числом безработных является уровень заработной платы по городу, но и эта связь недостаточно тесная и не обеспечивает статистическую значимость данной регрессионной модели:

Таблица 3

Оценки коэффициентов регрессионной модели и критерии их значимости

	Коэффициенты	Станд. ошибка	t-статистика
Y-пересечение	4815,26899	931,7689767	5,167879
зарплата	-0,0819416	0,064058829	-1,27916

Если доверять модели, описывающей линейную зависимость роста безработных в зависимости от уровня зарплаты, то основной вывод можно было бы сделать следующий: с уменьшением средней зарплаты на 0,08 руб. ряды ищущих достойную работу увеличиваются на 1 человека.

В 2012 – 2015 годах возможно несколько сценариев развития на рынке труда области. Варианты прогноза численности безработных и уровня регистрируемой безработицы зависят в первую очередь от продолжительности экономического кризиса. Вместе с тем необходимость комплексного решения проблем занятости и безработицы обусловлена рядом объективных причин: сложностью и масштабностью задач по созданию благоприятных условий для смягчения ситуации на рынке труда и необходимостью выполнения в рамках единой программы крупных по объему и требующих длительных сроков реализации мероприятий.

Адаптационные возможности рынка труда присутствуют, но ограни-чены, поэтому обоснованное вмешательство государства в процессы структурного формирования, перемещения и профессиональной адаптации трудовых ресурсов представляется неизбежным этапом развития российского рынка труда.

Список источников

1. Социально-экономическое положение Воронежа. Территориальный орган федеральной службы государственной статистики по Воронежской области – январь-декабрь 2006 [текст] / Воронеж – 2007.

2. Социально-экономическое положение Воронежа. Территориальный орган федеральной службы государственной статистики по Воронежской области – январь-декабрь 2008 [текст] / Воронеж – 2009.

3. Социально-экономическое положение Воронежа. Территориальный орган федеральной службы государственной статистики по Воронежской области – январь-декабрь 2010 [текст] / Воронеж – 2011.

ANALYSIS OF MAIN TRENDS OF THE YOUTH LABOR MARKET IN VORONEZH AND ADAPTIVE MODELING ITS INDICATORS

Kontsevaya Natalya Valeryevna,
Ph.D. of Economy, associate professorof the Chair of Economic and Mathematical Methods and Models of the All-Russian Correspondence Financial and Economic Institute; kontsevaya07@list.ru
Zvezdinskaya Anna Vladimirovna,
Degree-seeking student of the Chair of Bases of Labor and Management of Voronezh State University, zav1304@rambler.ru

In the article current state of the youth labor market in Voronezh is analyzed and its main trends are highlighted. The results of modeling and forecasting key market indicators based on adaptive methods and regression models are considered. Substantive interpretation is the results are carried out.
Keywords: analysis of trend, labor market, adaptive methods, Holt-Winters'es model, forecasting.

УДК 519.85

НЕЧЕТКАЯ МНОЖЕСТВЕННАЯ ЛИНЕЙНАЯ РЕГРЕССИОННАЯ МОДЕЛЬ ДЛЯ СИММЕТРИЧНЫХ НЕЧЕТКИХ ЧИСЕЛ L-R-ТИПА

Леденева Татьяна Михайловна,
профессор, доктор технических наук, заведующая кафедрой вычислительной математики и прикладных информационных технологий Воронежского государственного университета;
dean@amm.vsu.ru
Сапкина Наталья Владимировна,
аспирант кафедры вычислительной математики и прикладных информационных технологий Воронежского государ-ственного университета; natashasapkina@yandex.ru

В статье рассмотрена нечеткая множественная линейная регрессионная модель для симметричных нечетких чисел L-R-типа. Приведена оценка нечетких параметров модели методом наименьших квадратов. В заключении полученные теоретические результаты использованы для анализа информации.

Ключевые слова: метод наименьших квадратов, нечеткая множественная линейная регрессионная модель, операции с нечеткими числами, расстояние между двумя нечеткими числами, симметричные нечеткие числа.

Введение

Нечеткий регрессионный анализ – это нечеткий тип классического регрессионного анализа. Он применяется для нахождения функционального взаимодействия между зависимыми и независимыми переменными в нечеткой среде и позволяет решать различные задачи в ситуациях, когда традиционные методы неэффективны или совсем неприменимы из-за отсутствия достаточно точных сведений об исследуемых объектах.

Нечеткая информация принадлежит к данным естественной природы, подобно неценовой информации или данным со ссылкой на неопределенность, не относящейся к случайности. Такой тип данных легко найти в естественном языке, социологии, психометрии, окружающей среде, эконометрике и т. д. Для представления и моделирования нечеткой информации используются нечеткие числа.

Нечеткие числа и операции над ними

Пусть $\Re$ – это одномерное эвклидово пространство с нормой $\|\cdot\|$.

Нечетким числом называется полунепрерывная сверху выпуклая функция $F:\Re \to [0,1]$, где множество $\{x \in \Re \mid F(x)=1\}$ непусто. Другими словами, нечеткое число A определяется как выпуклое нормированное нечеткое множество вещественной прямой $\Re$, так что существует единственное $x_0 \in \Re$, для которого $F(x_0)=1$, причем функция принадлежности $F(x)$ является кусочно-непрерывной [3].

Определение 1.

Пусть L (а также R) – убывающая функция формы, действующая из $\Re^+$ в $[0,1]$, такая, что $L(0)=1$; $L(x)<1$ для всех $x>0$; $L(x)>0$ для всех $x<1$; $L(1)=0$ для всех x и $L(+\infty)=0$. Тогда нечеткое число A является *нечетким числом L-R-типа*, если для m, $\alpha>0$, $\beta>0$ из множества $\Re$ его функция принадлежности представима в виде:

$$\mu_A(x)=\begin{cases} L\left(\dfrac{m-x}{\alpha}\right), x \le m, \\ R\left(\dfrac{x-m}{\beta}\right), x \ge m, \end{cases}$$

где m – центральное (среднее) значение (мода) нечеткого числа, α и β – левый и правый коэффициенты нечеткости соответственно. Условно нечеткое число A обозначается в виде тройки параметров m, α, β. Если $\alpha=\beta$, то нечеткое число A называется *симметричным нечетким числом L-R-типа* и обозначается как $A=(m, \alpha)$ [3].

Определение 2.

Пусть $A=(m_a,\alpha_a)$ и $B=(m_b,\alpha_b)$ являются симметричными нечеткими числами L-R-типа. Тогда допустимы следующие операции:

1. $A+B=(m_a+m_b,\alpha_a+\alpha_b)$.
2. $\lambda A=\lambda(m_a,\alpha_a)=(\lambda m_a,\lambda\alpha_a)$, где $\lambda>0$.
3. $A-B=(m_a,\alpha_a)-(m_b,\alpha_b)=(m_a-m_b,\alpha_a+\alpha_b)$.

Определение 3 (формула расстояния Евклида).

Пусть $A=(m_a,\alpha_a)$ и $B=(m_b,\alpha_b)$ – два симметричных нечетких числа, тогда расстояние между ними определяется по формуле:

$$D=\sqrt{(m_a-m_b)^2+(\alpha_a-\alpha_b)^2}.$$

Нечеткая линейная множественная регрессионная модель

Исследуем зависимость переменной Y, представляющей собой симметричное нечеткое число L-R-типа, от объясняющих переменных $x_1, x_2, \ldots, x_n$, являющихся вещественными числами. Пусть i-е наблюдение зависимой переменной равно Y_i, а независимых – $x_{i1}, x_{i2}, \ldots, x_{in}$. Рассмотрим следующую обобщенную нечеткую линейную регрессионную модель [1, 4]:

$$Y_i = A_0 + A_1 x_{i1} + A_2 x_{i2} + \ldots + A_n x_{in} + E_i,\ i=\overline{1,m}, \quad (1)$$

где $x_{ik} \in R, k=\overline{1,n}$, $Y_i=(c_i,s_i)$, $i=\overline{1,m}$ – симметричные нечеткие числа с центром (модой) c_i и границей s_i; $A_j=(a_j,r_j)$, $j=\overline{0,n}$ – симметричные нечеткие регрессионные параметры, имеющие функцию принадлежности, подобную Y_i; $E_i=(e_i,\varepsilon_i)$, $i=\overline{1,m}$ – случайные ошибки, симметричные нечеткие числа; $i=\overline{1,m}$ – номер наблюдения, $k=\overline{1,n}$ – номер фактора.

Будем использовать метод наименьших квадратов для получения наиболее подходящих оценок нечетких регрессионных параметров A_j, $j=\overline{0,n}$.

Следует учитывать при этом, что $Y_i=(c_i,s_i)$, $i=\overline{1,m}$, и $A_j=(a_j,r_j)$, $j=\overline{0,n}$, имеют одинаковую функцию принадлежности, и после надлежащего преобразования данных можно сделать все $x_{ik}>0$, $i=\overline{1,m}$, $k=\overline{1,n}$.

Оценкой модели (1) является уравнение:

$$\tilde{Y}_i=\tilde{A}_0+\tilde{A}_1x_{i1}+\ldots+\tilde{A}_nx_{in},\ i=\overline{1,m}, \tag{2}$$

где $\tilde{Y}_i=(\tilde{c}_i,\tilde{s}_i)$, $i=\overline{1,m}$ – оценки зависимой переменной, $\tilde{A}_j=(\tilde{a}_j,\tilde{r}_j)$, $j=\overline{0,n}$ – оценки нечетких коэффициентов регрессии.

Так как $x_{ik}>0$, $i=\overline{1,m}$, $k=\overline{1,n}$, за счет преобразования данных, то уравнение (2) может быть записано в виде:

$$(\tilde{c}_i,\tilde{s}_i)=(\tilde{a}_0,\tilde{r}_0)+(\tilde{a}_1,\tilde{r}_1)x_{i1}+(\tilde{a}_2,\tilde{r}_2)x_{i2}+\ldots+(\tilde{a}_n,\tilde{r}_n)x_{in},\ i=\overline{1,m}. \tag{3}$$

Согласно формуле расстояния Евклида оценкой параметров a_j и r_j, основанной на методе наименьших квадратов, являются величины $\tilde{a}_j$ и $\tilde{r}_j$, минимизирующие величину D^2, где

$$D^2=\sum_{i=1}^{m}\left[\left(c_i-(\tilde{a}_0+\tilde{a}_1x_{i1}+\ldots+\tilde{a}_nx_{in})\right)^2+\left(s_i-(\tilde{r}_0+\tilde{r}_1x_{i1}+\ldots+\tilde{r}_nx_{in})\right)^2\right]. \tag{4}$$

Пусть $\|\vec{v}\|$ – длина вектора $\vec{v}$, тогда, используя векторные и матричные обозначения, величину D^2 можно переписать в виде $D^2=\|X\tilde{a}-c\|^2+\|X\tilde{r}-s\|^2$, где X – матрица объясняющих переменных (матрица плана) размером $m\times(n+1)$, $\tilde{a}=(\tilde{a}_0,\tilde{a}_1,\ldots,\tilde{a}_n)^T$, $\tilde{r}=(\tilde{r}_0,\tilde{r}_1,\ldots,\tilde{r}_n)^T$, $c=(c_1,c_2,\ldots,c_n)^T$ и $s=(s_1,s_2,\ldots,s_n)^T$. Тогда значения параметров $\tilde{a}$ и $\tilde{r}$, минимизирующих величину , находятся из условий:

$$\begin{cases}\dfrac{\partial D^2}{\partial\tilde{a}}=0,\\[2mm] \dfrac{\partial D^2}{\partial\tilde{r}}=0,\end{cases} \tag{5}$$

где $\dfrac{\partial D^2}{\partial\tilde{a}}=\left(\dfrac{\partial D^2}{\partial\tilde{a}_0},\dfrac{\partial D^2}{\partial\tilde{a}_1},\ldots,\dfrac{\partial D^2}{\partial\tilde{a}_n}\right)$, $\dfrac{\partial D^2}{\partial\tilde{r}}=\left(\dfrac{\partial D^2}{\partial\tilde{r}_0},\dfrac{\partial D^2}{\partial\tilde{r}_1},\ldots,\dfrac{\partial D^2}{\partial\tilde{r}_n}\right)$.

$$D^2=\|X\tilde{a}-c\|^2+\frac{1}{2}\|X\tilde{r}-s\|^2=(X\tilde{a}-c)^T(X\tilde{a}-c)+\frac{1}{2}(X\tilde{r}-s)^T(X\tilde{r}-s),$$

$$D^2=\tilde{a}^TX^TX\tilde{a}-c^TX\tilde{a}-\tilde{a}^TX^Tc+c^Tc+\frac{1}{2}\left(\tilde{r}^TX^TX\tilde{r}-s^TX\tilde{r}-\tilde{r}^TX^Ts+s^Ts\right)=$$

$$=\tilde{a}^TX^TX\tilde{a}-2\tilde{a}^TX^Tc+c^Tc+\frac{1}{2}\left(\tilde{r}^TX^TX\tilde{r}-2\tilde{r}^TX^Ts+s^Ts\right)$$

$$\begin{cases}\dfrac{\partial D^2}{\partial\tilde{a}}=2X^TX\tilde{a}-2X^Tc=0,\\[2mm] \dfrac{\partial D^2}{\partial\tilde{r}}=X^TX\tilde{r}-X^Ts=0.\end{cases} \tag{6}$$

В результате получим следующие оценки параметров регрессии:

$$\begin{cases} \tilde{a} = (X^T X)^{-1} X^T c, \\ \tilde{r} = (X^T X)^{-1} X^T s. \end{cases} \tag{7}$$

Из (7) следует, что оценки зависят только от наблюдаемых значений переменных. По свойствам действия с нечеткими числами уравнение (1) может быть представлено в виде системы:

$$\begin{cases} c_i = a_0 + a_1 x_{i1} + \ldots + a_n x_{in} + e_i, \\ s_i = r_0 + r_1 x_{i1} + \ldots + r_n x_{in} + \varepsilon_i. \end{cases} \tag{8}$$

Соотношения (8) удобно записать в матричной форме:

$$\begin{cases} c = Xa + e, \\ s = Xr + \varepsilon, \end{cases} \tag{9}$$

где X – матрица размером $m \times (n+1)$, $c = (c_1, c_2, \ldots, c_m)$, $s = (s_1, s_2, \ldots, s_m)$, $\tilde{a} = (\tilde{a}_0, \tilde{a}_1, \ldots, \tilde{a}_n)^T$, $\tilde{r} = (\tilde{r}_0, \tilde{r}_1, \ldots, \tilde{r}_n)^T$, $e = (e_1, e_2, \ldots, e_m)$, $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m)$.

Подставив (9) в (7), получим оценки параметров

$$\begin{cases} \tilde{a} = (X^T X)^{-1} X^T (Xa + e) = a + (X^T X)^{-1} X^T e, \\ \tilde{r} = (X^T X)^{-1} X^T (Xr + \varepsilon) = r + (X^T X)^{-1} X^T \varepsilon, \end{cases} \tag{10}$$

математическое ожидание которых равно:

$$\begin{cases} M(\tilde{a}) = a + (X^T X)^{-1} X^T M(e), \\ M(\tilde{r}) = r + (X^T X)^{-1} X^T M(\varepsilon). \end{cases} \tag{11}$$

Оценки, получаемые по методу наименьших квадратов, представляют собой несмещенные оценки параметров уравнения (1),

$$\begin{cases} M(\tilde{a}) = a, \\ M(\tilde{r}) = r, \end{cases} \tag{12}$$

тогда и только тогда, когда независимые переменные являются точными и математическое ожидание неучтенных переменных равно нулю [2]:

$$\begin{cases} M(e) = 0, \\ M(\varepsilon) = 0. \end{cases} \tag{13}$$

Найдем зависимость оцениваемых параметров модели (1) от погрешности вычисления. Из (7) и (10) имеем:

$$\begin{cases} a + (X^T X)^{-1} X^T e = (X^T X)^{-1} X^T c, \\ r + (X^T X)^{-1} X^T \varepsilon = (X^T X)^{-1} X^T s. \end{cases} \Rightarrow \begin{cases} a = (X^T X)^{-1} X^T (c - e), \\ r = (X^T X)^{-1} X^T (s - \varepsilon). \end{cases} \tag{14}$$

Метод, описанный выше, использует регрессию с центром и смещением. Оценка результатов не затрагивает порядковую функцию.

Анализ данных

Рассмотрим примеры, которые иллюстрируют полученные результаты [4]. Исходные данные представляют собой набор из десяти трехмерных векторов-входов с четкими компонентами и десяти соответствующих им нечетких выходов (табл. 1).

Таблица 1

Исходные данные

№ вектора	Переменные x_{i1}, x_{i2}, x_{i3}	Нечеткие выходы $Y_i = (c_i, s_i)$
1	(3, 5, 9)	(96, 42)
2	(14, 8, 3)	(120, 47)
3	(7, 1, 4)	(52, 33)
4	(11, 7, 3)	(106, 45)
5	(7, 12, 15)	(189, 79)
6	(8, 15, 10)	(194, 65)
7	(3, 9, 6)	(107, 42)
8	(12, 15, 11)	(216, 78)
9	(10, 5, 8)	(108, 52)
10	(9, 7, 4)	(103, 44)

Найдем оценки параметров $\tilde{A}_j$ модели (2) по формулам (7).

$$X^T X = \begin{pmatrix} 10 & 84 & 84 & 73 \\ 84 & 822 & 735 & 581 \\ 84 & 735 & 888 & 711 \\ 73 & 581 & 711 & 677 \end{pmatrix},$$

$$\left(X^T X\right)^{-1} = \frac{1}{14472471}\begin{pmatrix} 20385687 & -1422645 & 198675 & -1185903 \\ -1422645 & 167190 & -73857 & 87486 \\ 198675 & -73857 & 157364 & -123306 \\ -1185903 & 87486 & -123306 & 203670 \end{pmatrix},$$

$$\tilde{a} = (X^T X)^{-1} X^T c \approx \begin{pmatrix} -1{,}39 \\ 3{,}25 \\ 7{,}92 \\ 5{,}03 \end{pmatrix}, \tilde{r} = (X^T X)^{-1} X^T s \approx \begin{pmatrix} 8{,}01 \\ 1{,}64 \\ 1{,}2 \\ 2{,}85 \end{pmatrix} \qquad (15)$$

Для каждого номера i вычислим оценку $\tilde{Y}_i = (\tilde{c}_i, \tilde{s}_i)$:

$\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \tilde{A}_2 x_{i2} + \tilde{A}_3 x_{i3},\ i = \overline{1,10}$,

а также ошибку $E_i = (e_i, \varepsilon_i),\ i = \overline{1,m}$, где $e_i = c_i - \tilde{c}_i,\ \varepsilon_i = s_i - \tilde{s}_i,\ i = \overline{1,m}$.

Таблица 2

Полученные данные

№	$\tilde{c}_i$	$\tilde{s}_i$	e_i	ε_i	$\hat{e}_i$	$\hat{\varepsilon}_i$
1	93,23	44,58	2,77	-2,58	0,971	0,939
2	122,56	49,12	-2,56	-2,12	0,979	0,955
3	49,4	32,09	2,6	0,91	0,950	0,972
4	104,89	43	1,11	2	0,990	0,956
5	191,85	76,64	-2,85	2,36	0,985	0,970
6	193,71	67,63	0,29	-2,63	0,999	0,960
7	109,82	40,83	-2,82	1,17	0,974	0,972
8	211,74	77,04	4,26	0,96	0,980	0,988
9	110,95	53,21	-2,95	-1,21	0,973	0,977
10	103,42	42,57	-0,42	1,43	0,996	0,968

Точность полученных результатов (таблица 2) определим по формулам:

$$\hat{e}_i = 1 - \left|\frac{e_i}{c_i}\right|, \quad \hat{\varepsilon}_i = 1 - \left|\frac{\varepsilon_i}{s_i}\right|.$$

Данные таблицы свидетельствуют от том, что оценки (15) были вычислены с точностью 97% (в среднем).

Рассмотрим теперь пример нахождения параметров модели (1) с заранее заданной точностью 95%. Для этого воспользуемся формулой (14). Получим следующие результаты (таблица 3):

$$\left(X^T X\right)^{-1} = \frac{1}{14472471}\begin{pmatrix} 20385687 & -1422645 & 198675 & -1185903 \\ -1422645 & 167190 & -73857 & 87486 \\ 198675 & -73857 & 157364 & -123306 \\ -1185903 & 87486 & -123306 & 203670 \end{pmatrix},$$

$$a = (X^T X)^{-1} X^T (c - e) \approx \begin{pmatrix} -1,32 \\ 3,08 \\ 7,52 \\ 4,78 \end{pmatrix}, \quad r = (X^T X)^{-1} X^T (s - \varepsilon) \approx \begin{pmatrix} 7,61 \\ 1,56 \\ 1,14 \\ 2,71 \end{pmatrix}.$$

Таблица 3

Результаты вычислений с точностью 95%

№	e_i	ε_i	$\tilde{c}_i$	$\tilde{s}_i$	$c_i - \tilde{c}_i$	$s_i - \tilde{s}_i$	$\hat{e}_i$	$\hat{\varepsilon}_i$
1	4,8	2,1	91,2	39,9	7,46	-0,38	0,922	0,991
2	6	2,35	114	44,65	3,7	0,3	0,969	0,994
3	2,6	1,65	49,4	31,35	5,12	2,49	0,902	0,925
4	5,3	2,25	100,7	42,75	6,46	4,12	0,939	0,908
5	9,45	3,95	179,55	75,05	6,82	6,14	0,964	0,922
6	9,7	3,25	184,3	61,75	10,08	0,71	0,948	0,989
7	5,35	2,1	101,65	39,9	2,72	3,19	0,975	0,924
8	10,8	3,9	205,2	74,1	14,98	4,76	0,931	0,939
9	5,4	2,6	102,6	49,4	2,68	1,41	0,975	0,973
10	5,15	2,2	97,85	41,8	4,84	3,53	0,953	0,920

Проведя расчет точности найденных параметров по формулам:

$$\hat{e}_i = 1 - \left|\frac{c_i - \tilde{c}_i}{c_i}\right|, \quad \hat{\varepsilon}_i = 1 - \left|\frac{s_i - \tilde{s}_i}{s_i}\right|,$$

можно убедиться в том, что средняя погрешность вычислений действительно составляет 0,05%.

Заключение

В статье был рассмотрен нечеткий регрессионный анализ для множественной линейной регрессионной модели с нечеткими коэффициентами. Для оценки параметров модели был использован метод наименьших квадратов. Как показывают результаты практических экспериментов, нечеткая линейная регрессия позволяет найти неизвестные коэффициенты модели с достаточно высокой точностью, близкой к единице.

Список источников

1. Кремер, Н.Ш. Эконометрика : Учебник для студ. вузов [текст] / Н.Ш. Кремер, Б.А. Путко. – М.: ЮНИТИ, 2002. – 311 с.

2. Математико-статистические методы исследования взаимосвязей в экономике: Из теории и практики ГДР [текст] / Пер.с нем. А.Г.Закурдаева, Х.Н.Цаллагова; Нау.ред. К.Отто, В.В.Швыркова. – М.: Статистика, 1977. – 181 с.

3. Пегат, А. Нечеткое моделирование и управление [текст] / А. Пегат. – М.: Бином, 2009. – 798 с.

4. Tanaka, H. Fuzzy Data Analysis by Possibilistic Linear Models [текст] / H. Tanaka // Fuzzy Sets and Systems, 1987.

ILLEGIBLE MULTIPLE LINEAR REGRESSION MODEL FOR SYMMETRICAL L-R- ILLEGIBLE NUMBERS

Ledeneva Tatyana Mikhaylovna,
Dr. Sc. of Technical Science, Professor, Chief of the Chair of Calculus and Applied Information Technologies of Voronezh State University;
dean@amm.vsu.ru
Sapkina Natalya Vladimirovna,
Post-graduate student, Lecturer of the Chair of Calculus and Applied Information Technologies of Voronezh State University;
natashasapkina@yandex.ru

Illegible multiple linear regression model for symmetrical L-R- illegible numbers is introduced in this article. Least-squares method is applied for model parameters. The received theoretical results are used to analyze the data in the conclusion.
Keywords: distance between two illegible numbers, illegible multiple linear regression model, least-squares method, operations with illegible numbers, symmetrical illegible numbers.