سیزدهمین کنفرانس
آمار ایران
۲-۴ شهریور ۱۳۹۵

مدل‌های رگرسیون خطی فازی بر اساس روش‌های کمترین قدرمطلق و مربع خطا

سپیده مقدوری[۱]، محمدقاسم اکبری [۲]

[۱]کارشناس ارشد ، دانشگاه بیرجند ، گروه آمار

[۲]عضو هیأت علمی گروه آمار،

چکیده: در این مقاله به تخمین ضرایب رگرسیون خطی فازی در حالتی که خروجی و ضرایب فازی هستند، با استفاده از متر هاسدورف و مترهای پیشنهادی k و d که توسط نویسنده معرفی شده‌اند، به روش کمترین قدرمطلق خطا و کمترین مربعات، می‌پردازیم و نتایج حاصل را با روش محققین قبلی از نظر سه معیار نیکویی برازش، مقایسه می‌کنیم.

واژه‌های کلیدی: عدد فازی LR، رگرسیون فازی، نیکویی برازش، روش کمترین مربعات، روش کمترین قدرمطلق خطا.

کد موضوع‌بندی ریاضی (۲۰۱۰): 62J05، 62J86.

۱ مقدمه

نظریه مجموعه‌های فازی، توسط پروفسور لطفی زاده (۱۹۵۶)؛ دانشمند ایرانی و استاد دانشگاه برکلی آمریکا، ارائه شد. در نظریه مجموعه‌های معمولی، هر مجموعه با یک ویژگی خوش تعریف مشخص می‌شود. اگر یک شیء دارای آن ویژگی باشد، عضو مجموعه‌ی متناظر است و اگر نباشد، عضو آن نیست. مثلاً، اگر مجموعه مرجع X، فرض شود، می‌توان یک تابع نشانگر تعریف کرد که برای هر عضو متعلق به مجموعه A که با یک ویژگی خوش تعریف مشخص می شود، مقدار یک و برای هر عضو غیر متعلق به A، مقدار صفر بگیرد، یعنی

$$I_A(x) = \begin{cases} ۱ & x \in A, \\ ۰ & x \notin A. \end{cases}$$

بنا به پیشنهاد زاده، اگر برد تابع نشانگر را ازمجموعه دو عضوی $\{۰, ۱\}$ به بازه $[۰, ۱]$ توسعه دهیم، یک تابع عضویت خواهیم داشت که به هر x از X، عددی را از بازه $[۰, ۱]$ نسبت می‌دهد. برای تمایز یک مجموعه معمولی از یک مجموعه فازی، از علامت $\sim$ در بالای آن استفاده می‌شود. تابع عضویت مجموعه فازی $\tilde{A}$ را با نماد $\mu_{\tilde{A}}(x)$ و یا به طور مختصر به صورت $\tilde{A}(x)$ نمایش می‌دهیم. مجموعه فازی $\tilde{A}$ از مجموعه مرجع X به صورت زیر تعریف می‌شود:

$$\left\{(x, \tilde{A}(x)) : x \in X\right\}$$

[۲]سپیده مقدوری : sepidehmaghdoori@gmail.com

که $[\circ, 1]$ $\tilde{A}(x) : X \to$ تابع عضویت x در $\tilde{A}$ است.

۲ مفاهیم اولیه

تعریف ۱.۲. مجموعه همه عناصری از X را که درجه عضویت آنها در مجموعه فازی $\tilde{A}$ حداقل به بزرگی γ باشد، γ -برش $\tilde{A}$ (مجموعه تراز γام $\tilde{A}$) نامیده و با $\tilde{A}[\gamma]$ نشان داده می‌شود:

$$\tilde{A}[\gamma] = \{x \in X | \mu_{\tilde{A}}(x) \geq \gamma\}, \quad \circ < \gamma \leq 1.$$

تعریف ۲.۲. مجموعه فازی $\tilde{A}$ از R را یك عدد فازی گوییم، اگر

۱) $\tilde{A}$ تك نمائی باشد. یعنی یك $x_\circ \in R$ وجود داشته باشد که $\tilde{A}(x_\circ) = 1$.

۲) γ-برش‌های $\tilde{A}$، به ازای هر $\gamma \in (\circ, 1]$، بازه‌های بسته و کراندار باشند.

تعریف ۳.۲. فرض کنید $\tilde{A}$ یك عدد فازی روی مجموعه مرجع X باشد. $\tilde{A}$ را یك عدد فازی LR گویند، اگر تابع عضویت آن به صورت زیر باشد:

$$\mu_{\tilde{A}}(x) = \begin{cases} L(\dfrac{m-x}{\alpha}), & x \leq m, \\ R(\dfrac{x-m}{\beta}), & x > m, \end{cases}$$

که در آن L و R توابعی غیر صعودی از $\mathcal{R}^+$ به $[\circ, 1]$ هستند و $L(\circ) = R(\circ) = 1$ است. یک عدد فازی LR مانند $\tilde{A}$ را با نماد $\tilde{A} = (m, \alpha, \beta)_{LR}$ نشان می‌دهیم. عدد حقیقی m مقدار نما، α و β به ترتیب پهنای چپ و پهنای راست $\tilde{A}$ نامیده می‌شوند. همچنین $\tilde{A}$ را یك عدد فازی مثلثی نامند و با نماد $\tilde{A} = (m, \alpha, \beta)_T$ نشان می‌دهند، اگر: $L(x) = R(x) = max\{\circ, 1 - |x|\}$.

در صورتی که $L = R$ و $\alpha = \beta$، آنگاه عدد فازی را متقارن نامیده و با $\tilde{A} = (m, \alpha)_L$ نشان داده می‌شود.

تعریف ۴.۲. فرض کنید $\tilde{A} = (m, \alpha, \beta)_{LR}$ یك عدد فازی LR باشد آن‌گاه γ-برش $\tilde{A}$ به صورت زیر تعریف می‌شود:

$$\tilde{A}[\gamma] = [m - \alpha L^{-1}(\gamma), m + \beta R^{-1}(\gamma)]$$

قضیه ۵.۲. اگر $\tilde{A} = (m_a, \alpha_a, \beta_a)_{LR}$ و $\tilde{B} = (m_b, \alpha_b, \beta_b)_{LR}$ و $\lambda \in R$ آنگاه داریم:

$$\begin{aligned} \tilde{A} \oplus \tilde{B} &= (m_a, \alpha_a, \beta_a)_{LR} \oplus (m_b, \alpha_b, \beta_b)_{LR} \\ &= (m_a + m_b, \alpha_a + \alpha_b, \beta_a + \beta_b)_{LR}, \end{aligned}$$

$$\begin{aligned} \tilde{A} \ominus \tilde{B} &= (m_a, \alpha_a, \beta_a)_{LR} \ominus (m_b, \alpha_b, \beta_b)_{LR} \\ &= (m_a - m_b, \alpha_a + \beta_b, \beta_a + \alpha_b)_{LR}, \end{aligned}$$

$$\lambda \otimes \tilde{A} = \begin{cases} (\lambda m_a, \lambda\alpha_a, \lambda\beta_a)_{LR}, & \lambda > \circ, \\ (\lambda m_a, -\lambda\beta_a, -\lambda\alpha_a)_{RL}, & \lambda < \circ. \end{cases}$$

تعریف ۶.۲. لیو (۲۰۱۳) معیاری را برای مقایسه بین اعداد فازی و اعداد حقیقی معرفی کرده که از آن به عنوان تابع درجه اعتبار (میزان کوچکی عدد فازی $\tilde{A}$ نسبت به مقدار حقیقی x) یاد کرده و به صورت زیر نمایش داده:

$$C : F(\mathcal{R}) \to [\circ, 1]$$

که در آن

$$C(\tilde{A} \leq x) = \frac{sup_{y \leq x}\mu_{\tilde{A}}(y) + 1 - sup_{y > x}\mu_{\tilde{A}}(y)}{2}.$$

تعریف ۷.۲. کوچکترین کران بالای مجموعه تمام عناصری از $\tilde{A}[\circ]$ را که تابع درجه اعتبار آن‌ها حداقل به بزرگی α باشد، α-شک $\tilde{A}$ نامیده و با $\tilde{A}_\alpha$ نمایش می‌دهیم، یعنی :

$$\tilde{A}_\alpha = inf\{x \in \tilde{A}[\circ] : C\{\tilde{A} \leq x\} \geq \alpha\}$$

حال فرض کنید که $\tilde{A} = (a, l, r)_T$ و $\tilde{B} = (b, l', r')_T$ دو عدد فازی مثلثی باشند. در این صورت با استفاده از مفهوم α-شک، مترهای d و k را بین این دو عدد فازی به ترتیب به صورت زیر تعریف کرده و از آن‌ها برای تخمین پارامترهای مدل رگرسیون فازی استفاده می‌کنیم.

$$\begin{aligned}
d(\tilde{A}, \tilde{B}) &= \sqrt{\int_\circ^۱ (\tilde{A}_\alpha - \tilde{B}_\alpha)^۲ d\alpha} \\
&= \sqrt{\int_\circ^{\circ/۵} (\tilde{A}_\alpha - \tilde{B}_\alpha)^۲ d\alpha + \int_{\circ/۵}^۱ (\tilde{A}_\alpha - \tilde{B}_\alpha)^۲ d\alpha} \\
&= \sqrt{\int_\circ^{\circ/۵} (\tilde{A}_\alpha - \tilde{B}_\alpha)^۲ d\alpha + \int_\circ^{\circ/۵} (\tilde{A}_{۱-\alpha} - \tilde{B}_{۱-\alpha})^۲ d\alpha} \\
&= \sqrt{(a-b)^۲ + \frac{(a-b)}{۲}[(r-l) - (r'-l')] + \frac{(r-r')^۲ + (l-l')^۲}{۶}} \\
k(\tilde{A}, \tilde{B}) &= \int_\circ^۱ |\tilde{A}_\alpha - \tilde{B}_\alpha| d\alpha
\end{aligned}$$

تعریف ۸.۲. فرض کنید که $K_c(\mathcal{R}^n)$ خانواده‌ای از همه مجموعه‌های محدب فشرده ناتهی n بعدی روی $\mathcal{R}^n$ باشد، در این صورت متر هاسدورف بین دو مجموعه $A, B \in K_c(\mathcal{R}^n)$ به صورت زیر تعریف می‌شود:

$$d_H(A, B) = max\{\sup_{b \in B} \inf_{a \in A} \|a - b\|, \sup_{a \in A} \inf_{b \in B} \|a - b\|\}$$

که در آن $\|.\|$ نرم اقلیدسی است.

لم ۹.۲. اگر روی $\mathcal{R}$، $I_۱ = [a_۱, a_۲]$ و $I_۲ = [b_۱, b_۲]$ باشد، آنگاه:

$$\begin{aligned}
d_H(I_۱, I_۲) &= max\{|a_۱ - b_۱|, |a_۲ - b_۲|\} \\
&= |midI_۱ - midI_۲| + |sprI_۱ - sprI_۲|
\end{aligned}$$

که در آن $midI_۱ = \dfrac{a_۱ + a_۲}{۲}$ و $sprI_۱ = \dfrac{a_۲ - a_۱}{۲}$ است.

تعریف ۱۰.۲. متر هاسدورف تعمیم‌یافته بین اعداد فازی $\tilde{A}$ و $\tilde{B}$ به صورت زیر تعریف می‌شود:

$$D_p(\tilde{A}, \tilde{B}) = \begin{cases} (\int_\circ^۱ [d_H(\tilde{A}[\gamma], \tilde{B}[\gamma])]^p d\gamma)^{\frac{۱}{p}} & , p \in [۱, \infty) \\ \\ sup_{\gamma \in [\circ,۱]} d_H(\tilde{A}[\gamma], \tilde{B}[\gamma]) & , p = \infty \end{cases}$$

قضیه ۱۱.۲. فرض کنید که $\tilde{M} = (m, \lambda_m)_L$ و $\tilde{N} = (n, \lambda_n)_L$ اعداد فازی LR متقارن باشند، در این صورت با توجه به لم ۹.۲ متر هاسدورف بین این دو عدد فازی به صورت زیر می‌باشد:

$$d_H(\tilde{M}[\gamma], \tilde{N}[\gamma]) = |m - n| + L^{-۱}(\gamma)|\lambda_m - \lambda_n|$$

همچنین با استفاده از قضیه فوق، برای اعداد فازی LR متقارن $\tilde{M} = (m, \lambda_m)_L$ و $\tilde{N} = (n, \lambda_n)_L$ متر زیر به دست می‌آید:

$$D_۱(\tilde{M}, \tilde{N}) = |m - n| + L_۱|\lambda_m - \lambda_n|, \qquad L_۱ = \int_\circ^۱ L^{-۱}(\gamma) d\gamma.$$

۳ نیکویی برازش برای رگرسیون فازی

در بررسی مدل‌های رگرسیونی فازی، باید از معیار یا معیارهایی استفاده کرد تا مدل‌های مختلف را با هم مقایسه کرد. در راستای رسیدن به این هدف در این بخش دو معیار خطا و یک اندازه مشابهت را معرفی می‌کنیم.

معیار خطا

برای ارزیابی اعتبار یک مدل رگرسیون فازی، کیم و بیشو (۱۹۹۸) از تفاضل بین مقادیر عضویت پاسخ‌های فازی مشاهده شده و برآورد شده به عنوان یک معیار استفاده کردند که در زیر آن را معرفی می‌کنیم.

تعریف ۳.۱. خطای نسبی برآورد: فرض کنید $\tilde{y}_i$ و $\hat{\tilde{y}}_i$ به ترتیب عدد فازی پاسخ مشاهده شده و عدد فازی برآورد شده از یک مدل رگرسیونی باشد. تفاضل نسبی بین توابع عضویت این دو عدد به صورت زیر تعریف می شود:

$$E_1(\tilde{y}_i, \hat{\tilde{y}}_i) = \frac{\int_{S_{\tilde{y}_i} \cup S_{\hat{\tilde{y}}_i}} |\tilde{y}_i(x) - \hat{\tilde{y}}_i(x)| dx}{\int_{S_{\tilde{y}_i}} \tilde{y}_i(x) dx}$$

که در آن $S_{\tilde{y}_i}$ و $S_{\hat{\tilde{y}}_i}$ به ترتیب تکیه گاه $\tilde{y}$ و $\hat{\tilde{y}}_i$ هستند. از این معیار چن و دانگ (۲۰۰۸) و لیو و وانگ (۲۰۰۹) نیز برای نیکویی برازش مدل رگرسیون فازی خود استفاده کرده‌اند.

تعریف ۳.۲. خطای برآورد: این شاخص، تفاضل توابع عضویت متغیر پاسخ مشاهده شده و برآورد شده را محاسبه می‌کند و به صورت زیر تعریف می‌شود:

$$E_2(\tilde{y}_i, \hat{\tilde{y}}_i) = \int_{S_{\tilde{y}_i} \cup S_{\hat{\tilde{y}}_i}} |\tilde{y}_i(x) - \widehat{\tilde{y}}_i(x)| dx, \quad i = 1, \ldots, n$$

چن و دانگ (۲۰۰۸)، لیو و وانگ (۲۰۰۹)، کائو و چیو (۲۰۰۲) از این معیار برای نیکویی برازش مدل رگرسیون فازی خود استفاده کرده‌اند.

اندازه مشابهت

شاخص تطبیق یا اندازه مشابهت دو مجموعه فازی میزان مشابهت آن دو را به هم بیان می‌کند. یکی از اندازه مشابهت‌هایی که بین دوعدد فازی $\tilde{y}_i$ و $\hat{\tilde{y}}_i$ می‌توان تعریف کرد به صورت زیر است:

$$\begin{aligned} S(i) &= \frac{card(\tilde{y}_i(x) \cap \widehat{\tilde{y}}_i(x))}{card(\tilde{y}_i(x) \cup \widehat{\tilde{y}}_i(x))} \\ &= \frac{\int min\{\tilde{y}_i(x), \widehat{\tilde{y}}_i(x)\} dx}{\int max\{\tilde{y}_i(x), \widehat{\tilde{y}}_i(x)\} dx}, \quad i = 1, \ldots, n \end{aligned}$$

لازم به ذکر می‌باشد که مقادیری که $E_1(\tilde{y}_i, \widehat{\tilde{y}}_i)$ و $E_2(\tilde{y}_i, \widehat{\tilde{y}}_i)$ می‌توانند اختیار کنند $[0, \infty]$ است و برد S، $[0, 1]$ می‌باشد. بنابراین به منظور مقایسه این شاخص‌ها، با یک تغییر متغیر معیارهای جدیدی به صورت $G_1(\tilde{y}_i, \widehat{\tilde{y}}_i) = \frac{1}{1 + E_1(\tilde{y}_i, \widehat{\tilde{y}}_i)}$ و $G_2(\tilde{y}_i, \widehat{\tilde{y}}_i) = \frac{1}{1 + E_2(\tilde{y}_i, \widehat{\tilde{y}}_i)}$ بدست می‌آید که محدوده تغییر آنها تبدیل به بازه $[0, 1]$ می شود. در عمل به منظور ارزیابی نیکویی برازش مدل، از روابط زیر استفاده می‌کنیم:

$$G_1 = \sum_{i=1}^{n} \frac{G_1(i)}{n}, \quad G_2 = \sum_{i=1}^{n} \frac{G_2(i)}{n}, \quad S = \sum_{i=1}^{n} \frac{S(i)}{n}$$

مدلی که دارای G_1، G_2 و S بزرگتری باشد، برازش بهتری را به داده‌ها می‌دهد.

۴ رگرسیون خطی چندگانه بر پایه ورودی غیرفازی و خروجی فازی

در این بخش به سه روش متفاوت، پارامترهای فازی مدل را برآورد می‌کنیم .

الف) روش کمترین مربعات: مجموعه مقادیر مشاهده شده $(x_{i1}, ..., x_{ik}, \tilde{y}_i), i = 1, ..., n$ از جامعه مورد نظر را در نظر بگیرید. معادله رگرسیونی فازی زیر را به داده‌ها برازش می‌دهیم:

$$\tilde{y}_i = \tilde{\beta}_0 \oplus (\tilde{\beta}_1 \otimes x_{i1}) \oplus +... + (\tilde{\beta}_k \otimes x_{ik}), i = 1, ..., n \qquad (1.4)$$

که در آن:

$$\begin{cases} \tilde{y}_i = (y_i, l_i, r_i)_T & i = 1, ..., n, \\ \tilde{\beta}_j = (\beta_j, l_{\beta_j}, r_{\beta_j})_T & j = 0, 1, ..., k. \end{cases}$$

حال فرض کنید که $x_{ij} > 0$ بوده در این صورت معادله برآورد زیر را داریم:

$$\widehat{\tilde{y}_i} = (\sum_{j=0}^{k} \beta_j x_{ij}, \sum_{j=0}^{k} l_{\beta_j} x_{ij}, \sum_{j=0}^{k} r_{\beta_j} x_{ij})_T$$

مجموع مربعات خطای فازی بر اساس اندازه d^2 به صورت زیر بدست می‌آید:

$$\begin{aligned} SSE &= \sum_{i=1}^{n} d^2(\tilde{y}_i, \widehat{\tilde{y}_i}) \\ &= \sum_{i=1}^{n} \{(y_i - \sum_{j=0}^{k} \beta_j x_{ij})^2 \\ &+ \frac{1}{6}[(r_i - \sum_{j=0}^{k} r_{\beta_j} x_{ij})^2 \\ &+ (l_i - \sum_{j=0}^{k} l_{\beta_j} x_{ij})^2] \\ &+ \frac{1}{2}(y_i - \sum_{j=0}^{k} \beta_j x_{ij})[(r_i - l_i) \\ &- (\sum_{j=0}^{k} r_{\beta_j} x_{ij} - \sum_{j=0}^{k} l_{\beta_j} x_{ij})]\} \end{aligned}$$

حال برای اینکه بتوانیم به سادگی عبارت SSE را کمینه سازیم ماتریس‌های زیر را در نظر می‌گیریم:

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}, L_\beta = \begin{pmatrix} l_{\beta_0} \\ l_{\beta_1} \\ \vdots \\ l_{\beta_k} \end{pmatrix}, R_\beta = \begin{pmatrix} r_{\beta_0} \\ r_{\beta_1} \\ \vdots \\ r_{\beta_k} \end{pmatrix},$$

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, L = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}, R = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix},$$

و

$$X = \begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}.$$

در این صورت داریم:

$$\begin{aligned} SSE &= (Y - X\beta)'(Y - X\beta) + \frac{1}{2}(Y - X\beta)'(R - L) \\ &- \frac{1}{2}(Y - X\beta)'[XR_\beta - XL_\beta] \\ &+ \frac{1}{6}(R - XR_\beta)'(R - XR_\beta) \\ &+ \frac{1}{6}(L - XL_\beta)'(L - XL_\beta) \end{aligned}$$

با مشتق‌گیری نسبت به پارامترها و مینیمم کردن تابع فوق و در صورت وجود معکوس ماتریس $X'X$، داریم:

$$\widehat{\beta} = (X'X)^{-1}X'Y,\ \ \widehat{L}_\beta = (X'X)^{-1}X'L,\ \ \widehat{R}_\beta = (X'X)^{-1}X'R.$$

ب) روش کمترین قدرمطلق خطا با متر D_1: مدل رگرسیونی (۱.۴) را در نظر بگیرید. فرض کنید که :

$$\begin{cases} \tilde{y}_i = (y_i, l_i)_L & i = 1, ..., n, \\ \tilde{\beta}_j = (\beta_j, l_{\beta_j})_L & j = 0, 1, ..., k. \end{cases}$$

با فرض اینکه $x_{ij} > 0$ بوده، معادله برآورد زیر را داریم:

$$\widehat{\tilde{y}}_i = (\sum_{j=0}^{k}\beta_j x_{ij}, \sum_{j=0}^{k} l_{\beta_j}x_{ij})_L$$

حال به منظور برآورد پارامترها بر اساس روش چاچی و طاهری (۲۰۱۳) و با استفاده از متر D_1 مسئله برنامه‌ریزی غیر خطی زیر را در نظر می‌گیریم:

$$min_{\tilde{\beta}} \sum_{i=1}^{n}|y_i - \sum_{j=0}^{k}\beta_j x_{ij}| + L_1\sum_{i=1}^{n}|l_i - \sum_{j=0}^{k} l_{\beta_j}x_{ij}|$$

$$s.t. \quad l_{\beta_j} \in \mathcal{R}^+, \beta_j \in \mathcal{R},\ \ j = 0, 1, ..., k$$

لازم به ذکر است که چاچی و طاهری (۲۰۱۳) روش خود را با چندین محقق دیگر از جمله چوی و باکلی (۲۰۰۸)، کائو و چیو (۲۰۰۲) و تاناکا وهمکاران (۱۹۸۹) مقایسه کرده و ادعا کرده‌اند که روش آن‌ها بر اساس سه شاخص نیکویی برازش گفته شده، عملکرد بهتری داشته است. ما نشان خواهیم داد که مدل رگرسیونی که بر اساس متر پیشنهادی k بدست می‌آید، عملکرد بهتری در مقایسه با روش چاچی و طاهری (۲۰۱۳) خواهد داشت.

ج) روش کمترین قدرمطلق خطا با متر پیشنهادی k: مدل رگرسیونی (۱.۴) را در نظر بگیرید. فرض کنید که $\tilde{y}_i$ و $\tilde{\beta}_j$ها به ترتیب به صورت اعداد فازی مثلثی متقارن زیر باشند:

$$\begin{cases} \tilde{y}_i = (y_i, l_i)_T & i = 1, ..., n, \\ \tilde{\beta}_j = (\beta_j, l_{\beta_j})_T & j = 0, 1, ..., k. \end{cases}$$

حال فرض کنید که $x_{ij} > 0$ بوده در این صورت معادله برآورد زیر را داریم:

$$\widehat{\tilde{y}}_i = (\sum_{j=0}^{k}\beta_j x_{ij}, \sum_{j=0}^{k} l_{\beta_j}x_{ij})_T$$

به منظور برآورد پارامترها با متر k، به شیوه زیر عمل می‌کنیم:

$$min \sum_{i=1}^{n}\int_0^1 |y_i - \sum_{j=0}^{k}\beta_j x_{ij} + (1 - 2\alpha)(\sum_{j=0}^{k} l_{\beta_j}x_{ij} - l_i)|d\alpha$$

$$s.t. \quad l_{\beta_j} \in \mathcal{R}^+, \beta_j \in \mathcal{R},\ \ j = 0, 1, ..., k.$$

مثال ۱.۴. مجموعه داده‌های جدول ۱ با ورودی دقیق و خروجی فازی را که توسط تاناکا وهمکاران (۱۹۸۹) ارائه شده است، در نظر بگیرید.

جدول ۱: داده‌های مثال۱.۴

i	x_i	$\tilde{y}_i = (y_i, l_i)_T$
۱	۱	$(۸/۰, ۱/۸)_T$
۲	۲	$(۶/۴, ۲/۲)_T$
۳	۳	$(۹/۵, ۲/۶)_T$
۴	۴	$(۱۳/۵, ۲/۶)_T$
۵	۵	$(۱۳/۰, ۲/۴)_T$

مدل بهینه بر اساس سه روش گفته شده به صورت زیر بدست می‌آید.

۱- روش کمترین مربعات:

$$\widehat{\tilde{y}} = (۴/۹۵ + ۱/۷۱x, ۱/۸۴ + ۰/۱۶x)_T$$

۲- روش کمترین قدرمطلق خطا با متر $D_۱$:

$$\widehat{\tilde{y}} = (۶/۴۴۴ + ۱/۳۱۱۲x, ۱/۸ + ۰/۲x)_T$$

۳- روش کمترین قدرمطلق خطا با متر k:

$$\widehat{\tilde{y}} = (۶/۴۳۳۸ + ۱/۳۱۳۹x, ۱/۶۹۰۴ + ۰/۱۴۲۱x)_T$$

مقایسه مدل‌های ذکر شده از نظر شاخص‌های S، $G_۱$ و $G_۲$ در جدول ۲ آورده شده است. همان طور که مشاهده می‌کنید روش سوم از

جدول ۲: مقایسه روش‌ها

روش	S	$G_۱$	$G_۲$
۱- روش کمترین مربعات	$۰/۴۰۸۷_{(۳)}$	$۰/۵۵۳۳_{(۳)}$	$۰/۳۶۱۹_{(۳)}$
۲- روش کمترین قدرمطلق خطا با متر $D_۱$	$۰/۵۰۰۸_{(۲)}$	$۰/۶۲۰۹_{(۲)}$	$۰/۴۴۸۰_{(۲)}$
۳-روش کمترین قدرمطلق خطا با متر k	$۰/۵۱۴۴_{(۱)}$	$۰/۶۵۰۳_{(۱)}$	$۰/۵۰۴۴_{(۱)}$

نظر هر سه شاخص رتبه اول را کسب کرده و لذا بهترین برازش را به داده‌ها داده است.

مثال ۲.۴. جدول ۳ را که شامل داده‌های کیم و بیشو (۱۹۹۸) به صورت ورودی دقیق و خروجی فازی است، درنظر بگیرید. مدل بهینه بر اساس سه روش گفته شده به صورت زیر بدست می‌آید.

۱- روش کمترین مربعات:

$$\widehat{\tilde{y}} = -۱۶/۷۹۵۷ - ۱/۰۹۸۹x_۱ - ۱/۱۷۹۸x_۲ + (۱/۸۵۵۹, ۰/۲۰۴۰)x_۳$$

۲- روش کمترین قدرمطلق خطا با متر $D_۱$:

$$\widehat{\tilde{y}} = -۲/۸۲۷۳ - ۰/۳۸۷۸x_۱ - ۱/۰۱۲۵x_۲ + (۰/۶۱۸۵, ۰/۱۷۹۰)x_۳$$

جدول ۳: داده‌های مثال۲.۴

i	$\tilde{y}_i = (y_i, l_i)_T$	x_{i1}	x_{i2}	x_{i3}
۱	$(۵/۸۳, ۳/۵۶)_T$	۲/۰۰	۰/۰۰	۱۵/۲۵
۲	$(۰/۸۵, ۰/۵۲)_T$	۰/۰۰	۵/۰۰	۱۴/۱۳
۳	$(۱۳/۹۳, ۸/۵۰)_T$	۱/۱۳	۱/۵۰	۱۴/۱۳
۴	$(۴/۰۰, ۲/۴۴)_T$	۲/۰۰	۱/۲۵	۱۳/۶۳
۵	$(۱/۶۵, ۱/۰۱)_T$	۲/۱۹	۳/۷۵	۱۴/۷۵
۶	$(۱/۵۸, ۰/۹۶)_T$	۰/۲۵	۳/۵۰	۱۳/۷۵
۷	$(۸/۱۸, ۴/۹۹)_T$	۰/۷۵	۵/۲۵	۱۵/۲۵
۸	$(۱/۸۵, ۱/۱۳)_T$	۴/۲۵	۲/۰۰	۱۳/۵۰

جدول ۴: مقایسه روش‌ها

روش	S	G_1	G_2
۱- روش کمترین مربعات	$۰/۱۸۴۹_{(۳)}$	$۰/۳۶۷۴_{(۳)}$	$۰/۲۳۲۸_{(۳)}$
۲- روش کمترین قدرمطلق خطا با متر D_1	$۰/۳۶۳۲_{(۱)}$	$۰/۴۷۷۶_{(۲)}$	$۰/۳۵۳۴_{(۲)}$
۳-روش کمترین قدرمطلق خطا با متر k	$۰/۳۵۹۶_{(۲)}$	$۰/۵۵۸۰_{(۱)}$	$۰/۴۲۶۵_{(۱)}$

۳- روش کمترین قدرمطلق خطا با متر k:

$$\begin{cases} \widehat{y} = -۲/۱۴۳۷ - ۰/۲۸۹x_1 - ۰/۷۱۷x_2 + ۰/۵۰x_3 \\ \widehat{l} = ۰/۱۷۹۷ + ۰/۰۵۷۴x_1 + ۰/۰۰۲۱x_2 + ۰/۰۶۸۵x_3 \end{cases}$$

مقایسه مدل‌های ذکر شده از نظر شاخص‌های S، G_1 و G_2 در جدول ۴ آورده شده است. مدل پیشنهادی بر اساس متر k از نظر شاخص‌های G_1 و G_2 رتبه اول را کسب کرده است.

بحث و نتیجه‌گیری

با توجه به جداول ۲ و ۴ مشاهده می‌کنیم که متر k عملکرد بهتری را نسبت به سایر روش‌ها از نظر شاخص‌های گفته شده، داراست.

مراجع

Chachi J. and Taheri S. M. (2013), A least-absolutes regression model for imprecise response based on the generalized hausdorff-metric, *Journal of Uncertain Systems* 7, 256-276.

Chen S. P. and Dang J. F. (2008), A variable spread fuzzy linear regression model with higher explanatory power and forecasting accuracy, *Information Sciences*. 178, 3973-3988.

Choi S. H. and Buckley J. J. (2008), Fuzzy regression using least absolute deviation estimators, *Soft Computing*. 12, 257-263.

Kao C. and Chyu C. L. (2002), A fuzzy linear regression model with better explanatory power, *Fuzzy Sets and Systems*. 126, 401-409.

Kim B. and Bishu R. R. (1998), Evaluation of fuzzy regression models by comparing membership functions,*Fuzzy Sets and Systems*. 100, 343-352.

Liu B. (2013), Uncertainty Theory,4th ed, Springer-Verlag, Berlin.

Lu J. and Wang R. (2009), An enhanced fuzzy linear regression model with more exible spreads, *Fuzzy Sets and Systems*. 160, 2505-2523.

Tanaka H., Hayashi I. and Watada J.(1989), Possibilistic linear regression analysis for fuzzy data, *European Journal of Operational Research*. 40, 389-396.

Zadeh L. A.(1956), *Fuzzy sets*, Information and Control. 8, 338-353.