2

Regression and Self-regression Models with Fuzzy Coefficients

As the phenomenon in the world is complicated, at the time of carrying on statistic forecast, we will usually meet a type of fuzzy number that points are constant, the circle is changed, and vice versa. For such a case, an analytical problem needs considering in regression and self-regression under a fuzzy environment. In 1980, in this aspect, a regression analysis formulation was already developed according to a possible linear system [TUA80]. Hereafter the regression analysis was variously formed by means of fuzzy data analysis, and carried in extensive application [TUA82]. In 1989, based on the theory of Zadeh fuzzy sets [Zad65a], self-regression forecast model with fuzzy coefficients was advanced [cao89b][cao90].

This chapter introduces a regression and self-regression model containing $(\cdot, c)$ fuzzy coefficients, flat fuzzy coefficients as well as triangular fuzzy coefficients, concludes the regression analysis as a linear programming.

2.1 Regression Model with Fuzzy Coefficients

2.1.1 Introduction

Suppose a classical linear regression model to be

$$Y = A_1x_1 + A_2x_2 + \ldots + A_nx_n + \varepsilon,$$

where Y is a correlated variable, and x_i, A_i an independent variable and parameter, respectively, and ε an error.

Because problems within realistic world all contain a great quantity of fuzzyness, this section will consider a fuzzy model as follows:

$$\tilde{Y} = \tilde{A}_1x_1 + \tilde{A}_2x_2 + \cdots + \tilde{A}_nx_n + \varepsilon, \tag{2.1.1}$$

where $\tilde{Y}$ and $\tilde{A}_j (1 \leqslant j \leqslant n)$ are $(\cdot, c)$ fuzzy correlated variables and parameter; $x = (x_1, x_2, \ldots, x_n)^T$ is an independent variable vector, and independent variable $x_j (1 \leqslant j \leqslant n)$ in i-period changed backward, with ε being an error. We call (2.1.1) a regression model with fuzzy coefficients.

B.-Y. Cao: Optimal Models & Meth. with Fuzzy Quantities, STUDFUZZ 248, pp. 33–62.
springerlink.com

2.1.2 Definitions and Concepts of Fuzzy Parameters

Definition 2.1.1. Suppose $\mathscr{F}(\mathscr{R})$ is a fuzzy set, and $\tilde{A}_j \in \mathscr{F}(\mathscr{R})(j = 1, 2)$ denotes fuzzy parameters with its membership function (1.5.3). [TOA73] [TUA82]

Definition 2.1.2. Fuzzy number $\tilde{A}$ is a convex normalized fuzzy subset of a real axis satisfying

(i) $\exists x_0 \in \mathscr{R}$ and $\mu_{\tilde{A}}(x_0) = 0$;

(ii) $\tilde{A}$ is a piecewise continuous function.

The $\alpha-$cut set of $\tilde{A}$ is set $\tilde{A}_\alpha = \{x \in \mathscr{R}, \mu_{\tilde{A}}(x) \geqslant \alpha\}$, where $\alpha \in [0, 1]$.

Definition 2.1.3. If $\forall x, y, z \in \mathscr{R}$, and $x \leqslant y \leqslant z$, we have $\mu_{\tilde{A}}(y) \geqslant \mu_{\tilde{A}}(x) \wedge \mu_{\tilde{A}}(z)$, calling $\tilde{A}$ a normal fuzzy number.

As for relevant (a, c), the definitions and properties of fuzzy numbers, refer to the Ref. [TA84] and [Wat87].

Extension Principle: Suppose $\tilde{A}_1, \cdots, \tilde{A}_n$ to be (a, c) fuzzy numbers, mapping $f : \mathscr{R} \to \mathscr{R}$, i.e.,

$$f(x_1, x_2, \cdots, x_n) = x_1 * x_2 * \cdots * x_n.$$

Expand this operation '$*$' to fuzzy numbers, then rule

$$\begin{aligned} f(\tilde{A}_1, \tilde{A}_2 \cdots, \tilde{A}_n) &= \tilde{A}_1 * \tilde{A}_2 * \cdots * \tilde{A}_n \\ &= \int_{X_1 * \cdots * X_n} \frac{\min\{\tilde{A}_1(x_1), \cdots, \tilde{A}_n(x_n)\}}{f(x_1, \cdots, x_n)}, \end{aligned}$$

its membership function meaning

$$\mu_{f(\tilde{A}_1, \cdots, \tilde{A}_n)}(y) = \sup \min_{x_1, \cdots, x_n \in f^{-1}(y)} \{\mu_{\tilde{A}_1}(x_1), \cdots, \mu_{\tilde{A}_n}(x_n)\}.$$

By using α-cut sets, if $\tilde{B} = f(\tilde{A}_1, \cdots, \tilde{A}_n)$ means an image in $\tilde{A}_1, \cdots, \tilde{A}_n$, then

$$[f(\tilde{A}_1, \cdots, \tilde{A}_n)]_\alpha = f(A_{1\alpha}, \cdots, A_{n\alpha}) \Longleftrightarrow \forall y \in Y, \exists \bar{x}_1, \cdots, \bar{x}_n,$$

such that

$$\mu_{\tilde{B}}(y) = \mu_{(\tilde{A}_1 * \tilde{A}_2 * \cdots * \tilde{A}_n)}(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n).$$

Definition 2.1.4. Assume that two sets X and Y, $f : X \to Y$ denotes a function $Y = f(x, a)$, $f : X \to \mathscr{F}(y)$ denotes a fuzzy function $\tilde{Y} = f(x, \tilde{A})$, then the membership function of fuzzy set $\tilde{Y}$ denotes

$$\mu_{\tilde{Y}}(y) = \begin{cases} \max\limits_{\{a|y=f(x,a)\}} \mu_{\tilde{A}}(a), & \text{when} \quad \{a|y = f(x, a)\} \neq \phi, \\ 0, & \text{otherwise}, \end{cases}$$

where $x \in \mathscr{R}$, a parameter on the product space of $a = a_1 \times a_2 \times \cdots \times a_n$, where n is the number of independent variables. $\tilde{A}$ a fuzzy set, $\tilde{Y}$ a mapping of x in $\tilde{A}$, and $\mathscr{F}(y)$ a fuzzy-valued set.

Definition 2.1.5. The fuzzy parameter $\tilde{A}$ of fuzzy linear regression is defined from Cartesian product space $\mathscr{R}^n$ on the Cartesian product sets $\tilde{A} = \tilde{A}_1 \times \tilde{A}_2 \times \cdots \times \tilde{A}_n$, such as Figure 2.1.1 shows,

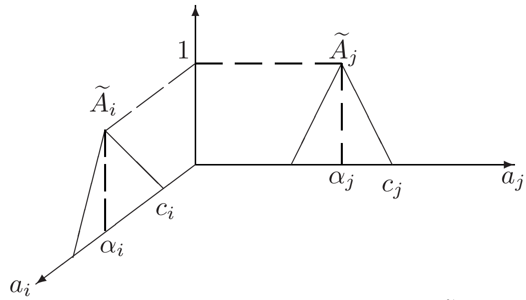

Fig. 2.1.1. Fuzzy Parameter $\tilde{A}$

its membership function is a triangle type, i.e.,

$$\mu_{\tilde{A}}(a) = \min_j \ \mu_{\tilde{A}_j}(a_j),$$

$$\mu_{\tilde{A}_j}(a_j) = \begin{cases} 1 - \dfrac{|\alpha_j - a_j|}{c_j}, & \text{when} \ \ \alpha_j - c_j \leqslant a_j \leqslant \alpha_j + c_j, \\ 0, & \text{otherwise,} \end{cases}$$

where $c_j > 0 (j = 1, 2, 3, \cdots, n)$.

Definition 2.1.6. Fuzzy regression parameter $\tilde{A}$ defined on the vector space $\mathscr{R}^n$ is written as a vector form

$$\tilde{A} = (\alpha, c), \alpha = (\alpha_1, \cdots, \alpha_n)^T, c = (c_1, \cdots, c_n)^T,$$

"T" means a transporting sign, α, c the center and the shape in $\tilde{A}$, respectively, and $\tilde{A}$ means "approximately A".

Suppose that $\tilde{Y}$ and $\tilde{A}$ are all convex, normalized fuzzy functions and fuzzy numbers below.

2.1.3 Establishment of Linear Regression Model

Suppose the linear regression model to be

$$\tilde{Y} = \tilde{A}_1 x_1 + \tilde{A}_2 x_2 + \cdots + \tilde{A}_n x_n = \tilde{A}x = (a^T x, c^T x), \tag{2.1.2}$$

where $\tilde{A}_j (j = 1, \cdots, n)$ is a waiting parameter.

Proposition 2.1.1. *The membership function of (2.1.2) is*

$$\mu_{\tilde{Y}}(y) = \begin{cases} 1 - \dfrac{|y-\alpha^T x|}{c^T|x|}, & x \neq 0, \\ 1, & x = 0, y = 0, \\ 0, & x = 0, y \neq 0, \end{cases}$$

where $|x| = (|x_1|, |x_2|, \cdots, |x_n|)^T$ *and* $\mu_{\tilde{Y}}(y) = 0$ *when* $c^T|x| \leqslant |y - \alpha^T x|$.

In fact, according to Definition 2.1.4 and stipulate, then

$$\begin{aligned} \mu_{\tilde{Y}}(y) &= \begin{cases} \bigvee\limits_{\{\alpha|\alpha^T x = y\}} \mu_{\tilde{A}}(\alpha), \ \{\alpha|\alpha^T x = y\} \neq \phi \\ 1, \qquad x = 0, y = 0 \\ 0, \qquad x = 0, y \neq 0 \end{cases} \\ &= \begin{cases} \bigvee\limits_{\{\alpha|\alpha^T x = y\}} \{\bigwedge\limits_{j=1}^{n} \mu_{\tilde{A}_p}(\alpha_j)\}, \ \{\alpha|\alpha^T x = y\} \neq \phi \\ 1, \qquad x = 0, y = 0 \\ 0, \qquad x = 0, y \neq 0 \end{cases} \\ &= \begin{cases} \bigvee\limits_{\{\alpha|\alpha^T x = y\}} \{\bigwedge\limits_{j=1}^{n} (1 - \dfrac{|\alpha_j - a_j|}{c_j})\}, \ \{\alpha|\alpha^T x = y\} \neq \phi \\ 1, \qquad x = 0, y = 0 \\ 0, \qquad x = 0, y \neq 0 \end{cases} \\ &= \begin{cases} 1 - \dfrac{|y-\alpha^T x|}{c^T|x|}, \ x \neq 0, \\ 1, \qquad x = 0, y = 0, \\ 0, \qquad x = 0, y \neq 0. \end{cases} \end{aligned}$$

Here, when $c^T|x| < |y - \alpha^T x|$, the above means a deviation between the calculation value of y and the actual value is bigger than a fuzzy shape in calculation values, then $\mu_{\tilde{Y}}(y) = 0$.

Take a sample $(y_i; x_{i1}, x_{i2}, \cdots, x_{in})$ for example, where the capacity is n, and $y_i = \alpha^T x_i (i = 1, 2, \cdots, n)$ is an observed value, $\hat{y}_i$ an estimation value, both of deviations are $\varepsilon_i = y_i - \hat{y}_i$, then $\tilde{Y} = (y, \varepsilon)$ (with correlated variable y being for center, the deviation being ε a shape) is a fuzzy correlated variable, and $\varepsilon = 0$ is non-fuzzy situation.

We aim at assuring fuzzy parameters according to observation value $\hat{\tilde{A}}_j (j = 1, 2, \cdots, n)$.

But adoption of classical least square method will meet trouble of whether $\hat{\tilde{A}}_j (j = 1, 2, \cdots, n)$ is differentiable. Hence, we determine $\hat{\tilde{A}}_j (j = 1, 2, \cdots, n)$ by use of methods below.

In order to measure the degree of fitting $\bar{h}$ between the observed data and the estimated one, a decision maker can choose a threshold value H. Here, H is selected by a person in experts' experience. The selection of H affects the width in fuzzy parameter c_j.

If compute $\max \bar{h} \geqslant H$, such that

$$\hat{Y}_i^H = \{y | \mu_{\hat{Y}_i}(y) \geqslant H\}, \tag{2.1.3}$$

then $\bar{h}$ is an optimal estimation of correlated variables in (2.1.2). The index of approximately degree in $\bar{h}$ shows as Figure 2.1.2:

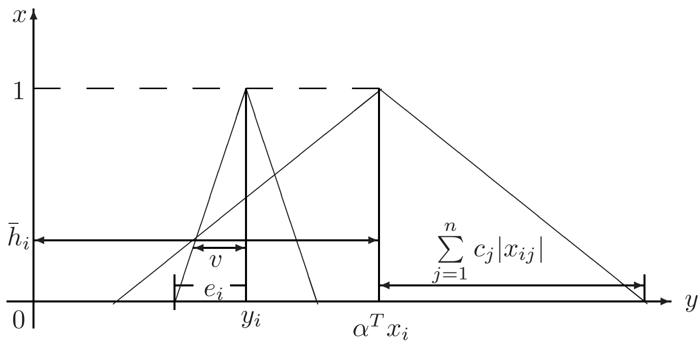

Fig. 2.1.2. The Index of Approximately Degree in $\bar{h}$

Theorem 2.1.1. *Assume that a fuzzy linear regression model as (2.1.2), then*

$$\max \bar{h} \geqslant H \Longleftrightarrow$$
$$\begin{cases} \alpha^T x_i + (1-H)\sum_{j=1}^{n} c_j |x_{ij}| \geqslant y_i + (1-H)\varepsilon_i, \\ -\alpha^T x_i + (1-H)\sum_{j=1}^{n} c_j |x_{ij}| \geqslant -y_i + (1-H)\varepsilon_i, \end{cases} \quad (i = 1, 2, \cdots, N).$$

Proof: Shown as Figure 2.1.2, $\bar{h}$ is derived as follow.

By using the similarity of the right triangles, then

$$\frac{v}{\varepsilon_i} = \frac{1-\overline{h}}{1}, \ v = \varepsilon_i(1-\overline{h}),$$

$$k = v + |y_i - \alpha^T x_i|, \ k = |y_i - \alpha^T x_i| + \varepsilon_i(1-\overline{h}).$$

Again by using the similarity of the right triangles, hence

$$\begin{aligned} \frac{1-\overline{h}}{1} &= \frac{k}{\sum_{j=1}^{n} c_j |x_{ij}|}, \\ \frac{1-\overline{h}}{1} &= \frac{|y_i - \alpha^T x_i| + \varepsilon_i(1-\overline{h})}{\sum_{j=1}^{n} c_j |x_{ij}|}. \end{aligned} \tag{2.1.4}$$

Find equation (2.1.4), then

$$1-\overline{h}=\frac{|y_i-\alpha^T x_i|}{\sum\limits_{j=1}^{n} c_j|x_{ij}|-\varepsilon_i},$$

i.e.,

$$\bar{h}_i=1-\frac{|y_i-\alpha^T x_i|}{\sum\limits_{j=1}^{n} c_j|x_{ij}|-\varepsilon_i}.$$

From (2.1.4), at $y_i-\alpha^T x_i\leqslant 0$, then

$$-\alpha^T x_i+(1-H)\sum_{j=1}^{n} c_j|x_{ij}|\geqslant -y_i+(1-H)\varepsilon_i,$$

at $y_i-\alpha^T x_i\geqslant 0$, we can get the same truth,

$$\alpha^T x_i+(1-H)\sum_{j=1}^{n} c_j|x_{ij}|\geqslant y_i+(1-H)\varepsilon_i(i=1,2,\cdots,N).$$

Combining two kinds of situations above, the theorem can be certificated.

Definition 2.1.7. The vagueness of the fuzzy linear model is denoted by $J(c)=\sum\limits_{j=1}^{n} c_j|x_{ij}|$, where x_{ij} is an observation datum, c_j a width in $\tilde{A}_j$.

Therefore, fuzzy parameter $\tilde{A}_j(j=1,\cdots,n)$ certainly is concluded to computation of an optimal solution $\hat{\tilde{A}}_j=(\alpha_j,c_j)$ in the following linear programming with parameter variables

$$\begin{aligned}
&\min\quad J(c)=\sum_{j=1}^{n} c_j|x_{ij}|\\
&\text{s.t.}\ \ \alpha^T x_i+(1-H)\sum_{j=1}^{n} c_j|x_{ij}|\geqslant y_i+(1-H)\varepsilon_i,\\
&\qquad -\alpha^T x_i+(1-H)\sum_{j=1}^{n} c_j|x_{ij}|\geqslant -y_i+(1-H)\varepsilon_i,\\
&\qquad c\geqslant 0, H\in[0,1],(i=1,\cdots,N).
\end{aligned} \tag{2.1.5}$$

Definition 2.1.8. Suppose the regression value of model is $\hat{\tilde{Y}}_i=(y_i,\varepsilon_i)$, but actually measure value is Y_i, then

$$RIC = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{N} y_i^2}}$$

is an accurate level measuring a forecast model, and $RIC \in [0, +\infty)$. When RIC=0, it is a perfect forecast.

Put $\hat{\hat{A}}_j$ into (2.1.2), that is, a fuzzy linear regression model with fuzzy coefficients is what we find. Obviously, $c = 0$ is a classical case.

The mold steps can be induced as follows.

Step 1. Put the collected data (ordinarily real data) into (2.1.3), and according to Theorem 2.1.1, change the solution to parameter $\tilde{A}_j$ into a solution to a linear programming.

Step 2. Find an optimal parameter solution $\hat{A}_j (j = 1, 2, \cdots, n)$ to (2.1.1), then we get a regression forecasting model

$$\hat{\hat{Y}}_i = \hat{\hat{A}}_1 x_{i1} + \hat{\hat{A}}_2 x_{i2} + \ldots + \hat{\hat{A}}_n x_{in} (i = 1, 2, \cdots, N).$$

Step 3. Obtain an accurate judgement in forecast model

a. The nearer RIC reaches zero, the nearer the value of $\hat{y}_i$ approaches y_i, which means the higher an accuracy of the forecasting value.

b. At RIC=0, this is a perfect forecast, here,

$$\hat{y}_i = (y_i + \varepsilon_i) \times 0.618 + (y_i - \varepsilon_i) \times 0.382.$$

Through judgement, the model passes through examination, then it can be thrown into a forecast.

c. The estimation of the forecast value range.

Suppose $\hat{y}_i = (y_i, \varepsilon_i) = \sum_{j=1}^{n} \hat{A}_j x_{ij}$, then take

$$\hat{y}_i^- = y_i - (1 - H)\varepsilon_i,\ \hat{y}_i^+ = y_i + (1 - H)\varepsilon_i,$$

hence $[\hat{y}_i^-, \hat{i}_t^+]$ is a forecast value range.

2.2 Self-regression Models with $(\cdot, c)-$Fuzzy Coefficients

2.2.1 Introduction

On the foundation of Ref.[Cao90],[Dia87] and [Wat87], we put another model into consideration, that is, self-regression model with $(\cdot, c)$ fuzzy coefficients.

It is used to generalize a fuzzy least squares system through a special example of T-fuzzy data, which will be more extensive in its application than a classical one.

2.2.2 Model

Let us consider the self-regression forecast model of classical n-order

$$Y_t = A_0 + A_1 Y_{t-1} + \cdots + A_n Y_{t-n} + e. \tag{2.2.1}$$

This means applying a fuzzy set theory to the expansion of (2.2.1), i.e.,

$$\widetilde{Y}_t = \widetilde{A}_0 + \widetilde{A}_1 Y_{t-1} + \cdots + \widetilde{A}_n Y_{t-n} + e_t, \tag{2.2.2}$$

calling (2.2.2) a self-regression model with $(\cdot, c)$ fuzzy coefficients, where parameter $\widetilde{A}_j (j = 0, 1, \cdots, n)$ to be estimated and dependent sequence $\widetilde{Y}_t$ are all $(\cdot, c)$ fuzzy numbers, and e_t is an error, t denotes benchmark time.

Assume $\widetilde{Y}_t$ and $\widetilde{A}_j (j = 0, 1, \cdots, n)$ to be all convex and normalized fuzzy numbers. In [Cao89b] appear an expansion principle and the conception of fuzzy numbers.

Definition 2.2.1. Let $f : \mathscr{R}^{n+1} \to \mathscr{F}(y)$ be a fuzzy function, $\widetilde{Y}_t = f(Y_{t-j}, \widetilde{A})(j = 1, 2, \cdots, n)$, where $Y_{t-j} \in \mathscr{R}, \widetilde{A}$ is a fuzzy set, $\mathscr{F}(y)$ represents all fuzzy subsets on $\mathscr{R}$ and the membership function of $\widetilde{Y}_t$ is

$$\mu_{\widetilde{Y}_t}(y) = \begin{cases} \max\limits_{\{a|y=f(y_{t-j},a)\}} \mu_{\widetilde{A}}(a), & \{a|y = f(y_{t-j}, a)\} \neq \phi, \\ 0, & \text{otherwise}. \end{cases}$$

Definition 2.2.2. Fuzzy self-regression parameter $\widetilde{A}$ is defined by a Cartesian product set

$$\widetilde{A} = \widetilde{A}_0 \times \widetilde{A}_1 \times \cdots \times \widetilde{A}_n,$$

which is on Cartesian product space $\mathscr{R}^{n+1}$. The membership function of $\widetilde{A}_j$ is

$$\mu_{\widetilde{A}_j}(a_j) = \begin{cases} \dfrac{1 - |\alpha_j - a_j|}{c_j}, & a_j \in [\alpha_j - c_j, \alpha_j + c_j], \\ 0, & \text{otherwise}, \end{cases}$$

where $a = \prod\limits_{j=0}^{n} a_j, \widetilde{A}_j = (\alpha_j, c_j)(j = 0, 1, \cdots, n), \alpha_j$ is the mean value of $\widetilde{A}_j$ and $c_j > 0$ is the width of $\widetilde{A}_j$.

Proposition 2.2.1. *Fuzzy self-regression model is*

$$\widetilde{Y}_t = \widetilde{A}_0 + \sum_{j=1}^{n} \widetilde{A}_j Y_{t-j} = \widetilde{A}Y = (\alpha^T Y, c^T Y), \tag{2.2.3}$$

where $Y = (1, Y_{t-1}, \cdots, Y_{t-n})^T, \alpha = (\alpha_0, \alpha_1, \cdots, \alpha_n)^T, c = (c_0, c_1, \cdots, c_n)^T$, *and the membership function of* $\widetilde{Y}_t$

$$\mu_{\widetilde{Y}_t}(y) = \begin{cases} 1 - \dfrac{|y - \alpha^T Y|}{c^T|Y|}, & Y \neq 0, \\ 0, & otherwise. \end{cases} \tag{2.2.4}$$

Proof: Applying Definition 2.2.1 and stipulating $\frac{0}{0}=0$,

$$\begin{aligned} \mu_{\widetilde{Y}_t}(y) &= \begin{cases} \bigvee\limits_{\{a|a^TY=y\}} \mu_{\widetilde{A}}(a), & \{a|a^TY=y\} \neq \phi \\ 0, & \text{otherwise} \end{cases} \\ &= \begin{cases} \bigvee\limits_{\{a|a^TY=y\}} \left\{ \bigwedge\limits_{j=1}^{n} (1 - \dfrac{|\alpha_j - a_j|}{c_j}) \right\}, & Y \neq 0 \\ 0, & \text{otherwise} \end{cases} \\ &= (2.2.4). \end{aligned}$$

A decision-maker choose threshold value H_0. If the degree of fitting H between forecast data and estimation value tallies with

$$\max H \geqslant H_0,$$

such that

$$\widetilde{Y}_t^{*H_0} = \{y|\mu_{\widetilde{Y}_t^*}(y) \geqslant H_0\}, \tag{2.2.5}$$

then we attain the best estimation of dependent variable of (2.2.3).

The approximate indicator of H is shown as in Figure 2.2.1:

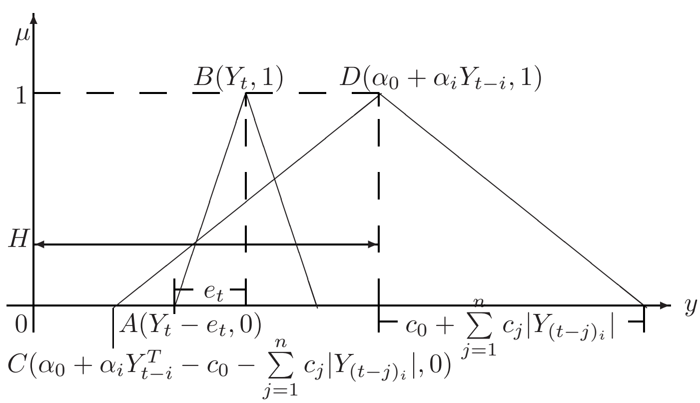

Fig. 2.2.1. The Approximate Indicator of H

Theorem 2.2.1. *Let fuzzy self-regression model be (2.2.2). Then*
$\max H \geqslant H_0$

$$\Longleftrightarrow \begin{cases} -\alpha_0 - \alpha_i Y_{t-i} + (1 - H_0)\left[c_0 + \sum\limits_{j=1}^{n} c_j |Y_{(t-j)_i}|\right] \\ \qquad \geqslant -Y_t + (1 - H_0)e_t, & (2.2.6) \\ \alpha_0 + \alpha_i Y_{t-i} + (1 - H_0)\left[c_0 + \sum\limits_{j=1}^{n} c_j |Y_{(t-j)_i}|\right] \\ \qquad \geqslant Y_t + (1 - H_0)e_t. & (2.2.7) \end{cases}$$

Proof: From Figure 2.2.1, at $Y_t - \alpha_i Y_{t-i} \leqslant 0$, the line segments $\overline{AB}$ and $\overline{CD}$ show separately below:

$$\begin{cases} x = e_t(y - 1) + Y_t, \\ y = x - (\alpha_0 + \alpha_i Y_{t-i}) + c_0 + \dfrac{\sum\limits_{j=1}^{n} c_j |Y_{(t-j)_i}|}{c_0 + \sum\limits_{j=1}^{n} c_j |Y_{(t-j)_i}|} \end{cases}$$

$$\Longleftrightarrow |H_0| = 1 - \frac{\alpha_0 + \alpha_i Y_{t-i} - Y_t}{c_0 + \sum\limits_{j=1}^{n} c_j |Y_{(t-j)_i}| - e_t},$$

where $e_t = y_t - \widehat{y}_t$ represents a deviation and $e_t = 0$ is a non-fuzzy state. Then by combining with (2.2.5), we can obtain (2.2.6).

When $y_t - \alpha_i Y_{t-i} \geqslant 0$, by the same method we can obtain (2.2.7).

If definition $J = \sum\limits_{j=1}^{n} c_j |Y_{(t-j)_i}|$ is the fuzzy degree of (2.2.2), we change finding parameter $\widehat{\widetilde{A}}_j (j = 0, 1, \cdots, n)$ into solving the optimal solution to

$$\min\{J = \sum_{j=1}^{n} c_j |Y_{(t-j)_i}| |(2.2.6), (2.2.7)\}. \tag{2.2.8}$$

Algorithm Steps

The built model steps of (2.2.2) are summed up as follows:

Step 1. By observation data, we work out a self-dependent sequence table.

Step 2. By

$$r_j = \frac{N \sum\limits_{i=1}^{N} Y_{(t-j)_i} Y_{t_i} - \sum\limits_{i=1}^{N} Y_{(t-j)_i} Y_{t_i}}{\sqrt{[N \sum\limits_{i=1}^{N} Y^2_{(t-j)_i} - (\sum\limits_{i=1}^{N} Y_{(t-j)_i})^2][N \sum\limits_{i=1}^{N} Y^2_{t_i} - (\sum\limits_{i=1}^{N} Y_{t_i})^2]}},$$

$$(j = 1, 2, \cdots, n \in \{j\}),$$

we calculate the self-dependent coefficients, where time moves backward by i.

Step 3. Determine $\widetilde{Y}_t = \widetilde{A}_0 + \sum_{j=1}^{n} \widehat{\widetilde{A}}_j Y_{(t-j)_i}$ to be the best fuzzy self-regression forecast model by taking $r_\kappa = \max\{r_j | j = 1, 2, \cdots, n\}$. Again according to Theorem 2.2.1, we solve $\widehat{\widetilde{A}}_j$ and obtain the self-regression equation

$$\widehat{\widetilde{y}}_t = (y_t, e_t) = \widehat{\widetilde{A}}_0 + \sum_{j=1}^{n} \widehat{\widetilde{A}}_j Y_{t-j}.$$

Step 4. Decision

Let

$$\widehat{y}_t = 0.618(y_t + e_t) + 0.382(y_t - e_t).$$

Then define

$$RIC = \sqrt{\frac{\sum_{i=1}^{N} (\widehat{y}_{t_i} - Y_{t_i})^2}{\sum_{i=1}^{N} Y_{t_i}^2}}, RIC \in [0, \infty).$$

The closer RIC approaches zero, the higher a precision of the forecast is. And $RIC = 0$ stands for a perfect forecast.

Step 5. Forecast

Let

$$\widehat{\widetilde{Y}}_{t+q} = \widehat{\widetilde{A}_0} + \sum_{j=1}^{n} \widehat{\widetilde{A}}_j Y_{t-(j+q)}.$$

Then the state at q moment can be forecasted, and the range of forecast value is estimated to be

$$Y^*_{t+j} \in [Y_{t+q} - (1 - H_0)e_{t+q}, Y_{t+q} + (1 - H_0)e_{t+q}].$$

2.2.3 Conclusion

In 1992, the author of Paper [Yin92] advances least fuzzy squares identification method by using the model of Papers [Cao89b][Cao90], calling fuzzy least squares systems. In south maintenance section in Zhengzhou Railroad Bureau, we analyzed the spectrum data sample of lubricate oil from BJ-type diesel locomotive. From 200 BJ-type motorcycles, we diagnosed 50 sets randomly by fuzzy least squares systems set up by the writer and knew that abnormal wear positions are generally exactitude, so are the diagnostics of total state and breakdown positions basically. Besides, a correct rate doubles than that is done by the methods of critical value or regression control finiteness. From this point, we can diagnose its breakdown without disassembly of diesel machines, therefore, acquisition of the economic profit is beyond estimation because of its convenience and practicality.

2.3 Exponential Model with Fuzzy Parameters

2.3.1 Introduction

Consider a model by Lenz, Isenson and Hartman as follows, where volume of information increases as time and factors concerned do, and we change it into a forecasting technique function before concluding it as the following mathematics model

$$\dot{Y}_t = kY_t \quad (k > 0), \tag{2.3.1}$$

where Y_t is a characteristic parameter; t is time; k a proportional constant; $\dot{Y}_t$ a relative increase rate and the solution to the equation (2.3.1) is an exponential one $Y_t = Y_0 e^{kt}$. Because the characteristic technology in long-distance telephone tallies with exponential regularity, we consider a more general exponential one as follows:

$$\hat{Y}(t) = A_1 A_2^t, \tag{2.3.2}$$

where A_1, A_2 are parameters to be estimated, and $\hat{Y}(t)$ denotes the evaluation in telephone amount during t years. Telephone amount fluctuates with various indeterminable factors. If we assume that the parameters waiting for evaluation in (2.3.2) are fuzzy numbers, the model will contain more information. Below, we fuzzify the parameters of model (2.3.2), based on Zadeh's fuzzy sets theory [Zad65a], establish a forecast model with the exponential type of the fuzzy parameter, and study the application of this model by practical example.

2.3.2 Exponential Model with Fuzzy Coefficients

Definition 2.3.1. Suppose $\mathscr{F}(\mathscr{R})$ to be a set of the whole fuzzy parameters, and $\tilde{A}_i \in \mathscr{F}(\mathscr{R})(i = 1, 2)$, we have

$$\tilde{Y}(t) = \tilde{A}_1 \tilde{A}_2^t, \tag{2.3.3}$$

where $\tilde{A}_1, \tilde{A}_2$ are flexibly fixed values in the closed intervals $[A_1^-, A_1^+], [A_2^-, A_2^+]$, respectively, $A_1^- < A_1^+, A_2^- < A_2^+$, and $A_1^-, A_1^+; A_2^-, A_2^+$ stand for all real numbers, $\tilde{Y}(t)$ denotes fuzzy telephone amount, t denotes time. We call (2.3.3) an exponential model with parameters.

Next, solutions are introduced to the Model (2.3.3).

1^0 Nonfuzzification

Theorem 2.3.1. *If the membership function $\phi : \mathscr{R} \to [0, 1]$ is a continuous and strictly monotone, then, the inverse function ϕ^{-1} exists, such that*

$$\phi(\tilde{A}_j) \geqslant \alpha \Rightarrow \tilde{A}_j \geqslant \phi^{-1}(\alpha), \alpha \in [0, 1], \ (j = 1, 2).$$

Proof: From definition of $\alpha-$cut set, the theorem appears obviously.

Let $\phi(\tilde{A}_j)$ like (1.5.3). If $\phi(\tilde{A}_j) \geqslant \alpha, \alpha \in [0, 1]$, then

$$\begin{aligned} \frac{\tilde{A}_j - A_j^-}{A_j^+ - A_j^-} \geqslant \alpha \Rightarrow \tilde{A}_j - A_j^- &\geqslant \alpha(A_j^+ - A_j^-) \\ \Rightarrow \tilde{A}_j &\geqslant A_j^- + \alpha(A_j^+ - A_j^-), (j = 1, 2). \end{aligned}$$

Take

$$\tilde{A}_1 \to A_1^- + \alpha(A_1^+ - A_1^-), \tilde{A}_2 \to A_2^- + \alpha(A_2^+ - A_2^-).$$

Put them into (2.3.3), then, $\tilde{Y}(t) \to Y(t,\alpha)$, and (2.3.3) becomes a crisp model

$$\hat{Y}(t,\alpha) = [A_1^- + \alpha(A_1^+ - A_1^-)][A_2^- + \alpha(A_2^+ - A_2^-)]^t, \alpha \in [0,1]. \qquad (2.3.4)$$

It is testified.

2^0 Linearizing

Let $A = A_1^- + \alpha(A_1^+ - A_1^-), B = A_2^- + \alpha(A_2^+ - A_2^-)$. Then change (2.3.4) into

$$\hat{Y}(t,\alpha) = AB^t. \qquad (2.3.5)$$

Linearize (2.3.5) by taking logarithm and we can get

$$\ln \hat{Y}(t,\alpha) = \ln A + t \ln B. \qquad (2.3.6)$$

3^0 Estimation parameters

Now coming next is estimation parameters A and B.

Theorem 2.3.2. *As for the given sample set* $\{Y(t_1,\alpha), Y(t_2,\alpha), \cdots, Y(t_N,\alpha)\}$, $\alpha \in [0,1]$, *the least squares estimator for parameters* A, B *with variable* α *are*

$$\hat{A} = \exp\left\{\frac{\sum_{k=1}^{N-1} \triangle t_k \ln Y(t_k,\alpha) - \sum_{k=1}^{N-1} t_k \ln \frac{Y(t_{k+1},\alpha)}{Y(t_k,\alpha)}}{\sum_{k=1}^{N-1} \triangle t_k^2}\right\}, \qquad (2.3.7)$$

$$\hat{B} = \exp\left\{\frac{\sum_{k=1}^{N-1} t_k \ln Y(t_k,\alpha) - \ln A \sum_{k=1}^{N} t_k}{\sum_{k=1}^{N} t_k^2}\right\}. \qquad (2.3.8)$$

Proof: a) Because of a sample set

$$\{Y(t_1,\alpha), \cdots, Y(t_N,\alpha)\} \to \{\ln Y(t_1,\alpha), \cdots, \ln Y(t_N,\alpha)\},$$

then for given sample points $\{\ln Y(t_k,\alpha)\}(k = 1,2,\cdots,N), \alpha \in [0,1]$, we take two near arbitrary sample points t_k and $t_{k+1}(k = 1,2,\cdots,N-1)$ into consideration from (2.3.6), then

$$\ln \hat{Y}(t_k,\alpha) = \ln A + t_k \ln B, \qquad (2.3.9)$$

$$\ln \hat{Y}(t_{k+1},\alpha) = \ln A + t_{k+1} \ln B. \qquad (2.3.10)$$

(2.3.9) $\times t_{k+1}$ − (2.3.10) $\times t_k$, we obtain

$$(t_{k+1} - t_k)\ln A = t_{k+1}\ln \hat{Y}(t_k,\alpha) - t_k \ln \hat{Y}(t_{k+1},\alpha). \tag{2.3.11}$$

b) Applying the least square method, we build an objective function by (2.3.11)

$$J_1 = \sum_{k=1}^{N-1}[t_{k+1}\ln Y(t_k,\alpha) - t_k \ln Y(t_{k+1},\alpha) - (t_{k+1}-t_k)\ln A]^2.$$

By combining (2.3.9), we build another objective function by the least square method

$$\begin{aligned} J_2 &= \sum_{k=1}^{N}[\ln Y(t_k,\alpha) - \ln \hat{Y}(t_k,\alpha)]^2 \\ &= \sum_{k=1}^{N}[\ln Y(t_k,\alpha) - (\ln A + t_k \ln B)]^2. \end{aligned}$$

To extract minimum of J_1 and J_2, we let $\dfrac{\partial J_1}{\partial \ln A} = 0$, $\dfrac{\partial J_2}{\partial \ln B} = 0$, and write down $\triangle t_k = t_{k+1} - t_k$ before obtaining

$$\begin{cases} \sum\limits_{k=1}^{N-1}[t_{k+1}\ln Y(t_k,\alpha) - t_k\ln Y(t_{k+1},\alpha)] \triangle t_k = \sum\limits_{k=1}^{N-1} \triangle t_k^2 \ln A, \\ 2\sum\limits_{k=1}^{N}[\ln Y(t_k,\alpha) - \ln A - t_k \ln B]t_k = 0. \end{cases} \tag{2.3.12}$$

Solve (2.3.12) and we get (2.3.7) and (2.3.8). It is certificated.

4^0 **Test**

Obviously, to a certain determination α, Model (2.3.5) is determined after two-step of linearized model, so is Model (2.3.3). From the principle, we get two determined values $\alpha_1, \alpha_2 (\in [0,1])$, such that

$$\hat{Y}(t_k,\alpha_1) \leqslant \hat{Y}(t_k,\alpha_2),$$

we get

$$\hat{Y}(t_k,\alpha) = \hat{Y}(t_k,\alpha_1) + 0.618 \times [\hat{Y}(t_k,\alpha_2), \hat{Y}(t_k,\alpha_1)]. \tag{2.3.13}$$

Again from formula

$$S = \sqrt{\frac{\sum\limits_{k=1}^{N}[Y(t_k,\alpha) - \hat{Y}(t_k,\alpha)]^2}{N}}, \tag{2.3.14}$$

$$E\% = \frac{1}{N}\sum_{k=1}^{N}\left|1 - \frac{Y(t_k,\alpha)}{\hat{Y}(t_k,\alpha)}\right| \times 100\%. \tag{2.3.15}$$

After finding a standard deviation S in a forecasting error and an average relative error percentage $E\%$, we determine fitting best for forecasting models when S and $E\%$ are smaller.

5^0 Model determination

Theorem 2.3.3. *Let $\phi : \mathscr{R} \to [0,1]$ be a membership function of continuous and strictly monotone. Then (2.3.3) $\Longleftrightarrow$ (2.3.6).*

Proof: From the discussion above, the result is obvious.

Put $\hat{A}, \hat{B}$ into (2.3.6), we obtain a crisp model

$$\ln \hat{Y}(t, \alpha) = \ln \hat{A} + t \ln \hat{B}.$$

Because of (2.3.6) $\Longleftrightarrow$ (2.3.5), hence

$$\hat{Y}(t_k, \alpha) = \hat{A}\hat{B}^t. \tag{2.3.16}$$

But (2.3.5) $\Longleftrightarrow$ (2.3.3), such that (2.3.3) $\Longleftrightarrow$ (2.3.6). It is certificated.

Therefore, we can design a controlling forecast system for telephone amount, with a classical system an exception.

If the above result strays away from practice, we can obtain $\hat{Y}(k, \alpha)$ by taking value from [0,1]. But if we do so, we may get infinite values. It is impossible for us to calculate infiniteness of $\hat{Y}$, so we calculate the value of $\hat{Y}(k, 0)$ by choosing $\alpha = 0$. Compare it with $\hat{Y}(k, 1)$, if $\hat{Y}(k, 0)$ is superior to $\hat{Y}(k, 1)$, then $\hat{Y}(k, 0)$ is the goal. Otherwise, we apply the 0.618 method for search until an optimal value of the problem is found.

Especially, when $t_k = k(k = 1, 2, \cdots, N)$, we have

$$\hat{Y}(t_k, \alpha) = \hat{Y}(k, \alpha),$$

where $\triangle t_k = t_{k+1} - t_k = 1$, and at this time, we change (2.3.7) and (2.3.8) into

$$\begin{aligned}\hat{A} &= \exp\left\{\frac{\sum\limits_{k=1}^{N-1} \ln Y(k, \alpha) - \sum\limits_{k=1}^{N-1} k[\ln Y(k+1, \alpha) - \ln Y(k, \alpha)]}{N-1}\right\} \\ &= \exp\left\{\frac{\sum\limits_{k=1}^{N-1} 2 \ln Y(k, \alpha) - (N-1)\ln Y(N, \alpha)}{N-1}\right\},\end{aligned} \tag{2.3.17}$$

$$\hat{B} = \exp\left\{\frac{6\sum\limits_{k=1}^{N} k \ln Y(k, \alpha) - 3N(N+1)\ln A}{N(N+1)(2N+1)}\right\}. \tag{2.3.18}$$

The models corresponding to (2.3.16) and (2.3.13) denote

$$\hat{Y}(k,\alpha) = \hat{A}\hat{B}^t \tag{2.3.19}$$

and

$$\hat{Y}(k,\alpha) = \hat{Y}(k,\alpha_1) + 0.618 \times [\hat{Y}(k,\alpha_2), \hat{Y}(k,\alpha_1)], \tag{2.3.20}$$

respectively.

Because of

$$\hat{A} = \hat{A}_1^- + \alpha(\hat{A}_1^+ - \hat{A}_1^-), \hat{B} = \hat{A}_2^- + \alpha(\hat{A}_2^+ - \hat{A}_2^-), \tag{2.3.21}$$

we compute those simultaneous equations (2.3.21) with a determination α, then $\hat{A}_1^-, \hat{A}_1^+, \hat{A}_2^-, \hat{A}_2^+$ are determined. Now,we synthesize again an exponential model below

$$\hat{Y}(k,\alpha) = \hat{A}_1\hat{A}_2^k, \tag{2.3.22}$$

such that an exponential model with fuzzy parameters can be obtained below

$$\tilde{Y}(k) = \tilde{A}_1\tilde{A}_2^k. \tag{2.3.23}$$

2.3.3 Practical Example

Example 2.3.1: The amount of long-distance telephone in China during 1980-1990 shows as follows.

Table 2.3.1. Amount of Long-distance Telephone in China

Year	No	Practical date	Year	No	Practical date	Year	No	Practical date
1980	1	[14940,21404]	1984	5	[31549,31553]	1988	9	[64615,64617]
1981	2	[18031,22049]	1985	6	[38250,38254]	1989	10	[78458,78462]
1982	3	[21760,23574]	1986	7	[42299,42303]	1990	11	[97932,106291]
1983	4	[26262,26556]	1987	8	[51521,51525]			

Forecast time for telephone by applying an exponential Model (2.3.23) with fuzzy parameters, and, from (2.3.11), we take $\alpha = 1$, with Formula (2.3.22) correspondingly being

$$\hat{Y}(k,1) = \hat{A}_1^+(\hat{A}_2^+)^k.$$

If by using (2.3.17) and (2.3.18), we can get parameters

$$\hat{A}_1^+ = 12380, \hat{A}_2^+ = 1.2069.$$

When $\alpha = \alpha_1 = \alpha_2 = 1$, from (2.3.20), then

$$\hat{Y}(k,1) = 12380 \times 1.2069^k (k = 1, 2, \cdots, 11).$$

Hence, telephone amount forecast value at $\alpha = 1$ is shown as Table 2.3.2.

Table 2.3.2. Amount of Long-distance Telephone at $\alpha = 1$ in China

Year	No	Practical date	Year	No	Practical date	Year	No	Practical date
1980	1	21404	1984	5	31553	1988	9	64617
1981	2	22049	1985	6	38254	1989	10	78462
1982	3	23574	1986	7	42303	1990	11	106291
1983	4	26556	1987	8	51525			

By a standard deviation formula (2.3.14),

$$S = \sqrt{\frac{\sum_{k=1}^{11}[Y(k,1) - \hat{Y}(k,1)]^2}{11}},$$

we can obtain $S = 4019$. Again, from formula (2.3.15) of percentage error

$$E\% = \frac{1}{11}\sum_{k=1}^{11}\left|1 - \frac{Y(k,1)}{\hat{Y}(k,1)}\right| \times 100\%,$$

we can get an average relative error to be 8.21%. While, by the aid of geometric average, we obtain $S = 9405, E\% = 19.78\%$, and $S = 4811, E\% = 9.74\%$ by average value exponential curve. Therefore, the fuzzy exponential forecast method mentioned here is superior to the above two [Zhe92].

Under the fiducial degree of 95%, the long-distance telephone in China varies at the following interval $\hat{Y} \pm 2S$. Hence, their forecast amount between 1980 – 1990 shows below.

Table 2.3.3. Forecast Amount of Long-distance Telephone in China

Year	No	Practical date	Year	No	Practical date	Year	No	Practical date
1980	1	[6910.6, 22972.2]	1984	5	[23670.5, 39732.1]	1988	9	[59230, 75291.6]
1981	2	[10002, 26063.6]	1985	6	[30229.5, 46291.1]	1989	10	[73146.3, 89207.9]
1982	3	[13733, 29794.6]	1986	7	[38145.6, 54207.2]	1990	11	[89941.9, 106003.4]
1983	4	[18236, 34297.5]	1987	8	[47699.5 63761]			

If we make use of the 0.618 method by selections of $\alpha(\in [0,1])$, and make use of (2.3.20) for search, we may acquire a better result.

2.3.4 Conclusion

The method in this section is an extension of fuzzy exponential forecast model. We can always change it into a series of determination forecast models for different α values ($\alpha \in [0,1]$), and then obtain a forecast value for linearized model respectively by adopting two-step of the least square method. Each forecast value $\hat{Y}$ fluctuates in the band region composed of $\hat{Y}^-$ and $\hat{Y}^+$, which presents us more information when we choose a satisfactory forecast result by 0.618. It is pointed out that the model here can be still expanded to contain situations with various fuzzy coefficients and even with the fuzzy variables. [Cao89b][Cao93e][DPr78][TUA82][Wat87][Zad82].

2.4 Regression and Self-regression Models with Flat Fuzzy Coefficients

2.4.1 Basic Properties

Definition 2.4.1. For $\forall x, y, z \in \mathscr{R}$ and $x \leqslant y \leqslant z$ satisfy

i) $\mu_{\tilde{A}}(y) \geqslant \mu_{\tilde{A}}(z) \wedge \mu_{\tilde{A}}(z)$,
ii) $\max\limits_{x\in\mathscr{R}} \mu_{\tilde{A}}(x) = 1$,

then call $\widetilde{A}$ a convex normal fuzzy number.

We also call $A_\alpha = \{x|\mu_{\tilde{A}}(x) \geqslant \alpha, 0 < \alpha \leqslant 1\}$ a platform of flat fuzzy number $\widetilde{A}$.

Proposition 2.4.1. *The flat fuzzy number $\widetilde{A}$ is convex $\Leftrightarrow$ A_α is all interval* $(0 < \alpha \leqslant 1)$.

Proof: "$\Rightarrow$ " If $\widetilde{A}$ is convex, then from Definition 2.4.1 we know, *i)* $y \in A_\alpha$, again according to randomness in x, y, z, we know, A_α is an interval necessarily.

"$\Leftarrow$ " If $\forall\alpha \in [0, 1], A_\alpha$ is an interval. Consider $x, z \in \mathscr{R}$, and let $\alpha_0 = \mu_{\tilde{A}}(x) = \mu_{\tilde{A}}(z)$, then A_{α_0} must be an interval. From $x \leqslant y \leqslant z$, $y \in A_{\alpha_0}$, then $\mu_{\tilde{A}}(y) \geqslant \alpha_0$, hence $\widetilde{A}$ is convex.

Again from Definition 2.4.1 we know that a flat fuzzy number necessarily satisfies $\max\limits_{x\in\mathscr{R}} \mu_{\tilde{A}}(x) = \bigvee\limits_{\alpha\in(A_j^-, A_j^+)} \mu_{\tilde{A}_j}(\alpha) = 1$, hence it is a convex normal fuzzy number.

2.4.2 Linear Regression Model with Flat Fuzzy Parameters

We always suppose that $\widetilde{A}_j$ is a convex and a normal fuzzy number, consider

$$\widetilde{Y} = \widetilde{A}_1 x_1 + \widetilde{A}_2 x_2 + \cdots + \widetilde{A}_n x_n \triangleq \widetilde{A}x, \tag{2.4.1}$$

where $\tilde{A} = (\tilde{A}_1, \tilde{A}_2, \cdots, \tilde{A}_n), x = (x_1, x_2, \cdots, x_n)^T$, which is a linear regression model. In the model, we call $y_j^* = (A_j^{*-}, A_j^{*+})x_j (j = 1, 2, \cdots, n)$ a regression value, $y_j = (A_j^{'-}, A_j^{'+})x_j$ an observation value, and $y_j - y_j^* = \varepsilon_j$ an observation error, ε_j a random variable with zero for the main value, $A_j^{*-} = A_j^{'-} \pm \varepsilon_j$ and $A_j^{*+} = A_j^{'+} \pm \varepsilon_j$.

Definition 2.4.2. Suppose $f : x \to \mathscr{F}(y)$ denotes fuzzy function $\widetilde{Y} = f(x, \widetilde{A})$, where $x \in \mathscr{R}, \mathscr{F}(y)$ is a fuzzy-valued set, the membership function in $\widetilde{Y}$ denotes

$$\mu_{\tilde{Y}}(y) = \begin{cases} \max\limits_{\{a|y=f(x,a)\}} \mu_{\tilde{A}}(a), & \{a|y=f(x,a)\} \neq \phi, \\ 0, & \text{otherwise.} \end{cases}$$

Definition 2.4.4. Suppose the quadruple parameter to be a flat fuzzy number $\widetilde{A}_j = (A_j^-, A_j^+, \sigma_{A_j^-}, \sigma_{A_j^+})$, then its membership function $\mu_{\tilde{A}_j}(a_j)$ is defined as

$$\mu_{\tilde{A}_j}(a_j) = \begin{cases} 1 - \dfrac{A_j^- - a_j}{\sigma_{A_j^-}}, & A_j^- - \sigma_{A_j^-} \leqslant a_j < A_j^-, \\ 1, & A_j^- \leqslant a_j \leqslant A_j^+, \\ 1 - \dfrac{a_j - A_j^+}{\sigma_{A_j^+}}, & A_j^+ < a_j \leqslant A_j^+ + \sigma_{A_j^+}, \\ 0, & \text{otherwise.} \end{cases}$$

Proposition 2.4.2. *Suppose regression coefficient* $\widetilde{A} = (A^-, A^+, \sigma_A^-, \sigma_A^+)$ *to be a flat fuzzy number, then the membership function in (2.4.1) is*

$$\mu_{\tilde{Y}}(y) = \begin{cases} 1 - \dfrac{A^- x^T - y}{\sigma_A^- x^T}, & (A^- - \sigma_A^-)x^T \leqslant y < A^- x^T, \\ 1, & A^- x^T \leqslant y \leqslant A^+ x^T, \\ 1 - \dfrac{y - A^+ x^T}{\sigma_A^+ x^T}, & A^+ x^T < y \leqslant (A^+ + \sigma_A^+)x^T, \\ 0, & \textit{otherwise}, \end{cases} \tag{2.4.2}$$

where $x = (x_1, x_2, \cdots, x_n)^T$.

Proof:

$$\begin{aligned} \mu_{\tilde{Y}}(y) &= \begin{cases} \bigvee\limits_{\{a|a^T x = y\}} \mu_{\tilde{A}}(a), & \{a|ax^T = y\} \neq \phi \\ 0, & \text{otherwise} \end{cases} \\ &= \begin{cases} \bigvee\limits_{\{a|a^T x = y\}} \{\bigwedge\limits_{j=1}^{n} \mu_{\tilde{A}_j}(a_j)\}, & \{ax^T = y\} \neq \phi \\ 0, & \text{otherwise} \end{cases} \\ &= \begin{cases} \bigvee\limits_{\{a|a^T x = y\}} \{\bigwedge\limits_{j=1}^{n} (1 - \dfrac{A_j^- - a_j}{\sigma_{A_j}^-})\}, & A_j^- - \sigma_{A_j^-} \leqslant a_j < A_j^- \\ 1, & A_j^- \leqslant a_j \leqslant A_j^+ \\ \bigvee\limits_{\{a|a^T x = y\}} \{\bigwedge\limits_{j=1}^{n} (1 - \dfrac{a_j - A_j^+}{\sigma_{A_j}^+})\}, & A_j^+ < a_j \leqslant A_j^+ + \sigma_{A_j}^+ \\ 0, & \text{otherwise} \end{cases} \\ &= (2.4.2). \end{aligned}$$

The proposition holds.

Suppose fuzzy linear regression model $\widetilde{Y}_i^* = \widetilde{A}_1^* x_{i1} + \widetilde{A}_2^* x_{i2} + \ldots + \widetilde{A}_n^* x_{in} \triangleq \widetilde{A}^* x_i (i = 1, 2, \cdots, N)$, where $\widetilde{A}^* = (\widetilde{A}_1^*, \widetilde{A}_2^*, \ldots, \widetilde{A}_n^*), x_i = (x_{i1}, x_{i2}, \ldots, x_{in})^T$. Then the membership function of $\widetilde{Y}_i^*$ is given by

$$\mu_{\tilde{Y}_i^*}(y) = 1 - \frac{|y_i - A_i^{*\pm} x_i|}{\sigma_{A_i^{*\pm}} |x_i|},$$

its degree of fitting estimation to the given data $Y_i = (y_i, \varepsilon_i)$ is measured by the following index $\overline{h}_l(l = 1, 2)$, which maximizes h subject to $Y_i^h \subset Y_i^{*h}(i = 1, 2, \cdots, N)$, where

$$Y_i^h = \{y|\mu_{\tilde{Y}_i}(y) \geqslant h\}, \qquad Y_i^{*h} = \{y|\mu_{\tilde{Y}_i^*}(y) \geqslant h\}$$

are $\overline{h}-$level sets, and index $\bar{h}$ is illustrated as Figure 2.4.1:

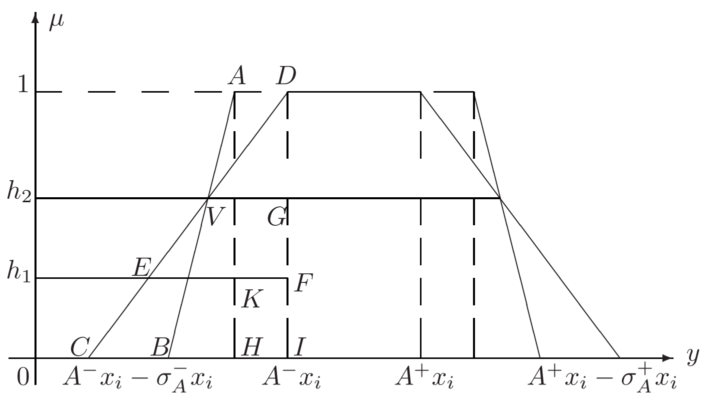

Fig. 2.4.1. Illustration for Membership Function of Regression Coefficient $\widetilde{A}$

The fitting degree of a fuzzy linear regression model to all data $\widetilde{Y}_1, \widetilde{Y}_2, \cdots, \widetilde{Y}_N$ are defined by $\min_l \{\overline{h}_l\}$.

Definition 2.4.4. Use

$$J^{(1)} = \sigma_{A_1}^- x_{i1} + \sigma_{A_2}^- x_{i2} + \cdots + \sigma_{A_n}^- x_{in},$$
$$J^{(2)} = \sigma_{A_1}^+ x_{i1} + \sigma_{A_2}^+ x_{i2} + \cdots + \sigma_{A_n}^+ x_{in} (i = 1, 2, \cdots, N)$$

to denote fuzzy degree of Model (2.4.1) in left and right shapes, respectively.

The problem is explained as fuzzy parameters $\tilde{A}^*$ being obtained, which minimize $(J^{(1)}, J^{(2)})$ subject to $\bar{h}_l \geqslant h$ for all l, where $h_l = (h_1, h_2)$ is a degree of the fitting in a fuzzy linear model chosen by decision makers.

Theorem 2.4.1. *Suppose the model with flat fuzzy data as (2.4.1), then*

$$\overline{h} = (\min \overline{h}_1, \min \overline{h}_2)^T \geqslant (h_1, h_2)^T \tag{2.4.3}$$

$$\Leftrightarrow \begin{cases} A_i^- x_i + (1 - h_1) \sum\limits_{j=1}^{n} \sigma_{A_j}^- |x_{ij}| \geqslant y_i^- + (1 - h_1)\varepsilon_i^- \\ -A_i^- x_i + (1 - h_1) \sum\limits_{j=1}^{n} \sigma_{A_j}^- |x_{ij}| \geqslant -y_i^- + (1 - h_1)\varepsilon_i^- \end{cases} \tag{2.4.4}$$

and

$$\begin{cases} A_i^+ x_i + (1-h_2)\sum\limits_{j=1}^{n}\sigma_{A_j}^+|x_{ij}| \geqslant y_i^+ + (1-h_2)\varepsilon_i^+, \\ -A_i^+ x_i + (1-h_2)\sum\limits_{j=1}^{n}\sigma_{A_j}^+|x_{ij}| \geqslant -y_i^+ + (1-h_2)\varepsilon_i^+. \end{cases} \tag{2.4.5}$$

Proof: Shown as Figure 2.4.1, because $\triangle ABH \sim \triangle AEG$, then

$$\frac{v}{\varepsilon_i^-} = \frac{1-\overline{h}_1}{1} \Longrightarrow \; v = \varepsilon_i^-(1-\overline{h}_1).$$

But $k = v + HI = V + |y_i^- - A_j^- x_i|$, again $\triangle CDI \sim \triangle EDF$, hence

$$\frac{1-\overline{h}_1}{1} = \frac{k}{\sum\limits_{j=1}^{n}\sigma_{A_j}^-|x_{ij}|} \Rightarrow 1-\overline{h}_1 = \frac{\varepsilon_i^-(1-\overline{h}_1) + |y_i^- - A_j^- x_{ij}|}{\sum\limits_{j=1}^{n}\sigma_{A_j}^-|x_{ij}|},$$

therefore,

$$\overline{h}_1 = 1 - \frac{|y_i^- - A_j^- x_i|}{\sum\limits_{j=1}^{n}\sigma_{A_j}^-|x_{ij}| - \varepsilon_i^-}. \tag{2.4.6}$$

The same truth is that we can get

$$\overline{h}_2 = 1 - \frac{|y_i^+ - A_j^+ x_i|}{\sum\limits_{j=1}^{n}\sigma_{A_j}^+|x_{ij}| - \varepsilon_i^+}. \tag{2.4.7}$$

Combine (2.4.3) and (2.4.6),(2.4.7), then

$$1 - \frac{|y_i^- - A_j^- x_i|}{\sum\limits_{j=1}^{n}\sigma_{A_j}^-|x_{ij}| - \varepsilon_i^-} \geqslant h_1,$$

$$1 - \frac{|y_i^+ - A_j^+ x_i|}{\sum\limits_{j=1}^{n}\sigma_{A_j}^+|x_{ij}| - \varepsilon_i^+} \geqslant h_2.$$

So that (2.4.4) and (2.4.5) are established, and the theorem is certificated.

Our problem is to determine parameter in (2.4.1) $\tilde{A}_j^* = (A_j^-, A_j^+, \sigma_{A_j}^-, \sigma_{A_j}^+)$, that is to find the minimum value of $J^{(1)}$ and $J^{(2)}$ under constraint $\overline{h} \geqslant (h_1, h_2)^T$, in order to solve a classical parameter programming as follows:

$$\min J^{(1)} \qquad \text{and} \qquad \min J^{(2)}$$

$$\begin{cases} \text{s.t. } (2.4.4) \\ \quad \sigma_{A_j}^- \geqslant 0, h_1 \in [0,1], \\ \quad (j = 1, 2, \cdots, n), \end{cases} \qquad \begin{cases} \text{s.t. } (2.4.5) \\ \quad \sigma_{A_j}^+ \geqslant 0, h_2 \in [0,1], \\ \quad (j = 1, 2, \cdots, n), \end{cases} \tag{2.4.8}$$

a simplex method or a dual simplex method is used to solve their optimal solution easily, obviously, in (2.4.8), the constraint condition of each problem, that is (2.4.4) and (2.4.5), all containing $2n$ constraint, its number is larger than a variables number, so to change them into a dual form is much easier than to find an optimal parameter solution $A_j^-, \sigma_{A_j}^-; A_j^+, \sigma_{A_j}^+$, synthesize it to a flat fuzzy number in sequence, record for $\tilde{A}_j = (A_j^-, A_j^+, \sigma_{A_j}^-, \sigma_{A_j}^+)(j = 1, 2, \cdots, n)$, thus the fuzzy parameters of (2.4.1) are acquired.

2.4.3. Precise Examination of Model and Modeling Method

For given data, by solving the classical parameter programming (2.4.8), a best fitting model can be obtained. Below we determine a judgement method to the forecast model in accuracy measurement.

Definition 2.4.5. Suppose fuzzy regression value of (2.4.1) is $\hat{\tilde{y}}_i^* = (y_i^{-*}, y_i^{+*}, \varepsilon_i^{-*}, \varepsilon_i^{+*})$, actually the value is denoted by y_i, then with

$$RIC = \sqrt{\frac{\sum\limits_{i=1}^{N} (\widehat{y}_i^* - y_i)^2}{\sum\limits_{i=1}^{N} y_i^2}} \tag{2.4.9}$$

being an accuracy degree's measure level in model (2.4.1), and $RIC \in [0, \infty)$.

1) At $RIC = 0$, it is a perfect forecast.

2) The more RIC approaches zero, the nearer $\widehat{y}_i^*$ value tends to y_i; it means a higher prediction.

According to the theories of optimization method, $\widehat{y}_i^*$ with y_i in (2.4.9) being defined below:

$$\widehat{y}_i^* = (y_i^{-*} - \varepsilon_i^{-*}) \times 0.382 + (y_i^{+*} + \varepsilon_i^{+*}) \times 0.618,$$
$$y_i = (y_i^- - \varepsilon_i^-) \times 0.382 + (y_i^+ + \varepsilon_i^+) \times 0.618.$$

After the model passes through prediction examination of (2.4.9), it can be thrown into forecast formally. Suppose that the forecast to acquire regression value is $\widehat{y}_{i+p}^* = (y_{i+p}^-, y_{i+p}^+, \varepsilon_{i+p}^-, \varepsilon_{i+p}^+)$, and take threshold value $h_0 (h_0 = h_1 \vee h_2)$, and then

$$\widehat{y}_{i+p}^{-*} = y_{i+p}^- - \varepsilon_{i+p}^-(1 - h_0), \quad \widehat{y}_{i+p}^{+*} = y_{i+p}^+ - \varepsilon_{i+p}^+(1 - h_0). \tag{2.4.10}$$

Hence, $y_{i+p}^* = [\widehat{y}_{i+p}^{-*}, \widehat{y}_{i+p}^{+*}]$ is a found forecast value in model (2.4.1).

Hereby, we can acquire steps of modeling.

I. According to collection of the data (ordinarily real data), substitution them (2.4.1). According to Theorem 2.4.1 and Definition 2.4.4, convert again the ordinarily linear programming (2.4.8) with parameter variables.

II. Solve two linear programming with parameter variables in the problem (2.4.8), respectively, a parameter optimal solution to (2.4.8) is found, that is, certain fuzzy regression parameters exist in (2.4.1).

III. Give a series of data, and the best fitting model is confirmed, making precise examination by (2.4.9).

IV. Forecasting

Let $Y_k = \sum_{i=1}^{N} \widehat{A}_i x_{ik}$. Then we can forecast status at time k. By using (2.4.10) again, we can ascertain the range in forecasting value.

2.4.4 Self-regression Forecasting Model with Flat Fuzzy Parameters

According to the above section theories in a fuzzy linear regression model, we can follow the Ref.[Cao89b], and induce fuzzy time series models from flat fuzzy numbers

$$\widetilde{Y}_t = \widetilde{A}_1 Y_{t-1} + \widetilde{A}_2 Y_{t-2} + \cdots + \widetilde{A}_n Y_{t-n}. \tag{2.4.11}$$

Definition 2.4.6. Consider Model (2.4.11), call it n-order self-regression model with flat fuzzy parameters, where $\tilde{Y}_t = (Y_t^-, Y_t^+, \sigma_t^-, \sigma_t^+)$.

According to observation data $Y_{(t-j)_i} (i = 1, 2, \cdots, N; j = 1, 2, \cdots, n)$, they are all ordinarily real numbers from the formula

$$\gamma_i = \frac{N \sum_{i=1}^{N} Y_{(t-j)_i} Y_t - \sum_{i=1}^{N} Y_{(t-j)_i} \sum_{i=1}^{N} Y_{t_i}}{\sqrt{[N \sum_{i=1}^{N} Y^2_{(t-j)_i} - (\sum_{i=1}^{N} Y_{(t-j)_i})^2][N \sum_{i=1}^{N} Y^2_{t_i} - (\sum_{i=1}^{N} Y_{t_i})^2]}}. \tag{2.4.12}$$

Calculate the self-related coefficient to change backward $i (i = 1, 2, \cdots, N)$ quarter. If we take $\gamma_q = \ max\{\gamma_i | i = 1, 2, \cdots, N\}$, then the model confirmed by $\widetilde{Y}_t = \sum_{j=1}^{n} \widetilde{A}_j Y_{(t-j)_q}$ is optimal.

Theorem 2.4.2. *Suppose n-order fuzzy self-regression model to be (2.4.11), then* $\min H_m \geqslant \beta_m, \beta_m \in [0, 1], (m = 1, 2)$

$$\Leftrightarrow \begin{cases} A_j^- Y_{t-i}^T + (1 - \beta_1) \sum_{j=1}^{n} \sigma_{A_j}^- |Y_{(t-j)_i}| \geqslant Y_t^- + (1 - \beta_1) e_t^- \\ -A_j^- Y_{t-i}^T + (1 - \beta_1) \sum_{j=1}^{n} \sigma_{A_j}^- |Y_{(t-j)_i}| \geqslant -Y_t^- + (1 - \beta_1) e_t^- \end{cases} \tag{2.4.13}$$

and

$$\begin{cases} A_j^+ Y_{t-i}^T + (1 - \beta_2) \sum_{j=1}^{n} \sigma_{A_j}^+ |Y_{(t-j)_i}| \geqslant Y_t^+ + (1 - \beta_2) e_t^+, \\ -A_j^- Y_{t-i}^T + (1 - \beta_2) \sum_{j=1}^{n} \sigma_{A_j}^+ |Y_{(t-j)_i}| \geqslant -Y_t^+ + (1 - \beta_2) e_t^+. \end{cases} \tag{2.4.14}$$

Proof: In Theorem 2.4.1, what needs is only to change y_i^-, y_i^+ into Y_t^-, Y_t^+; x_i, x_{ij} into $Y_{t-i}, Y_{(t-j)_i}$, respectively. Similar proof in Theorem 2.4.1, then this theorem holds true.

Definition 2.4.7. The fuzzy degree of left and right shapes in Model (2.4.11) are denoted by

$$s_1 = \sigma^-_{A_1} Y_{(t-j)_1} + \sigma^-_{A_2} Y_{(t-j)_2} \cdots + \sigma^-_{A_n} Y_{(t-j)_n},$$
$$s_2 = \sigma^+_{A_1} Y_{(t-j)_1} + \sigma^+_{A_2} Y_{(t-j)_2} \cdots + \sigma^+_{A_n} Y_{(t-j)_n}.$$

Then, the assurance of self-regression forecasting model (2.4.11) with flat fuzzy parameters comes to arbitrary m for finding $\min\limits_{m} s_m (m = 1, 2)$ under $\bar{h} \geqslant (\beta_1, \beta_2)^T$, that is to find an optimal solution to an ordinary parameter programming

$$\min s_1 \quad \text{and} \quad \min s_2$$
$$\begin{cases} \text{s.t. } (2.4.13), \\ \quad \sigma^-_{A_i} \geqslant 0, \beta_1 \in [0,1], \\ \quad (i = 1, \cdots, N), \end{cases} \qquad \begin{cases} \text{s.t. } (2.4.14), \\ \quad \sigma^+_{A_i} \geqslant 0, \beta_2 \in [0,1], \\ \quad (i = 1, \cdots, N), \end{cases} \tag{2.4.15}$$

where β_1, β_2 are the degree of the fitting of fuzzy self-regression models for decision makers to choose.

Obviously, modeling steps in (2.4.11) can be induced as follows:

I. The self-related sequence table is programmed according to collected data.

II. By use of (2.4.12), find self-related coefficient γ, and choose forecast model (2.4.10) from $\gamma_q = \max\{\gamma_i | i = 1, \cdots, N\}$.

III. Find an optimal parameter solution to (2.4.15), thus determine a fuzzy self-regression parameter.

IV. Give a list of data, the optimally fitting model is confirmed, making the accurate examination at the same time. Suppose

$$RIC = \sqrt{\frac{\sum\limits_{i=1}^{N} (\widehat{Y}_{t_i} - y_{t_i})^2}{\sum\limits_{i=1}^{N} y^2_{t_i}}},$$

where $\widehat{Y}_{t_i} = (\widehat{Y}^-_{t_i} - e^-_{t_i}) \times 0.382 + (\widehat{Y}^+_{t_i} + e^+_{t_i}) \times 0.618$; $y_{t_i} = (y^-_{t_i} - e^-_{t_i}) \times 0.382 + (y^+_{t_i} + e^+_{t_i}) \times 0.618$ is accurate to measurement level of forecast Model (2.4.11), and $RIC \in [0, \infty)$. Judge the following

1^0 At $RIC = 0$, it is a perfect forecast.

2^0 The more RIC approaches zero, the higher the forecast precision is, otherwise lower.

V. Forecast.

Let $\tilde{y}_{t+q} = \sum\limits_{j=1}^{n} \widetilde{\widehat{A}}_j Y_{t-(p_j+q)}$. Then we forecast the status at time q; its forecasting range is $Y^*_{t+q} = [\widehat{Y}^-_{t+q}, \widehat{Y}^+_{t+q}]$, where $\widehat{Y}^-_{t+q} = y^-_{t+q} - e^-_{t+q}(1 - H_0), \widehat{Y}^+_{t+q} = y^+_{t+q} - e^+_{t+q}(1 - H_0), H_0 = \beta_1 \vee \beta_2$ is a threshold value.

From (2.4.1) and (2.4.11) we know, when the spread of its parameter σ_A^-, σ_A^+ and spread of related fuzzy variables $\widetilde{Y}$ e^-, e^+ are all zero, (2.4.1) and (2.4.11) are exuviated into classical models in linear regression and self-regression.

2.5 Linear Regression with Triangular Fuzzy Numbers

This section presents a new definition on the distance between two triangular fuzzy numbers with respect to their parameter variables and it provides a new method to fuzzy linear regression problems.

2.5.1 Preliminary

In order to study the fuzzy linear regression with triangular fuzzy numbers, we introduce some basic knowledge as follows.

Definition 2.5.1. A fuzzy set $\tilde{A}$ is called a fuzzy number on $\mathscr{R}$ if it satisfies the following:

(1) There exists $x_0 \in \mathscr{R}$ such that $\mu_{\tilde{A}}(x_0) = 1$;
(2) $\forall \alpha \in [0,1], A_\alpha = \{x|\mu_{\tilde{A}}(x) \geqslant \alpha\} = [\underline{A}_\alpha, \overline{A}_\alpha]$ is a closed interval on $\mathscr{R}$.

Denote $\mathscr{F}(\mathscr{R})$ as the set of all fuzzy numbers on $\mathscr{R}$, and among $\mathscr{F}(\mathscr{R})$, we often use triangular fuzzy numbers.

Definition 2.5.2. If $A \in \mathscr{F}(\mathscr{R})$, it satisfies conditions
(1) $\forall \alpha \in [0,1], A_\alpha$ is a convex set on $\mathscr{R}$;
(2) Its membership function $\tilde{A}$ can be expressed as

$$\mu_{\tilde{A}}(x) = \begin{cases} \dfrac{x - A^L}{A^C - A^L}, & \text{when} \quad A^L \leqslant x \leqslant A^C, \\ \dfrac{x - A^R}{A^C - A^R}, & \text{when} \quad A^C \leqslant x \leqslant A^R, \\ 0, & \text{otherwise,} \end{cases}$$

then $\tilde{A}$ is called a triangular fuzzy number, where $\tilde{A} = (A^L, A^C, A^R)$, and $A^K(K = L, C, R)$ are called three parameter variables in $\tilde{A}$.

To the triangular fuzzy numbers, they satisfy the following properties.

Property 2.5.1. *Let* $\tilde{A} = (A^L, A^C, A^R), \tilde{B} = (B^L, B^C, B^R), k \in \mathscr{R}$. *Then*

(1) $\tilde{A} + \tilde{B} = (A^L + B^L, A^C + B^C, A^R + B^R)$;
(2) $\tilde{A} - \tilde{B} = (A^L - B^R, A^C - B^C, A^R + B^L)$;
(3) $k\tilde{A} = \begin{cases} (kA^L, kA^C, kA^R), \text{ when } & k \geqslant 0, \\ (kA^R, kA^C, kA^L), \text{ when } & k < 0. \end{cases}$

Besides the properties above, any two triangular fuzzy numbers can be compared with each other, that is

Definition 2.5.3. Let $\tilde{A} = (A^L, A^C, A^R), \tilde{B} = (B^L, B^C, B^R), k \in \mathscr{R}$, we have
(1) $\tilde{A} < \tilde{B}$ if only if $A^L < B^L, A^C < B^C$, and $A^R < B^R$;
(2) $\tilde{A} = \tilde{B}$ if only if $A^L = B^R, A^C = B^C$, and $A^R = B^L$;
(3) $A > B$ if only if $A^L > B^L, A^C > B^C$, and $A^R > B^R$.

2.5.2 Distance between Two Triangular Fuzzy Numbers

In order to estimate the regression parameters in the fuzzy linear regression models, first we introduce a new conception as follows.

Definition 2.5.4. Let $\tilde{A} = (A^L, A^C, A^R), \tilde{B} = (B^L, B^C, B^R), k \in \mathscr{R}$. Then we define

(1) Left distance: $d_L(\tilde{A}, \tilde{B}) = (A^L - B^L)^2$;
(2) Center distance: $d_C(\tilde{A}, \tilde{B}) = (A^C - B^C)^2$;
(3) Right distance: $d_R(\tilde{A}, \tilde{B}) = (A^R - B^R)^2$.

Obviously, from the definition above, we know that they are the distance square between the points that the three parameter variables correspond to the rectangular coordinate system in fact, so they are an ordinary distance. Thus follows the next.

Property 2.5.2. *Let $\tilde{A} = (A^L, A^C, A^R), \tilde{B} = (B^L, B^C, B^R), k \in \mathscr{R}$. Then*

(1) $d_L(\tilde{A}, \tilde{B}) \geqslant 0, d_C(\tilde{A}, \tilde{B}) \geqslant 0, d_R(\tilde{A}, \tilde{B}) \geqslant 0.$
(2) $d_L(\tilde{A}, \tilde{B}) = d_L(\tilde{B}, \tilde{A}), d_C(\tilde{A}, \tilde{B}) = d_C(\tilde{B}, \tilde{A}), d_R(\tilde{A}, \tilde{B}) = d_R(\tilde{B}, \tilde{A}).$
(Here we define $d_L(\tilde{B}, \tilde{A}) = (B^L - A^L)^2$*, and the same to the others).*
(3) $d_L(\tilde{A}, \tilde{B}) = d_C(\tilde{A}, \tilde{B}) = d_R(\tilde{A}, \tilde{B}) = 0.$
(4) $d_L(k\tilde{A}, k\tilde{B}) = k^2 d_L(\tilde{A}, \tilde{B}), d_C(k\tilde{A}, k\tilde{B}) = k^2 d_C(\tilde{A}, \tilde{B}),$
$d_R(k\tilde{A}, k\tilde{B}) = k^2 d_R(\tilde{A}, \tilde{B}), k \geqslant 0.$

Proof: From the Definition 2.5.1, (1) and (3) are obviously correct.

Now we only prove the left distance in (2) and (4), and the same to the others. Then, we have
(2) $d_L(\tilde{A}, \tilde{B}) = (A^L - B^L)^2 = (B^L - A^L)^2 = d_L(B, A)$.
(4) $d_L(k\tilde{A}, k\tilde{B}) = (kA^L - kB^L)^2 = (kB^L - kA^L)^2 = d_L(kB, kA)$.

2.5.3 Fuzzy Linear Regression

Now we consider the following fuzzy linear regression model

$$\tilde{y} = \tilde{a}_0 + \tilde{a}_1 x_1 + \tilde{a}_2 x_2, \ x_1, \ x_2 \geqslant 0, \tag{2.5.1}$$

where $\widetilde{y}, \widetilde{a}_0, \widetilde{a}_1$ and $\widetilde{a}_2$ are triangular fuzzy numbers with $\tilde{y} = (y^L, y^C, y^R), \tilde{a}_0 = (a_0^L, a_0^C, a_0^R)$, $\tilde{a}_1 = (a_1^L, a_1^C, a_1^R)$ and $\tilde{a}_2 = (a_2^L, a_2^C, a_2^R)$, and also all parameter variables are nonnegative real numbers.

Suppose x_{1i}, x_{2i} and $\widetilde{y}_i (i = 1, 2, \cdots, N)$ to be real input data and fuzzy output data, now we will calculate the estimated values of $\widetilde{a}_0, \widetilde{a}_1$ and $\widetilde{a}_2$ of Model (2.5.1). In many papers, the distance between two fuzzy numbers is mostly adopted by [Xu98], such that the optimal estimated values are got. Here we will introduce a new method.

To the fuzzy linear regression problems, we all know what most important is that we should make the error much less between the observed values and practical ones. But to Model (2.5.1), these data are all triangular fuzzy numbers, i.e., (y_i^L, y_i^C, y_i^R) and

$$(a_0^L + a_1^L x_{1i} + a_2^L x_{2i}, a_0^C + a_1^C x_{1i} + a_2^C x_{2i}, a_0^R + a_1^R x_{1i} + a_2^R x_{2i}).$$

According to the previous analysis,we can consider the corresponding parameter variables of the above. If the less errors between observed values and practical ones are, the less total error is. So the fuzzy linear regression problem is transformed to

$$\begin{cases} \min d_L(\widetilde{a}_0 + \widetilde{a}_1 x_1 + \widetilde{a}_2 x_2, \widetilde{y}) = \min \sum\limits_{i=1}^{N} (y_i^L - a_0^L - a_1^L x_{1i} - a_2^L x_{2i})^2, \\ \min d_C(\widetilde{a}_0 + \widetilde{a}_1 x_1 + \widetilde{a}_2 x_2, \widetilde{y}) = \min \sum\limits_{i=1}^{N} (y_i^C - a_0^C - a_1^C x_{1i} - a_2^C x_{2i})^2, \\ \min d_R(\widetilde{a}_0 + \widetilde{a}_1 x_1 + \widetilde{a}_2 x_2, \widetilde{y}) = \min \sum\limits_{i=1}^{N} (y_i^R - a_0^R - a_1^R x_{1i} - a_2^R x_{2i})^2. \end{cases}$$

According to the least square method, suppose

$$\begin{aligned} \frac{\partial d_L}{\partial a_l^L} &= \sum_{i=1}^{N} (y_i^L - a_0^L - a_1^L x_{1i} - a_2^L x_{2i}) = 0 (l = 0, 1, 2), \\ \frac{\partial d_C}{\partial a_l^C} &= \sum_{i=1}^{N} (y_i^C - a_0^C - a_1^C x_{1i} - a_2^C x_{2i}) = 0 (l = 0, 1, 2), \\ \frac{\partial d_R}{\partial a_l^R} &= \sum_{i=1}^{N} (y_i^R - a_0^R - a_1^R x_{1i} - a_2^R x_{2i}) = 0 (l = 0, 1, 2). \end{aligned} \tag{2.5.2}$$

For first formula of (2.5.2), we have

$$\begin{cases} n a_0^L + \sum\limits_{i=1}^{N} x_{1i} a_1^L + \sum\limits_{i=1}^{N} x_{2i} a_2^L = \sum\limits_{i=1}^{N} y_i^L, \\ \sum\limits_{i=1}^{N} x_{1i} a_0^L + \sum\limits_{i=1}^{N} x_{1i}^2 a_1^L + \sum\limits_{i=1}^{N} x_{1i} x_{2i} a_2^L = \sum\limits_{i=1}^{N} x_{1i} y_i^L, \\ \sum\limits_{i=1}^{N} x_{2i} a_0^L + \sum\limits_{i=1}^{N} x_{1i} x_{2i} a_1^L + \sum\limits_{i=1}^{N} x_{2i}^2 a_2^L = \sum\limits_{i=1}^{N} x_{2i} y_i^L, \end{cases} \tag{2.5.3}$$

when $\triangle = \begin{vmatrix} N & \sum\limits_{i=1}^{N} x_{1i} & \sum\limits_{i=1}^{N} x_{2i} \\ \sum\limits_{i=1}^{N} x_{1i} & \sum\limits_{i=1}^{N} x_{1i}^2 & \sum\limits_{i=1}^{N} x_{1i} x_{2i} \\ \sum\limits_{i=1}^{N} x_{2i} & \sum\limits_{i=1}^{N} x_{1i} x_{2i} & \sum\limits_{i=1}^{N} x_{2i}^2 \end{vmatrix} \neq 0$, by the aid of the Cramer rule, we have

$$a_0^L = \frac{\triangle_1}{\triangle}, a_1^L = \frac{\triangle_2}{\triangle}, a_2^L = \frac{\triangle_3}{\triangle}, \tag{2.5.4}$$

where $\triangle_j$ replaces the element of $j(j = 1, 2, 3)$ column in $\triangle$ into the term of equations $\sum_{i=1}^{N} y_i^L, \sum_{i=1}^{N} x_{1i}y_i^L, \sum_{i=1}^{N} x_{2i}y_i^L$, respectively.

Similarly, consider the second and third forms of (2.5.2), we can get

$$a_0^C = \frac{\triangle_1'}{\triangle}, a_1^C = \frac{\triangle_2'}{\triangle}, a_2^C = \frac{\triangle_3'}{\triangle}; a_0^R = \frac{\triangle_1''}{\triangle}, a_1^R = \frac{\triangle_2''}{\triangle}, a_2^R = \frac{\triangle_3''}{\triangle}, \tag{2.5.5}$$

where $\triangle_j'$ and $\triangle_j''$ are replacement of the element of $j(j = 1, 2, 3)$ column in $\triangle$ into the term of equations, respectively

$$\sum_{i=1}^{N} y_i^C, \sum_{i=1}^{N} x_{1i}y_i^C, \sum_{i=1}^{N} x_{2i}y_i^C \text{ and } \sum_{i=1}^{N} y_i^R, \sum_{i=1}^{N} x_{1i}y_i^R, \sum_{i=1}^{N} x_{2i}y_i^R.$$

Thus the estimated values of $\widetilde{a}_0$, $\widetilde{a}_1$ and $\widetilde{a}_2$ are

$$\widehat{\widetilde{a}}_0 = (a_0^L\ ,\ a_0^C\ ,\ a_0^R), \widehat{\widetilde{a}}_1 = (a_1^L\ ,\ a_1^C\ ,\ a_1^R), \widehat{\widetilde{a}}_2 = (a_2^L\ ,\ a_2^C\ ,\ a_2^R). \tag{2.5.6}$$

Definition 2.5.5. The parameter variables a_l^L , a_l^C and a_l^R $(l = 0, 1, 2)$ are called optimal estimated parameters of Model (2.5.1) if only if they satisfy (2.5.2), and the corresponding solutions (2.5.6) are called the optimal estimated values in (2.5.1).

So the estimated regression equation is

$$\widetilde{y} = \widehat{\widetilde{a}}_0 + \widehat{\widetilde{a}}_1 x_1 + \widehat{\widetilde{a}}_2 x_2\ ,\ x_1\ ,\ x_2 \geqslant 0. \tag{2.5.7}$$

Example 2.5.1 [Xu98]: The sales at a certain product on the market can be seen from Table 2.5.1

Table 2.5.1. Product Sales in Years

Year (x_i)	Amount of sales $(\widetilde{y}_i)$ (Unit:10^4 pieces)
1987	(228, 230, 231)
1988	(233, 236, 238)
1989	(239, 241, 244)

Try to estimate the amount of sales in 1990.

According to the formula (2.5.4) and (2.5.6), we can get $a_1^L = 222.3, a_2^L = 5.5; a_1^C = 224.7, a_2^C = 5.5; a_1^R = 224.7, a_2^R = 6.5$.

Therefore the regression equation is

$\widetilde{y} = (222.3, 224.7, 224.7) + (5.5, 5.5, 6.5)(x - 1986)$,

at $x = 1990$, we have $\widetilde{y} = (244.3, 246.7, 250.7)$.

2.5.4 Error Analysis

To Model (2.5.1), we get the following data $(\widetilde{y}_i, x_{1i}, x_{2i})(i = 1, 2, \cdots, N)$ by making observations. Then the practical values and observed values of $\widetilde{y}$ are $\widetilde{y}_i = (a_0^L + a_1^L x_{1i} + a_2^L x_{2i}, a_0^C + a_1^C x_{1i} + a_2^C x_{2i}, a_0^R + a_1^R x_{1i} + a_2^R x_{2i})$, and $\widehat{\widetilde{y}}_i = (y_i^L\ ,\ y_i^C\ ,\ y_i^R)$, respectively.

We have already got the estimated values of every parameter variable, now we analyze the left parameter variables, and the same to the others.

In fact, to Model (2.5.1), by Definition 2.5.3, we have

$$y^L = a_0^L + a_1^L x_1 + a_2^L x_2, \tag{2.5.8}$$

where y^L, a_l^L and $x_j (l = 0, 1, 2; j = 1, 2)$ are nonnegatively real numbers.

Obviously the Equation (2.5.8) can be regarded as an ordinary linear regression model, so we can estimate the previous values when entering ordinary cases. Thus, according to the properties of the ordinary linear regression, the estimated values a_0^L , a_1^L and a_2^L are unbiased estimations, and also the variance is the same to an ordinary case.

2.5.5 Comparison of Two-Kind Distance Formula

To the fuzzy linear regression (2.5.1), in most of papers, there is the following definition about the distance between two fuzzy numbers, that is

$$D^2(\tilde{A}, \tilde{B}) = \int_0^1 f(a) d^2(A_\alpha, B_\alpha) d\alpha, \tag{2.5.9}$$

where $d^2(A_\alpha, B_\alpha) = (\underline{A}_\alpha - \underline{B}_\alpha)^2 + (\overline{A}_\alpha - \overline{B}_\alpha)^2$, and $A_\alpha = [\underline{A}_\alpha, \overline{A}_\alpha], B_\alpha = [\underline{B}_\alpha, \overline{B}_\alpha]$. $f(\alpha)$ is a monotonously increasing function at $[0, 1]$, and

$$f(0) = 0, \int_0^1 f(\alpha) d\alpha = \frac{1}{2}.$$

If we use the distance above and by the differential, integral and the least-square method to the Model (2.5.1), and take $f(\alpha) = \alpha$, we can get [Lin01]

$$\sum_{i=1}^n x_{1i}^2 a_1^C + \sum_{i=1}^n x_{1i}^2 a_1^R = \sum_{i=1}^n x_{1i}(y_i^C + y_i^R), \tag{2.5.10}$$

$$\sum_{i=1}^n x_{1i}^2 a_1^L + \sum_{i=1}^n x_{1i}^2 a_1^C = \sum_{i=1}^n x_{1i}(y_i^L + y_i^C), \tag{2.2.11}$$

$$\sum_{i=1}^n x_{1i}^2 a_1^L + 6\sum_{i=1}^n x_{1i}^2 a_1^C + \sum_{i=1}^n x_{1i}^2 a_1^R = \sum_{i=1}^n x_{1i}(y_i^L + 6y_i^C + y_i^R). \tag{2.2.12}$$

Then according to the Cramer's rule, we can get values of a_1^L, a_1^C and a_1^R, and by using the same method, we can also calculate the others.

Compare (2.5.3) with (2.5.10) (2.5.11) and (2.5.12), obviously, in general, by using different distance, we can get different parameter estimated values. So, to the above distance (2.5.9), the process of calculation is more complex and the properties about parameter variables are not the same as this section. Therefore, the method of this section is more direct, above all it is provided with some practical values.