

К.С. Перфильева¹, А.А. Хоняков¹

¹Иркутский государственный университет путей сообщения, г. Иркутск, Российская Федерация

МОДЕЛИРОВАНИЕ ОБЪЕМА ПОГРУЗКИ НА ЖЕЛЕЗНОДОРОЖНОМ ТРАНСПОРТЕ

Аннотация. В статье описан процесс построения модели погрузки на железнодорожном транспорте РФ. Используются кусочно-линейные регрессионные соотношения с применением метода смешанного оценивания параметров. Производится сравнение полученных моделей.

Ключевые слова: линейная регрессия, кусочно-линейная регрессия, метод наименьших модулей, булевы переменные, задача частично булевого линейного программирования, метод смешанного оценивания, антиробастное оценивание.

К.С. Perfileva¹, А.А. Khonyakov¹

¹Irkutsk State Transport University, Irkutsk, Russia

MODELING OF LOADING VOLUME ON THE RAILWAY TRANSPORT

Annotation. The article considers an example of modeling the volume of loading of the main types of cargo on the railway transport of the Russian Federation. The simulation is performed in the form of a piecewise linear regression model, and a linear model is constructed using the method of mixed estimation of the parameters of regression equations. The obtained models are compared.

Keywords: linear regression, piecewise linear regression, method of least modules, BU-Lev the variables, the linear programming problem, the task of partial Boolean linear programming, mixed-method evaluation, antirebate evaluation.

Введение

Железнодорожный транспорт является важнейшей отраслью страны, играет ключевую роль в ее успешном развитии. От слаженной работы железнодорожного транспорта зависит экономическая и технологическая эффективность и функционирование отраслей промышленности, сельского хозяйства, деятельность всех структур с различной формой собственности. В конечном счете, транспорт обеспечивает жизнеспособность и жизнедеятельность общества, государства и его экономические отношения, а также взаимодействия с транспортом и народным хозяйством стран ближнего и дальнего зарубежья. Для исследователя железнодорожный транспорт представляет собой исключительно сложную многоуровневую иерархическую неоднородную систему, управление которой является уникальной проблемой. Грузооборот является обобщающим показателем, планируемым на всех уровнях и используется для определения потребности в подвижном составе и ремонтной базе, затратах труда, топлива, электроэнергии и т.д. В свою очередь, эффективность грузовой работы на железнодорожном транспорте в значительной мере зависит от его эксплуатационных параметров.

Регрессионный анализ [1-3, 5, 13] позволяет эффективно выполнять построение статистических моделей в различных областях знаний. Его суть состоит в определении аналитической формы связи, в которой изменение результивного признака обусловлено влиянием одного или нескольких факторных признаков.

Целью данной работы является моделирование объема погрузки основных видов грузов на железнодорожном транспорте с помощью метода смешанного оценивания параметров линейной регрессии (МСО) с использованием кусочно-линейных конструкций.

Для моделирования были использованы данные за 2005 – 2018 г.г. (таблица 1) на основе официальной статистики [12] по следующим показателям:

- y - объем погрузки основных видов грузов на железнодорожном транспорте, тыс. т;

- x_1 - протяженность путей сообщения (на конец года), тыс. км;
- x_2 - добыча топливно-энергетических полезных ископаемых (уголь), млн. т;
- x_3 - рабочий парк груженых железнодорожных вагонов (в среднем в сутки), тыс. штук;
- x_4 - вывозка древесины, млн. плотн.м³.

Таблица 1. Исходные данные

Год	y	x_1	x_2	x_3	x_4
2005	1273056,00	85,2	299	270	113,2
2006	1311068,99	85,2	310	276	117,6
2007	1344239,00	85,2	314	289	134,2
2008	1303740,00	85,6	329	287	108,2
2009	1108159,00	85,6	301	221	97,1
2010	1205769,00	85,7	322	328	117
2011	1241541,00	85,5	336	393	123
2012	1271855,00	85,6	357	393	122
2013	1226932,00	85,6	353	386	120
2014	1226932,00	86,3	357	420	123
2015	1217948,60	86,3	372	388	127
2016	1226951,09	86,4	386	389	137
2017	1266482,09	86,5	410	417	140
2018	1291504,70	86,6	439	461	149

Моделирование объема погрузки основных видов грузов на железнодорожном транспорте с помощью кусочно-линейной регрессионной модели

В работе [9] была рассмотрена кусочно-линейная регрессионная модель вида (1), представляющая собой симбиоз двух кусочно-линейных моделей с минимальными и максимальными вкладами независимых переменных [5-8]:

$$y_k = \min\{\alpha_1 x_{k1}, \alpha_2 x_{k2}, \dots, \alpha_m x_{km}\} + \max\{\beta_1 x_{k1}, \beta_2 x_{k2}, \dots, \beta_m x_{km}\} + \varepsilon_k, \quad k = \overline{1, n}. \quad (1)$$

Отыскание параметров такой модели сводится к задаче частично булевого линейного программирования (2-8):

$$z_{1k} + z_{2k} + u_k - v_k = y_k, \quad k = \overline{1, n}, \quad (2)$$

$$z_{1k} \leq \alpha_i x_{ki}, \quad k = \overline{1, n}, \quad i = \overline{1, m}, \quad (3)$$

$$z_{2k} \geq \beta_i x_{ki}, \quad k = \overline{1, n}, \quad i = \overline{1, m}, \quad (4)$$

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

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

$$\beta_i x_{ki} - z_{2k} \geq (-1 + \nu_{ki})M, \quad k = \overline{1, n}, \quad i = \overline{1, m}, \quad (7)$$

$$\sum_{i=1}^m \nu_{ki} = 1, \quad k = \overline{1, n}, \quad (8)$$

где M - заранее выбранное большое положительное число, а σ_{ki} и ν_{ki} - булевы переменные, $k = \overline{1, n}$, $i = \overline{1, m}$.

Для моделирования объема погрузки в модель (1) добавлен свободный член α_0 :

$$y_k = \alpha_0 + \min\{\alpha_1 x_{k1}, \alpha_2 x_{k2}, \dots, \alpha_m x_{km}\} + \max\{\beta_1 x_{k1}, \beta_2 x_{k2}, \dots, \beta_m x_{km}\} + \varepsilon_k, \quad k = \overline{1, n}. \quad (9)$$

При помощи специализированного программного комплекса [14] на основе исходных данных была получена кусочно-линейная регрессионная модель вида:

$$y = 655911,232 + \min\{3355,804 x_1; 863,104 x_2; 847,316 x_3; 2393,807 x_4\} + \max\{3355,804 x_1; 889,778 x_2; 847,316 x_3; 2569,853 x_4\}. \quad (10)$$

Средняя относительная ошибка аппроксимации для полученной модели составила 4,56%, что указывает на ее хорошее качество.

Для анализа влияния каждой независимой переменной построим вектора:

$$Ind^{\min} = (ind^{\min}_1, \dots, ind^{\min}_n), Ind^{\max} = (ind^{\max}_1, \dots, ind^{\max}_n) \quad (11)$$

где ind_k - номер переменной, значение которой с учетом параметра является наименьшим (наибольшим) в k -ом наблюдении выборки.

Для полученной модели эти вектора имеют вид:

$$Ind^{\min} = (3, 3, 3, 3, 3, 2, 1, 1, 1, 1, 1, 1, 1, 1), \quad (12)$$

$$Ind^{\max} = (4, 4, 4, 2, 1, 4, 3, 3, 3, 3, 2, 4, 2, 3). \quad (13)$$

Таким образом, вклад в функцию минимума не внесла переменная x_4 , в функцию максимума внесли свой вклад все переменные.

Моделирование объема погрузки основных видов грузов на железнодорожном транспорте с помощью МСО

Рассмотрим линейное регрессионное уравнение:

$$y_k = \sum_{i=1}^m \alpha_i x_{ki} + \varepsilon_k, k = \overline{1, n}, \quad (14)$$

где y - зависимая, а x_i - i -ая независимая переменная; α_i - i -ый подлежащий оцениванию параметр; ε - ошибки аппроксимации, k - номер наблюдения, n - их число.

Представим уравнение (14) в матричной форме:

$$y = X\alpha + \varepsilon, \quad (15)$$

где $y = (y_1, \dots, y_n)^T$, $a = (a_1, \dots, a_m)^T$, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$, $X = (n \times m)$ - матрица с компонентами x_{ki} .

Обширный класс методов оценивания параметров уравнения (15) связан с поиском L_ν - оценок посредством минимизации функций потерь вида [10]:

$$J_\nu^P(\alpha) = \sum_{k \in P} |\varepsilon_k|^\nu, P = \{1, 2, \dots, n\} \quad (16)$$

Такие функции характеризуются разной реакцией на наблюдения, не согласующиеся со всей выборкой. При этом, чем больше значение ν , тем сильнее L_ν - оценка реагирует на такие наблюдения. Методы оценивания, слабо реагирующие на подобные наблюдения, в регрессионном анализе называют устойчивыми, или робастными. При $\nu = 2$ это метод наименьших квадратов (МНК), при $\nu = 1$ - метод наименьших модулей (МНМ), при $\nu \rightarrow \infty$ - метод антиробастного оценивания (МАО).

В работах [8,10] предложен метод смешанного оценивания параметров линейной регрессии. Особенность этого метода состоит в поиске вектора параметров линейной регрессии с помощью минимизации суммы разных функций потерь на разных участках выборки.

Для реализации МСО необходимо разделить исходную выборку на две части, N_1 и N_2 , при этом разбиение должно удовлетворять правилу «длина выборки больше числа факторов». Сделаем это следующим образом:

$$N_1 = \{1, 2, 3, 4, 5, 6\}, N_2 = \{7, 8, 9, 10, 11, 12, 13, 14\}. \quad (17)$$

С помощью специализированного программного обеспечения [11] получена следующая линейная регрессионная модель:

$$y = 7135492,1 - 73029,85 x_1 + 20,981 x_2 - 170,544 x_3 + 3528,982 x_4. \quad (18)$$

Из полученной модели следует, что увеличение длины путей и количества вагонов вызывает уменьшение объёмов погрузки основных видов грузов на железнодорожном транспорте. Добыча топливно-энергетических полезных ископаемых и вывоз древесины, напро-

тив, ведут к увеличению объёмов погрузки. Первая часть этого вывода противоречит смыслу переменных, поэтому используем прием преобразования регрессоров с тем, чтобы добиться соответствия модели этому смыслу. Проведем «конкурс» моделей, состоящий в построении множества их альтернативных вариантов и выборе наиболее приемлемого. Таковым стало регрессионное уравнение:

$$y = 749660,33 + 1028,477 x_2 + 251,585 x_4 + 0,014 \frac{x_1^2 x_3^2}{t}, \quad (19)$$

где t – время (номер года, $t = 5$ для 2005 г).

Теперь в модели объема погрузки основных видов грузов на железнодорожном транспорте нет противоречий. Увеличение каждого из факторов влечет рост объема погрузки.

Значение средней ошибки аппроксимации уравнения (19) равно 2,14%.

Заключение

В данной работе было выполнено моделирование объема погрузки основных видов грузов на железнодорожном транспорте с помощью метода смешанного оценивания параметров линейной регрессии (МСО) и с использованием кусочно-линейных конструкций. Анализ характера вхождения независимых переменных в правые части уравнений для каждой из полученных моделей позволяет сделать вывод об их соответствии смыслу переменных. Каждая из моделей может быть использована для анализа общих закономерностей во взаимовлиянии выделенных переменных.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. – М.: Юнити, 1998. – 1022 с.
2. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Книга 1. В 2-х кн. – М.: Финансы и статистика, 1986. – 366 с.
3. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Книга 2. В 2-х кн. – М.: Финансы и статистика, 1986. – 351 с.
4. Доугерти К. Введение в эконометрику. – М.: ИНФРА-М, 2009. – 465 с.
5. Ильина Н.К., Лебедева С.А., Носков С.И. Идентификация параметров некоторых негладких регрессий//Информационные технологии и проблемы математического моделирования сложных систем. – 2016. – № 17. – С. 111.
6. Носков С.И., Лоншаков Р.В. Идентификация параметров кусочно-линейной регрессии//Информационные технологии и проблемы математического моделирования сложных систем. – 2008. – № 6. – С. 63-64.
7. Носков С.И. Идентификация параметров кусочно-линейной функции риска// Транспортная инфраструктура Сибирского региона. – 2017. – Т. 1. – С. 417-421.
8. Носков С.И. Технология моделирования объектов с нестабильным функционированием и неопределенностью в данных. – Иркутск: Облформпечать. – 1996. – 320 с.
9. Носков С.И., Хоняков А.А. Программный комплекс построения некоторых типов кусочно-линейных регрессий // «Информационные технологии и математическое моделирование в управлении сложными системами»: электрон.науч. журн. – 2019. – №3. – С. 47-55 – Режим доступа: <http://ismm-irgups.ru/toma/34-2019>, свободный. – Загл. с экрана. – Яз. рус., англ. (дата обращения: 20.11.2019)
10. Носков С.И. О методе смешанного оценивания параметров линейной регрессии// Информационные технологии и математическое моделирование в управлении сложными системами. – 2019. – №1. – С. 14-20.
11. Перфильева К.С., О методе смешанного оценивания параметров линейной регрессии // «Информационные технологии и математическое моделирование в управлении сложными системами»: электрон.науч. журн. – 2019. – №4. Режим доступа: <http://ismm-irgups.ru/toma/11-2019>, свободный. – Загл. с экрана. – Яз. рус., англ. (дата обращения: 20.11.2019)

12. Сайт Федеральной службы государственной статистики. - <https://www.gks.ru/>

13. Себер Дж. Линейный регрессионный анализ. – М.: Издательство «Мир», 1980. – 456

с.

14. Свидетельство о государственной регистрации программ для ЭВМ № 2019665873 Программный комплекс построения кусочно-линейных регрессий/ С.И.Носков, Хоняков А.А., Н.И. Глухов(Россия); Правообладатель: Федеральное государственное бюджетное образовательное учреждение высшего образования «Иркутский государственный университет путей сообщения» (ФГБОУ ВО ИРГУПС); заявка № 2019664699 19.11.19; дата регистр. 02.12.2019

REFERENCES

1. Ayvazyan S. A., Mkhitaryan V. S. Applied statistics and fundamentals of econometrics, Moscow: unity, 1998, 1022 p.

2. Draper N., Smith G. Applied regression analysis. Book 1. In 2 books-Moscow: Finance and statistics, 1986. - 366 p.

3. Draper N., Smith G. Applied regression analysis. Book 2. In 2 books-Moscow: Finance and statistics, 1986. - 351 p.

4. Dougherty K. Introduction to econometrics, Moscow: INFRA-M, 2009, 465 p.

5. Ilina N. K., Lebedeva S. A., Noskov S. I. Identification of parameters of some non-smooth regressions//Information technologies and problems of mathematical modeling of complex systems. - 2016. - № 17. - P. 111.

6. Noskov S. I., Lonshakov R. V. Identification of parameters of piecewise linear regression//Information technologies and problems of mathematical modeling of complex systems, 2008, no. 6, Pp. 63-64.

7. Noskov S. I. Identification of parameters of piecewise linear risk function// Transport infrastructure of the Siberian region, 2017, Vol. 1, Pp. 417-421.

8. Noskov S. I. Technology of modeling objects with unstable functioning and uncertainty in data. – Irkutsk: Allinternet. - 1996. - 320 p.

9. Noskov S. I., Khonyakov A. A. Software package for constructing some types of piecewise linear regressions // "Information technologies and mathematical modeling in the management of complex systems": electron. nauch. zhurn – - 2019. - №3. - Pp. 47-55-access Mode: <http://ismm-irgups.ru/toma/34-2019>, free. - Blank from the screen. – Yaz. Rus., eng. (accessed: 20.11.2019)

10. Noskov S. I. On the method of mixed estimation of linear regression parameters// Information technologies and mathematical modeling in the management of complex systems. - 2019. - №1. - P. 14-20.

11. Perfilova K. S., on the method of mixed estimation of linear regression parameters// Information technologies and mathematical modeling in the management of complex systems": electron. scientific journal-2019. - №4 - access Mode: <http://ismm-irgups.ru/toma/11-2019>, free. - Blank from the screen. - Yaz. Rus., eng.(accessed: 20.11.2019)

12. Website of the Federal state statistics service. - <https://www.gks.ru/>

13. Seber J. Linear regression analysis. - Moscow: Mir publishing House, 1980. - 456 p.

14. Certificate of state registration of computer programs no. 2019665873 Software complex for constructing piecewise linear regressions / S. I. Noskov, A. A. Honyakov, N. I. Glukhov(Russia); Copyright holder: Federal state budgetary educational institution of higher education "Irkutsk state University of railway transport" (FGBOU VO IRGUPS); application no. 2019664699 19.11.19; date register. 02.12.2019

Информация об авторах

Карина Сергеевна Перфильева – аспирант, кафедра «Информационные системы и защита информации», Иркутский государственный университет путей сообщения, г. Иркутск, e-mail: 552649-171233@mail.ru

Антон Андреевич Хоняков - аспирант, кафедра «Информационные системы и защита информации», Иркутский государственный университет путей сообщения, г. Иркутск, e-mail: anton_khonyakov@mail.ru

Authors

Karina Sergeevna Perfilieva - Postgraduate Student, «Information systems and information security», Irkutsk State Transport University, Irkutsk, e-mail: 552649-171233@mail.ru

Anton Andreyevich Khonyakov, Postgraduate Student, “Information systems and information security”, Irkutsk State Transport University, Irkutsk, e-mail: anton_khonyakov@mail.ru

Для цитирования

Перфильева К.С., Хоняков А.А. Моделирование объёма погрузки на железнодорожном транспорте [Электронный ресурс] // Молодая наука Сибири: электрон. науч. журн. – 2020. – №3(9). – Режим доступа: <http://mnv.irkups.ru/toma/39-2020>, свободный. – Загл. с экрана. – Яз. рус., англ. (дата обращения: 26.09.2020)

For citation

Perfileva K.S., Khonyakov A.A. Modeling of loading volume on the railway transport. *Molodaya nauka Sibiri: ehlektronnyj nauchnyj zhurnal* [Young science of Siberia: electronic scientific journal], 2020, no. 3(9). [Accessed 26/09/20] (in Russian)