ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2025 Управление, вычислительная техника и информатика № 73
Tomsk State University Journal of Control and Computer Science

ОБРАБОТКА ИНФОРМАЦИИ

DATA PROCESSING

Научная статья
УДК 519.862.6
doi: 10.17223/19988605/73/6

О связи между многослойной модульной регрессионной моделью и регрессией в виде производственной функции Леонтьева

Михаил Павлович Базилевский

Иркутский государственный университет путей сообщения, Иркутск, Россия, mik2178@yandex.ru

Аннотация. Предложена модель многослойной модульной регрессии, отличающаяся от известной спецификации тем, что в ней выходное значение с каждого слоя входит не только в модульную, но и в линейную часть последующего слоя. Задача оценки неизвестных параметров предложенной модели с помощью метода наименьших модулей сведена к задаче частично булевого линейного программирования. Доказано, что для любого порядка переменных качество аппроксимации предложенной модели всегда не хуже, чем качество линейной регрессии и производственной функции Леонтьева. Показано, что любая вложенная кусочно-линейная регрессия является разновидностью многослойной модульной регрессии. Проведены вычислительные эксперименты, подтверждающие корректность доказанной теоремы.

Ключевые слова: регрессионный анализ; многослойная модульная регрессия; производственная функция Леонтьева; метод наименьших модулей; задача частично-булевого линейного программирования.

***Для цитирования*:** Базилевский М.П. О связи между многослойной модульной регрессионной моделью и регрессией в виде производственной функции Леонтьева // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 73. С. 46–57. doi: 10.17223/19988605/73/6

Original article
doi: 10.17223/19988605/73/6

On the relationship between the multilayer modular regression model and regression in the form of Leontief's production function

Mikhail P. Bazilevskiy

Irkutsk State Transport University, Irkutsk, Russian Federation, mik2178@yandex.ru

Abstract. The article proposes a model of multilayer modular regression, which differs from the known specification in that the output value from each layer is included not only in the modular part, but also in the linear part of the subsequent layer. The problem of estimating unknown parameters of the proposed model using the least absolute deviations method is reduced to the problem of mixed 0-1 integer linear programming. It has been proved that, for any order of variables, the approximation quality of the proposed model is always better than the quality of linear regression or the Leontief production function. Any nested piecewise linear regression can be considered as a type of multilayer modular regression. Computational experiments have been carried out to confirm the correctness of this theorem.

Keywords: regression analysis; multilayer modular regression; Leontief production function; least absolute deviations method; mixed 0-1 integer linear programming problem.

***For citation*:** Bazilevskiy, M.P. (2025) On the relationship between the multilayer modular regression model and regression in the form of Leontief's production function. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitelnaja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 73. pp. 46–57. doi: 10.17223/19988605/73/6

Введение

Решение многих научных проблем в настоящее время сопряжено с установлением математической формы связи между выходной переменной и одной или несколькими входными переменными по имеющимся в наличии статистическим данным. Для построения таких зависимостей исследователи часто применяют инструменты регрессионного анализа [1, 2]. Так, например, в [3] с помощью регрессионного анализа установлено влияние на среднемесячный уровень воды озера Ловозеро трех предикторов – среднемесячных расходов воды притока, общей облачности и облачности нижнего яруса. В [4] построена зависимость валовой продукции сельского хозяйства Якутии от поголовья крупного рогатого скота и лошадей, площади пастбищ и сенокосов, количества сельскохозяйственных организаций и численности сельского населения. В [5] получена регрессионная модель удельных энергозатрат на тягу поездов от шести факторов – объемов перевозочной работы, эксплуатационной длины электрифицированных линий, средней массы поезда, среднесуточной производительности локомотива, средней участковой скорости, доли грузовой работы.

Объединяет работы [3–5] то, что в них для проведения регрессионного анализа выбрана модель множественной линейной регрессии, параметры которой довольно просто оценить с помощью метода наименьших квадратов (МНК). Известно, что МНК можно применять и при нелинейных преобразованиях входящих в регрессию переменных, например с помощью элементарных функций. Так, в [6, 7] с помощью МНК оцениваются параметры полиномиальной регрессии, а в [8, 9] – параметры степенной регрессии, также известной как производственная функция (ПФ) Кобба–Дугласа.

Представленные в работах [6–9] нелинейные регрессионные зависимости хорошо изучены и применяются в исследованиях не один десяток лет. Помимо МНК оценка неизвестных параметров регрессионных моделей может осуществляться методом наименьших модулей (МНМ), реализация которого возможна с помощью описанного в [10] алгоритма «итеративного МНК» (метод вариационно-взвешенных квадратичных приближений). В большинстве же исследований МНМ реализуется с использованием аппарата линейного программирования, что стало возможным благодаря работе [11]. А в результате развития аппаратных и программных средств решения задач частично целочисленного линейного программирования [12] на сегодняшний день разработаны и другие нелинейные спецификации регрессионных моделей, с помощью которых можно строить довольно точные математические зависимости с интересными свойствами.

К первому классу таких нелинейных спецификаций относятся модели, фундаментом для которых послужила ПФ Леонтьева [13], содержащая мультиарную операцию min. В монографии [14] задача МНМ-оценки таких регрессий сведена к задаче частично булевого линейного программирования (ЧБЛП). Позднее эти модели за счет использования мультиарных операций max и вложенности функций были обобщены и получили название «вложенные кусочно-линейные регрессии» (см., напр.: [15]). Ко второму классу нелинейных спецификаций относятся модели, содержащие операцию «модуль». Впервые задача их МНМ-оценки была сведена к задаче ЧБЛП в работе [16]. Позднее появились многослойные модульные регрессии [17], которые также называют «глубокие модульные регрессии», и «широкие модульные регрессии», представленные в статье [18]. Связь между вложенными кусочно-линейными и многослойными модульными регрессиями до сих пор не исследовалась.

Известно, что для любых чисел a и b бинарные операции min и max связаны с операцией модуль следующими соотношениями:

$$\min\{a,b\} = \frac{a+b-|a-b|}{2}, \tag{1}$$

$$\max\{a,b\} = \frac{a+b+|a-b|}{2}. \tag{2}$$

Выражения (1) и (2) позволяют предположить, что класс многослойных модульных регрессий включает в себя класс вложенных кусочно-линейных регрессий. Цель данной статьи состоит в том, чтобы доказать выдвинутое предположение на примере ПФ Леонтьева.

1. Исследование связи между ПФ Леонтьева и многослойной модульной регрессией

Пусть имеется выборка объема n, содержащая значения y_i, $i=\overline{1,n}$, выходной (объясняемой) переменной y и значения x_{ij}, $i=\overline{1,n}$, $j=\overline{1,l}$, l входных (объясняющих) переменных x_1, x_2, …, x_l. Рассмотрим модель множественной линейной регрессии с неизвестными параметрами α_0, α_1, …, α_l:

$$y_i=\alpha_0+\sum_{j=1}^{l}\alpha_j x_{ij}+\varepsilon_i, \quad i=\overline{1,n}, \tag{3}$$

где ε_i, $i=\overline{1,n}$, – ошибки регрессии.

Будем считать, что неизвестные параметры линейной регрессии (3) оцениваются с помощью МНМ, который состоит в минимизации суммы модулей ошибок, что означает решение задачи $\sum_{i=1}^{n}|\varepsilon_i|\to\min$. Решение этой задачи сводится к решению задачи линейного программирования [11]. Предположим, что с помощью МНМ найдены оценки неизвестных параметров модели (3). Обозначим сумму остатков этой модели как $\mathrm{SAE}_{\text{лин}}=\sum_{i=1}^{n}|e_i|$, где e_i – i-й остаток, равный разнице между фактическим и расчетным значением отклика y в i-м наблюдении.

Рассмотрим ПФ Леонтьева со свободным членом:

$$y_i=\alpha_0+\min\{\alpha_1 x_{i1},\alpha_2 x_{i2},...,\alpha_l x_{il}\}+\varepsilon_i, \quad i=\overline{1,n}, \tag{4}$$

где min – l-арная операция, возвращающая минимум из перечисленных в фигурных скобках {} l значений. Пусть модель (4) оценена с помощью МНМ в результате решения задачи ЧБЛП [14]. Обозначим сумму остатков этой модели как $\mathrm{SAE}_{\text{Леон}}$.

Теорема. Пусть $l\geq 3$, а все индексы входных переменных произвольно распределены по l элементам s_1, s_2, …, s_l вектора S. Предположим, что с помощью МНМ оценены неизвестные параметры α_0, α_1, …, α_n, β_1, …, β_n многослойной модульной регрессии

$$y_i=\alpha_0+z_{i,l-2}+\alpha_l x_{i,s_l}-\left|z_{i,l-2}+\beta_l x_{i,s_l}\right|+\varepsilon_i, \quad i=\overline{1,n}, \tag{5}$$

$$z_{i,k}=z_{i,k-1}+\alpha_{k+1}x_{i,s_{k+1}}-\left|z_{i,k-1}+\beta_{k+1}x_{i,s_{k+1}}\right|, \quad i=\overline{1,n},\ k=\overline{2,l-2}, \tag{6}$$

$$z_{i1}=\alpha_1 x_{i,s_1}+\alpha_2 x_{i,s_2}-\left|\beta_1 x_{i,s_1}+\beta_2 x_{i,s_2}\right|, \quad i=\overline{1,n}, \tag{7}$$

где $z_{i,k}$ – выходное значение на k-м слое в i-м наблюдении. Обозначим сумму остатков этой регрессии как $\mathrm{SAE}_{\text{мод}}$. Тогда будут справедливы неравенства

$$\mathrm{SAE}_{\text{мод}}\leq\mathrm{SAE}_{\text{лин}}, \tag{8}$$

$$\mathrm{SAE}_{\text{мод}}\leq\mathrm{SAE}_{\text{Леон}}. \tag{9}$$

Доказательство. Многослойная модульная регрессия (5)–(7) содержит l – 1 слоев: 1-й слой отражает равенство (7), промежуточные слои со 2-го до (l – 2)-го – равенства (6), выходной (l – 1)-й слой – равенство (5). Последовательно подставляя выходные значения z с нижних слоев в выражения для верхних слоёв, получим структурную спецификацию

$$y_i=\alpha_0+\sum_{j=1}^{l}\alpha_j x_{i,s_j}-F(\alpha,\beta,x)+\varepsilon_i, \quad i=\overline{1,n}, \tag{10}$$

где $F(\alpha,\beta,x)$ – функция нескольких переменных с неизвестными параметрами, содержащая модульные преобразования. Очевидно, что модель (10) есть расширенная версия линейной регрессии (3), откуда следует справедливость неравенства (8) для любого распределения индексов входных переменных по элементам вектора S.

Пусть $\beta_1=\alpha_1$, $\beta_j=-\alpha_j$, $j=\overline{2,l}$. Многослойная модульная регрессия (5)–(7) будет иметь вид:

$$y_i=\alpha_0+z_{i,l-2}+\alpha_l x_{i,s_l}-\left|z_{i,l-2}-\alpha_l x_{i,s_l}\right|+\varepsilon_i, \quad i=\overline{1,n}, \tag{11}$$

$$z_{i,k} = z_{i,k-1} + \alpha_{k+1} x_{i,s_{k+1}} - \left| z_{i,k-1} - \alpha_{k+1} x_{i,s_{k+1}} \right|, \quad i = \overline{1,n}, \; k = \overline{2,l-2}, \tag{12}$$

$$z_{i1} = \alpha_1 x_{i,s_1} + \alpha_2 x_{i,s_2} - \left| \alpha_1 x_{i,s_1} - \alpha_2 x_{i,s_2} \right|, \quad i = \overline{1,n}. \tag{13}$$

Модель (11)–(13) есть ограниченная версия модели (5)–(7), поэтому сумма ее остатков не меньше величины $\mathrm{SAE}_{\text{мод}}$. Из (1) следует, что $a + b - |a-b| = \min\{2a, 2b\}$. Учитывая это равенство, перепишем регрессию (11)–(13) в виде:

$$y_i = \alpha_0 + \min\left\{2 z_{i,l-2}, 2\alpha_l x_{i,s_l}\right\} + \varepsilon_i, \quad i = \overline{1,n}, \tag{14}$$

$$z_{i,k} = \min\left\{2 z_{i,k-1}, 2\alpha_{k+1} x_{i,s_{k+1}}\right\}, \quad i = \overline{1,n}, \quad k = \overline{2,l-2}, \tag{15}$$

$$z_{i1} = \min\left\{2\alpha_1 x_{i,s_1}, 2\alpha_2 x_{i,s_2}\right\}, \quad i = \overline{1,n}. \tag{16}$$

Последовательно подставляя для модели (14)–(16) выходные значения z с нижних слоев в выражения для верхних слоев, получим структурную спецификацию

$$y_i = \alpha_0 + \min\left\{2\min\left\{2\min\left\{2\alpha_1 x_{i,s_1}, 2\alpha_2 x_{i,s_2}\right\}, 2\alpha_3 x_{i,s_3}\right\}, \ldots\right\} + \varepsilon_i, \quad i = \overline{1,n}. \tag{17}$$

Многослойная конструкция (17), как показано в [19], может быть записана в виде:

$$y_i = \alpha_0 + \min\left\{2^{l-1}\alpha_1 x_{i,s_1}, 2^{l-1}\alpha_2 x_{i,s_2}, 2^{l-2}\alpha_3 x_{i,s_3}, \ldots, 2^2\alpha_{l-1} x_{i,s_{l-1}}, 2\alpha_l x_{i,s_l}\right\} + \varepsilon_i, \quad i = \overline{1,n}. \tag{18}$$

Сделав замены $\gamma_1 = 2^{l-1}\alpha_1$, $\gamma_j = 2^{l+1-j}\alpha_j$, $j = \overline{2,l}$, нетрудно видеть, что регрессия (18) есть ПФ Леонтьева (4).

Модель (18) эквивалентна многослойной модульной регрессии (11)–(13), сумма остатков которой не меньше величины $\mathrm{SAE}_{\text{мод}}$. Тогда для модели (5)–(7) справедливо неравенство (9) для любого распределения индексов входных переменных по элементам вектора *S*. ***Теорема доказана***.

Сделаем некоторые замечания касаемо доказанной теоремы.

Во-первых, доказанная теорема справедлива и при $l = 2$. Действительно, при $l = 2$ имеем однослойную конструкцию

$$y_i = \alpha_0 + \alpha_1 x_{i,s_1} + \alpha_2 x_{i,s_2} - \left|\beta_1 x_{i,s_1} + \beta_2 x_{i,s_2}\right| + \varepsilon_i, \quad i = \overline{1,n},$$

для которой по тем же причинам, что и при $l \ge 3$, справедливы неравенства (8) и (9).

Во-вторых, смысл доказанной теоремы состоит в том, что для любых статистических данных и для любого распределения индексов входных переменных по элементам вектора *S* качество аппроксимации оцененной с помощью МНМ многослойной модульной регрессии (5)–(7) всегда не хуже (на практике чаще лучше), чем качество линейной регрессии (3) либо ПФ Леонтьева (4).

В-третьих, если в многослойной модульной регрессии (5)–(7) заменить знаки «–», стоящие перед операциями модуль, на знаки «+», то аналогичным образом с использованием (2) можно доказать, что качество образованной модели всегда будет не хуже, чем качество линейной регрессии (3) либо регрессии [20] с *l*-арной операцией max вида:

$$y_i = \alpha_0 + \max\left\{\alpha_1 x_{i1}, \alpha_2 x_{i2}, \ldots, \alpha_l x_{il}\right\} + \varepsilon_i, \quad i = \overline{1,n}.$$

В-четвертых, при $l = 2$ величина $\mathrm{SAE}_{\text{мод}}$ многослойной модульной регрессии (5)–(7) не зависит от распределения индексов входных переменных по элементам вектора *S*, а при $l \ge 3$ – зависит. В любом случае справедливы неравенства (8) и (9).

Таким образом, из теоремы следует, что класс многослойных модульных регрессий включает в себя производственные функции Леонтьева.

Из доказательства теоремы также следует, что любую вложенную кусочно-линейную регрессию с помощью соотношений (1) и (2) всегда можно представить в виде многослойной модульной регрессии. Это можно сделать по следующему алгоритму.

1. Записать вложенную кусочно-линейную регрессию в виде многослойной конструкции только с бинарными операциями min и max так, как это показано в [19].

2. С использованием соотношений (1) и (2) перейти от бинарных операций min и max к спецификации с модулями.

3. В полученной конструкции обозначить все неизвестные параметры по-разному.

Например, пусть имеется вложенная кусочно-линейная регрессия вида:

$$y_i = \alpha_0 + \min\left\{\max\left\{\alpha_1 x_{i1}, \alpha_2 x_{i2}\right\}, \alpha_3 x_{i3}, \alpha_4 x_{i4}\right\} + \varepsilon_i, \quad i = \overline{1,n}. \tag{19}$$

В (19) операция max – бинарная, min – тернарная. Запишем модель (19) в виде многослойной конструкции только с бинарными операциями:

$$y_i = \alpha_0 + \min\left\{\min\left\{\max\left\{\alpha_1 x_{i1}, \alpha_2 x_{i2}\right\}, \alpha_3 x_{i3}\right\}, \alpha_4 x_{i4}\right\} + \varepsilon_i, \quad i = \overline{1,n}.$$

С помощью (1) и (2) перейдем от полученной конструкции к многослойной спецификации с модулями:

$$y_i = \alpha_0 + \frac{1}{2}\left(z_{i2} + \alpha_4 x_{i4} - \left|z_{i2} - \alpha_4 x_{i4}\right|\right) + \varepsilon_i, \quad i = \overline{1,n}, \tag{20}$$

$$z_{i2} = \frac{1}{2}\left(z_{i1} + \alpha_3 x_{i3} - \left|z_{i1} - \alpha_3 x_{i3}\right|\right), \quad i = \overline{1,n}, \tag{21}$$

$$z_{i1} = \frac{1}{2}\left(\alpha_1 x_{i1} + \alpha_2 x_{i2} + \left|\alpha_1 x_{i1} - \alpha_2 x_{i2}\right|\right), \quad i = \overline{1,n}. \tag{22}$$

Оцененная с помощью МНМ вложенная кусочно-линейная регрессия (19) и многослойная модульная регрессия (20)–(22) будут иметь одинаковые значения критерия SAE. В (20)–(22) всего 4 неизвестных параметра, которые дублируются. Обозначив же их по-разному, получим в общей сложности 8 неизвестных параметров. Тогда полученная многослойная модульная регрессия при оценивании с помощью МНМ в любом случае будет не хуже по критерию SAE, чем (20)–(22), а, следовательно, и (19).

Таким образом, любой вложенной кусочно-линейной регрессии по указанному алгоритму всегда можно поставить в соответствие многослойную модульную регрессию, которая будет не хуже, а за счет большего числа неизвестных параметров лучше первой. Тем самым многослойная модульная регрессия – гораздо более гибкий инструмент математического моделирования, чем вложенная кусочно-линейная регрессия.

2. Оценивание с помощью МНМ параметров многослойной модульной регрессии (5)–(7)

Многослойная модульная регрессия впервые была предложена в [17], и там же задача нахождения оценок ее параметров с помощью МНМ сведена к задаче ЧБЛП. Предлагаемая в данной статье модель (5)–(7) отличается от описанной ранее в работе [17] регрессии тем, что во второй на текущем слое выходные переменные z с предыдущего слоя используются только под знаком модуля. В многослойной модульной регрессии (5)–(7) выходные переменные с предыдущего слоя используются как под знаком модуля, так и в линейной части текущего слоя.

Используя некоторые математические приемы из [17], получим следующую задачу ЧБЛП для нахождения неизвестных параметров модели (5)–(7):

$$\sum_{i=1}^{n}\left(g_i + h_i\right) \to \min, \tag{23}$$

$$y_i = \alpha_0 + z_{i,l-2} + \alpha_l x_{i,s_l} - \left(u_{i,l-1} + v_{i,l-1}\right) + g_i - h_i, \quad i = \overline{1,n}, \tag{24}$$

$$u_{i,l-1} - v_{i,l-1} = z_{i,l-2} + \beta_l x_{i,s_l}, \quad i = \overline{1,n}, \tag{25}$$

$$z_{i,k} = z_{i,k-1} + \alpha_{k+1} x_{i,s_{k+1}} - \left(u_{i,k} + v_{i,k}\right), \quad i = \overline{1,n}, \quad k = \overline{2,l-2}, \tag{26}$$

$$u_{i,k} - v_{i,k} = z_{i,k-1} + \beta_{k+1} x_{i,s_{k+1}}, \quad i = \overline{1,n}, \quad k = \overline{2,l-2}, \tag{27}$$

$$z_{i1} = \alpha_1 x_{i,s_1} + \alpha_2 x_{i,s_2} - \left(u_{i1} + v_{i1}\right), \quad i = \overline{1,n}, \tag{28}$$

$$u_{i1} - v_{i1} = \beta_1 x_{i,s_1} + \beta_2 x_{i,s_2}, \quad i = \overline{1,n}, \tag{29}$$

$$u_{i,j} \le M \cdot \delta_{i,j}, \quad i = \overline{1,n}, \quad j = \overline{1,l-1}, \tag{30}$$

$$v_{i,j} \le M \cdot \left(1 - \delta_{i,j}\right), \quad i = \overline{1,n}, \quad j = \overline{1,l-1}, \tag{31}$$

$$\delta_{i,j} \in \{0,1\}, \quad i = \overline{1,n}, \quad j = \overline{1,l-1}, \tag{32}$$

$$u_{i,j} \ge 0, \; v_{i,j} \ge 0, \quad i = \overline{1,n}, \quad j = \overline{1,l-1}, \tag{33}$$

$$g_i \ge 0, \quad h_i \ge 0, \quad i = \overline{1,n}, \tag{34}$$

где M – заданное достаточно большое положительное число; если в i-м наблюдении на j-м слое под знаком модуля содержится положительное число, то оно присваивается переменной $u_{i,j}$, а $v_{i,j} = 0$; если в i-м наблюдении на j-м слое под знаком модуля содержится отрицательное число, то оно присваивается переменной $-v_{i,j}$, а $u_{i,j} = 0$; если в i-м наблюдении на верхнем слое ошибка регрессии ε_i есть положительное число, то оно присваивается переменной g_i, а $h_i = 0$; если в i-м наблюдении на верхнем слое ошибка регрессии ε_i есть отрицательное число, то оно присваивается переменной $-h_i$, а $g_i = 0$; булевы переменные $\delta_{i,j}$, $i = \overline{1,n}$, $j = \overline{1,l-1}$, удовлетворяют правилу

$$\delta_{i,j} = \begin{cases} 1, & \text{если в } i\text{-м наблюдении на } j\text{-м слое под знаком модуля положительное число,} \\ 0, & \text{в противном случае.} \end{cases}$$

Ограничения (24), (25) в сформулированной задаче соответствуют верхнему слою многослойной модульной регрессии, (26), (27) – внутренним слоям, (28), (29) – нижнему слою. Решение задачи ЧБЛП (23)–(34) дает оптимальные МНМ-оценки параметров многослойной модульной регрессии (5)–(7) при $l \ge 3$. При $l = 2$ модель будет состоять только из одного слоя, поэтому ее оценивание будет сводиться к решению более простой задачи ЧБЛП с целевой функцией (23), линейными ограничениями (29)–(34) и

$$y_i = \alpha_0 + \alpha_1 x_{i,s_1} + \alpha_2 x_{i,s_2} - \left(u_{i1} + v_{i1}\right) + g_i - h_i, \quad i = \overline{1,n}. \tag{35}$$

3. Вычислительные эксперименты

Для подтверждения корректности полученных в предыдущих разделах математических выводов на персональном компьютере с процессором AMD Ryzen 3 4300U и 16 Гб оперативной памяти проводились вычислительные эксперименты. Чем больше слоев в многослойной модульной регрессии, тем выше вычислительная сложность задачи ЧБЛП, поскольку в ней возрастает число булевых переменных. Поэтому для решения таких задач нужен эффективный решатель (солвер). В этой связи было решено использовать хорошо зарекомендовавший себя и представленный в открытом доступе решатель оптимизационных задач LPSolve [21]. Вместе с тем было решено использовать разработанный в Китае коммерческий продукт Cardinal Optimizer (COPT) [22]. Как заявляют разработчики, этот продукт в некоторых тестах превосходит другие коммерческие аналоги (CPLEX, Gurobi), которые сейчас недоступны в России, а также open-source солверы. Было интересно сравнить скорость работы LPSolve и COPT на примере построения многослойных модульных регрессий. Для управления солвером COPT был использован интерфейс Python.

Для проведения экспериментов использовались две выборки данных.

Первая выборка [23. С. 284] объема $n = 13$ содержит данные о результатах химических экспериментов (данные Хальда). Она включает в себя информацию по следующим переменным: y_1 – выделившееся тепло в калориях на грамм цемента; x_1 – количество трикальций-алюмината (% от веса клинкера); x_2 – количество трикальций-силиката (%); x_3 – количество тетракальций-алюминиум-феррита (%); x_4 – количество дикальций-силиката (%).

Вторая выборка объема $n = 23$ включает данные о новых единицах жилья в США и влияющих на них факторах. Эти данные встроены в эконометрический пакет Gretl (файл data4-3.gdt). В них содержится информация по следующим переменным: y_2 – общее количество новых единиц жилья, введенных в эксплуатацию (тыс.); x_1^* – население США (млн чел); x_2^* – валовой национальный продукт (млрд долл.); x_3^* – уровень безработицы (%); x_4^* – доходность ипотечных кредитов на новое жилье (%).

При проведении экспериментов менялся состав входящих в модель объясняющих переменных. Для этого были назначены следующие подмножества: {1, 2}, {1, 3}, {1, 4}, {2, 3}, {2, 4}, {3, 4}, {1, 2, 3}, {1, 2, 4}, {1, 3, 4}, {2, 3, 4} и {1, 2, 3, 4}. Например, подмножество {2, 3} для первой выборки означает, что объясняющими переменными выступают x_2 и x_3. Для каждой выборки и каждого состава переменных решались следующие задачи:

1) в пакете Gretl оценивались параметры линейной регрессии (3) и фиксировалось значение $SAE_{лин}$;

2) с помощью LPSolve оценивались параметры ПФ Леонтьева (4) и фиксировалось значение $SAE_{Леон}$;

3) с помощью LPSolve и COPT оценивались параметры многослойной модульной регрессии (5)–(7) с прямым порядком следования объясняющих переменных, фиксировались значение $SAE_{мод}$ и время решения задачи ЧБЛП (23)–(34) при $l \ge 3$ либо (23), (29)–(35) при $l = 2$;

4) с помощью LPSolve и COPT оценивались параметры многослойной модульной регрессии (5)–(7) с обратным порядком следования объясняющих переменных, фиксировались значение $SAE_{мод}$ и время решения задачи ЧБЛП (23)–(34) при $l \ge 3$ либо (23), (29)–(35) при $l = 2$.

Прямой порядок следования переменных означает, что если, например, выбран состав {1, 2, 3}, то первый элемент вектора *S* равен 1, второй – 2, третий – 3. При обратном порядке первый элемент вектора *S* равен 3, второй – 2, третий – 1.

Используемые солверы LPSolve и COPT содержат большое количество различных настраиваемых параметров. При решении задач ЧБЛП важно правильно установить такой параметр, как точность целого числа, которую обозначим IntTol. Если установить значение IntTol слишком большим, например 0,1, то значение 0,91 солвер будет идентифицировать как число «1», а значение 0,07 – как число «0». Такая настройка, возможно, увеличит скорость решения задачи, но результат может быть ошибочным. В LpSolve параметр IntTol по умолчанию равен $1e^{-7}$, а в COPT – $1\ e^{-6}$. Для организации равных условий работы пакетов параметру IntTol в COPT было присвоено значение $1e^{-7}$.

Хотелось бы обратить внимание на ограничения (30) и (31). Если точность целого числа IntTol в солверах выбрана $1e^{-7}$, то при очень большом числе *M* эти ограничения не будут гарантировать выполнение условий

$$u_{i,j} v_{i,j} = 0, \quad i = \overline{1,n}, \quad j = \overline{1,l-1}, \tag{36}$$

что может привести к ошибочным результатам. Поэтому условия (36) необходимо проверять после решения каждой задачи ЧБЛП. Если они не выполняются, то требуется перерешать задачу, уменьшив параметр IntTol.

Для данных Хальда параметр *M* был выбран 1 000, а для данных о жилье в США *M* = 10 000. Вместо переменной y_2 была использована переменная $y_2^* = y_2/M$. Ограничение по времени в солверах на решение одной задачи ЧБЛП составило 1 800 с (30 мин).

Результаты проведенных экспериментов представлены в таблице.

Условия (36) при IntTol = $1e^{-7}$ не выполнились только при обработке данных о жилье для составов {1, 2, 4} и {2, 3, 4}, поэтому в строках № 19 и 21 таблицы приведены результаты, полученные в LPSolve и COPT при IntTol = $1e^{-8}$.

В таблице символ «*» указывает на субоптимальные решения, найденные в установленный лимит времени 30 мин.

По таблице в результате обработки данных Хальда можно сделать следующие выводы.

1. Как линейная регрессия может превосходить по критерию SAE ПФ Леонтьева (например, строка № 1), так и ПФ может превосходить линейную регрессию (например, строка № 2).

2. При построении многослойной модульной регрессии абсолютно во всех экспериментах для любого порядка переменных оказались справедливы неравенства (8) и (9), что подтверждает корректность доказанной теоремы. Причем, только в 3 случаях из 22 значение $SAE_{мод}$ оказалось равно $SAE_{лин}$ или $SAE_{Леон}$, т.е. в большинстве случаев многослойная модульная регрессия оказалась лучше и линейной регрессии, и ПФ Леонтьева.

3. При изменении порядка следования переменных качество однослойных модульных регрессий по критерию SAE сохранилось, а качество двухслойных и трехслойных модульных регрессий изменилось.

4. В LPSolve для прямого и обратного порядков переменных среднее время построения однослойной модульной регрессии составило 0,06 с, двухслойной – 2,26 с, трехслойной – 25,51 с. Рост времени обусловлен увеличением в задаче ЧБЛП количества булевых переменных при увеличении числа слоев.

5. В COPT были получены точно такие же модели, что и в LPSolve. В COPT для прямого и обратного порядков переменных среднее время построения однослойной модульной регрессии составило 0,33 с, двухслойной – 0,59 с, трехслойной – 1,16 с. Получилось, что однослойные модульные регрессии в LPSolve и COPT оцениваются практически мгновенно, но COPT проигрывает по времени пакету LPSolve. Однако такой результат, вероятнее всего, можно объяснить тем, что при управлении COPT в Python на чтение оптимизационной задачи из программы уходит чуть больше времени, чем в LpSolve. Для двухслойных модульных регрессий время решения задач ЧБЛП в COPT оказалось в 3,84 раза, а для трехслойных – в 22,08 раза меньше, чем в LPSolve.

Результаты экспериментов

№	Состав	$SAE_{лин}$	$SAE_{Леон}$	Прямой порядок переменных				Обратный порядок переменных			
				LPSolve		COPT		LPSolve		COPT	
				$SAE_{мод}$	*t*, с	$SAE_{мод}$	*t*, с	$SAE_{мод}$	*t*, с	$SAE_{мод}$	*t*, с
Данные Хальда											
1	1, 2	23,126	56,167	23,126	0,048	23,126	0,33	23,126	0,047	23,126	0,36
2	1, 3	107,1	91,6	59,195	0,055	59,195	0,28	59,195	0,036	59,195	0,27
3	1, 4	23,339	107,3	21,972	0,064	21,972	0,33	21,972	0,063	21,972	0,35
4	2, 3	55,057	83,125	54,662	0,084	54,662	0,34	54,662	0,073	54,662	0,41
5	2, 4	82,957	72,421	70,347	0,070	70,347	0,35	70,347	0,085	70,347	0,35
6	3, 4	35,313	165	30,940	0,053	30,940	0,28	30,940	0,043	30,940	0,29
7	1, 2, 3	19,056	56,167	19,049	1,087	19,049	0,50	13,66	0,268	13,66	0,45
8	1, 2, 4	19,663	52	17,017	3,878	17,017	0,59	16,389	1,969	16,389	0,93
9	1, 3, 4	20,203	90,6	20,203	7,424	20,203	0,62	17,882	0,858	17,882	0,58
10	2, 3, 4	22,536	72,421	20,609	1,431	20,609	0,45	20,068	1,130	20,068	0,58
11	1, 2, 3, 4	18,834	52	12,459	40,144	12,459	1,01	13,809	10,876	13,809	1,30
Встроенные в Gretl данные data4-3											
12	1, 2	5,655	5,858	5,655	17,049	5,655	3,01	5,655	16,716	5,655	3,57
13	1, 3	5,896	5,773	5,000	4,503	5,000	0,54	5,000	4,716	5,000	0,49
14	1, 4	4,356	5,902	4,241	38,665	4,241	0,95	4,241	38,982	4,241	0,91
15	2, 3	5,853	5,797	5,164	4,672	5,164	0,60	5,164	4,922	5,164	0,75
16	2, 4	4,514	5,858	4,172	5,480	4,172	0,66	4,172	5,540	4,172	0,52
17	3, 4	6,055	6,053	5,004	5,077	5,004	0,67	5,004	4,921	5,004	0,69
18	1, 2, 3	5,199	5,773	4,764*	1800	4,145	4,77	4,758*	1800	4,308	40,22
19	1, 2, 4	4,346	5,855	3,818	1431,4	3,818	9,69	4,204*	1800	4,013	934,41
20	1, 3, 4	4,323	5,767	3,925*	1800	3,177	5,98	3,540	1785	3,540	5,92
21	2, 3, 4	4,176	5,784	3,343	813,87	3,343	32,03	4,056*	1800	3,713	3,75
22	1, 2, 3, 4	4,074	5,767	4,813*	1800	2,840	261,6	4,139*	1800	3,268	32,23

По таблице в результате обработки данных о жилье можно сделать следующие выводы.

1. Линейная регрессия и ПФ Леонтьева в зависимости от выбранных переменных могут превосходить друг друга по критерию SAE.

2. Не принимая во внимание субоптимальные решения, при построении многослойной модульной регрессии во всех экспериментах для любого порядка переменных определено, что оказались справедливы неравенства (8) и (9), и это вновь подтверждает корректность доказанной теоремы. И снова в большинстве случаев (в 20 из 22) многослойная модульная регрессия оказалась лучше как линейной регрессии, так и ПФ Леонтьева.

3. Изменение порядка следования переменных меняет значение критерия SAE для двухслойных и трехслойных модульных регрессий, но не меняет для однослойных.

4. Вторая выборка данных содержит больше наблюдений, чем первая, поэтому время решения задач ЧБЛП увеличилось. Причем LPSolve в установленный лимит времени 30 мин не справился с 7 задачами из 22. Так, в LPSolve для прямого и обратного порядков переменных среднее время построения однослойной модульной регрессии составило 12,6 с. Для двухслойной и трехслойной модульной регрессии – 1 628,78 и 1 800 с соответственно. Причем при снятии установленного лимита для последних двух спецификаций время будет еще больше.

5. Солвер COPT, в отличие от LPSolve, справился абсолютно со всеми задачами. Для прямого и обратного порядков переменных среднее время построения однослойной модульной регрессии составило 1,11 с, двухслойной – 129,59 с, трехслойной – 146,91 с. Среднее время оценки двухслойных регрессий оказалось высоким из-за эксперимента для обратного порядка переменных из множества {1, 2, 4}. Без учета этого эксперимента среднее время составит всего 14,62 с. Таким образом, для однослойных модульных регрессий время решения задач ЧБЛП в COPT оказалось в 11,32 раза, для двухслойных – как минимум в 12,57 раза, а для трехслойных – как минимум в 12,25 раза меньше, чем в LPSolve.

Рассмотрим внешний вид моделей с полным составом переменных, построенных по данным о рынке жилья. ПФ имеет вид:

$$\tilde{y}_2^* = 1{,}00054 + \min\left\{0{,}003032x_1^*, 0{,}000306x_2^*, 0{,}115717x_3^*, 0{,}097422x_4^*\right\}. \quad (37)$$

Как отмечено в строке № 22 таблицы, для регрессии (37) $SAE_{\text{Леон}} = 5{,}767$. При этом нулевыми оказались остатки в наблюдениях под номерами 3 и 21.

Трехслойная модульная регрессия для прямого порядка переменных имеет вид:

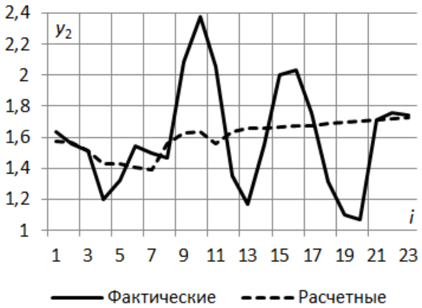
$$\tilde{y}_2^* = 3{,}46081 + 0{,}065492x_4^* + z_2 - \left|z_2 + 0{,}22326x_4^*\right|, \quad (38)$$

$$z_2 = z_1 - 0{,}029006x_3^* - \left|z_1 + 0{,}26546x_3^*\right|, \quad (39)$$

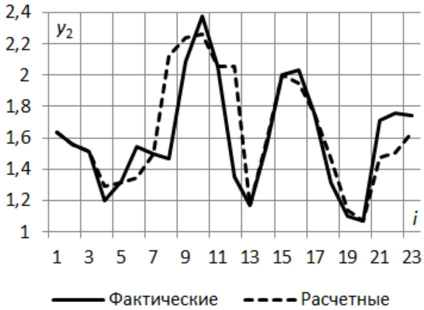
$$z_1 = -0{,}012955x_1^* + 0{,}000523x_2^* - \left|0{,}0041428x_1^* - 0{,}00038652x_2^*\right|. \quad (40)$$

Для регрессии (38)–(40) $SAE_{\text{мод}} = 2{,}840$. При этом нулевые остатки зафиксированы сразу в девяти наблюдениях – под номерами 1, 2, 3, 7, 11, 13, 15, 17 и 20.

Для наглядности на рис. 1, *a* изображены графики фактических и расчетных значений зависимой переменной y_2 для модели (37), а на рис. 1, *b* – для модели (38)–(40).



a



b

Рис. 1. Фактические и рассчитанные по моделям значения переменной y_2
Fig. 1. Actual and calculated values of the variable y_2 according to the models

По графикам на рис. 1 видно, что точность трехслойной модульной регрессии (38)–(40) существенно выше, чем точность ПФ Леонтьева (37).

Заключение

В статье предложена модель многослойной модульной регрессии (5)–(7), отличающаяся от разработанной в [17] модели тем, что выходное значение с каждого предыдущего слоя входит не только в модульную часть, но и в линейную часть последующего слоя. Доказано, что для любых статистических данных и для любого распределения индексов входных переменных по элементам вектора S качество аппроксимации оцененной с помощью МНМ многослойной модульной регрессии (5)–(7) всегда не хуже (на практике зачастую лучше), чем качество линейной регрессии (3) и ПФ Леонтьева (4). Показано, что класс многослойных модульных регрессий включает в себя класс вложенных кусочно-линейных регрессий. Разработан алгоритм трансформации любой вложенной кусочно-линейной регрессии в модель многослойной модульной регрессии, оценка которой с помощью МНМ гарантирует не ухудшение качества аппроксимации по критерию SAE. Задача нахождения точных МНМ-оценок неизвестных параметров предложенной многослойной модульной регрессии (5)–(7) сведена к задаче ЧБЛП. На двух разных выборках данных проведены вычислительные эксперименты, доказывающие корректность доказанной теоремы. В ходе экспериментов впервые проведен сравнительный анализ скорости солверов LPSolve и Cardinal Optimizer. Для первой выборки при оценке параметров однослойных модульных регрессий солверы решили задачи ЧБЛП практически мгновенно. При оценке двухслойных и трехслойных модульных регрессий Cardinal Optimizer оказался быстрее в 3,84 и 22,08 раза соответственно, чем LPSolve. Для второй выборки большего объема при оценке однослойных модульных регрессий Cardinal Optimizer оказался в 11,32 раза быстрее, чем LPSolve. При оценке двухслойных и трехслойных модульных регрессий в LPSolve оптимальное решение большинства задач ЧБЛП не было получено в установленный лимит времени 30 мин. При этом Cardinal Optimizer справился абсолютно со всеми задачами, оказавшись быстрее, чем LPSolve, как минимум в 12,57 и 12,25 раза для двухслойных и трехслойных модульных регрессий. Таким образом, выявлено существенное превосходство по скорости солвера Cardinal Optimizer перед LPSolve.

Список источников

1. Montgomery D.C., Peck E.A., Vining G.G. Introduction to linear regression analysis. New York : John Wiley and Sons, 2021. 704 p.
2. Baždarić K., Šverko D., Salarić I., Martinović A., Lucijanić M. The ABC of linear regression analysis: What every author and editor should know // European science editing. 2021. V. 47. doi: 10.3897/ese.2021.e63780
3. Батмазова А.А., Гайдукова Е.В., Дрегваль М.С. Особенности построения регрессионной модели для оценки и прогноза уровня озера Ловозеро // Международный научно-исследовательский журнал. 2024. № 7 (145). doi: 10.60797/IRJ.2024.145.67
4. Бястинова Л.М. Использование многофакторных регрессионных моделей при прогнозировании в сфере сельского хозяйства Якутии // Экономика и природопользование на Севере. 2023. № 4. С. 102–110. doi: 10.25587/2587-8778-2023-4-102-110
5. Каштанов А.Л., Комяков А.А., Никифоров М.М. Прогнозирование и верификация ключевых показателей энергетической эффективности железнодорожного транспорта // Вестник Уральского государственного университета путей сообщения. 2021. № 1. С. 46–54. doi: 10.20291/2079-0392-2021-1-46-54
6. Горшенин А.Ю., Васина Д.И. Сравнение используемых методов при прогнозировании выработки электроэнергии ветроэлектростанциями // Математические структуры и моделирование. 2023. № 3 (67). С. 36–42. doi:10.24147/2222-8772.2023.3.36-42
7. Брачунова У.В., Козловский В.Н., Шакурский М.В., Пантюхин О.В. Основные аспекты разработки математической модели и программного комплекса по оценке зарядного баланса легкового автомобиля // Известия Тульского государственного университета. Технические науки. 2023. № 5. С. 242–249. doi: 10.24412/2071-6168-2023-5-242-243
8. Гичиев Н.С. Анализ двухфакторной модели экономического роста на основе производственной функции Кобба–Дугласа // Региональные проблемы преобразования экономики. 2022. № 3 (137). С. 62–66. doi: 10.26726/1812-7096-2022-3-62-66
9. Ботнарюк М.В., Ксензова Н.Н. Производственная функция Кобба–Дугласа для оценки деятельности морского порта // Научные проблемы водного транспорта. 2022. № 74. С. 85–95. doi: 10.37890/jwt.vi74.348
10. Демиденко Е.З. Линейная и нелинейная регрессии. М. : Финансы и статистика, 1981. 303 с.
11. Wagner H.M. Linear programming techniques for regression analysis // Journal of the American Statistical Association. 1959. V. 54 (285). P. 206–212. doi: 10.1080/01621459.1959.10501506
12. Koch T., Berthold T., Pedersen J., Vanaret C. Progress in mathematical programming solvers from 2001 to 2020 // EURO Journal on Computational Optimization. 2022. V. 10. Art. 100031. doi: 10.1016/j.ejco.2022.100031
13. Клейнер Г.Б. Производственные функции: теория, методы, применение. М. : Финансы и статистика, 1986. 239 с.
14. Носков С.И. Технология моделирования объектов с нестабильным функционированием и неопределенностью в данных. Иркутск : Облинформпечать, 1996. 320 с.
15. Носков С.И. Оценивание параметров простой вложенной кусочно-линейной регрессии с линейной составляющей // Вестник Югорского государственного университета. 2024. Т. 20, № 1. С. 19–21. doi: 10.18822/byusu20240119-21
16. Базилевский М.П., Ойдопова А.Б. Оценивание модульных линейных регрессионных моделей с помощью метода наименьших модулей // Вестник Пермского национального исследовательского политехнического университета. Электротехника, информационные технологии, системы управления. 2023. № 45. С. 130–146. doi: 10.15593/2224-9397/2023.1.06

17. Базилевский М.П. Оценивание неизвестных параметров многослойной модульной регрессии методом наименьших модулей // Моделирование, оптимизация и информационные технологии. 2024. Т. 12, № 2 (45). doi:10.26102/2310-6018/2024.45.2.039
18. Базилевский М.П. Оценивание регрессионных моделей с регрессорами в виде модулей линейных комбинаций объясняющих переменных // System Analysis and Mathematical Modeling. 2024. Т. 6, № 3. С. 269–281. doi: 10.17150/2713-1734.2024.6(3).269-281
19. Базилевский М.П. Алгоритм оценивания неизвестных параметров одного вида многослойных неэлементарных линейных регрессий // Ученые записки Комсомольского-на-Амуре государственного технического университета. 2025. № 1 (81). С. 35–44.
20. 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.
21. Mixed Integer Linear Programming (MILP) solver – LPSolve // Sourceforge. URL: https://sourceforge.net/projects/lpsolve/ (accessed: 07.06.2025).
22. Ge D., Huangfu Q., Wang Z, Wu J., Ye Y. Cardinal Optimizer (COPT) user guide. URL: https://guide.coap.online/copt/en-doc (accessed: 07.06.2025).
23. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. М. : Финансы и статистка, 1987. Кн. 2. 353 с.

References

1. Montgomery, D.C., Peck, E.A. & Vining, G.G. (2021) *Introduction to Linear Regression Analysis*. New York: John Wiley and Sons.
2. Baždarić, K., Šverko, D., Salarić, I., Martinović, A. & Lucijanić, M. (2021) The ABC of linear regression analysis: What every author and editor should know. *European Science Editing.* 47. doi: 10.3897/ese.2021.e63780
3. Batmazova, A.A., Gaydukova, E.V. & Dregval, M.S. (2024) Osobennosti postroeniya regressionnoy modeli dlya otsenki i prognoza urovnya ozera Lovozero [Building a Regression Model for Assessing and Forecasting the Water Level of Lovozero Lake]. *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal*. 7(145). doi: 10.60797/IRJ.2024.145.67
4. Byastinova, L.M. (2023) Ispol'zovanie mnogofaktornykh regressionnykh modeley pri prognozirovanii v sfere sel'skogo khozyaystva Yakuti [Using Multivariate Regression Models for Forecasting in the Agricultural Sector of Yakutia]. *Ekonomika i prirodopol'zovanie na Severe.* 4. pp. 102–110. doi: 10.25587/2587-8778-2023-4-102-110
5. Kashtanov, A.L., Komyakov, A.A. & Nikiforov, M.M. (2021) Prognozirovanie i verifikatsiya klyuchevykh pokazateley energeticheskoy effektivnosti zheleznodorozhnogo transporta [Forecasting and Verification of Key Energy Efficiency Indicators of Railway Transport]. *Vestnik Ural'skogo gosudarstvennogo universiteta putey soobshcheniya.* 1. pp. 46–54. doi: 10.20291/2079-0392-2021-1-46-54
6. Gorshenin, A.Yu. & Vasina, D.I. (2023) Sravnenie ispol'zuemykh metodov pri prognozirovanii vyrabotki elektroenergii vetroelektrostantsiyami [Comparison of Methods Used in Forecasting Electricity Generation by Wind Power Plants]. *Matematicheskie struktury i modelirovanie.* 3(67). pp. 36–42. doi:10.24147/2222-8772.2023.3.36-42
7. Brachunova, U.V., Kozlovskiy, V.N., Shakurskiy, M.V. & Pantyukhin, O.V. (2023) Osnovnye aspekty razrabotki matematicheskoy modeli i programmnogo kompleksa po otsenke zaryadnogo balansa legkovogo avtomobilya [Main Aspects of Developing a Mathematical Model and Software Package for Assessing the Charge Balance of a Passenger Car]. *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki.* 5. pp. 242–249. doi: 10.24412/2071-6168-2023-5-242-243
8. Gichiev, N.S. (2022) Analiz dvukhfaktornoy modeli ekonomicheskogo rosta na osnove proizvodstvennoy funktsii Kobba–Duglasa [Analysis of a Two-Factor Economic Growth Model Based on the Cobb–Douglas Production Function]. *Regional'nye problemy preobrazovaniya ekonomiki.* 3(137). pp. 62–66. doi: 10.26726/1812-7096-2022-3-62-66
9. Botnaryuk, M.V. & Ksenzova, N.N. (2022) Proizvodstvennaya funktsiya Kobba–Duglasa dlya otsenki deyatel'nosti morskogo porta [The Cobb–Douglas Production Function for Assessing the Activities of a Seaport]. *Nauchnye problemy vodnogo transporta*. 74. pp. 85–95. doi: 10.37890/jwt.vi74.348
10. Demidenko, E.Z. (1981) *Lineynaya i nelineynaya regressii* [Linear and Nonlinear Regression]. Moscow: Finansy i statistika.
11. Wagner, H.M. (1959) Linear programming techniques for regression analysis. *Journal of the American Statistical Association.* 54(285). pp. 206–212. doi: 10.1080/01621459.1959.10501506
12. Koch, T., Berthold, T., Pedersen, J. & Vanaret, C. (2022) Progress in mathematical programming solvers from 2001 to 2020. *EURO Journal on Computational Optimization.* 10. Art. 100031. doi: 10.1016/j.ejco.2022.100031
13. Kleyner, G.B. (1986) *Proizvodstvennye funktsii: teoriya, metody, primenenie* [Production Functions: Theory, Methods, Application]. Moscow: Finansy i statistika.
14. Noskov, S.I. (1996) *Tekhnologiya modelirovaniya ob"ektov s nestabil'nym funktsionirovaniem i neopredelennost'yu v dannykh* [Modeling Technology for Objects with Unstable Operation and Data Uncertainty]. Irkutsk: Oblinformatpechat'.
15. Noskov, S.I. (2024) Otsenivanie parametrov prostoy vlozhennoy kusochno-lineynoy regressii s lineynoy sostavlyayushchey [Estimating Parameters of a Simple Nested Piecewise Linear Regression with a Linear Component]. *Vestnik Yugorskogo gosudarstvennogo universiteta*. 20(1). pp. 19–21. doi: 10.18822/byusu20240119-21
16. Bazilevskiy, M.P. & Oydopova, A.B. (2023) Otsenivanie modul'nykh lineynykh regressionnykh modeley s pomoshch'yu metoda naimen'shikh moduley [Estimating Modular Linear Regression Models Using the Least Absolute Deviations Method]. *Vestnik Permskogo natsional'nogo issledovatel'skogo politekhnicheskogo universiteta. Elektrotekhnika, informatsionnye tekhnologii, sistemy upravleniya.* 45. pp. 130–146. doi: 10.15593/2224-9397/2023.1.06

17. Bazilevskiy, M.P. (2024) Otsenivanie neizvestnykh parametrov mnogosloynoy modul'noy regressii metodom naimen'shikh moduley [Estimating Unknown Parameters of Multilayer Modular Regression Using the Least Absolute Deviations Method]. *Modelirovanie, optimizatsiya i informatsionnye tekhnologii*. 12(2(45)). doi: 10.26102/2310-6018/2024.45.2.039
18. Bazilevskiy, M.P. (2024) Otsenivanie regressionnykh modeley s regressorami v vide moduley lineynykh kombinatsiy ob"yasnyayushchikh peremennykh [Estimating Regression Models with Regressors in the Form of Modules of Linear Combinations of Explanatory Variables]. *System Analysis and Mathematical Modeling*. 6(3). pp. 269–281. doi: 10.17150/2713-1734.2024.6(3).269-281
19. Bazilevskiy, M.P. (2025) Algoritm otsenivaniya neizvestnykh parametrov odnogo vida mnogosloynykh neelementarnykh lineynykh regressiy [Algorithm for Estimating Unknown Parameters of One Type of Multilayer Non-elementary Linear Regressions]. *Uchenye zapiski Komsomol'skogo-na-Amure gosudarstvennogo tekhnicheskogo universiteta*. 1(81). pp. 35–44.
20. Noskov, S.I. & Khonyakov, A.A. (2019) Programmyy kompleks postroyeniya nekotorykh tipov kusochno-lineynykh regressiy [Software Package for Constructing Certain Types of Piecewise Linear Regressions]. *Informatsionnye tekhnologii i matematicheskoe modelirovanie v upravlenii slozhnymi sistemami*. 3(4). pp. 47–55.
21. *Mixed Integer Linear Programming (MILP) solver – LPSolve*. [Online] Available from: https://sourceforge.net/projects/lpsolve/ (Accessed: 7th June 2025).
22. Ge, D., Huangfu, Q., Wang, Z., Wu, J. & Ye, Y. (n.d.) *Cardinal Optimizer (COPT) user guide*. [Online] Available from: https://guide.coap.online/copt/en-doc (Accessed: 7th June 2025).
23. Draper, N. & Smith, G. (1987) *Prikladnoy regressionnyy analiz* [Applied Regression Analysis]. Vol. 2. Translated from English. Moscow: Finansy i statistika.

Информация об авторе:

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

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Bazilevskiy Mikhail P. (Associate Professor, Candidate of Technical Sciences, Irkutsk State Transport University, Irkutsk, Russian Federation). E-mail: mik2178@yandex.ru

The author declares no conflicts of interests.

Поступила в редакцию 07.06.2025; принята к публикации 02.12.2025

Received 07.06.2025; accepted for publication 02.12.2025