4. Причинный анализ

Существо проблемы причинного анализа можно представить следующим образом. Имеет место какое-либо социальное явление, которое характеризуется переменной ***x***. Оно зависит и причинно обусловлено другими данными социальными явлениями, характеризуемыми соответственно переменными ***y***, ***z*** и т. д. Требуется определить степень этой зависимости. Быть может, переменные ***y***, ***z*** не составляют все влияние на *x*, и тогда не учтенное в данном наблюдении влияние обозначим ***x.*** Переменные ***x*** могут сами, в свою очередь, быть связаны между собой. Быть может, что некоторые из переменных слабо связаны или не прямо связаны с ***x***. Необходимо оставить только существенные связи. В социологии проблему такого анализа эмпирических данных впервые начал решать Э. Дюркгейм, следуя миллевской традиции причинного вывода и весьма скрупулезно используя правило сопутствующих изменений из массы статистических данных о самоубийстве.

Между двумя переменными возможны такие структурные отношения;

$x \rightarrow y$ (x обусловливает y)
$x \leftarrow y$ (y обусловливает x)
$x \leftrightarrow y$ (взаимное воздействие)
$x — y$ (нет связи)

Структурные отношения для трех переменных см. на рис. 17.
Для последних двух случаев может оказаться, что

$r_{xy} \neq 0$, $a_{xy \cdot z} = 0$,

и из эмпирических данных нельзя будет решить, какая здесь структура.

Коэффициент корреляции между ***x*** и ***y*** не равен нулю **($r_{xy} \neq 0$)** и означает, что между *x* и *y* есть связь. Частный коэффициент корреляции между ***x*** *и* ***y*** при постоянном ***z*** равен нулю и означает, что связь между ***x*** и ***y*** обусловлена не их собственным воздействием, а действием переменной ***z*** (рис. 18).

Для этих случаев (δ) и (ξ) частный коэффициент корреляции

между ***x*** *и* ***y*** при постоянном ***z*** равен нулю. Случай (δ) есть так называемая ложная корреляция.

Только анализ причинных связей между переменными может позволить выявить структуру данной эмпирической системы переменных.

В первые десятилетия XX в. обостряется интерес к проблеме причинности. С одной стороны, он обусловлен развитием квантовой физики в связи с соотношением и статистических закономерностей и выявлением ограниченности лапласовского детерминизма в рамках действующей силовой причинности. С другой стороны, этот интерес вызван развитием эмпирических неэкспериментальных наук — демографии, эконометрики, социологии. Оно привело к расширению представлений о причинности, в какой-то мере — возврату к Аристотелю. В физике, естествознании причинность понималась как действующая причина, силовое взаимодействие. Эконометрика показала существование также иной причины, связанной с нормой, правилом, или «программной обусловленностью» [16].

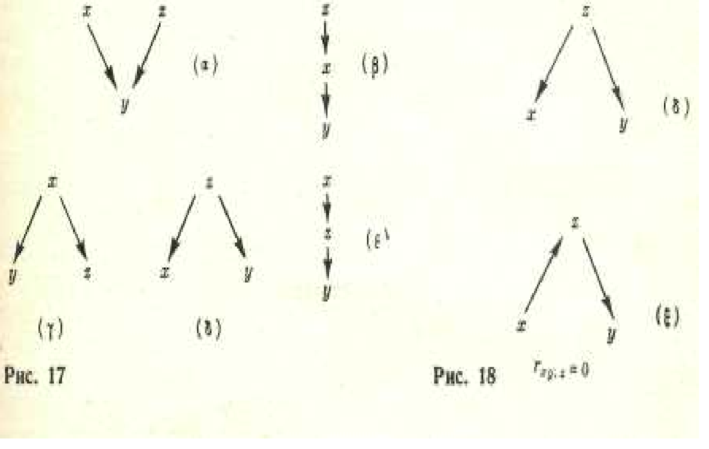

Рис. 17 Рис. 18

Эконометрика стала, по словам известного шведского специалиста Г. Уолда [17], пионером в изучении эмпирических неэкспериментальных данных.

Именно с эконометрикой, а затем с социологией наряду с квантовой механикой связано возрождение интереса к проблеме причинности. Вопрос, касающийся эмпирических данных в неэкспериментальной ситуации, впервые поставил А. Курно в отношении цены, функций спроса и предложений в условиях свободной конкуренции на рынке. Затем это получило развитие в системе уравнений равновесия Вальраса, «закона Парето», гарвардском барометре, большом числе работ по построению функций спроса и производственных функций. В этих экономических исследованиях анализировались эмпирические данные средствами математической статистики. Если при исследовании плодородия почвы или качества удобрений статистические методы применяются

[16] *Петров Ю. А.* Логическая функция категорий диалектики. М., 1972, с. 52.
[17] *Wold H.* Econometric as Pioneering in Nonexperimental Modelbuilding.— «Econometrics», 1969, v. 37, p. 319.

и были выработаны в экспериментальной ситуации, при которой специальным выбором исследуемых участков почвы можно было изолировать действие ряда фактов, чтобы проконтролировать исследуемый фактор, то в эконометрике столкнулись с эмпирической неэкспериментальной ситуацией. Проблема здесь упирается в наличие сложной системы множественных связей, каждая из которых не может быть изолирована и проконтролирована в эксперименте. Чтобы разрубить этот узел, эконометрика обратилась к понятию причинности только не в современном физическом смысле, а скорее, в аристотелевском понимании и ввела понятие причинных моделей [18] (первоначально задача стояла в выявлении связей между системой эмпирических данных). Можно написать структурные уравнения между этими переменными и полученную систему решать методом наименьших квадратов. Оказывается, что система структурных уравнений решается, если выделены так называемые экзогенные переменные, т. е. переменные, которые не определяются в данной системе, и так называемые эндогенные переменные, определяемые в данной системе [19].

Эконометрика связана с незкспериментальным построением моделей. Дуализм экспериментального и неэкспериментального построения уходит в глубь научного метода. Экспериментальные модели предиктивны. Неэкспериментальный метод, развитый в эконометрике, сталкивается с множеством проблем на всех уровнях, от уровня самых общих оснований научного метода до специальных технических проблем. Мы остановимся на двух ключевых проблемах. Первая относится к научной эволюции от детерминистских моделей к стохастическим. Статистическая картина иногда вызвана внешними причинами — ошибками наблюдений, иногда внутренними. В обоих случаях важна проблема «выбора регрессии». Проблема стохастических моделей возникла как проблема Макпранга. Проблема Макпранга разрешается, если ввести причинные отношения. Функция спроса есть причинно-следственное отношение с ценой в качестве причины и спросом в качестве следствия. Спрос обусловливается изменением цены и выражается уравнением .регрессии спроса на цену. Механизм ценообразования включает и спрос потребителя, и предложение производителя и не является просто обратным отношением функции спроса. В итоге выбор регрессии есть выбор между причинными моделями, а выбор между моделями определяет выбор между регрессиями.

[18] *Wold H.* Causality and Econometrics. — «Econometrica», 1954, v. 22, p. 162.
[19] Подробности см. в специальной литературе, например: *Тинтнер Г.* Введение в эконометрику. М., 1965.

Вторая проблема связана с различием причинных и непричинных моделей предсказания. Неэкспериментальные модели основываются на прошлых наблюдениях, и модель суммирует регулярности, наблюдаемые в прошлом. В прогнозной модели прошлые регулярности сохраняются в будущем. Прогноз в отношении неэкспериментального построения модели есть то же самое, что повторение в контролируемом эксперименте.

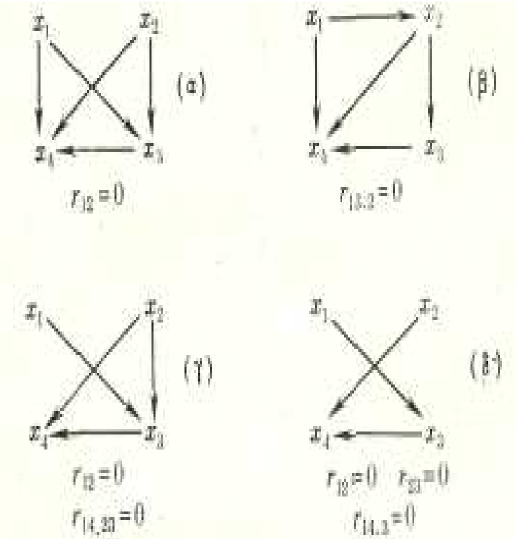

Рис. 19

Причинные связи между переменными стали изображать графически в виде диаграмм Райта, по имени биолога С. Райта, который в 20-х годах в изучении факторов наследственности применил подобные диаграммы и разработал специальный метод анализа причинных связей[20]. Если можно в какой-то степени понять антикаузализм Рассела в отношении причинности в классической физике, в которой она неотделима от экспериментальной ситуации и действующего закона, то в исследовании системы взаимосвязанных эмпирических переменных неэкспериментальных наук понятие причинности становится необходимым регулятивным и эвристическим принципом.

Для эмпирических данных определяются коэффициенты корреляции всех порядков. Мы говорим, что x_1 есть причины x_2, если с изменением x_1 изменяется в среднем x_2 при условии, что все остальные переменные постоянны. В этом случае на диаграмме эти две переменные связываются стрелкой от x_1 *к* x_2. Если же нет прямой связи между двумя переменными при условии постоянства всех других переменных, то на диаграмме переменные не связаны стрелкой. В этом случае частный коэффициент корреляции определенного порядка равен нулю (рис. 19).

[20] *Wright S.* Correlation and causality. — «J. Agric. Res.», 1921, v. 20, p. 557.

Априорно определяется возможный порядок воздействия переменных. В данных случаях принимается следующий порядок: воздействует на все переменные, на нее ни одна не воздействует: x_2 — на все, кроме x_1; x_3 — на x_4 ; x_4 не воздействует ни на одну переменную. Этот порядок обусловленности есть условие причинности.

Как только установлены по значению коэффициентов корреляции соответствующие причинные схемы, пишутся структурные уравнения в рекурсивной форме (или близкой к ней). Их коэффициенты дают меру причинного влияния переменных.

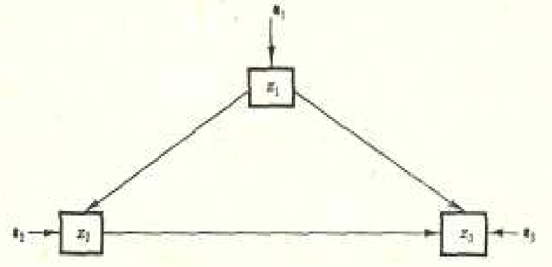

Рис. 20

Г. Саймон впервые применил анализ причинных связей, развитый в эконометрике к социологической проблематике[21].

Основная идея метода (обычно в социологической литературе он называется методом Саймона—Блэйлока) состоит в том, что хотя причинные отношения невозможно установить на основе данных о корреляциях, однако можно делать определенные выводы о причинных связях, рассматривая ряд альтернативных моделей и исключая те из них, предсказания по которым не согласуются с эмпирическими наблюдениями.

Такие медали включают: а) конечный набор явно определенных переменных, б) гипотезы о причинных взаимосвязях этих переменных и в) допущения о том, что возможное влияние внешних неучтенных переменных не нарушает наблюдаемую картину причинных связей между явными переменными. Метод Саймона—Блэйлока позволяет в ряде случаев предсказывать величины взаимных корреляций, давая тем самым эмпирически критерий оценки адекватности .причинной модели.

Например, причинную структуру, изображаемую графом связей на рис. 20, Саймон предлагает описывать следующей системой уравнений (которые в эконометрике называются структур-

[21] *Simon H.* Spurious Correlation: a Causal Interpretation.— «J. Amer. Statist. Assoc.», 1954, v. 49, p. 467.

ными уравнениями):

$x_1=\varepsilon_1;$

$x2=b_{2\ 1}x_1+\varepsilon_2$

$x2=b_{2\ 1}x_1+b_{3\ 2}x_2+\varepsilon_3,$ (1)

где x_1, x_2 *и* x_3—изучаемые признаки, а ε_1, ε_2 и ε_3—неявные факторы. При определенных допущениях коэффициенты **b_{ij}** равны соответствующим частным коэффициентам регрессии (например, $b_{3\ 2}=b_{3\ 2\cdot 1}$). При отсутствии какой-либо из связей в причинной структуре соответствующий коэффициент будет равен нулю и может служить эмпирическим критерием проверки адекватности модели.

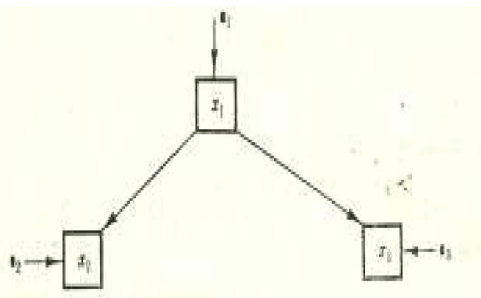

Рис. 21

Например, для графа, изображенного на рис. 21, коэффициент $b_{3\ 2}=b_{3\ 2\ \cdot 1}=0$, и, следовательно, частный коэффициент корреляции $r_{2\ 3\cdot 1}=0$; поскольку

$$r_{23\cdot 1}=\frac{r_{23}-r_{12}r_{13}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{13}^2}}$$

(2)

то $r_{23}=r_{1\ 2}r_{1\ 3}.$

Таким образом, если гипотетическая причинная структура подтверждается, то уравнение (2) должно удовлетворяться для эмпирических коэффициентов корреляции [22]. Аналогичным образом можно поступать в общем случае, т. е. частный коэффициент корреляции двух признаков, между которыми, по предположению, нет причинной связи, при фиксированных остальных должен быть равен нулю.

Не останавливаясь подробно на этих процедурах, следует сделать ряд общих замечаний относительно метода Саймона — Блэйлока.

[22] Поскольку измерения производятся с ошибкой, то, очевидно, $r_{2\ 3\cdot 1}$ приблизительно будет равен нулю, и следовательно, тождество (2) будет выполняться также лишь приблизительно.

Предложенный метод в определенных ситуациях действительно позволяет выделять из ряда альтернативных структур структуру связей, согласующуюся с эмпирическими данными. Тем не менее во многих случаях эти структуры остаются неразличимы. Кроме того, оказывается весьма неудовлетворительной интерпретация коэффициентов структурных уравнений.

В дальнейшем проблема интерпретации коэффициентов была решена Будоном[23], который показал связь между подходом Саймона—Блэйлока и работами С. Райта[24].

Самым простым подходом к описанию причинных связей между изучаемыми признаками является представление их системой линейных уравнений такого вида:

$$\begin{aligned} x_1 &= a_1 + b_{12}x_2 + \ldots + b_{1n}x_n + b_{1u}u_1; \\ x_2 &= a_2 + b_{21}x_1 + \ldots + b_{2n}x_n + b_{2u}u_2; \\ &\ldots\ldots\ldots\ldots \\ x_m &= a_m + b_{m1}x_1 + \ldots + b_{mn}x_n + b_{mu}u_m, \end{aligned} \qquad (3)$$

где x_i — переменные, соответствующие исследуемым признакам; u_i — дополнительные факторы, влияние которых необходимо учесть, но либо неизвестно, что это за факторы, либо неясны направления связей между ними и явными переменными x_i.

Решение этой системы уравнений связано с проблемой идентификации, т. е. встает вопрос, можно ли параметры системы определить из совместного наблюдения над переменными $x_1 \ldots x_n$. Эта задача достаточно подробно изучена в математической экономике, и мы коснемся ее лишь в самых общих чертах.

Вся эмпирическая информация о зависимостях между переменными $x_1 \ldots x_n$ суммируется в таблице эмпирических корреляций. Исходя из этой таблицы и необходимо получить оценки коэффициентов структурных уравнений (3). Нас будут интересовать условия, при которых могут быть получены оценки по методу наименьших квадратов, т. е. при каких условиях эти уравнения можно рассматривать как регрессионные уравнения.

Основные подходы и проблемы будут ясны при рассмотрении структуры связей простейшего типа.

Здесь $x_2, x_3 \ldots x_n$ — объясняющие переменные (их будем называть независимыми переменными).

x_1— признак, объясняемый непосредственным действием на него других явно выделенных признаков $x_2 \ldots x_n$ и неявно-

[23] *Fischer F.* The Indentification Problem in Econometrics. N. Y., 1965.

[24] Основные идеи метода измерения причинных связей изложены в ранних работах С. Райта, а наиболее полно — в статье «Корреляция и причинность» *(Wright S.* Correlation and Causality.— «J. Agric. Res.», 1921, v. 20, p. 557).

го фактора, рассматривается как обобщенный неявный фактор.

Предполагается, что $x_2 \; . \; . \; . \; x_n$ могут быть связаны ненаправленной корреляционной связью (это может осуществляться через посредство неявных факторов, действующих на $x_2 ... x_n$ и коррелирующих между собой).

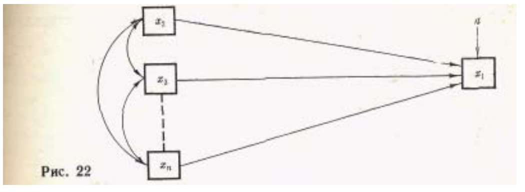

Рис. 22

Предполагая, что связи между переменными носят линейный характер, граф связей на рис. 22 эквивалентным образом можно описать линейным уравнением

$$x_1 = c_0 + c_{1\,2} x_2 + c_{1\,3} x_3 + \ldots + c_{1n} x_n + c_{1u} u \qquad (4)$$

Если u не коррелирует с $x_2 \; . \; . \; . \; x_n$, то для оценки параметров уравнения применим метод наименьших квадратов, и тогда неизвестные параметры будут совпадать с частными коэффициентами регрессии.

Однако коэффициенты регрессии зависят от единиц измерения, и поэтому оказывается невозможно прямо сравнить два коэффициента для различных переменных, если они неодинаково измерены. Отсюда вытекает трудность в получении сравнительных оценок действий объясняющих (причинных) переменных на объясняемую. Если, однако, переменные нормированы путем деления на стандартное отклонение, то мы получим стандартизованные коэффициенты регрессии, которые дают возможность прямого сравнения действий независимых переменных на зависимые. Это свойство и нашло широкое применение в моделях причинного анализа. (Однако тут же необходимо отметить, что стандартизованный коэффициент регрессии является функцией дисперсий зависимых и независимых переменных, что невыгодно отражается на его свойствах.) Введем стандартизованные переменные и коэффициенты следующим образом:

$$z_I = \frac{x_i - M_i}{\sigma_i} \text{ и } p_{1i} = c_{1i} \frac{\sigma_i}{\sigma_1},$$

где M_i — средняя i-й переменной; σ_i — стандартное отклонение i-й переменной[25].

Тогда уравнение (4) запишется в виде

$$z_1=p_{1\,2}z_2+p_{1\,2}z_3+\ldots+p_{1n}z_n+p_{1u}u$$

Коэффициент p_{1i} будем называть коэффициентом зависимо- : ста. Такое название оправдано его толкованием как доли стандартного отклонения зависимой переменной $z1$ (с соответствую\

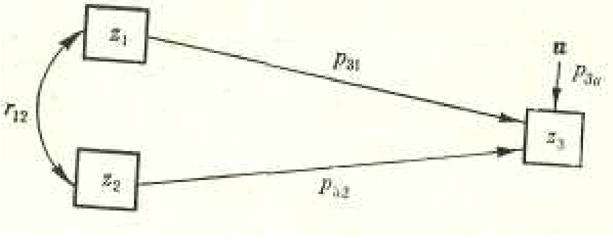

Рис. 23

щим знаком), непосредственно объясняемой фактором z_i, т. е, доли, которую можно было бы получить, если бы вариация этого фактора была идентична его действительно наблюдаемой дисперсии при фиксированном уровне остальных, в том числе и неявного фактора u. Таким образом, p_{1i} измеряет прямое влияние z_i на z_1 и если будет подтверждено, что рассматриваемый граф причинных связей соответствует реальности, то p_{1i} можно считать мерой причинного влияния z_i на z_1.

Простая система причинных связей дана на рис. 23.

Структурное уравнение имеет вид

$$z_3=p_{31}z_1+p_{32}z_2+p_{3u}u$$

Для оценки коэффициентов зависимости умножим обе части уравнения на z_i ($i=1$, 2, 3) и, вычисляя математическое ожидание от обеих частей уравнения, получим

$$M(z_3z_i)=p_{3\,1}M(z_1z_i)+p_{3\,2}M(z_2z_i)+p_{3\,4}M(uz_i).$$

Поскольку z_i стандартизованы, математические ожидания будут равны коэффициентам корреляции, т. е. $M(z_iz_j)=z_{ij}$. Величины $M(uz_1)$ и $M(uz_2)$ равны нулю, по предположению о не-

[25] Использование коэффициентов для анализа причинных связей впервые было изучено С. Райтом, который назвал их «path-коэффициентами». Это название закрепилось в американской литературе, и, более того, вся процедура исследования причинных связей с использованием этих стандартизованных коэффициентов получила это название.

коррелированности *u* с независимыми переменными и $M(uz_3)$ =

Таким образом, мы получаем простой способ нахождения соотношений для коэффициентов зависимости. Для нашего примера найденные уравнения имеют вид

$$r_{31}=p_{31}+p_{32}r_{21}; \qquad (5)$$

$$r_{32}=p_{32}+p_{31}r_{12}; \qquad (6)$$

$$r_{33}=1=p_{31}r_{31}+p_{32}r_{32}+p_{3u}^2, \qquad (7)$$

причем r_{3t}, r_{21}, r_{32} вычисляются из наблюдаемых данных. Уравнения (5) — (7) называются системой оценочных уравнений для коэффициентов зависимости. Из анализа этих уравнений можно сделать следующие выводы:

а) решая систему оценочных уравнений, найдем p_{31}:

$$p_{31}\frac{r_{31}-r_{32}\cdot r_{12}}{1-r_{12}^2}=$$

Сравнение этой оценки и оценки для c_{ij} уравнения (4) по методу наименьших квадратов показывает [26], что они эквивалентны. Этот результат верен и для n-мерного случая;

б) рассмотрим уравнение (5) из оценочной системы:

$$r_{31}=p_{31}+p_{32}r_{21}$$

Корреляцию между объясняемой и объясняющей переменной можно интерпретировать как сумму прямого воздействия (p_{31}) и косвенного $(p_{32}\ r_{21})$. Отсюда очевидна недостаточность выводов на основе одних лишь парных корреляций. Поскольку p_{31} измеряет действие г, на г, при фиксированных остальных переменных, а p_{32} — прямое действие z_2 на z_3 то за счет корреляции z_1 и z_2 парный коэффициент корреляции r_{31} может существенно искажать выводы; r_{31} может быть близок к нулю при разных знаках p_{31} и $p_{32}\cdot r_{21}$, однако реальное действие переменной z_1 на z_3 будет велико. Интерпретируя таким образом формулу (5), не следует забывать, что вся эта интерпретация имеет смысл лишь по отношению к постулируемой причинной структуре;

в) находя p_{3u}^2 из оценочной системы уравнений, нетрудно показать, что p_{3u}^2 выражается через коэффициент множественной корреляции, т. е.

$$P_{3u}^2=1-R^2$$

[26] См., например: *Кендэл М., Стюарт А.* Статистические выводы и связи. М., 1973, гл. 27.

г) для графа связей с га независимыми переменными общие формулы имеют вид, аналогичный вышеприведенным, и не нуждаются в пояснениях.

В общем случае граф причинных связей включает не одну, а несколько объясняемых переменных $z_1 \dots z_n$ и описывается системой линейных уравнений (3).

Пусть, как и раньше, переменные стандартизованы. Задача состоит в поиске условий, при которых параметры системы могут быть определены из наблюдаемых данных.

Важный класс систем, который имеет особую значимость для описания причинных структур, составляют так называемые рекурсивные системы, у которых матрицы коэффициентов треугольны, т. е.

$$
\begin{aligned}
z_2 &= p_{21} z_1 + p_{2u} u_2; \\
z_3 &= p_{31} z_1 + p_{32} z_2 + p_{3u} u_3; \\
&\dots\dots\dots\dots\dots\dots \\
z_n &= p_{n1} z_1 + p_{n2} z_2 + \dots + p_{nn-1} z_{n-1} + p_{nu} u_n.
\end{aligned}
$$

Если предположить, что u_i не коррелируют между собой, т. е. $cov(u_i u_j) = 0, (i \neq j)$, и не коррелируют с независимыми перемен-
$i=1$
ными, то рекурсивная система идентифицируема. Кроме того, к каждому уравнению может быть применена оценка по методу наименьших квадратов.

Система рекурсивных уравнений определяет однонаправленный причинный процесс — процесс, не содержащий явление взаимодействия следствия и причины. Этому случаю, очевидно, отвечает состояние динамического равновесия объектов исследования. И наоборот, если постулируемую систему связей рассматривать как однонаправленный процесс, то он может быть списан рекурсивной системой уравнений.

Для иллюстрации рассмотрим конкретный пример.

Проблема состоит в том, чтобы объяснить зависимую переменную — занятие домашним трудом z_9 — из ограниченного числа социально-демографических показателей, В частности, объяснение проводилось через такие характеристики, как пол (z_2), возраст (z_3), образование (z_4), профессия (z_5), заработная плата (z_6), семейное положение (z_7), наличие детей (z_8) и район проживания (z_1) .

В ходе исследования были сформулированы две альтернативные гипотезы относительно структуры причинных связей между рассматриваемыми признаками. Структура связей, получившая в конечном счете эмпирическое подтверждение, изображена на рис. 24 (такую сложную структуру удобнее записывать в виде матрицы коэффициентов).

Система структурных уравнений имеет следующий вид:

$$
\begin{aligned}
z_4 &= p_{41}z_1 + p_{42}z_1 + p_{43}z_3 + p_{4u}u_4;\\
z_5 &= p_{51}z_1 + p_{52}z_2 + p_{53}z_3 + p_{54}z_4 + p_{5u}u_5;\\
z_6 &= p_{63}z_3 + p_{64}z_4 + p_{6u}z_5 + p_{6u}u_6;\\
z_7 &= p_{73}z_3 + p_{76}z_6 + p_{7u}u_7;\\
z_8 &= p_{82}z_2 + p_{83}z_3 + p_{86}z_6 + p_{87}z_7 + p_{8u}u_8;\\
z_9 &= p_{91}z_1 + p_{92}z_2 + p_{93}z_3 + p_{94}z_4 + p_{95}z_5 + p_{96}z_6 +\\
&\quad + p_{97}z_7 + p_{98}z_8 + p_{9u}u_9.
\end{aligned}
$$

Оценочные уравнения получаются путем применения вышеописанной процедуры последовательно к каждому уравнению

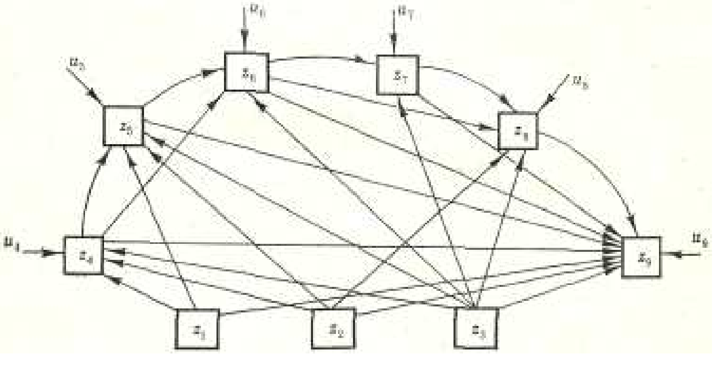

Рис. 24

системы. Заметим, что число получаемых оценочных уравнений здесь равно 33, а неизвестных коэффициентов — 24 (мы не рассматриваем уравнения для неявных факторов, поскольку они не влияют на оценку коэффициентов зависимости между явными переменными).

Переопределенную оценочную систему можно решить либо просто, как систему уравнений, либо — и это дает лучшее приближение к эмпирическим данным — методом наименьших квадратов. Полученные оценки повторной подстановкой в оценочные уравнения позволяют вычислить теоретические величины коэффициентов корреляции и путем их сравнения с эмпирическими проверить адекватность модели.

В табл. 15 приведены численные значения коэффициентов p_{ij} а также эмпирические и теоретические величины коэффици-

Таблица 15

№ связи	p_{ji}	Коэффициенты корреляции		№ связи	p_{ji}	Коэффициенты корреляции	
		теорети-ческие	эмпири-ческие			теоретические	эмпирические
91	0,0814	0,1653	0,0142	95	—0,0866	0,11	0,1
92	0,262	0,2587	0,263	96	—0,562	0,0375	0,0143
93	0,255	0,10	0,1326	97	0,821	0,4355	0,4454
94	—0,106	0,073	0,0748	98	0,314	0,3684	0,3305

ентов корреляции r_{ij} для связей объясняемого признака z_9 с объясняющими факторами $z_1 \ldots z_8$ [27].

По таблице можно проследить не только соотносительные силы влияния каждой из переменных на объясняемый признак, но и учесть перераспределение этого влияния по всем постулируемым связям. Например, поскольку прямое влияние переменной z_1 на z_9 гари фиксированных остальных измеряется величиной p_{ji}, а совместное действие всех переменных — коэффициентом корреляции r_{ji} , то разность $(r_{ji} — p_{ji})$ является мерой влияния переменных z_2 $z_3 \ldots z_8$ *на* z_j т. е. мерой общего косвенного воздействия этих *переменных*. Причем из оценочных уравнений можно вычислить опосредованное влияние каждой отдельной связи.

В общем случае модель изучаемой структуры связей может включать любое число объясняемых и объясняющих переменных при условии, что система структурных уравнений остается рекурсивной.

[27] Нереесова Е. Х. Выбор модели изучаемого явления методом зависимости.— В кн.: Динамика изменения положения женщин и семья, М., 1972, с. 114 (в работе Нерсесовой использован дихотомический аналог $r_{ij} — \Phi_{ij}$).

Заключение

Вся история конкретных социологических исследований в СССР за последние десять лет была связана со все более широким и более специализированным использованием математических методов сбора и обработки первичной социальной информации.

В первых конкретных социологических исследованиях при анализе социальных данных были взяты на вооружение простейшие математические и статистические методы — методы средних чисел, метод аналитических группировок, индексовый метод анализа, т. е. методы так называемой дескриптивной статистики.

По мере развития конкретных социологических исследований применялись все более точные математические методы анализа социальных данных и выборки. Оперирование с большими массивами социальной информации привело к проблеме использования вычислительной техники — счетно-перфорационных и электронно-вычислительных машин. Социологи-марксисты столкнулись с необходимостью измерения качественных социальных переменных и моделированием социальных процессов и явлений. В настоящее время перед марксистско-ленинской социологией стоит задача разработки методов измерения самых различных систем социальных показателей и индикаторов народнохозяйственного планирования, создания комплексных математических социально-экономических моделей и т. д.

Однако применение математики в исследовании некоторых социологических проблем, например проблем социальной структуры общества, пока идет в социально-экономическом направлении. Количественно изучается влияние материальных показателей: доход, заработная плата, жилая площадь, число школ, больниц, киноустановок и т. п. Все это, естественно, необходимо. Это — основа. Но вместе с тем необходимо смелее изучать собственно человека и переменные, характеризующие его развитие, изучать его социальные потребности, интересы и отношения в группах и территориальных общностях: его желания, мнения, установки, ориентации, симпатии, удовлетворенность жизнью, национальные чувства, физическое и духовное здоровье, т. е. необходимо решать проблему измерения в собственно социологических и социально-психологических аспектах исследования.

Ярким примером использования математики в социальных науках является экономика. Образец построения количественной модели экономики мы находим у К- Маркса, который во втором томе «Капитала» дал схемы расширенного производства. В настоящее время разработка проблем применения математических методов в экономике, внедрение количественных способов анализа экономических явлений не только обогащает экономическую теорию, но и стимулирует развитие самой математики [1]. Интересную мысль в отношении специфики использования математических методов в социальных науках высказали выдающийся математик современности Дж. фон Нейман и американский экономист О. Моргенштерн — авторы классической книги «Теория игр и экономическое поведение»: «Решающая фаза применения математики к физике — создание Ньютоном науки механики—происходила и едва ли могла быть отделена от открытия исчисления бесконечно малых. (Имеется еще несколько примеров, но ни один из них не сильнее этого.) Важность социальных явлений, богатство и множественность их проявлений по меньшей мере равны физическим. Следовательно, надо ожидать— или опасаться,—что потребуются математические открытия того же ранга, что исчисление бесконечно малых для того, чтобы произвести решительный переворот в этой области... невероятно, „что только повторение приемов, которые так хорошо служили в физике, даст что-то ценное для социальных явлений. В самом деле, вероятность очень мала, так как мы покажем, что в нашем обсуждении сталкиваемся с математическими задачами, которые совершенно отличны от задач, встретившихся в физике. Нужно иметь в виду эти наблюдения в связи с современным преувеличением роли математического анализа, дифференциальных уравнений и т. д. в качестве главных орудий математической экономики» [2].

Чрезвычайно важная задача — разработка и развитие специфических органически возникших из. потребностей социологии

[1] *Федоренко И. П.* Экономика и математика. М., 1967, с. 13.
[2] *Нейман Дж. фон, Моргенштерн О.* Теория игр и экономическое поведение. М., 1970, с. 7—9.

математических методов. В этой связи хотя бы в заключении мы не можем не упомянуть о важной группе методов — так называемых методах многомерного шкалирования, которые создаются и развиваются в процессе решения собственно социологических и психологических задач.

Представляет интерес то, что «методы многомерного шкалирования» сводятся к новым направлениям в области измерения в социологии и психологии. Эти методы имеют свои истоки в физике и психофизике. Как известно, звук можно рассматривать в одном измерении, например только по громкости или только по высоте. А можно определить сразу два этих параметра и соответственно найти точку в двухмерном пространстве. Этот вопрос возникает ,и имеет большое значение, когда нет предварительных сведений о числе измерений данного стимула или суждения. Например, мы спрашиваем «нравится —не нравится?», «красивая вещь или некрасивая?», «почему вещь нравится или кажется красивой?». И каждый ответ будет самостоятельным измерением красоты и привлекательности. Это связано с проблемой эквивалентности стимулов, что имеет большое распространение в изучении восприятия, мышления, обучения и измерения установок.

Предположим, что мы имеем четыре суждения — ABCD, которые зависят от двух измерений D_t и D_2. Если бы мы могли их изолировать, то могли бы сначала определить координаты *ABCD* по измерению D_1 а затем (изолируя D_1) — координаты по измерению D_2. Но это не всегда можно сделать, и если это сделать нельзя, то используют расстояния между стимулами. Эти расстояния выражают степень подобия или различия.

Если обнаружится, что AB + BC = AC, то можно считать, что имеет место одно измерение. Если AB+BC>AC, то имеет место два измерения. Три стимула могут быть расположены в пространстве не больше двух измерений (три точки определяют плоскость), четыре стимула могут быть в пространстве одного, двух, трех измерений. Наименьшее число измерений, в котором расположены все расстояния, и будет размерностью этого множества суждений (стимулов).

Л. Ричардсон в 1938 г. впервые для решения проблемы многомерного шкалирования ввел метод триад. Стимулы представляются группами по три, и в каждой группе надо определить, какие два стимула — самые близкие и какие — наиболее различаются. Эти три расстояния ранжируются. Затем применяется парное сравнение не к стимулам, а к расстояниям и соответственно применяется закон сравнительного суждения к расстояниям.

Позднее У. Торгерсон развил идеи Л. Ричардсона и тогда этот подход получил название метода многомерного шкалирования. Существенно, что в работах Ричардсона — Торгерсона размерность исследуемого явления определяется из рассмотрения совокупности эвклидовых расстояний между исходными стимулами-суждениями. Р. Шеппард, а затем Краскал начали использовать в качестве исходных данных не эвклидовы расстояния, а различные качественные меры близости между объектами. По своей сути подход Ричардсона-Торгерсона— метрическое многомерное шкалирование, а подход Шеппарда-Краскала — неметрическое многомерное шкалирование.

Работа по применению математических методов в социологии развивается в четырех направлениях: во-первых, это решение проблем методологии социологического исследования, а именно; выборки, анализа данных, измерения и моделирования; во-вторых, это расширение сферы и отраслей социологического знания, в которых возможно использование математических методов; в-третьих, это увеличение количества средств и методов и их модификаций из всевозможных разделов математики; и. наконец, это поиск на пути создания собственных математических методов в социологии. Эти четыре направления представляют единый процесс — ни одно не существует вне других. Действительно, поиск нового собственно социологического математического формализма происходит не абстрактно, а на базе уже имеющегося аппарата, необходимого в определенной области социологии и одновременно в ходе решения одной из методологических проблем социологического исследования — выборки, анализа данных, измерения или моделирования. Наконец, если мы возьмем, например, выборку или моделирование, то они осуществляются посредством уже имеющихся и апробированных математических средств во многих отраслях социологии и, естественно, для них возможен поиск новых путей.

Проникновение математики в социологию, в особенности в конкретные социальные исследования, происходит все более интенсивно, хотя и сопряжено с реальными ощутимыми трудностями. Одно из условий преодоления этих трудностей — их выявление и тем самым привлечение к ним внимания и математиков, и социологов. Совместная работа социологов и математиков, которая уже началась в нашей стране, будет содействовать развитию применения математики в социологии, выработке математических методов, специфических для социологии, и в частности методов социологического измерения.