

УДК 519.254 + 004.93'14

ББК 3.32.965.32.965.9

## **СТРУКТУРНАЯ ИДЕНТИФИКАЦИЯ СЛОЖНЫХ ОБЪЕКТОВ УПРАВЛЕНИЯ НА БАЗЕ МЕТОДОВ КУСОЧНОЙ АППРОКСИМАЦИИ <sup>1</sup>**

**Дорофеев Ю. А.<sup>2</sup>**

*(Учреждение Российской академии наук  
Институт проблем управления РАН, Москва)*

*Решается задача построения модели функционирования сложного объекта с помощью алгоритмов структурно-классификационного анализа и кусочной аппроксимации. Предлагается два подхода к решению этой задачи – при помощи итерационных алгоритмов, реализующих вариационный подход к задачам кусочной аппроксимации, и двухэтапных алгоритмов, в которых процессы структуризации пространства входных параметров и построения локальных регрессионных моделей разделены.*

Ключевые слова: классификационный анализ данных, структурная идентификация, кусочная аппроксимация сложных зависимостей, статистика Фишера-Чоу.

### **1. Постановка задачи**

В работе для простоты рассматривается статическая модель функционирования сложного объекта как модель зависимости выходного показателя  $y$  от вектора входных показателей

---

<sup>1</sup> Работа выполнена при частичной поддержке РФФИ, проекты №08-07-00347-а, №10-07-00210-а.

<sup>2</sup> Юлия Александровна Дорофеев, научный сотрудник (dorofeyuk\_julia@mail.ru).

$$(1) \quad y = F(x), \quad x = (x^{(1)}, \dots, x^{(k)}) \in X \subseteq R^k.$$

Такая модель строится по выборке из  $N$  векторов размерности  $(k + 1)$

$$(2) \quad (y_t, x_t) = (y_t, x_t^{(1)}, \dots, x_t^{(k)}) \in \tilde{X} = \mathbf{R}^{k+1}, \quad t = 1, \dots, N,$$

получаемых в режиме нормальной эксплуатации идентифицируемого объекта. Без особого труда можно показать, что предлагаемый далее подход может использоваться также для идентификации динамической модели достаточно общего вида

$$(3) \quad y(t) = F[x(t), x(t-1), x(t-2), \dots, x(t-m)],$$

где  $m$  – «глубина памяти» динамической модели. Другими словами, различие моделей (1) и (3) состоит только в размерности пространства входов  $X$ , которая увеличивается для (3) до  $km$ .

За критерий качества идентификации, как обычно, принимается среднеквадратичное отклонение выходного параметра у от аппроксимационной модели (функции аппроксимации):

$$(4) \quad J = \int_X [y - \tilde{F}(x)]^2 p(x) dx,$$

где  $p(x)$  – функция плотности распределения вероятности в пространстве  $X$ . Поставленная задача может быть решена при помощи классических статистических методов только в простых случаях. Например, если известно, что  $F(x)$  принадлежит некоторому параметрическому классу функций  $F(x, \alpha)$ , то соответствующая модель  $\tilde{F}(x)$  также может быть выбрана в этом классе  $\tilde{F}(x, \alpha)$ . В этом случае задача сводится к оценке вектора  $\alpha$  по имеющейся выборке наблюдений (2).

Однако в прикладных задачах информация подобного типа часто отсутствует. Более того, сложность функции  $F(x)$  не позволяет использовать классические методы математической статистики. Тем не менее, было замечено, что сложная во всем пространстве  $X$  функция  $F(x)$  может быть представлена в виде совокупности более простых «кусков», определённых на областях  $B_j$ . А именно, предлагается функциональный преобразователь  $F(x)$  (идентифицируемую модель) представлять в виде:

$$(5) \quad F(x) = \sum_{j=1}^r \varepsilon_j(x) F_j(x), \quad \varepsilon_j(x) = \begin{cases} 1, & \text{если } x \in B_j, \\ 0, & \text{если } x \notin B_j; \end{cases}$$

где  $r$  – число областей (классов);  $\varepsilon_j(x)$  – характеристические функции областей разбиения (классификации)

$$(6) \quad H = \{B_j \in X, \bigcup_{j=1}^r B_j = X\}.$$

Такое представление модели является основой метода кусочной аппроксимации. В этом случае аппроксимационная модель может быть представлена в виде

$$(7) \quad \tilde{F}(x) = \sum_{j=1}^r \varepsilon_j(x) \tilde{F}_j(x, \alpha_j),$$

где  $\tilde{F}_j(x, \alpha_j)$  – локальные функции аппроксимации в областях  $B_j$  из выбранного параметрического класса функций. В этом случае функционал (4), соответствующий идентифицируемой модели (7), записывается следующим образом:

$$(8) \quad J = \sum_{j=1}^r \int_{B_j} [y - \tilde{F}_j(x, \alpha_j)]^2 p(x) dx.$$

Тогда задача кусочной аппроксимации идентифицируемой модели состоит в нахождении такого разбиения на классы, для которого сумма квадратов невязок оценок локальных моделей всех классов была бы минимальна. Другими словами, необходимо найти такую классификацию (6) и такие значения векторных параметров  $\alpha_j$ , для которых функционал (8) принимал бы минимальное значение. Вообще говоря, параметр  $r$  (число областей  $B_j$ ) также должен участвовать в минимизации критерия (8). Однако для критерия в форме (8) такая минимизация даёт тривиальный результат – максимально возможное с точки зрения достоверной оценки коэффициентов регрессии  $\alpha_j$ . Очевидно, что это не соответствует интуитивному представлению об «оптимальном» числе областей.

## 2. Методы решения задачи структурной идентификации

Существует два подхода для решения поставленной задачи. Первый подход состоит в формальном рассмотрении функционала (8) и применении некоторого алгоритма его минимизации. Во втором подходе для нахождения областей разбиения (6) и локальных функций аппроксимации  $\tilde{F}_j(x, \alpha_j)$  используются методы распознавания образов и кластеризации.

### 2.1. ВАРИАЦИОННЫЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ СТРУКТУРНОЙ ИДЕНТИФИКАЦИИ

Для разработки алгоритма кусочной аппроксимации в соответствии с первым подходом необходимо рассмотреть первую вариацию функционала (8)  $\delta J$  и разработать алгоритм, обеспечивающий выполнение необходимого условия экстремума функционала  $J$ :  $\delta J = 0$ . Параметр  $r$  не участвует в минимизации функционала, т. е. число областей  $B_j$  задаётся заранее (например, экспертным путём).

Вариация  $\delta J$  разбивается на две независимые части:  $\delta J = \delta_1 J + \delta_2 J$ , где  $\delta_1 J$  – вариация по параметрам  $\alpha_j$  локальных регрессий;  $\delta_2 J$  – вариация по разбиению  $H$ , т. е. по границам областей  $B_j$ . В связи с тем, что вариации  $\delta_1 J$  и  $\delta_2 J$  берутся независимо, необходимое условие экстремума функционала  $J$  может быть переписано в следующей форме:  $\delta_1 J = 0 \cup \delta_2 J = 0$ .

Без ограничения общности, для более компактного формульного представления необходимые условия минимизации функционала далее будем рассматривать для  $r = 2$ .

$$(9) \int_{B_j} [y - \tilde{F}_j(x, \alpha_j)]^2 \nabla_{\alpha_j} \tilde{F}_j(x, \alpha_j) p(x) dx = 0, \quad j = 1, 2,$$

$$(10) \Phi(x, y) = [y - \tilde{F}_1(x, \alpha_1)]^2 - [y - \tilde{F}_2(x, \alpha_2)]^2 = 0, \quad x \in \Lambda,$$

где  $\nabla$  – градиентный оператор,  $\Lambda$  – кусочно-гладкая граница поверхности, разделяющей области  $B_1$  и  $B_2$ ,  $\Phi(x, y)$  – дискриминантная функция.

Для решения системы уравнений (9), (10) предлагается использовать итеративную процедуру типа стохастической аппроксимации [2]:

$$(11a) \quad \alpha_j(n+1) = \alpha_j(n) - \text{sign} \Phi[x(n+1), y(n+1)] \gamma_j(n+1) * \\ * \{y(n+1) - F_j[x(n+1), \alpha_j(n)]\} \nabla_{\alpha_j} \tilde{F}_j[x(n+1), \alpha_j(n)],$$

$$(11б) \quad \Phi[x(n+1), y(n+1)] = \{y(n+1) - \tilde{F}_1[x(n+1), \alpha_1(n)]\}^2 - \\ - \{y(n+1) - \tilde{F}_2[x(n+1), \alpha_2(n)]\}^2, \quad j = 1, 2.$$

Так как выражение (10) для дискриминантной функции  $\Phi(x, y)$  содержит выходной параметр  $y$ , который известен только для данной выборки наблюдений, это решающее правило не может быть использовано для прогнозирования. По этой причине дискриминантная функция должна быть построена как функция  $f(x)$ , зависящая только от входных параметров.

Для того чтобы построить аппроксимацию функции  $f(x)$ , можно использовать обычные алгоритмы распознавания образов с учителем [1]. В этом случае наблюдения (2) используются как обучающая выборка, а значения  $\text{sign} \Phi(x, y)$  рассматриваются как обучающие сигналы, содержащие информацию о том, где расположена точка  $x$ : в  $B_1$  (если  $\text{sign} \Phi(x, y) = 1$ ) или в  $B_2$  (если  $\text{sign} \Phi(x, y) = -1$ ).

Для аппроксимации функции  $f(x)$  можно использовать итерационный алгоритм, основанный на методе потенциальных функций [1, 3]. Этот алгоритм и уравнение (11a) фактически составляют адаптивный алгоритм кусочной аппроксимации.

## 2.2. ДВУХЭТАПНАЯ СХЕМА РЕШЕНИЯ ЗАДАЧИ СТРУКТУРНОЙ ИДЕНТИФИКАЦИИ

Как уже говорилось выше, при решении прикладных задач идентификации было замечено, что многие сложные объекты могут работать в нескольких технологических режимах, суще-

ственно различающихся своими моделями  $y = F_j(x)$ , где  $j$  – индекс режима [4]. При этом  $j$ -му режиму соответствует определённая область  $B_j$  в пространстве входных параметров  $X$ . В [5] для идентификации такого рода объектов впервые было предложено использовать методы кусочной аппроксимации. Обычно в качестве оценок локальных моделей  $\tilde{F}_j(x, \alpha_j)$  используются достаточно простые функции – линейные или даже константы.

В этом случае процедура кусочной аппроксимации состоит из двух этапов.

На первом этапе, используя выборку  $x_1, \dots, x_N$ , пространство  $X$  разбивается на  $r$  областей  $B_j$ , каждая из которых содержит только «близкие» наблюдения  $x_j$  (в соответствии с выбранным критерием близости). В качестве критерия близости обычно используется среднеквадратичное отклонение точек в области  $B_j$  [2, 6]:

$$(12) J = \sum_{j=1}^r \int_{B_j} (x - b_j)^2 p(x) dx,$$

где  $b_j$  – модель (эталон) области  $B_j$ . Для разбиения пространства  $X$  на области  $B_j$  обычно используются алгоритмы автоматической классификации (кластеризации) [4, 6].

На втором этапе по выборке (2) строятся локальные регрессионные модели  $\tilde{F}_j(x, \alpha_j)$ .

Особенность данного подхода состоит в том, что на первом этапе процедуры кусочной аппроксимации используется только информация о входных параметрах. Для большинства сложных объектов частота измерения входных параметров намного выше, чем выходных, поэтому количество входных параметров значительно превышает количество выходных. Классические алгоритмы идентификации, основанные на первом подходе к решению задачи кусочной аппроксимации, могут рассматривать только пары наблюдений  $(x_1, y_1), \dots, (x_N, y_N)$ , в то время как дополнительная информация о входных параметрах в этом случае не используется. В рамках же второго подхода был

разработан алгоритм, который позволяет использовать информацию о выходном параметре  $y$  уже после того, как получено разбиение  $\{B_j, j = 1, \dots, r\}$ . Основная идея этого алгоритма состоит в следующем.

Сначала пространство  $X$  разделяется на области  $B_j$ ,  $j = 1, \dots, l$ , где число  $l$  значительно больше, чем «реальное» число областей  $r$ . Для этой цели в работе использовался комплекс алгоритмов структурно-классификационного анализа [6], включающий алгоритмы:  $m$ -локальной оптимизации заданного критерия, выбора информативных параметров, выбора начального разбиения, выбора в определённом смысле «оптимально» числа классов, заполнения пропущенных наблюдений.

Далее производится пошаговое объединение областей  $B_j$  следующим образом. На каждом шаге находится ближайшая пара областей  $B_i$  и  $B_j$  – кандидатов на объединение. Затем проверяется гипотеза: «локальные модели аппроксимации  $\tilde{F}_i(x, \alpha_i)$  и  $\tilde{F}_j(x, \alpha_j)$  в областях  $B_i$  и  $B_j$  статистически неразличимы». Для этого вводится мера близости  $K(B_i, B_j)$  областей  $B_i$  и  $B_j$  [6] и применяется специальная процедура верификации этой гипотезы. В качестве оценок локальных моделей  $\tilde{F}_j(x, \alpha_j)$  обычно используются линейные функции. По этой причине далее рассматривается только кусочно-линейная модель (аппроксимация). В этом случае для верификации гипотезы использовалась статистика Фишера-Чоу [7]

$$F(k, N_i + N_j - 2k) = \frac{\sum_{s=1}^{N_i+N_j} \delta_s^2 + \sum_{p=1}^{N_i} \delta_p^2 + \sum_{l=1}^{N_j} \delta_l^2}{(N_i + N_j - 2k)^{-1} k \left[ \sum_{p=1}^{N_i} \delta_p^2 + \sum_{l=1}^{N_j} \delta_l^2 \right]},$$

$$(13) \quad \delta_p = [y(x_p) - \tilde{F}(x_p)], \quad x_p \in B_i, \quad \delta_l = [y(x_l) - \tilde{F}_j(x_l)], \quad x_l \in B_j, \\ \delta_s = [y(x_s) - \tilde{F}_{ij}(x_s)], \quad x_s \in B_i \cup B_j,$$

где  $k$  – размерность пространства  $X$ ;  $N_i$  и  $N_j$  – число наблюдений, попавшие в области  $B_i$  и  $B_j$  соответственно;  $\tilde{F}_{ij}(x)$  – локальная модель аппроксимации в объединенной области  $B_i \cup B_j$ . Таким образом,  $\delta$  – это разница между реальными и прогнозируемыми значениями выходного параметра  $y$  при условии, что  $x$  принадлежит к соответствующей области.

Если  $F \leq F_0$ , тогда гипотеза верна, где  $F_0$  – уровень значимости, в противном случае гипотеза отвергается, т. е. области  $B_i$  и  $B_j$  не объединяются.

Таким образом, алгоритм кусочно-линейной аппроксимации (идентификации) состоит в последовательном повторении следующей процедуры. На каждом шаге объединения необходимо найти ближайшую в определённом смысле пару областей  $B_i$  и  $B_j$ , для которых

$$(14) K(B_i, B_j) = \max_{l, p \neq l} K(B_l, B_p).$$

Найденные с помощью (14) области объединяются в новую область  $\hat{B}_i = B_i \cup B_j$ , если  $F \leq F_0$ , т. е. гипотеза «локальные модели аппроксимации  $\tilde{F}_i(x, \alpha_i)$  и  $\tilde{F}_j(x, \alpha_j)$  в областях  $B_i$  и  $B_j$  статистически эквивалентны» верна. Новая локальная модель аппроксимации в объединённой области  $\hat{B}_i$  обозначается как  $\tilde{F}_i^*(x) = \tilde{F}_{ij}(x)$ . Эта процедура повторяется для всех областей  $B_i$  и  $B_j$  (или  $B_i$  и  $\hat{B}_i$ ). В результате, возможно, получатся новые области  $\hat{B}_i$  и, соответственно, новые локальные модели аппроксимации  $\tilde{F}_i^*(x)$ , которые в совокупности дадут оценку итоговой кусочно-линейной модели идентифицируемого объекта, как только закончится процесс объединения областей.

Описанная процедура позволяет построить кусочно-линейную аппроксимацию неизвестной модели идентифицируемого объекта, учитывая геометрическую близость областей  $B_j$  в пространстве  $X$ , а также статистическую различимость

локальных регрессионных функций различных областей. Важное преимущество разработанной процедуры состоит в том, что число областей  $r$  при разбиении пространства  $X$  получается автоматически и оптимальным образом.

### 3. Заключение

Описанные алгоритмы кусочно-линейной аппроксимации были успешно использованы для идентификации сложных объектов управления во многих прикладных задачах. Во всех случаях разработанные алгоритмы показали свою высокую эффективность.

### Литература

1. АЙЗЕРМАН М.А., БРАВЕРМАН Э.М., РОЗОНОЭР Л.И. *Метод потенциалных функций в теории обучения машин.* – М.: «Наука», 1970. – 495 с.
2. БАУМАН Е.В., ДОРОФЕЮК А.А., ДОРОФЕЮК Ю.А. *Методы структурно-классификационного анализа, базирующиеся на процедурах стохастической аппроксимации* // Труды Второй международной конференции Управление развитием крупномасштабных систем (MLSD'2008). – М.: ИПУ РАН, 2008. – С. 192–200.
3. БРАВЕРМАН Э.М., МУЧНИК И.Б. *Структурные методы обработки эмпирических данных.* – М.: Наука, 1983. – 430 с.
4. ДОРОФЕЮК А.А., КАСАВИН А.Д., ТОРГОВИЦКИЙ И.Ш. *Применение методов автоматической классификации для построения статической модели объекта* // Автоматика и телемеханика. – 1970. – №2. – С. 34–40.
5. ДОРОФЕЮК А.А., ТОРГОВИЦКИЙ И.Ш. *Применение методов автоматической классификации данных в задаче контроля качества изделий* // Стандарты и качество. – 1967. – №4. – С. 25–30.

6. ДОРОФЕЮК Ю.А. *Комплекс алгоритмов структурно-классификационного анализа и его использование в задачах анализа и совершенствования крупномасштабных систем управления* // Труды Второй международной конференции Управление развитием крупномасштабных систем (MLSD'2008). Том I. – М.: ИПУ РАН, 2008. – С. 35–38.
7. CHOW G.C. *Tests of Equality between Sets of Coefficients in Two Linear Regressions* // *Econometrica*. – 1960. – Vol. 28, №3. – P. 79–86.

## COMPLEX CONTROL OBJECTS IDENTIFICATION ON THE BASE OF PIECEWISE APPROXIMATION METHODS

**Julia Dorofeyuk**, Institute of Control Sciences of RAS, Moscow, research assistant (dorofeyuk\_julia@mail.ru).

*Abstract: The problem of complex object functioning model development is solved with the aid of algorithms of structure-ranging analysis and piecewise approximation. Two approaches to the solution are considered – the first employs iterative algorithms which implement variational approach to piecewise approximation, while the second uses two-phase algorithms where the processes of input parameter space structuring and local regression model development are separated.*

Keywords: ranging data analysis, structure identification, piecewise approximation of complex dependence, Fisher-Chow statistic.

*Статья представлена к публикации членом редакционной коллегии А. С. Манделем*