سیستم‌های فازی و کاربردها
سال اول، شماره دوم (زمستان ۱۳۹۷) صص. ۳۵ تا ۴۴
DOR: 20.1001.1.27174409.1397.1.2.3.6

الگوریتم EM مبتنی بر مشاهدات فازی

عباس پرچمی

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

چکیده

الگوریتم EM ابزاری قدرتمند برای برآورد بیشترین درستنمایی مبتنی بر داده‌های ناقص است و در اغلب کتاب‌های استنباط آماری نیز مطرح شده است. در اینجا معنی کلمه «ناقص» حالتی کلی دارد و در موقعیت‌های متفاوت می‌تواند به معانی گوناگونی (مانند داده‌های گم شده، داده‌های بازه‌ای، مشاهدات سانسور شده و نظایر آنها) دلالت داشته باشد. این مقاله کاربرد جدیدی از الگوریتم EM را مطرح می‌سازد که در آن منظور از داده‌های ناقص داده‌های نادقیق/فازی/مبهم هستند. بر اساس این نوع داده‌های مبهم، در این مقاله برآورد بیشینه درستنماییِ پارامتر توزیع نمایی به کمک الگوریتم EM در قالب یک مثال عددی محاسبه شده است. این مثال می‌تواند به منظور فهم مطلب و نیز استفاده از الگوریتم EM در مثال‌ها/حالت‌های پیچیده‌تر، برای دانشجویان تحصیلات تکمیلی مفید باشد.

۱ مقدمه‌ای بر الگوریتم EM

فرض کنید $\mathbf{Y}$ بردار داده‌های مشاهده شده و $\mathbf{X}$ بردار داده‌های نامعلوم/ناقص باشند. همچنین فرض کنید θ پارامتر مجهول است که می‌خواهیم برآورد شود و $l_c(\theta)$ لگاریتم تابع درستنمایی بر اساس داده‌های کامل است که به ازای تمام مقادیر ممکن θ در فضای پارامتر Ω، تعریف می‌شود. الگوریتم

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

عبارات و کلمات کلیدی: الگوریتم EM، برآورد بیشینه درستنمایی، داده‌های فازی.

EM که نخستین بار توسط [۱] مطرح شد، با یک مقدار اولیه $\theta^{(\circ)} \in \Omega$ آغاز می‌شود و دو گام زیر را تا رسیدن به همگرایی، تکرار می‌کند:

گام E [۱]: محاسبه $l_c^{(j)}(\theta) = E_{\mathbf{X}|\mathbf{Y},\theta^{(j-1)}}[l_c(\theta)]$ که امید ریاضی با توجه به داده‌های کامل/بدون‌نقص $\mathbf{X}$ (که در دست نیستند و بطور کامل مشاهده نشده‌اند) به شرط داده‌های مشاهده شده‌ی ناقص $\mathbf{Y}$ گرفته می‌شود و باید توجه شود که مقدار $\theta^{(j-1)}$ در این امید ریاضی جایگذاری می‌شود.

گام M [۲]: یافتن $\theta^{(j)} \in \Omega$ بقسمی که $l_c^{(j)}(\theta)$ بیشینه شود.

تکرار دو گام فوق به ازای $j = ۱, ۲, \ldots$، منجر به همگرایی دنباله $\theta^{(۱)}, \theta^{(۲)}, \ldots$ در ماکزیمم موضعی لگاریتم درستنمایی کلیه داده‌های کامل می‌شود [۳] .

ثبت اطلاعات و داده‌ها در دنیای واقعی ممکن است با چالش‌هایی مواجه باشد و اطلاعات بطور تقریبی/نادقیق/فازی ثبت شوند. به عنوان نمونه می‌توان به مثال‌های زیر برای داده‌های فازی اشاره کرد:

۱. اندازه گیری حجم گاز متصاعد شده از دهانه یک آتشفشان در هر ساعت (یا از یک موتور جت در هر ثانیه تحت شرایطی مشخص)

۲. میزان علاقه/رضایت یک شاغل از شغلش و یا میزان رضایت/مطلوبیت از زندگی

۳. مدت زمان علاقمندی یک شاغل از شغلش از ابتدای شروع به کارش برحسب سال (مدتی که فرد با اشتیاق به کارش مشغول بود از بدو شروع به کار)

۴. فساد مواد غذایی تدریجی صورت می‌گیرد. بنابراین مدت زمانی که یک یخچال می‌تواند غذا/میوه‌ای خاص را سالم نگاه دارد

۵. درآمد ماهیانه یک راننده‌ی تاکسی

۶. ساعت طلوع/غروب خورشید. بنابراین، طول روز هم مقداری فازی است که ابهام آن در کشورهای قطبی مانند فنلاند نسبت به کشورهای استوایی بیشتر است

[1] Expectation

[2] Maximization

۷. طول عمر باطری

۸. آستانه‌ی تحمل یک بیمار

همان‌گونه که مطرح شد، الگوریتم EM ابزاری برای برآورد بیشینه درستنمایی مبتنی بر داده‌های ناقص است و داده‌های فازی می‌تواند یکی از مصداق‌های داده‌های ناقص به حساب آید، زیرا در این حالت بجای اینکه مقدار دقیقی به عنوان مشاهده به ثبت رسیده باشد تنها یک مقدار تقریبی (در قالب یک عدد فازی) مشاهده و ثبت شده است. لذا به منظور برآورد بیشینه درستنمایی مبتنی بر داده‌های فازی، می‌توان دو گام مطرح شده‌ی فوق را بصورت زیر بازنویسی نمود (در این حالت فرض بر آن است که تمامی داده‌های بصورت فازی مشاهده/ثبت شده‌اند):

گام E**:** محاسبه $l_c^{(j)}(\theta) = E_{\mathbf{X}|\tilde{\mathbf{X}},\theta^{(j-1)}}[l_c(\theta)]$ که امید ریاضی با توجه به داده‌های کامل/بدون‌نقص $\mathbf{X}$ (که در دست نیستند و مشاهده نشده‌اند) به شرط داده‌های مشاهده شده‌ی نادقیق/فازی $\tilde{\mathbf{X}}$ گرفته می‌شود.

گام M**:** یافتن $\theta^{(j)} \in \Omega$ بقسمی که $l_c^{(j)}(\theta)$ بیشینه شود.

مثال‌های اندکی از این وجه کاربردی الگوریتم EM که نخستین بار توسط [۲] پیشنهاد شد، در دست است. از این رو در بخش ۳ قصد داریم تا از الگوریتم EM در برآورد بیشترین درستنمایی پارامتر مجهول توزیع نمایی، در قالب یک مثال کاربردی استفاده کنیم. لذا پیش از آن، برخی از تعاریف و مفاهیم پیش‌نیاز را در بخش ۲ مطرح می‌سازیم.

۲ احتمال شرطی بر اساس مشاهدات فازی

تعریف ۱.۲. تابع چگالی/جرم احتمال شرطی X به شرط مشاهده‌ی فازی $\tilde{x}$ را با نماد $f_\theta(x|\ X \in \tilde{x})$ نمایش و بصورت زیر تعریف می‌کنیم

$$f_\theta(x|\ X \in \tilde{x}) = \frac{\tilde{x}(x)\ f_\theta(x)}{\int \tilde{x}(x)\ f_\theta(x)\ dx} \qquad (۱)$$

در حالت گسسته به جای انتگرال از مجموع‌یابی استفاده می‌شود. لازم بذکر است که [۴] تعریف مشابهی تحت عنوان چگالی X بر اساس اطلاعات فازی مطرح کرده است.

توجه ۲.۲. بدیهی است که تابع چگالی/جرم احتمال شرطی معرفی شده در تعریف ۱.۲، از خواص یک تابع چگالی/جرم احتمال مانند $f_\theta(x|\, X \in \tilde{x}) \geqslant ۰, \; \forall x \in R$ و $\int_{-\infty}^{\infty} f_\theta(x|\, X \in \tilde{x})\, dx = ۱$ برخوردار است.

تعریف ۳.۲. فرض کنید $\mathbf{X} = (X_۱, \ldots, X_n)$ نمونه‌ای تصادفی از یک جامعه با تابع چگالی/جرم احتمال $f_\theta(x)$ و همچنین $\tilde{\mathbf{x}} = (\tilde{x}_۱, \ldots, \tilde{x}_n)$ مقدار مشاهده‌شده‌ی نمونه باشد، بطوریکه $\tilde{x}_i$ مقدار مشاهده‌شده‌ی فازی برای متغیر X_i با تابع عضویت $\tilde{x}_i(x)$ به ازای $i = ۱, \ldots, n$ است. بعنوان تابعی از θ، تابع درستنمایی مبتنی بر مشاهدات فازی $\tilde{\mathbf{x}}$ برابر است با

$$
\begin{aligned}
L(\theta|\, \tilde{\mathbf{x}}) &= L(\theta|\, \tilde{x}_۱, \ldots, \tilde{x}_n) \\
&= \prod_{i=۱}^{n} P_\theta\left(X_i \in \tilde{x}_i\right) \\
&= \prod_{i=۱}^{n} \int_{-\infty}^{\infty} \tilde{x}_i(x)\, f_\theta(x)\, dx
\end{aligned}
\tag{۲}
$$

نتیجه ۴.۲. *احتمال شرطی پیشامد فازی $\tilde{x}'$، بر اساس احتمال پیشامد فازی [۵]، معادل است با*

$$
\begin{aligned}
P_\theta\left(X \in \tilde{x}'|\, X \in \tilde{x}\right) &= \int \tilde{x}'(x)\, f_\theta(x|\, X \in \tilde{x})\, dx \\
&= \frac{\int \tilde{x}'(x)\, \tilde{x}(x)\, f_\theta(x)\, dx}{\int \tilde{x}(x)\, f_\theta(x)\, dx} \\
&= \frac{P_\theta\left(X \in (\tilde{x}' \cap \tilde{x})\right)}{P_\theta\left(X \in \tilde{x}\right)}
\end{aligned}
\tag{۳}
$$

که درآن تابع عضویت اشتراک دو مجموعه فازی $\tilde{x}$ و $\tilde{x}'$ بصورت $(\tilde{x}' \cap \tilde{x})\,(x) = \tilde{x}'(x)\, \tilde{x}(x)$ تعریف شده است.

نتیجه ۵.۲. *بر اساس تابع چگالی/جرم احتمال شرطی معرفی شده در تعریف ۱.۲ ، امید ریاضی*

متغیر تصادفی X به شرط مشاهده‌ی فازی $\tilde{x}$ بصورت زیر بدست می‌آید (در حالت گسسته به جای انتگرال از مجموع‌یابی استفاده شود)

$$E_\theta\left(X|\; X \in \tilde{x}\right) = \int_{-\infty}^{\infty} x\; f_\theta(x|\; X \in \tilde{x})\; dx$$
$$= \frac{\int x\; \tilde{x}(x)\; f_\theta(x)\; dx}{\int \tilde{x}(x)\; f_\theta(x)\; dx} \qquad (۴)$$

۳ برآورد پارامتر توزیع نمایی به کمک الگوریتم EM مبتنی بر داده‌های ناقصِ فازی-مقدار

فرض کنید $X = (X_1, \ldots, X_n)$ یک نمونه تصادفی از توزیع نمایی با پارامتر λ باشد. حالتی را درنظر بگیرید که تمامی مشاهدات از نوع فازی بوده و در نتیجه بجای مشاهده‌ی مقادیر دقیق x_iها، یک تابع عضویت برای هر داده‌ی فازی به ثبت رسیده است. به عبارت دیگر، در این مثال قصد داریم تا برآورد بیشینه درستنمایی پارامتر مجهول λ را مبتنی بر مشاهدات/داده‌های فازی-مقدار $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$ به کمک الگوریتم EM محاسبه کنیم. تابع درستنمایی مبتنی بر این داده‌های فازی-مقدار برابر است با

$$L(\lambda|\tilde{\mathbf{x}}) = \prod_{i=1}^{n} P_\lambda\left(X_i \in \tilde{x}_i\right) = \prod_{i=1}^{n} \int_0^\infty \tilde{x}_i(x)\ f_\lambda(x)\ dx \qquad (5)$$

گرچه برای یافتن $MLE(\lambda)$ نمی‌توان از این تابع (و یا لگاریتم آن) نسبت به λ مشتق گرفت، اما می‌توان برای بیشینه‌سازی آن از روش‌های عددی (هم‌چون روش نیوتن رافسون) استفاده کرد. به منظور استفاده از الگوریتم EM، که رویکرد مورد نظر در این مقاله است، ابتدا تابع لگاریتم درستنمایی را مبتنی بر داده‌های کامل/حقیقی-مقدار محاسبه می‌کنیم

$$L_c(\lambda) = \lambda^n \exp(-\lambda \sum_{i=1}^{n} x_i) \qquad \Rightarrow \qquad l_c(\lambda) = n\ ln(\lambda) - \lambda \sum_{i=1}^{n} x_i \quad (6)$$

و با برابر صفر قرار دادن مشتق لگاریتم درستنمایی، $MLE(\lambda) = \frac{1}{\bar{x}} = \frac{n}{\sum_{i=1}^{n} x_i}$ بدست می‌آید. اما به یاد داشته باشیم که هدف محاسبه‌ی برآورد بیشینه درستنمایی بر اساس داده‌های فازی (و نه حقیقی) است.

طبق تعریف ۱.۲، تابع چگالی احتمال شرطی X_i به شرط مشاهده‌ی فازی $\tilde{x}_i$ برابر است با

$$\begin{aligned} f_\lambda(x_i|\ X_i \in \tilde{x}_i) &= \frac{\tilde{x}_i(x_i)\ f_\lambda(x_i)}{\int \tilde{x}_i(x)\ f_\lambda(x)\ dx} \\ &= \frac{\tilde{x}_i(x_i)\ \lambda\ \exp(-\lambda\ x_i)}{\int \tilde{x}_i(x)\ \lambda\ \exp(-\lambda\ x)\ dx} \\ &= \frac{\tilde{x}_i(x_i)\ \exp(-\lambda\ x_i)}{\int \tilde{x}_i(x)\ \exp(-\lambda\ x)\ dx}, \quad i = ۱, \ldots, n. \end{aligned} \tag{۷}$$

طبق الگوریتم مطرح‌شده در بخش ۱، در این مثال $\tilde{\mathbf{x}} = (\tilde{x}_۱, \ldots, \tilde{x}_n)$ داده‌های مشاهده‌شده‌ی ناقص و همچنین بردار $\mathbf{x} = (x_۱, \ldots, x_n)$ داده‌های کامل/بدون‌نقص (که در دست نیستند) می‌باشد. لذا لگاریتم درستنمایی jامین مرحله از الگوریتم مبتنی بر مشاهدات فازی (گام E) برابر

$$\begin{aligned} l_c^{(j)}(\lambda) &= E_{\mathbf{X}|\tilde{\mathbf{X}},\lambda^{(j-۱)}}\left[l_c(\lambda)\right] \\ &= E_{\lambda^{(j-۱)}}\left[l_c(\lambda)|X_۱ \in \tilde{x}_۱, \ldots,\ X_n \in \tilde{x}_n\right] \\ &= n\ln(\lambda) - \lambda \sum_{i=۱}^{n} E_{\lambda^{(j-۱)}}(X_i|\ X_۱ \in \tilde{x}_۱, \ldots, X_n \in \tilde{x}_n) \\ &= n\ln(\lambda) - \lambda \sum_{i=۱}^{n} E_{\lambda^{(j-۱)}}(X_i|\ X_i \in \tilde{x}_i) \end{aligned} \tag{۸}$$

است و مقدار بیشینه‌ی آن (گام M)، با ثابت فرض کردنِ $\lambda^{(j-۱)}$ (چون مقدارش در مرحله قبلی الگوریتم تعیین شده)، معادل است با

$$\frac{\partial l_c^{(j)}(\lambda)}{\partial \lambda} = ۰ \qquad \Rightarrow \qquad \hat{\lambda}^{(j)} = \frac{n}{\sum_{i=۱}^{n} E_{\lambda^{(j-۱)}}(X_i|\ X_i \in \tilde{x}_i)} \tag{۹}$$

که در آن، مطابق نتیجه ۵.۲، داریم

$$\begin{aligned} E_{\lambda^{(j-1)}}(X_i|\ X_i \in \tilde{x}_i) &= \int_{-\infty}^{\infty} x\ f_{\lambda^{(j-1)}}(x|\ X_i \in \tilde{x}_i)\ dx \\ &= \frac{\int_0^{\infty}\ x\ \ \tilde{x}_i(x)\ f_{\lambda^{(j-1)}}(x)\ dx}{\int_0^{\infty} \tilde{x}_i(x)\ f_{\lambda^{(j-1)}}(x)\ dx} \\ &= \left\{ \frac{\int\ x\ \ \tilde{x}_i(x)\exp(-\lambda\ x)\ dx}{\int \tilde{x}_i(x)\exp(-\lambda\ x)\ dx} \right\}_{\lambda=\lambda^{(j-1)}} \end{aligned} \tag{10}$$

اگر چه روش مطرح شده در این مقاله کلیت داشته و برای هر نوع عدد فازی برقرار است، اما به دلیل سادگی مطالب در ادامه از اعداد مثلثی فازی با نماد $\tilde{x} = T(c, s_l, s_r)$ و تابع عضویت زیر استفاده شده است که در آن c، s_l و s_r به ترتیب نشان دهنده‌ی هسته، پهنای چپ و پهنای راست عدد فازی $\tilde{x}$ هستند

$$\tilde{x}(x) = \begin{cases} \frac{x-c+s_l}{s_l} & \text{اگر} \quad c - s_l \leq x < c \\ \frac{c+s_r-x}{s_r} & \text{اگر} \quad c \leq x < c + s_r \\ 0 & \text{جاهای دیگر} \end{cases} \tag{11}$$

نمونه‌ای تصادفی از داده‌های مثلثی فازی زیر را برای برآورد پارامتر مجهول λ درنظر بگیرید (شکل ۱)

$$\tilde{x}_1 = T(1/019, 0/865, 1/374)$$

$$\tilde{x}_2 = T(1/141, 0/316, 1/548)$$

$$\tilde{x}_3 = T(0/806, 0/523, 0/944)$$

$$\tilde{x}_4 = T(0/679, 0/292, 0/428)$$

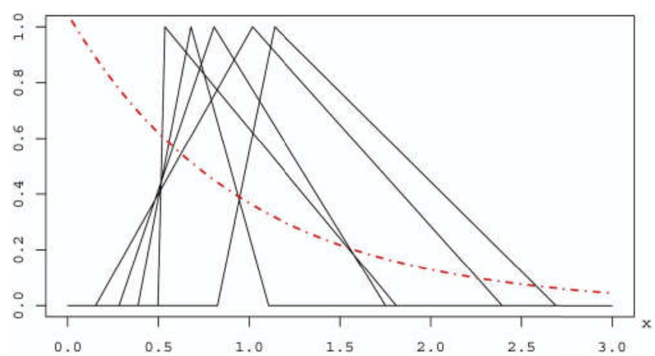
$$\tilde{x}_5 = T(0/536, 0/038, 1/274)$$

جدول ۱ تکرار الگوریتم و رسیدن به همگرایی آن را مبتنی بر این داده‌های فازی، به ازای دو مقدار

آغازین متفاوت $\lambda_1^{(0)}$ و $\lambda_2^{(0)}$ نشان می‌دهد که بر اساس هر دو مقدار آغازین، الگوریتم EM به یک همگرایی ($\hat{\lambda} = 1/042$) رسیده است.

جدول ۱: *روند همگرایی الگوریتم EM به ازای دو مقدار آغازین مختلف در مثال ۱.*

شماره تکرار الگوریتم (j)	$\lambda_1^{(j)}$	$\lambda_2^{(j)}$
۰	۵	۳۰
۱	۱/۳۹۷	۲/۰۳۱
۲	۱/۰۷۹	۱/۱۴۳
۳	۱/۰۴۵	۱/۰۵۲
۴	۱/۰۴۲	۱/۰۴۳
۵	۱/۰۴۲	۱/۰۴۲
۶		۱/۰۴۲



شکل ۱: *توابع عضویت داده‌های مثلثی فازی به همراه تابع چگالی احتمال نمایی برآوردشده.*

نتیجه‌گیری

الگوریتم EM یکی از روش‌های متداول برای برآورد بیشترین درستنمایی مبتنی بر داده‌های ناقص است. گر چه مقالات زیادی درباره‌ی این الگوریتم و کاربردهایش تا کنون نوشته شده است، اما مقاله‌ی حاضر کاربردی جدیدی از الگوریتم EM را مطرح می‌سازد که در آن منظور از داده‌های ناقص داده‌های فازی است. [۲] نخستین بار الگوریتم EM برای اعداد فازی را پیشنهاد کرد که

رویکردی کلی (رویکردی کارا برای حالت‌هایی که برخی از مشاهدات دقیق و برخی نیز فازی ثبت شده‌اند) و در عین حال پیچیده از نظر فرمولی/محاسباتی به شمار می‌آید. اما در این مقاله سعی شده است تا حالتی خاص از روش [۲] مورد بررسی قرار گیرد که تمامی داده‌های مشاهده/ثبت شده فازی هستند زیرا این حالت خاص منجر به فرمول‌های ساده‌تر شده و می‌تواند برای دانشجویان تحصیلات تکمیلی در حالت‌های پیچیده‌تر مفید باشد.

مراجع

[1] A. P. Dempster, N. M. Laird and D. B. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society* **39**(1977), 1-38.

[2] T. Denoeux, Maximum likelihood estimation from fuzzy data using the EM algorithm, *Fuzzy Sets and Systems* **183**(2011), 72-91.

[3] K. Knight, *Mathematical Statistics*, New York, Chapman & Hall, (2000).

[4] R. Pourmousa, On truncated measures of income inequality from a fuzzy perspective, *Iranian Journal of Fuzzy Systems* **15**(2018), 123-137.

[5] L. A. Zadeh, Probability measures of fuzzy events, *J. Math. Anal. Appl*, **23**(1968), 421-427.