

رگرسیون خطی فازی بردار پشتیبان با خطای فازی

اعظم مقدم، محسن عارفی* و محمدقاسم اکبری
گروه آمار، دانشکده علوم ریاضی و آمار، دانشگاه بیرجند

تاریخ دریافت: ۱۴۰۲/۷/۲۱

تاریخ پذیرش: ۱۴۰۲/۱۰/۱۵

نوع مقاله: علمی-پژوهشی

چکیده. در این مقاله، رویکردی برای برازش مدل رگرسیون خطی بر اساس بردارهای پشتیبان، زمانی که متغیر پاسخ، پارامترهای مدل و خطاها به صورت اعداد فازی باشند، ارائه شده است. در این روش، تابع هدف بر اساس مجموع قدرمطلق فواصل بین ابرصفحه‌های حاشیه‌ای غیرموازی بنا شده است. مدل ارائه شده نسبت به وجود داده‌های پرت از استواری مناسبی برخوردار است. مدل پیشنهادی بر اساس سه معیار نیکویی برازش با برخی مدل‌های دیگر مورد مقایسه قرار گرفته است.

۱. مقدمه

تحلیل رگرسیونی یکی از روش‌های معتبر در علم آمار برای بدست آوردن روابط بین مجموعه‌ای از متغیرهای مستقل و متغیر پاسخ است. رگرسیون فازی یکی از انواع رگرسیون‌ها است که هدف آن پیدا کردن ارتباط میان این متغیرها در فضای نادقیق است. اگر چه تلاش‌های زیادی در جهت برآورد پارامترهای مدل‌های رگرسیونی به شیوه‌های متفاوت صورت گرفته است اما همیشه مکمل‌هایی برای بهبود عملکرد وجود داشته است. گاهی این عملکردها بهینه‌تر از روش‌های موجود و گاهی نیز روشی موثر در گام‌های موفق در دیگر مقالات بوده است. شیوه

2010 Mathematics Subject Classification. 62J86, 03E72

* Corresponding author

E-mails: a.koohimoghadam@gmail.com, Arefi@Birjand.ac.ir, and g_z_akbari@birjand.ac.ir.

عبارات و کلمات کلیدی. بردار پشتیبان، رگرسیون فازی، خطای فازی، روش قدرمطلق خطا.

ارائه شده نیز می‌تواند علاوه بر بهبود عملکرد، شیوه‌ای موثر و تاثیرگذار در دیگر مقالات نیز باشد.

نخستین بار تاناکا و همکاران (۱۹۸۲) دیدگاه رگرسیون فازی خطی را گسترش داده‌اند. وپنیک (۱۹۹۵) اولین بار موضوع رگرسیون بردار پشتیبان را بر اساس تئوری یادگیری آماری معرفی نمود. زیربنای این روش مبتنی بر معرفی ساختار زیان و کمینه نمودن آن پایگذاری شده است. بسیاری از محققین تکنیک‌های متعددی را برای گنجاندن مفهوم نظریه مجموعه‌های فازی در مدل‌های رگرسیون آماری مرسوم پیشنهاد کرده‌اند. هانگ و وانگ (۲۰۰۳) به بررسی رگرسیون فازی خطی و غیرخطی بر مبنای بردارهای پشتیبان پرداخته‌اند. یائو و یو (۲۰۰۶) یک چارچوب اصلاحی برای ماشین بردار پشتیبان‌های متقارن طراحی کردند. ایده مورد نظر با کمینه کردن تابع هزینه ناشی از جمع تمام فواصل حاصل از هر نقطه مشاهدات تا خط حاشیه و خط مرکز پایگذاری می‌شود. رگرسیون فازی غیرخطی با ورودی دقیق و خروجی فازی توسط محققان زیادی با استفاده از بردار پشتیبان مورد بررسی قرار گرفته‌اند. از این موارد می‌توان به دیدگاه هائو و چیانگ (۲۰۰۸) و یانکویی (۲۰۱۲) اشاره نمود که به تابع هدف مورد نظر ابهامات ضرایب مدل نیز اضافه گردیده و علاوه بر آن از دیدگاه تاناکا و همکاران (۱۹۸۲) در تعریف محدودیت‌های مدل خود نیز از میزان درجه عضویت بهره گرفته‌اند. نقشی که روش بردار پشتیبان در مقایسه با دیگر مدل‌های رگرسیونی در تحلیل ایفا می‌کند، دارای مزیت‌هایی است که از آن جمله می‌توان به استوار بودن و افزایش سرعت همراه با دقت پیش‌بینی اشاره کرد. برای بررسی برخی مدل‌های رگرسیونی فازی در سال‌های اخیر به چن و نین (۲۰۲۰) و حسامیان و اکبری (۲۰۲۰ و ۲۰۲۱) مراجعه نمایید.

توسعه مدل‌های رگرسیون توسط انتقال به یک فضایی با ابعاد بالا که در آن داده‌های ورودی از طریق یک تابع خطی و یا غیرخطی نگاشت می‌شوند، کمک می‌کند تا بتوان پیچیدگی فضاهای موجود را کنترل نمود و با استفاده از تابع هسته به حل مسئله مورد نظر در رگرسیون خطی در این فضا پرداخت. برای بررسی برخی دیدگاه‌ها در رگرسیون فازی با استفاده از تابع هسته به چنگ و لی (۲۰۰۱)، باساک و همکاران (۲۰۰۷)، زائو و سان (۲۰۱۰)، وو و لائو (۲۰۱۰)، فرقانی و تبریزی (۲۰۱۱)، وانگ و زونگ (۲۰۱۴)، حسامیان و اکبری (۲۰۱۷)، بالاسندرام و منا (۲۰۱۸) و اسداللهی و همکاران (۲۰۲۱) مراجعه نمایید.

هدف اصلی این مطالعه معرفی یک مدل رگرسیون بردار پشتیبان (SVR) جدید مبتنی بر متغیرهای توضیحی و متغیر پاسخ فازی با حضور خطای فازی در مدل است. در مدل رگرسیون بردار پشتیبان خطا توسط کاربر تعیین می‌گردد. در این صورت کاربر باید به دنبال خطای بهینه

بگردد. اما در ایده مورد نظر، ابرصفحه‌های حاشیه‌ای برابر کران‌های متغیر پاسخ فازی برآورده شده فرض شده‌اند. از آنجایی که خطا یک مقدار فازی فرض شده است، به همین دلیل مقدار پهنای خطا یک پارامتر در نظر گرفته شده است. از طرفی برآورد فاصله بین دو ابرصفحه‌ای حاشیه‌ای کمک می‌کند تا دیگر همانند رگرسیون بردار پشتیبان نیاز به مشخص کردن خطا توسط کاربر نباشد و چون یک پارامتر است خود به خود به دست آید. همانطور که می‌دانیم در روش رگرسیون بردار پشتیبان، حداکثر خطا در نظر گرفته می‌شود که تمام داده‌ها را در بر گیرد، ولی در ایده بیان شده که به دیدگاه رگرسیون فاصله‌ای نیز نزدیک‌تر است، چون تفاوت ابرصفحه‌های حاشیه‌ای مبتنی بر پهنای خطا است و همچنین در حالت فازی به دنبال کاهش پهنای هستیم، بنابراین فاصله بین دو ابرصفحه حاشیه‌ای باید برخلاف روش بردار پشتیبان کمینه گردد. برای انجام برآورد مولفه‌های مجهول مدل رگرسیون چند متغیره فازی پیشنهادی توأماً از دو روش رگرسیون بردار پشتیبان و رگرسیون فاصله‌ای استفاده شد.

پیکربندی مقاله به صورت زیر است: در بخش دوم برخی از مفاهیم اساسی از مجموعه فازی و مدل‌های رگرسیون بردار پشتیبان ارائه شده است. در بخش سوم، روشی از مدل تلفیقی رگرسیون فاصله‌ای و رگرسیون بردار پشتیبان فازی همراه با خطای فازی، ضرایب فازی و مشاهدات پاسخ فازی پیشنهاد شده است. در بخش چهارم، برخی مثال‌های عددی برای ارزیابی اثربخشی و عملکرد روش پیشنهادی نسبت به روش‌های دیگر از نظر برخی معیارهای نیکویی برازش گزارش شده است. نهایتاً در بخش پنجم، برخی نتایج روش پیشنهادی به همراه آینده کار تحقیقاتی تشریح شده است.

۲. مفاهیم اولیه

در این بخش، برخی از مفاهیم اولیه مورد استفاده در مقاله معرفی می‌گردند (طاهری و ماشین چی (۱۳۹۲)). در ادامه به معرفی مدل‌های رگرسیون بردار پشتیبان در حالت دقیق پرداخته شده است.

۱.۲. مجموعه فازی.

تعریف ۱.۲. مجموعه فازی \tilde{A} از مجموعه مرجع X به صورت $\{(x, \mu_{\tilde{A}}(x)) : x \in X\}$ تعریف می‌شود که در آن $\mu_{\tilde{A}} : x \rightarrow [0, 1]$ تابع عضویت مجموعه فازی \tilde{A} است. بطور خلاصه تابع عضویت مجموعه فازی \tilde{A} به صورت $\tilde{A}(x) = \mu_{\tilde{A}}(x)$ نشان داده می‌شود.

تعریف ۲.۲. برای هر $\alpha \in (0, 1]$ ، آلفا-برش مجموعه فازی \tilde{A} به صورت زیر تعریف می‌گردد:

$$\tilde{A}[\alpha] = \{x \in X | \tilde{A}(x) \geq \alpha\}.$$

تعریف ۳.۲. مجموعه فازی \tilde{N} از X را یک عدد فازی گوئیم، اگر

- دقیقاً یک x ، وجود داشته باشد، بطوری که $\tilde{N}(x) = 1$.
- α -برش های \tilde{N} مجموعه‌های بسته و کراندار باشد.

تعریف ۴.۲. عدد فازی \tilde{N} را یک عدد فازی LR نامند، اگر تابع عضویت آن به شکل زیر باشد:

$$\tilde{N}(x) = \begin{cases} L\left(\frac{m-x}{\alpha}\right) & x \leq m, \\ R\left(\frac{x-m}{\beta}\right) & x > m, \end{cases}$$

که در آن L و R توابع غیر صعودی از مجموعه اعداد حقیقی مثبت به $[0, 1]$ و $L(0) = 1$ می‌باشد. در این صورت عدد فازی LR را به صورت $\tilde{N} = (m; \alpha, \beta)_{LR}$ نمایش می‌دهند. عدد m میانه و اعداد α و β که مثبت نیز هستند را پهنای چپ و پهنای راست می‌نامند.

ملاحظه ۵.۲. اگر تابع‌های L و R به صورت $L(x) = R(x) = \max\{0, 1 - x\}$ باشند، آن را عدد فازی مثلثی نامند و به صورت $\tilde{N} = (m, \alpha, \beta)_T$ نمایش می‌دهند.

ملاحظه ۶.۲. عدد فازی \tilde{N} را یک عدد فازی متقارن گوئید، هرگاه $L(x) = R(x)$ و $\alpha = \beta$ و به صورت $\tilde{N} = (m; \alpha)_L$ نمایش داده می‌شود.

بر اساس اصل گسترش، اعمال جبری ضرب و جمع بین دو عدد فازی LR به صورت زیر به دست می‌آید.

لم ۷.۲. اگر $\tilde{A} = (m; \alpha, \beta)_{LR}$ ، $\tilde{B} = (n; \delta, \lambda)_{LR}$ و $\lambda \in R$ آنگاه:

$$\lambda \otimes \tilde{A} = \begin{cases} (\lambda m; \lambda \alpha, \lambda \beta)_{LR}, & \lambda \geq 0, \\ (\lambda m; -\lambda \beta, -\lambda \alpha)_{RL}, & \lambda < 0, \end{cases}$$

و

$$\tilde{A} \oplus \tilde{B} = (m + n; \alpha + \delta, \beta + \lambda)_{LR}.$$

۲.۲. رگرسیون بردار پشتیبان. فرض کنید داده‌های $(x_i, y_i); i = 1, 2, \dots, n$ در دسترس باشند. می‌خواهیم مدل رگرسیونی خطی زیر را به داده‌ها برازش دهیم:

$$y = \beta \cdot x + b + \varepsilon,$$

که در آن b عرض از مبدا یا اربیبی، $\beta = (\beta_1, \dots, \beta_p)$ بردار ضرایب مدل و $x = (x_1, \dots, x_p)$ بردار متغیرهای مستقل را نشان می‌دهند. در روش رگرسیون بردار پشتیبان فاصله بین دو ابرصفحه حاشیه‌ای $\frac{2e}{\|\beta\|}$ را به گونه‌ای بیشینه می‌نمایند که خطاهای بین مقادیر متغیرهای پاسخ مشاهده شده و برآورد شده به حداقل مقدار خود برسند و خطای e حداکثر خطایی است که تمام داده‌ها داخل خطوط قرار گیرند. در این صورت، مسئله بهینه‌سازی زیر برای دستیابی به برآوردهای بهینه پارامترهای مدل را در نظر می‌گیریم:

$$\min_{\beta, b} \frac{1}{2} \|\beta\|^2,$$

s.t.

$$\begin{cases} y_i - \beta \cdot x_i - b \leq e, \\ \beta \cdot x_i + b - y_i \leq e, \quad i = 1, 2, \dots, n, \end{cases}$$

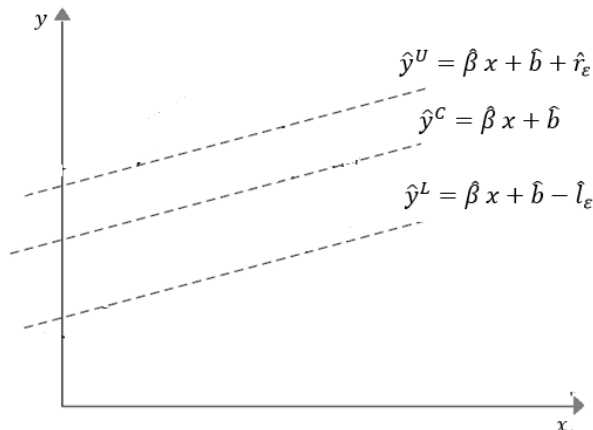
که مقدار e مقدار ثابتی است و توسط پژوهشگر تعیین می‌گردد. خط میانی در بین دو خط حاشیه مدل رگرسیون می‌باشد. نقاطی که روی دو خط حاشیه‌ای قرار گرفته است، بردارهای پشتیبان و یا نقاط پشتیبان گفته می‌شود که توسط آنها پارامترهای مدل تعیین می‌گردد.

۳. رگرسیون فازی بر اساس ماشین بردار پشتیبان با خطوط غیرموازی و با خطای فازی

مدل رگرسیون خطی چندمتغیره با خطای فازی و ضرایب فازی را به صورت زیر در نظر بگیرد:

$$\tilde{y}_i = (\tilde{\beta}_1 \otimes x_{i1}) \oplus (\tilde{\beta}_2 \otimes x_{i2}) \oplus \dots \oplus (\tilde{\beta}_p \otimes x_{ip}) \oplus b \oplus \tilde{\varepsilon}, \quad i = 1, \dots, n,$$

که در آن $\tilde{y}_i = (y_i; l_{y_i}, r_{y_i})_T$ متغیر پاسخ مشاهدات $x_i = (x_{i1}, \dots, x_{ip})$ ، $x_{ij} \geq 0, 1, \dots, n$ ، متغیر مستقل و $\tilde{\beta}_i = (\beta_i; l_{\beta_i}, r_{\beta_i})_T$ ضرایب مدل می‌باشند. مولفه خطا را نیز یک عدد فازی مثلثی به صورت $\tilde{\varepsilon} = (0; l_{\varepsilon}, r_{\varepsilon})_T$ در نظر می‌گیریم. خطوط حاشیه بالا و پایین نسبت به خط مدل رگرسیونی که می‌توان آنها را هم به نوعی ابرصفحه‌های



شکل ۱: نمایش خطوط مرکز و حاشیه با ضرایب مدل دقیق در حالت دو بعدی

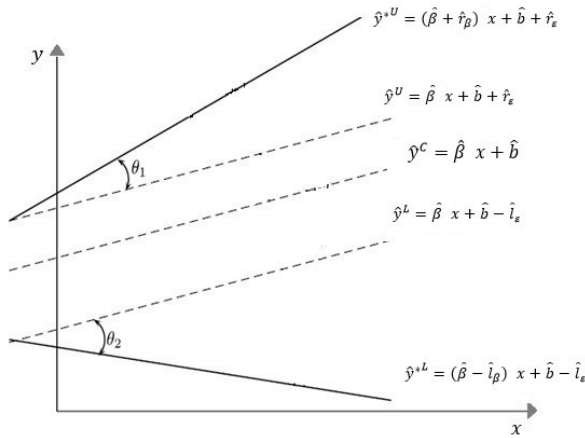
حاشیه‌ای در نظر گرفت، به صورت زیر نمایش داده می‌شود (شکل‌های ۱ و ۲ را در حالت دوبعدی ببینید):

$$\begin{cases} \hat{y}^{*C} = \hat{\beta}'\mathbf{x} + \hat{b}, \\ \hat{y}^{*U} = (\hat{\beta} + \hat{\mathbf{r}}_{\beta})'\mathbf{x} + (\hat{b} + \hat{r}_{\epsilon}), \\ \hat{y}^{*L} = (\hat{\beta} - \hat{\mathbf{l}}_{\beta})'\mathbf{x} + (\hat{b} - \hat{l}_{\epsilon}), \end{cases}$$

که در آن $\hat{\mathbf{r}}_{\beta} = (\hat{r}_{\beta_1}, \dots, \hat{r}_{\beta_p})$ و $\hat{\mathbf{l}}_{\beta} = (\hat{l}_{\beta_1}, \dots, \hat{l}_{\beta_p})$ ، $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_p)$ که در آن ضرایب مدل و پهناهای آن می‌باشند. اگر حالتی را در نظر بگیریم که پهناهای ضرایب مدل صفر باشد، در واقع ضرایب دقیق فرض شوند (شکل ۱)، در این صورت مدل رگرسیونی و حاشیه‌ها به صورت زیر نمایش داده می‌شود:

$$\begin{cases} \hat{y}^C = \hat{\beta}'\mathbf{x} + \hat{b}, \\ \hat{y}^U = \hat{\beta}'\mathbf{x} + (\hat{b} + \hat{r}_{\epsilon}), \\ \hat{y}^L = \hat{\beta}'\mathbf{x} + (\hat{b} - \hat{l}_{\epsilon}). \end{cases}$$

بدون آنکه به کلیت مسئله خللی وارد شود با قرار دادن $p = 1$ می‌توان شکل ۲ را در نظر گرفت. همانطور که در شکل مشاهده می‌شود \hat{y}^L موازی \hat{y}^U و همچنین خطوط \hat{y}^{*L} با خط \hat{y}^L زاویه θ_2 و \hat{y}^{*U} با خط \hat{y}^U زاویه θ_1 را ایجاد کرده است.



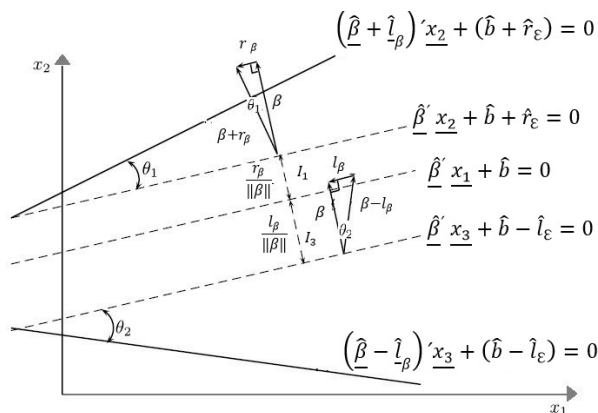
شکل ۲: نمایش خط مرکز و خطوط حاشیه‌ای در مدل با وجود پهنا در ضرایب مدل به صورت خطهای ممتد و مدل بدون پهنا در ضرایب به صورت خطهای بریده در حالت دو بعدی

اگر تمام مقادیر را به یک سمت انتقال دهیم، می‌توانیم معادلات بالا را به صورت زیر با نمادگذاری جدید بازنویسی نماییم (شکل ۳ را ملاحظه کنید):

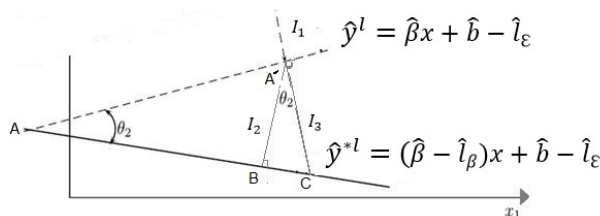
$$\begin{cases} V^C = \underline{\hat{\beta}}' \underline{\mathbf{x}}_1 + \hat{b} = 0, \\ V^U = (\underline{\hat{\beta}} + \underline{\hat{r}}_{\beta})' \underline{\mathbf{x}}_2 + (\hat{b} + \hat{r}_{\epsilon}) = 0, \\ V^L = (\underline{\hat{\beta}} - \underline{\hat{l}}_{\beta})' \underline{\mathbf{x}}_3 + (\hat{b} - \hat{l}_{\epsilon}) = 0, \end{cases}$$

که در آن $\underline{\hat{\beta}} = (1, \hat{\beta})$ ، $\underline{\mathbf{x}}_1 = (\hat{y}^{*C}, \mathbf{x})$ ، $\underline{\mathbf{x}}_2 = (\hat{y}^{*U}, \mathbf{x})$ ، $\underline{\mathbf{x}}_3 = (\hat{y}^{*L}, \mathbf{x})$ ، $\underline{\hat{\beta}} = (1, \hat{\beta})$ و $(1, \hat{r}_{\beta})$ در حالتی که ضرایب مدل دارای پهنای صفر باشند، در این صورت خط مرکز را با F^C ، خط برابر با کران پایین را با F^L و خط برابر با کران بالا را با F^U نمایش می‌دهیم. اکنون اگر بخواهیم فاصله یک موقعیت که با یک نقطه در فضای دو بعدی و به صورت (y_i, x_i) نمایش داده می‌شود را تا خطوط مرکز و حاشیه محاسبه کنیم می‌توانیم از فرمول فاصله یک نقطه تا خط استفاده کنیم. بنابراین فرض کنید نقطه مورد نظر $\underline{\mathbf{x}}_i = (y_i, x_i)$ بین دو خط V^C و V^L باشد، در این صورت فاصله $\underline{\mathbf{x}}_i$ تا خط مرکز برابر است با:

$$(1.3) \quad I_1 = \frac{|\underline{\hat{\beta}}' \underline{\mathbf{x}}_i + b|}{\|\underline{\hat{\beta}}\|}$$



شکل ۳: نمایش شکل قبل در حالت بازنویسی شده



شکل ۴: نمایش مثلثهای متشابه در شکل به صورت بزرگنمایی

از آنجایی که زاویه بین دو خط F^L و V^L برابر θ_2 در نظر گرفته شده است و همان گونه که در شکل ۴ دیده می شود دو مثلث $AA'C$ و $BA'C$ متشابه هستند، بنابراین زاویه \hat{A} با \hat{A}' برابر هستند که در این صورت بنا بر ضرب داخلی بردارها خواهیم داشت:

$$(۲.۳) \quad \cos \theta_2 = \frac{\underline{\hat{\beta}}'(\underline{\hat{\beta}} - \underline{\hat{l}}_{\beta})}{\|\underline{\hat{\beta}}\| \|\underline{\hat{\beta}} - \underline{\hat{l}}_{\beta}\|} = \frac{\underline{\hat{\beta}}' \underline{\hat{\beta}} - \underline{\hat{\beta}}' \underline{\hat{l}}_{\beta}}{\|\underline{\hat{\beta}}\| \|\underline{\hat{\beta}} - \underline{\hat{l}}_{\beta}\|} = \frac{\underline{\hat{\beta}}' \underline{\hat{\beta}}}{\|\underline{\hat{\beta}}\| \|\underline{\hat{\beta}} - \underline{\hat{l}}_{\beta}\|},$$

که تساوی آخر در رابطه (۲.۳) به علت متعامد بودن دو بردار $\underline{\hat{l}}_{\beta}$ و $\underline{\hat{\beta}}$ حاصل شده است (بردار پهنا بر شیب عمود است). بر اساس روابط مثلثاتی در دو مثلث قائم الزاویه داریم:

$$(۳.۳) \quad I_3 \cos \theta_2 = I_2.$$

از طرفی فاصله $\underline{\mathbf{x}}_i$ تا خط V^L برابر است با:

$$(۴.۳) \quad I_۲ = \frac{|(\hat{\underline{\beta}} - \hat{\underline{1}}_{\beta})' \underline{\mathbf{x}}_i + (\hat{b} - \hat{l}_{\varepsilon})|}{\|\hat{\underline{\beta}} - \hat{\underline{1}}_{\beta}\|}.$$

با جایگذاری روابط (۲.۳) و (۴.۳) در رابطه (۳.۳)، خواهیم داشت:

$$I_۳ = \frac{|(\hat{\underline{\beta}} - \hat{\underline{1}}_{\beta})' \underline{\mathbf{x}}_i + (\hat{b} - \hat{l}_{\varepsilon})|}{\|\hat{\underline{\beta}}\|}.$$

چون $\underline{\mathbf{x}}_i$ محدود بین دو خط V^L و خط مرکز می‌باشد، بنابراین برای آن نقطه داریم:

$$\begin{cases} (\hat{\underline{\beta}} - \hat{\underline{1}}_{\beta})' \underline{\mathbf{x}}_i + (\hat{b} - \hat{l}_{\varepsilon}) > ۰, \\ \hat{\underline{\beta}}' \underline{\mathbf{x}}_i + \hat{b} < ۰, \end{cases}$$

در این صورت جمع بین دو فاصله $I_۱$ و $I_۳$ برابر خواهد بود با:

$$(۵.۳) \quad I_۱ + I_۳ = \frac{|\hat{\underline{1}}_{\beta}' \underline{\mathbf{x}}_i + \hat{l}_{\varepsilon}|}{\|\hat{\underline{\beta}}\|}.$$

اکنون به همین شیوه، فاصله نقطه فرضی تا خط کران بالا مورد محاسبه قرار می‌گیرد و این مقدار به فرمول (۵.۳) اضافه می‌گردد. با همین روش می‌توان فاصله یک نقطه در میان خط وسط و خط حاشیه بالایی که قرار می‌گیرد را محاسبه نمود و فاصله نقطه تا خط مرکز و خط کران بالا را بدست آورد. در این صورت فاصله بدست آمده با نسبتی از فاصله نقطه فرضی تا کران پایین نیز محاسبه می‌گردد و با هم جمع می‌شود. در حالت متقارن این فاصله برای هر دو خط حاشیه یکسان خواهد بود و یک بار فاصله محاسبه می‌شود و در نهایت دو برابر خواهد شد، اگر چه که می‌توان برای بدست آوردن مدل رگرسیونی بهتر از پیش‌بینی کننده‌هایی با پهنای متغیر استفاده نمود. بنابراین در ادامه فاصله کران بالا تا کران پایین برای متغیر $\underline{\mathbf{x}}_i$ برابر است با:

$$\frac{|\hat{\underline{r}}_{\beta}' \underline{\mathbf{x}}_i + \hat{r}_{\varepsilon}|}{\|\hat{\underline{\beta}}\|} + \frac{|\hat{\underline{1}}_{\beta}' \underline{\mathbf{x}}_i + \hat{l}_{\varepsilon}|}{\|\hat{\underline{\beta}}\|}.$$

مسئله بهینه‌سازی برای این مدل رگرسیونی به گونه‌ای رقم می‌خورد که جمع همه‌ی فواصل به ازای هر مرکز مشاهده که می‌تواند نقش نقطه فرضی را داشته باشد، بایستی مقداری کمینه شود. محدودیت‌هایی همچون متعامد بودن دو بردار $\hat{\underline{1}}_{\beta}$ و $\hat{\underline{\beta}}$ ، $\hat{\underline{\beta}}$ و $\hat{\underline{r}}_{\beta}$ در محاسبه تابع هدف موردتوجه قرار گرفته است و به همین دلیل تنها محدودیتی همچون $[\hat{y}_i^{*L}, \hat{y}_i^{*U}] \subseteq [y_i^L, y_i^U]$ را بنابر دیدگاه فاصله‌ای در مسئله بهینه‌سازی در نظر می‌گیریم. در واقع بنابر دیدگاه فاصله‌ای

می‌بایست تمام پاسخ‌های مشاهدات در فاصله بین برآودهای پیش‌بینی قرار گیرد. بنابراین مسئله بهینه‌سازی به صورت زیر ارائه می‌گردد:

$$\min \sum_{i=1}^n \left(\frac{|\hat{\mathbf{r}}'_\beta \mathbf{x}_i + \hat{r}_\varepsilon|}{\|\hat{\underline{\beta}}\|} + \frac{|\hat{\mathbf{l}}'_\beta \mathbf{x}_i + \hat{l}_\varepsilon|}{\|\hat{\underline{\beta}}\|} \right),$$

s.t.

$$\begin{cases} (\hat{\beta} - \hat{\mathbf{l}}_\beta)' \mathbf{x}_i + (\hat{b} - \hat{l}_\varepsilon) \leq y_i - l_{y_i}, \\ (\hat{\beta} + \hat{\mathbf{r}}_\beta)' \mathbf{x}_i + (\hat{b} + \hat{r}_\varepsilon) \geq y_i + r_{y_i}, \\ \hat{r}_\varepsilon, \hat{l}_\varepsilon \geq 0, \quad i = 1, \dots, n. \end{cases}$$

با نرمال کردن پارامترهای مدل توسط تقسیم آنها بر $\|\hat{\underline{\beta}}\|$ در مسئله بهینه‌سازی و اضافه کردن متغیرهای کاهش مثبت یا کمبود با مقدار جریمه D می‌توان به صورت زیر آن را بازنویسی کرد:

$$\min \sum_{i=1}^n (|\hat{\mathbf{l}}'_\beta \mathbf{x}_i + \hat{l}_\varepsilon| + |\hat{\mathbf{r}}'_\beta \mathbf{x}_i + \hat{r}_\varepsilon|) + D \sum_{i=1}^n (\xi_i + \xi_i^*),$$

s.t.

$$\begin{cases} (\hat{\beta} - \hat{\mathbf{l}}_\beta)' \mathbf{x}_i + (\hat{b} - \hat{l}_\varepsilon) \leq y_i - l_{y_i} + \xi_i, \\ (\hat{\beta} + \hat{\mathbf{r}}_\beta)' \mathbf{x}_i + (\hat{b} + \hat{r}_\varepsilon) \geq y_i + r_{y_i} + \xi_i^*, \\ \hat{r}_\varepsilon, \hat{l}_\varepsilon \geq 0, \quad \xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, n, \end{cases}$$

که در آن $\hat{r}_\varepsilon = \frac{\hat{r}_\varepsilon}{\|\hat{\underline{\beta}}\|}$ و $\hat{\mathbf{l}}_\beta = \frac{\hat{\mathbf{l}}_\beta}{\|\hat{\underline{\beta}}\|}$, $\hat{\mathbf{r}}_\beta = \frac{\hat{\mathbf{r}}_\beta}{\|\hat{\underline{\beta}}\|}$, $\hat{\beta} = \frac{\hat{\beta}}{\|\hat{\underline{\beta}}\|}$, $\hat{b} = \frac{\hat{b}}{\|\hat{\underline{\beta}}\|}$, $\hat{l}_\varepsilon = \frac{\hat{l}_\varepsilon}{\|\hat{\underline{\beta}}\|}$ می‌گردد. از آن جایی که قدرمطلق را می‌توان به دو عامل d_i^- و d_i^+ و همچنین $d_i'^-$ و $d_i'^+$ نوشت، در این صورت خواهیم داشت:

$$\min \sum_{i=1}^n ((d_i^+ + d_i^-) + (d_i'^+ + d_i'^-)) + D \sum_{i=1}^n (\xi_i + \xi_i^*),$$

s.t.

$$\begin{cases} \hat{l}'_{\beta} \mathbf{x}_i + \hat{l}_{\varepsilon} = d_i^+ - d_i^-, \\ \hat{r}'_{\beta} \mathbf{x}_i + \hat{r}_{\varepsilon} = d_i'^+ - d_i'^-, \\ (\hat{\beta} - \hat{l}_{\beta})' \mathbf{x}_i + (\hat{b} - \hat{l}_{\varepsilon}) \leq y_i - l_{y_i} + \xi_i, \\ (\hat{\beta} + \hat{r}_{\beta})' \mathbf{x}_i + (\hat{b} + \hat{r}_{\varepsilon}) \geq y_i + r_{y_i} + \xi_i^*, \\ \hat{r}_{\varepsilon}, \hat{l}_{\varepsilon} \geq 0, \quad \xi_i, \xi_i^* \geq 0, \quad d_i^+, d_i^-, d_i'^+, d_i'^- \geq 0, \quad i = 1, \dots, n. \end{cases}$$

که در آن داریم:

$$\begin{cases} d_i^+ = \max\{0, \hat{l}'_{\beta} \mathbf{x}_i - \hat{l}_{\varepsilon}\}, & d_i'^+ = \max\{0, \hat{r}'_{\beta} \mathbf{x}_i + \hat{r}_{\varepsilon}\}, \\ d_i^- = \max\{0, -(\hat{l}'_{\beta} \mathbf{x}_i - \hat{l}_{\varepsilon})\}, & d_i'^- = \max\{0, -\hat{r}'_{\beta} \mathbf{x}_i - \hat{r}_{\varepsilon}\}, \end{cases}$$

ملاحظه ۱.۳. برای حل مسئله بهینه‌سازی از نرم افزار ممتتیکا و دستور NMinimize استفاده می‌کنیم.

ملاحظه ۲.۳. قابل توجه است که الزامی برای اینکه در قبال مشاهدات متقارن، مقادیر متغیر پاسخ برآوردشده متقارن باشد، وجود ندارد. در صورتی که بخواهیم پارامترهای برآورد شده فازی مثلثی متقارن محاسبه شوند، باید محدودیت برابری پهنای پارامتر را در مسئله در نظر بگیریم.

۱.۳. معیار نیکویی برازش مدل. برای ارزیابی مدل رگرسیون بردار پشتیبان فازی بیان شده، معیارهای نیکویی برازش زیر معرفی می‌گردد.

تعریف ۳.۳. مشاهدات و مقادیر پیش‌بینی متغیر پاسخ اعداد فازی مثلثی را به صورت $\tilde{y}_i = (y_i; l_{y_i}, r_{y_i})_T$ و $\hat{y}_i = (\hat{y}_i; l_{\hat{y}_i}, r_{\hat{y}_i})_T$ در نظر بگیرید. آنگاه معیار نیکویی برازش مدل رگرسیونی به صورت زیر تعریف می‌گردد:

$$I = \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{1 + d(\tilde{y}_i, \hat{y}_i)} \right)$$

که در آن داریم:

$$d(\tilde{y}_i, \hat{y}_i) = \frac{1}{3} (|\hat{y}_i - y_i| + |(\hat{y}_i - l_{\hat{y}_i}) - (y_i - l_{y_i})| + |(\hat{y}_i + r_{\hat{y}_i}) - (y_i + r_{y_i})|).$$

تعریف ۴.۳. مشاهدات و مقادیر پیش‌بینی متغیر پاسخ را به صورت $\tilde{y}_i = (y_i; l_{y_i}, r_{y_i})_T$ و $\hat{y}_i = (\hat{y}_i; l_{\hat{y}_i}, r_{\hat{y}_i})_T$ در نظر بگیرید. معیار نیکویی برازش بر اساس میانگین اندازه مشابهت به صورت زیر معرفی می‌گردد:

$$MSM = \frac{1}{n} \sum_{i=1}^n S_{PK}(\hat{y}_i, \tilde{y}_i),$$

که در آن $S_{PK}(\hat{y}_i, \tilde{y}_i) = \frac{card(\hat{y}_i \cap \tilde{y}_i)}{card(\hat{y}_i \cup \tilde{y}_i)}$ که در آن $Card(\tilde{N})$ به صورت $\int_{\mathbb{R}} \tilde{N}(x) dx$ تعریف می‌شود (پاپیس و کاراکاپلیدیز (۱۹۹۳)). همچنین داریم:

$$(\hat{y}_i \cap \tilde{y}_i)(x) = \min(\hat{y}_i(x), \tilde{y}_i(x)),$$

$$(\hat{y}_i \cup \tilde{y}_i)(x) = \max(\hat{y}_i(x), \tilde{y}_i(x))$$

تعریف ۵.۳. مشاهدات و مقادیر پیش‌بینی متغیر پاسخ را به صورت $\tilde{y}_i = (y_i; l_{y_i}, r_{y_i})_T$ و $\hat{y}_i = (\hat{y}_i; l_{\hat{y}_i}, r_{\hat{y}_i})_T$ در نظر بگیرید. معیار نیکویی برازش بر اساس خطای اندازه‌گیری به صورت زیر تعریف می‌شود:

$$G_1 = \frac{1}{n} \sum_{i=1}^n S_{CD}(\hat{y}_i, \tilde{y}_i),$$

که $S_{CD}(\hat{y}_i, \tilde{y}_i) = \frac{1}{1 + E_1(\hat{y}_i, \tilde{y}_i)}$ و $E_1(\hat{y}_i, \tilde{y}_i) = \frac{\int_{S_{\hat{y}_i} \cup S_{\tilde{y}_i}} |\hat{y}_i(x) - \tilde{y}_i(x)| dx}{\int_{S_{\hat{y}_i} \cap S_{\tilde{y}_i}} \tilde{y}_i(x) dx}$ در آن $S_{\hat{y}_i}$ و $S_{\tilde{y}_i}$ دامنه تغییرات در مجموعه فازی \hat{y}_i و \tilde{y}_i هستند (چن و دانگ (۲۰۰۸)).

۴. مثال‌های عددی

مثال ۱.۴. مجموعه داده با مقادیر متغیر مستقل دقیق و مقادیر پاسخ فازی را در جدول ۱ در نظر بگیرید (عارفی (۲۰۲۰)). این مدل پیشنهادی با ضرایب و خطای فازی متقارن فرض شده است. قابل توجه است که مدل رگرسیونی فازی اکثر مشاهدات متغیر پاسخ به همراه ابهام‌هایش را در بر گرفته است و در مورد مشاهداتی که خارج از محدوده هستند می‌توان با مقدار کمی جریمه مشاهده را به مجموعه داده فراخوانی نمود. مدل پیشنهادی با مدل‌های هانگ و وانگ (۲۰۰۳)، یائو و یو (۲۰۰۶) و یانکویی (۲۰۱۲) مورد مقایسه قرار گرفته است. نتایج ارزیابی

مدل و پارامترها در جدول ۲ خلاصه شده است. نمودار مربوط به مدل‌های رگرسیونی در شکل ۵ نمایش داده شده است.

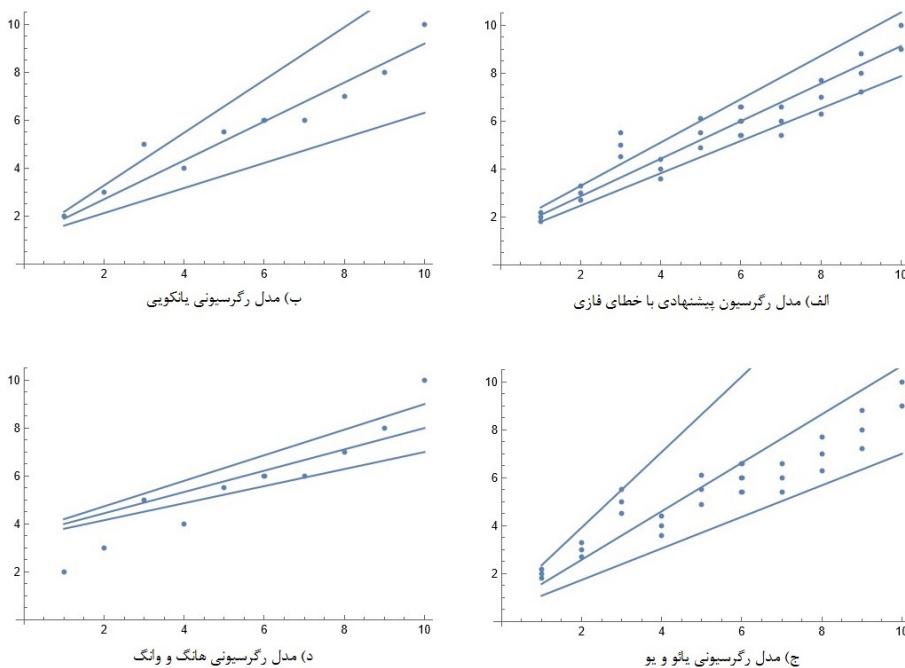
جدول ۱: مجموعه داده در مثال ۱.۴.

شماره	x_i	$(y_i; l_{y_i})_T$	$(\hat{y}_i; \hat{l}_{y_i}, \hat{r}_{y_i})_T$
۱	۱	$(۲; ۰/۲)_T$	$(۲/۰۸; ۰/۲۸, ۰/۳۱)_T$
۲	۲	$(۳; ۰/۳)_T$	$(۲/۸; ۰/۳۸, ۰/۴۳)_T$
۳	۳	$(۵; ۰/۵)_T$	$(۳/۶; ۰/۴۹, ۰/۵۵)_T$
۴	۴	$(۴; ۰/۴)_T$	$(۴/۴۳; ۰/۶۰, ۰/۶۷)_T$
۵	۵	$(۵/۵; ۰/۶)_T$	$(۵/۲۱; ۰/۷۱, ۰/۷۹)_T$
۶	۶	$(۶; ۰/۶)_T$	$(۵/۹۹; ۰/۸۲, ۰/۹۱)_T$
۷	۷	$(۶; ۰/۶)_T$	$(۶/۷۸; ۰/۹۳, ۱/۰۴)_T$
۸	۸	$(۷; ۰/۷)_T$	$(۷/۵۶; ۱/۰۳, ۱/۱۶)_T$
۹	۹	$(۸; ۰/۸)_T$	$(۸/۳۴; ۱/۱۴, ۱/۲۸)_T$
۱۰	۱۰	$(۱۰; ۱)_T$	$(۹/۱۳; ۱/۲۵, ۱/۴۰)_T$

جدول ۲: مقایسه مدل پیشنهادی با برخی مدل‌های دیگر

روش	تخمین پارامترهای مدل	MSM	G_1	I
روش پیشنهادی با خطای فازی	$\hat{\beta} = (۰/۷۸; ۰/۱۰, ۰/۱۲)_T, \hat{b} = ۱/۲۹$	۰/۳۴	۰/۴۷	۰/۷۰
روش یائو و یو	$\hat{\beta} = (۱/۰۱; ۰/۵۶, ۰/۳۵)_T, \hat{b} = ۰/۵۵$	۰/۱۳	۰/۱۶	۰/۳۴
روش یانگویی	$\hat{\beta} = (۰/۸۱; ۰/۲۸)_T, \hat{b} = ۱/۰۷$	۰/۲۹	۰/۳۴	۰/۵۶
روش هانگ و وانگ	$\hat{\beta}_1 = (-۰/۰۷; ۰/۴۰)_T, \hat{\beta}_2 = (۰; ۰/۵۲)_T,$ $\hat{\beta}_3 = (۰; ۰/۱)_T, \hat{\beta}_4 = (۰; ۰/۳۹)_T,$ $\hat{\beta}_5 = (۰; ۰/۰۱)_T, \hat{\beta}_6 = (۰; ۰/۰۰۶)_T,$ $\hat{\beta}_7 = (۰; ۰/۰۰۴)_T, \hat{\beta}_8 = (۰/۱۱; ۰/۰۱)_T,$ $\hat{\beta}_9 = (۰/۲۱; ۰/۰۲)_T, \hat{\beta}_{10} = (۰/۳۹; ۰/۰۳)_T,$ $\hat{b} = (۴; ۰/۶)_T$	۰/۲۸	۰/۴۸	۰/۶۱

همانطور که مشاهده می‌شود مدل پیشنهادی با خطای فازی توانسته است توسط هر سه اندازه مشابهت مقدار بالاتری نسبت به مدل‌های دیگر ارائه دهد. در هر چهار نمودار تمام داده در بین دو خط حاشیه قرار گرفته‌اند و آنهایی که در خارج هستند توسط مقادیر کمبود به مجموعه داده‌ها با مقدار جریمه اندکی فراخوانی می‌شوند. همانطور که دیده می‌شود در مدل



شکل ۵: مدل‌های رگرسیونی فازی در مثال ۱.۴

پیشنهادی خطوط حاشیه به خط مرکز متمایل‌تر از روش‌های دیگر می‌باشد و همچنین اندازه مشابهت نشان می‌دهد کمینه‌سازی بهتر انجام گرفته است.

مثال ۲.۴. فرض کنید بخواهیم ۱۰۰ داده با پاسخ فازی به صورت $\tilde{y}_i = (y_i, l_{y_i})_T$ شبیه‌سازی نماییم. شرایط زیر را در نظر می‌گیریم:

$$y_i = 3x_{1i} - 2x_{2i} + 5 + e_i,$$

که در آن

$$x_{1i} \sim U(-10, 10), \quad x_{2i} \sim N(3, 2), \quad e_i \sim N(0, 0.6),$$

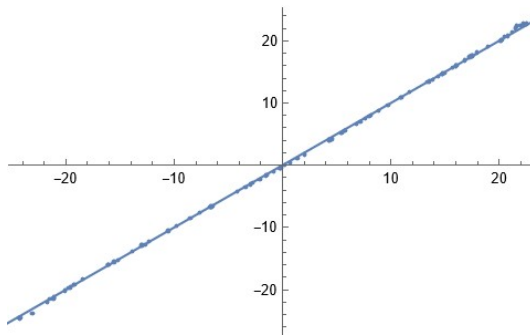
$$l_{y_i} \sim U(0, 0.8), \quad i = 1, 2, \dots, 100.$$

نمودار مراکز نقاط پیش‌بینی و متغیر پاسخ در مقابل خط نیمساز ربع اول و سوم به روش پیشنهادی در شکل ۶ به نمایش درآمده است. همانطور که در شکل ۶ دیده می‌شود اکثر نقاط

روی خط نیمساز قرار گرفته است و می‌توان گفت که تقریباً گویای تطابق زیاد مسئله می‌باشد. مقدار \hat{s} در جدول ۳ مقدار پهنای خط را در مسئله نمایش می‌دهد. همینطور اندازه مشابهت ارائه شده نیز نشان می‌دهد که روش پیشنهادی توانسته است به خوبی مدل رگرسیونی را برازش دهد.

جدول ۳: برآورد پارامترها و اندازه مشابهت برای داده‌های شبیه‌سازی مثال ۲.۴

$\hat{\beta}_1$	\hat{r}_{β_1}	\hat{l}_{β_1}	$\hat{\beta}_2$	\hat{r}_{β_2}	\hat{l}_{β_2}	\hat{b}	\hat{s}	I
۳/۰۹	۰/۱۳	۲/۳۴	-۱/۴	۰/۴	۱/۵	۳/۴	۱/۶	۰/۸۷



شکل ۶: نمودار مراکز مشاهدات متغیر پاسخ در مقابل مراکز مقادیر برآورد شده متغیر پاسخ نسبت به نیمساز ربع اول و سوم در مثال ۲.۴

مثال ۳.۴. مجموعه داده جدول ۴ ارائه شده توسط یائو و یو (۲۰۰۶) را در نظر بگیرید. ابتدا مدل رگرسیونی پیشنهادی در این مقاله و همچنین مدل ارائه شده توسط یائو و یو (۲۰۰۶) به داده‌ها بدون وجود داده پرت برازش شده است (شکل ۷ قسمتهای الف و ب). هر دو مدل برازش‌های مناسبی را دارند، ولی مدل رگرسیون پیشنهادی بر اساس معیارهای نیکویی برازش در جدول ۵ عملکرد بهتری را نشان می‌دهد. برای بررسی عملکرد این دو مدل رگرسیونی در مواجهه با داده پرت، مشاهده پاسخ شماره ۷ به صورت $T(30, 1)$ تغییر کرد. بر اساس معیارهای نیکویی برازش در جدول ۵ و همچنین شکل ۷ (قسمتهای ج و د)، هر دو مدل استواری خود را در مواجهه با داده پرت حفظ کرده‌اند و مدل پیشنهادی در این مقاله عملکرد بهتری را نیز داشته است.

جدول ۴: مجموعه داده در مثال ۳.۴

شماره	x_i	\tilde{y}_i
۱	۱	$(-۱/۶; ۰/۵)_T$
۲	۳	$(-۱/۸; ۰/۵)_T$
۳	۴	$(-۱; ۰/۵)_T$
۴	۵/۶	$(۱/۲; ۰/۵)_T$
۵	۷/۸	$(۲/۲; ۱)_T$
۶	۱۰/۲	$(۶/۸; ۱)_T$
۷	۱۱	$(۱۰; ۱)_T$
۸	۱۱/۵	$(۱۰; ۱)_T$
۹	۱۲/۷	$(۱۰; ۱)_T$

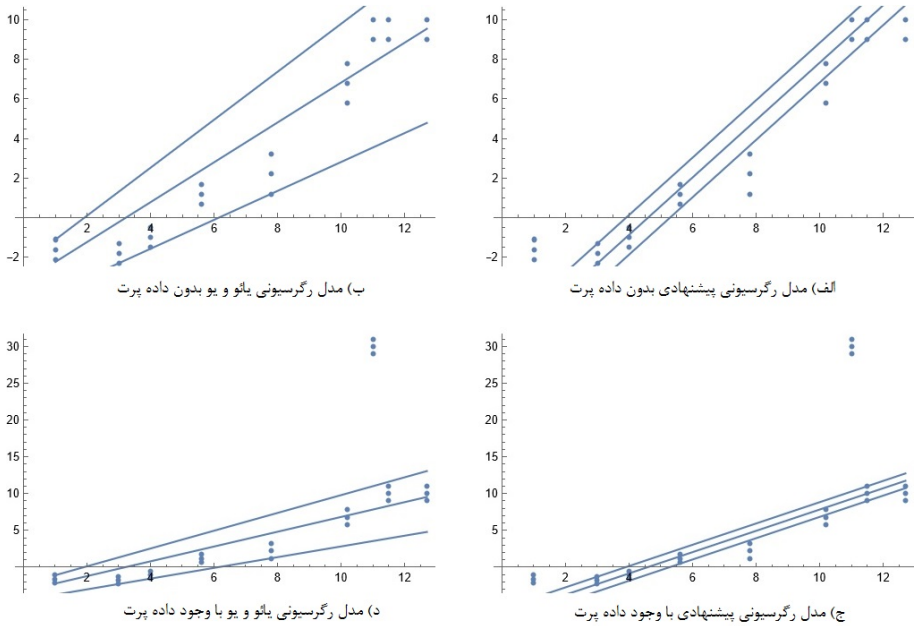
جدول ۵: مقایسه مدل پیشنهادی با مدل یائو و یو

روش	تخمین پارامترهای مدل	MSM	G_1	I
روش پیشنهادی بدون داده پرت	$\hat{\beta} = (۱/۴۵; ۷/۵۱ \times ۱۰^{-۱۱})_T, \hat{b} = -۶/۶$	۰/۲۸	۰/۴۵	۰/۵۶
روش یائو و یو بدون داده پرت	$\hat{\beta} = (۱/۰۰۵; ۰/۱۰, ۰/۱۳)_T, \hat{b} = -۳/۲$	۰/۲۵	۰/۳۳	۰/۲۷
روش پیشنهادی با وجود داده پرت	$\hat{\beta} = (۱/۴۵; ۱/۳۵ \times ۱۰^{-۹})_T, \hat{b} = -۶/۶$	۰/۲۵	۰/۴۳	۰/۴۹
روش یائو و یو با وجود داده پرت	$\hat{\beta} = (۱/۰۰۵; ۰/۱۰, ۰/۱۴)_T, \hat{b} = -۳/۲$	۰/۲۲	۰/۲۸	۰/۲۵

۵. نتیجه گیری

در این مقاله، یک رویکرد رگرسیونی مبتنی بر بردار پشتیبان با کمیت‌های نادقیق (فازی) و با وجود خطای فازی ارائه گردید. روش ارائه شده مبتنی بر خطوط حاشیه‌ای غیرموازی و با کمینه کردن قدرمطلق فاصله این خطوط بنا شده است. برخی از مزیت‌های روش پیشنهادی را می‌توان به صورت زیر بیان نمود:

- ۱- یک رویکرد جدید برای رگرسیون خطی با ضرایب فازی، مشاهدات پاسخ فازی و خطای فازی ارائه شده است.
- ۲- روش پیشنهادی یک تعمیم از رگرسیونی مبتنی بر بردارهای پشتیبان با ابرصفحه‌های حاشیه‌ای غیر موازی با کمیت‌های فازی است.
- ۳- روش پیشنهادی مبتنی بر مجموع قدرمطلق فواصل بین ابرصفحه‌های حاشیه‌ای غیرموازی



شکل ۷: مدل‌های رگرسیونی فازی در مثال ۳.۴

است و بنابراین تحت داده‌های پرت از استواری مناسبی برخوردار است. ارائه مدل‌های رگرسیونی با داده‌های توضیحی فازی و همچنین بسط و توسعه روش پیشنهادی با وجود خطاهای فازی متغیر، می‌تواند در آینده تحقیق مورد بررسی قرار گیرد.

مراجع

- [۱] طاهری، س.م. ماشین چپی، م. (۱۳۹۲) مقدمه‌ای بر احتمال و آمار فازی. انتشارات دانشگاه شهید باهنر کرمان.
- [2] Arefi, M. (2020). Quantile fuzzy regression based on fuzzy outputs and fuzzy parameters. *Soft Computing*, 24: 311-320.
- [3] Asadolahi, M., Akbari, M.G., Hesamian, G., Arefi, M. (2021). A robust support vector regression with exact predictors and fuzzy responses. *International Journal of Approximate Reasoning*, 132: 206-225.
- [4] Balasundaram, S., Meena, Y. (2018). Robust support vector regression in primal with asymmetric Huber loss. *Neural Processing Letters*, 49: 1399 -1431.
- [5] Basak, D., Pal, S., Ch, D., Patranabis, R. (2007). Support vector regression. *Neural Information Processing Letters and Reviews*, 11: 203-224.

- [6] Cheng, C., Lee, E.S. (2001). Fuzzy regression with radial basis function network. *Fuzzy Sets and Systems*, 119: 291-301.
- [7] Chen, L.H., Nien, S.H. (2020). A new approach to formulate fuzzy regression models, *Applied Soft Computing*, 86, Article No.: 105915.
- [8] Chen, S.P., Dang, J.F.,(2008). A variable spread fuzzy linear regression model with higher explanatory power and forecasting accuracy, *Information Sciences*, 178: 3973–3988.
- [9] Forghani, Y., Tabrizi, R. S., Yazdi, H. S., Akbarzadeh-T, M. R. (2011). Fuzzy support vector regression. In first International Conference on Computer and Knowledge Engineering (ICCKE) pp. 28-33.
- [10] Hao, P.Y., Chiang, J.H. (2008). Fuzzy regression analysis by support vector learning approach, *IEEE Transactions on Fuzzy Systems*, 16: 428–441.
- [11] Hesamian, G., Akbari, M.G. (2017). Nonparametric kernel estimation based on fuzzy random variables. *IEEE Transactions on Fuzzy Systems*, 25: 84-99.
- [12] Hesamian, G., Akbari, M.G. (2020). Fuzzy spline univariate regression with exact predictors and fuzzy responses, *J. Computational and Applied Mathematics*, 375: 1–13.
- [13] Hesamian, G., Akbari, M.G. (2021). A non-parametric model for fuzzy forecasting time series data. *Computational and Applied Mathematics*, 40: 10–23.
- [14] Hong, D. H., Hwang, C. (2003). Support vector fuzzy regression machines. *Fuzzy Sets and Systems*, 138: 271-281.
- [15] Pappis, C.P., Karacapilidis, N.I. (1993). A comparative assessment of measure of similarity of fuzzy values, *Fuzzy Sets and Systems*, 56: 171–174.
- [16] Tanaka, H., Uejima, S. and Asai, K. (1982). Linear regression analysis with fuzzy model. *IEEE Transactions on Systems, Man, and Cybernetics*, 12: 903-907.
- [17] Vapnic, V. (1995). *The Nature of Statistical Learning Theory*. Springer Verlag, New York.
- [18] Wang, K., Zhong, P. (2014). Robust non-convex least squares loss function for regression with outliers. *Knowledge-Based Systems*, 71: 290-302.
- [19] Wu, Q., Law, R. (2010). Fuzzy support vector regression machine with penalizing Gaussian noises on triangular fuzzy number space. *Expert Systems with Applications*, 37: 7788-7795.
- [20] Yao, C., Yu, P. (2006). Fuzzy regression based on asymmetric support vector machines. *Applied Mathematics and Computation*, 182: 175-193.
- [21] Yongqi, C. (2012). Least squares support vector fuzzy regression. *Energy Procedia*, 17: 711-716.
- [22] Zhao, Y. P., Sun, J. G. (2010). Robust truncated support vector regression. *Expert Systems with Applications*, 37: 5126-5133.