

ПРАКТИЧЕСКИЙ ОПЫТ

ОПЫТ ПРИМЕНЕНИЯ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА К ДАННЫМ О СОЦИАЛЬНО-ЭКОНОМИЧЕСКОМ РАЗВИТИИ ТЕРРИТОРИАЛЬНЫХ ОБЪЕКТОВ

Миркин Б. Г., Панфилова В. В.

(Москва)

В работе рассмотрена методика анализа многомерной слабо структурированной табличной информации о территориальных объектах. Эта методика основана на выделении групп признаков, отражающих укрупненные аспекты функционирования, и построении типологий объектов по каждой из этих групп с последующим сравнением типологий в статике и динамике. Получены некоторые выводы на реальных данных.

Длительная направленность хозяйственного механизма Московской области на игнорирование социальных приоритетов развития послужила первопричиной возникновения негативных тенденций в процессах воспроизводства населения. Это обуславливает необходимость целенаправленной социально-демографической политики в регионе. Однако ее выработка и реализация на практике сталкиваются с таким количеством непреодолимых проблем, что в настоящее время, пожалуй, единственный эффективный инструмент демографического регулирования в руках региональных органов управления, по крайней мере в Московской области, — привлечение дополнительных трудовых ресурсов. Понятно, что эффективность этого инструмента весьма условна, только в рамках «латания дыр». Что касается долговременных последствий использования привлекаемой рабочей силы, эффект также велик, но, как известно, имеет негативный характер.

Возникает идея проанализировать связи показателей естественного прироста населения с характером производства, вовлеченностью в него населения (трудовое поведение), развитием социальной инфраструктуры и т. п. с тем, чтобы, воздействуя на управляемые параметры, добиться желаемого развития демографической ситуации в регионе.

Для реализации этой идеи на материале Московской области были собраны данные о 57 территориальных объектах (ТО, все 39 районов и 18 находящихся в них относительно крупных городов) по 47 показателям. В число последних включены традиционные статистические индикаторы (коэффициенты рождаемости, смертности, брачности, прибытия, выбытия, фондоотдачи и фондовооруженности, доли детей, пенсионеров, женщин и специалистов среди занятых, семей в отдельных квартирах и общежитиях, экологически вредных выбросов, ручного труда, работающих в промышленности и сельском хозяйстве, детей, охваченных сетью дошкольных учреждений; объемы (на 1 жителя) капитальных вложений непроизводственного назначения, розничного товарооборота, денежных доходов и т. п.). Данные собирались за 2 периода: текущий (1988 г.) и прошлый (1979 г.) для сопоставительного анализа и прогнозирования.

Оказалось, что традиционные методы корреляционно-регрессионного анализа к полученной информации неприменимы. Какие-либо устойчивые зависимости показателей (не вытекающие из их определения, как, скажем, «коэффициент естественного прироста равен разности коэффициентов рождаемости и смертности») отсутствуют. В частности, несколько сотен попыток установления нетривиальных регрессионных линейных зависимостей (до 5—7 факторных переменных) не дали коэффициентов детерминации, больших 0,3—0,4, что, по нашему опыту, свидетельствует и об отсутствии сколь-нибудь выраженных нелинейных зависимостей.

В этих условиях была использована методика перехода от указанных показателей к «укрупненным аспектам» социально-экономического развития ТО с помощью построения разбиений множества объектов на классы, однородные по показателям, выражающим отдельные «укрупненные аспекты», и последующего анализа этих разбиений.

Дальнейшее представляет собой описание методики и некоторых результатов ее применения к данным о 57 ТО Московской области.

Методика использования многомерной группировки территории включает ряд этапов. Остановимся на этих этапах более подробно, сопровождая их примерами расчетов на реальных данных по ТО.

1. ВЫДЕЛЕНИЕ ГРУПП ПРИЗНАКОВ

Совокупность показателей состояния рассматриваемых объектов разбивается на группы, которые должны отражать отдельные укрупненные аспекты функционирования ТО. Цель такого разбиения — переход от исследования зависимостей между отдельными показателями к изучению связей между выделенными аспектами.

В данном случае было решено объединить показатели в 9 групп, в той или иной мере выражающих следующие укрупненные аспекты социально-экономического состояния ТО:

1. Естественный прирост (5 показателей).
2. Миграция (3).
3. Демография (7).
4. Трудовое поведение в народном хозяйстве (9).
5. Производство (5).
6. Экономика непроизводственной сферы (6).
7. Социальная инфраструктура (7).
8. Обеспеченность жилищем (3).
9. Экология (2).

Теперь следует сформировать типологии — группировки (разбиения) ТО по каждой группе показателей. Эти группировки должны репрезентировать укрупненные аспекты социально-экономического состояния и в этом качестве анализироваться.

2. ФОРМИРОВАНИЕ ТИПОЛОГИИ

1. Применение метода кластер-анализа. Этот метод построения многомерных типологий основан на формировании групп объектов, близких друг к другу в пространстве показателей, отобранных исследователем. Комбинации значений показателей, типичные для каждой группы, характеризуют ее как «тип», т. е. элемент конструируемой типологии. Это предопределяет выбор в качестве инструмента группировки так называемых «ядерных алгоритмов» кластер-анализа, позволяющих одновременно формировать и группы (кластеры), и типы (ядра) [1]. До последнего времени использование ядерных (как, впрочем, и других) алгоритмов

кластер-анализа было сопряжено с необходимостью априорного задания ряда экспертных характеристик, прежде всего числа групп и «предварительных претендентов» на роль их ядер. В настоящее время имеются модификации, работающие без использования этой информации [2]. Особенно обещающим для нашей работы является метод главных кластеров, основанный на специальной модели, представляющей исходные данные в терминах групп и их ядер [3], что открывает новые возможности их интерпретации и использования.

Обозначим через x_{ik} значение признака k , $k=1, \dots, n$, на объекте i , $i=1, \dots, N$, а через y_{ik} результат его стандартизации: $y_{ik} = (x_{ik} - \bar{x}_k) / \sigma_k$, где $\bar{x}_k = \sum_i x_{ik} / N$ — среднее значение, а $\sigma_k = (\sum_i (x_{ik} - \bar{x}_k)^2 / N)^{1/2}$ — средне-квадратическое отклонение признака k . Обозначим через S_t один из искомым кластеров (групп) объектов, $t=1, \dots, m$, причем число m кластеров может быть заранее не задано. Поставим в соответствие кластеру S_t два вектора: n -мерное ядро $c_t = (c_{tk})$ и N -мерный вектор принадлежности Z_t , где $z_{it} = 1$, если объект i принадлежит кластеру S_t , и $z_{it} = 0$ в противном случае.

В этих терминах модель главных кластеров может быть сформулирована как система равенств вида

$$y_{ik} = \sum_t c_{tk} z_{it} + \varepsilon_{ik}, \quad (1)$$

где ε_{ik} — «малые» невязки данного представления.

Согласно (1), стандартизованные значения признаков на произвольном объекте с точностью до невязок совпадают со значениями параметров кластера S_t , содержащего данный объект. Эти параметры c_{tk} удобно называть нагрузками признаков k на кластеры S_t (поскольку модель (1) с точностью до ограничений на значения z_{it} совпадает с моделями факторного анализа, где принята именно эта терминология).

Метод наименьших квадратов для (1) требует отыскания таких c_t , Z_t , которые минимизируют сумму квадратов невязок. Нетрудно видеть, в этом случае задача состоит в таком распределении объектов по кластерам S_t , при котором минимизируются квадраты евклидовых расстояний между векторами, представляющими данные объекты i , и центрами c_t , что при априорной неперекрываемости кластеров эквивалентно известному критерию минимизации средневзвешенной дисперсии признаков внутри кластеров [1]. Метод главных кластеров основан на последовательном исчерпании элементов y_{ik} суммой «кластерных характеристик» $c_{tk} z_{it}$ в (1). А именно, на каждом шаге t , $t=1, 2, \dots$, рассматриваются «остаточные» данные $Y_t = (y_{ik}^{(t)})$ (вначале используются исходные, т. е. $Y_1 = Y$), и для них решается задача отыскания единственного кластера (c_t, Z_t) по критерию минимизации суммы квадратов невязок. Иными словами, решается задача минимизации величины

$$\sum_{i,k} (y_{ik}^{(t)} - c_{tk} z_{it})^2$$

по всевозможным c_k и булевским z_i , а результат ее (не обязательно точного) решения $c = (c_k)$ и $Z = (z_i)$ принимается за искомые c_t и Z_t соответственно. Затем вычисляются следующие остаточные данные по формуле $y_{ik}^{(t+1)} = y_{ik}^{(t)} - c_{tk} z_{it}$, вытекающей из (1); t увеличивается на единицу, и процесс повторяется заново. При этом ни количество классов, ни исходные центры могут не задаваться заранее, а определяться по результатам расчетов.

Таблица 1

Характеристики	Кластер									
	1	2	3	4	5	6	7	8	9	10
Численность	6	16	13	6	5	2	5	1	2	1
Вес, %	38,4	21,4	14,2	2,2	2,0	0,4	0,4	0,1	0,1	0,01
Стандартизованные значения признаков										
1	2,3	-1,0	0,3	0,2	-0,4	0,3	-0,3	0,3	-0,1	0,1
2	1,2	-1,2	1,0	-0,3	0,0	0,5	-0,3	0,1	0,1	0
3	2,1	-0,1	-0,9	0,7	-0,7	-0,1	-0,1	0,2	-0,2	0,1

Задача отыскания оптимального кластера на каждом шаге указанного процесса имеет сложный комбинаторный характер. Для ее решения могут использоваться локально-оптимальные алгоритмы. Простейший из них, как показано в [3], является уточненной модификацией хорошо известного алгоритма «спектр» [4], согласно которому любой кластер формируется (начиная с некоторого объекта) путем последовательных присоединений объектов, в среднем наиболее близких к уже отобранному подмножеству. Уточнение относится к выбору начального объекта (самый дальний от начала координат) и к критерию остановки процесса присоединений (когда расстояние от объекта до центра кластера станет больше, чем до начала координат).

Соотношения (1) позволяют получить более тонкие средства интерпретации, чем простое перечисление элементов и ядер кластеров, даваемое традиционными кластерными алгоритмами. Справедливо следующее разложение квадратичного разброса данных [3]

$$\sum_{i,k} y_{ik}^2 = \sum_t \sum_k c_{tk}^2 \sum_i z_{it}^2 + \sum_{i,k} e_{ik}^2, \quad (2)$$

с помощью которого можно оценить вклады кластеров, а также признаков в кластерную структуру, причем величины $c_{tk}^2 \sum_i z_{it}^2$ задают весовые характеристики признаков для каждого кластера S_t .

Применение метода главных кластеров полезно, когда, как и здесь, нет никаких априорных соображений о конкретных типах объектов. Конечно, это не означает, что априорные соображения не используются вовсе. Модель главных кластеров подразумевает, во-первых, что объекты каждого кластера группируются «вокруг» его центра (в пространстве нормированных признаков) и, во-вторых, что порядок «отщепления» кластеров соответствует их вкладу в квадратичный разброс данных (так что последние из них представляют собой образования, несущественно отличающиеся от средних на всем массиве значений).

По каждой из 9 групп переменных на массивах 1979 и 1988 гг. с помощью метода главных кластеров были получены группировки объектов, включавшие 7—10 кластеров. Например, по трем показателям миграции (1 — коэффициент прибытия, 2 — коэффициент выбытия, 3 — сальдо миграции), входившим в группу 2, для данных 1988 г. выявлено 10 кластеров (см. табл. 1).

Здесь вес — это относительный вклад кластеров в квадратичный разброс данных, рассчитанный согласно формуле (2), а стандартизованные

Таблица 2

Характеристики	Кластер				
	1	2	3	4	5
Численность	8	21	15	6	7
Значения признаков					
1	1,7 48,4	-0,8 27,0	0,3 36,1	0,2 35,6	-0,3 30,9
2	0,9 32,8	-1,0 20,6	1,0 33,2	-0,3 24,9	0 27,0
3	1,6 15,6	-0,1 6,4	-0,8 2,9	0,7 10,7	-0,6 3,9

значения признаков — средние величин y_{ik} в каждом отдельном кластере, т. е. не что иное, как нагрузки c_{ik} .

Из таблицы видно, что ядро каждого следующего кластера ближе к центру тяжести всего множества объектов, чем у предыдущих. Это свойство позволяет контролировать «выбросы» в данных, проявляющиеся в том, что на первом шаге образуется кластер, содержащий 1—2 объекта, далеко отстоящие от всего «облака объектов». Иногда это отражает реальность, но чаще объясняется ошибкой информации.

Следует отметить еще одну особенность алгоритма главных кластеров, использовавшуюся в расчетах: как и в методе главных компонент, решение не инвариантно к выбору масштаба измерения признаков. Поэтому исходные данные предварительно приводятся к единому масштабу с помощью нормировки — деления значений признаков на их среднеквадратические отклонения. При этом неоправданно большой вес получают показатели с невысоким уровнем разброса, который может объясняться случайными причинами. Мы исключили те из них, для которых величина коэффициента вариации не превышала 10%, как, например, общий коэффициент рождаемости, равный 13,07 в 1979 г. (коэффициент вариации 7%) и 12,99 в 1988 г. (коэффициент вариации 10%).

2. *Укрупнение кластеров.* Как правило, метод главных кластеров дает слишком большое, с точки зрения наших целей, число кластеров. Поэтому надо объединить «близкие» из них так, чтобы довести их число до привычных 4—5 групп. В зависимости от задачи исследования это можно делать по-разному. Например, объединить последние кластеры как несущественно отличающиеся от центра тяжести всего множества в группу «средних» объектов. В данном случае было решено пойти по пути присоединения этих последних кластеров к наиболее сходным с ними «весомым» первым группам. Сходство определялось по наиболее «весомым» для кластеров показателям.

Так, в рассматриваемом примере кластер 9 имеет то же распределение нагрузок c_{ik} , что и кластер 5. В результате присоединения кластеров 8, 10 к кластеру 1, 7 к 2, 6 к 3, 9 к 5 получаем 5 групп (см. табл. 2).

В таблице указаны не только нагрузки, но и соответствующие значения признаков в реальном масштабе (в данном случае — в %).

3. *Описание типов.* Описание типов производится так же, как и интерпретация факторов в факторном анализе — с использованием нагрузок сформированных групп, причем весомость нагрузки определяется квадратом ее величины (см. (2)).

Типы миграционного поведения, зафиксированные в 5 классах рассматриваемого примера, могут быть описаны следующим образом (согласно табл. 2): А — высокий уровень всех трех показателей, особенно 1

и 3, Б — низкие значения показателей прибытия 1 и выбытия 2, В — высокий уровень коэффициента выбытия 2, низкий — сальдо 3, Г (кластер 5) — низкий уровень приживаемости 3 (остальные — на среднем уровне), Д (кластер 4) — относительно высокий уровень приживаемости 3 за счет пониженного коэффициента выбытия 2.

3. СРАВНЕНИЕ ТИПОЛОГИИ В СТАТИКЕ

1. *Таблица сопряженности.* Полученные на предыдущем этапе разбиения множества объектов по всем группам показателей (для конкретного периода) позволяют проанализировать попарные связи между выделенными аспектами рассматриваемого явления с помощью комбинационных группировок. Иными словами, для любых двух разбиений, S и S' , данного множества объектов, состоящих из классов S_i и $S_{i'}$, вычисляются численности $N_{ii'}$ их пересечений $S_i \cap S_{i'}$ и (при необходимости) средние значения признаков в группах $S_i \cap S_{i'}$. Указанные величины сводятся в таблицу сопряженности, по которой можно судить о степени взаимосвязи разбиений. Так, например, таблица сопряженности разбиений S и S' , полученных по показателям естественного прироста и демографии соответственно, имеет вид табл. 3 для данных 1979 г. и табл. 4 для 1988 г.

2. *Оценка связи между разбиениями.* Интуитивно ясно, что связь между разбиениями по табл. 3 выше, чем по табл. 4. Действительно, в первом случае 10 из 13 объектов первого, 11 из 13 второго, 8 из 12 третьего класса демографии попали в один и тот же класс естественного прироста (см. первые три строки табл. 3), тогда как в табл. 4 демографические классы распределились по классам естественного прироста значительно менее «кучно». Иными словами, из табл. 3 и 4 только первая позволяет использовать классы демографии для прогноза класса естественного прироста.

Для количественной оценки связи между разбиениями удобно применить модифицированный коэффициент Пирсона

$$\Phi = \frac{1}{m-1} \sum_{i,i'} (p_{ii'} - p_i p_{i'})^2 / p_i p_{i'}, \quad (3)$$

который выражает средний относительный уровень улучшения качества описания $(p_{ii'}/p_i p_{i'}) - 1$ классов i' одной классификации классами i другой [5]. Здесь использованы следующие обозначения: $p_{ii'}$ — доля объектов, одновременно попавших в класс S_i первого и класс $S_{i'}$ второго из разбиений, $p_i (p_{i'})$ — доля класса $S_i (S_{i'})$, m — общее количество классов S_i . Рассматривалась также аналогичная мера, предложенная в [5], как средний оптимальный относительный уровень улучшения качества описания

$$v = \frac{1}{m-1} \sum_i \max_{i'} \frac{p_{ii'}}{p_i} - 1.$$

Оказывается, для табл. 3 $\Phi = 0,224$, а для табл. 4 — $\Phi = 0,058$. Согласно правилам интерпретации в терминах модельной прогнозной ситуации [5, с. 81—82], первая величина позволяет прогнозировать классы одного разбиения по значениям второго с точностью 60%, а вторая — 40%. Несколько большие значения принимает мера v , дающая в табл. 3 и 4 соответственно 0,344 и 0,183.

3. *Анализ структуры связей между разбиениями.* Расчет меры Φ между всеми парами разбиений, выражающих укрупненные аспекты социально-экономического состояния территориальных объектов, привел к результатам, изображенным на рисунке. Здесь линиями соединены номера разбиений по тем укрупненным аспектам, величина связи которых

Таблица 3

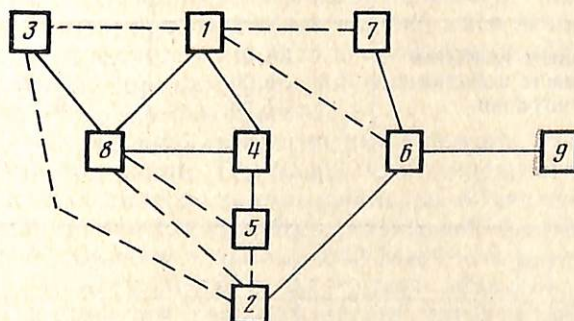
Классы <i>t</i>	Классы <i>t'</i>				
	1	2	3	4	Итого
1	10	0	3	0	13
2	0	11	1	1	13
3	0	8	3	1	12
4	4	0	2	2	8
5	2	1	6	1	10
Итого	16	20	15	5	56

Таблица 4

Классы <i>t</i>	Классы <i>t'</i>					Итого
	1	2	3	4	5	
1	2	0	6	0	0	8
2	3	2	7	7	2	21
3	2	5	3	2	3	15
4	1	1	2	1	1	6
5	1	0	2	3	1	7
Итого	9	8	19	14	7	57

по мере Φ превышала 0,1, причем сплошные линии соответствуют связям, которые были и в 1979, и в 1988 г., а штриховые — связям только 1979 г. По сравнению с 1979 г. никаких новых связей не добавилось, зато очень много «старых» исчезло. Картина 1979 г. более или менее соответствует интуитивно ожидавшимся связям между укрупненными аспектами социально-экономической активности. В частности, естественный прирост имеет непосредственное отношение ко всем рассматриваемым характеристикам, кроме производства, влияющего через социальную инфраструктуру и трудовое поведение населения, и миграции, также косвенно связанной с ним. Картина 1988 г. выглядит существенно более куцей. Исчезли связи миграции с производственным и демографическим блоками, естественный прирост теперь непосредственно соотносится только с социальной структурой и трудовым поведением, да и характер этих взаимоотношений иной. Так, из 19 объектов класса 3 разбиения по естественному приросту 12 объектов попадает в группу 2 разбиения по признакам социальной инфраструктуры, характеризующуюся низким уровнем всех семи показателей, особенно доли детей в детских учреждениях. Вообще, средняя доля детей (65,7%), посещающих детские учреждения, в этих территориальных объектах значительно ниже средней, равной 74%.

В целом сравнение структур связи рассматриваемых девяти укрупненных аспектов в 1979 и 1988 г. показывает, что в истекшем периоде зона действия естественных социально-экономических закономерностей оказалась суженной, проявилась тенденция к распаду естественного хода социального развития. Вероятно, сыграло свою роль увлечение «завозом»



Структура связей укрупненных характеристик социально-экономического развития ТО: 1 — естественный прирост, 2 — миграция, 3 — демография, 4 — трудовое поведение в народном хозяйстве, 5 — производство, 6 — экономика непродовольственной сферы, 7 — социальная инфраструктура, 8 — обеспеченность жилищем, 9 — экология

рабочей силы извне с соответствующим ослаблением работы по территориальному развитию.

Рассматривался также вопрос о связи построенных классификаций с разбиением территорий области по их удаленности от Москвы, часто используемым для социального и экономического управления. Оказалось, что эти связи в целом довольно слабы. Так, по данным 1988 г. связь, несколько большая 0,1, была зафиксирована только для трех групп признаков: производства, трудового поведения и социальной инфраструктуры. Это означает, что управление социально-демографическими процессами требует учета более тонких факторов, определяющих построенные разбиения.

4. АНАЛИЗ ДИНАМИКИ ПРОЦЕССА

1. Сравнение соответствующих разбиений. Данные 1979 и 1988 гг. обрабатывались независимо друг от друга. Для сопоставления полученных результатов были построены таблицы сопряженности разбиений по каждому отдельному укрупненному аспекту, полученные на этих двух массивах, и оценена их близость с использованием модифицированного коэффициента сопряженности (3). Оказалось, что соответствующие друг другу разбиения, хотя и отличаются, но все-таки достаточно близки. Так, связь разбиений по признакам миграции оказалась равной 0,231.

Различия в разбиениях определяются изменением характера миграционных процессов, о чем свидетельствует табл. 5.

Миграция за 9 лет в целом снизилась, хотя уровень сальдо остался почти на прежнем уровне (правда, стал более равномерным: нет случаев отрицательного сальдо, довольно частых в 1979 г.).

Подобные изменения произошли и для других показателей.

Связи между соответствующими друг другу классификациями двух массивов представлены в табл. 6.

Как видно, наиболее стабильными элементами региональной структуры являются в данном случае не демографические, а производственные аспекты: максимально сохранились разбиения по экологическим показателям производства (0,508), строительству жилищ (0,407), самому производству (0,337).

Таблица 5

Признак	Средние	
	1979	1988
Коэффициент прибытия	38,91	33,81
Коэффициент выбытия	31,00	26,85
Сальдо миграции	7,95	6,95

Таблица 6

Аспект	Связь	Аспект	Связь
Естественный прирост	0,171	Экономика непроизводственной сферы	0,122
Миграция	0,231	Социальная инфраструктура	0,245
Демография	0,158	Жилище	0,407
Трудовое поведение	0,281	Экология	0,508
Производство	0,337		

2. *Метод стабильных групп для сравнительного анализа типов.* Разбиение по миграции на данных 1988 г. рассмотрено в п. 2.3. В 1979 г. получилось также 5 классов, типологическое описание которых совпадает с описанием классов 1988 г. Это позволяет использовать для сравнительного анализа типов группы объектов, оставшиеся в том же классе («стабильные группы»). Так, типу А в 1979 г. принадлежит 11, а в 1988 г. — 8 объектов, причем 5 из них стабильны. Для типа Б аналогичные данные — 19, 21, 12, для типа В — 10, 15, 8. Наполнение типов Г и Д полностью сменилось: 2 и 3 стабильных объекта соответственно (причем объекты типа Г перешли в основном в тип Б).

Таким образом, наличие стабильных групп объектов позволяет сделать сравнительный анализ типов А, Б, В по их «устойчивым» элементам (см. табл. 7, в которой приведены некоторые наиболее «контрастные» характеристики).

Как видно, признаки 4, 12, 33, 34, 39 убывают при переходе от А к Б, тогда как 18, 22, 30, 31 возрастают. Показатели рождаемости 4, товарооборота 31 и жилплощади 33 в этих группах несколько выше среднего.

С точки зрения занятости группа А имеет непроемчивую, группа В — сельскохозяйственную, группа Б — промышленную ориентации. Сравнение производительности труда 22 и фондоотдачи 23 в группах А—В соответствует этому.

Относительно высокий уровень капитальных вложений в непроемчивую сферу 28 и средств, направляемых на социальное развитие 29, выделенных группе В в последнее время (в 1979 г. величина последнего показателя для всех групп была одинаковой на уровне порядка 96 руб. на человека), не может компенсировать неблагоприятные миграционные тенденции в этой группе.

Группа А (в основном города, но далеко не все 18, а только 5) имеет сравнительно хорошую обеспеченность жилплощадью 33, относительно высокий уровень естественного прироста 4 (за счет низкой смертности), улучшенное медицинское обслуживание 39. Вероятно, эти факторы важнее для приживаемости, чем некоторое отставание в денежных доходах 30 и уровне товарооборота 31.

Подобный анализ позволяет проследить определенные черты типов, которые могут оказаться полезными для предплановых проработок. Вместе с тем, он показывает, что рассматриваемая система показателей неудобна с точки зрения задач управления социальными процессами в регионе: она фиксирует состояние, но не факторы управления.

3. *Прогноз.* Как известно, прогноз, особенно долгосрочный, должен базироваться либо на адекватной модели процесса, либо на достаточно «длинном» временном ряде. В данном случае, имея всего две временные точки, естественно предположить, что полученные разбиения задают достаточно адекватную (правда, не количественную, а структурную) модель, и на основе этого делать прогноз.

В частности, мы рассчитали прогнозные значения показателей исходя из следующих соображений. Положим, что тенденции динамики будут «уловлены» в рамках классов территориальных объектов, однородных по динамике рассматриваемой группы показателей. Точнее, для каждого территориального объекта возьмем коэффициенты изменения значений его показателей, а затем сформируем группы объектов, однородные по коэффициентам изменения тех показателей, которые принадлежат этому укрупненному аспекту социально-экономического развития. Для каждой такой однородной группы в качестве прогнозных на 1995 г. значений берем те, которые получаются из данных 1988 г., умноженных на соответствующие коэффициенты, возведенные в степень 7/9 (7 лет с 1988 до 1995 г., отнесенные к 9-летнему периоду). Понятно, что при подобном

Характеристики	Коэффициент естественного прироста, %	Средний размер семьи, чел.	Доля специалистов среди занятых, %	Доля пенсионеров среди занятых, %	Доля занятых в непроизводственной сфере, %	Доля занятых в промышленности, %	Доля занятых в сельском хозяйстве, %	Доля занятых в строительстве, %	Удельный вес ручного труда, %
Номер признака	4	12	15	16	17	18	19	20	21
Тип									
А	2,38	3,1	37,9	36,7	56,0	22,6	3,8	5,7	37,1
Б	0,47	2,9	30,1	31,2	23,9	42,9	9,7	5,6	30,4
В	0,99	2,9	23,4	25,6	19,8	28,0	24,7	9,1	38,8
Среднее	1,29	3,0	31,2	32,2	30,5	33,1	13,7	7,0	36,0

«прогнозе» показатели, заключенные в определенных границах (например доля детей в яслях не может превышать 100%), могут выйти из этих границ, так, что необходимо «отслеживать» подобные ситуации.

Поскольку степень адекватности нашей структурной модели неизвестна, оценка качества прогноза — нелегкое дело. Существует комплексная территориально-отраслевая программа интенсификации социально-экономического развития Московской области на тринадцатую пятилетку, разрабатывавшаяся в 1989 г. Прямое сопоставление программы и прогноза возможно по трем показателям: объем платных услуг, объем жилплощади на 1 человека и число больничных коек. Ясно, что наш прогноз экстраполирует имеющиеся тенденции, тогда как данные комплексной программы используют плановые наметки инвестиций, строительства и пр. Оказалось, что по всем районам расчетные значения жилплощади существенно ниже, а объема платных услуг — значительно выше плановых наметок. С учетом имеющихся тенденций отставания строительства и девальвации цен не исключено, что прогнозные значения будут точнее.

*
* *

Таким образом, методика позволяет улавливать закономерности, недоступные традиционным линейным методам корреляционного, регрессионного и факторного анализов в типичной ситуации отсутствия сколько-нибудь выраженных линейных связей между показателями объектов.

В частности, выявлена устойчивая связь трех аспектов социально-экономического развития: естественного прироста населения, трудового поведения и социальной инфраструктуры. Естественные связи между другими аспектами социально-экономического поведения населения в период 1979—1988 гг. обнаружили тенденцию к разрушению. Косвенно этот вывод подтверждается дальнейшим ухудшением тенденций демографического воспроизводства области в 1989 г. (например, впервые коэффициент естественного прироста стал отрицательным).

Проведенный анализ показал также полную невозможность использования рассматриваемой системы показателей для управления демографическими процессами. По-видимому, отраженные в ней привычные категории экономической статистики фиксируют статический «портрет» территориальных объектов, вовсе не ориентированный на применение в

Таблица 7

Производительность труда, руб/чел	Фондоотдача, руб/руб	Капитальные вложе- ния в непроизводст- венную сферу, руб на 1 жителя	Средства на социаль- ное развитие, руб. на 1 жителя	Денежные доходы населения, руб. на 1 жителя	Товарооборот, руб. на 1 жителя	Размер жилой площа- ди, кв. м. на 1 жите- ля	Доля детей, охвачен- ных дошкольными учреждениями, %	Число врачей, чел на 10000 жителей	Число предприятий торговли, шт. на 10000 жителей
22	23	28	29	30	31	33	36	39	41
14 510	3,3	126,2	108,9	1543	1222	14,9	69,8	40,6	13,6
21 419	9,8	114,8	136,2	1706	1253	13,4	77,1	30,4	14,2
20 018	1,0	170,9	143,1	1637	1230	13,9	65,4	31,1	24,9
19 739	3,5	122,4	129,3	1607	1184	13,2	73,7	33,5	16,0

практике регионального управления. Для целей управления необходимо иметь входные и выходные показатели тех процессов и явлений, из которых и складывается демографическое поведение. К таким процессам можно отнести те, которые с известной долей условности обозначаются словами: «материнство», «детство», «семья», «пожилое население», «здоровье» и пр. Перестройка системы показателей невозможна без соответствующего изменения как системы территориального управления, так и научно-методологического инструментария. Первое требует структуризации целей и средств управления территориальными объектами, второе — выработку адекватной модели. Полезность описанной методики при решении этих вопросов представляется несомненной.

ЛИТЕРАТУРА

1. Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: классификация и снижение размерности. М.: Финансы и статистика, 1989.
2. Мандель И. Д. Кластерный анализ. М.: Финансы и статистика, 1989.
3. Миркин Б. Г. Метод главных кластеров//Автоматика и телемеханика, 1987. № 10.
4. Браверман Э. М., Мучник И. Б. Структурные методы обработки эмпирической информации. М.: Наука, 1983.
5. Миркин Б. Г. Группировки в социально-экономических исследованиях. М.: Финансы и статистика, 1985.

Поступила в редакцию
14 V 1990