

Нечеткое регрессионное моделирование в задачах управления аграрным производством

doi 10.26310/2071-3010.2019.249.7.013

Проблема учета и моделирования неопределенности в современных задачах управления является одной из самых насущных. Эффективность принимаемых решений существенно зависит от методов описания имеющейся в задаче неопределенности. Наибольшего развития в аграрной науке получили оптимизационные и эконометрические модели. Однако и те, и другие базируются на количественной детерминированной исходной информации и учете неопределенности как случайности, для описания которой применяются вероятностно-статистические методы. Между тем многие современные задачи принятия решений в планировании и управлении аграрным производством характеризуются наличием неопределенных факторов, а также наличием качественной, неточной или неполной информации. Для учета и описания такой неопределенности нужен подход альтернативной вероятностному подходу. Одним из наиболее эффективных математических инструментариев, направленных на формализацию и обработку неопределенной информации является метод теории нечетких множеств. Раздел эконометрики, связанный с применением теории нечетких множеств в регрессионном анализе, занимается разработкой методов нечеткого регрессионного моделирования. В данной статье рассматриваются возможности применения инструментария нечеткого регрессионного моделирования для анализа процессов управления аграрным производством.

Ключевые слова: нечеткое моделирование, регрессия, анализ, аграрное производство

Введение

Известно, что проблема учета неопределенности и неполноты данных занимает ключевое место при решении задач управления любым хозяйственным объектом, включая и управление в аграрно-промышленном комплексе (АПК). Посмотрим, как эта проблема решается при моделировании аграрных процессов. Наибольшее применение в аграрной науке и практике получили оптимизационные и эконометрические модели.

Разработанные к настоящему времени количественные методы принятия оптимальных решений в управлении АПК базируются на использовании количественной четкой информации и дают возможность выбирать наилучшие из множества возможных решений в двух случаях. Либо в условиях полной определенности (детерминированные модели). Либо в условиях одного конкретного вида неопределенности, а именно неопределенности вероятностного характера (стохастические модели). Между тем требование детерминированности моделей является упрощением реальности, так как в реальной жизни ситуации, лишённые неопределенности, скорее исключение, чем правило. Что касается учета неопределенности, то в процессе принятия решений возникают различные



В. Е. Парфенова,
д. э. н., профессор, Санкт-Петербургский
государственный аграрный университет
w.parfenova@mail.ru

виды неопределенности, часто не носящие вероятностного характера.

Многие современные задачи принятия решений в планировании и управлении аграрным производством характеризуются наличием неопределенных факторов, а также наличием качественной, неточной или неполной информации. Применение вероятностного подхода для описания неопределенности в этом случае не адекватно природе неопределенности. Нужен другой подход. Одним из наиболее эффективных математических инструментариев, направленных на формализацию и обработку неопределенной информации, является теория нечетких множеств. Данная теория позволяет с единых позиций рассмотреть различные виды неопределенности. Раздел математики, связанный с поиском оптимальных решений на базе нечеткой информации, получил название нечеткого математического программирования (НМП). Данный аппарат, как показано в [1, 2], находит широкое применение в постановке и решении задач экономики и менеджмента. Однако, применяют его в основном специалисты математических, информационных и технических специальностей. Возможности использования нечеткого инструментария в моделировании и решении оптимизационных задач аграрного производства нами были рассмотрены ранее [3].

Исследуемая проблема

Данная статья посвящена рассмотрению эконометрических методов, в частности методов регрессионного моделирования и анализа, получивших наряду с оптимизацией широкое распространение при исследовании аграрных процессов. Задача регрессионного анализа заключается в выявлении зависимости некоторого показателя от различных признаков и отображение этой взаимосвязи в форме модели. В классическом регрессионном анализе данные, как для зависимой переменной, так и независимых переменных являются числовыми. При этом для получения качественной регрессионной модели необходима точная числовая статистическая информация, так как неопределенность трактуется как случайность. В этом случае для получения оценок параметров модели применение статистических методов оправдано.

Однако, во многих современных прикладных задачах доступны лишь приближенные и нечеткие данные, или их недостаточно для установления достоверного вида вероятностного распределения. В этом случае неопределенность становится нечеткостью, для описания которой традиционные методы регрессионного моделирования не применимы. Наилучшим способом формализации приближенной и нечеткой информации, как уже отмечалось выше, является инструментарий теории нечетких множеств. Раздел эконометрики, связанный с построением регрессии в нечеткой форме, получил название нечеткого регрессионного моделирования. Методы нечеткого регрессионного моделирования на сегодняшний день достаточно разработаны [4-6]. Но для решения задач аграрной отрасли они пока не нашли должного применения.

Среди имеющихся к настоящему времени методов нечеткого регрессионного моделирования можно выделить две группы. Одна основывается на методе наименьших квадратов, другая — на линейном программировании. В данной статье предлагается построение нечеткой регрессионной модели для анализа производственного аграрного процесса, с применением метода линейного программирования.

В научной литературе рассматриваются различные варианты нечетких регрессионных моделей. Независимые переменные (регрессоры) регрессии могут быть как четкими, при нечетких коэффициентах так и нечеткими при четких коэффициентах, либо с нечеткими и регрессорами и коэффициентами модели. При этом данные для составления нечеткой регрессии

могут быть заданы интервально, нечетко или детерминировано. Большинство современных нечетких регрессионных моделей имеют линейную структуру [7, 8]. Мы рассмотрим моделирование нечеткой линейной множественной регрессии с нечеткими коэффициентами, четкими регрессорами и с детерминированными исходными данными.

Постановка задачи и алгоритм ее решения

Поставим следующую задачу нечеткой линейной регрессии. Имеем n результатов наблюдений зависимой переменной $y_j, j=1..n$, а также $m \times n$ значений независимых переменных $x_{ij}, i=1..m; j= 1..n$ (см. табл. 1).

Будем искать результирующую переменную y в виде нечеткого уравнения множественной регрессии:

$$y_{\text{неч } i} = A_{\text{неч } 0} + A_{\text{неч } 1} x_{1i} + \dots + A_{\text{неч } m} x_{mi}, \quad (1)$$

где x_i — детерминированные значения, а $y_{\text{неч } i}$ — нечеткие.

Задача построения нечеткой линейной регрессионной модели в этом случае состоит в подборе нечетких параметров $A_{\text{неч } j}, j=0..m$ так, чтобы удовлетворялись следующие два условия:

- (У1) получаемые на основе нечетких параметров нечеткие множества

$$Y_{\text{неч}} = A_{\text{неч } 0} + A_{\text{неч } 1} x_1 + \dots + A_{\text{неч } m} x_m$$

должны содержать наблюдаемые значения y_i со степенью достоверности не меньшей, чем некоторая заданная степень h , т. е. для каждого j должно выполняться неравенство $\mu_i(y_i) \geq h$, где μ_i — функция принадлежности нечеткого множества y_j , а h — заданный порог надежности;

- (У2) общая нечеткость модели должна быть минимальна.

Приведем понятия и арифметические операции теории нечетких множеств, используемые при построении такой модели [9].

Универсальное множества U — это полное множество, охватывающее всю проблемную область.

Нечетким множеством F на универсальном множестве U называется совокупность пар $(\mu_F(u), u)$, где $\mu_F(u)$ — функция принадлежности элемента $u \in U$ нечеткому множеству F .

Функция принадлежности $\mu_F(u)$ отражает степень принадлежности каждого элемента универсального множества нечеткому подмножеству F . Функция принадлежности принимает значения от 0 до 1.

Нечеткой величиной называется произвольное нечеткое множество A , заданное на множестве действительных чисел R .

Нечетким числом называется нечеткая величина, функция принадлежности которой является выпуклой и унимодальной.

Нечеткие величины типа «примерно k » задаются, как правило, нечеткими треугольными числами. Треугольным нечетким числом (ТНЧ) A называется

Таблица 1

Массив исходных данных

№ наблюдения	y	x_1	x_2	...	x_m
1	y_1	x_{11}	x_{21}	...	x_{m1}
2	y_2	x_{12}	x_{22}	...	x_{m2}
...
n	y_n	x_{1n}	x_{2n}	...	x_{mn}

нечеткое число, функция принадлежности которого имеет треугольный вид:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a}, & \text{если } x \in [a, b], \\ \frac{c-x}{c-b}, & \text{если } x \in [b, c], \\ 0, & \text{в противном случае.} \end{cases}$$

Треугольное нечеткое число A определяется тройкой действительных чисел (a, b, c) , где $a \leq b \leq c$, b – модальное значение, числа a и c характеризуют степень размытости числа b , при этом a – левый коэффициент нечеткости, c – правый коэффициент нечеткости. Если $b-a=c-b$, то ТНЧ называется симметричным нечетким числом.

Сумма треугольных нечетких чисел $A=(a^L; a; a^R)$ и $B=(b^L; b; b^R)$ – это треугольное нечеткое число: $A+B=(a^L+b^L; a+b; a^R+b^R)$.

Произведение треугольного нечеткого числа $A=(a-d; a; a+d)$ на четкое число l – это треугольное нечеткое число: $lA=(la-l|d; la; la+l|d)$.

Опишем алгоритм поиска коэффициентов $A_{\text{неч } j}$ регрессии (1).

Будем искать коэффициенты $A_{\text{неч } j}$ в виде симметричных треугольных нечетких чисел: $A_{\text{неч } j}=(a_j-b_j; a_j; a_j+b_j)$, $b_j \geq 0$. Тогда зависимые величины $y_{\text{неч } j}$ также являются симметричными нечеткими числами, которые определяются как $y_{\text{неч } j}=(c_i-d_i; c_i; c_i+d_i)$, $i=1...n$, где

$$c_i = a_0 + \sum_{j=1}^m a_j x_{ij}, \quad d_i = b_0 + \sum_{j=1}^m b_j |x_{ij}|.$$

Суммарная мера нечеткости равна

$$d = \sum_{j=1}^m d_i = b_0 + \sum_{i=1}^n \sum_{j=1}^m b_j |x_{ij}| \rightarrow \min.$$

Задача определения коэффициентов нечеткой регрессии сводится к постановке и решению задачи линейного программирования с критерием (У2) и ограничениями (У1).

Формально, задача состоит в поиске неизвестных коэффициентов $a_j, b_j, j=0...m$, при которых достигается:

$$\begin{aligned} (a_0 + \sum_{j=1}^m a_j x_{ij}) - (1-h) (b_0 + \sum_{j=1}^m b_j |x_{ij}|) &\leq y_i, \quad i=1...n, \\ (a_0 + \sum_{j=1}^m a_j x_{ij}) + (1-h) (b_0 + \sum_{j=1}^m b_j |x_{ij}|) &\geq y_i, \quad i=1...n, \\ b_j &\geq 0, \quad j=0...m. \end{aligned}$$

Пример

Проиллюстрируем применение изложенного метода на следующем примере. По 15 сельскохозяйственным предприятиям региона изучается зависимость годовой выручки от реализации на 1 га сельскохозяйственных угодий y (тыс. руб.) от величины затрат на реализованную продукцию на 1 га сельскохозяйственных угодий x_1 (тыс. руб.) и среднегодовой стоимости основных фондов на 1 га сельскохозяйственных угодий x_2 (тыс. руб.) (табл. 2). Требуется построить нечеткое уравнение множественной (двухфакторной) модели.

Нечеткая линейная двухфакторная регрессия имеет следующий вид:

$$Y_{\text{неч}} = (y_{\text{неч } 1}, y_{\text{неч } 2})^T = A_{\text{неч } 0} + A_{\text{неч } 1} x_1 + A_{\text{неч } 2} x_2. \quad (2)$$

Построение регрессии сводится к определению ее коэффициентов. Пусть коэффициенты регрессии (2) являются нечеткими симметричными треугольными числами: $A_{\text{неч } 1}=(a_1-b_1; a_1; a_1+b_1)$; $A_{\text{неч } 2}=(a_2-b_2; a_2; a_2+b_2)$. Тогда для каждого y_j ($i=1...15$) определена нечеткая величина $y_{\text{неч } i}=(c_i-d_i; c_i; c_i+d_i)$, где

$$c_i = a_0 + a_1 x_1 + a_2 x_2 = a_0 \sum_{j=1}^2 a_j x_i, \quad d_i = b_0 + b_1 x_1 + b_2 x_2 = b_0 \sum_{j=1}^2 b_j x_i.$$

Общая нечеткость

$$d = \sum_{i=1}^{15} d_i = 15b_0 + \sum_{i=1}^{15} \sum_{j=1}^2 b_j x_i.$$

Имеем следующую задачу линейного программирования:

$$15b_0 + \sum_{i=1}^{15} \sum_{j=1}^2 b_j x_i \rightarrow \min,$$

при ограничениях:

$$\begin{aligned} a_0 + \sum_{i=1}^2 a_i x_i - (1-h) (b_0 + \sum_{j=1}^2 b_j x_i) &\leq y_i, \quad i=1...15, \\ a_0 + \sum_{i=1}^2 a_i x_i + (1-h) (b_0 + \sum_{j=1}^2 b_j x_i) &\leq y_i, \quad i=1...15, \quad (3) \\ (b_0, b_1, b_2) &\geq 0. \end{aligned}$$

Задача (3) имеет 6 неизвестных и 33 ограничения. Решение задачи производилось с помощью встроенной функции «Поиск решения» в MS Excel.

Анализ результатов

Были рассмотрены значения $h=0; 0,25; 0,5; 0,75$ (см. табл. 3). В зависимости от заданного значения h найдены различные решения искомым па-

Таблица 2.
Исходные данные для построения нечеткой регрессии

№ предприятия	Выручка, тыс. руб., y	Затраты, тыс. руб., x_1	Среднегодовая стоимость основных фондов, тыс.руб., x_2
1	71	52,5	90,3
2	46,4	38,5	105,7
3	42,5	35,9	32,7
4	47,6	39,6	93,4
5	35,4	29,9	27,2
6	42,5	42,1	30,3
7	36,9	25,9	25,8
8	34,1	29,7	38,1
9	21,3	11	16,2
10	23,7	15,8	30
11	54,8	32,6	37,4
12	23,3	15,3	23,6
13	25,6	23,5	20,2
14	37,8	37,6	44,9
15	50,7	43,6	34,3

Таблица 3
Оценки нечетких коэффициентов

h	0	0,25	0,5	0,75	0
d	153,7	204,9	307,3	614,7	197,6
a_0	8,036	8,036	8,036	8,036	0
a_1	0,958	0,958	0,958	0,958	0,851
a_2	0,132	0,132	0,132	0,132	0,433
b_0	0	0	0	0	0
b_1	0,324	0,433	0,649	1,298	0
b_2	0	0	0	0	0,304

раметров нечеткой регрессии. Чем больше h , тем с большей достоверностью мы хотим получить значения величины результативного показателя. Размытость результативного показателя при этом возрастает.

В последнем столбце табл. 3 приведены коэффициенты нечеткой регрессии без учета неучтенных факторов ($a_0=0$). В этом случае общая неопределенность модели при таком же значении $h=0$ оказалась выше, чем при учете неучтенных факторов ($a_0 \neq 0$). Поэтому более предпочтительна модель при $a_0 \neq 0$.

Из табл. 3 видно, что нечетким коэффициентом в регрессии (2), связанным с нечеткостью зависимого показателя (выручка) при всех значениях h – степень достоверности, оказался коэффициент $A_{\text{неч}} 1$ при объясняющей переменной «объем затрат».

Выпишем регрессионную модель при $h=0, a_0 \neq 0$. Имеем

$$y_{\text{неч}} = 8,036 + 0,634; 0,958; 1,282 x_1 + 0,132 x_2. \quad (4)$$

Полученную модель (4) можно использовать для прогнозирования величины выручки. Для этого прогнозные значения факторов нужно подставить в полученное уравнение (4) и рассчитать левое, среднее и правое значения результативного показателя по прогнозным значениям факторов. В табл. 4 приведены результаты оценочной функции результативного показателя (выручки), полученные на основе нечеткой регрессии (2), а также прогнозные значения для контрольной выборки (приведены в трех последних строчках).

Данные табл. 4 показывают, что все исходные и прогнозные значения y попали в диапазон своих значений $y_{\text{неч}}$.

Выводы

Методы моделирования, основанные на теории нечетких множеств, дают более широкие возможности для анализа экономических процессов.

Нечеткая регрессия позволяет получить диапазон возможных значений, при этом обеспечивая охват области исходных данных.

Наличие интервального разброса результирующего признака позволяет учесть неопределенность исходной информации и может заменить доверительный интервал средней линии классической регрессии.

Нечеткий регрессионный анализ может применяться при задании исходных данных в форме четких, интервальных или нечетких чисел.

В приведенной модели для удобства и простоты вычислений функции принадлежности нечетких параметров оцениваемой переменной имели симметричную треугольную форму, однако для этого можно использовать различные виды функции принадлежности.

Таблица 4

Результаты анализа показателя выручки

$h=0$	Величина выручки, тыс. руб.				
	Факт	Нижняя граница диапазона	Середина нечеткого диапазона	Верхняя граница диапазона	Ширина нечеткого диапазона
Данные для прогноза	71	53,2	70,3	87,3	17,
	46,4	46,4	58,9	71,4	12,5
	42,5	35,1	46,8	58,4	11,6
	47,6	45,5	58,3	71,2	12,8
	35,4	30,6	40,3	50	9,7
	42,5	38,7	52,4	66	13,7
	36,9	27,9	36,3	44,7	8,4
	34,1	31,9	41,5	51,1	9,6
	21,3	17,1	20,7	24,3	3,6
	23,7	22	27,1	32,3	5,1
	54,8	33,6	44,2	54,8	10,6
	23,3	20,8	25,8	30,8	5
	25,6	25,6	33,2	40,8	7,6
	37,8	37,8	50	62,2	12,2
	50,7	40,2	54,3	68,5	14,1
Тест	35,9	27,4	34,8	42,3	
	47,7	40	52,6	65,2	
	31,3	29,5	39,6	49,8	

Список использованных источников

1. Ю. А. Зак. Принятие решений в условиях нечетких и размытых данных. М.: Изд. «Либроком», 2018. 352 с.
2. А. С. Птускин. Нечеткие модели и методы в менеджменте. М.: Изд. МГТУ им. Баумана, 2008. 216 с.
3. В. Е. Парфенова. Нечеткие модели принятия оптимальных решений в управлении аграрным производством//Иновации. № 10. 2018. С. 88-92.
4. J.-G. Lin, Q.-Y. Zhuang, C. Huang. Fuzzy Statistical Analysis of Multiple Regression with Crisp and Fuzzy Covariates and Applications in Analyzing Economic Data of China//Computational Economics. 2012. 39. P. 29-49.
5. Yun-Hsi O. Chang, Bilal M. Ayyub Fuzzy regression methods-a comparative assessment//Fuzzy Sets and Systems. Vol. 119 (2). 2001. P. 187-203.
6. M. Sakawa, H. Yano Multiobjective fuzzy linear regression analysis for fuzzy input-output data//Fuzzy Sets and Systems. 1992. V. 47. P. 173-181.
7. O. V. Seraya, D. A. Demin. Linear regression analysis of a small sample of fuzzy input data//Journal of Automation and Information Sciences. 2012. V. 44. Issue 7. P. 34-48.
8. H. Tanaka, J. Watada. Possibilistic linear systems and their application to the linear regression model//Fuzzy sets and systems. 1988. Vol. 27. № 3. P. 275-289.
9. Л. А. Заде. Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976. 168 с.

Fuzzy regression modeling in tasks of management of agrarian production

V. E. Parfenova, doctor of economical sciences, professor, Saint-Petersburg state agrarian university.

The problem of accounting for and modeling uncertainty in the modern tasks of management is one of the most pressing. The efficacy of decisions depends significantly on the methods of description available in the challenge of uncertainty. The greatest development in agricultural science has received optimization and econometric models. However, they both are based on quantitative deterministic baseline information and consideration uncertainty as randomness, for which apply probabilistic and statistical methods.

Meanwhile, many modern decision-making tasks in the planning and management of agricultural production are characterized by the presence of indeterminate factors and the availability of high-quality, inaccurate or incomplete information. To account for and describe this uncertainty needs probabilistic approach an alternative approach. One of the most effective mathematical tools aimed at formalizing and processing of uncertain information are the methods of theory of fuzzy sets. Section econometrics associated with the use of fuzzy set theory in regression analysis develops the methods of fuzzy regression modeling.

This article discusses the application possibilities of the fuzzy regression modeling for analysis of management processes in agricultural production.

Keywords: fuzzy modeling, regression analysis, agricultural production.

ТАЛОН ПОДПИСКИ ЖУРНАЛА



Подписка в редакции — это получение журнала сразу после тиража.

В редакции можно оформить подписку на 2019 год (с 1 по 12 номер) по льготной цене **18840 руб. 00 коп.** (Восемнадцать тысяч восемьсот сорок рублей 00 коп.), в том числе НДС — 1 712 руб. 73 коп.

Название организации _____

Фамилия, имя, отчество _____

Должность _____

Почтовый адрес (адрес доставки) _____

Просим высылать нам журнал «Иновации» в количестве _____ экземпляров.

Нами уплачена сумма _____

Платежное поручение № _____ от _____ 20 __ г.

Банковские реквизиты редакции:

ООО «ТРАНСФЕР-ИННОВАЦИИ», ИНН 7813280766, КПП 781301001
 р/с 40702810727000001308 ПАО «Банк Санкт-Петербург», г. Санкт-Петербург»,
 к/с 30101810900000000790, БИК 044030790

Дата заполнения талона подписки _____ Подпись _____

Подписка оформляется с любого номера.

Заполненный талон подписки мы принимаем по факсу: **(812) 234-09-18**

Контактное лицо: А. Б. Каминская.



ТАЛОН ПОДПИСКИ ЖУРНАЛА