Вин Ньи Ньи, В. Д. Чалый (д. т. н., профессор)
Московский инженерно-физический институт (государственный университет)

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЯХ

Известны методы использования динамических моделей для анализа процесса несанкционированного доступа в компьютерных сетях. В статье предлагаются статические модели в виде уравнений регрессии для описания и последующего исследования подобных задач.

Широкое использование информационных технологий невозможно без кибернетических методов построения статических моделей, полученных в результате проведения эксперимента. Для построения статических моделей в кибернетике широко используется при исследовании различных объектов и явлений метод «черного» ящика с применением моделей типа уравнения регрессии для многофакторных объектов:

$$Y = f(x_1, x_2, ..., x_n, \vec{E}), \qquad (1)$$

где Y — функция отклика, x_i — входные переменные, $\vec{E}$ — случайная помеха,

$$\text{или } Y = f(\vec{x}, \vec{B}) + \vec{E}, \qquad (2)$$

где $\vec{B}$ — вектор коэффициентов.

На рис. 1 показано представление черного ящика, т. е. исследуемого объекта или явления.

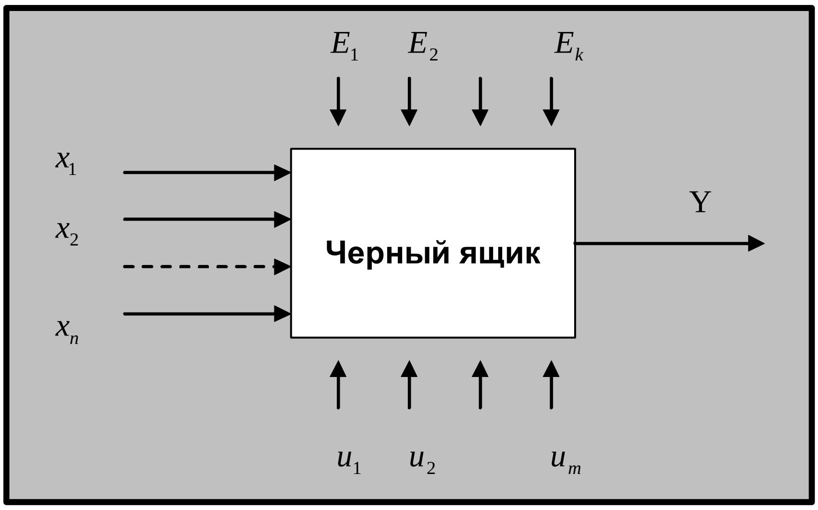

Рис. 1. Представление «черного ящика», где $u_1, u_2, ..., u_m$ — неуправляемые параметры

В последние годы при проведении эксперимента используется планирование входных переменных по различным планам с целью получения адекватной математической модели. В конечном итоге объект или явление описывается уравнением следующего вида:

$$Y = b_0 + \sum_{i=1}^{n} b_i x_i + \sum_{i<j}^{n} b_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i^2 + \sum_{i<j<k}^{n} b_{ijk} x_i x_j x_k + \sum_{i \neq j}^{n} b_{ijj} x_i x_j^2 + \sum_{i=1}^{n} b_{iii} x_i^3 + ... + \qquad (3)$$

Уравнение регрессии (3) может использоваться для построения линейных и нелинейных характеристик объекта. Целесообразно применять активный эксперимент, основанный на методах планирования эксперимента. Пассивный же эксперимент требует значительно большего количества опытов, поэтому признан малоэффективным.

Области применения подобного подхода: научные исследования, технические системы, информационные процессы и их безопасность, химические технологии, экология, медицина, летательные аппараты и др. Подобные методы целесообразно использовать в нанотехнологиях различных направлений.

Планирование эксперимента позволяет после получения математической модели проводить оптимизацию параметров, повышается производительность и качество, резко снижаются затраты на эксперимент. Объект должен быть управляемым и обеспечивать воспроизводимость эксперимента.

Известны методы планирования эксперимента типа полного факторного эксперимента (ПФЭ), дробного факторного эксперимента (ДФЭ), планы второго и третьего порядков. Построение планов эксперимента предлагается осуществлять на основе таких выбранных критериев, как композиционность, ортогональность, симметричность и близость плана к насыщенному. Этот подход позволяет последовательно повышать порядок статической математической модели (рис. 2).

Так, например, строится модель первого порядка, проверяется адекватность уравнения регрессии исследуемому объекту. Если модель адекватна, исследование прекращается. Если неадекватна, к точкам эксперимента первого порядка добавляется дополнительное ядро и вычисляются оценки коэффициентов модели второго порядка. Если модель адекватна, исследование прекращается. Если неадекватна, к предыдущим данным модели второго порядка добавляется еще одно ядро и образуется модель третьего порядка. Если модель третьего порядка адекватна, исследование прекращается, а если неадекватна, то область исследования разбивается на две области.

Таким образом, в подобной методике нет потери экспериментальных точек. Эти условия обеспечивает критерий композиционности. Критерий ортогональности матрицы значений функций независимых переменных обеспечивает простоту вычисления оценок коэффициентов полученной модели.

Близость к насыщенному плану сводит число экспериментальных точек к минимуму. Описанный подход в настоящее время является самым рациональным для тех объектов, где структура модель неизвестна. На рис. 3 показаны условия идентификации.

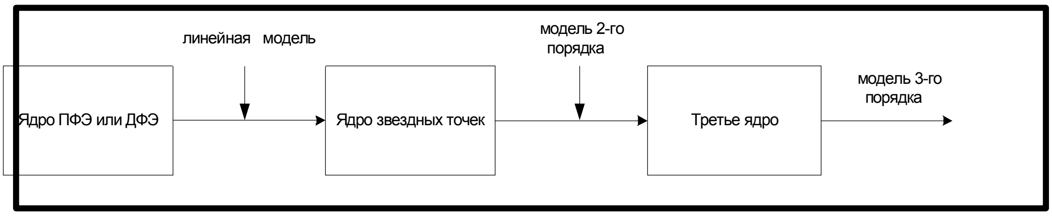

Рис. 2. Общая структурная схема планирования эксперимента

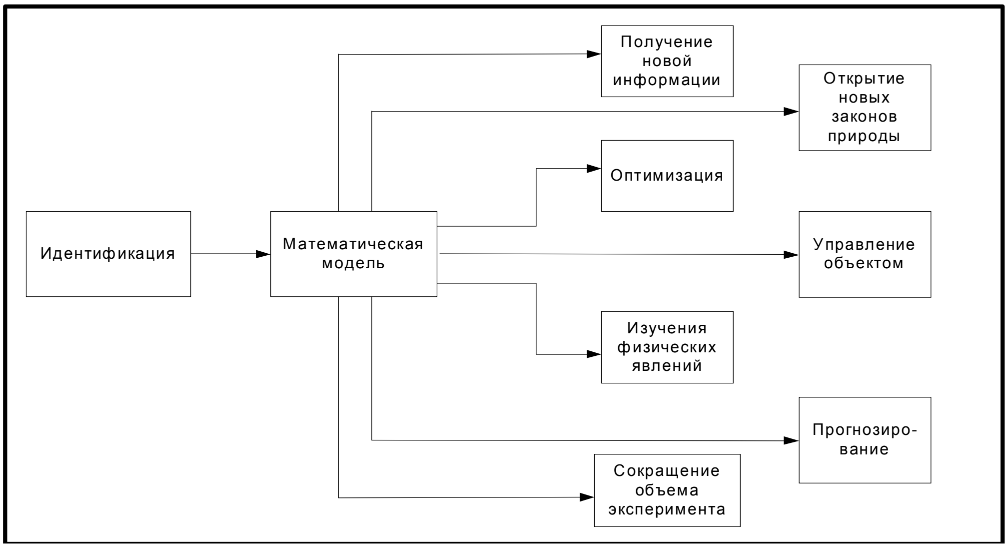

Рис. 3. Структурная схема использования методов планирования эксперимента (идентификации)

Ортогональные центральные композиционные планы (ОЦКП) разработаны Боксом и Уилсоном в 1951 г. для построения статических моделей второго порядка. Исследования показали, что в ОЦКП существует избыточность экспериментальных точек с увеличением числа факторов в объекте исследования. Число уровней варьирования факторов в ОЦКП различно, так, например, для числа управляемых факторов, равного двум, число уровней варьирования равно трем. Это минимальное число уровней варьирования, необходимое для построения модели второго порядка. Поэтому использование двухфакторного ОЦКП второго порядка для построения полной модели третьего порядка невозможно, так как уровней варьирования факторов должно быть не менее четырех.

Неполная модель третьего порядка для двух факторов записывается следующим образом:

$$Y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2 + b_{112}x_1^2x_2 + b_{221}x_2^2x_1 \qquad (4)$$

Для увеличения числа уровней варьирования факторов до пяти надо добавлять еще одно ядро полного факторного эксперимента с новым плечом, которое рассчитывается и отличается от основного уровня. При этих условиях получим полную модель третьего порядка:

$$Y=b_0x_0+b_1x_1+b_2x_2+b_{12}x_1x_2+b_{11}x_1^2+b_{22}x_2^2+b_{112}x_1^2x_2+b_{221}x_2^2x_1+b_{111}x_1^3+b_{222}x_2^3 \qquad (5)$$

Число опытов изменится для n = 2 и рассчитывается по формуле

$$N=2\cdot 2^n+2n+1 \qquad (6)$$

где N — число экспериментальных точек; n — число факторов; 2n — две пары звездных точек на каждый фактор; 2^n — центральная точка.

Таким образом, матрицы двухфакторных ортогональных планов третьего порядка принимают вид:

$$X=\begin{bmatrix} -1 & -1 \\ +1 & -1 \\ -1 & +1 \\ +1 & +1 \\ -0{,}8165 & -0{,}8165 \\ +0{,}8165 & -0{,}8165 \\ -0{,}8165 & +0{,}8165 \\ +0{,}8165 & +0{,}8165 \\ -1 & 0 \\ +1 & 0 \\ 0 & -1 \\ 0 & +1 \\ 0 & 0 \end{bmatrix} \qquad (7)$$

С учетом замены переменной матрица значений функций независимых переменных для неполной модели третьего порядка принимает следующий вид:

$$F = \begin{array}{c} \begin{array}{cccccccc} x_0 & x_1 & x_2 & x_1x_2 & x_1^2-\tfrac{2}{3} & x_2^2-\tfrac{2}{3} & (x_1^2-\tfrac{2}{3})x_2 & (x_2^2-\tfrac{2}{3})x_1 \end{array} \\ \begin{bmatrix} +1 & -1 & -1 & +1 & +0{,}333 & +0{,}333 & -0{,}333 & -0{,}333 \\ +1 & +1 & -1 & -1 & +0{,}333 & +0{,}333 & -0{,}333 & +0{,}333 \\ +1 & -1 & +1 & -1 & +0{,}333 & +0{,}333 & +0{,}333 & -0{,}333 \\ +1 & +1 & +1 & +1 & +0{,}333 & +0{,}333 & +0{,}333 & +0{,}333 \\ +1 & -1 & 0 & 0 & +0{,}333 & -0{,}667 & 0 & +0{,}667 \\ +1 & +1 & 0 & 0 & +0{,}333 & -0{,}667 & 0 & -0{,}667 \\ +1 & 0 & -1 & 0 & -0{,}667 & +0{,}333 & +0{,}667 & 0 \\ +1 & 0 & +1 & 0 & -0{,}667 & +0{,}333 & -0{,}667 & 0 \\ +1 & 0 & 0 & 0 & -0{,}667 & -0{,}667 & 0 & 0 \end{bmatrix} \end{array} \qquad (8)$$

В матрице (8) все вектор-столбцы являются ортогональными, поэтому коэффициенты уравнения регрессии оцениваются независимо друг от друга по формулам:

$$b_0^{'}=\frac{\sum_{u=1}^{N}x_{0u}y_u}{N}=\frac{\sum_{u=1}^{N}y_u}{N};\qquad b_1^{'}=\frac{\sum_{u=1}^{N}x_{1u}y_u}{\sum_{u=1}^{N}x_{1u}^2};\tag{9}$$

$$b_2^{'}=\frac{\sum_{u=1}^{N}x_{2u}y_u}{\sum_{u=1}^{N}x_{2u}^2};\qquad b_{12}^{'}=\frac{\sum_{u=1}^{N}x_{1u}x_{2u}y_u}{\sum_{u=1}^{N}x_{1u}^2x_{2u}^2};\tag{10}$$

$$b_{11}^{'}=\frac{\sum_{u=1}^{N}(x_{1u}^2-2/3)y_u}{\sum_{u=1}^{N}(x_{1u}^2-2/3)^2};\qquad b_{22}^{'}=\frac{\sum_{u=1}^{N}(x_{2u}^2-2/3)y_u}{\sum_{u=1}^{N}(x_{2u}^2-2/3)^2};\tag{11}$$

$$b_{112}^{'}=\frac{\sum_{u=1}^{N}(x_{1u}^2-2/3)x_{2u}y_u}{\sum_{u=1}^{N}[(x_{1u}^2-2/3)x_{2u}]^2};\qquad b_{221}^{'}=\frac{\sum_{u=1}^{N}(x_{2u}^2-2/3)x_{1u}y_u}{\sum_{u=1}^{N}[(x_{2u}^2-2/3)x_{1u}]^2};\tag{12}$$

Уравнение регрессии можно записать в виде:

$$\begin{aligned}Y&=b_0^{'}+b_1^{'}x_1+b_2^{'}x_2+b_{12}^{'}x_1x_2+b_{11}^{'}(x_1^2-2/3)+\\&+b_{22}^{'}(x_2^2-2/3)+b_{112}^{'}(x_1^2-2/3)x_2+b_{221}^{'}(x_2^2-2/3)x_1\end{aligned}\tag{13}$$

Раскрывая скобки и приводя подобные члены, получаем несмещенные оценки коэффициентов

$$b_0=b_0^{'}-2/3(b_{11}^{'}+b_{22}^{'});\quad b_1=b_1^{'}-2/3b_{122}^{'};\quad b_2=b_2^{'}-2/3b_{112}^{'};\quad b_{112}=b_{112}^{'};\quad b_{122}=b_{122}^{'};$$
$$b_{12}=b_{12}^{'};\quad b_{11}=b_{11}^{'};\quad b_{22}=b_{22}^{'};$$

Для (ОЦКП) второго порядка можно использовать для построения неполной модели третьего порядка вида:

$$\begin{aligned}Y&=b_0+b_1x_1+b_2x_2+b_3x_3+b_{12}x_1x_2+b_{13}x_1x_3+b_{23}x_2x_3+b_{11}x_1^2+b_{22}x_2^2+b_{33}x_3^2+\\&+b_{123}x_1x_2x_3+b_{111}x_1^3+b_{222}x_2^3+b_{333}x_3^3\end{aligned}\tag{14}$$

Для ортогонализации вводится замена переменных

$$x_i^3\rightarrow x_i^3-\eta x_i\tag{15}$$

Общая система уравнений для обеспечения ортогональности имеет следующий вид:

$$\begin{aligned}&\sum_{u=1}^{N}x_{ou}(x_{iu}^2-\xi)=\sum_{u=1}^{N}x_{iu}^2-N\xi=0\\&\sum_{u=1}^{N}(x_{iu}^2-\xi)(x_{ju}^2-\xi)=\sum_{u=1}^{N}x_{iu}^2x_{ju}^2-N\xi^2=0\\&\sum_{u=1}^{N}x_{iu}(x_{iu}^3-\eta x_{iu})=\sum_{u=1}^{N}x_{iu}^4-N\eta\xi=0\end{aligned}\tag{16}$$

Подставляя параметры плана в систему уравнений (16) и решая ее, получаем следующие значения параметров плана и параметров замен переменных:

$$\begin{aligned}\alpha&=1{,}215\\\xi&=0{,}730\\\eta&=1{,}1287\end{aligned}\tag{17}$$

Для (ОЦКП) второго порядка можно использовать для построения неполной модели третьего порядка вида:

$$Y=b_0+b_1x_1+b_2x_2+b_3x_3+b_4x_4+b_{12}x_1x_2+b_{13}x_1x_3+b_{14}x_1x_4+b_{23}x_2x_3+b_{24}x_2x_4+ \\ +b_{34}x_3x_4+b_{11}x_1^2+b_{22}x_2^2+b_{33}x_3^2+b_{44}x_2^4+b_{123}x_1x_2x_3+b_{124}x_1x_2x_4+b_{134}x_1x_3x_4+ \\ +b_{234}x_2x_3x_4+b_{111}x_1^3+b_{222}x_2^3+b_{333}x_3^3+b_{444}x_4^3 \quad (18)$$

Для ортогонализации вводится замена переменных:

$$x_i^3 \to x_i^3 - \eta x_i \quad (19)$$

По аналогии решая систему уравнений (16), получаем следующие значения параметров плана и параметров замен переменных:

$$a = 1{,}414214 \\ \zeta = 0{,}800 \\ \eta = 1{,}200 \quad (20)$$

Параметры планов и параметры замен переменных приведены в таблице 1. Для автоматизации оценивания коэффициентов полных и неполных математических моделей третьего порядка разработано программное обеспечение, которое прошло экспериментальную проверку и доказало свою работоспособность.

Таблица 1. Параметры планов и замен переменных.

n	Тип ядра	Число опытов	Число коэффициентов	Значения параметров		Параметры замены переменных	
				β	α	ζ	η
2	2^2	9	8	1	1	0,66667	0,89744
3	2^3	15	14	1	1,21541	0,73030	1,12871
4	2^4	25	23	1	1,414214	0,80000	1,20000
5	2^5	43	36	1	1,596007	0,862662	1,212495

В заключение следует отметить высокую эффективность повышения порядка уравнения регрессии с минимальным объемом экспериментальных точек.

СПИСОК ЛИТЕРАТУРЫ

1. *Налимов В. В., Чернова Н. А.* Статистические методы планирования экстремальных экспериментов. М., 1965.
2. *Марковский М. В., Чалый В. Д.* Технология идентификации и моделирования сложных нелинейных динамических систем // Приборы и системы управления. 1998. № 9.
3. *Бахтин А. В., Чалый В. Д.* Многокритериальные планы эксперимента для построения моделей объектов и процессов. М., 1995. – 116 с.
4. *Марковский М. В., Чалый В. Д.* Информационная технология идентификации динамических объектов: Учебное пособие. М., 1999.
5. *Чалый В. Д.* Планы эксперимента высоких порядков для идентификации объектов: Учебное пособие. М., 1987. – 64 с.
6. *Чалый В. Д.* Теоретические основы идентификации объектов: Учебное пособие. М., 1989. – 56 с.