

Т. В. Богачёв, Т. В. Алексейчик, И. А. Харитонов

**МЕТОДИКА АНАЛИЗА И ПРОГНОЗИРОВАНИЯ
ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ РЕГИОНАЛЬНЫХ ТРАНСПОРТНЫХ
СИСТЕМ МЕТОДОМ НЕЧЕТКОЙ ЛИНЕЙНОЙ РЕГРЕССИИ**

Аннотация

Предложена методика анализа объемов грузовых автомобильных перевозок в регионах Южного федерального округа (ЮФО) методом нечеткой линейной регрессии, которая позволяет определить вес соответствующего фактора для исследуемого показателя каждого региона округа.

Ключевые слова

Анализ транспортных систем, нечеткая линейная регрессия, порог надежности, задача линейного программирования.

T. V. Bogachiov, T. V. Alekseychik, I. A. Kharitonov

METHODS OF ANALYSIS AND FORECASTING OF ECONOMIC INDICATORS OF REGIONAL TRANSPORT SYSTEMS USING FUZZY LINEAR REGRESSION

Annotation

Method is proposed for analyzing the volume of road freight traffic in regions of Southern Federal Region (SFR) using fuzzy linear regression, which allows determining the weight of corresponding factor for the studied indicator of each region of SFR.

Keywords

Analysis of transportation systems, fuzzy linear regression, reliability threshold, linear programming problem.

Введение

В задачах, связанных с реальной жизнью в целом и экономикой в частности, данные поступают в недостаточном объеме или в искаженном виде, и, следовательно, содержат в себе неопределенность. При моделировании различных экономических процессов неопределенность может быть не только в анализируемых данных, но часто имеет место неопределенность, связанная с неучтенными факторами. В данной статье для анализа и прогнозирования объемов грузовых автомобильных перевозок в шести регионах Южного федерального округа (ЮФО) предлагается применить метод нечеткого моделирования с реализацией на языке программирования Python.

Южный федеральный округ выполняет значительную роль в экономическом развитии страны, что во многом определяется его географическим положением. Это обуславливает его значение в транспортной инфраструктуре и ресурсно-сырьевой базе Российской Федерации и доказывает актуальность темы исследования.

Рассмотрены следующие регионы ЮФО: Краснодарский край, Волгоградскую область, Ростовскую область, Астраханскую область, Республику Адыгея, Республику Калмыкия.

Материалы и методы

Для оценки состояния грузовых автомобильных перевозок в выбранных

регионах методом нечеткого моделирования в зависимости от факторов: плотность автомобильных дорог общего пользования с твердым покрытием, валовой региональный продукт на душу населения, индексы тарифов на грузовые перевозки, введем обозначения:

Y — объем перевозок грузов автомобильным транспортом организаций всех видов деятельности (млн т);

X_1 — плотность автодорог общего пользования с твердым покрытием (км дорог на 1000 кв. км территории);

X_2 — валовой региональный продукт на душу населения (тыс. руб.);

X_3 — индексы тарифов на грузовые перевозки (декабрь к декабрю предыдущего года, в %).

Данные по основным показателям развития транспортных систем регионов Южного федерального округа за 1996–2017 гг. взяты из [1].

Построим уравнение зависимости объема грузовых автомобильных перевозок от факторов X_1 , X_2 , X_3 методом нечеткой линейной регрессии, в котором обычно используются симметричные треугольные нечеткие числа. Отметим, что треугольное нечеткое число A задается тройкой чисел a_m , a , a_M , где

$$a_m \leq a \leq a_M. \quad (1)$$

При условии строгого неравенства (1) нечеткое число A имеет следующую функцию принадлежности:

$$\mu_A(x) = \begin{cases} 0, & x \notin [a_m; a_M]; \\ \frac{x - a_m}{a - a_m}, & a_m \leq x \leq a; \\ \frac{a_M - x}{a_M - a}, & a \leq x \leq a_M. \end{cases}$$

Число $A = \langle a_m, a, a_M \rangle$ называется симметричным, если выполнено условие $a - a_m = a_M - a = r$, где $r = 0$.

Сформулируем задачу нечеткой линейной регрессии следующим образом. Для данных k результатов наблюдений зависимой переменной y_j от n факторов

$$x_i = (x_{i1}, x_{i2}, \dots, x_{ik})^T, ,$$

где $i=1, \dots, n$, нужно найти нечеткие коэффициенты $A_0, A_1, \dots, A_n$ такие, чтобы выполнялись условия $\mu_j(Y_j) \geq h$, где $\mu_j(Y_j)$ — функция принадлежности нечеткого множества

$$Y_j = A_0 + A_1 x_{1j} + \dots + A_n x_{nj},$$

величина h — заданный порог надежности, причем неопределенность, связанная с этими коэффициентами, была бы минимальной. Пусть параметрами модели являются числа $A_i = \langle a_i - r_i, a_i, a_i + r_i \rangle$, где $a_i \in R$ и $r_i \geq 0$. В этом случае Y_j также является треугольным нечетким числом. В соответствии с [2, 3, 4] для нахождения a_i и r_i получаем задачу линейного программирования:

$$f = ka_0 + \sum_{j=1}^k \sum_{i=1}^n r_i x_{ij} \rightarrow \min \quad (2)$$

$$\begin{cases} y_j \geq a_0 + \sum_{i=1}^n a_i x_{ij} - (1-h) \left(r_0 + \sum_{i=1}^n r_i x_{ij} \right), \\ y_j \leq a_0 + \sum_{i=1}^n a_i x_{ij} + (1-h) \left(r_0 + \sum_{i=1}^n r_i x_{ij} \right), \\ r_i \geq 0, j=1 \dots k. \end{cases}$$

Для каждого исследуемого региона будем строить модели при пороге надежности, например, $h=0,4$, используя значения выбранных показателей за 1996–2017 гг., причем значения показателей последних четырех лет в дальнейшем используются для проверки

построенной модели на качество прогноза. Тогда в соответствии с (2) для нахождения a_i и r_i получаем задачу линейного программирования

$$f = 18a_0 + \sum_{j=1}^{18} \sum_{i=1}^3 r_i x_{ij} \rightarrow \min \quad (3)$$

$$\begin{cases} y_j \geq a_0 + \sum_{i=1}^3 a_i x_{ij} - (1-h) \left(r_0 + \sum_{i=1}^3 r_i x_{ij} \right), \\ y_j \leq a_0 + \sum_{i=1}^3 a_i x_{ij} + (1-h) \left(r_0 + \sum_{i=1}^3 r_i x_{ij} \right), \\ r_i \geq 0, j=1 \dots 18. \end{cases}$$

Данная задача линейного программирования решается для каждого региона ЮФО. Реализация решения задачи осуществлена на языке программирования Python с использованием библиотеки *cvxopt* [5] для самой оптимизации, библиотеки *pandas* [6] для вывода результатов в формате табличного

процессора Excel и библиотеки *numpy* [7] для проведения необходимых математических преобразований.

Результаты

Проиллюстрируем решение задачи (3) для Краснодарского края. Уравнение нечеткой регрессии для Краснодарского края имеет вид:

$$Y = 26,6072 + 0,09712X_1 - 0,00019X_2 + \langle 0,4871; 0,8921; 1,2971 \rangle X_3.$$

Нечетким параметром модели регрессии является только коэффициент при показателе индексы тарифов на грузовые перевозки (X_3), причем он имеет и больший вес в полученной регрессии. Учитывая, что $Y = \langle Y_d, Y_m, Y_u \rangle$, постро-

им график нечеткой линейной регрессии, который отражает нечеткий прогноз для исследуемого показателя перевозки грузов автомобильным транспортом в данном регионе (рис. 1).

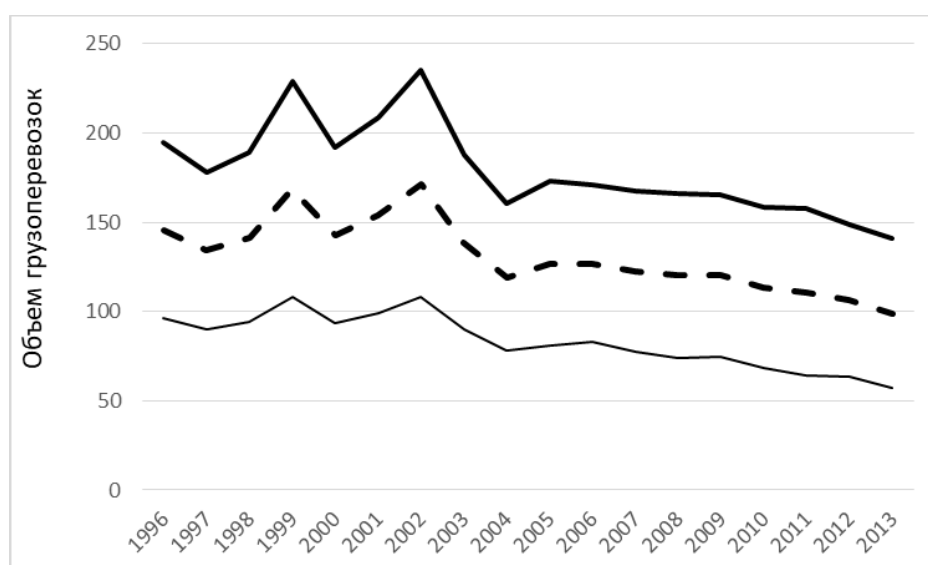


Рисунок 1 — График нечеткой линейной регрессии для показателя объема перевозки грузов автомобильным транспортом в Краснодарском крае организациями всех видов деятельности (Y , млн т), $h = 0,4$

Оценка результатов исследования влияния факторов X_1 , X_2 , X_3 на объем перевозок грузов автомобильным транспортом организаций всех видов

деятельности в Краснодарском крае выполнена с помощью контролирующей выборки за период 2014–2017 гг. (табл. 1).

Таблица 1 — Оценка результатов исследования Краснодарского края с помощью контролирующей выборки для модели с порогом надежности 0,4

Год	Ограничение слева	Ограничение справа	Y контролирующее	Принадлежность Y интервалу
2014	172,8605366	67,15986584	88,3	+
2015	152,3822849	57,05714003	84,6	+
2016	119,4483523	43,73943627	78,8	+
2017	133,52843	45,25540042	69,4	+

Как видно из данных таблицы 1, уравнение нечеткой линейной регрессии для Краснодарского края построено достаточно корректно и может быть использовано для прогноза.

$$Y = 7,80331 - 0,299235X_1 - 0,00017X_2 + \langle -0,06376101; 0,080601; 0,23356301 \rangle X_3.$$

В этом уравнении параметры при показателях X_1 и X_2 , а также свободный член являются четкими числами, а нечетким параметром модели регрессии является только коэффициент при показателе индексы тарифов на грузовые перевозки (X_3). Отметим, что больший вес

$$Y = \langle 123,691; 142,47; 161,249 \rangle - \langle 0,42252; 0,27852; 0,13452 \rangle X_1 - 0,00013X_2 - 0,0223X_3.$$

В этом уравнении свободный член и параметр при показателе X_1 являются нечеткими числами, а параметры при показателях X_2 и X_3 — четкие числа, причем больший вес в полученной регрессии имеет также показатель плот-

Представим решения задачи (3) для Волгоградской, Ростовской и Астраханской областей.

Для Волгоградской области уравнение нечеткой регрессии при $h = 0,4$ имеет вид:

в полученной регрессии имеет показатель плотность автомобильных дорог общего пользования с твердым покрытием (км дорог на 1000 кв. км территории).

Уравнение нечеткой регрессии при $h = 0,4$ для Ростовской области выглядит следующим образом:

ность автомобильных дорог общего пользования с твердым покрытием.

Для Астраханской области уравнение нечеткой регрессии при $h = 0,4$ запишем в виде:

$$Y = \langle 22,07; 24,508; 26,946 \rangle - 0,2768X_1 + \langle -0,01325; 0,001; 0,01526 \rangle X_3.$$

Больший вес в этом уравнении имеет показатель плотность автомобильных дорог общего пользования с твердым покрытием с четким параметром.

Полученные уравнения по аналогии с моделью для Краснодарского края позволяют построить нечеткий прогноз для показателя объема перевозки грузов автомобильным транспортом для каждой из трех рассмотренных областей.

Как и в таблице 1, оценка результатов исследования влияния факторов

X_1 , X_2 , X_3 на объем перевозок грузов автомобильным транспортом организаций всех видов деятельности для каждой области выполнена с помощью соответствующих контролирующих выборок. Полученные оценки подтверждают корректность построенных моделей и их применимость для прогнозирования.

Уравнение нечеткой линейной регрессии при $h = 0,4$ для Республики Адыгея:

$$Y = \langle -7,082; 24,508; 56,098 \rangle - 0,00256X_1 + 0,00259X_3.$$

Обратим внимание на то, что коэффициент при переменной X_2 равен нулю, а нечетким является только свободный

член. Если уменьшить порог надежности до величины $h = 0,2$, то уравнение нечеткой регрессии примет вид:

$$Y = \langle 0,816; 24,508; 48,2 \rangle + 0,00256X_1 + 0,0229X_3.$$

Эта модель является более адекватной, но имеет достаточно низкий

порог надежности. Так как только свободный член является нечетким, необ-

ходим более глубокий анализ факторов, влияющих на исследуемый показатель.

Для Республики Калмыкия уравнения нечеткой регрессии при $h = 0,2$ имеет вид:

$$Y = 14,565 - 0,14301X_1 + \langle -0,05535; -0,04275; -0,03015 \rangle X_3.$$

Так как нижняя граница построенного графика (отражающая пессимистический прогноз исследуемого показателя) (рис. 2) частично совпадает с осью абсцисс (в силу невозможности

принятия данным показателем отрицательных значений), то данная модель не применима для республики Калмыкия. Необходимо проанализировать факторы, входящие в модель.

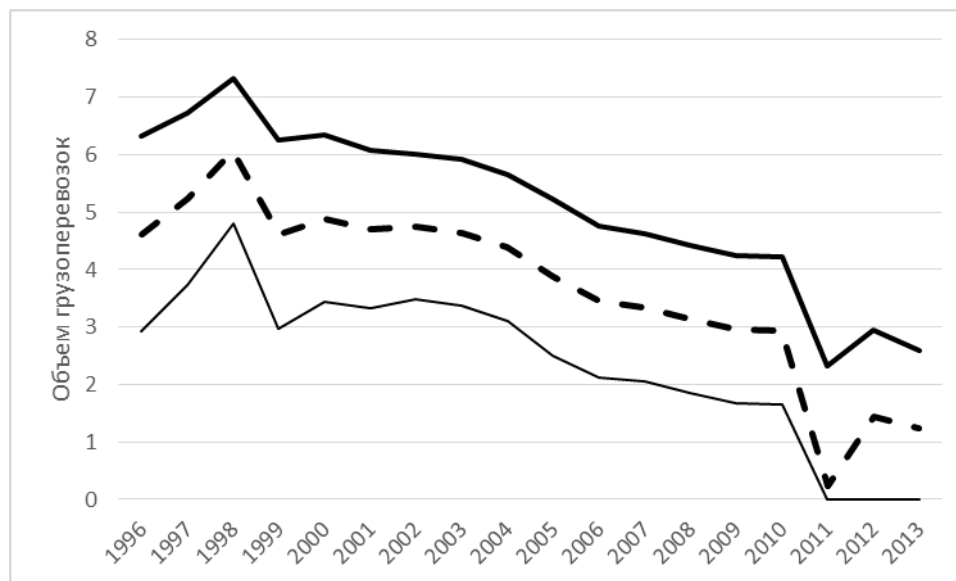


Рисунок 2 — График нечеткой линейной регрессии для показателя перевозки грузов автомобильным транспортом в республике Калмыкия организациями всех видов деятельности (Y , млн т), $h = 0,2$

Выводы

Проверка построенных нечетких регрессий для Волгоградской, Ростовской и Астраханской областей, а также для Краснодарского края подтвердила их адекватность, в том числе и с помощью контрольных выборок за последние 4 года. Необходимо отметить, что для Волгоградской области и Краснодарского края нечетким параметром в уравнении регрессии является параметр только для показателя индекса тарифов на грузовые перевозки, а параметры для плотности автомобильных дорог общего пользования с твердым покрытием и валового регионального продукта на душу населения, а также свободный

член являются четкими числами. Для Ростовской области нечетким параметром является только параметр при показателе плотности автомобильных дорог общего пользования с твердым покрытием, а также свободный член. Для Астраханской области нечетким параметром является параметр при показателе индекса тарифов на грузовые перевозки, а также свободный член.

Модель для Республики Адыгея может применяться при малых значениях порога надежности, причем нечетким параметром является только свободный член.

Построенная модель для Республики Калмыкия не является адекватной,

поэтому ее нельзя использовать для анализа и прогнозирования показателя объем перевозок грузов автомобильным транспортом. В таких случаях необходим более глубокий анализ факторов, влияющих на исследуемый показатель.

В работе проведен анализ объема грузовых автомобильных перевозок в шести регионах ЮФО. Построенные с помощью нечеткой линейной регрессии модели позволили исследовать зависимость объема перевозки грузов автомобильным транспортом в этих регионах от выбранных факторов в пяти регионах из шести. В результате анализа можно предложить следующий алгоритм анализа и прогнозирования экономических показателей с использованием метода нечеткой линейной регрессии.

1. Ввод данных для анализа экономических показателей за исследуемый период.

2. По каждому из исследуемых регионов решается соответствующая задача линейного программирования (3) с использованием программы на языке Python.

3. Анализ построенного уравнения нечеткой регрессии с выбранным порогом надежности.

4. Оценка результатов исследования с помощью контролирующей выборки для построенной модели.

5. Выявление на основе уравнения нечеткой регрессии степени влияния факторов на исследуемый показатель.

6. Разработка рекомендаций для улучшения деятельности исследуемого процесса.

Предложенный в работе метод целесообразно использовать для анализа и прогнозирования значений различных экономических показателей. Созданная программа на языке Python позволяет анализировать показатели различных отраслей с визуальным представлением данных не только региональной экономики.

Библиографический список

1. Регионы России. Социально-экономические показатели : стат. сб. [Электронный ресурс]. — Режим доступа : <https://www.gks.ru>.

2. Волкова, Е. С., Гисин, В. Б. Нечеткая линейная регрессия в модели роста технологических знаний // Вестник Финансового университета. — 2015. — № 5 (89). — С. 97–104.

3. Tanaka, H., Uejima, S., Asai, K. Linear regression analysis with fuzzy model // IEEE Transactions on Systems, Man and Cybernetics. — 1982. — № 12 (6). — P. 903–907.

4. Bogachiov, T., Alekseychik, T., Pushkar, O. Analysis of indicators of state of regional freight traffic by method of fuzzy linear regression // Computing with Words and Perceptions : 10th International Conf. on Theory and Application of Soft Computing. — Prague, 2019. — Vol. 1095. — P. 632–638.

5. Документация по библиотеке cvxopt [Электронный ресурс]. — Режим доступа : <http://cvxopt.org>.

6. Документация по библиотеке pandas. [Электронный ресурс]. — Режим доступа : <https://pandas.pydata.org>.

7. Документация по библиотеке numpy. [Электронный ресурс]. — Режим доступа : <https://docs.scipy.org>.

Bibliographic list

1. Region of Russia. Socio-economic indicators : stat. comp. [Electronic resource]. — Mode of access : <https://www.gks.ru>.

2. Volkova, E. S., Gisin V. B. Fuzzy linear regression in growth model of technological knowledge // Bulletin of Financial University. — 2015. — № 5 (89). — P. 97–104.

3. Tanaka, H., Uejima, S., Asai, K. Linear regression analysis with fuzzy model // IEEE Transactions on Systems, Man and Cybernetics. — 1982. — № 12 (6). — P. 903–907.

4. *Bogachiov, T., Alekseychik, T., Pushkar, O.* Analysis of indicators of state of regional freight traffic by method of fuzzy linear regression // Computing with Words and Perceptions : 10th International Conference on Theory and Application of Soft Computing. — Prague. — 2019. — Vol. 1095. — P. 632–638.

5. Cvxopt User's Guide [Electronic resource]. — Mode of access : cvxopt.org.

6. Pandas documentation [Electronic resource]. — Mode of access : <https://pandas.pydata.org>.

7. Numpy and Scipy Documentation [Electronic resource]. — Mode of access : <https://docs.scipy.org>.