2023 Электротехника, информационные технологии, системы управления № 45

Научная статья

DOI 10.15593/2224-9397/2023.1.06
УДК 519.862.6

М.П. Базилевский, А.Б. Ойдопова

Иркутский государственный университет путей сообщения,
Иркутск, Россия

ОЦЕНИВАНИЕ МОДУЛЬНЫХ ЛИНЕЙНЫХ РЕГРЕССИОННЫХ МОДЕЛЕЙ С ПОМОЩЬЮ МЕТОДА НАИМЕНЬШИХ МОДУЛЕЙ

Разработка новых видов математических моделей и методов их построения является актуальной научной задачей. Для построения математических моделей статистического типа активно применяются методы регрессионного анализа. Адекватная регрессионная модель позволяет не только прогнозировать значения зависимой переменной, но и объяснять механизм воздействия входных переменных на выходную, что повышает эффективность управленческих решений. Ранее авторами были предложены модели модульной линейной регрессии, но алгоритм их точного оценивания изобретен не был. **Цель исследования**: свести задачу точного оценивания модульных линейных регрессионных моделей с помощью метода наименьших модулей к задаче частично-булевого линейного программирования. **Результаты**: в данной работе для решения задачи оценивания модульной регрессионной модели с помощью метода наименьших модулей предложен алгоритм, согласно которому требуется решить серию задач частично-булевого линейного программирования и выбрать из полученных решений такое, у которого величина суммы модулей остатков минимальна. Предложенный подход был применен для моделирования грузовых железнодорожных перевозок Забайкальского края. Построенная модульная регрессия по величине суммы модулей остатков оказалась более чем в 3 раза лучше, чем оцененная с помощью метода наименьших модулей традиционная линейная регрессия. Дана интерпретация модульной регрессии. **Практическая значимость**: предложенный способ точного оценивания с помощью метода наименьших модулей модульных линейных регрессий может быть использован для решения конкретных прикладных задач анализа данных из любых предметных областей, что предполагает разработку соответствующего программного обеспечения. Модульные регрессии больше подойдут для моделирования процессов или явлений с переменными направлениями влияния входных переменных на выходную.

Ключевые слова: регрессионная модель, модульная регрессия, метод наименьших модулей, задача частично-булевого линейного программирования, отправление груза, экспорт.

M.P. Bazilevskiy, A.B. Oydopova

Irkutsk State Transport University, Irkutsk, Russian Federation

ESTIMATION OF MODULAR LINEAR REGRESSION MODELS USING THE LEAST ABSOLUTE DEVIATIONS

The development of new types of mathematical models and methods of their construction is an urgent scientific task. Regression analysis methods are actively used to build mathematical models of statistical type. An adequate regression model allows not only to predict the values of the dependent variable, but also to explain the mechanism of influence of input variables on the output, which increases the effectiveness of management decisions. Previously, the authors proposed modular linear regression models, but an algorithm for their accurate estimation was not invented. **The aim of the study** is to reduce the problem of accurate estimation of modular linear regression models using the least absolute deviations method to the mixed-integer 0-1 linear programming problem. **Results:** in this paper, in order to solve the problem of estimating a modular regression model using the least absolute deviations, an algorithm is proposed according to which it is required to solve a series of mixed-integer 0-1 linear programming problems and choose from the solutions obtained one in which the sum of the modules of the residuals is minimal. The proposed approach was applied to the simulation of freight rail transport in the Zabaikalsky kray. The constructed modular regression in terms of the sum of the modules of the residuals turned out to be more than 3 times better than the traditional linear regression estimated using the least absolute deviations. The interpretation of modular regression is given. **Practical significance:** the proposed method of accurate estimation using the least absolute deviations of modular linear regressions can be used to solve specific applied problems of data analysis from any subject areas, which involves the development of appropriate software. Modular regressions are more suitable for modeling processes or phenomena with variable directions of influence of input variables on output.

Keywords: regression model, modular regression, least absolute deviations, mixed-integer 0-1 linear programming, shipment, export.

Введение

В настоящее время для построения математических зависимостей между исследуемыми переменными (показателями, факторами) активно применяются методы регрессионного анализа [1, 2]. При этом, как отмечено в [3], ключевым этапом построения регрессионной модели является выбор её спецификации, т.е. состава переменных и формы связи между ними. К известным спецификациям относятся так называемые в экономике производственные функции [4]: линейная, Леонтьева, Кобба – Дугласа и т.д. Вместе с тем продолжается разработка новых спецификаций. Так, в [5] предложена кусочно-линейная регрессия с минимальным и максимальным вкладом независимых переменных, в [6] – кусочно-линейная авторегрессионная модель, в [7] – неэлементарная линейная регрессия, в [8] – индексная регрессия. Стоит отметить, что работы [5–8] объединяет то, что в них для оценки неизвестных параметров моделей применяется аппарат частично-булевого математического программирования, технология решения задач которого была существенно развита за последние пару десятков лет (см., например, [9–12]). Причем в работах [5, 6, 8] оценка моделей осуществляется с помощью метода наименьших модулей (МНМ) [13, 14], а в [7] – с помощью метода наименьших квадратов (МНК).

В работе [15] была предложена модель модульной линейной регрессии, для приближенной оценки которой с помощью МНК был разработан специальный алгоритм. Целью данной работы является разработка способа точного оценивания модульных регрессий с помощью МНМ.

1. МНМ-оценивание модульных регрессий

Важно отметить, что разработать описываемый в данном разделе способ МНМ-оценивания модульных регрессий помогли математические приемы, рассмотренные в монографии [13].

Модель модульной линейной регрессии [15] имеет вид:

$$y_i = \alpha_0 + \sum_{j=1}^{l} \alpha_j \left| x_{ij} - \lambda_j \right| + \varepsilon_i, \; i = \overline{1,n}, \tag{1}$$

где n – объем выборки; l – число объясняющих переменных; y_i, $j = \overline{1,n}$ – наблюдаемые значения объясняемой переменной y; x_{ij}, $i = \overline{1,n}$, $j = \overline{1,l}$ – наблюдаемые значения объясняющих переменных x_1, x_2, …, x_l; ε_i, $j = \overline{1,n}$ – ошибки аппроксимации; α_0, α_j, λ_j, $j = \overline{1,l}$ – неизвестные параметры.

Очевидно, что если $x_{ij} \geq 0$, а $\lambda_1 = \lambda_2 = \ldots = \lambda_l = 0$, то модульная регрессия (1) становится классической моделью множественной линейной регрессии. Для нахождения МНМ-оценок модульной регрессии (1) требуется решить оптимизационную задачу:

$$\sum_{i=1}^{n} \left| \varepsilon_i \right| = \sum_{i=1}^{n} \left| y_i - \alpha_0 - \sum_{i=1}^{l} \alpha_j \left| x_{ij} - \lambda_j \right| \right| \to \min. \tag{2}$$

Обозначим $\left| \alpha_j \right| = k_j \geq 0, \; j = \overline{1,l}.$ Тогда модель (1) можно записать в виде:

$$y_i = \alpha_0 + \sum_{j=1}^{l} (-1)^{\sigma j} \cdot k_j \left| x_{ij} - \lambda_j \right| + \varepsilon_i, \; i = \overline{1,n}, \tag{3}$$

где σ_j, $j=\overline{1,l}$ – бинарные переменные, заданные по правилу:

$$\sigma_j=\begin{cases}0, \text{ если } \alpha_j \ge 0,\\ 1, \text{ если } \alpha_j < 0.\end{cases}$$

Это означает, что если $\sigma_j = 0,$ то параметр α_j неотрицателен, а если $\sigma_j = 1,$ то отрицателен.

Используя известное свойство модуля $|a|\cdot|b|=|a\cdot b|$, перепишем выражение (3) в виде:

$$y_i=\alpha_0+\sum_{j=1}^{l}(-1)^{\sigma j}\left|k_j x_{ij}-\beta_j\right|+\varepsilon_i,\ i=\overline{1,n}, \tag{4}$$

где $\beta_j = k_j \cdot \lambda_j$, $j=\overline{1,l}.$

Рассмотрим случай, когда параметры α_j модульной регрессии неотрицательны, т.е. когда в модели (4) $\sigma_1=\sigma_2=...\sigma_l=0$:

$$y_i=\alpha_0+\sum_{j=1}^{l}\left|k_j x_{ij}-\beta_j\right|+\varepsilon_i,\ i=\overline{1,n}. \tag{5}$$

В выражении (5) раскроем модуль, введя переменные:

$$\mu_{ij}=\begin{cases}k_j x_{ij}-\beta_j, \text{ если } k_j x_{ij}-\beta_j \ge 0,\\ 0, \text{ если } k_j x_{ij}-\beta_j < 0,\end{cases}\quad i=\overline{1,n},\ j=\overline{1,l}.$$

$$v_{ij}=\begin{cases}0, \text{ если } k_j x_{ij}-\beta_j \ge 0,\\ k_j x_{ij}+\beta_j, \text{ если } k_j x_{ij}-\beta_j < 0,\end{cases}\quad i=\overline{1,n},\ j=\overline{1,l}.$$

Тогда будут справедливы следующие тождества:

$$k_j x_{ij}-\beta_j=\mu_{ij}-v_{ij},\ i=\overline{1,n},\ j=\overline{1,l}. \tag{6}$$

При этом выражение (5) примет вид:

$$y_i=\alpha_0+\sum_{j=1}^{l}(\mu_{ij}+v_{ij})+\varepsilon_i,\ i=\overline{1,n}. \tag{7}$$

Введем бинарные переменные δ_{ij}, $i=\overline{1,n}$, $j=\overline{1,l}$, по правилам:

$$\delta_{ij}=\begin{cases}0, \text{ если } k_j x_{ij}-\beta_j < 0,\\ 1, \text{ если } k_j x_{ij}-\beta_j \ge 0,\end{cases}\quad i=\overline{1,n},\ j=\overline{1,l}.$$

Это означает, что если $\delta_{ij} = 0$, то срабатывает переменная v_{ij} (модуль $\left|k_j x_{ij} - \beta_j\right|$ раскрывается со знаком «–», а если $\delta_{ij} = 1$, то переменная μ_{ij} (модуль $\left|k_j x_{ij} - \beta_j\right|$ раскрывается со знаком «+»).

Для обеспечения срабатывания для любого i и j либо переменной μ_{ij}, либо v_{ij} введем ограничения:

$$\mu_{ij} \le M \cdot \delta_{ij},\ i = \overline{1,n},\ \ j = \overline{1,l}, \tag{8}$$

$$v_{ij} \le M(1-\delta_{ij}),\ i = \overline{1,n},\ \ j = \overline{1,l}, \tag{9}$$

где M – большое положительное число.

Теперь раскроем модуль для функции потерь (2), введя переменные:

$$g_i = \begin{cases} y_j - \alpha_0 - \sum\limits_{j=1}^{l}(\mu_{ij} + v_{ij}), \text{ если } y_j - \alpha_0 - \sum\limits_{j=1}^{l}(\mu_{ij} - v_{ij}) \ge 0, \\ 0, \text{ в противном случае}, \end{cases} \ i = \overline{1,n}. \quad .$$

$$h_i = \begin{cases} -y_j + \alpha_0 + \sum\limits_{j=1}^{l}(\mu_{ij} + v_{ij}), \text{ если } y_j - \alpha_0 - \sum\limits_{j=1}^{l}(\mu_{ij} - v_{ij}) < 0, \\ 0, \text{ в противном случае}, \end{cases} \ i = \overline{1,n}.$$

Тогда выражение (7) нужно представить в виде:

$$y_i = \alpha_0 + \sum_{j=1}^{l}(\mu_{ij} + v_{ij}) + g_i - h_1,\ i = \overline{1,n}, \tag{10}$$

а целевую функцию (2) заменить функционалом:

$$\sum_{j=1}^{l}(g_i + h_i) \to \min. \tag{11}$$

Таким образом, задача МНМ-оценивания модульной регрессии (1) для случая $\sigma_1 = \sigma_2 = ...\sigma_l = 0$ свелась к задаче частично-булевого линейного программирования (ЧБЛП) с целевой функцией (11), линейными ограничениями (6), (8)–(10) и условиями неотрицательности переменных k_j, μ_{ij}, v_{ij}, g_i, h_i.

Введем ограничения:

$$y_i = \alpha_0 + \sum_{j=1}^{l}(-1)^{\sigma j}(\mu_{ij} + v_{ij}) + g_i - h_1,\ i = \overline{1,n}. \tag{12}$$

Тогда для получения общих МНМ-оценок модульной регрессии (1) необходимо решить 2^l задач ЧБЛП с целевой функцией (11), линейными ограничениями (6), (8), (9), (12) и условиями неотрицательности переменных k_j, μ_{ij}, v_{ij}, g_i, h_i, перебирая все возможные комбинации бинарных переменных σ_j в ограничениях (12), а затем выбрать лучшую модель с минимальной величиной суммы модулей ошибок.

2. Решение прикладной задачи

В наши дни внешние торговые связи играют важнейшую роль в экономике любого государства. От экспорта и импорта грузов зависит не только экономическое благополучие страны, но и возможность удовлетворять потребности населения в необходимых товарах. Доставка грузов невозможна без транспорта, особенно без железнодорожного.

Грузооборот железнодорожного транспорта является одним из важнейших оперативных индикаторов экономической активности в реальном секторе, в частности, экспортной активности. Внешние и внутренние факторы оказывают существенное влияние на динамику и структуру грузооборота железнодорожного транспорта. Главная задача всего этого процесса – формирование конкурентоспособных транспортных составляющих логистических систем экономических субъектов рынка (грузовладельцев), позволяющих удовлетворять их потребности не только с точки зрения возможности перемещения продукта, но и при обязательном выполнении требований, предъявляемых к качеству перевозки [16].

Забайкальский край расположен на границе с КНР. В настоящее время Забайкальский край – это динамично развивающийся край, крупный административный и индустриальный центр. Через Читу (столицу края) проходят все важнейшие воздушные трассы, Транссибирская железнодорожная магистраль и т.д. На сегодняшний день многие предприятия Забайкальского края вовлечены в производственно-торговые отношения с зарубежными предприятиями и организациями [16]. Математическое моделирование грузовых потоков в реальном времени требует большого количества вычислительных ресурсов, совершенствования математических моделей и численных методов, разработки современного программного обеспечения [17].

Отправление груза зависит от многих факторов (рис. 1).

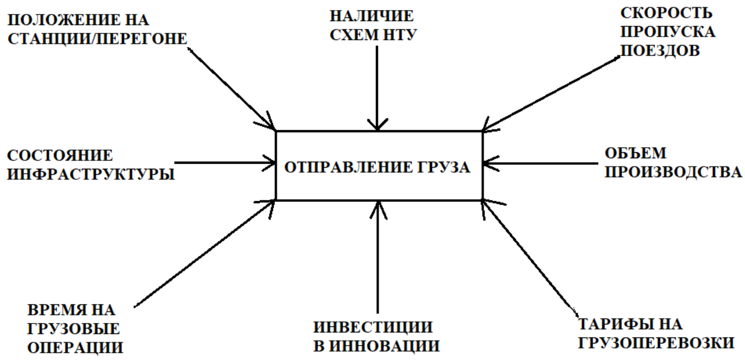


Рис. 1. Факторы, влияющие на отправление груза железнодорожным транспортом

Непосредственно количество отправленного груза связано с размером экспорта. Эта связь носит прямой характер. Рассмотрим также такой фактор, как наличие схем НТУ (непредусмотренные технические условия). Данная схема требуется для тех вагонов, в которых не предусмотрены техническими условиями способы размещения и крепления грузов для разовых либо нерегулярных перевозок грузов, имеющих строго определённые размеры и массу.

Схемы НТУ должны быть разработаны непосредственно к прибытию вагонов на станцию, но на практике наблюдается обратная ситуация – вагоны прибывают без схем. Они разрабатываются уже непосредственно после прибытия. На разработку и утверждение схемы затрачивается от трех дней (разработка схемы, создание реквизита крепления, работа бригады и т.д.). Соответственно возникают задержки при отправлении составов, которые затем суммируются и влияют на общее количество отправленного груза в год. Чем больше вагонов, требующих разработку схему НТУ, тем меньше отправленного груза [18]. Перейдем к построению математической модели [19–23].

Для построения модульной регрессии использовались статистические данные (табл. 1) по годам (с 2010 по 2020), полученные от сотрудников территориального центра фирменного транспортного обслуживания Забайкальской железной дороги для следующих переменных: y – отправление грузов железнодорожным транспортом (млн т); x_1 – вагоны, требующие разработки НТУ (ваг); x_2 – экспорт (ваг).

Таблица 1

Статистические данные

Год	y	x_1	x_2
2010	12169	6520	14523
2011	12375	6852	16240
2012	10895	7413	16542
2013	11383	7123	17890
2014	10326	6810	23650
2015	9455	7806	27406
2016	9531	9038	18252
2017	11473	7244	60080
2018	12690	13240	38018
2019	14876	4954	36423
2020	15288	5056	51778

Графики временных рядов для каждой переменной приведены на рис. 2, *а–в*.

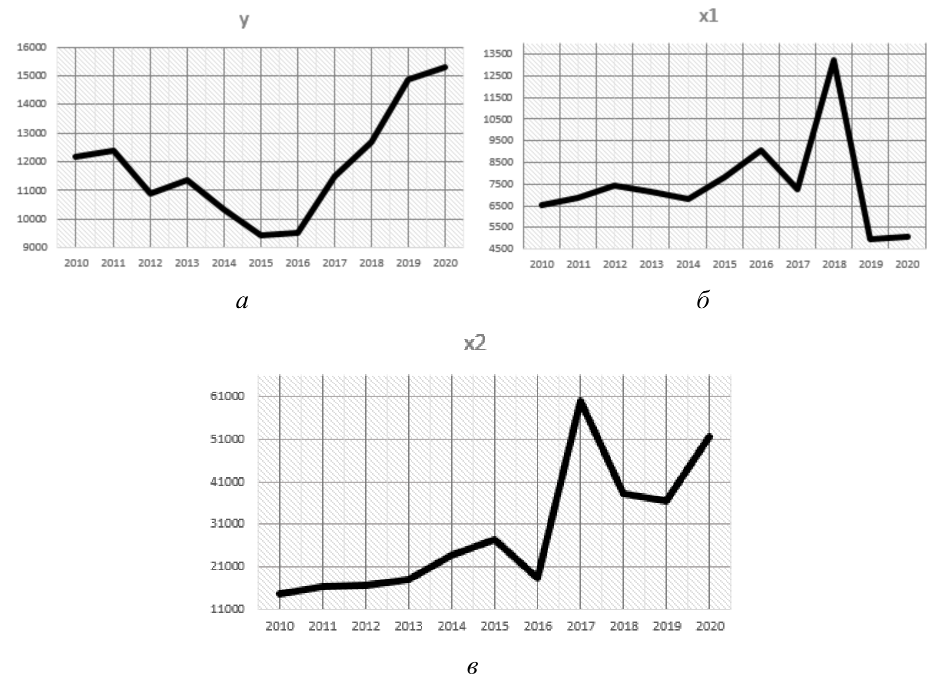


Рис. 2. Графики временных рядов

Анализируя графики, можно заметить, что отправление грузов железнодорожным транспортом снижалось до 2016 г. Связано это непосредственно с осложнением макроэкономической ситуации и внешними негативными геополитическими факторами. С 2017 г. отправление груза (как и размер экспорта) начало увеличиваться. На фоне девальвации и снижении внутреннего спроса начал расти экспорт. Перевозчики вышли на тесное сотрудничество с Европой и Азией, поставляя в соседние страны отечественные товары. Также с 2010 до 2019 г. наблюдается увеличение количества вагонов, требующих разработку схем НТУ. Это, в свою очередь, свидетельствует о том, что с 2010 г. возрастал спрос на негабаритные грузы, технику, оборудование. После 2019 г. количество вагонов резко сократилось из-за начавшейся пандемии коронавирусной инфекции, которая повлияла на снижение грузооборота в целом по сети. Корреляционная матрица для переменных y, x_1 и x_2 приведена в табл. 2.

Таблица 2

Корреляционная матрица

	y	x_1	x_2
y	1,0000	–0,3461	0,4670
x_1	–0,3461	1,0000	–0,0469
x_2	0,4670	–0,0469	1,0000

По корреляционной матрице видно, что объясняющие переменные слабо коррелируют между собой. Из этого следует, что во множественной регрессии будет отсутствовать эффект мультиколлинеарности.

С помощью МНМ [24, 25] в эконометрическом пакете Gretl было получено уравнение модульной регрессии:

$$\tilde{y} = 17351{,}8 - 0{,}9776 x_1 + 0{,}0556 x_2 . \tag{13}$$

Модель (13) можно интерпретировать следующим образом: с увеличением количества вагонов, требующих разработку НТУ x_1 на 1 вагон, отправление груза уменьшается на 0,977613 млн. т, а с увеличением размера экспорта – увеличится на 0,055603 млн. т.

Сумма модулей остатков регрессии (13) составила 13452,29.

Для получения общих МНМ-оценок модульной регрессии (1) с помощью LPSolve решалось $2^2 = 4$ задачи ЧБЛП с целевой функцией (11), линейными ограничениями (6), (8), (9), (12) и условиями неотрицательности переменных k_j, μ_{ij}, v_{ij}, g_i, h_i. Результаты решения задач оказались следующими:

1) при $\sigma_1 = 0$, $\sigma_2 = 0$ получено уравнение:

$$\tilde{y} = 5849,780 + |1,893x_1 - 18346,193| + |0,039x_2 - 1376,759|; \quad (14)$$

2) при $\sigma_1 = 0,\ \sigma_2 = 1$ получено уравнение:

$$\tilde{y} = 7647,793 + |1,523x_1 - 14954,445| - |0,012x_2 - 616,447|; \quad (15)$$

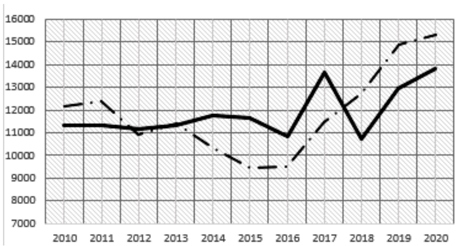
3) при $\sigma_1 = 1,\ \sigma_2 = 0$ получено уравнение:

$$\tilde{y} = 10453,929 - |0,028x_1 - 140,313| + |0,172x_2 - 4066,912|; \quad (16)$$

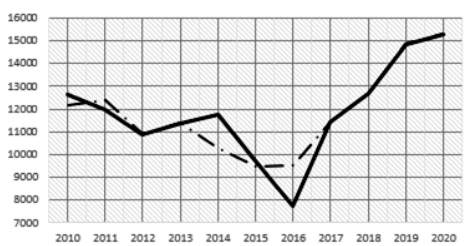
4) при $\sigma_1 = 1,\ \sigma_2 = 1$ получено уравнение:

$$\tilde{y} = 16370,777 - |0,303x_1 - 1533,022| - |0,166x_2 - 7504,607|. \quad (17)$$

Для регрессии (14) сумма модулей остатков оказалась равной 4380,225, для (15) – 4738,719, для (16) – 12383,272, для (17) – 10857,490. Таким образом, среди регрессий (14)–(17) самой лучшей по величине суммы модулей остатков оказалась модель (14). Причем эта величина оказалась в 3,071 меньше, чем для линейной регрессии (13). Превосходство модульной регрессии (14) над линейной регрессией (13) по качеству продемонстрировано на рис. 3, *а*, *б*, на каждом из которых приводятся фактические (штрихпунктирная линия) и расчетные (сплошная линия) по моделям значения зависимой переменной y.



а



б

Рис. 3. Графики фактических и расчетных значений зависимой переменной *y*: *а* – линейная регрессия (13); *б* – модульная регрессия (14)

Представим полученную модульную регрессию (14) в кусочно-заданном виде:

$$y = \begin{cases} -13873,172 + 1,893x_1 + 0,039x_2, \text{ при } x_1 \geq 9691,597, \text{ при } x_2 \geq 35301,513; \\ -11119,654 + 1,893x_1 - 0,039x_2, \text{ при } x_1 \geq 9691,597, \text{ при } x_2 < 35301,513; \\ 22819,214 - 1,893x_1 + 0,039x_2, \text{ при } x_1 < 9691,597, \text{ при } x_2 \geq 35301,513; \\ 25572,732 - 1,893x_1 - 0,039x_2, \text{ при } x_1 < 9691,597, \text{ при } x_2 < 35301,513. \end{cases}$$

Тогда модульную регрессию (14) можно интерпретировать следующим образом:

– если число вагонов x_1, требующих разработки НТУ, не меньше, чем 9691,597, а объемы экспорта x_2 не меньше, чем 35301,513, то на объемы грузоперевозок y переменные x_1 и x_2 влияют со знаком «+», причем с ростом x_1 на 1 вагон грузоперевозки y возрастают в среднем на 1,893 млн т, а с ростом x_2 на 1 вагон – возрастают в среднем на 0,039 млн т;

– если число вагонов x_1, требующих разработки НТУ, не меньше, чем 9691,597, а объемы экспорта x_2 меньше, чем 35301,513, то на объемы грузоперевозок y переменная x_1 влияет со знаком «+», а переменная x_2 со знаком «–», причем с ростом x_1 на 1 вагон грузоперевозки y возрастают в среднем на 1,893 млн т, а с ростом x_2 на 1 вагон – убывают в среднем на 0,039 млн т;

– если число вагонов x_1, требующих разработки НТУ, меньше, чем 9691,597, а объемы экспорта x_2 не меньше, чем 35301,513, то на объемы грузоперевозок y переменная x_1 влияет со знаком «–», а переменная x_2 со знаком «+», причем с ростом x_1 на 1 вагон грузоперевозки y убывают в среднем на 1,893 млн т, а с ростом x_2 на 1 вагон – возрастают в среднем на 0,039 млн т;

– если число вагонов x_1, требующих разработки НТУ, меньше, чем 9691,597, а объемы экспорта x_2 меньше, чем 35301,513, то на объемы грузоперевозок y переменные x_1 и x_2 влияют со знаком «–», причем с ростом x_1 на 1 вагон грузоперевозки y убывают в среднем на 1,893 млн т, а с ростом x_2 на 1 вагон – убывают в среднем на 0,039 млн т.

Таким образом, в модульных регрессиях, в отличие от традиционных линейных регрессий, знаки коэффициентов при объясняющих переменных не являются постоянными, а меняются в зависимости от выполненных условий переключения.

Заключение

В статье предложен математический аппарат для оценивания модульных регрессий с помощью МНМ. Решена задача моделирования железнодорожных грузовых перевозок Забайкальского края. В качестве объясняющих переменных выбраны количество вагонов, требующих разработку схем НТУ, и размер экспорта на отправление груза на Забайкальской железной дороге. По величине суммы модулей остатков модульная регрессия оказалась более чем в 3 раза лучше, чем традиционная линейная регрессия. Дана интерпретация модульной регрессии, которая позволила установить, что её отличие от линейной регрессии заключается в том, что знаки коэффициентов при объясняющих переменных в ней не являются постоянными, а меняются в зависимости от выполненных условий переключения, что в совокупности влияет на качество модели. Предложенный в статье математический аппарат будет использован в дальнейшем для разработки программного комплекса построения модульных регрессионных моделей.

Научная новизна работы состоит в том, что впервые предложен метод точной идентификации неизвестных параметров модульной регрессии. Модульные регрессии могут найти широкое применение при решении реальных задач анализа данных, поскольку их прогностические способности всегда не хуже, а зачастую лучше, чем прогностические способности традиционных моделей множественной линейной регрессии. В будущем также планируется сравнить результаты решения задач предложенным методом с другими методами анализа данных, основанных на машинном обучении.

Библиографический список

1. Montgomery D.C., Peck E.A., Vining G.G. Introduction to linear regression analysis. – John Wiley & Sons, March 2021. – 704 p.

2. Arkes J. Regression analysis: A practical introduction / Routledge. – February 2019. – 362 p.

3. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. – М.: ЮНИТИ, 1998. – 1000 с.

4. Клейнер Г.Б. Производственные функции: теория, методы, применение. – М.: Финансы и статистика, 1986. – 238 с.

5. Носков С.И., Хоняков А.А. Программный комплекс построения некоторых типов кусочно-линейных регрессий // Информационные технологии и математическое моделирование в управлении сложными системами. – 2019. – № 3 (4). –С. 47–55.

6. Носков С.И. Построение кусочно-линейной авторегрессионной модели произвольного порядка // Вестник Югорск. гос. ун-та. – 2022. – № 2 (65). – С. 89–94.

7. Базилевский М.П. Метод построения неэлементарных линейных регрессий на основе аппарата математического программирования // Проблемы управления. – 2022. – № 4. – С. 3–14.

8. Базилевский М.П. Оценивание индексных моделей регрессии с помощью метода наименьших модулей // Вестник Российского нового университета. Сер. Сложные системы: модели, анализ и управление. – 2020. – № 1. – С. 17–23.

9. Konno H., Yamamoto R. Choosing the best set of variables in regression analysis using integer programming // Journal of Global Optimization. – 2009. – Vol. 44, No. 2. – P. 273–282.

10. Park Y.W., Klabjan D. Subset selection for multiple linear regression via optimization // Journal of Global Optimization. – 2020. – Vol. 77, No. 3. – P. 543–574.

11. Chung S., Park Y.W., Cheong T. A mathematical programming approach for integrated multiple linear regression subset selection and validation // Pattern Recognition. – 2020. – Vol. 108. – P. 107565.

12. Bertsimas D., Li M.L. Scalable holistic linear regression // Operations Research Letters. – 2020. – Vol. 48, No. 3. – P. 203–208.

13. Носков С.И. Технология моделирования объектов с нестабильным функционированием и неопределенностью в данных. – Иркутск: Облинформпечать, 1996. – 320 с.

14. Bloomfield P., Steiger W. L. Least absolute deviations: theory, applications, and algorithms. – Boston: Birkhäuser, 1983. – 365 p.

15. Базилевский М.П., Ойдопова А.Б. Моделирование выбросов загрязняющих веществ в атмосферу Забайкальского края // Информационные технологии и математическое моделирование в управлении сложными системами. – 2022. – № 2 (14). – С. 8–18.

16. Железнов Д.В., Светлакова Е.Н., Долгополова М.А. Создание логистического центра по переработке скоропортящихся грузов на станции Забайкальск // Современные технологии. Системный анализ. Моделирование. – 2006. – № 4. – С. 204–207.

17. Введение в математическое моделирование транспортных потоков / А.В. Гасников, С.Л. Кленов, Е.А. Нурминский, Я.А. Холодов, Н.Б. Шамрай; приложения: М.Л. Бланк, Е.В. Гасникова, А.А. Замятин, В.А. Малышев, А.В. Колесников, А.М. Райгородский. – М.: Изд-во МФТИ, 2010. – 360 с.

18. Полярин Ю.А. О новых требованиях к размещению и креплению грузов в вагонах и универсальных контейнерах [Электронный ресурс] // Склад и техника. – 12 июля 2017. – №11/2004. – URL: https://sitmag.ru/article/10625-o-novyh-trebovaniyah-k-razmeshcheniyu-i-

krepleniyu-gruzov-v-vagonah-i-universalnyh-konteynerah (дата обращения: 18.10.2022).

19. Аюпов В.В. Математическое моделирование технических систем. – Пермь: ИПЦ «Прокрость», 2017. – 242 с.

20. Звонарев С.В. Основы математического моделирования. – Екатеринбург: Изд-во Урал. ун-та, 2019. – 112 с.

21. Хакимзянов Г.С., Чубаров Л.Б., Воронина П.В. Математическое моделирование. – Новосибирск: РИЦ НГУ, 2014. – 263 с.

22. Mansfield E.R., Helms B.P. Detecting multicollinearity // The American Statistician. – 1982. – Vol. 36, № 3a. – P. 158–160.

23. Самарский А.А., Михайлов А.П. Математическое моделирование: Идеи. Методы. Примеры. – М.: Физматлит, 2001. – 320 с.

24. Gorgees S.M. Statistical inference for least absolute deviation regression with autocorrelated errors // Journal of Southwest Jiaotong University. – 2022. – Vol. 55 (2).

25. Wang X. Group selection via adjusted weighted least absolute deviation regression // Journal of Computational and Applied Mathematics. – 2020. – Vol. 378. – P. 112924.

References

1. Montgomery D.C., Peck E.A., Vining G.G. Introduction to linear regression analysis. John Wiley & Sons, March, 2021, 704 p.

2. Arkes J. Regression analysis: A practical introduction. Routledge, February 2019, 362 p.

3. Aivazian S.A., Mkhitarian V.S. Prikladnaia statistika i osnovy ekonometriki [Applied Statistics and Fundamentals of Econometrics]. Moscow: IuNITI, 1998, 1000 p.

4. Kleiner G.B. Proizvodstvennye funktsii: teoriia, metody, primenenie [Production functions: theory, methods, application]. Moscow: *Finansy i statistika*, 1986, 238 p.

5. Noskov S.I., Khoniakov A.A. Programmnyi kompleks postroeniia nekotorykh tipov kusochno-lineinykh regressii [Software complex for constructing some types of piecewise linear regressions]. *Informatsionnye tekhnologii i matematicheskoe modelirovanie v upravlenii slozhnymi sistemami*, 2019, no. 3 (4), pp. 47-55.

6. Noskov S.I. Postroenie kusochno-lineinoi avtoregressionnoi modeli proizvol'nogo poriadka [Construction of a piecewise linear autoregressive

model of an arbitrary order]. *Vestnik Iugorskogo gosudarstvennogo universiteta*, 2022, no. 2 (65), pp. 89-94.

7. Bazilevskii M.P. Metod postroeniia neelementarnykh lineinykh regressii na osnove apparata matematicheskogo programmirovaniia [A method for constructing non-elementary linear regressions based on the apparatus of mathematical programming]. *Problemy upravleniia*, 2022, no. 4, pp. 3-14.

8. Bazilevskii M.P. Otsenivanie indeksnykh modelei regressii s pomoshch'iu metoda naimen'shikh modulei [Estimation of index regression models using the method of least modules]. *Vestnik Rossiiskogo novogo universiteta. Slozhnye sistemy: modeli, analiz i upravlenie*, 2020, no. 1, pp. 17-23.

9. Konno H., Yamamoto R. Choosing the best set of variables in regression analysis using integer programming. *Journal of Global Optimization*, 2009, vol. 44, no. 2, pp. 273-282.

10. Park Y.W., Klabjan D. Subset selection for multiple linear regression via optimization. *Journal of Global Optimization*, 2020, vol. 77, no. 3, pp. 543-574.

11. Chung S., Park Y.W., Cheong T. A mathematical programming approach for integrated multiple linear regression subset selection and validation. *Pattern Recognition*, 2020, vol. 108, 107565 p.

12. Bertsimas D., Li M.L. Scalable holistic linear regression. *Operations Research Letters*, 2020, vol. 48, no. 3, pp. 203-208.

13. Noskov S.I. Tekhnologiia modelirovaniia ob"ektov s nestabil'nym funktsionirovaniem i neopredelennost'iu v dannykh [Technology for modeling objects with unstable operation and uncertainty in data]. Irkutsk: Oblinformpechat', 1996, 320 p.

14. Bloomfield P., Steiger W. L. Least absolute deviations: theory, applications, and algorithms. Boston: Birkhäuser, 1983, 365 p.

15. Bazilevskii M.P., Oidopova A.B. Modelirovanie vybrosov zagriazniaiushchikh veshchestv v atmosferu Zabaikal'skogo kraia [Modeling of emissions of pollutants into the atmosphere of the Zabaikalsky krai]. *Informatsionnye tekhnologii i matematicheskoe modelirovanie v upravlenii slozhnymi sistemami*, 2022, no. 2 (14), pp. 8-18.

16. Zheleznov D.V., Svetlakova E.N., Dolgopolova M.A. Sozdanie logisticheskogo tsentra po pererabotke skoroportiashchikhsia gruzov na stantsii Zabaikal'sk [Creation of a logistics center for the processing of per-

ishable goods at the Zabaikalsk station]. *Sovremennye tekhnologii. Sistemnyi analiz. Modelirovanie*, 2006, no. 4, pp. 204-207.

17. Gasnikov A.V., Klenov S.L., Nurminskii E.A., Kholodov Ia.A., Shamrai N.B.; prilozheniia: Blank M.L., Gasnikova E.V., Zamiatin A.A., Malyshev V.A., Kolesnikov A.V., Raigorodskii A.M. Vvedenie v matematicheskoe modelirovanie transportnykh potokov [Introduction to mathematical modeling of traffic flows]. Moscow: Moskovskii fiziko-tekhnicheskii institut, 2010, 360 p.

18. Poliarin Iu.A. O novykh trebovaniiakh k razmeshcheniiu i krepleniiu gruzov v vagonakh i universal'nykh konteinerakh [On new requirements for the placement and securing of goods in wagons and universal containers]. *Sklad i tekhnika*, 12 июля 2017, no. 11/2004, available at: https://sitmag.ru/article/10625-o-novyh-trebovaniyah-k-razmeshcheniyu-i-krepleniyu-gruzov-v-vagonah-i-universalnyh-konteynerah (accessed 18 October 2022).

19. Aiupov V.V. Matematicheskoe modelirovanie tekhnicheskikh sistem [Mathematical modeling of technical systems]. Perm': IPTs “Prokrost'”, 2017, 242 p.

20. Zvonarev S.V. Osnovy matematicheskogo modelirovaniia [Fundamentals of mathematical modeling]. Ekaterinburg: Ural'skii universitet, 2019, 112 p.

21. Khakimzianov G.S., Chubarov L.B., Voronina P.V. Matematicheskoe modelirovanie [Mathematical modeling]. Novosibirsk: RITs Novosibirskogo gosudarstvennogo universiteta, 2014, 263 p.

22. Mansfield E.R., Helms B.P. Detecting multicollinearity. *The American Statistician*, 1982, vol. 36, no. 3a, pp. 158-160.

23. Samarskii A.A., Mikhailov A.P. Matematicheskoe modelirovanie: Idei. Metody. Primery [Mathematical modeling: Ideas. Methods. Examples]. Moscow: Fizmatlit, 2001, 320 p.

24. Gorgees S.M. Statistical inference for least absolute deviation regression with autocorrelated errors. *Journal of Southwest Jiaotong University*, 2022, vol. 55 (2).

25. Wang X. Group selection via adjusted weighted least absolute deviation regression. *Journal of Computational and Applied Mathematics*, 2020, vol. 378, 112924 p.

Сведения об авторах

Базилевский Михаил Павлович (Иркутск, Россия) – кандидат технических наук, доцент кафедры «Математика» Иркутского государственного университета путей сообщения (664074, Иркутск, ул. Чернышевского, 15, e-mail: mik2178@yandex.ru).

Ойдопова Аяна Батоевна (Иркутск, Россия) – аспирантка по направлению 1.2.2 «Математическое моделирование, численные методы и комплексы программ» Иркутского государственного университета путей сообщения (664074, Иркутск, ул. Чернышевского, 15, e-mail: aoydopova11@mail.ru).

About the authors

Mikhail P. Bazilevskiy (Irkutsk, Russian Federation) – Ph. D. in Technical Sciences, Associate Professor Department of Mathematics Irkutsk State Transport University (664074, Irkutsk, 15, Chernyshevsky str., e-mail: mik2178@yandex.ru).

Ayana B. Oydopova (Irkutsk, Russian Federation) – Graduate Student in the direction of 1.2.2 «Mathematical modeling, numerical methods and software packages» of Irkutsk State Transport University (664074, Irkutsk, 15, Chernyshevsky str., e-mail: aoydopova11@mail.ru).

Поступила: 06.11.2022. Одобрена: 15.02.2023. Принята к публикации: 01.04.2023.

Финансирование. Исследование не имело спонсорской поддержки.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов по отношению к статье.

Вклад авторов. Все авторы сделали эквивалентный вклад в подготовку публикации.

Просьба ссылаться на эту статью в русскоязычных источниках следующим образом:

Базилевский, М.П. Оценивание модульных линейных регрессионных моделей с помощью метода наименьших модулей / М.П. Базилевский, А.Б. Ойдопова // Вестник Пермского национального исследовательского политехнического университета. Электротехника, информационные технологии, системы управления. – 2023. – № 45. – С. 130–146. DOI: 10.15593/2224-9397/2023.1.06

Please cite this article in English as:

Bazilevskiy M.P., Oydopova A.B. Estimation of modular linear regression models using the least absolute deviations. *Perm National Research Polytechnic University Bulletin. Electrotechnics, information technologies, control systems*, 2023, no. 45, pp. 130-146. DOI: 10.15593/2224-9397/2022.3.06