سیستم‌های فازی و کاربردها
سال اول، شماره دوم (زمستان ۱۳۹۷) صص. ۱۰۵ تا ۱۱۷
DOR: 20.1001.1.27174409.1397.1.2.5.8

رویکردهای وزنی در برازش مدل‌های رگرسیون فازی

جلال چاچی و علیرضا چاجی

دانشگاه شهید چمران اهواز، دانشکده ریاضی و علوم کامپیوتر، گروه آمار
دانشگاه صنعتی شهدای هویزه، گروه برق

چکیده

در این مقاله توابع هدف مسائل بهینه‌سازی مرتبط با رگرسیون‌های فازی با معیارهای برازش کمترین مربعات خطا و کمترین قدرمطلق انحرافات (خطاها) را با توابع هدفی بر مبنای مجموع وزنی توابعی از خطاها (انحرافات یا باقیمانده‌ها) یا خطاهای مرتب شده، جایگزین می‌کنیم. با توجه به تعریف و شیوه انتخاب وزن‌های بهینه و نوع تابع اعمال شده بر خطاها، رویکردهای استوار و غیر استوار متداول در مدل‌های رگرسیون فازی بر اساس روابط جدیدی بر مبنای اینگونه معیارهای نیکویی برازش تعمیم داده می‌شود. همچنین اشاره‌ای کوتاه به محاسبات مربوط به مسائل بهینه‌سازی چنین مدل‌هایی خواهیم نمود. اینگونه رویکردها تاثیر زیادی در کاهش اثرات مخرب مشاهدات پرت در برآورد مدل بهینه دارند و جایگزین مناسبی برای مدل‌های رگرسیون فازی با توانایی تشخیص مشاهدات پرت هستند. این روش‌ها تعمیم مدل‌های رگرسیون فازی کمترین مربعات و کمترین قدرمطلق انحرافات هستند و می‌توانند برای مدل‌سازی هر نوع ترکیبی از مشاهدات ورودی-دقیق/فازی و خروجی-دقیق/فازی بکار برده شوند.

Mathematics Subject Classification (2010): 62A86; 62J86, **Email:** alichp@gmail.com, jalal.chachi@scu.ac.ir.

عبارات و کلمات کلیدی: مشاهده پرت، رگرسیون فازی وزنی، الگوریتم‌های وزنی تکرار شونده.

۱ مقدمه

در تحلیل مدل‌های رگرسیون فازی (شامل رویکردهای امکانی، رویکردهای مبتنی بر شیوه کمترین مربعات خطا و دیگر رویکردهای تلفیقی)، به منظور برآورد دقیق پارامترها، مجموعه مشاهدات نمونه باید با هدف شناسایی نقاط پرت مورد بررسی و تحلیل قرار گیرد. زیرا در صورت مشاهده نقطه (یا نقاط) پرت در مجموعه داده‌ها استفاده از روش‌های برآوردیابی استوار در برآورد پارامترها ارجحیت دارد [۳، ۶، ۸]. از این جهت تا کنون روش‌ها و رویکردهای مختلف و متنوعی در برآوردیابی استوار مدل‌های رگرسیون فازی معرفی و بررسی شده است [۱۳، ۱۶]. این تنوع در رویکرد از یک جهت به دلیل گستردگی در تنوع تعریف مشاهدات پرت در مطالعات مدل‌سازی رگرسیون در محیط فازی است [۱۱]، و از جهتی دیگر به نوع مشاهدات (فازی/دقیق) متغیرهای ورودی-خروجی نیز بستگی دارد [۱۰]. به طور عمده این روش‌ها به دو رده زیر دسته‌بندی می‌شوند:

۱. رویکردهای مبتنی بر تشخیص و شناسایی داده‌های پرت: اینگونه روش‌ها عمدتا مبتنی بر نمایش تابعی داده‌ها (به عنوان مثال نمودار پراکنش داده‌ها) و/یا روش‌های مبتنی بر تحلیل داده‌های فازی هستند [۱، ۹].

۲. رویکردهای مبتنی بر برآوردیابی استوار پارامترهای مدل رگرسیون فازی (اغلب در حضور داده‌های پرت و بدون حذف آنها): در اینگونه روش‌ها اغلب تابع هدف، که پارامترهای بهینه مدل از طریق کمینه کردن یا بیشینه کردن آن به‌دست می‌آیند، به گونه‌ای انتخاب می‌شود که نسبت به مشاهدات پرت استوار باشد [۱، ۳، ۶، ۸].

تشخیص و بررسی مشاهدات پرت (نقاط یا مجموعه مشاهداتی که به نوعی از مرکز مشاهدات اصلی فاصله دارند یا از مدل تبعیت نمی‌کنند) در مدل‌های رگرسیون فازی و غیر فازی اهمیت دارد. در برخی از رویکردها از جمله رویکردهای مبتنی بر روش برآوردیابی کمترین مربعات خطا، حتی یک مشاهده پرت به گونه‌ای در مدل اثر مخرب می‌گذارد که نتایج از واقعیت بسیار فاصله داشته باشند. در این شرایط مدل بهینه برآورد شده شامل خطای نامتعارفی است و به کلیت مشاهدات برازش نمی‌شود. البته مرزبندی بین نقاط خوب و بد (یا نقاط پرت) ممکن است مبهم باشد و نیاز به معیارهای خاصی داشته باشد. بسته به اینکه مدل برازش شده تا چه میزان به کلیت و مرکز

مشاهدات برازش می‌شود، می‌توان هر مشاهده را بر حسب درجه و یا وزنی به عنوان داده نامتعارف و یا پرت در نظر گرفت. ماهیت و مقادیر وزن‌ها به فاصله هر مشاهده از مدل برازش شده و/یا فاصله از مرکز داده‌ها بستگی دارد، که با توجه به میزان تاثیرگذاری آن در مدل و بر اساس روابطی خاص تعیین می‌شود. لذا با کاهش وزن مشاهداتی که توسط مدل برازش خوبی ندارند و یا تاثیر مخربی در برآوردیابی مدل دارند، از تاثیرپذیری مدل نسبت به اینگونه مشاهدات کاسته می‌شود [۳، ۴، ۵، ۷].

در این مقاله به معرفی برخی از رویکردهای برآوردیابی مدل‌های رگرسیون فازی بر اساس توابع نیکویی برازش مبتنی بر مجموع وزنی خطاها (انحرافات یا باقیمانده‌ها) یا مجموع وزنی خطاهای مرتب شده و یا در حالت کلی مجموع وزنی توابعی از خطاهای مرتب شده پرداخته می‌شود. این رویکردها را می‌توان برای کلیه مدل‌های رگرسیون وزنی فازی با داده‌های ورودی–دقیق/فازی و خروجی–دقیق/فازی در نظر گرفت. روش‌های معرفی شده در این مقاله کلی است و بر مبنای معیارهای برازش وزنی، در حضور مشاهدات پرت، برآوردهای استواری برای پارامترهای مدل ارائه خواهند داد. در این روش‌ها مقادیر بهینه پارامترها به همراه وزن‌های بهینه هر مشاهده نیاز به معرفی الگوریتم‌ها و روش‌های بهینه‌سازی خاصی از جمله الگوریتم‌های وزن‌دهی تکرار شونده، دارند. در کاربرد و محاسبات اینگونه الگوریتم‌ها مبتنی بر ساختار مدل رگرسیون فازی، نوع ورودی–خروجی مشاهدات و نحوه محاسبه مقادیر خطاها یا باقیمانده‌ها هستند و به‌گونه‌ای عمل می‌کنند که داده‌های پرت وزن‌های کوچکتر و داده‌های خوب وزن‌های بزرگتری در برازش مدل دارند.

۲ رگرسیون فازی

در حالت کلی، یک مدل رگرسیون فازی به صورت زیر

$$\widetilde{y} = \widetilde{f_\beta(x)} \qquad (۱)$$

برای مدل‌سازی متغیرهای خروجی–فازی ($\widetilde{y}$) و k متغیر ورودی–دقیق/فازی ($\widetilde{x}/x$) و پارامترهای دقیق/فازی ($\widetilde{\beta}/\beta$) تعریف می‌شود. مدل (۱) را می‌توان در حالات متنوع زیر (ولی نه فقط محدود به آنها) مورد بررسی و تحلیل قرار داد:

۱. مدل با ورودی‌های-دقیق (x) خروجی-فازی ($\widetilde{y}$) و پارامترهای فازی $\widetilde{\beta}$

$$\widetilde{f_\beta(x)} = \widetilde{\beta}_0 \oplus (\widetilde{\beta}_1 \otimes x_1) \oplus \ldots \oplus (\widetilde{\beta}_k \otimes x_k).$$

۲. مدل با ورودی‌های-فازی ($\widetilde{x}$) خروجی-فازی ($\widetilde{y}$) و پارامترهای دقیق β

$$\widetilde{f_\beta(x)} = \beta_0 \oplus (\beta_1 \otimes \widetilde{x}_1) \oplus \ldots \oplus (\beta_k \otimes \widetilde{x}_k).$$

۳. مدل با ورودی‌های-فازی ($\widetilde{x}$) خروجی-فازی ($\widetilde{y}$) که در آن فقط پارامتر عرض از مبدا فازی است و بقیه پارامترها دقیق هستند

$$\widetilde{f_\beta(x)} = \widetilde{\beta}_0 \oplus (\beta_1 \otimes x_1) \oplus \ldots \oplus (\beta_k \otimes x_k).$$

۴. مدل با ورودی‌های-فازی ($\widetilde{x}$) خروجی-فازی ($\widetilde{y}$) و پارامترهای فازی $\widetilde{\beta}$

$$\widetilde{f_\beta(x)} = \widetilde{\beta}_0 \oplus (\widetilde{\beta}_1 \otimes \widetilde{x}_1) \oplus \ldots \oplus (\widetilde{\beta}_k \otimes \widetilde{x}_k).$$

ملاحظه ۱.۲. برخی از رویکردهای امکانی در مدل‌بندی رگرسیون فازی معرفی شده‌اند که در آنها خروجی مقادیر دقیق دارند و دیگر کمیتها از قبیل ورودی‌ها و/یا پارامترهای مدل می‌تواند با توجه به ماهیت و ساختار مساله به صورت دقیق و یا فازی اختیار می‌شود [۱۴، ۱۵].

در ادامه هدف این است که در یک حالت کلی و جامع پارامترهای β بر اساس یک مساله بهینه‌سازی مبتنی بر مجموع وزنی تابعی از خطاها برآورد شوند. خطاها که به صورت فاصله یا تفاضل بین مقادیر مشاهده شده متغیر خروجی و مقادیر برآورد شده آن تعریف می‌شوند نقشی بسیار مهم و اساسی در برآورد پارامترهای یک مدل رگرسیون (کلاسیک و فازی) دارند. در این باره و در ادامه، به منظور محاسبه خطاها و همچنین ساختاردهی تابع برازش وزنی در مساله بهینه‌سازی، از یک فاصله مانند $\mathcal{D}$ بین اعداد فازی استفاده می‌شود که بر اساس آن، یک مساله بهینه‌سازی

ساختاربندی می‌شود. انتخاب‌های متنوع از فاصله $\mathcal{D}$ باعث می‌شود که نیکویی برازش مدل‌ها به حالات و روش‌های مختلفی اندازه‌گیری شود. برای مطالعه چند فاصله بین مجموعه‌های فازی و تعمیم‌های آن به [۲] مراجعه کنید. از جملهٔ متداول‌ترین معیارهای نیکویی برازش می‌توان به روش کمترین مربعات خطا و روش کمترین قدرمطلق انحرافات اشاره نمود که در آنها مساله بهینه‌سازی به صورت زیر در نظر گرفته می‌شود

$$\min_{\beta} \sum_{i=1}^{n} \mathcal{D}^2(\widetilde{y}_i, \widetilde{f_\beta(x_i)}), \tag{2}$$

$$\min_{\beta} \sum_{i=1}^{n} \mathcal{D}(\widetilde{y}_i, \widetilde{f_\beta(x_i)}). \tag{3}$$

در حالات متداول بالا مجموع فاصله‌ها (یا مربع فاصله‌ها) بین مقادیر مشاهده شده متغیر خروجی $\widetilde{y}$ و مقادیر برآورد شده آن توسط مدل $\widetilde{f_\beta(x)}$ کمینه می‌شود.

۳ مدل‌های رگرسیون فازی-وزنی

در رویکرد رگرسیون فازی وزنی هدف این است که با بررسی و تحلیل خطاها، وزن بهینه هر مشاهده متناسب با اهمیت و تاثیر آن مشاهده در برآوردیابی پارامترها، مشخص شود. تاثیر مشاهدات پرت در یک مدل رگرسیون (کلاسیک و یا فازی) اغلب با باقیمانده‌های بزرگی در معیار نیکویی برازش مدل همراه است. بردار وزن بر اساس رتبه یا مقدار هر باقیمانده (خطا) قادر به کنترل تاثیر آن مشاهده (یا مقدار خطای متناظر با آن) در برآورد مدل بهینه است. بر اساس وزن بهینه اطلاعات بسیار مهمی از قبیل پرت بودن یک مشاهده استخراج می‌شود. بنابراین با توجه به ماهیت مشاهدات پرت و نحوه تاثیرگذاری آنها در برآورد پارامترها و/یا مدل، الحاق وزنی خطاها بر اساس عملگرهای وزنی جایگزین بسیار مناسبی برای معرفی رویکردهای استوار در مدل‌های رگرسیون فازی است. برای این منظور، در ادامه به معرفی برخی روش‌ها و رویکردهای جایگزین و استوار برای الحاق و جمع‌بندی نیکویی برازش هر مشاهده در مدل با استفاده از عملگرهای وزنی پرداخته می‌شود. همچنین برخی روش‌های حل عددی مسایل بهینه‌سازی اینگونه مدل‌ها بیان می‌شود.

۱.۳ رگرسیون فازی کمترین مربعات خطای وزنی

در رگرسیون فازی کمترین مربعات خطای متداول، هدف این است که پارامترهای β بر اساس یک مساله بهینه‌سازی مبتنی بر جمع خطاها برآورد شوند. معمولا اینگونه روش‌ها به کمینه‌سازی مساله بهینه‌سازی (۲) طبق اعمال شرایطی می‌پردازند. برای برآورد بهینه پارامترهای β بر اساس نوع انتخاب فاصله بین اعداد فازی رویکرهای متنوع و مختلفی ارایه شده است. رویکرد کمترین مربعات خطا توسط چاچی [۳] به رگرسیون فازی وزنی تعمیم داده شد که در آن پارامترهای بهینه بر اساس کمینه‌سازی مساله بهینه‌سازی زیر طبق اعمال شرایطی به دست می‌آیند

$$\min_{\beta} \sum_{i=1}^{n} w_i \mathcal{D}^2(\widetilde{y}_i, \widetilde{f_\beta(x_i)}), \qquad (۴)$$

که در آن w_i وزن مشاهده/باقیمانده/خطای iام است. توجه کنید که در تابع برازش وزنی (۴) وزن‌ها وابسته به خطاها، خطاها وابسته به مقادیر پارامترهایی هستند که خود از تابع هدفی به‌دست می‌آیند که وابسته به وزن‌ها است. برآورد بهینه پارامتر β در تابع برازش وزنی (۴) بر اساس یک الگوریتم وزن–دهی تکرار شونده به دست می‌آید [۳].

۲.۳ رگرسیون فازی کمترین قدرمطلق انحرافات وزنی

در رگرسیون فازی کمترین قدرمطلق انحرافات، هدف این است که پارامترهای β بر اساس یک مساله بهینه‌سازی مبتنی بر جمع قدرمطلق انحرافات برآورد شوند. معمولا اینگونه روش‌ها به کمینه‌سازی یک مساله بهینه‌سازی خطی منجر می‌شوند که در آن پارامترهای بهینه از مساله بهینه‌سازی (۳) طبق اعمال شرایطی به دست می‌آیند. برای برآورد بهینه پارامترهای β بر اساس نوع انتخاب فاصله بین اعداد فازی رویکرهای مختلفی ارایه شده است. این رویکرد می‌تواند به رگرسیون فازی وزنی تعمیم داده شود که در آن پارامترهای بهینه بر اساس کمینه‌سازی مساله بهینه‌سازی زیر طبق اعمال شرایطی به دست می‌آیند [۴]

$$\min_{\beta} \sum_{i=1}^{n} w_i \mathcal{D}(\widetilde{y}_i, \widetilde{f_\beta(x_i)}),$$

۳.۳ رگرسیون فازی M-برآوردگرها

چاچی و همکاران [۷] پیشنهاد کردند که معیار زیر نسبت به مشاهدات پرت حساسیت کمتری دارد

$$\min_{\beta} \sum_{i=۱}^{n} \rho(\mathcal{D}(\widetilde{y}_i, \widetilde{f_\beta(x_i)})), \tag{۵}$$

که در آن ρ یکی از توابع معرفی شده توسط هوبر [۱۲] است و در شرایط زیر صدق می‌کند:

- $\rho(e) \geq ۰$،
- $\rho(۰) = ۰$،
- $\rho(e) = \rho(-e)$،
- $\rho(e) \leq \rho(e')$، برای هر $|e| \leq |e'|$.

وزن‌های بهینه به همراه مجموعه جوابهای بهینه یک دستگاه معادلات از پارامترها در مساله بهینه‌سازی (۵) بر اساس یک الگوریتم وزن-دهی تکرار شونده به دست می‌آیند.

ملاحظه ۱.۳. رویکردهای مختلفی از رگرسیون فازی M-برآوردگرها را می‌توان بر حسب انتخابهای متنوع تابع $\rho(\cdot)$، نوع مدل $\widetilde{f_\beta(x)}$ و تعریف فاصله $\mathcal{D}(\cdot,\cdot)$ بین اعداد فازی معرفی و بررسی نمود.

۴ مدل‌های رگرسیون فازی-وزنی با خطاهای مرتب شده

در ادامه به معرفی برخی روش‌ها و رویکردهای وزنی پرداخته می‌شود که برای الحاق و جمع‌بندی نیکویی برازش هر مشاهده در مدل عملگرهای وزنی را بر مقادیر خطاهای مرتب شده (یا توابعی متعارف از آنها) اعمال می‌کند. معمولا اینگونه روش‌ها به کمینه‌سازی یک مساله بهینه‌سازی با الگوریتم‌های تکرار شونده که اغلب پیچیده و زمان‌بر هستند، منجر می‌شوند. در اینگونه روش‌ها

پارامترهای بهینه از حل مساله بهینه‌سازی زیر طبق اعمال شرایطی به دست می‌آیند

$$\min_{\beta} \sum_{i=1}^{n} w_i \rho(e_{(i)}), \tag{6}$$

که در آن $e_{(1)} \leq e_{(2)} \leq \ldots \leq e_{(n)}$ مقادیر خطاهای مرتب شده هستند و

$$e_i = \mathcal{D}(\widetilde{y}_i, \widetilde{f_\beta(x_i)}), \qquad i = 1, \ldots, n.$$

رویکردهای مختلفی از رگرسیون فازی استوار را می‌توان به عنوان حالت خاصی از مساله بهینه‌سازی بالا تعریف و بررسی نمود. به این منظور چاچی و روزبه [۱، ۶] در مساله بهینه‌سازی (۶) با در نظر گرفتن $\rho(e) = e^2$ و اختیار بردار وزن‌ها به صورت

$$w_i = \begin{cases} 1 & i = 1, \ldots, h, \\ 0 & i = h+1, \ldots, n, \end{cases}$$

معیار نیکویی برازش زیر

$$\min_{\beta} \sum_{i=1}^{h} e_{(i)}^2, \tag{7}$$

که به رویکرد برآوریابی کمترین توان دوم خطاهای پیراسته شناخته شده است، را معرفی و بررسی نمودند. این معیار نسبت به مشاهدات پرت حساسیت خیلی کمی دارد و در آن h تعداد نقاط خوب و $n-h$ تعداد نقاط بد یا پرت می‌باشد و اغلب از طریق رابطه $[\frac{n+k+1}{2}]$ تعیین می‌شوند ($k+1$ تعداد پارامترهای مدل است). برای حل مساله بهینه‌سازی (۷) از روش‌های متداول بهینه‌سازی و بسته‌های رایج نرم‌افزاری استفاده شده است. کدهای برنامه‌نویسی اینگونه مسایل بهینه‌سازی بر مبنای الگوریتم زیر، معروف به روش انتخاب کلیه زیرمجموعه‌های به حجم h از داده‌ها، بنا نهاده شده اند:

۱. کلیه داده‌های به حجم h از n مشاهده را در نظر بگیرید.

۲. مساله رگرسیون کمترین مربعات خطای متداول (به عنوان مثال مساله بهینه‌سازی مربوط به رابطه (۲)) را برای هر مجموعه دادهٔ به حجم h حل کنید، و کلیه خطاهای $e_{(1)} \leq e_{(2)} \leq \ldots \leq e_{(n)}$ را محاسبه نمایید.

۳. مدلی که در بین تمام مدل‌های گام ۲ کمترین مقدار تابع هدف (۷) را داشته باشد به عنوان مدل بهینه معرفی می‌شود.

برای آشنایی بیشتر با روش‌ها و الگوریتم‌های سریع در حل اینگونه مسایل بهینه‌سازی به [۴] مراجعه کنید.

مساله بهینه‌سازی (۶) متناسب با انتخاب‌های خاصی از وزن‌های w_i، تابع $\rho(\cdot)$، نوع مدل $\widetilde{f_\beta(x)}$ و تعریف خاصی از فاصله $\mathcal{D}(\cdot,\cdot)$ بین اعداد فازی به حالات زیر تعمیم داده شده است:

۱. مدل کمترین قدرمطلق انحرافات پیراسته [۴]

$$\min_{\beta} \sum_{i=1}^{h} e_{(i)}. \qquad (8)$$

۲. مدل کمترین میانه قدرمطلق انحرافات [۴]

$$\min_{\beta} e_{(M)}, \qquad (9)$$

که در آن $e_{(M)} = median\{e_{(1)}, e_{(2)}, \ldots, e_{(n)}\}$.

۳. مدل کمترین میانه مربعات خطا [۱۰، ۱۱]

$$\min_{\beta} e^2_{(M)}, \qquad (10)$$

که در آن $e^2_{(M)} = median\{e^2_{(1)}, e^2_{(2)}, \ldots, e^2_{(n)}\}$.

۴. مدل کمترین چندک رتبه hام مربعات خطا [۵]

$$\min_{\beta} e^{2}_{(h)}, \qquad h = [\frac{n}{2}] + 1, \ldots, n. \qquad (11)$$

که در آن $e^{2}_{(1)} \leq e^{2}_{(2)} \leq \ldots \leq e^{2}_{(h)} \leq \ldots \leq e^{2}_{(n)}$.

۵. مدل کمترین چندک رتبه hام قدرمطلق انحرافات [۵]

$$\min_{\beta} e_{(h)}, \qquad h = [\frac{n}{2}] + 1, \ldots, n. \qquad (12)$$

که در آن $e_{(1)} \leq e_{(2)} \leq \ldots \leq e_{(h)} \leq \ldots \leq e_{(n)}$.

ملاحظه ۱.۴. در رویکردهای مربوط به مسایل بهینه‌سازی رابطه‌های (۸)، (۹)، (۱۰)، (۱۱)، (۱۲)، درصدی از داده‌ها (یا تعدادی از آنها) به عنوان مشاهداتی که قابلیت پرت بودن را دارا هستند، در برآورد مدل بهینه وارد نمی‌شوند. در رویکرد $\min_h e_{(h)}$ چندک رتبه hام خطاها کمینه می‌شوند. به عبارتی مدل بهینه بر اساس یک مجموعهٔ مشاهدات خوب hتایی از داده‌ها برآورد می‌شود. بنابراین اینگونه برآوردگرها تحت تاثیر حداقل $n-h$تا از داده‌های پرت قرار نمی‌گیرد. در محاسبات و کاربرد تعیین مقدار بهینه h یک مساله با اهمیت است که به دقت و درستی باید تعیین شود [۴، ۵]. در این رویکردها h را مقدار نقطه شکست[1] برآوردگرها تعریف می‌کنند، که عبارت است از کمترین تعداد مشاهدهٔ پرتی که با ورود به آماره (یا برآوردگر) مقدار آن را به شدت تغییر می‌هد. به عنوان مثال نقطه شکست میانگین برابر $h = 0$ است و برای میانه برابر $h = \frac{n}{2}$ است.

۵ نتیجه‌گیری

رویکردهای وزنی این توانایی را در برآورد بهینه پارامترهای مدل دارند که متناسب با مقادیر خطایی که هر مشاهده بوجود می‌آورد (یا متناسب با تابعی متعارف از مقادیر خطا) به تعیین وزن‌های نامنفی مربوط به هر مشاهده بپردازد. بنابراین در اینگونه از رویکردها به راحتی قادر هستیم که

[1]Breakdown Point

باقیمانده‌های با مقادیر بزرگ یا کوچک، و یا از جهتی آنهایی که در رتبه‌های خاصی هستند را در برآورد پارامترهای مدل کم اثر یا پر اثر کنیم. به عبارتی بردار وزن تاثیر هر مشاهده در برآوردیابی مدل را با توجه به مقدار خطا و رتبه خطا تعیین می‌کند. لذا بر این اساس می‌توان به شناسایی مشاهدات پرت پرداخت. واضح است که رویکردهای متداول کمترین مربعات خطای فازی یا کمترین انحراف باقیمانده‌های فازی و همچنین برخی از رویکردهای استوار مانند کمترین انحراف‌های پیراسته حالت خاصی از رویکرد وزنی (۶) هستند. در کاربرد بهتر است که در مجموعه مشاهداتی که مشکوک به وجود داده‌های پرت هستند از رویکرهای استوار وزنی معرفی شده در این مقاله استفاده شود.

مراجع

[۱] چاچي، ج. و روزبه، م. (۱۳۹۶). مدل‌سازی داده‌های مهندسی آب با استفاده از روش رگرسیون فازی استوار کمترین مربعات پیراسته، مجله مدل‌سازی پیشرفته ریاضی، دوره ۷، شماره ۱، بهار و تابستان ۱۳۹۶، صص ۱ تا ۱۸.

[۲] رضایی، ک. و رضایی، ح. (۱۳۹۷). بررسی معیارهای فاصله و شباهت برای مجموعه‌های فازی و برخی از توسعه‌های آنها، سیستم‌های فازی و کاربردها، دوره ۲، شماره ۱، پاییز وزمستان ۱۳۹۷، در حال چاپ.

[3] Chachi, J. (2019). A weighted least squares fuzzy regression for crisp input-fuzzy output data, IEEE Transactions on Fuzzy Systems, DOI: 438 10.1109/TFUZZ.2018.2868554.

[4] Chachi, J., Chaji, A. (2019). Detection of outliers problems and weighted fuzzy regression, Submitted.

[5] Chachi, J., Chaji, A. (2019). Quantile fuzzy regression and detection of outlier problems, Submitted.

[6] Chachi, J., Roozbeh, M. (2017). A fuzzy robust regression approach applied to bedload transport data, Communications in Statistics-Simulation and Computation, 47(3), 1703-1714.

[7] Chachi, J., Taheri, S. M., D'Urso, P., (2019). M-Estimates for Least-Squares Fuzzy Regression, Submitted.

[8] Chachi, J., Taheri, S. M., Fattahi, S. and Ravandi, S. A. H. (2016). Two robust fuzzy regression models and their application in predicting imperfections of cotton yarn, J. Textiles Polym., 4(2), 60-68.

[9] Coppi, R., D'Urso, P., Giordani, P., Santoro, A. (2006). Least squares estimation of a linear regression model with lr fuzzy response, Computational Statistics and Data Analysis, 51, 267-286.

[10] D'Urso, P., Massari, R. (2013). Weighted least squares and least median squares estimation for the fuzzy linear regression analysis, Metron, 71, 279-306.

[11] D'Urso, P., Massari, R., Santoro, A. (2011). Robust fuzzy regression analysis, Information Sciences, 181, 4154-4174.

[12] Huber, P., Ronchetti, E. M. (2009). Robust Statistics, 2ed., Wiley, NJ.

[13] Leski, J. M., Kotas, M. (2015). On robust fuzzy c-regression models, Fuzzy Sets and Systems, 279, 112-129.

[14] Tanaka, H., Hayashi, I., Watada, J. (1989). Possibilistic linear regression analysis for fuzzy data, European J. Operational Research, 40, 389-396.

[15] Tanaka, H., Uegima, S., Asai, K. (1982). Linear regression analysis with fuzzy model, IEEE Trans. Syst. Man Cybernet., 12, 903-907.

[16] Varga, S. (2007). Robust estimations in classical regression models versus robust estimations in fuzzy regression models, Kybernetika, 43, 503-508.