

ارائه یک مدل رگرسیونی جدید با رویکرد برنامه‌ریزی آرمانی فازی نوع-۲ بازه‌ای مقدار

میکائیل مختاری^۱، محمدحسن بهزادی^{۲*}، توفیق الهویرنلو^۳

(^۱) گروه آمار، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

(^۳) گروه ریاضی، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

تاریخ ارسال مقاله: ۱۴۰۰/۰۹/۲۱ تاریخ پذیرش مقاله: ۱۴۰۱/۱۱/۰۴

چکیده

استفاده از کلمات مبهم و غیردقیق (مانند تاحدی، کم و بیش، نسبتاً زیاد) در گفتار و قضاوت‌ها، تصادفی بودن برخی از وقایع، عدم توانایی ابزارهای فعلی در اندازه‌گیری متغیرهای کیفی (مانند رضایت، کیفیت، احساس) باعث شده تا بیشتر داده‌ها و اندازه‌گیری‌ها با درجاتی از عدم اطمینان همراه باشد. در این راستا، مدل‌های رگرسیون فازی با ترکیب روش‌های آماری با رویکردهای فازی از طریق ایجاد رابطه تابعی بین متغیرهای ورودی و خروجی، ابزارهای کارآمدی را برای تجزیه و تحلیل انواع مختلف عدم قطعیت (به ویژه عدم قطعیت احتمالی و عدم قطعیت فازی) فراهم می‌کنند. در مطالعه حاضر به منظور پر کردن شکاف رگرسیون فازی نوع-۱، یک مدل رگرسیونی (دو مرحله‌ای) با رویکرد برنامه‌ریزی آرمانی فازی نوع-۲ بازه‌ای ارائه می‌شود. برای برآورد ضرایب رگرسیون فازی، یک روش جدیدی برای مینیمم کردن قدرمطلق خطاها ارائه نموده و سپس از طریق متغیرهای کمکی، چهارچوب جدیدی برای حل مسائل برنامه‌ریزی آرمانی فازی نوع-۲ بازه‌ای پیشنهاد می‌کنیم. همچنین، برای ارزیابی عملکرد رویکرد پیشنهادی، چند معیار جدید را بر اساس اندازه‌های شباهت و فاصله فازی معرفی می‌کنیم. سرانجام، به منظور تشریح نتایج نظری روش پیشنهادی و تبیین چگونگی استفاده از آن برای بدست آوردن مدل رگرسیون فازی نوع-۲ بازه‌ای، دو مثال عددی ارائه می‌کنیم.

واژه‌های کلیدی: برنامه‌ریزی آرمانی فازی نوع-۲ بازه‌ای، قدر مطلق انحرافات فازی، ضرایب رگرسیونی فازی نوع-۲ بازه‌ای، مشاهدات فازی دور افتاده، عدم قطعیت احتمالی، عدم قطعیت فازی

۱- مقدمه

اغلب پدیده‌های جهان و یا شاید همه‌ی آن‌ها با نوعی از عدم قطعیت و نادقیقی همراه بوده و بر پایه دانش فعلی، تصادفی بودن (عدم قطعیت احتمالی^۲) و فازی بودن^۳ (عدم قطعیت فازی^۴) به عنوان دو منبع اصلی ایجاد عدم قطعیت و نادقیقی معرفی شده‌اند [۲۰۱]. عدم قطعیت احتمالی یک نوع عدم قطعیت عینی است و از تغییرپذیری یا غیر قابل پیش‌بینی بودن پدیده‌های تصادفی یا آزمایشات ناشی شده و با گذر زمان و یا از طریق انجام آزمایش کاملاً معین می‌شود در حالی که عدم قطعیت فازی هیچ ارتباطی با زمان و آزمایش نداشته و به مواردی نظیر تفاوت در برداشت معنایی، ناکافی بودن داده‌ها، ضعف دانش و ابزار بشری در شناخت پیچیدگی پدیده‌ها و همچنین عدم صراحت و عدم شفافیت در گفتار و قضاوت‌های انسانی (استفاده از کلمات نادقیق و مبهم^۵) برمی‌گردد [۳۰۱]. عدم قطعیت فازی با رویکردهای منطق فازی^۶ و عدم قطعیت احتمالی (تصادفی بودن^۷) با متدهای آمار و احتمال اندازه‌گیری، توصیف و تحلیل می‌شود. البته باید در نظر داشت که روش‌های آماری و فازی نه تنها در تقابل هم نیستند بلکه مکمل هم بوده و در اغلب مدل‌سازی‌ها، خصوصاً در مدل‌های رگرسیونی از تلفیق این دو رویکرد استفاده می‌شود. رگرسیون کلاسیک (که اولین بار توسط فرانسیس گالتون معرفی شده است) کارآمدترین و شناخته شده‌ترین مدل ریاضی است که از طریق ایجاد رابطه تابعی بین متغیر پاسخ و متغیرهای توضیحی، می‌تواند داده‌های تصادفی را مدل‌سازی و تحلیل نماید. با این حال، این مدل‌ها با وجود جامعیت نسبی و عملکرد قابل قبول، از انعطاف بالایی برخوردار نمی‌باشند [۵ و ۶] زیر:

۱. اجزاء تشکیل دهنده رگرسیون معمولی (یعنی متغیرهای مستقل، متغیرهای وابسته، پارامترها و جمله خطا) همگی باید کمیت‌های عددی دقیق و غیرفازی بوده و ثانیاً فرضیات زیر بنایی مدل در مورد داده‌های نمونه‌گیری شده، پارامترها و خطاها^۸ نیز باید برقرار بماند.

۲. اغلب داده‌ها و اندازه‌گیری‌هایی که منبای برنامه‌ریزی‌ها و تصمیم‌گیری‌ها قرار می‌گیرند علاوه بر اینکه ممکن است تصادفی باشند، در همین حال ممکن است نادقیق، غیر صریح، ناکافی و مبهم بوده و با درجه‌ای از عدم قطعیت فازی همراه باشند که رگرسیون معمولی از ابزارهای لازم برای مدل‌سازی و تجزیه و تحلیل این قبیل داده‌ها برخوردار نمی‌باشد.

از اینرو، در مواردی که فرضیات زیربنایی رگرسیون کلاسیک برقرار نبوده و یا حداقل یکی از اجزاء تشکیل دهنده آن فازی باشد، رگرسیون فازی جایگزین مناسبی محسوب می‌شود. رگرسیون فازی نوع-۱، برای اولین بار در سال ۱۹۸۲ توسط تاناکا و همکاران [۷]. معرفی شده و یک روش ناپارامتریک و تعمیم یافته رگرسیون معمولی است. رگرسیون فازی نوع-۱ با تلفیق تکنیک‌های آماری با رویکردهای فازی و از طریق ایجاد رابطه تابعی بین مجموعه‌ای از متغیرهای وابسته و مستقل، مجموعه‌ای از ابزارهای کارآمد را برای مدل‌سازی و تجزیه و تحلیل داده‌هایی که

^۲ Probabilistic uncertainty^۳ Fuzziness^۴ Fuzzy uncertainty^۵ کلمات نادقیقی نظیر تاحدودی، کم و بیش، نسبتاً زیاد، ناکافی، مناسب، کمی بهتر، و...، کلمات مبهم نامیده می‌شوند.^۶ در منطق فازی مجموعه‌ای از معلومات نادقیق، نامطمئن و سر بسته که به نحوی با قضاوت، تجربیات، دانسته‌ها و اندیشه خبرگان و تصمیم‌گیران سروکار داشته و ممکن است با الفاظ و کلمات محاوره‌ای تعریف شده باشند با استفاده از توابع عضویت به ارزشهای عددی در بازه [۰ و ۱]. تبدیل وقابل پردازش کامپیوتر می‌شوند. در سیستم‌های فازی نحوه پردازش اطلاعات، تقریباً شبیه به نظام پردازش اطلاعات در مغز انسان است [۲۰۱].^۷ آن دسته از پدیده‌هایی که قابل آزمایش و تجربه شدن بوده و نتایج حاصل از آزمایش و تجربه از روند خاصی پیروی نموده و بتوان پس از آزمایش‌های بسیار، نتیجه‌ی آزمایش جدید را پیش‌بینی نمود، به پدیده‌های تصادفی معروفند.^۸ در یک مدل رگرسیونی کلاسیک خطاها (تفاضل بین مقادیر مشاهده شده و برآورد شده متغیر پاسخ)، ناهمبسته و دارای توزیع نرمال با میانگین صفر و واریانس ثابت هستند

همزمان با درجه‌ای از عدم قطعیت تصادفی و فازی همراه هستند در اختیار ما قرار می‌دهد و از این طریق به بخشی از دغدغه‌ها و چالش‌های پیش روی محققین در خصوص همراه و هماهنگ ساختن ایده‌های نظری با تکنیک‌های عملی و نیز تبدیل ادراک کیفی نادقیق و سربسته به سنجش کمی راه حل‌های قابل قبولی را ارائه می‌کند [۴ و ۵]. یک مدل رگرسیونی فازی شامل ورودی (متغیرهای مستقل^۹ یا توضیحی^{۱۰})، خروجی (متغیر وابسته^{۱۱} یا پاسخ^{۱۲}) و ضرائب رگرسیونی است و بسته به اینکه هر یک از کمیت‌های مورد اشاره فازی یا غیر فازی^{۱۳} در نظر گرفته شوند سه رسته عمده برای تحلیل و مدل‌سازی رگرسیونی در محیط فازی به شرح زیر حاصل می‌شود:

- ۱- متغیرهای مستقل و وابسته فازی و پارامترهای مدل غیر فازی باشند.
 - ۲- متغیر وابسته و پارامترهای مدل فازی و متغیرهای مستقل غیر فازی باشند.
 - ۳- متغیرهای مستقل، وابسته و پارامترهای مدل همگی از نوع فازی باشند.
- علاوه بر تقسیم‌بندی فوق عوامل دیگری نیز به تنوع مدل‌های رگرسیونی در محیط فازی می‌افزاید که می‌توان به درجه مجموعه فازی^{۱۴}، نوع عدد فازی^{۱۵} و نیز روشی که برای برآورد پارامترهای مدل رگرسیون فازی به کار برده می‌شود اشاره نمود. از سال ۱۹۸۲ که رگرسیون فازی توسط تاناکا [۷] معرفی شده شیوه‌ها و تکنیک‌های متنوعی برای برآورد پارامترهای مدل رگرسیونی ارائه شده و این تکنیک‌ها از نظر دقت برازش، کارایی، استواری و سادگی محاسبات مورد مقایسه قرار گرفته و نهایتاً از نقطه نظر برآورد پارامترها در سه دسته عمده زیر گروه بندی شده اند [۸ و ۹].

۱. برآورد پارامترها با رویکرد مبتنی بر کمترین قدرمطلق انحرافات فازی^{۱۶} (FLAD)
 ۲. برآورد پارامترها با رویکرد مبتنی بر کمترین توان‌های دوم انحرافات فازی^{۱۷} (FLSD)
 ۳. برآورد پارامترها با رویکردهای ابتکاری^{۱۸} و استفاده همزمان از تکنیک FLAD و FLSD
- برآورد پارامترها با رویکرد کمترین قدرمطلق انحرافات فازی برگرفته از نظریه امکان و ریاضیات فازی بوده و اولین بار توسط تاناکا [۷] در سال ۱۹۸۲ معرفی شده است. در این روش ضرایب رگرسیونی مدل با استفاده از برنامه‌ریزی ریاضی (LP^{۱۹}، NLP^{۲۰}، GP^{۲۱} و ...) و از طریق کمینه کردن ابهام کل مدل برآورد می‌شوند. از سال ۱۹۸۲ به بعد، رویکرد FLAD با انواع مختلف ورودی و خروجی و با الگوریتم‌های گوناگون توسط محققین مورد نقد و بررسی قرار گرفته و مدل بهبود یافته این تکنیک اساس بسیاری از تحقیقات در حوزه تحلیل رگرسیون با داده‌های فازی نوع-۱ و ۲، فازی بازه‌ای مقدار نوع-۱ و ۲ و داده‌های فازی شهودی و... شده است که ذیلاً به برخی از این مطالعات اشاره شده است:

^۹ Independent variables

^{۱۰} Explanatory

^{۱۱} dependent variables

^{۱۲} Response

^{۱۳} Crisp or non fuzzy

^{۱۴} در حال حاضر مجموعه‌های فازی در سطح وسیعی تعمیم یافته که عمده ترین آنها عبارتند از مجموعه‌های فازی نوع-۱، نوع-۲، مجموعه‌های فازی بازه‌ای، مجموعه‌های فازی نوع-۲ کلی، مجموعه‌های فازی نوع-۲ بازه‌ای، مجموعه‌های فازی شهودی آتاناسوف نوع-۱ و نوع-۲ و مجموعه‌های فازی نوع-۲ مردد.

^{۱۵} هر کدام از مجموعه‌های فازی را می‌توان با توابع عضویت اعداد فازی نرمال، مثلثی، دوزنقه‌ای، سهموی و ... تعریف نمود که در این صورت نیز نحوه برآورد پارامترها و توصیف متغیرها بسته به اینکه کدام نوع از اعداد استفاده شده متفاوت خواهد بود.

^{۱۶} Fuzzy least absolute deviation

^{۱۷} Fuzzy least squar deviation

^{۱۸} Heuristic

^{۱۹} Linear programming

^{۲۰} Non linear programming

^{۲۱} Goal programming

در سال ۲۰۰۷ چوی و باکلی [۶] رویکرد FLAD را به منظور برآورد پارامترهای یک مدل رگرسیونی با ورودی-خروجی فازي مثلثی و پارامترهای غیر فازي بکار برد. در همان سال چن و هسویه [۵] و نیز شکوری و همکاران [۱۰] رویکرد برنامه‌ریزی آرمانی را جهت ارائه یک مدل رگرسیون فازي بر پایه برآوردگر FLAD پیشنهاد نمودند. در سال ۲۰۱۲ طاهری و کلکین نما [۱۱] با رویکرد کمترین قدرمطلق انحرافات فازي و از طریق کمینه کردن کل فاصله بین مقدار پاسخ‌های فازي مشاهده و برآورد شده، یک مدل رگرسیون فازي با ورودی-خروجی فازي مثلثی و پارامترهای قطعی معرفی کردند. لی و همکاران [۸] با همان روش یک مدل رگرسیونی با ورودی-خروجی فازي دوزنقه‌ای پیشنهاد دادند. همچنین در سال ۲۰۱۷ زنگ و همکاران [۱۲] با ورودی غیر فازي و ضرایب و خروجی فازي مثلثی یک مدل جدید ارائه نمودند. برخی محققین برای برآورد پارامترهای رگرسیون فازي برپایه روش FLAD، از رویکرد برنامه‌ریزی خطی فازي (LP)، برنامه‌ریزی آرمانی فازي (GP) و یا برنامه‌ریزی غیر خطی فازي (NLP) استفاده نموده‌اند. در این خصوص می‌توان به مطالعه حجتی و همکاران [۱۳] اشاره نمود که در سال ۲۰۰۵ با ورودی-خروجی و پارامترهای فازي مثلثی متقارن و با رویکرد برنامه‌ریزی آرمانی فازي و برپایه روش FLAD پارامترهای مدل رگرسیون فازي را برآورد نمودند. در سال ۲۰۱۱ حسن پور و همکاران [۱۴] رویکرد کمترین قدرمطلق انحرافات فازي و تکنیک برنامه‌ریزی آرمانی فازي را به منظور برآورد پارامترهای مدل رگرسیونی که در آن متغیرهای مستقل و وابسته اعداد فازي مثلثی بودند استفاده کردند. همچنین محققین دیگری همچون عارفی [۱۵]، لی و همکاران [۸] چن و نین [۹]، کلکین نما و طاهری [۱۶] نیز از متد FLAD استفاده نموده‌اند.

رویکرد مبتنی بر کمترین توان‌های دوم انحرافات فازي (FLSD)، که هدف آن کمینه کردن مجموع توان دوم خطای فازي بین مقدار پاسخ مشاهده شده و برآورد شده می‌باشد، نخستین بار در سال ۱۹۸۸ توسط دیاموند [۱۷] معرفی گردید و اخیراً رویکردهای بهبود یافته این روش توسط برخی از محققین با انواع مختلف ورودی/خروجی و با الگوریتم‌های متفاوت مورد مطالعه قرار گرفته و در عمل نیز بصورت موفقیت‌آمیز مورد استفاده واقع شده‌اند. مطالعه درسو و ماساری [۱۸] از جمله آنها است که با روش کمترین میانه وزنی و با برآوردگر FLSD، پارامترهای مدل رگرسیونی با ورودی و خروجی فازي را برآورد نموده و به منظور بررسی نیکویی برازش مدل معرفی شده یک روش ابتکاری معرفی نمودند. ربیعی و همکاران [۱۹] رویکرد حداقل مربعات را برای ارائه یک مدل رگرسیونی فازي با داده‌ها و پارامترهای بازه‌ای مقدار استفاده نمودند. همچنین محققین دیگری همچون رفیعی و قریشی [۲۰]، خمار و همکاران [۲۱] از متد FLSD استفاده نموده‌اند. با این حال، از نقطه نظر پایداری و استواری مدل، روش FLSD به نقاط پرت خیلی حساس بوده و در صورت وجود نقاط پرت در داده‌ها روش FLAD بر روش FLSD ترجیح داده می‌شود. البته می‌دانیم که در برخی موارد وجود حتی یک نقطه دور افتاده در مجموعه داده‌ها منجر به ایجاد نوسان در آماتیک در مقدار پارامترهای برآورد شده می‌شود.

در سال‌های اخیر رویکردهای ابتکاری نیز که اغلب با ترکیبی از روش‌های FLAD، FLSD و با بهره‌گیری از زبان ماشین انجام می‌شود از سوی محققین انجام و مورد نقد و بررسی قرار گرفته که از جمله آنها می‌توان به مطالعات حسین زاده و همکاران [۲۲]، وی و همکاران [۲۳] و حسامیان و همکاران [۲۴] اشاره کرد.

با مرور و بررسی دقیق و جامع مطالعات انجام شده در زمینه رگرسیون فازي (در داخل و خارج کشور)، برای مولفین مسجل گردید که هیچ مطالعه‌ای در مورد روش رگرسیون خطی فازي بر اساس داده‌های فازي نوع-۲ بازه‌ای مقدار و با رویکردی که ما ارائه خواهیم کرد، انجام نشده است. بر این اساس، در مطالعه حاضر یک مدل رگرسیون خطی که در آن متغیرهای توضیحی، متغیر پاسخ و ضرایب رگرسیونی از نوع اعداد فازي نوع-۲ بازه‌ای

مثلی هستند ارائه می‌شود. برتری و نکات قابل توجه این مطالعه در مقایسه با سایر مطالعات انجام شده در خصوص رگرسیون خطی فازی، به شرح زیر خلاصه شده است:

۱. در مطالعه حاضر عملگرهای جبری اعداد فازی نوع-۱ به اعداد فازی نوع-۲ بازه‌ای مقدار تعمیم داده شده و سپس برای اندازه‌گیری فاصله اعداد فازی نوع-۲ بازه‌ای مقدار و رتبه بندی آنها چندین فرمول جدید معرفی می‌شوند.

۲. در مطالعه حاضر شباهت‌ها و تفاوت‌های بین مسائل برنامه‌ریزی خطی/غیر خطی و نیز برنامه‌ریزی آرمانی در محیط‌های غیرفازی، فازی نوع-۱ و فازی نوع-۲ بازه‌ای مقدار بر جسته شده و به این سوال پاسخ داده می‌شود که چگونه هر یک از این آنها می‌تواند به دیگری منجر شوند و سپس یک چارچوب کاملاً جدید برای تخمین ضرایب رگرسیون فازی نوع-۲ بازه‌ای پیشنهاد می‌گردد.

۳. در این مطالعه، روش جدیدی را برای تعیین نقاط پرت فازی خفیف و شدید معرفی نموده و سپس روش مذکور را برای شناسایی و حذف نقاط پرت فازی به کار می‌بریم.

۴. برای ارزیابی کارایی مدل پیشنهادی و مقایسه آن با مدل‌های موجود، چنین معیار جدید را بر پایه فاصله و شباهت فازی معرفی می‌کنیم. برای انجام موارد فوق، در بخش ۲، ابتدا مفاهیم مجموعه‌های فازی نوع-۲ بازه‌ای مقدار و نیز عملیات جبری بر روی این مجموعه‌ها به طور مفصل ارائه شده و سپس یک رویکرد جدید برای اندازه‌گیری فاصله اعداد فازی نوع-۲ بازه‌ای و نحوه رتبه بندی آنها ارائه می‌شود. همچنین شباهت‌ها و تفاوت‌های بین مسائل برنامه‌ریزی خطی/غیر خطی و نیز برنامه‌ریزی آرمانی در محیط‌های غیرفازی، فازی نوع-۱ و فازی نوع-۲ بازه‌ای مقدار بر جسته شده و به این سوال پاسخ داده می‌شود که چگونه هر یک از آنها چگونه می‌تواند به دیگری منجر شوند و سپس یک چارچوب کاملاً جدید برای تخمین ضرایب رگرسیون فازی نوع-۲ بازه‌ای پیشنهاد می‌گردد. در بخش ۳، یک مدل جدید رگرسیون فازی نوع-۲ بازه‌ای مقدار بر پایه برنامه‌ریزی آرمانی نوع-۲ بازه‌ای و با رویکرد کمترین انحرافات فازی معرفی می‌شود. در بخش ۴، روش جدیدی را برای شناسایی و حذف نقاط پرت فازی ارائه نموده و سپس مدل نهایی را با داده‌های بدون نقطه پرت ارائه می‌کنیم. همچنین، برای نشان دادن نتایج نظری رویکرد پیشنهادی و ارزیابی عمل کردهای آن، دو مثال عددی ارائه نموده و در بخش ۵ یک نتیجه‌گیری جامع ارائه می‌کنیم.

۲. مبانی نظری

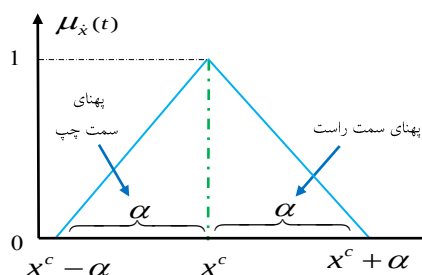
در این مطالعه نمادهای Ω و R به ترتیب بعنوان مجموعه مرجع و مجموعه اعداد حقیقی، F_1 بجای خانواده تمام اعداد فازی نوع-۱ در R ، F_2 بجای خانواده تمام اعداد فازی نوع-۲ بازه‌ای مقدار در R ، F_2^+ بجای اعداد فازی مثلی نوع-۲ بازه‌ای مقدار، $\oplus$ و $\otimes$ نیز به ترتیب برای جمع، تفریق و ضرب بین دو عدد فازی استفاده شده است.

۲-۱- مجموعه‌های فازی نوع-۲ و نوع-۲ بازه‌ای مقدار

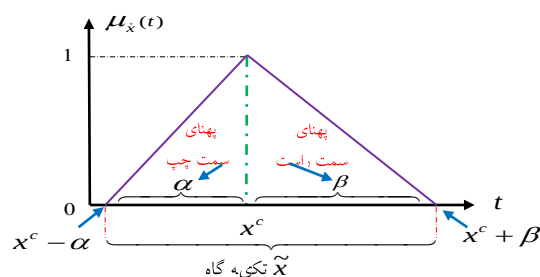
یک عدد فازی مثلی^{۲۲} مانند $\tilde{x}$ ، به صورت $\tilde{x} = (x^l, x^c, x^u)$ نمایش داده می‌شود که در آن x^u کران بالا و حداکثر مقداری است که عدد فازی $\tilde{x}$ می‌تواند اختیار کند. x^c محتمل‌ترین مقدار و x^l کران پایین و کمترین مقداری است که عدد فازی $\tilde{x}$ می‌تواند اختیار کند ضمن اینکه عدد فازی مثلی $\tilde{x}$ برحسب پهناهای چپ و راست به صورت $\tilde{x} = (x^c, \alpha, \beta)$ بیان می‌شود که α پهناهای چپ، β پهناهای راست و x^c مرکز $\tilde{x}$ هستند. در ضمن اگر $\alpha = \beta$ باشد عدد فازی مثلی را متقارن می‌نامند [۲۵]. تابع عضویت یک عدد فازی مثلی به شکل روبرو بیان می‌شود:

^{۲۲} Triangular fuzzy number

$$\dot{x}(t) = \begin{cases} \frac{t - (x^c - \alpha)}{\alpha}, & x^c - \alpha \leq t \leq x^c \\ 1, & t = x^c \\ \frac{x^c + \beta - t}{\beta}, & x^c \leq t \leq x^c + \beta \end{cases}$$



شکل ۲-۲ نمایش عدد فازی مثلثی متقارن



شکل ۲-۱ نمایش عدد فازی مثلثی نامتقارن

یک مجموعه فازی نوع-۲ متشکل از بی‌شمار مجموعه فازی نوع-۱ در یک مجموعه فازی است که با توابع عضویت اولیه و ثانویه تعریف می‌شوند [۲۵-۲۸] و چون خود توابع عضویت اولیه آنها یک مجموعه فازی نوع-۱ است به همین علت این مجموعه‌ها در مدل کردن عدم قطعیت‌های زبانی که برگرفته از قضاوت‌ها و ادراکات انسانی است از کارآمدی بسیار بالایی برخوردار هستند.

تعریف ۲-۱: اگر $\tilde{x}$ یک مجموعه فازی نوع-۲ روی مجموعه مرجع Ω باشد در این صورت:

$\tilde{x} = \int_{t \in \Omega} \int_{u(t) \in J_t \subseteq [0,1]} \mu_{\tilde{x}}(t, u(t)) / (t, u(t))$ بطوری که $t \in \Omega$ متغیر اولیه، J_t تابع عضویت اولیه t و $u \in J_t \subseteq [0,1]$ متغیر ثانویه و $\mu_{\tilde{x}}(t, u) \in [0,1]$ درجه عضویت ثانویه t می‌باشد [۲۵-۲۸]. اجتماع تمام توابع عضویت‌های اولیه یعنی $\bigcup_{t \in \Omega} J_t$ با یک ناحیه کراندار تحت عنوان اثر عدم قطعیت (FOU^{۲۳}) توصیف می‌شوند یعنی

$$FOU(\tilde{x}) = \bigcup_{t \in \Omega} J_t = \bigcup_{t \in \Omega} [\tilde{x}'(t), \tilde{x}''(t)]$$

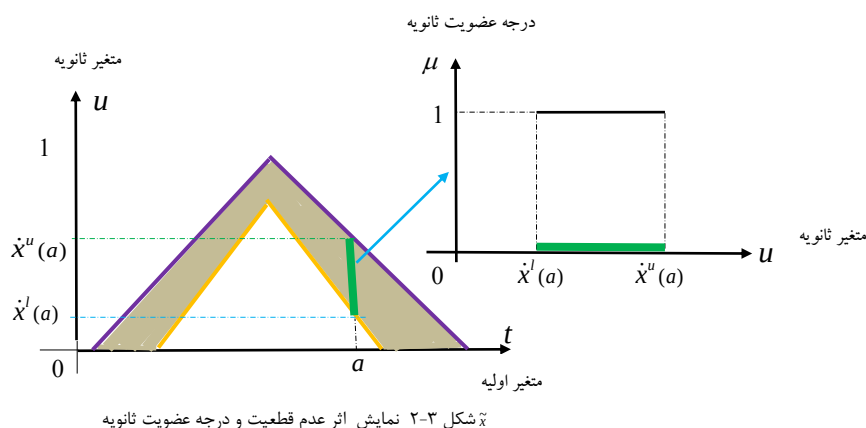
$\tilde{x}''$ تابع عضویت بالایی^{۲۴} و با نماد $\tilde{x}'' = UMF(\tilde{x})$ نمایش داده شده و $\tilde{x}'$ تابع عضویت پایینی^{۲۵} و با نماد $\tilde{x}' = LMF(\tilde{x})$ شناخته می‌شوند که هر دوی $\tilde{x}'$ و $\tilde{x}''$ مجموعه فازی نوع-۱ هستند. البته به دلیل وجود محاسبات پیچید در مجموعه‌های فازی نوع-۲، اغلب محققان ترجیح می‌دهند از فرم ساده شده آن یعنی مجموعه‌های فازی نوع-۲ بازه‌ای مقدار استفاده کنند [۲۵].

تعریف ۲-۲: اگر در یک مجموعه فازی نوع-۲ درجات عضویت ثانویه برابر یک در نظر گرفته شوند مجموعه حاصل فازی نوع-۲ بازه‌ای مقدار نامیده می‌شود. FOU مجموعه‌های فازی نوع-۲ بازه‌ای مقدار را بصورت کامل توصیف می‌کند، چون در این مجموعه‌ها درجات عضویت ثانویه مقدار ثابت ۱ داشته و اطلاعات جدیدی را منتقل نمی‌کند.

^{۲۳} Footprint of uncertainty

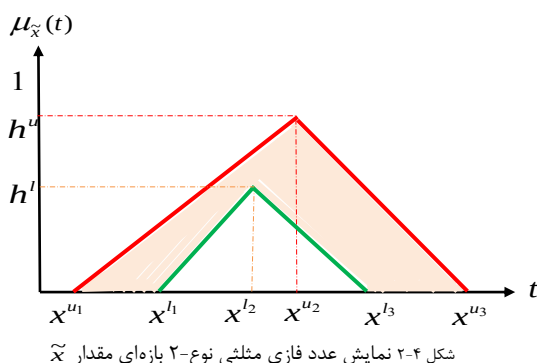
^{۲۴} Upper membership function(UMF)

^{۲۵} Lower membership function(LMF)



۲-۲- مجموعه‌های فازی مثلثی نوع-۲ بازه‌ای مقدار^{۲۶}

در مجموعه‌های فازی مثلثی نوع-۲ بازه‌ای مقدار با بهره‌گیری از انواع سیستم‌های رتبه‌بندی زبانی می‌توان ارزیابی‌های کلامی را به متغیرهای عددی عینی تبدیل نموده و زمینه را برای مدل کردن عدم قطعیت‌های زبانی که برگرفته از قضاوت‌ها و ادراکات انسانی است فراهم نمود نمونه‌ای از این رتبه‌بندی‌ها در جدول ۱ ارائه شده است. **تعریف ۲-۳:** اگر $\tilde{x}$ یک عدد فازی نوع-۲ باشد که هر دو تابع عضویت بالایی ($\tilde{x}^u$) و پائینی ($\tilde{x}^l$) آن اعداد فازی مثلثی باشند در این صورت $\tilde{x}$ عدد فازی مثلثی نوع-۲ بازه‌ای مقدار نامیده شده و به صورت $\tilde{x} = (\tilde{x}^u, \tilde{x}^l) = [(x^{u1}, x^{u2}, x^{u3}; h^u), (x^{l1}, x^{l2}, x^{l3}; h^l)]$ تعریف می‌شود.



تعریف ۲-۴: عدد فازی مثلثی نوع-۲ بازه‌ای مقدار $\tilde{x}$ نرمال کامل نامیده می‌شود اگر تابع عضویت بالایی ($\tilde{x}^u$) و پائینی ($\tilde{x}^l$) آن اعداد فازی مثلثی نرمال باشند که در این صورت $\tilde{x}$ به فرم زیر تعریف می‌شود:

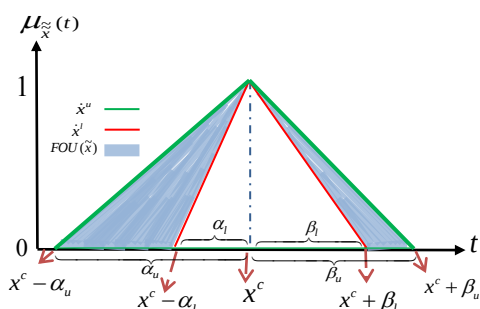
$$\tilde{x} = (\tilde{x}^u, \tilde{x}^l) = [(x^{u1}, x^{u2}, x^{u3}), (x^{l1}, x^{l2}, x^{l3})] = [(x^{u1}, x^c, x^{u3}), (x^{l1}, x^c, x^{l3})]$$

در این صورت تابع عضویت $\tilde{x}^u$ و $\tilde{x}^l$ به فرم زیر تعریف می‌شوند [۲۵ و ۲۸]:

$$\tilde{x}^u(t) = \begin{cases} \frac{t - x^{u1}}{x^c - x^{u1}}, & x^{u1} \leq t < x^c, \\ 1, & t = x^c, \\ \frac{x^{u2} - t}{x^{u2} - x^c}, & x^c < t \leq x^{u2}, \end{cases}, \quad \tilde{x}^l(t) = \begin{cases} \frac{t - x^{l1}}{x^c - x^{l1}}, & x^{l1} \leq t < x^c, \\ 1, & t = x^c, \\ \frac{x^{l2} - t}{x^{l2} - x^c}, & x^c < t \leq x^{l2}, \end{cases}$$

^{۲۶} Interval type - ۲ fuzzy number

تعریف ۲-۵: یک عدد فازی مثلثی نوع-۲ بازه‌ای مقدار نرمال^{۲۷} کامل بر پایه پهنای چپ و راست به فرم $\tilde{x} = [\tilde{x}^u, \tilde{x}^l] = [(x^c, \alpha_u, \beta_u), (x^c, \alpha_l, \beta_l)]$ بیان می‌شود که در آن α_u و β_u به ترتیب پهنای راست و چپ α_l و β_l پهنای چپ و راست $\tilde{x}^l$ و $\alpha_l, \beta_l, \alpha_u, \beta_u \geq 0$ می‌باشند [۲۵-۲۸].



شکل ۲-۵ نمایش عدد فازی مثلثی نوع-۲ بازه‌ای مقدار نامتقارن $\tilde{x}$

جدول ۱: متغیرهای زبانی و توابع عضویت مثلثی نوع-۲ بازه‌ای مقدار متناظر با آنها

مجموعه فازی	علامت اختصاری	عدد فازی	تابع عضویت $[(x_i^{u1}, x_i^{u2}), x^c, (x_i^{l1}, x_i^{l2})]$	بدینانه‌ترین حالت	ممکن‌ترین حالت	خوشبینانه‌ترین
خیلی خیلی کم	VVL	$\tilde{1}$	$[(0,0.4), 0.2, (0.1,0.3)]$	۰	۰.۲	۰.۴
خیلی کم	VL	$\tilde{2}$	$[(0.1,0.5), 0.3, (0.2,0.4)]$	۰.۱	۰.۳	۰.۶
کم	L	$\tilde{3}$	$[(0.2,0.6), 0.4, (0.3,0.5)]$	۰.۲	۰.۴	۰.۶
متوسط به پائین	ML	$\tilde{4}$	$[(0.3,0.7), 0.5, (0.4,0.6)]$	۰.۳	۰.۵	۰.۷
متوسط	M	$\tilde{5}$	$[(0.4,0.8), 0.6, (0.5,0.7)]$	۰.۴	۰.۶	۰.۸
متوسط به بالا	MH	$\tilde{6}$	$[(0.5,0.9), 0.7, (0.6,0.8)]$	۰.۵	۰.۷	۰.۹
زیاد (High)	H	$\tilde{7}$	$[(0.6,1), 0.8, (0.7,0.9)]$	۰.۶	۰.۷	۱
خیلی زیاد (Very high)	VH	$\tilde{8}$	$[(0.7,1), 0.9, (0.8,1)]$	۰.۷	۰.۸	۱
فوق‌العاده زیاد (Very Very High)	VVH	$\tilde{9}$	$[(0.8,1), 1, (0.9,1)]$	۰.۸	۰.۹	۱

۲-۳- عملیات جبری روی اعداد فازی مثلثی نوع-۲ بازه‌ای مقدار

تعریف ۲-۶: اگر $\tilde{x}_i, \tilde{y}_i \in \mathcal{F}_2$ در این صورت [۲۵-۲۸]:

$$\begin{aligned}
 ۱) \quad \tilde{x}_i \oplus \tilde{y}_i &= (\tilde{x}_i^u, \tilde{x}_i^l) \oplus (\tilde{y}_i^u, \tilde{y}_i^l) = [(\tilde{x}_i^u \oplus \tilde{y}_i^u), (\tilde{x}_i^l \oplus \tilde{y}_i^l)] \\
 &= [(x_i^{u1} + y_i^{u1}, x_i^{u2} + y_i^{u2}), x_i^c + y_i^c, (x_i^{l1} + y_i^{l1}, x_i^{l2} + y_i^{l2})] \quad (۱)
 \end{aligned}$$

$$\begin{aligned}
 ۲) \quad \tilde{x}_i \ominus \tilde{y}_i &= (\tilde{x}_i^u, \tilde{x}_i^l) \ominus (\tilde{y}_i^u, \tilde{y}_i^l) = [(\tilde{x}_i^u \ominus \tilde{y}_i^u), (\tilde{x}_i^l \ominus \tilde{y}_i^l)] \\
 &= [(x_i^{u1} - y_i^{u1}, x_i^{u2} - y_i^{u1}), x_i^c - y_i^c, (x_i^{l1} - y_i^{l1}, x_i^{l2} - y_i^{l1})] \quad (۲)
 \end{aligned}$$

^{۲۷} Perfectly normal triangular interval type-۲ fuzzy number

$$۳) \quad \tilde{x}_i = \begin{cases} [(wx_i^{u_1}, wx_i^{u_2}), wx_i^c, (wx_i^{l_1}, wx_i^{l_2})] & , w \geq 0 \\ [(wx_i^{u_2}, wx_i^{u_1}), wx_i^c, (wx_i^{l_2}, wx_i^{l_1})] & , w < 0 \end{cases} \quad (۳)$$

$$۴) \quad \frac{\tilde{x}_i}{w} = \begin{cases} [(\frac{x_i^{u_1}}{w}, \frac{x_i^{u_2}}{w}), \frac{x_i^c}{w}, (\frac{x_i^{l_1}}{w}, \frac{x_i^{l_2}}{w})] & , w > 0 \\ [(\frac{x_i^{u_2}}{w}, \frac{x_i^{u_1}}{w}), \frac{x_i^c}{w}, (\frac{x_i^{l_2}}{w}, \frac{x_i^{l_1}}{w})] & , w < 0 \end{cases}$$

$$۵) \quad \tilde{x}_i \otimes \tilde{y}_i = [(\dot{x}_i^u \otimes \dot{y}_i^u), (\dot{x}_i^l \otimes \dot{y}_i^l)] \\ = [(\dot{x}_i^{u_1} \times \dot{y}_i^{u_1}, \dot{x}_i^{u_2} \times \dot{y}_i^{u_2}), \dot{x}_i^c \times \dot{y}_i^c, (\dot{x}_i^{l_1} \times \dot{y}_i^{l_1}, \dot{x}_i^{l_2} \times \dot{y}_i^{l_2})] \quad \text{if } \dot{x}_i^{u_1}, \dot{y}_i^{u_1} \geq 0 \quad (۴)$$

۲-۴- فاصله فازی^{۲۸}

فاصله فازی در تعیین میزان تمایز بین مجموعه‌های فازی خصوصاً در موادی که نیاز به رتبه بندی و اولویت‌بندی اعداد فازی باشد نقش حیاتی دارد. تاکنون فرمول‌های مختلفی برای محاسبه فاصله فازی معرفی شده که فاصله فازی هاسدروفر^{۲۹}، همینگ^{۳۰} و اقلیدسی^{۳۱} جزو پرکاربردترین آنها بوده و در داده‌پردازی، استدلال تقریبی، تشخیص پزشکی، تجزیه و تحلیل محیط‌های تصمیم‌گیری، تحلیل رگرسیون و تحلیل ریسک در سطح وسیعی عملاً بکاربرده شده‌اند [۳۰ و ۳۱].

تعریف ۲-۷: فاصله فازی برای دو عدد فازی $\tilde{x}, \tilde{y} \in \mathcal{F}_2$ با نماد $d(\tilde{x}, \tilde{y})$ نشان داده می‌شود که $d: \mathcal{F}_2 \otimes \mathcal{F}_2 \rightarrow [0, \infty]$ تابع حقیقی d در صورتی متر نامیده می‌شود اگر برای هر $\tilde{x}, \tilde{y}, \tilde{z} \in \mathcal{F}_2$ اصول زیر برقرار باشد:

$$\begin{aligned} d(\tilde{x}, \tilde{y}) &\geq 0 & -۱ \\ d(\tilde{x}, \tilde{y}) &= d(\tilde{y}, \tilde{x}) & -۲ \\ d(\tilde{x}, \tilde{y}) + d(\tilde{y}, \tilde{z}) &\geq d(\tilde{x}, \tilde{z}) & -۳ \\ d(\tilde{x}, \tilde{x}) &= 0 & -۴ \end{aligned}$$

تعریف ۲-۸: فاصله همینگ (d_H) و فاصله اقلیدسی (d_E) بین دو عدد $\dot{x} = (x^l, x^c, x^u)$ و $\dot{y} = (y^l, y^c, y^2)$ به صورت زیر محاسبه می‌شود [۳۰]:

$$d_H(\dot{x}, \dot{y}) = \frac{1}{3} [|x^l - y^l| + |x^c - y^c| + |x^u - y^u|] \\ d_E(\dot{x}, \dot{y}) = \frac{1}{3} \sqrt{[(x^l - y^l)^2 + (x^c - y^c)^2 + (x^u - y^u)^2]}$$

اگر $\tilde{x} \in \mathcal{F}_2$ و $\tilde{y} \in \mathcal{F}_2$ باشند در این صورت: $d(\tilde{x}, \tilde{y}) = wd(\dot{x}^u, \dot{y}^u) + (1-w)d(\dot{x}^l, \dot{y}^l)$ که در آن $w \in (0, 1)$ بوده و توسط تصمیم‌گیر تعیین می‌شود. اگر $w \in (0, 0.5)$ در نظر گرفته شود به این معنی است که توابع عضویت پائینی با اهمیت‌تر از توابع عضویت بالایی است و $w \in (0.5, 1)$ به این معنی است که توابع عضویت بالایی با اهمیت‌تر از توابع عضویت پائینی است.

تعریف ۲-۹: فاصله همینگ (d_H) و اقلیدسی (d_E) بین دو مجموعه $\tilde{x}, \tilde{y} \in \mathcal{F}_2$ با $w = 0.5$ به صورت زیر تعریف می‌شود [۲۶ و ۳۰]:

$$d(\tilde{x}, \tilde{y}) = wd(\dot{x}^u, \dot{y}^u) + (1-w)d(\dot{x}^l, \dot{y}^l)$$

^{۲۸} Fuzzy distance measure^{۲۹} Hausdorff distance^{۳۰} Hamming distance^{۳۱} Euclidean distance

$$\begin{aligned}
 d_H(\tilde{x}', \tilde{y}') &= \frac{1}{3} [|x^h - y^h| + |x^c - y^c| + |x^{l^2} - y^{l^2}|] \\
 d_H(\tilde{x}, \tilde{y}) &= \frac{1}{2} [d_H(\tilde{x}^u, \tilde{y}^u) + d_H(\tilde{x}', \tilde{y}')] = \frac{1}{6} [\sum_{m=1}^2 (|x^{u_m} - y^{u_m}| + |x^{l_m} - y^{l_m}|) + 2|x^c - y^c|] \\
 d_E(\tilde{x}', \tilde{y}') &= \frac{1}{3} \sqrt{(x^h - y^h)^2 + (x^c - y^c)^2 + (x^{l^2} - y^{l^2})^2} \\
 d_E(\tilde{x}, \tilde{y}) &= \frac{1}{2} [d_E(\tilde{x}^u, \tilde{y}^u) + d_E(\tilde{x}', \tilde{y}')] \\
 &= \frac{1}{6} \{ \sqrt{(x^{u_1} - y^{u_1})^2 + (x^c - y^c)^2 + (x^{u_2} - y^{u_2})^2} + \sqrt{(x^h - y^h)^2 + (x^c - y^c)^2 + (x^{l^2} - y^{l^2})^2} \}
 \end{aligned} \tag{۵}$$

تعریف ۱۰-۲: اگر $\tilde{x} \in \mathcal{F}_2$ و $\tilde{0}$ مبدأ فازی باشد، فاصله علامت دار $\tilde{x}$ از $\tilde{0}$ به فرم زیر تعریف می‌شود [۳۰ و ۳۱]:

$$d_s(\tilde{x}, \tilde{0}) = \frac{1}{5} (x^{u_1} + x^h + x^c + x^{l^2} + x^{u_2})$$

و برای $\dot{x}^l = \dot{x}^u = \dot{x} = (x^l, x^c, x^u)$ فاصله از $\tilde{0}$ به صورت $d(\dot{x}, \tilde{0}) = \frac{1}{3} (x^l + x^c + x^u)$ تعریف می‌شود. ضمناً d_H یک تابع

$$d_H(w\tilde{x} \oplus \tilde{y}, \tilde{0}) = wd_H(\tilde{x}, \tilde{0}) \oplus d_H(\tilde{y}, \tilde{0}) \quad ; w \in R$$
 خطی است یعنی:

تعریف ۱۱-۲: برای $\tilde{x}, \tilde{y} \in \mathcal{F}_2$ ، رتبه براساس فاصله علامت‌دار به صورت زیر بیان می‌شود [۲۶ و ۳۰]:

$$\tilde{x} \leq \tilde{y} \quad \text{iff} \quad d_s(\tilde{x}, \tilde{0}) \leq d_s(\tilde{y}, \tilde{0}) \quad -1$$

$$\tilde{x} > \tilde{y} \quad \text{iff} \quad d_s(\tilde{x}, \tilde{0}) > d_s(\tilde{y}, \tilde{0}) \quad -2$$

$$\tilde{x} \approx \tilde{y} \quad \text{iff} \quad d_s(\tilde{x}, \tilde{0}) = d_s(\tilde{y}, \tilde{0}) \quad -3$$

۲-۵- متغیرهای آزاد در علامت در مسائل برنامه‌ریزی ریاضی

در فرمول‌بندی برنامه‌ریزی‌های ریاضی (LP, NLP و GP)، متغیرهای تصمیم نمی‌توانند مقادیر منفی بگیرند ولی اگر از نظر تصمیم‌گیرنده منفی شدن متغیر تصمیم امری مطلوب باشد باید بصورت متغیر آزاد در علامت معرفی و با دو متغیر کمکی غیر منفی جایگزین شود. یک متغیر آزاد در علامت f_i را می‌توان به شرح زیر تعریف نمود:

$$f_i^+ = \max\{0, f_i\} = \frac{f_i}{2} + \frac{|f_i|}{2}, \quad f_i^- = \max\{0, -f_i\} = \frac{-f_i}{2} + \frac{|-f_i|}{2}$$

$$f_i^+ + f_i^- = \left(\frac{f_i}{2} + \frac{|f_i|}{2}\right) + \left(\frac{-f_i}{2} + \frac{|f_i|}{2}\right) = |f_i|$$

$$f_i^+ - f_i^- = \left(\frac{f_i}{2} + \frac{|f_i|}{2}\right) - \left(\frac{-f_i}{2} + \frac{|f_i|}{2}\right) = f_i$$

در برنامه ریزی آرمانی به منظور برداشتن قدرمطلق می‌توان $|f_i|$ را با $f_i^+ + f_i^-$ و f_i را با $f_i^+ - f_i^-$ جایگزین

نموده و با افزودن قیدهای $f_i^+ \times f_i^- = 0$ ، $f_i^+ \geq 0$ و $f_i^- \geq 0$ مقدار f_i را بر حسب f_i^+ یا f_i^- به شرح زیر تعریف

نمود:

$$\text{if } f_i^+ = f_i^- = 0 \text{ then } f_i = 0,$$

$$\text{if } f_i^+ > 0 \text{ then } f_i = f_i^+,$$

$$\text{if } f_i^- > 0 \text{ then } f_i = -f_i^-.$$

تعریف ۱۲-۲: در یک مسئله برنامه ریاضی که پارامترها یا متغیرهای تصمیم و یا ترکیبی از هر دو فازی مثلثی

نوع-۲ بازه‌ای مقدار باشند، برنامه ریزی فازی نوع-۲ بازه‌ای مقدار می‌نامند [۲۵].

تعریف ۱۳-۲: فرم استاندارد یک مسئله برنامه ریزی خطی که تمام پارامترها و متغیرهای تصمیم آن فازی نوع-۲

بازه‌ای مقدار مثلثی باشند به صورت زیر تعریف می‌شود [۳۱-۳۶]:

$$\begin{aligned}
 (LP) \quad & \tilde{z} \approx \sum_{j=1}^m \tilde{c}_j \tilde{x}_j \\
 \text{s.t.} \quad & \sum_{j=1}^m \tilde{a}_{ij} x_j \approx \tilde{b}_i, \quad i=1, \dots, n, \\
 & \tilde{x}_j \geq 0, \quad j=1, \dots, m, \\
 & \tilde{c}_j, \tilde{a}_{ij}, \tilde{b}_i \in \mathcal{F}_2
 \end{aligned} \quad (V)$$

که $\tilde{x}_j$ عبارت است از j امین متغیر تصمیم (که به دلیل برخی محدودیت ممکن است بصورت $\tilde{x}_j \geq 0$ یا $\tilde{x}_j \leq 0$ و یا بصورت متغیر آزاد در علامت در نظر گرفته شوند که در این صورت می‌تواند مقادیر مثبت، منفی و یا صفر را انتخاب کنند)، $\tilde{Z}$ متغیر وابسته‌ای است که بوسیله متغیرهای مستقل $\tilde{x}_j$ طوری تعیین می‌شود که مقدار آن حداکثر یا حداقل شود. همچنین $\tilde{c}_j$ ، $\tilde{b}_i$ و $\tilde{a}_{ij}$ پارامترهای مدل بوده و مقادیر ثابت دارند که $\tilde{c}_j$ ضرایب هزینه یا سود، $\tilde{b}_i$ اعداد سمت راست و $\tilde{a}_{ij}$ ضرایب فنی نامیده می‌شوند.

تعریف ۲-۱۴: اگر $\tilde{S} = \{(\tilde{x}_1, \dots, \tilde{x}_m) \in (\mathcal{F}_2)^m : \sum_{j=1}^m \tilde{a}_{ij} \otimes \tilde{x}_j \approx \tilde{b}_i, \tilde{x}_j \geq 0, i=1, \dots, n, j=1, \dots, m\}$ مجموعه موجّه و $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_m) \in \tilde{S}$ نقطه موجّه مسئله LP باشند، نقطه $(\tilde{x}'_1, \dots, \tilde{x}'_m) \in (\mathcal{F}_2)^m$ جواب بهینه LP نامیده می‌شود اگر $\tilde{x}' = (\tilde{x}'_1, \dots, \tilde{x}'_p) \in \tilde{S}$ بوده و برای تمام $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_m) \in \tilde{S}$ داشته باشیم $\sum_{j=1}^m \tilde{c}_j \otimes \tilde{x}'_j \leq \sum_{j=1}^m \tilde{c}_j \otimes \tilde{x}_j$ [۳۸-۳۱].

تبصره ۲-۱: با توجه به اینکه $\sum_{j=1}^m \tilde{a}_{ij} \otimes \tilde{x}_j \approx \tilde{b}_i$ (یا همان

$$\sum_{j=1}^m a_{ij}^t x_j^t = b_i^t, t \in \{l_1, l_2, c, u_1, u_2\} \text{ را می‌توان به فرم } (\sum_{j=1}^m [(a_{ij}^{l_1} x_j^{l_1}, a_{ij}^{l_2} x_j^{l_2}), a_{ij}^c x_j^c, (a_{ij}^{u_1} x_j^{u_1}, a_{ij}^{u_2} x_j^{u_2})]) \approx [(b_i^{l_1}, b_i^{l_2}), b_i^c, (b_i^{u_1}, b_i^{u_2})])$$

نوشت، از اینرو مسأله (۷) را می‌توان به صورت زیر ارائه نمود:

$$\begin{aligned}
 (V-1) \quad & \text{Min} \quad \tilde{z} \approx \sum_{j=1}^m \tilde{c}_j \otimes \tilde{x}_j \\
 \text{s.t.} \quad & \sum_{j=1}^m [(a_{ij} x_j)^t = b_i^t, t \in \{l_1, l_2, c, u_1, u_2\}, \\
 & x_j^t \geq 0, \quad j=1, \dots, m; t \in \{l_1, l_2, c, u_1, u_2\}.
 \end{aligned}$$

قضیه ۲-۱: مسأله برنامه‌ریزی زیر را می‌توان به برنامه‌ریزی خطی (LP) تبدیل نمود.

$$\begin{aligned}
 \text{Min} \quad & d_s(\tilde{z}, \tilde{0}) = \sum_{j=1}^m d_s(\tilde{c}_j \otimes \tilde{x}_j, \tilde{0}) \\
 \text{s.t.} \quad & \sum_{j=1}^m [(a_{ij} x_j)^t = b_i^t, t \in \{u_1, l_1, c, l_2, u_2\}, \quad i=1, \dots, n, \quad (V-2) \\
 & d_s(\tilde{x}_j, \tilde{0}) \geq 0, \quad j=1, \dots, m, \\
 & \tilde{c}_j, \tilde{a}_{ij}, \tilde{b}_i \in \mathcal{F}_2.
 \end{aligned}$$

اثبات: فرض می‌کنیم که $\tilde{c}_j, \tilde{x}_j, \tilde{a}_{ij}, \tilde{b}_i \in \mathcal{F}_2$ و $\tilde{x}_j = [(x_j^{l_1}, x_j^{l_2}), x_j^c, (x_j^{u_1}, x_j^{u_2})] \in \mathcal{F}_2^+$

از تبصره ۲-۱ می‌دانیم که می‌توان $\sum_{j=1}^m \tilde{a}_{ij} \otimes \tilde{x}_j = \sum_{j=1}^m [(a_{ij}^{l_1} x_j^{l_1}, a_{ij}^{l_2} x_j^{l_2}), a_{ij}^c x_j^c, (a_{ij}^{u_1} x_j^{u_1}, a_{ij}^{u_2} x_j^{u_2})]$ را به صورت

$$\sum_{j=1}^m \tilde{c}_j \otimes \tilde{x}_j = \sum_{j=1}^m [(c_j^{l_1} x_j^{l_1}, c_j^{l_2} x_j^{l_2}), c_j^c x_j^c, (c_j^{u_1} x_j^{u_1}, c_j^{u_2} x_j^{u_2})]$$

نوشت در صورتی که $\sum_{j=1}^m a_{ij}^t x_j^t = b_i^t, t \in \{l_1, l_2, c, u_1, u_2\}$ و $\sum_{j=1}^m a_{ij}^t x_j^t = b_i^t, t \in \{l_1, l_2, c, u_1, u_2\}$ توابعی خطی بوده و $(x_j^{l_1}, x_j^{l_2}, x_j^c, x_j^{u_1}, x_j^{u_2}) \in (\mathbb{R}^+)^5, j=1, \dots, m$ حال از تعریف ۱۰-۲

و ۱۱-۲ داریم:

$$d_s(\sum_{j=1}^m \tilde{c}_j \otimes \tilde{x}_j, \tilde{0}) = \sum_{j=1}^m d_s(\tilde{c}_j \otimes \tilde{x}_j, \tilde{0}) = f(x) = \frac{1}{5} \sum_{j=1}^m (c_j^{u_1} x_j^{u_1} + c_j^{l_1} x_j^{l_1} + c_j^c x_j^c + c_j^{l_2} x_j^{l_2} + c_j^{u_2} x_j^{u_2})$$

و $d_s(\tilde{x}_j, \tilde{0}) \geq 0, j=1, \dots, m$ بنابراین، می‌توان مدل (۷-۳) را به فرم زیر تبدیل نمود:

$$\min f(x) = \frac{1}{5} \sum_{j=1}^m (c_j^{u_1} x_j^{u_1} + c_j^{l_1} x_j^{l_1} + c_j^c x_j^c + c_j^{l_2} x_j^{l_2} + c_j^{u_2} x_j^{u_2}) \quad (۷-۳)$$

s.t

$$\sum_{j=1}^m a_{ij}' x_j^t = b_i', t \in \{l_1, l_2, c, u_1, u_2\},$$

$$x_j^{u_1} \leq x_j^l \leq x_j^c \leq x_j^{l_2} \leq x_j^{u_2}$$

$$x_j^{u_1} \geq 0, j=1, \dots, m$$

واضح است که مسئله (۷-۳) برنامه‌ریزی خطی با متغیرهای تصمیم $x_j = (x_j^{u_1}, x_j^{l_1}, x_j^c, x_j^{l_2}, x_j^{u_2}) \in (R^+)^5, j=1, \dots, m$ می‌باشد.

قضیه ۲-۲: مسئله برنامه‌ریزی (۷-۳) و (۷) معادل هم هستند.

اثبات: فرض کنید $\tilde{x}_j = [(x_j^{l_1}, x_j^{l_2}), x_j^c, (x_j^{u_1}, x_j^{u_2})] \in \mathcal{F}_2^+, \tilde{c}_j, \tilde{a}_{ij}, \tilde{b}_i \in \mathcal{F}_2, \tilde{x}_j = [(x_j^{l_1}, x_j^{l_2}), x_j^c, (x_j^{u_1}, x_j^{u_2})] \in \mathcal{F}_2^+$ و

$$x_j = (x_j^{u_1}, x_j^{l_1}, x_j^c, x_j^{l_2}, x_j^{u_2}) \in (R_0^+)^5, j=1, \dots, m$$

$$\text{و } \tilde{S} = \{ \tilde{x} \in (\mathcal{F}_2^+)^m : \sum_{j=1}^m \tilde{a}_{ij} \otimes \tilde{x}_j \approx \tilde{b}_i, \tilde{x}_j \geq \tilde{0}, i=1, \dots, n, j=1, \dots, m \}$$

$$S = \{ x_j \in (R_0^+)^{5m} : \sum_{j=1}^m a_{ij}' x_j^t = b_i', t \in \{u_1, l_1, c, l_2, u_2\}, x_j^t \geq 0, i=1, \dots, n, j=1, \dots, m \}$$

شدنی مسئله (۷-۳) و (۷) باشند، در این صورت

$$(\tilde{x}_1, \dots, \tilde{x}_m) \in \tilde{S} \Leftrightarrow \sum_{j=1}^m \tilde{a}_{ij} \otimes \tilde{x}_j = \tilde{b}_i, i=1, \dots, n$$

$$\Leftrightarrow \sum_{j=1}^m a_{ij}' x_j^t = b_i', t \in \{l_1, l_2, c, u_1, u_2\}$$

$$\Leftrightarrow x_j = (x_j^{u_1}, x_j^{l_1}, x_j^c, x_j^{l_2}, x_j^{u_2}) \in S$$

حالا فرض می‌کنیم $\tilde{x}^* = (\tilde{x}_1^*, \dots, \tilde{x}_m^*) \in \tilde{S}$ یک جواب بهینه مسئله (۷) باشد که در این صورت برای کلیه برای تمام $\tilde{x}_j$ ها در $\tilde{S}$ خواهیم داشت $\sum_{j=1}^m \tilde{c}_j \otimes \tilde{x}_j \geq \sum_{j=1}^m \tilde{c}_j \otimes \tilde{x}_j^*$. حال با توجه به تعریف ۱۱-۲ و ۱۲-۲ داریم:

$$d_s(\sum_{j=1}^m \tilde{c}_j \otimes \tilde{x}_j, \tilde{0}) \geq d_s(\sum_{j=1}^m \tilde{c}_j \otimes \tilde{x}_j^*, \tilde{0}) \Leftrightarrow \frac{1}{5} \sum_{j=1}^m [\sum_{k=1}^2 (c_j^{l_k} x_j^{l_k} + c_j^{u_k} x_j^{u_k}) + c_j^c x_j^c] \geq \frac{1}{5} \sum_{j=1}^m [\sum_{k=1}^2 (c_j^{l_k} x_j^{l_k*} + c_j^{u_k} x_j^{u_k*}) + c_j^c x_j^{c*}]$$

روابط بالا به این معنی است که $x_j^* = (x_j^{u_1*}, x_j^{l_1*}, x_j^c, x_j^{l_2*}, x_j^{u_2*}) \in (R_0^+)^5, j=1, \dots, m$ یک جواب بهینه مسئله (۷-۳) می‌باشد. به همین طریق می‌توان ثابت کرد که هر جواب بهینه‌ای از (۷-۳) جواب بهینه برای (۷) نیز می‌باشد. بنابراین نتیجه می‌گیریم که مسئله (۷-۳) و (۷) معادلند.

تعریف ۱۵-۲: فرم کلی و استاندارد مسئله GP که در آن تمام متغیرهای تصمیم (متغیرهای مستقل)، متغیر وابسته و پارامترها، فازی نوع-۲ بازه‌ای مقدار باشند به فرم زیر بیان می‌شود [۳۱-۳۸]

$$(GP) \quad \min \tilde{D} = \sum_{i=1}^n |f_i(\tilde{x}_j) - \tilde{y}_i|$$

$$\text{s.t. } \tilde{x}_j \in \tilde{F}$$

۳- مدل پیشنهادی با متغیرها و پارامترهای فازی مثلثی نوع-۲ بازه‌ای مقدار

اگر $\tilde{x}_j, \tilde{y}_i, \tilde{b}_i \in \mathcal{F}_2, i=1, \dots, n, j=0, \dots, m$ ، مدل رگرسیونی فازی مثلثی نوع-۲ بازه‌ای مقدار را می‌توان به صورت زیر بیان نمود ($\tilde{x}_{0i} \approx \tilde{1}$):

$$\tilde{y}_i \approx \bigoplus_{j=0}^m \tilde{b}_j \otimes \tilde{x}_{ji} \approx \tilde{b}_0 \oplus \tilde{b}_1 \otimes \tilde{x}_{i1} \oplus \dots \oplus \tilde{b}_m \otimes \tilde{x}_{mi}; i=1, \dots, n. \quad (۸)$$

در مدل (۸)، $\tilde{y}_i \approx [(y_i^{u1}, y_i^{u2}), y_i^c, (y_i^{l1}, y_i^{l2})]$ ، متغیر پاسخ مشاهده شده در نمونه i ام، $\tilde{b}_j = [(b_j^{u1}, b_j^{u2}), b_j^c, (b_j^{l1}, b_j^{l2})]$ ، ضریب رگرسیونی و $\tilde{x}_{ji} \approx [(x_{ji}^{u1}, x_{ji}^{u2}), x_{ji}^c, (x_{ji}^{l1}, x_{ji}^{l2})]$ ، ضریب متغیر مستقل در نمونه i ام، بدون کاستن از کلیت، فرض می‌کنیم تمام $\tilde{x}_{ji}$ مقادیر مثبت باشند (در صورتی که $\tilde{x}_{ji} < 0$ باشد می‌توان با یک تبدیل ساده به حالت $\tilde{x}_{ji} > 0$ تغییر داد). فرم کلی یک مدل رگرسیونی فازی نوع-۲ بازه‌ای مقدار برآورد شده را می‌توان به فرم زیر ارائه نمود:

$$\hat{y}_i \approx \bigoplus_{j=0}^m \hat{b}_j \otimes \tilde{x}_{ji} \approx \hat{b}_0 \oplus \hat{b}_1 \otimes \tilde{x}_{i1} \oplus \dots \oplus \hat{b}_m \otimes \tilde{x}_{mi}; i=1, \dots, n. \quad (۹-۱)$$

در مدل (۹-۱)، $\hat{y}_i \approx [(\hat{y}_i^{u1}, \hat{y}_i^{u2}), \hat{y}_i^c, (\hat{y}_i^{l1}, \hat{y}_i^{l2})]$ ، متغیر پاسخ برآورد شده در نمونه i ام و $\hat{b}_j = [(\hat{b}_j^{u1}, \hat{b}_j^{u2}), \hat{b}_j^c, (\hat{b}_j^{l1}, \hat{b}_j^{l2})]$ ، ضریب رگرسیونی است.

۳-۱- برآورد ضرایب رگرسیون فازی نوع-۲ بازه‌ای مقدار

پارامترهای مدل پیشنهادی به شرح مراحل زیر برآورد شده‌اند:

مرحله اول: با توجه به اینکه $\hat{b}_j, j=0, \dots, m$ آزاد در علامت هستند لذا آنرا را به صورت $\hat{b}_j \approx \hat{b}_j^+ \oplus \hat{b}_j^-$ تعریف می‌کنیم که در آن $\hat{b}_j^+ \approx \max(0, \hat{b}_j) \approx \frac{|\hat{b}_j|}{2} \oplus \frac{\hat{b}_j}{2}$ (متغیر انحراف مثبت) و $\hat{b}_j^- \approx \max(0, -\hat{b}_j) \approx \frac{|\hat{b}_j|}{2} \oplus \frac{-\hat{b}_j}{2}$ (متغیر انحراف منفی) می‌باشد. از این رو مقدار $\hat{b}_j$ با دستور شرطی زیر تعیین می‌گردد:

$$\text{if } \hat{b}_j^{u1} > 0 \text{ then } \hat{b}_j \approx \hat{b}_j^+ \text{ else if } \hat{b}_j^{l1} > 0 \text{ then } \hat{b}_j \approx \hat{b}_j^- \text{ otherwise } \hat{b}_j \approx 0$$

حال با توجه به اینکه $\hat{b}_j^+ = [(\hat{b}_j^{l1+}, \hat{b}_j^{l2+}), \hat{b}_j^{c+}, (\hat{b}_j^{u1+}, \hat{b}_j^{u2+})]$ و $\hat{b}_j^- = [(\hat{b}_j^{l1-}, \hat{b}_j^{l2-}), \hat{b}_j^{c-}, (\hat{b}_j^{u1-}, \hat{b}_j^{u2-})]$ ، با استفاده از روابط (۱) و (۲) خواهیم داشت:

$$\hat{b}_j \approx \hat{b}_j^+ \oplus \hat{b}_j^- \approx [(\hat{b}_j^{u1+} - \hat{b}_j^{u2-}, \hat{b}_j^{u2+} - \hat{b}_j^{u1-}), \hat{b}_j^{c+} - \hat{b}_j^{c-}, (\hat{b}_j^{l1+} - \hat{b}_j^{l2-}, \hat{b}_j^{l2+} - \hat{b}_j^{l1-})] \quad (۹)$$

با استفاده از (۴) و جایگذاری از (۹) حاصل $\hat{b}_j \otimes \tilde{x}_{ji}$ بدین شرح بدست می‌آید:

$$\hat{b}_j \otimes \tilde{x}_{ji} \approx (\hat{b}_j^+ \oplus \hat{b}_j^-) \otimes \tilde{x}_{ji} \quad (۱۰)$$

$$\approx [(\hat{b}_j^{u1+} x_{ji}^{u1} - \hat{b}_j^{u2-} x_{ji}^{u1}, \hat{b}_j^{u2+} x_{ji}^{u2} - \hat{b}_j^{u1-} x_{ji}^{u2}), \hat{b}_j^{c+} x_{ji}^c - \hat{b}_j^{c-} x_{ji}^c, (\hat{b}_j^{l1+} x_{ji}^{l1} - \hat{b}_j^{l2-} x_{ji}^{l1}, \hat{b}_j^{l2+} x_{ji}^{l2} - \hat{b}_j^{l1-} x_{ji}^{l2})]$$

$$\hat{b}_0 \oplus \hat{b}_j \otimes \tilde{x}_{ji} \approx \hat{b}_0 \oplus (\hat{b}_j^+ \oplus \hat{b}_j^-) \otimes \tilde{x}_{ji} \quad (۱۱)$$

$$\approx [(\hat{b}_0^{u1} + \hat{b}_j^{u1+} x_{ji}^{u1} - \hat{b}_j^{u2-} x_{ji}^{u1}, \hat{b}_0^{u2} + \hat{b}_j^{u2+} x_{ji}^{u2} - \hat{b}_j^{u1-} x_{ji}^{u2}), \hat{b}_0^c + \hat{b}_j^{c+} x_{ji}^c - \hat{b}_j^{c-} x_{ji}^c, (\hat{b}_0^{l1} + \hat{b}_j^{l1+} x_{ji}^{l1} - \hat{b}_j^{l2-} x_{ji}^{l1}, \hat{b}_0^{l2} + \hat{b}_j^{l2+} x_{ji}^{l2} - \hat{b}_j^{l1-} x_{ji}^{l2})]$$

با استفاده از (۱۱) داریم:

$$\hat{y}_i^{u1} = \hat{b}_0^{u1} + \sum_{j=1}^m (\hat{b}_j^{u1+} x_{ji}^{u1} - \hat{b}_j^{u2-} x_{ji}^{u1}) - y_i^{u1} \quad (۱۲)$$

$$\hat{y}_i^{u2} = \hat{b}_0^{u2} + \sum_{j=1}^m (\hat{b}_j^{u2+} x_{ji}^{u2} - \hat{b}_j^{u1-} x_{ji}^{u2}), \quad (۱۳)$$

$$\hat{y}_i^c = \hat{b}_0^c + \sum_{j=1}^m (\hat{b}_j^{c+} x_{ji}^c - \hat{b}_j^{c-} x_{ji}^c), \quad (۱۴)$$

$$\hat{y}_i^{l1} = \hat{b}_0^{l1} + \sum_{j=1}^m (\hat{b}_j^{l1+} x_{ji}^{l1} - \hat{b}_j^{l2-} x_{ji}^{l1}), \quad (۱۵)$$

$$\hat{y}_i^{l2} = \hat{b}_0^{l2} + \sum_{j=1}^m (\hat{b}_j^{l2+} x_{ji}^{l2} - \hat{b}_j^{l1-} x_{ji}^{l2}), \quad (۱۶)$$

با استفاده از (۹) و جای گذاری (۱۲) تا (۱۶) در $[\hat{Y}_i^{u_1}, \hat{Y}_i^{u_2}], \hat{Y}_i^c, (\hat{Y}_i^{l_1}, \hat{Y}_i^{l_2})$ خواهیم داشت:

$$\hat{Y}_i \approx \{[(\hat{b}_0^{u_1} + \sum_{j=1}^m (\hat{b}_j^{u_1+} x_{ji}^{u_1} - \hat{b}_j^{u_1-} x_{ji}^{u_1})), \hat{b}_0^{u_2} + \sum_{j=1}^m (\hat{b}_j^{u_2+} x_{ji}^{u_2} - \hat{b}_j^{u_2-} x_{ji}^{u_2}), \hat{b}_0^c + \sum_{j=1}^m (\hat{b}_j^{c+} x_{ji}^c - \hat{b}_j^{c-} x_{ji}^c), \\ , [(\hat{b}_0^{l_1} + \sum_{j=1}^m (\hat{b}_j^{l_1+} x_{ji}^{l_1} - \hat{b}_j^{l_1-} x_{ji}^{l_1})), \hat{b}_0^{l_2} + \sum_{j=1}^m (\hat{b}_j^{l_2+} x_{ji}^{l_2} - \hat{b}_j^{l_2-} x_{ji}^{l_2})]]\} \quad (17)$$

مرحله دوم: برای محاسبه اختلاف بین مقدار برآورد شده و مشاهده شده متغیر پاسخ، فرمول فاصله فازی تعریف ۵ را استفاده می‌کنیم:

$$d_H(\hat{Y}_i, \tilde{Y}_i) = \frac{1}{6} [| \hat{Y}_i^{u_1} - Y_i^{u_1} | + | \hat{Y}_i^{u_2} - Y_i^{u_2} | + 2 | \hat{Y}_i^c - Y_i^c | + | \hat{Y}_i^{l_1} - Y_i^{l_1} | + | \hat{Y}_i^{l_2} - Y_i^{l_2} |] \\ = \frac{1}{6} [| \hat{b}_0^{u_1} + \sum_{j=1}^m (\hat{b}_j^{u_1+} x_{ji}^{u_1} - \hat{b}_j^{u_1-} x_{ji}^{u_1}) - Y_i^{u_1} | + | \hat{b}_0^{u_2} + \sum_{j=1}^m (\hat{b}_j^{u_2+} x_{ji}^{u_2} - \hat{b}_j^{u_2-} x_{ji}^{u_2}) - Y_i^{u_2} | + 2 | \hat{b}_0^c + \sum_{j=1}^m (\hat{b}_j^{c+} x_{ji}^c - \hat{b}_j^{c-} x_{ji}^c) - Y_i^c | \\ + | \hat{b}_0^{l_1} + \sum_{j=1}^m (\hat{b}_j^{l_1+} x_{ji}^{l_1} - \hat{b}_j^{l_1-} x_{ji}^{l_1}) - Y_i^{l_1} | + | \hat{b}_0^{l_2} + \sum_{j=1}^m (\hat{b}_j^{l_2+} x_{ji}^{l_2} - \hat{b}_j^{l_2-} x_{ji}^{l_2}) - Y_i^{l_2} |] \quad (18)$$

مرحله سوم: از فرمول (۱۸) و رویکرد برنامه‌ریزی آرمانی فازی نوع-۲ بازه‌ای مقدار به شرح مراحل زیر استفاده می‌کنیم:

$$\min_{i=1}^n d_H(\hat{Y}_i, \tilde{Y}_i) = \\ = \frac{1}{6} \min \sum_{i=1}^n [| \hat{b}_0^{u_1} + \sum_{j=1}^m (\hat{b}_j^{u_1+} x_{ji}^{u_1} - \hat{b}_j^{u_1-} x_{ji}^{u_1}) - Y_i^{u_1} | + | \hat{b}_0^{u_2} + \sum_{j=1}^m (\hat{b}_j^{u_2+} x_{ji}^{u_2} - \hat{b}_j^{u_2-} x_{ji}^{u_2}) - Y_i^{u_2} | + 2 | \hat{b}_0^c + \sum_{j=1}^m (\hat{b}_j^{c+} x_{ji}^c - \hat{b}_j^{c-} x_{ji}^c) - Y_i^c | \\ + | \hat{b}_0^{l_1} + \sum_{j=1}^m (\hat{b}_j^{l_1+} x_{ji}^{l_1} - \hat{b}_j^{l_1-} x_{ji}^{l_1}) - Y_i^{l_1} | + | \hat{b}_0^{l_2} + \sum_{j=1}^m (\hat{b}_j^{l_2+} x_{ji}^{l_2} - \hat{b}_j^{l_2-} x_{ji}^{l_2}) - Y_i^{l_2} |] \quad (19)$$

S.t.

$$\hat{b}_j^{u_1} \leq \hat{b}_j^{l_1} \leq \hat{b}_j^c \leq \hat{b}_j^{l_2} \leq \hat{b}_j^{u_2}, \quad j=1, \dots, m,$$

$$Y_i^{u_1} \leq Y_i^{l_1} \leq Y_i^c \leq Y_i^{l_2} \leq Y_i^{u_2}, \quad i=1, \dots, n.$$

$$0 \leq x_{ji}^{u_1} \leq x_{ji}^{l_1} \leq x_{ji}^c \leq x_{ji}^{l_2} \leq x_{ji}^{u_2}, \quad i=1, \dots, n, \quad j=1, \dots, p$$

قدم ۴: با استفاده از متغیر انحراف مثبت و منفی یعنی $e_i^{k+} = \max(0, \hat{Y}_i^k - Y_i^k) \geq 0$ و $e_i^{k-} = \max(0, Y_i^k - \hat{Y}_i^k) \geq 0$ و با

تبدیل $|\hat{Y}_i^k - Y_i^k|$ به $e_i^{k+} + e_i^{k-}$ به ازاء $k \in \{u_1, l_1, c, l_2, u_2\}; i=1, \dots, n$ ، قدر مطلق در (۱۹) به شرح زیر برداشته می‌شود:

$$e_i^{k+} = \max(0, \hat{Y}_i^k - Y_i^k) = \frac{|\hat{Y}_i^k - Y_i^k| + \hat{Y}_i^k - Y_i^k}{2}, \quad e_i^{k-} = \max(0, Y_i^k - \hat{Y}_i^k) = \frac{|Y_i^k - \hat{Y}_i^k| + Y_i^k - \hat{Y}_i^k}{2} \\ e_i^{k+} + e_i^{k-} = \frac{|\hat{Y}_i^k - Y_i^k| + \hat{Y}_i^k - Y_i^k}{2} + \frac{|Y_i^k - \hat{Y}_i^k| + Y_i^k - \hat{Y}_i^k}{2} = |Y_i^k - \hat{Y}_i^k| \\ e_i^{k+} + e_i^{k-} = |\hat{Y}_i^k - Y_i^k|, \quad (20)$$

$$e_i^{k+} - e_i^{k-} = \frac{|\hat{Y}_i^k - Y_i^k| + \hat{Y}_i^k - Y_i^k}{2} - \frac{|Y_i^k - \hat{Y}_i^k| + Y_i^k - \hat{Y}_i^k}{2} = \hat{Y}_i^k - Y_i^k \\ e_i^{k+} - e_i^{k-} = \hat{Y}_i^k - Y_i^k \Rightarrow \hat{Y}_i^k - e_i^{k+} + e_i^{k-} = Y_i^k, \quad (21)$$

$$k \in \{u_1, l_1, c, l_2, u_2\}; i=1, \dots, n.$$

با مینیمم کردن e_i^{k+} یا e_i^{k-} یا $e_i^{k+} + e_i^{k-}$ حاصل $|\hat{Y}_i^k - \tilde{Y}_i^k|$ مینیمم می‌شود. لذا اگر $\hat{Y}_i^k \geq Y_i^k$ مطلوب ما باشد مقدار e_i^{k-} را مینیمم می‌کنیم و زمانی که هدف حصول به $\hat{Y}_i^k \geq Y_i^k$ باشد e_i^{k+} را مینیمم می‌کنیم و اگر $Y_i^k \geq \hat{Y}_i^k$ حالت مطلوب باشد باید $e_i^{k+} + e_i^{k-}$ حداقل شود. در این خصوص با استفاده از (۲۰) و (۲۱) خواهیم داشت:

$$|\hat{Y}_i^k - Y_i^k| = e_i^{k+} + e_i^{k-} \quad (22)$$

$$\hat{Y}_i^k + e_i^{k-} - e_i^{k+} = Y_i^k \quad (23)$$

$$k \in \{u_1, l_1, c, l_2, u_2\}; i=1, \dots, n.$$

$$d_H(\hat{Y}_i, \tilde{Y}_i) = \frac{1}{6} [| \hat{Y}_i^{u_1} - y_i^{u_1} | + | \hat{Y}_i^{u_2} - y_i^{u_2} | + 2 | \hat{Y}_i^c - y_i^c | + | \hat{Y}_i^{l_1} - y_i^{l_1} | + | \hat{Y}_i^{l_2} - y_i^{l_2} |]$$

$$d_H(\hat{Y}_i, \tilde{Y}_i) = \frac{1}{6} \sum_{k=1}^2 [e_i^{u_k+} + e_i^{u_k-} + e_i^{l_k+} + e_i^{l_k-}] + \frac{1}{3} (e_i^{c+} + e_i^{c-}) \quad (24)$$

با جای گذاری (۲۲) و (۲۳) در (۱۸)، مسأله (۱۹) به فرم زیر تبدیل می‌شود:

$$\min \sum_{i=1}^n d_H(\hat{Y}_i, \tilde{Y}_i) = \frac{1}{6} \min \sum_{i=1}^n \{ \sum_{k=1}^2 [e_i^{u_k+} + e_i^{u_k-} + e_i^{l_k+} + e_i^{l_k-}] + 2(e_i^{c+} + e_i^{c-}) \} \quad (25)$$

s.t.

$$\hat{b}_0^{u_1} + [\sum_{j=1}^m (\hat{b}_j^{u_1+} \times x_{ji}^{u_1} - \hat{b}_j^{u_1-} \times x_{ji}^{u_1})] + e_i^{u_1-} - e_i^{u_1+} = y_i^{u_1}$$

$$\hat{b}_0^{u_2} + [\sum_{j=1}^m (\hat{b}_j^{u_2+} \times x_{ji}^{u_2} - \hat{b}_j^{u_2-} \times x_{ji}^{u_2})] + e_i^{u_2-} - e_i^{u_2+} = y_i^{u_2}$$

$$\hat{b}_0^c + [\sum_{j=1}^m (\hat{b}_j^{c+} \times x_{ji}^c - \hat{b}_j^{c-} \times x_{ji}^c)] + e_i^{c-} - e_i^{c+} = y_i^c$$

$$\hat{b}_0^{l_1} + [\sum_{j=1}^m (\hat{b}_j^{l_1+} \times x_{ji}^{l_1} - \hat{b}_j^{l_1-} \times x_{ji}^{l_1})] + e_i^{l_1-} - e_i^{l_1+} = y_i^{l_1}$$

$$\hat{b}_0^{l_2} + [\sum_{j=1}^m (\hat{b}_j^{l_2+} \times x_{ji}^{l_2} - \hat{b}_j^{l_2-} \times x_{ji}^{l_2})] + e_i^{l_2-} - e_i^{l_2+} = y_i^{l_2}$$

$$e_i^{k+} \times e_i^{k-} = 0, \quad (1-25)$$

$$e_i^{k+} \geq 0, e_i^{k-} \geq 0,$$

$$\hat{b}_0^k \in R, k \in \{u_1, l_1, c, l_2, u_2\},$$

$$\hat{b}_0^{u_1} \leq \hat{b}_0^{l_1} \leq \hat{b}_0^c \leq \hat{b}_0^{l_2} \leq \hat{b}_0^{u_2}$$

$$0 \leq \hat{b}_j^{u_1+} \leq \hat{b}_j^{l_1+} \leq \hat{b}_j^{c+} \leq \hat{b}_j^{l_2+} \leq \hat{b}_j^{u_2+},$$

$$0 \leq \hat{b}_j^{u_1-} \leq \hat{b}_j^{l_1-} \leq \hat{b}_j^{c-} \leq \hat{b}_j^{l_2-} \leq \hat{b}_j^{u_2-},$$

$$0 \leq x_{ji}^{u_1} \leq x_{ji}^{l_1} \leq x_{ji}^c \leq x_{ji}^{l_2} \leq x_{ji}^{u_2},$$

$$y_i^{u_1} \leq y_i^{l_1} \leq y_i^c \leq y_i^{l_2} \leq y_i^{u_2},$$

$$j = 1, \dots, m; i = 1, \dots, n.$$

در قضیه (۱-۳) ثابت می‌کنیم که در مسأله (۲۵) می‌توان قید (۱-۲۵) را حذف نمود.

قضیه ۳-۱: در مسأله (۲۵) به ازاء $m \in \{u_1, l_1, c, l_2, u_2\}, i = 1, \dots, n$ e_i^{m+} و e_i^{m-} نمی‌توانند هم زمان رخ دهند و

می‌توان قید $e_i^{m+} \times e_i^{m-} = 0$ را حذف نمود.

اثبات: در نظر بگیرید که $e_i^{\hat{m}-}$ و $e_i^{\hat{m}+}$, $m \in \{u_1, l_1, c, l_2, u_2\}, i = 1, \dots, n$ جواب بهینه پارتو مسأله (۲۵) باشند، یعنی

جواب دیگری مثل $e_i^{\bar{m}+}$ و $e_i^{\bar{m}-}$ وجود نداشته باشد که $e_i^{\bar{m}+} + e_i^{\bar{m}-} \leq e_i^{\hat{m}+} + e_i^{\hat{m}-}$ یا (حداقل برای یکی)

$e_i^{\bar{m}+} + e_i^{\bar{m}-} < e_i^{\hat{m}+} + e_i^{\hat{m}-}$ حال فرض کنید که $e_i^{\hat{m}+} \times e_i^{\hat{m}-} \neq 0$ بوده و برای $e_0^{m+} \geq 0$ و $e_0^{m-} \geq 0$ داشته باشیم

$e_0^{m+} \times e_0^{m-} = 0$ و $e_i^{\hat{m}-} - e_i^{\hat{m}+} = e_0^{m-} - e_0^{m+}$. در این صورت e_0^{m-} و e_0^{m+} در قید (۱-۲۵) صدق نموده و برای

$e_0^{m+} \times e_0^{m-} = 0$ دو وضعیت زیر پیش می‌آید:

الف- $e_0^{m+} = 0$ و $e_0^{m-} > 0$ باشد که در این صورت $e_i^{\hat{m}-} - e_i^{\hat{m}+} = e_0^{m-}$ و چون فرض کردیم $e_i^{\hat{m}+} \times e_i^{\hat{m}-} \neq 0$ پس

$$e_i^{\hat{m}+} > 0 \text{ و } e_i^{\hat{m}-} > e_i^{\hat{m}+} \text{ در نتیجه } e_i^{\hat{m}-} > e_i^{\hat{m}+}$$

ب- $e_0^{m-} = 0$ و $e_0^{m+} > 0$ باشد که در این صورت $e_i^{\hat{m}-} - e_i^{\hat{m}+} = -e_0^{m+}$ بوده و از $e_i^{\hat{m}+} > 0$ نتیجه می‌گیریم

که $e_i^{\hat{m}+} > e_0^{m+}$ است. حال از $e_i^{\hat{m}-} > e_i^{\hat{m}+}$ و $e_i^{\hat{m}+} > e_0^{m+}$ نتیجه می‌گیریم $e_i^{\hat{m}-} > e_0^{m+} + e_0^{m-}$ و این با جواب

بهینه پارتو بودن $e_i^{\hat{m}+}$ و $e_i^{\hat{m}-}$ در تناقض است، پس $e_i^{\hat{m}+}$ و $e_i^{\hat{m}-}$ نمی توانند هم زمان رخ دهند، لذا $e_i^{\hat{m}+} \times e_i^{\hat{m}-} = 0$ بوده و به علت اینکه از طریق سایر قیدها برآورد می شوند می توان آنرا حذف نمود.
با حذف قید (۲۵-۱)، مسأله (۲۵) به صورت زیر در می آید:

$$\min \sum_{i=1}^n d_h(\hat{y}_i, \tilde{y}_i) = \frac{1}{6} \min \sum_{i=1}^m [e_i^{u_1+} + e_i^{u_1-} + e_i^{u_2+} + e_i^{u_2-} + 2e_i^{c+} + 2e_i^{c-} + e_i^{l_1+} + e_i^{l_1-} + e_i^{l_2+} + e_i^{l_2-}] \quad (۲۶)$$

s.t.

$$\hat{b}_0^{u_1} + [\sum_{j=1}^p (\hat{b}_j^{u_1+} x_{ji}^{u_1} - \hat{b}_j^{u_2-} x_{ji}^{u_1})] + e_i^{u_1-} - e_i^{u_1+} = y_i^{u_1} \quad (۱-۲۶)$$

$$\hat{b}_0^{u_2} + [\sum_{j=1}^p (\hat{b}_j^{u_2+} x_{ji}^{u_2} - \hat{b}_j^{u_1-} x_{ji}^{u_2})] + e_i^{u_2-} - e_i^{u_2+} = y_i^{u_2} \quad (۲-۲۶)$$

$$\hat{b}_0^c + [\sum_{j=1}^p (\hat{b}_j^{c+} x_{ji}^c - \hat{b}_j^{c-} x_{ji}^c)] + e_i^{c-} - e_i^{c+} = y_i^c \quad (۳-۲۶)$$

$$\hat{b}_0^{l_1} + [\sum_{j=1}^p (\hat{b}_j^{l_1+} x_{ji}^{l_1} - \hat{b}_j^{l_2-} x_{ji}^{l_1})] + e_i^{l_1-} - e_i^{l_1+} = y_i^{l_1} \quad (۴-۲۶)$$

$$\hat{b}_0^{l_2} + [\sum_{j=1}^p (\hat{b}_j^{l_2+} x_{ji}^{l_2} - \hat{b}_j^{l_1-} x_{ji}^{l_2})] + e_i^{l_2-} - e_i^{l_2+} = y_i^{l_2} \quad (۵-۲۶)$$

$$e_i^{m+} \geq 0, \quad e_i^{m-} \geq 0, \quad (۶-۲۶)$$

$$\hat{b}_0^m \in R \quad (۷-۲۶)$$

$$\hat{b}_0^{u_1} \leq \hat{b}_0^{l_1} \leq \hat{b}_0^c \leq \hat{b}_0^{l_2} \leq \hat{b}_0^{u_2} \quad (۸-۲۶)$$

$$0 \leq \hat{b}_j^{u_1+} \leq \hat{b}_j^{l_1+} \leq \hat{b}_j^{c+} \leq \hat{b}_j^{l_2+} \leq \hat{b}_j^{u_2+}, \quad (۹-۲۶)$$

$$0 \leq \hat{b}_j^{u_1-} \leq \hat{b}_j^{l_1-} \leq \hat{b}_j^{c-} \leq \hat{b}_j^{l_2-} \leq \hat{b}_j^{u_2-}, \quad (۱۰-۲۶)$$

$$0 \leq x_{ji}^{u_1} \leq x_{ji}^{l_1} \leq x_{ji}^c \leq x_{ji}^{l_2} \leq x_{ji}^{u_2},$$

$$y_i^{u_1} \leq y_i^{l_1} \leq y_i^c \leq y_i^{l_2} \leq y_i^{u_2},$$

$$j=1, \dots, p; \quad m \in \{u_1, l_1, c, l_2, u_2\}; \quad i=1, \dots, n.$$

قضیه ۳-۲: فرض کنید $\tilde{y}_i \in \mathcal{F}_2$, $\tilde{b}_j \in \mathcal{F}_2$ و $\tilde{x}_{ij} \succeq \tilde{0}$ اگر $(\tilde{x}_{ij}, \tilde{y}_i)$ در مدل زیر صدق کند:

$$\tilde{y}_i \approx \bigoplus_{j=0}^p \tilde{b}_j \otimes \tilde{x}_{ji} = \tilde{b}_0 \oplus \tilde{b}_1 \otimes \tilde{x}_{i1} \oplus \dots \oplus \tilde{b}_p \otimes \tilde{x}_{ip}, \quad i=1, \dots, n \quad (۲۷)$$

در این صورت: الف- جواب بهینه مسأله (۲۶) شامل $\tilde{b} = (\tilde{b}_0, \tilde{b}_1, \dots, \tilde{b}_p)$ است.

ب- $\hat{y}_i \approx \tilde{y}_i$ یعنی خروجی مشاهده شده و برآورد شده با هم برابرند.

اثبات: ابتدا نشان می دهیم که $\tilde{b} = (\tilde{b}_0, \tilde{b}_1, \dots, \tilde{b}_p)$ یک جواب موجه مسأله (۲۶) است. با توجه به اینکه

$$\hat{b}_j \approx \tilde{b}_j^+ \oplus \tilde{b}_j^- \approx [(\hat{b}_j^{u_1+} - \hat{b}_j^{u_2-}, \hat{b}_j^{u_2+} - \hat{b}_j^{u_1-}, \hat{b}_j^{c+} - \hat{b}_j^{c-}, (\hat{b}_j^{l_1+} - \hat{b}_j^{l_2-}, \hat{b}_j^{l_2+} - \hat{b}_j^{l_1-})]$$

$$e_i^{k+} = e_i^{k-} = 0 \quad (۲۸)$$

$$\hat{b}_0^k = \tilde{b}_0^k \quad (۲۹)$$

$$\hat{b}_j^{k-} = \begin{cases} 0 & \text{if } b_j^k \geq 0 \\ -b_j^k & \text{if } b_j^k < 0 \end{cases} \quad (۳۰)$$

$$k \in \{u_1, l_1, c, l_2, u_2\}, \quad j=1, \dots, p$$

واضح است که (۲۸) تا (۳۰) در قیدهای (۶-۲۶) و (۷-۲۶) صدق می کنند. حال بدون کاستن از کلیات و برای سادگی فرض می کنیم $\tilde{b}_j \succeq \tilde{0}$ در این صورت:

$$\hat{b}_j^{k-} = 0, \quad (31)$$

$$\hat{b}_j^{k+} = b_j^k \quad ; j=1, \dots, p; k \in \{u_1, l_1, c, l_2, u_2\} \quad (32)$$

با جای گذاری (۳۱) و (۳۲) در (۱-۲۶) تا (۵-۲۶) خواهیم داشت:

$$\hat{y}_i^{u_1} + e_i^{u_1-} - e_i^{u_1+} = \hat{b}_0^{u_1} + \sum_{j=1}^p (\hat{b}_j^{u_1+} x_{ji}^{u_1} - \underbrace{\hat{b}_j^{u_1-} x_{ji}^{u_1}}_0) + e_i^{u_1-} - e_i^{u_1+} = b_0^{u_1} + \sum_{j=1}^p b_j^{u_1} x_{ji}^{u_1} = y_i^{u_1} \quad (1-33)$$

$$\hat{y}_i^{u_2} + e_i^{u_2-} - e_i^{u_2+} = \hat{b}_0^{u_2} + \sum_{j=1}^p (\hat{b}_j^{u_2+} x_{ji}^{u_2} - \underbrace{\hat{b}_j^{u_2-} x_{ji}^{u_2}}_0) + e_i^{u_2-} - e_i^{u_2+} = b_0^{u_2} + \sum_{j=1}^p b_j^{u_2} x_{ji}^{u_2} = y_i^{u_2} \quad (2-33)$$

$$\hat{y}_i^c + e_i^{c-} - e_i^{c+} = \hat{b}_0^c + \sum_{j=1}^p (\hat{b}_j^{c+} x_{ji}^c - \underbrace{\hat{b}_j^{c-} x_{ji}^c}_0) + e_i^{c-} - e_i^{c+} = b_0^c + \sum_{j=1}^p b_j^c x_{ji}^c = y_i^c \quad (3-33)$$

$$\hat{y}_i^{l_1} + e_i^{l_1-} - e_i^{l_1+} = \hat{b}_0^{l_1} + \sum_{j=1}^p (\hat{b}_j^{l_1+} x_{ji}^{l_1} - \underbrace{\hat{b}_j^{l_1-} x_{ji}^{l_1}}_0) + e_i^{l_1-} - e_i^{l_1+} = b_0^{l_1} + \sum_{j=1}^p b_j^{l_1} x_{ji}^{l_1} = y_i^{l_1} \quad (4-33)$$

$$\hat{y}_i^{l_2} + e_i^{l_2-} - e_i^{l_2+} = \hat{b}_0^{l_2} + \sum_{j=1}^p (\hat{b}_j^{l_2+} x_{ji}^{l_2} - \underbrace{\hat{b}_j^{l_2-} x_{ji}^{l_2}}_0) + e_i^{l_2-} - e_i^{l_2+} = b_0^{l_2} + \sum_{j=1}^p b_j^{l_2} x_{ji}^{l_2} = y_i^{l_2} \quad (5-33)$$

معادلات (۱-۳۳) تا (۵-۳۳) نشان می‌دهند که $\tilde{b} \approx (\tilde{b}_0, \tilde{b}_1, \dots, \tilde{b}_p)$ در قیود (۱-۲۶) تا (۵-۲۶) صدق می‌کنند یعنی $\tilde{b} \approx (\tilde{b}_0, \tilde{b}_1, \dots, \tilde{b}_p)$ جواب موجه مسأله (۲۶) است. برای اثبات بهینه بودن جواب، توجه می‌کنیم که مقدار تابع هدف مسئله (۲۶) به ازاء این جواب موجه صفر می‌شود که حداقل مقدار ممکن برای یک تابع نامنفی است. لذا مقادیر پیشنهادی یک جواب بهینه مسأله (۲۶) بوده و خواهیم داشت $\tilde{b}_j^+ \ominus \tilde{b}_j^- \approx \tilde{b}_j$ یعنی جواب بهینه مسأله (۲۶) شامل $\tilde{b} \approx (\tilde{b}_0, \tilde{b}_1, \dots, \tilde{b}_p)$ است.

اثبات قسمت دوم قضیه بلافاصله از (۱-۳۳) تا (۵-۳۳) حاصل می‌شود که به ازاء $i = 1, \dots, n$ داریم:

$$\hat{y}_i^{u_1} = b_0^{u_1} + \sum_{j=1}^p b_j^{u_1} \times x_{ji}^{u_1} = y_i^{u_1} \quad (1-34)$$

$$\hat{y}_i^{u_2} = b_0^{u_2} + \sum_{j=1}^p b_j^{u_2} \times x_{ji}^{u_2} = y_i^{u_2} \quad (2-34)$$

$$\hat{y}_i^c = b_0^c + \sum_{j=1}^p b_j^c \times x_{ji}^c = y_i^c \quad (3-34)$$

$$\hat{y}_i^{l_1} = b_0^{l_1} + \sum_{j=1}^p b_j^{l_1} \times x_{ji}^{l_1} = y_i^{l_1} \quad (4-34)$$

$$\hat{y}_i^{l_2} = b_0^{l_2} + \sum_{j=1}^p b_j^{l_2} \times x_{ji}^{l_2} = y_i^{l_2} \quad (5-34)$$

از (۱-۳۳) تا (۵-۳۳) نتیجه می‌گیریم که $\hat{y}_i \approx \tilde{y}_i$ و اثبات کامل می‌شود.

۱-۴ شناسایی و تحلیل مشاهدات فازی دور افتاده

یکی از اولین گام‌ها در جهت به دست آوردن یک تحلیل رگرسیونی معتبر، شناسایی مشاهدات دور افتاده به منظور حذف یا کاستن اثرات و یا بررسی علل رخ دادن این گونه مشاهدات می‌باشد. با اینکه نقاط دور افتاده اغلب به عنوان یک خطا یا اختلال در نظر گرفته شده و در برخی موارد حضور حتی یک نقطه دور افتاده در مجموعه داده‌ها منجر به برآوردهای دور از انتظار و غیرقابل اطمینان می‌شود، ضمن اینکه در مواردی همچون داده‌کاوی، با شناسایی و تحلیل نقاط پرت اطلاعات بسیار مهمی استخراج می‌شود.

یک مشاهده پرت یا دور افتاده فازی حداقل باید یکی از خصوصیات زیر را داشته باشد:

- ۱- در مقایسه با اکثریت داده‌ها دارای پهنای راست یا پهنای چپ (و یا هر دو) عریض‌تر داشته باشد.
- ۲- در مقایسه با اکثریت داده‌ها خیلی بزرگتر یا خیلی کوچکتر باشد.
- ۳- در مقایسه با اکثریت داده‌ها ارتباط بین ورودی-خروجی آن تفاوت معنی‌داری داشته باشد.

داده پرت ممکن است به علت بروز خطاهای ناخواسته در مرحله جمع آوری، ثبت و یا انتقال داده‌ها رخ دهد و یا اینکه ذاتاً از سایر داده‌ها متفاوت باشد. مثلاً وقتی که وزن نوزادان در بازه $[۲/۵-۴]$ کیلوگرم باشد ثبت وزن نوزادی با وزن $۵/۲$ یا $۱/۸$ کیلوگرم مشاهده پرت (غیر عادی) محسوب می‌شود حال وزن‌های غیر عادی ممکن است وزن واقعی یک نوزاد بوده و یا اینکه ناشی از بروز خطای ناخواسته (در مراحل اندازه‌گیری، ثبت و یا انتقال داده‌ها) باشد. در مطالعه حاضر به منظور مدیریت اثرات نقاط پرت، ابتدا یک مدل رگرسیونی را به کل داده‌های اصلی برازش داده و سپس مشاهداتی را که قدر مطلق خطای حاصل از تفاضل مقدار برآوردی و مشاهده شده آنها عددی بزرگ باشد را از مجموعه داده‌ها حذف و در مرحله دوم یک مدل رگرسیونی بر داده‌های بدون نقطه پرت برازش داده‌ایم.

۲-۴- اعتبارسنجی مدل پیشنهادی

اعتبارسنجی و ارزیابی میزان بهینه بودن یک مدل رگرسیونی معمولاً به دو روش: الف- ارزیابی براساس فرضیاتی که باید مدل در آنها صدق کند ب- ارزیابی براساس کارایی مدل در پیش‌بینی مقادیر جدید صورت می‌پذیرد در روش اول ارزیابی مدل فقط با داده‌هایی که مشاهده شده و در ساختن مدل به کار رفته‌اند انجام می‌شود لذا در این روش نمی‌توان کارایی و بهینگی مدل را برای داده‌های جدیدی که هنگام مدل‌سازی مشاهده نشده‌اند، ارزیابی نمود.

در روش دوم داده‌ها به دو قسمت داده‌های آموزشی^{۳۳} و آزمایشی^{۳۴} تقسیم شده و برای ارزیابی مدل از داده‌هایی استفاده می‌شود که مشاهده شده ولی در هنگام ساختن مدل به کار گرفته نشده‌اند. در مطالعه حاضر کارایی مدل پیشنهادی با استفاده از روش دوم و با رویکرد اعتبارسنجی متقابل^{۳۴} Leave-One-Out به شرح زیر ارزیابی و بررسی شده است:

در مرحله اول، اولین مشاهده یعنی $(\tilde{x}_1, \tilde{y}_1)$ را از مجموعه داده‌ها خارج نموده و پارامترهای مدل را با $n-1$ مشاهده باقی مانده برآورد نموده و با مدل جدید مقدار متغیر پاسخ را برای مشاهده خارج شده پیش‌بینی و آنرا $\hat{y}_{(-1)}$ می‌نامیم. در مرحله دوم مشاهده خارج شده در مرحله اول را به مجموعه داده‌ها برگردانده و این فرآیند را برای بقیه مشاهدات تکرار نموده و از این طریق $\hat{y}_{(-2)}, \dots, \hat{y}_{(-n)}$ محاسبه می‌کنیم. میانگین قدر مطلق خطای خارج از نمونه) حاصل از $R_E = \frac{1}{MAE} |MAE - MAE_{(-)}|$ را بدست می‌آوریم. اگر مقدار $MAE_{(-)}$ به مقدار MAE نزدیک بوده و R_E کوچک باشد توانایی و اعتبار مدل تأیید و مدل بهینه است.

۳-۴- معیارهای سنجش کارایی مدل‌های رگرسیون فازی

معیارهای فاصله و شباهت فازی به عنوان دو ابزار مهم در محیط فازی برای مقایسه مجموعه‌های فازی معرفی شده‌اند که در مدل‌سازی، برنامه‌ریزی، تصمیم‌گیری، استدلال و مواردی از این دست به صورت عملی به کار برده شده‌اند. برای ارزیابی مدل‌های رگرسیون فازی، چندین معیار فاصله و شباهت توسط محققان ارائه شده است [۴، ۸، ۱۸، ۲۴]. در این بخش، ما الگوریتم جدیدی برای تعیین فاصله و شباهت بین دو عدد فازی نوع-۱ مثلثی و نوع-۲ بازهای

^{۳۳} Training set

^{۳۳} Test set

^{۳۴} Cross validation

مقدار مثلثی را معرفی و آنها را جهت مقایسه خروجی‌های مشاهده شده و برآورد شده یک مدل رگرسیونی فازی به کار می‌بریم.

۴-۳-۱- اندازه فاصله فازی

اندازه فاصله فازی^{۳۵} به منظور تعیین میزان تفاوت دو مجموعه فازی محاسبه می‌شود. علت استفاده از فاصله فازی در یک مدل رگرسیونی اندازه‌گیری میزان تفاوت بین مقادیر پاسخ مشاهده شده و برآورد شده بوده و هر چقدر مقدار آن کوچکتر باشد نشانگر کارایی بالای مدل می‌باشد. در این قسمت بر پایه اندازه فاصله فازی، یک فرمول و چهارچوب جدید برای محاسبه میانگین میزان فاصله نسبی بین مقادیر مشاهده شده و برآورد شده متغیر خروجی یک مدل رگرسیونی فازی را (برای محیط‌های فازی نوع-۱ مثلثی و همچنین فازی نوع-۲ بازه‌ای مقدار مثلثی) به شرح زیر ارائه می‌کنیم:

$$1- \quad Md = \frac{1}{n} \sum_{i=1}^n d_i = \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{3} \sum_{k=1}^3 \left| \frac{\hat{y}_i^k - y_i^k}{y_i^k} \right| \right] \quad \text{که در آن} \quad d_i = \frac{1}{3} \sum_{k=1}^3 \left| \frac{\hat{y}_i^k - y_i^k}{y_i^k} \right| \quad \text{فاصله نسبی متغیر خروجی مشاهده شده}$$

و برآورد شده در مشاهده نام در محیط فازی نوع-۱ بوده، $d_i \in [0,1]$ و $Md \in [0,1]$. ضمناً $\hat{y}_i = [(y_i^1, y_i^2, y_i^3)]$ و $\tilde{y}_i = [(\hat{y}_i^1, \hat{y}_i^2, \hat{y}_i^3)]$ به ترتیب متغیرهای پاسخ مشاهده شده و برآورد شده در یک مدل رگرسیونی فازی نوع-۱ مثلثی می‌باشند.

$$2- \quad Md = \frac{1}{n} \sum_{i=1}^n d_i = \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{5} \sum_{k=1}^2 \left(\left| \frac{\hat{y}_i^{u_k} - y_i^{u_k}}{y_i^{u_k}} \right| + \left| \frac{\hat{y}_i^k - y_i^k}{y_i^k} \right| \right) + \left(\left| \frac{\hat{y}_i^c - y_i^c}{y_i^c} \right| \right) \right] \quad \text{که در آن،} \quad d_i \in [0,1] \quad \text{و} \quad Md \in [0,1]$$

و $\tilde{y}_i = [(y_i^1, y_i^2, y_i^3), y_i^c, (y_i^{u_1}, y_i^{u_2})]$ ضمناً $d_i = \frac{1}{5} \sum_{k=1}^2 \left(\left| \frac{\hat{y}_i^{u_k} - y_i^{u_k}}{y_i^{u_k}} \right| + \left| \frac{\hat{y}_i^k - y_i^k}{y_i^k} \right| \right) + \left(\left| \frac{\hat{y}_i^c - y_i^c}{y_i^c} \right| \right)$ و $\hat{y}_i = [(\hat{y}_i^1, \hat{y}_i^2), \hat{y}_i^c, (\hat{y}_i^{u_1}, \hat{y}_i^{u_2})]$ به ترتیب متغیرهای پاسخ مشاهده شده و برآورد شده در یک مدل رگرسیونی فازی نوع-۲ بازه‌ای مقدار مثلثی می‌باشند. لازم به توضیح است کوچک بودن مقدار Md بیانگر کارایی بالای مدل می‌باشد.

۴-۳-۲- اندازه تشابه فازی^{۳۶}

اندازه تشابه فازی به منظور تعیین میزان تشابه دو مجموعه فازی محاسبه می‌شود که مقدار آن یک عدد حقیقی در بازه $[0,1]$ می‌باشد. علت استفاده از اندازه تشابه فازی در یک مدل رگرسیونی فازی، اندازه‌گیری میزان تشابه بین مقادیر پاسخ مشاهده شده و برآورد شده بوده و هر چقدر مقدار آن به یک نزدیکتر باشد نشانگر کارایی بالای مدل رگرسیونی می‌باشد. در این قسمت از مطالعه، یک فرمول جدید برای محاسبه میانگین میزان تشابه بین مقادیر مشاهده شده و برآورد شده متغیر خروجی یک مدل رگرسیونی فازی (زمانی که داده‌ها فازی نوع-۲ بازه‌ای مثلثی باشند) به شرح زیر ارائه می‌شود:

$$1. \quad MS = \frac{1}{n} \sum_{i=1}^n S(\tilde{y}_i, \hat{y}_i) \quad \text{که در آن} \quad S(\hat{y}_i, \tilde{y}_i) = \frac{1}{3} \sum_{k=1}^3 \frac{\min(\hat{y}_i^k, y_i^k)}{\max(\hat{y}_i^k, y_i^k)} \quad MS \in [0,1] \quad \text{ضمن اینکه} \quad \hat{y}_i = [(y_i^1, y_i^2, y_i^3)] \quad \text{و}$$

$\tilde{y}_i = [(\hat{y}_i^1, \hat{y}_i^2, \hat{y}_i^3)]$ به ترتیب متغیرهای پاسخ مشاهده شده و برآورد شده در یک مدل رگرسیونی فازی نوع-۱ مثلثی می‌باشند.

$$2. \quad MS = \frac{1}{n} \sum_{i=1}^n S(\hat{y}_i, \tilde{y}_i) \quad \text{که در آن} \quad S(\hat{y}_i, \tilde{y}_i) = \frac{1}{5} \left[\sum_{k=1}^2 \left(\frac{\min(\hat{y}_i^{u_k}, y_i^{u_k})}{\max(\hat{y}_i^{u_k}, y_i^{u_k})} + \frac{\min(\hat{y}_i^k, y_i^k)}{\max(\hat{y}_i^k, y_i^k)} \right) + \frac{\min(\hat{y}_i^c, y_i^c)}{\max(\hat{y}_i^c, y_i^c)} \right] \quad MS \in [0,1] \quad \text{ضمن اینکه}$$

$\tilde{y}_i = [(y_i^1, y_i^2), y_i^c, (y_i^{u_1}, y_i^{u_2})]$ و $\hat{y}_i = [(\hat{y}_i^1, \hat{y}_i^2), \hat{y}_i^c, (\hat{y}_i^{u_1}, \hat{y}_i^{u_2})]$ به ترتیب متغیرهای پاسخ مشاهده شده و

^{۳۵} Fuzzy distance measure

^{۳۶} Fuzzy similarity measure

برآورد شده در یک مدل رگرسیون فازی نوع-۲ بازه‌ای مقدار مثلثی می‌باشند. لازم به توضیح است هر اندازه مقدار مقدار MS نزدیک به یک باشد بیانگر کارا و بهینه بودن مدل می‌باشد.

۴-۴ تحلیل مدل پیشنهاد شده با مثال‌های عدد

مثال ۱: در این مثال از طریق داده‌های ارائه شده در جدول ۲، پارامترهای مدل رگرسیونی فازی نوع-۲ بازه‌ای مقدار با رویکرد قدر مطلق انحرافات فازی و با بهره‌گیری از برنامه‌ریزی آرمانی فازی نوع-۲ بازه‌ای مقدار و با تعمیم فرمول فاصله فازی هامینگ به شرح زیر برآورد شده است:

$$\begin{aligned}\hat{b}_0 &= [(1/8, 8/76), 5/19, (3/7, 6/14)] \\ \hat{b}_1 &= [(0/8, 0/214), 0/198, (0/12, 0/207)] \\ \hat{y}_i &= \hat{b}_0 \oplus \hat{b}_1 \otimes \tilde{x}_i \\ &= [(\hat{b}_0^{u1} + \hat{b}_1^{u1} x_i^{u1}, \hat{b}_0^{u2} + \hat{b}_1^{u2} x_i^{u2}), \hat{b}_0^c + \hat{b}_1^c x_i^c, (\hat{b}_0^{l1} + \hat{b}_1^{l1} x_i^{l1}, \hat{b}_0^{l2} + \hat{b}_1^{l2} x_i^{l2})] \\ \hat{y}_i &= [(1/87 + 0/88 x_i^{u1}, 8/69 + 0/214 x_i^{u2}), 5/19 + 0/198 x_i^c, (3/61 + 0/12 x_i^{l1}, 7/44 + 0/207 x_i^{l2})], \quad i = 1, \dots, n \quad (35)\end{aligned}$$

به منظور شناسایی مشاهدات دور افتاده، در مدل (۳۵) مقدار خطا یعنی $|\tilde{y}_i - \hat{y}_i|$ برای $i = 1, \dots, n$ محاسبه و سومین و هیجدهمین داده به علت داشتن خطای بزرگتر در در مقایسه با بقیه داده‌ها، به عنوان داده پرت شناسایی و از مجموعه داده‌ها حذف و بعد از حذف نقاط پرت، مجدداً ضرایب رگرسیونی به شرح زیر برآورد گردید:

$$\begin{aligned}\hat{b}_0 &= [(1/83, 8/48), 5/14, (3/58, 7/38)] \\ \hat{b}_1 &= [(0/85, 0/216), 0/2, (0/13, 0/208)] \\ \hat{y}_i &= \hat{b}_0 \oplus \hat{b}_1 \otimes \tilde{x}_i \\ &= [(\hat{b}_0^{u1} + \hat{b}_1^{u1} x_i^{u1}, \hat{b}_0^{u2} + \hat{b}_1^{u2} x_i^{u2}), \hat{b}_0^c + \hat{b}_1^c x_i^c, (\hat{b}_0^{l1} + \hat{b}_1^{l1} x_i^{l1}, \hat{b}_0^{l2} + \hat{b}_1^{l2} x_i^{l2})] \\ \hat{y}_i &= [(1/83 + 0/85 x_i^{u1}, 8/48 + 0/216 x_i^{u2}), 5/14 + 0/2 x_i^c, (3/58 + 0/13 x_i^{l1}, 7/38 + 0/208 x_i^{l2})] \quad (36)\end{aligned}$$

به منظور ارزیابی کارایی مدل (۳۸)، مقادیر $|\tilde{y}_i - \hat{y}_i|$ و $|\tilde{y}_i - \hat{y}_{(-i)}|$ به ازاء $i = 1, \dots, n$ محاسبه و

$$M_{(-)} = \frac{1}{19} \sum_{i=1}^n |\tilde{y}_i - \hat{y}_{(-i)}| = 1/20, \quad MAE = \frac{1}{19} \sum_{i=1}^n AE_i = \frac{1}{19} \sum_{i=1}^n |\tilde{y}_i - \hat{y}_i| = 1/12$$

نزدیک MAE بوده و R_E خیلی کوچک است پس می‌توان گفت مدل (۳۶) معتبر و بهینه است.

در ستون اول و دوم جدول ۲، مقادیر ۲۲ مشاهده ورودی-خروجی فازی نوع-۲ بازه‌ای مقدار ارائه شده است. همچنین در ستون سوم این جدول مقادیر پیش‌بینی شده متغیر پاسخ و در ستون چهارم جدول ۲ فاصله فازی بین مقادیر مشاهده شده و پیش‌بینی شده متغیر پاسخ محاسبه شده است که نتایج این محاسبات بصورت بصری در شکل ۵-۱ نمایش داده شده است. در نمودار ۵-۱ محور x ‌ها بر حسب مقدار دیفازی شده متغیر فازی $\tilde{x}$ و محور y ‌ها نیز بر پایه مولفه‌های متغیر فازی $\tilde{y}$ (یعنی بر حسب مقادیر $y^{u1}, y^{u2}, y^c, y^{l1}, y^{l2}$) درجه‌بندی شده است.

همان‌طور که ملاحظه می‌شود مشاهده سوم (در ردیف سوم جدول ۲) و نیز مشاهده ۱۸ام به عنوان نقاط پرت فازی شناسایی و حذف شده‌اند. پس از حذف نقاط حذف و ایجاد مدل نهایی برپایه داده‌های عاری از نقاط پرت، نتایج حاصل از مدل نهایی، مجدداً مقادیر برآورد شده متغیر پاسخ و نیز فاصله بین متغیر پاسخ مشاهده شده و برآورد شده محاسبه و نتایج در ستونهای چهارم و پنجم جدول ۳ ارائه شده است.

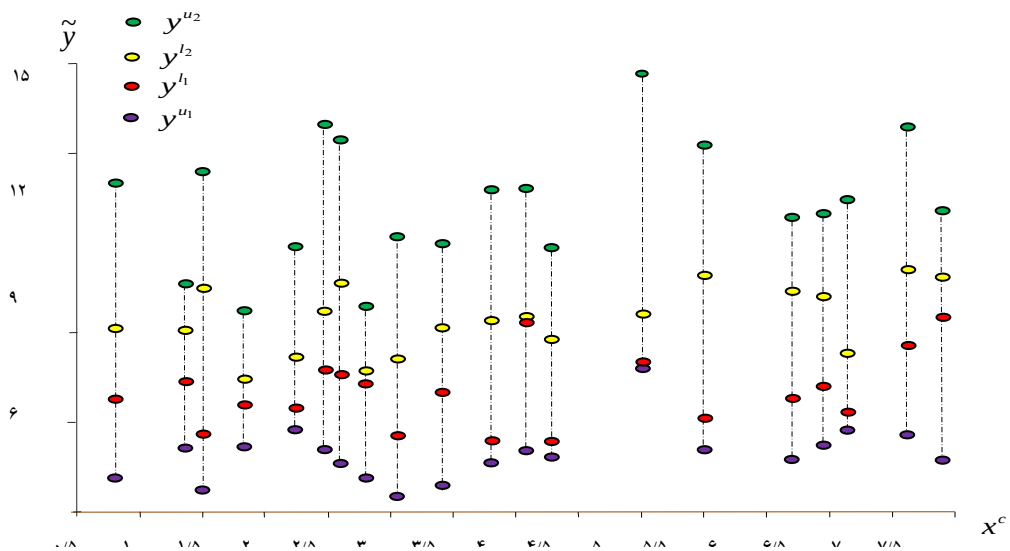
در نمودار ۵-۲ مدل برازش شده به داده‌های بدون نقطه پرت برازش شده است. در این نمودار محور محور X ها بر حسب مقادیر دیفازی شده متغیر فازی $\tilde{X}$ و محور Y ها نیز بر پایه مولفه‌های متغیر پاسخ برآورد شده (یعنی برحسب مقادیر $\hat{y}^{u_1}, \hat{y}^{u_2}, \hat{y}^c, \hat{y}^{l_1}, \hat{y}^{l_2}$) درجه‌بندی شده است.

جدول ۲ مقادیر متغیرهای ورودی- خروجی و فاصله فازی بین خروجی مشاهده شده و برآورد شده قبل از حذف نقاط پرت

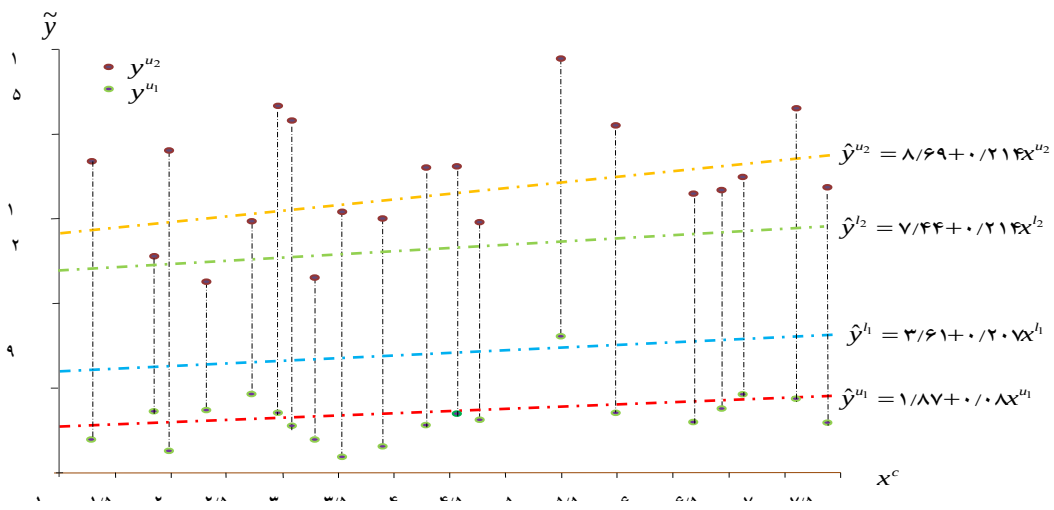
i	$\tilde{X}_i$	$\tilde{Y}_i$	$\hat{Y}_i$	$ \tilde{Y}_i - \hat{Y}_i $
۱	$[32, 7/24], 3/80, (4/50, 6/55)]$	$(1/71, 10/85), 6/48, (2/45, 6/66)]$	$1/98, 10/20, 5/94, (4/15, 8/80)]$	۱/۰۵
۲	$[(3/20, 8/4), 5/5, (4/7, 7/69)]$	$(2/14, 12/34), 7/99, (3/20, 11/2)]$	$(2/13, 10/5), 6/28, (4/17, 9/30)]$	۱/۳۴
۳	$1/18, 6/49, 5/12, (3/4, 5/29)]$	$5/12, 13/80, 6/32, (5/56, 11/20)]$	$[(1/96, 10/11), 6/2, (4/20, 8/56)]$	۲/۲۳
۴	$(3/50, 9/25), 6/20, (4/80, 7/6)]$	$[(1/82, 9/91), 7/45, (3/85, 9/74)]$	$(2/15, 10/7), 6/42, (4/19, 9/10)]$	۰/۶۴
۵	$[(0/31, 1/81)]$	$[(1/20, 11/107), 6/2, (3/82, 8/67)]$	$(1/89, 9/108), 5/35, (3/65, 7/81)]$	۰/۹۱
۶	$2/01, 5/13, 3/41, (2/5, 4/52)]$	$[(0/95, 9/103), 6/23, (4/06, 8/04)]$	$2/03, 9/79, 5/87, (3/91, 8/38)]$	۰/۵۴
۷	$[(0/9, 5/87), 2/8, (2/7, 4/8)]$	$[(1/21, 6/96), 4/79, (4/34, 6/52)]$	$(1/94, 9/95), 5/74, (3/93, 8/43)]$	۱/۴
۸	$0/24, 3/52, 1/83, (0/51, 3/4)]$	$[(2/25, 6/81), 4/52, (3/63, 6/59)]$	$(1/89, 9/44), 5/55, (3/67, 8/14)]$	۱/۱۲
۹	$1/2, 3/96, 2/47, (1/29, 3/69)]$	$[(2/14, 13/103), 6/79, (4/8, 8/59)]$	$[(1/96, 9/54), 5/8, (3/76, 8/2)]$	۱/۲۴
۱۰	$1/20, 13/64, 7/12, (4/47, 9/10)]$	$2/64, 12/95, 8/18, (5/63, 11/42)]$	$[(2/21, 11/60), 6/6, (4/15, 9/3)]$	۱/۳۹
۱۱	$1/5, 7/29, 4/28, (1/74, 6/34)]$	$[(1/9, 8/92), 5/85, (2/43, 8/51)]$	$[(1/9, 8/92), 5/85, (2/43, 8/51)]$	۰/۶۴
۱۲	$48, 10/81, 6/45, (4/35, 6/65)]$	$10/04, 7/27, (4/26, 8/24)]$	$11/0, 6/47, (4/13, 8/82)]$	۰/۵۲
۱۳	$1/6, 10/6, 5/1, (4/31, 7/87)]$	$(4/85, 14/71), 6/69, (5/08, 9/43)]$	$6/18, (4/12, 9/07)]$	۱/۶۷
۱۴	$1/2, 4/13, 2/24, (1/23, 3/26)]$	$[(2/82, 8/95), 5/26, (3/54, 8/26)]$	$(1/93, 9/57), 5/63, (3/76, 8/11)]$	۰/۴۵
۱۵	$(0/83, 3/76), 2/6, (2/57, 2/91)]$	$(1/68, 12/51), 7/73, (4/66, 8/84)]$	$(1/94, 9/49), 5/70, (3/92, 8/04)]$	۱/۳۷
۱۶	$0/05, 2/85, 1/5, (1/05, 2/08)]$	$[(0/8, 11/44), 7/55, (2/66, 8/87)]$	$[(1/87, 9/3), 5/49, (3/74, 7/87)]$	۱/۴۷
۱۷	$1/67, 9/53, 6/64, (4/27, 9/24)]$	$[(2/80, 10/51), 5/39, (3/40, 8/2)]$	$[(2/16, 10/7), 6/5, (4/12, 9/35)]$	۰/۷۷
۱۸	$1/62, 8/11, 4/51, (3/47, 5/56)]$	$[(4/8, 13/31), 7/5, (6/80, 9/01)]$	$(2/0, 10/42), 6/08, (4/02, 8/59)]$	۲/۰۶
۱۹	$1/8, 5/61, 3/05, (3/9, 4/71)]$	$[(0/59, 9/28), 5/19, (2/61, 8/08)]$	$(2/02, 9/89), 5/79, (4/03, 8/41)]$	۰/۸۸
۲۰	$3, 10/99, 7/40, (5/6, 10/83)]$	$(1/80, 10/15), 7/92, (6/58, 8/26)]$	$[(2/29, 11/0), 6/66, (4/28, 9/68)]$	۱/۲۷
۲۱	$(0/0, 6/54), 4/08, (4/91, 6/53)]$	$[(2/12, 10/89), 6/61, (6/4, 8/9)]$	$[(1/87, 10/11), 6/0, (3/72, 8/78)]$	۰/۸۹
۲۲	$1/8, 2/37, 1/36, (0/93, 1/93)]$	$[(2/2, 7/7), 6/15, (4/43, 7/63)]$	$[(1/88, 9/2), 5/46, (3/72, 7/84)]$	۰/۶۸

جدول ۳ مقادیر متغیرهای ورودی- خروجی و فاصله فازی بین روجی مشاهده شده و برآورد شده بعد از حذف نقاط پرت

i	$\tilde{X}_i$	$\tilde{Y}_i$	$\hat{Y}_i$	$ \tilde{Y}_i - \hat{Y}_i $
۱	$[32, 7/24], 3/80, (4/50, 6/55)]$	$(1/71, 10/85), 6/48, (2/45, 6/66)]$	$5/90, (4/17, 8/74)]$	۱/۰۸۳
۲	$[(3/20, 8/4), 5/5, (4/7, 7/69)]$	$(2/14, 12/34), 7/99, (3/20, 11/2)]$	$[(2/10, 10/3), 6/24, (4/19, 8/98)]$	۱/۴۰۵
۳	$(3/50, 9/25), 6/20, (4/80, 7/6)]$	$[(1/82, 9/91), 7/45, (3/85, 9/74)]$	$[(2/13, 10/5), 6/38, (4/20, 8/96)]$	۰/۶۲
۴	$[(0/31, 1/81)]$	$[(1/20, 11/107), 6/2, (3/82, 8/67)]$	$[(1/86, 8/88), 5/30, (3/62, 7/76)]$	۰/۹۷
۵	$2/01, 5/13, 3/41, (2/5, 4/52)]$	$[(0/95, 9/103), 6/23, (4/06, 8/04)]$	$(2/03, 9/79), 5/87, (3/91, 8/38)]$	۰/۴۹
۶	$[(0/9, 5/87), 2/8, (2/7, 4/8)]$	$[(1/21, 6/96), 4/79, (4/34, 6/52)]$	$[(1/91, 9/75), 5/70, (3/93, 8/38)]$	۱/۳۳
۷	$0/24, 3/52, 1/83, (0/51, 3/4)]$	$[(2/25, 6/81), 4/52, (3/63, 6/59)]$	$[(1/85, 9/24), 5/51, (3/65, 8/09)]$	۱/۰۶
۸	$1/2, 3/96, 2/47, (1/29, 3/69)]$	$[(2/14, 13/103), 6/79, (4/8, 8/59)]$	$[(1/93, 9/34), 5/63, (3/75, 8/15)]$	۱/۳۱
۹	$1/20, 13/64, 7/12, (4/47, 9/10)]$	$2/64, 12/95, 8/18, (5/63, 11/42)]$	$11/40, 6/56, (4/16, 9/25)]$	۱/۴۴
۱۰	$1/5, 7/29, 4/28, (1/74, 6/34)]$	$[(1/9, 8/92), 5/85, (2/43, 8/51)]$	$[(1/89, 10/10), 6/0, (3/81, 8/70)]$	۰/۵۷
۱۱	$48, 10/81, 6/45, (4/35, 6/65)]$	$10/04, 7/27, (4/26, 8/24)]$	$10/8, 6/43, (4/26, 8/24)]$	۰/۴۸
۱۲	$1/6, 10/6, 5/1, (4/31, 7/87)]$	$(4/85, 14/71), 6/69, (5/08, 9/43)]$	$6/14, (4/13, 9/02)]$	۱/۷۳
۱۳	$1/2, 4/13, 2/24, (1/23, 3/26)]$	$[(2/82, 8/95), 5/26, (3/54, 8/26)]$	$[(1/89, 9/37), 5/59, (3/74, 8/06)]$	۰/۴۲
۱۴	$(0/83, 3/76), 2/6, (2/57, 2/91)]$	$(1/68, 12/51), 7/73, (4/66, 8/84)]$	$[(1/90, 9/29), 5/66, (3/91, 7/99)]$	۱/۴۲
۱۵	$0/05, 2/85, 1/5, (1/05, 2/08)]$	$[(0/8, 11/44), 7/55, (2/66, 8/87)]$	$[(1/83, 9/11), 5/44, (3/72, 7/81)]$	۱/۵۲
۱۶	$1/67, 9/53, 6/64, (4/27, 9/24)]$	$[(2/80, 10/51), 5/39, (3/40, 8/2)]$	$[(2/14, 10/5), 6/47, (4/14, 9/30)]$	۰/۷۲
۱۷	$1/86, 5/61, 3/05, (3/47, 5/71)]$	$[(0/59, 9/28), 5/19, (2/61, 8/08)]$	$[(1/99, 9/69), 5/75, (4/03, 8/36)]$	۰/۸۲
۱۸	$3, 10/99, 7/40, (5/6, 10/83)]$	$(1/80, 10/15), 7/92, (6/58, 8/26)]$	$[(2/28, 10/9), 6/62, (4/31, 9/63)]$	۱/۳۳
۱۹	$(0/0, 6/54), 4/08, (4/91, 6/53)]$	$[(2/12, 10/89), 6/61, (6/4, 8/9)]$	$[(1/83, 9/89), 5/96, (3/70, 8/74)]$	۰/۹۶
۲۰	$1/8, 2/37, 1/36, (0/93, 1/93)]$	$[(2/2, 7/7), 6/15, (4/43, 7/63)]$	$[(1/85, 8/99), 5/41, (3/70, 7/78)]$	۰/۶۵



شکلا، ۵-۱ نمدها، $(x^c, \tilde{y})$



شکل ۲-۵ نمودار مولفه های $\hat{y}$

مثال ۲ (مطالعه مقایسه‌ای) در این مثال، قصد داریم عملکرد مدل پیشنهادی را با برخی از مدل‌های شناخته شده موجود مقایسه کنیم، اما هیچ یک از این مطالعات با داده‌های فازی نوع-۲ بازه‌ای مقدار مثلثی انجام نشده است. بنابراین، داده‌های فازی نوع-۱ مثلثی ارائه شده در ساکاو و یانو [۳۹] را با همان پشتیبان به داده‌های فازی نوع-۲ بازه‌ای مقدار مثلثی تبدیل نموده و مدل پیشنهادی را (براساس معیارهای سنجش کارایی ارائه شده در این مطالعه) با مطالعات ساکاو و یانو [۳۹]، چن و نین [۹]، دیاموند [۱۷] و حسن پور و همکاران [۱۴] مقایسه می‌کنیم. ما مقادیر پاسخ برآوردشده مدل‌های ارائه شده توسط چن و نین ($\hat{y}_{CN}$)، ساکاو و یانو ($\hat{y}_{SA}$)، دیاموند ($\hat{y}_D$) و حسن پور و همکاران ($\hat{y}_H$) در را به ترتیب در ستون‌های ۲ تا ۴ جدول ارائه کرده‌ایم.

جدول ۴ داده‌های ورودی-خروجی برگرفته از ساکاو- یانو و مقادیر پاسخ برآورد شده در مدل‌های [۹]، [۱۴]، [۱۷] و [۳۹]						
مشاهده	داده‌های ورودی-خروجی برگرفته از مطالعه ساکاو و یانو		مقدار پاسخ برآورد شده			
			$\hat{y}_{CN}$	$\hat{y}_{SA}$	$\hat{y}_D$	$\hat{y}_H$
۱	(۱/۲،۵/۲۰/۵)	(۳/۴،۵/۴۰/۵)	(۴/۸،۳/۵،۵/۳)	(۳/۴،۸/۵،۴/۰)	(۴/۴،۰/۵،۶/۲)	(۴/۴،۶/۵،۸/۱)
۲	(۳/۳۰،۰/۴،۵/۰)	(۵/۵۰،۰/۶،۵/۰)	(۵/۵۰،۰/۶،۵/۰)	(۴/۵،۵/۶،۲/۰)	(۴/۴،۸/۵،۵/۹)	(۵/۵،۳/۵،۵/۷)
۳	(۴/۵،۵/۶،۵/۵)	(۶/۷،۵/۸،۵/۵)	(۵/۶،۷/۷،۴/۱)	(۵/۶،۳/۷،۴/۷)	(۵/۶،۶/۷،۴/۲)	(۵/۶،۹/۶،۴/۸)
۴	(۶/۷،۵/۷،۰/۵)	(۶/۶۰،۰/۷،۵/۰)	(۶/۷،۵/۷،۰/۶)	(۶/۷،۳/۸،۳/۳)	(۶/۷،۴/۷،۲/۸)	(۶/۷،۸/۷،۱/۳)
۵	(۸/۸،۰/۹،۵/۰)	(۸/۸،۰/۹،۵/۰)	(۷/۷،۲/۸،۷/۲)	(۷/۸،۰/۹،۱/۳)	(۷/۸،۴/۸،۰/۶)	(۷/۷،۵/۷،۷/۹)
۶	(۹/۱۰،۵/۱۱،۵/۵)	(۷/۸۰،۰/۹،۰/۰)	(۷/۸،۹/۹،۶/۳)	(۷/۹،۸/۱۰،۳/۹)	(۹/۹،۲/۹،۰/۹)	(۸/۸،۲/۹،۶/۱)
۷	(۱۰/۱۱،۵/۱۱،۰/۵)	(۱۰/۱۰،۰/۱۱،۵/۰)	(۹/۸،۳/۹،۸/۳)	(۸/۹،۳/۱۰،۶/۹)	(۸/۹،۷/۹،۳/۹)	(۸/۸،۶/۹،۸/۱)
۸	(۱۲/۱۲،۰/۱۳،۵/۰)	(۹/۹۰،۰/۱۰،۵/۰)	(۹/۹۰،۰/۱۰،۵/۰)	(۹/۱۰،۰/۱۱،۴/۹)	(۹/۱۰،۵/۱۰،۱/۶)	(۹/۹،۳/۹،۵/۷)

در جدول ۵ داده‌های ورودی-خروجی ایجاد شده در این مطالعه (بر پایه داده‌های ساکاو-یانو) که با $(\tilde{X}_{our}, \tilde{Y}_{our})$ بیان شده و نیز مقادیر پاسخ برآورد شده‌ی مدل پیشنهادی $(\hat{Y}_{our})$ ارائه شده است.

جدول ۵ داده‌های ورودی-خروجی ایجاد شده در مطالعه حاضر و مقادیر پاسخ برآورد شده $(\hat{Y}_{our})$			
مشاهده	$\tilde{X}_{our}$	$\tilde{Y}_{our}$	$\hat{Y}_{our}$
۱	[(۱/۲،۷۵/۲۵)،۲/۰،(۱/۲،۵/۵)]	[(۳/۴،۷۵/۰)،۴/۲۵،(۳/۴،۵/۵)]	[(۴/۴،۲۱/۵۱)،۴/۳۷،(۳/۴،۹۶/۷۳)]
۲	[(۳/۳،۲۵/۷۵)،۳/۵،(۳/۴،۰/۰)]	[(۵/۵،۳۵/۷۵)،۵/۵،(۵/۶،۰/۰)]	[(۴/۵،۹۷/۲۸)،۵/۱۱،(۴/۵،۶۸/۴۹)]
۳	[(۵/۶،۰/۰)،۵/۵،(۴/۶،۵/۵)]	[(۷/۸،۰/۰)،۷/۵،(۶/۸،۵/۵)]	[(۵/۶،۸۵/۴۳)،۶/۰۹،(۵/۶،۴۲/۷۹)]
۴	[(۶/۷،۷۵/۲۵)،۷/۰،(۶/۷،۵/۵)]	[(۶/۶،۲۵/۷۵)،۶/۵،(۶/۷،۰/۰)]	[(۶/۷،۶۶/۰۱)،۶/۸۲،(۶/۷،۳۶/۲۸)]
۵	[(۸/۸،۲۵/۷۵)،۸/۵،(۸/۹،۰/۰)]	[(۸/۸،۲۵/۷۵)،۸/۵،(۱۰/۸،۰/۹)]	[(۷/۷،۴۲/۷۸)،۷/۵۶،(۷/۸،۰/۸/۰۴)]
۶	[(۱۰/۱۱،۰/۰)،۱۰/۵،(۹/۱۱،۵/۵)]	[(۷/۸،۵/۵)،۸/۰،(۷/۹،۰/۰)]	[(۸/۸،۲۸/۹۱)،۸/۵۴،(۷/۹،۸/۳۲)]
۷	[(۱۰/۱۱،۷۵/۲۵)،۱۱/۰،(۱۰/۱۱،۵/۵)]	[(۱۰/۱۰،۳۳/۷۳)،۱۰/۵،(۱۰/۱۰،۰/۹۸)]	[(۸/۹،۵۰/۰۳)،۸/۷۸،(۸/۹،۲۹/۳۳)]
۸	[(۱۲/۱۲،۲۵/۷۵)،۱۲/۵،(۱۲/۱۳،۰/۰)]	[(۹/۹،۲۵/۷۵)،۹/۵،(۹/۱۰،۰/۰)]	[(۹/۹،۳۸/۷۸)،۹/۵۲،(۹/۱۰،۰/۱۱)]

جدول ۶ مدل‌های برازش شده بر داده‌های ساکاو-یانو و مقایسه کارایی این مدل‌ها با مدل پیشنهادی ما			
ارائه دهنده مدل	مدل رگرسیون فاز ی برازش داده شده	معیار بکار برده شده	
		Md	MS
دیاموند	$\otimes \dot{X}_D (۰/۰،۴۴/۰،۴۴/۴۴) \oplus (۳/۳،۹۴/۳،۹۴/۹۴) \tilde{Y}_D =$	۰/۰۹۹	۰/۹۱۱
ساکاو و یانو	$\otimes \dot{X}_H (۰/۰،۵۰/۰،۵۸/۶۶) \oplus (۳/۳،۰۳/۳،۲۰/۳۷) \tilde{Y}_H =$	۰/۱۰۳	۰/۹۰۸
حسن پور و همکاران	$\otimes \dot{X}_H (۰/۰،۴۴/۰،۴۴/۴۴) \oplus (۳/۳،۹۴/۳،۹۴/۹۴) \tilde{Y}_H =$	۰/۱۰۴	۰/۹۰۸
چن و نین	$\otimes \dot{X}_{CN} (۰/۰،۴۴/۰،۴۴/۴۴) \oplus (۳/۳،۶۷/۴،۹۶/۲۲) \tilde{Y}_{CN} =$	۰/۰۹۸	۰/۹۱۵
مطالعه حاضر	$\otimes \tilde{X}_{PM} (۰/۰،۴۹/۵)،۰/۴۹،(۰/۰،۴۸/۵۱) \oplus [(۳/۳،۳۸/۴۱)،۳/۳۹،(۳/۳،۲۲/۴۵)] \hat{Y}_{PM} =$	۰/۰۹۴	۰/۹۱۹

با مراجعه به ستون‌های سوم و چهارم جدول فوق ملاحظه می‌کنیم که متوسط خطای نسبی مدل پیشنهادی ما در مقایسه با بقیه مدل‌ها کمتر و میانگین میزان تشابه در مدل پیشنهادی مطالعه ما از همه مدل‌ها بیشتر است که نشانگر بهینه و کارا بودن مدل پیشنهادی در این مطالعه می‌باشد.

۵ - نتیجه‌گیری

رگرسیون فازی نوع-۱ تعمیم یافته رگرسیون معمولی است که از طریق ترکیب تکنیک‌های آماری و رویکردهای فازی، مجموعه‌ای از ابزارهای کارآمد را برای مدل‌سازی و تجزیه و تحلیل داده‌هایی که همزمان با درجه‌ای از عدم قطعیت تصادفی و فازی همراه هستند در اختیار ما قرار می‌دهد. با این حال، رگرسیون فازی نوع-۱ در مواجهه با داده‌های با عدم قطعیت بالا (نظیر داده‌های برگرفته از دانش خبرگان و قضاوت‌ها و ادراکات انسانی) عملکرد قابل قبولی ندارد. در این زمینه، مجموعه‌های فازی نوع-۲ بازه‌ای مقدار با بهره‌گیری از انواع سیستم‌های رتبه‌بندی زبانی، ارزیابی‌های کلامی و داده‌های برگرفته از ادراکات انسانی و دانش خبرگان را به متغیرهای عددی عینی تبدیل نموده و زمینه را برای مدل کردن این قبیل داده فراهم می‌کنند. در مطالعه حاضر به منظور پر کردن بخشی از خلاء رگرسیون فازی نوع-۱، یک مدل رگرسیونی جدید بر پایه مجموعه‌های فازی نوع-۲ بازه‌ای و با رویکرد برنامه‌ریزی آرمانی فازی معرفی گردید. که ذیلاً به برخی از تازه‌ها و جنبه‌های مهم این مطالعه اشاره می‌شود:

۱. در این پژوهش ابتدا برخی از عملیات جبری فازی نوع-۱ به فازی نوع-۲ بازه‌ای تعمیم داده شده و سپس بر پایه فاصله هامینگ یک فرمول جدید برای محاسبه فاصله بین دو عدد فازی مثلثی نوع-۲ بازه‌ای مقدار ارائه شده است که در اندازه‌گیری میزان تمایز بین مجموعه‌های فازی (خصوصاً در موادی که رتبه بندی اعداد فازی مد نظر باشد) می‌تواند نقش موثری داشته باشد.

۲. در مطالعه حاضر یک فرآیند جدید برای شناسایی و تحلیل مشاهدات دور افتاده برای داده‌های فازی مثلثی نوع-۲ بازه‌ای مقدار ارائه شده است و ارائه این فرآیند جدید از آن جهت اهمیت دارد که شناسایی مشاهدات دور افتاده به منظور حذف یا کاستن اثرات و یا تحلیل علل رخ دادن آنها جزو مهم‌ترین گام‌ها در یک تجزیه و تحلیل یک مدل رگرسیونی معتبر محسوب می‌شود.

۳. ارزیابی اعتبار مدل پیشنهادی با یک متد جدید و با تعمیم رویکرد Leave One Out Cross Validation از محیط قطعی به محیط فازی نوع-۲ بازه‌ای مقدار (و از طریق محاسبه میانگین قدر مطلق خطای خارج از نمونه فازی و میانگین قدر مطلق خطای فازی ناشی از تفاضل بین مقادیر مشاهده شده و برآورد شده متغیر پاسخ) انجام گرفته است.

۴. در این مطالعه با برجسته نمودن شباهتها و تفاوت های Fuzzy GP, Fuzzy LP و Fuzzy NLP، برنامه‌ریزی آرمانی فازی نوع-۲ بازه‌ای مقدار معرفی و چگونگی تبدیل آن به GP و نهایتاً به LP تشریح شده است.

۵. اغلب مواردی که در این مطالعه برای اولین بار معرفی شده‌اند از طریق چند قضیه اثبات شده‌اند. البته در مدل پیشنهادی با افزایش تعداد متغیرهای ورودی و افزایش اندازه نمونه، تعداد قیدهای مدل نیز افزایش یافته و منجر به پیچیده شدن فرآیند برآورد ضرایب رگرسیونی می‌گردد. علاوه بر این با کوچکترین تغییر در متغیر پاسخ باید تمام قیدها نیز اصلاح شوند.

در خاتمه ارائه مدل‌های رگرسیون خطی و غیر خطی با رویکرد کمترین توان‌های دوم در محیط فازی نوع-۲ بازه‌ای مقدار و فازی نوع-۲ شهودی به عنوان موضوعات بالقوه برای کارهای آینده پیشنهاد می‌شود.

مراجع

- [۱] مختاری، میکائیل؛ حسین‌زاده لطفی، فرهاد؛ وحیدی، علیرضا (۱۳۸۸) ارائه یک مدل فازی برای اندازه‌گیری رضایت شغلی اساتید دانشگاه، تحقیق در عملیات در کاربردها، دوره ۶ شماره ۲۰، ۳۳-۴۷
- [۲] مختاری، میکائیل؛ کرد رستمی، سهراب (۱۳۹۰) ارائه یک مدل فازی برای ارزیابی عملکرد آموزشی اساتید، تحقیق در عملیات در کاربردها، شماره سوم، ۱۶۷-۱۷۸
- [۳] Allahviranloo T, (۲۰۲۰) Uncertain information and linear systems, Springer Nature Switzerland AG.
- [۴] Coopi R. (۲۰۰۸) Management of uncertainty in statistical reasoning: the case of regression analysis, International Journal of Approximate Reasoning ۴۷, ۲۸۴-۳۰۵
- [۵] Chen L.H, Hsueh C. (۲۰۰۷) A mathematical programming method for formulating a fuzzy regression model based on distance criterion, IEEE TRANSACTIONS ON SYSTEMS ۳۷.
- [۶] Choi S.H, Buckley J. (۲۰۰۷) Fuzzy regression using least absolute deviation estimators, Soft Comput ۱۲, ۲۵۷-۲۶۳
- [۷] Tanaka H, Uejima S, Asai K. (۱۹۸۲) Linear regression analysis with fuzzy model, IEEE Transactions on Systems, Man, and Cybernetics ۱۲, ۹۰۳-۹۰۷.
- [۸] Li J, Zeng W, Xie J, Yin Q. (۲۰۱۶) A new fuzzy regression model based on least absolute deviation, Engineering Applications of Artificial Intelligence ۵۲, ۵۴-۶۴.
- [۹] Chen L.H, Nien S.H. (۲۰۲۰) A new approach to formulate fuzzy regression models, Applied Soft Computing Journal, <https://doi.org/10.1016/j.asoc.2019.105915>.
- [۱۰] Shakouri H, Nadimi G. (۲۰۱۳) Outlier detection in fuzzy linear regression with crisp input-output by linguistic variable view, Applied Soft Computing ۱۳, ۷۳۴-۷۴۲.
- [۱۱] Taheri S. M. , Kelkinnama M. (۲۰۱۲) Fuzzy linear regression based on least absolute deviations, Iranian Journal of Fuzzy Systems ۹, ۱۲۱-۱۴۰.
- [۱۲] Zeng W, Fengb Q, Li J. (۲۰۱۷) Fuzzy least absolute linear regression, Applied Soft Computing ۵۲, ۱۰۰۹-۱۰۱۹
- [۱۳] Hojati M, Bector C.R, Smimou K. (۲۰۰۵) A simple method for computation of fuzzy linear regression, European Journal of Operational Research ۱۶۶, ۱۷۲-۱۸۴
- [۱۴] Hassanpour H, Maleki H, Yaghoobi M. (۲۰۱۱): A goal programming approach to fuzzy linear regression with fuzzy input-output data, Soft Computing, ۱۵, ۱۵۶۹-۱۵۸۰.
- [۱۵] Arefi M. (۲۰۲۰) Quantile fuzzy regression based on fuzzy outputs and fuzzy parameters, Soft Computing, <https://doi.org/10.1007/s00500-019-04424-2>
- [۱۶] Kelkinnama M, Taheri S.M. (۲۰۱۲) Fuzzy least-absolutes regression using shape-preserving operations, Inf Sci ۲۱۴, ۱۰۵-۱۲۰.
- [۱۷] Diamond P. (۱۹۸۸) Fuzzy Least Squares, Information Sciences ۴۶, ۱۴۱-۱۵۷.
- [۱۸] D'Urso P, Massari R. (۲۰۱۳) Weighted least squares and Least Median Squares estimation for the fuzzy linear regression analysis, METRON ۷۱, ۲۷۹-۳۰۶.
- [۱۹] Rabiei M.R, Arghami N.R, Taheri S. M, Sadeghpour Gildeh B. (۲۰۱۳) Least-squares approach to regression modeling in full, interval-valued fuzzy environment, Soft Compute, DOI 10.1007/s00500-013-1185-5
- [۲۰] Rafiei H, Ghoreyshi S. M. (۲۰۱۲) Symmetric fuzzy linear regression using multi-objective optimization, International Journal of Management Science and Engineering Management ۶, ۱۸۳-۱۹۱.
- [۲۱] Khammar A.H, Arefi M, Akbari M.G. (۲۰۲۰) A robust least squares fuzzy regression model based on kernel function, Iranian Journal of Fuzzy Systems ۱۷, ۱۰۵-۱۱۹

- [۲۲] Hosseinzadeh E, Hassanpour H, Arefi M, Aman M. (۲۰۱۶) A weighted goal programming approach to fuzzy linear regression with quasi type-۲ input-output data, *TWMS Journal of Applied and Engineering Mathematics* ۶، ۱۹۲-۲۱۳.
- [۲۳] Wei Y, Watada J, Pedrycz W. (۲۰۱۶) Design of a qualitative classification model through fuzzy support vector machine with type ۲ fuzzy expected regression classifier preset, *IEEJ Trans* ۱۱، ۳۴۸-۳۵۶.
- [۲۴] Hesamian G, Akbari M.G, Shams M. (۲۰۲۰) Parameter estimation in fuzzy partial univariate linear regression model with non-fuzzy inputs and triangular fuzzy outputs, *Iranian Journal of Fuzzy Systems*
- [۲۵] Srinivasan A, Geetharamani G. (۲۰۱۶) Linear programming problem with interval type-۲ fuzzy coefficients and an interpretation for its constraints, *Journal of Applied Mathematics*, ID ۸۴۹۶۸۱۲، ۱۱.
- [۲۶] Farhadinia B. (۲۰۱۴) Sensitivity analysis in interval valued trapezoidal fuzzy number linear programming problems, *Applied Mathematical Modeling* ۳۸، ۵۰-۶۲.
- [۲۷] Gong Y, Yang S, Dai L, Hu N. (۲۰۱۷) A new approach for ranking of interval type ۲ trapezoidal fuzzy numbers, *Journal of Intelligent & Fuzzy Systems* ۳۲، ۱۸۹۱-۱۹۰۲.
- [۲۸] Mendel J.M. (۲۰۱۷) Uncertain rule based fuzzy systems, DOI ۱۰.۱۰۰۷/۹۷۸-۳-۳۱۹-۵۱۳۷۰-۶_۶.
- [۲۹] Mishmast Nehi H, Keikha A (۲۰۱۶) TOPSIS and Choquet integral hybrid technique for solving MAGDM problems with interval type ۲ fuzzy numbers, *Journal of Intelligent & Fuzzy Systems* ۳۰، ۱۳۰۱-۱۳۱۰.
- [۳۰] Mishmast Nehi H, Keikha A (۲۰۱۶) TOPSIS and Choquet integral hybrid technique for solving MAGDM problems with interval type ۲ fuzzy numbers, *Journal of Intelligent & Fuzzy Systems* ۳۰، ۱۳۰۱-۱۳۱۰.
- [۳۱] Bhartia S.K, Singh S.R (۲۰۲۰) Interval-Valued Intuitionistic Fuzzy Linear Programming Problem, *New Mathematics and Natural Computation* ۱۶، ۵۳-۷۱.
- [۳۲] Sidhu S.K, Kumar A. (۲۰۱۸) Mehar method to find the fuzzy optimal solution of bounded fully fuzzy linear programs with symmetric trapezoidal fuzzy numbers, *Proc. Natl. Acad. Sci., India, Sect. A Phys* ۸۹، ۲۴۱-۲۴۸.
- [۳۳] Sun J, Lu Q. (۲۰۱۷) Regression analysis of a kind of trapezoidal fuzzy numbers based on a shape preserving operator, *Journal of Data Analysis and Information Processing* ۵، ۹۶-۱۱۴.
- [۳۴] Yang Z, Yin Y. (۲۰۱۳) Robust fuzzy varying coefficient regression analysis with crisp inputs and gaussian fuzzy output, *Journal of Computing Science and Engineering* ۷، ۲۶۳-۲۷.
- [۳۵] Zhang D, Lu Q. (۲۰۱۶) robust regression analysis with L-R type fuzzy input variables and fuzzy output variable, *Journal of Data Analysis and Information Processing* ۴، ۶۴-۸۰.
- [۳۶] Ebrahimnejad A. (۲۰۱۹) An effective computational attempt for solving fully fuzzy linear programming using MOLP problem, *Journal of Industrial and Production Engineering*, DOI: ۱۰.۱۰۸۰/۲۱۶۸۱۰۱۵، ۲۰۱۹، ۱۵۸۵۳۹۱.
- [۳۷] Jimenez M. (۲۰۲۰) Fuzzy pareto solutions in fully fuzzy multi objective linear programming, *AISC* ۹۹۱، ۵۰۹-۵۱۷.
- [۳۸] Pérez-Cañedo B, Rosete A, Verdegay J.L. (۲۰۲۰) A fuzzy goal programming approach to fully fuzzy linear regression, *CCIS* ۱۲۳۸، ۶۷۷-۶۸۸.
- [۳۹] Sakawa M, Yano H (۱۹۹۲) Multiobjective fuzzy linear regression analysis for fuzzy input-output data. *Fuzzy Sets Syst* ۴۷، ۱۷۳-۱۸۱