Interval regression by tolerance analysis approach

Milan Hladík[*] Michal Černý[†]

Abstract

In interval regression analysis, we are given crisp or interval data and we are to determine appropriate interval regression parameters. There exist different methods for dealing with this problem; many of them possess the property that some of the resulting regression parameters are crisp. This property is undesirable in a variety of applications. To overcome this drawback we propose a method motivated by tolerance analysis in linear systems. This method is not only computationally very cheap, but also yields intervals for regression parameters the widths of which are proportional to an in-advance given vector of parameters. For example, one choice of this vector allows to control relative tolerances and another leads to absolute tolerances. We show how the basic method, formulated for the model of crisp input – crisp output data, can be extended to interval data. For the interval-valued case we propose several formulations for the solution concept. We illustrate our approach by examples.

Keywords: *Possibilistic regression, fuzzy regression, interval regression.*

1 Introduction

In the traditional linear regression, we are to find a linear relationship between dependent and independent variables. Due to vague essence and uncertainties in many real-life problems, often it is more natural to seek a

[*]Charles University, Faculty of Mathematics and Physics, Department of Applied Mathematics, Malostranské nám. 25, 11800, Prague, Czech Republic, e-mail: `milan.hladik@matfyz.cz`

[†]University of Economics, Faculty of Computer Science and Statistics, nám. W. Churchilla 4, 13067, Prague, Czech Republic, e-mail: `cernym@vse.cz`.

relationship by means of interval or fuzzy parameters. Interval regression is a special case of more general fuzzy regression [16, 18]. In view of simplicity and ease of interpretation it was considered by many authors.

There are different concepts for dealing with interval regression models. The *possibilistic concept* developed by Tanaka et al. [18] and later studied e.g. in [16, 17] is the most popular one and was applied in a number of practical problems. For instance, let us mention applications in ergonomics [2], market sales forecasting [4], system identification [12], speech learning systems [14], or analytic hierarchy process [15].

There are different methods known for possibilistic regression. The basic one employs a linear programming formulation [8, 13, 16, 19]. It is quite simple and requires small computational effort; however, the resulting regression parameters are mostly crisp. To overcome this drawback, another approaches were developed using quadratic programming [17], neural networks [9] or support vector machines [3, 6, 7, 10]. The results are still not satisfactory and the methods are computationally more expensive.

Our approach is motivated by tolerance analysis in linear systems [5, 21, 22]. We suppose that we are given some initial crisp estimation of the regression parameters and we then extend them to intervals. This extension is done optimally and the widths of resulting intervals are proportional to a given vector. Moreover, the proposed method is very cheap to implement and calculate.

Let us introduce some notation. Intervals are written in boldface

$$\boldsymbol{a} := [\underline{a}, \overline{a}].$$

The corresponding center and radius are denoted respectively by $a^c := \frac{1}{2}.(\overline{a} + \underline{a})$ and $a^\Delta := \frac{1}{2}.(\overline{a} - \underline{a})$. Similarly for interval vectors and matrices. For interval arithmetic we refer readers to [1]. For a matrix M, by M_{i*} and M_{*i} we understand the i-th row and column, respectively.

The paper is organized as follows. In Section 2 we present our method on a model with *crisp input – crisp output* data. An extension to *crisp input – interval output* is done in Section 3, where several solution concepts are discussed. In Section 4, we introduce the most general *interval input – interval output* model. We generalize diverse solution concepts to this case and show that, under weak assumptions, the problems can be reduced to the simple crisp input – crisp output model. Finally, in Section 5, we demonstrate our approach by several examples.

2 Crisp input – crisp output

The concept of interval regression analysis was proposed by Tanaka et al. [18]. Therein, we are to determine interval regression parameters that comprise all possibilities determined by the model and data. A model in the interval linear regression analysis can be represented as

$$y = X_{*1}\boldsymbol{a}_1 + X_{*2}\boldsymbol{a}_2 + \cdots + X_{*1}\boldsymbol{a}_n = X\boldsymbol{a}$$

or,

$$y_j = x_{j,1}\boldsymbol{a}_1 + x_{j,2}\boldsymbol{a}_2 + \cdots + x_{j,n}\boldsymbol{a}_n = X_{j*}\boldsymbol{a}, \quad j = 1, \dots, p, \tag{1}$$

where $X \in \mathbb{R}^{p \times n}$ is an input matrix, $y \in \mathbb{R}^p$ is an output vector and $\boldsymbol{a}$ is an interval vector of regression parameters. Usually, the first column of X consists of all ones representing the absolute term. Interval coefficients $\boldsymbol{a}_i$, $i = 1, \dots, n$ can be expressed as

$$\boldsymbol{a}_i = [a_i - c_i, a_i + c_i].$$

The right-hand side of (1) is an interval which can be formulated as [6, 7, 13, 17]

$$\left[X_{j*}a - |X|_{j*}c, X_{j*}a + |X|_{j*}c\right]$$

where the absolute value $|\cdot|$ of a matrix is understood componentwise.

Suppose that the real vector $a \in \mathbb{R}^n$ is known. It can have been estimated, for instance, by ordinary least squares as $a = (X^TX)^{-1}X^Ty$ or by other estimators suitable for particular data. In our examples, we will always use the OLS estimator; however note that our method does not depend on the estimator used. Our aim is to determine a non-negative vector c, with as small entries as possible, such that *each equation in* (1) *is satisfiable.* We say that the j-th equation in (1) is *satisfiable* if there is $a' \in \boldsymbol{a}$ such that $y_j = X_{j*}a'$.

Suppose we are given non-negative tolerance rates $c^\Delta \in \mathbb{R}^n$. Typically, they are set as $c^\Delta := |a|$ for relative tolerances and $c^\Delta := 1$ for absolute tolerances. Then the problem is *to find a minimal tolerance quotient* $\delta \geq 0$ *such that* (1) *is satisfiable with* $\boldsymbol{a} := [a - \delta c^\Delta, a + \delta c^\Delta]$, i.e.

$$\forall j \in \{1, \dots, p\}\ \exists a' \in [a - \delta c^\Delta, a + \delta c^\Delta] : y_j = X_{j*}a'. \tag{2}$$

Note that if c^Δ contains zeros then it may happen that for all $\delta \geq 0$ some equation in (1) remains unsatisfiable. Nevertheless, as long as $c^\Delta > 0$, a minimal tolerance quotient exists.

Theorem 1. *If there is $j \in \{1, \dots, p\}$ such that $|X|_{j*}c^\Delta = 0$ and $y_j \neq X_{j*}a$ then there exists no δ satisfying* (2)*. Otherwise let*

$$\delta^* := \max_{j:|X|_{j*}c^\Delta > 0} \frac{|y_j - X_{j*}a|}{|X|_{j*}c^\Delta}, \tag{3}$$

where $\max \emptyset = 0$ *by definition. Then δ^* is the minimal tolerance quotient.*

Proof. Let $j \in \{1, \dots, p\}$ and $|X|_{j*}c^\Delta = 0$. Then for each $i \in \{1, \dots, n\}$ either $x_{ji} = 0$ or $c_i^\Delta = 0$. For every $\delta \geq 0$ and $a' \in [a - \delta c^\Delta, a + \delta c^\Delta]$ we have

$$X_{j*}a' = \sum_{i=1}^{n} x_{ji}a'_i = \sum_{i:c_i^\Delta = 0}^{n} x_{ji}a'_i = \sum_{i:c_i^\Delta = 0}^{n} x_{ji}a_i = X_{j*}a.$$

It means that if $y_j = X_{j*}a$ then δ can be any non-negative number, and if $y_j \neq X_{j*}a$ then the equation cannot be satisfied for any $\delta \geq 0$.

Let $j \in \{1, \dots, p\}$ and $|X|_{j*}c^\Delta > 0$. Define $a' \in \mathbb{R}^n$ as follows

$$a'_i := \begin{cases} a_i + \dfrac{y_j - X_{j*}a}{|X|_{j*}c^\Delta} c_i^\Delta & \text{if } x_{ji} \geq 0, \\ a_i - \dfrac{y_j - X_{j*}a}{|X|_{j*}c^\Delta} c_i^\Delta & \text{otherwise.} \end{cases}$$

Obviously, $a' \in [a - \delta^* c^\Delta, a + \delta^* c^\Delta]$, and

$$\begin{aligned} X_{j*}a' &= \sum_{i=1}^{n} x_{ji}a'_i = \sum_{i:x_{ji} \geq 0} x_{ji}a'_i + \sum_{i:x_{ji} < 0} x_{ji}a'_i \\ &= \sum_{i:x_{ji} \geq 0} x_{ji} \left(a_i + \frac{y_j - X_{j*}a}{|X|_{j*}c^\Delta} c_i^\Delta \right) + \sum_{i:x_{ji} < 0} x_{ji} \left(a_i - \frac{y_j - X_{j*}a}{|X|_{j*}c^\Delta} c_i^\Delta \right) \\ &= \sum_{i=1}^{n} x_{ji}a_i + \sum_{i=1}^{n} |x|_{ji} \frac{y_j - X_{j*}a}{|X|_{j*}c^\Delta} c_i^\Delta \\ &= X_{j*}a + \frac{y_j - X_{j*}a}{|X|_{j*}c^\Delta} \sum_{i=1}^{n} |x|_{ji} c_i^\Delta = y_j. \end{aligned}$$

Therefore (2) holds true.

To prove the minimality of δ^* we suppose that (3) is attained for some $j' \in \{1, \dots, p\}$. Without loss of generality assume that $y_{j'} > X_{j'*}a$. Define

$\delta^{\varepsilon} := \delta^* - \varepsilon$, where $\varepsilon > 0$ is arbitrarily small. We show that the j'-th equation in (1) is not satisfiable. Let $a' \in [a - \delta^{\varepsilon} c^{\Delta}, a + \delta^{\varepsilon} c^{\Delta}]$. Then

$$\begin{aligned} X_{j'*}a' &= \sum_{i=1}^{n} x_{j'i}a'_i \leq \sum_{i:x_{j'i}\geq 0} x_{j'i}(a_i + \delta^{\varepsilon} c_i^{\Delta}) + \sum_{i:x_{j'i}<0} x_{j'i}(a_i - \delta^{\varepsilon} c_i^{\Delta}) \\ &= X_{j'*}a + \sum_{i=1}^{n} |x|_{j'i}\delta^{\varepsilon} c_i^{\Delta} = X_{j'*}a + \left(\frac{y_{j'} - X_{j'*}a}{|X|_{j'*}c^{\Delta}} - \varepsilon\right) \sum_{i=1}^{n} |x|_{j'i}c_i^{\Delta} \\ &= y_{j'} - \varepsilon |X|_{j'*}c^{\Delta} < y_{j'}. \end{aligned}$$

Hence the equality can never be attained. $\square$

There are several advantages of our approach. Firstly, interval coefficients in $\boldsymbol{a}$ are not crisp and their widths are proportional to the prescribed tolerance rates c^{Δ}. Secondly, the formula for calculation of the intervals is simple and its computation is very cheap even for very large regression models. Indeed, the computational complexity is only linear with respect to input data. As shown in Example 2, the widths of the resulting intervals are easy to interpret.

Remark 1. As observed by many authors [7, 8, 10, 13], interval regression is sensitive to outliers. In our approach, it is very easy to deal with the case of data affected by outliers. Let us suppose that we expect or observe m outliers. Then we take the $(m+1)$-st greatest element in (3) instead of the maximal one. Thus, we eliminate m outliers. See also Example 3.

Similarly we can proceed in the problem considered by Lee and Tanaka [13]. They developed a method for inspecting only some fraction, say 60%, of center-located observations. In our approach, we omit 40% of the greatest elements in (3) to obtain a corresponding result for 60% of center-located observations.

Note that this approach to deal with outliers is sensitive to selection of a. We should use an appropriate regression method that is capable to handle outliers; see, for instance, [20].

3 Crisp input – interval output

For a data set with crisp input and interval output there are known several solution concepts, of which the best-known are the *possibility concept* [6, 13, 16, 17, 19] and the *necessity concept* [6, 16, 17, 19]. We will consider both even though the later is less applicable and can fail to find a solution. The third concept that we will discuss is a weaker variant of possibility [16].

Consider an interval linear regression model

$$\boldsymbol{y}_j = x_{j,1}\boldsymbol{a}_1 + x_{j,2}\boldsymbol{a}_2 + \cdots + x_{j,n}\boldsymbol{a}_n = X_{j*}\boldsymbol{a}, \quad j = 1, \ldots, p, \tag{4}$$

where x_{ji}, $j = 1, \ldots, p$, $i = 1, \ldots, n$ are reals and $\boldsymbol{y}_j = [\underline{y}_j, \overline{y}_j]$ are intervals.

In the same manner as in the previous section, we will seek for an interval vector $\boldsymbol{a}$ in the form of

$$\boldsymbol{a} = [a - \delta c^{\Delta}, a + \delta c^{\Delta}],$$

where a is an initial vector of parameters and c^{Δ} is a non-negative vector. Depending on the concept chosen we want to find minimal or maximal $\delta \geq 0$ such that the corresponding interval vector $\boldsymbol{a}$ satisfies certain conditions. Any $\delta \geq 0$ fulfilling the conditions will be called *admissible*.

The initial vector a can be, for example, calculated as the least-squares solution to $y^c = Xa$, that is, $a := (X^T X)^{-1} X^T y^c$. In all examples we will use this setting.

In the possibilistic concept, we have to find a vector of intervals $\boldsymbol{a}$ such that

$$\forall j = 1, \ldots, p \ \forall y'_j \in \boldsymbol{y}_j \ \exists a' \in \boldsymbol{a} : \ y'_j = X_{j*}a',$$

or, using interval arithmetics,

$$\forall j = 1, \ldots, p : \ \boldsymbol{y}_j \subseteq X_{j*}\boldsymbol{a}.$$

In other words, each equation and each output realization is fulfilled for some realization of the interval vector $\boldsymbol{a}$.

This problem can be reduced to the crisp input – crisp output model. Let us define $y_j \in \boldsymbol{y}_j$:

$$y_j = \begin{cases} \underline{y}_j & \text{if } |\underline{y}_j - X_{j*}a| \geq |\overline{y}_j - X_{j*}a|, \\ \overline{y}_j & \text{otherwise.} \end{cases} \tag{5}$$

Using the proposed method for the crisp input – crisp output model $y = X\boldsymbol{a}$ we obtain a quotient $\delta^* \geq 0$ from (3). The corresponding range of the right-hand side of (4) is

$$X_{j*}\boldsymbol{a} = [X_{j*}a - \delta^*|X|_{j*}c^{\Delta}, X_{j*}a + \delta^*|X|_{j*}c^{\Delta}].$$

It implies $|y_j - X_{j*}a| \le \delta^* |X|_{j*} c^\Delta$ and consequently for each $y'_j \in \boldsymbol{y}_j$ we have

$$|y'_j - X_{j*}a| \le |y_j - X_{j*}a| \le \delta^* |X|_{j*} c^\Delta.$$

Therefore $\boldsymbol{y}_j \subseteq X_{j*}\boldsymbol{a}$ and δ^* is admissible. The minimality follows from its minimality for $y = X\boldsymbol{a}$.

Under the necessity concept, we want to determine an interval vector $\boldsymbol{a}$, being as wide as possible, such that

$$\forall j = 1, \ldots, p \ \forall a' \in \boldsymbol{a} \ \exists y'_j \in \boldsymbol{y}_j : \ y'_j = X_{j*}a',$$

or, using interval arithmetics,

$$\forall j = 1, \ldots, p : \ \boldsymbol{y}_j \supseteq X_{j*}\boldsymbol{a}.$$

Note that in this case the optimization direction is the opposite one: we are maximizing the width of the resulting intervals $\boldsymbol{a}_i$, $i = 1, \ldots, n$. If $X_{j*}a \notin \boldsymbol{y}_j$ for some $j \in \{1, \ldots, p\}$ then the problem has no solution. Otherwise, we always achieve an optimal solution. The problem can be rewritten as

$$\begin{aligned} X_{j*}\boldsymbol{a} &\le \overline{y}_j, \quad j = 1, \ldots, p, \\ X_{j*}\boldsymbol{a} &\ge \underline{y}_j, \quad j = 1, \ldots, p. \end{aligned}$$

which is a standard tolerance analysis problem [5, 21, 22]. Therefore the optimal tolerance is

$$\delta^* := \min_{j : |X|_{j*} c^\Delta > 0} \left\{ \frac{\overline{y}_j - X_{j*}a}{|X|_{j*} c^\Delta}, \ \frac{X_{j*}a - \underline{y}_j}{|X|_{j*} c^\Delta} \right\}.$$

A weaker possibilistic concept was concerned in Tanaka [16]. Therein, one seeks for an interval vector $\boldsymbol{a}$ such that

$$\forall j = 1, \ldots, p \ \exists y'_j \in \boldsymbol{y}_j \ \exists a' \in \boldsymbol{a} : \ y'_j = X_{j*}a',$$

or, using interval arithmetics,

$$\forall j = 1, \ldots, p : \ \boldsymbol{y}_j \cap X_{j*}\boldsymbol{a} \ne \emptyset.$$

It says that each equation is fulfilled for some realization of the interval output $\boldsymbol{y}$ and the interval vector $\boldsymbol{a}$. Even this problem can be reduced to the crisp input – crisp output model by defining $y_j \in \boldsymbol{y}_j$ as

$$y_j = \begin{cases} X_{j*}a & \text{if } X_{j*}a \in \boldsymbol{y}_j, \\ \underline{y}_j & \text{if } \underline{y}_j > X_{j*}a, \\ \overline{y}_j & \text{otherwise.} \end{cases}$$

The calculated quotient $\delta^* \geq 0$ of that crisp input – crisp output model is the minimal one. This is easy to observe since for each $y'_j \in \boldsymbol{y}_j$ we have, by definition of y_j,

$$|y'_j - X_{j*}a| \geq |y_j - X_{j*}a| = \delta^*|X|_{j*}c^\Delta.$$

Thus if we decrease the value of δ^* the system will not be satisfiable.

4 Interval input – interval output

In this section, we discuss a more general case with not only interval output, but also interval input. Since this case has not been investigated in standard interval regression analysis, we introduce some solution concepts extending the concepts from the previous section.

Consider an interval linear regression model

$$\boldsymbol{y}_j = \boldsymbol{x}_{j,1}\boldsymbol{a}_1 + \boldsymbol{x}_{j,2}\boldsymbol{a}_2 + \cdots + \boldsymbol{x}_{j,n}\boldsymbol{a}_n = \boldsymbol{X}_{j*}\boldsymbol{a}, \quad j = 1, \ldots, p,$$

where $\boldsymbol{x}_{ji} = [\underline{x}_{ji}, \overline{x}_{ji}]$, $j = 1, \ldots, p$, $i = 1, \ldots, n$ and $\boldsymbol{y}_j = [\underline{y}_j, \overline{y}_j]$ are given intervals. Again, we will seek for an interval vector $\boldsymbol{a}$ in the form of

$$\boldsymbol{a} = [a - \delta c^\Delta, a + \delta c^\Delta],$$

where a is an initial vector of parameters and c^Δ is a given non-negative vector. We try to compute the minimal tolerance quotient $\delta \geq 0$ satisfying conditions of the concept used. The initial vector a can be calculated e.g. as the least-squares solution to $y^c = X^c a$, that is, $a := (X^{cT}X^c)^{-1}X^{cT}y^c$, but not always it is the best choice.

Weak possibilistic solution

The first concept of *weak possibility* is an adaptation of [16]. We aim at finding a vector of intervals $\boldsymbol{a}$ such that

$$\forall j = 1, \ldots, p \; \exists X'_{j*} \in \boldsymbol{X}_{j*} \; \exists y'_j \in \boldsymbol{y}_j \; \exists a' \in \boldsymbol{a} : \; y'_j = X'_{j*}a',$$

or, using interval arithmetics,

$$\forall j = 1, \dots, p: \ \boldsymbol{y}_j \cap \boldsymbol{X}_{j*}\boldsymbol{a} \neq \emptyset.$$

In other words, each equation and each output realization is fulfilled for some realization of the interval input and the interval vector $\boldsymbol{a}$. We propose the following reduction to the crisp input – crisp output model. For each $j \in \{1, \dots, p\}$ do the following: If $\boldsymbol{y}_j \cap \boldsymbol{X}_{j*}a \neq \emptyset$ then omit the j-th entry. If $\overline{y}_j < \underline{\boldsymbol{X}_{j*}a}$ then put

$$y'_j := \overline{y}_j, \quad x'_{ji} := \begin{cases} \underline{x}_{ji} & \text{if } a_i \geq 0, \\ \overline{x}_{ji} & \text{otherwise.} \end{cases}$$

Analogously, if $\underline{y}_j > \overline{\boldsymbol{X}_{j*}a}$ we put

$$y'_j := \underline{y}_j, \quad x'_{ji} := \begin{cases} \overline{x}_{ji} & \text{if } a_i \geq 0, \\ \underline{x}_{ji} & \text{otherwise.} \end{cases}$$

This way we get the crisp model $y' = X'\boldsymbol{a}$ which is solved by the method of Section 2. The resulting tolerance quotient δ^* is always admissible by construction of the reduced model, but it need not be minimal; see Example 1. However, under a certain assumption on sign invariancy, it is minimal. If $c^{\Delta} = |a|$, this assumption is usually satisfied since otherwise the tolerance quotient must be at least one. Such a high value would indicate that the model chosen is probably not appropriate and other model should be used.

Proposition 1. *Suppose that $\boldsymbol{a}_i = [a_i - \delta^* c_i^{\Delta}, a_i + \delta^* c_i^{\Delta}]$ does not contain zero in its interior for every $i = 1, \dots, n$. Then δ^* is minimal.*

Proof. The quotient δ^* is given by (3). The maximum in (3) is attained for some $j' \in \{1, \dots, p\}$. Assume that $\overline{y}_{j'} < \underline{\boldsymbol{X}_{j'*}a}$; the case $\underline{y}_{j'} > \overline{\boldsymbol{X}_{j'*}a}$ is analogous. The quotient δ^* was constructed such that the equation $y'_{j'} = \overline{y}_{j'} = \underline{X'_{j'*}\boldsymbol{a}}$ holds. Due to the sign invariancy assumption we have

$$\begin{aligned} \underline{X'_{j'*}\boldsymbol{a}} &= \sum_{a_i \geq 0,\, \underline{x}_{j'i} \geq 0} \underline{x}_{j'i}(a_i - \delta^* c_i^{\Delta}) + \sum_{a_i \geq 0,\, \underline{x}_{j'i} < 0} \underline{x}_{j'i}(a_i + \delta^* c_i^{\Delta}) \\ &+ \sum_{a_i < 0,\, \overline{x}_{j'i} \geq 0} \overline{x}_{j'i}(a_i - \delta^* c_i^{\Delta}) + \sum_{a_i < 0,\, \overline{x}_{j'i} < 0} \overline{x}_{j'i}(a_i + \delta^* c_i^{\Delta}) \\ &= X'_{j'*}a - \sum_{a_i \geq 0} |\underline{x}|_{j'i} \delta^* c_i^{\Delta} - \sum_{a_i < 0} |\overline{x}|_{j'i} \delta^* c_i^{\Delta}. \end{aligned} \tag{6}$$

We will prove that $\underline{X'_{j'*}\boldsymbol{a}} = \underline{\boldsymbol{X}_{j'*}\boldsymbol{a}}$. It suffices to show that $\underline{X'_{j'*}\boldsymbol{a}} \leq \underline{X_{j'*}\boldsymbol{a}}$ for any $X_{j'*} \in \boldsymbol{X}_{j'*}$. Let $X_{j'*} \in \boldsymbol{X}_{j'*}$ and consider the following cases:

1. If $a_i \geq 0$ and $\underline{x}_{j'i} \geq 0$ then
$$\underline{x}_{j'i}(a_i - \delta^* c_i^{\Delta}) \leq x_{j'i}(a_i - \delta^* c_i^{\Delta}).$$
2. If $a_i \geq 0$ and $\underline{x}_{j'i} < 0$ then
$$\underline{x}_{j'i}(a_i + \delta^* c_i^{\Delta}) \leq \begin{cases} x_{j'i}(a_i + \delta^* c_i^{\Delta}) & \text{if } x_{j'i} < 0 \\ x_{j'i}(a_i - \delta^* c_i^{\Delta}) & \text{if } x_{j'i} \geq 0. \end{cases}$$
3. If $a_i < 0$ and $\overline{x}_{j'i} \geq 0$ then
$$\overline{x}_{j'i}(a_i - \delta^* c_i^{\Delta}) \leq \begin{cases} x_{j'i}(a_i - \delta^* c_i^{\Delta}) & \text{if } x_{j'i} \geq 0 \\ x_{j'i}(a_i + \delta^* c_i^{\Delta}) & \text{if } x_{j'i} < 0. \end{cases}$$
4. If $a_i < 0$ and $\overline{x}_{j'i} < 0$ then
$$\overline{x}_{j'i}(a_i + \delta^* c_i^{\Delta}) \leq x_{j'i}(a_i + \delta^* c_i^{\Delta}).$$

Thus we can write

$$\begin{aligned}
\underline{X'_{j'*}\boldsymbol{a}} &\leq \sum_{a_i \geq 0,\, x_{j'i} \geq 0} x_{j'i}(a_i - \delta^* c_i^{\Delta}) + \sum_{a_i \geq 0,\, x_{j'i} < 0} x_{j'i}(a_i + \delta^* c_i^{\Delta}) \\
&+ \sum_{a_i < 0,\, x_{j'i} \geq 0} x_{j'i}(a_i - \delta^* c_i^{\Delta}) + \sum_{a_i < 0,\, x_{j'i} < 0} x_{j'i}(a_i + \delta^* c_i^{\Delta}) \\
&= \underline{X_{j'*}\boldsymbol{a}}.
\end{aligned}$$

Now, we have $\overline{y}_{j'} = \underline{X'_{j'*}\boldsymbol{a}} = \underline{\boldsymbol{X}_{j'*}\boldsymbol{a}}$. If

$$\sum_{a_i \geq 0} |\underline{x}|_{j'i} c_i^{\Delta} + \sum_{a_i < 0} |\overline{x}|_{j'i} c_i^{\Delta} = 0$$

then $\delta^* = 0$. Otherwise, in view of (6), any arbitrarily small decrease of δ^* would imply

$$\overline{y}_{j'} < \underline{X'_{j'*}\boldsymbol{a}} = \underline{\boldsymbol{X}_{j'*}\boldsymbol{a}},$$

and the tolerance quotient would not be admissible. $\square$

Possibilistic solution

The *possibilistic* solution $\boldsymbol{a}$ should satisfy

$$\forall j = 1, \ldots, p \; \forall y'_j \in \boldsymbol{y}_j \; \exists X'_{j*} \in \boldsymbol{X}_{j*} \; \exists a' \in \boldsymbol{a}: \; y'_j = X'_{j*} a',$$

or, using interval arithmetics,

$$\forall j = 1, \ldots, p: \; \boldsymbol{y}_j \subseteq \boldsymbol{X}_{j*} \boldsymbol{a}.$$

In this case, we reduce the model to two crisp input – crisp output models

$$y^1 = X^1 \boldsymbol{a}^1, \quad \text{and} \quad y^2 = X^2 \boldsymbol{a}^2,$$

calculate both the corresponding quotients δ_1^* and δ_2^*, and then take the larger one $\delta^* := \max\{\delta_1^*, \delta_2^*\}$. The first reduction proceeds as follows. For each $j \in \{1, \ldots, p\}$ do: If $\overline{y}_j \in \boldsymbol{X}_{j*} a$ then omit the j-th entry. If $\overline{y}_j < \underline{\boldsymbol{X}_{j*} a}$ then put

$$y_j^1 := \overline{y}_j, \quad x_{ji}^1 := \begin{cases} \underline{x}_{ji} & \text{if } a_i \geq 0, \\ \overline{x}_{ji} & \text{otherwise,} \end{cases}$$

otherwise put

$$y_j^1 := \overline{y}_j, \quad x_{ji}^1 := \begin{cases} \overline{x}_{ji} & \text{if } a_i \geq 0, \\ \underline{x}_{ji} & \text{otherwise.} \end{cases}$$

The second reduction is similar. If $\underline{y}_j \in \boldsymbol{X}_{j*} a$ then omit the j-th entry. If $\underline{y}_j < \underline{\boldsymbol{X}_{j*} a}$ then put

$$y_j^2 := \underline{y}_j, \quad x_{ji}^2 := \begin{cases} \underline{x}_{ji} & \text{if } a_i \geq 0, \\ \overline{x}_{ji} & \text{otherwise,} \end{cases}$$

otherwise put

$$y_j^2 := \underline{y}_j, \quad x_{ji}^2 := \begin{cases} \overline{x}_{ji} & \text{if } a_i \geq 0, \\ \underline{x}_{ji} & \text{otherwise.} \end{cases}$$

In this method, the resulting tolerance quotient δ^* is admissible, but is minimal only under the sign invariance assumption as in the case of weak possibility.

Proposition 2. *Suppose that* $\boldsymbol{a}_i = [a_i - \delta^* c_i^\Delta, a_i + \delta^* c_i^\Delta]$ *does not contain zero in its interior for every* $i = 1, \ldots, n$. *Then* δ^* *is minimal.*

Proof. For every $j = 1, \dots, p$ the interval inclusion

$$\boldsymbol{y}_j \subseteq \boldsymbol{X}_{j*}\boldsymbol{a}$$

holds if and only if both $\underline{y}_j \in \boldsymbol{X}_{j*}\boldsymbol{a}$ and $\overline{y}_j \in \boldsymbol{X}_{j*}\boldsymbol{a}$ are true. When the output is crisp then weak possibility and possibility concepts coincide. Thus we reduced the problem to finding a simultaneous weak possibilistic solution to $\underline{y}_j = \boldsymbol{X}_{j*}\boldsymbol{a}$ and $\overline{y}_j = \boldsymbol{X}_{j*}\boldsymbol{a}$. This proves admissibility and minimality, too. $\square$

Strong possibilistic solution

The third concept introduced is *strong possibility.* Herein, we want to compute $\boldsymbol{a}$ such that

$$\forall j = 1, \dots, p \ \forall y'_j \in \boldsymbol{y}_j \ \forall X'_{j*} \in \boldsymbol{X}_{j*} \ \exists a' \in \boldsymbol{a} : \ y'_j = X'_{j*}a',$$

or, using interval arithmetics,

$$\forall j = 1, \dots, p \ \forall X'_{j*} \in \boldsymbol{X}_{j*} : \ \boldsymbol{y}_j \subseteq X'_{j*}\boldsymbol{a}.$$

Similarly as in the previous concept, we reduce the model to two crisp input – crisp output models

$$y^1 = X^1\boldsymbol{a}^1 \quad \text{and} \quad y^2 = X^2\boldsymbol{a}^2, \tag{7}$$

and we choose the largest tolerance quotient of both models. The crisp models are constructed as follows. For each $j \in \{1, \dots, p\}$ put

$$y^1_j := \underline{y}_j, \quad x^1_{ji} := \begin{cases} \overline{x}_{ji} & \text{if } a_i \geq 0, \\ \underline{x}_{ji} & \text{otherwise,} \end{cases}$$
$$y^2_j := \overline{y}_j, \quad x^2_{ji} := \begin{cases} \underline{x}_{ji} & \text{if } a_i \geq 0, \\ \overline{x}_{ji} & \text{otherwise.} \end{cases}$$

Generally, the resulting tolerance quotient δ^* need not be admissible, but under the sign invariancy assumption it is admissible and minimal.

Proposition 3. *Suppose that* $\boldsymbol{a}_i = [a_i - \delta^* c^\Delta_i, a_i + \delta^* c^\Delta_i]$ *does not contain zero in its interior for every* $i = 1, \dots, n$. *Then* δ^* *is admissible and minimal.*

Proof. Let $j \in \{1, \dots, n\}$ and $X'_{j*} \in \boldsymbol{X}_{j*}$. We show that $\boldsymbol{y}_j \subseteq X'_{j*}\boldsymbol{a}$, or, in other words, $\underline{y}_j \geq \underline{X'_{j*}\boldsymbol{a}}$ and $\overline{y}_j \geq \overline{X'_{j*}\boldsymbol{a}}$. By construction of $\delta^* = \max\{\delta_1^*, \delta_2^*\}$ we have

$$
\begin{aligned}
\underline{y}_j &\geq \underline{X^1_{j*}\boldsymbol{a}^1} = \underline{X^1_{j*}[a - \delta_1^* c^\Delta, a + \delta_1^* c^\Delta]} \\
&= \sum_{a_i \geq 0,\, \overline{x}_{ji} \geq 0} \overline{x}_{ji}(a_i - \delta_1^* c_i^\Delta) + \sum_{a_i \geq 0,\, \overline{x}_{ji} < 0} \overline{x}_{ji}(a_i + \delta_1^* c_i^\Delta) \\
&+ \sum_{a_i < 0,\, \underline{x}_{ji} \geq 0} \underline{x}_{ji}(a_i - \delta_1^* c_i^\Delta) + \sum_{a_i < 0,\, \underline{x}_{ji} < 0} \underline{x}_{ji}(a_i + \delta_1^* c_i^\Delta) \\
&\geq \sum_{a_i \geq 0,\, x'_{ji} \geq 0} x'_{ji}(a_i - \delta_1^* c_i^\Delta) + \sum_{a_i \geq 0,\, x'_{ji} < 0} x'_{ji}(a_i + \delta_1^* c_i^\Delta) \\
&+ \sum_{a_i < 0,\, x'_{ji} \geq 0} x'_{ji}(a_i - \delta_1^* c_i^\Delta) + \sum_{a_i < 0,\, x'_{ji} < 0} x'_{ji}(a_i + \delta_1^* c_i^\Delta) \\
&\geq \underline{X'_{j*}\boldsymbol{a}^1} \geq \underline{X'_{j*}\boldsymbol{a}}
\end{aligned}
$$

Analogously $\overline{y}_j \geq \overline{X'_{j*}a}$ can be proven by exhibiting the second reduced model in (7). The minimality follows from the fact that the tolerance quotient must be so large that both reduced models are admissible. The reduced models are of the crisp input – crisp output type and for that case the minimality of the tolerance quotient was proven. □

Example 1. This example illustrates that tolerance quotients calculated by reduction methods presented in this section need not be minimal. Consider a model involving the equation

$$30 = [7, 9]\boldsymbol{a}_1 - [1, 3]\boldsymbol{a}_2.$$

Let $a = (1, 1)^T$, $c^\Delta = (1, 1)^T$. Since the output is crisp, the weak possibility and possibility concepts coincide. It is easy to see that the optimal tolerance is $\delta^* = 2$ with the model

$$y = \boldsymbol{x}_1[-1, 3] + \boldsymbol{x}_2[-1, 3].$$

However, our method for weak possibility reduces the model to the crisp one

$$30 = 9\boldsymbol{a}_1 - \boldsymbol{a}_2$$

yielding tolerance quotient 2.2. The method for possibility constructs two crisp models, both of which are the same as the previous one, and hence the resulting tolerance is also 2.2.

The tolerance for strong possibility is computed as follows. The interval model is reduced to two crisp models

$$30 = 9\boldsymbol{a}_1^1 - \boldsymbol{a}_2^1 \quad \text{and} \quad 30 = 7\boldsymbol{a}_1^2 - 3\boldsymbol{a}_2^2.$$

The former gives tolerance 2.2 and the latter 2.6. However, the larger of them, 2.6, is not admissible, since e.g.

$$30 = 7\boldsymbol{a}_1 - \boldsymbol{a}_2$$

requires the tolerance quotient to be at least 3.

5 Numerical examples

Example 2. Consider the house price model from [13, 17]

$$y = \boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \boldsymbol{a}_3 x_3 + \boldsymbol{a}_4 x_4.$$

The data are displayed in Table 1, where x_1 represents an absolute term, x_2 stands for quality of material, x_3 is the area of the first floor (m^2) and x_4 is the area of the second floor (m^2). The output in y is the sale price (10000 JPY).

In [17], the following possibility model was obtained by using a linear programming-based method

$$y = [0,\ 0] + [207.533,\ 282.801]x_2 + [5.853,\ 5.853]x_3 + [4.786,\ 4.786]x_4.$$

The quadratic programming-based method yields

$$y = [-8.813,\ 8.813]+[209.337,\ 258.965]x_2+[6.137,\ 6.137]x_3+[4.413,\ 5.387]x_4.$$

In both models, there are some crisp parameters.

Our method proceeds as follows. First, we estimate a as

$$a := (X^T X)^{-1} X^T y = (-239.2748, 264.8359, 7.4659, 6.7569)^T$$

and put $c^\Delta := |a|$, for instance. By Theorem 1, we calculate $\delta^* = 0.0482$. It means that in the resulting regression intervals $\boldsymbol{a}_i$, $i = 1, \ldots, 4$, all entries of a perturb within 4.82% tolerance. Numerically, the model reads

$$y = [-250.817,\ -227.732]+[252.060,\ 277.612]x_2+[7.105,\ 7.826]x_3+[6.431,\ 7.083]x_4.$$

j	y	x_1	x_2	x_3	x_4
1	606	1	1	38.09	36.43
2	710	1	1	62.10	26.50
3	808	1	1	63.76	44.71
4	826	1	1	74.52	38.09
5	865	1	1	75.38	41.10
6	852	1	2	52.99	26.49
7	917	1	2	62.93	26.49
8	1031	1	2	72.04	33.12
9	1092	1	2	76.12	43.06
10	1203	1	2	90.26	42.64
11	1394	1	3	85.70	31.33
12	1420	1	3	95.27	27.64
13	1601	1	3	105.98	27.64
14	1632	1	3	79.25	66.81
15	1699	1	3	120.5	32.25

Table 1: (Example 2) House price model data.

We see that the widths of the resulting intervals are proportional to the entries of a. The result is easy to interpret, and all widths are determined by one quotient: 4.82%.

Example 3. This example is adopted from [8] and shows how our method can handle outliers. Data are displayed in Table 2. The basic data set is indexed be $j = 1, \ldots, 8$ and the last two entries represent outliers.

j	1	2	3	4	5	6	7	8	9	10
x_j	2	4	6	8	10	12	14	16	2	16
y_j	14	16	14	18	18	22	18	22	4	32

Table 2: (Example 3) Crisp input – crisp output with two outliers (column 9 and 10).

The interval regression method based on linear programming yields a model

$$y = [11,\, 13] + [0.5,\, 0.75]x \tag{8}$$

for the data without outliers ($j = 1, \ldots, 8$), and

$$y = [1.667,\ 11.429] + [1.166,\ 1.286]x$$

for the data with outliers ($j = 1, \ldots, 10$). The method by Ishibuchi and Tanaka [8] applied to the whole data set recognizes the two outliers. Hence it gives narrower interval parameters in the resulting model, the same as in (8).

For the case with no outliers ($j = 1, \ldots, 8$), our method starts with an estimation

$$a := (X^T X)^{-1} X^T y = (12.9286, 0.5357)^T$$

and $c^\Delta := |a|$. The optimal tolerance is $\delta^* = 0.1365$ and the corresponding model is illustrated in Figure 1 and is described by

$$y = [11.163,\ 14.694] + [0.462,\ 0.609]x.$$

When we involve outliers ($j = 1, \ldots, 10$), the initial least squares estimation is

$$a := (8.1233, 1.0752)^T.$$

With $c^\Delta := |a|$ we calculate the optimal tolerance $\delta^* = 0.6107$. The resulting model reads

$$y = [3.163,\ 13.084] + [0.419,\ 1.732]x.$$

In Figure 2 we can see that the range is unnecessarily wide. It is caused by the proportionality of intervals and by their symmetry along the entries of a.

Let us improve the model affected by outliers. along Remark 1. By taking the third highest value of (3) we obtain a significantly smaller tolerance $\delta^* = 0.2878$ with the model

$$y = [5.785,\ 10.461] + [0.766,\ 1.385]x.$$

Note that this result can be further improved since the initial estimation a was corrupted by the outliers. As a consequence we recognized only one outlier, the second one is wrong; see Figure 3. If we consider the L_1-norm estimate of a (the “least abscissae method”, which is reducible to linear programming) resulting in

$$a := (12.8571, 0.5714)^T$$

instead of the L_2-norm (least squares method), then the resulting tolerance is $\delta^* = 0.1404$ and the model is

$$y = [11.052,\ 14.662] + [0.491,\ 0.652]x;$$

see Figure 4. Thus to obtain a good model we should choose an appropriate initial estimation a that is more robust if outliers are present.

Example 4. Consider the crisp input – interval output example from [17] with data displayed in Table 3. In the model, data are approximated by a quadratic function

$$y = \boldsymbol{a}_1 + \boldsymbol{a}_2 x + \boldsymbol{a}_3 x^2.$$

j	1	2	3	4	5	6	7	8
x_j	1	2	3	4	5	6	7	8
$\boldsymbol{y}_j$	[15,30]	[20,37.5]	[15,35]	[25,60]	[25,55]	[40,65]	[55,95]	[70,100]

Table 3: (Example 4) Crisp input – interval output.

The quadratic programming method [17] results in the possibility model

$$y = [-2.778,\ 23.704] + [3.704,\ 7.592]x + [0.37,\ 0.37]x^2,$$

whereas the necessity model reads

$$y = [9.259,\ 11.667] + [4.629,\ 6.667]x + [0.37,\ 0.37]x^2.$$

Our method starts with an initial estimate

$$a := (X^T X)^{-1} X^T y^c = (25.1562, -2.4033, 1.2574)^T.$$

Put $c^\Delta := |a|$. According to (5), we set

$$y := (15, 37.5, 15, 60, 25, 40, 95, 70)^T$$

and we get the optimal tolerance $\delta^* = 0.4434$. That is, entries of the vector a may vary within 44.34% tolerance and the resulting possibility model is

$$y = [14.002,\ 36.311] + [-3.469,\ -1.338]x + [0.700,\ 1.815]x^2.$$

Similarly we obtain $\delta^* = 0.106$ for the necessity model

$$y = [22.489,\ 27.824] + [-2.658,\ -2.148]x + [1.124,\ 1.391]x^2.$$

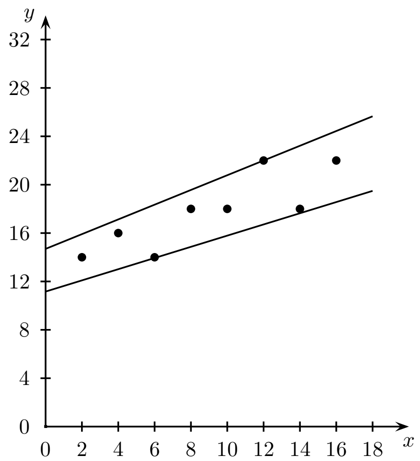

Figure 1: (Example 3) Basic interval regression model without outliers.

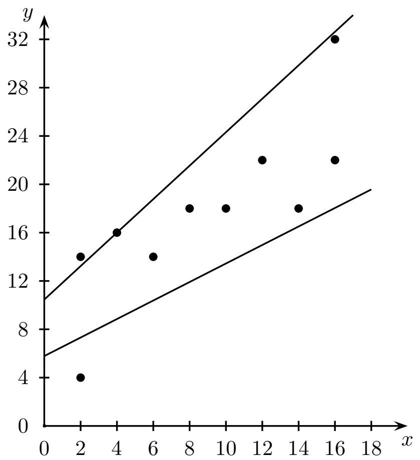

Figure 3: (Example 3) Improved interval regression model; a is obtained by least squares.

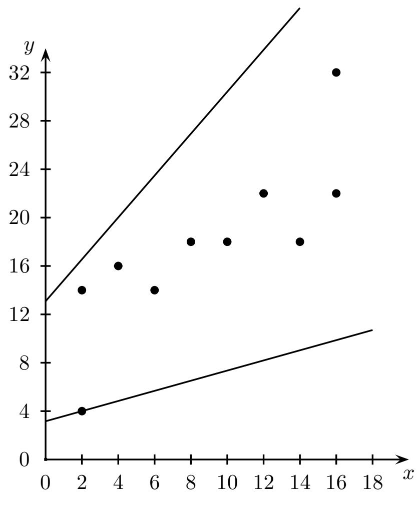

Figure 2: (Example 3) Basic interval regression model with outliers.

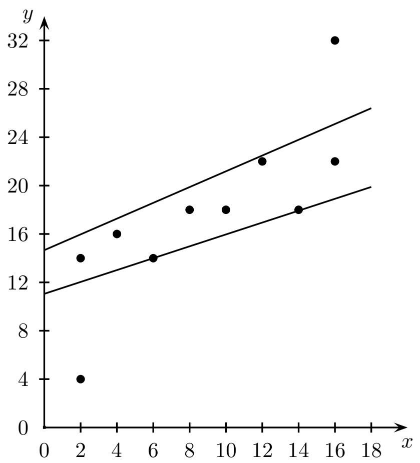

Figure 4: (Example 3) Improved interval regression model; a is obtained by L_1 regression.

Example 5. Consider an interval input – interval output regression model from [11]

$$\boldsymbol{y} = \boldsymbol{a}_1 + \boldsymbol{a}_2\boldsymbol{x}_2 + \boldsymbol{a}_3\boldsymbol{x}_3,$$

where data are displayed in Table 4.

j	$\boldsymbol{y}$	x_1	$\boldsymbol{x}_2$	$\boldsymbol{x}_3$
1	[78,80]	1	[5.5,8.5]	[22,30]
2	[72,76]	1	[0.5,1.5]	[27,31]
3	[100,108]	1	[10,12]	[51,61]
4	[81,95]	1	[9,13]	[28,34]
5	[94,98]	1	[5,9]	[50,54]
6	[105,113]	1	[8,14]	[55,55]
7	[103,103]	1	[3,3]	[69,73]
8	[72,74]	1	[1,1]	[28,34]
9	[93,93]	1	[1,3]	[51,57]
10	[112,120]	1	[18,24]	[45,49]
11	[81,87]	1	[0.5,1.5]	[32,48]
12	[112,114]	1	[7,15]	[60,72]
13	[107,111]	1	[6.5,13.5]	[61,75]

Table 4: (Example 5) Interval input – interval output.

In accordance with [11] we define

$$a := (X^{cT}X^c)^{-1}X^{cT}y^c = (53.1765, 1.4635, 0.6514)^T$$

and set $c^\Delta := |a|$. The tolerance quotient for weak possibility is zero since $\boldsymbol{y}_j \cap \boldsymbol{X}_{j*}\boldsymbol{a} \neq \emptyset$ for every $j \in \{1, \dots, 13\}$.

To determine the tolerance quotient for the possibility concept we have to solve two reduced crisp input – crisp output interval regression models. The first one yields the tolerance quotient $\delta_1^* = 0.0321$ and the second one yields $\delta_1^* = 0.0424$. Thus the resulting tolerance quotient is $\delta^* = 0.0424$ and the corresponding model is

$$\boldsymbol{y} = [50.922,\ 55.431] + [1.401,\ 1.526]\boldsymbol{x}_2 + [0.624,\ 0.679]\boldsymbol{x}_3.$$

The strong possibility concept also requires to solve two reduced models. The first one gives the tolerance quotient $\delta_1^* = 0.1415$ and the second one gives $\delta_2^* = 0.1639$. The larger is the true value and the resulting model is

$$\boldsymbol{y} = [44.463,\ 61.890] + [1.224,\ 1.703]\boldsymbol{x}_2 + [0.545,\ 0.758]\boldsymbol{x}_3.$$

Note that the assumptions of Propositions 2 and 3 are satisfied, so all the presented tolerance quotients are admissible and minimal.

6 Conclusion

In this paper, we adapted tolerance analysis for interval regression analysis. The proposed method is more flexible and the widths of resulting interval parameters are proportional to the 'demand' c^{Δ}. The computational cost is very low and hence the method can by used for huge data sets. The interval parameters are easy to understand and interpret for a user or a decision maker.

There are still some open problems and challenges. As we observed in Example 3, the resulting interval parameters can be slightly wider than it is necessary. It is due to the proportionality to c^{Δ} and symmetry of intervals along a. More narrow intervals may be obtained if we omit symmetry. To develop an effective method for this case is a task for future research.

The presented method is applicable for the crisp input – crisp output case as well as for particular solution concepts in the crisp input – interval output cases. In the more general interval input – interval output case the method works under some general assumptions. To develop a method avoiding these assumptions remains as an open problem.

References

[1] G. Alefeld and J. Herzberger. *Introduction to interval computations*. Computer Science and Applied Mathematics. Academic Press, New York, 1983.

[2] P.-T. Chang, E. S. Lee, and S. A. Konz. Applying fuzzy linear regression to VDT legibility. *Fuzzy Sets and Systems*, 80(2):197–204, 1996.

[3] P.-Y. Hao. Interval regression analysis using support vector networks. *Fuzzy Sets and Systems*, 160(17):2466–2485, 2009.

[4] B. Heshmaty and A. Kandel. Fuzzy linear regression and its applications to forecasting in uncertain environment. *Fuzzy Sets and Systems*, 15(2):159–191, 1985.

[5] M. Hladík. Tolerance analysis in linear programming. Technical report KAM-DIMATIA Series (2008-901), Department of Applied Mathematics, Charles University, Prague, 2008.

[6] C.-H. Huang and H.-Y. Kao. Interval regression analysis with soft-margin reduced support vector machine. *Lecture Notes in Computer Science*, 5579 LNAI:826–835, 2009.

[7] C. Hwang, D. H. Hong, and K. Ha Seok. Support vector interval regression machine for crisp input and output data. *Fuzzy Sets and Systems*, 157(8):1114–1125, 2006.

[8] H. Ishibuchi and H. Tanaka. Several formulations of interval regression analysis. In *Proceedings of Sino-Japan Joint Meeting on Fuzzy Sets and Systems*, pages (B2–2)1–4, Beijing, China, 1990.

[9] H. Ishibuchi, H. Tanaka, and H. Okada. An architecture of neural networks with interval weights and its application to fuzzy regression analysis. *Fuzzy Sets and Systems*, 57(1):27–39, 1993.

[10] J.-T. Jeng, C.-C. Chuang, and S.-F. Su. Support vector interval regression networks for interval regression analysis. *Fuzzy Sets and Systems*, 138(2):283–300, 2003.

[11] G. Jun-peng and L. Wen-hua. Regression analysis of interval data based on error theory. In *Proceedings of 2008 IEEE International Conference on Networking, Sensing and Control, ICNSC*, pages 552–555, Sanya, China, 2008.

[12] M. Kaneyoshi, H. Tanaka, M. Kamei, and H. Furuta. New system identification technique using fuzzy regression analysis. In *Proceedings of the First International Symposium on Uncertainty Modeling and Analysis*, pages 528–533, Baltimore, MD, USA, 1990.

[13] H. Lee and H. Tanaka. Upper and lower approximation models in interval regression using regression quantile techniques. *Eur. J. Oper. Res.*, 116(3):653–666, 1999.

[14] P. Liu. Study on a speech learning approach based on interval support vector regression. In *Proceedings of 4th International Conference on Computer Science & Education*, pages 1009–1012, Nanning, China, 2009.

[15] K. Sugihara, H. Ishii, and H. Tanaka. Interval priorities in AHP by interval regression analysis. *Eur. J. Oper. Res.*, 158(3):745–754, 2004.

[16] H. Tanaka. Fuzzy data analysis by possibilistic linear models. *Fuzzy Sets and Systems*, 24(3):363–375, 1987.

[17] H. Tanaka and H. Lee. Interval regression analysis by quadratic programming approach. *IEEE Transactions on Fuzzy Systems*, 6(4):473–481, 1998.

[18] H. Tanaka, S. Uejima, and K. Asai. Linear regression analysis with fuzzy model. *IEEE Trans. Syst., Man Cybern.*, SMC-12(6):903–907, 1982.

[19] H. Tanaka and J. Watada. Possibilistic linear systems and their application to the linear regression model. *Fuzzy Sets and Systems*, 27(3):275–289, 1988.

[20] J. A. Víšek. The least trimmed squares. Part I: Consistency. *Kybernetika*, 42(1):1–36, 2006.

[21] J. E. Ward and R. E. Wendell. Approaches to sensitivity analysis in linear programming. *Ann. Oper. Res.*, 27:3–38, 1990.

[22] R. E. Wendell. Linear programming. III: The tolerance approach. In Gal, Tomas et al., editor, *Advances in sensitivity analysis and parametric programming*, chapter 5, pages 1–21. Kluwer Academic Publishers, Dordrecht, 1997.