Информация об авторах

Краковский Юрий Мечеславович - д. т. н., профессор кафедры «Информационные системы и защита информации», Иркутский государственный университет путей сообщения, г. Иркутск, e-mail: kum@stranzit.ru

Лузгин Александр Николаевич - к. т. н., преподаватель кафедры «Информационные технологии», Иркутский государственный университет, г. Иркутск, e-mail: alexln@mail.ru

Authors

Yuri Mecheslavovich Krakovsky – Doctor of Engineering Science, Prof., the Subdepartment of Information Systems and Information Protection, Irkutsk State Transport University, Irkutsk, e-mail: kum@stranzit.ru

Alexander Nikolaevich Luzgin – Ph.D. in Engineering Science, Member of the Subdepartment of Information Technologies, Irkutsk State University, Irkutsk, e-mail: alexln@mail.ru

Для цитирования

Краковский Ю. М. Интервальное прогнозирование динамических показателей на основе логистических регрессионных моделей / Ю. М. Краковский, А. Н. Лузгин // Современные технологии. Системный анализ. Моделирование. - 2017. - Т. 56, № 4. - С. 122–131. - DOI: 10.26731/1813-9108.2017.4(56).122-131.

For citation

Krakovsky Y. M., Luzgin A. N. Interval'noe prognozirovanie dinamicheskikh pokazatelei na osnove logisticheskikh regressionnykh modelei [The interval forecasting of dynamic indicators based on logistic regression models]. *Sovremennye tekhnologii. Sistemnyi analiz. Modelirovanie* [*Modern Technologies. System Analysis. Modeling*], 2017. Vol. 56, No. 4, pp. 122–131. DOI: 10.26731/1813-9108.2017.4(56).122-131.

***УДК** 519.237.5* ***DOI:** 10.26731/1813-9108.2017.4(56).131-138*

М. П. Базилевский

Иркутский государственный университет путей сообщения, г. Иркутск, Российская Федерация
***Дата поступления:** 29 сентября 2017 г.*

РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ ОЦЕНИВАНИЯ ПАРАМЕТРОВ АДДИТИВНОЙ СТЕПЕННОЙ РЕГРЕССИИ

Аннотация. *Регрессионный анализ является признанным инструментом построения математических моделей статистического типа. Методы регрессионного анализа находят применение в различных областях: в экономике, технике, образовании, медицине и др. Основными этапами построения регрессионной модели являются: идентификация переменных, сбор статистических данных, спецификация модели, т. е. выбор математической формы связи между переменными, идентификация параметров модели, верификация модели, т. е. определение степени соответствия построенной модели реальному объекту исследования, и интерпретация результатов, заключающаяся в прогнозировании, принятии управленческих решений и т. д.*

Статья посвящена проблеме выбора структурной спецификации регрессионной модели. Предложены нелинейные по параметрам аддитивные степенные регрессии, представляющие более гибкий инструмент моделирования, чем аналогичные степенные модели с мультипликативными независимыми переменными. Для оценивания неизвестных параметров предложенных аддитивных степенных регрессий были разработаны 3 специальных алгоритма, основу которых составляет нелинейный метод наименьших квадратов. С использованием эконометрического пакета Gretl было проведено исследование разработанных алгоритмов. При этом для оценки неизвестных параметров нелинейных моделей в Gretl был использован алгоритм Левенберга - Марквардта. Наилучшие результаты показал алгоритм с предварительным выбором начального приближения. Показано, что несколько первых шагов этого алгоритма представляют собой однокритериальный «конкурс» степенных регрессионных моделей. Проведен численный эксперимент, доказывающий рациональность использования алгоритма оценивания аддитивных степенных регрессий с предварительным выбором начального приближения при организации «конкурса» регрессионных моделей.

Ключевые слова: *регрессионная модель, аддитивная степенная регрессия, нелинейный метод наименьших квадратов, алгоритм Левенберга - Марквардта, «конкурс» моделей.*

M. P. Bazilevsky

Irkutsk State Transport University, Irkutsk, the Russian Federation
***Received**: September 29, 2017*

THE DEVELOPMENT AND RESEARCH OF ALGORITHMS OF ESTIMATION OF PARAMETERS OF THE ADDITIVE POWER REGRESSION

Abstract. *Regression analysis is a recognized tool for constructing mathematical models of statistical type. Methods of the regression analysis are used in various fields: in economics, technology, education, medical field, etc. The main stages in constructing the regression model are: identifying variables, collecting statistical data, specifying the model, i.e. choosing the mathematical form of the relationship between the variables, identification of model parameters, model verification. In other words, determining the degree of* *conformity of the constructed model to the real object of study, and interpretation of the results, consisting in forecasting, making managerial decisions, etc.*

The article is devoted to the problem of choosing the structural specification of the regression model. Additive power regressions that are non-linear in the parameters are presented, representing a more flexible modeling tool than similar power models with multiplicative independent variables. To estimate the unknown parameters of the proposed additive power regressions, 3 special algorithms were developed, based on the nonlinear least squares method. Using the Gretl econometric package, a study of the developed algorithms was carried out. In this case, the Levenberg-Marquardt algorithm was used to estimate the unknown parameters of nonlinear models in Gretl. The best results were shown by an algorithm with a preliminary choice of the initial approximation. It is shown that the first few steps of this algorithm represent a one-criteria "contest" of power regression models. A numerical experiment is performed proving the rationality of using the algorithm of estimating additive power regressions with a preliminary choice of the initial approximation in organizing the "competition" of the regression models.

Keywords: *regression model, additive power regression, nonlinear least squares method, Levenberg-Marquardt algorithm, «competition» of models.*

Введение

Одной из главных проблем регрессионного моделирования является проблема спецификации модели, связанная с определением списка экзогенных и эндогенных переменных и выбором математической формы связи между ними [1,2]. Число таких спецификаций, используемых на практике, так велико, что не представляется возможным даже перечислить все из них. Описание основных форм связи между переменными в регрессионных моделях можно найти, например, в работах [3–5].

В данной статье исследуются вопросы оценивания нелинейной аддитивной степенной регрессии.

Аддитивная степенная регрессия

Прежде чем ввести понятие аддитивной степенной регрессии, отметим, что на практике исследователи зачастую пользуются линейной моделью регрессии вида:

$$y_i = a_0 + \sum_{j=1}^{m} a_j x_{ij} + \varepsilon_i, \quad i = \overline{1,n}, \qquad (1)$$

где n – объем выборки, y – зависимая (объясняемая) переменная $x_1, x_2, ..., x_m$ – независимые (объясняющие) переменные $a_0, a_1, ..., a_m$ – неизвестные параметры. Высокая распространенность модели (1) обусловлена легкостью оценивания неизвестных параметров, которые могут быть найдены с помощью обычного метода наименьших квадратов (МНК), а также возможностью их интерпретации.

Но реальные социально-экономические процессы редко носят линейный характер, поэтому эксперты в области регрессионного моделирования прибегают к построению нелинейных моделей. К числу таких нелинейных регрессий относится степенная функция, называемая также функцией Кобба - Дугласа [3].

Степенная модель регрессии с мультипликативной ошибкой имеет вид:

$$y_i = a_0 \prod_{j=1}^{m} x_{ij}^{a_j} \varepsilon_i, \quad i = \overline{1,n}, \qquad (2)$$

а с аддитивной ошибкой - вид:

$$y_i = a_0 \prod_{j=1}^{m} x_{ij}^{a_j} + \varepsilon_i, \quad i = \overline{1,n}. \qquad (3)$$

В регрессиях (2) и (3) величины a_j, $j = \overline{1,m}$, называются коэффициентами эластичности, и они показывают, на сколько процентов меняется значение переменной y с изменением x_j на 1 % (при неизменных остальных независимых переменных). Обе модели являются нелинейными по параметрам. Однако модель (2) можно линеаризовать с помощью логарифмирования, т. е. свести к линейному по параметрам виду [3], а затем оценить её по МНК. А вот модель (3) линеаризовать нельзя, поэтому для оценки её параметров требуется привлечение методов нелинейного оценивания (Левенберга - Марквардта, Гаусса - Ньютона и др.). Поэтому на практике исследователи отдают предпочтение регрессии (2). Понятно, что оценки параметров моделей (2) и (3) могут существенно различаться.

Степенные регрессии (2) и (3) являются мультипликативными по независимым переменным $x_1, x_2, ..., x_m$ (МСР). Введем аддитивные по независимым переменным степенные регрессии (АСР) с мультипликативной и аддитивной ошибкой:

$$y_i = a_0 + \sum_{j=1}^{m} a_j x_{ij}^{b_j} \varepsilon_i, \qquad i = \overline{1,n}, \qquad (4)$$

$$y_i = a_0 + \sum_{j=1}^{m} a_j x_{ij}^{b_j} + \varepsilon_i, \quad i = \overline{1,n}. \qquad (5)$$

АСР (4) и (5) также являются нелинейными по параметрам. Их достоинством является то, что они содержит $2m+1$ неизвестных параметров, в отличие от моделей (2) и (3), содержащих $m+1$ параметр. Следовательно, АСР (4) и (5) представляют собой более гибкий инструмент моделирования, чем регрессии (1)-(3). С другой стороны, увеличение числа параметров существенно усложняет процедуру их оценивания.

Дальнейшее изложение посвящено исследованию алгоритмов оценивания АСР с аддитивной ошибкой (5), которая при $b_1 = b_2 = \ldots = b_m = 1$ является частным случаем линейной регрессии (1).

Алгоритмы оценивания АСР

Для исследования алгоритмов оценивания нелинейных по параметрам АСР (5) был использован бесплатный пакет эконометрического моделирования Gretl [6], в котором реализованы все основные эконометрические процедуры и функции.

Пакет Gretl позволяет оценивать нелинейные регрессии по МНК с использованием алгоритма Левенберга - Марквардта [7]. Для этого пользователю необходимо ввести спецификацию регрессии и задать начальные значения параметров. При этом рекомендуется указывать ещё аналитические выражения для первых производных функции регрессии по каждому неизвестному параметру, что призвано усилить качественные свойства алгоритма оценивания.

Оценивание нелинейной регрессии в пакете Gretl осуществляется итерационно. Процесс прекращается, когда выполняется критерий сходимости или когда достигается максимальное количество итераций. Пусть k – количество оцениваемых параметров в модели. Если заданы аналитические выражения для первых производных, то максимальное число итераций будет $100(k+1)$, а если не заданы, то $200(k+1)$.

Пусть ε – заданное малое число. Считается, что итерация в пакете Gretl сходится, если выполняется хотя бы 1 из следующих условий:

– и фактическое, и предсказанное относительное уменьшение суммы квадратов ошибок не превосходит ε;

– относительная ошибка между двумя последовательными итерациями не превосходит ε.

По умолчанию в пакете Gretl значение $\varepsilon = 1{,}82 \cdot 10^{-12}$. Но это значение можно изменять вручную.

Для оценивания параметров АСР (5) были предложены и исследованы следующие алгоритмы.

1. Простейший алгоритм, в котором начальные приближения параметров задаются случайно, после чего находятся оценки методом Левенберга - Марквардта.

2. Алгоритм метода простой итерации, представленный на рис. 1.

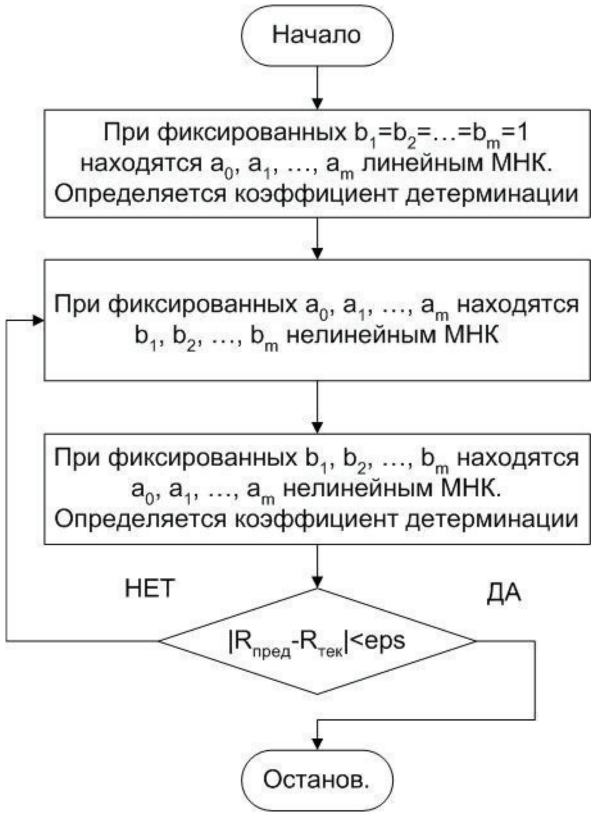

Рис. 1. Алгоритм метода простой итерации

В алгоритме на рис. 1 параметры $R_{пред}$ и $R_{тек}$ означают предыдущее и текущее значения критерия детерминации, а eps – заданная точность сходимости.

3. Алгоритм с предварительным выбором начального приближения, представленный на рис. 2. Если число независимых переменных в АСР (5) равно m, а число разбиений j-го отрезка $\left[b_j^-, b_j^+\right]$, как показано на рис. 2, равно p_j, то для получения начального приближения требуется оценить $\prod_{j=1}^{m}\left(p_j + 1\right)$ линейных моделей и выбрать из них регрессию с наименьшей величиной суммы квадратов ошибок.

Рис. 2. Алгоритм с предварительным выбором начального приближения

Многочисленные эксперименты с различными статистическими данными в Gretl показали, что первый из этих трех алгоритмов, т.е. простейший алгоритм, практически никогда не сходится. Причиной этого, вероятнее всего, является высокая степень нелинейности и чрезвычайно большое количество оцениваемых параметров АСР (5), что делает практически невозможным случайное угадывание хорошего начального приближения.

Второй алгоритм метода простой итерации в ходе экспериментов продемонстрировал чрезвычайно низкую скорость сходимости, а иногда и вовсе не сходился по истечении десятков тысяч итераций. Но всё же он гораздо чаще оказывался сходящимся, чем первый. По крайней мере, на любой итерации есть возможность прервать этот алгоритм и получить хоть какое-то неоптимальное решение, которое можно использовать в качестве начального приближения для оценивания регрессии по методу Левенберга - Марквардта.

Наилучшие результаты показал третий алгоритм с предварительным выбором начального приближения. Только в редких случаях он не оказывался сходящимся. Но всё же главной его проблемой является выбор ограничений b_j^- и b_j^+ на оценки параметров b_j и числа разбиений отрезков $\left[b_j^-, b_j^+\right]$. В ходе экспериментов ограничения задавались $b_j^- = -10$, $b_j^+ = 10$, а число разбиений $p_j = 20 - 1 = 19$ (исключена точка $b_j^{10} = 0$).

Стоит отметить, что МНК-оценки для нелинейной АСР (5) теоретически вообще могут не существовать, либо функция суммы квадратов ошибок может иметь несколько точек локальных минимумов, что затрудняет поиск точки глобального минимума [7]. Поэтому алгоритм оценивания АСР (5) с предварительным выбором начального приближения не всегда гарантирует нахождение глобального минимума, но даже в самом худшем случае за счет встроенных в него переборных процедур дает весьма приемлемое решение задачи.

Связь алгоритма с «конкурсом» моделей

Для решения задачи выбора структурной спецификации регрессии целесообразно пользоваться технологией организации «конкурса» моделей [8–10]. Её суть заключается в построении множества альтернативных вариантов регрессий и последующем многокритериальном выборе наиболее приемлемого уравнения.

«Конкурс» моделей можно представить в виде последовательности этапов:

1) формирование альтернативных уравнений;
2) оценивание параметров регрессий;
3) выбор наилучшего варианта.

В качестве альтернативных вариантов уравнений в «конкурсе» моделей могут принимать участие аддитивные зависимости:

$$y_i = a_0 + \sum_{j=1}^{m} a_j f_{hj}(x_{ij}) + \varepsilon_i, \quad i = \overline{1,n}, \qquad (6)$$

где f_{hj} – преобразование j-й переменной, выбираемое из набора $F(x) = \{f_1(x), f_2(x), ..., f_l(x)\}$. В качестве преобразований могут использоваться любые элементарные функции, например показательные 2^x, $e^{0{,}04x}$, степенные $\frac{1}{x}$, x^2, $\sqrt{x}$, тригонометрические $\sin(x)$, $\cos(x)$ и др.

Будем использовать в качестве преобразований переменных в регрессии (6) только степенные функции $f_{hj}\left(x_{ij}\right) = x_{ij}^{b_j}$. Тогда аддитивная зависимость (6) трансформируется в АСР (5) с аддитивной ошибкой. Из этого следует, что в алгоритме оценивания АСР (5) с предварительным выбором начального приближения первые 3 блока на рис. 2 представляют собой не что иное, как «конкурс» моделей с выбором наилучшего уравнения только по одному критерию – величине суммы квадратов ошибок. Но алгоритм на рис. 2 содержит еще один дополнительный блок – нахождение оценок с помощью нелинейного МНК, что не всегда, но часто приводит к улучшению качественных свойств регрессионных моделей. Таким образом, интеграция этого дополнительного блока в технологию «конкурса» степенных моделей гарантирует, что результаты моделирования не ухудшаются.

Численный эксперимент

Для ответа на вопросы об эффективности предложенного алгоритма оценивания АСР, степени его влияния на технологию «конкурса» моделей был проведен численный эксперимент.

Для проведения эксперимента были использованы статистические данные о работе выпарного аппарата на большом промышленном предприятии из монографии [11]. Выборка содержит 25 наблюдений. Зависимой переменной является переменная y – количество используемого пара (фунтов в месяц). Независимыми переменными выступают:

x_1 – количество жирной кислоты (фунтов в месяц);

x_2 – количество глицерина (фунтов в месяц);

x_3 – средняя скорость ветра (миль в час);

x_5 – число рабочих дней в месяце;

z – число дней с температурой не ниже 32 градусов по Фаренгейту;

x_7 – средняя температура воздуха по Фаренгейту.

В монографии [11] по этим данным была построена линейная модель:

$$\hat{y} = \underset{(8,276)}{9{,}127} + \underset{(4,431)}{0{,}203}\,x_5 - \underset{(-9,05)}{0{,}0724}\,x_7 . \qquad (7)$$

В регрессии (7) в скобках указаны t-статистики. Критерий детерминации этой модели $R^2 = 0{,}849$, Фишера $F = 61{,}904$, Дарбина - Уотсона $DW = 2{,}195$, сумма квадратов ошибок $RSS = 9{,}629$.

В работе [12] с использованием технологии «конкурса» моделей была найдена более адекватная, чем регрессия (7), зависимость:

$$\hat{y} = \underset{(19,5)}{23{,}106} + \underset{(6,396)}{0{,}00435}\cdot 2^{0,4x_5} - \underset{(-13,12)}{3{,}876}\ln x_7 . \qquad (8)$$

Критерии адекватности модели (8): $R^2 = 0{,}908$, $F = 108{,}357$, $DW = 1{,}998$, $RSS = 5{,}881$.

Перейдем к описанию численного эксперимента. Требовалось оценить двухфакторную АСР (5) по алгоритму с предварительным выбором начального приближения (рис. 2). Требование вхождения в искомую модель ровно двух независимых переменных продиктовано необходимостью сохранения в ней числа степеней свободы тем же, что и в регрессиях (7) и (8), что позволит корректно сравнивать их между собой.

Стоит отметить, что в работе [12] предложены 2 стратегии построения аддитивных зависимостей (6): когда каждая независимая переменная входит в модель ровно 1 раз или произвольное число раз. В данном численном эксперименте рассматривались АСР только с единственным вхождением в них каждой независимой переменной.

Поскольку общее число независимых переменных, участвующих в процессе построения АСР, равно 6, а модели должны быть двухфакторными, то пришлось работать с каждой парой независимых переменных в отдельности, т. е. для каждой пары независимых переменных, общее число которых $C_6^2 = 15$, реализовывать алгоритм с предварительным выбором начального приближения.

Согласно алгоритму на рис. 2, на первом этапе для каждой из 15 пар независимых переменных были заданы ограничения на параметры $-5 \le b_j \le 5$, $j = \overline{1,2}$. Затем каждый такой отрезок был разбит на 9 отрезков точками: -5, -4, -3, -2, -1, 1, 2, 3, 4, 5. После чего оценивалось 100 линейных моделей, и выбиралась лучшая из них с наименьшей величиной суммы квадратов ошибок. Далее с использованием значений оценок лучшей регрессии в качестве начального приближения, находились оценки АСР методом Левенберга - Марквардта. И на последнем шаге решалась задача многокритериального выбора наилучшего из 15 регрессий уравнения, т. е. организовывался «конкурс» моделей.

Результаты оценивания АСР представлены в табл. 1. В ней в первом столбце указан номер пары переменных, во втором – обозначения переменных, в третьем – этапы алгоритма (рис. 2), первые три из которых обозначены а), а четвертый – б). В четвертом столбце приводятся оцененные модели согласно этапам алгоритма. При этом для этапа б) справа от регрессии в скобках указано число итераций. В пятом столбце указаны критерии адекватности соответствующих регрессий. Если сходимость алгоритма не была достигнута, то в соответствующих ячейках таблицы ставился символ «–».

Таблица 1

Результаты оценивания АСР

№	Переменные	Этап	Оцененные модели	Критерии адекватности (R^2; F; RSS; DW)
1	x_1, x_2	а)	$\hat{y} = \underset{(0,167)}{0,5} + \underset{(2,941)}{1,889}\, x_1 - \underset{(-2,109)}{6,39 \cdot 10^{-5}}\, x_2^5$	$(0,290;\ 4,500;\ 45,288;\ 0.964)$
		б)	$\hat{y} = \underset{(-1,747)}{-13,530} + \underset{(2,932)}{12,337}\, x_1^{0,4055} - \underset{(-2,039)}{0,00024}\, x_2^{4,3989}$ (151)	$(0,293;\ 4,564;\ 45,103;\ 0,982)$
2	x_1, x_3	а)	$\hat{y} = \underset{(3,927)}{6,227} + \underset{(3,192)}{0,918}\, x_1 - \underset{(-4,436)}{57,698}\, x_3^{-2}$	$(0,549;\ 13,426;\ 28,738;\ 1,735)$
		б)	$\hat{y} = \underset{(7,115)}{7.664} + \underset{(3,207)}{0,240}\, x_1^{1,5503} - \underset{(-4,517)}{86,210}\, x_3^{-2,3189}$ (47)	$(0,551;\ 13,506;\ 28,645;\ 1,695)$
3	x_1, x_5	а)	$\hat{y} = \underset{(8,627)}{12,610} - \underset{(-0,635)}{183,814}\, x_1^{-4} - \underset{(-1,699)}{56,844}\, x_5^{-1}$	$(0,304;\ 4,811;\ 44,397;\ 1,047)$
		б)	$\hat{y} = \underset{(5,738)}{14,291} - \underset{(-0,623)}{333,353}\, x_1^{-4,5028} - \underset{(-1,700)}{35,310}\, x_5^{-0,6814}$ (36)	$(0,304;\ 4,814;\ 44,389;\ 1,045)$
4	x_1, z	а)	$\hat{y} = \underset{(11,28)}{7,440} + \underset{(5,115)}{0,112}\, x_1^2 - \underset{(-7,038)}{1,019 \cdot 10^{-7}}\, z^5$	$(0,731;\ 29,881;\ 17,171;\ 2,427)$
		б)	$\hat{y} = \underset{(13,25)}{7,796} + \underset{(5,13)}{0,0598}\, x_1^{2,2964} - \underset{(-7,082)}{9,44 \cdot 10^{-9}}\, z^{5,6947}$ (55)	$(0,731;\ 29,936;\ 17,148;\ 2,436)$
5	x_1, x_7	а)	$\hat{y} = \underset{(8,626)}{4,091} + \underset{(4,919)}{0,0086}\, x_1^3 + \underset{(11,67)}{179,346}\, x_7^{-1}$	$(0,875;\ 77,376;\ 7,943;\ 1,643)$
		б)	$\hat{y} = \underset{(-9,769)}{-33,236} + \underset{(5,283)}{0,0021}\, x_1^{3,7260} + \underset{(12,28)}{60,649}\, x_7^{-0,0980}$ (142)	$(0,884;\ 84,094;\ 7,382;\ 1,896)$
6	x_2, x_3	а)	$\hat{y} = \underset{(11,57)}{8,662} + \underset{(2,873)}{0,0422}\, x_2^2 - \underset{(-4,547)}{218,225}\, x_3^{-3}$	$(0,516;\ 11,752;\ 30,854;\ 1,678)$
		б)	$\hat{y} = \underset{(8.612)}{8.209} + \underset{(2.851)}{0.145}\, x_2^{1.4979} - \underset{(-4.483)}{134.359}\, x_3^{-2.641}$ (76)	$(0.518;\ 11.803;\ 30.784;\ 1.721)$
7	x_2, x_5	а)	$\hat{y} = \underset{(4.275)}{5.365} - \underset{(-1.406)}{2.82 \cdot 10^{-5}}\, x_2^5 + \underset{(3.277)}{0.011}\, x_5^2$	$(0.336;\ 5.557;\ 42.396;\ 1.418)$
		б)	$\hat{y} = \underset{(3,796)}{5,061} - \underset{(-1,443)}{4,57 \cdot 10^{-9}}\, x_2^{8,8589} + \underset{(3,377)}{0,0218}\, x_5^{1,7780}$ (67)	$(0,341;\ 5,703;\ 42,026;\ 1,352)$
8	x_2, z	а)	$\hat{y} = \underset{(13,69)}{14.506} - \underset{(-3,785)}{24,881}\, x_2^{-1} - \underset{(-5,403)}{8,62 \cdot 10^{-5}}\, z^3$	$(0,642;\ 19,730;\ 22,843;\ 2,54)$
		б)	–	–
9	x_2, x_7	а)	$\hat{y} = \underset{(2,940)}{2,408} + \underset{(4,447)}{0,456}\, x_2^1 + \underset{(11,09)}{179,11}\, x_7^{-1}$	$(0,862;\ 68,930;\ 8,782;\ 1,80)$
		б)	$\hat{y} = \underset{(-5,484)}{-7,993} + \underset{(4,616)}{0,101}\, x_2^{1,5663} + \underset{(11,47)}{42,482}\, x_7^{-0,263}$ (68)	$(0,867;\ 71,974;\ 8,46;\ 1,967)$
10	x_3, x_5	а)	$\hat{y} = \underset{(13,.17)}{8,773} - \underset{(-3,964)}{178,207}\, x_3^{-3} + \underset{(3,139)}{4,44 \cdot 10^{-7}}\, x_5^5$	$(0,541;\ 12,953;\ 29,306;\ 1,807)$
		б)	$\hat{y} = \underset{(13,09)}{8,767} - \underset{(-3,958)}{174,486}\, x_3^{-2,9848} + \underset{(3,137)}{5,1 \cdot 10^{-7}}\, x_5^{4,9559}$ (24)	$(0,541;\ 12,953;\ 29,306;\ 1,809)$
11	x_3, z	а)	$\hat{y} = \underset{(24,32)}{10,973} - \underset{(-1,265)}{85,521}\, x_3^{-3} - \underset{(-2,057)}{0,0018}\, z^2$	$(0,442;\ 8,727;\ 35,585;\ 2,822)$
		б)	$\hat{y} = \underset{(22,44)}{11,219} - \underset{(-1,383)}{50,355}\, x_3^{-2,5248} - \underset{(-2,015)}{0,0105}\, z^{1,4999}$ (223)	$(0,443;\ 8,754;\ 35,535;\ 2,827)$
12	x_3, x_7	а)	$\hat{y} = \underset{(6,621)}{6,091} - \underset{(-0,556)}{6,956}\, x_3^{-2} + \underset{(5,852)}{165,464}\, x_7^{-1}$	$(0,742;\ 31,679;\ 16,447;\ 2,64)$
		б)	$\hat{y} = \underset{(6,359)}{6,124} - \underset{(-0,558)}{5,253}\, x_3^{-1,7518} + \underset{(5,864)}{165,79}\, x_7^{-1,0007}$ (52)	$(0,742;\ 31,684;\ 16,446;\ 2,641)$

№	Пере-ре-мен-ные	Этап	Оцененные модели	Критерии адекватности $(R^2; F; RSS; DW)$
13	x_5, z	а)	$\hat{y} = \underset{(15,58)}{8,822} + \underset{(3,98)}{4,9 \cdot 10^{-7}}\, x_5^5 - \underset{(-5,329)}{8,3 \cdot 10^{-5}}\, z^3$	$(0,656;\ 21;\ 21,93;\ 2,635)$
		б)	$\hat{y} = \underset{(16,24)}{8,971} + \underset{(4,014)}{1,7 \cdot 10^{-7}}\, x_5^{5,325} - \underset{(-5,387)}{0,0004}\, z^{2,5342}$ (24)	$(0,657;\ 21,106;\ 21,864;\ 2,597)$
14	x_5, x_7	а)	$\hat{y} = \underset{(26,43)}{11,887} + \underset{(6,295)}{4,2 \cdot 10^{-7}}\, x_5^5 - \underset{(-12,16)}{0,078}\, x_7$	$(0,898;\ 96,906;\ 6,505;\ 2,128)$
		б)	$\hat{y} = \underset{(17,11)}{33,4} + \underset{(6,467)}{1,43 \cdot 10^{-8}}\, x_5^{6,0545} - \underset{(-13,14)}{13,898}\, x_7^{0,154}$ (251)	$(0,908;\ 109,12;\ 5,844;\ 2,031)$
15	z, x_7	а)	$\hat{y} = \underset{(2,081)}{2,917} + \underset{(2,076)}{0,0608}\, z + \underset{(6,329)}{242,829}\, x_7^{-1}$	$(0,781;\ 39,331;\ 13,947;\ 2,104)$
		б)	$\hat{y} = \underset{(-4,064)}{-19.,3} + \underset{(2.317)}{0,0077}\, z^{1,5961} + \underset{(6.419)}{60,98}\, x_7^{-0,2055}$ (85)	$(0,79;\ 41,349;\ 13,409;\ 2.236)$

По табл. 1 видно, что алгоритм с предварительным выбором начального приближения позволяет всегда получить приемлемые оценки АСР (5), но, как уже отмечалось, не гарантирует их оптимальности. Благодаря процедуре предварительного выбора начального приближения из 15 моделей только для одной регрессии с переменными x_2, z метод Левенберга - Марквардта оказался не сходящимся, что может быть связано с неудачным выбором отрезка и точек разбиения для параметров b. Но даже в этом случае были определены неплохие оценки АСР.

Как было отмечено выше, первые 3 блока в алгоритме на рис. 2 представляют собой однокритериальный «конкурс» моделей. Результаты в табл. 1 демонстрируют, что использование метода Левенберг - Марквардта для выигравшей в «конкурсе» регрессии, как минимум, не ухудшает её аппроксимационные свойства, а в большинстве случаев улучшает их. Правда, такое улучшение в среднем не оказалось значительным. Например, максимальный рост значения критерия детерминации составил всего 0,01 для модели с переменными x_5, x_7. Но при этом значительно выросло значение критерия Фишера – на 12,214, снизилась величина суммы квадратов ошибок – на 0,661, улучшилось значение критерия Дарбина - Уотсона– с 2,128 до 2,031.

Проведенный по результатам таблицы 1 многокритериальный «конкурс» моделей позволил определить наилучшую АСР:

$$\hat{y} = \underset{(17,11)}{33,4} + \underset{(6,467)}{1,43 \cdot 10^{-8}}\, x_5^{6,0545} - \underset{(-13,14)}{13,898}\, x_7^{0,154}. \qquad (9)$$

Критерии адекватности АСР (9): $R^2 = 0,908$, $F = 109,12$, $RSS = 5,844$, $DW = 2,031$. Как видно, качество АСР (9) оказалось несколько выше, чем для регрессии (8). Таким образом, интеграция алгоритма оценивания АСР (5) в технологию «конкурса» моделей не ухудшает, а зачастую и вовсе улучшает результаты моделирования.

Заключение

Разработанный алгоритм оценивания АСР с предварительным выбором начального приближения может быть использован в сочетании в «конкурсом» моделей для улучшения качественных свойств получаемых регрессий. Его главным недостатком является отсутствие гарантии нахождения точки глобального минимума. Дальнейшим направлением этого исследования является модернизация или разработка нового алгоритма, гарантирующего достижение оптимального решения задачи. Помимо этого, научный интерес представляет исследование АСР (4) с мультипликативной ошибкой, а также АСР, в которые независимые переменные могут входить произвольное число раз.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Айвазян С.А. Методы эконометрики. М. : Магистр ; ИНФРА-М, 2010. 512 с.
2. Доугерти К. Введение в эконометрику. М. : ИНФРА-М, 2009. 465 с.
3. Клейнер Г.Б. Производственные функции. М. : Финансы и статистика, 1986. 239 с.
4. Клейнер Г.Б., Смоляк С.А. Эконометрические зависимости: принципы и методы построения. М. : Наука, 2000. 104 с.
5. Greene W.H. Econometric analysis. New York University, 2002. 994 p.
6. Gnu Regression, Econometrics and Time-series Library : site. URL: http://gretl.sourceforge.net/ru.html (access date: 27.10.17).
7. Демиденко Е.З. Линейная и нелинейная регрессия. М. : Финансы и статистика, 1981. 303 с.
8. Носков С.И. Технология моделирования объектов с нестабильным функционированием и неопределенностью в данных. Иркутск : Облинформпечать, 1996. 321 с.

9. Базилевский М.П., Носков С.И. Методические и инструментальные средства построения некоторых типов регрессионных моделей // Системы. Методы. Технологии. 2012. № 1. С. 80–87.
10. Базилевский М.П., Носков С.И. Технология организации конкурса регрессионных моделей // Информационные технологии и проблемы математического моделирования сложных систем. 2009. № 7. С. 77–84.
11. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. М. : Вильямс, 2007. 912 с.
12. Базилевский М.П. Программно-математическое обеспечение автоматизации многокритериального выбора регрессионных моделей : автореф. дис. … канд. техн. наук. Иркутск, 2012. 19 с.

REFERENCES

1. Aivazyan S.A. Metody ekonometriki [Methods of Econometric]. Moscow: Magistr ; INFRA-M Publ., 2010, 512 p.
2. Dougherty Ch. Introduction to Econometrics. Oxford University Press, 2007, 480 p. (Russ. ed.: Dougerti K. Vvedenie v ekonometriku. Moscow: INFRA-M Publ., 2009, 465 p.).
3. Kleiner G.B. Proizvodstvennye funktsii [Production functions]. Moscow: Finansy i statistika Publ., 1986, 239 p.
4. Kleiner G.B., Smolyak S.A. Ekonometricheskie zavisimosti: printsipy i metody postroeniya [Econometric dependencies: principles and methods of construction]. Moscow: Nauka Publ., 2000, 104 p.
5. Greene W.H. Econometric analysis. New York University, 2002, 994 p.
6. Gnu Regression, Econometrics and Time-series Library : site. URL: http://gretl.sourceforge.net/ru.html (access date: 27.10.17).
7. Demidenko E.Z. Lineinaya i nelineinaya regressiya [Linear and nonlinear regression]. Moscow: Finansy i statistika Publ., 1981, 303 p.
8. Noskov S.I. Tekhnologiya modelirovaniya ob"ektov s nestabil'nym funktsionirovaniem i neopredelennost'yu v dannykh [The technology of modeling objects with unstable functioning and uncertainty in the data]. Irkutsk: Oblinformpechat' Publ., 1996, 321 p.
9. Bazilevskii M.P., Noskov S.I. Metodicheskie i instrumental'nye sredstva postroeniya nekotorykh tipov regressionnykh modelei [Methodical and instrumental means for constructing some types of regression models]. *Sistemy. Metody. Tekhnologii* [*Systems. Methods. Technologies*], 2012, No. 1, pp. 80–87.
10. Bazilevskii M.P., Noskov S.I. Tekhnologiya organizatsii konkursa regressionnykh modelei [Technology of the competition of regression models]. *Informatsionnye tekhnologii i problemy matematicheskogo modelirovaniya slozhnykh system* [*Information technologies and problems of mathematical modeling of complex systems*], 2009. No. 7, pp. 77–84.
11. Draper N.R., Smith G. Applied Regression Analysis. Wiley-Interscience, 1998, 736 p. (Russ. ed.: Dreiper N., Smit G. Prikladnoi regressionnyi analiz. Moscow: Vil'yams Publ., 2007, 912 p.).
12. Bazilevskii M.P. Programmno-matematicheskoe obespechenie avtomatizatsii mnogokriterial'nogo vybora regressionnykh modelei : avtoref. dis. … kand. tekhn. nauk [Programmatic and mathematical support for the automation of multi-criteria choice of regression models: author's abstract. Ph.D. (Engineering) thesis]. Irkutsk, 2012, 19 p.

Информация об авторах

Базилевский Михаил Павлович - к. т. н., доцент кафедры «Математика», Иркутский государственный университет путей сообщения, г. Иркутск, e-mail: mik2178@yandex.ru

Authors

Mikhail Pavlovich Bazilevsky – Ph.D. in Engineering Science, Assoc. Prof., the Subdepartment of Mathematics, Irkutsk State Transport University, Irkutsk, e-mail: mik2178@yandex.ru

Для цитирования

Базилевский М. П. Разработка и исследование алгоритмов оценивания параметров аддитивной степенной регрессии // Современные технологии. Системный анализ. Моделирование. - 2017. - Т. 56, № 4. - С. 131–138. - DOI: 10.26731/1813-9108.2017.4(56).131-138.

For citation

Bazilevsky M. P. The development and research of algorithms of estimation of parameters of the additive degree regression. *Sovremennye tekhnologii. Sistemnyi analiz. Modelirovanie* [*Modern Technologies. System Analysis. Modeling*], 2017. Vol. 56, No.4, pp. 131–138. DOI: 10.26731/1813-9108.2017.4(56).131-138.

***УДК** 519.876.2, 004.421.2:519.178, 351.861(094.5)* ***DOI:** 10.26731/1813-9108.2017.4(56).138-144*

В. С. Асламова, Е. А. Темникова, В. Е. Гозбенко

Иркутский государственный университет путей сообщения, г. Иркутск, Российская Федерация
***Дата поступления:** 22 сентября 2017 г.*

АВТОМАТИЗАЦИЯ РАСЧЕТА КРАТЧАЙШЕГО ПУТИ ЭВАКУАЦИИ НАСЕЛЕНИЯ НА ТРАНСПОРТНОЙ СЕТИ С ЦИКЛОМ

Аннотация. *В статье приведены основные изменения, введенные в постановление Правительства РФ № 61 «Правила эвакуации населения, материальных и культурных ценностей в безопасные районы»: исключен термин «загородная зона», уточнены термины «рассредоточение», «зона возможных опасностей», определены полномочия руководителей гражданской обороны всех уровней, которые теперь обязаны осуществлять общее руководство проведением эвакуации и могут самостоятельно принимать решения о районах размещения рассредоточиваемых персонала и населения. Представлены математическая модель, алгоритм и программа автоматизированного расчета кратчайшего пути заблаговременной эвакуации населения, материальных и культурных ценностей в безопасный район с использованием личного, городского грузового и пассажирского транспорта. Математическая модель оптимизационной задачи сформулирована в виде линейной сетевой модели (графа), представляющей маршрут передвижения транспорта по существующей сети автомобильных дорог с циклом и имеющий наименьшую протяженность. Исходной вершиной маршрута служит сборный эвакуационный пункт либо место проживания,*