An understandable way to discover methods to model interval input–output samples

Shan Xu[1] · **Shenggang Li**[1] · **Heng Liu**[2] · **Harish Garg**[3] · **Xuqin Jin**[1] · **Jingjing Zhao**[4]
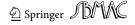
Abstract

This paper shows an application of plausible reasoning methods (mainly, specialization and analogous) in mathematical modeling. Our attention is how a practitioner to determine analogously a more balanced scientific model to assist the desire during solving the entire problem. Taking interval samples modeling as a problem, we exemplify (with consideration paid to the motivation and course of discovering) how to discover, based on the classical corresponding methods, three linear regression models and two linear-like interpolation models relying on n-variable interval input-1-variable interval output samples. The rationality of these recommended models are proved, and applications of them are illuminated in detail by examples. Strategies to model further interval samples towards satisfactory are also exposed.

Keywords Plausible reasoning · Interval number · Interval input–output sample modeling · Linear regression · Linear-like interpolation

Mathematics Subject Classification 00A08 · 00A30 · 00A35 · 54A40

---

[1] We also make no difference between an interval number $\mathbb{a}$ and a closed interval $[\dot{a}, \ddot{a}]$ of R for convenience; notice that $\dot{a} = p_1(\mathbb{a})$ and $\ddot{a} = p_2(\mathbb{a})$, where p_1 (resp., p_2) is the first (resp., the second) projection from R^2 to R. We will also write $\mathbb{a} = \langle \dot{a}, \ddot{a} \rangle \prec \langle \dot{b}, \ddot{b} \rangle = \mathbb{b}$ if $\dot{a} < \dot{b}$ and $\ddot{a} < \ddot{b}$.

✉ Shenggang Li
  shengganglinew@126.com

✉ Harish Garg
  harishg58iitr@gmail.com; harish.garg@thapar.edu

[1] College of Mathematics and Statistics Science, Shaanxi Normal University, Xi'an 710062, China

[2] School of Science, Guangxi University of Nationalities, Nanning 530006, China

[3] School of Mathematics, Thapar Institute of Engineering and Technology, Deemed University, Patiala, Punjab 147004, India

[4] School of Psychology, Shaanxi Normal University, Xi'an 710062, China

1 Introduction and preliminaries

An interval number (Moore 1966) $\mathbb{a}$ is a pair $\langle \dot{a}, \ddot{a} \rangle \in R^2$ with $\dot{a} \leq \ddot{a}$, where R is the set of all real numbers with the ordinary partial order. In this paper, we write $\mathbb{a} = \langle a, a \rangle$ as $a^!$ and identify[1] it with the corresponding real number a (therefore, interval numbers can be looked as an extension of real numbers) and designate the number $\langle \dot{a}, \ddot{a} \rangle^c = 0.5(\dot{a} + \ddot{a})$ (resp., $\langle \dot{a}, \ddot{a} \rangle^r = 0.5(\ddot{a} - \dot{a})$) the center or midpoint (resp., the radius) of the interval number $\langle \dot{a}, \ddot{a} \rangle$ (thus $\langle \dot{a}, \ddot{a} \rangle$ can also be written as $\lceil \langle \dot{a}, \ddot{a} \rangle^c, \langle \dot{a}, \ddot{a} \rangle^r \rceil$ or $\langle \dot{a}, \ddot{a} \rangle^c \pm \langle \dot{a}, \ddot{a} \rangle^r$). We use $\mathbb{R}$ to denote the set of all interval numbers. We define $a \vee b = \max\{a, b\}, a \wedge b = \min\{a, b\}$, and $\lfloor a, b \rfloor = \langle a \wedge b, a \vee b \rangle$ for any $a, b \in R$. The four elementary operations on R have also been extent to $\mathbb{R}$ (where $\mathbb{a} = \langle \dot{a}, \ddot{a} \rangle, \mathbb{b} = \langle \dot{b}, \ddot{b} \rangle \in \mathbb{R}$, see Moore 1966): $\langle \dot{a}, \ddot{a} \rangle \oplus \langle \dot{b}, \ddot{b} \rangle = \langle \dot{a} + \dot{b}, \ddot{a} + \ddot{b} \rangle$, $\langle \dot{a}, \ddot{a} \rangle \ominus \langle \dot{b}, \ddot{b} \rangle = \lfloor \dot{a} - \dot{b}, \ddot{a} - \ddot{b} \rfloor$, $\langle \dot{a}, \ddot{a} \rangle \odot \langle \dot{b}, \ddot{b} \rangle = \lfloor \dot{a}\dot{b}, \ddot{a}\ddot{b} \rfloor$, $\langle \dot{a}, \ddot{a} \rangle \oslash \langle \dot{b}, \ddot{b} \rangle = \lfloor \frac{\dot{a}}{\dot{b}}, \frac{\ddot{a}}{\ddot{b}} \rfloor$ $(0 \notin [\dot{b}, \ddot{b}])$. Moreover, a total order (see Hu and Wang 2006; Xu and Yager 2006) $\leq$ can be defined on $\mathbb{R}$ by putting

$$\langle \dot{a}, \ddot{a} \rangle \leq \langle \dot{b}, \ddot{b} \rangle \iff \langle \dot{a}, \ddot{a} \rangle^c < \langle \dot{b}, \ddot{b} \rangle^c, \text{ or } \langle \dot{a}, \ddot{a} \rangle^c = \langle \dot{b}, \ddot{b} \rangle^c \text{ but } \ddot{a} - \dot{a} \geq \ddot{b} - \dot{b}.$$

Interval data arise in many cases (such as numerical analysis—managing rounding errors, computer-assisted proofs, global optimization, individually, modeling uncertainty) because the data included there can not be exactly expressed in real numbers but can be revealed in interval numbers. For instance, consider the following problems with interval numbers as input and output (briefly, interval input–output) sample set [for more details, we refer to see also Table 4 in Inuiguchi and Mizoshita (2012)][2]:
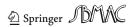
(I) The pattern recognition problem including interval samples. Astragali Radix is a medicinal and edible plant of the same origin that can regulate the body's immune function and is perfect for endangered patients. Usually, Astragali Radixes are distributed (based on test and measure data) into 5 grades: 1 (the lowest grade), . . ., 5 (the highest grade). The following Table 1 (taken from Zhang et al. 2020) dispenses some useful samples, where $\mathbb{x}_1$ stands for length (cm) of Astragali Radix, $\mathbb{x}_2$ stands for head diameter (cm) of Astragali Radix, $\mathbb{x}_3$ stands for tail diameter (cm) of Astragali Radix, and $\mathbb{y}$ stands for grade of Astragali Radix. For a given Astragali Radix $\mathbb{t}_0 = (35.8 \pm 5.4, 1.7 \pm 0.2, 0.8 \pm 0.1)$ (whose grade cannot be resolved instantly by using Pharmacopoeia of the People's Republic of China 1992), try to match it.

(II) A control problem involving interval samples:

$$\frac{d\boldsymbol{x}}{dt} = f(\boldsymbol{x}) + \boldsymbol{u}(t) \tag{1}$$

where $\boldsymbol{x}(t) = (x_1(t), x_2(t))^{\mathrm{T}} \in U$ (a compact set in R^2) represents the state at time t, $f(\boldsymbol{x})$ is a binary continuous function on U which is just observable (i.e., we can get the approximate value, an **interval number** in general, for each $\boldsymbol{x}$), and $\boldsymbol{u}$ is the controller to be designed.

Just like succeeding a decimal by an integer (so-called rounding up or down), practitioners can restore each interval number $\mathbb{a}$ in a sample set by its center $\mathbb{a}^c$ to process the new sample using some known processes. However, this is an unacceptable method due to the loss of the information. The other way to deal with it while solving the practical problem is the use of a real input-interval output model (i.e. real number input-interval

---

[2] The interval input–output data sets are also often used in real world applications. For examples, the measurements obtained by same/similar measuring instruments, health parameters of different patients in statistical databases, etc.

Table 1 Sample information of Astragali Radix

x_1	x_2	x_3	y
44.5 ± 5.9	1.9 ± 0.2	1.4 ± 0.1	5 ± 0.1
52.2 ± 6.6	1.7 ± 0.2	1.2 ± 0.1	4 ± 0.1
58.5 ± 7.1	1.4 ± 0.1	1.0 ± 0.2	3 ± 0.1
42.6 ± 9.6	1.2 ± 0.2	0.8 ± 0.1	2 ± 0.1
44.5 ± 9.2	1.0 ± 0.1	0.7 ± 0.1	1 ± 0.1
42.9 ± 9.6	2.2 ± 0.3	1.6 ± 0.3	5 ± 0.1
36.7 ± 10.4	1.7 ± 0.2	1.2 ± 0.2	4 ± 0.1
38.2 ± 7.4	1.5 ± 0.2	1.1 ± 0.1	3 ± 0.1
40.9 ± 8.0	1.3 ± 0.1	0.9 ± 0.2	2 ± 0.1
43.4 ± 9.2	1.1 ± 0.2	0.7 ± 0.1	1 ± 0.1
31.5 ± 5.5	1.7 ± 0.1	1.3 ± 0.2	5 ± 0.1
25.5 ± 2.9	1.4 ± 0.2	1.1 ± 0.1	4 ± 0.1
27.7 ± 4.6	1.2 ± 0.1	0.9 ± 0.1	3 ± 0.1
28.7 ± 5.3	1.1 ± 0.1	0.8 ± 0.1	2 ± 0.1
25.6 ± 4.9	0.9 ± 0.1	0.7 ± 0.1	1 ± 0.1
32.3 ± 4.9	1.9 ± 0.2	1.4 ± 0.1	5 ± 0.1
31.2 ± 4.9	1.5 ± 0.1	1.2 ± 0.1	4 ± 0.1
29.7 ± 4.3	1.4 ± 0.1	1.1 ± 0.1	3 ± 0.1
32.4 ± 4.6	1.2 ± 0.1	0.9 ± 0.1	2 ± 0.1
29.2 ± 3.8	1.0 ± 0.1	0.8 ± 0.1	1 ± 0.1
36.5 ± 9.0	1.3 ± 0.2	0.9 ± 0.1	4 ± 0.1
32.4 ± 5.5	1.0 ± 0.1	0.8 ± 0.1	3 ± 0.1
30.4 ± 5.4	0.9 ± 0.1	0.7 ± 0.1	2 ± 0.1
32.9 ± 6.1	0.8 ± 0.1	0.6 ± 0.1	1 ± 0.1

number output, called also crisp input-interval output); for more details, see (Hwang et al. 2006; Ishibuchi and Tanaka 1990; Jeng et al. 2003; Lee and Tanaka 1999) and the following Example 1.1 and Remark 1.2.

Example 1.1 Consider a real input–output (i.e. real number input-real number output) sample (an n-element set) $S = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ from a continuous function $f(x)$. Without loss of generality we assume $S = \{(-1, 0), (0, 1), (1, 0)\}$ (a 3-element set). Firstly, we get the linear interpolation $f_{1,3,0}(x)$ of f based on S

$$f_{1,3,0}(x) = \begin{cases} 1 + x, & x \in [-1, 0), \\ 1 - x, & x \in [0, 1]. \end{cases}$$

Secondly, we get the first real input-interval output model $f_{1,3,\varepsilon}(x) = \langle f_{1,3,\varepsilon}^-(x), f_{1,3,\varepsilon}^+(x) \rangle = f_{1,3,0}(x) \pm \varepsilon$; clearly, $f_{1,3,\varepsilon}^-(x) = f_{1,3,0}(x) - \varepsilon$ and $f_{1,3,\varepsilon}^+(x) = f_{1,3,0}(x) + \varepsilon$ are continuous. Similarly, we get $f_{1,n,0}(x)$ and $f_{1,n,\varepsilon}(x) = \langle f_{1,n,\varepsilon}^-(x), f_{1,n,\varepsilon}^+(x) \rangle = f_{1,n,0}(x) \pm \varepsilon$ $(n > 3, \varepsilon > 0)$. Notice $f_{1,n,0}(x_i) \in f_{1,n,\varepsilon}(x_i)$ $(i = 1, 2, \ldots, n)$, we have $\lim_{\varepsilon \to 0} f_{1,n,\varepsilon}(x_i) = \langle f_{1,n,0}(x_i), f_{1,n,0}(x_i) \rangle = f_{1,n,0}(x_i)$ $(i = 1, 2, \ldots, n)$, and $\lim_{n \to +\infty} f_{1,n,\varepsilon}(x) = \langle f_{1,n,0}(x), f_{1,n,0}(x) \rangle = f_{1,n,0}(x)$ $(x \in [a, b])$ if $a = x_1 < x_2 < \cdots x_n = b$ is an equal-length partition. Thirdly, we get the second real input-interval output

model $f_{2,3,\varepsilon}(x) = \langle f_{2,3,\varepsilon}^-(x), f_{2,3,\varepsilon}^+(x) \rangle (\varepsilon > 0)$, where

$$
f_{2,3,\varepsilon}^-(x) = \begin{cases} 1 + x - \varepsilon, & x \in [-1, -0.5\varepsilon), \\ \sqrt{0.5\varepsilon^2 - x^2} + 1 - 2\varepsilon, & x \in [-0.5\varepsilon, 0.5\varepsilon), \\ 1 - x - \varepsilon, & x \in [0.5\varepsilon, 1] \end{cases}
$$

and

$$
f_{2,3,\varepsilon}^+(x) = \begin{cases} 1 + x + \varepsilon, & x \in [-1, -0.5\varepsilon), \\ \sqrt{0.5\varepsilon^2 - x^2} + 1, & x \in [-0.5\varepsilon, 0.5\varepsilon), \\ 1 - x + \varepsilon, & x \in [0.5\varepsilon, 1] \end{cases}
$$

are apparently differentiable. Similarly, we get $f_{2,n,\varepsilon}(x) = \langle f_{2,n,\varepsilon}^-(x), f_{2,n,\varepsilon}^+(x) \rangle$ $(n > 3)$. Notice $f_{1,n,0}(x_i) \in f_{2,n,\varepsilon}(x_i)$ $(i = 1, 2, \ldots, n)$, $\lim_{\varepsilon \to 0} f_{2,n,\varepsilon}(x_i) = f_{2,n,0}(x_i)$ $(i = 1, 2, \ldots, n)$, and $\lim_{n \to +\infty} f_{2,n,\varepsilon}(x) = f_{2,n,0}(x)$ $(x \in [a, b])$ if $a = x_1 < x_2 < \cdots x_n = b$ is an equal-length partition.

Remark 1.2 An interval linear regression model for a real input–output sample can also be written as $Y(\boldsymbol{x}) = A_0 \oplus (A_1 \otimes x_1) \oplus (A_2 \otimes x_2) \oplus \cdots \oplus (A_n \otimes x_n) = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{x} = (1, x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ is a real input column vector, $Y(\boldsymbol{x})$ is the corresponding estimated interval, and $\boldsymbol{A} = (A_0, A_1, A_2, \ldots, A_n)$ is an interval coefficient row vector with $A_i = \lceil a_i, c_i \rceil$ $(i = 0, 1, 2, \ldots, n)$, the optimal interval coefficients. Moreover, the optimal interval coefficients is a solution of the following quadratic programming problem [see (Chukhrova and Johannssen 2019; Tanaka and Lee 1998) and the references here in]

$$
\min_{\boldsymbol{a}, \boldsymbol{c}} J = \boldsymbol{c}^{\mathrm{T}} \left(\sum_{j=1}^{m} |\boldsymbol{x}_j|^{\mathrm{T}} |\boldsymbol{x}_j| \right) \boldsymbol{c} + \xi \boldsymbol{a}^{\mathrm{T}} \boldsymbol{a},
$$
$$
\text{subject to } \boldsymbol{a}^{\mathrm{T}} \boldsymbol{x}_j - \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}_j \leq y_j \leq \boldsymbol{a}^{\mathrm{T}} \boldsymbol{x}_j + \boldsymbol{c}^{\mathrm{T}} |\boldsymbol{x}_j|,
$$
$$
c_i \geq 0 \ (i = 0, 1, \ldots, n; \ j = 1, 2, \ldots, p),
$$

where $\boldsymbol{a} = (a_0, a_1, \ldots, a_n)^{\mathrm{T}}$, $\boldsymbol{c} = (c_0, c_1, \ldots, c_n)^{\mathrm{T}}$, $\boldsymbol{c}^{\mathrm{T}}$ is the transpose of $\boldsymbol{c}$, $(\boldsymbol{x}_j^{\mathrm{T}}, y_j) = (x_{j1}, x_{j2}, \ldots, x_{jn}; y_j)$ is j-th sample, $|\boldsymbol{x}_j|^{\mathrm{T}} = (|x_{j1}|, |x_{j2}|, \ldots, |x_{jn}|)$, and $\xi > 0$ (very small).

Thus, the main concern of this paper is how practitioners to model interval input–output samples. There are some available work in the literature related to the topic (we refer to Boukezzoula et al. 2011, 2018, 2020; Chuang 2008; Hladíka and Černý 2012); but we extend the method as presented by Hladíka and Černý (2012) because the readers probably see from it not only the motivation the authors propose their method but also the course the method is formed [see Polya's famous book (Polya 1954) for more in this direction]. The present paper is a sequel of these works which exemplifies, in the light of Polya's idea, how to restore or approximate the true function (in form of interval input–output function) from the interval input–output samples based on some commonly used methods. For problem (I), we first concentrate the original problem by assuming that each interval datum has a radius 0 and thus get a new sample set and find a solution: Determine the grade of $\mathbf{t}_0$ using the classical linear regression function obtained relying on the new sample set. Then we investigate the generalization of this strategy in the case of the interval sample. Analogously, for problem (II), we first use an interval input–output function (which can be formulated) to approximate f (by observing the analogy between the case of interval input–output sample and the case of the real input–output sample) and then consider the establishment of the new system.

To make our discussion more reasonable, let us recall three most usually used metrics on the n-dimension Euclidean space R^n $(n \geq 1)$ which are defined by $(\boldsymbol{a} = (a_1, \ldots, a_n), \boldsymbol{b} = (b_1, \ldots, a_n))$ $\underset{n}{\rho}(\boldsymbol{a}, \boldsymbol{b}) = \max\{|a_1 - b_1|, \ldots, |a_n - b_n|\}$ (called the Chebyshev metric), $\rho_n(\boldsymbol{a}, \boldsymbol{b}) = \sqrt{\sum_{i=1}^n |a_i - b_i|^2}$ (called the Euclidean metric), and $\bar{\rho}_n(\boldsymbol{a}, \boldsymbol{b}) = \sum_{i=1}^n |a_i - b_i|$ (called the Manhattan metric or the city block metric).

The rest of the paper consists of four sections. Section 2 exemplifies how to discover, based on the classical linear regression, methods to model interval input–output samples. Section 3 investigates, based on the classical linear interpolation, the same problem as in Sect. 2. Section 4 exemplifies applications of the models proposed. Conclusions, discussion, and strategies for further generalization are given in Sect. 5.

2 Three models stemmed from the classical linear regression

In this section, the following interval input–output sample set, consisting of $n+1$-dimension row vectors of interval numbers (each interval number contains the true datum), will be considered:

$$\mathbb{S} = (\mathbb{x}_{1,1}, \mathbb{x}_{2,1}, \ldots, \mathbb{x}_{n,1}; \mathbb{y}_1), (\mathbb{x}_{1,2}, \mathbb{x}_{2,2}, \ldots, \mathbb{x}_{n,2}; \mathbb{y}_2),$$
$$\ldots, (\mathbb{x}_{1,m}, \mathbb{x}_{2,m}, \ldots, \mathbb{x}_{n,m}; \mathbb{y}_m) \tag{2}$$

We first present three linear regression models, each is stemmed from the corresponding linear regression modal with real input–output sample. Then we prove that each modal is an universal approximator (Ying 2015) to the corresponding linear regression modal based on the true sample.

Remark 2.1 (1) If $m = n + 1 = 2$, $(\mathbb{x}_{1,1}, \mathbb{y}_1) = (x_1, y_1)$, and $(\mathbb{x}_{1,2}, \mathbb{y}_2) = (x_2, y_2)$ $(x_1 \neq x_2)$. Then the linear regression function based on $\mathbb{S}$ is a classical one: $f_c(x) = \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2} + \frac{y_1 - y_2}{x_1 - x_2} x = a_0 + a_1 x$.

(2) If $m = n + 1 = 2$, $(\mathbb{x}_{1,1}, \mathbb{y}_1) = (\lceil x_1, \delta \rceil, \lceil y_1, \delta \rceil)$, and $(\mathbb{x}_{1,2}, \mathbb{y}_2) = (\lceil x_2, \delta \rceil, \lceil y_2, \delta \rceil)$ $(x_1 \neq x_2, \delta > 0)$. It is most possible for us to get two functions: one is the classical linear regression function

$$\underline{f}(t) = \frac{(x_1 - \delta)(y_2 - \delta) - (x_2 - \delta)(y_1 - \delta)}{x_1 - x_2} + \frac{y_1 - y_2}{x_1 - x_2} t$$
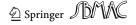
based on the real input–output sample set $\{(x_1 - \delta, y_1 - \delta), (x_2 - \delta, y_2 - \delta)\}$, another is the classical linear regression function

$$\bar{f}(t) = \frac{(x_1 + \delta)(y_2 + \delta) - (x_2 + \delta)(y_1 + \delta)}{x_1 - x_2} + \frac{y_1 - y_2}{x_1 - x_2} t$$

based on the real input–output sample set $\{(x_1 + \delta, y_1 + \delta), (x_2 + \delta, y_2 + \delta)\}$. It is also natural for us to take the linear regression function based on the above interval input–output sample set $\mathbb{S}$ as $\hat{f}_\delta(\mathbb{x}) = a_0 \oplus [a_1 \otimes \mathbb{x}]$ or $\check{f}_\delta(\mathbb{x}) = \lfloor \underline{f}(\dot{x}), \bar{f}(\ddot{x}) \rfloor$, where

$$a_0 = \left\lfloor \frac{(x_1 - \delta)(y_2 - \delta) - (x_2 - \delta)(y_1 - \delta)}{x_1 - x_2}, \frac{(x_1 + \delta)(y_2 + \delta) - (x_2 + \delta)(y_1 + \delta)}{x_1 - x_2} \right\rfloor,$$

$a_1 = \frac{y_1 - y_2}{x_1 - x_2}$, and $\mathbb{x} = \langle \dot{x}, \ddot{x} \rangle \in \mathbb{R}$. It can be easily seen that $\lim_{\delta \to 0} \hat{f}_\delta(\mathbb{x}) = \lim_{\delta \to 0} \check{f}_\delta(\mathbb{x}) = f(x)$ $(\forall \mathbb{x} \in \mathbb{R})$. This confirms the rationality that we take the linear regression function based on the above interval input–output sample set $\mathbb{S}$ as $\hat{f}_\delta$ or $\check{f}_\delta$.

Springer

(3) Notice that the radii of interval numbers in a data set are not the same in general in practice problems. So we should consider a little big generalization of (2): $m = n + 1 = 2$, $(\mathbb{x}_{1,1}, \mathbb{y}_1) = (\lceil x_1, \delta_{1,1} \rceil, \lceil y_1, \delta_{1,2} \rceil)$, and $(\mathbb{x}_{1,2}, \mathbb{y}_2) = (\lceil x_2, \delta_{2,1} \rceil, \lceil y_2, \delta_{2,2} \rceil)$ $(x_1 \neq x_2, \delta_{i,j} > 0, i, j = 1, 2)$. Analogously to (2), we can get the linear regression function $\hat{f}_{\boldsymbol{\delta}}(\mathbb{x}) = \langle \dot{a}_0, \ddot{a}_0 \rangle \oplus \langle \dot{a}_1, \ddot{a}_1 \rangle \mathbb{x}$ or $\check{f}_{\boldsymbol{\delta}}(\mathbb{x}) = \; \lfloor \underline{f}(\dot{x}), \overline{f}(\ddot{x}) \rfloor \; (\mathbb{x} = \langle \dot{x}, \ddot{x} \rangle \in \mathbb{R})$ based on this interval input–output sample set $\mathbb{S}$, where

$$\dot{a}_0 = \frac{(x_1 - \delta_{1,1})(y_2 - \delta_{2,2}) - (x_2 - \delta_{2,1})(y_1 - \delta_{1,2})}{(x_1 - x_2) + (\delta_{1,1} - \delta_{2,1})} \wedge \frac{(x_1 + \delta_{1,1})(y_2 + \delta_{2,2}) - (x_2 + \delta_{2,1})(y_1 + \delta_{1,2})}{(x_1 - x_2) + (\delta_{1,1} - \delta_{2,1})},$$

$$\ddot{a}_0 = \frac{(x_1 - \delta_{1,1})(y_2 - \delta_{2,2}) - (x_2 - \delta_{2,1})(y_1 - \delta_{1,2})}{(x_1 - x_2) + (\delta_{1,1} - \delta_{2,1})} \vee \frac{(x_1 + \delta_{1,1})(y_2 + \delta_{2,2}) - (x_2 + \delta_{2,1})(y_1 + \delta_{1,2})}{(x_1 - x_2) + (\delta_{1,1} - \delta_{2,1})},$$

$$\dot{a}_1 = \frac{(y_1 - y_2) + (\delta_{2,2} - \delta_{1,2})}{(x_1 - x_2) + (\delta_{2,1} - \delta_{1,1})} \wedge \frac{(y_1 - y_2) + (\delta_{1,2} - \delta_{2,2})}{(x_1 - x_2) + (\delta_{1,1} - \delta_{2,1})},$$

$$\ddot{a}_1 = \frac{(y_1 - y_2) + (\delta_{2,2} - \delta_{1,2})}{(x_1 - x_2) + (\delta_{2,1} - \delta_{1,1})} \vee \frac{(y_1 - y_2) + (\delta_{1,2} - \delta_{2,2})}{(x_1 - x_2) + (\delta_{1,1} - \delta_{2,1})},$$

$$\underline{f}(t) = \frac{(x_1 - \delta_{1,1})(y_2 - \delta_{2,2}) - (x_2 - \delta_{2,1})(y_1 - \delta_{1,2})}{(x_1 - x_2) + (\delta_{2,1} - \delta_{1,1})} + \frac{(y_1 - y_2) + (\delta_{2,2} - \delta_{1,2})}{(x_1 - x_2) + (\delta_{2,1} - \delta_{1,1})} t$$

is the classical linear regression function based on the real input–output sample set $\{(x_1 - \delta_{1,1}, y_1 - \delta_{1,2}), (x_2 - \delta_{2,1}, y_2 - \delta_{2,2})\}$, and

$$\overline{f}(t) = \frac{(x_1 + \delta_{1,1})(y_2 + \delta_{2,2}) - (x_2 + \delta_{2,1})(y_1 + \delta_{1,2})}{(x_1 - x_2) + (\delta_{1,1} - \delta_{2,1})} + \frac{(y_1 - y_2) + (\delta_{1,2} - \delta_{2,2})}{(x_1 - x_2) + (\delta_{1,1} - \delta_{2,1})} t$$

is the classical linear regression function based on the real input–output sample set $\{(x_1 + \delta_{1,1}, y_1 + \delta_{1,2}), (x_2 + \delta_{2,1}, y_2 + \delta_{2,2})\}$. In addition, both $\hat{f}_{\boldsymbol{\delta}}(\mathbb{x}) \longrightarrow f(x)$ and $\check{f}_{\boldsymbol{\delta}}(\mathbb{x}) \longrightarrow f(x)$ hold for all $\mathbb{x} \in \mathbb{R}$ as $\boldsymbol{\delta} = (\delta_{1,1}, \delta_{1,2}, \delta_{2,1}, \delta_{2,2}) \longrightarrow (0, 0, 0, 0)$. This further support us to take the linear regression function based on the above interval input–output sample set $\mathbb{S}$ as $\hat{f}_{\boldsymbol{\delta}}$ or $\check{f}_{\boldsymbol{\delta}}$.

(4) If $\delta_{1,1} = \delta_{1,2} = \delta_{2,1} = \delta_{2,2}$ does not hold, then $a_1 = \frac{y_1 - y_2}{x_1 - x_2}$ does not hold (which can be seen from (1)–(3)) and the output should be $f_c(x) + r$, where r should be a continuous function of variables $\delta_{1,1}, \delta_{1,2}, \delta_{2,1}$, and $\delta_{2,2}$ satisfying $r(0, 0, 0, 0) = 0$. Thus we can take r as the second kind linear regression function (i.e. the linear regression function having no the constant term and obtaining by the classical least square estimation method) $r(\delta) = \frac{\delta_{1,1}\delta_{1,2} + \delta_{2,1}\delta_{2,2}}{\delta_{1,1}^2 + \delta_{2,1}^2} \delta$ based on the real input–output sample set $\{(\delta_{1,1}, \delta_{1,2}), (\delta_{2,1}, \delta_{2,2})\}$. Therefore, we can also take the linear regression function based on the interval input–output sample set $\mathbb{S}$ as $\tilde{f}_{\boldsymbol{\delta}}(\mathbb{x}) = \lceil f_c(\mathbb{x}^c), r(\mathbb{x}^r) \rceil = f_c(\mathbb{x}^c) \pm r(\mathbb{x}^r)$, where f_c is the linear regression function based on $\{(\mathbb{x}_{1,1}^c, \mathbb{y}_1^c), (\mathbb{x}_{1,2}^c, \mathbb{y}_2^c)\}$.

Example 2.2 Consider the interval input–output sample set $\mathbb{S} = \{(\mathbb{x}_{1,i}, \mathbb{x}_{2,i}, \mathbb{y}_i) \mid i = 1, 2, \ldots, 10\} = \{(\lceil x_{1,i}, \delta_{1,i} \rceil, \lceil x_{2,i}, \delta_{2,i} \rceil, \lceil y_i, r_i \rceil,) \mid i = 1, 2, \ldots, 10\}$ in Table 2.

(1) Similar to Remark 2.1(2) and Remark 2.1(3), we have $\underline{f}(u, v) = \underline{a}_0 + \underline{a}_1 u + \underline{a}_2 v = 95.9639 + 0.02u - 7.3671v$, $\overline{f}(u, v) = \bar{a}_0 + \bar{a}_1 u + \bar{a}_2 v = 126.6484 + 0.0088u - 6.9587v$, and thus we obtain the first model $\hat{f}(\mathbf{x}) = \hat{f}(\mathbb{x}_1, \mathbb{x}_2) = \mathbb{a}_0 \oplus [\mathbb{a}_1 \otimes \mathbb{x}_1] \oplus [\mathbb{a}_2 \otimes \mathbb{x}_2] = \langle 95.9639, 126.6484 \rangle \oplus [\langle 0.0088, 0.02 \rangle \otimes \mathbb{x}_1] \ominus [\langle 6.9587, 7.3671 \rangle \otimes \mathbb{x}_2]$ and the second model $\check{f}(\mathbf{x}) = \check{f}(\mathbb{x}_1, \mathbb{x}_2) = \lfloor \underline{f}(\dot{x}_1, \dot{x}_2), \overline{f}(\ddot{x}_1, \ddot{x}_2) \rfloor = \lfloor 95.9639 + 0.02\dot{x}_1 - 7.3671\dot{x}_2, 126.6484 + 0.0088\ddot{x}_1 - 6.9587\ddot{x}_2 \rfloor$. Similar to Remark 2.1(4), we have $f_c(\mathbb{x}_1^c, \mathbb{x}_2^c) = f_c(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 = 111.6918 + 0.0143x_1 - 7.1882x_2$ (i.e. the classical linear regression function or the first kind linear regression function based on the sample set $\{(\mathbb{x}_1^c, \mathbb{x}_2^c, \mathbb{y}^c) \mid (\mathbb{x}_1, \mathbb{x}_2, \mathbb{y}) \in \mathbb{S}\}$). We chose, by the classical least square estimation method, a linear regression function

Table 2 The interval input–output data

	$[x_1, \delta_1]$	$[x_2, \delta_2]$	$[y, r]$
1	1000	$[5, 1]$	$[100, 20]$
2	600	$[7, 1]$	75
3	$[1200, 20]$	$[6, 1]$	80
4	500	$[6, 1]$	70
5	$[300, 10]$	$[8, 1]$	$[50, 10]$
6	$[400, 10]$	$[7, 1]$	$[65, 10]$
7	1300	$[5, 1]$	90
8	1100	$[4, 1]$	100
9	1300	$[3, 1]$	110
10	$[300, 10]$	$[9, 1]$	$[60, 10]$

$r(\delta_1, \delta_2) = 0.\dot{1}\delta_1 + 4.\dot{4}\delta_2$ from the set of all 2-variable linear functions without constant terms (i.e. the second kind linear regression function)[3] As a result, we obtain the third model $\tilde{f}(\mathbf{x}) = \tilde{f}([x_1, \delta_1], [x_2, \delta_2]) = [111.6918 + 0.0143x_1 - 7.1882x_2, 0.\dot{1}\delta_1 + 4.\dot{4}\delta_2] = (111.6918 + 0.0143x_1 - 7.1882x_2) \pm (0.\dot{1}\delta_1 + 4.\dot{4}\delta_2)$.

(2) Assume $s_0 = \big(1000, 600, 1200, 500, 300, 400, 1300, 1100, 1300, 300; 5, 7, 6, 6, 8, 7, 5, 4, 3, 9; 100, 75, 80, 70, 50, 65, 90, 100, 110, 60\big) \in \mathbb{x}_{1,1} \times \mathbb{x}_{1,2} \times \cdots \times \mathbb{x}_{1,10} \times \mathbb{x}_{2,1} \times \mathbb{x}_{2,2} \times \cdots \times \mathbb{x}_{2,10} \times \mathbb{y}_1 \times \mathbb{y}_2 \times \cdots \times \mathbb{y}_{10}$ is the true data. Then $f(\mathbf{x}) = f(x_1, x_2) = 111.6918 + 0.0143x_1 - 7.1882x_2$ is the linear regression function based on the real input-real output data set s_0. For each $\mathbf{s} = \big(1000 + a, 600 + a, 1200 + a, 500 + a, 300 + a, 400 + a, 1300 + a, 1100 + a, 1300 + a, 300 + a; 5 + b, 7 + b, 6 + b, 6 + b, 8 + b, 7 + b, 5 + b, 4 + b, 3 + b, 9 + b; 100 + c, 75 + c, 80 + c, 70 + c, 50 + c, 65 + c, 90 + c, 100 + c, 110 + c, 60 + c\big) \in \mathbb{x}_{1,1} \times \mathbb{x}_{1,2} \times \cdots \times \mathbb{x}_{1,10} \times \mathbb{x}_{2,1} \times \mathbb{x}_{2,2} \times \cdots \times \mathbb{x}_{2,10} \times \mathbb{y}_1 \times \mathbb{y}_2 \times \cdots \times \mathbb{y}_{10}$, let $f_{abc}(\mathbf{x}) = f_{abc}(x_1, x_2) = \beta_a + \beta_b x_1 + \beta_c x_2$ be the linear regression function based on the real input–output data set $\mathbf{s}$, $\rho_3^*(f, f_{abc}) = \rho_3((111.6918, 0.0143, 7.1882), (\beta_a, \beta_b, \beta_c))$, $\rho_3^*(f, f_{abc}) = \rho_3((111.6918, 0.0143, 7.1882), (\beta_a, \beta_b, \beta_c))$, and $\bar{\rho}_3^*(f, f_{abc}) = \bar{\rho}_3((111.6918, 0.0143, 7.1882), (\beta_a, \beta_b, \beta_c))$. Then we have the computing results as shown in Table 3.

Supported by Remark 2.1 and Example 2.2, we have reasons to propose the following Algorithm 2.3:

---

3 Similar to obtaining a linear regression function $\tilde{r}(\mathbf{x}) = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + \tilde{\beta}_2 x_2 + \cdots + \tilde{\beta}_n x_n$ (by the classical least square estimation method) from the set of all n-variable linear functions, we can also obtain a linear regression function $r(\mathbf{x}) = r(x_1, x_2, \ldots, x_n) = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$ from the set of all n-variable linear functions without the constant term by the classical least square estimation method, i.e. by solving the linear system of linear equations $\frac{\partial e}{\partial \beta_1} =$

$\frac{\partial e}{\partial \beta_2} = \cdots \frac{\partial e}{\partial \beta_n} = 0$ or $\begin{cases} (\mathbf{x}_1 \mathbf{x}_1^\mathsf{T})\beta_1 + (\mathbf{x}_1 \mathbf{x}_2^\mathsf{T})\beta_2 + \cdots + (\mathbf{x}_1 \mathbf{x}_n^\mathsf{T})\beta_n = \mathbf{x}_1 \mathbf{y}^\mathsf{T}, \\ (\mathbf{x}_2 \mathbf{x}_1^\mathsf{T})\beta_1 + (\mathbf{x}_2 \mathbf{x}_2^\mathsf{T})\beta_2 + \cdots + (\mathbf{x}_2 \mathbf{x}_n^\mathsf{T})\beta_n = \mathbf{x}_2 \mathbf{y}^\mathsf{T}, \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ (\mathbf{x}_n \mathbf{x}_1^\mathsf{T})\beta_1 + (\mathbf{x}_n \mathbf{x}_2^\mathsf{T})\beta_2 + \cdots + (\mathbf{x}_n \mathbf{x}_n^\mathsf{T})\beta_n = \mathbf{x}_n \mathbf{y}^\mathsf{T}, \end{cases}$ where $e(\beta_1, \beta_2, \ldots, \beta_n) =$

$\sum_{i=1}^m [y_i - r(x_{1,i}, x_{2,i}, \ldots, x_{n,i})]^2$, and $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \mathbf{y}\} = \{(x_{1,1}, \ldots, x_{1,m}), (x_{2,1}, \ldots, x_{2,m}), \ldots, (x_{n,1}, \ldots, x_{n,m}), (y_1, \ldots, y_m)\}$ is the real input-real output data set (having $n + 1$ element and) consisting of m-dimension row vectors of real numbers. General speaking, $\beta_1 \neq \tilde{\beta}_1, \beta_2 \neq \tilde{\beta}_2, \ldots, \beta_n \neq \tilde{\beta}_n$. It can be easily seen from above linear system of linear equations that $\beta_1 = \frac{(\mathbf{x}_1 \mathbf{y}^\mathsf{T})\|\mathbf{x}_2\|^2 - (\mathbf{x}_2 \mathbf{y}^\mathsf{T})(\mathbf{x}_1 \mathbf{x}_2^\mathsf{T})}{\|\mathbf{x}_1\|^2 \|\mathbf{x}_2\|^2 - 2(\mathbf{x}_1 \mathbf{x}_2^\mathsf{T})}$ and

$\beta_2 = \frac{\|\mathbf{x}_1\|^2(\mathbf{x}_2 \mathbf{y}^\mathsf{T}) - (\mathbf{x}_1 \mathbf{y}^\mathsf{T})(\mathbf{x}_1 \mathbf{x}_2^\mathsf{T})}{\|\mathbf{x}_1\|^2 \|\mathbf{x}_2\|^2 - 2(\mathbf{x}_1 \mathbf{x}_2^\mathsf{T})}$ when $n = 2$.

Table 3 Comparison of distances between f and f_{abc} under three commonly used metrics

(a, b, c)	$f_{abc}(\boldsymbol{x})$	$\underline{\rho}_3^*(f, f_{abc})$	$\rho_3^*(f, f_{abc})$	$\bar{\rho}_3^*(f, f_{abc})$
$(0.033, 0.001, 0.0015)$	$111.7000 + 0.0143x_1 - 7.1882x_2$	0.0082	0.0082	0.0082
$-(0.033, 0.001, 0.0015)$	$111.6836 + 0.0143x_1 - 7.1882x_2$	0.0082	0.0082	0.0082
$(0.03, 0.001, 0.0015)$	$111.7001 + 0.0143x_1 - 7.1881x_2$	0.0083	0.0083	0.0083
$-(0.031, 0.001, 0.0015)$	$111.6836 + 0.0143x_1 - 7.1884x_2$	0.0082	0.0082	0.0082
$(0.029, 0.001, 0.0015)$	$111.7001 + 0.0143x_1 - 7.1881x_2$	0.0083	0.0083	0.0083
$-(0.03, 0.001, 0.0015)$	$111.6836 + 0.0143x_1 - 7.1884x_2$	0.0082	0.0082	0.0082
$(0.033, 0.00098, 0.0015)$	$111.6999 + 0.0143x_1 - 7.1881x_2$	0.0081	0.0081	0.0081
$-(0.033, 0.00099, 0.0015)$	$111.6837 + 0.0143x_1 - 7.1884x_2$	0.0081	0.0081	0.0081
$(0.033, 0.00097, 0.0015)$	$111.6998 + 0.0143x_1 - 7.1881x_2$	0.008	0.008	0.008
$-(0.033, 0.00097, 0.0015)$	$111.6838 + 0.0143x_1 - 7.1884x_2$	0.008	0.008	0.008
$(0.033, 0.001, 0.0012)$	$111.6997 + 0.0143x_1 - 7.1881x_2$	0.0079	0.0079	0.0079
$-(0.033, 0.001, 0.0013)$	$111.6838 + 0.0143x_1 - 7.1884x_2$	0.008	0.008	0.008
$(0.033, 0.001, 0.001)$	$111.6995 + 0.0143x_1 - 7.1881x_2$	0.0077	0.0077	0.0077
$-(0.033, 0.001, 0.0011)$	$111.6840 + 0.0143x_1 - 7.1884x_2$	0.0078	0.0078	0.0078

Algorithm 2.3 The linear regression functions based on the interval input–output sample set $\mathbb{S}$ in equality (2) can be given by the following ways:

(1) **Step 1** Compute the ordinary linear regression function $\underline{f}(t) = \underline{f}(t_1, \ldots, t_n) = \underline{a}_0 + \underline{a}_1 t_1 + \cdots + \underline{a}_n t_n$ $(\forall t \in R^n)$, based on the real input–output sample set $\underline{S} = \{(\dot{x}_{1,1}, \dot{x}_{2,1}, \ldots, \dot{x}_{n,1}; \dot{y}_1), (\dot{x}_{1,2}, \dot{x}_{2,2}, \ldots, \dot{x}_{n,2}; \dot{y}_2), \ldots, (\dot{x}_{1,m}, \dot{x}_{2,m}, \ldots, \dot{x}_{n,m}; \dot{y}_m)\}$, using the classical method.

Step 2 Compute the ordinary linear regression function $\bar{f}(t) = \bar{f}(t_1, \ldots, t_n) = \bar{a}_0 + \bar{a}_1 t_1 + \cdots + \bar{a}_n t_n$ $(\forall t \in R^n)$, based on the real input–output sample set $\bar{S} = \{(\ddot{x}_{1,1}, \ddot{x}_{2,1}, \ldots, \ddot{x}_{n,1}; \ddot{y}_1), (\ddot{x}_{1,2}, \ddot{x}_{2,2}, \ldots, \ddot{x}_{n,2}; \ddot{y}_2), \ldots, (\ddot{x}_{1,m}, \ddot{x}_{2,m}, \ldots, \ddot{x}_{n,m}; \ddot{y}_m)\}$, using the classical method.

Step 3 Obtain the first linear regression function $\hat{f}(\mathbf{x}) = \hat{f}(\mathbb{x}_1, \mathbb{x}_2, \ldots, \mathbb{x}_n) = \mathbb{a}_0 \oplus [\mathbb{a}_1 \otimes \mathbb{x}_1] \oplus [\mathbb{a}_2 \otimes \mathbb{x}_2] \oplus \cdots \oplus [\mathbb{a}_n \otimes \mathbb{x}_n]$ by computing $\mathbb{a}_0 = \lfloor \underline{a}_0, \bar{a}_0 \rfloor, \mathbb{a}_1 = \lfloor \underline{a}_1, \bar{a}_1 \rfloor$, $\ldots, \mathbb{a}_n = \lfloor \underline{a}_n, \bar{a}_n \rfloor$.

Step 4 Obtain the second linear regression function $\check{f}(\mathbf{x}) = \check{f}(\mathbb{x}_1, \ldots, \mathbb{x}_n) = \lfloor \underline{f}(\dot{x}_1, \ldots, \dot{x}_n), \bar{f}(\ddot{x}_1, \ldots, \ddot{x}_n) \rfloor$.

(2) **Step 1** Compute the ordinary linear regression function $f_c(t) = f_c(t_1, \ldots, t_n) = a_0 + a_1 t_1 + \cdots + a_n t_n$ $(\forall t \in R^n)$, based on the real input–output sample set

$$S^c = \left\{ \left(\frac{\ddot{x}_{1,1} + \dot{x}_{1,1}}{2}, \frac{\ddot{x}_{2,1} + \dot{x}_{2,1}}{2}, \ldots, \frac{\ddot{x}_{n,1} + \dot{x}_{n,1}}{2}; \frac{\ddot{y}_1 + \dot{y}_1}{2} \right), \left(\frac{\ddot{x}_{1,2} + \dot{x}_{1,2}}{2}, \frac{\ddot{x}_{2,2} + \dot{x}_{2,2}}{2}, \ldots, \frac{\ddot{x}_{n,2} + \dot{x}_{n,2}}{2}; \frac{\ddot{y}_2 + \dot{y}_2}{2} \right), \right.$$
$$\left. \ldots, \left(\frac{\ddot{x}_{1,m} + \dot{x}_{1,m}}{2}, \frac{\ddot{x}_{2,m} + \dot{x}_{2,m}}{2}, \ldots, \frac{\ddot{x}_{n,m} + \dot{x}_{n,m}}{2}; \frac{\ddot{y}_m + \dot{y}_m}{2} \right) \right\},$$

using the classical method.

Step 2 Compute the ordinary linear regression function $r(\boldsymbol{\delta}) = r(\delta_1, \ldots, \delta_n) = b_1 \delta_1 + \cdots + b_n \delta_n$ $(\forall \boldsymbol{\delta} \in R^n)$, based on the real input–output sample set

$$S^r = \left\{ \left(\frac{\ddot{x}_{1,1} - \dot{x}_{1,1}}{2}, \frac{\ddot{x}_{2,1} - \dot{x}_{2,1}}{2}, \ldots, \frac{\ddot{x}_{n,1} - \dot{x}_{n,1}}{2}; \frac{\ddot{y}_1 - \dot{y}_1}{2} \right), \left(\frac{\ddot{x}_{1,2} - \dot{x}_{1,2}}{2}, \frac{\ddot{x}_{2,2} - \dot{x}_{2,2}}{2}, \ldots, \frac{\ddot{x}_{n,2} - \dot{x}_{n,2}}{2}; \frac{\ddot{y}_2 - \dot{y}_2}{2} \right), \right.$$
$$\left. \ldots, \left(\frac{\ddot{x}_{1,m} - \dot{x}_{1,m}}{2}, \frac{\ddot{x}_{2,m} - \dot{x}_{2,m}}{2}, \ldots, \frac{\ddot{x}_{n,m} - \dot{x}_{n,m}}{2}; \frac{\ddot{y}_m - \dot{y}_m}{2} \right) \right\},$$

using the least-square method. Then $\tilde{f}(\mathbf{x}) = \tilde{f}(\mathbb{x}_1, \ldots, \mathbb{x}_n) = \lceil f_c(\mathbb{x}_1^c, \ldots, \mathbb{x}_n^c), r(\mathbb{x}_1^r, \ldots, \mathbb{x}_n^r) \rceil = f_c(\mathbb{x}_1^c, \mathbb{x}_2^c, \ldots, \mathbb{x}_n^c) \pm r(\mathbb{x}_1^r, \mathbb{x}_2^r, \ldots, \mathbb{x}_n^r)$ is the third linear regression function.

The rationality of Algorithm 2.3 is guaranteed by the following

Theorem 2.4 *Let* $\boldsymbol{s} = ((x_{1,1}, x_{2,1}, \ldots, x_{n,1}; y_1), (x_{1,2}, x_{2,2}, \ldots, x_{n,2}; y_2), \ldots, (x_{1,m}, x_{2,m}, \ldots, x_{n,m}; y_m))$ *be a real input–output datum from* $\boldsymbol{S} = (\mathbb{x}_{1,1} \times \mathbb{x}_{2,1} \times \cdots \times \mathbb{x}_{n,1} \times \mathbb{y}_1) \times (\mathbb{x}_{1,2} \times \mathbb{x}_{2,2} \times \cdots \times \mathbb{x}_{n,2} \times \mathbb{y}_2) \times \cdots \times (\mathbb{x}_{1,m} \times \mathbb{x}_{2,m} \times \cdots \times \mathbb{x}_{n,m} \times \mathbb{y}_m)$ *(which is a rearrangement of the sample set in equality (2)) and* $f_{\boldsymbol{s}}(\mathbf{x}) = \beta_0(\boldsymbol{s}) + \beta_1(\boldsymbol{s}) x_1 + \cdots + \beta_n(\boldsymbol{s}) x_n$ *the linear regression function based on the real input–output sample set* $\boldsymbol{s}$. *Then*

(1) $\beta_0(\boldsymbol{s}), \beta_1(\boldsymbol{s}), \ldots, \beta_n(\boldsymbol{s})$ *are continuous functions from* $\left(R^{2m(n+1)}, \rho \right)$ *to* (R, ρ_1) $(\rho \in \{\underline{\rho}_{2m(n+1)}, \rho_{2m(n+1)}, \bar{\rho}_{2m(n+1)}\})$.

(2) *The mapping* $g : \left(R^{m(n+1)}, \rho_{m(n+1)} \right) \longrightarrow \left(R^{n+1}, \rho \right)$ $(\rho \in \{\underline{\rho}_{n+1}, \rho_{n+1}, \bar{\rho}_{n+1}\})$, *defined by* $g(\boldsymbol{s}) = (\beta_0(\boldsymbol{s}), \beta_1(\boldsymbol{s}), \ldots, \beta_n(\boldsymbol{s}))$, *is continuous.*

(3) *Let* $f(\mathbf{x}) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n$ *be the linear regression function based on the true sample* $\boldsymbol{t} = ((x_{1,1}^0, x_{2,1}^0, \ldots, x_{n,1}^0; y_1^0), (x_{1,2}^0, x_{2,2}^0, \ldots, x_{n,2}^0; y_2^0), \ldots, (x_{1,m}^0, x_{2,m}^0, \ldots, x_{n,m}^0; y_m^0))$ *in* $\boldsymbol{S}$. *Then, for each* $\varepsilon > 0$, *there exists a* $\boldsymbol{\delta} = (\delta_0, \delta_1, \delta_2, \ldots, \delta_n)^T > \boldsymbol{0}$ *(i.e.* $\delta_0 > 0, \delta_1 > 0, \delta_2 > 0, \ldots, \delta_n > 0$) *and a* $\delta > 0$, *such that (where* $\rho \in \{\underline{\rho}_{n+1}, \rho_{n+1}, \bar{\rho}_{n+1}\}$*)*

i) $\rho\big(g(\boldsymbol{s}), (\beta_0, \beta_1, \dots, \beta_n)\big) < \varepsilon$ *if* $\boldsymbol{S}$ *is a* $\boldsymbol{\delta}$-*sample set, i.e. it satisfies* $\max\big\{\mathbb{y}_1^r, \mathbb{y}_2^r, \dots,$
$\mathbb{y}_m^r\big\} \leq \delta_0,\ \max\big\{\mathbb{x}_{1,1}^r, \mathbb{x}_{1,2}^r, \dots, \mathbb{x}_{1,m}^r\big\} \leq \delta_1,\ \max\big\{\mathbb{x}_{2,1}^r, \dots, \mathbb{x}_{2,m}^r\big\} \leq \delta_2,\ \dots,$
$\max\big\{\mathbb{x}_{n,1}^r, \mathbb{x}_{n,2}^r, \dots, \mathbb{x}_{n,m}^r\big\} \leq \delta_n.$

ii) $\rho\big(g(\boldsymbol{s}), (\beta_0, \beta_1, \dots, \beta_n)\big) < \varepsilon$ *if* $\boldsymbol{S}$ *is a* δ-*sample set, i.e. it satisfies* $\max\{\mathbb{y}_1^r, \mathbb{y}_2^r, \dots,$
$\mathbb{y}_m^r; \mathbb{x}_{1,1}^r, \mathbb{x}_{1,2}^r, \dots, \mathbb{x}_{1,m}^r; \mathbb{x}_{2,1}^r, \mathbb{x}_{2,2}^r, \dots, \mathbb{x}_{2,m}^r; \dots, \mathbb{x}_{n,1}^r, \mathbb{x}_{n,2}^r, \dots, \mathbb{x}_{n,m}^r\} \leq \delta.$

(4) $\hat{f}$ *is a universal approximator to* f, *i.e. for each compact set* $U \subseteq R^{n+1}$ *and each* $\varepsilon > 0$,
there exists a $\delta(U, \varepsilon) > 0$ *such that* $\sup_{\boldsymbol{x} \in U} |(a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n) - f(\boldsymbol{x})| < \varepsilon$
for each $\boldsymbol{a} = (a_0, a_1, a_2, \dots, a_n) \in \mathbb{a}_0 \times \mathbb{a}_1 \times \mathbb{a}_2 \times \dots \times \mathbb{a}_n$ *and each* $\delta(U, \varepsilon)$-*sample
set* $\boldsymbol{S}$ *(containing* $\boldsymbol{t}$*), where* $\hat{f}(\mathbb{x}) = \hat{f}(\mathbb{x}_1, \mathbb{x}_2, \dots, \mathbb{x}_n) = \mathbb{a}_0 \oplus [\mathbb{a}_1 \otimes \mathbb{x}_1] \oplus [\mathbb{a}_2 \otimes \mathbb{x}_2] \oplus$
$\dots \oplus [\mathbb{a}_n \otimes \mathbb{x}_n]$ *is the first linear regression function based on* $\boldsymbol{S}$.

(5) $\check{f}$ *is a universal approximator to* f, *i.e. for each compact set* $U \subseteq R^{n+1}$ *and each* $\varepsilon > 0$,
there exists a $\delta(U, \varepsilon) > 0$ *such that* $\sup_{\boldsymbol{x} \in U} |\underline{f}(\boldsymbol{x}) - f(\boldsymbol{x})| < \varepsilon$ *and* $\sup_{\boldsymbol{x} \in U} |\bar{f}(\boldsymbol{x}) - f(\boldsymbol{x})| < \varepsilon$
for each $\delta(U, \varepsilon)$-*sample set* $\boldsymbol{S}$ *(containing* $\boldsymbol{t}$*), where* $\check{f}(\mathbb{x}) = \check{f}(\mathbb{x}_1, \mathbb{x}_2, \dots, \mathbb{x}_n)$ *is the
third linear regression function based on* $\boldsymbol{S}$.

(6) $\tilde{f}$ *is a universal approximator to* f, *i.e. for each compact set* $U \subseteq R^{n+1}$ *and
each* $\varepsilon > 0$, *there exists a* $\delta(U, \varepsilon) > 0$ *such that* $\sup\{r(\boldsymbol{x}) \mid \boldsymbol{x} \in U\} < \varepsilon$
for each $\delta(U, \varepsilon)$-*sample set* $\boldsymbol{S}$ *(containing* $\boldsymbol{t}$*), where* $\tilde{f}(\mathbb{x}) = \tilde{f}(\mathbb{x}_1, \mathbb{x}_2, \dots, \mathbb{x}_n) =$
$\langle f(\mathbb{x}_1^c, \mathbb{x}_2^c, \dots, \mathbb{x}_n^c), r(\mathbb{x}_1^r, \mathbb{x}_2^r, \dots, \mathbb{x}_n^r)\rangle$ *is the second linear regression function based
on* $\boldsymbol{S}$.

Proof

Step 1   For two k-variables polynomials $P(\boldsymbol{x})$ and $Q(\boldsymbol{x})$ $(k \geq 1)$, the function $\frac{P(\boldsymbol{x})}{Q(\boldsymbol{x})}$ is
continuous in $R^k - E_Q$, where E_Q is the zero-point set of $Q(\boldsymbol{x})$ (and thus a finite
set). Moreover, the inequalities $\underline{\rho}_{k+1}(\boldsymbol{x}, \boldsymbol{y}) \leq \rho_{k+1}(\boldsymbol{x}, \boldsymbol{y}) \leq \bar{\rho}_{k+1}(\boldsymbol{x}, \boldsymbol{y})$ hold for
any $\{\boldsymbol{x}, \boldsymbol{y}\} \subseteq R^{k+1}$.

Step 2   As $\underline{\rho}_{2m(n+1)}$, $\rho_{2m(n+1)}$, and $\bar{\rho}_{2m(n+1)}$ induce the same topology (i.e. the Euclidean
topology) on $R^{2m(n+1)}$, it can be easily seen from computing formulae of clas-
sical linear regression and Step 1 that (1) is true. By (1), $p_i \circ g(\boldsymbol{s}) = \beta_i(\boldsymbol{s}) :$
$\big(R^{m(n+1)}, \rho_{m(n+1)}\big) \longrightarrow (R, \rho_1)$ is a continuous function (where $p_i : R^{n+1} \longrightarrow R$
is the i-th projection, $i = 1, 2, \dots, n + 1$). Thus $g(\boldsymbol{s})$ is a continuous function i.e.
(2) is true (equivalently, (3) is true). (4)–(6) follow from (3).   □

Example 2.5   Consider the interval sample in Table 4.

As $\underline{f}(t) = 52.0962 + 1.3713t_2 - 0.7425t_3$ and $\bar{f}(t) = 56.5474 + 1.4580t_2 -$
$0.5399t_3$ $(\forall t = (t_1, t_2, t_3) \in R^3)$, $\hat{f}(\mathbb{x}) = \langle 52.0962, 56.5474 \rangle \oplus [\langle 1.3713, 1.4580 \rangle \otimes \mathbb{x}_2] \ominus$
$[\langle 0.5399, 0.7425 \rangle \otimes \mathbb{x}_3]$ and

$$\check{f}(\mathbb{x}_1, \mathbb{x}_2, \mathbb{x}_3) = \lfloor\ 52.0962 + 1.3713\dot{\mathbb{x}}_2 - 0.7425\dot{\mathbb{x}}_3,\ 56.5474 + 1.4580\ddot{\mathbb{x}}_2 - 0.5399\ddot{\mathbb{x}}_3\ \rfloor$$
$$(\forall \mathbb{x} = (\mathbb{x}_1, \mathbb{x}_2, \mathbb{x}_3) \in \mathbb{R}^3).$$

As $f_c(\mathbb{x}_1^c, \mathbb{x}_2^c, \mathbb{x}_3^c) = 53.1765 + 1.4635\mathbb{x}_2^c + 0.6514\mathbb{x}_3^c$ and $r(\mathbb{x}_1^r, \mathbb{x}_2^r, \mathbb{x}_3^r) = 0.7668\mathbb{x}_2^r +$
$0.1887\mathbb{x}_3^r$ $(\forall(\mathbb{x}_1, \mathbb{x}_2, \mathbb{x}_3) \in \mathbb{R}^3)$, $\tilde{f}(\mathbb{x}_1, \mathbb{x}_2, \mathbb{x}_3) = \lceil 53.1765 + 1.4635\mathbb{x}_2^c + 0.6514\mathbb{x}_3^c, 0.7668\mathbb{x}_2^r$
$+ 0.1887\mathbb{x}_3^r \rceil$ $(\forall \mathbb{x} = (\mathbb{x}_1, \mathbb{x}_2, \mathbb{x}_3) \in \mathbb{R}^3)$. Notice that two models obtained in (Baczyński et al
2008) for above sample are $\mathbb{y} = \langle 50.922, 55.431 \rangle \oplus [\langle 1.401, 1.526 \rangle \otimes \mathbb{x}_2] \oplus [\langle 0.624, 0.679 \rangle$
$\otimes \mathbb{x}_3]$ and $\mathbb{y} = \langle 44.463, 61.89 \rangle \oplus [\langle 1.224, 1.703 \rangle \otimes \mathbb{x}_2] \oplus [\langle 0.545, 0.758 \rangle \otimes \mathbb{x}_3]$.

Table 4 Interval sample from Hladíka and Černý ([2012](#))

	$\mathbb{x}_1$	$\mathbb{x}_2$	$\mathbb{x}_3$	$\mathbb{y}$
1	1	$\langle 5.5, 8.5 \rangle$	$\langle 22, 30 \rangle$	$\langle 78, 80 \rangle$
2	1	$\langle 0.5, 1.5 \rangle$	$\langle 27, 31 \rangle$	$\langle 72, 76 \rangle$
3	1	$\langle 10, 12 \rangle$	$\langle 51, 61 \rangle$	$\langle 100, 108 \rangle$
4	1	$\langle 9, 13 \rangle$	$\langle 28, 34 \rangle$	$\langle 81, 95 \rangle$
5	1	$\langle 5, 9 \rangle$	$\langle 50, 54 \rangle$	$\langle 94, 98 \rangle$
6	1	$\langle 8, 14 \rangle$	55	$\langle 105, 113 \rangle$
7	1	$\langle 3, 3 \rangle$	$\langle 69, 73 \rangle$	$\langle 103, 103 \rangle$
8	1	1	$\langle 28, 34 \rangle$	$\langle 72, 74 \rangle$
9	1	$\langle 1, 3 \rangle$	$\langle 51, 571 \rangle$	$\langle 93, 93 \rangle$
10	1	$\langle 18, 24 \rangle$	$\langle 45, 49 \rangle$	$\langle 112, 120 \rangle$
11	1	$\langle 0.5, 1.5 \rangle$	$\langle 32, 48 \rangle$	$\langle 81, 87 \rangle$
12	1	$\langle 7, 15 \rangle$	$\langle 60, 72 \rangle$	$\langle 112, 114 \rangle$
13	1	$\langle 6.5, 13.5 \rangle$	$\langle 61, 75 \rangle$	$\langle 107, 111 \rangle$

Table 5 Special interval sample set

	$\lceil x_1, \delta_1 \rceil$	$\lceil x_2, \delta_2 \rceil$	$\lceil y, r \rceil$
1	$\lceil 1000, 10 \rceil$	$\lceil 5, 2 \rceil$	$\lceil 100, 5 \rceil$
2	$\lceil 600, 10 \rceil$	$\lceil 7, 2 \rceil$	$\lceil 75, 5 \rceil$
3	$\lceil 1200, 20 \rceil$	$\lceil 6, 2 \rceil$	$\lceil 80, 5 \rceil$
4	$\lceil 500, 10 \rceil$	$\lceil 6, 2 \rceil$	$\lceil 70, 5 \rceil$
5	$\lceil 300, 10 \rceil$	$\lceil 8, 2 \rceil$	$\lceil 50, 5 \rceil$

Example 2.6 Consider the special interval input–output sample set $\mathbb{S} = \{(\mathbb{x}_{1,i}, \mathbb{x}_{2,i}, \mathbb{y}_i) \mid i = 1, 2, 3, 4, 5\}$ in Table 5 (where $\delta_2 \equiv 2$, $r \equiv 5$, and δ_1 is almost equal to 10).

Then $\underline{f}(\boldsymbol{t}) = 111.7393 + 0.0124t_1 - 11.4787t_2$, $f_c(\boldsymbol{t}) = f_c(t_1, t_2) = 139.57317 + 0.0124t_1 - 11.4787t_2$, and $\bar{f}(\boldsymbol{t}) = \bar{f}(t_1, t_2) = 167.407 + 0.0124t_1 - 11.4787t_2$ ($\forall \boldsymbol{t} = (t_1, t_2) \in R^2$). By Algorithm 2.3, $\hat{f}(\mathbb{t}) = \hat{f}(\lceil t_1, 10\rceil, \lceil t_2, 2\rceil) = \langle 111.7393 + 0.0124t_1 - 11.4787t_2, 167.407 + 0.0124t_1 - 11.4787t_2 \rangle = \check{f}(\mathbb{t})$ ($\forall \mathbb{t} = (\lceil t_1, 10 \rceil, \lceil t_2, 2 \rceil) \in \mathbb{R}$). As $r(\delta_1, \delta_2) = 2.5\delta_2 = 5$ (if $\delta_2 = 2$), $\tilde{f}(\mathbb{t}) = \lceil 0.0124t_1 - 11.4787t_2 + 139.5732, 5 \rceil = f_c(\boldsymbol{t}) \pm 5 = f_c(t_1, t_2) \pm 5$ ($\mathbb{t} = (\lceil t_1, 10 \rceil, \lceil t_2, 2 \rceil) \in \mathbb{R}$). This motivates the following easy-to-use Corollary 2.7:

Corollary 2.7 *For a special sample (or data set)* $\mathbb{S} = \big\{(\lceil x_{1,1}, \varepsilon_1\rceil, \lceil x_{2,1}, \varepsilon_2\rceil, \ldots, \lceil x_{n,1}, \varepsilon_n\rceil;$ $\lceil y_1, \varepsilon_0 \rceil), (\lceil x_{1,2}, \varepsilon_1 \rceil, \lceil x_{2,2}, \varepsilon_2 \rceil, \ldots, \lceil x_{n,2}, \varepsilon_n \rceil; \lceil y_2, \varepsilon_0 \rceil), \ldots, (\lceil x_{1,m}, \varepsilon_1 \rceil, \lceil x_{2,m}, \varepsilon_2 \rceil, \ldots,$ $\lceil x_{n,m}, \varepsilon_n \rceil; \lceil y_m, \varepsilon_0 \rceil)\big\}$, *the first two linear regression functions in Algorithm 2.3 are the same:* $\hat{f}(\mathbb{t}) = \hat{f}(\lceil t_1, \varepsilon_1 \rceil, \lceil t_2, \varepsilon_2 \rceil, \ldots, \lceil t_n, \varepsilon_n \rceil) = \downarrow f_c(\boldsymbol{t}) - o, f_c(\boldsymbol{t}) + o \downharpoonleft = \check{f}(\mathbb{t})$, *it equals exactly to* $\lceil f_c(t_1, \ldots, t_n), |o| \rceil = f_c(t_1, \ldots, t_n) \pm |o|$, *which can be looked to be* $\tilde{f}(\mathbb{t})$ *(because* $r(\boldsymbol{\delta})$ *in Algorithm 2.3 has infinite many chooses, including* $|o|$*), here* $\boldsymbol{t} = (t_1, t_2, \ldots, t_n)$ *and* $o = \varepsilon_0 - a_1\varepsilon_1 - a_2\varepsilon_2 - \cdots - a_n\varepsilon_n$.

Proof For two column vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)^{\mathrm{T}}$ and $\boldsymbol{z} = (z_1, z_2, \ldots, z_m)^{\mathrm{T}}$, we write $\bar{x} = \frac{1}{m}(x_1 + x_2 + \cdots + x_m)$ and $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{z} = x_1 z_1 + x_2 z_2 + \cdots + x_m z_m$. Using $\boldsymbol{x}_1$ (resp., $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_n, \boldsymbol{y}$) to denote the column vector $(x_{1,1}, x_{1,2}, \ldots, x_{1,m})^{\mathrm{T}}$ (resp., $(x_{2,1}, x_{2,2}, \ldots, x_{2,m})^{\mathrm{T}}, \ldots, (x_{n,1}, x_{n,2}, \ldots, x_{n,m})^{\mathrm{T}}, (y_1, y_2, \ldots, y_m)^{\mathrm{T}}$) and $\varepsilon^{\downarrow}$ to denote a

column vector with the constant coordinates ε, we get $\overline{(\boldsymbol{y} - \varepsilon_0^!)^{\mathrm{T}} (\boldsymbol{x}_i - \varepsilon_i^!)} - m \overline{\boldsymbol{y} - \varepsilon_0^!} \, \overline{\boldsymbol{x}_i - \varepsilon_i^!}$

$= \left(\overline{\boldsymbol{y}^{\mathrm{T}} \boldsymbol{x}_i - m \varepsilon_i \bar{\boldsymbol{y}} - m \varepsilon_0 \bar{\boldsymbol{x}}_i + \varepsilon_0^{!\mathrm{T}} \varepsilon_i^!} \right) - m \overline{\boldsymbol{y} - \varepsilon_0^!} \, \overline{\boldsymbol{x}_i - \varepsilon_i^!} = \overline{\boldsymbol{y}^{\mathrm{T}} \boldsymbol{x}_i} - m \bar{\boldsymbol{y}} \bar{\boldsymbol{x}}_i \ (i = 1, 2, \ldots, n).$

Analogously, $\overline{(\boldsymbol{x}_i - \varepsilon_i^!)^{\mathrm{T}} (\boldsymbol{x}_j - \varepsilon_j^!)} - m \overline{\boldsymbol{x}_i - \varepsilon_i^!} \, \overline{\boldsymbol{x}_j - \varepsilon_j^!} = \overline{\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j} - m \bar{\boldsymbol{x}}_i \bar{\boldsymbol{x}}_j \ (i, j = 1, 2, \ldots, n).$

This implies $\underline{a}_i = a_i \ (i, j = 1, 2, \ldots, n)$, and thus $\underline{a}_0 = \overline{\boldsymbol{y} - \varepsilon_0^!} - \underline{a}_1 \overline{\boldsymbol{x}_1 - \varepsilon_1^!} - \underline{a}_2 \boldsymbol{x}_2 - \varepsilon_2^! - \cdots - \underline{a}_n \overline{\boldsymbol{x}_n - \varepsilon_n^!} = \overline{\boldsymbol{y} - \varepsilon_0^!} - a_1 \overline{\boldsymbol{x}_1 - \varepsilon_1^!} - a_2 \overline{\boldsymbol{x}_2 - \varepsilon_2^!} - \cdots - a_n \overline{\boldsymbol{x}_n - \varepsilon_n^!} = (\bar{\boldsymbol{y}} - a_1 \bar{\boldsymbol{x}}_1 - a_2 \bar{\boldsymbol{x}}_2 - \cdots - a_n \bar{\boldsymbol{x}}_n) - (\varepsilon_0 - a_1 \varepsilon_1 - a_2 \varepsilon_2 - \cdots - a_n \varepsilon_n) = a_0 - o.$ Therefore, $\underline{f}(\boldsymbol{t}) = f_c(\boldsymbol{t}) - o$; analogously, $\bar{f}(\boldsymbol{t}) = f_c(\boldsymbol{t}) + o.$ $\qquad \square$

Remark 2.8 In practice we can ask people to collect sample as that in Corollary 2.7 so as to use the simplified models in Corollary 2.7.

3 Two models stemmed from the classical linear interpolations

In this section, we first present two linear-like interpolation models with interval input–output sample, each is stemmed from the classical linear interpolation modal with real input–output sample. Then we prove that each modal is an universal approximator to the corresponding linear interpolation modal.

Example 3.1 (1) For a real input–output data set $S = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$ satisfying $x_{i+1} - x_i = h_i > 0 \ (i = 1, 2, \ldots, m-1, \ m > 3)$, the continuous function $\hat{f}(x) = y_1 l(x_1)(x) + y_2 l(x_2)(x) + \cdots + y_m l(x_m)(x)$ (called the 1-variable linear interpolation function) interpolates the given sample, i.e. it satisfies $\hat{f}(x_i) = y_i \ (i = 1, 2, \ldots, m)$, where

$$l(x_1)(x) = \begin{cases} \frac{x - x_1}{h_1} + 1, & x \in [x_1 - h_1, x_1), \\ \frac{x_1 - x}{h_1}, & x \in [x_1, x_2), \\ 0, & \text{otherwise}, \end{cases}$$
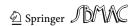
$$l(x_m)(x) = \begin{cases} \frac{x - x_{m-1}}{h_{m-1}}, & x \in [x_{m-1}, x_m), \\ \frac{x_m - x}{h_m} + 1, & x \in [x_m, x_m + h_{m-1}], \\ 0, & \text{otherwise}, \end{cases}$$

and

$$l(x_i)(x) = \begin{cases} \frac{x - x_i}{h_{i-1}}, & x \in [x_{i-1}, x_i), \\ \frac{x_i - x}{h_i}, & x \in [x_i, x_{i+1}), \\ 0, & \text{otherwise}, \end{cases} \quad (i = 2, 3, \ldots, m-1)$$

are called basis functions relaying on S.

(2) Let $\{(\boldsymbol{x}_0, f(\boldsymbol{x}_0)) \mid \boldsymbol{x}_0 \in S\}$ be a multi-input-one output sample set (from an n-variable function f), where $S = S_1 \times S_2 \times \cdots \times S_n$, $S_1 = \left\{ x_1^{(1)}, x_1^{(2)}, \ldots, x_1^{(m_1)} \right\}$ with $x_1^{(i+1)} - x_1^{(i)} = h_1^{(i)} > 0 \ (i = 1, 2, \ldots, m_1 - 1)$, $S_2 = \left\{ x_2^{(1)}, x_2^{(2)}, \ldots, x_2^{(m_2)} \right\}$ with $x_2^{(i+1)} - x_2^{(i)} = h_2^{(i)} > 0 \ (i = 1, 2, \ldots, m_2 - 1)$, $\ldots$, $S_n = \left\{ x_n^{(1)}, x_n^{(2)}, \ldots, x_n^{(m_n)} \right\}$ with $x_n^{(i+1)} - x_n^{(i)} = h_n^{(i)} > 0 \ (i = 1, 2, \ldots, m_n - 1)$, and $m = \min\{m_1, m_2, \ldots, m_n\} > 3$. Then the n-variable linear-like interpolation function $\hat{f}(\boldsymbol{x}) = \sum_{\boldsymbol{x}_0 \in S} f(\boldsymbol{x}_0) l(\boldsymbol{x}_0)(\boldsymbol{x})$ ($\forall \boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in R^n$) is continuous and satisfies $\hat{f}(\boldsymbol{x}_0) = f(\boldsymbol{x}_0) \ (\forall \boldsymbol{x}_0 \in$

S), where $l(\boldsymbol{x}_0)(\boldsymbol{x}) = l(x_1^{(k_1)})(x_1) \times l(x_2^{(k_2)})(x_2) \times \cdots \times l(x_n^{(k_n)})(x_n)$ if $\boldsymbol{x}_0 = \left(x_1^{(k_1)}, x_2^{(k_2)}, \ldots, x_n^{(k_n)}\right)$, $l(x_1^{(k_1)})(x_1)$ is a base function relying on the sample set S_1, $l(x_2^{(k_2)})(x_2)$ is a base function relying on the sample set S_2, ..., $l(x_n^{(k_n)})(x_n)$ is a base function relying on the sample set S_n.

Motivated by Example 3.1, we have (as in Sect. 2) the following Algorithm 3.2:

Algorithm 3.2 Consider an interval number sample set (from an n-variable function f, $n > 1$) $\{(\mathbf{x}_0, f_\wr(\mathbf{x}_0)) \mid \mathbf{x}_0 \in \mathbb{S}\}$, where $(\mathbf{x}_0, f_\wr(\mathbf{x}_0))$ is the observe value of the true sample $(x_0, f(x_0))$, $\mathbb{S} = \mathbb{S}_1 \times \mathbb{S}_2 \times \cdots \times \mathbb{S}_n$ (the set of all interpolation nodes), $\mathbb{S}_1 = \left\{ \mathbb{x}_1^{(1)}, \mathbb{x}_1^{(2)}, \ldots, \mathbb{x}_1^{(m_1)} \right\}$ with $\mathbb{x}_1^{(2)} \ominus \mathbb{x}_1^{(1)} = h_1^{(1)} > 0$, $\mathbb{x}_1^{(3)} \ominus \mathbb{x}_1^{(2)} = h_1^{(2)} > 0$, ..., and $\mathbb{x}_1^{(m_1)} \ominus \mathbb{x}_1^{(m_1-1)} = h_1^{(m_1-1)} > 0$, $\mathbb{S}_2 = \left\{ \mathbb{x}_2^{(1)}, \mathbb{x}_2^{(2)}, \ldots, \mathbb{x}_2^{(m_2)} \right\}$ with $\mathbb{x}_2^{(2)} \ominus \mathbb{x}_2^{(1)} = h_2^{(1)} > 0$, $\mathbb{x}_2^{(3)} \ominus \mathbb{x}_2^{(2)} = h_2^{(2)} > 0$, ..., and $\mathbb{x}_2^{(m_2)} \ominus \mathbb{x}_2^{(m_2-1)} = h_2^{(m_2-1)} > 0$, ..., $\mathbb{S}_n = \left\{ \mathbb{x}_n^{(1)}, \mathbb{x}_n^{(2)}, \ldots, \mathbb{x}_n^{(m_n)} \right\}$ with $\mathbb{x}_n^{(2)} \ominus \mathbb{x}_n^{(1)} = h_n^{(1)} > 0$, $\mathbb{x}_n^{(3)} \ominus \mathbb{x}_n^{(2)} = h_n^{(2)} > 0$, ..., and $\mathbb{x}_n^{(m_n)} \ominus \mathbb{x}_n^{(m_n-1)} = h_n^{(m_n-1)} > 0$. Then the n-variable linear-like interpolation function of f, based on the data set $\{(\mathbf{x}_0, f_\wr(\mathbf{x}_0)) \mid \mathbf{x}_0 \in \mathbb{S}\}$, can be given by the following ways:

(1) **Step 1** Compute the n-variable linear-like interpolation function $\underline{f}(\boldsymbol{t}) = \sum_{\mathbf{x}_0 \in \mathbb{S}} p_1 (f_\wr(\mathbf{x}_0)) l(\dot{\mathbb{x}}_1, \dot{\mathbb{x}}_2, \ldots, \dot{\mathbb{x}}_n)(\boldsymbol{t})$ $(\forall \boldsymbol{t} = (t_1, t_2, \ldots, t_n) \in R^n)$ based on the real input–output sample set $\{((\dot{\mathbb{x}}_1, \dot{\mathbb{x}}_2, \ldots, \dot{\mathbb{x}}_n), p_1(f_\wr(\mathbf{x}_0))) \mid \mathbf{x}_0 = (\mathbb{x}_1, \mathbb{x}_2, \ldots, \mathbb{x}_n) \in \mathbb{S}\}$, using the method given in Example 3.1(2).

Step 2 Similarly, compute the n-variable linear-like interpolation function $\bar{f}(\boldsymbol{t}) = \sum_{\mathbf{x}_0 \in \mathbb{S}} p_2(f_\wr(\mathbf{x}_0)) l(\ddot{\mathbb{x}}_1, \ddot{\mathbb{x}}_2, \ldots, \ddot{\mathbb{x}}_n)(\boldsymbol{t})$ $(\forall \boldsymbol{t} = (t_1, t_2, \ldots, t_n) \in R^n)$ based on the real input–output sample set $\{((\ddot{\mathbb{x}}_1, \ddot{\mathbb{x}}_2, \ldots, \ddot{\mathbb{x}}_n), p_2(f_\wr(\mathbf{x}_0))) \mid \mathbf{x}_0 = (\mathbb{x}_1, \mathbb{x}_2, \ldots, \mathbb{x}_n) \in \mathbb{S}\}$, using the method given in Example 3.1(2).

Step 3 Obtain, based on Steps 1–2, the first n-variable linear-like interpolation function $\check{f}(\mathbf{x}) = \lfloor \underline{f}(\dot{\mathbb{x}}_1, \dot{\mathbb{x}}_2, \ldots, \dot{\mathbb{x}}_n), \bar{f}(\ddot{\mathbb{x}}_1, \ddot{\mathbb{x}}_2, \ldots, \ddot{\mathbb{x}}_n) \rfloor$ $(\forall \mathbf{x} = (\mathbb{x}_1, \mathbb{x}_2, \ldots, \mathbb{x}_n) \in \mathbb{R}^n)$.

(2) **Step 1** Compute the n-variable linear-like interpolation function $f_c(\boldsymbol{t}) = \sum_{\mathbf{x}_0 \in \mathbb{S}} f_\wr(\mathbf{x}_0)^c l(\mathbb{x}_1^c, \mathbb{x}_2^c, \ldots, \mathbb{x}_n^c)(\boldsymbol{t})$ $(\forall \boldsymbol{t} = (t_1, t_2, \ldots, t_n) \in R^n)$ based on the real input–output sample set $\{((\mathbb{x}_1^c, \mathbb{x}_2^c, \ldots, \mathbb{x}_n^c), f_\wr(\mathbf{x}_0)^c) \mid \mathbf{x}_0 = (\mathbb{x}_1, \mathbb{x}_2, \ldots, \mathbb{x}_n) \in \mathbb{S}\}$, using the method given in Example 3.1(2).

Step 2 Compute the ordinary linear regression function (without the term b_0)[4] $r(\boldsymbol{\delta}) = r(\delta_1, \ldots, \delta_n) = b_1\delta_1 + \cdots + b_n\delta_n$ $(\forall \boldsymbol{\delta} \in R^n)$, based on the real input–output sample set $\{((\mathbb{x}_1^r, \mathbb{x}_2^r, \ldots, \mathbb{x}_n^r), f_\wr(\mathbf{x}_0)^r) \mid \mathbf{x}_0 = (\mathbb{x}_1, \mathbb{x}_2, \ldots, \mathbb{x}_n) \in \mathbb{S}\}$. Then $\tilde{f}(\mathbf{x}) = f_c(\mathbb{x}_1^c, \mathbb{x}_2^c, \ldots, \mathbb{x}_n^c) \pm r(\mathbb{x}_1^r, \mathbb{x}_2^r, \ldots, \mathbb{x}_n^r) = \lceil f_c(\mathbb{x}_1^c, \mathbb{x}_2^c, \ldots, \mathbb{x}_n^c), r(\mathbb{x}_1^r, \mathbb{x}_2^r, \ldots, \mathbb{x}_n^r) \rceil$ is the second n-variable linear-like interpolation function.

The rationality of Algorithm 3.2 is guaranteed by the following Theorem 3.3. For a vector $\mathbf{x} = (\mathbb{x}_1, \mathbb{x}_2, \ldots, \mathbb{x}_n)$ with interval number coordinates, we write $\boldsymbol{t} \in \mathbf{x}$ if $\boldsymbol{t} = (t_1, t_2, \ldots, t_n)$ and $(t_1, t_2, \ldots, t_n) \in \mathbb{x}_1 \times \mathbb{x}_2 \times \cdots \times \mathbb{x}_n$.

Theorem 3.3 (1) *For each real input–output sample* $\boldsymbol{s} = \left\{(\mathbf{x}_0^\bullet, f_\wr(\mathbf{x}_0)^\bullet) \mid \mathbf{x}_0 \in \mathbb{S}\right\}$ *(where $\mathbf{x}_0^\bullet \in \mathbf{x}_0$ and $f_\wr(\mathbf{x}_0)^\bullet \in f_\wr(\mathbf{x}_0)$), let $f_{\boldsymbol{s}}(\boldsymbol{t}) = f_{\boldsymbol{s}}(t_1, \ldots, t_n)$ be the n-variable linear-like interpolation function, based on the real input–output sample set $\boldsymbol{s}$ and obtained by using the method given in Example 3.1(2). Then $f_{\boldsymbol{s}}(\boldsymbol{t})$ is also continuous on $\boldsymbol{s}$.*

(2) *If $\left\{(\mathbf{x}_0^\bullet, f_\wr(\mathbf{x}_0)^\bullet) \mid \mathbf{x}_0 \in \mathbb{S}\right\}$ (where $\mathbf{x}_0^\bullet \in \mathbf{x}_0$ and $f_\wr(\mathbf{x}_0)^\bullet \in f_\wr(\mathbf{x}_0)$) is the true sample set. Then, for each $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f_{\boldsymbol{s}}(x_0) - f_\wr(\mathbf{x}_0)^\bullet| < \varepsilon$ $(\forall \boldsymbol{x}_0 \in$*

---

4 We can also take $r(\boldsymbol{\delta})$ to be the cubic spline interpolation function (see Example 3.5(1)).

Table 6 Special interval sample set

$\mathbb{X}_1$ ╲ $\mathbb{X}_2$	$\lceil 5, 0.1 \rceil$	$\lceil 6, 0.1 \rceil$	$\lceil 7, 0.1 \rceil$
$\lceil 300, 10 \rceil$	$\lceil 50, 1 \rceil$	$\lceil 60, 1 \rceil$	$\lceil 70, 1 \rceil$
$\lceil 600, 10 \rceil$	$\lceil 60, 1 \rceil$	$\lceil 70, 1 \rceil$	$\lceil 80, 1 \rceil$
$\lceil 900, 10 \rceil$	$\lceil 70, 1 \rceil$	$\lceil 80, 1.1 \rceil$	$\lceil 100, 1 \rceil$

$\mathbf{x}_0$, $\forall \mathbf{x}_0 = (\mathbb{x}_1, \mathbb{x}_2, \ldots, \mathbb{x}_n) \in \mathbb{S}$) *whenever* $\max \left\{ \mathbb{x}_1^r, \mathbb{x}_2^r, \ldots, \mathbb{x}_m^r, f_{\lambda}(\mathbf{x}_0)^r \right\} < \delta$, *where* $f_{\boldsymbol{s}}(\boldsymbol{t}) = f_{\boldsymbol{s}}(t_1, \ldots, t_n)$ *is the n-variable linear-like interpolation function, based on the real input–output sample set* $\boldsymbol{s} = \left\{ (\mathbf{x}_0^{\bullet}, f_{\lambda}(\mathbf{x}_0)^{\bullet}) \mid \mathbf{x}_0 \in \mathbb{S} \right\}$ *(where* $\mathbf{x}_0^{\bullet} \in \mathbf{x}_0$ *and* $f_{\lambda}(\mathbf{x}_0)^{\bullet} \in f_{\lambda}(\mathbf{x}_0)$) *and obtained by using the method given in Example* 3.1(2).

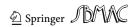**Proof** Similarly to that of Theorem 2.4. □

Remark 3.4 In practical problem if the set $\mathbb{S}$ of all interpolation nodes (related to the n-variable function f) is not an n-dimension cub-like set, then we consider the new set $\mathbb{\check{S}}$ of all interpolation nodes (related to the n-variable function $\check{f}$ which is an extension of f) and give the two n-variable linear-like interpolation functions ($\check{f}$ and $\tilde{f}$) of $\check{f}$ using Algorithm 3.2 (and thus Theorem 3.3 holds for $\check{f}$, particularly, for f), where $\mathbb{\check{S}} = \mathbb{S}_1 \times \mathbb{S}_2 \times \cdots \times \mathbb{S}_n$, $\mathbb{S}_i = \{ P_i(\boldsymbol{s}) \mid \boldsymbol{s} \in \mathbb{S} \}$ $(i = 1, 2, \ldots, n)$, $P_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ $(i = 1, 2, \ldots, n)$ is the $i - th$ projection, and the value of $\check{f}(\boldsymbol{s})$ will be given by experts using arithmetic average operator, weighted average operator, or other aggregation operators (Beliakov et al. 2007; Cubillo et al. 2015; Deschrijver and Kerre 2008) ($\boldsymbol{s} \in \mathbb{\check{S}} - \mathbb{S}$).

Example 3.5 (1) The 1-variable linear-like interpolation functions are simple. Consider the interval input–output sample set $\mathbb{S} = \{ (\mathbb{x}_i, \mathbb{y}_i) \mid i = 1, 2, \ldots, 41 \}$ in (Chuang 2008, Table 2). By Algorithm 3.2, $\check{f}(\mathbf{x}_i) = \mathbb{y}_i$ $(i = 1, 2, \ldots, 41)$. By Algorithm 3.2(2), we have $[\tilde{f}(\mathbf{x}_i)]^c = [\mathbb{y}_i]^c$ $(i = 1, 2, \ldots, 41)$ and $r(\delta) = 25.0891\delta$, thus the minimum of the lengths of all intervals $\tilde{f}(\mathbf{x}_i) \cap \mathbb{y}_i$ $(i = 1, 2, \ldots, 41)$ is 0.0302. However, if we take $r(\delta)$ to be the cubic spline interpolation function, then $\tilde{f}(\mathbf{x}_i) = \mathbb{y}_i$ $(i = 1, 2, \ldots, 41)$, and thus the minimum of the lengths of all intervals $\tilde{f}(\mathbf{x}_i) \cap \mathbb{y}_i$ $(i = 1, 2, \ldots, 41)$ is 0.9123.

(2) Consider the special interval input–output sample set $\mathbb{S} = \{ (\mathbb{x}_{1,i}, \mathbb{x}_{2,j}, \mathbb{y}_{i,j}) \mid i, j = 1, 2, 3 \}$ as shown in Table 6 (where $\delta_1 \equiv 10$, $\delta_2 \equiv 0.1$, and r almost takes the same value 1).

Step 1 $l\left(x_1^{(1)}\right)(t_1) = \max \left\{ 1 - \frac{1}{300}|t_1 - 300|, 0 \right\}$, $l\left(x_1^{(2)}\right)(t_1) = \max \left\{ 1 - \frac{1}{300}|t_1 - 600|, 0 \right\}$, and $l\left(x_1^{(3)}\right)(t_1) = \max \left\{ 1 - \frac{1}{300}|t_1 - 900|, 0 \right\}$. Similarly, $l\left(x_2^{(1)}\right)(t_2) = \max\{1 - |t_2 - 5|, 0\}$, $l\left(x_2^{(2)}\right)(t_2) = \max\{1 - |t_2 - 6|, 0\}$, and $l\left(x_2^{(3)}\right)(t_2) = \max\{1 - |t_2 - 7|, 0\}$. Thus

$$
\begin{aligned}
f_c(\boldsymbol{t}) &= y_{1,1} l\left(x_1^{(1)}\right)(t_1) l\left(x_2^{(1)}\right)(t_2) + y_{1,2} l\left(x_1^{(1)}\right)(t_1) l\left(x_2^{(2)}\right)(t_2) \\
&\quad + y_{2,1} l\left(x_1^{(2)}\right)(t_1) l\left(x_2^{(1)}\right)(t_2) + y_{2,2} l\left(x_1^{(2)}\right)(t_1) l\left(x_2^{(2)}\right)(t_2) \\
&= 50\left(2 - \frac{t_1}{300}\right)(6 - t_2) + 60\left(2 - \frac{t_1}{300}\right)(t_2 - 5) \\
&\quad + 60\left(\frac{t_1}{300} - 1\right)(6 - t_2) + 70\left(\frac{t_1}{300} - 1\right)(t_2 - 5) \\
&= \frac{1}{30} t_1 + 10 t_2 - 10
\end{aligned}
$$

if $t = (t_1, t_2) \in [300, 600) \times [5, 6)$; $f_c(t) = 60\left(2 - \frac{t_1}{300}\right)(7 - t_2) + 70\left(2 - \frac{t_1}{300}\right)(t_2 - 6) +$

$70\left(\frac{t_1}{300} - 1\right)(7 - t_2) + 80\left(\frac{t_1}{300} - 1\right)(t_2 - 6) = \frac{1}{30}t_1 + 10t_2 - 10$ if $t \in [300, 600) \times [6, 7)$; $f_c(t)$

$= 60\left(3 - \frac{t_1}{300}\right)(6 - t_2) + 70\left(3 - \frac{t_1}{300}\right)(t_2 - 5) + 50\left(\frac{t_1}{300} - 2\right)(6 - t_2) + 80\left(\frac{t_1}{300} - 2\right)(t_2 - 5) =$

$\frac{1}{30}t_1 + 10t_2 - 10$ if $t \in [600, 900) \times [5, 6)$; $f_c(t) = 70\left(3 - \frac{t_1}{300}\right)(7 - t_2) + 80\left(3 - \frac{t_1}{300}\right)(t_2 -$

$6) + 80\left(\frac{t_1}{300} - 2\right)(7 - t_2) + 100\left(\frac{t_1}{300} - 2\right)(t_2 - 6) = \frac{1}{30}t_1t_2 - \frac{1}{6}t_1 - 10t_2 + 110$ if $t \in$

$[600, 900) \times [6, 7)$; $f_c(t) = y_{1,1}l\left(x_1^{(1)}\right)(t_1)l\left(x_2^{(1)}\right)(t_2) = \frac{1}{6}t_1t_2 - \frac{2}{3}t_1$ if $t \in [0, 300) \times [4, 5)$;

$f_c(t) = y_{1,1}l\left(x_1^{(1)}\right)(t_1)l\left(x_2^{(1)}\right)(t_2) + y_{1,2}l\left(x_1^{(1)}\right)(t_1)l\left(x_2^{(2)}\right)(t_2) = \frac{1}{30}t_1t_2$ if $t \in [0, 300) \times$

$[5, 6)$; $f_c(t) = y_{1,2}l\left(x_1^{(1)}\right)(t_1)l\left(x_2^{(2)}\right)(t_2) + y_{1,3}l\left(x_1^{(1)}\right)(t_1)l\left(x_2^{(3)}\right)(t_2) = \frac{1}{30}t_1t_2$ if $t \in$

$[0, 300) \times [6, 7)$; $f_c(t) = y_{1,3}l\left(x_1^{(1)}\right)(t_1)l\left(x_2^{(3)}\right)(t_2) = -\frac{7}{30}t_1t_2 + \frac{28}{15}t_1$ if $t \in [0, 300) \times [7, 8]$;

$f_c(t) = y_{1,1}l\left(x_1^{(1)}\right)(t_1)l\left(x_2^{(1)}\right)(t_2) + y_{2,1}l\left(x_1^{(2)}\right)(t_1)l\left(x_2^{(1)}\right)(t_2) = \frac{1}{30}t_1t_2 - \frac{2}{15}t_1 + 40t_2 - 160$

if $t \in [300, 600) \times [4, 5)$; $f_c(t) = y_{1,3}l\left(x_1^{(1)}\right)(t_1)l\left(x_2^{(3)}\right)(t_2) + y_{2,3}l\left(x_1^{(2)}\right)(t_1)l\left(x_2^{(3)}\right)(t_2) =$

$-\frac{1}{30}t_1t_2 + \frac{4}{15}t_1 - 60t_2 + 480$ if $t \in [300, 600) \times [7, 8]$; $f_c(t) = y_{2,1}l\left(x_1^{(2)}\right)(t_1)l\left(x_2^{(1)}\right)(t_2) +$

$y_{3,1}l\left(x_1^{(3)}\right)(t_1)l\left(x_2^{(1)}\right)(t_2) = \frac{1}{30}t_1t_2 - \frac{2}{15}t_1 + 40t_2 - 160$ if $t \in [600, 900) \times [4, 5)$;

$f_c(t) = y_{2,3}l\left(x_1^{(2)}\right)(t_1)l\left(x_2^{(3)}\right)(t_2) + y_{3,3}l\left(x_1^{(3)}\right)(t_1)l\left(x_2^{(3)}\right)(t_2) = -\frac{1}{15}t_1t_2 + \frac{8}{15}t_1 - 40t_2 + 320$

if $t \in [600, 900) \times [7, 8]$; $f_c(t) = y_{3,1}l\left(x_1^{(3)}\right)(t_1)l\left(x_2^{(1)}\right)(t_2) = -\frac{7}{30}t_1t_2 + \frac{14}{15}t_1 + 280t_2 - 1120$

if $t \in [900, 1200] \times [4, 5)$; $f_c(t) = y_{3,1}l\left(x_1^{(3)}\right)(t_1)l\left(x_2^{(1)}\right)(t_2) + y_{3,2}l\left(x_1^{(3)}\right)(t_1)l\left(x_2^{(2)}\right)(t_2) =$

$-\frac{1}{30}t_1t_2 - \frac{1}{15}t_1 + 40t_2 + 80$ if $t \in [900, 1200] \times [5, 6)$; $f_c(t) = y_{3,2}l\left(x_1^{(3)}\right)(t_1)l\left(x_2^{(2)}\right)(t_2) +$

$y_{3,3}l\left(x_1^{(3)}\right)(t_1)l\left(x_2^{(3)}\right)(t_2) = -\frac{1}{15}t_1t_2 + \frac{2}{15}t_1 + 80t_2 - 160$ if $t \in [900, 1200] \times [6, 7)$;

$f_c(t) = y_{3,3}l\left(x_1^{(3)}\right)(t_1)l\left(x_2^{(3)}\right)(t_2) = \frac{1}{3}t_1t_2 - \frac{8}{3}t_1 - 400t_2 + 3200$ if $t \in [900, 1200] \times [7, 8]$.

In a words, $f_c(t) = \begin{cases} \frac{1}{30}t_1 + 10t_2 - 10, & t \in [300, 600) \times [5, 7), \\ \frac{1}{30}t_1 + 10t_2 - 10, & t \in [600, 900) \times [5, 6), \\ \frac{1}{30}t_1t_2 - \frac{1}{6}t_1 - 10t_2 + 110, & t \in [600, 900) \times [6, 7), \\ \frac{1}{6}t_1t_2 - \frac{2}{3}t_1, & t \in [0, 300) \times [4, 5), \\ \frac{1}{30}t_1t_2, & t \in [0, 300) \times [5, 7), \\ -\frac{7}{30}t_1t_2 + \frac{28}{15}t_1, & t \in [0, 300) \times [7, 8], \\ \frac{1}{30}t_1t_2 - \frac{2}{5}t_1 + 40t_2 - 160, & t \in [300, 600) \times [4, 5), \\ -\frac{1}{30}t_1t_2 + \frac{4}{5}t_1 - 60t_2 + 480, & t \in [300, 600) \times [7, 8], \\ \frac{1}{30}t_1t_2 - \frac{2}{15}t_1 + 40t_2 - 160, & t \in [600, 900) \times [4, 5), \\ -\frac{1}{15}t_1t_2 + \frac{8}{15}t_1 - 40t_2 + 320, & t \in [600, 900) \times [7, 8], \\ -\frac{7}{30}t_1t_2 + \frac{14}{15}t_1 + 280t_2 - 1120, & t \in [900, 1200] \times [4, 5), \\ -\frac{1}{30}t_1t_2 - \frac{1}{15}t_1 + 40t_2 + 80, & t \in [900, 1200] \times [5, 6), \\ -\frac{1}{15}t_1t_2 + \frac{2}{15}t_1 + 80t_2 - 160, & t \in [900, 1200] \times [6, 7), \\ \frac{1}{3}t_1t_2 - \frac{8}{3}t_1 - 400t_2 + 3200, & t \in [900, 1200] \times [7, 8], \\ 0, & \text{otherwise.} \end{cases}$

Step 2 $l\big(x_1^{(1)}-\delta_1\big)(t_1) = \max\{1-\frac{1}{300}|t_1-290|,\,0\}$, $l\big(x_1^{(2)}-\delta_1\big)(t_1) = \max\{1-\frac{1}{300}|t_1-590|,\,0\}$, $l\big(x_1^{(3)}-\delta_1\big)(t_1) = \max\{1-\frac{1}{300}|t_1-890|,\,0\}$, $l\big(x_2^{(1)}-\delta_2\big)(t_2) = \max\{1-|t_2-4.9|,\,0\}$, $l\big(x_2^{(2)}-\delta_2\big)(t_2) = \max\{1-|t_2-5.9|,\,0\}$, $l\big(x_2^{(3)}-\delta_2\big)(t_2) = \max\{1-|t_2-6.9|,\,0\}$. Thus

$$\underline{f}(t) = \begin{cases}
\frac{1}{30}t_1 + 10t_2 - \frac{29}{3} & t \in [290, 590) \times [4.9, 6.9), \\[4pt]
-\frac{1}{3000}t_1 t_2 + \frac{1049}{30000}t_1 + \frac{3059}{300}t_2 - \frac{31891}{3000}, & t \in [590, 890) \times [4.9, 5.9), \\[4pt]
\frac{101}{3000}t_1 t_2 - \frac{4969}{30000}t_1 - \frac{2959}{300}t_2 + \frac{32317}{3000}, & t \in [590, 890) \times [5.9, 6.9), \\[4pt]
\frac{49}{300}t_1 t_2 - \frac{637}{1000}t_1 + \frac{49}{30}t_2 - \frac{637}{100}, & t \in [0, 290) \times [3.9, 4.9), \\[4pt]
\frac{1}{30}t_1 t_2 + \frac{1}{3}t_2, & t \in [0, 290) \times [4.9, 6.9), \\[4pt]
-\frac{23}{100}t_1 t_2 + \frac{1817}{1000}t_1 - \frac{23}{10}t_2 + \frac{1817}{100}, & t \in [0, 290) \times [6.9, 7.9], \\[4pt]
\frac{1}{30}t_1 t_2 - \frac{13}{100}t_1 + \frac{118}{3}t_2 - \frac{767}{5}, & t \in [290, 590) \times [3.9, 4.9), \\[4pt]
-\frac{1}{30}t_1 t_2 + \frac{79}{300}t_1 - \frac{178}{3}t_2 + \frac{7031}{15}, & t \in [290, 590) \times [6.9, 7.9), \\[4pt]
\frac{1}{30}t_1 t_2 - \frac{13}{100}t_1 + \frac{118}{3}t_2 - \frac{767}{5}, & t \in [590, 890) \times [3.9, 4.9), \\[4pt]
-\frac{1}{15}t_1 t_2 + \frac{79}{150}t_1 - \frac{119}{3}t_2 + \frac{9401}{30}, & t \in [590, 890) \times [6.9, 7.9), \\[4pt]
-\frac{23}{100}t_1 t_2 + \frac{897}{1000}t_1 + \frac{2737}{10}t_2 - \frac{106743}{100}, & t \in [890, 1190) \times [3.9, 4.9), \\[4pt]
-\frac{33}{1000}t_1 t_2 - \frac{683}{10000}t_1 + \frac{3927}{100}t_2 + \frac{81277}{1000}, & t \in [890, 1190) \times [4.9, 5.9), \\[4pt]
-\frac{67}{1000}t_1 t_2 + \frac{1323}{10000}t_1 + \frac{7973}{100}t_2 - \frac{157437}{1000}, & t \in [890, 1190) \times [5.9, 6.9), \\[4pt]
\frac{33}{100}t_1 t_2 - \frac{2607}{1000}t_1 - \frac{3927}{10}t_2 + \frac{310233}{100}, & t \in [890, 1190) \times [6.9, 7.9], \\[4pt]
0, & \text{otherwise.}
\end{cases}$$

Step 3 $l\big(x_1^{(1)}+\delta_1\big)(t_1) = \max\{1-\frac{1}{300}|t_1-310|,\,0\}$, $l\big(x_1^{(2)}+\delta_1\big)(t_1) = \max\{1-\frac{1}{300}|t_1-610|,\,0\}$, $l\big(x_1^{(3)}+\delta_1\big)(t_1) = \max\{1-\frac{1}{300}|t_1-910|,\,0\}$, $l\big(x_2^{(1)}+\delta_2\big)(t_2) = \max\{1-|t_2-5.1|,\,0\}$, $l\big(x_2^{(2)}+\delta_2\big)(t_2) = \max\{1-|t_2-6.1|,\,0\}$, $l\big(x_2^{(3)}+\delta_2\big)(t_2) = \max\{1-|t_2-7.1|,\,0\}$. Thus

$$\bar{f}(t) = \begin{cases}
\frac{1}{30}t_1 + 10t_2 - \frac{31}{3}, & t \in [310, 610) \times [5.1, 7.1), \\[4pt]
\frac{1}{3000}t_1 t_2 + \frac{949}{30000}t_1 + \frac{2939}{300}t_2 - \frac{27889}{3000}, & t \in [610, 910) \times [5.1, 6.1), \\[4pt]
\frac{33}{1000}t_1 t_2 - \frac{5029}{30000}t_1 - \frac{1013}{100}t_2 + \frac{336769}{3000}, & t \in [610, 910) \times [6.1, 7.1), \\[4pt]
\frac{17}{100}t_1 t_2 - \frac{697}{1000}t_1 - \frac{17}{10}t_2 + \frac{697}{100}, & t \in [0, 310) \times [4.1, 5.1), \\[4pt]
\frac{1}{30}t_1 t_2 - \frac{1}{3}t_2, & t \in [0, 310) \times [5.1, 7.1), \\[4pt]
-\frac{71}{300}t_1 t_2 + \frac{1917}{1000}t_1 + \frac{71}{30}t_2 - \frac{1917}{100}, & t \in [0, 310) \times [7.1, 8.1], \\[4pt]
\frac{1}{30}t_1 t_2 - \frac{41}{300}t_1 + \frac{122}{3}t_2 - \frac{2501}{15}, & t \in [310, 910) \times [4.1, 5.1), \\[4pt]
-\frac{1}{30}t_1 t_2 + \frac{27}{100}t_1 - \frac{182}{3}t_2 + \frac{2457}{5}, & t \in [310, 610) \times [7.1, 8.1], \\[4pt]
\frac{1}{30}t_1 t_2 - \frac{41}{300}t_1 + \frac{122}{3}t_2 - \frac{2501}{15}, & t \in [610, 910) \times [4.1, 5.1), \\[4pt]
-\frac{1}{15}t_1 t_2 + \frac{27}{50}t_1 - \frac{121}{3}t_2 + \frac{3267}{10}, & t \in [610, 910) \times [7.1, 8.1], \\[4pt]
-\frac{71}{300}t_1 t_2 + \frac{2911}{3000}t_1 + \frac{8591}{30}t_2 - \frac{352231}{300}, & t \in [910, 1210) \times [4.1, 5.1), \\[4pt]
-\frac{101}{3000}t_1 t_2 - \frac{1949}{30000}t_1 + \frac{12221}{300}t_2 + \frac{235829}{3000}, & t \in [910, 1210) \times [5.1, 6.1), \\[4pt]
-\frac{199}{3000}t_1 t_2 + \frac{1343}{10000}t_1 + \frac{24079}{300}t_2 - \frac{162503}{1000}, & t \in [910, 1210) \times [6.1, 7.1), \\[4pt]
\frac{101}{300}t_1 t_2 - \frac{2727}{1000}t_1 - \frac{12221}{30}t_2 + \frac{329967}{100}, & t \in [910, 1210) \times [7.1, 8.1], \\[4pt]
0, & \text{otherwise.}
\end{cases}$$

Table 7 Comparison with true sample

$(x_{1,i}, x_{2,i})$	$y_{i,j}$	$y_{i,j}^-$	$y_{i,j}^+$
$\langle 300 \pm 10^{-2}, 5 \pm 10^{-5}\rangle$	50	50.3329	44.3720
$\langle 300 \pm 10^{-2}, 6 \pm 10^{-5}\rangle$	60	60.3329	58.0021
$\langle 300 \pm 10^{-2}, 7 \pm 10^{-5}\rangle$	70	62.4004	67.6691
$\langle 600 \pm 10^{-2}, 5 \pm 10^{-5}\rangle$	60	60.3326	54.6009
$\langle 600 \pm 10^{-2}, 6 \pm 10^{-5}\rangle$	70	70.3632	69.6671
$\langle 600 \pm 10^{-2}, 7 \pm 10^{-5}\rangle$	80	71.7002	79.6671
$\langle 900 \pm 10^{-2}, 5 \pm 10^{-5}\rangle$	70	67.6592	63.6010
$\langle 900 \pm 10^{-2}, 6 \pm 10^{-5}\rangle$	80	78.2155	79.7541
$\langle 900 \pm 10^{-2}, 5 \pm 10^{-5}\rangle$	100	86.1339	98.3772

Step 4 $\check{f}(\mathbf{t}) = \lfloor \underline{f}(\dot{t}_1, \dot{t}_2), \bar{f}(\ddot{t}_1, \ddot{t}_2)\rfloor \ (\forall \mathbf{t} = (t_1, t_2) \in \mathbb{R}^2)$. Using the classical least square estimation method we can obtain an linear regression function[5] $r(\boldsymbol{\varepsilon}) = r(\varepsilon_1, \varepsilon_2)$ (it has no the constant term) based on the sample set

$$\Big\{ \overbrace{(10, 0.1, 1), \ldots, (10, 0.1, 1)}^{7\text{times}}, (10, 0.1, 1.1), (10, 0.1, 1) \Big\}.$$
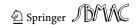
Calculations show that $r(\boldsymbol{\varepsilon}) = r(\varepsilon_1, \varepsilon_2) = 0.1011\varepsilon_1$. Therefore, $\tilde{f}(\mathbf{t}) = \lceil f_c(t_1^c, t_2^c), 0.1011\varepsilon_1\rceil = f_c(t_1^c, t_2^c) \pm 0.1011\varepsilon_1$. For an $\mathbf{s} = (\lceil 400, 10\rceil, \lceil 7, 0.1\rceil) \in \mathbb{R}$, as $\underline{f}(\dot{s}_1, \dot{s}_2) = \underline{f}(390, 6.9) = 72.\dot{3}$, $f_c(s_1^c, s_2^c) = f_c(400, 7) = 73.\dot{3}$, and $\bar{f}(\ddot{s}_1, \ddot{s}_2) = \bar{f}(410, 7.1) = 74.\dot{3}$, $\check{f}(\mathbf{s}) = \langle 72.\dot{3}, 74.\dot{3}\rangle$ and $\tilde{f}(\mathbf{s}) = \lceil 73.\dot{3}, 1.011\rceil = \langle 72.3223, 74.3443\rangle$.

Step 5 If the true sample set $S = \{(x_{1,i}, x_{2,j}, y_{i,j}) \mid i, j = 1, 2, 3\}$ consists of centers of interval numbers in $\mathbb{S}$, then the compute results and the true sample are list in Table 7, where $y_{i,j}^c = f_c(x_{1,i}, x_{2,j})$, $y_{i,j}^- = \underline{f}(x_{1,i} - 10^{-2}, x_{2,j} - 10^{-5})$, and $y_{i,j}^+ = \bar{f}(x_{1,i} + 10^{-2}, x_{2,j} + 10^{-5})$.

Example 3.6 Consider the special interval input–output sample set $\mathbb{S} = \big\{(300\pm10, 5\pm2, 50\pm 5), (300\pm10, 7\pm2, 70\pm5), (600\pm10, 6\pm2, 70\pm5), (900\pm10, 5\pm2, 70\pm5), (900\pm10, 7\pm2, 90\pm5)\big\}$ (from a function f) in Table 8 and its completion $\tilde{\mathbb{S}} = \{(x_{1,i}, x_{2,j}, y_{i,j}) \mid i, j = 1, 2, 3\}$ (from the corresponding function $\acute{f}$) in Table 9, where the unknown sample are determined using an arithmetic average operator (of course, we can also use an appropriate aggregation operator, a t-norm, a t-conorm, etc.): $* = 0.5(\lceil 50, 1\rceil \oplus \lceil 70, 1\rceil) = \lceil 60, 1\rceil$, $** = 0.5(\lceil 50, 1\rceil \oplus \lceil 70, 1\rceil) = \lceil 60, 1\rceil$, $\star = 0.5(\lceil 70, 1\rceil \oplus \lceil 90, 1\rceil) = \lceil 80, 1\rceil$, $\star\star = 0.5(\lceil 70, 1\rceil \oplus \lceil 90, 1\rceil) = \lceil 80, 1\rceil$. Let $f_c, \underline{f}, \bar{f}, \check{f}$, and $\tilde{f}$ be as in Example 3.5(2), then

$$f_c(\mathbf{t}) = \begin{cases} \frac{1}{30}t_1 + 10t_2 - 10, & \mathbf{t} \in [600, 900) \times [5, 7), \\ -\frac{1}{30}t_1 t_2 + \frac{4}{15}t_1 - 60t_2 + 480, & \mathbf{t} \in [600, 900) \times [7, 8], \\ -\frac{1}{30}t_1 t_2 - \frac{1}{15}t_1 + 40t_2 + 80, & \mathbf{t} \in [900, 1200) \times [5, 7), \\ \frac{3}{10}t_1 t_2 - \frac{12}{5}t_1 - 360t_2 + 2880, & \mathbf{t} \in [900, 1200) \times [7, 8], \\ f_c(\mathbf{t}), & \text{otherwise,} \end{cases}$$

---

[5] Actually, we obtain a plane determined by three points $\{(0, 0, 0), (10, 0.1, 1), (10, 0.1, 1.1)\}$.

Table 8 Special interval input–output sample

$\mathbb{X}_1$ \ $\mathbb{X}_2$	$\lceil 5, 0.1 \rceil$	$\lceil 6, 0.1 \rceil$	$\lceil 7, 0.1 \rceil$
$\lceil 300, 10 \rceil$	$\lceil 50, 1 \rceil$	$*$	$\lceil 70, 1 \rceil$
$\lceil 600, 10 \rceil$	$**$	$\lceil 70, 1 \rceil$	$\star$
$\lceil 900, 10 \rceil$	$\lceil 70, 1 \rceil$	$\star\star$	$\lceil 90, 1 \rceil$

Table 9 Completion of Table 8

$\mathbb{X}_1$ \ $\mathbb{X}_2$	$\lceil 5, 0.1 \rceil$	$\lceil 6, 0.1 \rceil$	$\lceil 7, 0.1 \rceil$
$\lceil 300, 10 \rceil$	$\lceil 50, 1 \rceil$	$\lceil 60, 1 \rceil$	$\lceil 70, 1 \rceil$
$\lceil 600, 10 \rceil$	$\lceil 60, 1 \rceil$	$\lceil 70, 1 \rceil$	$\lceil 80, 1 \rceil$
$\lceil 900, 10 \rceil$	$\lceil 70, 1 \rceil$	$\lceil 80, 1 \rceil$	$\lceil 90, 1 \rceil$

$$
\underline{f}(t) = \begin{cases}
\frac{1}{30}t_1 + 10t_2 - \frac{29}{3}, & t \in [590, 890) \times [4.9, 6.9), \\
-\frac{1}{30}t_1 t_2 + \frac{79}{30}t_1 - \frac{178}{3}t_2 + \frac{7031}{15}, & t \in [590, 890) \times [6.9, 7.9), \\
-\frac{1}{30}t_1 t_2 - \frac{1}{15}t_1 + \frac{119}{3}t_2 + \frac{238}{3}, & t \in [890, 1190) \times [4.9, 6.9), \\
\frac{89}{300}t_1 t_2 - \frac{7031}{3000}t_1 - \frac{10591}{30}t_2 + \frac{836689}{300}, & t \in [890, 1190) \times [6.9, 7.9), \\
f(t), & \text{otherwise},
\end{cases}
$$

$$
\bar{f}(t) = \begin{cases}
\frac{1}{30}t_1 + 10t_2 - \frac{31}{3}, & t \in [610, 910) \times [5.1, 7.1), \\
-\frac{1}{30}t_1 t_2 + \frac{27}{300}t_1 - \frac{182}{3}t_2 + \frac{2457}{5}, & t \in [610, 910) \times [7.1, 8.1), \\
-\frac{1}{30}t_1 t_2 - \frac{1}{15}t_1 + \frac{121}{3}t_2 + \frac{242}{3}, & t \in [910, 1210) \times [5.1, 7.1), \\
\frac{91}{300}t_1 t_2 - \frac{2457}{1000}t_1 - \frac{11011}{30}t_2 + \frac{297297}{100}, & t \in [910, 1200) \times [7.1, 8.1), \\
f(t), & \text{otherwise}.
\end{cases}
$$

By Algorithm 3.2, $\check{f}(\mathbf{t}) = \check{f}(\mathbb{t}_1, \mathbb{t}_2) = \lfloor \underline{f}(\check{\imath}_1, \check{\imath}_2), \bar{f}(\ddot{\imath}_1, \ddot{\imath}_2) \rfloor$ and $\tilde{f}(\mathbf{t}) = f_c(\mathbb{t}_1^c, \mathbb{t}_2^c) \pm 1$ $(\mathbf{t} \in \mathbb{R}^2)$.

4 Applications

Example 4.1 Now we give detailed solutions to problem (I) in Sect. 1.

(1) To determine the grade by using the linear regression model presented in Sect. 2. Firstly, compute the 3-variable linear regression functions. $\underline{f}(t) = -1.5253 - 0.0141t_1 + 1.4017t_2 + 3.5535t_3$, $f_c(t) = -1.514 - 0.0114t_1 + 0.1363t_2 + 4.6991t_3$, $\bar{f}(t) = -0.8411 - 0.0259t_1 + 1.893t_2 + 1.917t_3$, and $r(\boldsymbol{\delta}) = 0.0075\delta_1 + 0.1585\delta_2 - 0.169\delta_3$. By Algorithm 2.3, $\hat{f}(\mathbf{t}) = \langle -1.5253, -0.8411 \rangle \oplus \langle -0.0266, -0.0141 \rangle t_1 \oplus \langle 1.4017, 1.893 \rangle \mathbb{t}_2 \oplus \langle 1.917, 3.5535 \rangle \mathbb{t}_3$, $\tilde{f}(\mathbf{t}) = \lceil -1.514 - 0.0114\mathbb{t}_1^c + 0.1363\mathbb{t}_2^c + 4.6991\mathbb{t}_3^c, 0.0075\delta_1 + 0.1586\delta_2 - 0.169\delta_3 \rceil$ and $\check{f}(\mathbf{t}) = \lfloor \underline{f}(\check{\imath}_1, \check{\imath}_2, \check{\imath}_3), \bar{f}(\ddot{\imath}_1, \ddot{\imath}_2, \ddot{\imath}_3) \rfloor$ $(\forall \mathbf{t} = (\mathbb{t}_1, \mathbb{t}_2, \mathbb{t}_3) \in \mathbb{R}^3)$. Secondly, $\hat{f}(\mathbf{t}_0) = \langle 1.4635, 4.7158 \rangle$, $\check{f}(\mathbf{t}_0) = \lfloor 2.983, 3.1963 \rfloor$, and $\tilde{f}(\mathbf{t}_0) = \lceil 2.0694, 0.0551 \rceil$. Finally, determine the grade of $\mathbf{t}_0$ by experts based on the computation results.

(2) To determine the grade by using the linear-like interpolation model presented in Sect. 3. Firstly, supplement Table 9 (which we will call Table 10). In Table 10, only 2^3 samples are need because, for the computation of each of $\{\check{f}(\mathbf{t}_0), \tilde{f}(\mathbf{t}_0)\}$, only the corresponding

2^3 basis functions relaying the sample will be used. Secondly, compute $\check{f}(\mathbf{t}_0)$ or $\tilde{f}(\mathbf{t}_0)$ (and thus determine the grade). We omit the computation course (including Table 10) here because it is tedious (unless a matched program is given).

(3) To determine the grade by using some easy-to-think-of methods (which can also be discovered in plausible manner). For example, take three Astragali Radix $\mathbf{t}_1 = (\lceil 36.5, 9 \rceil, \lceil 1.3, 0.2 \rceil, \lceil 0.9, 0.1 \rceil)$, $\mathbf{t}_2 = (\lceil 36.7, 10.4 \rceil, \lceil 1.7, 0.2 \rceil, \lceil 1.2, 0.2 \rceil)$, and $\mathbf{t}_3 = (\lceil 32.4, 5.5 \rceil, \lceil 1, 0.1 \rceil, \lceil 0.8, 0.1 \rceil)$ (such that each is much similar to $\mathbf{t}_0$ in one aspect) in Table 9 first. Then compute

$$
\begin{aligned}
\mathrm{d}_1 &= d(\mathbf{t}_0, \mathbf{t}_1) = ||(35.8 - 5.4) - (36.5 - 9)|, |(35.8 + 5.4) - (36.5 + 9)|| \\
&\quad \oplus ||(1.7 - 0.27) - (1.3 - 0.2)|, |(1.7 + 0.27) - (1.3 + 0.2)|| \\
&\quad \oplus ||(0.8 - 0.1) - (0.9 - 0.1)|, |(0.8 + 0.1) - (0.9 + 0.1)|| \\
&= \langle 2.9, 4.3 \rangle \oplus \langle 0.33, 0.47 \rangle \oplus \langle 0.1, 0.1 \rangle = \langle 3.33, 4.87 \rangle, \\
\mathrm{d}_2 &= d(\mathbf{t}_0, \mathbf{t}_2) = ||(35.8 - 5.4) - (36.7 - 10.4)|, |(35.8 + 5.4) - (36.7 + 10.4)|| \\
&\quad \oplus ||(1.7 - 0.27) - (1.7 - 0.2)|, |(1.7 + 0.27) - (1.7 + 0.2)|| \\
&\quad \oplus ||(0.8 - 0.1) - (1.2 - 0.2)|, |(0.8 + 0.1) - (1.2 + 0.2)|| \\
&= \langle 4.1, 5.9 \rangle \oplus \langle 0.07, 0.07 \rangle \oplus \langle 0.3, 0.5 \rangle = \langle 4.47, 6.47 \rangle, \\
\mathrm{d}_3 &= d(\mathbf{t}_0, \mathbf{t}_3) = ||(35.8 - 5.4) - (32.4 - 5.5)|, |(35.8 + 5.4) - (32.4 + 5.5)|| \\
&\quad \oplus ||(1.7 - 0.27) - (1 - 0.1)|, |(1.7 + 0.27) - (1 + 0.1)|| \\
&\quad \oplus ||(0.8 - 0.1) - (0.8 - 0.1)|, |(0.8 + 0.1) - (0.8 + 0.1)|| \\
&= \langle 3.3, 3.5 \rangle \oplus \langle 0.53, 0.87 \rangle \oplus \langle 0.1, 0.1 \rangle = \langle 3.83, 4.37 \rangle.
\end{aligned}
$$

Finally, as $\mathrm{d}_1 \leq \mathrm{d}_3 \leq \mathrm{d}_2$, $\mathbf{t}_0$ is most similar to $\mathbf{t}_1$, and thus the grade of $\mathbf{t}_0$ can be judged as 4.

Example 4.2 In a practical problem with an interval input–output sample and a given interval input $\mathbf{t}$, how to obtain the interval output? the suggested strategies are as follows:
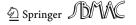
(1) If the sample is managed so carefully that are as in Examples 2.6 and 3.1 (i.e., the radii are the same in the same variable data), then we can compute the output of $\mathbf{t}$ just using one of the five kinds of functions (i.e., the linear regression functions or the line-like interpolation functions given in this paper, see Corollary 2.7 and Examples 3.5 and 3.6 for details); of course, we can also take the output of $\mathbf{t}$ to be an aggregation (for example, a simple weighted average) of these outputs.

(2) In other cases, we can compute the output of $\mathbf{t}$ using one of the three kinds of linear regression functions given in this paper, see Example 4.1(1) for details); we can also compute the output of $\mathbf{t}$ as in the last part of Example 4.1(2). Of course, we can also take the output of $\mathbf{t}$ to be an appropriate aggregation of these outputs.

Example 4.3 Consider the control problem $\frac{d\mathbf{x}}{dt} = (f_1(\mathbf{x}), f_2(\mathbf{x}))^T + \mathbf{u}(t)$ in (II) of Sect. 1. Assume the sample obtained on $y_1 = f_1(\mathbf{x})$ is as the Table 4 in Example 2.6, and the sample obtained on $y_2 = f_2(\mathbf{x})$ is as the Table 7 in Example 3.6. Replacing $(f_1(\mathbf{x}), f_2(\mathbf{x}))$ by $(f_{1,c}(\mathbf{x}), f_{2,c}(\mathbf{x}))$ (resp, $(\underline{f}_1(\mathbf{x}), \underline{f}_2(\mathbf{x}))$, $(\bar{f}_1(\mathbf{x}), \bar{f}_2(\mathbf{x}))$) (see Example 3.5(2)), we obtain the following three systems:

$$
\frac{d\mathbf{x}}{dt} = \begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} f_{1,c}(\mathbf{x}) \\ f_{2,c}(\mathbf{x}) \end{pmatrix} + \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} \tag{3}
$$

$$
\frac{d\mathbf{x}}{dt} = \begin{pmatrix} \underline{f}_1(\mathbf{x}) \\ \underline{f}_2(\mathbf{x}) \end{pmatrix} + \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} \tag{4}
$$

$$
\frac{d\mathbf{x}}{dt} \begin{pmatrix} \bar{f}_1(\mathbf{x}) \\ \bar{f}_2(\mathbf{x}) \end{pmatrix} + \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} \tag{5}
$$

We declare that system (3) is asymptotically stable if we take $(u_1(t), u_2(t))^{\mathrm{T}} = \big(-k_1 x_1(t) - f_{1,c}(\boldsymbol{x}(t)), -k_2 x_2(t) - f_{2,c}(\boldsymbol{x}(t))\big)^{\mathrm{T}}$ $(k_1 > 0, k_2 > 0)$. Indeed, define the Lyapunov function as $2V(\boldsymbol{x}) = x_1^2 + x_2^2$. Then $\frac{\mathrm{d}V}{\mathrm{d}t} = x_1 \frac{\mathrm{d}x_1}{\mathrm{d}t} + x_2 \frac{\mathrm{d}x_2}{\mathrm{d}t} = x_1[f_{1,c}(\boldsymbol{x}) + u_1] + x_2[f_{2,c}(\boldsymbol{x}) + u_2] = -k_1 x_1^2 - k_2 x_2^2 \leq 0$, which implies that system (3) is asymptotically stable. Similarly, systems (4) and (5) are also asymptotically stable. The idea and method used here can be used immediately to control chaotic fractional-order neural networks (Han et al. 2020).

5 Conclusions

Intelligent doctors, plus the Tradition Chinese Medicine (TCM for short), can resist having diarrhea, cold, being ill with a fever, and tonsillitis completely (even coronavirus disease-2019, COVID-19 for short, in a degree actually) with a smaller side effect. The present paper confirms that interval input–output samples (which are major samples or data in TCM all the time) can always be disposed towards satisfactory with the help of models (stemmed from the classical ones) and practitioners' cooperation. Given an n-variable interval input-1-variable interval output sample set $\mathbb{S}$, we have illustrated in much detail how to establish, in Polya's discovering pattern (mainly using specialization and analogous towards exploring solutions to problems), three linear regression models and two linear-like interpolation models relaying on $\mathbb{S}$. Each model is proved to be a universal approximator to the corresponding model based on the true samples under some easy-to-be-satisfied conditions. As the computations only involve centers, left endpoints, or right endpoints of interval numbers in the sample, off-the-shelf software can be utilized. Practitioners can optimize these models in the following ways:

(1) Collect directly or obtain such kind of interval samples that have the same radius r and r is as small as possible [this can be realized by practitioners themselves directly or by experts indirectly using three-way decision theory (Yao 2017)].
(2) Take $g_1(\mathbf{x}) = \mathrm{w}_1 \hat{f}(\mathbf{x}) \oplus \mathrm{w}_2 \check{f}(\mathbf{x}) \oplus \mathrm{w}_3 \tilde{f}(\mathbf{x})$ to be the new model in the case of linear regression model, where $\{\mathrm{w}_1^c, \mathrm{w}_2^c, \mathrm{w}_3^c\} \subseteq (0, 1)$ satisfying $\mathrm{w}_1^c + \mathrm{w}_2^c + \mathrm{w}_3^c \approx 1$ may be determined by experts; take $g_2(\mathbf{x}) = \mathrm{w}_1 \check{f}(\mathbf{x}) \oplus \mathrm{w}_2 \tilde{f}(\mathbf{x})$ to be the new model in the case of linear-like interpolation model, where $\{\mathrm{w}_1^c, \mathrm{w}_2^c\} \subseteq (0, 1)$ satisfying $\mathrm{w}_1^c + \mathrm{w}_2^c \approx 1$ may be determined by experts.

Notice that the models disposing of interval samples in this paper are stemmed from the classical linear regression model or the classical linear interpolation model, thus practitioners can also make out similarly other new models disposing of interval samples based on other classical models disposing of real number samples. Theoretically, these work can be generalized further (e.g., replace the operations by demand-oriented aggregation-like operation, t-norm-like operation, or t-conorm-like operation to be coined by experts), if necessary. Finally, some new universal approximators can also be discovered and used to control chaotic fractional-order neural networks (Han et al. 2020) based on this research (which will be our future work).

Availability of data and materials Not applicable.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Human and animal rights statement This article does not contain any studies with human participants performed by any of the authors.

References

Beliakov G, Pradera A, Calvo T (2007) Aggregation functions: a guide for practitioners. Springer, Berlin

Boukezzoula R, Galichet S, Bisserier A (2011) A Midpoint-Radius approach to regression with interval data. Int J Approx Reason 52:1257–1271

Boukezzoula R, Galichet S, Coquin D (2018) From fuzzy regression to gradual regression: interval-based analysis and extensions. Inf Sci 441:18–40

Boukezzoula R, Coquin D, Jacq K (2020) A possibilistic regression based on gradual interval B-splines: application for hyperspectral imaging lake sediments. Inf Sci 510:183–204

Chuang CC (2008) Extended support vector interval regression networks for interval input–output data. Inf Sci 178:871–891

Chukhrova N, Johannssen A (2019) Fuzzy regression analysis: systematic review and bibliography. Appl Soft Comput 84:105708

Cubillo S, Hernández P, Torres-Blanc C (2015) Examples of aggregation operators on membership degrees of type-2 fuzzy sets. In: Proceedings of IFSA-EUSFLAT 2015, Gijón, Spain, pp 719–726

Deschrijver G, Kerre E (2008) Aggregation operation in interval-valued fuzzy and Atanassov's intuitionistic fuzzy set theory. In: Bustince H et al (eds) Fuzzy sets and their extensions: representation, aggregation and models. Springer, Berlin, pp 183–203

Han Z, Li SG, Liu H (2020) Composite learning sliding mode synchronization of chaotic fractional-order neural networks. J Adv Res 25: 87–96. https://doi.org/10.1016/j.jare.2020.04.006

Hladíka M, Černý M (2012) Interval regression by tolerance analysis approach. Fuzzy Sets Syst 193:85–107

Hu BQ, Wang S (2006) A novel approach in uncertain programming part I: new arithmetic and order relation for interval numbers. J Ind Manag Optim 2:351–371

Hwang C, Hong HD, Seok KH (2006) Support vector interval regression machine for crisp input and output data. Fuzzy Sets Syst 157:1114–1125

Inuiguchi M, Mizoshita F (2012) Qualitative and quantitative data envelopment analysis with interval data. Ann Oper Res 195:189–220

Ishibuchi H, Tanaka H (1990) Several formulations of interval regression analysis. In: Proceedings of Sino-Japan joint meeting on fuzzy sets and systems, Beijing, China, (B2-2), pp 1–4

Jeng JT, Chuang CC, Su SF (2003) Support vector interval regression networks for interval regression analysis. Fuzzy Sets Syst 138:283–300

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

Moore RE (1966) Interval analysis. Prentice-Hall, Hoboken

Polya G (1954) Mathematics and plausible reasoning, Vol. I, induction and analogy in mathematics. Princeton University Press, Hoboken

Tanaka H, Lee H (1998) Interval regression analysis by quadratic programming approach. IEEE Trans Fuzzy Syst 6:473–481

Xu ZS, Yager RR (2006) Some geometric aggregation operators based on intuitionistic fuzzy sets. Int J Gen Syst 35:417–433

Yao YY (2017) Interval sets and three-way concept analysis in incomplete contexts. Int J Mach Learn Cybern 8:3–20

Ying H (2015) A sufficient condition on a general class of interval type-2 Takagi-Sugeno fuzzy systems with linear rule consequent as universal approximators. J Intell Fuzzy Syst 29:1219–1227

Zhang R, Cao QW, Li K, Qin XM (2020) Analysis of correlation between commercial traits and chemical characteristics and absolute growth years of Astragali Radix. Chin Tradit Herb Drugs 51:451–460 in Chinese