

ОБ ОДНОЙ ОСОБЕННОСТИ ЗАДАЧИ ИДЕНТИФИКАЦИИ В УСЛОВИЯХ НЕПАРАМЕТРИЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ

Н. А. Дунаева, А. В. Медведев, А. Н. Сергеев, А. И. Шестернев

Сибирский государственный аэрокосмический университет имени академика

М.Ф. Решетнёва

Красноярск, Россия

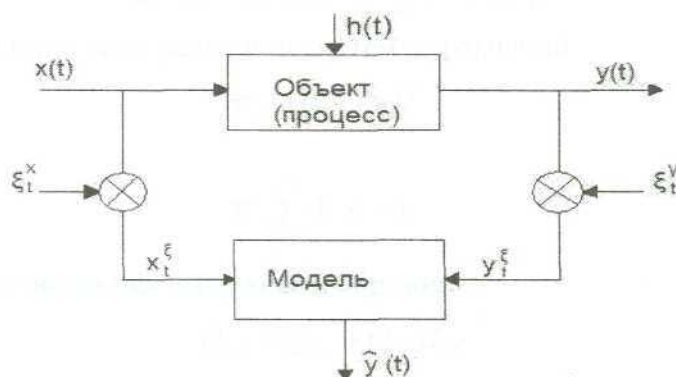
E-mail: saor_medvedev@sibsau.ru

Рассматривается задача идентификации многомерных систем при различном уровне априорной информации. Обсуждаются методы параметрической и непараметрической идентификации процессов трубчатой структуры. Этот класс процессов обусловлен тем, что часто приходится иметь дело с статистически зависимыми компонентами вектора входных переменных. Последнее порождает некоторые особенности при решении задач идентификации многомерных систем. Приведены результаты исследования непараметрических моделей объектов трубчатой структуры методом статистического моделирования.

Ключевые слова: идентификация, непараметрические оценки, непараметрические модели, непараметрические комбинированные модели, процессы трубчатой структуры.

Задача идентификации систем обычно разбивается на два основных этапа [1,2]: определение структуры модели с точностью до вектора параметров и оценка этих параметров по наблюдениям «вход-выход». Эта задача возникает при изучении свойств и особенностей объектов с целью последующего управления ими, либо при создании адаптивных систем, в которых на основе идентификации объекта вырабатываются оптимальные управляющие воздействия.[2]. В зависимости от наличия исходной (априорной) информации об изучаемом процессе различают задачи идентификации в узком и широком смыслах. При моделировании разнообразных дискретно-непрерывных процессов в настоящее время доминирует теория идентификации в узком смысле [1,2].

Прежде, чем изучать этот вопрос, покажем схему наблюдения переменных процесса.



На рисунке 1 $x(t)$ и $y(t)$ – соответственно векторы входных и выходных переменных процесса, $h(t)$ – случайная помеха, действующая на процесс, ξ_t^x и ξ_t^y – случайные помехи, действующие в каналах измерения, x_t^ξ и x_t^ζ – измерения со случайными ошибками соответственно входных и выходных переменных, t – дискретное время. Обозначим оператор, преобразующий $x(t)$ и $h(t)$ в $y(t)$ следующим образом

$$y(t) = A(x(t), h(t)). \quad (1)$$

Задача идентификации в узком смысле состоит в выборе параметризованной структуры модели

$$\hat{y}(t) = A^a(x(t), a) \quad (2)$$

и последующем оценивании параметров a . Класс операторов (моделей) A^a с точностью до набора параметров a .

В зависимости от метода оценивания a в ряде случаев требуется информация о случайных помехах ξ и h . Таким образом, видно, что решение задачи идентификации в узком смысле требует довольно много априорной информации.

Априорная информация об объекте при идентификации в широком смысле существенно беднее. В этом случае известны лишь качественные свойства оператора A , в частности для статических систем с запаздыванием известно, что характеристика объекта однозначна по входным переменным, либо A линейный оператор системы и др. Можно выделить следующие три уровня доступной исследователю априорной информации, от которых во многом зависит вид и способ решения задачи идентификации:

- Параметрический уровень априорной информации предполагает наличие параметрической структуры (2) и плотностей вероятности случайных помех или сведений о различных моментах последних.

- Непараметрический уровень априорной информации не предполагает наличия параметризованной модели типа (2), но требует некоторых сведений качественного характера о процессе (1), например, однозначность, либо неоднозначность его характеристик, для динамических процессов – линейность или тип нелинейности оператора A .

- И наконец, важными, с точки зрения практики, являются задачи идентификации многосвязных систем в условиях, когда объем исходной информации не соответствует ни одному из вышеописанных уровней. Например, для отдельных характеристик многомерного процесса на основании физико-химических закономерностей, энергетических соотношений могут быть выведены параметрические закономерности, а для других нет. Таким образом, мы находимся в ситуации, когда задача идентификации формулируется одновременно в условиях и параметрической, и непараметрической априорной информации.

При непараметрическом подходе для аппроксимации экспериментальных (статистических) данных используются регрессионные модели. Непараметрическая оценка условного математического ожидания $M\{y|x\}$ имеет вид [3]:

$$y_s = \sum_{i=1}^s y_i \Phi\left(\frac{x-x_i}{C_s}\right) / \sum_{i=1}^s \Phi\left(\frac{x-x_i}{C_s}\right), \quad (3)$$

где колоколообразная функция $\Phi(\cdot)$ должна удовлетворять следующим условиям:

$$\Phi\left(\frac{x-t}{C_s}\right) < \infty, \forall x \in \Omega(x), \quad \int_{\Omega(x)} \Phi\left(\frac{x-t}{C_s}\right) dx = 1, \quad \lim_{s \rightarrow \infty} \frac{1}{C_s} \Phi\left(\frac{x-t}{C_s}\right) = \delta(x-t), \quad (4)$$

здесь $\delta(x)$ – дельта-функция Дирака.

На параметр C_s накладываются следующие ограничения, обеспечивающие сходимость непараметрической процедуры:

$$C_s > 0, \lim_{s \rightarrow \infty} C_s = 0, \lim_{s \rightarrow \infty} s C_s = \infty. \quad (5)$$

Для многомерного случая аналог оценки (3) имеет вид

$$y_s(x_1, \dots, x_k) = \sum_{i=1}^s y_i \prod_{j=1}^k \Phi\left(\frac{x^j - x_i^j}{C_s^j}\right) / \sum_{i=1}^s \prod_{j=1}^k \Phi\left(\frac{x^j - x_i^j}{C_s^j}\right), \quad (6)$$

где $x = (x_1, \dots, x_n)$. В этом случае

$$\lim_{s \rightarrow \infty} s C_s^k = \infty. \quad (7)$$

В качестве колоколообразной функции $\Phi\left(\frac{x^j - x_i^j}{C_s^j}\right)$ может быть принята усеченная парабола:

$$\Phi\left(\frac{x^j - x_i^j}{C_s^j}\right) = \begin{cases} 0.335 - 0.067 \left(\frac{x^j - x_i^j}{C_s^j}\right)^2, & \text{если } \left(\frac{x^j - x_i^j}{C_s^j}\right)^2 \leq 5 \\ 0, & \text{если } \left(\frac{x^j - x_i^j}{C_s^j}\right)^2 > 5 \end{cases}, \quad j = \overline{1, n}. \quad (8)$$

Выбор параметра размытости C_s , осуществляется путем минимизации среднеквадратичного критерия

$$W(C_s) = \frac{1}{s} \sum_{j=1}^s (y_j - y_{s_j}(x, C_s))^2 \rightarrow \min_{C_s}, \quad (9)$$

где $i \neq j$. Индекс i фигурирует в оценке (6)

ИДЕНТИФИКАЦИЯ ПРОЦЕССОВ ТРУБЧАТОЙ СТРУКТУРЫ

Известно, что для всякого реального процесса определена область существования входных, выходных переменных $\Omega(x, y)$. Наиболее интересным в данной ситуации представляется рассмотрение процесса, представляющего собой некоторую «объемную полосу» [4], показанную на рисунке 2.

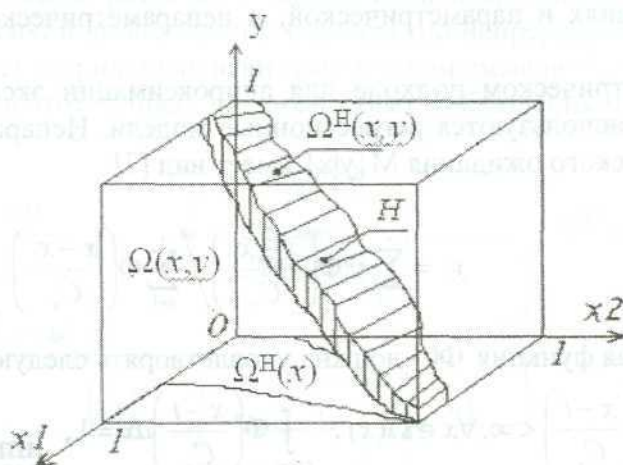


Рис. 2.

Такие процессы назовем Н – процессами и, соответственно, модели таких процессов назовем Н – моделями [4]. Из рисунка 2 видно, что значения Н – процесса принадлежат области $\Omega^H(x, y)$, причем $\Omega^H(x, y) \subset \Omega(x, y)$. Строго говоря область существования переменных (x, y) должна определяться не областью $\Omega(x, y)$, как это сегодня имеет место в теории идентификации, а областью $\Omega^H(x, y)$. Отыскание этой области представляет собой одну из новых задач идентификации. Другой новой задачей идентификации является построение Н – моделей, а не параметрических моделей типа (2), которые обычно рассматриваются при моделировании разнообразных процессов.

При моделировании разнообразных реальных дискретно-непрерывных процессов, мы часто сталкиваемся с ситуацией, когда имеет место зависимость компонент вектора x . Собственно это обстоятельство и порождает трубчатый характер протекания процесса в области $\Omega(x, y)$. Последнее обусловлено тем, что если например x_1 приняло некоторое фиксированное значение, то другие компоненты вектора x могут принять значения во вполне определенной подобласти $\Omega(x)$, т.е. $\Omega^H(x) \subset \Omega(x)$. Таким образом на практике мы часто имеем дело с Н-процессами и, соответственно, с необходимостью строить Н – модели.

Параметрическая идентификация

В задаче параметрической идентификации или идентификации в узком смысле предполагается представление модели объекта с точностью до вектора параметров α , т.е.

$$\hat{y} = \hat{f}(x, \alpha), \quad (10)$$

где $\hat{y}$ – выход модели $\forall x \in \Omega(x) \subset R^n$. При этом $\Omega(x)$ известна но $\Omega^H(x)$ никогда не известна. Методы параметрической идентификации чрезвычайно разнообразны [1,2]. Но при использовании модели (10) или в более общем случае (2) мы сталкиваемся с трудностями, описанными выше. Соответствующая (10) Н-модель имеет вид [5]

$$\hat{y} = \hat{f}(x, \alpha) \Theta^H(x), \quad (11)$$

где $\Theta^H(x)$ – индикаторная функция, например, в качестве оценки $\Theta^H(x)$ может быть

$$\Theta_s^H(x) = \text{sgn} \sum_{i=1}^s \prod_{v=1}^n \Phi \left(\frac{x_v - x_v^i}{C_s} \right), \quad (12)$$

где финитные колоколообразные функции те же что и ранее.

В более общем случае

$$\hat{y}(t) = A^\alpha(x(t), \alpha) \Theta^H(x(t)). \quad (13)$$

Это наиболее общий вид Н-моделей как для статических так и для динамических процессов.

ОЦЕНКА ОБЪЕМА $\Omega^H(X, Y)$

Для оценки объема $\Omega^H(x, y)$ используем идею метода Монте-Карло. Зададимся следующими начальными условиями. Имеется выборка наблюдений за процессом $(x_i, y_i), i = \overline{1, s}$, где s – объем выборки, x – входные переменные $x \in R^k$, y – выходная переменная. Кроме того имеется информация о диапазоне изменения входных переменных – компонент вектора $x = (x_1, x_2, \dots, x_k)$, то есть имеется набор границ областей для каждой из компонент вектора: $\Omega(x_j) = (a_j, b_j), j = \overline{1, k}$.

Для того, чтобы выделить реальную область изменения входных и выходных параметров процесса из общего объема области $\Omega(x, y)$, проведем заданное количество испытаний, каждое из которых описывается следующей схемой:

1. Сгенерируем s значений случайного вектора $x = (x_1, x_2, \dots, x_k)$, которые обозначим $x^i = (x_1^i, x_2^i, \dots, x_k^i)$, $i = \overline{1, n}$. Таким образом, получена равномерная многомерная сетка в $\Omega(x)$.

2. Для выделения трубчатой области $\Omega^H(x, y)$ и оценки ее объема найдем значения индикаторной функции для каждого из сгенерированных в п.1 случайных векторов, в качестве которой можно использовать, например, следующую функцию:

$$\Theta(x^j) = \text{sgn} \sum_{i=1}^s \prod_{v=1}^k \Phi \left(\frac{x_v^j - x_v^i}{C_{s,v}} \right), j = \overline{1, n}. \quad (14)$$

Функция $\Phi(\cdot)$ и параметр размытости $C_{s,v}$ удовлетворяют условиям (4), (7), параметр размытости $C_{s,v}$ определяется также как и в оценке (6)

Объем многомерного куба определяемого областью $\Omega(x, y)$ обозначим V . Поскольку интервалы изменения (x, y) известны, то V легко вычислить. Необходимо найти объем $\Omega^H(x, y)$, который обозначим V^H . Приблизительно V^H можно оценить по формуле

$$V_n^H = \frac{s_0}{s} V, \quad (15)$$

где s_0 – число ненулевых значений $\Theta(x^j)$, $j = \overline{1, n}$.

Помимо нахождения объема области реального протекания процесса, индикаторная функция и использование идеи метода Монте-Карло позволяют непосредственно оценить область $\Omega^H(x, y)$, за пределы которой, процесс не может распространяться.

ПОСТАНОВКА ЗАДАЧИ МОДЕЛИРОВАНИЯ И ИССЛЕДОВАНИЯ СТАТИЧЕСКОГО ОБЪЕКТА ТРУБЧАТОЙ СТРУКТУРЫ

Пусть имеется выборка наблюдений за объектом $(x_i, y_i), i = \overline{1, s}$, где s – объем выборки, x – входные переменные $x \in R^k$, y – выходная переменная. Кроме того имеется информация о диапазоне изменения входных переменных – компонент вектора $x = (x_1, x_2, \dots, x_k)$, то есть имеется набор границ областей для каждой из компонент вектора x : $\Omega(x_j) = (a_j, b_j), j = \overline{1, k}$.

Моделирования и исследования многомерного статического объекта предполагает, в частности, решение следующих задач:

1. Построение модели исследуемого процесса с применением теории комбинированных непараметрических моделей.

2. Исследование с помощью метода Монте-Карло области $\Omega(x, y)$ и выделить из этой области область реального распространения параметров объекта. По результатам проведенных исследований сделать вывод о принадлежности исследуемого объекта к группе Н-объектов и о виде области распространения входных и выходных параметров объекта.

ИМИТАЦИЯ ОБЪЕКТА МОДЕЛИРОВАНИЯ

Для проведения исследований и получения выборок наблюдений за объектом $(x_i, y_i), i = \overline{1, s}$ объект имитировался с помощью генерирования входных и выходных переменных, связи между которыми описываются следующей системой уравнений:

$$\begin{cases} y = (x_1 + 0.5 \exp(0.5x_2) - 0.25 \exp(-2x_3) + 1.5(0.7x_4)^2 + 3.6x_5x_7 + 1.5x_6 + 0.4x_8^2 - 0.8x_9x_{10}) \cdot 0.15, \\ x_2 = 2.3(0.8x_1)^2, \\ x_3 = -0.8(1.12x_1)^2, \\ x_4 = 2.9 \exp(x_1), \\ x_5 = 0.4(1.3x_1)^2, \\ x_6 = -2.5 \exp(0.7x_1), \\ x_7 = 4.8x_1, \\ x_8 = -3.5x_1, \\ x_9 = -0.6(1.25x_1)^3, \\ x_{10} = 2(0.9x_1)^2. \end{cases}$$

ПРОЦЕСС МОДЕЛИРОВАНИЯ

В рамках рассматриваемого подхода построения комбинированных моделей, для проведения численных исследований было предложено использовать комбинированную модель следующего вида:

$$\hat{y}_s = \frac{\sum_{j=1}^s y_j \prod_{i=1}^{10} \Phi\left(\frac{x_i - x'_i}{c'_i}\right)}{\sum_{j=1}^s \prod_{i=1}^{10} \Phi\left(\frac{x_i - x'_i}{c'_i}\right)},$$

$$x_2 = a_2(b_2x_1)^2,$$

$$x_3 = a_3(b_3x_1)^2,$$

$$x_4 = a_4 \exp(b_4x_1),$$

$$x_5 = a_5(b_5x_1)^2,$$

$$x_6 = a_6 \exp(b_6x_1),$$

$$x_7 = a_7x_1,$$

$$x_8 = a_8x_1,$$

$$x_9 = a_9(b_9x_1)^3,$$

$$x_{10} = a_{10}(b_{10}x_1)^2.$$

При моделировании $\Omega(x)$ была принята в виде куба, при этом все компоненты вектора x принадлежали замкнутому интервалу $[0; 3]$.

Построение непараметрической части модели проводилось с использованием элементов теории непараметрической идентификации. Настройка параметров размытости проводилась методом золотого сечения на предварительных этапах перед проведением

основной части численных исследований эффективности построенной комбинированной модели.

Для настройки коэффициентов параметрического блока использовались рекуррентные процедуры типа стохастических аппроксимаций, в частности алгоритмы Литвакова и Кестена [2]. Настройка коэффициентов параметрической части модели проводилась на этапе, предваряющим основные численные исследования.

В качестве критерия адекватности модели для оценки качества настройки коэффициентов, входящих в параметрическую часть модели, и для настройки параметра размытости, входящего в непараметрическую часть модели, использовалась абсолютная ошибка моделирования.

ОПРЕДЕЛЕНИЕ ОБЪЕМА ОБЛАСТИ ПРОТЕКАНИЯ ПРОЦЕССА

В рамках рассматриваемого подхода были проведены исследования, направленные на исследование области распространения процесса и определение ее объема относительно объема гипотетического куба (области), которая рассматривалась бы без предположения о трубчатой структуре исследуемой области.

Исследования были проведены с применением колоколообразных функций двух видов – треугольной и параболической. Исследования проводились при уровне помех в 0, 5, 10, 15 процентов.

Результаты проведенных численных исследований приведены в таблицах 1, 2.

Таблица 1

Результаты численных исследований при использовании колоколообразной функции треугольного вида

Объем выборки	Уровень помех, %	Соотношение объемов
50000	0	0,00023
	5	0,00068
	10	0,00096
	15	0,00167
100000	0	0,00022
	5	0,00071
	10	0,00105
	15	0,00178

Таблица 2

Результаты численных исследований при использовании параболической колоколообразной функции

Объем выборки	Уровень помех, %	Соотношение объемов
50000	0	0,00020
	5	0,00073
	10	0,00108
	15	0,00164

100000	0	0,00019
	5	0,00065
	10	0,00101
	15	0,00172

Изложенные выше модели, описывающие процессы трубчатого характера, могут быть использованы при создании различных систем [6], в том числе адаптивных и обучающихся компьютерных систем моделирования и управления дискретно-непрерывными процессами.

ВЫВОДЫ

Результаты проведенных исследований позволяют сформулировать следующие выводы:

1. Протекание исследуемых процессов в трубке обусловлено стохастической зависимостью компонент вектора входных переменных. На практике это означает, что в случае если одна из компонент вектора x приняла определенное значение, то другие компоненты могут принять значение только из области $\Omega^H(x, y)$. При исследовании методом статистического моделирования H -моделей этот факт использовался следующим образом: значение выходной переменной объекта зависело от всех десяти входных переменных. Зависимости $x_2, \dots, x_{10}$ от x_1 выбирались произвольным образом. На полученные в результате вычислений значения выходной переменной накладывалась аддитивная помеха.

2. Учет информации о принадлежности процесса к процессам с трубчатой структурой распространения входных и выходных параметров позволяет значительно (в несколько сотен и, при небольшой помехе, тысяч раз) сократить область исследования и соответственно объем выборок наблюдений, требуемых для получения приемлемого результата моделирования.

3. Увеличение уровня помехи, накладываемой на выход объекта, приводит к расширению трубчатой области протекания процесса.

ЛИТЕРАТУРА

1. Льюнг, Л. Идентификация систем. Теория для пользователя: Пер. с англ./ Под ред. Я. З. Цыпкина. – М.: Наука, 1991. – 432 с.
2. Цыпкин, Я. З. Основы информационной теории идентификации / Я. З. Цыпкин. – М.: Наука, 1984. – 320 с.
3. Надарая, Э. А. Непараметрические оценки кривой регрессии / Э. А. Надарая // Некоторые вопросы теории вероятностных процессов, Выпуск 5 – Тбилиси: АН Грузинской ССР. 1965. С. 56-68
4. Медведев, А. В. Анализ данных в задаче идентификации. В кн. Компьютерный анализ данных моделирования. Т.2.- Минск: БГУ, 1995, С.201-206.
5. Medvedev, A.V. Non-parametric Identification: Robustness and Computer Intensive Methods / A. V. Medvedev // Computer data analysis and modeling. Minsk: BSU, 2001. С 123-130.
6. Медведев А. В. Непараметрические системы адаптации / А.В. Медведев. – Новосибирск: Наука, 1983. – 174 с.