

معدل‌گیری از داده‌های نادقیق

ماشالله ماشین‌چی*

چکیده

در زندگی روزمره با اطلاعاتی مواجه می‌شویم که در قالب مفاهیم نادقیق بیان شده‌اند و نه با اعداد. در صورتی که اطلاعات برحسب اعداد باشند، می‌توان با روشهای رایج آماری آنها را معدل‌گیری کرد. در این مقاله دو روش برای معدل‌گیری داده‌های نادقیق ارائه می‌کنیم. در یک روش، داده‌های نادقیق را به صورت عناصری از یک شبکه تلفی کرده و سپس عمل معدل‌گیری را توسط یک همسوساز در این شبکه انجام می‌دهیم. در روش دیگر، ابتدا داده‌های نادقیق را در قالب مجموعه‌های مشکک مدل‌سازی کرده و سپس آنها را با استفاده از اصل توسیع معدل‌گیری می‌کنیم. در هر دو روش برای معدل‌گیری به یک همسوساز مناسب نیاز داریم که این مفهوم را در اینجا معرفی کرده و خواصی از آن را بررسی می‌کنیم. با ارائه چند مثال به بیان دو روش می‌پردازیم تا عیب و حسن و همین‌طور چگونگی کاربرد روشها آشکار گردد.

۱ مقدمه

در بسیاری از مسائل آماری، اقتصادی و روزمره با اطلاعاتی مواجه می‌شویم که نیاز به معدل‌گیری آنها داریم. در صورتی که این داده‌ها به صورت عدد بیان شوند توسط معیارهای تمرکز مثل میانگین حسابی، میانگین هندسی، میانگین هارمونیک و غیره، که در علم آمار به خوبی شناخته شده‌اند، می‌توان دادها را به اصطلاح همسوز کرد و عمل معدل‌گیری را انجام داد [۱]. اما در بسیاری از موارد داده‌های به دست آمده از نوع عددی نیستند و اغلب

به صورت مفاهیم نادقیق [۲ و ۳] مثل: خیلی خوب، متوسط، حول و حوش ۱۸ و غیره در دسترس قرار می‌گیرند. به طور نمونه این مطلب در مکالمات روزمره بسیار مشهود است و بیشتر اطلاعاتی که بین افراد رد و بدل می‌شود از نوع نادقیق‌اند. با این همه ما از این اطلاعات هم استنتاجهای مناسبی را انجام می‌دهیم. مورد دیگری که می‌توان اشاره کرد، این است که اصولاً جمع‌آوری داده‌های عددی ممکن نیست، مانند آمار بلایای طبیعی در یک منطقه دور افتاده یا در مواردی که پاسخ دهندگان رغبت کمتری به دادن اطلاعات دقیق عددی داشته باشند مانند تهیه آمار درباره میزان درآمد افراد. سرانجام در مواردی تهیه داده‌های نادقیق آسانتر و کم‌خرج‌تر است. لذا در این مقاله فرض می‌کنیم که داده‌های مفروض نادقیق‌اند و می‌خواهیم این داده‌ها را همسوز کرده و معدل آنها را به دست آوریم. بنابراین سؤال این است که چگونه می‌توان این معدل‌گیری را انجام داد؟

در این مقاله ما دو روش برای انجام معدل‌گیری ارائه می‌کنیم. در یک روش فرض می‌کنیم که داده‌های ما دارای یک ساختار شبکه‌ای [۱۵] مناسب‌اند یا از یک شبکه مناسب بایستی انتخاب شوند. سپس با ارائه مفهوم یک همسوساز روی این شبکه‌ها معدل‌گیری را انجام می‌دهیم. لذا در این روش داده‌های نادقیق صرفاً به عنوان عناصری از یک شبکه تلفی می‌شوند. البته روش استفاده از شبکه‌ها برای استنتاج از مفاهیم نادقیق اخیراً توسط برخی نویسندگان [۵، ۴] ارائه شده و نشان داده شده که چنانچه این مفاهیم نادقیق به طریق مناسبی انتخاب شوند، دارای ساختارهای شبکه‌ای جالبی به نام جبر پرچین (hedge algebra) می‌باشند. ما در اینجا به نحوی از این ایده استفاده کرده‌ایم. در روش دیگر مفاهیم نادقیق را

* دکتر ماشالله ماشین‌چی، دانشکده ریاضی و کامپیوتر، دانشگاه شهید باهنر کرمان

تذکر ۱.۲ اگر فرض کنیم که $h: L^1 \rightarrow L$ یک عمل دوتایی شرکتپذیر روی L باشد، آنگاه با فرض

$$h(a, b, c) = h(a, h(b, c)), \forall a, b, c \in L$$

می‌توان h را به یک همسوساز روی L توسعه داد.

طبیعی است که هر عمل n -تایی h روی L نمی‌تواند یک همسوساز مناسب برای عمل معدل‌گیری باشد. برای به دست آوردن یک همسوساز جالب بایستی شرایطی معقول را بر h اعمال کرد. در اینجا با الهام گرفتن از خواص میانگین حسابی به عنوان یک همسوساز، شرایط زیر را اعمال می‌کنیم که در صورت نیاز به هر کدام از آنها در جای خود به ذکر آن می‌پردازیم.

$$\text{ش ۱. } h(1, 1, \dots, 1, 1) = 1, h(0, 0, \dots, 0, 0) = 0$$

ش ۲. اگر $\underline{a} = (a_1, \dots, a_n) \in L^n$ و $\underline{b} = (b_1, \dots, b_n) \in L^n$ به طوری که $\underline{a} \leq \underline{b}$ یعنی $a_i \leq b_i$ برای $i = 1, 2, \dots, n$ آنگاه $h(\underline{a}) \leq h(\underline{b})$ به عبارتی h به طور یکنوا صعودی است.

ش ۳. h متقارن است. یعنی برای هر $(a_1, \dots, a_n) \in L^n$ داریم

$$h(a_1, \dots, a_n) = h(a_{p(1)}, \dots, a_{p(n)})$$

که در آن p یک تابع جایگشت روی مجموعه $\{1, 2, \dots, n\}$ است.

ش ۴. h خود توان^۲ است. یعنی به ازای هر $a \in L$ داریم

$$h(a, \dots, a, a) = a$$

تذکر ۲.۲ با توجه به ش ۴ می‌بینیم که در این حالت ش ۱ نتیجه می‌شود. به عنوان مثال، تعداد زیادی از همسوسازهای محسوس را می‌توان ارائه کرد. از جمله میانگین تعمیم یافته که در بخش ۱.۳ و معدل‌گیر وزنی که در بخش ۲.۳ ارائه شده‌اند. اما دسته دیگری از همسوسازها را که کاربرد زیادی نیز دارند، و به نام نرم‌های مثلثی شبکه‌ای نیز معروف‌اند، به لحاظ نیاز در تعریف زیر ارائه می‌کنیم.

تعریف ۲.۲ تابع $f: L \times L \rightarrow L$ را در یک L -نرم گوئیم، اگر دارای خواص زیر باشد

$$(۱) f(x, 1) = x \quad \forall x \in L \text{ به ازای هر}$$

$$(۲) f(y, x) = f(x, y) \quad \forall x, y \in L \text{ به ازای هر}$$

$$(۳) \text{ اگر } y \leq z \text{ آنگاه } f(x, y) \leq f(x, z) \text{ به ازای هر } x, y, z \in L$$

$$(۴) f \text{ شرکتپذیر است، یعنی } f(f(x, y), z) = f(x, f(y, z)) \text{ به ازای هر } x, y, z \in L$$

به صورت مجموعه‌های مشکک [۲] مدل‌سازی می‌کنیم و سپس با استفاده از اصل توسیع به همسوسازی اطلاعات و معدل‌گیری آنها می‌پردازیم. در هر حال در هر دو روش نیازمندیم تا مفهوم همسوساز را معرفی کنیم. لذا در این مقاله ابتدا همسوسازهای شبکه‌ای را معرفی کرده و سپس برخی همسوسازهای خاص مثل میانگین تعمیم یافته، معدل‌گیری وزنی و معدل‌گیر ترتیبی را نیز ارائه می‌کنیم. این همسوسازهای اخیراً به طور وسیعی مورد مطالعه و استفاده قرار گرفته‌اند [۶-۱۳]. در این مقاله با ارائه چند مثال به بیان دو روش می‌پردازیم تا عیب و حسن و همین طور چگونگی کاربرد دو روش را آشکار نماییم.

۲ عملگر همسوساز شبکه‌ای^۱

فرض کنید از سه استاد در مورد وضعیت تحصیلی یک دانشجو سؤال کنیم و آنان پاسخهای خود را به صورتهای نادقیق: خوب، تقریباً خوب و یا کم و بیش خوب بیان کنند. حال می‌خواهیم بدانیم که وضعیت این دانشجو روی هم رفته چطور است؟ برای بررسی مسائلی نظیر این موضوع نیاز به عملگرهای همسوساز شبکه‌ای داریم که در این بخش به معرفی آنها می‌پردازیم.

فرض کنید $(L, \leq, \wedge, \vee)$ یک شبکه [۱۵] باشد، که در صورت نیاز و بدون تذکر فرض می‌کنیم L دارای کوچکترین عنصر^۰ و بزرگترین عنصر^۱ است. به عنوان مثال $(p(X), \subseteq, \cap, \cup)$ یک شبکه است، که در آن X یک مجموعه نا تهی، $p(X)$ مجموعه همه زیرمجموعه‌های X است و به علاقه $\subseteq, \cap, \cup$ به ترتیب نمادهای زیرمجموعه بودن، اشتراک و اجتماع هستند. بدین ترتیب در این شبکه عنصر^۰ همان مجموعه تهی $\emptyset$ و عنصر^۱ همان مجموعه X است. خواننده برای مطالعه در خصوص شبکه‌ها و مثالهای دیگر می‌تواند به [۱۷، ۱۵] مراجعه کند.

مشبه $(L^n, \leq, \wedge, \vee)$ را که در آن $L^n = L \times \dots \times L$ حاصلضرب دکارتی n نسخه، n عددی طبیعی، از L است می‌توان به طور طبیعی ساخت. بنابراین ترتیب روی L^n از ترتیب روی L به طور مؤلفه‌ای القاء می‌شود و به ازای هر $\underline{a} = (a_1, \dots, a_n) \in L^n$ و $\underline{b} = (b_1, \dots, b_n) \in L^n$ به ترتیب $\wedge, \vee$ را به صورت $(a_1 \wedge b_1, \dots, a_n \wedge b_n)$ و $(a_1 \vee b_1, \dots, a_n \vee b_n)$ تعریف می‌کنیم.

تعریف ۱.۲ هر تابع $h: L^n \rightarrow L$ را یک عملگر همسوساز وابسته به شبکه L ، یا به طور مختصر یک همسوساز گوئیم.

در واقع یک همسوساز عبارت است از عملی n -تایی روی شبکه L .

1) Lattice aggregation operator

2) idempotent

$a_* = h(a_*, \dots, a_*) \leq h(a_1, \dots, a_n) \leq h(a^*, \dots, a^*) = a^*$
یعنی h در نامساوی (۱) صدق می کند. به علاوه اگر یک تابع همسوساز
در نامساوی (۱) صدق کند، آنگاه داریم

$$a = a \wedge a \wedge \dots \wedge a \leq h(a, \dots, a) \leq a \vee a \vee \dots \vee a = a$$

بنابراین $h(a, \dots, a) = a$ و لذا خود توان است و ش ۴ برقرار است.

تذکر ۲.۳ با توجه به قضیه ۱.۲ ملاحظه می شود که کرانهای بالا و پایین
نامساوی (۱) در قضیه ۲.۲ تنها عملگرهای همسوساز در کلاس نرمهای
مثلثی شبکه ای هستند که دارای هر چهار شرط ش ۱ - ش ۴ می باشند.
به علاوه اگر قرار دهیم

$$\text{Min}(a_1, \dots, a_n) = a_1 \wedge a_2 \wedge \dots \wedge a_n,$$

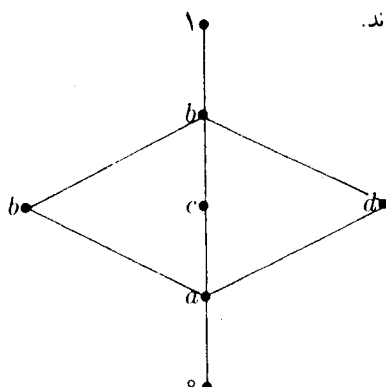
$$\text{Max}(a_1, \dots, a_n) = a_1 \vee a_2 \vee \dots \vee a_n$$

آنگاه با توجه به خاصیت شبکه ها، $\text{Min} \leq \text{Max}$ و به علاوه با توجه به
قضیه ۲.۲ هر همسوساز h که بین دو همسوساز Min, Max قرار گیرد حتماً
خودتوان است. بنابراین هر همسوساز روی L از Min قویتر و از Max
ضعیفتر است.

به منظور کاربرد همسوساز شبکه ای به ارائه دو مثال می پردازیم.

مثال ۱.۲ فرض کنید نظر سه استاد را در مورد وضعیت یک دانشجو
سؤال کرده ایم. این استادان پاسخهای خود را به صورت نادقیق: خوب
(d)، تقریباً خوب (c) و کم و بیش خوب (b) بیان کرده اند. حال می خواهیم
بدانیم وضعیت تحصیلی این دانشجو روی هم رفته چگونه است؟

فرض کنید نظر هر سه استاد هم ارزش است و به علاوه به طور ذهنی
فرض می کنیم که نمرات ارائه شده عناصری از یک شبکه اند. در اینجا با
کمی دقت این طور استنباط می کنیم که نمرات b, c, d در این شبکه قابل
مقایسه نیستند. لذا با این برداشتها شبکه L را که در شکل ۱ آمده است
برای مطالعه این مسأله چنین در نظر می گیریم که در آن مابقی عناصر به
صورت بسیار خوب (۱)، خوب (e)، کمی خوب (a) و چنین نیست خوب
(۰) تعبیر شده اند.



شکل ۱. شبکه نمرات L که بر اساس داده ها ساخته شده است.

در تعریف فوق f را یک L -نرم نرم گوئیم، اگر به جای خاصیت
(۱) فرض زیر برقرار باشد:

$$f(x, 0) = x \quad (۱')$$

L -نرم ها و L -نرم نرم را نرمهای مثلثی شبکه ای گوئیم.

حال با توجه به شرکت پذیری یک نرم مثلثی شبکه ای و تذکر

۱.۲ می توان از هر نرم مثلثی شبکه ای یک همسوساز ساخت

که در ش ۱ - ش ۳ صدق می کند.

به عنوان مثال، چنانچه شبکه ما $([0, 1], \leq, \min, \max)$ باشد، در این
صورت اگر f یکی از اعمال دوتایی $\min$ (مینیم)، $\max$ (ماکسیم)،
 prod (حاصل ضرب) روی بازه $[0, 1]$ باشد، آنگاه f در هر حالت یک نرم
مثلثی روی بازه حقیقی $[0, 1]$ است که در آن $\leq$ رابطه کوچکتر یا مساوی
معمولی روی اعداد حقیقی است. البته مثالهای زیاد دیگری از نرمهای
مثلثی می توان ارائه کرد که علاقه مندان می توانند به [۱۴] مراجعه نمایند. برای
حالتی که شبکه L زنجیر شکل ۲ است جدول ۱ مثالی از یک نرم مثلثی
روی L است که متفاوت با $\wedge, \vee$ است. خواننده می تواند مثالهای متفاوت
دیگری را نیز به دست آورد، چنانچه یک نرم مثلثی شبکه ای در ش ۳
صدق کند آنگاه قضیه زیر را داریم.

قضیه ۱.۲ فرض کنید f یک نرم مثلثی شبکه ای باشد. آنگاه
 $f(a, a) = a$ به ازای هر $a \in L$ اگر و فقط اگر $f = \wedge$ یا $f = \vee$.

اصولاً نرمهای مثلثی دسته خاصی از همسوسازها هستند که کاربرد
وسعی در تعریف اعمال اجتماع و اشتراک روی مجموعه های مشکک دارند
[۱۴]. اما در حالت کلی می توان قضیه جالب زیر را در مورد هر همسوساز
دلخواه ارائه کرد که این همسوساز را به دو همسوساز خاص $\wedge, \vee$ روی شبکه
 L محدود می کند.

قضیه ۲.۲ فرض کنید $h: L^n \rightarrow L$ یک همسوساز باشد. اگر h
در ش ۲ و ش ۴ صدق کند آنگاه به ازای هر $a_1, \dots, a_n \in L$

$$a_1 \wedge a_2 \wedge \dots \wedge a_n \leq h(a_1, a_2, \dots, a_n) \leq a_1 \vee a_2 \vee \dots \vee a_n \quad (۱)$$

به علاوه چنانچه یک همسوساز h در نامساوی (۱) صدق کند آنگاه h حتماً
دارای خاصیت خودتوانی ش ۴ است.

برهان. فرض کنید h در ش ۲ و ش ۴ صدق کند. آنگاه با فرض

$$a^* = a_1 \vee a_2 \vee \dots \vee a_n, \quad a_* = a_1 \wedge a_2 \wedge \dots \wedge a_n$$

به راحتی می توان دید که

جدول ۱

h	$\circ$	a	b	c	d	$\backslash$
$\circ$	$\circ$	$\circ$	$\circ$	$\circ$	$\circ$	$\circ$
a	$\circ$	a	a	c	c	a
b	$\circ$	a	b	c	c	b
c	$\circ$	c	c	c	c	c
d	$\circ$	c	c	c	d	d
$\backslash$	$\circ$	a	b	c	d	$\backslash$

حال با توجه به شرکتپذیری عمل h روی L می‌توانیم به راحتی آن را به یک همسوساز $L : L^2 \rightarrow L$ توسیع دهیم. این همسوساز دارای همه خواص ش ۱ - ش ۴ است و با همسوسازهای Min و Max فرق می‌کند. مثلاً $\text{Min}(a, c) = a$ در حالی که $h(a, d) = c$. اما به هر حال طبق قضیه ۲.۲ داریم $\text{Min} \leq h \leq \text{Max}$. حال با توجه به h معدل نظرات برابر است با:

$$h(b, c, d) = h(b, h(c, d)) = h(b, c) = c = \text{کمی خوب}$$

که نتیجه معقولی است. اما روی هم رفته معدل‌گیری توسط این همسوساز همه جا رضایت‌بخش نیست. به عنوان مثال معدل ضعیف، متوسط، و کمی خوب به صورت زیر است

$$h(a, b, c) = h(a, h(b, c)) = h(a, c) = c = \text{کمی خوب}$$

که شاید نتیجه جالبی نباشد.

تذکر ۲.۴ همسوسازهای ارائه شده در مثالهای ۱.۲ و ۲.۲ ممکن است قانع‌کننده نباشند ولی ارائه آنها نشان دهنده روش استفاده از آنهاست و همان‌طور که ملاحظه شد ما با داده‌های نادقیق صرفاً به صورت عناصر یک شبکه برخورد کرده‌ایم. البته انتخاب این شبکه‌ها می‌تواند بسیار منطقی صورت گیرد. چنانچه اخیراً در مقالات [۵، ۴] این روش برای استدلال تقریبی توسط مفاهیم نادقیق به کار رفته و نشان داده شده که مفاهیم نادقیق مورد بحث در صورتی که مناسب انتخاب شوند دارای ساختار شبکه‌ای جالبی به نام جبر پرچین^۱ هستند. در اینجا ذکر یک نکته لازم است که هنوز برنگارنده کاملاً روشن نیست که اگر داده‌های نادقیق هموزن نباشند آنگاه انتخاب h در مثالهای ۲.۲ و ۱.۲ چگونه بایستی صورت گیرد؟ که خود نیاز به تحقیق بیشتر دارد.

حال مسأله برمی‌گردد به اینکه یک همسوساز $L : L^2 \rightarrow L$ که مناسب این مدل‌سازی باشد، به دست آوریم به طوری که حتی‌الامکان در همه شرایط ش ۱ - ش ۴ صدق کند. اگر فرض کنیم که $h = \text{Min}$ در این صورت

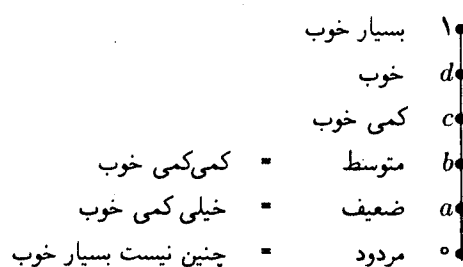
$$h(b, c, d) = \text{Min}(d, \text{Min}(c, d))$$

$$= \text{Min}(b, a) = a = \text{کمی خوب}$$

در این مسأله می‌توان گفت که همسوساز Min یک انتخاب بدبینانه است زیرا که ما در معدل‌گیری بدترین حالت را مدنظر قرار داده‌ایم و در مقابل همسوساز Max یک انتخاب خوشبینانه است، زیرا که در این حالت متمایل به ملاحظه حالت‌های بهتر هستیم. لذا ممکن است انتخاب‌های دیگری از همسوسازها، مدل‌سازی بهتری از مسأله را ارائه دهد.

در مثال ۱.۲ ما ارائه نمره توسط سه استاد را آزاد قرار دادیم. در حالی که ممکن است شبکه نمره‌های خود را از ابتدا مشخص کنیم و سپس هر نمره‌ای را تنها به عنوان انتخابی از این شبکه در نظر بگیریم. در مثال بعد ما این روش را به کار می‌بریم.

مثال ۲.۲ فرض کنید که نمرات فقط می‌توانند از یک شبکه، که تشکیل یک زنجیر متناهی [۱۵] می‌دهند، انتخاب شوند. یعنی همه نمرات با هم قابل مقایسه‌اند. ما در این مثال زنجیر L را به صورت شکل ۲ در نظر گرفته‌ایم.



شکل ۲. شبکه نمرات L که از قبل ساخته شده است.

حال مانند مثال ۱.۲ از سه استاد می‌خواهیم تا نظرشان را در مورد وضعیت تحصیلی یک دانشجو بیان کنند ولی این دفعه مقید به انتخاب از عناصر زنجیر L اند.

فرض کنید پاسخ آنان به صورت خوب (d)، کمی خوب (c) و متوسط (b) باشد. حال می‌خواهیم معدل نظریات این استادان را در مورد این دانشجو به دست آوریم. مجدداً بایستی یک همسوساز مناسب برای این معدل‌گیری انتخاب کنیم. در اینجا با قدری کاوش عمل دوتایی $L : L^2 \rightarrow L$ را به صورت زیر به دست می‌آوریم:

1) hedge algebra

۳ برخی همسوسازهای خاص

اگر L شبکه اعداد حقیقی R^+ یا بازه $[0, 1]$ باشد، که رابطه $\leq$ روی L همان ترتیب معمولی روی اعداد حقیقی است و $\vee$ نیز همان توابع $\min$ و $\max$ روی اعداد حقیقی اند، آنگاه نماد L_* را به جای L به کار می بریم. با این فرض داده های بیان شده در L_* به صورت اعداد حقیقی خواهند بود که برای معدل گیری آنها می توان از معدل گیرهای معمولی استفاده کرد. در این بخش ما به شرح این گونه همسوسازها می پردازیم. خواننده علاقه مند می تواند به مقاله های [۶-۱۳] و کتاب [۱۴] مراجعه کند.

۱.۳ میانگین تعمیم یافته^۱

تعریف ۳.۱.۱ فرض کنید همسوساز $h_\alpha : L_*^n \rightarrow L_*$ به صورت زیر تعریف شود:

$$h_\alpha(a_1, \dots, a_n) = \left(\frac{a_1^\alpha + \dots + a_n^\alpha}{n} \right)^{1/\alpha} \quad (2)$$

که در آن پارامتر $\alpha \in R$ و در حالتی که $\alpha \neq 0$ فرض می کنیم که همه a_i ها مخالف صفرند. در این صورت h_α را میانگین تعمیم یافته گوئیم. به علاوه h_α دارای خواص زیر است:

(۱) اگر $\alpha < 0$ ، $a_i \rightarrow 0$ برای هر $i = 1, 2, \dots, n$ آنگاه $h_\alpha(a_1, \dots, a_n)$

(۲) اگر $\alpha \rightarrow 0$ آنگاه $\lim_{\alpha \rightarrow 0} h_\alpha(a_1, \dots, a_n) = (a_1 \dots a_n)^{1/n}$ یعنی در این حالت مقدار حد همان میانگین هندسی است:

$$\lim_{\alpha \rightarrow -\infty} h_\alpha(a_1, \dots, a_n) = \min(a_1, \dots, a_n) \quad (3)$$

$$\lim_{\alpha \rightarrow +\infty} h_\alpha(a_1, \dots, a_n) = \max(a_1, \dots, a_n) \quad (4)$$

(۵) اگر $\alpha = -1$ آنگاه $h_{-1}(a_1, \dots, a_n) = \frac{n}{1/a_1 + \dots + 1/a_n}$ میانگین هارمونیک است:

(۶) اگر $\alpha = 1$ آنگاه $h_1(a_1, \dots, a_n) = \frac{a_1 + \dots + a_n}{n}$ میانگین حسابی است:

(۷) h_α در ش ۴ صدق می کند.

۲.۳ معدل گیری وزنی^۲

دسته ای از همسوسازها به نام معدل گیر وزنی اند. برای تعریف این معدل گیرها همان طور که از نام آنها برمی آید، باید وابسته به یک بردار وزن باشند.

تعریف ۱.۲.۳ فرض کنید $\underline{w} = (w_1, \dots, w_n) \in [0, 1]^n$ اگر $\sum_{i=1}^n w_i = 1$ آنگاه $\underline{w}$ را یک بردار وزنی گوئیم.

تعریف ۲.۲.۳ فرض کنید $\underline{w} \in [0, 1]^n$ یک بردار وزنی مفروض باشد. همسوساز $h_{\underline{w}} : L_*^n \rightarrow L_*$ را که به صورت زیر تعریف می شود یک معدل گیر وزنی گوئیم:

$$h_{\underline{w}}(a_1, \dots, a_n) = \sum_{i=1}^n w_i a_i$$

دسته خاصی از معدل گیرهای وزنی را می توان در قضیه زیر مشخص کرد.

قضیه ۱.۲.۳ فرض کنید همسوساز $h : L_*^n \rightarrow L_*$ پیوسته باشد و در ش ۴ و شرط مرزی $h(0, 0, \dots, 0) = 0$ صدق کند. به علاوه فرض کنید برای هر $a_i, b_i \in L_*$ که $a_i + b_i \in L_*$ داریم:

$$h(a_1 + b_1, \dots, a_n + b_n) = h(a_1, \dots, a_n) + h(b_1, \dots, b_n)$$

آنگاه h یک معدل گیر وزنی است.

برهان $h_i : L_* \rightarrow L_*$ را به صورت $h_i(x) = h(0, \dots, 0, x, \dots, 0)$ برای هر $i = 1, 2, \dots, n$ که در آن x در i امین مکان است تعریف می کنیم. تابع h_i به لحاظ پیوستگی h پیوسته است و به ازای هر $x, y \in L_*$ داریم:

$$h_i(x + y) = h_i(x) + h_i(y) \quad (3)$$

حال با استفاده از شرط مرزی روی h داریم $h_i(0) = 0$. معادله (۳) یک معادله تابعی کوشی [۱۶] است که از حل آن داریم:

$$h_i(x) = w_i x$$

که در آن $x \in L_*$ و $w_i = h_i(1)$ و $0 < w_i$ به ازای هر $i = 1, 2, \dots, n$ لذا

$$\begin{aligned} h(a_1, \dots, a_n) &= h(a_1, 0, \dots, 0) + h(0, a_2, 0, \dots, 0) \\ &\quad + \dots + h(0, 0, \dots, 0, a_n) \\ &= h_1(a_1) + h_2(a_2) + \dots + h_n(a_n) \\ &= w_1 a_1 + \dots + w_n a_n \\ &= \sum_{i=1}^n w_i a_i \end{aligned}$$

1) generalized mean

2) weighted averaging

۴ مدل سازی مفاهیم نادقیق توسط

مجموعه های مشکک و همسوسازی آنها

در بخش ۲، داده ها به عنوان عناصری صرف از یک شبکه تلقی شدند و سپس روی آنها عمل همسوسازی انجام شد. برخی اوقات می توان ابتدا داده ها را به صورت مجموعه های مشکک مدل سازی کرد و سپس روی آنها عمل همسوسازی را انجام داد [۳،۲]. برای این منظور ابتدا یک تعریف، برخی نمادها و سپس اصل توسعه را ارائه می کنیم.

تعریف ۱.۴ $\bar{A} : X \rightarrow L$ را که در آن X مجموعه ای ناتهی و L یک شبکه است یک L - زیرمجموعه X گوئیم. $\bar{A}(x)$ را به عنوان میزان عضویت x در $\bar{A}$ تعبیر می کنیم. در حالتی که L شبکه $[0, 1]$ باشد، $\bar{A}$ را یک زیرمجموعه مشکک X گوئیم.

در صورتی که $X = \{x_1, \dots, x_n\}$ متناهی باشد، L -مجموعه $\bar{A}$ را به صورت زیر نمایش می دهیم:

$$\bar{A} = \frac{\bar{A}(x_1)}{x_1} + \dots + \frac{\bar{A}(x_n)}{x_n} \quad (4)$$

که در آن $\frac{\bar{A}(x)}{x}$ بدین معنی است که میزان عضویت x در $\bar{A}$ برابر با $\bar{A}(x)$ است و معنی تقسیم نمی دهد. همین طور علامت $+$ در (۴) برای نشان دادن تمامی عناصر و میزان عضویت آنها به کار رفته و به معنی جمع بین عناصر نیست. به علاوه فرض می کنیم:

$$F_L(X) = \{\bar{A} | \bar{A} \text{ یک } L\text{-مجموعه } X = L^X\}$$

و اگر L شبکه $[0, 1]$ باشد آنگاه می نویسیم $F(X)$.

اصل توسعه. فرض کنید $f : X \times X \rightarrow X$ یک عمل دوتایی شرکتپذیر روی مجموعه X و $h : L^n \rightarrow L$ یک همسوساز باشد. f یک تابع $\bar{f}_h$ به صورت زیر القا می کند:

$$\begin{aligned} \bar{f}_h : L^X \times \dots \times L^X &\rightarrow L^X, \\ (\bar{A}_1, \dots, \bar{A}_n) &\mapsto \bar{f}_h(\bar{A}_1, \dots, \bar{A}_n) \bar{f}_h(\bar{A}_1, \dots, \bar{A}_n)(z) \\ &= \begin{cases} \sup_{z=f(x_1, \dots, x_n), x_i \in X} h(\bar{A}_1(x_1), \dots, \bar{A}_n(x_n)) & z \in X \\ 0 & \text{جاهای دیگر} \end{cases} \end{aligned}$$

این تابع جدید $\bar{f}_h$ را توسعه f گوئیم. روند فوق را اصل توسعه نامیم. در

چون h در ش ۴ صدق می کند بنابراین $\underline{w}$ یک بردار وزنی است زیرا به ازای $a \neq 0$ داریم

$$a = h(a, \dots, a) = \sum_{i=1}^n w_i a = a \sum_{i=1}^n w_i$$

پس $\sum_{i=1}^n w_i = 1$ و بنابراین $h = h_{\underline{w}}$ معدل گیری وزنی وابسته به بردار $\underline{w}$ است.

۳.۳ معدل گیری وزنی ترتیبی

در این حالت بایستی معدل گیر به دست آمده اولاً وزنی باشد و ثانیاً یک نوع ترتیب در آن ملاحظه شود. برای این منظور تعریف دسته ای از این معدل گیرها را به صورت زیر ارائه می دهیم:

تعریف ۱.۳.۳ فرض کنید $\underline{w} = (w_1, \dots, w_n) \in [0, 1]^n$ یک بردار وزنی باشد. همسوساز $h_{\underline{w}} : L^n \rightarrow L$ را معدل گیری وزنی ترتیبی گوئیم، اگر

$$h_{\underline{w}}(a_1, \dots, a_n) = w_1 b_1 + \dots + w_n b_n$$

که در آن بردار $(b_1, \dots, b_n)$ بردار نزولی به دست آمده از بردار $(a_1, \dots, a_n)$ است. یعنی b_i ها همان a_i ها مرتب شده اند که $b_1 \geq b_2 \geq \dots \geq b_n$. برای روشن شدن تعریف مثلاً اگر

$$\underline{w} = (0.3, 0.1, 0.2, 0.4)$$

آنگاه

$$h_{\underline{w}}(6, 9, 2, 7) = 0.3 \times 9 + 0.1 \times 7 + 0.2 \times 6 + 0.4 \times 2$$

$h_{\underline{w}}$ دارای خواص زیر است:

(۱) اگر $\underline{w} = (0, \dots, 0, 1)$ آنگاه

$$h_{\underline{w}}(a_1, \dots, a_n) = \min(a_1, \dots, a_n)$$

(۲) اگر $\underline{w} = (1, 0, \dots, 0)$ آنگاه

$$h_{\underline{w}}(a_1, \dots, a_n) = \max(a_1, \dots, a_n)$$

(۳) اگر $\underline{w} = (1/n, \dots, 1/n)$ آنگاه

$$h_{\underline{w}}(a_1, \dots, a_n) = \frac{a_1 + \dots + a_n}{n}$$

(۴) اگر $\underline{w}$ یک بردار وزنی دلخواه باشد آنگاه $h_{\underline{w}}$ در ش ۴ صدق می کند.

بنابراین می توان، با توجه به مجموعه مشکک (۶)، گفت که نمره این دانشجو حول و حوش ۱۷ است.

مثال ۲.۴ در مثال قبل مدل سازی را در $F(X)$ انجام دادیم. فرض کنید که $\bar{A}, \bar{B} \in F_L(x)$ که در آن $X = [0, 20]$ و L شبکه مثال ۲.۲ است که در شکل ۲ ارائه شده است. حال $\bar{A}, \bar{B}$ را به صورت زیر بیان می کنیم:

$$\bar{A} = \frac{b}{17} + \frac{d}{18} + \frac{c}{19}, \quad \bar{B} = \frac{a}{15} + \frac{d}{16} + \frac{c}{17}$$

فرض کنید که $f(x, y) = \frac{x+y}{2}$. به علاوه فرض کنید که همسوزا h روی L همانند مثال ۲.۲ باشد. در این صورت با استفاده از (۵) به دست می آوریم:

$$\frac{\bar{A} + \bar{B}}{2} = \bar{f}_h(\bar{A}, \bar{B}) = \frac{a}{16} + \frac{c}{16.5} + \frac{c}{17} + \frac{c}{17.5} + \frac{c}{18}$$

لذا نتیجه ای که می توان ارائه کرد این است که با اطمینان نسبتاً خوب، نمره دانشجو در فاصله $[16.5, 18]$ قرار دارد.

نتیجه گیری

با توجه به مثالهای ارائه شده می توان یکی از عیبهای روش استفاده از مدل سازی مفاهیم نادقیق توسط L مجموعه ها و سپس معدل گیری آنها را در این دید که اصولاً معدل گیری توسط یک L مجموعه در عمل مشکل است و بستگی به شخص و محیط تصمیم گیری دارد که مدل سازی را ناپایدار می کند. عیب دوم اینکه $\bar{f}_h$ خود توان نیست، یعنی همواره $\bar{A} = \bar{f}_h(\bar{A}, \dots, \bar{A})$ نیست که این برخلاف انتظار است. این موضوع را می توان به راحتی در دو مثال ۱.۴ و ۲.۴ تحقیق کرد. اما در روش استفاده از عملگرهای همسوزا شبکه ای و تلقی داده ها به صورت عناصر یک شبکه با دو عیب ذکر شده قبل مواجه نیستیم. یعنی مشکل ناپایداری مدل سازی را نداریم و به علاوه با فرض خود توانی h داریم که $h(a, a) = a$. در عوض عیب این روش این است که ساختن خود h مشکل است و به علاوه تاکنون روشی برای وزن دهی به داده ها موجود نیست و نیاز به تحقیق دارد.

اصل توسعه ملاحظه می شود که برای به دست آوردن $\bar{f}_h$ نیاز به یک عملگر همسوزا داریم. چنانچه L شبکه $[0, 1]$ باشد، در این صورت بسته به نوع مسأله می توان h را میانگین تعمیم یافته در بخش ۱.۳ یا معدل گیر وزنی در بخش ۲.۳ انتخاب کرد. مثلاً ما در مثال ۱.۴ همسوزا خود را ساده ترین حالت یعنی $h(x, y) = \min(x, y)$ انتخاب کرده ایم. در حالی که در مثال ۲.۴ همسوزا h را همانند جدول ۱ گرفته ایم. اما در هر دو حالت خواننده می تواند همسوزاهای دیگری را جایگزین کند. حال برای روشن شدن مطلب به ارائه مثال می پردازیم.

مثال ۱.۴ فرض کنید نظر دو استاد را در مورد وضعیت تحصیلی یک دانشجو سؤال کرده ایم. پاسخ آنان به صورت: خیلی خوب ($\bar{A}$) و متوسط ($\bar{B}$) بیان شده است. حال سؤال می شود معدل تحصیلی این دانشجو چیست؟

برای پاسخ به این سؤال ابتدا فرض می کنیم که $\bar{A}, \bar{B} \in F(X)$ که در آن $X = [0, 20]$ فضای نمره است. یعنی در اینجا نمره ها به صورت نادقیق اند اما مدل سازی آنها توسط مجموعه های مشکک انجام شده که به صورت زیرند:

$$\bar{A} = \frac{0.8}{17} + \frac{1}{18} + \frac{0.9}{19}, \quad \bar{B} = \frac{0.5}{15} + \frac{1}{16} + \frac{0.5}{17}$$

ملاحظه می کنیم که مدل سازی این مفاهیم گرچه با عقل سلیم مطابقت دارد ولی کاملاً ذهنی اند و بستگی به شخص و محیط تصمیم گیری دارند. مثلاً در اینجا $\bar{A}$ به این معنی است که نمره دانشجو حول و حوش ۱۸ است و $\bar{B}$ به معنی اینکه نمره دانشجو حول و حوش ۱۶ است. حال برای معدل گیری $\bar{A}, \bar{B}$ به صورت $\frac{\bar{A} + \bar{B}}{2}$ از اصل توسعه استفاده می کنیم. اگر فرض کنیم که $f(x, y) = \frac{x+y}{2}$ ، $h(x, y) = \min(x, y)$ و $L = [0, 1]$ آنگاه با محاسبه (۵) به راحتی می توان دید که

$$\begin{aligned} \frac{\bar{A} + \bar{B}}{2} &= \bar{f}_h(\bar{A}, \bar{B}) \\ &= \frac{0.5}{16} + \frac{0.8}{16.5} + \frac{1}{17} + \frac{0.9}{17.5} + \frac{0.5}{18} \end{aligned}$$

مراجع

- [۱] جواد بهبودیان، آمار و احتمال مقدماتی، بنیاد فرهنگی رضوی، ۱۳۶۸.
- [۲] سید محمود طاهری، آشنایی با نظریه مجموعه های فازی، انتشارات جهاد دانشگاهی، ۱۳۷۵.
- [۳] ماشاءالله ماشین چی، ریاضیات مفاهیم نادقیق و سیستمهای هوشمند، گزارش کامپیوتر، شماره ۱ سال هجدهم شماره پیاپی ۱۳۰، صفحات ۲۵-۳۰.
- [4] N. Cat Ho and W. Wechler, Hedge algebras: An algebraic approach to structure of set of linguistic truth values, Fuzzy Set and Systems 35 (1990) 281-293.
- [5] N. Cat Ho and W. Wechler, Extended hedge algebras and their application to fuzzy logic, Fuzzy Sets and Systems 52 (1992) 259-281.

- [6] R. R. Yager, Constrained OWA aggregation, Fuzzy Sets and Systems 81 (1996) 89-101.
- [7] R. R. Yager, Aggregation operators and fuzzy system modeling, Fuzzy Sets and Systems 67 (1994) 129-125.
- [8] R. R. Yager, Families OWA operators, Fuzzy Sets and Systems 59 (1993) 125-148.
- [9] R. R. Yager, Connective and quantifiers in fuzzy sets, Fuzzy Sets and Systems 40 (1991) 39-76.
- [10] R. R. Yager, On ordered weighted averaging aggregation in multicriteria decision making IEEE Tran. Systems, Man Cybernet 18 (1988) 183-190.
- [11] R. R. Yager, On mean type aggregation IEEE Tran. Systems, Man Cybernet Part B, 26 (1996) 209-326.
- [12] H. Dyckhoff and W. Pedrycz, Generalized means as model of compensative connectives, Fuzzy Sets and Systems 14 (1984) 143-154.
- [13] A. L. Ralescu and D. A. Ralescu, New concepts of fuzzy aggregation, Fuzzy Sets and Systems, to appear.
- [14] G. J. Klir and Bo Yuan, Fuzzy sets and fuzzy logic, Theory and Applications, Prentice Hall, 1995.
- [15] G. Birkhoff, Lattice theory, American Mathematical Society, Providence, 1967.
- [16] J. Aczel, Lectures on functional equation and their applications, Academic Press, New York, 1966.
- [17] Vijay K. Khanna, Lattices and Boolean algebras, Vikas Publishing House PVT 1994.

این مقاله با حمایت مالی مرکز پژوهشی ریاضی ماهانی دانشگاه شهید باهنر کرمان تهیه شده است.
