

FUZZY-РЕГРЕССИОННЫЕ МОДЕЛИ В УСЛОВИЯХ НАЛИЧИЯ В СТАТИСТИЧЕСКОЙ ВЫБОРКЕ НЕЧИСЛОВОЙ ИНФОРМАЦИИ

Ю.А. ЗАК

Аннотация. Рассмотрены алгоритмы решения задач нечеткого регрессионного анализа в условиях, когда входные и выходная переменные представлены Fuzzy-множествами, определенными с точностью до неизвестных параметров, а коэффициенты регрессии — действительные числа. Предложены некоторые новые критерии аппроксимации, основанные на сравнении свертки длин сечений и координат центров тяжести функций принадлежности Fuzzy-множеств, которые могут быть использованы для нечетких множеств переменных задачи общего вида. Описаны алгоритмы преобразования переменных, представленных терминами лингвистической переменной или параметрами числовых шкал, в нечеткие множества и использования этих данных в задачах Fuzzy-регрессионного анализа. Полученные результаты позволят решать многие прикладные проблемы в экономике, логистике, социологии и маркетинге.

Ключевые слова: Fuzzy-регрессионного анализ, нечисловая статистика, лингвистические переменные, параметры числовых шкал, Fuzzy-множества, критерии аппроксимации, метод наименьших квадратов.

ВВЕДЕНИЕ

В эконометрике, маркетинговых и социологических исследованиях при построении математических моделей широко используется аппарат математической статистики и построения регрессионных моделей [1]. Широкое распространение здесь получили методы нечисловой статистики. В экономике доля нечисловых данных существенно больше, чем в технике и технологии, и неопределенность приходится описывать в терминах теории нечеткости или математики и статистики интервальных данных.

В ситуациях, когда многие входные факторы модели могут быть представлены лишь булевыми, лингвистическими или нечеткими данными, либо некоторыми градациями числовых шкал, в качестве альтернативных подходов могут использоваться методы Fuzzy-регрессионного анализа. Результатом расчета на основе математических моделей Fuzzy-регрессионного анализа является некоторое нечеткое множество с функцией принадлежности непрерывного вида, которое определяет диапазон возможных значений выходной переменной и оценку (некоторый аналог вероятности) получения этого значения в пределах данного диапазона.

СОСТОЯНИЕ РАЗРАБОТОК В ОБЛАСТИ FUZZY-РЕГРЕССИОННОГО АНАЛИЗА

Математическим моделям и алгоритмам решения задач Fuzzy-регрессионного анализа посвящено большое количество публикаций. Функции регрес-

сии в общем случае, когда входные, выходная переменные и коэффициенты модели — нечеткие множества, рассмотрены в работах Ю.П. Зайченко [2, 3, 5, 20], где предложены методы решения таких задач алгоритмами нечеткого МГУА. В работах С.Д. Штовбы [8], Х. Танака (1982 г.) [10, 11], как и в статьях [13–15] и многих других публикациях, также рассмотрена модель линейной регрессии с нечеткими коэффициентами. В работах [10, 11] для определения значений этих коэффициентов, минимизирующих суммарную средневзвешенную размытость параметров функции принадлежности и рассматриваемую в различных метриках, предложены методы линейного программирования. В 1987 г. А. Селминс и П. Даймонд [14], а также Янг и Лиу в 2003 г. [19] предложили методику построения моделей нечеткой регрессии, основанной на методе наименьших квадратов [13–15, 18, 19]. Для построения критериев аппроксимации (целевой функции нечеткой идентификации) использовались различные метрики, среди которых наибольшее распространение получили показатели α -сечений нечетких множеств. В ряде случаев сформулированная оптимизационная задача становится нелинейной и многоэкстремальной. Для решения ее применялись градиентные, поисковые методы и генетические алгоритмы (см., например, [15]). Робастые регрессионные модели представлены в работах [9, 17, 18]. Решения для частных случаев функции принадлежности треугольного типа рассмотрены в работах [4, 10–15, 17–19]. Формулирование и решение задачи нечеткого регрессионного анализа в виде многокритериальной задачи линейного программирования описаны в статье [4]. Решение задачи нечеткого регрессионного анализа представлено в условиях, когда входные и выходная переменные — Fuzzy-множества, а коэффициенты регрессии — действительные числа. Свободный член уравнения регрессии — нечеткое множество. Рассмотрены некоторые новые критерии аппроксимации, основанные на сравнении свертки длин сечений и координат центров тяжести функций принадлежности Fuzzy-множеств. Разработаны детерминированные эквиваленты сформулированных задач и алгоритмы расчета параметров критериев аппроксимации, т.е. детерминированных значений коэффициентов уравнения регрессии и свободного члена, представленного Fuzzy-множеством, определенным с точностью до неизвестных параметров. Предложены критерии адекватности Fuzzy-регрессионных моделей. Хотя рассматриваемая постановка является частным случаем моделей [2, 3, 5, 20], полученные в работе и перечисленные результаты имеют большое прикладное значение и позволят решить многие прикладные проблемы в экономике, логистике, социологии и маркетинге.

ПОСТАНОВКА И МАТЕМАТИЧЕСКАЯ ФОРМУЛИРОВКА ЗАДАЧИ

Рассматривается следующая постановка задачи Fuzzy-регрессионного анализа. Пусть матрица наблюдений (XY), каждая строка которой — комплект входной и выходной информации, представлена нечеткими множествами. Таблица исходных данных содержит N комплектов экспериментальных данных (строк). Необходимо найти нечеткую линейную регрессионную модель вида

$$\bar{Y} = a_1 \otimes \bar{X}_1 + a_2 \otimes \bar{X}_2 + \dots + a_j \otimes \bar{X}_j + \dots + a_n \otimes \bar{X}_n + \bar{a}_0, \quad (1)$$

где $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_j, \dots, \bar{X}_n$; \bar{a}_0 — некоторые нечеткие множества с заданными функциями принадлежности (в частном случае некоторые из них — действительные числа), а коэффициенты модели $a_1, a_2, \dots, a_j, \dots, a_n$ — некоторые действительные числа.

Отрицательное влияние некоторых входных факторов на выходную переменную можно представить также в виде соотношения (1), заменив их сопряженными Fuzzy-числами. Причем в Fuzzy-регрессионных моделях знаки этих коэффициентов в результате расчетов не меняются.

Ограничимся рассмотрением треугольных и трапецевидных функций принадлежности LR-представления, где (λ_i, m_i, μ_i) и $(\lambda_i, m_i^1, m_i^2, \mu_i)$ — соответственно параметры этой функции принадлежности. Рассмотрим математические модели этой задачи в случае, когда Fuzzy-множества входных и выходных переменных представлены треугольными функциями принадлежности. Центральные точки функций принадлежности $\mu(\bar{X}_{ij})$ и $\mu(\bar{Y}_i)$ соответственно входных \bar{X}_{ij} и выходных переменных \bar{Y}_i этих Fuzzy-множеств обозначим соответственно m_{ij} и \bar{m}_i , левые крайние точки — λ_{ij} , λ_i , а правые крайние точки — γ_{ij} и γ_i , где $j=1, \dots, n$, $i=1, \dots, N$.

Критерий качества аппроксимации — минимум средневзвешенной суммы квадратов отклонений расчетных параметров выходной переменной по Fuzzy-регрессионной модели от их фактических значений:

$$\begin{aligned} \Phi_1 = \eta_1 \sum_{i=1}^N \left[\bar{m}_i - \left(\sum_{j=1}^n a_j m_{ij} + a_0 \right) \right]^2 + \eta_2 \sum_{i=1}^N \left[\bar{\lambda}_i - \left(\sum_{j=1}^n a_j \lambda_{ij} + a_0 \right) \right]^2 + \\ + \eta_3 \sum_{i=1}^N \left[\bar{\gamma}_i - \left(\sum_{j=1}^n a_j \gamma_{ij} + a_0 \right) \right]^2 \rightarrow \min. \end{aligned} \quad (2)$$

В случае трапецевидных функций принадлежности входных \bar{X}_{ij} и выходных переменных \bar{Y}_i обозначим соответственно m_{ij}^1 , m_{ij}^2 и \bar{m}_i^1 , \bar{m}_i^2 , левые крайние точки — λ_{ij} , λ_i , а правые крайние точки — γ_{ij} и γ_i , где $j=1, \dots, n$, $i=1, \dots, N$. Критерий аппроксимации в этом случае имеет вид

$$\begin{aligned} \Phi_2 = \eta_1 \sum_{i=1}^N \left[\bar{m}_i^1 - \left(\sum_{j=1}^n a_j m_{ij}^1 + a_0 \right) \right]^2 + \eta_2 \sum_{i=1}^N \left[\bar{m}_i^2 - \left(\sum_{j=1}^n a_j m_{ij}^2 + a_0 \right) \right]^2 + \\ + \eta_3 \sum_{i=1}^N \left[\bar{\lambda}_i - \left(\sum_{j=1}^n a_j \lambda_{ij} + a_0 \right) \right]^2 + \eta_4 \sum_{i=1}^N \left[\bar{\gamma}_i - \left(\sum_{j=1}^n a_j \gamma_{ij} + a_0 \right) \right]^2 \rightarrow \min. \end{aligned} \quad (3)$$

В уравнениях (2), (3) η_1, η_2, η_3 или $\eta_1, \eta_2, \eta_3, \eta_4$, где $0 < \eta_r \leq 1$, $r=1, 2, 3, 4$ — весовые коэффициенты, определяющие важность значения каждого из параметров, причем в выражении (2) $\eta_1 + \eta_2 + \eta_3 = 1$, а в выражении (3) — $\eta_1 + \eta_2 + \eta_3 + \eta_4 = 1$.

Координата абсцисс центра тяжести нечеткого множества вычисляется по формуле

$$G(\bar{X}_{ij}) = \frac{\int_{-\infty}^{\infty} \bar{X}_{ij} \mu_i(\bar{X}_{ij}) d\bar{X}_{ij}}{\int_{-\infty}^{\infty} \mu_i(\bar{X}_{ij}) d\bar{X}_{ij}}; \quad G(\bar{Y}_i) = \frac{\int_{-\infty}^{\infty} \bar{Y}_i \mu_i(\bar{Y}_i) d\bar{Y}_i}{\int_{-\infty}^{\infty} \mu_i(\bar{Y}_i) d\bar{Y}_i},$$

$$j = 1, \dots, n; \quad i = 1, \dots, N,$$

которая для функций принадлежности треугольного типа имеет вид

$$G(\bar{X}_{ij}) = \frac{1}{3}(\lambda_{ij} + m_{ij} + \gamma_{ij}); \quad G(\bar{Y}_i) = \frac{1}{3}(\bar{\lambda}_i + \bar{m}_i + \bar{\gamma}_i), \quad j = 1, \dots, n; \quad i = 1, \dots, N;$$

$$\Phi_3 = \sum_{i=1}^N \left[G(\bar{Y}_i) - \sum_{j=1}^n a_j \cdot G(\bar{X}_{ij}) + a_0 \right]^2 =$$

$$= \sum_{i=1}^N \left[(\bar{\lambda}_i + \bar{m}_i + \bar{\gamma}_i) - \sum_{j=1}^n a_j (\lambda_{ij} + m_{ij} + \gamma_{ij}) + a_0 \right]^2 \rightarrow \min. \quad (4)$$

Выражение (4) представим в виде $\Phi_4 = \sum_{i=1}^N \left[Y_i - \sum_{j=1}^n a_j z_{ij} + a_0 \right]^2 \rightarrow \min,$

где для треугольных и трапецевидных функций принадлежности значения $z_{ij} \quad j = 1, \dots, n; \quad i = 1, \dots, N$, вычисляются соответственно по формулам:

$$z_{ij} = \sum_{p=1}^P 0,5 \omega_p \{ [m_{ij} - (m_{ij} - \lambda_{ij})(1 - \beta_p)] + [m_{ij} + (\gamma_{ij} - m_{ij})(1 - \beta_p)] \},$$

$$z_{ij} = \sum_{p=1}^P 0,5 \omega_p \{ [m_{ij}^1 - (m_{ij}^1 - \lambda_{ij})(1 - \beta_p)] + [m_{ij}^2 + (\gamma_{ij} - m_{ij}^2)(1 - \beta_p)] \},$$

где $0 \leq \omega_p \leq 1, \quad p = 1, \dots, P$ — весовые коэффициенты, удовлетворяющие соотношению $\sum_{p=1}^P \omega_p = 1; \quad 0 \leq \beta_p \leq 1,0, \quad p = 1, \dots, P$ — значения функции принадлежности в различных сечениях.

Критерий аппроксимации (4) является частным случаем критерия аппроксимации (2).

АЛГОРИТМЫ ПОСТРОЕНИЯ FUZZY-РЕГРЕССИОННЫХ МОДЕЛЕЙ

Необходимыми и достаточными условиями достижения локального минимума значения критериев Φ_1 и Φ_2 является система линейных уравнений вида

$$\frac{d\Phi_p}{da_0} = 0, \quad \frac{d\Phi_p}{da_j} = 0, \quad p = 1, 2, 3, 4; \quad j = 1, \dots, n. \quad (5)$$

Вычислив частные производные по каждому из оптимизируемых параметров, после некоторых алгебраических преобразований получаем, например, для критерия (3) следующую систему линейных алгебраических уравнений для определения значений коэффициентов $a_1, a_2, \dots, a_j, \dots, a_n$ —

$$\sum_{j=1}^n \beta_{kj} a_j + a_0 = B_k, \quad k = 1, \dots, n.$$

Как и в классических методах регрессионного анализа, перейдем к нормированным показателям. Вычислим

$$M(z_j) = \frac{1}{N} \sum_{i=1}^N z_{ij}, \quad \sigma^2(z_j) = \frac{1}{N-1} \sum_{i=1}^N \left[\sum_{i=1}^N z_{ij} - M(z_j) \right]^2, \quad j = 1, \dots, n; \quad (6)$$

$$M(Y) = \frac{1}{N} \sum_{i=1}^N Y_i, \quad \sigma^2(Y) = \frac{1}{N-1} \sum_{i=1}^N \left[\sum_{i=1}^N Y_i - M(Y) \right]^2; \quad (7)$$

$$\rho_{kj} = \frac{\frac{1}{N} \sum_{i=1}^N [z_{ki} - M(z_k)][z_{ji} - M(z_j)]}{\sigma(z_k) \sigma(z_j)}, \quad j, k = 1, \dots, n, \quad k \neq j;$$

$$\rho_{kk} = 1; \quad \rho(y, z_k) = \frac{\frac{1}{N} \sum_{i=1}^N [z_{ki} - M(z_k)][Y_i - M(Y)]}{\sigma(z_k) \cdot \sigma(Y)}, \quad k = 1, \dots, n,$$

и перейдем к решению следующей системы n линейных алгебраических уравнений относительно переменных $\delta_j, j = 1, \dots, n$:

$$\sum_{j=1}^n \rho_{kj} \delta_j = \rho(y, z_k), \quad k = 1, \dots, n.$$

Вычислив значения коэффициентов $\delta_j, j = 1, \dots, n$, определим значения коэффициентов $\bar{a}_j, j = 1, \dots, n$, в нормальном масштабе измерения по формулам $\bar{a}_j = \delta_j \cdot \frac{\sigma(Y)}{\sigma(z_j)}, j = 1, \dots, n$.

Значение свободного члена уравнения нечеткой регрессии a_0 в виде нечеткого множества \bar{A}_0 , определенного с точностью до неизвестных параметров, который для критериев $\Phi_1 - \Phi_3$ не обязательно должен совпадать с видом нечетких множеств выходной переменной, вычисляется по приведенным ниже формулам.

Рассмотрим в качестве примеров следующий случай.

Параметры нечеткого множества треугольного вида

$$\lambda_0 = \lambda(\bar{A}_0), \quad \gamma_0 = \gamma(\bar{A}_0) \quad \text{и} \quad m = m(\bar{A}_0) \quad (\text{или} \quad m_0^1 = m_0^1(\bar{A}_0) \quad \text{и} \quad m_0^2 = m_0^2(\bar{A}_0))$$

определяются в результате минимизации одного из следующих критериев:

$$F_1 = \min_{(\lambda_0, m_0, \gamma_0)} \left\{ \eta_1 \sum_{i=1}^N \left[\bar{m}_i - \left(\sum_{j=1}^n \bar{a}_j m_{ij} + m_0 \right) \right]^2 + \right.$$

$$\begin{aligned}
 & + \eta_2 \sum_{i=1}^N \left[\bar{\lambda}_i - \left(\sum_{j=1}^n \bar{a}_j \lambda_{ij} + \lambda_0 \right) \right]^2 + \eta_3 \sum_{i=1}^N \left[\bar{\gamma}_i - \left(\sum_{j=1}^n a_j \gamma_{ij} + \gamma_0 \right) \right]^2 \Bigg\}; \\
 & F_2 = \\
 & = \min_{(\lambda_0, m_0^1, m_0^2, \gamma_0)} \left\{ \begin{aligned} & \eta_1 \sum_{i=1}^N \left[\bar{m}_{ij}^1 - \left(\sum_{j=1}^n \bar{a}_j m_{ij}^1 + m_0^1 \right) \right]^2 + \eta_2 \cdot \sum_{i=1}^N \left[\bar{m}_{ij}^2 - \left(\sum_{j=1}^n \bar{a}_j m_{ij}^2 + m_0^2 \right) \right]^2 + \\ & + \eta_3 \sum_{i=1}^N \left[\bar{\lambda}_i - \left(\sum_{j=1}^n \bar{a}_j \lambda_{ij} + \lambda_0 \right) \right]^2 + \eta_4 \cdot \sum_{i=1}^N \left[\bar{\gamma}_i - \left(\sum_{j=1}^n \bar{a}_j \gamma_{ij} + \gamma_0 \right) \right]^2 \end{aligned} \right\}; \\
 & F_3 = \min_{(\lambda_0, m_0, \gamma_0)} \left\{ \sum_{i=1}^N \left[Y_i - \sum_{j=1}^n \bar{a}_j z_{ij} + z_0(\lambda_0, m_0, \gamma_0) \right]^2 \right\},
 \end{aligned}$$

где значения z_{ij} и Y_i вычисляются соответственно по формулам соответственно (5) или (6) и (7), а $z_0(\lambda_0, m_0, \gamma_0)$ или $z_0(\lambda_0, m_0^1, m_0^2, \gamma_0)$ определяется в виде

$$\begin{aligned}
 & z_0(\lambda_0, m_0, \gamma_0) = \\
 & = \sum_{p=1}^P 0,5 \omega_p \{ [m_0 + (m_0 - \lambda_0)(1 - \beta_p)] + [m_0 + (\gamma_0 - m_0)(1 - \beta_p)] \} \quad (8)
 \end{aligned}$$

или

$$\begin{aligned}
 & z_0(\lambda_0, m_0^1, m_0^2, \gamma_0) = \\
 & = \sum_{p=1}^P 0,5 \omega_p \{ [m_0^1 + (m_0^1 - \lambda_0)(1 - \beta_p)] + [m_0^2 + (\gamma_0 - m_0^2)(1 - \beta_p)] \}. \quad (9)
 \end{aligned}$$

Кроме того, могут использоваться критерии аппроксимации

$$F_4 = \min_{(\lambda_0, m_0, \gamma_0)} \left\{ \sum_{i=1}^N \left[(\bar{\lambda}_i + \bar{m}_i + \bar{\gamma}_i) - \sum_{j=1}^n \bar{a}_j (\lambda_{ij} + m_{ij} + \gamma_{ij}) + (\lambda_0 + m_0 + \gamma_0) \right]^2 \right\} \quad (10)$$

или

$$\begin{aligned}
 & F_4 = \min_{(\lambda_0, m_0^1, m_0^2, \gamma_0)} \times \\
 & \times \left\{ \sum_{i=1}^N \left[(\bar{\lambda}_i + \bar{m}_i + \bar{\gamma}_i) - \sum_{j=1}^n \bar{a}_j (\lambda_{ij} + m_{ij} + \gamma_{ij}) + (\lambda_0 + m_0^1 + m_0^2 + \gamma_0) \right]^2 \right\}. \quad (11)
 \end{aligned}$$

Решение каждой из оптимизационных задач (8)–(11) может быть получено решением системы линейных алгебраических уравнений, полученной аналогично (5).

ОЦЕНКА АДЕКВАТНОСТИ FUZZY-РЕГРЕССИОННОЙ МОДЕЛИ

Результат расчета выходных показателей на основе нечеткой регрессионной модели — Fuzzy-множество треугольного или трапецевидного типа. В качестве детерминированного эквивалента прогнозируемой величины может быть принята координата абсцисс центра тяжести нечеткого множества, полученного в результате расчета по Fuzzy-регрессионной модели.

Как оценки качества прогнозирования на основе нечеткой регрессионной модели могут рассматриваться следующие два показателя:

- среднеквадратическое значение суммы отклонений фактической координаты абсцисс центра тяжести нечеткого множества выходной переменной $G(Y_i^\Phi)$ от соответствующего значения, рассчитанного по Fuzzy-регрессионной математической модели нечеткого множества $G(Y_i^P)$:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N [G(Y_i^P) - G(Y_i^\Phi)]^2};$$

- среднеквадратическое значение суммы отклонений расчетного и фактического значений функций принадлежности одних и тех же значений выходной переменной:

$$\Delta = \sqrt{\frac{1}{N-1} \left(\sum_{i=1}^N \sum_{t=1}^T w^{t,-} [\{Y_i^{\Phi,t,-} | \mu_i^t(Y_i) = \beta^t\} - \{\bar{Y}_i^{P,t,-} | \mu_i^t(Y_i) = \beta^t\}]^2 + \sum_{i=1}^N \sum_{t=1}^T w^{t,+} [\{Y_i^{\Phi,t,+} | \mu_i^t(Y_i) = \beta^t\} - \{\bar{Y}_i^{P,t,+} | \mu_i^t(Y_i) = \beta^t\}]^2 \right)}.$$

Здесь

$\{Y_i^{\Phi,t,-} | \mu_i^t(Y_i) = \beta^t\}$ и $\{Y_i^{\Phi,t,+} | \mu_i^t(Y_i) = \beta^t\}$ — соответственно t -е фактическое минимальное и максимальное значение выходной величины в i -м комплекте контрольной выборки статистических данных, соответствующие значению $\mu_i^t(Y_i) = \beta^t$;

$\{\bar{Y}_i^{P,t,-} | \mu_i^t(Y_i) = \beta^t\}$ и $\{\bar{Y}_i^{P,t,+} | \mu_i^t(Y_i) = \beta^t\}$ — соответственно t -е минимальное и максимальное значение выходной величины в i -м комплекте контрольной выборки статистических данных, соответствующие значению $\mu_i^t(Y_i) = \beta^t$ и рассчитанные по уравнению Fuzzy-регрессионного анализа;

$0 < w^{t,-} < 1$, и $0 < w^{t,+} < 1$, $t = 1, \dots, T$ — значения весовых коэффициентов, определяющих степень важности учета каждого из значений отклонений, причем $\sum_{t=1}^T (w^{t,-} + w^{t,+}) = 1$.

Значения $\{Y_i^{t,-} | \mu_i^t(Y_i) = \beta^t\}$ и $\{Y_i^{t,+} | \mu_i^t(Y_i) = \beta^t\}$ определяются по формулам:

- для треугольных функций принадлежности

$$\{Y_i^{t,-} | \bar{\mu}_i(Y_i^{t,-}) = \beta^t\} =$$

$$= \begin{cases} 0, & \text{если } Y_i^{t,-} \leq m_i(Y_i^{t,-}) - \lambda(Y_i^{t,-})_i \text{ или } Y_i^{t,-} \geq m_i(Y_i^{t,-}) + \gamma(Y_i^{t,-})_i, \\ \lambda_i(Y_i^t) + \beta^t [m_i(Y_i^{t,-}) - \lambda_i(Y_i^{t,-})], & \text{если } m_i(Y_i^{t,-}) - \lambda_i(Y_i^{t,-}) > Y_i^{t,-}, \\ & \text{и } Y_i^{t,-} \leq m_i(Y_i^{t,-}), \end{cases}$$

- для трапецевидных функций принадлежности

$$\{Y_i^{t,+} \mid \bar{\mu}_i(Y_i^{t,-}) = \beta^t\} = \begin{cases} 0, & \text{если } Y_i^{t,+} \leq m_i(Y_i^{t,+}) - \lambda(Y_i^{t,-})_i \text{ или } Y_i^{t,-} \geq m_i(Y_i^{t,+}) + \gamma(Y_i^{t,+})_i, \\ \gamma_i(Y_i^{t,-}) + \beta^t [\gamma_i(Y_i^{t,-}) - m_i(Y_i^{t,-})], & \text{если } m_i(Y_i^{t,+}) - \gamma_i(Y_i^{t,+}) > Y_i^{t,-}, \\ & \text{и } Y_i^{t,-} \leq m_i(Y_i^{t,-}). \end{cases}$$

С достаточной для практических приложений точностью в большинстве случаев могут использоваться построенные Fuzzy-регрессионные модели для прогнозирования значения переменной составляющей затрат, если справедливы следующие показатели их адекватности:

$$\sigma \leq \vartheta \sqrt{\frac{1}{N-1} \sum_{i=1}^N [G(Y_i^\Phi) - M\{G(Y_i^\Phi)\}]^2},$$

где $M\{G(Y_i^\Phi)\} = \frac{1}{N} \sum_{i=1}^N G(Y_i^\Phi)$, а $\vartheta \leq 0,1$, либо $\vartheta \Psi \leq \Delta$, где

$$\Psi = \sqrt{\frac{1}{N-1} \sum_{i=1}^N \sum_{t=1}^T \{w^{t,-} [M(Y_i^{\Phi,-}) - Y_i^{\Phi,t,-}]^2 + w^{t,+} [M(Y_i^{\Phi,+}) - Y_i^{\Phi,t,+}]^2\}},$$

$$M(Y_i^{\Phi,t,-}) = \frac{1}{N} \sum_{i=1}^N Y_i^{\Phi,t,-}, \quad M(Y_i^{\Phi,t,+}) = \frac{1}{N} \sum_{i=1}^N Y_i^{\Phi,t,+}.$$

ЗАКЛЮЧЕНИЕ

В условиях, когда в статистической выборке некоторые или все входные и выходная переменные представлены нечисловой информацией (булевы, лингвистические переменные, данные числовых шкал, нечеткие множества и т.п.), использование Fuzzy-регрессионных моделей является эффективной альтернативой получения количественных зависимостей, установленных экспертами качественных закономерностей изучаемых явлений.

В отличие от традиционных регрессионных зависимостей в Fuzzy-регрессионных моделях рассчитанное значение выходной переменной представлено в виде некоторого диапазона возможных значений с оценкой веса каждого из этих значений в пределах этого диапазона. Это позволит в ряде случаев более объективно оценить риск принимаемых решений на основе полученных результатов расчета.

Предложенные в работе ранее не описанные в литературе критерии аппроксимации, алгоритмы построения нечетких регрессионных моделей и методов преобразования многих видов нечисловой информации в действительное число или нечеткое множество, а также использование этой информации в вычислительных схемах расчета детерминированных коэффициен-

тов уравнения регрессии расширят область применения полученных результатов в экономике, социологии, маркетинге и других приложениях.

ЛИТЕРАТУРА

1. *Расин Дж.* Непараметрическая эконометрика: вводный курс / Расин Дж. // Квантиль. — 2008. — № 4. — С. 7–56.
2. *Зайченко Ю.П.* Нечеткий метод группового учета аргументов при неопределенных входных данных / Ю.П. Зайченко // Системні дослідження та інформаційні технології. — 2007. — № 3. — С. 100–112.
3. *Зайченко Ю.П.* Нечеткие модели и методы в интеллектуальных системах / Ю.П. Зайченко. — К.: Изд. Дом «Слово», 2008. — 354 с.
4. *Зак Ю.А.* Математические модели прогнозирования затрат времени и стоимости перевозки грузов / Ю.А. Зак // Логистика сегодня, Grebennikov. — 2015. — № 1. — С. 162–172.
5. *Згуровский М.* Модели и методы принятия решений в нечетких условиях / М. Згуровский, Ю. Зайченко. — К.: Наук. думка, 2013. — 275 с.
6. *Зак Ю.А.* Принятие решений в условиях размытых и нечетких данных / Ю.А. Зак // URSS, М., 2013. — 352 с.
7. *Ярушкина Н.Г.* Основы теории нечетких и гибридных систем / Н.Г. Ярушкина. — М.: Финансы и статистика, 2004. — 320 с.
8. *Штовба С.Д.* Нечеткая идентификация на основе регрессионных моделей параметрической функции принадлежности / С.Д. Штовба // Проблемы управления и информатики. — 2006. — № 6. — С. 1–8.
9. *Грицюк В.И.* Нечеткий робастный регрессионный анализ для нечетких входных и выходных данных / В.И. Грицюк // Технологический аудит и резервы производства. — 2015. — № 6. — С. 4–8.
10. *Tanaka H.* Linear regression analysis with fuzzy model / H. Tanaka, S. Uejima, K. Asai // IEEE Trans. Systems Man Cybernet. — 1982. — 12, N 6. — P. 903–907.
11. *Tanaka H.* Possibilistic linear system and their application to the linear regression model / H. Tanaka, J. Warada // Fuzzy Sets and Systems. — 1988. — N 27. — P. 275–289.
12. *Diamond P.* Fuzzy least squares / P. Diamond // Information Sciences. — 1988. — N 46. — P. 141–157.
13. *Celmins A.* Least Squares model fitting to fuzzy vector data / A. Celmins // Fuzzy Sets and System. — 1987. — Vol. 22. — P. 260–269.
14. *Diamond P.* Least squares fitting of several Fuzzy variables / P. Diamond // Proceedings of Second IFSA Congress. — Tokyo, 1987. — P. 20–25.
15. *Aliev R.* Genetic algorithms-based fuzzy regression analysis / R. Aliev, B. Fazlollahi, R. Vahidov // Soft Computing. — 2002. — N 6. — P. 470–475.
16. *Papadopoulos B.* Similarities and distances in fuzzy regression modeling / B. Papadopoulos, M. Sirpi // Soft Computing. — 2004. — № 8. — P. 556–561.
17. *D-Urso P.* Robust fuzzy regression analysis / P. D-Urso, R. Massan, A. Santoro // Informations Sciences. — 2011. — Vol. 181. — P. 1154–1174.
18. *Rousseeuw P.* Applying robust regression to insurance / P. Rousseeuw // Insurance: Mathematics and Economics. — 1984. — Vol. 3, N 1. — P. 67–72.
19. *Yang M-S.* Fuzzy Least Squares algorithms for Interactive fuzzy linear regression models / M-S. Yang, H.H. Liu // Fuzzy Sets and Systems. — 2003. — Vol. 135, N 2. — P. 305–316.
20. *Zaychenko Yu.* The Fuzzy group method of data handling and Its Application for Economical Processes forecasting / Yu. Zaychenko // Scientific Inquiry. — 2006. — Vol. 7, N 1. — P. 83–98.

Поступила 16.06.2016