

روش‌های مقابله با هم‌خطی چندگانه در مدل‌های رگرسیون خطی فازی با ورودی و خروجی فازی

علیرضا عرب پور، مهدیه امینی

بخش آمار، دانشکده ریاضی و کامپیوتر، دانشگاه شهید باهنر کرمان

تاریخ پذیرش: ۱۴۰۰/۲/۶

تاریخ دریافت: ۱۳۹۹/۷/۱۱

نوع مقاله: علمی-پژوهشی

چکیده

وجود هم‌خطی چندگانه در مدل‌های رگرسیون چندگانه برآورد ضرایب رگرسیونی را تحت تاثیر قرار میدهد به همین علت تفسیر مناسب و گویایی از مدل رگرسیونی بدست نمی‌آید. در این مقاله از روش رگرسیون مؤلفه‌های اصلی فازی برای مواجهه با مشکل هم‌خطی چندگانه در مدل‌های رگرسیون فازی با ورودی و خروجی فازی استفاده می‌کنیم. همچنین روشهای مقابله با هم‌خطی چندگانه را معرفی می‌کنیم و در نهایت مثالهای عددی ارائه می‌دهیم که تاثیر به‌کارگیری روشهای مقابله با هم‌خطی چندگانه را نشان می‌دهد.

مشکل هم‌خطی چندگانه در مدل‌های رگرسیون ضرایب رگرسیونی نادرستی را نتیجه می‌دهد. برای مواجهه با مشکل هم‌خطی چندگانه روش‌های متعددی معرفی شده که یکی از آنها رگرسیون مؤلفه‌های اصلی می‌باشد. در این مقاله، رگرسیون مؤلفه‌های اصلی فازی برای مواجهه با مشکل هم

Mathematics Subject Classification (2010): 62A86, **Email:** arabpour@uk.ac.ir.

عبارات و کلمات کلیدی: رگرسیون فازی، هم‌خطی چندگانه، مولفه‌های اصلی

۱۳۹۹ (انجمن سیستم‌های فازی ایران)

خطی چندگانه در مدل های رگرسیون فازی با ورودی و خروجی فازی معرفی می شود. ابتدا، متغیرهای ورودی فازی به مؤلفه های اصلی فازی تبدیل می شوند، سپس از عملگرهای حافظ شکل برای رگرسیون مؤلفه های اصلی فازی استفاده میکنیم. برای مقابله با مشکل هم خطی چند گانه در رگرسیون خطی فازی، وانگ و تسار^۱ [۵۷] شاخص اطمینان جزئی (IC)^۲ را پیشنهاد دادند که زیر مجموعه ی ورودی مطلوب را انتخاب می کند. روش محاسبه ی شاخص اطمینان جزئی به صورت زیر است:

مدل رگرسیون خطی فازی با دو متغیر ورودی به صورت $\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 x_1 + \tilde{A}_2 x_2$ است. ضریب شاخص اطمینان جزئی بین \tilde{Y} و x_1 به شرط x_2 که با نماد $IC_{\tilde{Y} | x_1 | x_2}$ مشخص می شود، به صورت زیر بدست می آید:

$$IC_{\tilde{Y} | x_1 | x_2} = \frac{SSE(x_2) - SSE(x_1, x_2)}{SSE(x_2)} = \frac{SSR(x_1 | x_2)}{SSE(x_2)} \quad (1)$$

پس الگوریتم های گزینشی پیشرو به کار برده می شود که وارد کردن یک متغیر ورودی را بر طبق شاخص اطمینان جزئی بزرگتر پیشنهاد می دهد.

از معادله (۱) می توان دید که شاخص اطمینان جزئی مشابه ضریب تعیین جزئی می باشد. اگر چه می دانیم که ضریب تعیین جزئی برای سنجش اثر متغیر ورودی جدید استفاده می شود تا اینکه اثر هم خطی چندگانه.

پس از تشخیص وجود هم خطی چندگانه در مدل رگرسیون فازی، ابتدا مؤلفه های اصلی فازی را محاسبه می نماییم، سپس با استفاده از روش برنامه ریزی ریاضی تعمیم یافته مؤلفه های اصلی فازی را رگرسیون می کنیم. دو روش برای به دست آوردن مؤلفه های اصلی فازی در زیر معرفی شده که در بخش های بعد به طور کامل شرح داده می شوند. این دو روش عبارتند از:

۱- روش شبکه ی عصبی خود هم بسته

۲- روش آنالیز مؤلفه ی اصلی مراکز فازی

¹ Wang and Tsaur

² IC: Index of confidence

۱ آنالیز مؤلفه‌ی اصلی اطلاعات فازی با استفاده از شبکه‌ی عصبی

خود هم بسته

جهت آشنایی با آنالیز مؤلفه‌ی اصلی و شبکه‌ی خطی خودهم بسته، ابتدا مقدمه‌ی ای را به تفکیک برای هر کدام از آنها در حالت غیر فازی بیان می‌کنیم، سپس مفاهیم بیان شده را به حالتی که اطلاعات فازی باشند، گسترش می‌دهیم. سپس آنالیز مؤلفه‌ی اصلی با استفاده از شبکه‌ی عصبی خودهم بسته را برای اطلاعاتی با ورودی و خروجی فازی دوزنقه‌ی ای اجرا می‌نماییم.

۱.۱ آنالیز مؤلفه‌ی اصلی

در این بخش برخی از مفاهیم پایه‌ی ای را در مورد آنالیز مؤلفه‌ی اصلی (*PCA*) شرح می‌دهیم. برای اطلاعات بیشتر در این زمینه به [۶] مراجعه شود.

فرض کنید x_1, \dots, x_n ، بردار حقیقی d -بعدی باشند که مجموعه‌ی داده‌ها را تشکیل می‌دهند، با در نظر گرفتن اصل کلی، فرض می‌کنیم که اطلاعات مرکزی شده باشند. آنالیز مؤلفه‌ی اصلی می‌کوشد که زیر فضای q -بعدی L از \mathbb{R}^d را طوری پیدا کند که تصاویر متعامد $P_L x^p$ از این n نقطه روی L واریانس بیشینه داشته باشند. فرض کنید L بردار یکه باشد، تصویر $x \in \mathbb{R}^d$ روی L به صورت زیر می‌باشد:

$$P_L x = (u'x) u$$

که علامت پریم، ترانپاز را مشخص می‌کند.

واریانس این اطلاعات در جهت L می باشد، بنابراین

$$\begin{aligned} \frac{1}{n} \sum_{p=1}^n (u'x_p)^2 &= \frac{1}{n} \sum_{p=1}^n u'x_p x_p' u \\ &= u' \frac{1}{n} \left(\sum_{p=1}^n x_p x_p' \right) u = u' S u \end{aligned}$$

که S ماتریس کواریانس نمونه ای اطلاعات می باشد. بنابراین آنالیز مؤلفه ی اصلی، بردار u^* را طوری پیدا می کند که $u' S u$ را تحت قید $\|u\| = 1$ ماکزیم کند.

به آسانی می توان نشان داد که جواب، بردار ویژه ی نرمال شده ی u_1 از ماتریس S مربوط به بزرگترین مقدار ویژه λ_1 می باشد و

$$u_1' S u_1 = \lambda_1 u_1' u_1 = \lambda_1$$

در اینجا هیچ مشکلی برای گسترش دادن استدلال پیدا کردن زیر فضای q -بعدی L روی نقاط تصویر شده ی $P_L x_p$ که واریانس بیشینه دارند، وجود ندارد. برای سادگی محاسبات، فرض می کنیم که S پرتبه است و مقادیر ویژه ی آن به صورت $\lambda_1 > \lambda_2 > \dots > \lambda_d$ با بردارهای ویژه ی نرمال شده ی متناظر u_1, \dots, u_d باشند. آنگاه قضیه ی زیر برقرار می باشد:

قضیه ۰.۱.۱ [۶] در بین همه ی زیر فضاهای q -بعدی L از R^d ، آن زیر فضایی که بوسیله اولین q تا بردار ویژه u_1, \dots, u_q از S گسترش یافته، عبارت $\sum_{p=1}^n \|P_L x_p\|^2$ را بیشینه می کند. یا به طور معادل، عبارت $\sum_{p=1}^n \|x_p - P_L x_p\|^2$ را کمینه می کند.

□

اثبات. [۶]

خطوطی که به وسیله ی بردارهای ویژه u ساخته می شوند، محورهای اصلی اطلاعات می باشند

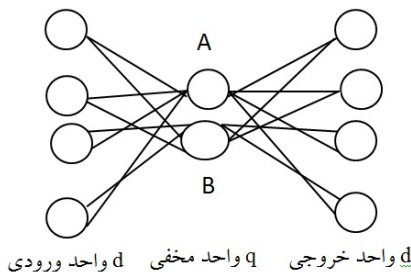
و $y_j = u'_j x$ یک تصویر q بعدی از بردار x در امتداد محورهای اصلی می باشد، که آنها را مؤلفه های اصلی می نامیم.

بردار y_p مؤلفه های اصلی برای هر بردار الگوی اصلی ($x_p, p = 1, 2, \dots, n$) در فرم ماتریسی به صورت $y_p = U'_q x_p$ بدست می آید، $U_q = [u_1, \dots, u_q]$ یک ماتریس $d \times q$ بعدی می باشد که q بردار ویژه ی نرمال شده ی S را به عنوان ستونهایش در نظر میگیریم. واریانس j امین مؤلفه ی اصلی به صورت زیر بدست می آید:

$$\frac{1}{n} \sum_{p=1}^n u'_j x_p x'_p u_j = u'_j S u_j = \lambda_j.$$

۲.۱ شبکه های خطی خود هم بسته

رابطه ی بین آنالیز مولفه ی اصلی و روش پرسپترون چند لایه ی خودهم بسته (IC)^۳ برای نخستین بار توسط بورلارد و کمپ (IC)^۴ [۱۰] مورد بررسی قرار گرفت. در شکل ۱ ساختار یک شبکه ی عصبی پیش خورانی سه لایه با یک لایه ی ورودی از d واحد، یک لایه ی مخفی با اندازه ی q و یک لایه ی خروجی از d واحد نشان داده شده است. فرض کنید A یک ماتریس $q \times d$ از وزن های ورودی به مخفی و B ماتریس $d \times q$ از وزنها ی



شکل ۱: ساختار یک شبکه ی عصبی

³ Autoassociative multilayer perceptrons

⁴ Bourlard and kamp

مخفی به خروجی باشد. واحدهای مخفی و خروجی طوری در نظر گرفته شده اند که توابع انتقال همانی دارند، بنابراین خروجی شبکه (z) از بردار ورودی (x) به صورت زیر محاسبه میشود:

$$z = BAx$$

همچنین فرض کنید این شبکه به طریقه ی خودهم بسته تعلیم داده شده است، یعنی ورودی ها را به عنوان خروجی های هدف به کار میگیرد و سپس نگاشت همانی را تقریب میزند. اگر $q < d$ باشد، چنین وظیفه ای به سیستم تحمیل خواهد شد که راه های مؤثری برای اختصار اطلاعات موجود در الگوهای ورودی پیدا کند. این شبکه به عنوان یک استخراج کننده ی خصوصیات بی سرپرست کار خواهد کرد.

کاربرد چنین طرحی برای اختصار اطلاعات و کاهش ابعاد در ابتدا به وسیله روملهارت (IC)^۵ [۱۰] و همکاران [۴۹] پیشنهاد شد. این موضوع به وسیله بورلارد و کمپ [۱۰] با بکارگیری مفهوم تجزیه ی مقدار منفرد ماتریس ها مورد بررسی قرار گرفت. بالدی و هرنیک (IC)^۶ [۸، ۹] توصیف کاملی از سطوح خطای شبکه های خطی چند لایه (شبکه های خودهم بسته با یک لایه ی مخفی) ارائه دادند.

خلاصه ای از یافته های بالدی و هرنیک [۸، ۹] به شرح زیر می باشد:

فرض کنید $E(A, B)$ تابع خطای درجه دوم باشد که به صورت زیر تعریف می شود:

$$E(A, B) = \sum_{p=1}^n e(x_p, z_p),$$

که n بردار d بعدی ورودی، $p = 1, 2, \dots, n$ بردار d بعدی z_p بردار d بعدی خروجی و $e(x_p, z_p)$ خطای مدل را برای الگوی p مشخص میکند.

$$e(x_p, z_p) = \|x_p - z_p\|^2 = \sum_{k=1}^d (x_{kp} - z_{kp})^2, \quad p = 1, \dots, n.$$

⁵ Rumelhart

⁶ Baldi and Hornik

خطای مجموع ممکن است به صورت تابعی از نگاشت سراسری $W = BA$ بیان شود، که برحسب بیشترین رتبه q مقید می شود. برای هر ماتریس $q \times q$ وارون پذیر C داریم:

$$E(A, B) = E(CA, BC^{-1})$$

بنابراین نگاشت سراسری W با یک خانواده Y نامتناهی از ماتریسهای وزن (A, B) متناظر است.

همانند قبل فرض کنید S ماتریس کواریانس نمونه ای اطلاعات باشد، همچنین فرض کنید S بر رتبه با مقادیر ویژه $\lambda_d > \dots > \lambda_2 > \lambda_1$ و بردارهای ویژه متناظر یک متعامد u_1, \dots, u_d باشد.

قضیه Y زیر توسط بالدی و هرنیک مطرح شده است:

قضیه ۲.۱. [۹]: خطای E بیان شده به صورت تابعی از نگاشت سراسری W یک مینیمم موضعی و فرا موضعی منحصر به فرد به شکل $W = BA$ دارد با

$$A = CU'_q \quad (۲)$$

$$B = U_q C^{-1} \quad (۳)$$

که U_q مانند قبل ماتریس $[u_1, \dots, u_q]$ و C یک ماتریس $q \times q$ وارون پذیر دلخواه هستند.

□

اثبات. [۹]

نگاشت بهینه $W = U_q U'_q$ بدین معنی است که تصویر متعامد P_L روی زیر فضای L به وسیله Y اولین بردارهای ویژه Y ماتریس کواریانس اطلاعات S گسترده شده است. آنالیز مؤلفه Y اصلی وقتی که C یک ماتریس همانی باشد به عنوان یک مورد خاص در نظر گرفته می شود. در این صورت، فعالیت ها در لایه Y مخفی کاملاً منطبق با مؤلفه های اصلی اطلاعات می باشند. اگر تابع خطا به وسیله Y یک الگوریتم تکراری مینیمم شود، راه حل ویژه Y به دست می آید، اگر چه

ماتریس C دلخواه می باشد. بردار فعالیت های واحد مخفی به صورت زیر است:

$$Ax = CU'_q x = Cy$$

که y بردار مؤلفه های اصلی برای ورودی x است. بنابراین فعالیت های واحد مخفی يك تبدیل خطی دلخواه از مؤلفه های اصلی هستند. اگر چه ممکن است به عنوان يك راه حل عمومی برخی مزیت ها از جمله استواری را از نقطه نظر اختصار اطلاعات داشته باشیم [۸]، اما نمایشهای درونی تولید شده مستقیماً قابل استفاده در آنالیز اطلاعات نیست، زیرا مقیاسگذاری فعالیت های لایه ی مخفی کاملاً اختیاری است. يك راه برای رفع این ابهام این است که قید $A' = B$ را معرفی کنیم، در این صورت از معادلات (۲) و (۳) نتیجه می شود:

$$U_q C' = U_q C^{-1}$$

چون (u_1, \dots, u_q) به شکل يك پایه ی یکه متعامد می باشد، داریم $U'_q U_q = I_q$ ، که I_q ماتریس همانی با بعد q است. با ضرب نمودن هر دو طرف معادله ی بالا از سمت چپ در U'_q ، به این نتیجه میرسیم که $C' = C^{-1}$ یا به طور معادل، $C' C = I_q$ است. از این رو، C ماتریس متعامدی است که دلالت دارد بر اینکه فعالیت های واحد مخفی و مؤلفه های اصلی به وسیله يك تبدیل متساوی (گروهی از تبدیلات هم اندازه، دوران ها و چرخش ها را شامل می شود) مرتبط شده اند. يك راه ساده برای تحمیل کردن قید $A' = B$ این است که معادله ی گسترش را به صورت زیر بازنویسی کنیم:

$$z = BB'x \quad (۴)$$

که در حالت يك بعدی به صورت زیر است:

$$z_k = \sum_{j=1}^q B_{kj} \sum_{i=1}^d B_{ij} x_i \quad k = 1, \dots, d.$$

۳.۱ گسترش به اطلاعات فازی

فرض کنیم که مجموعه ای از n موضوع توصیف شده به وسیله d صفت که مقادیری را در مجموعه $F(\mathbb{R})$ اعداد فازی حقیقی می گیرد، داریم. این اطلاعات به شکل $\tilde{x}_1, \dots, \tilde{x}_n$ هستند که هر $\tilde{x}_p \in F(\mathbb{R}^d)$ یک بردار از d عدد فازی می باشد که به صورت $(\tilde{x}_{ip})_{1 \leq i \leq d}$ مشخص می شود.

میخواهیم این اطلاعات را به اعداد فازی با ابعاد کوچکتری مثل $\tilde{y}_1, \dots, \tilde{y}_n$ با $\tilde{y}_p \in F(\mathbb{R}^d)$ که $p = 1, 2, \dots, n$ و $q < d$ خلاصه کنیم. حال این سوال مطرح می شود که چگونه میتوان چنین رهیافتی را در مورد اطلاعات فازی تعمیم داد؟ یک جواب مناسب برای این سوال ممکن است با پیاده سازی شبکه ی عصبی آنالیز مؤلفه ی اصلی که در بخش قبل شرح داده شد، پیدا شود.

یک بار دیگر شبکه ی سه لایه با واحدهای مخفی و خروجی خطی که در شکل ۱ نمایش داده شده را مورد ملاحظه قرار می دهیم و فرض می کنیم که یک بردار \tilde{x} از d عدد فازی به لایه ی ورودی خورانده شود. خروجی شبکه با بکارگیری اصل گسترش زاده در معادله ی (۴) محاسبه می شود. k امین مؤلفه ی بردار خروجی فازی \tilde{z}_k ، یعنی \tilde{z}_k برای ورودی فازی \tilde{x} به صورت زیر تعریف می شود:

$$\mu_{\tilde{z}_k}(u) = \sup_{v_1, \dots, v_d} \min_{1 \leq i \leq d} \mu_{\tilde{x}_i}(v_i), \quad \forall u \in \mathbb{R} \quad (5)$$

این سوپریم تحت قید زیر بدست می آید:

$$u = \sum_{j=1}^q B_{kj} \sum_{i=1}^d B_{ij} v_i. \quad (6)$$

را بطه (۵) به صورت زیر نوشته می شود:

$$\tilde{z}_k = \sum_{j=1}^q B_{kj} \sum_{i=1}^d B_{ij} \tilde{x}_i, \quad k = 1, 2, \dots, d.$$

محاسبه ی کاربردی \tilde{z}_k برای اعداد فازی ذوزنقه ای در بخش های بعدی شرح داده می شود. با استفاده از نمادسازی ماتریسی، معادله (۶) به طور طبیعی به صورت زیر نوشته می شود:

$$\tilde{z} = BB' \tilde{x}$$

که حالت فازی معادله ی (۴) است.

بردار فعالیت های واحد مخفی فازی را به وسیله \tilde{y} مشخص می نمایم که به صورت $\tilde{y} = B' \tilde{x}$ تعریف می شود. این بردار یک نمایش داخلی از الگوی ورودی \tilde{x} تشکیل می دهد. نمایش داخلی n بردار تعلیمی \tilde{x}_p به وسیله تعمیم دادن ضابطه ی خطای درمعادلات تابع خطای درجه دوم $E(A, B)$ داده شده، خروجی های فازی و مقادیر هدف فازی را به وجود می آورد. برای اطلاعات بیشتر در این زمینه مرجع [۳۰] و [۲۴] را مطالعه نمایید) فرض کنید $e(\tilde{x}, \tilde{z})$ به عنوان یک مقیاس از عدم تجانس بین اعداد فازی \tilde{x} و \tilde{z} باشد.

آنگاه، خطای مدل برای الگوی p به صورت زیر تعریف می شود:

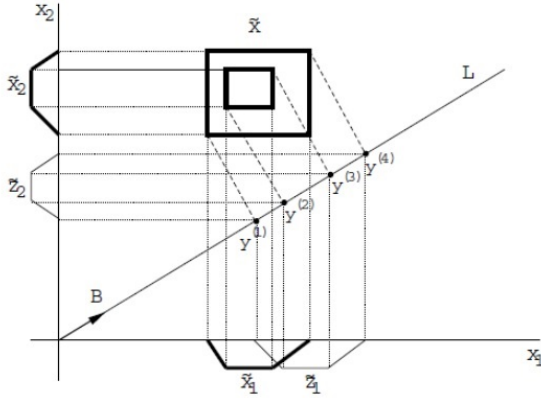
$$e(\tilde{x}, \tilde{z}) = \sum_{k=1}^d e(\tilde{x}_k, \tilde{z}_k),$$

و خطای مجموع روی مجموعه ی تعلیمی به صورت زیر می باشد:

$$E(B) = \sum_{k=1}^d e(\tilde{x}_p, \tilde{z}_p) = \sum_{p=1}^n \sum_{k=1}^d e(\tilde{x}_{kp}, \tilde{z}_{kp}). \quad (7)$$

با بکارگرفتن یک روش پارامتری کردن مناسب برای اعداد فازی، آنچه با مراجعه به عناصر B آشکار می شود این است که $E(B)$ غیر خطی می باشد. فرض کنید B^* یک راه حل برای مسئله ی بهینه سازی غیر خطی $\min_B E(B)$ باشد. این مسئله را می توان با بکارگیری یک روش مدرج تکراری استاندارد حل نمود. بردارهای خروجی لایه ی مخفی $\tilde{y}_p = B^* \tilde{x}_p$ -نمایش q -بعدی بهینه از بردارهای ورودی اصلی \tilde{x}_p را تشکیل می دهد، به این مفهوم که بردارهای خروجی لایه ی

مخفی یک تبدیل خطی از بردارهای ورودی اصلی با کمترین خطا می باشند.



شکل ۲: بیان هندسی آنالیز مؤلفه ی اصلی فازی در دو بعد با یک واحد مخفی

شکل ۲ نمای هندسی از یک شبکه ی خود هم بسته برای ورودیهای ۲-بعدی ($d = 2$) و یک واحد مخفی ($q = 1$) را نشان می دهد، در اینجا، ماتریس وزن B با ابعاد 1×2 است و به عنوان برداری از R^2 در نظر گرفته می شود. اعمال اجرا شده بر روی یک بردار ورودی فازی $\tilde{x} = (\tilde{x}_1, \tilde{x}_2)$ با مؤلفه های فازی دوزنقه ای را در نظر بگیرید. واحد مخفی، حاصل ضرب عددی بین \tilde{x} و B را محاسبه میکند، در حقیقت این واحد \tilde{x} را روی خط L گسترده شده به وسیله B تصویر می کند. نتیجه ی فعالیت واحد مخفی یک عدد فازی \tilde{y} است که α -برشهای آن برای هر $\alpha \in]0, 1]$ تصویری از مجموعه $(\tilde{x}_1)_\alpha \times (\tilde{x}_2)_\alpha$ روی L می باشد و $(\tilde{x}_1)_\alpha$ و $(\tilde{x}_2)_\alpha$ به ترتیب α -برشی از \tilde{x}_1 و \tilde{x}_2 را مشخص می کند. خروجی های شبکه از برگشت دادن تصویر \tilde{y} روی دو محور آغازین x_1 و x_2 بدست می آیند. مرحله یادگیری شامل تعیین بردار وزن B است که مقیاس اختلاف ($E(B)$) بین الگوهای اصلی \tilde{x}_p و بردارهای نوسازی شده \tilde{x}_p را مینیمم کند.

یادآوری ۱-۲: میتوان شبکه ی خودهم بسته ای را که قبلاً شرح داده شد، طوری تعمیم دهیم که وزن های فازی را بپذیرد. چنین شبکه عصبی فازی به وسیله ی مؤلفین مختلف درگرسیون فازی یا طبقه بندی پیشنهاد شده است (منابع [۳۷, ۳۶, ۱۱] را مشاهده نمایید).

روش های مقابله با هم خطی چندگانه در مدل های رگرسیون فازی ————— ۴۰

برای چنین شبکه ای تفسیر نمایش های ارائه شده در مورد فعالیت های واحد مخفی خیلی دشوار می باشد پس از جذابیت کاربرد چنین رهیافتی برای آنالیز اطلاعات اکتشافی کاسته می شود. به همین علت، دربرگیری لایه های مخفی اضافی و غیر خطی بودن که برای اختصار اطلاعات ممکن و مفید می باشد، در این قسمت مورد رسیدگی واقع نشده است.

یادآوری ۲-۲: یک راه ساده تر برای توسعه ی آنالیز مؤلفه ی اصلی برای اطلاعات فازی به صورت زیر می باشد:

(۱) غیرفازی نمودن اطلاعات (یعنی با تبدیل کردن هر عدد فازی به عدد غیر فازی [۳۰])

(۲) اجرا کردن آنالیز مؤلفه ی اصلی استاندارد روی این اعداد حقیقی

(۳) تصویر کردن اطلاعات فازی روی زیرفضای گسترده شده به وسیله ی اولین q بردار ویژه ی بدست آمده در مرحله ی قبلی

درواقع این اندیشه پیش تر از ”روش مرکزی” به وسیله ی کازس^۷ و همکاران [۱۲] در مورد اطلاعات با مقادیر بازه ای (فاصله ای) پیشنهاد شد.

یادآوری ۳-۲: در تابع خطای (۷)، همه ی جمله های خطا دارای وزن یکسان هستند. در مورد اعداد ناهمگن منطقی نیست که همه ی جمله های خطا وزن یکسانی داشته باشند، فرض کنید \tilde{x} عدد فازی باشد که α -برش های آن بازه های بسته ای به صورت $[(\tilde{x})_\alpha^L; (\tilde{x})_\alpha^U]$ باشند. طبق تعریف دلگادو^۸ [۱۹]، ابهام \tilde{x} یا $Amb(\tilde{x})$ را به صورت زیر تعریف میکنیم:

$$Amb(\tilde{x}) = \int_0^1 [(\tilde{x})_\alpha^U - (\tilde{x})_\alpha^L] d\alpha, \quad (۸)$$

حال فرض کنید $[0, 1] \rightarrow \mathbb{R} : \varphi$ یک تابع نزولی باشد. آنگاه تابع خطای $E(B)$ در معادله ی (۸) به صورت زیر در می آید [۲۰]:

$$E_\alpha(B) = \sum_{p=1}^n \sum_{k=1}^d \varphi[Amb(\tilde{x}_{kp})] e(\tilde{x}_{kp}, \tilde{z}_{kp}).$$

⁷Cazes

⁸ Delgado

۲ مؤلفه های اصلی برای مشاهدات دوزنقه ای

فرض کنید مؤلفه های بردار ورودی $\tilde{x} \in F(R^d)$ اعداد فازی دوزنقه ای به صورت $\tilde{x}_i = (x_i^{(l)}, x_i^{(1)}, x_i^{(2)}, x_i^{(r)})$, $i = 1, 2, \dots, d$ فازی نسبت به عمل جمع، تفریق و ضرب در یک عدد حقیقی، بسته است [۲۵]. در نتیجه خروجی های (\tilde{z}_k) این شبکه اعداد فازی دوزنقه ای هستند. نحوه ی محاسبه ی خروجی های شبکه و تابع خطا در مورد اعداد فازی دوزنقه ای در ادامه شرح داده می شود:

در ابتدا خروجی واحد مخفی j ام (\tilde{y}_j) را به صورت زیر تعریف میکنیم:

$$\tilde{y}_j = \sum_{i=1}^d B_{ij} \tilde{x}_i = (y_j^{(l)}, y_j^{(1)}, y_j^{(2)}, y_j^{(r)}), \quad j = 1, \dots, q.$$

با بکارگیری خواص اعداد فازی دوزنقه ای، \tilde{y}_j به صورت زیر محاسبه می شود:

$$y_j^{(l)} = \sum_{\substack{i=1 \\ B_{ij} > 0}}^d B_{ij} x_i^{(l)} + \sum_{\substack{i=1 \\ B_{ij} < 0}}^d B_{ij} x_i^{(r)}$$

$$y_j^{(1)} = \sum_{\substack{i=1 \\ B_{ij} > 0}}^d B_{ij} x_i^{(1)} + \sum_{\substack{i=1 \\ B_{ij} < 0}}^d B_{ij} x_i^{(2)}$$

$$y_j^{(2)} = \sum_{\substack{i=1 \\ B_{ij} > 0}}^d B_{ij} x_i^{(2)} + \sum_{\substack{i=1 \\ B_{ij} < 0}}^d B_{ij} x_i^{(1)}$$

$$y_j^{(r)} = \sum_{\substack{i=1 \\ B_{ij} > 0}}^d B_{ij} x_i^{(r)} + \sum_{\substack{i=1 \\ B_{ij} < 0}}^d B_{ij} x_i^{(l)}$$

به طور مشابه، k امین خروجی به صورت زیر تعریف می شود:

$$\tilde{z}_k = \sum_{j=1}^q B_{kj} \tilde{y}_j = (z_k^{(l)}, z_k^{(1)}, z_k^{(2)}, z_k^{(r)}), \quad k = 1, \dots, d,$$

بنابراین داریم :

$$z_k^{(l)} = \sum_{\substack{j=1 \\ B_{kj} > 0}}^q B_{kj} y_j^{(l)} + \sum_{\substack{j=1 \\ B_{kj} < 0}}^q B_{kj} y_j^{(r)}$$

$$z_k^{(1)} = \sum_{\substack{j=1 \\ B_{kj} > 0}}^q B_{kj} y_j^{(1)} + \sum_{\substack{j=1 \\ B_{kj} < 0}}^q B_{kj} y_j^{(2)}$$

$$z_k^{(2)} = \sum_{\substack{j=1 \\ B_{kj} > 0}}^q B_{kj} y_j^{(2)} + \sum_{\substack{j=1 \\ B_{kj} < 0}}^q B_{kj} y_j^{(1)}$$

$$z_k^{(r)} = \sum_{\substack{j=1 \\ B_{kj} > 0}}^q B_{kj} y_j^{(r)} + \sum_{\substack{j=1 \\ B_{kj} < 0}}^q B_{kj} y_j^{(l)}$$

قابل ذکر است بیان تحلیلی تابع خطا در [۲۰] با انتخاب یک متر مناسب میسر میشود. همچنین هر عدد فازی دوزنقه ای به عنوان یک نقطه در فضای اقلیدسی چهار بعدی مورد بررسی واقع می شود [۲۳، ۳۰].

خطای k امین مؤلفه (\tilde{x}_k) را به صورت زیر تعریف می کنیم:

$$e(\tilde{x}_k, \tilde{z}_k) = (z_k^{(l)} - x_k^{(l)})^2 + (z_k^{(1)} - x_k^{(1)})^2 + (z_k^{(2)} - x_k^{(2)})^2 + (z_k^{(r)} - x_k^{(r)})^2$$

و

$$E(B) = \sum_{p=1}^n \sum_{k=1}^d e(\tilde{x}_{kp}, \tilde{z}_{kp}).$$

مینیم کردن $E(B)$ با استفاده از روش کاهش گرادیان انجام می شود. الگوریتمی برای محاسبه ی مشتقات $E(B)$ در [۴۹] شرح داده شده است.

۳ مؤلفه های اصلی برای مشاهدات مثلثی

فرض کنید مؤلفه های بردار ورودی $\tilde{x} \in F(\mathbb{R}^d)$ اعداد فازی مثلثی به صورت $\tilde{x}_i = (x_i^{(l)}, x_i, x_i^{(r)})$, $i = 1, \dots, d$. در نتیجه خروجی های (\tilde{z}_k) این شبکه، اعداد فازی مثلثی می باشند. محاسبات خروجی های شبکه و تابع خطا در مورد اعداد فازی مثلثی، به صورت زیر می باشد: در ابتدا خروجی واحد مخفی j ام (\tilde{y}_j) را به صورت زیر تعریف میکنیم: $\tilde{y}_j = \sum_{i=1}^d B_{ij} \tilde{x}_i = (y_j^{(l)}, y_j^{(c)}, y_j^{(r)})$, $j = 1, 2, \dots, q$. با بکارگیری خواص اعداد فازی مثلثی، \tilde{y}_j به صورت زیر محاسبه می شود:

$$y_j^{(l)} = \sum_{\substack{i=1 \\ B_{ij} > 0}}^d B_{ij} x_i^{(l)} + \sum_{\substack{i=1 \\ B_{ij} < 0}}^d B_{ij} x_i^{(r)}$$

$$y_j = \sum_{\substack{i=1 \\ B_{ij} > 0}}^d B_{ij} x_i + \sum_{\substack{i=1 \\ B_{ij} < 0}}^d B_{ij} x_i = \sum_{i=1}^d B_{ij} x_i$$

$$y_j^{(r)} = \sum_{\substack{i=1 \\ B_{ij} > 0}}^d B_{ij} x_i^{(r)} + \sum_{\substack{i=1 \\ B_{ij} < 0}}^d B_{ij} x_i^{(l)}$$

همانطور که قبلاً گفتیم فعالیت های لایه ی مخفی، مؤلفه های اصلی فازی در نظر گرفته می شوند، در نتیجه \tilde{y}_j به عنوان مؤلفه ی اصلی شبکه ی عصبی خود هم بسته می باشد. به طور مشابه، k امین خروجی به صورت زیر تعریف می شود:

$$\tilde{z}_k = \sum_{j=1}^q B_{kj} \tilde{y}_j = (z_k^{(l)}, z_k^{(c)}, z_k^{(r)}), \quad k = 1, 2, \dots, d,$$

بنابراین داریم :

$$z_k^{(l)} = \sum_{\substack{j=1 \\ B_{kj} > 0}}^q B_{kj} y_j^{(l)} + \sum_{\substack{j=1 \\ B_{kj} < 0}}^q B_{kj} y_j^{(r)}$$

$$z_k = \sum_{\substack{j=1 \\ B_{kj} > 0}}^q B_{kj} y_j + \sum_{\substack{j=1 \\ B_{kj} < 0}}^q B_{kj} y_j = \sum_{j=1}^q B_{kj} y_j$$

$$z_k^{(r)} = \sum_{\substack{j=1 \\ B_{kj} > 0}}^q B_{kj} y_j^{(r)} + \sum_{\substack{j=1 \\ B_{kj} < 0}}^q B_{kj} y_j^{(l)}$$

خطای بازسازی شده ی k امین مؤلفه (\tilde{x}_k) را به صورت زیر تعریف میکنیم:

$$e(\tilde{x}_k, \tilde{z}_k) = (z_k^{(l)} - x_k^{(l)})^2 + (z_k - x_k)^2 + (z_k^{(r)} - x_k^{(r)})^2, \quad k = 1, \dots, d$$

و

$$E(B) = \sum_{p=1}^n \sum_{k=1}^d e(\tilde{x}_{kp}, \tilde{z}_{kp}).$$

۴ روش آنالیز مؤلفه ی اصلی مراکز فازی

یکی دیگر از روش های بدست آوردن مؤلفه های اصلی فازی روش آنالیز مؤلفه ی اصلی مراکز فازی می باشد. یک ماتریس فازی $\tilde{x} = [x^l, x^c, x^r]$ را در نظر بگیرید، ماتریس مراکز \tilde{x} به

صورت زیر می باشد:

$$x^c = \begin{pmatrix} x_{11}^c & \dots & x_{1p}^c \\ \vdots & \ddots & \vdots \\ x_{n1}^c & \dots & x_{np}^c \end{pmatrix}$$

و ماتریس پهناهای چپ و راست \bar{x} به ترتیب به صورت زیر نمایش داده می شوند [۳۵]:

$$x^a = \begin{pmatrix} x_{11}^a & \dots & x_{1p}^a \\ \vdots & \ddots & \vdots \\ x_{n1}^a & \dots & x_{np}^a \end{pmatrix}$$

و

$$x^b = \begin{pmatrix} x_{11}^b & \dots & x_{1p}^b \\ \vdots & \ddots & \vdots \\ x_{n1}^b & \dots & x_{np}^b \end{pmatrix}$$

که

$$x_{ij}^a = x_{ij}^c - x_{ij}^l, \quad x_{ij}^b = x_{ij}^r - x_{ij}^c, \quad i = 1, \dots, n, \quad j = 1, \dots, p.$$

برای بدست آوردن مؤلفه های اصلی فازی با استفاده از این روش ابتدا باید مقادیر ویژه و بردارهای ویژه ی هر يك از ماتریس های مراکز و پهناهای چپ و راست را به طور مجزا محاسبه کنیم. با استفاده از روش آنالیز مؤلفه اصلی که برای اعداد غیر فازی استفاده میشود، مقادیر ویژه و بردارهای ویژه را برای هر يك از این ماتریس ها به شیوه ی زیر محاسبه می نماییم:

۱- عناصر ماتریس را مرکزی می کنیم. یعنی:

$$(x_{ij} - \bar{x}_{0j}), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, p.$$

۲- ماتریس $\sum = X'X$ را محاسبه می نماییم.

۳- با استفاده از معادله ی زیر مقادیر ویژه را بدست می آوریم:

$$|\sum - \lambda I| = 0$$

۴- بردار ویژه ی متناظر با هر مقدار ویژه با استفاده از رابطه ی زیر بدست می آید:

$$\sum e_k = \lambda_k e_k, \quad k = 1, 2, \dots, q, \quad q \leq p.$$

مؤلفه های اصلی مراکز فازی به صورت زیر محاسبه می شوند:

$$z_k^c = e_{k1}^c x_1^c + \dots + e_{kp}^c x_p^c, \quad k = 1, 2, \dots, q, \quad q \leq p,$$

$$z_k^a = e_{k1}^a x_1^a + \dots + e_{kp}^a x_p^a, \quad k = 1, 2, \dots, q, \quad q \leq p,$$

$$z_k^b = e_{k1}^b x_1^b + \dots + e_{kp}^b x_p^b, \quad k = 1, 2, \dots, q, \quad q \leq p.$$

حال اگر مقدار z_k به دست آمده برای پهناها منفی باشد مقدار مثبت آن را در نظر می گیریم.

۵ مثال های عددی

برای آشنایی با روشهای اجرای رگرسیون مؤلفه ی اصلی به عنوان يك راه حل برای مواجهه با مشکل هم خطی چندگانه در مدل های رگرسیون فازی، در این فصل در مثال ۶-۱ داده های ورودی فازی (شامل قیمت انتشار روزنامه و کوپن) و داده های خروجی فازی (میزان فروش محصول در هر ماه) در ۱۲ شعبه در جدول ۶-۱ آورده شده است. هدف این است که به رابطه ی بین قیمت انتشار روزنامه، کوپن و میزان فروش محصول دست یابیم.

نخست جهت آشنایی با اجرای روش برنامه ریزی تعمیم یافته ی تاناکا برای مدل های رگرسیونی با ورودی فازی، ضرایب رگرسیونی را برآورد نموده سپس با استفاده از روش های شناسایی هم خطی چندگانه تشخیص می دهیم که مدل رگرسیونی بدست آمده برای این داده ها، درگیر چنین مشکلی می باشد یا خیر.

پس از تشخیص وجود مشکل هم خطی چندگانه در مدل رگرسیون فازی بدست آمده، در پی اجرای رگرسیون مؤلفه های اصلی فازی جهت رفع این مشکل هستیم. روشهای شبکه ی عصبی خودهم بسته و آنالیز مؤلفه ی اصلی مراکز فازی را جهت دستیابی به مؤلفه های اصلی به ترتیب در مثال

جدول ۱: داده های خام فازی مثال عددی ۱.۶

	فروش شعبه	d	z _۱	d	z _۲	d
۱	۱۰۴	۳	۴۷	۳	۲۲	۳
۲	۷۰	۳	۴۶	۳	۲۰	۳
۳	۹۰	۳	۳۷	۲	۱۳	۲
۴	۵۶	۳	۲۴	۱	۲	۱
۵	۸۴	۳	۴۳	۳	۱۷	۲
۶	۱۲۰	۳	۵۴	۴	۲۹	۳
۷	۶۲	۳	۳۵	۲	۷	۱
۸	۷۶	۳	۳۹	۲	۱۴	۲
۹	۶۶	۳	۳۱	۲	۶	۱
۱۰	۹۶	۳	۴۹	۳	۲۶	۳
۱۱	۷۰	۳	۴۵	۳	۱۹	۲
۱۲	۱۱۴	۳	۵۱	۴	۲۴	۳

دوم و سوم اجرا می نماییم و سپس در هر يك از این مثال ها با استفاده از روش برنامه ریزی تعمیم یافته ی تاناکا، مدل رگرسیونی را با استفاده از مؤلفه های اصلی حاصل بدست می آوریم و در انتها با استفاده از روش های شناسایی هم خطی چندگانه نشان می دهیم که آیا روش به کار گرفته شده مشکل هم خطی چندگانه را رفع کرده است یا خیر.

لازم به ذکر است، برنامه های این بخش با استفاده از نرم افزار آماری R نوشته شده، جهت برنامه نویسی با این نرم افزار از پکیج آلاباما^۹ استفاده شده است.

مثال ۱.۵. مجموعه ی اطلاعات فازی شامل قیمت انتشار روزنامه (\tilde{x}_1) ، کوپن (\tilde{x}_2) و مقدار فروش هر ماه (\tilde{y}) در ۱۲ شعبه در جدول ۱ نشان داده شده است. توجه داشته باشید نماد d پهنای هر متغیر فازی است. داده های فازی، به صورت اعداد فازی مثلثی متقارن در نظر گرفته می شوند. با فرض $L^{-1}(h) = \circ/۸$ ضرایب رگرسیونی و مدل رگرسیون فازی به صورت زیر بدست می آید:

$$\tilde{y} = (۱۲۶/۹۳۱۳, \circ) + (-۳/۱۲۸۹, \circ/۴۵۲۰)\tilde{x}_1 + (۵/۲۲۰۱, ۱/۲۳۷۶)\tilde{x}_2$$

^۹alabama

روش های مقابله با هم خطی چندگانه در مدل های رگرسیون فازی ————— ۴۸

سپس، جهت بررسی مناسب بودن مدل رگرسیونی فوق می توانیم ضرایب همبستگی فازی را به صورت زیر محاسبه کنیم. ضرایب همبستگی رگرسیونی به شرح زیر است:

$$R_{\bar{x}_1, \bar{y}} = [0/5098, 0/9555],$$

$$R_{\bar{x}_2, \bar{y}} = [0/5670, 0/9653],$$

$$R_{\bar{x}_1, \bar{x}_2} = [0/7442, 0/9999].$$

چون ضریب همبستگی بین \bar{y} و \bar{x}_1 مثبت می باشد، منفی بودن مرکز ضریب رگرسیون فازی \bar{x}_1 در مدل رگرسیون فازی فوق، غیر منطقی به نظر می رسد، همچنین ضریب همبستگی بزرگ بین \bar{x}_1 و \bar{x}_2 ($[0/7442, 0/9999]$) نشان دهنده ی وجود مشکل هم خطی چندگانه در مدل رگرسیون فازی است.

پس از تشخیص وجود مشکل هم خطی چندگانه در این مدل رگرسیونی، درصدد هستیم که با استفاده از روش های بیان شده در این مقاله این مشکل را رفع نماییم.

مثال ۲.۵. (پیاده سازی روش شبکه ی عصبی خودهم بسته) در جدول ۲ مؤلفه های اصلی فازی را به روش شبکه ی عصبی خود هم بسته برای داده های جدول ۱ با استفاده از روابط بالا محاسبه می نمایم که نتایج به صورت زیر است:

در ادامه با توجه به مؤلفه های اصلی فازی بدست آمده در این جدول، مانند قسمت قبل ضرایب رگرسیون فازی را برآورد می نمایم. سپس جهت بررسی مناسب بودن این روش برای مواجهه با مشکل هم خطی چندگانه، ضریب همبستگی بین \bar{z}_1 و \bar{z}_2 را محاسبه می کنیم، که نتایج به شرح زیر است:

$$\bar{y} = (135/6604, 0) + (-3/5389, 0/4469)\bar{z}_1 + (5/7271, 1/2382)\bar{z}_2$$

جدول ۲: مؤلفه های اصلی فازی بدست آمده به روش شبکه عصبی خودهمبسته

	فروش شعبه	d	z_1	d	z_2	d
۱	۱۰۴	۳	۴۷/۰۰۰۰۰۲	۳/۰۰۰۰۰۰۲	۲۱/۹۹۹۹۹۷	۳/۰۰۰۰۰۰۲
۲	۷۰	۳	۴۶/۰۰۰۰۰۱	۳/۰۰۰۰۰۰۲	۱۹/۹۹۹۹۹۷	۳/۰۰۰۰۰۰۲
۳	۹۰	۳	۳۷/۰۰۰۰۰۱	۲/۰۰۰۰۰۰۱	۱۲/۹۹۹۹۹۷	۲/۰۰۰۰۰۰۱
۴	۵۶	۳	۲۴/۰۰۰۰۰۰۱	۱/۰۰۰۰۰۰۱	۱/۹۹۹۹۹۸۳	۱/۰۰۰۰۰۰۱
۵	۸۴	۳	۴۳/۰۰۰۰۰۱	۳/۰۰۰۰۰۰۱	۱۶/۹۹۹۹۹۷	۲/۰۰۰۰۰۰۲
۶	۱۲۰	۳	۵۴/۰۰۰۰۰۲	۴/۰۰۰۰۰۰۲	۲۸/۹۹۹۹۹۶	۳/۰۰۰۰۰۰۳
۷	۶۲	۳	۳۵/۰۰۰۰۰۰۵	۲/۰۰۰۰۰۰۱	۶/۹۹۹۹۹۷۹	۱/۰۰۰۰۰۰۱
۸	۷۶	۳	۳۹/۰۰۰۰۰۱	۲/۰۰۰۰۰۰۱	۱۳/۹۹۹۹۹	۲/۰۰۰۰۰۰۱
۹	۶۶	۳	۳۱/۰۰۰۰۰۰۴	۲/۰۰۰۰۰۰۱	۵/۹۹۹۹۹۷۹	۱/۰۰۰۰۰۰۱
۱۰	۹۶	۳	۴۹/۰۰۰۰۰۰۲	۳/۰۰۰۰۰۰۲	۲۵/۹۹۹۹۹۷	۳/۰۰۰۰۰۰۲
۱۱	۷۰	۳	۴۵/۰۰۰۰۰۱	۳/۰۰۰۰۰۰۱	۱۸/۹۹۹۹۹۷	۲/۰۰۰۰۰۰۲
۱۲	۱۱۴	۳	۵۱/۰۰۰۰۰۲	۴/۰۰۰۰۰۰۲	۲۳/۹۹۹۹۹۶	۳/۰۰۰۰۰۰۳

همچنین ضریب همبستگی بین \tilde{z}_1 و \tilde{z}_2 به صورت زیر است:

$$\tilde{R}_{\tilde{z}_1, \tilde{z}_2} = [0/7442, 0/9999]$$

از ضریب همبستگی بدست آمده می توان فهمید که مؤلفه های اصلی فازی که از روش شبکه ی عصبی خودهم بسته بدست آمده اند، با یکدیگر هم بسته می باشند. در نتیجه، همچنان در مدل رگرسیونی مشکل هم خطی چندگانه وجود دارد و این روش برای رفع مشکل هم خطی چندگانه برای داده های جدول ۱ مناسب نمی باشد.

مثال ۳.۵. (پیاده سازی روش آنالیز مؤلفه ی اصلی مراکز فازی) در جدول ۳ مؤلفه های اصلی فازی برای داده های جدول ۱ به روش آنالیز مؤلفه ی اصلی مراکز فازی بدست آمده است. مدل رگرسیون فازی با استفاده از مؤلفه های اصلی بدست آمده به روش آنالیز مؤلفه ی اصلی مراکز فازی به صورت زیر است:

$$\tilde{y} = (79/8011, 24/8952) + (1/9236, 0)\tilde{z}_1 + (5/8379, 0)\tilde{z}_2$$

جدول ۳: مؤلفه های اصلی فازی با بکارگیری آنالیز مؤلفه ی اصلی مراکز فازی

	فروش شعبه	d	z_1	d	z_2	d
۱	۱۰۴	۳	۷/۵۳۸	۰/۸۱۱	۲۸۵.۰	۰/۳۴۸
۲	۷۰	۳	۴۳۳/۵	۸۱۱/۰	-۴۶۹/۰	۳۸۴/۰
۳	۹۰	۳	-۹۰۹/۵	۶۰۲/۰	۹۶۴/۰	۳۳۱/۰
۴	۵۶	۳	-۹۰۷۲۲	۲/۰۱۵	۱/۷۳۰	۰/۲۷۸
۵	۸۴	۳	۱/۱۹۱	۰/۱۳۱	-۰/۵۶۳	۰/۳۴۹
۶	۱۲۰	۳	۱۷/۴۳۵	۱/۵۴۴	۰/۵۰۵	۰/۲۹۶
۷	۶۲	۳	-۱۱/۵۰۲	۱/۲۸۲	-۲/۲۶۰	۰/۴۰۱
۸	۷۶	۳	-۳/۷۷۳	۰/۶۰۲	۰/۰۳۴	۰/۳۳۱
۹	۶۶	۳	-۱۵/۰۸۴	۱/۲۸۲	-۰/۲۱۷	۰/۴۰۱
۱۰	۹۶	۳	۱۱/۴۷۸	۰/۸۱۱	۱/۷۹۴	۰/۳۸۴
۱۱	۷۰	۳	۴/۰۱۹	۰/۱۳۱	-۰/۵۰۰	۰/۳۴۹
۱۲	۱۱۴	۳	۱۱/۸۱۱	۱/۵۴۴	-۱/۰۳۴	۰/۲۹۶

همچنین ضرایب همبستگی به شرح زیر است:

$$R_{\tilde{z}_1, \tilde{y}} = [۰/۷۰۶۹, ۰/۹۰۶۸],$$

$$R_{\tilde{z}_2, \tilde{y}} = [-۰/۲۴۰۵, ۰/۴۵۰۷],$$

$$R_{z_1, z_2} = [-۰/۳۳۴۶, ۰/۳۲۰۹].$$

با توجه به بازه ی ضریب همبستگی بین \tilde{z}_1 و \tilde{z}_2 میتوان نتیجه گرفت که مرکز این بازه تقریباً صفر است و این مطلب نشان دهنده ی این است که \tilde{z}_1 و \tilde{z}_2 ناهم بسته هستند. پس با این روش توانستیم مشکل هم خطی چندگانه را برای داده های جدول ۱ از بین ببریم. با استفاده از مدل رگرسیونی بدست آمده از مؤلفه های اصلی که به روش آنالیز مؤلفه ی اصلی مراکز فازی محاسبه شده اند، میتوان به نتایج زیر دست یافت: چون ضرایب رگرسیون فازی \tilde{z}_1 و \tilde{z}_2 در مدل رگرسیونی فازی مثال ۳.۶ مثبت است، لذا قیمت انتشار روزنامه و کوپن همواره اثرات مثبتی بر میزان فروش محصول خواهند داشت. همچنین مدل رگرسیونی بدست آمده معقول و منطقی به

نظر می رسد.

در نهایت، از نتایج آنالیز مؤلفه ی اصلی فازی می توان نتیجه گرفت که \tilde{x}_1 و \tilde{x}_2 اثرات مثبتی روی مقدار فروش محصول دارند.

۶ نتیجه گیری

در این مقاله روش شبکه ی عصبی خودهم بسته و روش آنالیز مؤلفه ی اصلی مراکز فازی برای تبدیل متغیرهای اصلی فازی به مؤلفه های اصلی فازی استفاده شد، سپس این مؤلفه های اصلی به عنوان متغیرهای ورودی فازی جدید در نظر گرفته شده و با استفاده از روش برنامه ریزی تعمیم یافته تاناکا رگرسیون می شوند. با استفاده از نتایج به دست آمده، نتیجه می گیریم که آنالیز مؤلفه ی اصلی مراکز فازی، مؤلفه های اصلی فازی دقیق تری را نسبت به روش شبکه ی عصبی خودهم بسته به وجود می آورد و بر مشکل هم خطی چندگانه در مدل رگرسیون فازی غلبه می نماید.

مراجع

- [۱] حسین یزدی، م. حسن پور، ح. و ماشین چی، م، آشنایی با حساب اعداد فازی، فرهنگ و اندیشه ریاضی، شماره ۳۸، (۱۳۸۶) ۴۹-۶۲.
- [۲] فرشادفر، ع، اصول و روش های پیشرفته آماری (تجزیه رگرسیون)، انتشارات طاق بستان (۱۳۸۰).
- [۳] ماشین چی، م، مجموعه های مشکک، انتشارات دانشگاه شهید باهنر کرمان (۱۳۷۹).
- [۴] مونت گمری، د.س. و پک، ا، مقدمه ای بر تحلیل رگرسیون خطی، مترجم: رضوی س.ا، انتشارات دانشگاه شهید باهنر کرمان، (۱۳۸۸).
- [5] Abdalla, A. and Buckley, J.J., Monte carlo methods in fuzzy regression, Soft Comput, 12(2008)463-468.

- [6] Anderson, T.W., An Introduction to Multivariate Statistical Analysis, Wiley, New-York, 1984.
- [7] Arabpour, A.R. and Tata, M., Estimating the parameters of a fuzzy linear regression model, Iranian Journal of Fuzzy Systems, 5(2)(2008)1-19.
- [8] Baldi, P. and Hornik, k., Learning in linear neural networks: A survey, IEEE Transaction on Neural Networks, 6(1995)837-858.
- [9] Baldi, P. and Hornik, K., Neural networks and principal component analysis: Learning from examples without local minima. Neural Networks, 2(1989)53-58.
- [10] Bourlard, H. and Kamp, Y., Auto-association by Multilayer Perceptrons and Singular Value Decomposition, Biological cybernetics, 59(1988)291-294.
- [11] Buckley, J.J. and Hayashi, Y., Fuzzy neural networks: A survey. Fuzzy sets and systems, 66(1994)1-13.
- [12] Cazes, P., Chouakria, A., Diday, E. and Schektman, Y., Extension de lanalyse en components principales a des donnees de type intervalle. Revue de Statistique Appliquee, 14(1997)5-24.
- [13] Celmins, A., Least squares model fitting to fuzzy vector data, Fuzzy Sets and Systems, 22(1987)245-269.
- [14] Celmins, A., Multidimensional least-squares fitting of fuzzy models, Mathematical Modeling, 9(1987)669-690.
- [15] Chang, P.T. and Lee, E.S., A generalized fuzzy weighted least squares regression, Fuzzy Sets and Systems, 82(1996)289-298.

- [16] Chang, P.T. and Lee, E.S., Fuzzy linear regression with spreads unrestricted in sign, *Computers and Mathematics with Applications*, 28(1994)61-70.
- [17] Chang, Y.H., Hybrid fuzzy least squares regression analysis and its reliability measures, *Fuzzy Sets and Systems*, 119(2001)225-246
- [18] Coppi, R., D'Urso, P., Giordani, p. and Santoro, A. least squares estimation of a linear regression model with L-R fuzzy response, *Computational Statistics and Data Analysis*, 51(2006)267-286.
- [19] Delgado, M., Vila, M.A. and Voxman, W., On a canonical representation of fuzzy numbers, *Fuzzy Sets and Systems*, 93(1998)125-135.
- [20] Denoeux, T. and Masson, Multidimensional scaling of interval-valued dissimilarity data. *Pattern Recognition Letters*, 21(2000)83-92.
- [21] Denoeux, T. and Masson, M., Principle Component Analysis of Fuzzy Data Using Autoassociative Neural Network, *IEEE Transaction on Fuzzy systems*, 12(2004),336-349.
- [22] Diamond, P., Fuzzy least squares, *Information Sciences*, 46(1988)141-157.
- [23] Diamond, p. and Tanaka, H., Fuzzy regression analysis. In Slowinski, R., editor, *Fuzzy sets in decision analysis, operations research and statistics*, Pages 349-387. Kluwer Academic Publishers, Boston, 1998.
- [24] Dubois, D., Kerre, E., Mesiar, R. and Prade, H., Fuzzy interval analysis. In Dubois, D. and Prade, H., editors, *Fundamentals of Fuzzy sets*, peage 483-581. Kluwer Academic Publishers, Boston, 2000.

- [25] Dubois, D. and Prade, H., Possibility Theory: An approach to computerized processing of uncertainty. Plenum Press, New-York, 1988.
- [26] Dubois, D. and Prade, H., Theory and applications: Fuzzy Sets and Systems, Academic press, New York,(1980).
- [27] D'Urso, P. and Gastaldi, T., A least squares approach to fuzzy linear regression analysis, Computational Statistics and Data Analysis, 34(2000)427-440.
- [28] D'Urso, P., Linear Regression Analysis for Input and Fuzzy/Crisp Output Data, Computational Statistics and Data Analysis, 42(2003)47-72.
- [29] Hassanpur, H., Malaki, H.R. and Yaghoobi, M.A., A goal programming approach for fuzzy linear regression with non-fuzzy input and fuzzy output data, Asia Pacific Journal of operational research, 26(2009)584-604.
- [30] Heilpern, S., Representation and application of fuzzy numbers, Fuzzy Sets and Systems, 91(1997)259-268
- [31] Hojati, M., Bector, C.R. and Smimou, K., A simple method for computation of fuzzy linear regression, European Journal of Operational Research, 166(2005)172-184.
- [32] Hong, D.H., Do, H.Y, Song, J., Fuzzy least-squares linear regression analysis Using Shape-Preserving Operations, Information Sciences, 138(2001)185-193.
- [33] Hong, D.H., Do, H.Y., Fuzzy System reliability analysis by the Use of (the weakest t-norm) on fuzzy number arithmetic operations, Fuzzy Sets and Systems, 90(1997)307-316.

- [34] Hong, D.H., Lee, S., and Do, H.Y., Fuzzy linear regression analysis for Fuzzy Data Using Shape-Preserving Operations, *Fuzzy Sets and Systems*, (2001),513-526.
- [35] Huang, J.J., Tzeng, G.H. and Ong, C.S., Fuzzy Principal Component Regression (FPCR) For Fuzzy Input and Output Data, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 14, 1(2006)87-100.
- [36] Ishibuchi, H., Kwon, K. and Tanaka, H., A Learning algorithm of fuzzy neural networks with triangular fuzzy weights, *Fuzzy Sets and Systems*, 71(1995)277-293.
- [37] Ishibuchi, H., Marioka, K., and Turksen, I.B., Learning by fuzzified neural networks. *International Journal of Approximate Reasoning*, 13(1995)327-358.
- [38] Jajuga, K., Linear fuzzy regression, *Fuzzy Sets and Systems*, 20(1986)343-453.
- [39] Kao, C. and Chyu, C., A fuzzy linear regression model with better explanatory power, *Fuzzy Sets and Systems*, 126(2002)401-409.
- [40] Kim, B. and Bishu, R.R., Evaluation of fuzzy linear regression models by comparing membership functions, *Fuzzy Sets and Systems*, 100(1988)343-353.
- [41] Kim, K.J., Modkowitz, H., Koksalan, M., Fuzzy Versus Statistical Linear Regression, *European Journal of Operational Research*, 92(1996)417-434.
- [42] Kolesa rova, A., Additive preserving the linearity of fuzzy intervals, *Tatra Mountains math. Publ.* 6(1995)75-81.
- [43] Ling, C.H., Representation of associative functions, *Publ. Math. Debrecen*, 12(1965)189-212.

- [44] Ma, M., Freedman, M. and Kandel, A., General fuzzy least squares, *Fuzzy Sets and Systems*, 88(1977)107-118.
- [45] Mesiar, R., Shape preserving addition of fuzzy intervals, *Fuzzy Sets and Systems*, 86(1997)73-78.
- [46] Neguyen, H.T. and Walker, E.A., *A first course in fuzzy logic*, Department of Mathematical Science, third edition. Chapman and Hall, London(2006).
- [47] Peters, G., Fuzzy linear regression with fuzzy intervals, *Fuzzy Sets and Systems*, 63(1994)45-55.
- [48] Redden, D.T., Woodall, W.H., Properties of certain fuzzy linear regression methods, *Fuzzy Sets and Systems*, 64(1994)361-375.
- [49] Rumelhart, D.E., Hinton, G.E. and Williams, R.J., Learning internal representations by error propagation. In Rumelhart, D.E., McClelland, J.L. and PDP Research Group, editors, *Parallel Distributed Processing: Explorations in the microstructure of Cognition*, 1(1988)318-362.
- [50] Sakawa, M. and Yano, H., Multiobjective fuzzy linear regression analysis for fuzzy input-output data, *Fuzzy Sets and Systems*, 47(1992)173-181.
- [51] Sanchez, J.D.A. and Gomez, A.T., Estimating a fuzzy term structure of interest rates using fuzzy regression techniques, *European Journal of Operational Research*, 154(2004)804-818.
- [52] Tanaka, H., Fuzzy data analysis by Possibilistic linear model, *Fuzzy Sets and Systems*, 24(1987)363-375.

- [53] Tanaka, H., Hayashi, I. and Watada, J., possibilistic linear regression analysis for fuzzy data, *European Journal of Operational Research*, 40(1989)389-396.
- [54] Tanaka, H., Uegimisis, S. and Asai, K., Linear regression analysis with fuzzy model, *IEEE Transaction on Systems Man and Cybernetics*, 12(1982)903-907.
- [55] Tanaka, H., and Watada, J., Possibilistic systems and their application to the linear regression model, *Fuzzy Sets and Systems*, 27(1988)275-289.
- [56] Valentine, D., Abdi, H., O'Toole, A.J. and Cottrell, G.W., Connectionist models of face Processing: A survey. *Pattern Recognition*, 27(1994)1209-1230
- [57] Wang, H.F., and Tsuar R.C., Insight of a fuzzy regression model, *Fuzzy Sets and Systems*, 112(2000)355-369