Archive of SID

دسترسی در سایتِ http://jnrm.srbiau.ac.ir
سال چهارم، شماره شانزدهم، زمستان ۱۳۹۷
شماره شاپا: ۲۵۸۸-۵۸۸X

پژوهش‌های نوین در ریاضی

دانشگاه آزاد اسلامی، واحد علوم و تحقیقات

یک روش ماتریسی جهت برآورد ضرایب رگرسیون خطی مبتنی بر اعداد-زاده

سمیه ایزدی[۱]، توفیق الهویرنلو[*۲]

[(۱)] گروه ریاضی، دانشکده علوم پایه، واحد همدان، دانشگاه آزاد اسلامی، همدان، ایران
[(۲)] گروه ریاضی کاربردی، دانشکده علوم پایه، واحد علوم تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

تاریخ ارسال مقاله: ۹۷/۰۶/۲۷ تاریخ پذیرش مقاله: ۹۷/۰۷/۲۲

چکیده

در این مقاله، یک روش جدید جهت برآورد تقریب ضرایب رگرسیون خطی مبتنی بر اعداد- زاده ارائه می‌شود. در این مدل مشاهدات اعداد حقیقی هستند و ضرایب رگرسیون و متغیر وابسته (y) مقادیری با ارزشگذاری اعداد-زاده هستند. جهت برآورد ضرایب این مدل ابتدا مدل رگرسیون خطی مبتنی بر اعداد-زاده را به دو مدل رگرسیون خطی فازی تبدیل کرده سپس دو مدل را به صورت دستگاه Ax=y تبدیل می‌کنیم که در آن A ضریب رگرسیون خطی و x متغیر مستقل و y متغیر وابسته است. در نهایت برای مینیمم‌سازی دستگاه فوق از روش مجموع مربعات خطا مبتنی بر فاصله d استفاده می‌کنیم. طی دو مثال نشان می‌دهیم که روش پیشنهادی از دقت مطلوبی جهت برآورد رگرسیون بر اساس اعداد فازی برخوردار است.

واژه‌های کلیدی: اعداد- زاده، رگرسیون خطی مبتنی بر اعداد-زاده، روش ماتریسی.

*. عهده‌دار مکاتبات: Email: tofigh@allahviranloo.com

www.SID.ir

Archive of SID

۱- مقدمه

فرم کلی رگرسیون خطی را در نظر بگیرید.

$$Y_k = A_0 + A_1 x_1 + \cdots + A_n x_n \qquad (۱)$$

که در آن مشاهدات، ضرایب و نتیجه مشاهدات حقیقی هستند. اما با این حال، در بسیاری از شرایط واقعی زندگی، که در آن پیچیدگی سیستم فیزیکی دیکته شده است اتخاذ یک دیدگاه کلی‌تر لازم است که در آن متغیرهای رگرسیون به عنوان سازمان‌های غیر عددی مانند متغیرهای زبانی داده شده‌اند [۱]. متأسفانه، از جمله شرایط واقعی زندگی اغلب در خارج از محدوده از تجزیه و تحلیل رگرسیون کلاسیک هستند [۲و۳] پس از معرفی مفهوم مجموعه های فازی توسط زاده در سال ۱۹۶۵ [۶-۴] محققان مختلف اقدام به گسترش تجزیه و تحلیل رگرسیون پرداختند. رگرسیون خطی فازی (FLR) برای اولین بار توسط تاناکا و همکاران پیشنهاد شد [۷]. فرمی از رگرسیون خطی فازی که در آن مشاهدات حقیقی هستند، ضرایب و نتیجه مشاهدات در محیطی مبهم و نامعلوم به صورت زیر است.

$$\tilde{y}_k = \tilde{A}_{k0} + \tilde{A}_{k1} x_1 + \cdots + \tilde{A}_{kn} x_n \qquad (۲)$$

روش‌های مختلفی جهت این گونه از مسائل ارائه شده است [۱۲-۸]. اما برای اینکه این اطلاعات مفید باشند باید قابل اعتماد باشند. انسان‌ها دارای یک ظرفیت آشکار برای اخذ تصمیمات منطقی بر پایه اطلاعات مبهم، غیر دقیق و یا ناکامل می‌باشند. فرمالیزاسیون این ظرفیت حداقل تا درجاتی چالشی است که به سختی برآورد می‌شود. زاده موضوعی را به نام عدد-زاده پیشنهاد کرد که یک جفت از اعداد فازی $(\tilde{A}, \tilde{B})$ می‌باشد [۱۳]. اولین جزء، $\tilde{A}$، یک محدودیت برای مقادیری است که متغیر نامشخص X که دارای ارزش یا مقدار حقیقی است می‌تواند داشته باشد. جزء دوم یعنی $\tilde{B}$ مقیاسی از قابلیت اطمینان مربوط به اولین جزء است. معمولاً A و B با زبان طبیعی توصیف می‌شوند. مثلاً: (حدود ۴۵ دقیقه، بسیار مطمئن است). در سال‌های اخیر تحقیقات متعددی در خصوص مسائل مبتنی بر اعداد -زاده و همچنین کاربرد این اعداد صورت گرفته است [۲۳-۱۴]. آنالیز رگرسیون خطی مبتنی بر اعداد- زاده (ZLR) بسطی از آنالیز رگرسیون فازی است که برخی از عنصرها در این مدل که با اطلاعات مبهم نشان داده شده‌اند دارای درجه اطمینان هستند که در کار قبلی [۱۴] این نوع رگرسیون را معرفی کرده‌ایم.

فرم کلی رگرسیون خطی مبتنی بر اعداد زاده را به صورت زیر در نظر بگیرید.

$$[Y_i]^z = (A_Y, B_Y), \qquad (۳)$$

که در آن $[.]^z$ نماد یک مقدار مبتنی بر اعداد-زاده است. A_Y قسمت محدودیت مقدار Y_i و B_Y قسمت احتمال (احتمال این که Y_i ،A_Y است) می‌باشد. رابطه (۳) را می‌توان به فرم زیر هم نوشت

$$[Y_i]^z = [A_0]^z + [A_1]^z x_{i1} + \cdots + [A_n]^z x_{in} \quad (۴)$$

که در آن $x_{i1} \in R$ و $[A_0]^z, [A_1]^z, [A_2]^z, \dots, [A_n]^z$ و $[Y_i]^z$ اعداد-زاده هستند.

فرض کنید $[A]^z$ و $[y]^z$ به ترتیب به صورت $[A]^z = (\tilde{A}_k, \tilde{A}_H)$ و $[y]^z = (\tilde{y}_k, \tilde{y}_H)$ باشند که $\tilde{y}_k, \tilde{y}_H, \tilde{A}_k, \tilde{A}_H$ اعداد فازی هستند. $\tilde{A}_k$ و $\tilde{y}_k$ نقش محدودیت و $\tilde{A}_H$ و $\tilde{y}_H$ نقش درجه اطمینان محدودیت‌ها را به ترتیب برای $[A]^z$ و $[y]^z$ ایفاد می‌کنند. بحث را به حالتی که $\tilde{y}_k, \tilde{y}_H, \tilde{A}_k, \tilde{A}_H$ اعداد فازی مثلثی متقارن باشند محدود می‌کنیم. در کار قبلی حالت‌های مختلفی از این نوع رگرسیون را مطرح کردیم و برای هر حالت روشی جهت برآورد مقدار آن ارائه شد. اما هیچ یکی از روش‌ها دارای جواب یکتا نبودند در این مقاله برای حالتی از مسئله که در آن مشاهدات حقیقی هستند اما ضرایب و نتیجه مشاهدات نامعلوم و به شکل ارزش گذاری-Z می‌باشند روش ماتریسی ارائه می‌شود که نشان می‌دهیم با استفاده از این روش یک جواب یکتا برای مسئله رگرسیون خطی مبتنی بر اعداد- زاده بدست می‌آید. این روش ابتدا توسط سو ولی [۱۰] جهت برآورد مقدار رگرسیون خطی فازی ارائه شده است سپس محمدی و طاهری [۹] روش سو و لی را تعمیم دادند تا نواقص روش آنها را پوشش دهند. که ما نیز از این روش ماتریسی جهت برآورد مقدار رگرسیون خطی

www.SID.ir

Archive of SID

مبتنی بر اعداد –زاده استفاده می‌کنیم. این مقاله به شرح زیر است. در بخش ۲، مفاهیم پایه و ضروری بیان شده است. در بخش ۳، برآورد رگرسیون خطی مبتنی بر اعداد–زاده با استفاده از ماتریس بیان شده است. در بخش ۴، مثال عددی آمده است. در بخش ۵، نتیجه‌گیری بیان شده است. در بخش ۶، مراجع آمده است.

۲– مفاهیم پایه و ضروری

تعریف ۱،۲. فرم پارامتری یک عدد فازی را با یک زوج مرتب از توابع به صورت $(a_1(\alpha), b_1(\alpha))$ که $0 \le \alpha \le 1$ معرفی می‌شود و شرایط زیر برقرار است

۱) $a_1(\alpha)$ یک تابع صعودی کراندار و از راست روی بازه $[0, 1]$ پیوسته است.

۲) $b_1(\alpha)$ یک تابع نزولی کراندار و از چپ روی بازه $[0, 1]$ پیوسته است.

تعریف ۲،۲. اگر A و B اعداد فازی و به صورت:

$$[A]_\alpha = [a_1(\alpha), a_2(\alpha)]$$

و

$$[B]_\alpha = [b_1(\alpha), b_2(\alpha)]$$

و $\alpha \in [0, 1]$ باشند، در این صورت اعمال فازی بین آنها این گونه تعریف می‌شوند [۱۱]:

$$[A + B]_\alpha = [a_1(\alpha) + b_1(\alpha), a_2(\alpha) + b_2(\alpha)]$$
$$[-A]_\alpha = [-a_2(\alpha), -a_1(\alpha)]$$
$$[A - B]_\alpha = [a_1(\alpha) - b_2(\alpha), a_2(\alpha) - b_1(\alpha)]$$
$$[\lambda A]_\alpha = [\lambda a_1(\alpha), \lambda a_2(\alpha)], \qquad \lambda \ge 0$$
$$[\lambda A]_\alpha = [\lambda a_2(\alpha), \lambda a_1(\alpha)], \qquad \lambda < 0$$

تعریف ۳،۲. فاصله بین دو عدد فازی $\widetilde{A}$ و $\widetilde{B}$ بر پایه‌ی تابع وزنی $f(\alpha)$ به صورت زیر تعریف می‌شود.

$$d(\widetilde{A}, \widetilde{B}) = \left[\int_0^1 f(\alpha) d^2(\widetilde{A}_\alpha, \widetilde{B}_\alpha) d_\alpha\right]^{\frac{1}{2}}$$
$$d^2(\widetilde{A}_\alpha, \widetilde{B}_\alpha) = [a_1(\alpha) - b_1(\alpha)]^2 + [a_2(\alpha) - b_2(\alpha)]^2$$

که در آن $\widetilde{A}$ و $\widetilde{B}$ دو عدد فازی هستند،

$$\widetilde{A}_\alpha = [a_1(\alpha), a_2(\alpha)]$$

و

$$\widetilde{B}_\alpha = [b_1(\alpha), b_2(\alpha)]$$

به ترتیب α–برش‌های $\widetilde{A}$ و $\widetilde{B}$ هستند. $f(\alpha)$ یک تابع صعودی روی بازه [۰،۱] است که برای آن $f(0) = 0$ و $\int_0^1 f(\alpha) d\alpha = \frac{1}{2}$ است. مقدار $d(\widetilde{A}_\alpha, \widetilde{B}_\alpha)$ فاصله بین α–برش اعداد فازی از $\widetilde{A}$ و $\widetilde{B}$ را اندازه می‌گیرد. می‌توان تابع $f(\alpha)$ را به عنوان وزن $d^2(\widetilde{A}_\alpha, \widetilde{B}_\alpha)$ تعبیر کرد.

تعریف ۴،۲. عدد فازی $\widetilde{A}$ یک عدد فازی مثلثی است اگر تابع عضویت آن به صورت زیر باشد:

$$\widetilde{A}(x) = \begin{cases} \dfrac{x - (a - s_a^L)}{s_a^L} & a - s_a^L \le x \le a \\ \dfrac{(a + s_a^R) - x}{R} & a \le x \le a + s_a^R \\ 0 & \text{otherwise} \end{cases}$$

می‌نویسیم $\widetilde{A} = (a, s_a^L, s_a^R)_T$، a مقدار $\widetilde{A}$ است. s_a^L و s_a^R به ترتیب پنهای چپ و راست. در حالت خاص، اگر $s_a^L = s_a^R$ باشد آنگاه $\widetilde{A}$ عدد فازی متقارن نامیده می‌شود و می‌نویسیم $\widetilde{A} = (a, s_a)$.

قضیه ۵،۲. فرض کنید $\widetilde{A} = (a, s_a)_T$ و $\widetilde{B} = (b, s_b)_T$ دو عدد فازی متقارن باشند. آنگاه بر پایه‌ی تابع وزنی $f(\alpha) = \alpha$،

$$d^2(\widetilde{A}, \widetilde{B}) = (a - b)^2 + \frac{1}{6}(s_a - s_b)^2$$

اثبات. [۹].

تعریف ۶،۲ تعریف عدد–زاده

در رفرنس [۹]، زاده ایده عدد–زاده همراه با یک متغیر نامعلوم X معرفی کرد. یک عدد–زاده یک زوج مرتب از اعداد فازی (A, B) است. در اینجا A یک زیرمجموعه فازی از محدودیت‌هایی است که مقادیر x می‌تواند داشته باشد و B یک زیر مجموعه فازی از مقیاس اطمینان مولفه A است. زاده (x, A, B) را به عنوان یک ارزشگذاری– Z معرفی کرد و نشان داد این مقدار معادل با این است که x برابر است با (A, B). در اینجا Z اطلاعات راجع به مقدار متغیر x را فراهم می‌آورد.

مثال‌هایی از این ارزشگذاری z به صورت زیر است.

www.SID.ir

Archive of SID

مثال: (حدود ۴۵ دقیقه، خیلی مطمئن است)، (حدود ۳۰ دقیقه، مطمئن است).
این ارزشگذاری برای Z طبق پیشنهاد زاده به عنوان یک محدودیت در x مشاهده و به صورت زیر تفسیر گردید.

$$\text{Prob}\,(x \text{ is } A) \text{ is } B$$

در واقع بدان معنی است که

$$R(y):\ y \ \ is \ \ A \rightarrow Poss(y=u)=\mu_A(u)$$
$$P(y \ \ is \ \ A)=\int_R \mu_A(u) P_y(u) du \ \ is \ \ B$$

که μ_A تابع عضویت مجموعه فازی A و u مقداری از y است. $P_y(u)$ تابع چگالی احتمال y و $P(y=u)$ تابع احتمال y است. در جاییکه توزیع احتمال پایه را نمیدانیم از این اطلاعات واضح است که توزیع احتمال خود عدد فازی است.
تعریف ۷،۲. تعریف. فرض کنید $\tilde{Z}_1=(\tilde{A},\tilde{B})$ و $\tilde{Z}_2=(\tilde{A}',\tilde{B}')$ دو عدد زی–نامبری باشند بطوریکه

$$\tilde{A}=(a,s_a)_T\,,\ \ \tilde{A}'=(a',s'_a)_T$$

و

$$\tilde{B}=(b,s_b)_T,\ \ \tilde{B}'=(b',s'_b)_T$$

فازی متقارن باشند. آنگاه بر پایه‌ی تابع وزنی $f(\alpha)=\alpha$، [۱۴]

$$d^2(A,A')=(a-a')^2+\frac{1}{6}(s_a-s'_a)^2$$
$$d^2(B,B')=(b-b')^2+\frac{1}{6}(s_b-s'_b)^2$$

قضیه ۸،۲. فرض x یک ماتریس باشد و $A=x'x$ که در آن x' ترانهاده ماتریس x باشد. اگر رتبه ماتریس x برابر با $n+1$ باشد آنگاه ماتریس A یک ماتریس معین مثبت است [۱۰].
با توجه به مدل رگرسیون خطی فازی در رابطه ۲،۱ با توجه به قضیه ۵،۲ و قضیه ۸،۲ قضیه زیر را داریم.
قضیه ۹،۲. [۱۰] اگر رتبه ماتریس x برابر با $n+1$ باشد، $A^{-1}S \geq 0$ آن گاه مسئله مینیمم سازی

$$\begin{cases} Aa=y \\ A\sigma=S \end{cases}$$

دارای جواب یکتا است که از رابطه

$$\begin{cases} a=A^{-1}y \\ \sigma=A^{-1}S \end{cases}$$

بدست می‌آید (شرط $A^{-1}S \geq 0$ در قضیه ۹،۲ تضمین می‌کند که مقادیر σها که پهنای ضرایب فازی مدل هستند، نامنفی باشند).

۳. برآورد رگرسیون خطی مبتنی بر اعداد زاده (رگرسیون زی–نامبری (ZLR))

برای برآورد رگرسیون خطی زی–نامبری بر اساس تعریف ۲–۷ ابتدا مسئله را به دو قسمت تقسیم می‌کنیم که قسمت اول بر اساس محدودیت و قسمت دوم بر اساس مقیاس اطمینان قسمت اول است. لذا رابطه (۴) را می‌توان به فرم دو رابطه (۵) و (۶) نوشت. یعنی

$$\tilde{y}_k=\tilde{A}_{k0}+\tilde{A}_{k1}x_1+\cdots+\tilde{A}_{kn}x_n \qquad (۵)$$
$$(\tilde{y}_k,x_{i1},x_{i2},\ldots,x_{in});\quad x_{ij}\in R,$$
$$i=1,\ldots,m,\quad j=1,\ldots,n.$$

و

$$\tilde{y}_H=\tilde{A}_{H0}+\tilde{A}_{H1}x_1+\cdots+\tilde{A}_{Hn}x_n \qquad (۶)$$
$$(\tilde{y}_H,x_{i1},x_{i2},\ldots,x_{in});\quad x_{ij}\in R,$$
$$i=1,\ldots,m,\quad j=1,\ldots,n.$$

برای برآورد رگرسیون ZLR که به صورت‌های فوق تجزیه شده است از روش مجموع مربعات خطا استفاده می‌شود [۱۰] که این کار را با توجه به فاصله d بیان شده در تعریف ۳،۲ می‌توان نوشت

$$SEE(Z_0,Z_1,\ldots,Z_n)= \qquad (۷)$$
$$=\begin{cases} SEE(\tilde{A}_0,\tilde{A}_1,\ldots,\tilde{A}_n)=\sum_{i=1}^m d^2\left(\tilde{Y}_{Ai},\tilde{y}_{Ai}\right) \\ \quad =\sum_{i=1}^m d^2\left(\tilde{A}_0+\tilde{A}_1x_{i1}+\cdots+\tilde{A}_nx_{in},\tilde{y}_{Ai}\right) \\ SEE(\tilde{B}_0,\tilde{B}_1,\ldots,\tilde{B}_n)=\sum_{i=1}^m d^2\left(\tilde{Y}_{Bi},\tilde{y}_{Bi}\right) \\ \quad =\sum_{i=1}^m d^2\left(\tilde{B}_0+\tilde{B}_1x_{i1}+\cdots+\tilde{B}_nx_{in},\tilde{y}_{Bi}\right) \end{cases}$$

از طرفی فرمول رگرسیون خطی مبتنی بر اعداد زاده که

www.SID.ir

Archive of SID

ضرایب آن مقادیری زی-نامبری دارند را می‌توان به صورت زیر هم نوشت

$$Z_0+Z_1x_{i1}+\cdots+Z_nx_{in}= \\ (\tilde{A}_0,\tilde{B}_0)+\cdots+(\tilde{A}_n,\tilde{B}_n)x_{in}= \\ \left((\tilde{A}_0+\cdots+\tilde{A}_nx_{in}),\ (\tilde{B}_0+\cdots+\tilde{B}_nx_{in})\right) \quad (۸)$$

اگر $\tilde{A}_i=(a_j,\sigma_j)$ ، $\tilde{B}_i=(b_j,\delta_j)$ ، $\tilde{y}_{Ai}=(y_{Ai},S_{Ai})$ و $\tilde{y}_{Bi}=(y_{Bi},S_{Bi})$ باشند. رابطه (۷) به صورت زیر در می‌آید.

$$\begin{cases} Minimize\ SEE(\tilde{A}_0,\tilde{A}_1,\ldots,\tilde{A}_n) \\ =\sum_{i=1}^{m}\ (a_0+\cdots+a_nx_{in}-y_{Ai})^2 \\ +\frac{1}{6}\sum_{i=1}^{m}(\sigma_0+\cdots+\sigma_nx_{in}-S_{Ai})^2 \\ \\ Minimize\ \ SEE(\tilde{B}_0,\tilde{B}_1,\ldots,\tilde{B}_n) \\ =\sum_{i=1}^{m}\ (b_0+\cdots+b_nx_{in}-y_{Bi})^2 \\ +\frac{1}{6}\sum_{i=1}^{m}(\delta_0+\cdots+\delta_nx_{in}-S_{Bi})^2 \end{cases} \quad (۹)$$

برای مینیمم سازی عبارت بالا، از اولین تابع بیان شده در رابطه (۹) نسبت به a_j و σ_j ، از دومین تابع بیان شده در رابطه (۹) نسبت به b_j و δ_j به ازای $j=0,1,\ldots,n$ مشتق می‌گیریم. از برابر صفر قرار دادن عبارت‌های حاصله، معادلات زیر، به ازای $j=0,1,\ldots,n$ نتیجه می‌شوند

$$\begin{cases} \sum_{i=1}^{m}\ (a_0+a_0x_{i1}+\cdots+a_nx_{in}-y_{Ai})\,x_{ij}=0 \\ \sum_{i=1}^{m}(\sigma_0+\sigma_1x_{i1}+\cdots+\sigma_nx_{in}-S_{Ai})\,x_{ij}=0 \\ \sum_{i=1}^{m}\ (b_0+b_0x_{i1}+\cdots+b_nx_{in}-y_{Bi})\,x_{ij}=0 \\ \sum_{i=1}^{m}(\delta_0+\delta_1x_{i1}+\cdots+\delta_nx_{in}-S_{Bi})\,x_{ij}=0 \end{cases}$$

که در آن، به ازای $i=1,\ldots,m$ ، $x_{i0}=1$ قرار داده شده است. از معادلات بالا روابط زیر، به ازای $j=0,1,\ldots,n$ ، به دست می‌آیند:

$$\begin{cases} a_0\sum_{i=1}^{m}x_{i0}\,x_{ij}+a_1\sum_{i=1}^{m}x_{i1}\,x_{ij}+\cdots+a_n\sum_{i=1}^{m}x_{in}\,x_{ij} \\ \quad =\sum y_{Ai}x_{ij}\,,\ \ j=0,1,\ldots,n. \\ \sigma_0\sum_{i=1}^{m}x_{i0}\,x_{ij}+\sigma_1\sum_{i=1}^{m}x_{i1}\,x_{ij}+\cdots+\sigma_n\sum_{i=1}^{m}x_{in}\,x_{ij} \\ \quad =\sum S_{Ai}x_{ij}\,,\ \ j=0,1,\ldots,n. \\ b_0\sum_{i=1}^{m}x_{i0}\,x_{ij}+b_1\sum_{i=1}^{m}x_{i1}\,x_{ij}+\cdots+b_n\sum_{i=1}^{m}x_{in}\,x_{ij} \\ \quad =\sum y_{Bi}x_{ij}\,,\ \ j=0,1,\ldots,n. \\ \delta_0\sum_{i=1}^{m}x_{i0}\,x_{ij}+\delta_1\sum_{i=1}^{m}x_{i1}\,x_{ij}+\cdots+\delta_n\sum_{i=1}^{m}x_{in}\,x_{ij} \\ \quad =\sum S_{Bi}x_{ij}\,,\ \ j=0,1,\ldots,n. \end{cases} \quad (۱۰)$$

با استفاده از تکنیک ارائه شده برای برآورد رگرسیون فازی با استفاده از روش ماتریس توسط سو و لی [۱۰] دستگاه زیر را تشکیل داده و با حل آن مقادیر a، σ، b و σ به دست می‌آیند

$$\begin{cases} Aa=y_A \\ A\sigma=S_A \\ Ab=y_B \\ A\sigma=S_B \end{cases} \quad (۱۱)$$

برای حل دستگاه فوق قرار داده می‌شود

$$A=x'x\,, \quad (۱۲)$$

$$x=\begin{bmatrix} 1 & x_{11} & \cdots & x_{1m} \\ 1 & x_{21} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{m1} & \cdots & x_{mm} \end{bmatrix}_{m(n+1)}$$

که در آن ماتریس x' ترانهاده ماتریس x است. در نتیجه

$$\begin{cases} a=(a_0,a_1,\ldots,a_n)^T, \\ y_A=[\sum_{i=1}^{m}y_{Ai}x_{i0}\,,\sum_{i=1}^{m}y_{Ai}x_{i1}\,,\ldots,\sum_{i=1}^{m}y_{Ai}x_{in}]^T \\ \sigma=(\sigma_0,\sigma_1,\ldots,\sigma_n)^T, \\ S_A=[\sum_{i=1}^{m}S_{Ai}x_{i0}\,,\sum_{i=1}^{m}S_{Ai}x_{i1}\,,\ldots,\sum_{i=1}^{m}S_{Ai}x_{in}]^T \\ b=(b_0,b_1,\ldots,b_n)^T, \\ y_B=[\sum_{i=1}^{m}y_{Bi}x_{i0}\,,\sum_{i=1}^{m}y_{Bi}x_{i1}\,,\ldots,\sum_{i=1}^{m}y_{Bi}x_{in}]^T \\ \delta=(\delta_0,\delta_1,\ldots,\delta_n)^T, \\ S_B=[\sum_{i=1}^{m}S_{Bi}x_{i0}\,,\sum_{i=1}^{m}S_{Bi}x_{i1}\,,\ldots,\sum_{i=1}^{m}S_{Bi}x_{in}]^T \end{cases} \quad (۱۳)$$

بر پایه‌ی قضیه ۸،۲ و با توجه به اینکه اگر A همیشه مثبت باشد دارای معکوس A^{-1} است، نتیجه زیر به دست می‌آید.

www.SID.ir

Archive of SID

نتیجه ۱،۳ اگر رتبه ماتریس x برابر با $n+1$ باشد آن گاه دستگاه (۱۱) دارای جواب یکتا به صورت زیر است.

$$\begin{cases} a = A^{-1}y_A \\ \sigma = A^{-1}S_A \\ b = A^{-1}y_B \\ \delta = A^{-1}S_B \end{cases} \qquad (۱۴)$$

اثبات. برای هر کدام از روابط (۵) و (۶) با توجه به روابط (۷) و (۱۰) به ترتیب داریم

$$\begin{cases} Aa = y_A \\ A\sigma = S_A{}' \end{cases} \qquad (۱۵)$$

$$\begin{cases} Ab = y_B \\ A\sigma = S_B \end{cases} \qquad (۱۶)$$

با توجه به قضیه ۸،۲ و قضیه ۹،۲ هر دو دستگاه (۱۵) و (۱۶) دارای جواب یکتا هستند و این یعنی که دستگاه (۱۱) دارای جواب یکتا است.

بر پایه‌ی نتیجه‌ی بالا، می‌توان قضیه زیر را در مورد مینیمم سازی عبارت (۳،۳) اثبات کرد.

قضیه ۲،۳. اگر رتبه ماتریس x برابر با $n+1$ باشد، $\mathrm{A}^{-1}S_A \geq 0$ و $A^{-1}S_B \geq 0$ آن گاه مسئله مینیمم سازی عبارت (۳،۳) دارای جواب یکتا است که از رابطه (۱۰،۳) بدست می‌آید (شرطهای $\mathrm{A}^{-1}S_A \geq 0$ و $A^{-1}S_B \geq 0$ در قضیه ۲،۳ تضمین می‌کند که مقادیر σ_iها و δ_iها، که پهناهای ضرایب فازی مدل هستند، نامنفی باشند).

در صورتی که $\mathrm{A}^{-1}S_A < 0$ یا $A^{-1}S_B < 0$ باشند. دیگر روش سو و لی [۱۰] جوابگو نیست و برای رفع این مشکل از تکنیکی که محمدی و طاهری در برآورد رگرسیون فازی با استفاده از ماتریس‌ها ارائه کردند [۹] استفاده می‌شود. یعنی فرض کنید در مدلی که طبق روش قبل به داده‌ها برازش می‌دهیم، یک یا چند σ_i ها یا δ_iها منفی به دست می‌آید. در این صورت مجدداً یک مدل جدید به داده‌ها برازش می‌دهیم، با این تفاوت که این بار ضرایب متغیرهایی را که برای آنها $\sigma_i < 0$ و یا $\delta_i < 0$ به دست آمده است، غیر فازی فرض می‌کنیم. به بیان دیگر پهنای ضرایب فازی این متغیرها را صفر قرار می‌دهیم [۹]. بنابراین در این حالت، مقادیر a_iها و b_iها (مراکز ضرایب فازی) مشابه حالت قبل به ترتیب از روابط $a = A^{-1}y_A$ و $b = A^{-1}y_B$ به دست می‌آیند. ولی برای مقادیر σ_iها و یا δ_iها (پهنای ضرایب فازی مدل) باید به ترتیب از روابط $\sigma^* = \mathrm{A}^{*-1}S_A$ و $\delta^* = \mathrm{A}^{*-1}S_B$ استفاده کرد. در این روابط σ^* و δ^* به ترتیب بردارهایی مانند σ و δ هستند با این تفاوت که درایه‌های متناظر با ضرایب غیر فازی، حذف شده‌اند. همچنین A^* ماتریسی مانند A است با این تفاوت که سطرهای متناظر با ضرایب غیر فازی، حذف شده‌اند.

۴. نتایج عددی

مثال ۴-۱: برآورد رگرسیون خطی زی-نامبری مربوط با داده‌های جدول (۱) با روش پیشنهادی به صورت زیر است (بخشی از اطلاعات جدول (۱) از مرجع [۱۴] گرفته شده است).

ماتریس x و x' به صورت زیر است

$$x = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}_{5\times 2} \quad ,$$

$$x' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}_{2\times 5}$$

مقدار ماتریس A به صورت زیر به دست می‌آید.

$$A = x'x = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 5 & 15 \\ 15 & 55 \end{bmatrix}_{2\times 2}$$

مقادیر $\mathrm{y_A}$، $\mathrm{y_B}$، $\mathrm{S_A}$ و $\mathrm{S_B}$ و همچنین مقادیر A^{-1}، a، σ، b و δ به صورت‌های زیر به دست می‌آیند.

www.SID.ir

Archive of SID

$$a = (a_0, a_1)^T,$$

$$y_A = \left[\sum_{i=1}^{24} y_i x_{i0}, \sum_{i=1}^{24} y_i x_{i1}\right]^T = (50.4, 68.3)^T$$

$$\sigma = (\sigma_0, \sigma_1)^T,$$

$$S_A = \left[\sum_{i=1}^{24} S_i x_{i0}, \sum_{i=1}^{24} S_i x_{i1}\right]^T = (11.6, 36.4)^T$$

$$b = (b_0, b_1)^T,$$

$$y_B = \left[\sum_{i=1}^{m} y_{Bi} x_{i0}, \sum_{i=1}^{m} y_{Bi} x_{i1}\right]^T = (\,2.375, 7.385)^T$$

$$\delta = (\delta_0, \delta_1)^T,$$

$$S_B = \left[\sum_{i=1}^{m} S_{Bi} x_{i0}, \sum_{i=1}^{m} S_{Bi} x_{i1}\right]^T = (2.168\,, 6.522\,)^T$$

$$A^{-1} = \frac{1}{50}\begin{bmatrix} 55 & -15 \\ -15 & 5 \end{bmatrix}$$

$$a = A^{-1}\, y_A = \frac{1}{50}\begin{bmatrix} 55 & -15 \\ -15 & 5 \end{bmatrix}\begin{bmatrix} 50.4 \\ 168.3 \end{bmatrix} = \begin{bmatrix} 4.95 \\ 1.71 \end{bmatrix} = [4.95 \quad 1.71]^T$$

$$\sigma = A^{-1}\, S_A = \frac{1}{50}\begin{bmatrix} 55 & -15 \\ -15 & 5 \end{bmatrix}\begin{bmatrix} 11.6 \\ 36.4 \end{bmatrix} = \begin{bmatrix} 1.84 \\ 0.16 \end{bmatrix} = [1.84 \quad 0.16]^T$$

$$b = A^{-1}\, y_B = \frac{1}{50}\begin{bmatrix} 55 & -15 \\ -15 & 5 \end{bmatrix}\begin{bmatrix} 2.375 \\ 7.385 \end{bmatrix} = \begin{bmatrix} 0.397 \\ 0.026 \end{bmatrix} = [0.397 \quad 0.026]^T$$

$$\delta = A^{-1} S_B = \frac{1}{50}\begin{bmatrix} 55 & -15 \\ -15 & 5 \end{bmatrix}\begin{bmatrix} 2.168 \\ 6.522 \end{bmatrix} = \begin{bmatrix} 0.4282 \\ 0.0018 \end{bmatrix} = [0.4282 \quad 0.0018]^T$$

در نتیجه

$$y = \tilde{A}_0 + \tilde{A}_1 x_{i1} = (4.95, 1.84)(3.97, 0.4282) + (1.71, 0.16)(0.026, 0.0018) x_{i1}$$

که می‌توان جواب بدست آمده در این مثال را با مقدار بدست آمده در [۱۴] مقایسه کرد.

جواب بدست آمده در [۱۴] :

$$\mathrm{y} = (4.95, 1.84)(0.398, 0.427) + (1.70, 0.16)(0.025, 0.002)x$$

مثال۴-۲: برآورد رگرسیون خطی زی-نامبری مربوط با داده‌های جدول (۲) با روش پیشنهادی به صورت زیر است (بخشی از اطلاعات جدول (۲) از مرجع [۱۰] گرفته شده است).

ماتریس x و x' به صورت زیر است

$$x = \begin{bmatrix} 1 & 0.78 \\ \dots & \dots \\ 1 & 1.13 \end{bmatrix}_{24\times 2}, \quad x' = \begin{bmatrix} 1 & \dots & 1 \\ 1 & \dots & 1.13 \end{bmatrix}_{2\times 24}$$

مقدار ماتریس A به صورت زیر به دست می‌آید.

$$A = x'x = \begin{bmatrix} 1 & \dots & 1 \\ 1 & \dots & 1.13 \end{bmatrix}_{2\times 24}\begin{bmatrix} 1 & 0.78 \\ \dots & \dots \\ 1 & 1.13 \end{bmatrix}_{24\times 2} = \begin{bmatrix} 24 & 20.9 \\ 20.9 & 26.81 \end{bmatrix}_{2\times 2}$$

مقادیر y_A ، y_B ، S_A ، S_B و همچنین مقادیر A^{-1}، a، σ، b و δ به صورت‌های زیر به دست می‌آیند.

جدول ۱: ورودی حقیقی، خروجی مبتنی بر اعداد-زاده

i	Xi	$[\mathrm{Y_i}]^z = (Yi\,, \underline{e}i, \overline{e}i\,)_T$
1	1	(8, 1.8)(0.495,0.492)
2	2	(6.4, 2.2)(0.365,0.370)
3	3	(9.5, 2.6)(0.444,0.400)
4	4	(13.5, 2.6)(0.527,0.440)
5	5	(13, 2.4)(0.544,0.466)

www.SID.ir

Archive of SID

$$\begin{cases} a = (a_0, a_1)^T, \\ y_A = \left[\sum_{i=1}^{24} y_i x_{i0}, \sum_{i=1}^{24} y_i x_{i1}\right]^T = (164.09, 202.37)^T \\ \sigma = (\sigma_0, \sigma_1)^T, \\ S_A = \left[\sum_{i=1}^{24} S_i x_{i0}, \sum_{i=1}^{24} S_i x_{i1}\right]^T = (16.41, 20.23)^T \\ b = (b_0, b_1)^T, \\ y_B = \left[\sum_{i=1}^{m} y_{Bi} x_{i0}, \sum_{i=1}^{m} y_{Bi} x_{i1}\right]^T = (\ 9.01, 8.61)^T \\ \delta = (\delta_0, \delta_1)^T, \\ S_B = \left[\sum_{i=1}^{m} S_{Bi} x_{i0}, \sum_{i=1}^{m} S_{Bi} x_{i1}\right]^T = (12.74, 11\ )^T \end{cases}$$

$$A^{-1} = \frac{1}{206.63}\begin{bmatrix} 26.81 & -20.9 \\ -20.9 & 24 \end{bmatrix}$$

$$a = A^{-1}\, y_A = \frac{1}{206.63}\begin{bmatrix} 26.81 & -20.9 \\ -20.9 & 24 \end{bmatrix}\begin{bmatrix} 164.09 \\ 202.37 \end{bmatrix} = \begin{bmatrix} 0.82 \\ 6.90 \end{bmatrix} = [0.82 \quad 6.90]^T$$

$$\sigma = A^{-1}\, S_A = \frac{1}{206.63}\begin{bmatrix} 26.81 & -20.9 \\ -20.9 & 24 \end{bmatrix}\begin{bmatrix} 16.41 \\ 20.23 \end{bmatrix} = \begin{bmatrix} 0.082 \\ 6.090 \end{bmatrix} = [0.082 \quad 0.69]^T$$

$$b = A^{-1}\, y_B = \frac{1}{206.63}\begin{bmatrix} 26.81 & -20.9 \\ -20.9 & 24 \end{bmatrix}\begin{bmatrix} 9.01 \\ 8.61 \end{bmatrix} = \begin{bmatrix} 0.29 \\ 0.08 \end{bmatrix} = [0.29 \quad 0.08]^T$$

$$\delta = A^{-1} S_B = \frac{1}{206.63}\begin{bmatrix} 26.81 & -20.9 \\ -20.9 & 24 \end{bmatrix}\begin{bmatrix} 12.74 \\ 11 \end{bmatrix} = \begin{bmatrix} 0.54 \\ -0.01 \end{bmatrix} = [0.54 \quad -0.01]^T$$

از طرفی باید $A^{-1}S_B \geq 0 \Longrightarrow \delta \geq 0.$ بنابراین

$$\delta^* = A^{*-1} S_B = \frac{1}{20.9}[12.47] = \begin{bmatrix} 0.54 \\ -0.01 \end{bmatrix} = 0.59$$

در نتیجه

$$y = \tilde{A}_0 + \tilde{A}_1 x_{i1} = (0.82, 0.082)(0.29, 0.59) + (6.90, 0.69)(0.08) x_{i1}$$

به ازای $\mathrm{x}_{i1} = 0.75$ مقدار رگرسیون زی–نامبری به صورت زیر بدست می‌آید.

$$y = \big(0.82 + (6.9 \times 0.75), 0.082 + (0.69 * 0.75)\big), (0.29 + (0.08 * 0.75), 0.59)$$

$$y = (5.9, 0.59)(0.35, 0.59)$$

جواب بدست آمده در [۱۴] :

$$\mathrm{y} = (5.7, 0.6)(0.36, 0.51)$$

جدول ۲: داده‌های عددی و خروجی مبتنی بر اعداد –زاده.

شماره مشاهدات	x_i	$[y_i]^z = (y_{ki}, s_i)(y_{Hi}, c_i)$
1	0.78	(3.08, 0.31)(0.23, 0.39)
2	0.64	(2.86, 0.29)(0.22, 0.37)
3	0.62	(6.25, 0.63)(0.40, 0.59)
4	0.49	(4.11, 0.41)(0.29, 0.48)
5	1.10	(1.04, 0.10)(0.08, 0.16)
6	0.61	(2.71, 0.27)(0.21, 0.36)
7	0.74	(4.45, 0.45)(0.31, 0.50)
8	1.15	(6.92, 0.69)(0.43, 0.62)
9	1.08	(7.41, 0.74)(0.45, 0.63)
10	0.38	(9.08, 0.91)(0.50, 0.67)
11	0.61	(6.56, 0.66)(0.41, 0.60)
12	0.98	(5.05, 0.51)(0.34, 0.53)

www.SID.ir

Archive of SID

13	0.71	(5.23, 0.52)(0.35, 0.54)
14	0.51	(5.16, 0.52) (0.35, 0.54)
15	0.77	(11.10, 1.11)(0.56, 0.69)
16	0.99	(4.47, 0.45)(0.31, 0.50)
17	3.56	(28.84, 2.88) (0.69, 0.55)
18	0.86	(9.43, 0.94) (0.52, 0.67)
19	0.61	(4.50, 0.45)(0.31, 0.50)
20	0.64	(9.30, 0.94)(0.51, 0.67)
21	0.71	(9.48, 0.95)(0.52, 0.67)
22	0.61	(3.65, 0.37)(0.26, 0.44)
23	0.63	(10.14, 1.01)(0.54, 0.68)
24	1.13	(3, 0.3)(0.22, 0.39)

نتیجه‌گیری

در این مقاله، سعی کردیم با استفاده از ماتریس‌ها روشی جهت برآورد رگرسیون خطی مبتنی بر اعداد–زاده که در آن متغیرهای مستقل مقدار حقیقی و ضرایب و متغیرهای وابسته مقادیری مبتنی بر اعداد–زاده هستند را معرفی کنیم. برای حل این مسئله با توجه به این که متغیرهای وابسته و ضرایب مقادیری بر اساس اعداد– زاده دارند مدل رگرسیون خطی مبتنی بر اعداد–زاده را به دو مدل رگرسیون خطی فازی تبدیل می‌کنیم سپس با استفاده از روش ماتریسی مبتنی بر مینیمم‌سازی مجموع مربعات خطا با توجه به فاصله d ضرایب هر یک از دو مدل رگرسیون خطی فازی را برای حالتی که $\mathrm{A}^{-1}S_A \geq 0$ و $A^{-1}S_B \geq 0$ باشند بدست می‌آوریم. و برای حالتی که $\mathrm{A}^{-1}S_A < 0$ یا $A^{-1}S_B < 0$ باشند از روش محمدی و طاهری برای تقریب ضرایب دو مدل استفاده کردیم. که در نهایت با جایگذاری ضرایب دو مدل رگرسیون فازی در مدل اصلی یعنی رگرسیون خطی مبتنی بر اعداد–زاده ضرایب مسئله ما بدست می‌آید. نشان دادیم که اگر رتبه ماتریس متغیرهای مستقل برابر با $n+1$ و $\mathrm{A}^{-1}S_A \geq 0$ و $A^{-1}S_B \geq 0$ باشند آن گاه مسئله مینیمم سازی رابطه ۷،۳ دارای جوای یکتا است. طی یک مثال جواب بدست آمده از روش پیشنهادی را با جواب بدست آمده در کار قبلی خودمان که تنها روش موجود برای این مدل از رگرسیون است مقایسه کردیم.

www.SID.ir

Archive of SID

فهرست منابع

[1] C.B. Cheng, E.S. Lee, Fuzzy regression with radial basis function network, Fuzzy Sets and Systems 119 (2) (2001) 291–301.

[2] A. Bardossy, Note on fuzzy regression, Fuzzy Sets and Systems 37 (1990) 65–75.

[3] A. Bardossy, I. Bogardi, L. Duckstein, Fuzzy regression in hydrology,Water Resources Res. 26 (1990) 1497–1508.

[4] L.A. Zadeh, Fuzzy sets, Inform. and Control 8 (1965) 338–353.

[5] L.A. Zadeh, Fuzzy sets and information granularity, in: M.M. Gupta, R.K. Ragade, R.R. Yager (Eds.), Advances in Fuzzy Set Theory and Applications, North Holland, Amsterdam, (1979), pp. 3–18.

[6] L.A. Zadeh, Fuzzy logic = computing with words, IEEE Trans. Fuzzy Systems 4 (2) (1996) 103–111.

[7] H. Tanaka, Fuzzy data analysis by possibilistic linear models, Fuzzy Sets and Systems, 24(1987), 363-375.

[8] H. Tanaka, I. Havashi and J. Watada,. Possibilistic Liner regression analysis for fuzzy data, European J. Oper. Res., (1989) 40: 389-396.

[9] J. Mohammadi, S. M. Taheri, Pedomodels fitting with fuzzy least squares regression, Iraninan J. Fuzzy Systems, (2004) 1 (2): 45-61.

[10] R. Xu, C. Li, Multidimentional least-squares fitting with a fuzzy model, fuzzy Sets and Systems, (2001)119: 215-223.

[11] M. Hojati, C. R. Bector, Smimou K. A Simple Method of Fuzzy Linear Regression. European Journal of Operational Research (2005); 166; 172-184.

[12] G. Peters, Fuzy Linear Regression with fuzzy intervals. Fuzzy Sets and Systems (1994); 63; 45-55.

[13] L.A. Zadeh, A Note on Z-numbers, Information Sciences 181 (2011) 2923–2932.

[14] S. Ezadi, T. Allahviranloo, Numerical solution of linear regression based on Z-numbers by improved neural network, paper accept.

[15] S. Ezadi, T. Allahviranloo, New multi-layer method for Z-number ranking using Hyperbolic Tangent function and convex combination, Intelligent Automation Soft Computing., (2017), 1-7.

[16] S. Ezadi, T. Allahviranloo, Two new methods for ranking of Z-numbers based on sigmoid function and sign method, International Journal of Intelligent Systems., (2018), 1-12.

[17] B. Kang, D. WEI, Y. LI and Y. DENG, Decision Making Using Z-numbers under Uncertain Environment, Journal of Computational Information Systems, 7 (2012) 2807–2814.

[18] B. Kang, D. Wei, Y. Li, Y. Deng, A method of converting Z-number to classical fuzzy number, Journal of Information and Computational Scienc., 3 (2012), 703-709.

[19] R.A. Alive, A.V. Alizadeh, O.H. Huseynov, The arithmetic of discrete Z-numbers, Inform. Sciences., 290 (2015) 134-155.

www.SID.ir

Archive of SID

[20] R.A. Alive, O.H. Huseynov, R.R. Alive, A.V. Alizadeh, The arithmetic of Z-numbers. Theory and Applications, World Scientific, Singapore, (2015).

[21] R.A. Alive, O.H. Huseynov, and R. Serdaroglu, Ranking of Z-numbers, and its Application in Decision Making. Int.J.

[22] ASA. Bakar, A. Gegov, (2015) Multi-layer decision methodology for ranking Z-numbers. Int J Comput Intell Syst, 8:395–406.

[23] H. J. Zimmermann, Fuzzy set theory and its applications, Kluwer Academic, Boston, (1991).

www.SID.ir

Archive of SID

www.SID.ir