DOI: 10.47026/1810-1909-2024-4-75-83

УДК 519.852
ББК 22.18

С.И. НОСКОВ, А.Р. ЧЕКАЛОВА

ПОСТРОЕНИЕ ОДНОРОДНОЙ ВЛОЖЕННОЙ КУСОЧНО-ЛИНЕЙНОЙ РЕГРЕССИИ С ЗАПАЗДЫВАЮЩИМИ ПЕРЕМЕННЫМИ

Ключевые слова: *однородная вложенная кусочно-линейная регрессия с запаздывающими переменными, идентификация параметров, метод наименьших модулей, задача линейно-булева программирования, объем добычи нефти, капитальные вложения, ввод новых скважин.*

При построении регрессионных моделей объектов любой природы часто возникает необходимость задействовать нелинейные аппроксимирующие конструкции, в том числе кусочно-линейные, при этом исследуемый процесс может иметь выраженный динамический характер, поэтому в качестве регрессоров могут быть использованы запаздывающие (лаговые) переменные.
Цель исследования *– разработка алгоритмического способа идентификации параметров однородной вложенной кусочно-линейной регрессии с запаздывающими переменными.*
Материалы и методы. *Для достижения цели использовались предложенные ранее одним из авторов приемы сведения задач оценивания параметров вложенных кусочно-линейных моделей к задачам линейно-булева программирования. Был применен также известный в регрессионном анализе метод наименьших модулей. В качестве объекта моделирования был принят объем добычи нефти в Российской Федерации с использованием статистических исходных данных за 2013–2022 гг.*
Результаты исследования. *Разработан алгоритмический способ построения однородной вложенной кусочно-линейной регрессии с запаздывающими переменными, сводящийся к решению задачи линейно-булева программирования. Он применен для построения модели для определения возможного объема добычи нефти в Российской Федерации. При этом в качестве независимых переменных использовались данные об объемах капитальных вложений российских вертикально-интегрированных нефтяных компаний и о вводе в действие новых скважин.*
Выводы. *Разработанный способ построения однородной вложенной кусочно-линейной регрессии с запаздывающими переменными при использовании метода наименьших модулей сводится к задаче линейно-булева программирования. Подобные модели позволяют выявлять лимитирующие значения зависимых переменных факторов с учетом возможных задержек во влиянии на внешнем и внутреннем уровнях.*

Введение. При построении регрессионных моделей объектов любой природы часто возникает необходимость использовать нелинейные аппроксимирующие конструкции, в том числе кусочно-линейные. При этом, если исследуемый процесс имеет выраженный динамический характер, в качестве регрессоров могут быть задействованы запаздывающие (лаговые) переменные. Так, в работе [22] рассматриваются пороговые модели авторегрессии и модели авторегрессии с функциональными коэффициентами в качестве особых случаев, но с дополнительными преимуществами, такими как отображение более тонкой структуры базовой динамики и лучшая производительность прогнозирования после выборки. Также предлагается новый бутстреп-тест на предмет оценки качества моделей и селектор полосы пропускания, основанный на недавно определенной перекрестной проверке ожидаемых ошибок прогнозирования.

В [20] параметрические и непараметрические модели используются для прогнозирования ежемесячных колебаний уровня воды в озере Урмия (Иран). Одиннадцать предыдущих уровней воды в виде данных с запаздыванием помесячно используются в качестве известных независимых переменных модели, в то время как уровень воды в озере на двенадцатом месяце рассматривается как неизвестная зависимая переменная, подлежащая прогнозированию. Параметрические модели, используемые при моделировании, включают мультилинейную регрессию, аддитивную и мультипликативную нелинейную регрессию и дерево решений с нейронной сетью прямого и обратного распространения ошибки. Исследование [14] посвящено оценке нелинейной модели с распределенным запаздыванием на основе деревьев байесовской аддитивной регрессии. Этот метод использует набор деревьев регрессии, каждое из которых предполагает кусочно-постоянные отношения в пространстве воздействия и времени. Показано, что полученная модель превосходит модели на основе сплайнов, когда поверхность времени экспозиции не является гладкой, в то время как оба метода работают одинаково в условиях, когда истинная поверхность гладкая. В статье [15] исследуется влияние изменений цен на нефть на уровень безработицы в Нигерии с использованием реальных цен на нефть марок Brent и West Texas International, с применением методов оценки линейной и нелинейной авторегрессии с распределенным лагом. В работе [17] отмечается, что в регрессии временных рядов, где один выброс может появляться в векторе регрессоров несколько раз из-за присутствия запаздывающих переменных, устойчивость способа оценки к выбросам может вызывать серьезную озабоченность. Показано, что высокая устойчивость S-оценок в поперечной регрессии распространяется и на временные ряды. В [18] предлагается подход на основе нелинейной векторной авторегрессионной нейронной сети для прогнозирования потоков авиапассажиров. В публикации [10] описаны статистические методы, которые могут охарактеризовать лагированные эффекты и представить структурированный подход к анализу данных с целью разработки соответствующей модели. В статье [9] показана тестовая статистика на основе регрессии двойной длины для тестирования нелинейности и/или пространственного лага. В частности, выводятся тесты для совместной проверки линейных или лог-линейных моделей без зависимости от пространственного запаздывания по сравнению с общей моделью Бокса-Кокса с зависимостью от него. В [11] рассматривается проблема идентификации и оценки в моделях дискретного выбора панельных данных, когда набор объясняющих переменных включает строго экзогенные переменные, лаги эндогенной зависимой переменной, а также ненаблюдаемые индивидуальные эффекты. В [4] для оценивания времени запаздывания в регрессионной модели предлагается использовать подход, основанный на генетическом алгоритме. В публикации [21] дисперсия оценивается как точка пересечения линейной регрессии с лаговой оценкой дисперсии типа Гассера в качестве зависимой переменной. В [8] с помощью векторной авторегрессии изучаются связи между ценами на акции и ключевыми макроэкономическими индикаторами: инфляцией, промышленным производством, процентными ставками, денежной массой и некоторыми взаимодействиями между последней группой переменных.

Приведенный краткий обзор публикаций указывает на актуальность проблемы расширения арсенала используемых при регрессионном моделировании форм связи между переменными путем разработки новых типов кусочно-линейных конструкций.

Цель исследования состоит в разработке алгоритмического способа идентификации параметров однородной вложенной кусочно-линейной регрессии с запаздывающими переменными.

Материалы и методы. Для достижения цели использовались предложенные ранее одним из авторов приемы сведения задач оценивания параметров вложенных кусочно-линейных моделей к задачам линейно-булева программирования (ЛБП). Был применен также известный в регрессионном анализе метод наименьших модулей. В качестве объекта моделирования был принят объем добычи нефти в Российской Федерации с использованием статистических исходных данных за 2013–2022 гг.

Результаты исследования. Пусть для исследуемого динамического процесса задана выборка данных (X, y), где X − матрица наблюдений независимых переменных с компонентами x_{ki} размерности $(n \times m)$; $y = (y_1, \dots, y_n)^T$ – вектор наблюдений зависимой переменной; n – длина выборки; m – число предикторов (независимых переменных); k – номер наблюдения (года, месяца, квартала и т.д.).

При регрессионном моделировании динамических процессов эффективны зависимости с запаздывающими переменными:

$$y_k = \alpha_0 + \sum_{i=1}^{m} \alpha_i \, x_{(k-\tau_i),i} + \varepsilon_k, \tag{1}$$

где τ_i – лаги запаздывания ($i = \overline{1, m}$); ε_k – ошибки аппроксимации.

Линейная модель учитывает то обстоятельство, что на зависимую переменную могут оказывать влияние независимые переменные в прошлые по отношению к моменту времени k периоды.

В работе [6] одним из авторов введены вложенные кусочно-линейные регрессии:

– простая вложенная кусочно-линейная регрессия первого типа

$$y_k = \min\{ \min_{i \in I}\{\alpha_i x_{ki}\}, \max_{i \in J}\{\beta_i x_{ki}\}\} + \varepsilon_k, \tag{2}$$

– простая вложенная кусочно-линейная регрессия второго типа

$$y_k = \max\{ \min_{i \in I}\{\alpha_i x_{ki}\}, \max_{i \in J}\{\beta_i x_{ki}\}\} + \varepsilon_k, \tag{3}$$

– однородная вложенная кусочно-линейная регрессия первого типа

$$y_k = \min\{ \min_{i \in I^1}\{\alpha_i^1 x_{ki}\}, \dots, \min_{i \in I^G}\{\alpha_i^G x_{ki}\}\} + \varepsilon_k, \tag{4}$$

– однородная вложенная кусочно-линейная регрессия второго типа

$$y_k = \max\{ \max_{i \in J^1}\{\beta_i^1 x_{ki}\}, \dots, \max_{i \in J^H}\{\beta_i^H x_{ki}\}\} + \varepsilon_k, \tag{5}$$

где индексные множества $I^i, i = \overline{1, G}, J^i, i = \overline{1, H}$ являются подмножествами множества номеров независимых переменных $\{1,2,\dots,m\}$:

$$I, J, I^i \subseteq \{1,2, \dots, m\}, i = \overline{1, G}, J^i \subseteq \{1,2,\dots, m\}, i = \overline{1, H}.$$

При этом допускаются их всевозможные непустые попарные пересечения.

В случае, когда в качестве функции потерь выбрана соответствующая методу наименьших модулей сумма абсолютных ошибок аппроксимации

$$\Delta = \sum_{k=1}^{n} |\varepsilon_k|, \tag{6}$$

задачи оценивания параметров вложенных кусочно-линейных регрессий (2)–(5) могут быть сведены к задачам ЛБП (см., например, [5]).

Поставим задачу построения однородной вложенной кусочно-линейной регрессии с запаздывающими переменными:

$$y_k = a_0 + \min\{\min_{i\in I^1}\{\alpha_i^1 x_{k-\tau_1^j,1}\}, \dots, \min_{i\in I^m}\{\alpha_i^m x_{k-\tau_m^j,m}\}\} + \varepsilon_k, \quad k = \overline{\bar{\tau}+1, n}, \tag{7}$$

где $\tau_i^j \in I^i \subseteq \{0,1,\dots,\bar{\tau}_i\}$; $\bar{\tau}_i = \max\limits_{j=\overline{1,|I^i|}} \tau_i^j$, $i = \overline{1,m}$; $\bar{\tau} = \max\limits_{i=\overline{1,m}} \bar{\tau}_i$; $|I^i|$ – мощность множества I^i.

Введем в рассмотрение векторы

$$q^i = (q_1^i, \dots, q_{|I^i|}^i),\ i = \overline{1,m},$$

компоненты которых равны элементам множеств I^i.

Введем также матрицы $\bar{X}^i$ ($i=\overline{1,m}$) размерности $(n-\bar{\tau}) \times |I^i|$ по правилу

$$\bar{x}_{\bar{\tau}+k,j}^i = x_{\bar{\tau}-q_j^i+k,i},\ k = \overline{1, n-\bar{\tau}},\ j = \overline{1,|I^i|},\ i = \overline{1,m}.$$

Кроме того, введем переменные

$$v_{ki} = \min_{j=\overline{1,|I^i|}} \{a_i^j \bar{x}_{kj}^i\}, k = \overline{\bar{\tau}+1, n},\ i = \overline{1,m},$$

$$w_k = \min_{i=\overline{1,m}} v_{ki}, k = \overline{\bar{\tau}+1, n}.$$

Тогда, по аналогии с примененными в [5] приемами, задача идентификации параметров $a_0, a_i^j, j = \overline{1,|I^i|}$ $(i = \overline{1,m})$ по методу наименьших модулей сводится к следующей задаче ЛБП:

$$v_{ki} \le a_i^j \bar{x}_{kj}^i,\ k = \overline{\bar{\tau}+1, n},\ j = \overline{1,|I^i|}, i = \overline{1,m}, \tag{8}$$

$$a_i^j \bar{x}_{kj}^i - v_{ki} \le (1 - s_{kij}) M_1,\ k = \overline{\bar{\tau}+1, n},\ j = \overline{1,|I^i|},\ i = \overline{1,m}, \tag{9}$$

$$\sum_{j=1}^{|I^i|} s_{kij} = 1, k = \overline{\bar{\tau}+1, n}, i = \overline{1,m}, \tag{10}$$

$$w_k \le v_{ki},\ k = \overline{\bar{\tau}+1, n},\ i = \overline{1,m}, \tag{11}$$

$$v_{ki} - w_k \le (1 - r_{ki}) M_2,\ k = \overline{\bar{\tau}+1, n},\ i = \overline{1,m}, \tag{12}$$

$$\sum_{i=1}^{m} r_{ki} = 1, k = \overline{\bar{\tau}+1, n}, \tag{13}$$

$$a_0 + w_k + \bar{u}_k - \bar{v}_k = y_k,\ k = \overline{\bar{\tau}+1, n}, \tag{14}$$

$$\bar{u}_k \ge 0, \bar{v}_k \ge 0, k = \overline{\bar{\tau}+1, n}, \tag{15}$$

$$s_{kij} \in \{0,1\}, k = \overline{\bar{\tau}+1, n},\ j = \overline{1,|I^i|},\ i = \overline{1,m}, \tag{16}$$

$$r_{ki} \in \{0,1\}, k = \overline{\tau + 1, n},\ i = \overline{1, m}, \tag{17}$$

$$\sum_{k=\overline{\tau}+1}^{n} (\overline{u}_k + \overline{v}_k) + \rho \sum_{i=1}^{m} \sum_{j=1}^{|I^i|} a_i^j \to \min, \tag{18}$$

где M_1 и M_2 – наперед заданные большие положительные числа; ρ – малое положительное число.

Применим описанный выше способ оценивания параметров однородной вложенной кусочно-линейной регрессии с запаздывающими переменными для построения модели объема добычи нефти в Российской Федерации. Введем следующие переменные:

y – объем добычи нефти, млн т;

x_1 – объем капитальных вложений российских вертикально-интегрированных нефтяных компаний, трлн руб.;

x_2 – ввод новых скважин, тыс. шт.

Заметим, что именно эти предикторы являются определяющими при моделировании объема добычи нефти (см., в частности, [7]).

В таблице приведены статистические данные для указанных показателей за 2013–2022 гг., взятые из официальных источников [1–3, 12, 13, 16, 19].

Статистические данные

Год	y	x_1	x_2
2013	523,3	0,8962	6,454
2014	526,8	0,93	6,065
2015	534,1	1,07	6,261
2016	547,5	1,19	7,141
2017	546,8	1,39	8,184
2018	555,8	1,45	7,746
2019	560,2	1	7,861
2020	512,7	1,3	6,96
2021	524	1,494	7,4
2022	534	1,818	7,9

Сформируем дополнительную информацию:

$$I^1 = \{0, 1\}, I^2 = \{2, 3\},$$

$$\overline{X}^1 = \begin{pmatrix} 1,19 & 1,07 \\ 1,39 & 1,19 \\ 1,45 & 1,39 \\ 1 & 1,45 \\ 1,3 & 1 \\ 1,494 & 1,3 \\ 1,818 & 1,494 \end{pmatrix}, \overline{X}^2 = \begin{pmatrix} 6,065 & 6,454 \\ 6,261 & 6,065 \\ 7,141 & 6,261 \\ 8,184 & 7,141 \\ 7,746 & 8,184 \\ 7,861 & 7,746 \\ 6,96 & 7,861 \end{pmatrix}.$$

В результате решения задачи ЛБП (8)–(18) получим следующую однородную вложенную кусочно-линейную модель с запаздывающими переменными:

$$y_k = 485,47 +$$

$$+ \min\{\min\{74,73x_{k1}, 51,53x_{k-1,1}\}, \min\{9,85x_{k-2,2}, 11,23x_{k-3,2}\}\} + \varepsilon_k, \tag{19}$$

$$k = \overline{4, 10}.$$

Приведем также оптимальные значения других переменных задачи:

$$V = \begin{pmatrix} 55{,}15 & 59{,}73 \\ 61{,}33 & 61{,}66 \\ 71{,}64 & 70{,}33 \\ 74{,}73 & 80{,}21 \\ 51{,}53 & 76{,}29 \\ 67 & 77{,}42 \\ 76{,}99 & 65{,}55 \end{pmatrix}, W = \begin{pmatrix} 55{,}15 \\ 61{,}33 \\ 70{,}33 \\ 74{,}73 \\ 51{,}53 \\ 67 \\ 68{,}55 \end{pmatrix},$$

$$S_{i=1} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, S_{i=2} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}, r = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \overline{u} = \begin{pmatrix} 6{,}88 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \overline{v} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 24{,}31 \\ 28{,}47 \\ 20{,}02 \end{pmatrix}.$$

Таким образом, сумма модулей ошибок модели

$$\sum_{k=4}^{10} |\varepsilon_k| = \sum_{k=4}^{10} (\overline{u}_k + \overline{v}_k) = 79{,}68,$$

а, например, для 2022 г. внешний минимум в модели (19) «сработал» на втором внутреннем минимуме (поскольку $r_{10,1} = 0, r_{10,2} = 1$), а в нем – на переменной $x_{k-2,2}$ (так как $s_{10,2,1} = 1$, $s_{10,2,2} = 0$). То есть ограничивающим объем добычи нефти в этом году фактором стало число введенных в эксплуатацию скважин в 2020 г. Аналогичным образом можно выявить лимитирующие факторы для остальных шести лет.

Заметим, что время решения задачи ЛБП при расчете параметров модели (19) с помощью размещенной в сети Интернет в свободном доступе программы LPsolve составило 0,033 с, что вполне приемлемо для реальных размерностей.

Выводы. При построении регрессионных моделей объектов любой природы часто возникает необходимость использовать нелинейные аппроксимирующие конструкции, в том числе кусочно-линейные. При этом, если исследуемый процесс имеет выраженный динамический характер, в качестве регрессоров могут быть задействованы запаздывающие (лаговые) переменные. Предложен алгоритмический способ идентификации параметров однородной вложенной кусочно-линейной регрессии с запаздывающими переменными, сводящийся к решению задачи ЛБП. Его существенным достоинством является возможность построения модели указанного класса по реальным данным с допустимой размерностью за приемлемое время. Построена модель этого типа для объема добычи нефти в Российской Федерации. В качестве независимых переменных задействованы объемы капитальных вложений российских вертикально-интегрированных нефтяных компаний и ввод новых скважин. Построенная за приемлемое время модель позволяет адекватно выявить лимитирующие объем добычи нефти в каждом году факторы с учетом их запаздывания, что позволяет констатировать достижение цели исследования.

Литература

1. Глава Минэнерго отметил рост капитальных вложений в добычу нефти [Электронный ресурс] // Прайм. 2023. 16 марта. URL: https://1prime.ru/20230316/840089260.html (дата обращения: 30.06.2024).

2. Нефть и капитал [Электронный ресурс]. URL: https://oilcapital.ru/news/2015-02-20/vink-uvelichili-investitsii-v-neftedobychu-v-rf-v-2014-g-na-3-8-do-930-mlrd-rub-minenergo-878817 (дата обращения: 30.06.2024).

3. *Новак А.* Инвестиции ВИНК в нефтедобычу в РФ 2016 году выросли на 10% – до 1,19 трлн рублей [21.12.2016] // Рамблер/Финансы: [сайт]. URL: https://finance.rambler.ru/-economics/35647334/?utm_content=finance_media&utm_medium=read_more&utm_source=copylink (дата обращения: 30.06.2024).

4. *Новиков Е.И.* Оценивание параметров линейных регрессионных моделей с учетом запаздывания влияния факторов на зависимую переменную // Интернет-журнал «Науковедение». 2016. Т. 8, № 4(35). С. 74.

5. *Носков С.И., Белинская С.И.* Вычисление оценок параметров однородной вложенной кусочно-линейной регрессии // Вестник Дагестанского государственного технического университета. Технические науки. 2023. Т. 50, № 4. С. 115–120.

6. *Носков С.И.* Некоторые формы вложенной кусочно-линейной регрессии // Известия Тульского государственного университета. Технические науки. 2023. № 3. С. 467–469.

7. Факторы развития нефтесервисного рынка России / *И.В. Филимонова, В.Ю. Немов, А.В. Комарова, С.В. Кожевина* // Нефтегазовая вертикаль. 2020. № 21-22. С. 6–14.

8. *Camilleri S.J., Scicluna N., Ye Bai.* Do stock markets lead or lag macroeconomic variables? Evidence from select European countries. *The North American Journal of Eco nomics and Finance*, 2019, vol. 48, pp. 170–186.

9. *Dong Li, Canh Le.* Nonlinearity and Spatial Lag Dependence: Tests Based on Double-Length Regressions. *Journal of Time Series Econometrics*, 2010, vol. 2, no. 1.

10. *Heagerty P.J., Comstock B.A.* Exploration of Lagged Associations Using Longitudinal Data. *Biometrics*, 2013, vol. 69(1), pp. 197–205.

11. *Honoré B.E., Kyriazidou E.* Panel Data Discrete Choice Models with Lagged Dependent Variables. *Econometrica*, 2003, vol. 68, no. 4, pp. 839–874.

12. iFinance. Available at: http://global-finances.ru/dobyicha-nefti-v-rossii-po-godam/ (Access Date: 2024, June 30).

13. INFOline. Available at: https://infoline.spb.ru/news/?news=160790 (Access Date: 2024, June 30).

14. *Mork D., Wilson A.* Treed distributed lag nonlinear models. *Biostatistics*, 2022, vol. 23(3), pp. 754–771.

15. *Raifu I.A., Aminu A., Folawewo A.O.* Investigating the relationship between changes in oil prices and unemployment rate in Nigeria: linear and nonlinear autoregressive distributed lag approaches. *Future Business Journal*, 2020, vol. 6, 28. DOI: 10.1186/s43093-020-00033-w.

16. RCC. Available at: http://rcc.ru/article/obem-kapvlozheniy-rossiyskih-vink-v-neftedobychu-v-2019-godu-snizilsya-na-27-75072 (Access Date: 2024, June 30).

17. *Sakata S., White H.* S-estimation of nonlinear regression models with dependent and heterogeneous observations. *Journal of Econometrics*, 2001, vol. 103(1-2), pp. 5–72.

18. *Shaolong Sun, Hongxu Lu, Kwok-Leung Tsui, Shouyang Wang*. Nonlinear vector auto-regression neural network for forecasting air passenger flow. *Journal of Air Transport Management*, 2019, vol. 78, pp. 54–62.

19. Statista. Available at: https://www.statista.com/statistics/1318266/new-oil-wells-russia/ (Access Date: 2024, June 30).

20. *Vaheddoost B., Aksoy H., Abghari H.* Prediction of Water Level using Monthly Lagged Data in Lake Urmia. Iran Water Resources Management, 2016, vol. 30(13), pp. 4951–4967.

21. *WenWu Wang, Lu Lin, Li Yu.* Optimal variance estimation based on lagged second-order difference in nonparametric regression. *Computational Statistics*, 2017, vol. 32, pp. 1047–1063.

22. *Zongwu Cai, Jianqing Fan, Qiwei Yao*. Functional-Coefficient Regression Models for Nonlinear Time Series. *Journal of the American Statistical Association*, 2000, vol. 95, no. 451, pp. 941–956.

НОСКОВ СЕРГЕЙ ИВАНОВИЧ – доктор технических наук, профессор, кафедра информационных систем и защиты информации, Иркутский государственный университет путей сообщения, Россия, Иркутск (sergey.noskov.57@mail.ru; ORCID: https://orcid.org/0000-0003-4097-2720).

ЧЕКАЛОВА АЛЕКСАНДРА РОМАНОВНА – магистрант кафедры информационных систем и защиты информации, Иркутский государственный университет путей сообщения, Россия, Иркутск (chekalova49@gmail.com; ORCID: https://orcid.org/0009-0009-3811-9051).

Sergei I. NOSKOV, Aleksandra R. CHEKALOVA

CONSTRUCTION OF A HOMOGENEOUS NESTED PIECEWISE LINEAR REGRESSION WITH LAGGING VARIABLES

Key words: *homogeneous nested piecewise linear regression with lagging variables, parameter identification, method of least modules, linear Boolean programming problem, oil production volume, capital investments, commissioning of new wells.*

When constructing regression models of objects of any nature, it is often necessary to use nonlinear approximating constructions, including piecewise linear ones, while the process under study can have a pronounced dynamic nature, therefore, lagging (lag) variables can be used as regressors.
The research purpose *is to develop an algorithmic method of identifying the parameters of a homogeneous nested piecewise linear regression with lagging variables.*
Materials and methods. *To achieve the goal, the methods of reducing the problems of estimating the parameters of nested piecewise linear models to linear Boolean programming problems proposed earlier by one of the authors were used. The least absolute values method, known in regression analysis, was also applied. The volume of oil production in the Russian Federation was adopted as the modeling object using statistical initial data in 2013–2022.*
Research results. *An algorithmic method of constructing a homogeneous nested piecewise linear regression with lagging variables has been developed, which is reduced to solving a linear Boolean programming problem. It has been applied to construct a model of the possible volume of oil production in the Russian Federation. In this case, data on the volume of capital investments of Russian vertically integrated oil companies and on the commissioning of new wells have been used as independent variables.*
Conclusions. *The developed method of constructing a homogeneous nested piecewise linear regression with lagging variables using the least absolute values method is reduced to a linear Boolean programming problem. Such models allow us to identify limiting values of dependent variables, taking into account possible delays in influence at external and internal levels.*

References

1. *Glava Minenergo otmetil rost kapital'nyh vlozhenii v dobychu nefti* [The head of the Ministry of Energy noted the increase in capital investments in oil production]. *Prime*, 2023, March 16. Available at: https://1prime.ru/20230316/840089260.html (Accessed Date: 2024, Jun. 30).

2. *Neft' i kapital* [Oil and capital]. Available at: https://oilcapital.ru/news/2015-02-20/vink-uvelichili-investitsii-v-neftedobychu-v-rf-v-2014-g-na-3-8-do-930-mlrd-rub-minenergo-878817 (Accessed Date: 2024, Jun. 30).

3. Novak A. *Investicii VINK v neftedobychu v RF 2016 godu vyrosli na 10% – do 1,19 trln rublei* [Investments of vertically integrated oil companies in oil production in the Russian Federation in 2016 increased by 10% – to 1.19 trillion rubles]. Available at: https://finance.rambler.ru/economics/35647334-investitsii-vink-v-neftedobychu-v-rf-2016-godu-vyrosli-na-10-do-1-19-trln-rubley/ (Accessed Date: 2024, Jun. 30).

4. Novikov E.I. *Otsenivanie parametrov lineinykh regressionnykh modelei s uchetom zapazdyvaniya vliyaniya faktorov na zavisimuyu peremennuyu* [Estimation of the parameters of linear regression models, taking into account the delay in the influence of factors on the dependent variable]. *Internet-zhurnal «Naukovedenie»*, 2016, vol. 8, no. 4 (35), p. 74.

5. Noskov S.I., Belinskaya S.I. *Vychislenie otsenok parametrov odnorodnoi vlozhennoi kusochno-lineinoi regressii* [Calculation of parameter estimates for homogeneous nested piecewise linear regression]. *Vestnik Dagestanskogo gosudarstvennogo tekhnicheskogo universiteta. Tekhnicheskie nauki*, 2023, vol. 50, no. 4, pp. 115–120.

6. Noskov S.I. *Nekotorye formy vlozhennoi kusochno-lineinoi regressii* [Some forms of nested piecewise linear regression]. *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki*, 2023, no. 3, pp. 467–469.

7. Filimonova I.V., Nemov V.Yu., Komarova A.V., Kozhevina S.V. *Faktory razvitiya nefteservisnogo rynka Rossii* [Factors of development of the Russian oilfield services market]. *Neftegazovaya vertikal'*, 2020, no. 21-22, pp. 6–14.

8. Camilleri S.J., Scicluna N., Ye Bai. Do stock markets lead or lag macroeconomic variables? Evidence from select European countries. *The North American Journal of Eco nomics and Finance*, 2019, vol. 48, pp. 170–186.

9. Dong Li, Canh Le. Nonlinearity and Spatial Lag Dependence: Tests Based on Double-Length Regressions. *Journal of Time Series Econometrics*, 2010, vol. 2, no. 1.

10. Heagerty P.J., Comstock B.A. Exploration of Lagged Associations Using Longitudinal Data. *Biometrics*, 2013, vol. 69(1), pp. 197–205.

11. Honoré B.E., Kyriazidou E. Panel Data Discrete Choice Models with Lagged Dependent Variables. *Econometrica*, 2003, vol. 68, no. 4, pp. 839–874.

12. iFinance. Available at: http://global-finances.ru/dobyicha-nefti-v-rossii-po-godam/ (Access Date: 2024, June 30).

13. INFOline. Available at: https://infoline.spb.ru/news/?news=160790 (Access Date: 2024, June 30).

14. Mork D., Wilson A. Treed distributed lag nonlinear models. *Biostatistics*, 2022, vol. 23(3), pp. 754–771.

15. Raifu I.A., Aminu A., Folawewo A.O. Investigating the relationship between changes in oil prices and unemployment rate in Nigeria: linear and nonlinear autoregressive distributed lag approaches. *Future Business Journal*, 2020, vol. 6, 28. DOI: 10.1186/s43093-020-00033-w.

16. RCC. Available at: http://rcc.ru/article/obem-kapvlozheniy-rossiyskih-vink-v-neftedobychu-v-2019-godu-snizilsya-na-27-75072 (Access Date: 2024, June 30).

17. Sakata S., White H. S-estimation of nonlinear regression models with dependent and heterogeneous observations. *Journal of Econometrics*, 2001, vol. 103(1-2), pp. 5–72.

18. Shaolong Sun, Hongxu Lu, Kwok-Leung Tsui, Shouyang Wang. Nonlinear vector auto-regression neural network for forecasting air passenger flow. *Journal of Air Transport Management*, 2019, vol. 78, pp. 54–62.

19. Statista. Available at: https://www.statista.com/statistics/1318266/new-oil-wells-russia/ (Access Date: 2024, June 30).

20. Vaheddoost B., Aksoy H., Abghari H. Prediction of Water Level using Monthly Lagged Data in Lake Urmia. Iran Water Resources Management, 2016, vol. 30(13), pp. 4951–4967.

21. WenWu Wang, Lu Lin, Li Yu. Optimal variance estimation based on lagged second-order difference in nonparametric regression. *Computational Statistics*, 2017, vol. 32, pp. 1047–1063.

22. Zongwu Cai, Jianqing Fan, Qiwei Yao. Functional-Coefficient Regression Models for Nonlinear Time Series. *Journal of the American Statistical Association*, 2000, vol. 95, no. 451, pp. 941–956.

SERGEY I. NOSKOV – Doctor of Technical Sciences, Professor, Department of Information Systems and Information Security, Irkutsk State Transport University, Russia, Irkutsk (sergey.noskov.57@mail.ru; ORCID: https://orcid.org/0000-0003-4097-2720).

ALEKSANDRA R. CHEKALOVA – Master's Program Student, Department of Information Systems and Information Security, Irkutsk State Transport University, Russia, Irkutsk (chekalova49@gmail.com; ORCID: https://orcid.org/0009-0009-3811-9051).

Формат цитирования: *Носков С.И., Чекалова А.Р.* Построение однородной вложенной кусочно-линейной регрессии с запаздывающими переменными // Вестник Чувашского университета. 2024. № 4. С. 75–83. DOI: 10.47026/1810-1909-2024-4-75-83.