УДК 510:514; 519.237.8

ФРАКТАЛЬНЫЙ ПОДХОД К АНАЛИЗУ ДАННЫХ В МОДЕЛЯХ МНОГОМЕРНОЙ КЛАССИФИКАЦИИ

Буховец Алексей Георгиевич,
доктор технических наук, доцент, профессор кафедры прикладной математики и математических методов в экономике;
abuhovets@mail.ru
Бирючинская Татьяна Яковлевна,
старший преподаватель кафедры прикладной математики и математических методов в экономике; abuhovets@mail.ru

В статье рассматриваются методы построения фрактальных множеств посредством систем итерируемых функций. Делается вывод о возможности использования фрактальных методов для выделения типичных объектов в задачах типологического анализа. Приводится практический пример, позволяющий оценить предложенный подход.
Ключевые слова: фрактальный анализ, классификационные модели.

Введение

В последние десятилетия внимание специалистов в области математического моделирования все больше привлекают социально-экономические системы. И если раньше основная роль в создании таких моделей принадлежала математикам, то в настоящее время кажется, что ситуация меняется зеркально – уже математики и физики стремятся найти (осмыслить и понять) новые идеи в процессах моделирования и исследования социально-экономических процессов. Сейчас уже практически можно считать общепризнанным, что именно эти системы являются наиболее сложным из всего, что известно науке. В процессах формирования и функционирования таких систем задействовано большое число факторов, как внешних, так и внутренних, их поведение представляет собой непростую смесь стохастических и детерминированных закономерностей. Вследствие этого при исследовании такого рода моделей и возникает надежда (возможно, призрачная) на то, что открытие общих законов эволюции сложных систем будет найдено именно в этой области. В подтверждение этого тезиса укажем, что в июне 2009 года в Москве прошел I первый Всероссийский конгресс по эконофизике (экономической физике). Появление новых математических методов и моделей позволяет по-новому взглянуть на уже ставшими традиционные подходы, и в свете этого получить новые интерпретации результатов.

Алгоритмы построения фрактальных множеств

Понятие фрактальных множеств (фрактальных структур) было введено сравнительно недавно, около 1975 г., американским математиком польского происхождения Б. Мандельбротом. Не вдаваясь в утомительные подробности строгого математического определения, мы лишь отметим, что для нашей работы вполне достаточно будет рассмотрения фрактальных объектов, представленных совокупностью точек некоторого признакового пространства. Такие пространства составляют методическую основу работ, где используются методы многомерного статистического анализа: факторного, кластерного или регрессионного. В этом случае фрактальное множество обычно представимо в виде совокупности точек, обладающей некоторыми свойствами, основными из которых являются дробная (фрактальная) размерность множества, меньшая, чем размерность признакового пространства, и наличие свойства самоподобия.

Большая часть работ Б. Мандельброта так или иначе связана с изучением фрактальных структур, их свойств, областей их приложения. В частности, приведем цитату из его широко известной книги [1, с.127]: «Я считаю совершенно необходимым – параллельно с продолжением попыток объяснить кластеризацию – найти способ описать ее и смоделировать реальность чисто геометрическими средствами (курсив автора книги- Б.М.)». Фактически здесь речь идет об одной важной проблеме современной науки – о необходимости не только описывать результаты классификации (кластеризацию), но и о необходимости моделировать эти структуры. Современный подход к исследованию каких-либо структур в качестве основного требования выдвигает построение описания процессов, порождающих эти структуры.

В задачах моделирования социально-экономических процессов, в большинстве случаев, производится приближение одного математического объекта другим. Например, задача регрессионного анализа может быть интерпретирована как задача наилучшей аппроксимации множества точек многомерного пространства некоторой гиперповерхностью, обычно, гиперплоскостью в случае построения множественной линейной регрессии. В классификационных задачах полученные классы аппроксимируются обычно функциями плотностей [4]. На наш взгляд было бы более естественно приближать одно множество точек $\{\mathrm{Y_i}\}$ другим множеством точек $\{\mathrm{X_j}\}$, особенно в том случае, когда свойства и характер поведения второго множества исследованы и могут быть достаточно хорошо воспроизведены.

В качестве аппроксимирующего множества нами предлагается использовать множества, которые возникают в результате выполнения итеративных процедур. Как будет показано в дальнейшем, за счет выбора значений набора $\{P_i\}$ и параметра μ, при заданных значениях $\{\mathrm{Z_i}\}$, имеется возможность моделировать структуры многомерных данных.

Для генерирования фрактальных структур предлагается использовать следующую процедуру [2]:

1. Пусть, в признаковом пространстве R^P задано некоторое множество точек $Z=\{Z_i\}$, называемое протофракталом. В этом же признаковом пространстве произвольным образом определяется начальная точка $X_0 \in R^P$. Для получения точек моделируемого множества последовательно выполняются следующие операции:

2. Случайным образом выбирается точка протофрактала, $Z_j^{(R)} \in Z$.

3. Вычисляются координаты точки $X_1=(x_{11}, x_{12}, \ldots, x_{1P})$ по формуле

$$x_{1i}=\frac{x_{0i}+\mu z_i^{(R)}}{1+\mu}, \quad (i=1,\ldots, P) \tag{1}$$

В дальнейшем полученная точка X_1 принимается за исходную, и пункты 2-3 повторяются столько раз, сколько точек протофрактала необходимо получить. В результате выполнения такой процедуры будет построено множество точек $E_n=\{X_1, X_2, \ldots, X_n\}$, которое будем называть предфракталом.

Рассмотренная выше процедура может быть отнесена к классу рандомизированных итерированных функциональных систем, т.н. RIFS (Random Iterated Function System), в которых множества, имеющие фрактальную структуру, порождаются детерминированными правилами, выполняемыми случайным образом. Полученное в результате выполнения этой процедуры множество называют аттрактором системы.

Построение аттрактора системы можно выполнить, изменив порядок действий. Обозначим $X_0=\{x_0\}$ начальную точку процедуры. Тогда

$$x_1=\frac{x_0+\mu z_1^{(R)}}{1+\mu}=\frac{1}{1+\mu}\left(x_0+\mu z_1^{(R)}\right)=\xi\alpha_0, \text{ где } \xi=\frac{1}{1+\mu},\ \alpha_0=x_0+\mu z_1^{(R)},$$

$$x_2=\frac{x_1+\mu z_i^{(R)}}{1+\mu}=\frac{1}{1+\mu}\left(x_1+\mu z_i^{(R)}\right)=\xi(\xi\alpha_0+\alpha_1), \text{ где } \alpha_1=\mu z_i^{(R)};$$

$$x_3=\frac{x_2+\mu z_i^{(R)}}{1+\mu}=\frac{1}{1+\mu}\left(x_2+\mu z_i^{(R)}\right)=\xi(\xi(\xi\alpha_0+\alpha_1)+\alpha_2), \text{ где } \alpha_3=\mu a_j^{(R)};$$

………………………………………………………..…………………

$$x_n=\xi(\ldots\xi(\xi(\xi\alpha_0+\alpha_1)+\alpha_2)+\ldots+\alpha_{n-1}).$$

Полученное выражение представляет собой запись схемы Горнера для многочлена степени p от переменной ξ без свободного члена с коэффициентами $\alpha_{n-1}, \alpha_{n-2}, \ldots, \alpha_0$, которые представляют собой случайным образом выбранные из совокупности точки протофрактала $Z=\{z_k\}_{k=1}^{K}$ на i-ом шаге, умноженные на коэффициент μ , т.е. $a_i=\mu z_i^{(R)}$, $i=1,\ldots,p$, $a_0=x_0+\mu z_1^{(R)}$. Другими словами имеет место следующее равенство

$$x_n = \alpha_0 \xi^n + \alpha_1 \xi^{n-1} + \ldots + \alpha_{n-1}\xi = \xi \sum_{m=0}^{n-1} \alpha_m \xi^{n-m-1}$$.

Если изменить порядок суммирования и нумерацию коэффициентов $\left(\beta_i = \alpha_{n-1-i}, i = 1, \ldots, n-1\right)$, то можно записать

$x_n = \xi \sum_{i=0}^{n-1} \beta_i \xi^i = \frac{1}{1+\mu} \sum_{i=0}^{n-1} \mu z_{(R)}^{n-1-i} \le \frac{\mu}{1+\mu} M \sum_{i=0}^{n-1} \xi^i$, где $M = \max_i \left| z_R^{(i)} \right|$.

Тогда, рассматривая x_n как частичную сумму ряда, обозначим

$$x^* = \lim_{n\to\infty} x_n = \frac{1}{1+\mu} \sum_{i=0}^{\infty} \beta_i \xi^i \le \frac{\mu}{1+\mu} M \sum_{i=0}^{\infty} \xi^i ,$$

и учитывая, что $\xi = \frac{1}{1+\mu}$, $0 < \xi < 1$, получаем, что все возможные значения x^* ограничены суммой бесконечно убывающей геометрической прогрессией.

Для значения x_n можно получить более точную оценку, перегруппировав члены ряда

$$x_p = \frac{1}{1+\mu} \sum_{i=0}^{p} \beta_i \xi^i = \frac{1}{1+\mu} (z_1 \sum_{i''} \xi^{i'} + z_2 \sum_{i''} \xi^{i'} + \ldots + z_K \sum_{i'''} \xi^{i''}),$$

или,

$$x^* = z_1 \frac{\mu}{1+\mu} \sum_{i'} \xi^{i'} + z_2 \frac{\mu}{1+\mu} \sum_{i''} \xi^{i'} + \ldots + z_K \frac{\mu}{1+\mu} \sum_{i'''} \xi^{i''} = \sum_{i=1}^{K} \alpha_i z_i , \qquad (2)$$

где $\alpha_i = \frac{\mu}{1+\mu} \sum_{i'} \xi^{i'}$, причем $\alpha_i > 0, \sum_{i=1}^{K} \alpha_i = 1$, а суммирование в отдельных слагаемых ведется по количеству выбранных в п.2 точкам протофрактала.

Следовательно, предельное значение x^* является, во–первых, конечным значением, а, во–вторых, представляет собой выпуклую комбинацию точек протофрактала $\mathbf{Z} = \{Z_i\}_{i=1}^{K}$. Это будет означать ограниченность множества получаемых точек, а учитывая, что построение проводится в конечномерном пространстве – его компактность.

Рассмотренное выше выполнение процедуры (2), позволяет представить ее в несколько ином виде:

1. Суммируется конечное число n первых случайно взятых слагаемых из n-ой частичной суммы ряда $\sum_{i=1}^{n} \frac{\mu}{(1+\mu)^i} \approx 1$ и вычисляются элементы матрицы $A = \left\| a_{ij} \right\|_{N\times K}$, где N – число требуемых точек предфрактала, K – число точек протофрактала. Правило отнесения i–го члена ряда к какому-либо слагаемому может быть, вообще говоря, произвольным.

2. В результате вычислений будет сформирована матрица A. После умножения на матрицу $Z = \left\| z_{kp} \right\|_{K\times P}$, т.е. выполнения операции $X = A * Z$,

будет получен список координат точек предфрактала X в заданном пространстве R^p. Таким образом, если сформирована матрица $A = \|\alpha_{ij}\|_{N \times K}$, то построение нового предфрактала, полученного за счёт изменения конфигурации точек множества $Z = \|z_{kp}\|_{K \times P}$ при сохранении его численности не вызывает особых затруднений.

Основные свойства фрактальных множеств

Можно показать, что полученное в результате выполнения процедуры (1-м или 2-м способом) множество точек En обладает следующими свойствами [3]:

1. Множество имеет нулевую лебеговскую меру. Практически это означает, генерируемое множество состоит из отдельных точек, т.е. представляет собой объединение одноточечных множеств.

2. Предельное множество X является совершенным, т.е. замкнуто и не содержит изолированных точек.

3. Множество X вполне несвязно (вполне разрывно). Такое множество представляет собой, объединение отдельных одноточечных множеств.

4. Предельное множество X может быть гомеоморфно отображено на канторово множество. Другими словами, получаемое множество топологически эквивалентно канторовому множеству.

5. Построенные множества имеют сингулярную функцию распределения. Фактически это означает, что строить оценку плотности распределения для таких множеств не имеет смысла.

6. Построенные множества являются самоподобными, т.е. существуют части множества (реплики), каждая из которых получается из целой части посредством преобразования подобия, в данном случае линейного сжатия (см (1)).

Теперь предположим, что некоторая совокупность данных была расклассифицирована. В ходе построения классификации были выделены классы $S_1, S_2, \ldots, S_K$, и получены их относительные частоты $P_1, P_2, \ldots, P_K$, где K – число классов. Будем предполагать, что имеет место фрактальная структура с постоянным значением параметра μ. В этом случае в соответствии с гипотезой о существовании указанных классов и их относительных численностях, формируется матрица A, число строк N которой равно числу классифицируемых объектов, а число столбцов K равно числу выделенных классов. Элементы строки матрицы $\grave{a}_{ij}$ $(i = 1, \ldots, N)$ вычисляются как суммы случайно отобранных членов ряда

$$\mu \sum_{\mathrm{i}=1}^{\infty} \xi^{\mathrm{i}} = \sum_{\mathrm{i}=1}^{\infty} \frac{\mu}{(1+\mu)^{\mathrm{i}}} = 1 . \tag{3}$$

Можно показать, что при таком построении вероятность отнесения члена ряда ξ^i в сумму a_{ij} является постоянной и не зависит от значения параметра μ, а также от вида закона распределения суммы a_{ij}.

После того, как будет сформирована матрица A, появляется возможность определения типичных объектов исходного множества X. Для этого необходимо решить уравнение $X=A*Z$ относительно Z. Решение $Z*$ может быть получено, поскольку матрица А имеет полный ранг по столбцам

$$Z^* = A^+ Y = (A^T A)^{-1} A^T Y \quad , \qquad (4)$$

где $A^+ = (A^T A)^{-1} A^T$ – матрица обобщенного преобразования Мура – Пенроуза (МП – преобразования [5]). Как известно, для любой матрицы A МП – матрица A^+ существует и единственна (см. [5, с. 59]).

После того, как будет определена матрица $Z*$, строки которой представляют оценки координат точек протофрактала, можно построить множество точек в пространстве R^P, необходимое для исследования структурных особенностей исследуемой системы или ее эволюции.

Представленная выше методика построения фрактальных моделей была апробирована нами при определении типичных объектов в классификационной задаче регионов России, а также при решении задач типологизации стран мировой экономической системы.

Использование фрактального подхода в социально-экономических исследованиях

Рассмотрим в качестве примера использование фрактального подхода задачу определения типичных представителей. Как известно, эта задача играет большую роль в построении классификационной (районированной) выборки, широко использующейся в социологических обследованиях.

В современной литературе различают два понятия типичных объектов [6]: 1) типичный объект рассматривается как предельное среднее. 2) типичный объект – как выражение предельного своеобразия объектов. В первом случае типичным является объект со свойствами, близкими по своей выраженности к среднему значению класса, во втором — с максимально выраженными свойствами. В случае если классы анализируются в рамках системного подхода, т.е. как некоторые паттерны целостной совокупности, то в качестве типичного объекта предпочтительнее взять объект, значения признаков которого максимально выражают свойства выделенного класса.

Работа по построению типологии была проведена совместно с Е.О. Лавровой – магистром ГУ-ВШЭ.

Исходя из анализа теоретических работ, посвященных типологии российских регионов были выбраны следующие блоки показателей:

- уровень жизни в регионе;
- инвестиционная активность;
- экономический потенциал.

При построении типологии, были получены статистические данные по 83 регионам [7]. Каждый регион характеризовался следующими восьмью показателями:

• доля населения с доходами ниже прожиточного минимума (индикатор бедности);

• отношение среднедушевых доходов к прожиточному минимуму;

• отношение среднедушевых расходов к прожиточному минимуму;

• доля инвестиций в валовом региональном продукте;

• темпы роста инвестиций в 2009 г. по отношению к среднероссийскому уровню на соответствующий период 2008 г.;

• отношение иностранных инвестиций к валовому региональному продукту;

• уровень безработицы в регионе (доля безработных от экономически активного населения региона);

• темпы роста валового регионального продукта по отношению к ВВП.

Все показатели, кроме прожиточного минимума (2010 г.) и доли инвестиций в основной капитал (2007 г.), были рассчитаны за 2008 г. (ввиду того, что большинство из них имеют привязку к ВРП, а последние значения ВРП, имевшиеся в нашем распоряжении, были за 2008 г.). Отобранные статистические данные были предварительно стандартизированы.

Для построения классификации регионов были использованы методы кластерного анализа: алгоритмы иерархической классификации и метод k-средних. Была использована невзвешанная евклидова метрика. Представленную структуру данных хорошо можно было представить по результатам метода k-средних. Исследование структуры данных дополнялось изучением проекций данных в пространстве первых 2-х главных компонент, которые представляли 67,4% обобщенной дисперсии. Краткое описание результатов классификационных построений представлены ниже.

Итоговое классификационное разбиение содержит 8 кластеров, среднее значение показателей кластеров приведено в таблице 1.

Кластер 1 (9 / 11) Московская область, <u>г. Санкт – Петербург</u>, республика Татарстан, Самарская область, **Свердловская область**, Тюменская область, Ханты – Мансийский автономный округ, Ямало-Ненецкий автономный округ, Челябинская область.

Кластер 2 (3 / 4) **Амурская область**, <u>Приморский край</u> и республика Якутия (Саха).

Кластер 3 (2 / 2) **республика Ингушетия** и **Чеченская республика**.

Кластер 4 (1 / 1) **Ненецкий автономный округ**.

Кластер 5 (27 / 33) Владимирская область, Воронежская область, Ивановская область, Костромская область, Рязанская область, республика Карелия, республика Калмыкия, Кабардино-Балкария, Карачаево-Черкессия, республика Марий Эл, республика Мордовия, <u>**Чувашская республика**</u>, Кировская область, Саратовская область, **Ульяновская область**, республика Алтай, республика Бурятия, республика Тыва, <u>республика Хакасия</u>, Алтайский край, **Забайкальский край**, Иркутская область, Томская область, Камчатский край, Хабаровский край, Магаданская область, Еврейская область.

Кластер 6 (7 / 8) Чукотский автономный округ, Сахалинская область, Вологодская область, Архангельская область, Липецкая область, **<u>Калужская область</u>**, Белгородская область.

Кластер 7 (33 / 40) Брянская область, Курская область, Орловская область, **<u>Смоленская область</u>**, Тамбовская область, Тверская область, Тульская область, Ярославская область, республика Коми, Калининградская область, Ленинградская область, Мурманская область, Новгородская область, Псковская область, республика Адыгея, Краснодарский край, Астраханская область, Волгоградская область, Ростовская область, республика Дагестан, республика Северная Осетия-Алания, Ставропольский край, республика Башкортостан, Удмуртская республика, **Пермский край**, Нижегородская область, Оренбургская область, Пензенская область, Курганская область, Красноярский край, Кемеровская область, Новосибирская область, Омская область.

Кластер 8 (1 / 1) **г. Москва.**

Жирным шрифтом выделены регионы, типичные в смысле близости к средним значениям показателей кластера. <u>Подчеркиванием</u> выделены регионы, типичность которых была установлена в рамках фрактального подхода – речь об этом пойдет ниже.

Обобщенные характеристики кластеров приведены в табл. 1.

Таблица 1

Средние значения показателей кластеров

№ кластера	Численности кластеров	Доля населения с доходами ниже прожиточного минимума	Отношение среднедушевого дохода к прожиточному минимуму	Отношение среднедушевых расходов к прожиточному минимуму	Темпы роста инвестиций в %% к 2008 г. по сравнению со среднерос. уровнем, 2009
1	9	10,59	3,25	2,55	96,27
2	3	20,67	1,71	1,31	196,86
3	2	36,10	1,10	0,52	169,99
4	1	10,20	1,74	1,23	43,44
5	27	21,19	1,69	1,36	100,56
6	7	12,49	2,29	1,83	96,80
7	33	15,08	2,23	1,88	109,75
8	1	10,00	4,17	3,84	92,84

Окончание табл. 1

№ кластера	Доля инвестиций в основной капитал в ВРП_2007	Отношение иностранных инвестиций к ВРП	Отношение темпов роста ВРП и ВВП	Уровень безработицы
1	27,80	0,07	0,03	6,46
2	35,47	0,05	0,01	9,07
3	70,05	0,00	0,00	43,95
4	92,70	0,37	0,00	9,70
5	27,07	0,02	0,00	10,81
6	34,69	0,27	0,01	6,59
7	26,88	0,03	0,01	9,17
8	11,50	0,12	0,20	2,70

Ниже приведены краткие характеристики выделенных кластеров.

1-й кластер условно можно назвать «самым благополучным», если исходить из высокого уровня жизни и экономического потенциала Ближе всего к кластерному центру находится Свердловская область.

2-й кластер составили регионы, которые можно определить как «потенциально неблагополучные». Эти регионы быстрыми темпами приближаются к «неблагополучным». Самый яркий представитель – Амурская область.

3-й кластер образуют два региона, которые условно можно назвать «самыми неблагополучными». Здесь самая высокая доля населения с доходами ниже прожиточного минимума, отношение среднедушевых доходов и расходов также самое низкое, причем отношение расходов к прожиточному минимуму меньше 1.

4-й кластер состоит из одного региона, который выделен в отдельный кластер из-за низких темпов роста инвестиций и высокой доли иностранных инвестиций.

5-й кластер один из самых многочисленных. Регионы этого кластера имеют довольно высокую долю населения с доходами ниже прожиточного минимума, низкие среднедушевые доходы и расходы. Темпы роста инвестиций в основной бюджет по сравнению со среднероссийским уровнем невысоки. Если сравнивать положение регионов данного кластера с другим, можно заметить, что он схож со 2-м кластером («потенциально неблагополучные»), за тем лишь исключением, что во втором кластере наблюдается другая структура распределения данных, и эта структура, видимо, обусловлена главным различием этих кластеров – высокими темпами роста во втором кластере. Эти регионы определяются как «неблагополучные». Самые яркие представители «неблагополучных» регионов – Забайкальский край, Чувашская республика и Ульяновская область.

6-й кластер составили регионы, которые условно можно определить как «потенциально благополучные». Самый типичный пример «потенциально благополучных» регионов – Калужская область.

7-й кластер самый многочисленный Регионы 7-го кластера занимают положение между «неблагополучными» и «потенциально благополучными» регионами, т.е. можно сказать, что это «средние» по уровню развития регионы без ярко выраженных тенденций перехода к неблагополучным или благополучным регионам. Среднедушевые доходы и расходы выше среднего уровня по России, но ниже, же у «благополучных» и «потенциально благополучных» регионов. Типичными представителями кластера служат Смоленская область и Пермский край.

В отдельный – 8-й – кластер выделен г. Москва. В Москве наблюдается самый высокий уровень жизни и экономического потенциала.

Для определения типичных объектов в рамках фрактального подхода была построена матрица A размером 83х8 при значении параметра процедуры, равного $\mu = 17{,}8$. Грубая оценка параметра μ была получена методом моментов. Окончательное значение параметра выбиралось экспертным путем.

С помощью матрицы A по формуле (4) были получены точки признакового пространства, представляющие протофрактал. Следует отметить, что при формировании данной матрицы необходимо передерживаться следующего правила [3;с.118]: максимальный элемент i –й строки должен находиться на k-м месте, если объект Y_i принадлежит классу S_{k}. Другими словами, элемент α_{jk} должен быть модой i-й строки генерируемой матрицы.

Каждая строка этой матрицы Z^* может быть соотнесена с ближайшим объектом классификации

$$Z = \begin{pmatrix} 10.276 & 3.313 & 2.59 & 94.907 & 27.899 & 0.065 & 0.035 & 6.253 \\ 20.638 & 1.688 & 1.289 & 200.858 & 34.995 & 0.046 & 0.006 & 8.455 \\ 37.262 & 0.442 & 0.164 & 173.095 & 71.949 & -0.003 & -0.001 & 45.915 \\ 9.955 & 1.709 & 1.191 & 39.852 & 96.355 & 0.389 & -0 & 9.738 \\ 21.437 & 1.667 & 1.335 & 100.208 & 26.923 & 0.023 & 0.002 & 10.876 \\ 12.33 & 2.296 & 1.827 & 96.794 & 35.277 & 0.279 & 0.007 & 6.367 \\ 15.02 & 2.23 & 1.883 & 109.989 & 26.779 & 0.029 & 0.007 & 9.177 \\ 9.695 & 4.28 & 3.953 & 91.895 & 10.636 & 0.125 & 0.211 & 2.321 \end{pmatrix}.$$

Точки выделенного протофрактала отражают структуру распределения объектов – регионов России – в пространстве выбранных для классификации признаков. Результаты вычислений для сравнения их со средними значениями приводится в табл. 2.

Таблица 2

Сравнение значений признаков средних значений (левая колонка) и типичных объектов, полученных в рамках фрактальной теории

№ клас-тера	Доля населения с доходами ниже прожиточного минимума		Отношение среднедушевого дохода к прожиточному минимуму		Отношение среднедушевых расходов к прожиточному минимуму		Темпы роста инвестиций в %% к 2008 г. по сравнению со среднероссийск. уровнем 2009 г.		Доля инвестиций в основной капитал в ВРП_2007	
1	10,59	10,29	3,25	3,3	2,55	2,59	96,27	95,03	27,80	27,81
2	20,67	20,87	1,71	1,69	1,31	1,29	196,86	201,2	35,47	35,98
3	36,10	36,64	1,10	0,55	0,52	0,21	169,99	174,1	70,05	70,29
4	10,20	9,62	1,74	1,75	1,23	1,23	43,44	35,51	92,70	95,5
5	21,19	21,35	1,69	1,67	1,36	1,34	100,56	100,3	27,07	27,04
6	12,49	12,09	2,29	2,31	1,83	1,84	96,80	96,16	34,69	34,78
7	15,08	15,02	2,23	2,23	1,88	1,88	109,75	109,6	26,88	26,8
8	10,00	9,75	4,17	4,27	3,84	3,94	92,84	91,99	11,50	10,71

Окончание табл. 2

№ клас-тера	Отношение иностранных инвестиций к ВРП		Отношение темпов роста ВРП и ВВП		Уровень безработицы	
1	0,065	0,067	0,033	0,034	6,46	6,17
2	0,047	0,049	0,006	0,006	9,07	8,97
3	0,00	0,001	0,001	0,001	43,95	44,91
4	0,371	0,386	0,002	0,003	9,70	9,65
5	0,025	0,023	0,003	0,002	10,81	10,87
6	0,266	0,279	0,006	0,007	6,59	6,19
7	0,029	0,029	0,007	0,007	9,17	9,17
8	0,121	0,125	0,204	0,21	2,70	2,37

Выводы и заключение.

Ближайшие к точкам протофрактала объекты кластеров в некоторых случаях отличаются от объектов, ближайших к кластерным центрам (см. табл. 2). Так, самым ярким представителем «самых благополучных» регионов с точки зрения фрактального подхода является г. Санкт-Петербург (была Свердловская область), «потенциально благополучных» – Приморский край (была Амурская область), «неблагополучных» – республика Хакасия и Чувашская республика (был Забайкальский край и Чувашская республика). В свою очередь, в кластерах «потенциально благополучных» и «средних» типичные представители не изменились (Калужская и Смоленская области соответственно).

Как можно видеть при сравнении таблиц, в ходе интерпретации кластеров в рамках фрактального подхода складывается следующая ситуация: значения точек протофрактала, соответствующие кластерам, где средние значения высокие, становятся еще выше, а в тех кластерах, где они низкие – еще ниже. Так, в данной задаче, по совокупности показателей типичным представителем благополучных регионов в социально-экономическом плане, будет тот регион, который должен обладать стабильно растущим доходом, развитыми рынками и качественным жильем, а так же занимать лидирующее положение по остальным показателям. Ситуация с неблагополучными районами, противоположная. Представителем данного класса будет регион с набором самых низких показателей. В данном экономическом неравенстве, среднюю позицию будет представлять регион с средними значениями показателей. Однако, такая упрощенная ситуация при анализе классификационного разбиения складывается не часто. Обычно кластеров бывает больше трех, и главной задачей аналитика при содержательной интерпретации является сравнение кластеров между собой, т.е. определение отношения кластеров друг к другу, их взаимосвязи. В нашем случае использование фрактального подхода позволяет оценить различия в структуре кластеров, а также обнаружить скрытые закономерности в общей структуре, и оценить, насколько «потенциальные» регионы отличаются от более однородных.

Как видно, различия в положении средних выборочных значений кластеров и вычисленных значений признаков точек протофракалов связано с тем положением, которое занимает кластер во всей исследуемой системе объектов.

В заключение еще раз отметим, что рассмотрение (исследование) эмпирических данных с точки зрения фрактального подхода позволяет практически реализовать метод выделения типичных объектов, определяемых как объекты, наиболее ярко и полно выражающих свойства классов.

Список источников

1. Мандельброт, Б. Фрактальная геометрия природы [текст] / Б. Мандельброт. – М., 2002. – 656 с.

2. Буховец, А.Г. Моделирование фрактальных структур данных [текст] / А.Г. Буховец, Е.А. Буховец // Системы управления и информационные технологии. – 2008. – № 3(33). – С. 4 – 7.

3. Буховец, А.Г. Использование фрактальных моделей в задачах классификации [текст] / А.Г. Буховец, Т.Я. Бирючинская, Е.А. Буховец // Системы управления и информационные технологии. – 2009. – 3.1 (37), С. 117 – 121.

4. Буховец, А.Г. Математические модели и методы типологического анализа. – Социологические методы в современной исследовательской практике. Материалы Всероссийской научной конференции [текст] / А.Г. Буховец. – М., 2007. – С. 16 – 31.

5. Магнус, Я.Р. Матричное дифференциальное исчисление с приложениями к статистике и эконометрике [текст] / Я.Р. Магнус, Х. Нейдеккер. – М.: ФИЗМАТЛИТ, 2002. – 495 с.

6. Ганзен, В.А., Фомин А.А. О понятии типа в психологии [текст] / В.А. Ганзен, А.А. Фомин // Вестник СПбГУ. – 1993. – вып. 1. – №6.

7. Федеральная служба государственной статистики [электронный ресурс]. – URL: http://www.gks.ru.

FRACTAL APPROACH TO DATA ANALYSIS IN MODELS OF MULTIDIMENSIONAL CLASSIFICATION

Bukhovets Aleksey Georgievich,
Dr. Sc. Of Technical Science, Associate Professor, Professor of the Chair of Applied Mathematics and Mathematical Methods in Economy of Voronezh State Agricultural University; abuhovets@mail.ru
Biruchinskaya Tatyana Yakovlevna,
Senior Lecturer of the Chair of Applied Mathematics and Mathematical Methods in Economy of Voronezh State Agricultural University; abuhovets@mail.ru

The construction methods of fractal sets by means of systems of iterated functions are consideredin the article. The conclusion about use possibility fractal methods for detection of typical objects in typological analysis is made. The practical example is resulted, allowing to estimate the offered approach.
Keywords: fractal analyses, multidimensional classification.