Лекция 3. РЕГРЕССИОННЫЕ МОДЕЛИ РОСТА ДЕРЕВЬЕВ И ДРЕВОСТОЕВ

Результаты любого эксперимента в виде измерений или наблюдений следует рассматривать как случайную выборку из общей (генеральной) совокупности. Количество информации, методы ее сбора и обработки должны обеспечивать необходимую точность результатов. Никакие математические методы и ЭВМ не могут добавить точности исходным данным, увеличить количество содержащейся в них информации и в конечном счете повысить надежность результатов; они могут лишь помочь строго оценить эту надежность и придать результатам наиболее целесообразную форму. Отсюда следует, что математические методы применимы не к любым хаотическим данным, а только к тем, которые собраны и обработаны с применением достаточно строгих требований математического планирования эксперимента и статистического анализа данных [1, 13].

3.1. Множественные регрессионные модели роста деревьев и древостоев

В моделировании хода роста насаждений и разработке имитационных моделей строения и производительности древостоев широко используются множественные регрессионные модели [1, 6, 9]. Математическое описание функций системы (биогеоценоза, насаждения и т. д.) в целом и функций связи отдельных элементов системы можно выполнить в виде обобщенного дискретного полинома Колмогорова–Габора [23]:

$$Y = b_0 + \sum b_4 x_n + \sum \sum b_{n_1 n_2} x_{n_1} x_{n_2} + ... + \sum b_{n_1} x_{n_1}^m \tag{46}$$

При двух факторах (x_1,x_2) линейная модель первой степени имеет вид:

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_1 x_1 x_2, \tag{47}$$

где b_0, b_1, b_2, b_3 – коэффициенты регрессии.

Линейная модель второй степени имеет уже 11членов:

$$\begin{gathered} Y = b_0 + b_2 x_2 + b_3 x_1 x_2 + b_4 x_1^2 + b_5 x_2^2 + b_6 x_1^2 x^2 + b_7 x_1 x_2 + b_8 x_1^2 x_2^2 + \\ + b_9 x_1 x_1^2 x_2^2 + b_{10} x_2 x_1^2 x_2^2. \end{gathered} \tag{48}$$

Количество членов уравнения быстро растет с увеличением числа аргументов (факторов). Так, модель второй степени при

4 факторах включает 70 членов. Объем наблюдений возрастает также с увеличением числа переменных, так как число наблюдений должно быть в 5–7 раз больше числа факторов. При разработке модели (48) необходимо провести эксперимент объемом 50–70 наблюдений. Для формального решения задачи объем наблюдений с ростом числа аргументов становится практически необозрим.

В уравнении (48) можно выделить три качественно отличные части: 1) линейную – с коэффициентом при аргументах в степени единица (b_1x_2 и b_2x_2); 2) нелинейную – с коэффициентами при аргументах в степени m>1 ($b_4 x_1^2$ и $b_5 x_2^2$); 3) неаддитивную – с коэффициентами при произведениях аргументов по два, три и более ($b_3x_1x_2$, $a_6 x_1^2 x_2$ и т. д.).

Практика применения регрессионного анализа показывает, что нет необходимости рассматривать в уравнениях слишком высокие степени и произведения многих аргументов. На линейную часть уравнения часто приходится наибольшая информация (70–90%), а вклад нелинейной и неаддитивной частей сравнительно невелик. Следовательно, сначала необходимо описать объект системой множественных линейных регрессионных моделей, а затем оценить, насколько улучшается аппроксимация функции, если дополнительно вводятся в уравнение нелинейные и неаддитивные члены.

Линейные и нелинейные регрессионные уравнения, применяемые в моделировании хода роста древостоев, можно представить в виде трех типов:

1) линейная регрессия по x и коэффициентами a, b, т. е. $Y=a + bx$;
2) нелинейное уравнение по x и линейное по коэффициентам a, b, c, d, т. е. уравнение $Y = a + bx + cx^2 + dx^3$;
3) нелинейная модель как по x_i, так и по коэффициентам a, b, c, d, т. е. регрессия $Y = ax_1^b + cx_2^{dx^3}$.

Нелинейные модели, которые можно привести к линейному виду называются внутренне линейными. Путем преобразования и замены переменных внутренне линейные модели приводятся к линейной множественной регрессии вида

$$Y = b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n. \tag{49}$$

Парабола 3-го порядка $Y = a + bx + cx^{2} + dx^{3}$ путем замены переменных $X_1=X$; $X_2=X^2$; $X_3=X^3$ и коэффициентов $a=b_0$; $b_1=b$; $b_2=c$; $b_3=d$ приводится к линейному виду (49):

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3.$$

Известное в лесной таксации уравнение Корсуня

$Y = \dfrac{x^2}{a + bx + cx^2}$ преобразуется путем замены $Y' = \dfrac{x^2}{Y}$; $x_1 = x$; $x_2 = x^2$; $b_0 = a$; $b_1 = b$; $b_2 = c$. Тогда получим $Y' = b_1 + b_1 x_1 + b_2 x_2$.

Мультипликативная модель $Y = a x_1^b x_2^c$ приводится к линейному виду логарифмированием: $\ln Y = \ln a + b \ln x_1 + c \ln x_2$ и заменой переменных $Y' = \ln Y$; $b_0 = \ln a$; $b_1 = b$; $b_2 = c$; $x_1^1 = \ln x_1$; $x_2^1 = \ln x_2$. Отсюда имеем линейную регрессию $Y' = b_0 + b_1 x_1^1 + b_2 x_2^1$.

Экспоненциальная модель $Y = 1/(1 - \mathrm{e}^{a+bx})$ преобразуется к линейной:

$$1 - \frac{1}{Y} = \mathrm{e}^{a+bx}; \ln\left(1 - \frac{1}{Y}\right) = a + bx; Y' = b_0 + b_1 x_1, \qquad (50)$$

где $X^t = \ln(1 - \dfrac{1}{X})$.

Линейные модели второй и более высоких степеней типа (48) также приводятся к линейному виду путем замены переменных $x_3 = x_1 x_2$; $x_4 = x_1^2$; $x_5 = x_2^2$; $x_6 = x_1^2 x_2$ и т. д.

Линеаризация нелинейных по параметрам зависимостей может привести к резко искаженным величинам дисперсий и оценок параметров регрессионной модели. С другой стороны, задача оценивания параметров нелинейным регрессионным анализом может быть сложной и иногда вообще нерешаемой. Поэтому следует особое внимание обратить на статистические предпосылки и методические аспекты регрессионного анализа [26].

Статистические методы отбора переменных в регрессионном анализе обычно противоположны по характеру критерия выбора «наилучшего» уравнения: 1) стремление включить в модель по возможности больше факторов, чтобы надежнее определить прогнозируемые величины; 2) из-за затрат, связанных с получением данных при большом числе независимых переменных, мы должны стремиться к тому, чтобы уравнение было простым и включало как можно меньше переменных. Нет однозначного статистического метода для выполнения этого выбора, субъективные суждения являются составной частью любого метода. Одним из наиболее эффективных методов является шаговый регрессионный анализ, который широко применяется в технических исследованиях. При этом методе производится автоматический отбор переменных с помощью вычислительной техники на основе оценки значимости коэффициентов и параметров регрессии [26]. В лесохозяйственных исследованиях мы имеем дело с «пассивным» экспериментом, многие независимые переменные значительно варьируют и сильно коррелированы. Это может привести к смещению в оценках коэффициентов регрессии так, что регрессионная модель потеряет всякий смысл [1, 25].

Классический регрессионный анализ предполагает ряд постулатов в основном статистическом анализе. Эти постулаты гласят, что регрессия представляет собой линейную комбинацию некоторых линейно независимых базисных функций от факторов с неизвестными коэффициентами (параметрами). Факторы (x_1, x_2 и т. д.) должны быть детерминированными и измеренными точно. В отношении зависимых переменных (Y) считается, что это равноточные (с одинаковой дисперсией) некоррелированные случайные величины, имеющие нормальное распределение. Такие предпосылки регрессионного анализа позволяют получить несмещенные и эффективные оценки коэффициентов регрессии методом наименьших квадратов и осуществить проверки основных статистических гипотез относительно уравнения.

В дальнейшем появился целый набор методов и процедур для анализа или замены некоторых предпосылок классического регрессионного анализа. Так, если зависимые переменные неравноточны и коррелированы, то для оценки коэффициентов регрессии применяют взвешенный метод наименьших квадратов. Мощным средством обнаружения некоторых отклонений от исходных предпосылок регрессионного анализа является анализ остатков от регрессии [25, 26].

Мультиколлинеарность, или коррелируемость, независимых переменных считается сильной, если коэффициент парной корреляции независимых переменных больше коэффициента множественной корреляции. При этом оценки коэффициентов регрессии методом наименьших квадратов становятся настолько неудовлетворительными, что даже знаки перед ними часто не соответствуют истинным. Эффективными оценками при сильной мультиколлинеарности являются гребневые, или ридж-оценки, коэффициентов регрессии [26].

В соответствии с общими предпосылками регрессионного анализа, при моделировании хода роста древостоев можно принять следующие основные постулаты: 1) регрессионная модель должна объяснять не менее 90% вариации зависимой переменной (коэффициент детерминации $R \geq 0{,}90$); 2) достоверность регрессии оценивается по F-критерию Фишера; 3) коэффициенты регрессии должны быть значимы по t-критерию Стьюдента на 5%-ном уровне значимости; 4) измерения (оценки) зависимых переменных равноточны и некоррелированы; 5) мультиколлинеарность независимых переменных (факторов) незначительная; 6) регрессионная модель отличается относительной ошибкой – менее 10%-ного среднего значения предсказываемой зависимой переменной; 7) остатки от регрессии должны быть без заметной автокорреляции, нормально распределены и без систематической составляющей [25].

Регрессионный анализ тесно связан с дисперсионным анализом. Общая дисперсия в комплексе равна $\text{у}_y^2 = \sum (y_i - \bar{y})^2$. Дисперсия зависимой переменной, обусловленная регрессией, будет $\text{у}_{y/x}^2 = \sum \left(\tilde{y} - \bar{y}\right)^2$. Остаточная (случайная) дисперсия $\text{у}_0^2 = \sum \left(\tilde{y} - y_i\right)^2$. В приведенных формулах оценки дисперсий: у_i – измеренная i-тая зависимая случайная величина; $\tilde{y}$ – значение зависимой переменной по уравнению регрессии; $\bar{y}$ – среднеарифметическое значение зависимой случайной величины в выборке. Коэффициент детерминации или квадрат коэффициента множественной корреляции равен $R^2 = \text{у}_{yx}^2 / \text{у}_y^2$. Коэффициент детерминации определяет вариацию зависимой переменной (высот, диаметров, запасов древостоев) относительно среднего уровня (тренда) линии регрессии хода роста древостоев, т. е. изменения их высот, диаметров, запасов с возрастом. Практически аппроксимацией полинома высокой степени (модель 46) можно достигнуть $R^2 \approx 1$, что вовсе не подтверждает адекватность модели реальному процессу.

Критерий Фишера (F-критерий) используется как общий критерий оценки достоверности регрессии с определенным уровнем вероятности. Особое внимание следует уделить исследованию остатков [26]. Остатки есть разность между фактическими наблюдениями и значениями зависимой переменной величины, предсказанными по регрессионному уравнению. В общей математической модели временного ряда хода роста древостоев (76) остатки представляют собой случайную составляющую (U_t). Если регрессионная модель признана правильной, то остатки от регрессии представляют собой ошибки измерений . При проведении регрессионного анализа мы делаем некоторые предположения в признании ошибок независимыми, имеющими нулевое среднее, одинаковую (постоянную) дисперсию и подчиняющимися нормальному распределению. Последнее предположение необходимо для применения F-критерия Фишера.

Исследование остатков проводят графическим способом. Основные виды графиков остатков: общий график, в зависимости временной шкалы, от предсказываемых значений зависимой переменной или независимых переменных и др. [26]. Подходящий график временной последовательности остатков должен иметь вид горизонтальной полосы. Примеры ситуаций, характеризующих неудовлетворительное распределение остатков, даны на рис. 10.

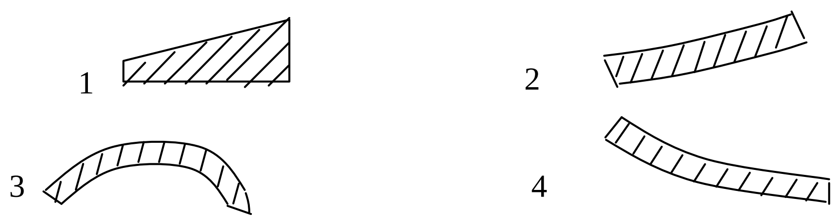


Рис. 10. Неудовлетворительное распределение остатков от регрессии

Неудовлетворительное распределение остатков может обусловливаться рядом причин:

1) дисперсия остатков не постоянна, а растет со временем, поэтому надо применить взвешенный метод наименьших квадратов; 2) остатки имеют положительную систематическую составляющую, и модель завышает результаты; 3) в модель следует включить линейный и квадратичный члены от времени; 4) регрессия характеризуется отрицательной систематической составляющей. Если дисперсия остатков не постоянная, распределение остатков отличается от нормального, то следует ввести дополнительные члены уравнения или произвести преобразования наблюдений до регрессионного анализа.

В регрессионном анализе, когда по (n) наблюдениям оцениваются (p) параметры (коэффициенты) регрессионной модели (n), остатки связаны лишь с (n–p) степенями свободы. Следовательно, остатки не могут быть независимыми и между ними существует корреляция [26].

Регрессионный анализ, проверка адекватности регрессионной модели, исследование остатков от регрессии возможны при достаточном числе наблюдений. При малом числе наблюдений трудно, а иногда невозможно оценить лучшее уравнение регрессии.

Множественный регрессионный анализ зависимости диаметров и высот деревьев выполнен нами по данным перечислительной таксации чистых одновозрастных насаждений [25].

Опытные данные представлены в виде материалов таксации на двух пробных площадях сплошной рубки деревьев, заложенных в березняках кисличных и черничных, составляющих более половины березовых лесов по суходолу. Пробная площадь 1 (Городокский лесхоз) – березняк кисличный семенного происхождения, Ia класса бонитета, состав 9Б1Ос (55 лет), запас – 256 м^3/га, средний диаметр – 22,6 см, высота – 26,9 м, площадь сечения – 20,5 м^2/га, полнота – 0,7; на пробе срублено и протаксировано 243 ствола березы бородавчатой. Пробная площадь 2 (Вилейский лесхоз) – березняк черничный семенного происхождения, I класса бонитета, состав 10Б (60 лет), запас – 193 м^3/га, средний диаметр – 19,0 см, средняя высота – 22,2 м, площадь сечения – 19,5 м^2/га, полнота – 0,7; срублено и протаксировано 236 стволов березы. На каждом стволе диаметры на высоте 1,3 м измерялись с точностью до 0,1 см, высоты деревьев до 0,1 м.

Путем аналитического анализа моделей связи диаметров и высот деревьев в древостое отобраны уравнения параболического, логарифмического и экспоненциального типов с преобразованием и без преобразования зависимой переменной.

Параболического типа:

$H = a + bD + cD^2;$ (51)

$H = a + bD + cD^2 + eD^3;$ (52)

$H = a + bD + cD^2\ eD^3 + fD^4;$ (53)

$H = a + bD^{-2};$ (54)

$H = a + bD + cD^{1/2}H;$ (55)

$H = a + bD^{-1} + cD^{-2};$ (56)

$H = a + bD^{1/2} + cD + eD^2;$ (57)

$H = a + bD + cD^2 + eD^{1/2} + fD^{-1/2} + gD^{-1} + kD^{-2}.$ (58)

Логарифмического типа:

$H = a + b\lg D;$ (59)

$H = a + bD + s\lg D.$ (60)

Уравнения с преобразованием переменной:

$\lg H = a + b\lg D;$ (61)

$\lg H = a + b\lg D + c\lg^2 D;$ (62)

$D/H = a + bD;$ (63)

$$D^2/H = a + bD + cD^2; \quad (64)$$
$$D/H = a + bD + cD^{-1}; \quad (65)$$
где H – высота; D – диаметр; a, b, c, e, f, g, k – коэффициенты уравнений.

Обработка опытных данных и оценка коэффициентов регрессий методом наименьших квадратов производилась на ЕС ЭВМ по стандартной программе множественного регрессионного анализа [24].

Коэффициент детерминации (R^2) определяет вариацию высот деревьев относительно среднего уровня линии регрессии. В березняке кисличном (проба 1) вариация высот выше, чем в черничном, и уравнения объясняют только 63–73% вариации высот деревьев (табл. 1).

Критерий Фишера используется как общий критерий значимости регрессии. Вычисленные критерии во всех случаях значительно превышают критическое (табличное) значение, что указывает на достигнутое снижение общей вариации высот, предсказанных по регрессионным моделям.

Оценки коэффициентов регрессий, как выборочных изменчивых показателей, получены с точностью 5–10% и по t-критерию значимы на принятом уровне. Свободный член (a) уравнения (65) является недостоверным для данных пробных площадей и может быть исключен из модели. Стандартные ошибки (S) оценки высот по уравнениям в березняке кисличном составляют 2–3 м, или 9–12%, в черничном 0,5–1,0 м, или 2–4% среднего значения предсказываемых высот. Максимальные отклонения (Δ_{max}) опытных значений высот от регрессий достигают 5–10 м. Размерность стандартной ошибки и остатков от регрессии зависит от преобразования зависимой переменной (табл. 1).

Аппроксимация уравнений (51 – 65) к данным таксации диаметров и высот в березняке черничном (проба 2) характеризуется более высокими коэффициентами детерминации, меньшей стандартной ошибкой и суммой квадратов остатков или отклонений от линии регрессии ($\sum\Delta^2$). С другой стороны, регрессии имеют значительную автокорреляцию в остатках (R_1=0,6–0,7), что можно объяснить влиянием временного фактора (возраста деревьев) на вариацию высот. Введение дополнительных независимых переменных в модель (58) значительно уменьшает автокорреляцию в остатках (R_1=0,233).

Уравнения параболического типа характеризуются автокорреляцией в остатках, распределение которых значительно отличается от нормального. Парабола II порядка (51) и модель (58) являются наиболее подходящими для интерполяции связи диаметров и высот деревьев в березовых древостоях. Уравнения логарифмического типа (59) и (60) дают систематическое занижение в оценке высот деревьев и лучше подходят для определения высот тонкомерных стволов. Преобразование зависимой переменной (lg*H*) приводит к нормальному распределению остатков (модели (61, 62)). В уравнении Корсуня (64) представлены наиболее подходящие показатели модели связи диаметров и высот деревьев: высокий коэффициент детерминации даже для данных пробы 1, стандартная ошибка в пределах 0,6–2,0 м (или 2–7% среднего значения), меньшая сумма квадратов отклонений, незначительная автокорреляция в остатках и приблизительно нормальное их распределение.

Проведенные исследования моделей связи показывают, что реальные выборочные данные таксации диаметров и высот деревьев часто не соответствуют предпосылкам метода наименьших квадратов и регрессионного анализа. Нарушение предпосылок связано с наличием корреляции между независимыми переменными в моделях (51 – 65). Это затрудняет проведение регрессионного анализа и построение моделей связи; усложняется процесс выделения наиболее существенных факторов, искажаются оценки и смысл коэффициентов регрессии при попытке их лесоводственной интерпретации, возникают осложнения вычислительного характера.

Необходимо также отметить, что для аналитического описания связи диаметров и высот деревьев нецелесообразно использовать многочлены высокой степени или другие уравнения, содержащие большое число параметров, так как полученные модели связи (особенно при малом числе наблюдений) будут отражать случайные колебания, а не основную тенденцию развития явления. По данным пробной площади 1, например, при оценке параметров регрессии (58) получена вырожденная матрица переменных, т. е. имеются осложнения вычислительного характера. Полиномы подходят для выравнивания и интерполяции (определения промежуточных значений) данных в пределах имеющегося опытного материала и непригодны для прогноза роста древостоев.

Таблица 1. **Показатели регрессий зависимости диаметров и высот деревьев в березовых древостоях**

№ урав-	Пробная площадь 1			Пробная площадь 2		
	R^2	S	Остатки	R^2	S	Остатки

не-ния			$\sum\Delta^2$	Δ_{max}	R_1			$\sum\Delta^2$	Δ_{max}	R_1
51	0,717	2,54	1542,3	9,30	–0,119	–	–	–	–	–
52	0,726	2,50	1495,0	9,14	–0,181	0,965	0,748	132,0	2,24	0,657
53	0,726	2,50	1493,0	9,20	–0,250	0,970	0,698	114,3	2,55	0,567
54	0,631	2,89	2011,0	10,86	0,541	0,930	1,058	266,2	5,14	0,770
55	0,629	2,85	1943,7	10,06	0,717	–	–	–	–	–
56	–	–	–	–	–	0,967	0,732	126,9	2,33	0,596
57	0,723	2,51	1506,3	9,12	–0,131	–	–	–	–	–
58	–	–	–	–	–	0,979	0,582	79,0	2,19	0,233
59	0,707	2,57	1593,7	8,82	0,146	0,942	0,963	220,6	–	–
60	0,723	2,51	1510,4	9,08	–0,093	0,963	0,772	141,3	2,44	0,647
61	0,670	0,056	0,747	0,29	0,056	0,899	0,029	0,206	–	–
62	0,712	0,052	0,652	0,26	–0,133	0,965	0,017	0,071	0,052	0,715
63	0,633	0,104	2,57	0,57	–0,008	–	–	–	–	–
64	0,959	1,91	869,9	10,53	–0,034	0,995	0,596	84,1	2,28	0,461
65	0,651	1,101	2,46	0,53	–0,136	0,954	0,031	0,225	0,104	0,612

В качестве моделей связи диаметров и высот деревьев в высокопродуктивных березовых древостоях для прогноза роста насаждений рекомендуется применять уравнения Корсуня

$$H=\frac{D^2}{0{,}558+0{,}121D+0{,}031D^2} \qquad (66)$$

и Бакмана $\lg H = -0{,}782 + 2{,}879\lg D - 0{,}944\lg^2 D$. (67)

3.4. Структурные модели

Регрессионные модели, построенные на базе полиномов, носят, как правило, формальный характер. Их используют для описания изучаемых объектов, относительно которых нет достаточно четких количественных представлений о закономерности динамики процесса. Исследователей чаще интересуют содержательные, физические модели, отражающие механизм, сущность явлений.

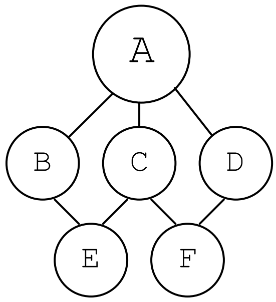
Методы регрессионного анализа позволяют вскрыть зависимости, не поддающиеся непосредственному измерению, однако они не объясняют внутреннюю структуру отношений между компонентами биогеоценоза. При разработке множественной регрессии предполагается отсутствие теории о структуре изучаемого объекта [27]. Исследователь выделяет зависимые и независимые переменные, указывает возможный вид зависимости, т. е. констатирует наличие количественной зависимости результативного признака от факторов. Очень часто, кроме информации о входах и выходах системы, исследователь имеет априорную информацию о логически обоснованной цепи связей между переменными, т. е. о внутренней структуре объекта. Для математического описания и анализа внутренней структуры отношений между компонентами биогеоценоза и его количественными признаками разрабатывают структурные модели [27].

Структурная модель включает систему уравнений, описывающих все связи между переменными. Для разработки структурной модели биогеоценоза необходимо выполнить системный анализ и построить логическую основу модели. При этом удобно воспользоваться матрицами и диаграммами. Понятие «система» главным образом применяется при описании системы как единого целого и связей между частями (элементами) системы, а не описание свойств этих элементов.

Представление системы в виде диаграммы, или графа (рис. 11), показывает, что определенные элементы являются предшествующими (E,F), другие – результативными (B,C,D). Матрица полезна при описании больших систем. Она позволяет выполнить алгебраический анализ системы (сложение и умножение матриц).

Вершины графа обозначают соответствующую переменную, а стрелки указывают направления связей между переменными. Односторонние связи – одна стрелка, двухсторонние связи – стрелки в двух направлениях, т. е. трудно выделить предшествующую и результирующую переменные.

При построении системы структурных уравнений различают: 1) эндогенные переменные, значения которых определяются в результате одновременного взаимодействия переменных; 2) экзогенные переменные, значения которых определяются вне модели. Система структурных уравнений должна содержать столько уравнений, сколько имеется эндогенных переменных, однако число экзогенных переменных может быть больше или меньше числа эндогенных переменных [27].



диаграмма

	A	B	C	D	E	F
A	1					
B	1					
C	1					
D	1					
E		1	1			
F			1			

матрица

Рис. 11. Представление системы в виде диаграммы (графа) и матрицы

В модели (68) имеется только одна результирующая переменная (x_3), которая является зависимой от двух независимых переменных (x_1, x_2). Ошибку (ξ) можно также отнести к независимой переменной. Специальный класс таких моделей:

модель в виде одного уравнения

$$x_3 = a + bx_1 + cx_2 + \xi; \tag{68}$$

рекурсивные модели

$$\begin{aligned} x_1 &= a_{10} + \xi_1; \\ x_2 &= a_{20} + a_{21}x_1 + \xi_2; \\ x_3 &= a_{30} + a_{31}x_1 + a_{32}x_2 + \xi_3; \end{aligned} \tag{69}$$

одновременная модель

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \xi_1 = 0; \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \xi_2 = 0; \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \xi_3 = 0. \end{aligned} \tag{70}$$

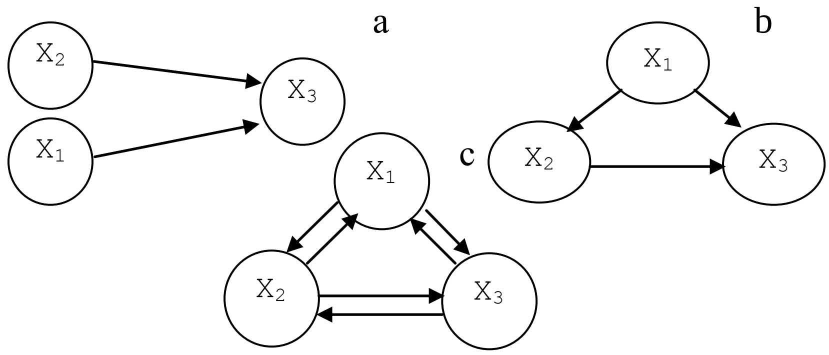


Рис. 12. Структурные модели и графы моделей: a) модель в виде одного уравнения; b) рекурсивные модели; c) одновременная модель

P. Riihinen разделил структурные модели, применяемые в эконометрии, на следующие категории: 1) модель в виде одного уравнения; 2) модель в виде системы уравнений (рекурсивные и одновременные). P. Kilkki выполнил анализ применения этих моделей в лесной таксации для математического описания лесного биогеоценоза и разработки унифицированной системы информации [27].

Модели, включающие только const и (ξ), дают маргинальное распределение зависимых переменных.

В рекурсивной модели (69) одна переменная (x_2) может быть зависимой переменной в одном уравнении и независимой в другом, однако связи всегда односторонние. Одна и та же переменная не может быть одновременно предшествующей и результативной к другой переменной.

В одновременной (совместной) модели (70) одна переменная (x_1) может быть предшествующей и результативной к другой переменной (x_2). Такие переменные являются эндогенными. Переменные, которые не являются исходными (предшествующими) в одновременной модели, называются экзогенными переменными. Значения экзогенных переменных получают вне модели, эндогенные переменные определяются в результате взаимодействия соотношений модели.

Структурная модель может рассматриваться как причинная, если моделируемые связи вскрывают сущность явления и не противоречат его теоретическому объяснению. Для оценивания параметров структурных моделей применяются специальные методы: косвенный метод наименьших квадратов, двух– и трехшаговый методы наименьших квадратов и др. [26]. Американский биолог S. Wrigh в 1934 году предложил метод путевого анализа для исследования причинных связей в структурных моделях. В отличие от корреляционного и регрессионного анализов путевой анализ позволяет выявить не только тесноту связи между различными признаками, но и проанализировать взаимоотношения между признаками, объединенными в определенную систему.