ISSN 2782-2001
Системы анализа и обработки данных
том 90, № 2, 2023, с. 7–22

http://journals.nstu.ru/vestnik
Analysis and data processing systems
Vol. 90, No. 2, 2023, pp. 7–22

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ

INFORMATION TECHNOLOGIES AND TELECOMMUNICATIONS

УДК 519.237

DOI: 10.17212/2782-2001-2023-2-7-22

Идентификация локально-адаптивных регрессионных моделей с треугольными индикаторными функциями*

А.А. ПОПОВ

630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет

a.popov@corp.nstu.ru

Основная идея построения локально-адаптивных регрессионных моделей (LAR-моделей) состоит в использовании регрессоров, определенных на локальных подобластях значений факторов. Принадлежность значений факторов той или иной локальной подобласти задается индикаторными функциями. Индикаторные функции по своей природе близки к известным понятиям функций принадлежности из теории нечетких систем (Fuzzy Systems). Как правило, для обеспечения необходимой гладкости искомой зависимости отклика от действующих факторов такие локальные подобласти задаются с перекрытием в виде так называемых нечетких партиций. Тип или вид индикаторных функций может быть самым различным: треугольные, трапецеидальные, нелинейные. Задание того или иного вида индикаторной функции определяет схему взвешивания локальных моделей. Каждая индикаторная функция должна быть определена на всей области действия соответствующего фактора. В качестве индикаторных в работе используются функции треугольного типа. В качестве локальных моделей рассматриваются линейные по факторам модели. Отмечается, что в исходном виде предлагаемые LAR-модели не идентифицируемы. Рассматривается вопрос идентификации подобных моделей в случае совместного оценивания всех параметров. Вводится процедура редукции модели. Результирующая модель выписывается в пространстве функций, допускающих оценку. Для случая разбиения области определения фактора на две, три или четыре нечеткие партиции предлагается базис функций, допускающих оценку. Приводятся результаты вычислительного эксперимента по восстановлению регрессионной зависимости обычными полиномами различной степени и LAR-моделями. Отмечается эффективность LAR-моделей в сравнении с полиномами 3-й и 4-й степеней.

Ключевые слова: локально-адаптивные регрессионные модели, локальные модели, индикаторные функции, треугольные индикаторные функции, нечеткие партиции, функции допускающие оценку, базис функций, допускающих оценку, идентификация

---

* *Статья получена 06 февраля 2023 г.*

ВВЕДЕНИЕ

В реальных ситуациях знания о модели объекта, как правило, не полны. При этом часто применяют методику моделирования по принципу «от простого к сложному», например, от линейной модели к квадратичной и т. д. При отсутствии априорных знаний о структуре модели объекта можно использовать непараметрическую регрессию (см., например, [1]). Одновременно с этим имеется подход, позволяющий, оставаясь в рамке линейной параметрической модели, получать достаточно сложные зависимости отклика от входных факторов. Это технология локально-адаптивных регрессионных моделей, когда поведение объекта в различных частях факторного пространства моделируется отдельными локальными моделями [2–10]. Данная технология имеет важное отличие от известного подхода, связанного с построением кусочно-линейных моделей [11]. При использовании кусочно-линейных моделей поведение отклика описывается набором локальных моделей, жестко состыкованных в точках их переключения. В результате полученная зависимость действительно имеет кусочно-линейный вид. Идейным основанием локально-адаптивного моделирования можно считать подход Takagi и Sugeno (TS-модели) [2]. В этом случае за счет использования функций принадлежности переход от одной локальной модели к другой удается сгладить. Следует отметить, что технология TS-моделирования напрямую никак не связана с первоначальной концепцией Fuzzy Systems, использующей понятия нечетких чисел (см., например, [12]). При использовании технологии локально-адаптивного моделирования результаты наблюдений за входом и выходом объекта считаются четкими величинами. Имеющуюся неопределенность в поведении отклика описывают в терминах теории вероятности. Первоначальная проблема использования локально-адаптивных моделей состоит в том, что они не идентифицируемы [13], что проявляется в невозможности получения однозначных оценок параметров, входящих в модель. Решение этой проблемы возможно через проведение редукции исходных моделей к моделям полного ранга.

1. ЛОКАЛЬНО-АДАПТИВНОЕ РЕГРЕССИОННОЕ МОДЕЛИРОВАНИЕ

Рассмотрим методологию локально-адаптивного регрессионного моделирования.

Пусть модель наблюдения имеет вид

$$y = f^T(x)\theta + e = \sum_{l}^{m} f_l(x)\theta_l + e\,, \tag{1}$$

где $f^T(x) = (f_1(x),\ldots,f_m(x))$ – вектор функций от факторов $x = (x_1,\ldots,x_k)^T$; $\theta = (\theta_1,\ldots,\theta_m)^T$ – неизвестные параметры; e – ошибка, имеющая постоянную дисперсию и нулевое математическое ожидание; y – значение зависимой переменной. В классических методах регрессионного моделирования исходят из того, что векторная функция $f^T(x) = (f_1(x),\ldots,f_m(x))$ априори известна и на ее основе можно адекватно восстановить зависимость. На практике это далеко не всегда выполняется, и приходится прибегать к тем или иным непараметрическим моделям. Основная идея построения локально-адаптивных регрессионных моделей (LAR-моделей) состоит в том, чтобы использовать регрессоры, определенные на локальных подобластях действия факторов. Как правило, для обеспечения необходимой гладкости искомой зависимости отклика от действующих факторов такие локальные подобласти задаются с перекрытием. Пусть области определения факторов $x_1,\dots,x_k$ разбиты на $s_1,\dots,s_k$ локальных подобластей. Для фактора x_j локальные подобласти будем называть нечеткими партициями и обозначать как $\Re_{j1},\Re_{j2},\dots,\Re_{js_j}$, $j=1,\dots,k$. Принадлежность заданного значения фактора той или иной нечеткой партиции будем определять через индикаторные функции $\mu_i(x_1)\in[0,\ 1]$, $i=1,\dots,s_1$; $\mu_j(x_2)\in[0,\ 1]$, $j=1,\dots,s_2$; $\mu_l(x_k)\in[0,\ 1]$, $l=1,\dots,s_k$. Введенное понятие индикаторной функции близко к понятию функции принадлежности для лингвистических переменных [8]. Здесь мы намеренно вместо понятия «функция принадлежности» будем использовать термин «индикаторная функция». Делается это для того, чтобы при рассмотрении локально-адаптивных регрессионных моделей можно было избежать неверную их ассоциацию с нечеткими системами (Fuzzy systems). Однако в ранних работах мы использовали понятие нечетких регрессионных моделей (см., например, [9]), что приводило к некоторой путанице, поскольку, по сути, концепция нечетких чисел и лингвистических переменных не использовалась. Тип (или вид) индикаторных функций может быть самым различным: треугольные, трапецеидальные, нелинейные [8]. Задание того или иного вида индикаторной функции определяет схему взвешивания локальных моделей. Каждая индикаторная функция должна быть определена на всей области действия соответствующего фактора. Как правило, индикаторные функции зависят от ряда параметров, которые можно настраивать. Для равномерного взвешивания всех локальных моделей при получении итоговой регрессионной зависимости необходимо выполнить нормировку индикаторных функций, состоящую в том, что в любой точке факторного пространства выполняются условия:

$$\sum_{i=1}^{s_1}\mu_i(x_1)=1,\quad \sum_{j=1}^{s_2}\mu_j(x_2)=1,\dots,\quad \sum_{l=1}^{s_k}\mu_l(x_k)=1\,. \tag{2}$$

Для простоты изложения ограничимся рассмотрением однофакторных зависимостей. В качестве индикаторных функций будем использовать треугольные. На рис. 1 и 2 приведен вид таких индикаторных функций для случаев разбиения области определения фактора на две и три партиции. Видим, что графическое представление таких индикаторных функций может иметь вид обрезанного или полного треугольника. При разбиении области определения на три и более партиции M применялось правило: ширина отрезка [–1, +1], равная двум, делилась на число половинных оснований треугольных индикаторных функций, равное $2(M-2)$. Полученная величина $\Delta=2/(2(M-2))$ есть ширина основания неполных треугольников, которые располагаются на левом и правом краях отрезка [–1, +1].

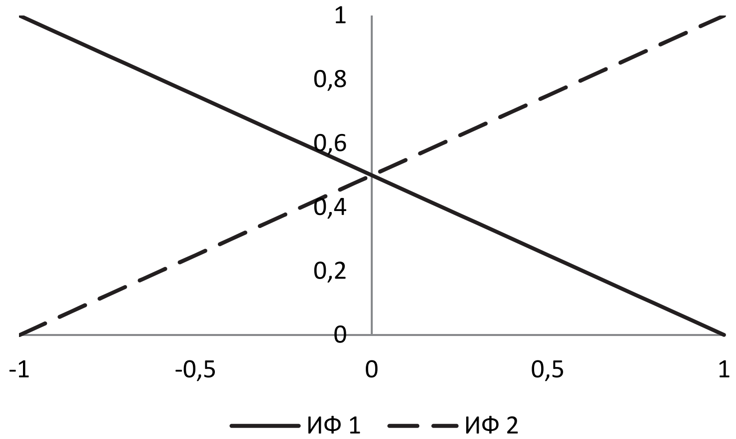


Рис. 1. Разбиение области действия фактора на две пересекающиеся партиции с двумя индикаторными функциями ИФ 1 и ИФ 2

Fig. 1. Partitioning the factor domain into two overlapping partitions with two indicator functions IF 1 and IF 2

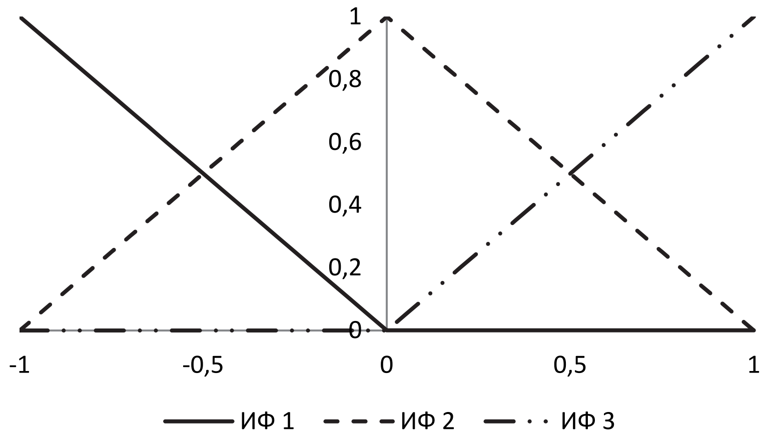


Рис. 2. Разбиение области действия фактора на три пересекающиеся партиции с тремя индикаторными функциями ИФ 1, ИФ 2, ИФ 3

Fig. 2. Partitioning the factor domain into three overlapping partitions with three indicator functions IF 1, IF 2, IF 3

В теории Fuzzy Systems при использовании функций принадлежности треугольного вида в качестве локальных часто используется модель синглтона, состоящая только из свободного члена. Это дает возможность строить LAR-модель в виде нечеткого дерева решений. Этот простейший случай мы сейчас не будем рассматривать. В настоящей работе в качестве локальных будем рассматривать линейные модели. Дерево решений в этом случае можно представить как совокупность правил

$$\Pi_i : If \ \ (x \in \ \Re_i), \ \text{тогда} \ \ y_i' = a_i + b_i x, \ \ i = \overline{1, M}, \tag{3}$$

где M – число нечетких партиций. Объединяя совокупность отдельных правил из (3) через их взвешивание, получим модель наблюдения

$$y_j = \sum_{i=1}^{M} \mu_i(x_j) a_i + \sum_{i=1}^{M} \mu_i(x_j) x_j b_i + \varepsilon_j, \quad j = \overline{1,n}, \tag{4}$$

где n – число наблюдений за входом x_j и выходом y_j. По имеющимся наблюдениям необходимо оценить вектор параметров $\theta = (a_1, a_2, \ldots, a_M, b_1, b_2, \ldots, b_M)$. После оценивания параметров θ предсказывать значение отклика в заданной точке можно по модели

$$\hat{y}(x) = \sum_{i=1}^{M} \mu_i(x) \hat{a}_i + \sum_{i=1}^{M} \mu_i(x) x \hat{b}_i .$$

В матричном виде модель (4) имеет вид

$$y = X\theta + \varepsilon, \tag{5}$$

где $X - (n \times p)$ – расширенная матрица наблюдения; $\theta - (p \times 1)$ – вектор неизвестных параметров; $p = 2M$. Условия (2) по своей природе похожи на соотношения, которые можно наблюдать в моделях дисперсионного анализа, когда качественные факторы представляются в виде совокупности фиктивных переменных [10]. Именно поэтому можно ожидать, что матрица X будет неполного столбцового ранга. Для оценивания всех параметров θ в модели (5) по методу наименьших квадратов в этом случае придется прибегать к использованию обобщенного обращения матрицы $X^T X$. Возможность использовать обычное обращение матриц может появиться, если удастся редуцировать исходную модель (4) к модели полного ранга. Исследуем вопрос идентифицируемости LAR-моделей вида (4) с треугольными индикаторными функциями.

2. ВОПРОСЫ ИДЕНТИФИЦИРУЕМОСТИ ЛОКАЛЬНО-АДАПТИВНЫХ РЕГРЕССИОННЫХ МОДЕЛЕЙ

Столбцы матрицы X соответствуют регрессорам $\mu_1(x)$, $\mu_2(x), \ldots,$ $\mu_M(x)$, $\mu_1(x)x$, $\mu_2(x)x, \ldots,$ $\mu_M(x)x$. Обозначим столбцы матрицы X аналогично, но с нижним подчеркиванием: $\underline{\mu_1(x)}$, $\underline{\mu_2(x)}, \ldots,$ $\underline{\mu_M(x)}$, $\underline{\mu_1(x)x}$, $\underline{\mu_2(x)x}, \ldots,$ $\underline{\mu_M(x)x}$.

Рассмотрение вопроса идентифицируемости LAR-моделей начнем с самого простого случая, когда область определения фактора разбита на две партиции, как на рис. 1. Индикаторные функции, определенные на отрезке $[-1, +1]$, задаются выражениями $\mu_1(x) = (1-x)/2$, $\mu_2(x) = (1+x)/2$. Можно видеть, что между столбцами матрицы X имеются следующие закономерности:

$$\begin{aligned} &\underline{\mu_1(x)} + \underline{\mu_2(x)} = \underline{1}, \\ &\underline{\mu_1(x)x} + \underline{\mu_2(x)x} = \underline{x}, \end{aligned} \tag{6}$$

где $\underline{1}$ – столбец из единиц, $\underline{x}$ – столбец из значений фактора x. При этом вторая зависимость есть следствие первой. Сами по себе зависимости (6) не влияют на ранг матрицы *X*, поскольку столбцы $\underline{1}$ и $\underline{x}$ в нее не входят. Тем не менее матрица X будет иметь неполный столбцовый ранг. Причиной возникновения дефекта ранга матрицы X является тот факт, что индикаторные функции треугольного вида линейно связаны с фактором x. В данном случае это проявляется в том, что $\underline{\mu_2(x)} - \underline{\mu_1(x)} = \underline{x}$. В результате можем зафиксировать, что между столбцами матрицы X имеется зависимость вида $\underline{\mu_2(x)} - \underline{\mu_1(x)} = \underline{\mu_1(x)x} + \underline{\mu_2(x)x}$. Появление этой зависимости напрямую связывается с особенностями треугольных индикаторных функций. При использовании, например, трапециевидных индикаторных функций подобные зависимости не встречаются [13].

Имеется возможность проведения редукции исходной модели наблюдения к модели полного ранга. Способ такой редукции подробно описан в работе [14]. Редукция применялась также и при решении проблемы индентифицируемости деревьев решений с лингвистическими переменными [15]. Редуцирование модели (5) к модели полного ранга проводится через факторизацию матрицы $X = X_1 A$, где X_1 – матрица полного столбцового ранга:

$$y = X\theta + \varepsilon = X_1 A\theta + \varepsilon = X_1\overline{\theta} + \varepsilon\,. \tag{7}$$

В выражении (7) матрица $A = (I_r,\, \tilde{A})$, матрица $X = (X_1,\, X_2)$, где X_1 содержит все линейно независимые столбцы, число которых равно *r*, а X_2 – линейно зависимые; $\tilde{A} = \left(X_1^T X_1\right)^{-1} X_1^T X_2$ является решением матричного уравнения $X_2 = X_1\tilde{A}$. Вектор $\overline{\theta} = A\theta$ есть линейные комбинации исходных параметров. Важно то, что его можно несмещенно оценить по имеющимся наблюдениям. Параметрические функции из числа $\overline{\theta} = A\theta$ называют функциями, допускающими оценку(ФДО) [14]. Совокупность линейно независимых ФДО образует их базис. Важным моментом анализа линейных моделей неполного ранга считается априорное определение базиса ФДО. В случае двух партиций с зависимостями (6) в качестве зависимого столбца матрицы X выберем столбец $\underline{\mu_2(x)}$. Выразим его через другие столбцы: $\underline{\mu_2(x)} = \underline{\mu_1(x)} + \underline{\mu_1(x)x} + \underline{\mu_2(x)x}$. Учитывая данное соотношение, выпишем выражение для A:

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Теперь можем записать следующее утверждение.

У т в е р ж д е н и е 1. Для однофакторной LAR-модели с двумя партициями базис ФДО составляют параметрические функции:

$$\bar{a}_1 = a_1 + a_2,$$
$$\bar{b}_1 = b_1 + a_2, \tag{8}$$
$$\bar{b}_2 = b_2 + a_2.$$

Рассмотрим следующий вариант LAR-модели. Пусть область определения фактора разбита на три партиции, как на рис. 2. Индикаторные функции задаются следующими выражениями:

$$\mu_1(x) = \begin{cases} -x, & -1 \le x \le 0, \\ 0, & x > 0; \end{cases}$$
$$\mu_2(x) = \begin{cases} x+1, & -1 \le x \le 0, \\ 1-x, & 0 < x \le 1; \end{cases} \tag{9}$$
$$\mu_3(x) = \begin{cases} 0, & x \le 0, \\ x, & x > 0. \end{cases}$$

Между столбцами матрицы X для данной модели имеются следующие зависимости:

$$\underline{\mu_1(x)} + \underline{\mu_2(x)} + \underline{\mu_3(x)} = \underline{1},$$
$$\underline{\mu_1(x)x} + \underline{\mu_2(x)x} + \underline{\mu_3(x)x} = \underline{x}, \tag{10}$$
$$\underline{\mu_3(x)} - \underline{\mu_1(x)} = \underline{x}.$$

Возьмем в качестве линейно зависимого столбец $\underline{\mu_3(x)}$. Выразим его через другие линейно независимые: $\underline{\mu_3(x)} = \underline{\mu_1(x)} + \underline{\mu_1(x)x} + \underline{\mu_2(x)x} + \underline{\mu_3(x)x}$. Матрица $\tilde{A}$ будет иметь вид столбца $\tilde{A}^T = (1, 0, 1, 1, 1)$. Учитывая структуру матрицы $A = (I_r, \tilde{A})$, можно сформулировать следующее утверждение.

У т в е р ж д е н и е 2. Для однофакторной LAR-модели с тремя партициями (9) базис ФДО составляют параметрические функции:

$$\begin{aligned} &\bar{a}_1 = a_1 + a_3, \quad \bar{a}_2 = a_2, \\ &\bar{b}_1 = b_1 + a_3, \quad \bar{b}_2 = b_2 + a_3, \quad \bar{b}_3 = b_3 + a_3. \end{aligned} \tag{11}$$

Продолжим рассмотрение вариантов LAR-моделей. Пусть область определения фактора разбита на четыре партиции:

$$
\begin{aligned}
&\mu_1(x)=\begin{cases}(-1/3-x)/(2/3), & -1\le x\le -1/3,\\ 0, & x>-1/3;\end{cases}\\
&\mu_2(x)=\begin{cases}(x+1)/(2/3), & -1\le x\le -1/3,\\ (1/3-x)/(2/3), & -1/3<x\le -1/3,\\ 0 & x>1/3;\end{cases}\\
&\mu_3(x)=\begin{cases}(x+1/3)/(2/3), & -1/3\le x\le 1/3,\\ (1-x)/(2/3) & 1/3<x\le 1,\\ 0 & x<-1/3;\end{cases}\\
&\mu_4(x)=\begin{cases}0, & -1\le x\le 1/3,\\ (x-1/3)/(2/3), & 1/3<x\le 1.\end{cases}
\end{aligned}
\tag{12}
$$

Вид индикаторных функций при разбиении на четыре партиции представлен на рис. 3.

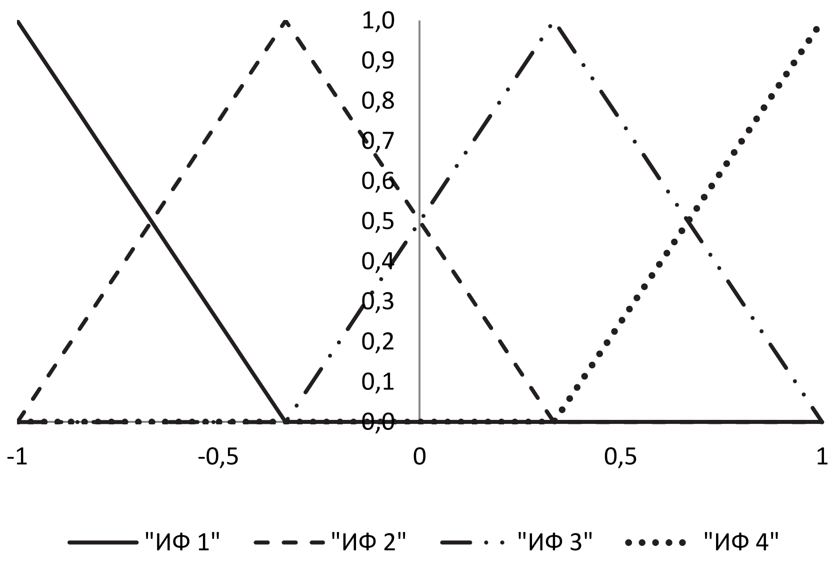


Рис. 3. Разбиение области действия фактора на четыре пересекающиеся партиции с четырьмя индикаторными функциями ИФ 1, ИФ 2, ИФ 3, ИФ 4

Fig. 3. Partitioning the factor domain into four overlapping partitions with four indicator functions IF 1, IF 2, IF 3, IF 4

Между столбцами матрицы X для данной модели имеются следующие зависимости:

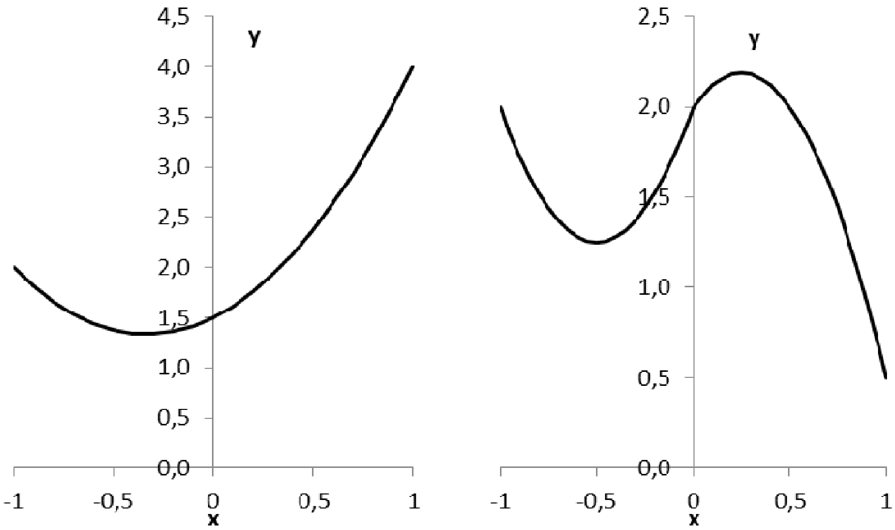
$$\begin{aligned}&\underline{\mu_1(x)}+\underline{\mu_2(x)}+\underline{\mu_3(x)}+\underline{\mu_4(x)}=\underline{1},\\&\underline{\mu_1(x)x}+\underline{\mu_2(x)x}+\underline{\mu_3(x)x}+\underline{\mu_4(x)x}=\underline{x},\\&-\underline{\mu_1(x)}-(1/3)\underline{\mu_2(x)}+(1/3)\underline{\mu_3(x)}+\underline{\mu_4(x)}=\underline{x}.\end{aligned} \tag{13}$$

С учетом этого справедливо следующее утверждение.

Утверждение 3. Для однофакторной LAR-модели с четырьмя партициями (12) базис ФДО составляют параметрические функции:

$$\begin{aligned}&\bar{a}_1=a_1+a_4,\quad \bar{a}_2=a_2+(1/3)a_4,\quad \bar{a}_3=a_3-(1/3)a_4,\\&\bar{b}_1=b_1+a_4,\quad \bar{b}_2=b_2+a_4,\quad \bar{b}_3=b_3+a_4,\quad \bar{b}_4=b_4+a_4.\end{aligned} \tag{14}$$

Рассматривать LAR-модели с числом партиций более четырех нецелесообразно. Описательные возможности LAR-моделей с двумя, тремя и четырьмя партициями не меньше, чем у полиномов степени 2, 3 и 4. На рис. 4 приведены примеры зависимостей, описываемые LAR-моделями с двумя и тремя партициями. Будем обозначать такие модели как LAR-2, LAR-3.



а *б*

Рис. 4. Зависимость, описываемая моделью LAR-2 (*а*); зависимость, описываемая моделью LAR-3 (*б*)

Fig. 4. Dependence described by the model LAR-2 (*a*); dependence described by the model LAR-3 (*b*)

Наилучшие линейные оценки для ФДО с учетом (7) вычисляются как

$$\hat{\bar{\theta}}=\left(X_1^T X_1\right)^{-1} X_1^T y\,. \tag{15}$$

Важный положительный момент использования модели, выписанной в терминах ФДО, состоит в том, что оценивание всех параметрических функций производится совместно. Это позволяет получать их оценки с учетом явной коррелированности. Последний момент не учитывается, если проводить оценку параметров каждой локальной модели в (3) по отдельности.

3. ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

Приведем пример моделирования и оценивания регрессионной LAR-модели с одним фактором, заданным на интервале $[-1,\ +1]$.

Пусть модель с четырьмя партициями LAR-4, порождающая данные, имеет вид

$$E(y/x)=\sum_{i=1}^{M}\mu_i(x)a_i+\sum_{i=1}^{M}\mu_i(x)xb_i=1\mu_1(x)+2.3\mu_2(x)+1.8\mu_3(x)+$$

$$+0.6\mu_4(x)-1\mu_1(x)x+2\mu_2(x)x-0.8\mu_3(x)x-0.1\mu_4(x)x. \qquad (16)$$

График зависимости выхода от входа по данной модели представлен на рис. 5. В модели (16) имеется 8 параметров, но несмещенно оценить можно только 7 из них, представляющих собой ФДО (14). Результаты оценивания по незашумленным и зашумленным данным с дисперсией помехи $\sigma^2=0.1$ и $\sigma^2=0.05$ показаны в табл. 1, в которой для редуцированной модели приведены регрессоры, оценка ФДО для каждого из них и для сравнения истинное значение ФДО.

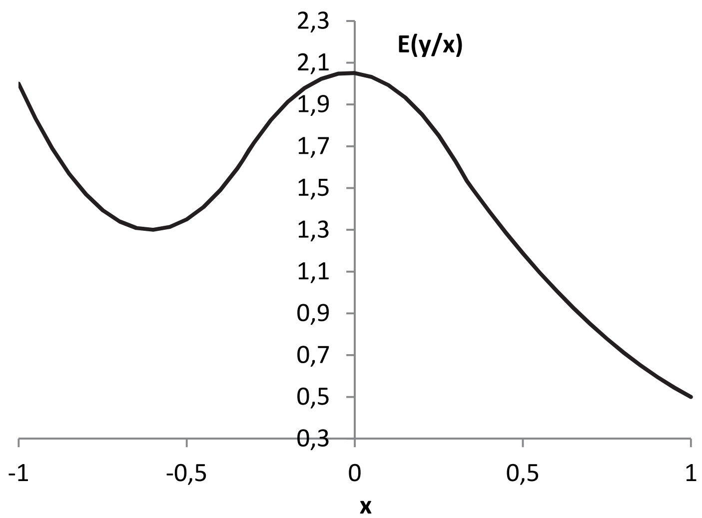


Рис. 5. График зависимости по модели (16)

Fig. 5. Dependence graph for the model (16)

Таблица 1

Table 1

Оценка параметров редуцированной модели

Estimation of the reduced model parameters

Регрессор	Оценка ФДО по незашумленным данным	Оценка ФДО по зашумленным данным, $\sigma^2 = 0.1$	Оценка ФДО по зашумленным данным, $\sigma^2 = 0.05$	Истинное значение ФДО
$\mu_1(x)$	1.6	2.5939	1.87	1 + 0.6 = 1.6
$\mu_2(x)$	2.5	2.6388	2.56	2.3 + 0.2 = 2.5
$\mu_3(x)$	1.6	1.5068	1.63	1.8 – 0.2 = 1.6
$\mu_1(x)x$	–0.4	0.5405	–0.13	–1 + 0.6 = –0.4
$\mu_2(x)x$	2.6	3.3100	2.83	2 + 0.6 = 2.6
$\mu_3(x)x$	–0.2	0.0350	–0.337	–0.8 + 0.6 = –0.2
$\mu_4(x)x$	0.5	0.4388	0.558	–0.1+0.6=0.5

Восстановленная по зашумленным данным с $\sigma^2 = 0.1$ модель в виде графика представлена на рис. 6, где точками обозначены значения отклика.

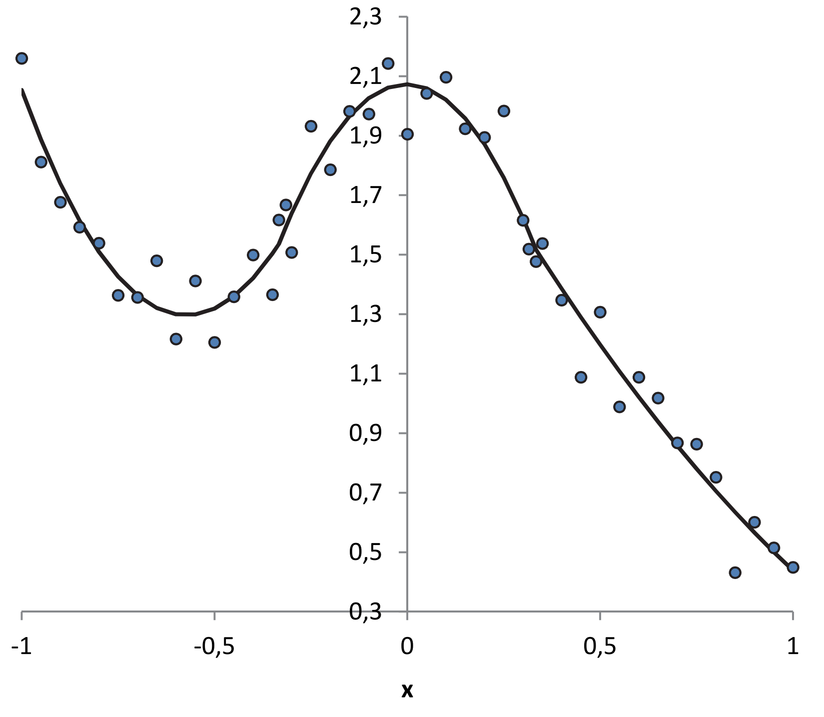


Рис. 6. График восстановленной зависимости

Fig. 6. Graph of the restored dependence

Для сравнения данные, порожденные моделью (16), были аппроксимированы различными полиномами и LAR-моделями. В табл. 2 приведены значения остаточных сумм квадратов при использовании различных моделей.

Таблица 2

Table 2

Остаточные суммы квадратов для различных моделей

Residual sums of squares for various models

Полином степени q				LAR-3	LAR-4
$q = 1$	$q = 2$	$q = 3$	$q = 4$		
6.04	2.98	2.42	0.465	0.497	0.127

По результатам, представленным в табл. 2, можно говорить о том, что классические полиномы не справляются с задачей аппроксимации данных, порождаемых моделью LAR-4 вида (16). Вводя LAR-модель (3), мы отмечали, что в теории Fuzzy Systems большое распространение получили модели синглтона. Если использовать такую модель на четырех партициях для описания данных, порождаемых моделью (16), то остаточная сумма квадратов будет равна 2.48, что очень близко к значению, достигаемому при использовании полинома степени 3. В то же время модель синглтона значительно проигрывает модели LAR-4. В работе [16] с использованием концепции локально-адаптивных моделей решена практическая задача построения стойкостной модели сверления. Полученные результаты убедительно показывают преимущество построенной LAR-модели перед ранее использованными полиномами степени 2.

ЗАКЛЮЧЕНИЕ

В работе впервые отмечен факт неидентифицируемости LAR-моделей с треугольными индикаторными функциями, который возникает при попытке провести оценивание параметров всех локальных моделей одновременно. Данный эффект не имеет места быть в случае использования индикаторных функций трапециевидного или нелинейного вида. Для LAR-моделей с линейными локальными моделями и числом партиций от двух до четырех предложен базис ФДО, которые можно несмещенно оценивать. Это позволяет исходные LAR-модели априори выписывать в пространстве ФДО и проводить идентификацию таких моделей. В вычислительном эксперименте продемонстрировано, что качество аппроксимации данных подобными LAR-моделями превышает качество аппроксимации обычными полиномами степеней 2, 3 и 4, а также LAR-моделями в виде синглтона.

СПИСОК ЛИТЕРАТУРЫ

1. *Vapnik V.* Statistical Learning Theory. – New York: John Wiley, 1998. – 736 p.

2. *Takagi T., Sugeno M.* Fuzzy identification of systems and its applications to modeling and control // IEEE Transactions on Systems, Man and Cybernetics. – 1985. – Vol. 15, N 1. – P. 116–132.

3. *Dickerson J.A., Kosko B.* Fuzzy function approximation with ellipsoidal rules // IEEE Transaction on Fuzzy Systems. – 1996. – Vol. 26, N 4. – P. 542–560.

4. *Kosko B.* Fuzzy systems as universal approximators // Proceedings First IEEE International Conference on Fuzzy Systems. – San Diego, CA, USA, 1992. – P. 1153–1162. – DOI: 10.1109/FUZZY.1992.258720.

5. *Butkiewicz B., Rutkowski L., Kacprzyk J.* Simple modification of Takagi – Sugeno model // Neural Networks and Soft Computing. – 2003. – Vol. 11. – P. 504–509.

6. *Babuška R.* Fuzzy modeling for control. – London; Boston: Kluwer Academic Publ., 1998. – 257 p.

7. *Lilly J.H.* Fuzzy control and identification. – Hoboken, NJ: Wiley, 2010. – 231 p.

8. *Пегат А.* Нечеткое моделирование и управление: пер. с англ. – 2-е изд. – М.: Бином, 2013. – 798 с.

9. *Попов А.А.* Регрессионное моделирование на основе нечетких правил // Сборник научных трудов НГТУ. – 2000. – № 2 (19). – С. 49–57.

10. *Ходашинский И.А., Сарин К.С.* Методика построения компактных и точных нечетких систем типа Такаги – Сугено // Доклады ТУСУР. – 2014. – Т. 19, № 1. – С. 50–56.

11. *Котюков В.И.* Многофакторные кусочно-линейные модели. – М.: Финансы и статистика, 1984. – 216 с.

12. *Yen K.K., Ghoshray S., Roig G.* A linear regression model using triangular fuzzy number coefficients // Fuzzy Sets and Systems. – 1999. – Vol. 106, iss. 2. – P. 167–177.

13. *Попов А.А.* Идентификация локально адаптивных регрессионных моделей // Обработка информации и математическое моделирование: материалы конференции, Новосибирск, 23–24 апр. 2020 г. – Новосибирск, 2020. – С. 155–160. – URL: https://sibsutis.ru/workgroups/w/group/46/files/Материалы%20конференций/РНТК-2020 (дата обращения: 29.05.2023).

14. *Попов А.А.* Конструирование дискретных и непрерывно-дискретных моделей регрессионного типа // Сборник научных трудов НГТУ. – 1996. – № 1 (3). – С. 21–30.

15. *Попов А.А.* Построение деревьев решений для прогнозирования количественного признака на классе логических функций от лингвистических переменных // Научный вестник НГТУ. – 2009. – № 3 (36). – С. 77–86.

16. *Попов А.А., Карманов В.С.* Построение стойкостной модели сверления с использованием локально адаптивных регрессионных моделей // Динамика систем, механизмов и машин. – 2020. – Т. 8, № 1. – С. 141–146. – DOI: 10.25206/2310-9793-8-1-141-146.

Попов Александр Александрович, доктор технических наук, профессор кафедры теоретической и прикладной информатики Новосибирского государственного технического университета. Основное направление научных исследований – статистические методы анализа данных, оптимальное планирование экспериментов, методы машинного обучения. Имеет более 250 публикаций, в том числе 3 монографии. E-mail: a.popov@corp.nstu.ru

Popov Alexander Alexandrovich, D.Sc. (Eng.), Professor, Department of Theoretical and Applied Computer Science, Novosibirsk State Technical University. The main area of his research is statistical methods of data analysis, optimal design of experiments, and methods of machine learning. He has more than 250 publications, including 3 monographs. E-mail: a.popov@corp.nstu.ru

DOI: 10.17212/2782-2001-2023-2-7-22

Identification of locally-adaptive regression models with triangular indicator functions[*]

A. A. POPOV

Novosibirsk State Technical University, 20 K. Marx Prospekt, Novosibirsk, 630073, Russian Federation

a.popov@corp.nstu.ru

Abstract

The basic idea of constructing locally-adaptive regression models (LAR models) consists in the use of regressors defined on the local subregions of factor values. The belonging of factor values to a particular local subdomain is set by indicator functions. Indicator functions by their nature are close to the well-known concepts of membership functions from the theory of fuzzy systems (Fuzzy Systems). As a rule, to provide the required smoothness of the required dependence of the response on the acting factors such local subdomains are defined with overlapping - in the form of the so-called fuzzy partitions. Type or type of indicator functions may be very different: triangular, trapezoidal, and non-linear. Specifying one or another type of indicator function determines the scheme of weighing local models. Each indicator function must be defined for the entire range of the corresponding factor. Triangular-type functions are used as indicator functions in this work. Linear factor models are considered as local models. It is noted that in their original form the proposed LAR models are not identifiable. The issue of identification of such models in the case of joint estimation of all parameters is considered. The procedure of model reduction is introduced. The resulting model is written out in the space of functions that allow estimation. In the case of dividing the domain of factor determination into two, three or four fuzzy partitions we propose the basis of functions allowing evaluation. The results of computational experiment on regression dependence reconstruction by ordinary polynomials of different degrees and by LAR models are given. The efficiency of LAR models in comparison with polynomials of degree 3 and 4 is noted.

Keywords: local-adaptive regression models, local models, indicator functions, triangular indicator functions, fuzzy partitions, estimated functions, basis of estimated functions, identification

REFERENCES

1. Vapnik V. *Statistical Learning Theory*. New York, John Wiley, 1998. 736 p.

2. Takagi T., Sugeno M. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics*, 1985, vol. 15, no. 1, pp. 116–132.

3. Dickerson J.A., Kosko B. Fuzzy function approximation with ellipsoidal rules. *IEEE Transaction on Fuzzy Systems*, 1996, vol. 26, no. 4, pp. 542–560.

4. Kosko B. Fuzzy systems as universal approximators. *Proceedings First IEEE International Conference on Fuzzy Systems*, San Diego, CA, USA, 1992, pp. 1153–1162. DOI: 10.1109/FUZZY.1992.258720.

5. Butkiewicz B., Rutkowski L., Kacprzyk J. Simple modification of Takagi-Sugeno model. *Neural Networks and Soft Computing*, 2003, no. 11, pp. 504–509.

6. Babuška R. *Fuzzy modeling for control*. London, Boston, Kluwer Academic Publ., 1998. 257 p.

7. Lilly J.H. *Fuzzy control and identification*. Hoboken, NJ, Wiley, 2010. 231 p.

8. Piegat A. *Fuzzy modeling and control*. Heidelberg, Physica-Verlag, 2001 (Russ. ed.: Pegat A. *Nechetkoe modelirovanie i upravlenie*. 2nd ed. Moscow, Binom Publ, 2013. 798 p.).

[*] *Received 06 February 2023.*

9. Popov A.A. Regressionnoe modelirovanie na osnove nechetkikh pravil [Regression modeling based on fuzzy rules]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta = Transaction of scientific papers of the Novosibirsk state technical university*, 2000, no. 2 (19), pp. 49–57.

10. Hodashinsky I.A., Sarin K.S. Metodika postroeniya kompaktnykh i tochnykh nechetkikh sistem tipa Takagi–Sugeno [Technique for designing accurate and compact Takagi–Sugeno fuzzy systems]. *Doklady TUSUR = Proceedings of TUSUR University*, 2014, vol. 19, no. 1, pp. 50–56.

11. Kotyukov V.I. *Mnogofaktornye kusochno-lineinye modeli* [Multifactor piecewise linear models]. Moscow, Finansy i statistika Publ., 1984. 216 p.

12. Yen K.K., Ghoshray S., Roig G. A linear regression model using triangular fuzzy number coefficients. *Fuzzy Sets and Systems*, 1999, vol. 106, iss. 2, pp. 167–177.

13. Popov A.A. [Identification of locally adaptive regression models]. *Obrabotka informatsii i matematicheskoe modelirovanie: materialy konferentsii* [Information processing and mathematical modeling]. Conference proceedings, Novosibirsk, 23–24 April, 2020, pp. 155–160. (In Russian). Available at: https://sibsutis.ru/workgroups/w/group/46/files/Материалы%20конференций/РНТК-2020 (accessed 29.05.2023).

14. Popov A.A. Konstruirovanie diskretnykh i nepreryvno-diskretnykh modelei regressionnogo tipa [Construction of discrete and continuous-discrete models such as regression]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta = Transaction of scientific papers of the Novosibirsk state technical university*, 1996, no. 1 (3), pp. 21–30.

15. Popov A.A. Postroenie derev'ev reshenii dlya prognozirovaniya kolichestvennogo priznaka na klasse logicheskikh funktsii ot lingvisticheskikh peremennykh [Construction of decision trees to predict the quantitative trait in the class of logic functions of linguistic variables]. *Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta = Science bulletin of the Novosibirsk state technical university*, 2009, no. 3 (36), pp. 77–86.

16. Popov A.A., Karmanov V.S. Postroenie stoikostnoi modeli sverleniya s ispol'zovaniem lokal'no adaptivnykh regressionnykh modelei [Building a wear resistance model of drilling operation using locally adaptive regression models]. *Dinamika sistem, mekhanizmov i mashin = Dynamics of Systems, Mechanisms and Machines*, 2020, vol. 8, no. 1, pp. 141–146. DOI: 10.25206/2310-9793-8-1-141-146.

Для цитирования:

Попов А.А. Идентификация локально-адаптивных регрессионных моделей с треугольными индикаторными функциями // Системы анализа и обработки данных. – 2023. – № 2 (90). – С. 7–22. – DOI: 10.17212/2782-2001-2023-2-7-22.

For citation:

Popov A.A._Identifikatsiya lokal'no-adaptivnykh regressionnykh modelei s treugol'nymi indikatornymi funktsiyami [Identification of locally-adaptive regression models with triangular indicator function]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2023, no. 2 (90), pp. 7–22. DOI: 10.17212/2782-2001-2023-2-7-22.