

Разработка кусочно-линейной регрессионной модели сталелитейной компании с помощью непрерывной формы метода максимальной согласованности

С.И. Носков, Ю.А. Бычков

Иркутский государственный университет путей сообщения

Аннотация: В работе представлен краткий обзор публикаций, описывающих опыт применения методов математического моделирования для решения различных задач. Построена многофакторная кусочно-линейная регрессионная модель сталелитейной компании с помощью непрерывной формы метода максимальной согласованности. Для оценки адекватности модели использованы критерии: средняя относительная ошибка аппроксимации, непрерывный критерий согласованности поведения, сумма модулей ошибок аппроксимации. Сделан вывод о том, что полученная модель обладает достаточной точностью и может быть использована для прогнозирования.

Ключевые слова: математическое моделирование, кусочно-линейная регрессия, метод наименьших модулей, непрерывная форма метода максимальной согласованности, сталелитейная компания.

Математическое моделирование является эффективным инструментом изучения и прогнозирования явлений и процессов в разных областях знаний [1, 2]. Среди методов математического моделирования широкое распространение получил регрессионный анализ, который позволяет исследовать взаимосвязи между переменными и строить прогностические модели на основе имеющихся данных [3, 4]. В зависимости от целей исследования и совокупности исходной информации исследователь может выбирать различные виды регрессий - линейная [5, 6], кусочно-линейная [7, 8] и нелинейная [9, 10].

Рассмотрим производственную функцию некоторого объекта, выраженную в виде кусочно-линейной регрессионной модели (1):

$$y_k = \min\{a_1x_{k1}, a_2x_{k2}, \dots, a_mx_{km}\} + \varepsilon_k, k = \overline{1, n}. \quad (1)$$

Здесь y – зависимая переменная (обычно выпуск продукции), $x_i, i = \overline{1, m}$ – независимые переменные (ресурсные показатели), $a_i, i = \overline{1, m}$ – подлежащие оцениванию параметры, $\varepsilon_k, k = \overline{1, n}$ – ошибки аппроксимации и n – количество наблюдений.

Как отмечено в работе [11], аппроксимирующая модель (1) обладает особенностью – значение выходного показателя зависит от лимитирующего входного фактора, при этом увеличение других факторов не приводит к увеличению выхода.

Идентификация параметров подобной регрессионной модели может быть осуществлена с помощью непрерывной формы метода максимальной согласованности [12, 13], которая сводится к следующей задаче линейно-булевого программирования:

$$r \sum_{k=1}^n (u_k - v_k) + (1 - r) \sum_{k=1}^{n-1} \sum_{s=k+1}^n l_{ks} \rightarrow \min, \quad (2)$$

$$z_k + u_k - v_k = y_k, k = \overline{1, n}, \quad (3)$$

$$z_k \leq a_i x_{ki}, k = \overline{1, n}, i = \overline{1, m}, \quad (4)$$

$$a_i x_{ki} - z_k \leq (1 - \sigma_{ki})M, k = \overline{1, n}, i = \overline{1, m}, \quad (5)$$

$$\sum_{i=1}^m \sigma_{ks} = 1, k = \overline{1, n}, \quad (6)$$

$$\omega_{ks}(z_k - z_s) + l_{ks} \geq 0, k = \overline{1, n-1}, s = \overline{k+1, n}, \quad (7)$$

$$\sigma_{ki} \in \{0, 1\}, k = \overline{1, n}, i = \overline{1, m}. \quad (8)$$

где M - большое положительное число, $r \in (0, 1]$ – выбираемое исследователем число, устанавливающее компромисс между методом наименьших модулей (далее – МНМ) и непрерывной формой метода максимальной согласованности:

$$\omega_{ks} = \begin{cases} 1, & y_k - y_s > 0 \\ -1, & y_k - y_s < 0 \\ 0, & y_k - y_s = 0, \end{cases}$$

$$k = \overline{1, n-1}, s = \overline{k+1, n}.$$

Особенности выбора параметра r при построении регрессионных моделей отражены в работе [14].

Верифицировать построенные варианты регрессионной модели будем с помощью следующих критериев адекватности:

1. Средняя относительная ошибка аппроксимации E :

$$E = \frac{1}{n} \sum_{k=1}^n \left| \frac{y_k - \min\{a_1 x_{k1}, a_2 x_{k2}, \dots, a_m x_{km}\}}{y_k} \right| 100\%.$$

2. Непрерывный критерий согласованности поведения L :

$$L = \sum_{k=1}^{n-1} \sum_{s=k+1}^n l_{ks}.$$

3. Сумма модулей ошибок аппроксимации M :

$$M = \sum_{k=1}^n |y_k - \min\{a_1 x_{k1}, a_2 x_{k2}, \dots, a_m x_{km}\}|.$$

В качестве объекта исследования выделено ПАО «Северсталь» (далее – Компания) – крупная отечественная вертикально-интегрированная сталелитейная и горнодобывающая компания, создающая новые продукты и комплексные решения из стали вместе с клиентами и партнерами. Компания контролирует все этапы создания стоимости – от добычи железной руды и выплавки стали до производства готовой продукции, дистрибуции, создания готовых комплексных решений для потребителей, а также сервисного обслуживания. Продукция компании включает в себя высококачественный металлопрокат, в том числе переработанный, стальные трубы, метизы, используемые в различных отраслях промышленности, а также железорудное сырье и здания из металлоконструкций полного металлургического цикла, специализирующееся на выпуске листового и сортового проката и стальных труб черных металлов широкого ассортимента, а также метизной и штамповочной продукции.

Введем следующие обозначения:

y – выручка, млн. долларов,

x_1 – горячекатаный лист, тыс. т.,

x_2 – холоднокатаный лист, тыс. т.,

x_3 – оцинкованный лист и лист с другим металлическим покрытием, тыс. т.,

x_4 – сортовой прокат, тыс. т.,

x_5 – трубы большого диаметра, тыс. т.,

x_6 – прочие трубные изделия, профили, тыс. т.

Исходные данные для моделирования отражены в табл.1 и получены из официальных отчётных документов Компании «Операционные результаты «Северстали» за четвёртый квартал и двенадцать месяцев года» (использовались отчеты с 2010 по 2021 год) [15].

Таблица 1.

Исходные данные для моделирования

год	y	x_1	x_2	x_3	x_4	x_5	x_6
2009	13050	5977	2599	2228	759	390	435
2010	13573	6403	2252	1733	762	471	490
2011	15812	6053	1916	1991	805	501	636
2012	14104	6439	1835	2071	819	318	636
2013	9434	4206	1419	615	889	321	741
2014	8296	3662	1443	592	1168	400	862
2015	6396	3988	1336	624	1269	548	830
2016	5916	4080	964	560	1401	389	839
2017	7848	3949	1324	654	1406	416	830
2018	8580	3895	1286	835	1460	440	762
2019	8157	4382	977	934	1234	383	789
2020	6870	4442	918	939	668	188	997
2021	11638	4815	1092	1015	651	343	758

Моделирование осуществлялось с помощью специализированной программы [16]. Результаты моделирования отражены в табл. 2.

Таблица 2.

Варианты регрессионные модели

1	Построенная модель	
	Значения критериев адекватности	
$r = 0,1$	$y = \min (2,203x_1; 8,703x_2; 13,626x_3; 17,498x_4; 41,930x_5; 30,00x_6)$ $E = 11,875; L = 3203,316; M = 12936,439;$ $\lambda = (6, 4, 1, 5, 3, 3, 3, 3, 1, 1, 2, 5, 2)$	
$r = 0,2$	$y = \min (2,242x_1; 8,859x_2; 13,871x_3; 17,812x_4; 42,682x_5; 30,00x_6)$ $E = 11,447; L = 3260,807; M = 12646,524;$ $\lambda = (6, 4, 1, 5, 3, 1, 3, 3, 1, 1, 2, 5, 2)$	
$r = 0,5$	$y = \min (2,330x_1; 8,349x_2; 12,947x_3; 17,812x_4; 44,352x_5; 30,00x_6)$ $E = 10,393; L = 4363,724; M = 11258,273;$ $\lambda = (6, 4, 1, 5, 3, 3, 3, 3, 3, 1, 2, 2, 2)$	

1	2
$r = 0,7$	$y = \min (2,341x_1; 8,349x_2; 12,472x_3; 17,812x_4; 44,352x_5; 30,00x_6)$
	$E = 10,152; L = 4876,496; M = 10936,110;$ $\lambda = (6, 4, 1, 5, 3, 3, 3, 3, 3, 1, 2, 2, 2)$
$r = 0,8$	$y = \min (2,369x_1; 8,349x_2; 12,283x_3; 17,812x_4; 44,352x_5; 30,00x_6)$
	$E = 9,983; L = 5519,213; M = 10755,577;$ $\lambda = (6, 4, 4, 5, 3, 3, 3, 3, 3, 1, 2, 2, 2)$
$r = 1$	$y = \min (2,369x_1; 8,349x_2; 12,00x_3; 17,812x_4; 44,352x_5; 30,00x_6)$
	$E = 9,858; L = 6335,380; M = 10577,191;$ $\lambda = (6, 4, 4, 5, 3, 3, 3, 3, 3, 1, 2, 2, 2)$

Здесь λ – вектор срабатываний с компонентами, рассчитанными по правилу: $\lambda_k = s$, если $\min (\alpha_1 x_{k1}, \alpha_2 x_{k2}, \alpha_3 x_{k3}) = \alpha_s x_{ks}$.

На основе анализа таблицы 2. можно сделать следующие выводы:

1. Построенные варианты кусочно-линейной модели обладают достаточно высокой точностью. Значение E не превышает 12%, что является хорошим показателем для негладких моделей.

2. С увеличением r улучшаются и значения критериев адекватности, при $r = 0,8$ они достигают своего оптимального значения.

3. Вектор срабатывания λ с большей частотой содержит показатель x_3 , что свидетельствует о том, что именно он является проблемным фактором и для наращивания выручки необходимо, в первую очередь, увеличивать именно его.

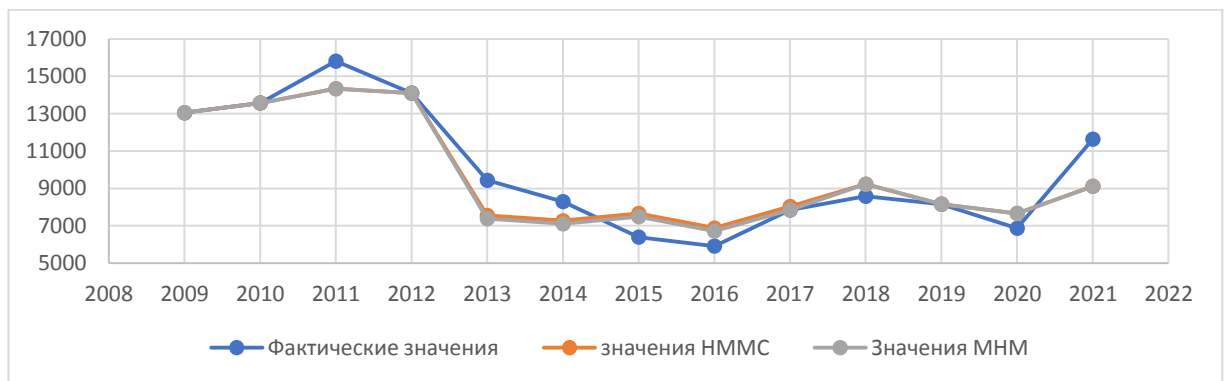


Рис. 1. Фактические и расчетные значения переменной y .

На рис. 1 представлены фактические и вычисленные по модели с $r = 0,8$, а также с соответствующим МНМ $r = 1$ значения y .

Таким образом, именно модель при $r = 0,8$ и подлежит дальнейшему использованию для решения с ее помощью различных практических проблем анализа и прогнозирования.

Литература

1. Платонов А. А., Терновская О. В. Математическая модель оценки качества работ по расчистке территорий от нежелательной растительности // Инженерный вестник Дона. 2022. № 6. URL: ivdon.ru/ru/magazine/archive/n6y2022/7746.

2. Зятева О. А. Имитационное моделирование показателей научной деятельности вузов // Инженерный вестник Дона. 2022. № 5. URL: ivdon.ru/ru/magazine/archive/n5y2022/7670.

3. Rishi T. Proceedings of the Jepson Undergraduate Conference on International Economics, 2022, Vol. 4, Article 4. URL: scholarworks.uni.edu/jucie/vol4/iss1/4.

4. Lee S. Y., Lee J. J., Lee H. Frontiers in Public Health, 2022, Vol. 10. URL: frontiersin.org/journals/public-health/articles/10.3389/fpubh.2022.1024751/full.

5. Гурова Л. П. Особенности линейной парной регрессии и её применение в экономике // Аллея науки. 2020. Т. 1. № 7(46). С. 278-282.

6. Иващук О. Д., Пузырев Н. С., Родионов А. Ю. Разработка системы прогнозирования цен акций с использованием LSTM-сетей и метода линейной регрессии // Информационные системы и технологии. 2023. № 6(140). С. 47-52.

7. Жижин К. С., Благородова Н. В. Использование кусочно-линейной регрессии в прогнозировании чрезвычайных ситуаций // Международный журнал экспериментального образования. 2016. № 5-3. С. 337-338.

8. Гордиенко А. С., Лозинская А. М., Тетерина Д. В., Шенкман Е. А. Исследование зависимости потребления электроэнергии и температуры в России: региональный разрез // Известия Российской академии наук. Энергетика. 2019. № 1. С. 15-27.

9. Князева С. В., Никитина А. Д., Белова Е. И., Плотникова А. С., Подольская Е. С., Ковганко К. А. Методические подходы к оценке характеристик лесов по данным спутниковой съемки сверхвысокого пространственного разрешения в оптическом диапазоне // Лесоведение. 2021. № 6. С. 645-672.

10. Синдеева Л. В., Медведева Н. Н., Николаев В. Г., Стрелкович Н. Н., Орлова И. И. Применение методов регрессионного анализа в биомедицинских исследованиях // Вестник новых медицинских технологий. 2013. Т. 20, № 2. С. 216-219.

11. Иванова Н. К., Лебедева С. А., Носков С. И. Идентификация параметров некоторых негладких регрессий // Информационные технологии и проблемы математического моделирования сложных систем. 2016. № 17. С. 107-110.

12. Носков С. И. Применение непрерывного критерия согласованности поведения при построении регрессионных моделей // Известия Тульского государственного университета. Технические науки. 2021. № 6. С. 74-78.

13. Носков С. И., Бычков Ю. А. Регрессионная модель динамики дорожно-транспортных происшествий в Российской Федерации // Южно-Сибирский научный вестник. 2023. № 1(47). С. 37-42.

14. Носков С. И., Бычков Ю. А. Вычислительные эксперименты с непрерывной формой метода максимальной согласованности в регрессионном анализе // Вестник Воронежского государственного технического университета. 2022. Т. 18, № 2. С. 7-12.



15. ПАО «Северсталь»: официальный сайт. URL: severstal.com/rus/ir/indicators-reporting/operational-results.

16. Носков С. И., Бычков Ю. А. Программа оценивания параметров кусочно-линейной регрессионной модели на основе применения непрерывной формы метода максимальной согласованности. Свидетельство о государственной регистрации программы для ЭВМ № 2023619993, 2023. URL: elibrary.ru/item.asp?id=53820507.

References

1. Platonov A. A., Ternovskaya O. V. Inzhenernyj vestnik Dona. 2022. № 6. URL: ivdon.ru/ru/magazine/archive/n6y2022/7746.
2. Zyateva O. A. Inzhenernyj vestnik Dona. 2022. № 5. URL: ivdon.ru/ru/magazine/archive/n5y2022/7670.
3. Rishi T. Proceedings of the Jepson Undergraduate Conference on International Economics, 2022, Vol. 4, Article 4. URL: scholarworks.uni.edu/jucie/vol4/iss1/4.
4. Lee S. Y., Lee J. J., Lee H. Frontiers in Public Health, 2022, Vol. 10. URL: frontiersin.org/journals/public-health/articles/10.3389/fpubh.2022.1024751/full.
5. Gurova L. P. Alleya nauki. 2020. T. 1. № 7(46). pp. 278-282.
6. Ivashchuk O. D., Puzyrev N. S., Rodionov A. Yu. Informatsionnye sistemy i tekhnologii. 2023. № 6(140). pp. 47-52.
7. Zhizhin K. S., Blagorodova N. V. Mezhdunarodnyj zhurnal eksperimental'nogo obrazovaniya. 2016. № 5-3. pp. 337-338.
8. Gordienko A. S., Lozinskaya A. M., Teterina D. V., Shenkman E. A. Izvestiya Rossiyskoy akademii nauk. Energetika. 2019. № 1. pp. 15-27.
9. Knyazeva S. V., Nikitina A. D., Belova E. I., Plotnikova A. S., Podol'skaya E. S., Kovganko K. A. Lesovedenie. 2021. № 6. pp. 645-672.



10. Sindeeva L. V., Medvedeva N. N., Nikolaev V. G., Strelkovich N. N., Orlova I. I. Vestnik novykh meditsinskikh tekhnologiy. 2013. T. 20, № 2. pp. 216-219.

11. Ivanova N. K., Lebedeva S. A., Noskov S. I. Informatsionnye tekhnologii i problemy matematicheskogo modelirovaniya slozhnykh sistem. 2016. № 17. pp. 107-110.

12. Noskov S. I. Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki. 2021. № 6. pp. 74-78.

13. Noskov S. I., Bychkov Yu. A. Yuzhno-Sibirskiy nauchnyj vestnik. 2023. № 1. pp. 37-42.

14. Noskov S. I., Bychkov Yu. A. Vestnik Voronezhskogo gosudarstvennogo tekhnicheskogo universiteta. 2022. T. 18, № 2. pp. 7-12.

15. PAO «Severstal'»: ofitsial'nyj sayt. [PJSC Severstal: official website]. URL: severstal.com/rus/ir/indicators-reporting/operational-results.

16. Noskov S. I., Bychkov Yu. A. Svidetel'stvo o gosudarstvennoy registratsii programmy dlya EVM № 2023619993, 2023. [A program for measuring the parameters of a piecewise linear regression model based on the application of the continuous form of the method of higher consistency]. URL: elibrary.ru/item.asp?id=53820507.

Дата поступления: 19.02.2023

Дата публикации: 21.03.2024