Dual models for possibilistic regression analysis

Peijun Guo\textsuperscript{a},*, Hideo Tanaka\textsuperscript{b}

\textsuperscript{a}Faculty of Economics, Kagawa University, Takamatsu, Kagawa 760-8523, Japan
\textsuperscript{b}Department of Kansei Information, Hiroshima International University, Gakuendai 555-36, Kurosecho, Higashi-Hiroshima 724-0695, Japan

Available online 2 May 2006

Abstract

Upper and lower regression models (dual possibilistic models) are proposed for data analysis with crisp inputs and interval or fuzzy outputs. Based on the given data, the dual possibilistic models can be derived from upper and lower directions, respectively, where the inclusion relationship between these two models holds. Thus, the inherent uncertainty existing in the given phenomenon can be approximated by the dual models. As a core part of possibilistic regression, firstly possibilistic regression for crisp inputs and interval outputs is considered where the basic dual linear models based on linear programming, dual nonlinear models based on linear programming and dual nonlinear models based on quadratic programming are systematically addressed, and similarities between dual possibilistic regression models and rough sets are analyzed in depth. Then, as a natural extension, dual possibilistic regression models for crisp inputs and fuzzy outputs are addressed.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Possibilistic regression analysis; Linear programming; Quadratic programming; Rough sets

1. Introduction

Regression analysis is a fundamental analytic tool in many disciplines. The method analyzes the relationship between dependent and independent variables based on the given data from the statistical viewpoint, that is, the observation data are random with some measure errors or noise. On the other hand, the relation between the dependent and independent variables can be characterized by a fuzzy functional relationship for the given data that may be fuzzy or crisp. That is to say, an uncertain phenomenon should be modeled by a fuzzy functional relationship. Based on this idea, Tanaka et al. proposed a possibilistic regression model where a fuzzy linear system was used as a regression model (Tanaka, 1987; Tanaka et al., 1989, 1982; Tanaka and Watada, 1988). For crisp input–output data, the models have been developed in several directions. Firstly, different kinds of possibility distributions, such as quadratic and exponential possibility distributions have been used to formulate possibilistic regression models for considering interactive relations of fuzzy coefficients (Tanaka and Ishibuchi, 1991; Tanaka et al., 1995). Secondly, quadratic programming (QP) problems and nonlinear regression models are used for obtaining better fitness and more diverse spread coefficients (Tanaka and Lee, 1998). Thirdly, some technique has been used to overcome the shortcoming of possibilistic regression sensitive to

* Corresponding author. Tel.: +81 86 8321909; fax: +81 86 8321909.
E-mail address: guo@ec.kagawa-u.ac.jp (P. Guo).
outliers (Chen, 2001; Lee and Tanaka, 1999). Moreover, Wang et al. gave some explanation for possibilistic regression models (Wang and Tsaur, 2000). For the fuzzy output case, there are mainly two kinds of methods. One is the least squares approach where the diversity between estimated fuzzy output and given fuzzy output is minimized according to different kinds of distances between two fuzzy numbers (Celmins, 1987; Diamond, 1988; Kao and Chyu, 2003; Tran and Duckstein, 2002; D’Urso, 2003; Wu, 2003). The other method is possibilistic regression (Tanaka and Guo, 1999, 2002a,b; Tanaka et al., 1998), where inclusion relations between given outputs and estimated outputs play a critical role being very similar to the rough set concept (Pawlak, 1984). Using this method, the given fuzzy outputs are approximated by two fuzzy outputs from inside and outside directions. The idea for analyzing the unknown from upper and lower directions has been used for identifying the possibility distributions (Guo and Tanaka, 2003a,b), for obtaining the fuzzy decision variables in upper decision level (Guo et al., 2000), for obtaining interval weights of priorities in the analytic hierarchy process (AHP) (Sugihara et al., 2004), and for obtaining fuzzy component scores in component analysis (D’Urso and Giordani, 2005). Moreover, neural networks have been used for getting nonlinear fuzzy regression models (Cheng and Lee, 2001; Dunyk and Wunsch, 2000; Ishibushi and Nii, 2001).

In this paper, dual possibilistic regression models with crisp inputs and interval outputs are deeply analyzed from Sections 2 to 5, then dual possibilistic regression models with crisp inputs and fuzzy outputs are addressed. In Section 2, the basic dual linear models based on linear programming (LP) are addressed. In Section 3, nonlinear dual models based on LP are proposed. In Section 4, nonlinear dual models based on QP are introduced. In Section 5, similarities between dual regression models and rough sets are discussed. In Section 6, dual possibilistic regression models with crisp inputs and fuzzy outputs are addressed, which are natural extensions of the interval output case. Finally, some concluding remarks are included.

2. Dual linear possibilistic regression models based on LP

An interval linear model is expressed as

$$Y = A_1 x_1 + \cdots + A_n x_n = A^t x,$$

where \(x_i\) is an input variable, \(A_i\) is an interval denoted as \(A_i = (a_i, c_i)I\) with center \(a_i\) and spread \(c_i\), \(Y\) is an estimated interval, \(x = [x_1, \ldots, x_n]^t\) is an input vector and \(A = [A_1, \ldots, A_n]^t\) is an interval coefficient vector.

The interval output in (1) can be obtained as follows (Tanaka and Guo, 1999):

$$Y(x) = (a^t x, c^t |x|)_I.$$  \hspace{1cm} (2)

When the given outputs are intervals but the given inputs are crisp, we can consider two regression models, namely, an upper regression model and a lower regression model, where the estimated interval outputs approximate to the given outputs from upper and lower directions as shown below. These two regression models are called dual possibilistic models. The given data are denoted as \((Y_j, x_{j1}, \ldots, x_{jn}) = (Y_j, x_j^I)\) where \(Y_j\) is an interval output denoted as \((y_j, e_j)^I\).

The dual possibilistic models are denoted respectively as follows:

$$Y^*_j = A^*_1 x_{j1} + \cdots + A^*_n x_{jn} \quad \text{(upper regression model)},$$  \hspace{1cm} (3)

$$Y^*_{s,j} = A^*_{s1} x_{j1} + \cdots + A^*_{sn} x_{jn} \quad \text{(lower regression model)}.$$  \hspace{1cm} (4)

The dual possibilistic regression models are described as follows:

Upper regression model: the problem here is to satisfy

$$Y_j \subseteq Y^*_j, \quad j = 1, \ldots, m,$$

and to find the interval coefficients \(A^*_i = (a^*_i, c^*_i)_I\) that minimize the sum of the spreads of the estimation intervals, that is,

$$J^* = \sum_{j=1}^{m} c^*_{i,j} |x_j|,$$  \hspace{1cm} (6)
where the minimization stems from the inclusion relations (5). Since the constraint conditions \( Y_j \subseteq Y^*_j \) can be written as
\[
y_j - e_j \geq a^{st}x_j - c^{st}|x_j|, \\
y_j + e_j \leq a^{st}x_j + c^{st}|x_j|, \tag{7}
\]
where \( a^* = [a^*_1, \ldots, a^*_n]^t \) and \( c^* = [c^*_1, \ldots, c^*_n]^t \), the problem for obtaining the interval coefficients \( A^*_j \) can be described as the following LP problem:
\[
\begin{align*}
\min_{a^*, c^*} & \quad \sum_{j=1,...,m} c^{st}|x_j| \\
\text{s.t.} & \quad y_j - e_j \geq a^{st}x_j - c^{st}|x_j|, \quad j = 1, \ldots, m, \\
& \quad y_j + e_j \leq a^{st}x_j + c^{st}|x_j|, \quad j = 1, \ldots, m, \\
& \quad c^* \geq 0. \tag{8}
\end{align*}
\]

**Lower regression model:** the problem here is to satisfy
\[
Y_{sj} \subseteq Y_j, \quad j = 1, \ldots, m \tag{9}
\]
and to find the interval coefficients \( A_{si} = (a_{si}, c_{si})_I \) that maximize the sum of the spreads of the estimation intervals:
\[
J_a = \sum_{j=1,...,m} c^i_s|x_j|, \tag{10}
\]
where the maximization stems from the inclusion relations (9). Since the constraint conditions \( Y_{sj} \subseteq Y_j \) can be written as
\[
y_j - e_j \leq a^i_sx_j - c^i_s|x_j|, \\
y_j + e_j \geq a^i_sx_j + c^i_s|x_j|, \quad j = 1, \ldots, m, \tag{11}
\]
the problem for obtaining the interval coefficients \( A_{si} \) can be described as the following LP problem:
\[
\begin{align*}
\max_{a^i_s, c^i_s} & \quad \sum_{j=1,...,m} c^i_s|x_j| \\
\text{s.t.} & \quad y_j - e_j \leq a^i_sx_j - c^i_s|x_j|, \quad j = 1, \ldots, m, \\
& \quad y_j + e_j \geq a^i_sx_j + c^i_s|x_j|, \quad j = 1, \ldots, m, \\
& \quad c^i_s \geq 0. \tag{12}
\end{align*}
\]
It can be seen that the estimated intervals from upper and lower regression models satisfy inclusion relations \( Y_{sj} \subseteq Y_j \subseteq Y^*_j \), \( j = 1, \ldots, m \).

Let us show the validity of the above formulations. Assume that the given data \( \left( Y^0_j, x^0_j \right), \quad j = 1, \ldots, m \), satisfy the linear interval system
\[
Y^0_j = A^0_{1x_1j} + \cdots + A^0_{nxnj} = A^0|x^0_j, \tag{13}
\]
where \( A^0 = (a^0, c^0)_I \).
Theorem 1. If the given data \((Y^0_j, x^0_j)\), \(j = 1, \ldots, m\), satisfy (13), the interval vector \(A^*\) and \(A^s\) obtained from (8) and (12), respectively, are the same as \(A^0\). Thus, we have

\[
A^* = A^s = A^0, \quad Y^*_j = Y^s_j = Y^0_j, \quad j = 1, \ldots, m. \tag{14}
\]

Proof. Let us prove only \(A^* = A^0\) in the upper regression model. Since \((Y^0_j, x^0_j)\) satisfies (13), we have

\[
e^0_j = c^0 |x^0_j|, \tag{15}
\]
\[
y^0_j = a^0 x^0_j. \tag{16}
\]
Substituting (15) into the constraint conditions of (8) yields that

\[
y^0_j \geq a^* x^0_j - c^* |x^0_j| + c^0 |x^0_j|, \quad y^0_j \leq a^* x^0_j + c^* |x^0_j| - c^0 |x^0_j|. \tag{17}
\]
Setting \(a^* = a^0\) and \(c^* = c^0\), \((a^0, c^0)^T\) is a feasible solution of the LP problem (8). If there is another solution \(c'\) such that

\[
\sum_{j=1}^{m} c'_t |x^0_j| < \sum_{j=1}^{m} c^0_t |x^0_j|, \tag{18}
\]
thus, for some \(l\) we have

\[
c'_t |x^0_l| < c^0_t |x^0_l|. \tag{19}
\]
The \(l\)th constraint condition of (8) can be rewritten as

\[
y^0_j \geq a^* x^0_j + k_l, \quad y^0_j \leq a^* x^0_j - k_l, \tag{20}
\]
where \(k_l\) is as follows:

\[
k_l = (c^0 - c')^t |x^0_l| > 0. \tag{21}
\]
It is obvious from the contradiction of (20) that (19) cannot hold. Thus, the optimal solution \(c^*\) should be \(c^0\). Moreover, it follows from (17) with \(c^* = c^0\) that \(y^0_j = a^* x^0_j\). Thus, \(a^*\) is equal to \(a^0\). \(\square\)

Theorem 2. There always exists an optimal solution in the upper regression model (8) while it is not assured that there is always an optimal solution in the lower regression model (12) for interval linear systems.

Proof. In the upper regression model, there is an admissible set of the constraint conditions (8) if a sufficient large positive vector is taken for \(c^*\). On the contrary, there is a case where there is no admissible set of (12) even if a zero vector is taken for \(c^*\). \(\square\)

In order to check whether there is a lower regression model, let us introduce the following conjunction model:

\[
\min_{a,c} J = \sum_{j=1,...,m} c^t |x_j| \quad \text{s.t. } Y(x_j) \cap Y_j \neq \emptyset, \quad j = 1, \ldots, m, \\
\quad c \geq 0, \tag{22}
\]
where $\Phi$ is the empty set. The constraint conditions in (22) can be rewritten as

\[a^t x_j - c^t |x_j| \leq y_j + e_j,\]
\[y_j - e_j \leq a^t x_j + c^t |x_j|,\]
\[c \geq 0.\]

(23)

**Theorem 3.** There exists a lower regression model (12) if and only if $J = 0$ in the conjunction model (22).

**Proof.** Assume that there is a lower regression model (12). This assumption can lead to $c = 0$. Thus, we have $J = 0$. Conversely, if we assume $J = \sum c^t |x_j| = 0$, the constraint conditions in (23) become

\[y_j - e_j \leq a^t x_j \leq y_j + e_j.\]

(24)

It follows from (24) that there is an admissible set in the lower regression model (12). Thus, there exists a lower regression model (12). □

To obtain the upper and lower regression models simultaneously, the following LP problem can be considered by combining (8) and (12).

\[
\min_{a^*, c^*, a_s, c_s} \sum_{j=1,\ldots,m} c^t_j |x_j| - \sum_{j=1,\ldots,m} c^t_j |x_j|
\]

s.t.  
\[Y_j^* \supseteq Y_j, \quad j = 1, \ldots, m,\]
\[Y_{s,j} \subseteq Y_j, \quad j = 1, \ldots, m,\]
\[A^*_i \supseteq A_{s,i}, \quad i = 1, \ldots, n,\]
\[c^* \geq 0,\]
\[c_s \geq 0,\]

(25)

which is also the LP problem. Thus, it is easy to obtain the interval coefficients $A^*_i, A_{s,i}, \ i = 1, \ldots, n$ by solving the LP problem (25). $Y^*(x)$ and $Y_s(x)$ are used to denote interval estimation outputs from the upper and lower regression models for any vector $x$, respectively.

**Theorem 4.** The optimal upper and lower estimations $Y^*(x)$ and $Y_s(x)$ in the LP problem (25) have the inclusion relation such that $Y^*(x) \supseteq Y_s(x)$ for any $x$.

**Proof.** Recalling (3) and (4), for any $x = [x_1, \ldots, x_l]^t$, $Y^* = A^*_{1} x_1 + \cdots + A^*_{n} x_n$ and $Y_s = A_{s,1} x_1 + \cdots + A_{s,n} x_n$ hold. Without loss of generality, $Y^* = A^*_{1} x_1 + A^*_{2} x_2$ and $Y_s = A_{s,1} x_1 + A_{s,2} x_2$ are considered. From (25) it is known that $A^*_1 \supseteq A_{s,1}$ and $A^*_2 \supseteq A_{s,2}$ hold. According to interval arithmetic, if $A^*_1 \supseteq A_{s,1}$, then $A^*_1 x_1 \supseteq A_{s,1} x_1$ and if $A^*_2 \supseteq A_{s,2}$, then $A^*_2 x_2 \supseteq A_{s,2} x_2$, and if $A^*_1 x_1 \supseteq A_{s,1} x_1$ and $A^*_2 x_2 \supseteq A_{s,2} x_2$, then $A^*_1 x_1 + A^*_2 x_2 \supseteq A_{s,1} x_1 + A_{s,2} x_2$. That is $Y^*(x) \supseteq Y_s(x)$. □

3. **Dual nonlinear possibilistic regression models based on LP**

As described in Section 2 the upper and lower regression models can be obtained for the interval outputs. Unfortunately, the solution of the lower regression model (12) does not always exist for a linear regression model. Let us consider the following polynomial as a regression model:

\[Y = A_0 + \sum A_i x_i + \sum A_{il} x_i x_l + \sum A_{iljk} x_i x_l x_k + \cdots,\]

(26)

where $A_0, A_i, A_{il}$ and $A_{iljk}$ are intervals.

In general, it is well-known that any functional relation between inputs and outputs can be approximated by a proper polynomial. Therefore it is assured that there exists a lower regression model as (26). It should be noted that (26) is a linear system with regard to interval coefficients. Thus, interval coefficients can be obtained by solving the LP problem.
We can check the existence of a lower regression model by the conjunction model (22) with Theorem 3. If there is no lower regression model, increase the number of terms of the polynomial (26) until a solution is found. Furthermore, let us define a measure of fitness for the \( j \)th data unit as follows:

\[
\varphi_Y(z_j) = \frac{d^*_t |z_j|}{d^*_t |z_j|},
\]

where the terms \( x_i, x_i x_l, \ldots \) in (26) are regarded as new variables whose values for \( j \)th input data unit form vector \( z_j \), and the spreads of whose upper interval coefficients and lower interval coefficients form vector \( d^*_t \) and \( d^*_t \), respectively. \( \varphi_Y(z_j) \) indicates how close the upper output is to the lower output for the \( j \)th input unit. Then, the measure of fitness \( \varphi_Y \) for all data units can be defined as

\[
\varphi_Y = \frac{\sum_{j=1}^{m} \varphi_Y(z_j)}{m} = \left( \frac{\sum_{j=1}^{m} d^*_t |z_j|}{d^*_t |z_j|} \right) / m,
\]

where \( 0 \leq \varphi_Y \leq 1 \). Even if a lower regression model does not exist, we can set \( 0 \) for \( d^*_t \). The larger the value of \( \varphi_Y \), the more the model is fitting to the data. If the given input–output data satisfy a regression system (26), then we can obtain the upper and lower regression models which are identical. In this case, \( \varphi_Y \) becomes 1. The procedure for obtaining upper and lower polynomial models is as follows:

**Step 1:** Take a linear function as a regression model:

\[
Y = A_0 + \sum A_i x_i. \tag{29}
\]

**Step 2:** Solve the conjunction problem (22). If there is an optimal solution with \( J = 0 \), go to Step 4. Otherwise, go to Step 3.

**Step 3:** Increase the terms of the polynomials, for example,

\[
Y = A_0 + \sum A_i x_i + \sum A_{il} x_i x_l. \tag{30}
\]

Go to Step 2.

**Step 4:** Solve the LP problem (25) and calculate the measure of fitness \( \varphi_Y \). If \( \varphi_Y \geq \omega \), then go to Step 5 where \( 0 < \omega \leq 1 \) is a tolerance limit given by an analyst. Otherwise, go to Step 3.

**Step 5:** End the procedure.

### Numerical example.

The data of crisp inputs and interval outputs are shown in Table 1. The tolerance limit \( \omega \) is set as 0.25 because we want to obtain such upper and lower regression models that the spread of estimated output from the lower model is one-fourth of the spread of estimated output from the upper model. Let us explain our proposed algorithm as the following sequence.

Assume the following linear function as a regression model:

\[
Y = A_0 + A_1 x. \tag{31}
\]

Solve the conjunction problem (22) using the given data. The value of the objective function does not satisfy \( J = 0 \). Thus increase the terms of the polynomials as

\[
Y = A_0 + A_1 x + A_2 x^2. \tag{32}
\]
Solve the conjunction problem (22). Also the condition of $J = 0$ is not satisfied. Therefore, increase the terms of polynomial again as
\[ Y = A_0 + A_1 x + A_2 x^2 + A_3 x^3. \] (33)

Solving the conjunction problem (22), we obtain the model as
\[ \hat{Y} = (10.8889, 0)I + (6.5992, 0)I x - (1.6032, 0)I x^2 - (0.1151, 0)I x^3, \] (34)
where $J = 0$. Using the model (34), we solve the LP problem (25). Then we obtain the upper regression model $Y^*(x)$ and the lower regression model $Y_*(x)$ as follows:
\[ Y^* = (6.9150, 1.8850)I + (9.2274, 0)I x - (2.1884, 0)I x^2 + (0.1598, 0.0012)I x^3, \] (35)
\[ Y_* = (7.4894, 0.5356)I + (9.2274, 0)I x - (2.1884, 0)I x^2 + (0.1589, 0.0003)I x^3. \] (36)

The upper and lower models are depicted in Fig. 1. The value of $\varphi_Y(x_j)$ for each $x_j$ is shown in Table 2. The measure of fitness is greater than the tolerance limit, that is, $\varphi_Y(=0.2812) > \omega(=0.25)$. Thus, we accept (35) and (36) as optimal models, which satisfy $Y^*(x) \supseteq Y_*(x)$ for any $x$.

This example shows that we can obtain suitable upper and lower regression models, whose measure of fitness $\varphi_Y$ is larger than the given tolerance limit $\omega$, by increasing the terms of the polynomials.

4. Dual nonlinear possibilistic regression models based on QP

We solve the upper and the lower regression models (3) and (4) by QP problems with crisp inputs and interval outputs. In order to guarantee that $Y_*(x_j) \subseteq Y_j \subseteq Y^*(x_j)$ for any arbitrary $x_j$, it is convenient to assume that
\[ A^*_i = (a_i, c_i + d_i)I \quad \text{and} \quad A_{si} = (a_i, c_i)I, \quad i = 0, \ldots, n. \] (37)
The objective function is introduced as the following quadratic function:

\[ J = \sum_{j=1}^{m} (d^t|x_j|)^2 = d^t \left( \sum_{j=1}^{m} |x_j||x_j|^t \right) d, \] (38)

which is the sum of squared spread differences between upper and lower regression model. Therefore, interval regression analysis is to determine the interval coefficients \( A^*_i \) and \( A^*_{i,0} \), \( i = 0, \ldots, n \) that minimize the objective function (38) and satisfy inclusion relations \( Y^*_i (x_j) \subseteq Y_j \subseteq Y^* (x_j) \), \( j = 1, \ldots, m \), which can be described as the following QP problem:

\[
\begin{align*}
\min_{a, c, d} & \quad J = d^t \left( \sum_{j=1}^{m} |x_j||x_j|^t \right) d + \xi (a^t a + c^t c) \\
\text{s.t.} & \quad a^t x_j + c^t x_j + d^t|x_j| \geq y_j + e_j, \quad j = 1, \ldots, m, \\
& \quad a^t x_j - c^t x_j - d^t|x_j| \leq y_j - e_j, \quad j = 1, \ldots, m, \\
& \quad a^t x_j + c^t x_j \leq y_j + e_j, \quad j = 1, \ldots, m, \\
& \quad a^t x_j - c^t x_j \geq y_j - e_j, \quad j = 1, \ldots, m, \\
& \quad c \geq 0,
\end{align*}
\] (39)

where \( \xi \) is a small positive number. \( \xi (a^t a + c^t c) \) is added into the objective function (38) so that (39) becomes a strictly convex quadratic programming. It should be noted that \( d \geq 0 \) is not included in the constraints of (39) because it is redundant.

As mentioned before, there is always an upper linear model but maybe not lower linear model for a given crisp input and interval output data. In the case that no lower linear regression model exists, we can take the following steps for obtaining the upper and lower polynomial regression models:

**Step 1:** Take a linear function as a regression model:

\[ Y = A_0 + \sum A_i x_i. \] (40)

**Step 2:** Solve the problem (39) and calculate the measure of fitness \( \varphi_Y \) defined in (28). For the case of no existence of feasible solution in (39), \( \varphi_Y \) is set as 0. If \( \varphi_Y \geq \omega \), then go to Step 4 where \( \omega \) is a tolerance limit given by an analyst. Otherwise, go to Step 3.

**Step 3:** Increase the terms of the polynomials, for example,

\[ Y = A_0 + \sum A_i x_i + \sum A_{ij} x_i x_j. \] (41)

Go to Step 2.

**Step 4:** End the procedure.

**Numerical example.** The data set of crisp inputs and interval outputs is shown in Table 3.

The linear interval model for the given input and output is assumed as

\[ Y(x) = A_0 + A_1 x, \] (42)
and the tolerance limit is set as 0.25. Using (39), we obtained the upper regression model $Y^*(x)$ and the lower regression model $Y_*(x)$ as

$$Y^*(x) = (7.311, 11.856) + (8.371, 2.462)x, \quad (43)$$

$$Y_*(x) = (7.311, 0.168) + (8.371, 0.514)x, \quad (44)$$

which are depicted in Fig. 2 where two outer lines represent the upper model $Y^*(x)$ and two inner lines represent the lower model $Y_*(x)$. The measure of fitness $\phi_Y$ is obtained as 0.101, which is smaller than $\omega = 0.25$. Let us change the linear model (42) into the following nonlinear model:

$$Y(x) = A_0 + A_1x + A_2x^2. \quad (45)$$

Using (39) again, the upper regression model $Y^*(x)$ and the lower regression model $Y_*(x)$ are obtained as

$$Y^*(x) = (10.463, 13.241) + (5.648, 1.944)x + (0.370, 0)x^2, \quad (46)$$

$$Y_*(x) = (10.463, 1.204) + (5.648, 1.019)x + (0.370, 0)x^2, \quad (47)$$

which are depicted in Fig. 3 where two outer lines represent the upper model $Y^*(x)$ and two inner lines represent the lower model $Y_*(x)$. In this case, the measure of fitness $\phi_Y$ is obtained as 0.252. It is easy to see that from Figs. 2 and 3 that polynomial (45) can make the upper and lower regression models closer than the linear function (42).
Table 4
Similarities between rough sets and dual regression models

<table>
<thead>
<tr>
<th>Rough sets</th>
<th>Dual regression models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper approximation: $R^*(Z)$</td>
<td>Upper interval output: $Y^*(x)$</td>
</tr>
<tr>
<td>Lower approximation: $R_*(Z)$</td>
<td>Lower interval output: $Y_*(x)$</td>
</tr>
<tr>
<td>Cardinality of $R^<em>(Z)$: Card($R^</em>(Z)$)</td>
<td>Spread of $Y^<em>(x)$: $e_{t</em>}</td>
</tr>
<tr>
<td>Cardinality of $R_<em>(X)$: Card($R_</em>(Z)$)</td>
<td>Spread of $Y_<em>(x)$: $e_{t</em>}</td>
</tr>
<tr>
<td>Inclusion relation $R^<em>(Z) \subseteq Z \subseteq R_</em>(Z)$</td>
<td>Inclusion relation $Y^<em>(x) \supseteq Y(x) \supseteq Y_</em>(x)$</td>
</tr>
<tr>
<td>Accuracy measure of a set $Z \alpha(Z)$</td>
<td>Measure of fitness: $\frac{1}{m} \sum_{i=1}^{m} (e_{t*}^i</td>
</tr>
</tbody>
</table>

5. Similarities between dual regression models and rough sets

Rough set theory has been proposed by Pawlak and extensively applied to classification problems, machine learning, and decision analysis, etc. (Pawlak, 1984). For comparing dual possibilistic regression models with rough sets, the basic notions of rough sets are introduced below.

Let $U$ be the universe of objects and $R$ be an equivalence relation in $U$. Then by $U/R$ we mean the family of all equivalence classes of $R$. Equivalence classes of the relation $R$ are called elementary sets. Any finite union of elementary sets is said to be a definable set. Given a set $Z$, the upper and lower approximations of $Z$, denoted as $R^*(Z)$ and $R_*(Z)$, respectively are two definable sets defined as follows:

$$R^*(Z) = \{ Y \in U/R : Y \cap Z \neq \emptyset \},$$ (48)

$$R_*(Z) = \{ Y \in U/R : Y \subseteq Z \},$$ (49)

where $\emptyset$ is the empty set. An accuracy measure of a set $Z$, denoted as $\alpha(Z)$, is defined as

$$\alpha(Z) = \frac{\text{Card}(R_*(Z))}{\text{Card}(R^*(Z))},$$ (50)

where $\text{Card}(R_*(Z))$ and $\text{Card}(R^*(Z))$ are the cardinalities of $R_*(Z)$ and $R^*(Z)$.

In fact, the upper and lower approximations of $Z$ can be regarded as the optimal solutions of the following optimization problem.

$$\max_{R_u(Z), R_l(Z)} \alpha(Z) = \frac{\text{Card}(R_l(Z))}{\text{Card}(R_u(Z))}$$

$$\text{s.t. } R_l(Z) \subseteq Z \subseteq R_u(Z),$$ (51)

where $R_u(Z)$ and $R_l(Z)$ are two definable sets by $U/R$. It is straightforward that maximizing $\alpha(Z)$ leads to maximizing $\text{Card}(R_l(Z))$ and minimizing $\text{Card}(R_u(Z))$ constrained by $R_l(Z) \subseteq Z \subseteq R_u(Z)$. As a result, the optimal solution of $R_u(Z)$ is the least definable set containing the set $Z$ and $R_l(Z)$ is the greatest definable set contained in $Z$, that is, the upper and lower approximations of $Z$, $R^*(Z)$ and $R_*(Z)$, respectively. Likewise, $R^*(X)$ and $R_*(X)$ are corresponding to the upper and lower estimated interval outputs $Y_*(x)$ and $Y_*(x)$ in dual possibilistic regression models. Moreover, finer the equivalence classes are, better the approximation quality is in rough sets. Similarly, more the number of terms of polynomials is, better the measure of fitness in dual regression models is. Therefore it can be concluded that polynomials in dual regression models play the similar role of an equivalence relation in rough sets.

Lastly, we can summarize the similarities between our models and rough sets in Table 4.

6. Dual possibilistic regression models with fuzzy outputs

From Sections 2 to 5, as a core part of possibilistic regression, dual regression models for crisp input and interval output have been analyzed in depth. From now, let us consider how to approximate the functional relations between inputs and outputs from lower and upper directions with fuzzy outputs.
**Definition 1.** A fuzzy number \( A \) is called an \( L-R \) fuzzy number and denoted as \( (a, c_l, c_r)_{LR} \) if its membership function is defined by

\[
\Pi_A(x) = \begin{cases}
L \left( \frac{(a-x)}{c_l} \right), & x \leq a, \\
1, & x = a, \\
R \left( \frac{(x-a)}{c_r} \right), & x \geq a,
\end{cases}
\]

where \( c_l > 0, c_r > 0 \) and reference functions \( L : [0, +\infty) \to [0, 1] \) and \( R : [0, +\infty) \to [0, 1] \) are strictly decreasing functions with \( L(0) = 1 \) and \( R(0) = 1 \).

An \( n \)-dimensional vector \( x = [x_1, \ldots, x_n]^t \) can be fuzzified as an \( L-R \) fuzzy vector \( A \) whose membership function is defined as

\[
\Pi_A(x) = \Pi_{A_1}(x_1) \land \cdots \land \Pi_{A_n}(x_n),
\]

where \( \land \) is min operator and \( \Pi_{A_i}(x_i) \) is the membership function of an \( L-R \) fuzzy number, denoted as \( (a_i, c_{li}, c_{ri})_{LR} \). The \( n \)-dimensional \( L-R \) fuzzy vector is denoted as \( A = (a, c_l, c_r)_{LR} \) with \( a = [a_1, \ldots, a_n]^t \), \( c_l = [c_{l1}, \ldots, c_{ln}]^t \) and \( c_r = [c_{r1}, \ldots, c_{rn}]^t \).

Consider a fuzzy linear system

\[
Y = A_1 x_1 + \cdots + A_n x_n = A^t x,
\]

where \( x_i \) is a real number and \( A \) is an \( n \)-dimensional \( L-R \) fuzzy vector whose element is \( (a_i, c_{li}, c_{ri})_{LR} \). From the extension principle, it is known that \( Y \) is an \( L-R \) fuzzy number as follows (Tanaka and Guo, 1999):

\[
Y = \left( \sum_{i=1}^{n} x_i a_i, \sum_{i=1}^{n} x_i c_{li}, \sum_{i=1}^{n} x_i c_{ri} \right)_{LR} = (a^t x, c^t_{l} | x |, c^t_{r} | x |)_{LR}.
\]

Its \( h \)-level set, denoted as \([Y]_h\), is as follows:

\[
[Y]_h = \left[ a^t x - L^{-1}(h)c^t_{l} | x |, a^t x + R^{-1}(h)c^t_{r} | x | \right],
\]

where \( 0 < h \leq 1 \).

When the given outputs denoted as \( Y_j = (y_j, e_{lj}, e_{rj})_{LR}, \ j = 1, \ldots, m \) are fuzzy numbers and the given inputs \( x_j = [x_{j1}, \ldots, x_{jn}]^t, \ j = 1, \ldots, m \) are crisp, we can also consider upper and lower regression models as follows:

\[
Y^*_j = A^*_1 x_{j1} + \cdots + A^*_n x_{jn}, \\
Y^*_s = A^*_s x_{j1} + \cdots + A^*_s x_{jn},
\]

where \( A^*_s \) and \( A^*_s \) are \( L-R \) fuzzy numbers denoted as \( A^*_j = (a^*_j, c^*_{lj}, c^*_{rj})_{LR} \) and \( A^*_s = (a^*_s, c^*_{ls}, c^*_{rs})_{LR} \).

It is natural to extend dual regression models with interval outputs into fuzzy output based on the following assumptions.

1. The fuzzy output \( Y^*_j \) from the upper model (57) includes the given fuzzy output \( Y_j \) in its \( h \)-level set.
2. The fuzzy output \( Y^*_s \) from the lower model (58) is included by the given fuzzy output \( Y_j \) in its \( h \)-level set.
3. The sum of spreads of fuzzy outputs of (57) is minimized.
4. The sum of spreads of fuzzy outputs of (58) is maximized.

**Upper regression model:** The problem is to obtain the optimal coefficients \( A^*_j = (a^*_j, c^*_{lj}, c^*_{rj})_{LR}, i = 1, \ldots, n \) that minimize

\[
J^* = \sum_{j=1}^{m} (c^*_{lj} + c^*_{rj})^t | x_j |
\]

s.t. \([Y^*_j]_h \subseteq [Y_j]_h, \quad j = 1, \ldots, m\).
where \( c_i^* = [c_{i1}, \ldots, c_{im}] \), \( c_r^* = [c_{r1}, \ldots, c_{rn}] \) and \( 0 < h \leq 1 \). The constraint in (59) can be rewritten as
\[
\begin{align*}
\left[ y_j - L^{-1}(h)e_{lj}, y_j + R^{-1}(h)e_{rj} \right] \subseteq \left[ a^{st}x_j - L^{-1}(h)c_i^{st}, a^{st}x_j + R^{-1}(h)c_r^{st} \right],
\end{align*}
\]
which yields the following inequalities:
\[
\begin{align*}
y_j - L^{-1}(h)e_{lj} \geq a^{st}x_j - L^{-1}(h)c_i^{st} |x_j|, \\
y_j + R^{-1}(h)e_{rj} \leq a^{st}x_j + R^{-1}(h)c_r^{st} |x_j|,
\end{align*}
\]
where \( a^* = [a_1^*, \ldots, a_m^*] \). Thus, the LP problem for obtaining the optimal fuzzy coefficients can be described as
\[
\begin{align*}
\min_{a^*, c_i^*, c_r^*} \quad & J^* = \sum_{j=1}^{m} (c_i^* + c_r^*) |x_j| \\
\text{s.t.} \quad & y_j - L^{-1}(h)e_{lj} \geq a^{st}x_j - L^{-1}(h)c_i^{st} |x_j|, \quad j = 1, \ldots, m, \\
& y_j + R^{-1}(h)e_{rj} \leq a^{st}x_j + R^{-1}(h)c_r^{st} |x_j|, \quad j = 1, \ldots, m, \\
& c_i^* \geq 0, \\
& c_r^* \geq 0.
\end{align*}
\]

**Lower regression model:** The problem is to obtain the optimal coefficients \( A_{kl} = (a_{kl}, c_{kl}, c_{rl})_{LR} \), \( i = 1, \ldots, n \) that maximize
\[
\begin{align*}
J_a = \sum_{j=1}^{m} (c_{si} + c_{sr}) |x_j| \\
\text{s.t.} \quad & [Y_{sj}]_h \subseteq [Y_j]_h, \quad j = 1, \ldots, m,
\end{align*}
\]
where \( c_{si} = [c_{si1}, \ldots, c_{sinn}] \), \( c_{sr} = [c_{sr1}, \ldots, c_{srn}] \) and \( 0 < h \leq 1 \). The constraint in (63) can be rewritten as
\[
\begin{align*}
y_j - L^{-1}(h)e_{lj} \leq a_i^{st}x_j - L^{-1}(h)c_i^{st} |x_j|, \\
y_j + R^{-1}(h)e_{rj} \geq a_i^{st}x_j + R^{-1}(h)c_r^{st} |x_j|,
\end{align*}
\]
where \( a_i = [a_{i1}, \ldots, a_{im}] \). Thus, the LP problem for obtaining the optimal fuzzy coefficients can be described as
\[
\begin{align*}
\max_{a_i, c_{si}, c_{sr}} \quad & J_a = \sum_{j=1}^{m} (c_{si} + c_{sr}) |x_j| \\
\text{s.t.} \quad & y_j - L^{-1}(h)e_{lj} \leq a_i^{st}x_j - L^{-1}(h)c_i^{st} |x_j|, \quad j = 1, \ldots, m, \\
& y_j + R^{-1}(h)e_{rj} \geq a_i^{st}x_j + R^{-1}(h)c_r^{st} |x_j|, \quad j = 1, \ldots, m, \\
& c_{si} \geq 0, \\
& c_{sr} \geq 0.
\end{align*}
\]

According to the formulations of upper and lower models, it can be seen that
\[
[Y_{sj}]_h \subseteq [Y_j]_h \subseteq [Y_j^*]_h,
\]
which shows that the upper model can be characterized as the smallest upper bound, while the lower model can be characterized as the greatest lower bound for the \( h \)-level sets. It is easy to understand that the dual linear models with fuzzy outputs are a natural extension of the dual linear models with interval outputs with considering the \( h \)-level sets of fuzzy numbers. Similarly, it is not difficult to build dual nonlinear possibilistic regression models with fuzzy output by LP and QP problems.

Lastly, we can mention a remark on how to determine the value of \( h \)-level as follows. If you think the given data is enough to reflect all of possible cases, you set \( h \) as zero so that you can obtain dual possibilistic regression models.
with the narrowest spreads of outputs. If you think the given data is not enough to reflect all of possible cases, you set \( h \) be higher so that you can obtain dual possibilistic regression models with wider spreads of outputs, which means the wider range should be considered. The details on this relation between \( h \)-level and the obtained width are given in Tanaka and Guo (1999).

7. Conclusions

In this paper, firstly possibilistic regression for crisp inputs and interval outputs is considered where the basic dual linear models based on LP, dual nonlinear models based on LP and dual nonlinear models based on QP are systematically addressed. Then dual possibilistic regression models with crisp inputs and fuzzy outputs are proposed as a natural extension of the interval output case. The obtained upper and lower regression models are used to characterize the uncertain functional relation between crisp inputs and interval (or fuzzy) outputs. The given outputs include the lower estimated outputs, but are included by the upper estimated outputs. From this aspect, the lower possibilistic regression corresponds to necessity and the upper possibilistic regression corresponds to possibility (Tanaka and Guo, 1999). In other words, if the given output is regarded as a kind of necessity, then the estimated output from the lower possibilistic regression model should be recommended; if the given output is regarded as a kind of possibility, then the estimated output from the upper possibilistic regression model should be recommended. Linear possibilistic models are simple. However, there is no guarantee that the lower possibilistic model always exists. Nonlinear possibilistic regression can give the upper and lower possibilistic regression models by adding the terms of polynomials. Inclusion relationship between the given output and the estimated output is the core for possibilistic regression analysis, which is very different from the fuzzy least squares approach minimizing the fuzzy distance. However, dual possibilistic regression models have strong common points with rough sets. Both of them characterize the uncertainty from upper and lower directions by a pair of well-known systems. For possibilistic regressions, a functional relation is used. For rough sets, an equivalence relation is used. Possibilistic regression models suggest solutions from possibility and necessity viewpoints for the given uncertain output, which are very useful for analyzing the data with uncertainty extensively existing in business and economics society.

References