-0255\(02\)00280-3](https://doi.org/10.1016/S0020-0255(02)00280-3)
30. *Berg D. B., Popkov V. V.* General numerical model of the competition life cycle: from physics to economy // *Physica A: Statistical Mechanics and its Applications*. 2003. Vol. 324, Issue 1–2. Pp. 167–173. [https://doi.org/10.1016/S0378-4371\(02\)01892-7](https://doi.org/10.1016/S0378-4371(02)01892-7)
31. *Berg D., Demina M., Isaichik K., Panachev A., Popkov V., Parusheva S.* Competition of payment systems in Russia: Numerical analysis // *AIP Conference Proceedings*. 2018. Vol. 2040, Issue 1. 050014. <https://doi.org/10.1063/1.5079112>
32. *Thorndike R.* Who Belongs in the Family? // *Psychometrika*. 1953. Vol. 18. No. 4. Pp. 267–276. <https://doi.org/10.1007/BF02289263>
33. *Eltibi M. F., Ashour W. M.* Initializing K-Means Clustering Algorithm using Statistical Information // *International Journal of Computer Applications (IJCA)*. 2011. Vol. 29. Pp. 51–55. <https://doi.org/10.5120/3573-4930>
34. *Likas A. C., Vlassis N. A., Verbeek J. J.* The global k-means clustering algorithm // *Pattern Recognition*. 2003. Vol. 36, Issue 2. Pp. 451–461. [https://doi.org/10.1016/S0031-3203\(02\)00060-2](https://doi.org/10.1016/S0031-3203(02)00060-2)
35. *Коломак Е. А.* Межрегиональное неравенство в России: экономический и социальный аспект // *Пространственная экономика*. 2010. № 1. С. 26–35. <http://dx.doi.org/10.14530/se.2010.1.026-035>
36. *Найден С. Н., Белоусова А. В.* Методический инструментарий оценки благосостояния населения: межрегиональное сопоставление // *Экономика региона*. 2018. Т. 14, № 1. С. 53–68. <http://dx.doi.org/10.17059/2018-1-5>

ИНФОРМАЦИЯ ОБ АВТОРАХ

Берг Дмитрий Борисович

Доктор физико-математических наук, профессор, базовая кафедра аналитики больших данных и методов видеоанализа Института радиоэлектроники и информационных технологий Уральского федерального университета имени первого Президента России Б. Н. Ельцина, г. Екатеринбург, Россия (620002, г. Екатеринбург, ул. Мира, 19); ORCID <https://orcid.org/0000-0002-7703-9750> e-mail: bergd@mail.ru

Балунгу Даниель Мусафири

Ассистент базовой кафедры аналитики больших данных и методов видеоанализа Института радиоэлектроники и информационных технологий Уральского федерального университета имени первого Президента России Б. Н. Ельцина, г. Екатеринбург, Россия (620002, г. Екатеринбург, ул. Мира, 19); ORCID <https://orcid.org/0009-0001-5098-7603> e-mail: danielbal03.db@gmail.com

Шеломенцев Андрей Геннадьевич

Доктор экономических наук, профессор, главный научный сотрудник Института социально-экономических исследований — обособленное структурное подразделение Федерального государственного бюджетного научного учреждения Уфимского федерального исследовательского центра Российской академии наук (ИСЭИ УФИЦ РАН), г. Уфа, Россия (450054, г. Уфа, Проспект Октября, 71); ORCID <https://orcid.org/0000-0003-1904-9587> e-mail: a.shelom@yandex.ru

Гончарова Ксения Сергеевна

Кандидат экономических наук, научный сотрудник Института экономики Уральского отделения РАН, г. Екатеринбург, Россия (620014, г. Екатеринбург, ул. Московская, 29); ORCID <https://orcid.org/0000-0003-2381-3322> e-mail: ksenia.gon4arowa@gmail.com

БЛАГОДАРНОСТИ

Исследование выполнено при поддержке гранта Российского научного фонда № 22-28-01702 «Экспериментальные траектории процессов пространственной конвергенции и дивергенции доходов населения регионов России в условиях их адаптации к динамичным изменениям».






ДЛЯ ЦИТИРОВАНИЯ

Берг Д. Б., Балунгу Д. М., Шеломенцев А. Г., Гончарова К. С. Экспериментальные траектории процессов конвергенции и дивергенции неравномерности доходов населения регионов России // Journal of Applied Economic Research. 2024. Т. 23, № 2. С. 364–393. <https://doi.org/10.15826/vestnik.2024.23.2.015>

ИНФОРМАЦИЯ О СТАТЬЕ

Дата поступления 4 декабря 2023 г.; дата поступления после рецензирования 10 апреля 2024 г.; дата принятия к печати 24 апреля 2024 г.

Experimental Trajectories of Convergence and Divergence Processes of Russian Regions Population Incomes Inequality

Dmitry B. Berg¹  , Daniel M. Balungu¹ , Andrei G. Shelomentsev² ,
Kseniya S. Goncharova³ 

¹ Ural Federal University
named after the First President of Russia B. N. Yeltsin,
Yekaterinburg, Russia

² Institute of Socio-Economic Research,
Ufa, Russia

³ Institute of Economics, The Ural Branch of Russian Academy of Sciences,
Yekaterinburg, Russia

 bergd@mail.ru

Abstract. The study is devoted to the problem of differentiation of incomes of the population of Russia's regions. The goal of the work is to develop a methodology for analyzing the processes of differentiation of incomes of the population in Russia's regions based on the theory of dynamic systems and machine learning, and its testing on recent analytical data. The research hypothesis lies in the assumption of the simultaneous coexistence of processes of convergence and divergence of processes of differentiation of incomes of the population of Russia's regions, depending on external and internal factors. These processes are the subject of the research. The information base is the Federal State Statistics Service's data on the values of the Gini index of 80 regions for the period from 1995 to 2018. To construct experimental trajectories, in addition to the Gini index, two independent dynamic variables were used – its first and second derivatives, which made it possible to construct 3 different spaces of states (from one-dimensional to three-dimensional). Using the k-means clustering method, the entire observed set of states was divided into 5 clusters, the number of which was previously determined by the “elbow method” test. As a result of the calculations, the predominance of convergent processes over divergent ones during the studied period was proven. It was found that the individual trajectories of certain regions in the space of state differ significantly: the trajectories of some regions can be localized within only one cluster, while some parts of the trajectories of others can belong to up to 4 clusters. The majority of the trajectories are located within 2–3 clusters. The theoretical significance of the results obtained lies in deepening the understanding of the regional specifics of the dynamic changes in the differentiation of income of the population of the constituent entities of the Russian Federation. The practical significance of the research results lies in the expansion of instrumental support for decision-making in the implementation of the state policy in the field of regulating the differentiation of income of the population at the regional level.

Key words: machine learning; clustering; theory of dynamic systems; space of states; experimental trajectories; territorial disproportions; differentiation of living standards; convergence/divergence of incomes.

JEL R1, C61

References

1. Sala-i-Martin, X. (1996). The Classical Approach to Convergence Analysis. *The Economic Journal*, Vol. 106, Issue 437, 1019–1036. <https://doi.org/10.2307/2235375>
2. Barro, R.J., Sala-i-Martin, X. (2004). *Economic Growth*. Second Edition. Cambridge, MIT Press. Available at: <http://piketty.pse.ens.fr/files/BarroSalaIMartin2004.pdf>

3. Kuznets, S. (1955). Economic Growth and Income Inequality. *American Economic Review*, Vol. 45, No. 1, 1–28. Available at: <https://www.aeaweb.org/aer/top20/45.1.1-28.pdf>
4. Malkina, M.Yu. (2015). Assessment of factors of intra-regional differentiation of incomes of the population of the Russian Federation. *Spatial Economics*, No. 3, 97–119. (In Russ.). <http://dx.doi.org/10.14530/se.2015.3.097-119>
5. Yurevich, M.A. (2019). Social Inequality, Investment, and Economic Growth. *Journal of Economic Regulation*, Vol. 10, No. 4, 35–46. (In Russ.). <https://doi.org/10.17835/2078-5429.2019.10.4.035-046>
6. Malkina, M.Yu. (2017). The relationship between types of inequality and indicators of living standards and well-being of the population in Russian regions. *Terra Economicus*, Vol. 15, No. 4, 46–63. (In Russ.). <https://doi.org/10.23683/2073-6606-2017-15-4-46-63>
7. Zubarevich, N.V., Safronov, S.G. (2019). People and money: incomes, consumption and financial behavior of the population of the Russian regions in 2000–2017. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Geograficheskaya*, Vol. 5, 3–17. (In Russ.). <https://doi.org/10.31857/S2587-5566201953-17>
8. Ovcharova, L.N., Popova, D.O., Rudberg, A.M. (2016). Decomposition of income inequality factors in modern Russia. *Journal of the New Economic Association*, No. 3, 170–185. (In Russ.). <https://doi.org/10.31737/2221-2264-2016-31-3-8>
9. Liubimov, I.L. (2017). Inequality and economic growth: challenges for the Russian economy. *Russian Journal of Entrepreneurship*, Vol. 17, No. 1, 11–22. (In Russ.). <http://dx.doi.org/10.18334/rp.17.1.2194>
10. Grigoriev, R.A., Kramin, M.V., Kramin, T.V., Timiryasova, A.V. (2015). Inequality in income distribution and economic growth in the regions of Russia in the post-crisis period. *Economy of Region*, No. 3, 102–113. (In Russ.). <https://doi.org/10.17059/2015-3-9>
11. Niea, H., Xingb, C. (2019). Education expansion, assortative marriage, and income inequality in China. *China Economic Review*, Vol. 55, 37–51. <https://doi.org/10.1016/j.chieco.2019.03.007>
12. Altunbas, Y., Thorntonab, J. (2019). The impact of financial development on income inequality: Aquantile regression approach. *Economics Letters*, Vol. 175, 51–56. <https://doi.org/10.1016/j.econlet.2018.12.030>
13. Liu, C., Jiang, Y., Xie, R. (2019.) Does income inequality facilitate carbon emission reduction in the US? *Journal of Cleaner Production*, Vol. 217, 380–387. <https://doi.org/10.1016/j.jclepro.2019.01.242>
14. Ibragimova, Z.F., Frants, M.V. (2018). Income inequality, its subjective perception and influence on the population's psychosocial well-being. *Statistics and Economics*, Vol. 15, No. 4, 52–60. (In Russ.). <http://dx.doi.org/10.21686/2500-3925-2018-4-52-60>
15. Afonina, T.N., Lomanov, A.O., Shumetov, V.G. (2018). Assessment of the quality of life of the population of the regions of Central Russia, taking into account the differentiation of monetary incomes. *Regional Economics: Theory and Practice*, Vol. 16, Issue 10, 1976–1994. (In Russ.). <https://doi.org/10.24891/re.16.10.1976>
16. Voinova, V.I. (2021). Approaches to decomposition of income inequality: differences and conditions of application. *Accounting and Statistics*, No. 3, 88–98. (In Russ.). <https://doi.org/10.54220/1994-0874.2021.63.3.010>
17. Repiská, R., Grisáková, N., Štetka, P. (2022). Hierarchical Clustering of EU Countries Based on HDI and EPI Index. *Proceedings from the EDAMBA 2021 Conference*, 420–430. <https://doi.org/10.53465/EDAMBA.2021.9788022549301.420-430>
18. Fahmiyah, I., Ningrum, R. (2023). Human Development Clustering in Indonesia: Using K-Means Method and Based on Human Development Index Categories. *Journal of Advanced Technology and Multidiscipline*, Vol. 2, No. 1, 27–33. <https://doi.org/10.20473/jatm.v2i1.45070>
19. Larasati, S., Nisa, K., Herawati, N. (2021). Robust Principal Component Trimmed Clustering of Indonesian Provinces Based on Human Development Index Indicators.

Journal of Physics: Conference Series, Vol. 1751, 012021. <http://dx.doi.org/10.1088/1742-6596/1751/1/012021>

20. Sihombing, S., Sembiring, R., Irawan, E. (2022). Application of the K-Means Clustering Algorithm in Grouping Regencies/Cities in North Sumatra Province Based on Human Development Index Indicators. *Jurnal Mandiri IT*, Vol. 11, No. 1, 20–25. <https://doi.org/10.35335/mandiri.v11i1.124>

21. Olimpia, N. (2019). Disparities Regarding Competitiveness, Human Capital and Inclusive Development in The EU: A Cluster Analysis. *Annals — Economy Series*, Vol. 1, 61–71. Available at: <https://econpapers.repec.org/RePEc:cbu:jrnlec:y:2019:v:1:p:61-71>

22. Kharlamova, G., Stavitskiy, A., Zarotiadis, G. (2018). The impact of technological changes on income inequality: the EU states case study. *Journal of International Studies*, Vol. 11, No. 2, 76–94. <https://doi.org/10.14254/2071-8330.2018/11-2/6>

23. Panzera, D., Postiglione, P. (2019). Measuring the Spatial Dimension of Regional Inequality: An Approach Based on the Gini Correlation Measure. *Social Indicators Research*, Vol. 148, 379–394. <https://doi.org/10.1007/s11205-019-02208-7>

24. Espoir, D.K. (2022). Convergence or divergence patterns in income distribution across countries: A new evidence from a club clustering algorithm. *Cogent Economics & Finance*, Vol. 10, Issue 1, 2025667. <https://doi.org/10.1080/23322039.2022.2025667>

25. Ogundari, K. (2023). Club Convergence in Income Inequality in Africa. *Social Indicators Research*, Vol. 167, 319–337. <https://doi.org/10.1007/s11205-023-03108-7>

26. Kydland, F.E., Prescott, E.C. (1982). Time to Build and Aggregate Fluctuations. *Journal of Econometrics*, Vol. 50, No. 6, 1345–1370. <https://doi.org/10.2307/1913386>

27. Dinopoulos, E., Segerstrom, P. (2010). Intellectual property rights, multinational firms and economic growth. *Journal of Development Economics*, Vol. 92, Issue 1, 13–27. <https://doi.org/10.1016/j.jdevco.2009.01.007>

28. Kenett, D.Y., Havlin, S. (2015). Network science: a useful tool in economics and finance. *Mind & Society*, Vol. 14, 155–167. <https://doi.org/10.1007/s11299-015-0167-y>

29. Tesfatsion, L. (2003). Agent-based computational economics: modeling economies as complex adaptive systems. *Journal of Information Sciences*, Vol. 149, Issue 4, 262–268. [https://doi.org/10.1016/S0020-0255\(02\)00280-3](https://doi.org/10.1016/S0020-0255(02)00280-3)

30. Berg, D.B., Popkov, V.V. (2003). General numerical model of the competition life cycle: from physics to economy. *Physica A: Statistical Mechanics and its Applications*, Vol. 324, Issue 1–2, 167–173. [https://doi.org/10.1016/S0378-4371\(02\)01892-7](https://doi.org/10.1016/S0378-4371(02)01892-7)

31. Berg, D., Demina, M., Isaichik, K., Panachev, A., Popkov, V., Parusheva, S. (2018). Competition of payment systems in Russia: Numerical analysis. *AIP Conference Proceedings*, Vol. 2040, Issue 1, 050014. <https://doi.org/10.1063/1.5079112>

32. Thorndike, R. (1953). Who Belongs in the Family? *Psychometrika*, Vol. 18, No. 4, 267–276. <https://doi.org/10.1007/BF02289263>

33. Eltibi, M.F., Ashour, W.M. (2011). Initializing K-Means Clustering Algorithm using Statistical Information. *International Journal of Computer Applications (IJCA)*, Vol. 29, 51–55. <https://doi.org/10.5120/3573-4930>

34. Likas, A.C., Vlassis, N.A., Verbeek, J.J. (2003). The global k-means clustering algorithm. *Pattern Recognition*, Vol. 36, Issue 2, 451–461. [https://doi.org/10.1016/S0031-3203\(02\)00060-2](https://doi.org/10.1016/S0031-3203(02)00060-2)

35. Kolomak, E.A. (2010). Inter-regional inequality in Russia: economic and social aspect. *Spatial Economics*, No. 1, 26–35. (In Russ.). <http://dx.doi.org/10.14530/se.2010.1.026-035>

36. Naiden S.N., Belousova A. V. (2018). Methodological tools to assess the population welfare: Interregional comparison. *Economy of Region*, Vol. 14, No. 1, 53–68. <http://dx.doi.org/10.17059/2018-1-5>

INFORMATION ABOUT AUTHORS

Dmitry Borisovich Berg

Doctor of Physics and Mathematics, Professor, Basic Department of Big Data Analytics and Video Analysis Methods, Institute of Radioelectronics and Information Technologies, Ural Federal University named after the first President of Russia B. N. Yeltsin, Yekaterinburg, Russia (620002, Yekaterinburg, Mira street, 19); ORCID <https://orcid.org/0000-0002-7703-9750> e-mail: bergd@mail.ru

Daniel Musafiri Balungu

Assistant, Basic Department of Big Data Analytics and Video Analysis Methods, Institute of Radioelectronics and Information Technologies, Ural Federal University named after the first President of Russia B. N. Yeltsin, Yekaterinburg, Russia (620002, Yekaterinburg, Mira street, 19); ORCID <https://orcid.org/0009-0001-5098-7603> e-mail: danielbal03.db@gmail.com

Andrei Gennad'evich Shelomentsev

Doctor of Economics, Professor, Chief Researcher, The Institute of Socio-Economic Research — A Separate Structural Unit of the Federal State Budgetary Scientific Institution of the Ufa Federal Research Center of the Russian Academy of Sciences (ISEI UFITs RAS), Ufa, Russia (450054, Ufa, Oktyabrya Avenue, 71); ORCID <https://orcid.org/0000-0003-1904-9587> e-mail: a.shelom@yandex.ru

Kseniya Sergeevna Goncharova

Candidate of Economic Sciences, Researcher, Institute of Economics, The Ural Branch of Russian Academy of Sciences, Yekaterinburg, Russia (620014, Yekaterinburg, Moskovskaya street, 29); ORCID <https://orcid.org/0000-0003-2381-3322> e-mail: ksenia.gon4arowa@gmail.com

ACKNOWLEDGMENTS

The study was financially supported by the Russian Science Foundation as part of a research project № 22-28-01702 «Experimental trajectories of convergence and divergence of Russian population's spatial income differentiation in the context of adaptation to dynamic changes».

FOR CITATION

Berg, D.B., Balungu, D.M., Shelomentsev, A.G., Goncharova, K.S. (2024). Experimental Trajectories of Convergence and Divergence Processes of Russian Regions Population Incomes Inequality. *Journal of Applied Economic Research*, Vol. 23, No. 2, 364–393. <https://doi.org/10.15826/vestnik.2024.23.2.015>

ARTICLE INFO

Received December 4, 2023; Revised April 10, 2024; Accepted April 24, 2024.

