

یادگیری مبتنی بر کلاس بندی c- میانگین با استفاده از توزیع آمیخته مقیاسی نرمال با اطلاعات گمشده

حمید بیگدلی* و سعید دریجانی

دانشگاه فرماندهی و ستاد آجا، تهران، ایران

تاریخ پذیرش: ۱۴۰۱/۹/۳

تاریخ دریافت: ۱۴۰۱/۲/۷

نوع مقاله: علمی-پژوهشی

چکیده

یکی از پرکاربردترین مدل‌ها برای رده‌بندی یا کلاس‌بندی داده‌ها که در سال‌های اخیر مورد توجه بسیاری از پژوهشگران قرار گرفته است، مدل‌های مخلوط متناهی است. بطور کلی رده‌بندی به فرایندی گفته می‌شود که در آن هر یک از مشاهدات به یکی از گروه‌های مشخص شده تعلق گرفته می‌شود. گرچه ایده اصلی در مدل‌های مخلوط بر اساس توزیع نرمال بوده است، اما در سال‌های اخیر با معرفی توزیع‌های دیگر مدل‌های مخلوط بر اساس این توزیع‌ها مورد توجه بسیاری از محققین بوده است. در مقالات از الگوریتم EM و گسترش‌های آن برای برآوردیابی استفاده شده است. با این حال این امکان وجود دارد که الگوریتم EM نتایج مناسبی برای کلاس‌بندی ارائه ندهد، زیرا در این روش هر عضو مشاهدات متعلق به یک کلاس است. این محدودیت باعث استفاده از رویکرد کلاس‌بندی فازی در این نوع مسائل شد. در این مقاله یک الگوریتم کلاس‌بندی براساس توزیع مخلوط متناهی آمیخته مقیاسی نرمال ارائه شده است. در این الگوریتم برای کلاس‌بندی از روش یادگیری فازی c- میانگین استفاده شده است. برای بررسی تاثیر مقادیر گمشده بر کلاس‌بندی داده‌ها، داده گمشده نیز در نظر گرفته شده است. از ساختار توزیع مخلوط متناهی آمیخته مقیاسی نرمال برای بررسی داده‌های گمشده و کلاس‌بندی داده‌ها (ادامه دارد)

استفاده می‌شود. در انتها نیز با استفاده از مثال واقعی و داده‌های شبیه‌سازی شده، مقایسه بین الگوریتم LB-FCM و EM صورت می‌گیرد. از این مقایسه نتیجه شده است که استفاده از این الگوریتم برای کلاس‌بندی داده‌ها مناسبتر است.

۱ مقدمه

مدل مخلوط متناهی یک روش معمول برای کلاس‌بندی داده‌های چند متغیره است. توزیع مخلوط متناهی آمیخته مقیاسی نرمال^۱ برای آنالیز کلاس‌بندی داده‌ها مورد استفاده قرار می‌گیرد. استفاده از داده گمشده در این روش ضرورت دارد زیرا ممکن است که استنباط بر روی این داده‌ها دچار مشکل و نقص شود. لین و همکاران^۲ [۷] با استفاده از الگوریتم EM و توزیع مخلوط متناهی نرمال در حضور داده گمشده، مدل مخلوط متناهی را در نظر گرفتند. در ادامه وانگ و همکاران^۳ [۹] این مدل را با استفاده از توزیع تی تعمیم دادند که اگر درجه آزادی در توزیع تی افزایش یابد به مدل نرمال نزدیک می‌شود. اخیرا نیز وانگ و همکاران [۱۰] توزیع آمیخته متناهی مخلوط مقیاسی-شکلی از توزیع چوله نرمال را در نظر گرفته و برآورد پارامترهای آن را با استفاده از الگوریتم EM بدست آوردند.

به هر حال استفاده از الگوریتم EM ممکن است که نتایج را برای کلاس‌بندی داده‌ها دچار انحراف کند، زیرا در این روش هر نقطه از داده‌ها دقیقا متعلق به یکی از گروه‌ها و کلاس‌ها خواهد بود. برای حل این مشکل، روش کلاس‌بندی فازی مورد استفاده قرار گرفت. زاده^۴ [۱۲] مجموعه‌های فازی را همراه با تابع عضویت فازی معرفی کرد که از این تابع عضویت فازی در مشخص کردن کلاس داده‌ها در حالت فازی استفاده می‌شود. گوستاوسن و کسل^۵ [۴] یک روش توسعه یافته از کلاس‌بندی فازی ارائه دادند و نام آن را فازی c- میانگین (FCM) نامیدند. یانگ و ناتالیانی^۶ [۱۱] یک توسعه نیرومند از FCM با اضافه کردن چندین جمله آنتروپی ارائه دادند. برای مطالعه بیشتر پیرامون

¹finite mixture of scale mixture of normal distribution

²Lin and et al.

³Wang and et al.

⁴Zadeh

⁵Gustafson and Kessel

⁶Yang and Nataliani.

کلاس‌بندی فازی می‌توان به منابع [۳]، [۲]، [۶]، [۸] و [۵] مراجعه کرد.

هدف اصلی از این مقاله توسعه کلاس بندی فازی براساس توزیع مخلوط متناهی آمیخته مقیاسی نرمال (FM-SMN) همراه با داده گمشده است. داده‌های گمشده امکان دارد در نقاط پرت اتفاق افتاده باشند و استنباط بدون در نظر گرفتن داده گمشده می‌تواند این آنالیز را دچار مشکل و انحراف کند.

با انگیزه‌ای که از این نتایج حاصل می‌شود، یک روش FCM توسعه یافته برای کلاس‌بندی داده‌های چند متغیره FM-SMN همراه با اطلاعات گمشده ارائه می‌دهیم. برای برآورد پارامترهای از الگوریتم یادگیری نیرومند FCM (LB-FCM) استفاده می‌کنیم. همچنین مشاهدات گمشده وابسته به خود مشاهدات هستند و هیچگونه وابستگی به دلیل گمشدگی آنها ندارند. برای سادگی در محاسبات، دو ماتریس معرفی می‌شوند که می‌توانند قسمت مشاهده شده و گمشده داده‌ها را از هم جدا کنند. این روش‌ها برای تحلیل داده‌های واقعی و شبیه سازی شده به کار می‌روند.

این مقاله به صورت زیر مرتب شده است. در بخش دوم توزیع FM-SMN معرفی می‌شود و حالت‌های خاص آن با ذکر بعضی خواص بیان می‌گردند. در بخش سوم پارامترهای مدل با روش LB-FCM برآورد شده است. در بخش چهارم و پنجم به ترتیب این روش‌ها برای شبیه‌سازی و یک مثال واقعی مورد استفاده و کارایی آن‌ها مورد بررسی قرار می‌گیرند.

۲ مروری بر توزیع آمیخته مقیاسی نرمال و مخلوط متناهی آن همراه با داده گمشده

در این بخش برخی نتایج اولیه و نمایش تصادفی از توزیع آمیخته مقیاسی نرمال را ارائه می‌دهیم. برای توضیحات بیشتر به اندرسون و مالوس^۷ [۱] مراجعه کنید.

متغیر تصادفی X دارای توزیع آمیخته مقیاسی نرمال SMN با پارامتر مکان μ ، ماتریس واریانس-کوواریانس Σ و پارامتر ν که با نماد $SMN(\mu, \Sigma, \nu)$ معرفی

⁷ Andrews and Mallows

می‌شود، هرگاه نمایش تصادفی آن به صورت زیر باشد:

$$X \stackrel{d}{=} \mu + W^{-1/2} X_1, \quad X_1 \perp W, \quad (1)$$

که در آن X_1 یک بردار تصادفی دارای توزیع نرمال استاندارد p -بعدی و W یک متغیر تصادفی مثبت مستقل از X_1 بوده که دارای تابع توزیع $H(\cdot; \nu)$ است. بنابراین می‌توان تابع چگالی بردار تصادفی X را به صورت

$$f_{SMN}(x; \mu, \Sigma, \nu) = \int_0^\infty \phi_p(x; \mu, w^{-1}\Sigma) dH(w; \nu) \quad x \in \mathbb{R}^p,$$

بدست آورد. با توزیع‌های مختلفی که برای W می‌توان در نظر گرفت موارد خاص از توزیع کلی SMN را بدست می‌آید. در این مقاله روی چند مورد رایج از موارد خاص توزیع SMN متمرکز هستیم.

- توزیع نرمال (N): اگر در کلاس توزیع SMN، توزیع W با احتمال ۱ برابر با ۱ باشد یعنی توزیع W تباہ در نقطه ۱ باشد توزیع SMN همان توزیع نرمال است.
- توزیع تی (t): اگر در کلاس توزیع SMN، توزیع W گامای با پارامتر مقیاس $\nu/2$ و پارامتر شکل $\nu/2$ باشد، آنگاه توزیع SMN همان توزیع تی است. توزیع تی را با نماد $X \sim t(\mu, \Sigma, \nu)$ نمایش می‌دهیم.
- توزیع اسلش (SL): اگر در کلاس توزیع SMN، توزیع W بتا با پارامتر ν و ۱ ($Beta(\nu, 1)$) باشد، آنگاه توزیع SMN همان توزیع اسلش است. توزیع اسلش را با نماد $X \sim SL(\mu, \Sigma, \nu)$ نمایش می‌دهیم.
- توزیع نرمال آلوده شده (CN): اگر در کلاس توزیع SMN، توزیع W دارای تابع چگالی زیر باشد،

$$h(w; \nu, \gamma) = \nu I_\gamma(w) + (1 - \nu) I_1(w),$$

بطوریکه $I_A(\cdot)$ تابع نشانگر روی مجموعه A باشد، آنگاه توزیع SMN

همان توزیع نرمال آلوده شده است. توزیع نرمال آلوده شده را با نماد $X \sim CN(\mu, \Sigma, \nu, \gamma)$ نمایش می‌دهیم.

برای مشاهده تابع چگالی‌های مورد استفاده به پیوست مراجعه شود. حال متغیر تصادفی X دارای توزیع مخلوط از توزیع SMN است هرگاه تابع چگالی آن به صورت زیر باشد:

$$f(x_i|\Theta) = \sum_{k=1}^c \pi_k f_{SMN}(x_i; \theta_k), \quad i = 1, \dots, n, \quad (2)$$

که در آن $\pi = (\pi_1, \dots, \pi_c)^T$ بردار وزن‌های آمیخته‌ی متناظر، با شرط $\sum_{i=1}^c \pi_i = 1$ بوده و برای $\theta_k = (\mu_k, \Sigma_k, \nu_k)$ داریم $\Theta = (\pi_1, \dots, \pi_{c-1}, \theta_1, \dots, \theta_c)$. در ادامه، این مدل به صورت FM-SMN نشان داده خواهد شد.

برای سادگی در محاسبات دو ماتریس $O_j(p_j^o \times p)$ و $M_j((p - p_j^o) \times p)$ معرفی می‌شوند بطوریکه $X_j^o = O_j X_j$ و $X_j^m = M_j X_j$ به ترتیب قسمت مشاهده شده و قسمت گمشده داده‌ها باشند. به راحتی می‌توان نشان داد که

$$O_j^T X_j^o + M_j^T X_j^m = X_j, \quad O_j^T O_j + M_j^T M_j = I,$$

است. ابتدا متغیر تخصیص^۸ را به صورت $Z_j = (Z_{1j}, \dots, Z_{cj})$ در نظر می‌گیریم که در آن برای $j = 1, \dots, n$ هر یک از z_{ij} برابر ۱ خواهد بود اگر X_j متعلق به i امین جامعه باشد و در غیر این صورت $Z_{ij} = 0$. در نتیجه می‌توان بیان کرد که Z_j ها مستقل از یکدیگر دارای توزیع چند جمله‌ای^۹ با یک آزمایش و احتمالات (π_1, \dots, π_c) خواهد بود که با $Z_j \sim \text{Multi}(1; \pi_1, \dots, \pi_c)$ نشان داده می‌شود. بنابراین می‌توان نمایش تصادفی زیر را برای مدل (۲) در نظر گرفت:

$$X_j | (Z_{ij} = 1) \sim \text{SMN}(\mu_i, \Sigma_i, \nu_i), \quad Z_i \sim \text{Multi}(1, \pi_1, \dots, \pi_c).$$

قضیه زیر برای امید ریاضی‌های شرطی استفاده شده در الگوریتم LB-FCM ضروری و

⁸Component

⁹Multinomial distribution

مفید است.

قضیه ۱.۲. در نظر می گیریم که مدل (۲) برقرار باشد. آنگاه داریم:

(۱) توزیع شرطی X_j^o به شرط $W_j = w_j$ و $Z_{ij} = 1$ به صورت

$$X_j^o | (w_j, Z_{ij} = 1) \sim N_{p_j^o}(\mu_{ij}^o, w_j^{-1} \Sigma_{ij}^{oo}),$$

است، به طوریکه $\mu_{ij}^o = O_j \mu_i$ و $\Sigma_{ij}^{oo} = O_j \Sigma_i O_j^T$ می باشند.

(۲) توزیع حاشیه ای X_j^o به صورت زیر حاصل می شود:

$$f(x_j^o | \Theta) = \sum_{i=1}^c \pi_i f_{SMN}(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo}, \nu), \quad j = 1, \dots, n,$$

(۳) توزیع شرطی X_j^m به شرط $X_j^o = x_j^o$ و $Z_{ij} = 1$ به صورت

$$X_j^m | (x_j^o, w_j, Z_{ij} = 1) \sim N_{p-p_j^o}(\mu_{ij}^{m.o}, u_j^{-1} \Sigma_{ij}^{mm.o}),$$

است، که در آن $\mu_{ij}^{m.o} = M_j (\mu_i + \Sigma_i S_{ij}^{oo} (y_j - \mu_i))$ و $\Sigma_{ij}^{mm.o} = M_j (I_p - \Sigma_i S_{ij}^{oo}) \Sigma_i M_j^T$ می باشند.

همچنین توزیع شرطی W_j به شرط $X_j^o = x_j^o$ و $Z_{ij} = 1$ و امید ریاضی شرطی های مورد نیاز برای هر یک از موارد خواص توزیع SMN در پیوست الف ارائه شده است.

۳ الگوریتم کلاس بندی LB-FCM

در این بخش یک نوع الگوریتم کلاس بندی تعمیم یافته FCM را برای توزیع SMN ارائه خواهیم داد.

مطابق یانگ و ناتالیانی^{۱۰} [۱۱]، تابع هدف را برای توزیع SMN برای داده های

¹⁰Yang and Nataliani

کامل به صورت

$$\begin{aligned}
 J(U, \Theta) &= \sum_{i=1}^c \sum_{j=1}^n u_{ij} \left(\frac{p}{\nu} \log \nu \pi + \frac{1}{\nu} \log |\Sigma_i^{-1}| + \frac{p-1}{\nu} \hat{\kappa}_{ij}^{(k)} \right) \\
 &+ \frac{1}{\nu} \text{tr} \left(\Sigma_i^{-1} \hat{\Psi}_{ij}^{(k)} \right) - \hat{\varphi}_{ij}^{(k)} \Big) - r_3 n \sum_{i=1}^c \alpha_i \log \alpha_i \\
 &- \sum_{i=1}^c \sum_{j=1}^n (r_1 u_{ij} \log \alpha_i + r_2 u_{ij} \log u_{ij}), \tag{۳}
 \end{aligned}$$

بیان می‌کنیم، بطوریکه $r_1, r_2, r_3 > 0$

$$\begin{aligned}
 \hat{\Psi}_{ij}^{(k)} &= \hat{w}_{ij}^{(k)} (\hat{x}_{ij}^{(k)} - \mu_i) (\hat{x}_{ij}^{(k)} - \mu_i)^\top + (I_p - \hat{\Sigma}_i^{(k)} \hat{S}_{ij}^{oo(k)}) \hat{\Sigma}_i^{(k)}, \\
 \hat{x}_{ij}^{(k)} &= \hat{\mu}_i^{(k)} + \hat{\Sigma}_i^{(k)} \hat{S}_{ij}^{oo(k)} (x_j - \hat{\mu}_i^{(k)}), \quad \hat{w}_{ij}^{(k)} = E(W_j | x_j^o, Z_{ij} = 1, \hat{\Theta}^{(k)}), \\
 \hat{\varphi}_{ij}^{(k)} &= E(h(w_j; \nu_i) | x_j^o, Z_{ij} = 1, \hat{\Theta}^{(k)}), \quad \hat{\kappa}_{ij}^{(k)} = E(\log(W_j) | x_j^o, Z_{ij} = 1, \hat{\Theta}^{(k)}),
 \end{aligned}$$

و u_{ij} تابع عضویت فازی باشد که شرایط زیر را نیز دارا است:

$$0 \leq u_{ij} \leq 1, \quad \forall i, j, \quad \sum_{i=1}^c u_{ij} = 1, \quad 0 < \sum_{j=1}^n u_{ij} < n. \tag{۴}$$

تابع لاگرانژ مربوط به (۳) به صورت

$$\begin{aligned}
 \tilde{J}(U, \tau, \Theta) &= \sum_{i=1}^c \sum_{j=1}^n u_{ij} \left(\frac{p}{\nu} \log \nu \pi + \frac{1}{\nu} \log |\Sigma_i^{-1}| + \frac{p-1}{\nu} \hat{\kappa}_{ij}^{(k)} \right) \\
 &+ \frac{1}{\nu} \text{tr} \left(\Sigma_i^{-1} \hat{\Psi}_{ij}^{(k)} \right) - \hat{\varphi}_{ij}^{(k)} \Big) - r_1 \sum_{i=1}^c \sum_{j=1}^n u_{ij} \log \alpha_i \\
 &- r_3 n \sum_{i=1}^c \alpha_i \log \alpha_i, - \sum_{j=1}^n \tau_j \left(\sum_{i=1}^c u_{ij} - 1 \right) \\
 &- \tau_{n+1} \left(\sum_{i=1}^c \alpha_i - 1 \right) + r_2 \sum_{i=1}^c \sum_{j=1}^n u_{ij} \log u_{ij}, \tag{۵}
 \end{aligned}$$

است که در آن $\tau = (\tau_1, \dots, \tau_{n+1})$ ضرایب لاگرانژ می باشند. مقادیر به روز شده تابع عضویت فازی u_{ij} با مشتق گیری از (۵) و برابری با صفر، به صورت زیر محاسبه خواهند شد:

$$\hat{u}_{ij}^{(k)} = \exp \left\{ \left(\hat{d}_{ij}^{(k)} + r_1 \log \hat{\alpha}_i^{(k)} \right) / r_2 \right\} / \sum_{i=1}^c \exp \left\{ \left(\hat{d}_{ij}^{(k)} + r_1 \log \hat{\alpha}_i^{(k)} \right) / r_2 \right\}, \quad (6)$$

بطوریکه

$$\hat{d}_{ij}^{(k)} = \left(\frac{p}{r_1} \log 2\pi + \frac{1}{r_1} \log |\Sigma_i^{-1}| + \frac{p-1}{r_1} \hat{\kappa}_{ij}^{(k)} + \frac{1}{r_1} \text{tr} \left(\Sigma_i^{-1} \hat{\Psi}_{ij}^{(k)} \right) - \hat{\varphi}_{ij}^{(k)} \right). \quad (7)$$

با مینیمم سازی (۵) نسبت به α_i با در نظر گرفتن شرط $\sum_{i=1}^c \alpha_i = 1$ داریم:

$$\hat{\alpha}_i^{(new)} = \frac{1}{n} \sum_{j=1}^n \hat{u}_{ij} + \frac{r_2}{r_1} \hat{\alpha}_i^{(old)} \left(\log \hat{\alpha}_i^{(old)} - \sum_{i=1}^c \hat{\alpha}_i^{(old)} \log \hat{\alpha}_i^{(old)} \right).$$

علاوه بر این، تعداد جدید کلاس ها به صورت زیر بدست می آیند:

$$c^{(new)} = c^{(old)} - |\{ \hat{\alpha}_i^{(old)} : \hat{\alpha}_i^{(old)} < 1/n, i \in \{1, \dots, c^{(old)}\} \}|. \quad (8)$$

حال برای اینکه $\hat{\alpha}_i^{(new)}$ و \hat{u}_{ij} بدست آمده شرایط اصلی خود را بر اساس تعداد کلاسهای جدید دارا باشند، تغییر زیر را روی آنها اعمال می کنیم.

$$\hat{\alpha}_i^{(new)} = \frac{\hat{\alpha}_i^{(new)}}{\hat{\alpha}_1^{(new)} + \dots + \hat{\alpha}_{c^{(new)}}^{(new)}} \quad (9)$$

و

$$\hat{u}_{ij} = \frac{\hat{u}_{ij}}{\hat{u}_{1j} + \dots + \hat{u}_{c^{(new)}j}}, \quad (10)$$

بطوریکه $\hat{u}_{1j} + \dots + \hat{u}_{c^{(new)}j} = 1$ و $\hat{\alpha}_1^{(new)} + \dots + \hat{\alpha}_{c^{(new)}}^{(new)} = 1$ و ناتالیانی [۱۱] استفاده از $r_1^{(k)} = e^{-k/10}$ ، $r_2^{(k)} = e^{-k/10}$ و

$$r_3 = \min \left[\frac{1}{c} \sum_{i=1}^c \exp\{-\eta n |\hat{\alpha}_i^{(new)} - \hat{\alpha}_i^{(old)}|\}, \frac{1 - \max_{1 \leq i \leq c} \left(\frac{1}{n} \sum_{j=1}^n \hat{u}_{ij} \right)}{-\hat{\pi}_{(c)}^{(old)} E} \right], \quad (11)$$

را توصیه کردند به طوری که $\hat{\alpha}_{(c)}^{(old)} = E = \sum_{i=1}^c \hat{\alpha}_i^{(old)} \log \hat{\alpha}_i^{(old)}$ و $\max(\hat{\alpha}_1^{(old)}, \dots, \hat{\alpha}_c^{(old)})$ و $\eta = \min\{1, \delta^{\lfloor p/2-1 \rfloor}\}$ با a که قسمت صحیح عدد a است.

با مینیم سازی (۵) نسبت به μ_i داریم:

$$\hat{\mu}_i^{(k+1)} = \frac{\sum_{j=1}^n \hat{u}_{ij}^{(k)} \hat{x}_{ij}^{(k)}}{\sum_{j=1}^n \hat{u}_{ij}^{(k)} \hat{w}_{ij}^{(k)}}. \quad (12)$$

با مینیم سازی تابع هدف (۵) نسبت به پارامتر Σ_i و به ازای مقدار به روز شده $\hat{\mu}_i^{(k+1)}$ برآورد Σ_i به صورت

$$\hat{\Sigma}_i^{(k+1)} = \frac{\sum_{j=1}^n \hat{u}_{ij}^{(k)} \hat{\Psi}_{ij}^{(k)}}{\sum_{j=1}^n \hat{u}_{ij}^{(k)}}. \quad (13)$$

حاصل می شود. پارامتر ν_i نیز با در نظر گرفتن مقادیر $\hat{\mu}_i^{(k+1)}$ و $\hat{\Sigma}_i^{(k+1)}$ و با مینیم سازی تابع هدف (۵) به صورت زیر به روزرسانی می شود. بنابراین داریم

$$\hat{\nu}_i^{(k+1)} = \arg \min_{\nu_i} \tilde{J}(U, \tau, \Theta) \Big|_{\mu_i = \hat{\mu}_i^{(k+1)}, \Sigma_i = \hat{\Sigma}_i^{(k+1)}}. \quad (14)$$

برای حالتی که $\nu_1 = \dots = \nu_c = \nu$ هستند پارامتر ν با ماکزیمم سازی تابع زیر انجام می شود.

$$\hat{\nu}^{(k+1)} = \arg \max_{\nu} \sum_{j=1}^n \log \left[\sum_{i=1}^c \hat{\alpha}_i^{(k+1)} f_{SMN_{p_j}^o}(x_j; \hat{\mu}_i^{(k+1)}, \hat{\Sigma}_i^{(k+1)}, \nu) \right]$$

نقاط اولیه $c^{(o)} = n$ و $\alpha_i = 1/n$ را در نظر می گیریم. همچنین برای سایر پارامترها با استفاده از یانگ و همکاران [۳] داریم:

$$\mu_i^{(o)} = x_j, \quad \Sigma_i^{(o)} = d_{i(\lfloor \sqrt{c^{(o)}} \rfloor)}^{\chi} I_d, \quad (15)$$

به طوری که I_d یک ماتریس $d \times d$ و

$$D_i = \text{sort}\{d_{ij}^{\chi} = \|x_j - \mu_i\|^2 : d_{ij}^{\chi} > 0, \\ i \neq j, 1 \leq i, j \leq n\} = \{d_{i(1)}^{\chi}, d_{i(2)}^{\chi}, \dots, d_{i(n)}^{\chi}\}.$$

است. الگوریتم LB-FCM را می توان به صورت زیر مرتب و اجرا کرد.

الگوریتم LB-FCM:

(۱) در نظر می گیریم $\epsilon = 10^{-5}$ و مجموعه مقادیر اولیه را به صورت

$$c^{(o)} = n, \alpha_i = 1/n, \mu_i^{(o)} = x_i, \Sigma_i^{(o)}, r_1^{(o)} = r_p^{(o)} = r_{\bar{p}}^{(o)} = 1,$$

می باشند. توجه داشته باشید که مقدار $\nu_i^{(o)}$ مطابق هر توزیع خاص در توزیع SMN در نظر گرفته می شود.

(۲) محاسبه $\hat{U}^{(o)} = \{\hat{u}_{ij}^{(o)}\}$ با استفاده از (۶)، $\hat{w}^{(o)} = \{\hat{w}_{ij}^{(o)}\}$ و $\hat{\kappa}^{(o)} = \{\hat{\kappa}_{ij}^{(o)}\}$ با استفاده از پیوست الف با $\hat{\Theta}^{(o)}$ و $k = 1$.

(۳) به روز رسانی $\hat{\Theta}^{(k)}$ با استفاده از $(\hat{U}^{(k-1)}, \hat{w}^{(k-1)}, \hat{\kappa}^{(k-1)}, \hat{x}^{(k-1)})$.

(۴) محاسبه $\hat{u}_{ij}^{(k)}$ با استفاده از $(c^{(k-1)}, \hat{\alpha}_i^{(k-1)}, \hat{\mu}_i^{(k-1)}, \hat{\Sigma}_i^{(k-1)}, \hat{\nu}_i^{(k-1)}, r_1^{(k-1)})$.

$$r_{\psi}^{(k-1)}.$$

(۵) محاسبه $\hat{\alpha}^{(k)}$ از (۶) با استفاده $\hat{U}^{(k-1)}$ ، $\hat{\alpha}^{(k-1)}$ ، $\hat{r}_{\psi}^{(k-1)}$ و $\hat{r}_{\psi}^{(k-1)}$.

(۶) به روز رسانی $r_{\psi}^{(k)} = e^{-k/100}$ ، $r_{\psi}^{(k)} = e^{-k/10}$ و $r_{\psi}^{(k)}$ با استفاده از $\hat{\alpha}^{(k)}$ و $\hat{\alpha}^{(k-1)}$ در رابطه (۱۱).

(۷) به روز رسانی $c^{(k)}$ با استفاده از $\hat{\alpha}_i^{(k)} < 1/n$ که $\hat{\alpha}_i^{(k)}$ و $\hat{u}_{ij}^{(k)}$ قبلا به ترتیب با استفاده از روابط (۸) و (۱۰) سازگار شده‌اند. اگر $k \geq 60$ و $c^{(k-60)} - c^{(k)} = 0$ آنگاه $\hat{r}_{\psi}^{(k)} = 0$ در نظر می‌گیریم.

(۸) به روز رسانی $\hat{U}^{(k)}$ با استفاده از (۶)، $\hat{w}^{(k)}$ و $\hat{\kappa}^{(k)}$ با استفاده از پیوست الف با $\hat{\Theta}^{(k)}$.

(۹) به روز رسانی $\hat{\Theta}^{(k+1)}$ با استفاده از $\hat{U}^{(k)}$ ، $\hat{w}^{(k)}$ ، $\hat{\kappa}^{(k)}$ ، $\hat{x}^{(k)}$ و روابط (۹) و (۱۲)–(۱۴).

(۱۰) اگر $\max_{1 \leq i \leq c^{(k)}} \|\mu_i^{(k+1)} - \mu_i^{(k)}\| < \epsilon$ باشد حلقه را متوقف کرده و در غیر این صورت با افزایش یک مقدار به k به مرحله ۵ برمی‌گردیم و این حلقه را تا حصول نتیجه دلخواه ادامه می‌دهیم.

به عنوان نتیجه ای از مراحل بالا می‌توان پیش‌بینی نقاط گمشده را با استفاده از فرمول زیر بدست آورد:

$$\hat{x}_j^m = M_j \sum_{i=1}^c \hat{\alpha}_i \left(\hat{\mu}_i + \sum_i \hat{S}_{ij}^{oo} (x_j - \hat{\mu}_i) \right) \quad (16)$$

۴ مطالعه شبیه‌سازی

در این قسمت، سه سناریو شبیه‌سازی برای بررسی عملکرد مدل FM-SMN و همچنین برآورد پارامترها براساس الگوریتم LB-FCM در نظر گرفته شده است. همچنین برای هر مدل و در هر الگوریتم مورد استفاده، دو شاخص نرخ بد رده‌بندی^{۱۱} (MCR) و شاخص

^{۱۱} Misclassification rate

رتبه‌ی اصلاح شده^{۱۲} (ARI) را مورد محاسبه قرار داده‌ایم. لازم به توضیح است که هرچه معیار MCR کوچکتر و ARI بزرگتر باشد، گروه‌بندی انجام شده به وسیله‌ی مدل برازش داده شده به گروه‌بندی واقعی داده‌ها شبیه‌تر خواهد بود.

۱.۴ شبیه‌سازی ۱

در این بخش با استفاده از شبیه‌سازی نشان می‌دهیم الگوریتم LB-FCM کلاس‌بندی داده‌ها را با وجود داده پرت به خوبی نسبت به الگوریتم EM مشخص می‌کند. برای این شبیه‌سازی، به تعداد ۱۰۰ بار نمونه‌هایی به حجم ۳۰۰ از توزیع نرمال دومتغیره با دو جزء و با پارامترهای $\mu_1 = (0, 4)^T$ ، $\mu_2 = (0, 7)^T$ ، $\Sigma_1 = \Sigma_2 = \begin{pmatrix} 1 & 0.2 \\ 0.2 & 1 \end{pmatrix}$ و $\pi_1 = \pi_2 = 1/2$ تولید شده است. برای وجود داده پرت در این داده‌ها، در هر تکرار به داده‌های تولید شده دو سطح ۲٪ و ۴٪ از توزیع یکنواخت در بازه $U = ([3, 4], [5, 6])^T$ تولید و به داده‌های تولید شده اضافه می‌کنیم. همچنین به این داده‌ها ۲۰٪ داده گمشده با روش MAR می‌دهیم. هر یک از توزیع‌های SMN بر روی این داده‌ها با استفاده از الگوریتم LB-FCM پیاده‌سازی شدند و شاخص‌های ARI و MCR آن‌ها مورد بررسی قرار گرفت. همچنین تعداد تکرار الگوریتم برای رسیدن به تعداد کلاس صحیح نیز مورد بررسی قرار گرفت. جدول ۱ نتایج مربوط به میانگین شاخص‌های ARI و MCR را برای ۱۰۰ تکرار نشان می‌دهد. با توجه به جدول می‌توان نتیجه گرفت که توزیع تی بهتر از توزیع‌های این کلاس رتبه بندی صحیح را انجام داده است و در مقایسه با حالتی که داده پرت وجود ندارد شاخص‌های ARI و MCR تفاوت زیادی با دو حالت دیگر ندارند که این نشان می‌دهد در مواجهه با داده پرت به خوبی این الگوریتم برخورد می‌کند. همچنین جدول ۲ میانگین تعداد تکرار لازم برای رسیدن به تعداد کلاس صحیح را نشان داده است. از این جدول نیز نتیجه می‌شود توزیع نرمال آلوده شده زودتر از سایر توزیع‌ها به تعداد کلاس صحیح رسیده است. علاوه بر آن دیده می‌شود که الگوریتم LB-FCM به تعداد تکرار کمتری برای بدست آوردن تعداد کلاس صحیح نسبت به الگوریتم EM نیاز دارد. در انتها نیز شکل ۱ آورده شده است که برای یکی از داده‌های تولید شده کلاس‌بندی داده‌ها را

¹² Adjusted rank index

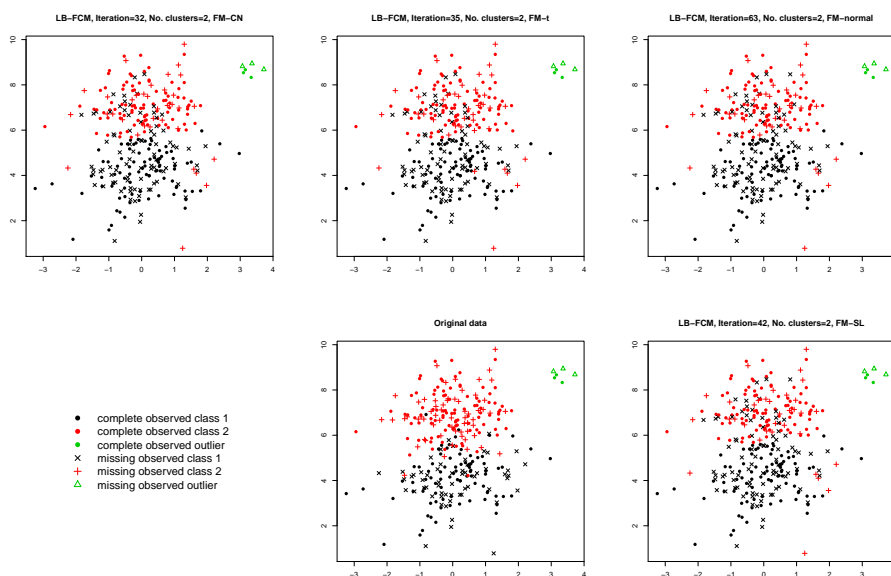
نشان می‌دهد.

جدول ۱: میانگین شاخص‌های *ARI* و *MCR* برای الگوریتم‌های *EM* و *LB-FCM*.

FM-CN		FM-SL		FM-t		FM-normal		معیار کارایی	درصد داده پرت
LB-FCM	EM	LB-FCM	EM	LB-FCM	EM	LB-FCM	EM		
۰٫۷۷۹	۰٫۷۴۵	۰٫۷۹۱	۰٫۷۵۰	۰٫۸۰۷	۰٫۷۷۳	۰٫۷۸۱	۰٫۷۴۰	ARI	۰٪
۰٫۳۰۵	۰٫۳۱۰	۰٫۲۹۷	۰٫۳۰۸	۰٫۲۲۶	۰٫۲۸۶	۰٫۲۹۳	۰٫۳۳۴	MCR	
۰٫۷۵۵	۰٫۶۹۶	۰٫۷۳۳	۰٫۷۰۲	۰٫۷۷۰	۰٫۷۲۱	۰٫۷۱۲	۰٫۶۸۱	ARI	۲٪
۰٫۳۴۴	۰٫۳۹۰	۰٫۳۲۵	۰٫۳۷۸	۰٫۲۷۲	۰٫۳۳۹	۰٫۴۰۳	۰٫۴۴۵	MCR	
۰٫۷۰۳	۰٫۶۵۲	۰٫۷۱۸	۰٫۶۷۹	۰٫۷۵۹	۰٫۷۱۱	۰٫۶۹۳	۰٫۶۵۲	ARI	۴٪
۰٫۳۷۴	۰٫۴۳۳	۰٫۳۸۰	۰٫۴۱۵	۰٫۳۲۲	۰٫۳۷۵	۰٫۴۶۸	۰٫۵۱۲	MCR	

جدول ۲: میانگین تعداد تکرار تا حصول کلاس صحیح برای الگوریتم‌های *EM* و *LB-FCM*.

FM-CN		FM-SL		FM-t		FM-normal		درصد داده پرت
LB-FCM	EM	LB-FCM	EM	LB-FCM	EM	LB-FCM	EM	
۳۵	۳۳۹	۵۳	۳۱۴	۴۲	۲۹۰	۵۷	۳۸۳	۰٪
۳۹	۳۷۴	۵۹	۳۷۳	۴۷	۳۱۰	۶۶	۴۲۲	۲٪
۵۵	۴۶۱	۷۶	۵۱۰	۶۶	۴۸۰	۷۹	۵۳۵	۴٪



شکل ۱: نتایج کلاس‌بندی برای یکی از نمونه‌ها با استفاده از الگوریتم *LB-FCM*.

۲.۴ شبیه‌سازی ۲

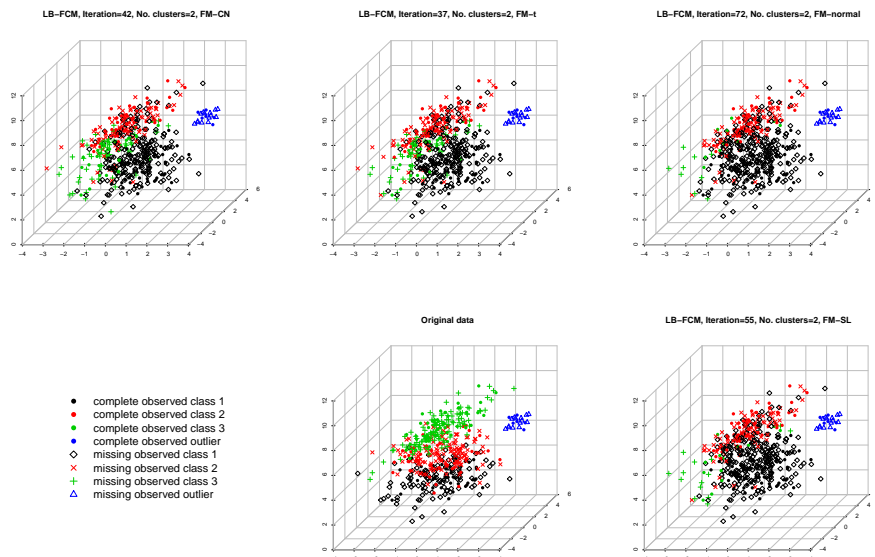
در این بخش با استفاده از شبیه‌سازی نشان می‌دهیم الگوریتم LB-FCM کلاس‌بندی داده‌ها را با وجود نرخ بالاتر داده گمشده و پرت به خوبی نسبت به الگوریتم EM در مواجهه با داده‌ها تعداد زیاد مشخص می‌کند. برای این شبیه‌سازی، به تعداد ۱۰۰ بار نمونه‌هایی به حجم ۵۰۰ از توزیع نرمال سه متغیره با سه جزء و با پارامترهای $\mu_1 = (0, 0, 4)^T$ ، $\mu_2 = (0, 0, 6)^T$ ، $\mu_3 = (0, 0, 8)^T$ ، $\Sigma_1 = \Sigma_2 = \begin{pmatrix} 1 & 0.5 & 0 \\ 0.5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ، $\Sigma_3 = \begin{pmatrix} 1 & 0.2 & 0.8 \\ 0.2 & 1 & 0.5 \\ 0.8 & 0.5 & 1 \end{pmatrix}$ و $\pi_1 = \pi_2 = \pi_3 = 1/3$ تولید شده است. برای وجود داده پرت در این داده‌ها، در هر تکرار به داده‌های تولید شده دو سطح ۶٪ و ۱۰٪ از توزیع یکنواخت در بازه $U = ([4, 5], [6, 7], [8, 9])^T$ تولید و به داده‌های تولید شده اضافه می‌کنیم. همچنین به این داده‌ها ۳۰٪ داده گمشده با روش MAR می‌دهیم. هر یک از توزیع‌های SMN بر روی این داده‌ها با استفاده از الگوریتم LB-FCM پیاده‌سازی شدند و شاخص‌های ARI و MCR آنها مورد بررسی قرار گرفت. همچنین تعداد تکرار الگوریتم برای رسیدن به تعداد کلاس صحیح نیز مورد بررسی قرار گرفت. جدول ۵ نتایج مربوط به میانگین شاخص‌های ARI و MCR را برای ۱۰۰ تکرار نشان می‌دهد. با توجه به جدول می‌توان نتیجه گرفت که توزیع نرمال آلوده شده بهتر از توزیع‌های این کلاس رتبه بندی صحیح را انجام داده است و در مقایسه با حالتی که داده پرت وجود ندارد شاخص‌های ARI و MCR تفاوت زیادی با دو حالت دیگر ندارند که این نشان می‌دهد در مواجهه با داده پرت به خوبی این الگوریتم برخورد می‌کند. همچنین جدول ۴ میانگین تعداد تکرار لازم برای رسیدن به تعداد کلاس صحیح را نشان داده است. از این جدول نیز نتیجه می‌شود توزیع تی شده زودتر از سایر توزیع‌ها به تعداد کلاس صحیح رسیده است. علاوه بر آن دیده می‌شود که الگوریتم LB-FCM به تعداد تکرار کمتری برای بدست آوردن تعداد کلاس صحیح نسبت به الگوریتم EM نیاز دارد. در انتها نیز شکل ۲ آورده شده است که برای یکی از داده‌های تولید شده کلاس‌بندی داده‌ها را نشان می‌دهد.

جدول ۳: میانگین شاخص‌های *ARI* و *MCR* برای الگوریتم‌های *EM* و *LB-FCM*.

FM-CN		FM-SL		FM-t		FM-normal		معیار کارایی	درصد داده پرت
LB-FCM	EM	LB-FCM	EM	LB-FCM	EM	LB-FCM	EM		
۰٫۶۳۰	۰٫۵۵۶	۰٫۵۵۹	۰٫۵۱۲	۰٫۶۰۲	۰٫۵۴۰	۰٫۵۳۰	۰٫۵۰۳	<i>ARI</i>	۰٪
۰٫۳۷۸	۰٫۴۳۰	۰٫۴۱۵	۰٫۴۴۴	۰٫۳۹۹	۰٫۴۱۳	۰٫۴۳۷	۰٫۴۶۲	<i>MCR</i>	
۰٫۵۷۳	۰٫۴۵۹	۰٫۴۸۵	۰٫۴۰۱	۰٫۵۲۰	۰٫۴۳۵	۰٫۴۱۴	۰٫۳۷۲	<i>ARI</i>	۶٪
۰٫۳۹۲	۰٫۴۷۱	۰٫۴۷۳	۰٫۵۳۶	۰٫۴۱۸	۰٫۵۱۲	۰٫۵۱۹	۰٫۵۵۱	<i>MCR</i>	
۰٫۵۵۳	۰٫۴۳۱	۰٫۴۶۸	۰٫۳۹۰	۰٫۵۰۴	۰٫۴۱۹	۰٫۳۹۰	۰٫۳۲۶	<i>ARI</i>	۱۰٪
۰٫۴۱۰	۰٫۴۸۶	۰٫۴۸۰	۰٫۵۶۱	۰٫۴۳۹	۰٫۵۵۳	۰٫۵۷۳	۰٫۶۱۲	<i>MCR</i>	

جدول ۴: میانگین تعداد تکرار تا حصول کلاس صحیح برای الگوریتم‌های *EM* و *LB-FCM*.

FM-CN		FM-SL		FM-t		FM-normal		درصد داده پرت
LB-FCM	EM	LB-FCM	EM	LB-FCM	EM	LB-FCM	EM	
۴۳	۳۸۸	۵۱	۳۵۸	۳۵	۳۷۵	۶۰	۵۳۱	۰٪
۴۹	۴۱۷	۵۷	۳۹۶	۴۳	۴۰۱	۶۸	۵۷۰	۶٪
۷۱	۴۸۵	۷۴	۴۸۳	۶۸	۵۲۰	۸۲	۷۳۰	۱۰٪



شکل ۲: نتایج کلاس‌بندی برای یکی از نمونه‌ها با استفاده از الگوریتم *LB-FCM*.

۳.۴ شبیه‌سازی ۳

در این قسمت، به منظور بررسی خواص مجانبی برآوردگرهای بیان شده با استفاده از شبیه‌سازی توزیع آمیخته‌تی دو متغیره با دو جزء داده تولید شده و دو ویژگی ارببی و میانگین مربعات خطا مورد بررسی قرار گرفته است. در این مطالعه مقادیر پارامتر

به صورت زیر فرض شده است:

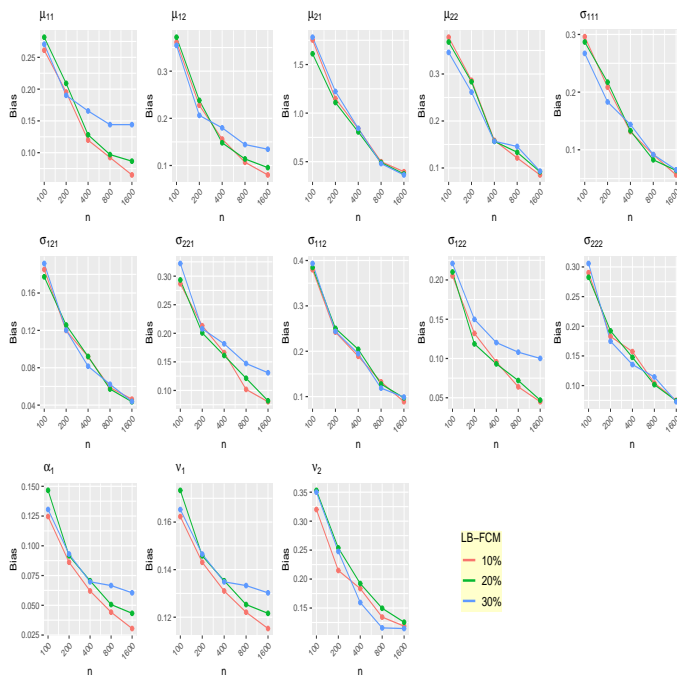
$\Sigma_2 = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}, \Sigma_1 = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}, \mu_2 = (5, 7)^T, \mu_1 = (2, 3)^T$
 $\nu_1 = 5, \nu_2 = 13$ و $\pi_1 = \pi_2 = 1/2$. دو معیار اریبی (Bias) و میانگین مربعات
 خطای (MSE) پارامترها برای ۱۰۰ تکرار به صورت زیر محاسبه شده‌اند:

$$\text{Bias} = \frac{1}{100} \sum_{i=1}^{100} (\hat{\theta}_i - \theta_{true}) \quad \text{و} \quad \text{MSE} = \frac{1}{100} \sum_{i=1}^{100} (\hat{\theta}_i - \theta_{true})^2,$$

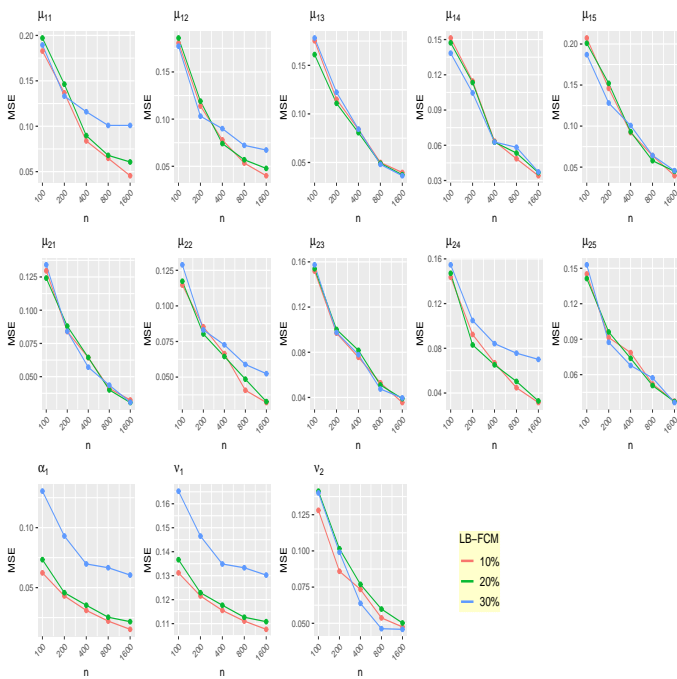
که در آن $\hat{\theta}_i$ نشانگر برآورد پارامتر مورد نظر در تکرار i می‌باشد. حال برای اندازه
 نمونه‌های $n = 100, 200, 400, 800, 1600, 5000$ به تعداد ۱۰۰ بار از توزیع تی آمیخته
 با پارامترهای بیان شده داده تولید شده و مقادیر اریبی و میانگین مربعات خطا برای هر
 یک از پارامترها محاسبه شده است. همچنین سه سطح ۱۰٪، ۲۰٪ و ۳۰٪ برای داده
 گمشده در نظر گرفته شده است. در شکل‌های ۳ و ۴ نتایج این شبیه‌سازی خلاصه شده
 است. همانطور که از این شکل‌ها ملاحظه می‌شود، با افزایش حجم نمونه مقدار Bias و
 MSE کاهش پیدا می‌کند.

۵ بازسازی تصویر

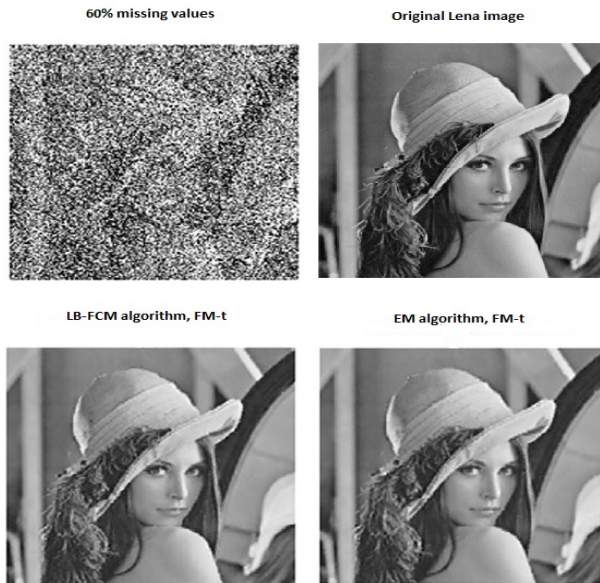
در این بخش الگوریتم بیان شده را برای توزیع SMN معرفی شده را در مدل‌سازی داده‌های
 واقعی مورد بررسی قرار خواهیم داد. همچنین هدف اصلی در این بخش کلاس‌بندی
 صحیح داده‌های چند متغیره در حضور داده گمشده است تا کارایی الگوریتم بیان شده
 مورد بررسی قرار گیرد. داده‌های مربوط به تبدیل عکس لنا به پیکسل‌های آن می‌باشند.
 این داده‌ها شامل 512×512 پیکسل هستند که با استفاده از نرم افزار Matlab آماده
 شده‌اند. در این داده‌ها ۶۰٪ داده گمشده به روش MAR به وجود آمده است. این داده‌ها
 به صورت یک نمونه 4096 تایی از بردارهایی با بعد ۶۴ نمایش داده می‌شوند. در این
 جا هر یک از توزیع‌ها با دو الگوریتم EM و LB-FCM بر روی این داده‌ها پیاده‌سازی
 می‌شوند. در بازسازی تصویر با استفاده از دو روش EM و LB-FCM بهترین توزیع
 توزیع تی است. شکل ۵ نشان می‌دهد که در این توزیع استفاده از روش LB-FCM بهتر



شکل ۳: میانگین مقدار اریبی برای هر یک از پارامترها با استفاده از الگوریتم $LB-FCM$ برای توزیع تی.



شکل ۴: میانگین میانگین مربعات خطا برای هر یک از پارامترها با استفاده از الگوریتم LB-FCM برای توزیع تی.



شکل ۵: بازسازی تصویر لنا با ۶۰ درصد داده گمشده با دو روش EM و $LB-FCM$ برای توزیع تی.

از روش EM است زیرا سرعت به دست آوردن تعداد گروه‌های مناسب سریع تر و شکل به دست آمده نیز بهتر است. از معایب روش EM این است که برای داده‌های که اصلاً هیچ اطلاعی از تعداد دسته‌های صحیح نداریم، بدست آوردن تعداد مناسب دسته‌ها کار سختی است ولی در روش $LB-FCM$ الگوریتم به طور هوشمند این کار را انجام می‌دهد.

۶ نتیجه‌گیری

در این مقاله ضمن بیان کلاس توزیع‌های SMN و بررسی برخی ویژگی‌های آن‌ها، برآورد پارامتر و کلاس‌بندی داده‌ها بر اساس الگوریتم $LB-FCM$ در حضور داده گمشده بیان و ویژگی‌های آنها مورد تجزیه و تحلیل قرار گرفت. همانطور که با چند مثال واقعی و شبیه‌سازی نشان داده شده است، استفاده از الگوریتم $LB-FCM$ می‌تواند به عنوان یک رقیب و جایگزین برای سایر الگوریتم‌های پیشنهادی برای کلاس‌بندی مطرح شده در سال‌های اخیر باشد. همچنین مطالعات شبیه‌سازی انجام شده تحت این الگوریتم در مقابل الگوریتم EM و مدل‌های دیگر نشان می‌دهد که کلاس‌بندی با این شیوه دارای

ویژگی‌های مطلوب در برازش کلاس صحیح داده‌ها می‌باشد. همان طور که بیان شد استفاده از الگوریتم EM در حالتی که تعداد کلاس صحیح داده‌ها مشخص نباشد مناسب نیست زیرا در این الگوریتم نیاز است با شاخص‌هایی ابتدا تعداد کلاس صحیح را بدست آورد که این موضوع نیاز به وقت زیاد دارد. از محاسن استفاده از الگوریتم LB-FCM می‌توان به سرعت زیاد بدست آوردن تعداد کلاس صحیح در مقایسه با الگوریتم EM اشاره کرد. برای آینده تحقیق می‌توان این الگوریتم را بر روی توزیع مخلوط متناهی آمیخته مقیاسی چوله نرمال پیاده‌سازی کرد. حالت خاص این کلاس از توزیع‌ها همان FM-SMN است در صورتی که پارامتر چولگی برابر با صفر باشد.

مراجع

- [1] D.R. Andrews and C.L. Mallows, Scale mixture of normal distributions, J. Roy. Stat. Soc. B 36, 1974, pp. 99–102
- [2] LA. García-Escudero, F. Greselin and AM. Iscar. “Robust, fuzzy, and parsimonious clustering, based on mixtures of factor analyzers”. Inter. Jou. Appr. Reas., vol. 94, 2018, pp. 60–75.
- [3] LA. García-Escudero, D. Rivera-García, AM. Iscar and J. Ortega. “Cluster analysis with cellwise trimming and applications for the robust clustering of curves”. Info. Scie. vol. 573, 2021, pp. 100–124.
- [4] D. E. Gustafson and W.C. Kessel. “Fuzzy clustering with a fuzzy covariance matrix”. IEEE conf. on deci. and conto. inclu. the 17th sympo. on adap. proc., 1979, pp. 761–766.
- [5] F. Hashemi, M. Naderi and M. Mashinchi. “Clustering right-skewed data stream via Birnbaum–Saunders mixture models: A flexible approach based on fuzzy clustering algorithm”. Appl. Sof. Compu., vol. 82, 2019, 105539.

- [6] Z. Ju and H. Liu. “Fuzzy gaussian mixture models”. *Patt. Recog.* vol. 45, 2012, pp. 1146–1158.
- [7] T. I. Lin, J. C. Lee and H. J. Ho. “On fast supervised learning for normal mixture models with missing information”. *Patt. Recog.*, vol. 39, 2006, pp. 1177–1187.
- [8] Quost, B., Denoeux, T., 2016. “Clustering and classification of fuzzy data using the fuzzy EM algorithm”. *Fuz. Set. and Sys.*, vol. 286, 2016, pp. 134–156.
- [9] H. Wang, Q. bing Zhang, B. Luo and S. Wei. “Robust mixture modelling using multivariate t-distribution with missing information”. *Patt. Recog.*, vol. 25, 2004, pp. 701–710.
- [10] WL. Wang, A. Jamalizadeh, and TI, Lin. “Finite mixtures of multivariate scale-shape mixtures of skew-normal distributions”. *Stat Papers*, vol. 61, 2020, pp. 2643–2670.
- [11] M. S. Yang, and Y. Nataliani. “Robust-learning fuzzy c-means clustering algorithm with unknown number of clusters”. *Patt. Recog.*, vol. 71, 2017, pp. 45–59.
- [12] L. A. Zadeh. “Fuzzy sets”. *Info. and cont.* vol. 8, 1965, pp. 338–353.

پیوست الف

و تابع چگالی توزیع تی به صورت زیر حاصل می شود:

$$f_{t_p}(x; \mu, \Sigma, \nu) = \frac{(\pi\nu)^{-p/2} \Gamma\left(\frac{\nu+p}{2}\right) |\Sigma|^{1/2}}{\Gamma(\nu/2) [1 + \delta(x; \mu, \Sigma)/\nu]^{(\nu+p)/2}}, \quad (17)$$

به طوری که $\delta(x; \mu, \Sigma) = (x - \mu)^T \Sigma^{-1} (x - \mu)$ است. باید توجه داشت که در توزیع تی اگر $\nu = 1$ باشد، آنگاه توزیع کوشی بدست می آید.

و تابع چگالی توزیع اسلش به صورت زیر حاصل می شود:

$$f_{SL}(x; \mu, \Sigma, \nu) = \int_0^1 \nu w^{p+\nu-1} \phi_p(wx; w\mu, \Sigma) dw \quad x \in \mathbb{R}^p.$$

فرض کنید $W_j \sim \text{Gamma}(\nu/2, \nu/2)$ و $X_j^o \sim t(\mu_{ij}^o, \Sigma_{ij}^{oo}, \nu)$ باشد. در این صورت W_j به شرط $X_j^o = x_j^o$ و $Z_{ij} = 1$ دارای توزیعی $\text{Gamma}(\frac{\nu+p_j^o}{\nu+\delta(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo})})$ است. بطور دقیق تر داریم:

$$f(w_j | x_j^o, Z_{ij} = 1) = \frac{w_j^{\frac{\nu_i+p_j^o}{2}-1}}{\Gamma(\frac{\nu_i+p_j^o}{2})} \left(\frac{\nu_i + \delta(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo})}{2} \right)^{\frac{\nu_i+p_j^o}{2}} \\ \times \exp \left\{ -\frac{w_j}{2} (\nu_i + \delta(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo})) \right\}.$$

علاوه بر آن امید ریاضی های مورد نیاز در الگوریتم LB-FCM به صورت زیر حاصل می شوند:

$$E(W_j | (x_j^o, Z_{ij} = 1)) = \frac{\nu_i + p_j^o}{\nu_i + \delta(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo})}, \\ E(\log(W_j) | (x_j^o, Z_{ij} = 1)) = \text{DG} \left(\frac{\nu_i + p_j^o}{2} \right) \\ - \log \left(\frac{\nu_i + \delta(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo})}{2} \right),$$

به طوری که $\text{DG}(w) = \frac{d}{dw} \log \Gamma(w)$ است.

فرض کنید توزیع W بتا با پارامتر ν و 1 باشد و $X_j^o \sim \text{SL}(\mu_{ij}^o, \Sigma_{ij}^{oo}, \nu)$ باشد. در این صورت W_j به شرط $X_j^o = x_j^o$ و $Z_{ij} = 1$ دارای توزیعی بریده شده گاما به صورت $\text{Gamma}(\frac{\nu+p_j^o+1}{\nu+\delta(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo})}) I_{[0,1]}$ است. علاوه بر آن امید ریاضی های مورد نیاز در الگوریتم LB-FCM به صورت عددی با انتگرال گیری بدست می آید. و تابع چگالی توزیع نرمال آلوده شده به صورت زیر حاصل می شود:

$$f_{CN}(x; \mu, \Sigma, \nu, \gamma) = \nu \phi_p(x; \mu, \Sigma) + (1 - \nu) \phi_p(x; \mu, \gamma \Sigma).$$

فرض کنید $f_W(w_j) = h(w; \nu, \gamma) = \nu I_\gamma(w_j) + (1 - \nu) I_1(w_j)$ و $X_j^o \sim$ $CN(\mu_{ij}^o, \Sigma_{ij}^{oo}, \gamma, \nu)$ باشد. در این صورت W_j به شرط $X_j^o = x_j^o$ و $Z_{ij} = 1$ دارای تابع چگالی زیر است:

$$f(w_j | x_j^o, Z_{ij} = 1) = \frac{(\nu I_\gamma(w_j) + (1 - \nu) I_1(w_j)) * \phi(x_j^o; \mu_{ij}^o, w_j^{-1} \Sigma_{ij}^{oo})}{\nu \phi(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo}) + (1 - \nu) \phi(x; \mu, \gamma \Sigma)}.$$

علاوه بر آن امید ریاضی‌های مورد نیاز در الگوریتم LB-FCM به صورت زیر حاصل می‌شود.

$$E(W_j | (x_j^o, Z_{ij} = 1)) = \frac{\nu_i * \phi(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo})}{\nu_i \phi(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo}) + (1 - \nu_i) \phi(x_j^o; \mu_{ij}^o, \gamma \Sigma_{ij}^{oo})} + \frac{(1 - \nu) * \phi(x_j^o; \mu_{ij}^o, \gamma \Sigma_{ij}^{oo})}{\nu_i \phi(x_j^o; \mu_{ij}^o, \Sigma_{ij}^{oo}) + (1 - \nu_i) \phi(x_j^o; \mu_{ij}^o, \gamma \Sigma_{ij}^{oo})}.$$

همچنین امید ریاضی $E(\log(W_j) | (x_j^o, Z_{ij} = 1))$ به صورت عددی حاصل می‌شود.