پنجمین کنفرانس سیستمهای فازی ایران
هفدهم الی نوزدهم شهریورماه ۱۳۸۳

روشی برای دسته‌بندی و اولویت‌بندی فازی بر اساس کارائی

صابر ساعتی*[۱]، عزیزاله معماریانی*[۲]، سعید محرابیان*[۳]
گروه ریاضی، دانشگاه آزاد اسلامی واحد تهران ـ شمال
ssaatim@yahoo.com
*** گروه OR، مرکز مطالعات بنیادی، دانشگاه امام حسین (ع)**

چکیده

برای دسته‌بندی واحدهای تصمیم‌گیری، روشهای مختلفی ارائه شده است. از مواردی که در این روشها به آن توجه نمی‌شود، اولویت بین دسته‌ها و اولویت بین واحدهای تصمیم‌گیری واقع در هر دسته است. در این مقاله، با استفاده از تحلیل پوششی داده‌های فازی، یک روش دسته‌بندی بر اساس رتبه‌بندی واحدهای تصمیم‌گیری ارائه می‌شود. در این روش، علاوه بر دسته‌بندی واحدها، اولویت بین دسته‌ها و همچنین اولویت بین واحدهای واقع در هر دسته نیز تعیین می‌گردد.

واژه‌های کلیدی: دسته‌بندی ـ اولویت‌بندی ـ رتبه‌بندی ـ تحلیل پوششی داده‌ها ـ اعداد فازی

مقدمه

به کمک تحلیل ممیز (Discriminant Analysis)، می‌توان بر اساس مشخصه‌ای، مشاهدات را به دو یا چند گروه دسته‌بندی کرد. روش‌های متعددی برای دسته‌بندی مشاهدات وجود دارد (برای مطالعه بیشتر به Sharma [۳] مراجعه شود). روش مقدار برش (Cutoff-value)، روش نظریه تصمیم‌گیری آماری (Statistical Decision Theory)، روش تابع دسته‌بندی (Classification Function) و روش فاصله Mahalanobis از جمله روش‌هائی هستند که به طور متداول از آنها استفاده می‌گردد. در سه روش اول، بر اساس تابع ممیز (Discriminant Function) ضرایبی بدست می‌آید که به کمک آنها، شاخص‌های مشاهدات با هم ترکیب شده و تبدیل به یک عدد می‌گردد که این عدد را میزان ممیز متوسط (Average Discriminant Score) می‌نامند. سپس، بر اساس ترکیبی از این مقادیر (که بسته به نوع روش، این ترکیب متفاوت است) و مقایسه میزان ممیز متوسط هر مشاهده با آن، اعضاء دسته‌ها مشخص می‌گردد. در روش فاصله Mahalanobis نیز با تعریف یک تابع فاصله، فاصله مشاهدات از گروه‌ها مشخص شده و بدین ترتیب مشاهدات دسته‌بندی می‌شوند.

همانطوریکه اشاره شد، هر مشاهده‌ای دارای یک سری شاخص‌ها است. از یک دیدگاه، می‌توان شاخص‌ها را به دو گروه ورودی (Input) و خروجی (Output) تقسیم کرد. هر شاخصی که زیاد بودن آن برای مشاهدات مفید نباشد، ماهیت ورودی و هر شاخصی که زیاد بودن آن برای مشاهدات امتیاز باشد، ماهیت خروجی دارد. با این دیدگاه، استفاده از روش‌های مذکور برای دسته‌بندی مشاهداتی که دارای شاخص‌های ورودی و خروجی هستند نمی‌تواند مناسب باشد. حتی تبدیل ورودی‌ها و

۱- استادیار گروه ریاضی، دانشگاه آزاد اسلامی واحد تهران ـ شمال
۲- دانشیار، بخش مهندسی صنایع، دانشگاه بوعلی‌سینا همدان
۳- استادیار گروه ریاضی، دانشگاه تربیت معلم تهران

خروجی‌ها به یکدیگر با اعمال تغیراتی (مانند عکس کردن)، بدلیل اینکه مشاهدات را تغییر می‌دهد، روشی مناسبی نیست. در چنین حالتی، روش‌های مذکور که بر اساس مفهوم فاصله و بعبارت دیگر با استفاده از یک معیار absolute اقدام به دسته‌بندی مشاهدات می‌کنند، نمی‌تواند نتایج مناسبی ارائه دهد.

تحلیل پوششی داده‌ها (Data Envelopment Analysis) روشی است که امروزه در زمینه‌های مختلف، کاربرد وسیعی پیدا کرده است. بر اساس گزارش Tavares [۱]، پس از اولین مقاله Charnes و همکاران [۲] در سال ۱۹۷۸ در مورد تحلیل پوششی داده‌ها، بیش از ۳۲۰۳ مقاله، کتاب، گزارش و پایان‌نامه تا پایان سال ۲۰۰۱ توسط ۲۱۵۲ مولف در رابطه با این روش بچاپ رسیده است. معمول‌ترین کاربرد تحلیل پوششی داده‌ها، تعیین مقداری است که توسط آن، کارائی نسبی یک واحد تصمیم‌گیری (Decision Making Unit) در مقایسه با سایر واحدهای تصمیم‌گیری مشخص می‌گردد. بدین ترتیب، واحدهای تصمیم‌گیری به دو گروه کارا و ناکارا دسته‌بندی می‌شوند.

اخیراً، روش‌های مختلفی برای دسته‌بندی واحدهای تصمیم‌گیری بکمک تحلیل پوششی داده‌ها ارائه شده است (برای مطالعه بیشتر به [۴،۵،۶،۷،۸،۹،۱۰] مراجعه شود). نکته قابل توجه در رابطه با این روش‌ها این است که هیچ یک از آنها قادر به دسته‌بندی واحدهای تصمیم‌گیری با داده‌های فازی نیستند. از طرف دیگر، همانند روش‌های کلاسیک دسته‌بندی، اولویت بین دسته‌ها و واحدهای واقع در دسته‌ها مشخص نمی‌شود. از محدودیت‌های دیگر روش‌های مذکور، این است که تنها قادرند واحدهای تصمیم‌گیری را به دو گروه کارا و ناکارا دسته‌بندی کنند.

در این مقاله، با ارائه روشی جدید، دسته‌بندی واحدهای تصمیم‌گیری با داده‌های فازی انجام می‌گیرد. اساس روش پیشنهادی، رتبه‌بندی واحدهای تصمیم‌گیری با داده‌های فازی است. بکمک این روش، می‌توان واحدها را به بیش از دو دسته نیز دسته‌بندی کرد.

ساختار مقاله بدین شکل است که در بخش بعدی، روش پیشنهادی ارائه می‌گردد. سپس، در بخش سوم، برای روشن‌تر شدن روش پیشنهادی، یک مثال عددی در نظر گرفته شده است. نهایتاً، یک نتیجه‌گیری در انتهای مقاله آمده است.

روش پیشنهادی

روشهای دسته‌بندی، مشاهدات را بر اساس مشخصه‌ای از قبل تعیین شده به گروههای جدا از هم تقسیم می‌کنند. این مشخصه، می‌تواند رنگ، اندازه، کیفیت، فاصله یا مقدار مشاهدات باشد. در اکثر روشهای کلاسیک، از معیار فاصله برای دسته‌بندی مشاهدات استفاده می‌شود که این معیار، یک معیار absolute است. در این بخش، روشی متفاوت برای دسته‌بندی مشاهدات ارائه می‌شود. در روش پیشنهادی، مشاهدات بر اساس رتبه، دسته‌بندی می‌گردند. بدین ترتیب، معیار دسته‌بندی، یک معیار rational خواهد بود.

فرض کنیم مشاهدات، n واحد تصمیم‌گیری باشند که هر یک m ورودی را برای تولید s خروجی مصرف می‌کنند. بعبارت دیگر، واحد تصمیم‌گیری j ام $(j \in J = \{1,...,n\})$، m ورودی $(i \in I = \{1,...,m\})\, x_{ij}$ را مصرف کرده و s خروجی $(r \in O = \{1,...,s\})\, y_{rj}$ را تولید می‌کند. با توجه به اینکه در مسائل زندگی روزمره، بدلیل پیچیدگی مسائل و ضعف ابزاری، بسیاری از کمیت‌ها، بصورت فازی و یا کیفی بیان می‌شوند، فرض می‌کنیم، تمام ورودی‌ها و خروجی‌ها فازی باشند. این فرض مشکلی در محاسبات ایجاد نمی‌کند. زیرا، هر عدد قطعی را نیز می‌توان بصورت یک عدد فازی بیان کرد.

روش کار بدین شکل می‌باشد که ابتدا، واحدهای تصمیم‌گیری بر اساس کارائی‌شان رتبه بندی می‌گردند. واحدهائی که دارای کارائی بزرگتر یا مساوی یک باشند، در دسته اول قرار می‌گیرند. در گام بعدی، این واحدها از مجموعه واحدهای تصمیم‌گیری حذف شده و رتبه‌بندی برای بقیه واحدها انجام می‌گیرد. در این مجموعه، واحدهائی که دارای کارائی بزرگتر یا مساوی یک هستند، در دسته دوم قرار می‌گیرند. این واحدها را از مجموعه واحدهای تصمیم‌گیری باقیمانده حذف کرده و روش را برای بقیه واحدها تکرار می‌کنیم. این کار تا زمانی ادامه می‌یابد که تنها یک واحد تصمیم‌گیری باقی مانده باشد.

روش‌های دسته‌بندی موجود، اولویت بین دسته‌ها را نمی‌توانند تعیین کنند. از مزایای روش پیشنهادی، می‌توان به تعیین اولویت بین دسته‌ها اشاره کرد. دسته‌ها با ترتیب بوجود آمدن، اولویت‌بندی می‌شوند. بدین شکل که اولین دسته، دارای اولویت اول، دومین دسته، دارای اولویت دوم و الی آخر. این روش علاوه بر تعیین اولویت دسته‌ها، مشاهدات داخل هر دسته را نیز اولویت‌بندی می‌کند. اولویت مشاهدات در دسته‌ها، بر اساس رتبه آنها در دسته مشخص می‌شود.

در این مقاله، از روش ارائه شده توسط ساعتی و همکاران [۱۱]، برای رتبه‌بندی واحدهای تصمیم‌گیری با داده‌های فازی استفاده می‌گردد. بدلیل کاربردهای فراوان اعداد فازی مثلثی، فرض می‌کنیم ورودی‌ها و خروجی‌های واحدهای تصمیم‌گیری بصورت اعداد فازی مثلثی باشند یعنی، $\widetilde{x}_{ij} = (x_{ij}^m, x_{ij}^l, x_{ij}^u)\ (\forall i, j)$ و $\widetilde{y}_{rj} = (y_{rj}^m, y_{rj}^l, y_{rj}^u)\ (\forall r, j)$ که در آن، m، l و u بترتیب، نشان دهنده مقدار میانی، کران پایین و کران بالای اعداد فازی هستند. در این صورت، مدل ارائه شده در [۱۱]، بفرم زیر خواهد بود:

$$
\begin{aligned}
\min \quad & \theta \\
\text{s.t}: \quad & \theta(\alpha x_{ip}^m + (1-\alpha)x_{ip}^l) \geq \sum_{j\in J} \lambda_j (\alpha x_{ij}^m + (1-\alpha)x_{ij}^u) && \forall i \in I, \\
& \alpha y_{rp}^m + (1-\alpha) y_{rp}^u \leq \sum_{j\in J} \lambda_j (\alpha y_{rj}^m + (1-\alpha) y_{rj}^l) && \forall r \in O, \\
& \lambda_j \geq ۰ && \forall j \in J.
\end{aligned}
\qquad (۱)
$$

همانطوریکه مشخص است، در روش پیشنهادی از قبل نمی‌توان تعداد دسته‌ها را مشخص کرد و پس از اجرای روش، تعداد آنها تعیین می‌گردند.

الگوریتم دسته‌بندی

روش دسته‌بندی ارائه شده در بخش قبل را می‌توان به صورت الگوریتم زیر بیان نمود:

گام ۱: J را مجموعه اندیس تمام واحدهای تصمیم‌گیری در نظر بگیرید.

گام ۲: قرار دهید $k = ۰$ (شماره دسته).

گام ۳: قرار دهید $M = \phi$ (اندیس واحدهای دسته‌بندی شده).

گام ۴: مدل (۱) را برای تمام واحدهای تصمیم‌گیری که اندیس آنها در J قرار دارد، اجرا کنید.

گام ۵: قرار دهید $k = k + ۱$.

گام ۶: واحدهای تصمیم‌گیری با θ بهین بزرگتر یا مساوی ۱ را در دسته k ام قرار دهید.

گام ۷: اندیس واحدهای واقع در دسته k ام را به مجموعه M اضافه کنید.

گام ۸: قرار دهید $J = J - M$.

گام ۹: اگر $J = \phi$ الگوریتم خاتمه می‌یابد. در غیر اینصورت به گام ۴ بروید.

مثال عددی

در این بخش، برای توضیح بیشتر روش دسته‌بندی پیشنهادی، یک مثال عددی ارائه می‌گردد. داده‌های این مثال از [۱۱] انتخاب شده‌اند. ۱۰ واحد تصمیم‌گیری را با دو ورودی فازی و دو خروجی فازی، مطابق جدول زیر در نظر می‌گیریم.

جدول ۱ـ داده‌های مثال عددی

	I_1	I_2	O_1	O_2
D_1	(۷/۰،۶/۰،۸/۰)	(۳۰/۰،۲۹/۰،۳۲/۰)	(۳۸/۰،۳۵/۵،۴۱/۰)	(۴۱۱/۰،۴۰۹/۰،۴۱۶/۰)
D_2	(۶/۰،۵/۵،۶/۵)	(۳۵/۰،۳۳/۰،۳۶/۵)	(۴۰/۰،۳۹/۵،۴۳/۰)	(۴۸۰/۰،۴۷۸/۰،۴۸۴/۰)
D_3	(۹/۰،۷/۵،۱۰/۵)	(۴۵/۰،۴۳/۰،۴۸/۰)	(۳۵/۰،۳۲/۵،۳۸/۰)	(۲۹۹/۰،۲۹۷/۰،۳۰۱/۰)
D_4	(۸/۰،۷/۰،۱۰/۰)	(۳۹/۰،۳۷/۵،۴۲/۰)	(۳۱/۰،۲۸/۰،۳۱/۰)	(۳۵۲/۰،۳۴۷/۵،۳۶۰/۰)
D_5	(۱۱/۰،۹/۰،۱۲/۰)	(۴۴/۰،۴۳/۰،۴۵/۰)	(۳۵/۰،۳۳/۰،۳۸/۰)	(۴۱۱/۰،۴۰۶/۰،۴۱۵/۰)
D_6	(۱۰/۰،۱۰/۰،۱۰/۰)	(۵۵/۰،۵۳/۰،۵۷/۵)	(۳۸/۰،۳۶/۰،۴۰/۰)	(۲۸۶/۰،۲۸۲/۰،۲۸۹/۰)
D_7	(۱۲/۰،۱۰/۰،۱۴/۰)	(۱۱۰/۰،۱۰۷/۰،۱۱۳/۰)	(۳۶/۰،۳۴/۵،۳۸/۰)	(۴۰۰/۰،۳۹۶/۰،۴۰۵/۰)
D_8	(۱۳/۰،۹/۰،۱۶/۰)	(۱۰۰/۰،۹۵/۰،۱۰۱/۰)	(۴۱/۰،۳۷/۰،۴۶/۰)	(۳۹۳/۰،۳۸۷/۰،۴۰۲/۰)
D_9	(۱۴/۰،۱۲/۰،۱۵/۰)	(۱۲۵/۰،۱۲۰/۰،۱۳۱/۰)	(۲۷/۰،۲۴/۰،۲۸/۰)	(۴۰۴/۰،۴۰۰/۰،۴۰۶/۰)
D_{10}	(۸/۰،۵/۰،۱۰/۰)	(۳۸/۰،۳۵/۰،۳۹/۰)	(۵۰/۰،۴۸/۰،۵۱/۰)	(۴۷۰/۰،۴۷۰/۰،۴۷۰/۰)

در این جدول، برخی از ورودی‌ها و خروجی‌های واحدهای تصمیم‌گیری اعداد قطعی هستند. این نوع داده‌ها را با مساوی قرار دادن مقدار میانی و کران‌های بالا و پایین بصورت اعداد فازی مثلثی نمایش داده‌ایم.

الگوریتم پیشنهادی را بازاء $\alpha = ۰/۵$ اجرا می‌کنیم. با اجرای مدل (۱) در تکرار اول، نتایج جدول ۲ حاصل می‌شود.

جدول ۲ـ نتایج تکرار اول الگوریتم

واحد تصمیم‌گیری	D_{10}	D_2	D_1	D_3	D_5	D_4	D_6	D_8	D_7	D_9
رتبه	۱/۲۴	۱/۱۴	۱/۱۰	۰/۷۳	۰/۷۲	۰/۷۱	۰/۶۴	۰/۶۳	۰/۵۳	۰/۴۱

همانطوریکه از نتایج جدول ۲ مشخص است، θ بهین واحدهای تصمیم‌گیری D_{10}، D_2 و D_1 بزرگتر از یک است. لذا، این واحدها در دسته اول قرار می‌گیرند. با حذف اندیس واحدهای تصمیم‌گیری واقع در دسته اول از مجموعه اندیس J و اجرای مجدد مدل (۱) ، نتایج جدول ۳ بدست می‌آید.

جدول ۳ـ نتایج تکرار دوم الگوریتم

واحد تصمیم‌گیری	D_4	D_3	D_5	D_8	D_6	D_7	D_9
رتبه	۱/۲۵	۱/۲۳	۱/۱۴	۱/۱۲	۱/۰۷	۰/۹۹	۰/۸۰

با توجه به نتایج جدول ۳، اعضای دسته دوم به ترتیب بصورت D_4، D_3، D_5، D_8 و D_6 مشخص می‌گردد. از مجموعه اندیس J مجموعه اندیس واحدهای واقع در دسته دوم یعنی $\{۳،۴،۵،۶،۸\}$ را کم کرده و مدل (۱) را برای تعیین وضعیت واحدهای تصمیم‌گیری D_7 و D_9 اجرا می‌کنیم. نتایج این اجرا در جدول ۴ نشان داده شده است.

جدول ۴ـ نتایج تکرار سوم الگوریتم

واحد تصمیم‌گیری	D_7	D_9
رتبه	۱/۲۴	۱/۰۹

نهایتاً، نتایج اجرای روش پیشنهادی برای داده‌های جدول (۱)، بازاء $\alpha = ۰/۵$، در جدول ۵ ارائه شده است.

جدول ۵ ـ دسته‌بندی داده‌های مثال عددی

دسته اول	دسته دوم	دسته سوم
$D_{۱۰}$ = ۱/۲۴	$D_{۴}$ = ۱/۲۵	$D_{۷}$ = ۱/۲۴
$D_{۲}$ = ۱/۱۴	$D_{۳}$ = ۱/۲۳	$D_{۹}$ = ۱/۰۲
$D_{۱}$ = ۱/۱۰	$D_{۵}$ = ۱/۱۴	
	$D_{۸}$ = ۱/۱۲	
	$D_{۶}$ = ۱/۰۷	

همانطوریکه مشاهده می‌شود، ۱۰ واحد تصمیم‌گیری مثال، به سه دسته تقسیم شده‌اند. اولویت بین دسته‌ها بترتیب بصورت دسته اول، دوم و سوم است. در دسته اول، سه واحد تصمیم‌گیری $D_{۱}$، $D_{۲}$ و $D_{۱۰}$ قرار دارد که اولویت آنها بترتیب $D_{۱۰}$، $D_{۲}$ و $D_{۱}$ است.

نتیجه‌گیری

دسته‌بندی موضوعات، واحدها، افراد و غیره، یاز مسائلی است که در اکثر سیستم‌ها مورد توجه بوده و دارای اهمیت است. دسته‌بندی معمولاً بر اساس یک یا چند صفت مشخص انجام می‌گیرد. این دسته‌بندی ممکن است ساده بوده (مانند دسته‌بندی افراد بر حسب جنس) و ممکن است بسیار پیچیده و مشکل باشد (مانند دسته‌بندی افراد بر حسب کارائی آنها در انجام کاری). در این مقاله، یک روش دسته‌بندی مبتنی بر رتبه واحدها ارائه گردیده است. در این روش، برای دسته‌بندی واحدها، از یک معیار نسبی، یعنی کارائی استفاده شده است. با این تعریف، روش پیشنهادی قادر است که واحدهائی با شاخص‌های متفاوت چه از نظر ماهیت (ورودی یا خروجی) و چه از نظر نوع (کمی، کیفی یا فازی) را در دسته‌های مختلف، دسته‌بندی کند. این توانائی روش، باعث گسترده شدن دامنه کاربرد آن در مسائل جهان واقعی مانند دسته‌بندی مدارس، دانشگاه‌ها، مقالات، اساتید، دانشجویان، شعب بیمه، بانک‌ها و ... می‌گردد.

از مواردی که در این روش در نظر گرفته نشده و می‌توان با تغییراتی آنها را اعمال نمود، می‌توان به استفاده از انواع داده‌ها از قبیل مقادیر دوتائی، غیر مطلق، مقادیر صحیح و ... در مدل اشاره نمود. با توجه به اینکه اساس روش پیشنهادی، برنامه‌ریزی خطی است، استفاده از این نوع داده‌ها، زیاد بعید بنظر نمی‌رسد. از موارد دیگر، مشخص نبودن تعداد دسته‌ها تا اتمام الگوریتم است. این مورد نیز با توجه به کاربرد آن در برخی مسائل، بایستی در مدل لحاظ شود.

مراجع

[1] Tavares، G.، (2002)، "A Bibliography of Data Envelopment Analysis"، Rutcor research report، Rutgers University.

[2] Charnes، A.، W. W. Cooper and E. Rhodes، (1978)، "Measuring the Efficiency of Decision Making Units"، European Journal of Operational Research، Vol. 2، Pages 429-444.

[3] Sharma، S.، (1996)، "Applied Multivariate Techniques"، John Wiley & Sons، Inc.

[4] Sinuany-Stern، Z.، A. Mehrez and A. Barboy، (1994)، "Academic Departments Efficiency Via DEA"، Computers & Operations Research، Vol. 21، Pages 543-556.

[5] Golany، B. and Yu، G.، (1995)، "A Goal Programming-Discriminant Function Approach to the Estimation of an Empirical Production Function Based on DEA Results"، The Journal of Productivity Analysis، Vol. 6، Pages171-186.

[6] Retzlaff-Roberts، D.L.، (1996)، "Relating Discriminant Analysis and Data Envelopment to One Another"، Computers & Operations Research، Vol. 23، Pages 311-322.

[7] Retzlaff-Roberts، D.L.، (1996)، "Classification in Automobile Insurance Using a DEA and Discriminant Analysis Hybrid"، The Journal of Productivity Analysis، Vol. 7، Pages 417-427.

[8] Retzlaff-Roberts، D.L.، (1997)، "A Data Envelopment Analysis Approach to Discriminant Analysis"، Annals of Operations Research، Vol. 73، Pages 299-321.

[9] Sinuany-Stern، Z. and L. Friedman، (1998)، "DEA and the Discriminant Analysis of Ratios for Ranking Units"، European Journal of Operational Research، Vol. 111، Pages 470-478.

[10] Sueyoshi ،T.، (1999)، "DEA-Discriminant Analysis in the View of Goal Programming"، European Journal of Operational Research، Vol. 115، Pages 564-582.

[11] Saati M.، S.، A. Memariani and G. R. Jahanshahloo، (2002)، "Efficiency Analysis and Ranking of DMUs with Fuzzy Data"، Fuzzy Optimization and Decision Making، Vol. 1، No. 3، Pages 255-267.