Regression with Evidential Coefficients

Alexander Lepskiy(✉)

Higher School of Economics, 20 Myasnitskaya Ulitsa, 101000 Moscow, Russia
alepskiy@hse.ru
https://www.hse.ru/en/org/persons/10586209

Abstract. This article considers a regression model with coefficients that are specified by bodies of evidence defined on the numerical axis. Optimization problems of finding such evidential coefficients have been set. The relationship between the evidential formulation of the problem and some formulations of fuzzy regression problems is shown. The advantages of evidential regression compared to possibilistic fuzzy regression (better robustness, lower degree of fuzziness of coefficients) are demonstrated using a numerical example.

Keywords: Evidential regression · Fuzzy regression

1 Introduction

Regression data analysis is one of the central machine learning techniques. It involves finding a relationship between input and output observed variables (sample data). There are many techniques for estimating the regression function. First, there are methods of mathematical statistics (for example, the method of least squares), based on the assumption that the input and output variables are sample values of random variables. Secondly, these are machine learning methods [6]: K-nearest neighbor smoother, SVM regression [5], regularization-based methods (ridge regression [7], Lasso method [15]), splines, etc.

Classical regression methods assume that the data sources are reliable, and the data itself is accurate. However, in a few problems this may be far from the case. Therefore, the problem of regression analysis with imprecise and uncertain (unreliable) data is relevant.

Data inaccuracy can be modeled by fuzzy sets. As a result, fuzzy regression methods were developed. First of all, possibilistic [14] and metric [4] methods are distinguished.

Uncertainty in data can be modeled using evidence theory [2,13], which has many applications, including in forecasting problems (see, for example, [9–11]). The EVREG (EVidential REGression) method [12] was one of the first methods

This article is an output of a research project “Models and methods for decision-making under deep uncertainty” implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE University).

C. Kahraman et al. (Eds.): INFUS 2025, LNNS 1531, pp. 1–8, 2025.
https://doi.org/10.1007/978-3-031-98304-7_24

of regression analysis. This method is implemented based on the K-nearest neighbor method and belongs to the group of local (nonparametric) regression analysis methods. For each neighboring element, a simple body of evidence is formed, associated with the corresponding output observable value (singleton) and with a mass depending on the distance between the neighboring and the current element. After which, all the simple bodies of evidence of the K neighbors are aggregated using Dempster's rule into one common body of evidence, which determines the predicted value. Recently [3] a new regression analysis model called ENNreg (Evidential Neural Network regression) was proposed, which computes a prediction in the form of a so-called Gaussian random fuzzy number.

In this paper, a new approach will be proposed that develops the possibilistic model [14] for finding fuzzy linear regression coefficients. At the same time, this model will reflect information about the degree of belief in the found imprecise (interval or fuzzy) regression coefficients.

The remainder of the article is structured as follows. Section 2 provides the necessary background from evidence theory and fuzzy sets. In Sect. 3, two evidential regression models are introduced. The first is a model with consistent unblurred evidential coefficients (Subsect. 3.1). The second is a model with fuzzy focal elements (Subsect. 3.2). Section 4 presents a numerical example.

2 Necessary Background from the Theory of Evidence

Below we will consider bodies of evidence defined on some subset $X \subseteq \mathbb{R}^n$. Let 2^X be the set of all subsets of X. A pair $F = (\mathcal{A}, m)$ is called a body of evidence [13] on the X if the $\mathcal{A}$ is a finite set of subsets of X, which are called focal elements; $m : 2^X \to [0, 1]$ is a mass function satisfying the conditions: $m(A) > 0 \Leftrightarrow A \in \mathcal{A}$, $\sum_{A \in \mathcal{A}} m(A) = 1$. Below we will consider both non-blurred and fuzzy focal elements $\tilde{A}$, which are normal, i. e. $h_{\tilde{A}} = \sup_x \mu_{\tilde{A}}(x) = 1$, where $\mu_{\tilde{A}}$ is the membership function of the fuzzy set $\tilde{A}$.

A body of evidence $F_A = (\{A\}, 1)$ with one focal element is called categorical. In particular, F_X is a vacuous body of evidence. An arbitrary body of evidence $F = (\mathcal{A}, m)$ can be represented as a convex sum of categorical bodies of evidence $F = \sum_{A \in \mathcal{A}} m(A) F_A$. We will consider the simple bodies of evidence of the form $F_A^\alpha = (1 - \alpha) F_A + \alpha F_X$, $\alpha \in [0, 1]$. The body of evidence $F = (\mathcal{A}, m)$ is called consonant if $A' \subseteq A''$ or $A'' \subseteq A'$ is true for any $A', A'' \in \mathcal{A}$.

The belief function $Bel(A) = \sum_{B \subseteq A} m(B)$ and the plausibility function $Pl(A) = \sum_{A \cap B \neq \emptyset} m(B)$ are assigned to the body of evidence $F = (\mathcal{A}, m)$. These functions are the lower and upper estimates of the probability of an event. The function $Pl(x) = \sum_{B : x \in B} m(B)$ is called the contour function (for evidence bodies with normal fuzzy focal elements $Pl(x) = \sum_{\tilde{B}} m(\tilde{B}) \mu_{\tilde{B}}(x)$). It is known that the contour function coincides with the possibility distribution function for consonant bodies of evidence. In this case, the contour function can be considered as a membership function of a fuzzy set.

The degree of uncertainty of the body of evidence $F = (\mathcal{A}, m)$ is characterized using the functional H [1]. Below we will use the functional $H(A) =$

$\sum_{A\in\mathcal{A}} m(A)\lambda(A)$, where λ is the Lebesgue measure (if $A = \tilde{A}$ is a fuzzy set with an integrable membership function $\mu_{\tilde{A}}$, then $\lambda(\tilde{A}) = \int_X \mu_{\tilde{A}}(t)\lambda(dt)$).

Evidence theory provides great opportunities for aggregating bodies of evidence. Below we will use only the conjunctive rule of combination, which forms a new body of evidence $F = (\mathcal{A}, m) = \otimes_{k=1}^{l} F_k$ from k bodies of evidence $F_k = (\mathcal{A}_k, m_k)$, $k = 1,\ldots,l$ according to the rule:

$$m(A) = \sum_{\substack{A_1\cap\ldots\cap A_l = A,\\ A_k\in\mathcal{A}_k, k=1,\ldots,l}} m_1(A_1)\cdot\ldots\cdot m_l(A_l). \tag{1}$$

For non-conflicting bodies of evidence (i. e. $A_1 \cap \ldots \cap A_l \neq \emptyset\ \forall A_k \in \mathcal{A}_k$, $k = 1,\ldots,l$), which are the only ones considered in this article, this rule coincides with Dempster's rule [2]. There are various ways of generalizing Dempster's rule to the case of bodies of evidence with fuzzy focal elements. Below we will consider the approach of Ishizuka [8] (it is assumed that all fuzzy sets are normal):

$$m(\tilde{A}) = \frac{1}{k} \sum_{\substack{\tilde{A}_1\cap\ldots\cap\tilde{A}_l = \tilde{A},\\ \tilde{A}_k\in\mathcal{A}_k, k=1,\ldots,l}} h_{\tilde{A}_1\cap\ldots\cap\tilde{A}_l} m_1(\tilde{A}_1)\cdot\ldots\cdot m_l(\tilde{A}_l), \tag{2}$$

where $k = \sum_{\tilde{A}_k\in\mathcal{A}_k, k=1,\ldots,l} h_{\tilde{A}_1\cap\ldots\cap\tilde{A}_l} m_1(\tilde{A}_1)\cdot\ldots\cdot m_l(\tilde{A}_l)$.

3 Statement of the Evidential Regression Problem

We consider the problem of approximating point data $\{(x_i, y_i)\}_{i=1}^N$ by a function $f(x; A_0,\ldots,A_n)$, where A_j, $j = 0,\ldots,n$ are bodies of evidence. For simplicity, we will further consider only the case of paired linear regression:

$$f(x; A_0, A_1) = A_0 + A_1 x.$$

In general, we will assume that evidential coefficients are simple bodies of evidence, each of which is determined by one focal element (an interval or a fuzzy number) and the degree of confidence that the true value of the coefficient belongs to this element.

3.1 Evidential Regression with Interval Coefficients

We will assume that evidential coefficients are simple bodies of evidence $A_j^\alpha = (1-\alpha)F_{[a_j,b_j]} + \alpha F_{X_j}$, where $X_j \subseteq \mathbb{R}$, $j = 0, 1$. Note that this can be written as

$$F_D^\alpha = (1-\alpha)F_D + \alpha F_\Pi, \tag{3}$$

where Π, D are some rectangles in $\mathbb{R}^2$, $D \subseteq \Pi$ defined by lines with the desired parameters, possible parameters and sample data.

Without loss of generality, we can assume that all points are ordered by the first coordinate in the sample $\{(x_i, y_i)\}_{i=1}^N$: $x_1 \leq \ldots \leq x_N$. Let $L_{\mathbf{w}}(x) = L_{(w_0,w_1)}(x) = w_0 + w_1 x$. The model will consist of the following steps.

1. To determine X_0, X_1 (or Π) from the sample $\{(x_i, y_i)\}_{i=1}^N$, we find the regression line $L_{\mathbf{c}}(x)$, the parameters of which $\mathbf{c} = (c_0, c_0)$ (assuming that the errors are distributed according to the normal law $N(0, \sigma^2)$) are calculated using the formulas:
$$c_1 = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2}, \ c_0 = \bar{y} - c_1\bar{x}, \ \bar{x} = \frac{1}{N}\sum_{i=1}^N x_i, \ \bar{y} = \frac{1}{N}\sum_{i=1}^N y_i.$$
Next, we find the projections of the points $\{(x_i, y_i)\}_{i=1}^N$ onto the line $L_{\mathbf{c}}$. As a result, we obtain the points $P'(x_i', y_i') \in L_{\mathbf{c}}$, $i = 1, \ldots, N$. Let $\underline{P}' = \arg\min_{P'_i} x_i'$, $\bar{P}' = \arg\max_{P'_i} x_i'$ and $\underline{L}'$, $\bar{L}'$ be two lines passing through points $\underline{P}'$, $\bar{P}'$ and perpendicular to $L_{\mathbf{c}}$. Similarly, two other parallel lines $\underline{L}''$, $\bar{L}''$ are defined, parallel to each other and parallel to the line $L_{\mathbf{c}}$. The lines $\underline{L}'$, $\bar{L}'$, $\underline{L}''$, $\bar{L}''$ will limit the rectangle Π, containing all sample points, and located "along" the regression line.
2. In the new coordinate system (x', y'), connected with the line $L_{\mathbf{c}}$ ($O' = L_{\mathbf{c}} \cap \underline{L}'$ is the origin of coordinates, the axis $O'x'$ coincides with the line $L_{\mathbf{c}}$), we find two lines of "tolerances" of the form $L_{\mathbf{\Delta}}^{\pm}(x') = \pm\Delta_0 \pm \Delta_1 x'$ which, together with the rectangle Π', limit the domain $D'_{\mathbf{\Delta}} = \left\{(x', y') : L_{\mathbf{\Delta}}^{-}(x') \le y' \le L_{\mathbf{\Delta}}^{+}(x') \ \forall x' \in \Pi'_{x'}\right\}$ and satisfy the condition $\frac{1}{N}\left|\{i : ({x'}_i, {y'}_i) \in {D'}_{\mathbf{\Delta}}\}\right| \ge 1 - \alpha$. Note that in the general case we have $D'_{\mathbf{\Delta}} \not\subseteq \Pi'$.
We will assume with respect to the "tolerances" that $\Delta_0 \ge 0$, $\Delta_1 \in \mathbb{R}$ are true, but in any case we have $L_{\mathbf{\Delta}}^{-}(x') \le y' \le L_{\mathbf{\Delta}}^{+}(x') \ \forall x' \in {\Pi'}_{x'}$.
We will consider the "tolerances" $\mathbf{\Delta} = (\Delta_0, \Delta_1)$ to be optimal, which minimize the area of the domain $D'_{\mathbf{\Delta}}$: $\mathbf{\Delta}^{(opt)} = \arg\min_{\mathbf{\Delta}} S(D'_{\mathbf{\Delta}})$, $D' = \arg\min_{{D'}_{\mathbf{\Delta}}} S(D'_{\mathbf{\Delta}})$. If $D' \not\subseteq \Pi'$, then we "expand" Π' in a minimal way along the axis $O'y'$ so that the inclusion is satisfied. Returning to the original coordinate system, we obtain the desired rectangles Π, D and the corresponding parameters of the evidential regression lines.
3. The problem of step 2 is solved for l values $0 < \alpha_1 < \ldots < \alpha_l \le 1$. As a result, we obtain l simple bodies of evidence $F_{D_k}^{\alpha_k}$, $k = 1, \ldots, l$ of the form (3), where $\Pi = \bigcup_{k=1}^l \Pi_k$.
4. Simple bodies of evidence $F_{D_k}^{\alpha_k}$, $k = 1, \ldots, l$ are aggregated using the conjunctive rule $F = \otimes_{k=1}^l F_{D_k}^{\alpha_k}$ according to formula (1). As a result, we will obtain the final body of evidence, which will determine the evidential regression coefficients.

3.2 Conjunctive Aggregation of Jointly Consonant Bodies of Evidence

The simple bodies of evidence $F_{D_k}^{\alpha_k}$, $k = 1, \ldots, l$ obtained in Subsect. 3.1 may turn out to be jointly consonant, i. e. there exists a permutation of indices such that $D_{i_1} \subseteq \ldots \subseteq D_{i_l} \subseteq \Pi$. In this case, the aggregation of such bodies of evidence is significantly simplified.

Proposition 1. *The body of evidence* $F = (\mathcal{A}, m) = \otimes_{k=1}^{l} F_{D_k}^{\alpha_k}$ *obtained as a result of conjunctive aggregation (1) of jointly consonant simple bodies of evidence* $F_{D_k}^{\alpha_k}$, $k = 1, \ldots, l$ *will be consonant,* $\mathcal{A} = \{D_1, \ldots, D_l, \Pi\}$. *Moreover, if* $D_{i_1} \subseteq \ldots \subseteq D_{i_l} \subseteq \Pi$, *then (assuming* $\alpha_{i_0} = 1$*) we have*

$$m(D_{i_k}) = (1 - \alpha_{i_k}) \alpha_{i_0} \alpha_{i_1} \ldots \alpha_{i_{k-1}}, \; k = 1, \ldots, l, \; m(\Pi) = \alpha_{i_1} \ldots \alpha_{i_l}.$$

Corollary 1. *If* $F_{D_k}^{\alpha_k}$, $k = 1, \ldots, l$ *are jointly consonant simple bodies of evidence and* Pl *is the plausibility function of their conjunctive aggregation (1)* $\otimes_{k=1}^{l} F_{D_k}^{\alpha_k}$, *then for ordering* $\emptyset = D_{i_0} \subseteq D_{i_1} \subseteq \ldots \subseteq D_{i_l} \subseteq D_{i_{l+1}} = \Pi$ *we have:*

$$Pl(x) = \alpha_{i_0} \ldots \alpha_{i_{k-1}} \text{ if } x \in D_{i_k} \backslash D_{i_{k-1}}, \; k = 1, \ldots, l+1, \; \alpha_{i_0} = 1.$$

3.3 Evidential Regression with Fuzzy Coefficients

Let us consider the case when the regression coefficients can be fuzzy numbers. For example, these could be triangular fuzzy numbers $\tilde{a}_j = (c_j - \Delta_j, c_j, c_j + \Delta_j)$, $\Delta_j \geq 0$, $j = 0, 1$. Then the information that the regression coefficients are such fuzzy numbers with belief level $1 - \alpha$ can be represented by a simple body of evidence

$$A_j^{\alpha} = (1 - \alpha) F_{(c_j - \Delta_j, c_j, c_j + \Delta_j)} + \alpha F_{X_j}, \; j = 0, 1,$$

where $X_j \subseteq \mathbb{R}$, $j = 0, 1$ are universal sets on which triangular symmetric fuzzy numbers are considered; X_j must be consistent with the data sample $\{(x_i, y_i)\}_{i=1}^{N}$ (for example, they can be defined as in Subsect. 3.1).

Let us set the problem of finding triangular fuzzy numbers (parameters c_j, $\Delta_j \geq 0$, $j = 0, 1$) for which the uncertainty functional $\Phi(H(A_0), H(A_1))$ (Φ is the convolution of two uncertainties) of the bodies of evidence would be minimal and (by analogy with Tanaka's approach [14]) at least $(1 - \alpha)N$ sample points would fall into a given h-cut of the model solution $(\tilde{a}_0)_h + (\tilde{a}_1)_h x$. It is easy to show that the last requirement is equivalent to the condition

$$|\{i : |c_0 + c_1 x_i - y_i| \leq (1 - h)(\Delta_0 + \Delta_1 |x_i|)\}| \geq (1 - \alpha)N. \tag{4}$$

If we use the value $H(A) = \sum_{\tilde{a} \in \mathcal{A}} m(\tilde{a}) |\tilde{a}|$, as a measure of the uncertainty of the body of evidence $A = \sum_{\tilde{a} \in \mathcal{A}} m(\tilde{a}) F_{\tilde{a}}$, then

$$H(A_j^{\alpha}) = (1 - \alpha)\Delta_j + \alpha |X_j|, \; j = 0, 1.$$

Since for fixed X_0, X_1 and α the minimization of $\Phi(H(A_0^{\alpha}), H(A_1^{\alpha}))$ is reduced to the minimization of the functional $\tilde{\Phi}(\Delta_0, \Delta_1)$, which depends only on Δ_0, Δ_1, then for a stable solution we add to the minimized function the mean squared error (MSE) between the model values $c_0 + c_1 x_i$ with coefficients from the kernels of fuzzy numbers and the sample values y_i: $\sum_{i \in I} (c_0 + c_1 x_i - y_i)^2$, where I is the set of indices satisfying the condition $I = \{i : |c_0 + c_1 x_i - y_i| \leq (1 - h)(\Delta_0 + \Delta_1 |x_i|)\}$.

Then the function to be minimized takes the form

$$\sum_{i \in I} (c_0 + c_1 x_i - y_i)^2 + \theta \tilde{\Phi}(\Delta_0, \Delta_1) \text{ under condition (4)},$$

where $\theta > 0$. If we solve this problem for l values $0 < \alpha_1 < \ldots < \alpha_l \leq 1$, we obtain l pairs of simple bodies of evidence $A_0^{\alpha_k}$, $A_1^{\alpha_k}$, $k = 1, \ldots, l$ with fuzzy focal elements, which can then be aggregated using the conjunctive rule (2). As a result, we obtain evidential coefficients $A_j = \otimes_{k=1}^{l} A_j^{\alpha_k}$, $j = 0, 1$.

4 Numerical Example

Let us present the results of evidential regression on synthetic (heteroscedastic) data $\{(x_i, y_i)\}_{i=1}^{30}$: $x_i \sim N(\frac{1}{3}i + 1, 0.0009)$, $i = 1, \ldots, 30$; $y_i \sim N(1 + 2x_i, 16)$, $i = 1, \ldots, 10$ and $y_i \sim N(1 + 2x_i, 4)$, $i = 11, \ldots, 30$.

Evidential Regression with Interval Coefficients. For three values $\alpha_1 = 0.1$, $\alpha_2 = 0.3$, $\alpha_3 = 0.5$, we find the optimal domains (see Subsect. 3.1) $\Pi = D_0 \supseteq D_1 \supseteq D_2 \supseteq D_3$ (see Fig. 1(a)), the boundaries of which correspond to the boundary values of the intervals of the regression coefficients. Note that the intersection of the domain D_i with the previous domain D_{i-1} without changing the set of points belonging to each of the regions was considered instead of the region D_i itself in some cases. As a result, we obtain jointly consonant bodies of evidence, from which a contour function can be constructed (see corollary 1):

$$Pl(x, y) = \begin{cases} 1, & (x, y) \in D_3, \\ 0.5, & (x, y) \in D_2 \backslash D_3, \\ 0.15, & (x, y) \in D_1 \backslash D_2, \\ 0.015, & (x, y) \in \Pi \backslash D_1. \end{cases}$$

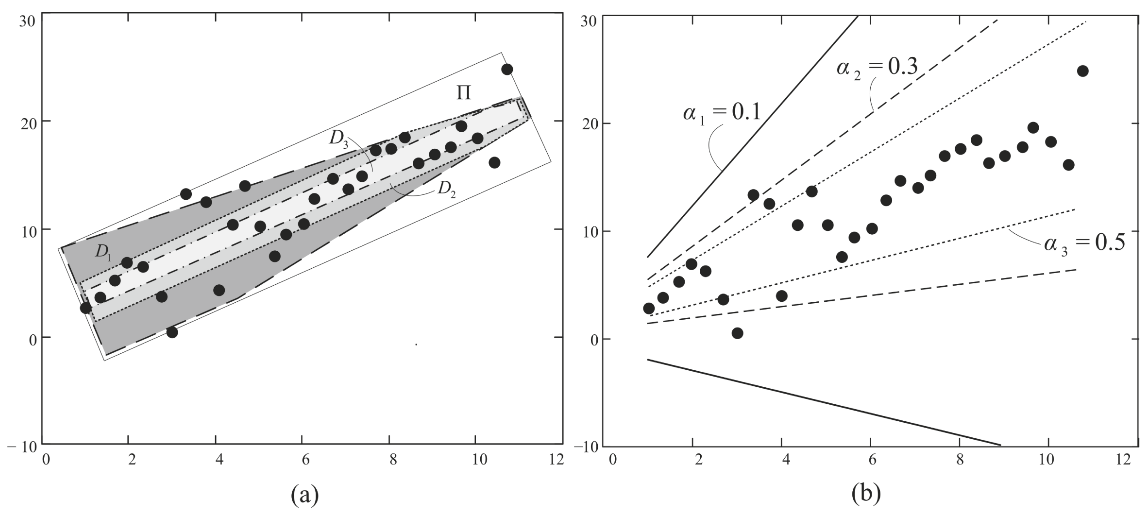

Fig. 1. (a) Optimal domains for evidential regression; (b) Evidential regression with fuzzy coefficients.

This function can be viewed as a membership function of a fuzzy set that determines the regression coefficients.

Evidential Regression with Fuzzy Coefficients. The results of evidential regression with fuzzy coefficients (see Subsect. 3.2; only the support boundaries are visualized) for values $\alpha_1 = 0.1$, $\alpha_2 = 0.3$, $\alpha_3 = 0.5$ and $h = 0.7$, $\Phi(t, s) = 0.6t + 0.4s$ are shown in Fig. 1(b). The coefficients are triangular fuzzy numbers $\tilde{a}_0$ and $\tilde{a}_1$, the parameters of which for three values α are given in Table 1.

Table 1. Fuzzy evidential regression coefficients.

	$\tilde{a}_0$	$\tilde{a}_1$
$\alpha_1 = 0.1$	$(-0.89,\ 1.02,\ 2.94)$	$(-1.02,\ 1.85,\ 4.72)$
$\alpha_2 = 0.3$	$(0.62,\ 1.45,\ 2.28)$	$(0.56,\ 1.81,\ 3.06)$
$\alpha_3 = 0.5$	$(1.38,\ 1.88,\ 2.37)$	$(1.02,\ 1.76,\ 2.51)$

We will perform aggregation $A_j = \otimes_{k=1}^{3} A_j^{\alpha_k}$, $j = 0, 1$ of simple bodies of evidence-coefficients using Dempster's rule (2) for fuzzy focal elements. As a result, we will obtain the final evidential coefficients. The graph of the contour function Pl_1 of the evidential coefficient A_1 is shown in Fig. 2. Note that the focal fuzzy elements are not strictly consonant (though they are close to being so). So, we have $\max_t Pl_1(t) \approx 0.991 < 1$. At the same time, the degrees of fuzziness and ambiguity in this case will be significantly lower than in the case of fuzzy regression, which would result in triangular fuzzy coefficients. The appearance of the contour function can be used to judge the presence of outliers and/or heteroscedasticity of the data.

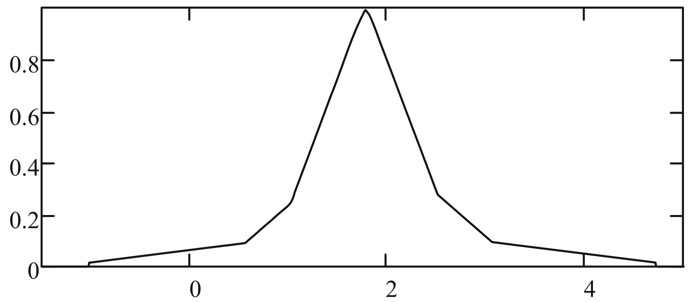

Fig. 2. Contour function of evidential coefficient A_1.

5 Conclusion

The article considers two models of evidential regression. The first model is a model with interval focal elements, and the second is a model with fuzzy focal elements. New models allow us to obtain more complete information about the desired coefficients with a lower degree of blurring compared to fuzzy regression models. In addition, these models will be more robust compared to possibility models of fuzzy regression, since the method does not require that all sample elements (including outliers) belong to a given cutting set.

Generalization of the proposed methodology to the case of multiple regression (including the development of computational procedures) is an important problem for further research.

References

1. Bronevich, A., Lepskiy, A.: Imprecision indices: axiomatic, properties and applications. Int. J. Gen. Syst. **44**(7–8), 812–832 (2015)
2. Dempster, A.P.: Upper and lower probabilities induced by multivalued mapping. Ann. Math. Stat. **38**, 325–339 (1967)
3. Denœux, T.: Quantifying prediction uncertainty in regression using random fuzzy sets: the ENNreg model. IEEE Trans. Fuzzy Syst. **31**(10), 3690–3699 (2023)
4. Diamond, P.: Fuzzy least squares. Inf. Sci. **46**, 141–157 (1988)
5. Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V.: Support vector regression machines. In: Mozer, M., Jordan, M., Petsche, T. (eds.) Advances in Neural Information Processing Systems, vol. 9, pp. 155–161. MIT Press, Cambridge (1997). https://doi.org/10.5555/2998981.2999003
6. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer, Heidelberg (2001)
7. Hoerl, A.E., Kennard, R.W.: Ridge regression: biased estimation for nonorthogonal problems. Technometrics **12**(1), 55–67 (1970)
8. Ishizuka, M.: An extension of Dempster & Shafer's theory to fuzzy set for constructing expert systems. Seisan-Kenkyu **34**(7), 312–315 (1982)
9. Lepskiy, A., Suevalov, A.: Application of the belief function theory to the development of trading strategies. Procedia Comput. Sci. **162**, 235–242 (2019)
10. Denœux, T., Lefèvre, E., Liu, Z., Pichon, F. (eds.): BELIEF 2021. LNCS (LNAI), vol. 12915. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-88601-1
11. Lepskiy, A.: On optimal blurring of point expert estimates and their aggregation in the framework of evidence theory. Procedia Comput. Sci. **214**, 573–580 (2022)
12. Petit-Renaud, S., Denœux, T.: Nonparametric regression analysis of uncertain and imprecise data using belief functions. Intern. J. Approx. Reas. **35**(1), 1–28 (2004)
13. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
14. Tanaka, H.: Fuzzy data analysis by possibilistic linear models. Fuzzy Sets Syst. **24**, 363–375 (1987)
15. Tibshirani, R.: Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Series B (Methodological) **58**(1), 267–288 (1996)