УДК 519.246.8

ОЦЕНКА СОСТОЯНИЯ ОБЪЕКТА ПО РЕГРЕССИОННЫМ ЗАВИСИМОСТЯМ ПРИ ПРОГНОЗИРОВАНИИ ВХОДНЫХ ПАРАМЕТРОВ

Ю.Е. Кувайскова, Д.С. Бубырь, В.Н. Клячкин

Ульяновский государственный технический университет

Поступила в редакцию 28.05.2014

Предложена методика оценки выходных параметров, характеризующих состояние технического объекта, на основе построения кусочно-линейных регрессионных зависимостей от входных характеристик объекта, прогнозируемых с использованием адаптивного динамического регрессионного моделирования на основе смешанной модели авторегрессии-скользящего среднего, комплекса авторегрессионных моделей с условной гетероскедастичностью, или последовательного применения этих подходов.

Ключевые слова: *моделирование, технический объект, прогнозирование, кусочно-линейная регрессия, локальность модели*

Моделирование по поступающим результатам измерений и прогнозирование состояния технического объекта по полученным моделям с формированием сигнала предупреждения о возможной аномальной ситуации, при которой характеристики процесса выходят за допустимые пределы, является важной задачей [1-3]. Решение задачи прогнозирования состояния объекта по регрессионным моделям предполагает, что будущие значения входных параметров (регрессоров) известны. Однако на практике при исследовании технических объектов значения входных и выходных параметров определяются по результатам измерений в режиме реального времени. Таким образом, при прогнозировании состояния объекта значения входных параметров неизвестны. Для решения этой задачи предлагается алгоритм прогнозирования состояния технического объекта по кусочно-линейным регрессионным зависимостям выходных характеристик объекта с использованием спрогнозированных значений входных параметров, получаемых по моделям временных рядов.

Рассмотрим технический объект *Y*, состояние которого зависит от некоторых входных параметров, значения которых регистрируются через определенные промежутки времени и образуют временные ряды $X_1, X_2, ..., X_m$, где m – количество параметров. Алгоритм моделирования и прогнозирования состояния технического объекта состоит из следующих этапов:

– построение моделей входных параметров объекта, образующих систему временных рядов $X_1, X_2, ..., X_m$, на основе методологии адаптивного динамического регрессионного моделирования [4-6];

– прогнозирование по полученным моделям временных рядов значений входных параметров исследуемого объекта;

– построение регрессионной зависимости $f(X_1, X_2, \ldots, X_m)$, характеризующей связь между состоянием объекта *Y* и его входных параметров X_j, $j = \overline{1, m}$;

– использование полученной регрессии и спрогнозированных значений входных параметров для прогнозирования состояния технического объекта.

Согласно методологии динамического регрессионного моделирования временной ряд представляется в виде суммы составляющих:

$$y(t) = f(t) + g(t) + \psi(t) + \varepsilon(t) \ , \qquad (1)$$

где $y(t)$ – значения временного ряда, фиксируемые в моменты времени $t = 1, 2, \ldots, n$; $f(t)$ – неслучайная (долговременная) функция тренда; $g(t)$ – неслучайная периодическая функция; $\psi(t)$ – случайная с элементами регулярности функция; $\varepsilon(t)$ – нерегулярная компонента (случайная величина, ошибка).

Кувайскова Юлия Евгеньевна, кандидат технических наук, доцент кафедры прикладной математики и информатики. E-mail: u.kuvaiskova@mail.ru
Бубырь Дмитрий Сергеевич, аспирант
Клячкин Владимир Николаевич, доктор технических наук, профессор кафедры прикладной математики и информатики. E-mail: v_kl@mail.ru

При выявлении значимой регулярности процесса методами фрактального и мультифрактального анализов [7] выделяется функция тренда $f(t)$. Далее остатки исследуются на существование периодических компонент при помощи гармонического анализа и других методов. Остаточные колебания после выделения неслучайных функций сглаживаются случайной с элементами регулярности функцией $\psi(t)$. В качестве данной функции могут выступать либо авторегрессионная модель подходящего порядка, либо смешанная модель авторегрессии-скользящего среднего (АРСС), либо одна из комплекса авторегрессионных моделей с условной гетероскедастичностью (ARCH) [8], либо последовательное применение этих подходов.

После структурно-параметрической идентификации модели временного ряда проверяется соблюдение условий применения регрессионного анализа: постоянство дисперсии, независимость регрессоров, нормальность распределения ошибок, нулевое значение математического ожидания ошибок, независимость ошибок. При нарушении этих условий используются специальные методы адаптации [4]. Если основные условия соблюдаются, построенная комплексная модель временного ряда может быть использована для прогнозирования. После получения предсказанных значений входных параметров по моделям временных рядов, строятся регрессионные зависимости выходных характеристик объекта в зависимости от входных (регрессоров).

При неоднородности физических свойств объекта на области значений регрессоров, «глобальные» модели часто обладают недостаточно высокой точностью. В этом случае предлагается применять принцип «кусочности» или локальности модели, то есть вариации её параметров по области значений регрессоров. В качестве прогнозирующей зависимости можно использовать кусочно-линейную регрессию с разрывом по отклику и с добавлением элемента авторегрессии для учёта значения отклика в предыдущие дни. При этом свободный коэффициент в модель предполагается не включать. В этом случае регрессионная модель представляется в виде:

$$\begin{aligned} Y(t) &= (\beta_{11}X_1 + \beta_{21}X_2 + \ldots + \beta_{m1}X_m + \beta_{m+1,1}Y(t-1))(Y(t) \le c) + \\ &(\beta_{12}X_1 + \beta_{22}X_2 + \ldots + \beta_{m2}X_m + \beta_{m+1,2}Y(t-1))(Y(t) > c) \end{aligned}, \qquad (2)$$

где m – количество регрессоров модели; k – порядок авторегрессии; c – точка разрыва; $(Y(t) \le c), (Y(t) > c)$ – логические выражения, принимающие значения: 1 – если истинно, 0 – если ложно; $Y(t\text{-}1)\ldots\ Y(t\text{-}k)$ – значение отклика в предыдущие моменты времени. Фактически данная модель состоит из двух регрессий, каждая из которых применяется для определённого уровня отклика.

Одним из ключевых моментов при использовании кусочных регрессий является способ отнесения текущих данных к нужному «куску». Предлагается следующий алгоритм этой процедуры. Вначале вычисляется прогноз по обеим регрессиям. Затем полученные результаты сравниваются со значением в точке разрыва. Если прогноз по первому куску меньше либо равен значению в точке разрыва, а прогноз по второму куску – больше, то за итоговый результат принимается среднее арифметическое данных чисел. Если хотя бы для одного куска наблюдается «нарушение», то итоговым значением считается прогноз по тому куску, для которого выполняется неравенство. Если же условия нарушаются для обоих случаев, то также вычисляется среднее арифметическое.

Например, имеется кусочно-линейная регрессия:

$$Y = Y_1 \cdot (Y \le c) + Y_2 \cdot (Y > c), \qquad (3)$$

где $Y_1 = (b_{01} + b_{11} \cdot X_1 + \ldots + b_{m1} \cdot X_m)$; $Y_2 = (b_{02} + b_{12} \cdot X_1 + \ldots + b_{m2} \cdot X_m)$; c – точка разрыва.

Пусть на основе регрессий Y_1 и Y_2 получены прогнозы $\hat{Y}_1$ и $\hat{Y}_2$, причём $\hat{Y}_1 \le c$ и $\hat{Y}_2 \le c$. В данном случае наблюдается нарушение для второй регрессии, поэтому за итоговый прогноз берётся $\hat{Y}_1$. По полученной модели кусочно-линейной регрессии с добавлением элемента авторегрессии и спрогнозированным значениям входных параметров прогнозируется будущее состояние технического объекта с целью формирования сигнала о возможной аварийной ситуации. Эффективность разработанной методики прогнозирования состояния технического объекта иллюстрируется на примере моделирования выходной характеристики объекта Y и множества входных параметров, образующих временные ряды $X_1, X_2, \ldots, X_7$ по результатам 35 наблюдений. Для системы временных рядов (входных параметров объекта) согласно методологии адаптивного динамического регрессионного моделирования построены следующие модели:

$$X_1(t) = 0{,}47715 \cdot X_1(t-1) + 0{,}60453 \cdot X_1(t-2),$$

$$X_2(t) = 0{,}89898 \cdot X_2(t-1) + 0{,}09956 \cdot X_2(t-2),$$
$$X_3(t) = 0{,}79444 \cdot X_3(t-1) + 0{,}17374 \cdot X_3(t-2),$$
$$X_4(t) = 0{,}57316 \cdot X_4(t-1) + 0{,}42619 \cdot X_4(t-2),$$
$$X_5(t) = 0{,}87087 \cdot X_5(t-1) + 0{,}12289 \cdot X_5(t-2),$$
$$X_6(t) = 0{,}57535 \cdot X_6(t-1) + 0{,}41826 \cdot X_6(t-2),$$
$$X_7(t) = 0{,}90417 \cdot X_7(t-1) + 0{,}09523 \cdot X_7(t-2).$$

По полученным моделям временных рядов выполнен прогноз на четыре наблюдения. Для выходного параметра объекта Y построена кусочно-линейная регрессия с авторегрессией 1-го порядка (КЛР-АР(1)):

$$\begin{aligned} Y(t) = {} & (-0{,}002228 \cdot X_1 - 0{,}034335 \cdot X_2 + 0{,}028101 \cdot X_3 - 0{,}295028 \cdot X_4 + 0{,}505128 \cdot X_5 + \\ & + 0{,}031676 \cdot X_6 + 0{,}433677 \cdot X_7 + 0{,}787707 \cdot Y(t-1)) \cdot (Y(t) \le 1{,}0109) + (-0{,}028096 \cdot X_1 - \\ & - 0{,}049394 \cdot X_2 + 0{,}046701 \cdot X_3 - 0{,}208433 \cdot X_4 + 1{,}182003 \cdot X_5 + 0{,}010884 \cdot X_6 + \\ & + 0{,}520216 \cdot X_7 - 0{,}110927 \cdot Y(t-1)) \cdot (Y(t) > 1{,}0109) \end{aligned} \quad (4)$$

С использованием полученной регрессии и спрогнозированных значений входных параметров $X_1, X_2, ..., X_7$ выполнен прогноз выходной характеристики объекта Y на четыре наблюдения. На рис. 1 показан смоделированный временной ряд, а на рис. 2 – соответствующий прогноз на 4 измерения. Сплошной линией показаны опытные данные, штриховой – расчетные.

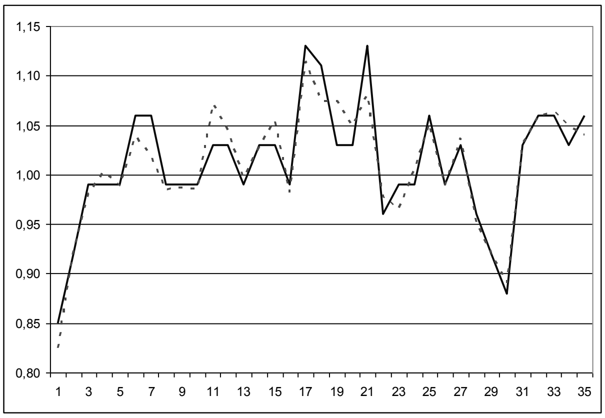

Рис. 1. Моделирование на основе КЛР-АР(1)

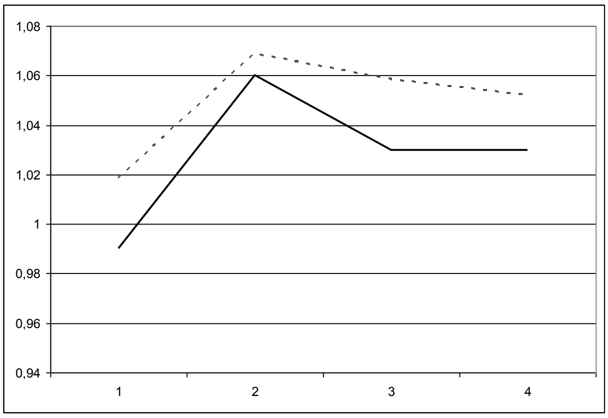

Рис. 2. Прогнозирование состояния объекта на основе КЛР-АР(1)

Для сравнительного анализа эффективности предлагаемой методики прогнозирования выходных характеристик объекта по той же выборке были построены модели множественной линейной регрессии (ЛР) и множественной линейной регрессии с добавлением элемента авторегрессии 1-го порядка (ЛР-АР(1)):

$$\begin{aligned} Y(t) = {} & -13{,}604 + 0{,}0112 \cdot X_1 - 0{,}4476 \cdot X_2 + 0{,}0422 \cdot X_3 - 0{,}3325 \cdot X_4 + 1{,}1932 \cdot X_5 \\ & + 0{,}0216 \cdot X_6 + 4{,}5121 \cdot X_7 \end{aligned}, \quad (5)$$

$$\begin{aligned} Y(t) = {} & -19{,}6682 - 0{,}0011 \cdot X_1 - 0{,}6002 \cdot X_2 + 0{,}0367 \cdot X_3 - 0{,}2679 \cdot X_4 + X_5 \\ & + 0{,}0229 \cdot X_6 + 6{,}0737 \cdot X_7 + 0{,}3024 \cdot Y(t-1) \end{aligned}. \quad (6)$$

В табл. 1 представлены результаты сравнения кусочно-линейной модели с ЛР и ЛР-АР(1) по точности прогнозирования. В первом столбце таблицы представлены исходные значения отклика Y; в следующих столбцах для каждой модели представлены значения спрогнозированных значений отклика Y (прогноз), а также ошибка прогноза в процентах, вычисляемая по формуле (7).

$$e = \frac{|Y - \hat{Y}|}{Y} \cdot 100\%, \quad (7)$$

где Y – исходное значение отклика; $\hat{Y}$ – прогнозное значение отклика.

По данным таблицы следует, что при использовании модели кусочно-линейной регрессии с авторегрессией 1-го порядка точность прогнозирования существенно повышается по сравнению с моделями обычной множественной линейной регрессии и множественной линейной регрессии с добавлением элемента авторегрессии 1-го порядка.

Таблица 1. Сравнение моделей по точности прогнозирования

Исходные значения отклика Y	**Модель регрессии**					
	КЛР-АР(1)		ЛР		ЛР-АР(1)	
	прогноз	ошибка, %	прогноз	ошибка, %	прогноз	ошибка, %
0,99	1,02	2,87	0,97	1,75	1,27	28,77
1,06	1,07	0,84	0,60	43,82	0,81	24,05
1,03	1,06	2,76	0,58	43,48	0,64	37,44
1,03	1,05	2,14	0,61	40,75	0,63	39,24

Выводы: использование кусочно-линейных моделей регрессии позволяет повысить точность прогнозирования состояния технического объекта. Безусловным преимуществом кусочно-линейных регрессий относительно обычных линейных регрессий является учёт неоднородности свойств объекта за счёт своей «кусочности» и вариации уровней отклика. Это позволяет более точно моделировать и прогнозировать процесс, в особенности, когда поступающие данные отличаются значительной нестабильностью значений.

Работа выполнена в рамках задания Минобрнауки России №2014/232.

СПИСОК ЛИТЕРАТУРЫ:

1. *Клячкин, В.Н.* Информационно-математическая система раннего предупреждения об аварийной ситуации / *В.Н. Клячкин, Ю.Е. Кувайскова, А.А. Алешина, Ю.А. Кравцов* // Известия Самарского научного центра РАН. 2013. №4(4). С. 919-923.
2. *Клячкин, В.Н.* Моделирование вибраций гидроагрегата на основе адаптивных динамических регрессий / *В.Н. Клячкин, Ю.Е. Кувайскова, А.А. Алешина* // Автоматизация и современные технологии. 2014. №1. С. 30-34.
3. *Клячкин, В.Н.* Диагностика состояния объекта по наличию неслучайных структур на контрольных картах / *В.Н. Клячкин, Ю.А. Кравцов* // Приборы и системы. Управление, контроль, диагностика. 2013. №5. С.44-50.
4. *Валеев С.Г.* Регрессионное моделирование при обработке наблюдений. М. : Наука, 1991. 272 с.
5. *Валеев, С.Г.* Особенности построения регрессионных моделей при многомерном контроле технологического процесса / *С.Г. Валеев, В.Н. Клячкин* // Радиоэлектроника. Информатика. Управление. 2002. №1. С. 48-52.
6. *Валеев, С.Г.* Программное обеспечение обработки временных рядов техногенных характеристик / *С.Г. Валеев, Ю.Е. Кувайскова* // Обозрение прикладной и промышленной математики. 2009. Т. 16, выпуск 6. С. 1037-1038.
7. *Валеев, С.Г.* Применение мультифрактального анализа при описании временных рядов в технике и экономике / *С.Г. Валеев, Ю.Е. Кувайскова, С.А. Губайдуллина* // Вестник Ульяновского государственного технического университета. 2008. №2. С. 23-27.
8. *Валеев, С.Г.* Использование ARCH-структур и фильтра Калмана для моделирования динамики технико-экономических показателей / *С.Г. Валеев, Ю.Е. Кувайскова* // Вестник УлГТУ. 2007. №2. С. 29-33.
9. *Кувайскова, Ю.Е.* Прогнозирование состояния технического объекта на основе мониторинга его параметров / *Ю.Е. Кувайскова, В.Н. Клячкин, Д.С. Бубырь* // XII Всероссийское совещание по проблемам управления. Институт проблем управления им. Трапезникова РАН [Электронный ресурс] URL: http://vspu2014.ipu.ru/node/2940 (дата обращения: 16.05.2014).
10. *Крашенинников, В.Р.* Кусочно-квадратичное моделирование регрессионных зависимостей при оценке качества / *В.Р. Крашенинников, Д.С. Бубырь* // Мат-лы 3-й науч.-практ. internet-конференции «Междисциплинарные исследования в области математического моделирования и информатики» 20-21 февраля 2014 г. / отв. ред. *Ю.С. Нагорнов*. – Ульяновск: SIMJET, 2014. С. 233-236.
11. *Васильев, К.К.* Статистический анализ многомерных изображений / *К.К. Васильев, В.Р. Крашенинников*. – Ульяновск: УлГТУ, 2007. 170 с.

FORECASTING OF THE TECHNICAL OBJECT STATE BASED ON THE PIECEWISE LINEAR REGRESSION MODELS

© 2014 Yu.E. Kuvayskova, D.S. Bubyr, V.N. Klyachkin

Ulyanovsk State Technical University

In this article the method of forecasting output parameters characterizing the state of the technical object, based on the construction of piecewise linear regression dependencies on input characteristics of the object.

Key words: *modeling, technical object, forecasting, piecewise linear regression, locality of the model*

Kuvayskova, Candidate of Technical Sciences, Associate Professor at the Department "Applied Mathematics and Computing Science". E-mail: u.kuvaiskova@mail.ru; Dmitriy Bubyr, Post-Graduate Student; Vladimir Klyachkin, Doctor of Technical Sciences, Professor at the Department "Applied Mathematics and Computing Science". E-mail: v_kl@mail.ru.