

2. *Chen Zesheng, Gao Lixin, Kwait Kevin*. Modeling the spread of active worms // INFOCOM 2003 [Электронный ресурс]. – 2003. – Режим доступа: http://www.ieee-infocom.org/2003/papers/46_03.PDF. – Дата доступа: 19.03.2010 г.
3. *Williamson M.M., Leveille J*. An epidemiological model of virus spread and cleanup // HPL-2003-39 [Электронный ресурс]. – 2003. – Режим доступа: <http://www.hpl.hp.com/techreports/2003/HPL-2003-39.pdf>. – Дата доступа: 19.03.2010 г.
4. *Ревюз Д.* Цепи Маркова. – М.: РФФИ, 1997.

Статья представлена к публикации членом редколлегии Чье Ен Уном.

E-mail:

Далингер Яков Михайлович – iakovdalinger@gmail.com;

Бабанин Дмитрий Владимирович – iakovdalinger@gmail.com;

Бурков Сергей Михайлович – burkov@khb.ru.

УДК 681.51

© 2012 г. **Г.Б. Диго,**

Н.Б. Диго,

А.Ю. Торгашов, д-р техн. наук

(Институт автоматизации и процессов управления ДВО РАН, Владивосток),

И.С. Можаровский

(Владивостокский государственный университет экономики и сервиса)

АНАЛИЗ ИДЕНТИФИЦИРУЕМОСТИ НЕЛИНЕЙНЫХ ОБЪЕКТОВ УПРАВЛЕНИЯ СЛАБО ФОРМАЛИЗОВАННОЙ СТРУКТУРЫ*

Предлагается методика определения идентифицируемости сложных нелинейных объектов при их неизвестной структуре на основе алгоритма чередующихся условных математических ожиданий. Приводятся примеры, иллюстрирующие ее применение.

Ключевые слова: модель, идентифицируемость, виртуальный анализатор, массообменные процессы, алгоритм чередующихся условных математических ожиданий.

Введение

Подходы, обычно используемые при идентификации объектов и систем, основаны на предположении о возможности получения аналитически заданной функциональной зависимости с последовательным уточнением значений ее коэффициентов. Несмотря на их постоянное усовершенствование, они применимы только к хорошо формализуемым объектам.

* Работа выполнена при частичной поддержке грантов ДВО РАН № 12-I-П17-02 и № 12-I ОЭММПУ-04 и гранта ДВО-РФФИ № 11-08-98500-р_восток_a.

В реальных условиях большинство объектов является слабо формализуемыми из-за недостаточности имеющихся знаний о них и среде, в которой они функционируют. Управление ими представляет большой интерес, поэтому актуально направление научных исследований, связанное с изучением и построением сложных и слабо формализуемых управляемых объектов и систем.

Если методы, разработанные для моделирования и качественного анализа хорошо формализованных сложных объектов, основаны на алгебраических, логико-лингвистических и нечетко-логических моделях принятия решения, то при применении к слабо формализованным системам необходимо их определенное развитие с учетом особенностей конкретных объектов. Так, используемые в настоящее время численные методы анализа параметров моделей, как правило, носят апостериорный характер, включают полученные по данным эксперимента численные оценки значений параметров, хотя иногда обращаются и к теоретическим оценкам. Следует отметить, что вопросам их идентифицируемости (возможности оценить значения параметров математической модели объекта в условиях заданных ограничений) посвящено очень ограниченное число работ [1, 2]. Именно поэтому оказывается актуальной проблема разработки численного алгоритма анализа идентифицируемости параметров. Она соответствует практическим нуждам исследователей при разработке моделей физико-химических систем, – например, в установлении единственности числового значения требуемого параметра безотносительно к тому, допускают ли определение остальные параметры.

При не известных структуре модели и, соответственно, количестве переменных, включаемых в модель (например, при идентификации моделей промышленных массообменных процессов для *прогноза* показателей качества продуктов [3]), возникает нестандартная задача идентифицируемости, для которой отсутствуют методы решения и более приемлемым становится термин «идентифицируемость объекта». При этом под идентифицируемостью объекта понимается возможность построения его модели на основе имеющихся данных вход-выход.

В статье предложена методика определения идентифицируемости сложных нелинейных процессов при неизвестной структуре модели на основе алгоритма чередующихся условных математических ожиданий (alternating conditional expectation – ACE) [4, 5]. Представлены примеры, иллюстрирующие применение разработанной методики.

Постановка задачи

Рассматривается объект, функционирующий в условиях неопределенности с m входными контролируруемыми переменными x_1, \mathbf{K}, x_m и выходом y , связанными между собой некоторой функциональной зависимостью

$$y = F(\mathbf{X}, \mathbf{B}) + e, \quad (1)$$

где $\mathbf{X} = (x_1, \mathbf{K}, x_m)$ – вектор входных контролируемых переменных; e – действующее на выходе неконтролируемое возмущение; $\mathbf{B} = (b_1, \mathbf{K}, b_m)$ – вектор коэффициентов. Пусть о виде истинной зависимости (структуры) F в (1) ничего не

известно. В такой ситуации обычно приходится обращаться к регрессионным моделям и выбирать структуру, опираясь на имеющуюся априорную информацию об исследуемом объекте. Очевидно, что это связано с вычислительными проблемами, возникающими при формировании и анализе нескольких вариантов обычно многомерных нелинейных моделей, – например, по параметрам вида

$$\hat{y}_j = b_0 + \sum_{i=1}^p b_i f_{ij}, \quad j = 1, \mathbf{K}, n, \quad (2)$$

где \hat{y}_j – предсказанное значение, соответствующее j -му значению y_j выходной переменной y ; $f_{ij} = f_{ij}(x_1, \mathbf{K}, x_m)$, $i = 1, \mathbf{K}, p$, $j = 1, \dots, n$ – некоторые функции входных контролируемых переменных x_1, \mathbf{K}, x_m ; $b_0, b_1, \mathbf{K}, b_p$ – оценки коэффициентов из (1). В общем случае их количество может быть бесконечным, но заранее неизвестно, удастся ли выбрать среди них модель, адекватно описывающую зависимость (1) в условиях заданных ограничений на функционирование объекта.

Ставится задача разработки алгоритма, оценивающего идентифицируемость сложных нелинейных объектов при неизвестной структуре.

Анализ задачи

Очевидно, что одним из путей отбора входных переменных, влияющих на значение выхода, и выбора структуры адекватной модели является применение корреляционного и регрессионного анализа для оценки корреляционных зависимостей между рассматриваемыми переменными. Однако для слабо формализованных нелинейных объектов он обычно не вносит ясности в выбор структуры модели. После анализа имеющихся в литературе численных методов решения поставленной задачи и результатов, полученных при построении виртуальных анализаторов для массообменных процессов с помощью алгоритма АСЕ [6], для разработки методики оценивания идентифицируемости объектов выбран подход на основе алгоритма АСЕ. Это обосновано, в первую очередь, тем, что оптимальные преобразования в нем не требуют априорных предположений о функциях, связывающих выход и входные переменные, а поэтому он является мощным средством при исследовании данных и их взаимосвязей.

При использовании алгоритма АСЕ переходим от модели вида (2) к выражению

$$q(Y) = a + \sum_{i=1}^p \Phi_i(X_i) + e, \quad (3)$$

где q – функция выходной переменной U ; Φ_i – функции входов X_i , $i = 1, \dots, p$; e – ошибка, т.е. переходим от классического уравнения регрессии (2), линейного по параметрам, к уравнению типа (3).

Алгоритм АСЕ для заданного набора данных, состоящего из выходной переменной U и входных переменных X_1, X_2, \dots, X_p , начинает свою работу с вычисления некоторых начальных (базовых) преобразований $q(Y)$, $\Phi_1(X_1), \dots, \Phi_p(X_p)$. По ним вычисляется

$$e^2(q, \Phi_1, \dots, \Phi_p) = E\{[q(Y) - \sum_{i=1}^p \Phi_i(X_i)]\}^2 \quad (4)$$

и находятся оптимальные преобразования зависимых и независимых переменных, обеспечивающие максимальную корреляцию между ними (преобразованными зависимыми и независимыми переменными). Они получаются минимизацией по $\Phi_1(X_1), \dots, \Phi_p(X_p)$ и $q(Y)$ остаточной дисперсии e^2 из (4) линейной зависимости между преобразованной зависимой переменной и суммой преобразованных независимых переменных

$$\Phi_i(X_i) = E[q(Y) - \sum_{j=1}^p \Phi_j(X_j) | X_i], \quad i = 1, 2, \dots, p,$$

$$q(Y) = E[\sum_{i=1}^p \Phi_i(X_i) | Y] / \left\| E[\sum_{i=1}^p \Phi_i(X_i) | Y] \right\|,$$

последовательно по каждой из переменных, для которых вычисляются условные математические ожидания [2]. В результате получаем оптимальные преобразования $\Phi_i^*(X_i), i = 1, \dots, p$ и $q^*(Y)$, связанные выражением

$$q^*(Y) = \sum_{i=1}^p \Phi_i^*(X_i) + e^*,$$

где e^* – ошибка, не устраняемая использованием алгоритма АСЕ.

Алгоритм оценки идентифицируемости нелинейного объекта

Исходной информацией для оценки идентифицируемости служат базовая матрица (матрица данных, сформированная из выборки, содержащей значения входных и выходной переменных), число M возмущенных матриц данных, получаемых из базовой матрицы путем добавления к ее элементам малых случайных чисел.

Алгоритм идентифицируемости объекта:

1) базовую матрицу преобразуем в расширенную матрицу данных размера $(K \times (n+2))$, включая в нее некоррелируемый с выходом дополнительный нормально распределенный вход ξ с параметрами $0; 1$ ($\xi \in N(0, 1)$);

2) получаем базовый набор векторов оптимальных преобразований Φ по каждому входу исследуемого объекта, применяя алгоритм АСЕ к расширенной матрице данных

$$\Phi_i^b(X_i) = (\Phi_i^{b,1}, \dots, \Phi_i^{b,j}, \dots, \Phi_i^{b,K})^T, \quad i = 1, \dots, n+2, \quad X_{n+1} = x, \quad X_{n+2} = Y$$

и вектор разностей

$$\Delta \Phi_i^b = (\Delta \Phi_i^{b,1}, \dots, \Delta \Phi_i^{b,k}, \dots, \Delta \Phi_i^{b,K-1})^T, \quad (5)$$

в котором $\Delta \Phi_i^{b,k} = \Phi_i^{b,k+1} - \Phi_i^{b,k}, \quad k = 1, \dots, K-1$;

3) из базовой матрицы формируем набор матриц размера $(K \times (n+2))$ для получения векторов оптимальных преобразований, используя возмущающие воздействия. Для этого к переменным $X_i, i = 1, 2, \dots, n, n+2$ добавляем малые случай-

ные числа $a_i^q = e_i^q \in N(0,1)$, $i=1,2,\dots,n,n+2$, $q=1,\dots,M$;

4) находим множество векторов оптимальных преобразований и разностей

$$\Phi_i^q(X_{a,i}) = (\Phi_{a,i}^{q,1}, \dots, \Phi_{a,i}^{q,j}, \dots, \Phi_{a,i}^{q,K})^T, \quad i=1,\dots,n+2, \quad q=1,\dots,M,$$

$$\Delta\Phi_{a,i}^q = (\Delta\Phi_{a,i}^{q,1}, \dots, \Delta\Phi_{a,i}^{q,k}, \dots, \Delta\Phi_{a,i}^{q,K-1})^T, \quad (6)$$

где $X_{a,i} = X_i + a_i^q$, $a_{n+1}^q = 0$, $\Delta\Phi_{a,i}^{q,k} = \Phi_{a,i}^{q,k+1} - \Phi_{a,i}^{q,k}$, $k=1,\dots,K-1$.

5) преобразуем полученные векторы разностей (5) и (6) к виду

$$\Delta\Phi_{m,i}^b = (\Delta\Phi_{m,i}^{b,1}, \dots, \Delta\Phi_{m,i}^{b,k}, \dots, \Delta\Phi_{m,i}^{b,K-1})^T, \quad (7)$$

$$\Delta\Phi_{m,i}^q = (\Delta\Phi_{m,i}^{q,1}, \dots, \Delta\Phi_{m,i}^{q,k}, \dots, \Delta\Phi_{m,i}^{q,K-1})^T, \quad q=1,\dots,M, \quad (8)$$

где $\Delta\Phi_{m,i}^{b,k} = \Delta\Phi_{a,i}^{b,k} / S_i^b$, $\Delta\Phi_{m,i}^{q,k} = \Delta\Phi_{a,i}^{q,k} / S_i^q$; индекс m – признак усреднения разностей $\Delta\Phi_{a,i}^{b,k}$ и $\Delta\Phi_{a,i}^{q,k}$;

$$S_i^b = ((\Delta\Phi_{a,i}^b - \overline{\Delta\Phi_{a,i}^b})^2 / (K-2))^{1/2};$$

$$S_i^q = ((\Delta\Phi_{a,i}^q - \overline{\Delta\Phi_{a,i}^q})^2 / (K-2))^{1/2};$$

$$\overline{\Delta\Phi_{a,i}^b} = \sum_{k=1}^{K-1} (\Delta\Phi_{a,i}^{b,k}) / (K-1), \quad \overline{\Delta\Phi_{a,i}^q} = \sum_{k=1}^{K-1} (\Delta\Phi_{a,i}^{q,k}) / (K-1);$$

б) находим отклонения разностей (7) базовых оптимальных преобразований от разностей (8) для каждого $q=1,\dots,M$

$$\Delta V_i^{q,k} = \Delta\Phi_{m,i}^{b,k} - \Delta\Phi_{m,i}^{q,k}, \quad i=1,\dots,n+2, \quad k=1,\dots,K-1,$$

из которых формируем последовательность векторов

$$\Delta V_i^q = (\Delta V_i^{q,1}, \dots, \Delta V_i^{q,k}, \dots, \Delta V_i^{q,K-1})^T, \quad i=1,\dots,n+2; \quad (9)$$

7) получаем количественную оценку отклонений ΔV_i^q из (9)

$$\Delta E_i^q = \sum_{k=1}^{K-1} |\Delta V_i^{q,k}|, \quad i=1,\dots,n+2, \quad q=1,\dots,M; \quad (10)$$

8) определяем показатель идентифицируемости по i -й переменной

$$H_i = \Delta E_{m,n+1} / \Delta E_{m,i}, \quad i=1,\dots,n+2, \quad (11)$$

где $\Delta E_{m,i} = \sum_{q=1}^M \Delta E_i^q / M$, а ΔE_i^q вычисляется по (10);

9) сравниваем полученные показатели идентифицируемости H_i из (11) с соответствующим пороговым значением H_p . Если $H_i > H_p$, то объект идентифицируем, в противном случае он не идентифицируем на основе представленных данных.

Заметим, что единой методики для определения порогового значения H_p не существует. При неизвестной структуре модели для его задания приходится учитывать особенности конкретных объектов или процессов и условия их функционирования. Вычислительные эксперименты показали, что точность показателя идентифицируемости H_i зависит от числа M возмущенных матриц, устанавливае-

мого опытным путем. Он рассчитывается для каждого набора данных, зависит от случайных возмущений, используемых в алгоритме АСЕ, а его окончательное значение находится усреднением полученных величин.

Тестовые примеры

Предложенная методика иллюстрируется двумя тестовыми примерами.

Пример 1. Пусть объект задается функциональной зависимостью вида

$$Y = X_1^2 + \sin(3X_2) + (X_3 + 3) + \ln(X_4). \quad (12)$$

По уравнению (12), входным переменным $X_i, i=1,2,3,4$ и ограничениям $-2.5 \leq X_1 \leq 2.5, 0 \leq X_2 \leq 1, 0 \leq X_3 \leq 1, 0 < X_4 \leq 2$ формируется выборка объема $K = 1000$, представляющая собой матрицу размера $(K \times 5)$. Она преобразуется в расширенную выборку включением дополнительного некоррелированного с выходом Y входа X_5 с параметрами $0;1$ ($X_5 \in N(0, 1)$). Анализ коэффициентов парной корреляции и корреляционных отношений, полученных по исходной выборке (табл. 1), не позволяет сделать вывод о возможной структуре модели.

Таблица 1

Переменные	X_1	X_2	X_3	X_4	X_5	Y
Корреляционные связи						
Коэффициенты парной корреляции	0.0507	-0.0042	0.2665	0.3866	0.0167	1.0000
Корреляционные отношения	0.7621	0.2703	0.3119	0.3982	0.1148	1.0000

Применяя алгоритм АСЕ к расширенной выборке, формируем базовый набор векторов оптимальных преобразований $\Phi_i^b(X_{a,i})$, представленный на рис. 1 и свидетельствующий о достаточно точно найденной структуре модели.

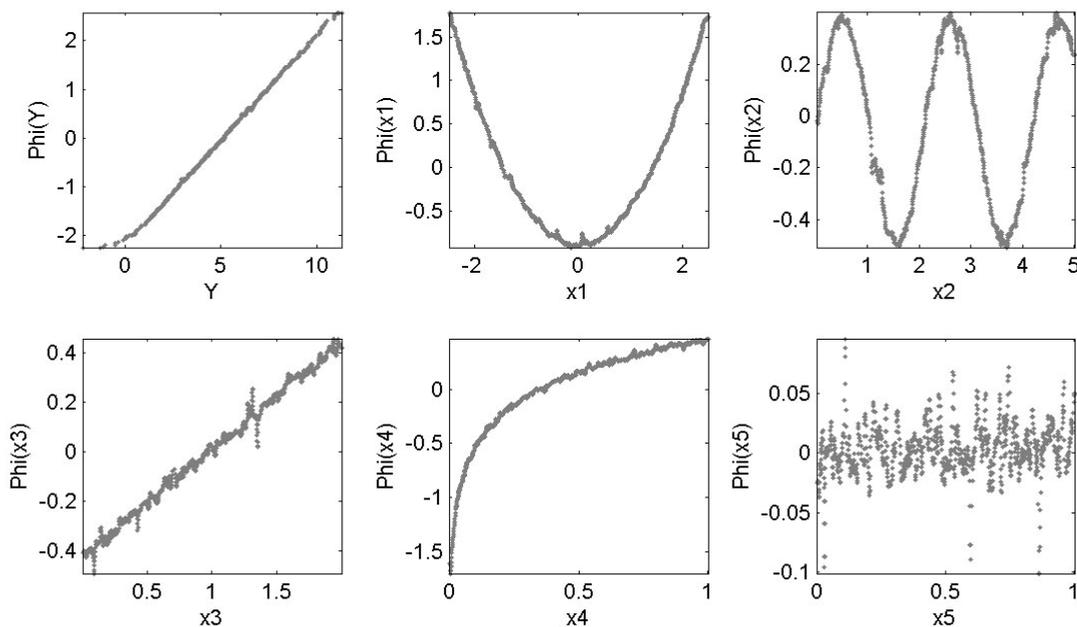


Рис. 1. Результаты применения алгоритма АСЕ к элементам базовой матрицы.

Для оценки идентифицируемости выбрано $M = 20$ получено 20 векторов оптимальных преобразований, проведено их сравнение с базовыми оценками модели (рис. 2).

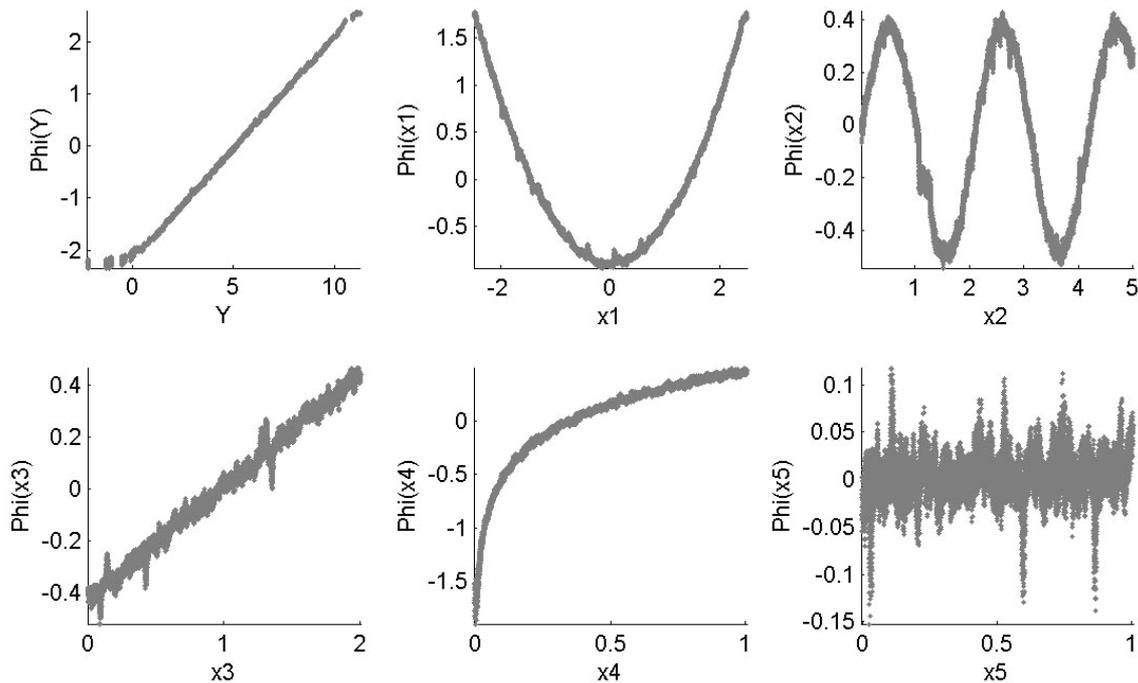


Рис. 2. Результаты применения алгоритма ACE к сформированным матрицам.

В табл. 2 сведены результаты применения предложенного подхода.

Таблица 2

Параметр	X1	X2	X3	X4	X5	Y
$\Delta E_{m,i}$	0.043	0.107	0.132	0.113	1.45	0.032
N_i	33.76	13.59	10.95	12.83	1.00	45.65

Значение $\Delta E_{m,5} = 1.45$ для входа X5 соответствует его неидентифицируемости. Остальные значения $\Delta E_{m,i}$ (средние суммы расстояний между точками базовой оценки модели от текущей оценки модели для выхода и каждого входа) полностью подтверждают существование нелинейной модели для исследуемого объекта и могут служить признаком его идентифицируемости. Величины N_i отражают вклад каждой переменной относительно неидентифицируемого дополнительного входа, при этом вход X1 дает наибольший вклад в модель, вход X3 – наименьший, и полученные результаты не противоречат описанию (12). Таким образом, объект идентифицируем, поскольку значение показателя идентифицируемости для выхода $Z_X = 45.65$, что значительно больше заданного порогового значения $N_p = 5$.

Пример 2. В предыдущем примере для оценки идентифицируемости объекта использовалась его сложная нелинейная структура. Для подтверждения правильной работы предлагаемой методики рассматривается пример неидентифицируемого объекта.

Строим хаотичную матрицу данных размера (1000×5), где Y принимает случайные значения $0 \leq Y \leq 100$, входы в модель $X_i, i=1,2,3,4$ не связаны между собой и принимают значения соответственно из диапазонов: $-2.5 \leq X_1 \leq 2.5$, $0 \leq X_2 \leq 5$, $0 \leq X_3 \leq 2$, $0 \leq X_4 \leq 1$. Преобразуем ее в расширенную матрицу, включая дополнительный вход $X_5 \in N(0;1)$, не коррелированный с Y .

Анализ коэффициентов парной корреляции и корреляционных отношений, полученных по исходной выборке (табл. 3), указывает на отсутствие каких-либо взаимосвязей между данными.

Таблица 3

Переменные	X_1	X_2	X_3	X_4	X_5	Y
Корреляционные связи						
Коэффициенты парной корреляции	0.0114	-0.0107	-0.0259	-0.0281	0.0135	1.0000
Корреляционные отношения	0.0328	0.0921	0.0832	0.1051	0.0907	1.0000

Применяя алгоритм ACE к расширенной выборке, формируем базовый набор векторов оптимальных преобразований $\Phi_i^b(X_{a,i})$, графически представленный на рис. 3.

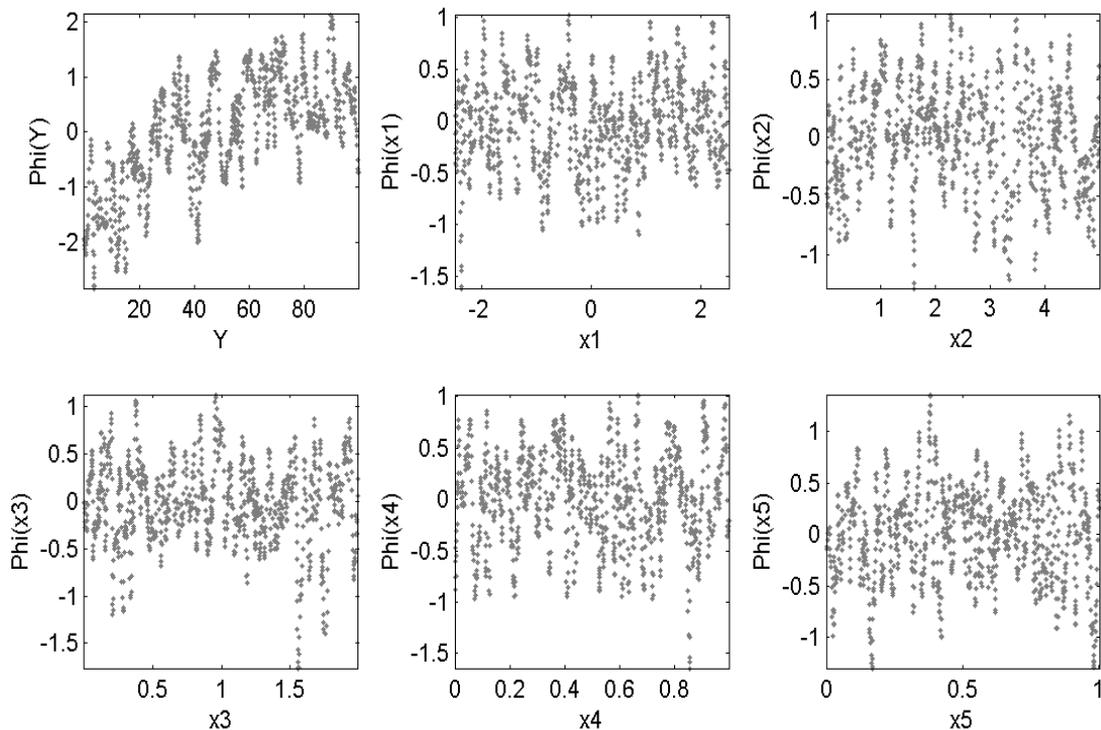


Рис. 3. Результаты применения алгоритма ACE к элементам базовой матрицы.

Как и в предыдущем примере, для оценки идентифицируемости выбрано $M = 20$, по сформированным выборкам получено 20 векторов оптимальных преобразований (рис. 4).

Анализ полученных результатов позволяет сделать вывод об отсутствии

связей между данными исходной матрицы и неидентифицируемости объекта, что подтверждается близкими к единице значениями показателей идентифицируемости H_i в табл. 4, полученными по формуле (11), при пороговом числе $H_p = 5$.

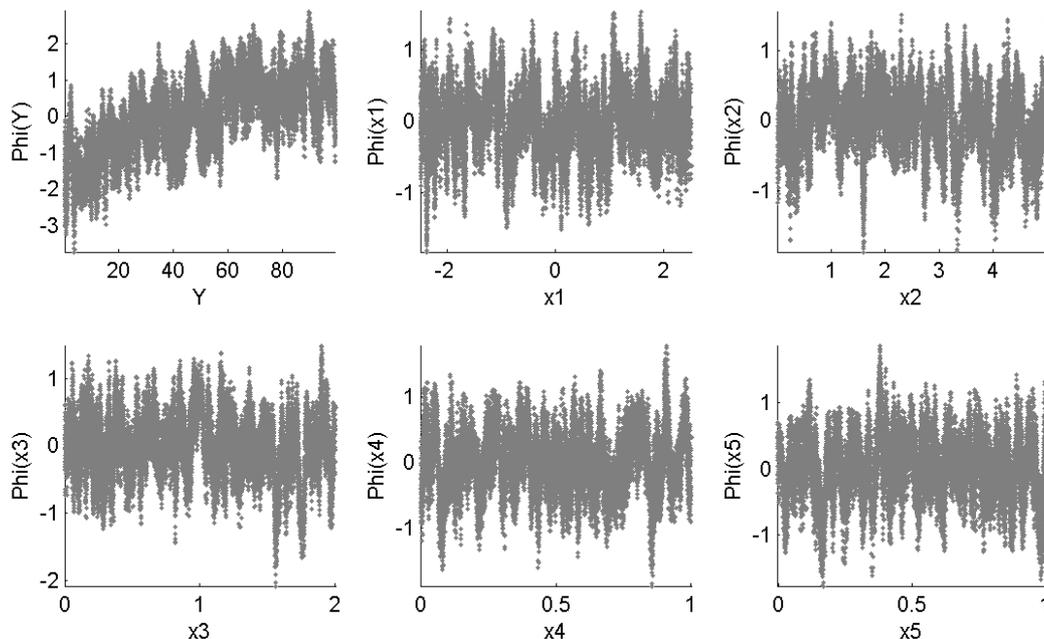


Рис. 4. Результаты применения алгоритма ACE к сформированным матрицам.

Таблица 4

Параметр	X_1	X_2	X_3	X_4	X_5	Y
$\Delta E_{m,i}$	1.29	1.18	1.27	1.43	1.34	0.98
H_i	1.04	1.14	1.06	0.94	1.00	1.38

Оценка идентифицируемости моделей виртуальных анализаторов, построенных на данных феноменологической модели ректификационной колонны

При построении модели ВА качества выходного продукта РК по доле изопентана Y (моль/кг) использовались данные с реального объекта: X_1 – температура вверху РК ($^{\circ}\text{C}$); X_2 – температура верха выходного продукта ($^{\circ}\text{C}$); X_3 – температура входа продукта ($^{\circ}\text{C}$); X_4 – поток сырья на входе в РК ($\text{м}^3/\text{ч}$); X_5 – температура выхода продукта ($^{\circ}\text{C}$).

Поскольку для применения алгоритма ACE их оказалось недостаточно, были получены дополнительные данные по феноменологической (физико-химической) модели в программной среде Aspen+ для формирования выборки объема $K = 7500$ в виде матрицы размера $(K \times 6)$. Ее элементы по каждой из переменных близки к значениям из диапазонов параметров промышленной РК.

Приведенные в табл. 5 коэффициенты парной корреляции и корреляционные отношения, рассчитанные по расширенной выборке, указывают на достаточно сильные связи всех входов, кроме дополнительного входа $X_6 \in N(0;1)$ и выхода, поэтому возможно построение ВА.

Таблица 5

Переменные	X_1	X_2	X_3	X_4	X_5	X_6	Y
Корреляционные связи							
Коэффициенты парной корреляции	-0.6930	0.6930	-0.9696	-0.9552	-0.4884	-0.0039	1.0000
Корреляционные отношения	0.6906	0.6919	0.9628	0.9610	0.5299	0.0144	1.0000

Результаты применения разработанной методики оценки идентифицируемости представлены в табл. 6 и на рис. 5, 6.

Таблица 6

Параметр	X_1	X_2	X_3	X_4	X_5	X_6	Y
$\Delta E_{m,i}$	0.0442	0.0554	0.1259	0.0705	0.2401	1.7365	0.1235
H_i	39.300	31.358	13.796	24.644	7.233	1.000	14.064

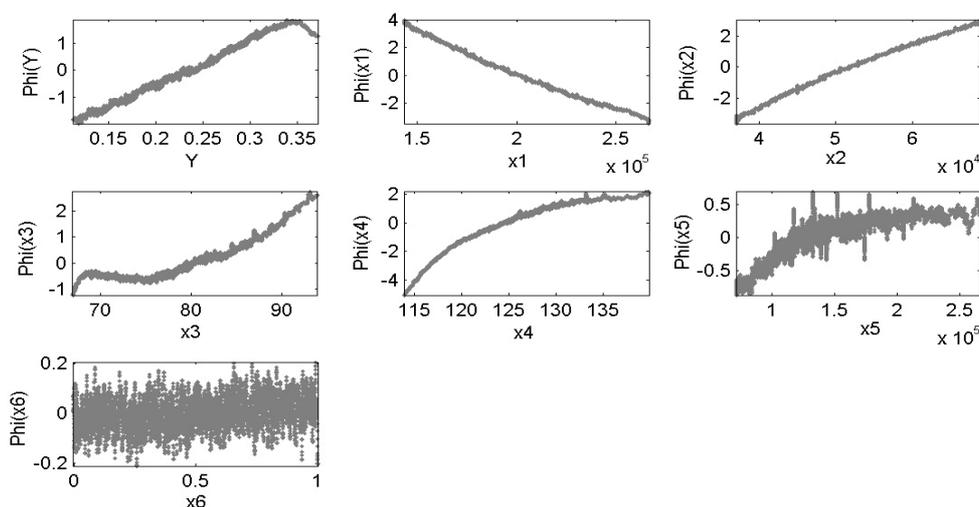


Рис. 5. Результаты применения алгоритма ACE к элементам базовой матрицы изопентана.

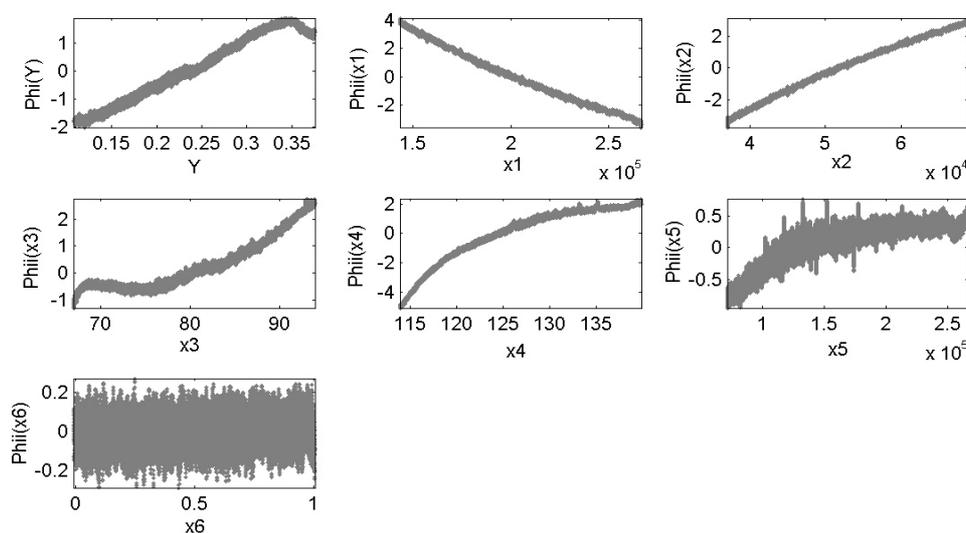


Рис. 6. Результаты применения алгоритма ACE к сформированным матрицам по доле изопентана.

Из их анализа следует, что модель идентифицируема, поскольку значение коэффициента идентифицируемости для выхода $H_Y = 14.06$, что значительно больше заданного порогового значения $H_p = 5$. Кроме того, видно, что выход Y имеет более слабые связи с X_5 , который при выборе более высокого порогового значения H_p может оказаться вообще исключенным из модели.

Оценка идентифицируемости моделей виртуальных анализаторов для промышленной ректификационной колонны

При оценивании идентифицируемости моделей ВА массообменных процессов промышленной РК в исходную выборку данных, исходя из физического смысла массообменного процесса, проходящего в колонне, были взяты входы $X_1 - X_5$, использованные в предыдущем примере, вход X_6 , количество производимого продукта (т/ч). Из нее были сформированы выборки по долям изопентана и бензолобразующих, расширенные за счет дополнительного входа $X_7 \in N(0;1)$.

Парные коэффициенты и корреляционные отношения, полученные по этим выборкам, приведены в табл. 7.

Таблица 7

Переменные	X_1	X_2	X_3	X_4	X_5	X_6	X_7	Y
Корреляционные связи								
Коэффициенты парной корреляции по доле изопентана	0.1950	0.2391	0.0137	0.5130	0.2105	-0.8844	0.0579	1.0000
Корреляционные отношения по доле изопентана	0.3997	0.6761	0.4559	0.6205	0.3140	0.8883	0.2935	1.0000
Коэффициенты парной корреляции по доле бензолобразующих	-0.5509	0.2944	-0.0705	-0.3748	-0.4585	0.8310	-0.0228	1.0000
Корреляционные по доле бензолобразующих	0.5402	0.6577	0.6037	0.5511	0.5025	0.8846	0.3275	1.0000

Из их анализа следует, что между обоими выходами и входом X_6 существует хорошая корреляционная зависимость, между обоими выходами и входами X_3, X_7 отсутствует линейная связь, а по остальным переменным корреляционная связь невелика, поэтому нет оснований использовать полученные корреляционные отношения для определения нелинейных зависимостей.

При использовании предложенной методики получены результаты, представленные по доле изопентана и бензолобразующих в табл. 8 и на рис. 7 и 8 соответственно.

Таблица 8

Параметр	X_1	X_2	X_3	X_4	X_5	X_6	X_7	Y
$\Delta E_{m,i \text{ изоп}}$	0.465	0.139	0.805	0.388	0.176	0.102	0.839	0.0792
$H_{i, \text{изоп}}$	1.805	6.035	1.043	2.162	4.773	8.229	1.000	10.594
$\Delta E_{m,i \text{ бенз}}$	0.791	0.176	0.672	0.320	0.278	0.112	1.082	0.070
$H_{i, \text{бенз}}$	1.369	6.138	1.611	3.382	3.898	9.642	1.000	15.386

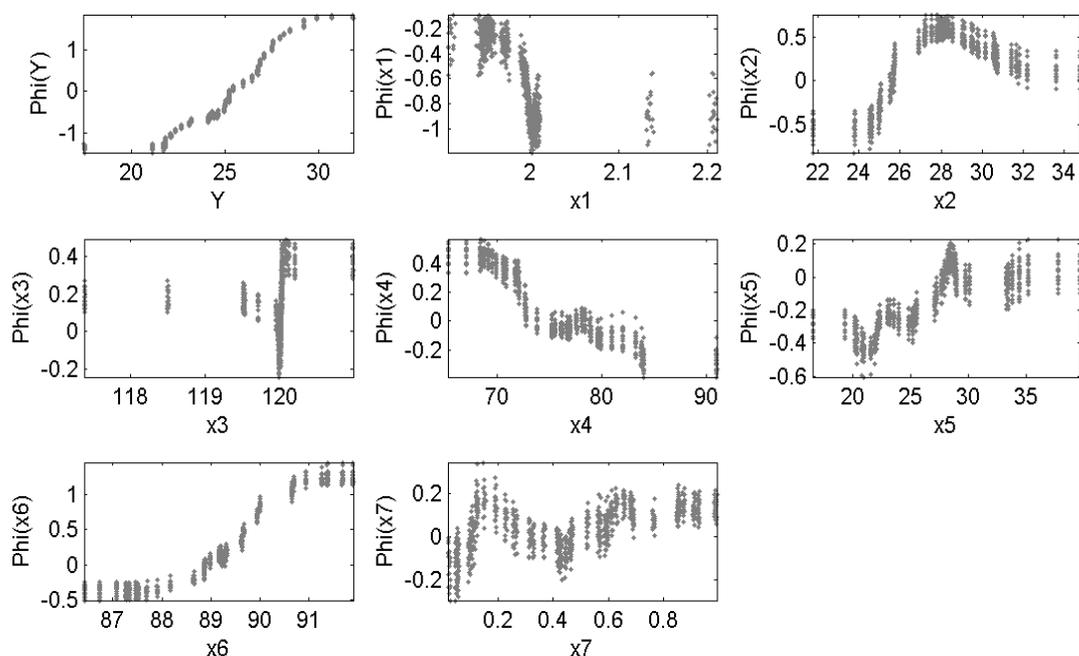


Рис. 7. Результаты применения алгоритма ACE к сформированным матрицам по доле изо-пентана.

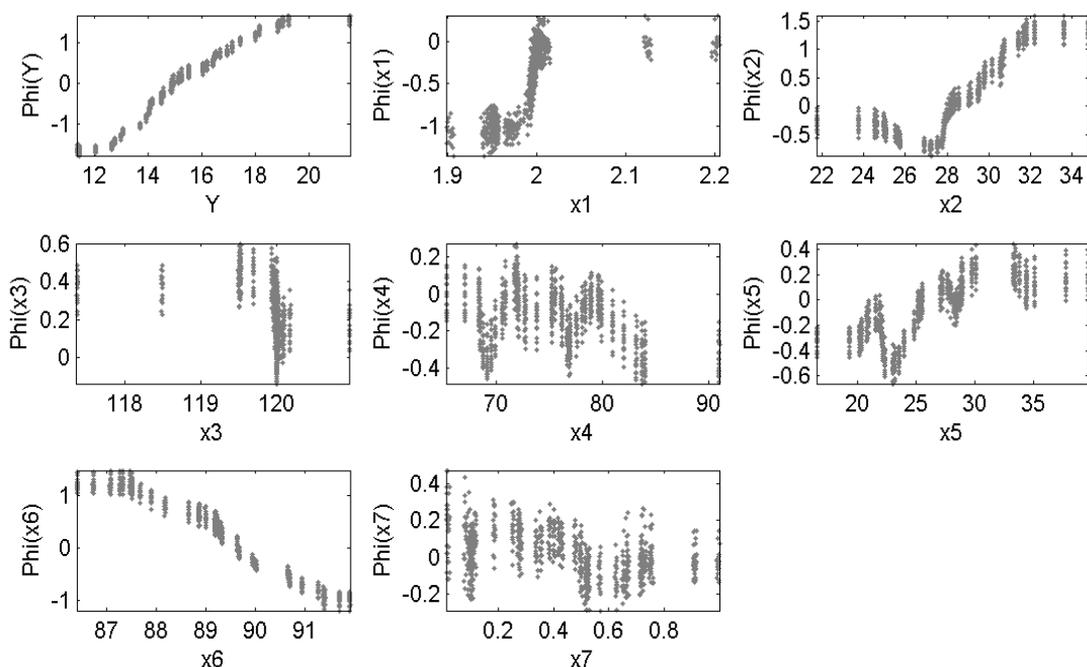


Рис. 8. Результаты применения алгоритма ACE к сформированным матрицам по доле бензообразующих.

Их анализ позволяет сделать вывод, что модели ВА идентифицируемы, поскольку значения их показателей идентифицируемости значительно превышают выбранное пороговое значение $N_p = 5$ ($N_{изоп} = 10.594$ и $N_{бенз} = 15.386$), а входные переменные X_1 , X_3 малоинформативны. При проведении вычислительных экспериментов было установлено, что, используя предлагаемую методику, можно

достичь идентифицируемости моделей ВА реальных массообменных процессов для порогового значения $N_p > 5$ и числе возмущенных матриц $M \geq 50$.

Заключение

Предложенная методика определения идентифицируемости сложных нелинейных слабо формализованных объектов на основе алгоритма ACE и дополнительного входа проверена на тестовых примерах и применена для определения идентифицируемости моделей виртуального анализатора качества выходного продукта на промышленной ректификационной колонне по долям изопентана и бензолобразующих компонентов.

Проведенные исследования показали, что разработанный подход позволяет определять идентифицируемость объектов при заданных условиях и различных выборках данных, выявлять неинформативные или слабо информативные входы, быстро, надежно, просто и наглядно оценивать их идентифицируемость на основе любых выборок достаточного объема.

ЛИТЕРАТУРА

1. Фишер Ф. Проблема идентификации в эконометрике. – М.: Статистика, 1978.
2. Горский В.Г., Кацман Е.А., Клебанова Ф.Д., Григорьев А.А. Численное исследование идентифицируемости параметров нелинейных моделей // Теоретическая и экспериментальная химия. – 1987. – №2. – С. 191-197.
3. Дуго Г.Б., Дуго Н.Б., Можаровский И.С., Торгашов А.Ю. Исследование моделей виртуальных анализаторов массообменного технологического процесса ректификации // Информатика и системы управления. – 2011. – №4(30). – С. 17-27.
4. Breiman L., Friedman J. Estimating optional transformations for multiple regression and correlation // Journal of the American Statistical Association. – 1985. – Vol. 80. – P. 580-598.
5. Hengl S., Kreuz C., Timmer J., Maiwald T. Data-based identifiability analysis of non-linear dynamical models // Bioinformatics. – 2007. – Vol. 23, No. 19. – P. 2612–2618.
6. Wang D., Murphy M. Estimating optimal transformations for multiple regression using the ACE algorithm // Journal of Data Science. – 2004. – Vol. 2. – P. 329-346.

Статья представлена к публикации членом редколлегии О.В. Абрамовым.

E-mail:

Дуго Галина Борисовна – bernatsk@iacp.dvo.ru;

Дуго Наталья Борисовна – digo@iacp.dvo.ru;

Можаровский Игорь Сергеевич – studvvsu@gmail.com;

Торгашев Андрей Юрьевич torgashov@iacp.dvo.ru.