KYBERNETIKA — VOLUME 42 (2006), NUMBER 4, PAGES 423 – 440

INTERVAL LINEAR REGRESSION METHODS BASED ON MINKOWSKI DIFFERENCE – A BRIDGE BETWEEN TRADITIONAL AND INTERVAL LINEAR REGRESSION MODELS

Masahiro Inuiguchi and Tetsuzo Tanino

In this paper, we extend the traditional linear regression methods to the (numerical input)-(interval output) data case assuming both the observation/measurement error and the indeterminacy of the input-output relationship. We propose three different models based on three different assumptions of interval output data. In each model, the errors are defined as intervals by solving the interval equation representing the relationship among the interval output, the interval function and the interval error. We formalize the estimation problem of parameters of the interval function so as to minimize the sum of square/absolute interval errors. Introducing suitable interpretation of minimization of an interval function, each estimation problem is well-formulated as a quadratic or linear programming problem. It is shown that the proposed methods have close relation to both traditional and interval linear regression methods which are formulated in different manners.

Keywords: interval linear regression analysis, least squares method, minimum absolute deviations method, Minkowski difference

AMS Subject Classification: 65G40, 62J05, 26E25, 37E05

1. INTRODUCTION

In order to know a relationship between the dependent and independent variables based on the observed data, function fitting techniques have been developed. The traditional one is the regression analysis. In the regression analysis, a certain functional dependency is assumed and the parameters are estimated so as to minimize a total deviation between observed values and estimated values. The least squares method is the most well-known and popular method for such a parameter estimation. The minimum absolute deviations method [9] is also proposed as a parameter estimation method, which has some robust property against a few outlying data. In the traditional regression method, deviations between observed and estimated data are assumed to be observation/measurement errors.

On the other hand, fuzzy/possibilistic regression method proposed by Tanaka et al. [12, 13], it is assumed that deviations between observed and estimated data are caused by the fuzziness/possibility of the relationship between the dependent (output) and independent (input) variables. Namely, the functional relationship itself is assumed to include some indeterminacy so that observed output data may fluctuate. From this point of view, fuzzy/interval parameters of the assumed function dependency are estimated so as to minimize the sum of widths of estimated fuzzy/interval dependent (output) values under a condition that all estimated fuzzy/interval output values include the corresponding observed data.

This approach has been extended to a case when observed output data are not real numbers but fuzzy/interval values. In this case, three estimation methods were proposed: Possibility (Pos) Problem, Necessity (Nes) Problem and Complement-Necessity (C-Nes) Problem [10]. In Pos Problem, the fuzzy/interval parameters are estimated so as to minimize the sum of widths of estimated fuzzy/interval dependent (output) values under a condition that all estimated fuzzy/interval values intersect the corresponding observed fuzzy/interval data. In Nes Problem, which is called 'the possibility regression model' later in [3], the fuzzy/interval parameters are estimated so as to minimize the sum of widths of estimated fuzzy/interval dependent (output) values under a condition that all estimated fuzzy/interval values include the corresponding observed fuzzy/interval data. Whereas Pos and Nes Problems are extensions of the fuzzy/possibilistic regression model above in case of usual data, C-Nes Problem is not. In C-Nes Problem, which is called 'the necessity regression model' later in [3], the fuzzy/interval parameters are estimated so as to maximize the sum of widths of estimated fuzzy/interval dependent (output) values under a condition that all estimated fuzzy/interval values are included in the corresponding observed fuzzy/interval data.

Moreover, based on ideas of goodness-of-fit, the least squares approach to regression by functional relationship with fuzzy/interval valued parameters has been developed under the presence of fuzzy data, i. e., (numerical input)-(fuzzy output) data and (fuzzy input)-(fuzzy output) data by Diamond et al. [2]. To do this, they defined a metric distance between two fuzzy numbers. This approach includes the traditional regression method, i. e., the least squares method as a special case, whereas Tanaka's approach does not.

In this paper, we propose interval linear regression methods by (numerical input)-(interval output) data based on Minkowski difference [1]. In our approaches, we assume that the functional dependency between input and output variables itself includes some indeterminacy as Tanaka's model does and, at the same time, the observed output data include some errors as the traditional model does. Such concepts are modeled by three kinds of interval equations each of which is a straightforward extension of the traditional linear regression model. The multiplicity of the extension is caused by the property of interval calculations. Solving each interval equation, we can define an interval error and conceptually formulate an interval linear regression problem as a problem of minimizing the sum of square interval errors or a problem of minimizing the sum of absolute interval errors. However, since objective functions of the formulated problems are interval functions, the problem is still ill-posed. To clarify the problems, the lexicographical minimization of the lower and upper bounds of an interval objective function is introduced in two of three models and the minimization of the upper bound of an interval objective function is introduced in the other model.

It is shown that all linear regression problems are reduced to quadratic or linear programming problems. The close connection with Tanaka's Nes and C-Nes models are revealed in two of the proposed models. The similarity of the third model to Diamond's approach is discussed. Moreover, from the model derivation, the proposed interval linear regression methods are extensions of the traditional linear regression methods, i. e., the least squares method and the minimum absolute deviations method. Indeed, two of the proposed methods reduced to the traditional linear regression method when given data are usual (numerical input)-(numerical output) data. We give numerical examples in order to show the similarity and the difference between the previous and the proposed interval regression models.

2. LINEAR REGRESSION METHODS

2.1. Traditional linear regression method

In the traditional linear regression analysis, the output (dependent) variable y is assumed to be represented as $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x}$ with input (independent) variables x_j, $j = 1, 2, \ldots, m$ and a suitable coefficient vector $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_m)^{\mathrm{T}}$, where $\boldsymbol{x} = (1, x_1, x_2, \ldots, x_m)^{\mathrm{T}}$. Given a set of observed input-output data, $(\boldsymbol{x}_i, y_i)$, $i \in N = \{1, 2, \ldots, n\}$, the problem is to estimate the suitable $\boldsymbol{\alpha}$. In the regression analysis, each observed value y_i is assumed to include an error e_i so that we have the following model:

$$y_i = \boldsymbol{\alpha}\boldsymbol{x}_i + e_i. \tag{1}$$

The smaller each absolute value of error $|e_i|$ is, the more suitable $\boldsymbol{\alpha}$ is. Thus, $\boldsymbol{\alpha}$ has been estimated so as to minimize the sum of square errors,

$$\text{minimize } \sum_{i=1}^{n} {e_i}^2, \tag{2}$$

or to minimize the sum of absolute errors,

$$\text{minimize } \sum_{i=1}^{n} |e_i|. \tag{3}$$

(2) is more frequently used and called the least squares method, on the other hand, (3) is called the minimum absolute deviations method. Moreover, whereas (2) provides a maximum likelihood estimator under the normal distribution assumption, (3) has less effects of the outlying data and is regarded as a kind of robust regression method.

2.2. Previous interval linear regression methods

Tanaka et al. [12] assumed that the data fluctuations are not caused by errors but the fuzziness/possibility of the input-output relationship itself. From this point of view, he proposed a fuzzy linear regression which is also called possibilistic linear regression later in [13]. In this regression analysis, we assume that coefficient vector $\boldsymbol{\alpha}$ of the relationship $y = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x}$ is a possibilistic variable vector restricted by an $(m+1)$-dimensional fuzzy set $\boldsymbol{A} = (A_0, A_1, \ldots, A_m)^{\mathrm{T}}$, where each A_j is assumed to be a symmetric L-L fuzzy number with center and spread parameters. Thus, given an input values $\boldsymbol{x}$, we obtain a possible range of output value y as a symmetric L-L fuzzy number $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}$, where $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}$ is obtained by the extension principle [4].

Under this assumption, the regression problem is reduced to estimate $\boldsymbol{A}$, i. e., center and spread parameters. Since the data fluctuations are assumed to be caused by fuzziness/possibility of the input-output relationship, every output data y_i should be located in the possible range $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$ to a certain extent. Under such constraints, the center and spread parameters are estimated so as to minimize the sum of spreads of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$'s. As is shown in Tanaka et al. [10] and Diamond and Tanaka [3], this approach is also made in the case of interval linear regression, since interval is a special case of symmetric L-fuzzy number. In this paper, we discuss interval linear regression analyses. Thus, we concentrate the further explanation of Tanaka's approach on interval case.

Let a_j^{C} and a_j^{W} be the center and width of an interval $A_j = [a_j^{\mathrm{L}}, a_j^{\mathrm{R}}]$, i. e., $a_j^{\mathrm{C}} = \frac{1}{2}(a_j^{\mathrm{L}} + a_j^{\mathrm{R}})$ and $a_j^{\mathrm{W}} = a_j^{\mathrm{R}} - a_j^{\mathrm{L}}$. By a_j^{C} and a_j^{W}, an interval A_j is denoted as $\langle a_j^{\mathrm{C}}, a_j^{\mathrm{W}} \rangle$. By the interval computation [8] or equivalently by the extension principle, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}$ is obtained as

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x} = \left\langle \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x},\ \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}| \right\rangle, \tag{4}$$

where $\boldsymbol{a}^{\mathrm{C}} = (a_0^{\mathrm{C}}, a_1^{\mathrm{C}}, \ldots, a_m^{\mathrm{C}})^{\mathrm{T}}$, $\boldsymbol{a}^{\mathrm{W}} = (a_0^{\mathrm{W}}, a_1^{\mathrm{W}}, \ldots, a_m^{\mathrm{W}})^{\mathrm{T}}$ and $|\boldsymbol{x}| = (1, |x_1|, |x_2|, \ldots \ldots, |x_m|)^{\mathrm{T}}$. Hence, the estimation problem of $\boldsymbol{A}$ is formulated as the following linear programming problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i|, \\ \text{subject to} \quad & \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x}_i - \tfrac{1}{2}\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i| \le y_i,\ i \in N, \\ & \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x}_i + \tfrac{1}{2}\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i| \ge y_i,\ i \in N, \\ & \boldsymbol{a}^{\mathrm{W}} \ge \mathbf{0}. \end{aligned} \tag{5}$$

The objective function shows the sum of widths of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$'s and the constraints show $y_i \in \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$, $i \in N$. Since a linear programming problem has its optimal solution at an extreme point of the feasible set, it is easy for variables a_j^{C} and a_j^{W} in (5) to be zero at the optimum. In order to have more non-zero a_j^{C}'s and a_j^{W}'s, Tanaka and Lee [11] have proposed to use the following objective function:

$$\text{minimize} \quad \sum_{i=1}^{n} (\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i|)^2. \tag{6}$$

In this case, the problem becomes a quadratic programming problem. The other modifications are done also in Tanaka and Lee [11].

Tanaka et al. [13] has also considered a case where the observed output data are given as fuzzy numbers or intervals but input data are still given as real numbers. Such cases are simply called (numerical input)-(fuzzy output) case and (numerical input)-(interval output) case. Such vague output data can be obtained when output data are observed several times at the same input values and take the range or when output value is concerned with human factors such as utility, subjective worth and some other things obtained through subjective and psychological evaluation.

Since approaches are the same between both cases, we concentrate again in (numerical input)-(interval output) case. Let Y_i be the observed output data with respect to input data $\boldsymbol{x}_i$. In this case, two extensions of the constraints of (6), i. e., $y_i \in \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$, are conceivable. One is $Y_i \cap \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i \neq \emptyset$ and the other is $Y_i \subseteq \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$. As can be seen easily, those conditions are reduced to $y_i \in \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$ when Y_i degenerates to a singleton $\{y_i\}$. Replacing $y_i \in \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$ with $Y_i \cap \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i \neq \emptyset$ and with $Y_i \subseteq \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$, we have two different problems. The former problem is called Pos Problem and the latter is called Nes Problem. Nes Problem is more meaningful than Pos Problem when we estimate the possible range of output values. Let $Y_i = [y_i^{\mathrm{L}}, y_i^{\mathrm{R}}]$, $i \in N$, then Nes Problem is formulated as the following linear programming problem:

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i|, \\
\text{subject to} \quad & \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x}_i - \tfrac{1}{2}\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i| \leq y_i^{\mathrm{L}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x}_i + \tfrac{1}{2}\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i| \geq y_i^{\mathrm{R}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{W}} \geq \boldsymbol{0}.
\end{aligned} \tag{7}$$

We can use quadratic objective function (6) instead of the linear objective function in (7) so that the reduced problem becomes a quadratic programming problem. We call the Nes Problem with the linear objective function L-Nes Problem which stands for the linear Nes Problem, on the other hand, the one with quadratic objective function Q-Nes Problem which stands for the quadratic Nes Problem.

One of the interesting facts in (numerical data)-(interval data) case is that we have the converse approach. In Nes Problem, we minimize the sum of widths of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$'s under constraints $Y_i \subseteq \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$, $i \in N$. As the converse problem, we may maximize the sum of the widths of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$'s under constraints $Y_i \supseteq \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$, $i \in N$. In spite that the converse problem does not always have a solution but it has a significant meaning that we may estimate the narrowest ranges of possibility of the output value as $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}$. The converse problem is called C-Nes Problem. Linear C-Nes Problem (L-C-Nes Problem) formulated as the following linear programming problem:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{n} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i|, \\
\text{subject to} \quad & \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x}_i - \tfrac{1}{2}\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i| \geq y_i^{\mathrm{L}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x}_i + \tfrac{1}{2}\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i| \leq y_i^{\mathrm{R}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{W}} \geq \boldsymbol{0}.
\end{aligned} \tag{8}$$

Quadratic C-Nes Problem (Q-C-Nes Problem) can be considered by replacing the linear objective function with a quadratic objective function of (6). However, the problem becomes a non-convex programming problem (a convex maximization problem) and is relatively difficult to solve.

On the other hand, Diamond [2] proposed a least squares approach to the fuzzy linear regression analysis under a situation (numerical input)-(fuzzy output) data or (fuzzy input)-(fuzzy output) data are given. In the rest of this section, restricting ourselves into a special case, i. e., the (numerical input)-(interval output) data case, the method is described briefly.

Given two intervals, $Z = [z^{\rm L}, z^{\rm R}]$ and $W = [w^{\rm L}, w^{\rm R}]$, L_2-metric D_2 is defined by

$$D_2(Z, W)^2 = (z^{\rm L} - w^{\rm L})^2 + (z^{\rm R} - w^{\rm R})^2. \tag{9}$$

Based on the metric D_2, the regression problem of fitting an interval function $\boldsymbol{A}^{\rm T}\boldsymbol{x}$ to given (numerical input)-(interval output) data $(\boldsymbol{x}_i, Y_i)$ is formulated as

$$\text{minimize} \sum_{i=1}^{n} D_2(\boldsymbol{A}^{\rm T}\boldsymbol{x}_i, Y_i)^2. \tag{10}$$

Using parameters $\boldsymbol{a}^{\rm C}$ and $\boldsymbol{a}^{\rm W}$, we may write this problem as

$$\underset{\boldsymbol{a}^{\rm C}, \boldsymbol{a}^{\rm W}}{\text{minimize}} \quad \sum_{i=1}^{n} \left({\boldsymbol{a}^{\rm C}}^{\rm T}\boldsymbol{x}_i - \tfrac{1}{2}{\boldsymbol{a}^{\rm W}}^{\rm T}|\boldsymbol{x}_i| - y_i^{\rm L} \right)^2 + \left({\boldsymbol{a}^{\rm C}}^{\rm T}\boldsymbol{x}_i + \tfrac{1}{2}{\boldsymbol{a}^{\rm W}}^{\rm T}|\boldsymbol{x}_i| - y_i^{\rm R} \right)^2. \tag{11}$$

This problem is called a least squares problem or LS Problem for short.

For intervals $Z = [z^{\rm L}, z^{\rm R}]$ and $W = [w^{\rm L}, w^{\rm R}]$, we may define L_1-metric D_1 by

$$D_1(Z, W) = |z^{\rm L} - w^{\rm L}| + |z^{\rm R} - w^{\rm R}|. \tag{12}$$

As a minimum absolute deviations counterpart, the following regression problem is conceivable:

$$\underset{\boldsymbol{a}^{\rm C}, \boldsymbol{a}^{\rm W}}{\text{minimize}} \quad \sum_{i=1}^{n} \left| {\boldsymbol{a}^{\rm C}}^{\rm T}\boldsymbol{x}_i - \tfrac{1}{2}{\boldsymbol{a}^{\rm W}}^{\rm T}|\boldsymbol{x}_i| - y_i^{\rm L} \right| + \left| {\boldsymbol{a}^{\rm C}}^{\rm T}\boldsymbol{x}_i + \tfrac{1}{2}{\boldsymbol{a}^{\rm W}}^{\rm T}|\boldsymbol{x}_i| - y_i^{\rm R} \right|. \tag{13}$$

This problem is called a minimum absolute deviations problem or MAD problem for short.

3. THE PROPOSED METHODS

3.1. Minkowski difference model

Suppose the input-output dependency itself has some indeterminacy so as to be expressed by an interval linear function and the given observed data $(\boldsymbol{x}_i, Y_i)$, $i \in N$ include errors, we have the following model as a direct extension of (1):

$$Y_i = \boldsymbol{A}\boldsymbol{x}_i + E_i. \tag{14}$$

Note that the model (14) includes not only interval coefficients $\boldsymbol{A}$ but also an interval error E_i in the expression of the interval output value Y_i.

Let $E_i = [e_i^{\mathrm{L}}, e_i^{\mathrm{R}}]$ with $e_i^{\mathrm{L}} \leq e_i^{\mathrm{R}}$. Then (14) is reduced to

$$\begin{cases} e_i^{\mathrm{L}} = y_i^{\mathrm{L}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i + \frac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i|, \\ e_i^{\mathrm{R}} = y_i^{\mathrm{R}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i - \frac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i|. \end{cases} \tag{15}$$

We have $e_i^{\mathrm{L}} \leq e_i^{\mathrm{R}}$ if $\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| \leq y_i^{\mathrm{R}} - y_i^{\mathrm{L}}$. Thus the model (14) implicitly assumes the width of estimated interval $\boldsymbol{A}^{\mathrm{T}} \boldsymbol{x}_i$ is not larger than that of the given interval output datum Y_i. Using Minkowski difference $\ominus_{\mathrm{M}}$, we can represent E_i as $E_i = Y_i \ominus_{\mathrm{M}} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{x}_i$, where for intervals Z and W, $Z \ominus_{\mathrm{M}} W$ is defined by $Z \ominus_{\mathrm{M}} W = \bigcap_{r \in W} (Z - r)$ and we have $Z \ominus_{\mathrm{M}} W = [z_i^{\mathrm{L}} - w_i^{\mathrm{L}}, z_i^{\mathrm{R}} - w_i^{\mathrm{R}}]$ when $z_i^{\mathrm{R}} - z_i^{\mathrm{L}} \geq w_i^{\mathrm{R}} - w_i^{\mathrm{L}}$. From this fact, the model (14) is called Minkowski Difference Model or MD Model for short.

In analogy with (2) and (3), we have

$$\text{minimize} \quad \sum_{i=1}^{n} {E_i}^2, \tag{16}$$

$$\text{minimize} \quad \sum_{i=1}^{n} |E_i|, \tag{17}$$

where the square and absolute values of an interval are defined by

$$Z^2 = \begin{cases} [(\max(z^{\mathrm{L}}, -z^{\mathrm{R}}))^2, (\max(-z^{\mathrm{L}}, z^{\mathrm{R}}))^2] & \text{if } z^{\mathrm{L}} \cdot z^{\mathrm{R}} \geq 0, \\ [0, (\max(-z^{\mathrm{L}}, z^{\mathrm{R}}))^2] & \text{if } z^{\mathrm{L}} \cdot z^{\mathrm{R}} < 0, \end{cases} \tag{18}$$

$$|Z| = \begin{cases} [\max(z^{\mathrm{L}}, -z^{\mathrm{R}}), \max(-z^{\mathrm{L}}, z^{\mathrm{R}})] & \text{if } z^{\mathrm{L}} \cdot z^{\mathrm{R}} \geq 0, \\ [0, \max(-z^{\mathrm{L}}, z^{\mathrm{R}})] & \text{if } z^{\mathrm{L}} \cdot z^{\mathrm{R}} < 0. \end{cases} \tag{19}$$

Since (16) and (17) are minimization of an interval function, they are ill-posed problems. As is in interval programming literature [7], some interpretation of problems should be introduced such as minimizing the lower bound, minimizing the upper bound and so on. We interpret problems (16) and (17) as lexicographical minimization problems of the lower and upper bounds of interval functions. Let $|E_i|$ have $|e|_i^{\mathrm{L}}$ and $|e|_i^{\mathrm{R}}$ as the lower and upper bounds, respectively, i. e., $|E_i| = [|e|_i^{\mathrm{L}}, |e|_i^{\mathrm{R}}]$. (16) and (17) are formulated as the following problems:

$$\text{minimize} \quad \sum_{i=1}^{n} {|e|_i^{\mathrm{L}}}^2 + \epsilon \sum_{i=1}^{n} {|e|_i^{\mathrm{R}}}^2, \tag{20}$$

$$\text{minimize} \quad \sum_{i=1}^{n} |e|_i^{\mathrm{L}} + \epsilon \sum_{i=1}^{n} |e|_i^{\mathrm{R}}, \tag{21}$$

where ϵ is sufficiently small positive number (mathematically speaking, non-Archimedean number such that $\epsilon < r$, $\forall\, r > 0$).

Those problems are reduced to the following quadratic and linear programming problems:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n} \left({d_i^{\mathrm{L}}}^2 + {d_i^{\mathrm{R}}}^2 \right) + \epsilon \sum_{i=1}^{n} {d_i}^2, \\
\text{subject to} \quad & \boldsymbol{a}^{\mathrm{C}\,\mathrm{T}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}\,\mathrm{T}} |\boldsymbol{x}_i| + d_i^{\mathrm{L}} \ge y_i^{\mathrm{L}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{C}\,\mathrm{T}} \boldsymbol{x}_i + \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}\,\mathrm{T}} |\boldsymbol{x}_i| - d_i^{\mathrm{R}} \le y_i^{\mathrm{R}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{C}\,\mathrm{T}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}\,\mathrm{T}} |\boldsymbol{x}_i| - d_i \le y_i^{\mathrm{L}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{C}\,\mathrm{T}} \boldsymbol{x}_i + \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}\,\mathrm{T}} |\boldsymbol{x}_i| + d_i \ge y_i^{\mathrm{R}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{W}\,\mathrm{T}} |\boldsymbol{x}_i| \le y_i^{\mathrm{R}} - y_i^{\mathrm{L}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{W}},\ \boldsymbol{d}^{\mathrm{L}},\ \boldsymbol{d}^{\mathrm{R}},\ \boldsymbol{d} \ge \boldsymbol{0},
\end{aligned}
\tag{22}
$$

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n} \left(d_i^{\mathrm{L}} + d_i^{\mathrm{R}} \right) + \epsilon \sum_{i=1}^{n} d_i, \\
\text{subject to} \quad & \text{constraints of (22)},
\end{aligned}
\tag{23}
$$

where $\boldsymbol{d}^{\mathrm{L}} = (d_1^{\mathrm{L}}, d_2^{\mathrm{L}}, \ldots, d_n^{\mathrm{L}})^{\mathrm{T}}$, $\boldsymbol{d}^{\mathrm{R}} = (d_1^{\mathrm{R}}, d_2^{\mathrm{R}}, \ldots, d_n^{\mathrm{R}})^{\mathrm{T}}$, $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)^{\mathrm{T}}$. Each of those problems is called Minkowski difference problem or MD Problem for short. Moreover, (22) is called a quadratic Minkowski difference problem (Q-MD Problem) whereas (23) is called a linear Minkowski difference problem (L-MD Problem). Note that, for $i \in N$, the following equalities are satisfied at optimal solutions of (22) and (23).

$$
\begin{aligned}
& d_i = \max \left(\boldsymbol{a}^{\mathrm{C}\,\mathrm{T}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}\,\mathrm{T}} |\boldsymbol{x}_i| - y_i^{\mathrm{L}}, y_i^{\mathrm{R}} - \boldsymbol{a}^{\mathrm{C}\,\mathrm{T}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}\,\mathrm{T}} |\boldsymbol{x}_i| \right), \\
& d_i^{\mathrm{L}} \cdot d_i^{\mathrm{R}} = 0, \quad |E_i| = [d_i^{\mathrm{L}} + d_i^{\mathrm{R}}, d_i].
\end{aligned}
\tag{24}
$$

Practically (22) and (23) can be solved by quadratic and linear programming techniques with setting ϵ as a sufficiently small positive number such as 0.001. However, there is no guarantee that the selected ϵ is sufficiently small to obtain real optimal solutions of those problems. In case of (23), we can utilize multi-phase linear programming [6] in order to obtain an optimal solution without specification of ϵ. On the other hand, an optimal solution to (24) can be obtained without specification of ϵ by solving two quadratic programming problems sequentially.

Theorem 1. $d_i^{\mathrm{L}} + d_i^{\mathrm{R}} = 0$ if and only if $\boldsymbol{A}^{\mathrm{T}} \boldsymbol{x}_i \subseteq Y_i$.

Proof. It is obvious from the constraints of (22). □

Theorem 1 shows that constraints of C-Nes Problem (8) are satisfied with $(\hat{\boldsymbol{a}}^{\mathrm{C}}, \hat{\boldsymbol{a}}^{\mathrm{W}})$ when the first term of the objective function of (22) (or (23)) is zero at the optimal solution. From this fact, we can recognize the similarity between C-Nes Problem and MD Problem. Whereas C-Nes Problem yields a solution only when $\boldsymbol{A}^{\mathrm{T}} \boldsymbol{x}_i \subseteq Y_i$,

$i \in N$ are satisfied, MD Problem yields a solution even when $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i \subseteq Y_i$, $i \in N$ cannot be satisfied.

The similarity between L-C-Nes and L-MD Problems is stronger than that between Q-C-Nes and Q-MD Problems as shown in Theorem 2.

Theorem 2. Assume there is a feasible solution to (8), i. e., $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i \subseteq Y_i$, $i \in N$. Let $q_i^{\mathrm{L}} = \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x}_i - \frac{1}{2}\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i| - y_i^{\mathrm{L}} \geq 0$ and $q_i^{\mathrm{R}} = y_i^{\mathrm{R}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x}_i - \frac{1}{2}\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i| \geq 0$. Then, L-C-Nes Problem is equivalent to a problem minimizing $\sum_{i=1}^{n}(q_i^{\mathrm{L}} + q_i^{\mathrm{R}})$ subject to constraints of (8). Similarly, L- and Q-MD Problems are reduced to problems minimizing $\sum_{i=1}^{n} \max(q_i^{\mathrm{L}}, q_i^{\mathrm{R}})$ and $\sum_{i=1}^{n} \max(q_i^{\mathrm{L}}, q_i^{\mathrm{R}})^2$ subject to constraints of (8), respectively.

P r o o f . The first assertion of the theorem is easily obtained from $\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i| + q_i^{\mathrm{L}} + q_i^{\mathrm{R}} = y_i^{\mathrm{R}} - y_i^{\mathrm{L}}$ under the assumption. The second assertion comes from the fact that $E_i = [-q_i^{\mathrm{L}}, q_i^{\mathrm{R}}]$, $i \in N$ implies $|E_i| = [0, \max(q_i^{\mathrm{L}}, q_i^{\mathrm{R}})]$ under the assumption. □

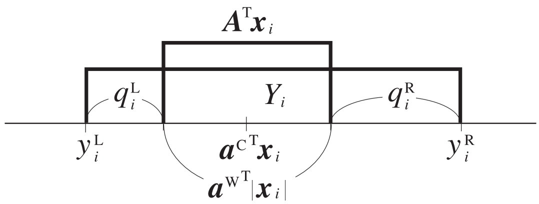


Fig. 1. q_i^{L} and q_i^{R} in MD Problem.

Variables q_i^{L} and q_i^{R} are depicted in Figure 1. Under the assumption of Theorem 2, the sum $q_i^{\mathrm{L}} + q_i^{\mathrm{R}}$ in L-C-Nes Problem is replaced with the maximum value $\max(q_i^{\mathrm{L}}, q_i^{\mathrm{R}})$ in L-MD Problem. This shows that centers of the estimated intervals obtained from L-MD Problem lie in more central places of the given interval output data than those obtained from L-C-Nes Problem.

From the equivalence between maximizing $\sum_{i=1}^{n} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i|$ and minimizing the expression $\sum_{i=1}^{n}(q_i^{\mathrm{L}} + q_i^{\mathrm{R}})$, one may consider a problem minimizing $\sum_{i=1}^{n}({q_i^{\mathrm{L}}}^2 + {q_i^{\mathrm{R}}}^2)$, which is equivalent to (10), under constraints of (8) as a quadratic counterpart of L-C-Nes Problem other than Q-C-Nes Problem. This quadratic counterpart is called $\mathrm{Q_2}$-C-Nes Problem. We have a similar relationship between $\mathrm{Q_2}$-C-Nes and Q-MD Problems to that between L-C-Nes and L-MD Problems. The difference of objective functions of Q- and $\mathrm{Q_2}$-C-Nes Problems can be seen by

$$\begin{aligned}(\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i|)^2 &= (y_i^{\mathrm{R}} - y_i^{\mathrm{L}} - q_i^{\mathrm{L}} - q_i^{\mathrm{R}})^2 \\ &= (y_i^{\mathrm{R}} - y_i^{\mathrm{L}})^2 + {q_i^{\mathrm{L}}}^2 + {q_i^{\mathrm{R}}}^2 - 2(y_i^{\mathrm{R}} - y_i^{\mathrm{L}})(q_i^{\mathrm{L}} + q_i^{\mathrm{R}}) + 2q_i^{\mathrm{L}} \cdot q_i^{\mathrm{R}}. \quad (25)\end{aligned}$$

Since $(y_i^{\mathrm{R}} - y_i^{\mathrm{L}})$ is constant, Q-C-Nes Problem is a linear combination between L-C-Nes Problem and another quadratic counterpart called $\mathrm{Q_3}$-C-Nes Problem which minimizes $\sum_{i=1}^{n}(q_i^{\mathrm{L}}+q_i^{\mathrm{R}})^2$ under constraints of (8). It should be noted that Q_2- and Q_3-C-Nes Problems are convex quadratic programming problems while Q-C-Nes Problem is a non-convex one.

Theorem 3. In MD Problems, $d_i = 0$ if and only if $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i = Y_i$.

Proof. It is obvious from $|e|_i^{\mathrm{R}} \le d_i$ for any feasible solution to MD Problems. □

From Theorem 3, we can obtain an interval linear function totally fitted to all given data when the optimal objective function value of an MD Problem is zero. Thus, in MD Problem, $\boldsymbol{A}$ is estimated so as first to satisfy $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i \subseteq Y_i$, $i \in N$ and then to make $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$'s as closer to Y_i's as possible under constraints $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i \subseteq Y_i$, $i \in N$.

Theorem 4. When $y_i^{\mathrm{L}} = y_i^{\mathrm{R}}$, $i \in N$, i.e., interval output data Y_i's are degenerated to usual data y_i, Q- and L-MD Problems are reduced to the traditional least squares and minimum absolute deviations problems, respectively.

Proof. It is trivial. □

Theorem 4 reconfirm that MD Problems are extensions of the traditional regression problems.

3.2. Converse Minkowski difference model

In MD Model, we implicitly assume that the width of Y_i is wider than that of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$ because Y_i is assumed to be a sum of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$ and an interval error E_i. This is reasonable when output values are recorded as wider ranges during observation. We can imagine also the opposite case when output values are recorded as narrower ranges. Thus, the opposite assumption is also conceivable, i.e., we may assume that the width of Y_i is narrower than that of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$. To treat this case, we extend a model which we get by transposing e_i to the left-hand side in (1), i.e., $y_i - e_i = \boldsymbol{\alpha}\boldsymbol{x}_i$. To do this, we have

$$Y_i - E_i = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i. \tag{26}$$

In this case, the interval error $E_i = [e_i^{\mathrm{L}}, e_i^{\mathrm{R}}]$ is defined as follows when $\boldsymbol{a}^{\mathrm{W}\,\mathrm{T}}|\boldsymbol{x}_i| \ge y_i^{\mathrm{R}} - y_i^{\mathrm{L}}$ is satisfied:

$$\begin{cases} e_i^{\mathrm{L}} = y_i^{\mathrm{R}} - \boldsymbol{a}^{\mathrm{C}\,\mathrm{T}}\boldsymbol{x}_i - \frac{1}{2}\boldsymbol{a}^{\mathrm{W}\,\mathrm{T}}|\boldsymbol{x}_i|, \\ e_i^{\mathrm{R}} = y_i^{\mathrm{L}} - \boldsymbol{a}^{\mathrm{C}\,\mathrm{T}}\boldsymbol{x}_i + \frac{1}{2}\boldsymbol{a}^{\mathrm{W}\,\mathrm{T}}|\boldsymbol{x}_i|. \end{cases} \tag{27}$$

Define a sum $Z \oplus_{\mathrm{M}} W$ by $Z \oplus_{\mathrm{M}} W = \bigcap_{r\in W}(Z + r)$. Then we have $Z \oplus_{\mathrm{M}} W = [z_i^{\mathrm{L}} + w_i^{\mathrm{R}}, z_i^{\mathrm{R}} + w_i^{\mathrm{L}}]$ when $z_i^{\mathrm{R}} - z_i^{\mathrm{L}} \ge w_i^{\mathrm{R}} - w_i^{\mathrm{L}}$. Using $\oplus_{\mathrm{M}}$, (26) is equivalent to

$$Y_i = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i \oplus_{\mathrm{M}} E_i. \tag{28}$$

We call this model Converse Minkowski Difference Model or C-MD Model for short.

To C-MD Model, (16) and (17) are considered. In the same manner as MD Model, those ill-posed problems are treated as (20) and (21). It is different from the case of MD Model that, for any given (numerical input)-(interval output) data, $|e|_i^{\mathrm{L}} = 0$, $i \in N$ are satisfied with sufficiently large a_j^{W}'s in C-MD Model. Thus, (20) and (21) are reduced to the following quadratic and linear programming problems, respectively:

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n} {d_i}^2, \\
\text{subject to} \quad & \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| \le y_i^{\mathrm{L}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i + \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| \ge y_i^{\mathrm{R}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| + d_i \ge y_i^{\mathrm{L}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i + \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| - d_i \le y_i^{\mathrm{R}},\ i \in N, \\
& \boldsymbol{a}^{\mathrm{W}},\ \boldsymbol{d} \ge \boldsymbol{0},
\end{aligned} \tag{29}$$

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n} d_i, \\
\text{subject to} \quad & \text{constraints of (29).}
\end{aligned} \tag{30}$$

Each of those problems is called converse Minkowski difference problem or C-MD Problem for short. Moreover, (29) is called a quadratic converse Minkowski difference problem (Q-C-MD Problem) whereas (30) is called a linear Minkowski difference problem (L-C-MD Problem). Note that, for $i \in N$, the following equalities are satisfied at optimal solutions of (29) and (30).

$$\begin{aligned}
& d_i = \max\left(y_i^{\mathrm{L}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i + \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i|, \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| - y_i^{\mathrm{R}} \right), \\
& |E_i| = [0, d_i].
\end{aligned} \tag{31}$$

Since the first two constraints of (29) are the same as those of Nes Problem (7), C-MD Problem is similar to Nes Problem. Moreover, when $y_i^{\mathrm{L}} = y_i^{\mathrm{R}}$, i. e., interval output data Y_i's are degenerated to usual data y_i, C-MD problem is similar to Problem (5).

A stronger similarity between L-Nes and L-C-MD Problems is found in Theorem 5.

Theorem 5. Let $g_i^{\mathrm{L}} = y_i^{\mathrm{L}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i + \frac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| \ge 0$ and $g_i^{\mathrm{R}} = \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i + \frac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| - y_i^{\mathrm{R}} \ge 0$. Then, L-Nes Problem is equivalent to a problem minimizing $\sum_{i=1}^{n} (g_i^{\mathrm{L}} + g_i^{\mathrm{R}})$ subject to constraints of (7). Similarly, L- and Q-C-MD Problems are reduced to problems minimizing $\sum_{i=1}^{n} \max(g_i^{\mathrm{L}}, g_i^{\mathrm{R}})$ and $\sum_{i=1}^{n} \max(g_i^{\mathrm{L}}, g_i^{\mathrm{R}})^2$ subject to constraints of (7), respectively.

P r o o f. It can be proved in the same manner as Theorem 2. □

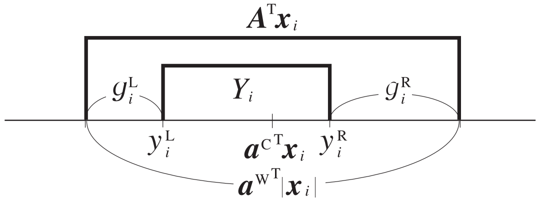


Fig. 2. $g_i^{\rm L}$ and $g_i^{\rm R}$ in C-MD Problem.

Variables $g_i^{\rm L}$ and $g_i^{\rm R}$ are depicted in Figure 2. From Theorem 5, the sum $g_i^{\rm L}+g_i^{\rm R}$ in L-Nes Problem is replaced with the maximum value $\max(g_i^{\rm L}, g_i^{\rm R})$ in L-C-MD Problem. This shows that centers of the estimated intervals obtained from L-C-MD Problem lie in more central places of the given interval output data than those obtained from L-Nes Problem.

From the equivalence between minimizing $\sum_{i=1}^n \boldsymbol{a}^{\rm W\,T}|\boldsymbol{x}_i|$ and minimizing the expresssion $\sum_{i=1}^n (g_i^{\rm L}+g_i^{\rm R})$, one may consider a problem minimizing $\sum_{i=1}^n ({g_i^{\rm L}}^2+{g_i^{\rm R}}^2)$, which is equivalent to (10), under constraints of (7) as a quadratic counterpart of L-Nes Problem other than Q-Nes Problem. This quadratic counterpart is called $\rm Q_2$-Nes Problem. We have a similar relationship between $\rm Q_2$-Nes and Q-C-MD Problems to that between L-Nes and L-C-MD Problems. The difference of objective functions of Q- and $\rm Q_2$-Nes Problems can be seen by

$$\begin{aligned}(\boldsymbol{a}^{\rm W\,T}|\boldsymbol{x}_i|)^2 &= (y_i^{\rm R}-y_i^{\rm L}+g_i^{\rm L}+g_i^{\rm R})^2\\ &= (y_i^{\rm R}-y_i^{\rm L})^2+{g_i^{\rm L}}^2+{g_i^{\rm R}}^2+2(y_i^{\rm R}-y_i^{\rm L})(g_i^{\rm L}+g_i^{\rm R})+2g_i^{\rm L}\cdot g_i^{\rm R}. \quad (32)\end{aligned}$$

Since $(y_i^{\rm R}-y_i^{\rm L})$ is constant, Q-Nes Problem is an intermediate problem between L-Nes Problem and another quadratic counterpart called $\rm Q_3$-Nes Problem which minimizes $\sum_{i=1}^n (g_i^{\rm L}+g_i^{\rm R})^2$ under constraints of (7). Q-, $\rm Q_2$- and $\rm Q_3$-Nes Problems are all convex quadratic programming problems.

Theorem 6. In C-MD Problems, $d_i = 0$ if and only if $\boldsymbol{A}^{\rm T}\boldsymbol{x}_i = Y_i$.

P r o o f. It is obvious from $|e|_i^{\rm R} \le d_i$ for any feasible solution to C-MD Problems. □

From Theorem 6, we can obtain an interval linear function totally fitted to all given data when the optimal objective function value of an C-MD Problem is zero. Thus, in C-MD Problem, $\boldsymbol{A}$ is estimated so as to make $\boldsymbol{A}^{\rm T}\boldsymbol{x}_i$'s as closer to Y_i's as possible under constraints $\boldsymbol{A}^{\rm T}\boldsymbol{x}_i \supseteq Y_i$, $i \in N$.

3.3. Symmetric Minkowski difference model

We assumed that the width of Y_i is narrower than that of $\boldsymbol{A}^{\rm T}\boldsymbol{x}_i$ in C-MD Model while we assumed that the width of Y_i is wider in MD Model. When output values are recorded as appropriate ranges during observation, both models may be applicable from Theorems 3 and 6. Moreover, we do not need to assume those assumptions in this case. In order to treat such a case, the following model is conceivable:

$$Y_i - E_i^{\mathrm{LS}} = \boldsymbol{A}^{\mathrm{T}} \boldsymbol{x}_i + E_i^{\mathrm{RS}}, \tag{33}$$

where superscripts 'LS' and 'RS' stand for 'left-hand side' and 'right-hand side', respectively. In this model, we impose complementary conditions on interval errors E_i^{LS} and E_i^{RS} such that $E_i^{\mathrm{LS}} = [0,0]$ when $y_i^{\mathrm{R}} - y_i^{\mathrm{L}} \geq \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i|$ and $E_i^{\mathrm{RS}} = [0,0]$ otherwise. Let $E_i = E_i^{\mathrm{LS}} + E_i^{\mathrm{RS}}$. Then $E_i = [e_i^{\mathrm{L}}, e_i^{\mathrm{R}}]$ is obtained as

$$e_i^{\mathrm{L}} = \min\left(y_i^{\mathrm{L}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i + \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i|, y_i^{\mathrm{R}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| \right), \tag{34}$$

$$e_i^{\mathrm{R}} = \max\left(y_i^{\mathrm{L}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i + \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i|, y_i^{\mathrm{R}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| \right). \tag{35}$$

The model (33) is called Symmetric Minkowski Difference Model or S-MD Model for short.

To S-MD Model, (16) and (17) are considered again. In this case, the lower bound of $|E_i|$, i. e., $|e|_i^{\mathrm{L}}$ is not a convex function of parameters $\boldsymbol{a}^{\mathrm{C}}$ and $\boldsymbol{a}^{\mathrm{W}}$. Thus, it is not easy to solve (20) and (21). However, fortunately, the upper bound of $|E_i|$, i. e., $|e|_i^{\mathrm{R}}$ is a convex function. Then, we solve the following problems instead of (20) and (21), respectively:

$$\text{minimize} \quad \sum_{i=1}^{n} |e|_i^{\mathrm{R}^2}, \tag{36}$$

$$\text{minimize} \quad \sum_{i=1}^{n} |e|_i^{\mathrm{R}}, \tag{37}$$

Those problems are reduced to the following quadratic and linear programming problems, respectively:

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} {d_i}^2, \\ \text{subject to} \quad & \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| + d_i \geq y_i^{\mathrm{L}}, \ i \in N, \\ & \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i - \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| - d_i \leq y_i^{\mathrm{L}}, \ i \in N, \\ & \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i + \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| + d_i \geq y_i^{\mathrm{R}}, \ i \in N, \\ & \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}} \boldsymbol{x}_i + \tfrac{1}{2} \boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}} |\boldsymbol{x}_i| - d_i \leq y_i^{\mathrm{R}}, \ i \in N, \\ & \boldsymbol{a}^{\mathrm{W}}, \ \boldsymbol{d} \geq \boldsymbol{0}, \end{aligned} \tag{38}$$

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} d_i, \\ \text{subject to} \quad & \text{constraints of (38)}. \end{aligned} \tag{39}$$

Note that $|e|_i^{\mathrm{R}} = d_i$ at optimal solutions of (38) and (39). Each of those problems is called Symmetric Minkowski difference Problem or S-MD Problem, for short. More

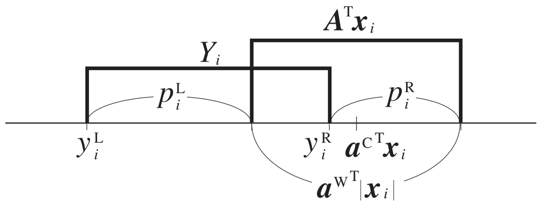


Fig. 3. p_i^{L} and p_i^{R} in S-MD Problem.

specifically, (38) is called a quadratic symmetric Minkowski difference problem (Q-S-MD Problem) and (39) is called a linear symmetric Minkowski difference problem (L-S-MD Problem).

Theorem 7. In S-MD Problems, $d_i = 0$ if and only if $Y_i = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_i$.

Proof. It is evident from $|e|_i^{\mathrm{R}} \leq d_i$ for any feasible solution to S-MD Problems. □

We can find the similarity between LS (resp. MAD) Problem and Q-S-MD (resp. L-S-MD) Problem as in Theorem 8.

Theorem 8. Let $p_i^{\mathrm{L}} = \left|y_i^{\mathrm{L}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x}_i + \frac{1}{2}\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i|\right| \geq 0$ and $p_i^{\mathrm{R}} = \left|y_i^{\mathrm{R}} - \boldsymbol{a}^{\mathrm{C}^{\mathrm{T}}}\boldsymbol{x}_i - \frac{1}{2}\boldsymbol{a}^{\mathrm{W}^{\mathrm{T}}}|\boldsymbol{x}_i|\right| \geq 0$. Then, LS and MAD Problems are problems minimizing $\sum_{i=1}^{n}({p_i^{\mathrm{L}}}^2 + {p_i^{\mathrm{R}}}^2)$ and $\sum_{i=1}^{n}(p_i^{\mathrm{L}} + p_i^{\mathrm{R}})$, respectively. Similarly, Q- and L-S-MD Problems are reduced to problems minimizing $\sum_{i=1}^{n}\max(p_i^{\mathrm{L}}, p_i^{\mathrm{R}})^2$ and $\sum_{i=1}^{n}\max(p_i^{\mathrm{L}}, p_i^{\mathrm{R}})$, respectively.

Proof. The first assertion is obvious from the definitions of problems. The second assertion comes from $d_i = \max(p_i^{\mathrm{L}}, p_i^{\mathrm{R}})$ at the optimal solutions. □

Variables p_i^{L} and p_i^{R} are depicted in Figure 3. From Theorem 8, the square sum ${p_i^{\mathrm{L}}}^2 + {p_i^{\mathrm{R}}}^2$ in LS Problem is replaced with the square maximum $\max(p_i^{\mathrm{L}}, p_i^{\mathrm{R}})^2$ in Q-S-MD Problem. On the other hand, the sum $p_i^{\mathrm{L}} + p_i^{\mathrm{R}}$ in MAD Problem is replaced with the maximum value $\max(p_i^{\mathrm{L}}, p_i^{\mathrm{R}})$ in L-S-MD Problem. These facts show that centers of the estimated intervals obtained from Q- and L-S-MD Problems lie in more central places of the given interval output data than those obtained from LS and MAD Problems, respectively. Moreover, all differences between L-C-Nes and L-MD Problems, between L-Nes and L-C-MD Problems, and between MAD and L-S-MD Problems are the same, i. e., difference between sum and maximum of p_i^{L} and p_i^{R}.

As in MD Problems, we have the following theorem.

Theorem 9. When $y_i^{\mathrm{L}} = y_i^{\mathrm{R}}$, $i \in N$, i. e., interval output data Y_i's are degenerated to usual data y_i, Q- and L-S-MD Problems are reduced to the traditional least squares and minimum absolute deviations problems, respectively.

Proof. It is trivial. □

4. A NUMERICAL EXAMPLE

In order to see the similarity and difference between the previous and the proposed approaches to interval linear regression, we use data of feed speed and surface roughness in [3] (see Table 1). The following interval functional relationship is assumed:

$$Y = A_0 + A_1 x + A_2 x^2. \tag{40}$$

Table 1. Data.

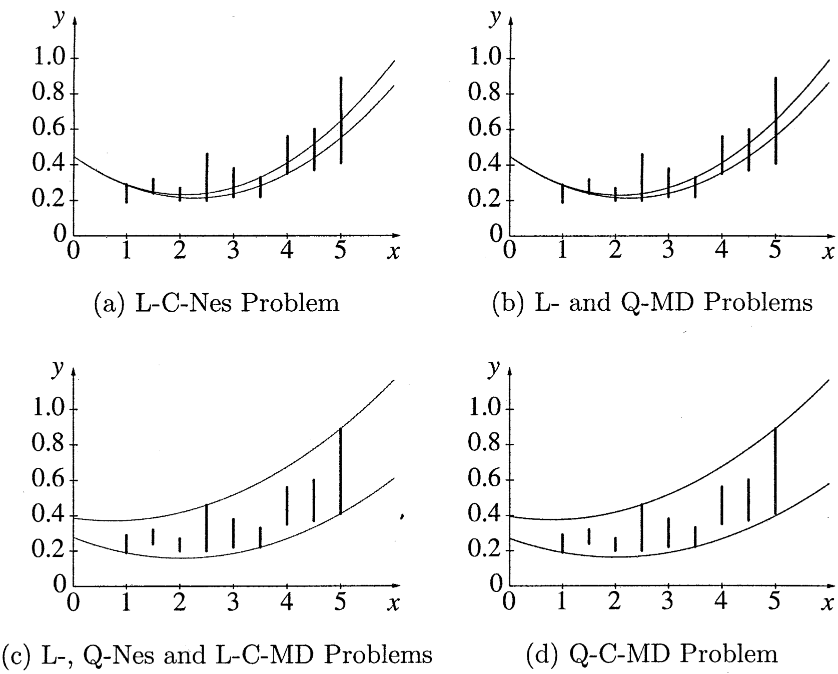
x_i	Y_i
1	$[0.19, 0.29]$
1.5	$[0.24, 0.32]$
2	$[0.2, 0.27]$
2.5	$[0.2, 0.46]$
3	$[0.22, 0.38]$
3.5	$[0.22, 0.33]$
4	$[0.35, 0.56]$
4.5	$[0.37, 0.6]$
5	$[0.41, 0.89]$

Table 2. Results of Interval Linear Regression Methods.

problem	A_0	A_1	A_2
L-C-Nes	$\langle 0.446516, 0\rangle$	$\langle -0.205806, 0\rangle$	$\langle 0.047355, 0.00387\rangle$
L-MD Q-MD	$\langle 0.449000, 0\rangle$	$\langle -0.209000, 0\rangle$	$\langle 0.048222, 0.003556\rangle$
L-Nes Q-Nes L-C-MD	$\langle 0.333333, 0.110000\rangle$	$\langle -0.080000, 0.074000\rangle$	$\langle 0.028667, 0\rangle$
Q-C-MD	$\langle 0.332964, 0.125312\rangle$	$\langle -0.077027, 0.059310\rangle$	$\langle 0.027815, 0.002882\rangle$
MAD	$\langle 0.363000, 0.034000\rangle$	$\langle -0.135000, 0\rangle$	$\langle 0.038000, 0.014000\rangle$
LS	$\langle 0.366000, 0.058110\rangle$	$\langle -0.132000, 0\rangle$	$\langle 0.036667, 0.012260\rangle$
L-S-MD	$\langle 0.325000, 0.058728\rangle$	$\langle -0.118333, 0\rangle$	$\langle 0.036667, 0.009454\rangle$
Q-S-MD	$\langle 0.395246, 0.081168\rangle$	$\langle -0.157264, 0\rangle$	$\langle 0.041643, 0.008464\rangle$

Assuming the interval output data are recorded as wider intervals, we apply C-Nes and MD Models. By solving L-C-Nes, L-MD and Q-MD Problems, we obtain A_i, $i = 1, 2, 3$ as in Table 2. On the other hand, assuming the interval output data are recorded as narrower intervals, we apply Nes and C-MD Models. By solving L-Nes, Q-Nes, L-C-MD and Q-C-MD Problems, we obtain A_i, $i = 1, 2, 3$ as in Table 2. Moreover, assuming the interval output data are recorded as appropriate intervals, we apply metric and S-MD Models. By solving LS, MAD, L-S-MD and Q-S-MD Problems, we obtain A_i, $i = 1, 2, 3$ as in Table 2.

From Table 2, we can recognize the difference among obtained results. However, results between C-Nes and MD Models, between Nes and C-MD Models and between metric and S-MD Models are quite similar. The similarities can also be seen in Figure 4.



(a) L-C-Nes Problem

(b) L- and Q-MD Problems

(c) L-, Q-Nes and L-C-MD Problems

(d) Q-C-MD Problem

Fig. 4. Graphs of the obtained interval functions.

5. CONCLUDING REMARKS

We have proposed three kinds of interval linear regression models based on Minkowski difference. The regression problems were formulated so as to minimize the interval deviations. Then the idea of formulation is same as the traditional linear regression methods. However two of formulated interval linear regression problems are similar to Nes and C-Nes problems in interval linear regression methods proposed by Tanaka et al. [10]. From this fact the proposed methods can be seen as a bridge of the traditional linear regression methods and the previous interval linear regression methods.

Since the proposed methods is based on minimization of deviations, we can introduce the M-estimator developed in robust regression methods [5] to interval linear regression problems. This is a future research topic.

ACKNOWLEDGMENT

The authors express our appreciations to Mr. Haruki Fujita for his help in computations of numerical examples. The part of this research is supported by the Japan Society for the Promotion of Science under the Grant-in-Aid for Scientific Research No. 16510114.

(Received September 1, 2006.)

REFERENCES

[1] J.-P. Aubin and H. Frankowska: Set-Valued Analysis. Birkhäuser, Boston 1990.

[2] P. Diamond: Fuzzy least squares. Inform. Sci. *46* (1988), 141–157.

[3] P. Diamond and H. Tanaka: Fuzzy regression analysis. In: Fuzzy Sets in Decision Analysis, Operations Research and Statistics (R. Słowinski, ed.), Kluwer, Boston 1988, pp. 349–387.

[4] D. Dubois and H. Prade: Fuzzy numbers: An overview. In: Analysis of Fuzzy Information, Vol. I: Mathematics and Logic (J. C. Bezdek, ed.), CRC Press, Boca Raton 1987, pp. 3–39.

[5] P. J. Huber: Robust statistics. Ann. Math. Statist. *43* (1972), 1041–1067.

[6] J. P. Ignizio: Linear Programming in Single- & Multiple-Objective Systems. Prentice-Hall, Englewood Cliffs, NJ 1982.

[7] M. Inuiguchi and Y. Kume: Goal programming problems with interval coefficients and target ontervals. European J. Oper. Res. *52* (1991), 345–360.

[8] R. E. Moore: Methods and Applications of Interval Analysis. SIAM, Philadelphia 1979.

[9] W. F. Shape: Mean-absolute-deviation characteristic lines for securities and portfolios. Management Sci. *18* (1971), 2, B1–B13.

[10] H. Tanaka, I. Hayashi, and K. Nagasaka: Interval regression analysis by possibilistic measures (in Japanese). Japan. J. Behaviormetrics *16* (1988), 1, 1–7.

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

[12] H. Tanaka, S. Uejima, and K. Asai: Linear regression analysis with fuzzy model. IEEE Trans. Systems Man Cybernet. *12* (1982), 903–907.

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

Masahiro Inuiguchi, Department of Systems Innovation, Graduate School of Engineering Science, Osaka University, 1-3, Machikaneyama-cho, Toyonaka, Osaka 560-8531. Japan.

e-mail: inuiguti@sys.es.osaka-u.ac.jp

Tetsuzo Tanino, Department of Electrical, Electronic and Information Engineering, Graduate School of Engineering, Osaka University, 2-1, Yamada-oka, Suita, Osaka 565-0871. Japan.
e-mail: tanino@eei.eng.osaka-u.ac.jp