Int. J. Fuzzy Syst.
https://doi.org/10.1007/s40815-019-00692-0

Ridge Fuzzy Regression Model

Seung Hoe Choi[1] · Hye-Young Jung[2] · Hyoshin Kim[3]

Received: 15 March 2019 / Revised: 12 June 2019 / Accepted: 27 June 2019

Abstract Ridge regression model is a widely used model with many successful applications, especially in managing correlated covariates in a multiple regression model. Multicollinearity represents a serious threat in fuzzy regression models as well. We address this issue by combining ridge regression with the fuzzy regression model. Our proposed algorithm uses the α-level estimation method to evaluate the parameters of the ridge fuzzy regression model. Two examples are given to illustrate the ridge fuzzy regression model with crisp input/fuzzy output and fuzzy coefficients.

Keywords Ridge regression · Multicollinearity · Ridge fuzzy regression model · Fuzzy multiple linear regression model

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea Government (MSIT) (Nos. 2017R1C1B1005069, 2017R1D1A1B03029559).

✉ Hye-Young Jung
hyjunglove@snu.ac.kr

Seung Hoe Choi
shchoi@kau.ac.kr

Hyoshin Kim
hyoshinkim@snu.ac.kr

[1] School of Liberal Arts and Science, Korea Aerospace University, Seoul, South Korea

[2] Faculty of Liberal Education, Seoul National University, Seoul, South Korea

[3] Department of Statistics, Seoul National University, Seoul, South Korea

1 Introduction

Linear regression models and ridge regression models are statistical methods which have been widely used in many instances. They have proven to be strong in applications of precise datasets. However, the data types we meet are not always precise. We often encounter non-precise data such as linguistic data or imprecise data, the data structures in which classical statistical theory cannot manage. In order to handle these types of data, the fuzzy regression model has been proposed. Fuzzy regression model was originally developed by Tanaka et al. [1] and has been since then utilized to analyze unconventional datasets [1–20]. As in multiple regression models in traditional statistical theory, multicollinearity is also an important issue in fuzzy multiple regression models. We introduce the ridge fuzzy regression model which combines ridge regression with the fuzzy regression model in order to reduce the effect of multicollinearity when it is present. To construct the ridge fuzzy regression model, we propose the α-level estimation algorithm based on the α-level ridge loss function.

This paper is organized as follows: Sect. 2 introduces the linear regression model and the ridge regression model. Section 3 describes the fuzzy multiple regression model and the algorithm for the proposed ridge fuzzy regression model using the α-level estimation method. In Sect. 4, we illustrate the ridge fuzzy regression model with crisp input/fuzzy output and fuzzy coefficients. The performance of the ridge fuzzy regression model is shown in comparison with fuzzy multiple regression and other existing methods, with applications to two examples. Finally, Sect. 5 concludes the paper.

Published online: 17 July 2019

2 Ridge Regression Model

2.1 Linear Regression Model

The multiple linear regression model is formulated as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i, i = 1, \ldots, n. \quad (1)$$

The least squares estimator for the model is

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} \quad (2)$$

where $\mathbf{X} = (\boldsymbol{X}_1', \ldots, \boldsymbol{X}_n')'$ and $\boldsymbol{Y} = (y_1, \ldots, y_n)$.

When $p > n$, or when in the presence of multicollinearity, that is, two or more of the covariates in a multiple linear regression model are highly correlated, $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ fails to exist. In such cases, the least squares estimator is not unique. The ridge regression model can effectively handle this problem.

2.2 Ridge Regression Model

The ridge regression model was originally introduced by Hoerl and Kennard [21] to resolve the problem of the least square estimator when $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ does not exist. The motivation for the ridge estimator is to replace $\boldsymbol{X}'\boldsymbol{X}$ by $\boldsymbol{X}'\boldsymbol{X} + \lambda\boldsymbol{I}$, where λ is a positive constant. This makes the matrix $\boldsymbol{X}'\boldsymbol{X} + \lambda\boldsymbol{I}$ invertible.

Ridge regression minimizes the conventional criterion of least squares in the following way [21]:

$$\hat{\beta}^{\text{ridge}} = \min_{\beta}\left[\sum_{i=1}^{n}\left(y_i - \beta_0 - \sum_{k=1}^{p} x_{ik}\beta_k\right)^2 + \lambda\sum_{k=1}^{p}\beta_k^2\right] \quad (3)$$

where λ is a positive tuning parameter pre-defined by the researcher. Another explanation as to why we use the ridge regression model can be given in terms of the bias–variance trade-off. As λ increases, coefficients become close to zero as to derive the penalty term small. Flexibility decreases, and as a result, variance decreases and bias increases. Therefore, we can have better results than least squares estimation by significantly reducing variance at the cost of some loss in bias.

3 Ridge Fuzzy Regression Model

In real-world problems, we come across linguistic data or imprecise data, for example, words such as "young," "tall," or "high." Fuzzy sets introduced by Zadeh [22] have been applied to represent such data types. Fuzzy regression models making use of fuzzy data, represented by fuzzy sets, have been studied over the last decades. Its applications have shown to be favorable in various fields of study. Similar to traditional multiple regression methods, we often encounter the multicollinearity phenomenon in fuzzy multiple regression models as well. Thus, the ridge fuzzy regression model which joins together the ridge regression model with fuzzy data is needed to lessen the problems associated with multicollinearity.

3.1 Fuzzy Numbers

Following Zadeh [22–24], we briefly summarize important notions regarding fuzzy numbers and the distance between them. A fuzzy set is a set of ordered pairs $A = \{(x, \mu_A(x)) : x \in X\}$ where $\mu_A(x) : X \to [0, 1]$ is a membership function representing the degree of membership of x in a set A. The support of a fuzzy set is defined by $S(A) = \{x \in R | \mu_A(x) > 0\}$. For any α in [0,1], the α-level set of a fuzzy set A is a crisp (classical) set $A_\alpha = \{x \in R^1 : \mu_A(x) \geq \alpha\}$ which contains all the elements in X, while its membership value in A is greater than or equal to α. The α-level set of a fuzzy set A can be represented by $A(\alpha) = [l_A(\alpha), r_A(\alpha)]$.

When A is a crisp set, its membership function can take on only the values zero or one depending on whether x does or does not belong to A, respectively. $\mu_A(x)$ reduces to the indicator function $I_A(x)$ of a set A in such cases.

It is assumed that a fuzzy number, A, is a normal and convex fuzzy subset of the real line, R, with bounded support. Often the following parametric class of fuzzy numbers, the so-called LR-fuzzy numbers, is used as a special case.

$$\mu_A(x) = \begin{cases} L\left(\dfrac{m-x}{l}\right) & \text{if} \quad x \leq m, \\ R\left(\dfrac{x-m}{l}\right) & \text{if} \quad x > m, \end{cases} \text{for } x \in R \quad (4)$$

Here, $L, R : R \to [0, 1]$ are fixed, left-continuous, and non-increasing functions with $R(0) = L(0) = 1$, and $R(1) = L(1) = 0$. L and R are the left and right shape functions of A, respectively. m is the mode of A. l and $r > 0$ are the left and right spreads of A, respectively. The spreads l and r represent the fuzziness of a fuzzy number and may be symmetric or non-symmetric. When $l = r = 0$, a fuzzy number becomes a precise real number with no fuzziness. Thus, a precise real number can be considered as a special case of a fuzzy number. For a precise observation $a \in R$, its corresponding membership function is $\mu_A(a) = 1$.

The resolution identity theorem proposed by Zadeh [24] states that a fuzzy set can be represented by an α-level set or by its membership function. Let A be a fuzzy number with a membership function $\mu_A(x)$ and an α-level set $A(\alpha)$. Then, we have that $\mu_A(x) = \text{Sup}\{\alpha \cdot I_{A(\alpha)}(x) : \alpha \in [0, 1]\}$.

In the fuzzy set theory, triangular and trapezoidal fuzzy numbers are special cases of LR-fuzzy numbers. They are used extensively throughout various researches [25]. In this paper, we adopt a triangular fuzzy number in our proposed algorithm. The membership function of a triangular fuzzy number $A = (a_l, a_m, a_r)_{TR}$ is given by

$$\mu_A(x) = \begin{cases} \dfrac{x - a_l}{a_m - a_l} & \text{if} \quad x \leq m, \\ \dfrac{a_l - x}{a_r - a_m} & \text{if} \quad x > m, \end{cases} \quad \text{for} \quad x \in R \tag{5}$$

where $a_l, a_m,$ and a_r are the left end-point, mid-point, and right end-point, respectively.

3.2 Fuzzy Regression Model

The fuzzy multiple linear regression model is given as follows:

$$Y_i = A_0 \oplus A_1 \odot X_{i1} \oplus \cdots \oplus A_p \odot X_{ip} \oplus E_i, \quad i = 1, \ldots, n \tag{6}$$

where $\oplus$ and $\odot$ are addition and multiplication between two fuzzy numbers, respectively. For the arithmetic operations, see [25]. Since the crisp data are a special case of fuzzy data, the ordinary regression model is also a special case of the fuzzy regression model.

Many estimation methods exist which estimate the parameters for the fuzzy regression model. In this paper, we modify the α-level estimation method proposed by Choi et al. [15]. By convention, we usually assume that the mid-point y_{im} of Y_i has been centered and $X_{ik} (k = 1, 2, \ldots, p)$ have been standardized.

3.3 Ridge Fuzzy Regression Model

We use the α-level ridge loss function to estimate parameters of out ridge fuzzy regression model. It reduces the ordinary ridge loss function if we use the ordinary ridge regression model. The α-level estimation algorithm based on the α-level ridge loss function is as follows:

Step 1 Create an α-level of the dependent variable Y. For any positive integer k, the set of an α-level is given by $A = \left\{\alpha_j : \alpha_j \in (0, 1), j = 1, 2, \ldots, s\right\} \cup \{0, 1\}$.

Step 2 Use the ridge estimation method to find the estimator $\widehat{l_{A_k}}(1)$ and $\widehat{r_{A_k}}(1)$ of $l_{A_k}(1)$ and $r_{A_k}(1)$ by minimizing the following α-level ridge loss functions

$$\sum_{i=1}^{n}\left(l_{Y_i}(1) - \sum_{k=1}^{p} l_{A_k}(1) l_{X_{ik}}(1)\right)^2 + \lambda \sum_{k=1}^{p} l_{A_k}(1)^2 \tag{7}$$

and

$$\sum_{i=1}^{n}\left(r_{Y_i}(1) - \sum_{k=1}^{p} r_{A_k}(1) r_{X_{ik}}(1)\right)^2 + \lambda \sum_{k=1}^{p} r_{A_k}(1)^2, \tag{8}$$

respectively.

Step 3 For given set A in Step 1, let $\alpha^* = \max_{\alpha_j} A$. Then, find the intermediate estimators $\overline{l_{A_k}}(\alpha^*)$ and $\overline{r_{A_k}}(\alpha^*)$ by minimizing

$$\sum_{i=1}^{n}\left(l_{Y_i}(\alpha^*) - \sum_{k=1}^{p} l_{A_k}(\alpha^*) l_{X_{ik}}(\alpha^*)\right)^2 + \lambda \sum_{k=1}^{p} l_{A_k}(\alpha^*)^2 \tag{9}$$

and

$$\sum_{i=1}^{n}\left(r_{Y_i}(\alpha^*) - \sum_{k=1}^{p} r_{A_k}(\alpha^*) r_{X_{ik}}(\alpha^*)\right)^2 + \lambda \sum_{k=1}^{p} r_{A_k}(\alpha^*)^2, \tag{10}$$

respectively. We then obtain the estimators $\widehat{l_{A_k}}(\alpha^*)$ and $\widehat{r_{A_k}}(\alpha^*)$ by modifying the intermediate estimators $\overline{l_{A_k}}(\alpha^*)$ and $\overline{r_{A_k}}(\alpha^*)$ so that the estimated coefficients obtained from Step 3 form a pre-defined shape of the membership function. For this, the following *min* and *max* operators are used.

$$\widehat{l_{A_k}}(\alpha^*) = \min\left\{\overline{l_{A_k}}(\alpha^*), \widehat{l_{A_k}}(1)\right\} \tag{11}$$

and

$$\widehat{r_{A_k}}(\alpha^*) = \max\{\overline{r_{A_k}}(\alpha^*), \widehat{r_{A_k}}(1)\}. \tag{12}$$

Step 4 For any $\alpha_j \in (0, 1)$, find the intermediate estimators $\overline{l_{A_k}}(\alpha_j)$ and $\overline{r_{A_k}}(\alpha_j)ofl_{A_k}(\alpha_j)andr_{A_k}(\alpha_j)$, using the ridge loss functions given in Step 3. Next, find the estimators $\widehat{l_{A_k}}(\alpha_j)$ and $\widehat{r_{A_k}}(\alpha_j)$ of $l_{A_k}(\alpha_j)$ and $r_{A_k}(\alpha_j)$ by modifying the intermediate estimators so that the intermediate estimators obtained above form a pre-defined shape of the membership function.

$$\widehat{l_{A_k}}(\alpha_j) = \begin{cases} \mathrm{Max}\left\{\widehat{l_{A_k}}(\alpha^*), \mathrm{Min}\left\{\overline{l_{A_k}}(\alpha_j), \widehat{l_{A_k}}(1)\right\}\right\} & \text{if}\, \alpha_j \leq \alpha^*, \\ \mathrm{Min}\left\{\overline{l_{A_k}}(\alpha_j), \widehat{l_{A_k}}(\alpha^*)\right\} & \text{if}\, \alpha_j > \alpha^*. \end{cases} \tag{13}$$

and

$$\widehat{r_{A_k}}(\alpha_j) = \begin{cases} \mathrm{Min}\{\widehat{r_{A_k}}(\alpha^*), \mathrm{Max}\{\overline{r_{A_k}}(\alpha_j), \widehat{r_{A_k}}(1)\}\} & \text{if}\, \alpha_j \leq \alpha^*, \\ \mathrm{Max}\{\overline{r_{A_k}}(\alpha_j), \widehat{r_{A_k}}(\alpha^*)\} & \text{if}\, \alpha_j > \alpha^*. \end{cases} \tag{14}$$

Step 5 Repeat the same process to find $\widehat{l_{A_k}}(0)$ and $\widehat{r_{A_k}}(0)$ of $l_{A_k}(0)$ and $r_{A_k}(0)$.

Table 1 Input–output data concerning house prices

No.	$Y = (y, s)$	x_1	x_2	x_3
1	(606, 100)	38.09	36.43	5
2	(710, 50)	62.10	26.50	6
3	(808, 100)	63.76	44.71	7
4	(826, 150)	74.52	38.09	8
5	(865, 250)	75.38	41.40	7
6	(852, 200)	52.99	26.49	4
7	(917, 200)	62.93	26.49	5
8	(1031, 250)	72.04	33.12	6
9	(1092, 600)	76.12	43.06	7
10	(1203, 100)	90.26	42.64	7
11	(1394, 350)	85.70	31.33	6
12	(1420, 250)	95.27	27.64	6
13	(1601, 300)	105.98	27.64	6
14	(1632, 500)	79.25	66.81	6
15	(1699, 650)	120.50	32.25	6

y house price (100,000 yen), x_1 first-floor space (m^2), x_2 second-floor space (m^2), x_3 number of rooms

Table 2 Fitted values of house prices

No.	Fitted values		
	$Y = (y, s)$	$\widehat{Y_{reg}}$	$\widehat{Y_{\text{ridge}}}$
1	(606, 100)	(611.08, 523.15)	(649.93, 456.95)
2	(710, 50)	(740.94, 596.52)	(791.36, 525.15)
3	(808, 100)	(843.94, 748.89)	(882.84, 656.15)
4	(826, 150)	(782.98, 764.85)	(852.90, 672.21)
5	(865, 250)	(1005.96, 796.06)	(1028.82, 699.16)
6	(852, 200)	(908.83, 539.24)	(904.17, 474.02)
7	(917, 200)	(919.96, 601.65)	(931.63, 529.73)
8	(1031, 250)	(1003.14, 710.54)	(1016.43, 625.10)
9	(1092, 600)	(1040.75, 813.65)	(1057.95, 714.40)
10	(1203, 100)	(1285.09, 899.16)	(1273.84, 790.84)
11	(1394, 350)	(1221.09, 782.36)	(1210.32, 689.69)
12	(1420, 250)	(1341.92, 813.70)	(1320.32, 718.66)
13	(1601, 300)	(1531.15, 880.95)	(1487.21, 778.69)
14	(1632, 500)	(1571.18, 1018.42)	(1485.94, 890.67)
15	(1699, 650)	(1847.99, 1008.06)	(1762.34, 890.87)

Step 6 Find the membership functions $\mu_{A_k}(x)$ for the fuzzy regression coefficients $A_k(k = 0, 1, \ldots, p)$, by performing linear regression on the estimated α-level sets $\widehat{A_k}(\alpha_j) = \left[\widehat{l_{A_k}}(\alpha_j), \widehat{r_{A_k}}(\alpha_j)\right](j = 1, 2, \ldots, s)$, $\widehat{A_k}(0)$, and $\widehat{A_k}(1)$. Give a constraint so that the top of the predefined membership function satisfies the condition of α-level 1. For example, give the following constraint

$$\mu_{\widehat{A_k}}\left(\widehat{l_{A_k}}(1)\right) = \mu_{\widehat{A_k}}(\widehat{r_{A_k}}(1)) = 1. \tag{15}$$

4 Numerical Examples

We illustrate the performance of our ridge fuzzy regression model through two examples. We use the following fuzzy multiple regression model with crisp input/fuzzy output and fuzzy coefficients:

$$Y_i = A_0 \oplus A_1 x_{i1} \oplus \cdots \oplus A_p x_{ip} \oplus E_i, \quad i = 1, \ldots, n. \tag{16}$$

The mid-point of Y_i is assumed to be centered, and covariates $x_{ik}(k = 1, 2, \ldots, p)$ are assumed to be standardized.

To show the performance of the estimated ridge fuzzy regression model, we apply the following two performance measures RMSE$_F$ (root mean square error of fuzzy numbers) and MAPE$_F$ (mean absolute percentage error of fuzzy numbers). For this, we use the Diamond distance [11], a popularly used measure to define distance between fuzzy numbers. Denote the observed value as $Y_i = (y_{il}, y_{im}, y_{ir})_{TR}$ and the fitted value as $\widehat{Y_i} = (\widehat{y_{il}}, \widehat{y_{im}}, \widehat{y_{ir}})_{TR}$. RMSE$_F$ and MAPE$_F$ are defined as follows.

$$\begin{aligned}\text{RMSE}_F &= \frac{1}{n}\sum_{i=1}^{n} d^2\left(Y_i, \widehat{Y_i}\right) \\ &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left\{(y_{il} - \widehat{y_{il}})^2 + (y_{im} - \widehat{y_{im}})^2 + (y_{ir} - \widehat{y_{ir}})^2\right\}}\end{aligned} \tag{17}$$

and

$$\text{MAPE}_F = \frac{100\%}{n}\sum_{i=1}^{n}\left(\left|\frac{y_{il} - \widehat{y_{il}}}{y_{il}}\right| + \left|\frac{y_{im} - \widehat{y_{im}}}{y_{im}}\right| + \left|\frac{y_{ir} - \widehat{y_{ir}}}{y_{ir}}\right|\right). \tag{18}$$

In the two examples, symmetric fuzzy numbers are used. Thus, we apply slight modification to the algorithm presented in Sect. 3. In Step 4 of our algorithm, we choose $\min\{\widehat{l_{A_k}}(0), \widehat{r_{A_k}}(0)\}$ and $\min\{\widehat{l_{A_k}}(\alpha_j), \widehat{r_{A_k}}(\alpha_j)\}(j = 1, 2, \ldots, s)$ based on RMSE$_F$. Then, the opposite left end-point or the right end-point is obtained symmetrically to make a symmetric fuzzy number. The remaining steps of the algorithm are the same.

4.1 Example 1

We apply our proposed ridge fuzzy regression model and the fuzzy multiple linear regression model to the house price data taken from Tanaka [1, 3]. The house price data are shown in Table 1. The vagueness s (1000 yen) is assigned by the authors. For both models, the α-levels were

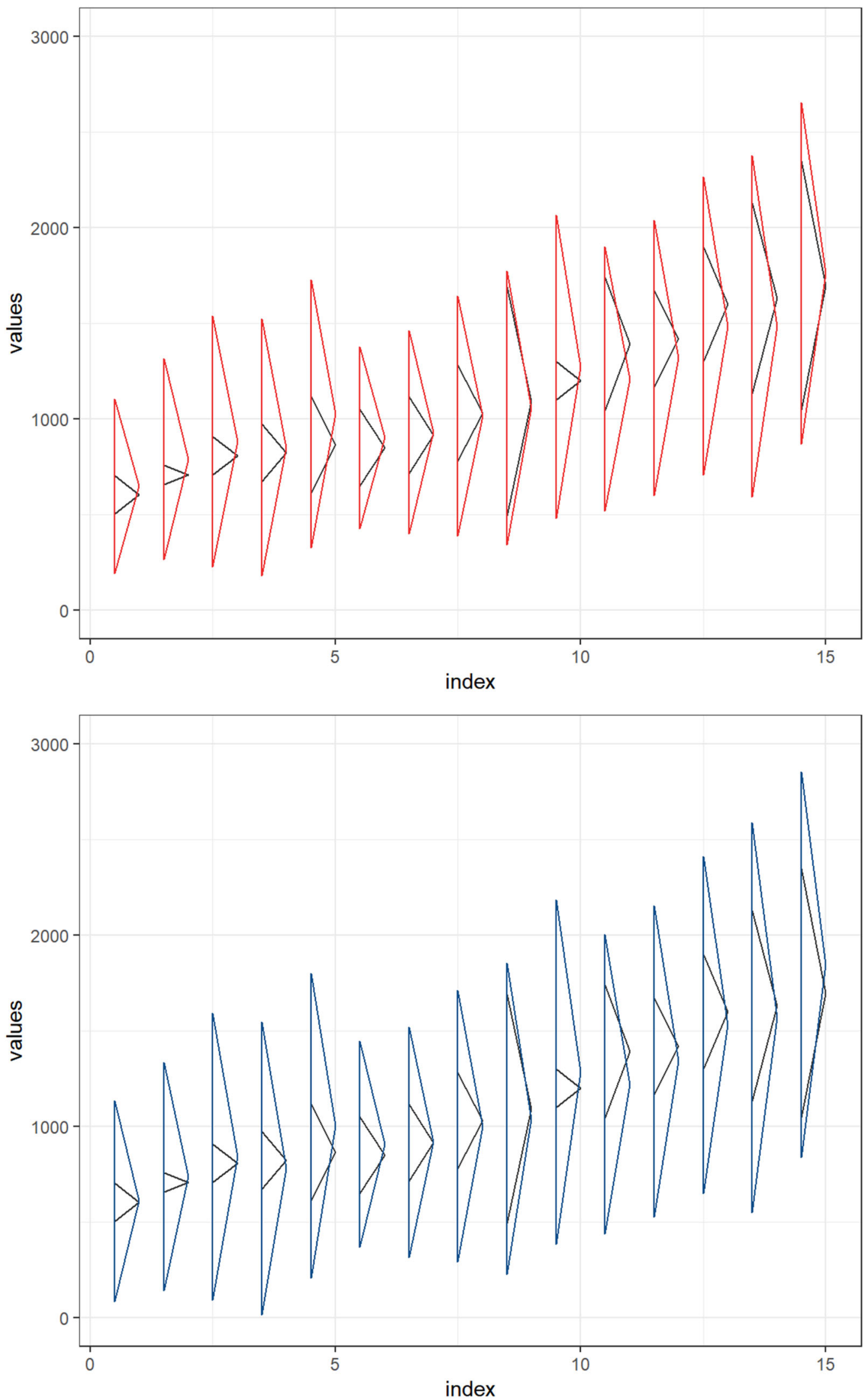

Fig. 1 The observed and estimated values of the dependent variable. Above: ridge fuzzy regression model; below: fuzzy multiple linear regression model

Table 3 Modified coefficients $\widehat{l_{A_k}}(\alpha_j)$ and $\widehat{r_{A_k}}(\alpha_j)$ obtained in Steps 3–6 of the proposed algorithms

α-levels	Optimal lambda	$\widehat{l_{A_0}}(\alpha)$	$\widehat{l_{A_1}}(\alpha)$	$\widehat{l_{A_2}}(\alpha)$	$\widehat{l_{A_3}}(\alpha)$
1	32.95	307.36	15.58	10.60	− 127.44
0.75	30.13	307.36	14.21	8.96	− 127.44
0.5	27.31	307.36	12.85	7.34	− 127.44
0.25	24.49	307.36	11.52	5.74	− 127.44
0	105.37	307.36	7.41	2.11	− 127.44

α-levels	Optimal lambda	$\widehat{r_{A_0}}(\alpha)$	$\widehat{r_{A_1}}(\alpha)$	$\widehat{r_{A_2}}(\alpha)$	$\widehat{r_{A_3}}(\alpha)$
1	32.95	307.36	15.58	10.60	− 127.44
0.75	35.77	307.36	16.97	12.26	− 127.44
0.5	38.59	307.36	18.37	13.93	− 127.44
0.25	41.41	307.36	19.78	15.61	− 127.44
0	44.23	307.36	21.20	17.30	− 127.44

α-levels	$\widehat{l_{A_0}}(\alpha)$	$\widehat{l_{A_1}}(\alpha)$	$\widehat{l_{A_2}}(\alpha)$	$\widehat{l_{A_3}}(\alpha)$
1	284.07	17.67	13.08	− 164.49
0.75	284.07	16.10	11.13	− 164.49
0.5	284.07	14.53	9.18	− 164.49
0.25	284.07	12.96	7.23	− 164.49
0	284.07	11.39	5.28	− 164.49

α-levels	$\widehat{r_{A_0}}(\alpha)$	$\widehat{r_{A_1}}(\alpha)$	$\widehat{r_{A_2}}(\alpha)$	$\widehat{r_{A_3}}(\alpha)$
1	284.07	17.67	13.08	− 164.49
0.75	284.07	19.24	15.03	− 164.49
0.5	284.07	20.81	16.98	− 164.49
0.25	284.07	22.38	18.93	− 164.49
0	284.07	23.95	20.87	− 164.49

Above: ridge fuzzy regression model; below: fuzzy multiple linear regression model

set as 1, 0.75, 0.5, 0.25, and 0. The fitted values for both methods are given in Table 2. Results show the fitted values for the ridge fuzzy regression method more accurately describe the original data than the fuzzy multiple linear regression approach. This is clarified in addition in Fig. 2. In the plot of the observed-versus-fitted values, a comparison of the ridge fuzzy regression model and the fuzzy multiple linear regression model is shown. The black triangles correspond to the observed values, and the red and blue triangles to the fitted values from ridge fuzzy multiple regression, and the fuzzy multiple regression, respectively. Both methods estimated the mid-points well. The spread lengths, however, are shorter for our ridge fuzzy regression model than the others (Fig. 1).

An analysis of the α-level dependent regression coefficients is shown as follows. In Steps 3 through 6 of our algorithm, we obtained the estimators $\widehat{l_{A_k}}(\alpha_j)$ and $\widehat{r_{A_k}}(\alpha_j)$, $k = 0, 1, \ldots, p$, $j = 1, 2, \ldots, s$ by modifying the intermediate estimators so that the coefficients form a pre-defined shape of the membership function. The estimated coefficients have been altered so that each of the estimators for both methods is symmetric fuzzy numbers. The estimators $\widehat{l_{A_k}}(\alpha_j)$ and $\widehat{r_{A_k}}(\alpha_j)$ are shown in Table 3. Note that the ridge fuzzy regression coefficients for $k = 1, \ldots, p$ have shrunk toward 0 compared to the fuzzy linear regression estimators for all α-levels.

We made a slight modification to the algorithm of Sect. 3 to obtain the final estimators $\widehat{l_{A_k}}(\alpha_j)$ and $\widehat{r_{A_k}}(\alpha_j)$. We performed linear regression on the estimated α-level sets $\widehat{A_k}(\alpha_j) = \left[\widehat{l_{A_k}}(\alpha_j), \widehat{r_{A_k}}(\alpha_j)\right](j = 1, 2, \ldots, s)$, $\widehat{A_k}(0)$, and $\widehat{A_k}(1)$. Then in Step 4 of our algorithm, we chose $\min\{\widehat{l_{A_k}}(0), \widehat{r_{A_k}}(0)\}$ and

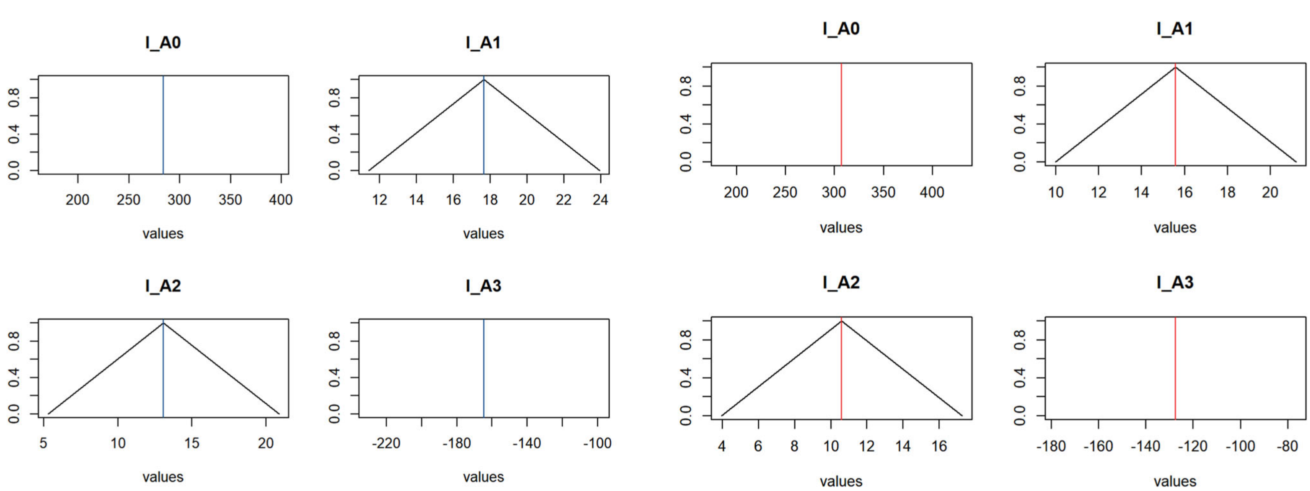

Fig. 2 The plot of fuzzy coefficients. Above: ridge fuzzy regression model; below: fuzzy multiple linear regression model

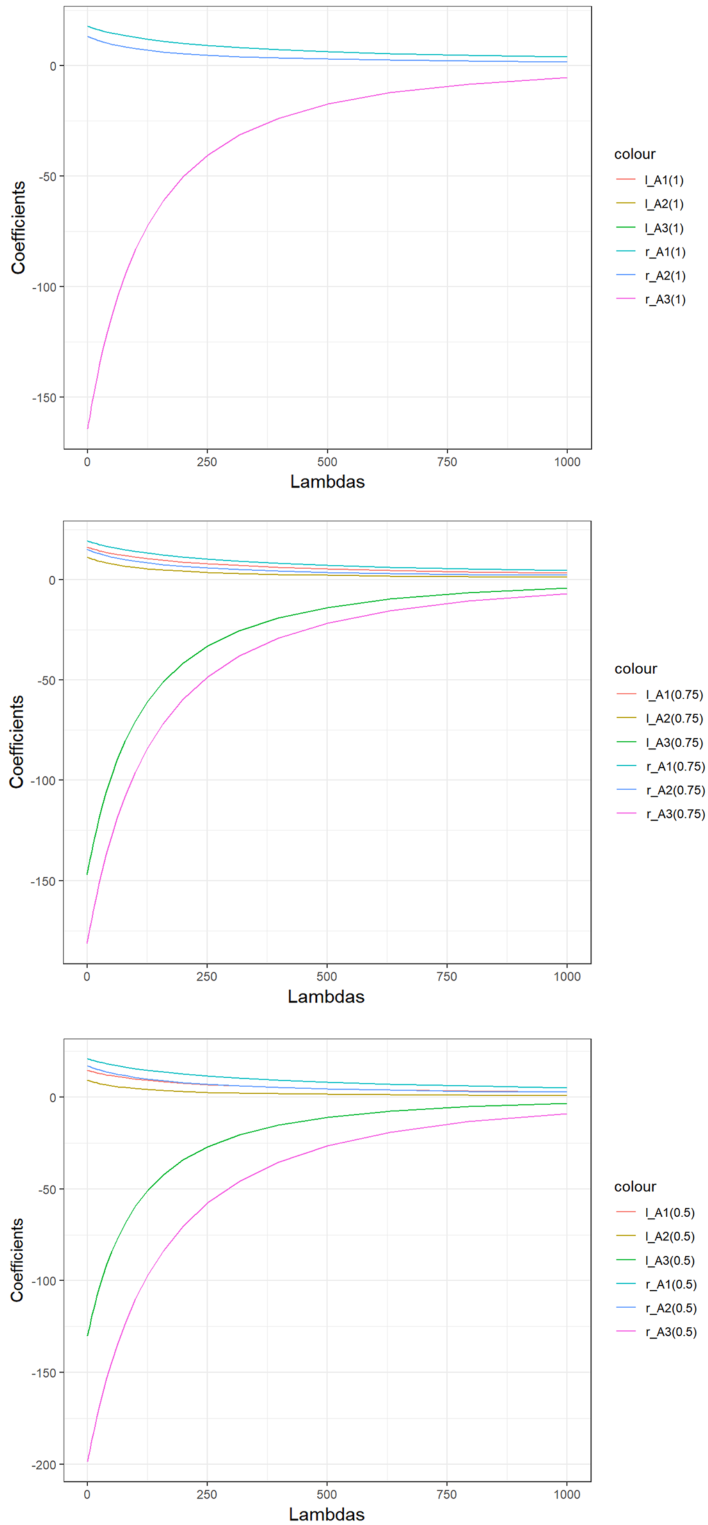

Fig. 3 Coefficient profiles for α-levels: 1, 0.75, 0.5, 0.25, and 0

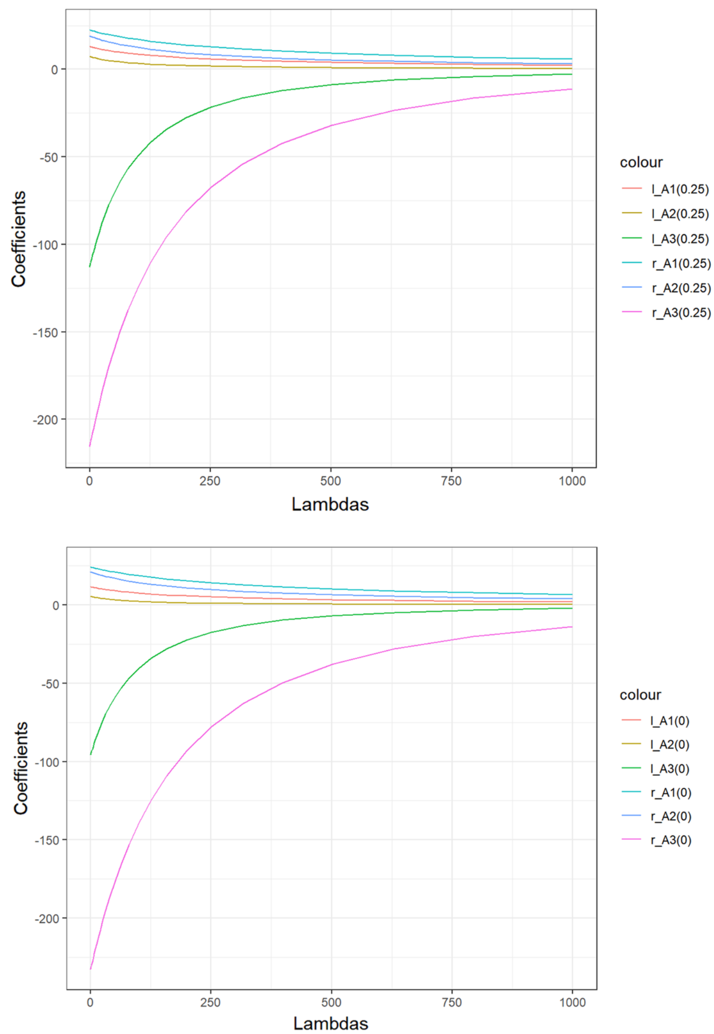

Fig. 3 continued

$\min\{\widehat{l_{A_k}}(\alpha_j), \widehat{r_{A_k}}(\alpha_j)\}(j = 1, 2, \ldots, s)$ based on which left or right slope of the pre-defined membership function improved the performance of RMSE_F. The fitted equation based on the estimators $\widehat{l_{A_k}}(\alpha_j)$ and $\widehat{r_{A_k}}(\alpha_j)$ of the ridge fuzzy regression and the fuzzy multiple linear regression method is given as follows, respectively.

$$\begin{aligned}\widehat{Y_{\text{ridge}}} = {} & (307.36,\ 307.36,\ 307.36) \\ & + (9.98,\ 15.58,\ 21.19)x_1 \\ & + (3.92,\ 10.60,\ 17.29)x_2 \\ & + (-127.44,\ -127.44,\ -127.44)x_3.\end{aligned} \tag{19}$$

$$\begin{aligned}\widehat{Y_{\text{reg}}} = {} & (284.07,\ 284.07,\ 284.07) \\ & + (11.39,\ 17.67,\ 23.95)x_1 \\ & + (5.28,\ 13.08,\ 20.87)x_2 \\ & + (-164.49,\ -164.49,\ -164.49)x_3.\end{aligned} \tag{20}$$

Table 4 RMSE and MAPE

	RMSE_F	MAPE_F (%)
Ridge fuzzy regression	624.62	92
Fuzzy regression	743.26	109
SVM regression	1178.04	137

To give a better picture of the obtained estimators, the plot of fuzzy coefficients for the ridge fuzzy regression and the fuzzy multiple linear regression, respectively, is shown in Fig. 2.

To check whether the ridge fuzzy regression method behaves as in the classical ridge regression model, we performed the ridge fuzzy regression analysis over a range of lambda values at each of the α-levels 1, 0.75, 0.5, 0.25, and 0. Figure 3 shows that all α-level intermediate

Table 5 The perfect multicollinearity problem

Data			Fitted values
$Y = (y, s)$	x_1	x_2	$\widehat{Y_{\text{ridge}}}$
(12.5, 5.0)	2.2	3.6	(13.48, 4.71)
(17.0, 5.0)	3.2	5.4	(17.74, 4.86)
(18.5, 5.0)	3.5	6.0	(19.16, 4.91)
(22.0, 5.0)	4.2	7.4	(22.48, 5.02)
(24.0, 5.0)	4.6	8.2	(24.38, 5.09)
(26.5, 5.0)	5.1	9.2	(26.75, 5.17)
(28.0, 5.0)	5.4	9.8	(28.17, 5.22)
(31.5, 5.0)	6.1	11.2	(31.49, 5.34)
(33.5, 5.0)	6.5	12.0	(33.39, 5.41)
(37.0, 5.0)	7.2	13.4	(36.70, 5.52)
(38.5, 6.0)	7.5	14.0	(38.13, 5.57)
(41.5, 6.0)	8.1	15.2	(40.97, 5.67)
(43.5, 6.0)	8.5	16.0	(42.87, 5.74)
(46.5, 6.0)	9.1	17.2	(45.71, 5.84)
(49.0, 6.0)	9.6	18.2	(48.08, 5.92)

coefficients $\overline{l_{A_k}}(\alpha_j)$ and $\overline{r_{A_k}}(\alpha_j)$ of the ridge fuzzy regression model shrink to zero as lambda increases.

Finally, the performance measures introduced previously were checked. Table 4 shows the RMSE_F and MAPE_F for the ridge fuzzy regression model compared to the fuzzy multiple linear regression model and the fuzzy SVM regression proposed by Hao and Chiang [26]. Both performance measures are greatly reduced for our proposed ridge fuzzy regression model compared to the other methods.

4.2 Example 2

Next, we show that the proposed ridge fuzzy regression model can handle the perfect multicollinearity problem. The dataset shown in Table 5 is taken from Hong et al. [27]. The data are an example of a situation in which one variable is a perfect linear combination of the other variable. Note that in this case, the fuzzy multiple linear regression coefficients and its fitted values cannot be computed due to perfect collinearity.

The fitted values for the ridge fuzzy regression are given in Table 5. Results show the fitted values for the ridge

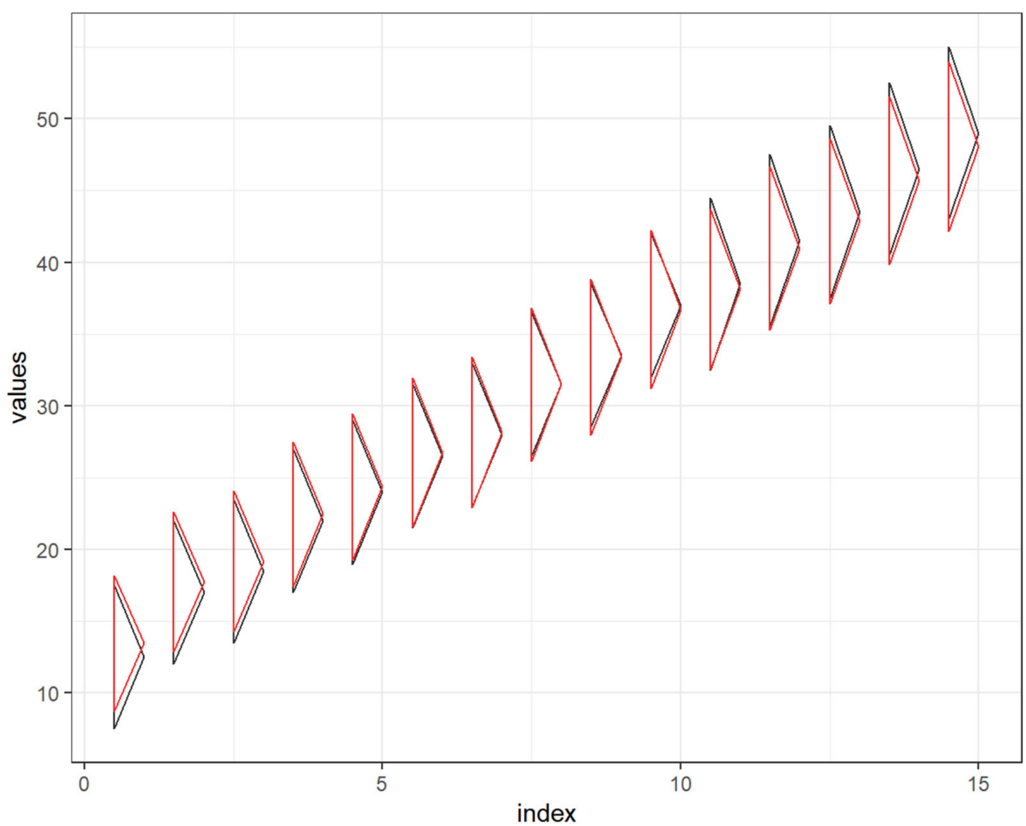

Fig. 4 The observed and fitted values of the dependent variable, ridge fuzzy regression method

fuzzy regression method accurately describe the output, with near-accurate prediction of spread lengths. Figure 4 shows the plot of the observed-versus-fitted values for the proposed model. The black triangles correspond to the observed values, and the red triangles to the fitted data. Both the mid-points and the spread lengths for the fitted values are almost identical to the original observed values.

The ridge fuzzy regression coefficients $\widehat{l_{A_k}}(\alpha_j)$ and $\widehat{r_{A_k}}(\alpha_j)$, $k = 0, 1, \ldots, p$, $j = 1, 2, \ldots, s$ obtained in Steps 3 through 6 of our algorithm are shown in Table 6. As in Example 1, the ridge fuzzy regression coefficients were obtained by modifying the intermediate estimators $\overline{l_{A_k}}(\alpha_j)$ and $\overline{r_{A_k}}(\alpha_j)$. They have been modified so that estimators form symmetric fuzzy numbers. Note that here the estimators for the fuzzy multiple linear regression coefficients could not be obtained due to perfect multicollinearity. The fitted equation based on the final obtained coefficients $\widehat{l_{A_k}}(\alpha_j)$ and $\widehat{r_{A_k}}(\alpha_j)$ is given as follows. Here, the final coefficients were obtained through a slight modification in the algorithm of Sect. 3. Linear regression was performed on the estimated α-level sets $\widehat{A_k}(\alpha_j) = \left[\widehat{l_{A_k}}(\alpha_j), \widehat{r_{A_k}}(\alpha_j)\right]$ $(j = 1, 2, \ldots, s)$, $\widehat{A_k}(0)$, and $\widehat{A_k}(1)$, and we chose $\min\{\widehat{l_{A_k}}(0), \widehat{r_{A_k}}(0)\}$ and $\min\{\widehat{l_{A_k}}(\alpha_j), \widehat{r_{A_k}}(\alpha_j)\}(j = 1, 2, \ldots, s)$ based on which left or right slope of the triangular membership function improved RMSE_F performance.

$$\begin{aligned}\widehat{Y_{\text{ridge}}} &= (-0.61,\ 3.76,\ 8.12) \\ &\quad + (2.29,\ 2.38, 2.46)x_1 \\ &\quad + (1.14,\ 1.18, 1.22)x_2\end{aligned} \tag{21}$$

We clarify the form of the pre-defined membership function presented by showing the plot of fuzzy coefficients for the ridge fuzzy regression in Fig. 5.

The intermediate coefficients $\overline{l_{A_k}}(\alpha_j)$ and $\overline{r_{A_k}}(\alpha_j)$ of the fuzzy ridge regression model were checked by an analysis

Table 6 Modified coefficients $\widehat{l_{A_k}}(\alpha_j)$ and $\widehat{r_{A_k}}(\alpha_j)$ obtained in Steps 3–6 of the ridge fuzzy regression model

α-levels	Optimal lambda	$\widehat{l_{A_0}}(\alpha)$	$\widehat{l_{A_1}}(\alpha)$	$\widehat{l_{A_2}}(\alpha)$
1	1.20	3.76	2.38	1.18
0.75	1.19	2.67	2.35	1.17
0.5	1.18	1.57	2.33	1.16
0.25	1.17	0.48	2.31	1.15
0	1.15	− 0.61	2.29	1.14
α-levels	Optimal lambda	$\widehat{r_{A_0}}(\alpha)$	$\widehat{r_{A_1}}(\alpha)$	$\widehat{r_{A_2}}(\alpha)$
1	1.20	3.76	2.38	1.18
0.75	1.21	4.85	2.40	1.19
0.5	1.22	5.94	2.42	1.20
0.25	1.23	7.03	2.44	1.21
0	1.24	8.12	2.46	1.22

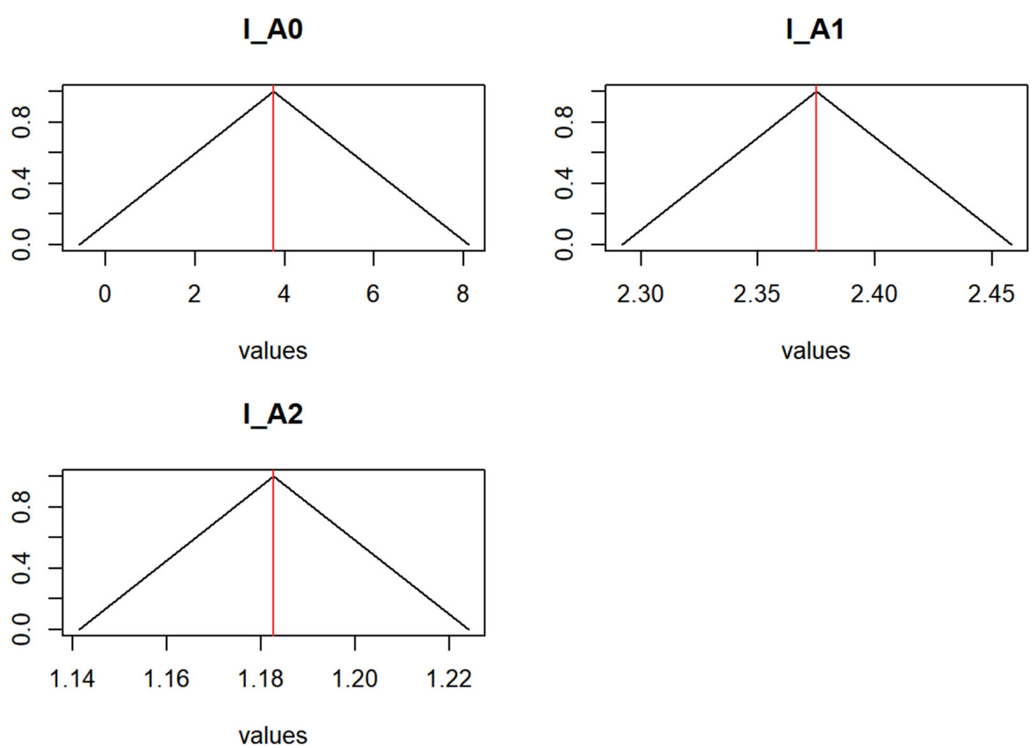

Fig. 5 The plot of fuzzy coefficients for the ridge fuzzy regression model

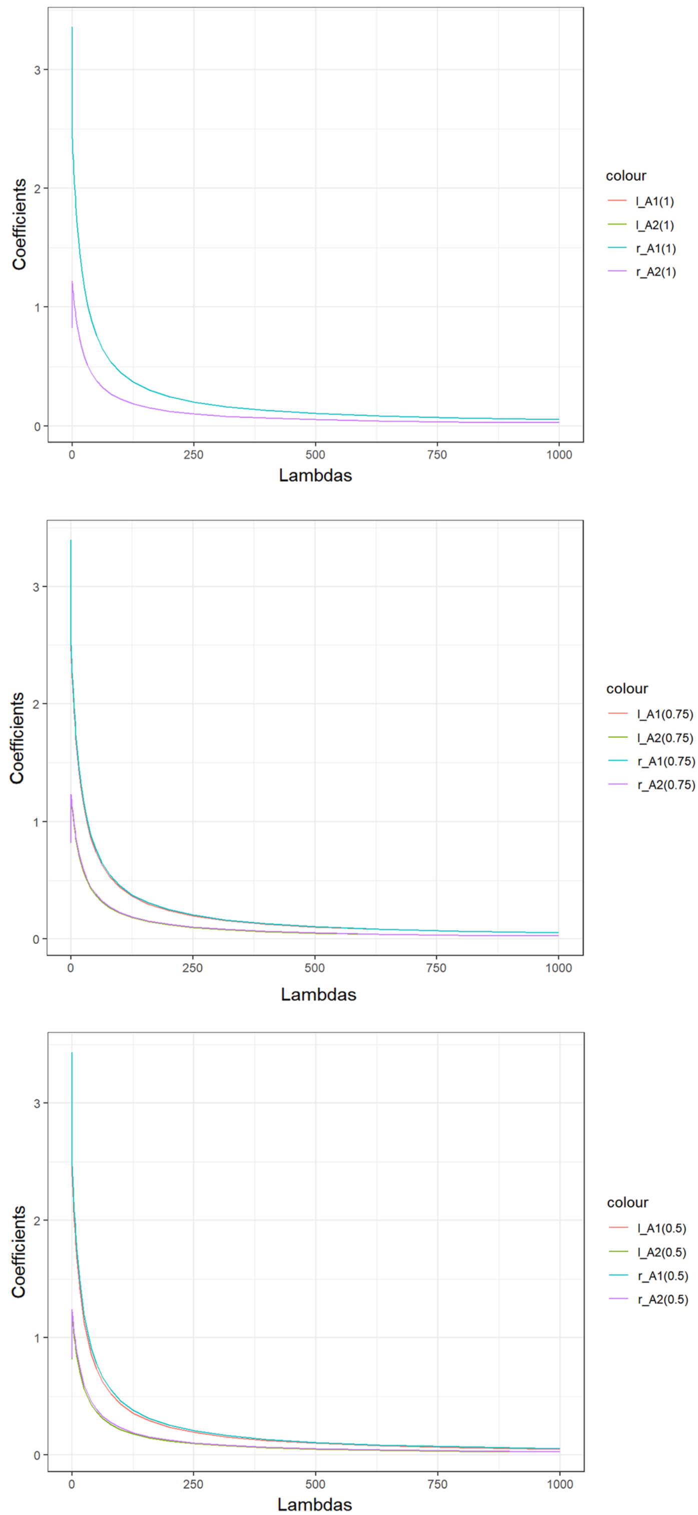

Fig. 6 Coefficient profiles for α-levels 1, 0.75, 0.5, 0.25 and 0

over a range of lambda values at each of the α-levels 1, 0.75, 0.5, 0.25, and 0. All α-level coefficients of our method shrunk to zero as lambda increases, as shown in Fig. 6.

Table 7 shows the $RMSE_F$ and $MAPE_F$ for the ridge fuzzy regression model in comparison with the fuzzy regression model. The $RMSE_F$ and $MAPE_F$ could not be

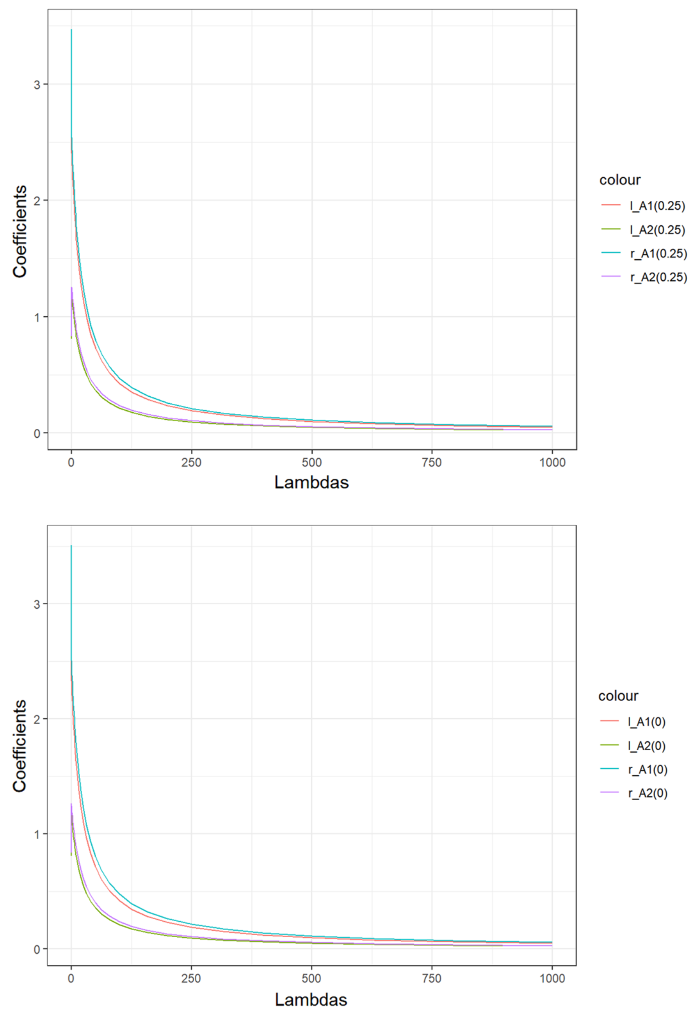

Fig. 6 continued

Table 7 RMSE and MAPE

	$RMSE_F$	$MAPE_F$
Ridge fuzzy regression	1.06	7%
Fuzzy regression	NA	NA

computed for the fuzzy multiple linear regression method due to collinearity between predictors.

5 Conclusion

Handling multicollinearity in multiple linear regression models is an important topic in statistics. The phenomenon can lead to inaccurate estimates of the regression coefficients, inflate their standard errors, give non-significant p-values, and degrade the predictability of the fitted model.

Ridge regression was originally motivated by Hoerl and Kennard [21] to reduce the effect of multicollinearity in its presence. When two or more of the covariates in a multiple linear regression model are highly correlated, the columns of the design matrix become linearly dependent. Thus, the ordinary regression parameter cannot be estimated. With ridge regression, one can also improve model performance by lowering model complexity. As models become more complex, insignificant local structures can be picked up, known as overfitting. In such cases, coefficient estimates suffer from high variance as more terms are included in the model. Ridge regression lowers the parameter dimension, allowing some bias to achieve a lower variance, thus improving overall model performance.

The fuzzy regression model is one of the most widely used statistical models in fuzzy statistical analyses. As in traditional regression models, multicollinearity poses a problem in fuzzy regression models as well. In this paper, we propose the α-level estimation algorithm based on the α-level ridge loss function to estimate the parameters for the ridge fuzzy regression model. The fuzzy ridge regression model constructs a fuzzy linear function that minimizes the sum of squared error term and penalty term, whereas a fuzzy multiple linear regression model builds a fuzzy linear function that minimizes the squared error. To show the performance of the proposed α-level estimation algorithm, two examples are presented. The house price data in Example 1 are one of the most famous datasets in fuzzy regression analysis. The second dataset in Example 2 gives an example of perfect multicollinearity. Both results show the proposed ridge fuzzy regression model is superior to the fuzzy multiple regression model, as well as previous methods.

The novelty of our paper is that we suggested a method of ridge fuzzy regression which takes into account the centering and scaling issues frequently found in ridge-type regressions. For both the ridge fuzzy regression and the fuzzy multiple linear regression, the estimated fuzzy coefficients can be reversely estimated. That is, the signs of $\widehat{l_{A_k}}(\alpha)$ and $\widehat{r_{A_k}}(\alpha)$ may be switched. This is a phenomenon common among variations of the fuzzy linear regression model, not just our proposed methodology. Our model successfully handles this situation by modifying the estimated coefficients to form a pre-defined shape of the membership function. Most importantly, our model performance is comparable with, or in some cases even better than, that of existing models.

In our future studies, we plan to apply the proposed ridge fuzzy regression model to correlated genetic datasets. In addition, we will construct the algorithm for the ridge fuzzy regression model based on other distances between fuzzy numbers.

References

1. Tanaka, H., Uejima, S., Asai, K.: Linear regression analysis with fuzzy model. IEEE Trans. Syst. Man Cybernet. **12**(6), 903–907 (1982)
2. Kim, B., Bishu, R.R.: Evaluation of fuzzy linear regression models by comparison membership function. Fuzzy Sets Syst. **100**, 343–352 (1998)
3. Tanaka, H., Lee, H.: Interval regression analysis by quadratic programming approach. IEEE Trans. Fuzzy Syst. **6**(4), 473–481 (1998)
4. Jung, H.Y., Yoon, J.H., Choi, S.H.: Fuzzy linear regression using rank transform method. Fuzzy Sets Syst. **274**, 97–108 (2015)
5. Kim, I.K., Lee, W.J., Yoon, J.H., Choi, S., Chio, H.: Fuzzy regression model using trapezoidal fuzzy numbers for re-auction data. Int. J. Fuzzy Log. Intell. Syst. **16**(1), 72–80 (2016)
6. Buckley, J.J., Feuring, T.: Linear and non-linear fuzzy regression: evolutionary algorithm solutions. Fuzzy Sets Syst. **112**, 381–394 (2000)
7. Buckley, J.J., Qu, Y.: On using α-cuts to evaluate fuzzy functions. Fuzzy Sets Syst. **38**, 309–312 (1990)
8. Lee, W.J., Jung, H.Y., Yoon, J.H., Choi, S.H.: Analysis of variance for fuzzy data based on permutation method. Int. J. Fuzzy Log. Intell. Syst. **17**(1), 43–50 (2017)
9. Ming, M., Friedman, M., Kandel, A.: General fuzzy least squares. Fuzzy Sets Syst. **88**, 107–118 (1997)
10. Draper, N.R., Smith, H.: Applied Regression Analysis. Wiley, Hoboken (1981)
11. Diamond, P.: Fuzzy least squares. Inf. Sci. **46**, 141–157 (1988)
12. Hao, P.Y., Chiang, J.H.: A fuzzy model of support vector regression machine. Int. J. Fuzzy Syst. **9**(1), 45–50 (2007)
13. Giachetti, R.E., Young, R.E.: A parametric representation of fuzzy numbers and their arithmetic operators. Fuzzy Sets Syst. **91**(2), 185–202 (1997)
14. Yager, R.R.: Using trapezoids for representing granular objects: applications to learning and OWA aggregation. Inf. Sci. **178**(2), 363–380 (2008)
15. Choi, S.H., Jung, H.Y., Lee, W.J., Yoon, J.H.: Fuzzy regression model with monotonic response function. Commun. Kor. Math. Soc. **33**(3), 973–983 (2018)
16. Choi, S.H., Yoon, J.H.: General fuzzy regression using least squares method. Int. J. Syst. Sci. **41**, 477–485 (2010)
17. Choi, S.H., Buckley, J.J.: Fuzzy regression using least absolute deviation estimators. Soft. Comput. **12**, 257–263 (2008)
18. Lee, W.J., Jung, H.Y., Yoon, J.H., Choi, S.H.: The statistical inferences of fuzzy regression based on bootstrap techniques. Soft. Comput. **19**, 883–890 (2015)
19. Lee, W.J., Jung, H.Y., Yoon, J.H., Chio, S.H.: A novel forecasting method based on F-transform and fuzzy time series. Int. J. Fuzzy Syst. **19**(6), 1793–1802 (2017)
20. Zhang, X., Omachi, S., Aso, H.: Fuzzy regression analysis using RFLN and its application. Proc. FUZZ-IEEE'97 **1**, 51–56 (1997)
21. Hoerl, A.E., Kennard, R.W.: Ridge regression: biased estimation for nonorthogonal problems. Technometrics **12**, 55–67 (1970)
22. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
23. Zadeh, L.A.: Probability measures of fuzzy events. J. Math. Anal. Appl. **23**(2), 421–427 (1968)
24. Zadeh, L.A.: The concept of linguistic variable and its application to approximate reasoning, I, II and III. Inf. Sci. **8**, 199–249 (1975). **(8, pp. 301–357, and 9, pp. 43–80)**
25. Dubois, D., Prade, H.: Fuzzy Sets and Systems Theory and Applications. Academic Press, New York (1980)
26. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, Berlin (2001)
27. Hong, D.H., Hwang, C., Ahn, C.: Ridge estimation for regression models with crisp inputs and Gaussian fuzzy output. Fuzzy Sets Syst. **142**(2), 307–331 (2004)

Seung Hoe Choi obtained his Ph.D. degree in Mathematical Statistics from Yonsei University Korea in 1994. Since 1996, he is currently full professor of Korea Aerospace University. His main research interests are mathematical prediction method using the soft computing and statistical prediction in sports like soccer, baseball and basketball.

Hye-Young Jung is a teaching professor at Seoul National University, South Korea. She received the Ph.D. degree in mathematics (major: statistics) from Yonsei University, South Korea, in 2014 and performed postdoctoral research in statistics at Seoul National University from 2014 to 2016. Her research interest lies in fuzzy theory, rough theory, belief theory, bioinformatics, statistical modeling, and time-series analysis.

Hyoshin Kim received the B.Ec. degree in Statistics from Sungkyunkwan University, Seoul, South Korea, in 2017, and the M.S. degree in Statistics from Seoul National University, Seoul, South Korea, in 2019. Her research interests include spatial statistics, bayesian analysis, and big data.