

بسم الله الرحمن الرحيم



فرهنگ و اندیشه ریاضی

(علمی-ترویجی)

ISSN: 1022-6443

سال ۳۵، پاییز و زمستان ۱۳۹۵

(تاریخ انتشار: زمستان ۱۳۹۵)

شماره پیاپی: ۵۹

صاحب امتیاز: انجمن ریاضی ایران
مدیر مسئول: سید منصور واعظپور
سردبیر: روح الله جهانیپور
ویراستار ارشد: سعید مقصودی
مدیر اجرایی: مژگان محمودی

هیأت تحریریه:

jamali@khu.ac.ir
jahanipu@kashanu.ac.ir
Rafiepour@uk.ac.ir ، drafiepour@gmail.com
mmahmoudi@sharif.ir
mohamadian_r@scu.ac.ir
m-mahmoudi@sbu.ac.ir
s_maghsodi@znu.ac.ir
a-nazari@araku.ac.ir
hoseynhashemi@gmail.com

علیرضا جمالی، دانشگاه خوارزمی
روح الله جهانیپور، دانشگاه کاشان
ابوالفضل رفیعپور، دانشگاه شهید باهنر کرمان
محمد غلامزاده محمودی، دانشگاه صنعتی شریف
رستم محمدیان، دانشگاه شهید چمران اهواز
مژگان محمودی، دانشگاه شهید بهشتی
سعید مقصودی، دانشگاه زنجان
علی محمد نظری، دانشگاه اراک
بهنام هاشمی، دانشگاه صنعتی شیراز
حروف چینی: X-Persian - دفتر انجمن ریاضی
تهیه و تنظیم: فریده صمدیان

۸۸۸۰۷۷۷۵، ۸۸۸۰۷۷۹۵، ۸۸۸۰۸۸۵۵
iranmath@ims.ir
http://www.ims.ir
mct.iranjournals.ir

نشانی: تهران- صندوق پستی ۴۱۸-۱۳۱۴۵
شماره های تلفن دفتر انجمن ریاضی ایران:
رایانامه نشریه:
نشانی اینترنتی انجمن ریاضی ایران:
سامانه اینترنتی نشریه:

راهنمای علاقه‌مندان به ارسال مقاله برای «فرهنگ و اندیشه ریاضی»

«فرهنگ و اندیشه ریاضی»، نشریه علمی-ترویجی انجمن ریاضی ایران است و هدف از انتشار آن، آشنا کردن خوانندگان با شاخه‌های گوناگون علوم ریاضی، فرهنگ، تاریخ و فلسفه ریاضی است به گونه‌ای که برای دانشجویان، دبیران ریاضی، اعضای هیأت علمی گروه‌های ریاضی دانشگاه‌ها و سایر علاقه‌مندان به ریاضی قابل استفاده باشد. «فرهنگ و اندیشه ریاضی»، از مقالات در زمینه‌های علوم ریاضی، تدریس، یادگیری و آموزش ریاضی، تاریخ و فلسفه ریاضی و کاربردهای ریاضی در علوم دیگر که در چارچوب زیر نوشته شده باشند، استقبال می‌کند:

- ارائه موضوعی به‌روز و مطرح در ریاضیات چنان‌که علاقه‌مندان به زمینه‌های پژوهشی را برای پیگیری آن موضوع آماده سازد؛
- ترجمه مقاله‌هایی از نوع یاد شده در بالا یا ترجمه مقاله‌های کلاسیک ریاضی با رعایت اصول ترجمه؛
- ارائه موضوعات آموزشی حاوی نکته‌های درسی سودمند برای کلاس درس و برهان‌هایی ساده‌تر یا کوتاه‌تر یا زیباتر از آنچه در متون کلاسیک موجود است؛
- نقد کتاب‌های منتشر شده به زبان فارسی در زمینه علوم ریاضی اعم از تألیف یا ترجمه؛
- مقالات بایستی با استفاده از نرم‌افزار «زی‌پرشین» تهیه شده و نسخه پی‌دی‌اف آن از طریق سامانه اینترنتی نشریه ارسال شود. در صورتی که مقاله فرستاده شده، پس از طی مراحل داوری، برای چاپ پذیرفته شود، نویسندگان یا مترجم باید فایل «زی‌پرشین» مقاله را از طریق سامانه نشریه ارسال کنند؛
- فرستادن اصل مقاله ترجمه شده همراه با ترجمه آن الزامی است؛
- نام و نشانی اصل مقاله ترجمه شده باید به‌صورت پاورقی در صفحه اول ترجمه مقاله ذکر شود؛
- اصطلاحات ریاضی به‌کار رفته باید بر طبق واژه‌نامه ریاضی و آمار انجمن ریاضی ایران، چاپ مرکز نشر دانشگاهی باشد و اگر لغتی در این واژه‌نامه نیست، معادل انگلیسی آن در پاورقی ذکر شود؛
- مقاله ارسالی به «فرهنگ و اندیشه ریاضی» نباید برای بررسی و چاپ به مجله‌های دیگر فرستاده شده باشد؛
- مقاله ارسالی باید دربرگیرنده نام و نام خانوادگی، مرتبه علمی، آدرس کامل پستی و رایانامه نویسنده(گان) یا مترجم(ان) باشد؛
- در ابتدای مقاله، چکیده‌ای در حداکثر ۲۰۰ کلمه شامل نکات اصلی بیان شده در مقاله، آورده شود؛
- پس از چکیده، واژه‌ها و اصطلاحات کلیدی (حداکثر ۵ تا) مشخص شود؛
- رده‌بندی موضوعی اولیه و ثانویه مقاله، بر مبنای رده‌بندی موضوعی MSC^{۲۰۱۰} ذکر شود؛
- اسامی افراد خارجی در متن مقاله باید به‌صورت فارسی و در پاورقی به زبان اصلی نوشته شود؛
- به‌منظور تسریع در ویرایش و چاپ مقالات، لطفاً اصول سجاوندی، نگارش و ویرایش در نوشتن مقالات را بر مبنای شیوه‌نامه نگارش مرکز نشر دانشگاهی، رعایت کنید؛
- فهرست مراجع را در انتهای مقاله، به‌ترتیب حروف الفبای نام خانوادگی افراد تنظیم کنید؛ نه به ترتیبی که در مقاله ارجاع می‌دهید. ضمناً در نوشتن مراجع، سبک زیر را رعایت کنید:

۱. اگر مرجع ذکر شده، مقاله لاتین است، عنوان مقاله به‌صورت رومن، نام مجله به‌صورت ایتالیک، شماره مجلد آن به‌صورت سیاه، سال چاپ داخل پرانتز و در پایان، صفحه‌هایی از مجله که مقاله مرجع در آنها چاپ شده است، تایپ شود؛
۲. اگر مرجع ذکر شده، کتاب لاتین است، نام کتاب به‌صورت ایتالیک تایپ شود؛
- تعداد صفحه‌های مقاله ارسالی در قالب صفحه‌بندی مجله، ترجیحاً از ۲۵ صفحه بیشتر نشود؛
- هیأت تحریریه در رد، قبول و ویرایش مقالات آزاد است.

«فرهنگ و اندیشه ریاضی» امسال در دو شماره (بهار و پاییز) منتشر و برای اعضای حقیقی، حقوقی و مشترکین انجمن ریاضی ایران ارسال می‌شود. شماره‌های قبلی مجله را می‌توانید از دبیرخانه انجمن ریاضی ایران خریداری فرمایید. نمونه فایل مقاله با قالب تهیه شده برای مجله، بر روی سامانه موجود است.

فهرست مطالب

نوع مقاله‌های «فرهنگ و اندیشه ریاضی»	
روح‌الله جهانی پور	۱
سیر تاریخی فلسفه ریاضی	
سیاوش شهشپهانی	۷
میراث فوریه	
ژان-پیر کاهان، برگردان: سعید مقصودی	۲۹
مقایسه‌ای از رویکردهای نیمن-پیرسون، بیز و کم-بیشینه در آزمون فرضیه‌های فازی	
عباس پرچمی، سید محمود طاهری، ناصر رضا ارقامی	۴۳
تجزیه نامنفی ماتریسی: روشی برای تحلیل داده‌های نامنفی	
مهسا یوسفی	۷۱
حجم گوی‌های واحد تعمیم‌یافته	
زیانفو وانگ، برگردان: رضا کهکشانی و رسول کاظمی نجف‌آبادی	۹۱
رابطه‌ای معادل با رابطه فیثاغورس و سه‌تایی فیثاغورسی وارون	
جواد بهبودیان، صدیقه بلادی، مهرداد سیمکانی	۹۹



عکس روی جلد: دکتر سیاوش شهشپهانی، استاد دانشگاه صنعتی شریف

نوع مقاله‌های «فرهنگ و اندیشه ریاضی»*

روح‌الله جهانی‌پور

سی و چهار سال از انتشار نخستین شماره مجله «فرهنگ و اندیشه ریاضی» می‌گذرد. در این سال‌ها، پیشامدهای گوناگونی هم در بخش فرهنگ و هم در حوزه اندیشه در ریاضیات کشور رخ داده است، اما چنین می‌نماید که فرهنگ اندیشیدن در ریاضیات، به فراموشی سپرده شده یا از اهمیت آن کاسته شده است. هدف کلی بنیانگذاران «فرهنگ و اندیشه ریاضی»، نه چاپ مقاله‌های پژوهشی یا شبه‌پژوهشی در شاخه‌های خاص ریاضی که آنها را به پیروی از گردانندگان مجله «جنگ ریاضی دانشجو»، مقاله‌های فنی می‌نامیم، بلکه کوشش برای رشد و گسترش فرهنگ ریاضی‌خوانی، ریاضی‌نویسی و اندیشیدن در ریاضیات به معنای عام، از مسیر انتشار مقاله‌های توصیفی به شیوه کلی‌نگرانه بوده است که طیف گسترده‌ای از خوانندگان از دانشجویان گرفته تا استادان ریاضی دانشگاه‌ها با مطالعه آنها، با بخش‌های گوناگون ریاضیات آشنا شوند و یا به دانش خود ژرفای بیشتری ببخشند. گرچه تاکنون مقاله‌های خوبی در این چارچوب در «فرهنگ و اندیشه ریاضی» چاپ شده است، گویا برای برخی از دوستان دلسوزی که می‌خواهند خود را از این فضای پرهیاهوی نوشتن و چاپ مقاله‌های پژوهشی و گرفتن امتیاز آنها! دور کنند و با نوشتن یا ترجمه مقاله‌های ترویجی و مروری، «فرهنگ و اندیشه ریاضی» را در رسیدن به هدفش یاری رسانند، هنوز روشن نیست که چه نوع مقاله‌هایی در «فرهنگ و اندیشه ریاضی» چاپ می‌شود و به‌ویژه منظور از مقاله توصیفی چیست. تفاوت مقاله توصیفی با مقاله فنی در محتوا، سبک نگارش و هدف آن دو است. مقاله فنی درباره (گوشه‌ای از) یک موضوع بسیار ویژه از یک شاخه خاص ریاضی نگاشته می‌شود، در حالی که نویسنده مقاله توصیفی، سیر رشد و تحول اندیشه در یک موضوع کلی یا بخشی بزرگ (کلاسیک یا جدید) از ریاضیات را به لحاظ تاریخی، فلسفی و شخصیت‌شناسی شرح می‌دهد. در مقاله فنی، پس از یک بیان تاریخی کوتاه درباره کارهایی که پژوهشگران پیشین انجام داده‌اند، محتوا به شیوه «تعریف-قضیه-برهان» ارائه می‌شود، در حالی که در مقاله توصیفی بنا بر ارائه مطلب به سبکی زنده، هیجان‌انگیز و خودکفا است

* سخن سردبیر.

به‌طوری که شمار بسیاری از خوانندگان (حتی با کمترین سطح از دانش مربوط به موضوع مقاله) مطالعه مقاله را تا پایان با علاقه دنبال و احساس کنند بر دانش‌شان درباره آن موضوع افزوده شده است. گذشته از این‌ها، خواننده یک مقاله فنی، معمولاً متخصص در بخش ویژه‌ای از ریاضیات است که درونمایه مقاله با آن بخش پیوند دارد. اما خواننده مقاله توصیفی، برحسب سطح دشواری یا آسانی مطلب، ممکن است دانشجوی سال اول دانشگاه در رشته ریاضی یا معلم ریاضی یا حتی دانش‌آموز دبیرستان هم باشد. بنابراین بسیاری از مقاله‌هایی که می‌توان آنها را خلاصه پایان‌نامه‌های کارشناسی ارشد یا گوشه‌ای از پایان‌نامه دکتری دانشجویان رشته ریاضی دانست، از دسته توصیفی نیستند. آن مقاله‌هایی هم که با بیان و اثبات چند قضیه می‌خواهند برخی از ویژگی‌های موجوداتی خاص را که تنها در خلوتگاه ذهن متخصصان زیست می‌کنند، بر خواننده آشکار سازند، به‌روشنی در چارچوب مقاله‌های توصیفی جای نمی‌گیرند.

برای اینکه بنویسیم، نخست باید بخوانیم! هر اندازه بکوشیم تعریفی فراگیر از مقاله توصیفی ارائه کنیم، تأثیرش در خواننده آن گونه نیست که از خواندن یک مقاله توصیفی واقعی به‌دست می‌آید. افزون بر «فرهنگ و اندیشه ریاضی»، «نشر ریاضی» مرکز نشر دانشگاهی و «جنگ ریاضی دانشجویی» دانشگاه تهران، مجله‌هایی بودند که در گذشته‌های نه‌چندان دور منتشر می‌شدند و انتشار آنها نقشی به یاد ماندنی در گسترش فرهنگ خواندن و نوشتن مقاله‌های توصیفی ریاضی و رشد آگاهی‌های دانشجویان رشته ریاضی آن دوران در مورد محتوای درس‌هایی که می‌خواندند و انگیزه پیدایش موضوع‌های گوناگون ریاضی داشته است. خوشبختانه فایل‌های چاپی بسیاری از شماره‌های پیشین «فرهنگ و اندیشه ریاضی» به‌صورت رایگان روی صفحه اصلی سامانه اینترنتی این مجله قرار دارد و خواننده علاقه‌مند به آشنایی با مقاله توصیفی، می‌تواند از خواندن مقاله‌هایی که تاکنون در این دسته نگاشته و در «فرهنگ و اندیشه ریاضی» چاپ شده‌اند، بهره‌برد و اگر موفق شود به شماره‌های دو مجله دیگری که نام بردیم نیز دسترسی پیدا کند و آثار تألیفی و ترجمه‌ای زیبایی را که از بزرگان و پیشکسوتان باتجربه جامعه ریاضی کشور در آنجا چاپ شده است، بخواند، بی‌تردید مسیر این آشنایی را درست پیموده است. برای نمونه، یکی از این گونه مقاله‌ها را با نام *سیر تاریخی فلسفه ریاضیات* نوشته دکتر سیاوش شهشهانی که در شماره اول «فرهنگ و اندیشه ریاضی» در سال ۱۳۶۱ چاپ شده است، با ویرایش جدید در این شماره بازنشر می‌کنیم تا هم خاطره گذشته را در ذهنمان زنده کند، هم از خواندن دوباره آن درس بگیریم و هم با شیوه نگارش مقاله توصیفی آشنا شویم.

سواي خواندن مقاله‌های توصیفی، برای نوشتن یک مقاله توصیفی خوب، حتی در بخشی کوچک از ریاضیات، باید هرچه بیشتر ریاضیات بخوانیم! متأسفانه در سالهای اخیر با گسترش بدون کیفیت دوره‌های دکتری ریاضی در دانشگاه‌ها، کاسته شدن از طول دوره دکتری و کم شدن تعداد واحدها در بخش آموزشی این دوره و گرایش به افزایش آهنگ زمانی نوشتن و چاپ مقاله، بسیاری از دانشجویان و دانش‌آموختگان

دکتری که باید در آینده در جایگاه عضو هیأت علمی دانشگاه، آموزشگر و گسترش‌دهنده فرهنگ ریاضی باشند، تک‌بعدی و جزئی‌نگر پرورش می‌یابند و نه تنها با تاریخ، انگیزه‌ها و فلسفه پیدایش موجودات ریاضی (حتی در حوزه تخصصی خودشان) آشنایی ندارند، به پیوندهای میان شاخه‌های گوناگون ریاضی هم آگاه نیستند. این در حالی است که با نگاهی به فهرست نام‌های ریاضیدانانی که در سال‌های اخیر موفق به گرفتن مدال فیلدز شده‌اند، به‌روشنی درمی‌یابیم که آن کارهای پژوهشی شایسته اعطای این مدال بوده‌اند که موضوعی پژوهشی در میانه دو یا چند شاخه از ریاضیات را به پیش رانده‌اند. مقصود از بیان این مطلب این نیست که جهت‌گیری کارهای پژوهشی‌مان را از چاپ مقاله یک‌باره به سوی دریافت مدال فیلدز بگردانیم! یا آنقدر ریاضیات بلد باشیم که بتوانیم مانند هرمان وایل، ریاضیات نیمه اول قرن بیستم را توصیف کنیم چرا که در پایان قرن بیستم هم کسی پیدا نشد که کاری مشابه برای نیمه دوم آن قرن انجام دهد. می‌خواهیم بگوییم که برای به‌دست آوردن دریافتی ژرف از گستره پهنای ریاضیات و پیوندهای میان شاخه‌های گوناگون آن و سپس نوشتن مقاله توصیفی درباره آن، لازم است از همه جای ریاضیات و علوم دیگری که انگیزه‌بخش مسئله‌های ریاضی هستند، به‌ویژه فیزیک، اندیشمندانه مطلب بخوانیم.

اگر کسی را توان نوشتن نیست ولی علاقه‌مند به ترجمه است، می‌تواند مقاله‌های توصیفی مناسبی را که در برخی نشریه‌های معتبر خارجی^۱ چاپ می‌شوند، انتخاب و پس از معرفی به هیأت تحریریه «فرهنگ و اندیشه ریاضی» و گرفتن تأییدیه، ترجمه کند. برای اینکه خوب ترجمه کنیم، با فرض اینکه بر موضوع مقاله آگاهی داریم، در گام نخست باید هم بر دستور و اصول نگارش زبانی که مقاله با آن نوشته شده است مسلط باشیم و هم بر دستور و نگارش زبان فارسی و در گام دوم، با روش‌ها و ظرافت‌های کار مشکل ترجمه نیز آشنا باشیم. به‌ویژه حفظ/مانت در ترجمه که از مؤلفه‌های اصلی فرایند ترجمه به‌شمار می‌آید، به معنای ترجمه کلمه‌به‌کلمه نیست، بلکه مترجم باید ضمن پایبندی به سبک نگارش نویسنده و هماهنگی متن ترجمه‌شده با متن اصلی، تلاش برای فهم درست خواننده را از محتوای مقاله و مقصود نویسنده، وظیفه خود بداند. برای خوب ترجمه کردن، در دست داشتن ابزارها به‌تنهایی کافی نیست. باید ترجمه‌های خوب را هم خواند و راهنمای علاقه‌مندان در این مورد نیز همان است که در بند پیش گفتیم.

دسته‌ای دیگر از مقاله‌های مناسب برای چاپ در «فرهنگ و اندیشه ریاضی»، با فرهنگ خواندن و نقد کتاب پیوند دارد. البته مقصود از نقد، بررسی همه جانبه و بیان نقاط قوت و یادآوری نقاط ضعف اثر است نه نفی و حمله به آن. کتاب‌های یکدیگر را، چه تألیف باشد چه ترجمه، حتی اگر نمی‌خواهیم نقد کنیم یا برای تدریس مناسب نمی‌دانیم، خوب است که بخوانیم و اشکالات و کاستی‌های آنها را به نویسنده یا مترجم یادآور شویم تا با اصلاح آنها در ویرایش‌های بعدی، کتاب هم از دیدگاه علمی و به‌لحاظ نگارشی، بهبود یابد و رفته رفته ماندگار شود. قرن‌ها است نتایج پژوهش‌ها در حوزه‌های علوم دینی نه در قالب مقاله

^۱ مثلاً American Mathematical Monthly یا Mathematics Magazine که از مجموعه انتشارات MAA هستند.

(های ISI یا رده‌بندی‌های غیرعلمی دیگری از این دست)، بلکه در قالب کتاب منتشر می‌شود و به دست دیگران می‌رسد. پژوهشگرانِ دیگر کتاب را می‌خوانند، درس می‌دهند، بر آن نقد و حاشیه می‌نویسند و پس از مدتی، کار به جایی می‌رسد که صاحب اثر نه با نام خودش، بلکه با نام کتابش شناخته می‌شود. چه خوب است در کنار روش‌های کنونی نشر نتایج کارهای پژوهشی، این سنت دیرینه را که از نگاه علمی، آزمون‌شده است، یاد بگیریم و در مجموعه علوم دانشگاهی نیز به کار ببریم و فرهنگ آن را گسترش دهیم. سال‌ها است بر پژوهشگران آموزش در علم، آشکار شده است که آشنایی با تاریخ و چگونگی رشد و نمو یک شاخه از علم، تأثیری ژرف در آموزش و شناخت درست آن دارد. به‌ویژه آگاهی از زندگی‌نامه و کوشش‌های علمی و شیوه اندیشیدنِ چهره‌های برجسته آن شاخه علمی، نه تنها به ما می‌آموزد که برای کسب پیروزی‌های علمی، زندگی خویش را چگونه برنامه‌ریزی کنیم، بلکه رهنمودی برای دوستداران پژوهش در آن حوزه نیز هست. آنهایی که سابقه‌ای طولانی در مطالعه، تدریس و پژوهش در ریاضیات دارند، بی‌گمان با نام، زندگی‌نامه و فعالیت‌های پژوهشی ریاضیدانان بزرگ، دست‌کم در محدوده تخصص خود، آشنایی دارند. اما لازم است دانشجویان و دانش‌آموختگان ریاضی نیز این مسیر را بپیمایند. البته روشن است که بخش بزرگی از ریاضیات کنونی، از جنبه‌های محتوایی، فلسفی و نگارشی، دستاوردی ایرانی نیست و از این رو تاریخ این علم، به‌ویژه از قرن هفدهم میلادی به این سو، بخشی از تاریخ علم غرب به‌شمار می‌آید. با وجود این، هر ایرانی دوستدارِ ریاضی کنج‌کاو است از زندگی و کارهای علمی ریاضیدانان ایرانی دوران گذشته و کنونی که بخشی از فرهنگ این مرز و بوم است، آگاه شود. گرچه افرادی همچون مرحوم ابوالقاسم قربانی، عمر خود را در راه شناساندن ریاضیدانان ایرانی-اسلامی و کاوش در کارهای پژوهشی آنان که واقعاً وقت، دقت و حوصله بسیار می‌خواهد، صرف کردند و آثاری به‌یاد ماندنی از خود بر جای گذاشتند، هنوز باید کارهای بسیاری در این زمینه انجام شود. بنابراین «فرهنگ و اندیشه ریاضی» بر آن است که مقاله‌هایی با موضوع زندگی‌نامه و شرح پژوهش‌های برگزیده ریاضیدانان برجسته و تأثیرگذار در شاخه‌های گوناگون ریاضی را نیز، ایرانی باشند یا غیر ایرانی، چاپ کند.

نوعی مقاله کوتاه هم وجود دارد که در آن، یک مطلب درسی کلاسیک به گونه‌ای دیگر آموزش داده می‌شود؛ یا برای قضیه‌ای شناخته شده، اثباتی کوتاه‌تر، زیباتر یا ساده‌تر ارائه می‌شود؛ یا مسئله‌ای زیبا طرح و حل می‌شود چنان‌که هم دانشجویان می‌توانند آن مقاله را بخوانند و لذت ببرند و هم استادان ریاضی در کلاس‌های درسشان از آن بهره گیرند؛ یا نویسنده تجربه خود را در تدریس نوآورانه یک موضوع ریاضی دانشگاهی به دانشجویان، برای خوانندگان شرح می‌دهد تا آنها هم آن روش را در کلاس درس خود بیازمایند. در «فرهنگ و اندیشه ریاضی» از این دست مقاله‌ها نیز، چه تألیف باشند چه ترجمه، استقبال می‌شود.

هیأت تحریریه دوره جدید «فرهنگ و اندیشه ریاضی» که کار خود را از پاییز سال ۱۳۹۵ آغاز کرده است، بر خود لازم می‌داند از زحمات هیأت تحریریه دوره قبل به‌ویژه آقایان دکتر احمد صفاپور از

دانشگاه ولی‌عصر رفسنجان، دکتر احسان ممتحن از دانشگاه یاسوج، دکتر رشید زارع نهندی از دانشگاه تحصیلات تکمیلی علوم پایه زنجان، دکتر عبدالعزیز عبداللهی از دانشگاه شیراز، دکتر عین‌الله پاشا از دانشگاه خوارزمی و خانم دکتر شیوا زمانی از دانشگاه صنعتی شریف که جمع ما را ترک کردند، تشکر و قدردانی نماید. همچنین زحمات بی‌دریغ خانم فریده صمدیان را که با صبر و حوصله فراوان کار سخت آماده‌سازی فنی مجله را بر عهده داشته و دارند، ارج می‌نهم. ضمناً یکی از اهدافی که هیأت تحریریه دوره قبل پیش روی خود قرارداد، انتشار ویژه‌نامه‌هایی با موضوع‌های به‌روز و زنده در ریاضیات بود که پیامد آن انتشار ویژه‌نامه آموزش ریاضی به‌مناسبت درگذشت زنده‌یاد پرویز شهریاری، معلم بزرگ ریاضی کشور، با همکاری و تلاش تحسین برانگیز خانم دکتر زهرا گویا از دانشگاه شهید بهشتی، در نیمه نخست سال ۱۳۹۵ بود. در دوره کنونی نیز این هدف دنبال خواهد شد و امیدواریم با همکاری جامعه ریاضی کشور بتوانیم ویژه‌نامه‌های دیگری با کیفیت بالا منتشر کنیم.

سیر تاریخی فلسفه ریاضیات*

سیاوش شهشپانی

چکیده

در این مقاله، تاریخ فلسفه ریاضی را در سه دوره مورد مطالعه قرار می‌دهیم. دوره اول، از ادوار باستانی تا اواخر قرن هفدهم میلادی؛ دوره دوم که با ابداع حساب دیفرانسیل و انتگرال توسط نیوتن و لایبنیتس آغاز می‌شود و حدوداً دو قرن به درازا می‌کشد؛ و دوره سوم، از زمان دقیق‌سازی مبانی آنالیز ریاضی توسط وایشراس و دیگران تا به امروز. همچنین سه مکتب فلسفی منطق‌گرایی، صورت‌گرایی و شهودگرایی در ریاضیات را نقد و نقاط ضعف و قوت آنها را بررسی می‌کنیم. نشان خواهیم داد که این سه مکتب غالب در فلسفه ریاضی، هر یک با روی گردانیدن از تجربه عینی ریاضی‌کاران، به مقوله‌ای مستقل از ریاضیات تبدیل شده است که جوابگوی کنجکاوی‌های فلسفی ریاضیدانان حرفه‌ای نیست.

۱. سرآغاز

«... روش توصیف تاریخی ... روش محبوب سخنرانان وحشت‌زده است که محدودیت‌های آن را حربه‌ای برای مهار کردن لفاظی حرص‌آلود خود می‌یابند.»^۱ اما بی‌شک اگر بررسی تاریخی، کمکی به کاهش لفاظی باشد، باید از آن بهره جست. ولی در مطالعه فلسفه ریاضیات، دلایل مثبت‌تری نیز برای به‌کارگرفتن بررسی تاریخی وجود دارد. یکی از مضمون‌هایی که سعی خواهیم کرد در این سخنرانی بپرورانم،

عبارات و کلمات کلیدی. فلسفه ریاضی، صورت‌گرایی، منطق‌گرایی، شهودگرایی، استدلال در ریاضیات.
* این مقاله بر اساس یک سخنرانی در «سمینار فلسفه، منطق و روش‌شناسی ریاضی» که در اردیبهشت‌ماه سال ۱۳۵۹ ایراد شده، تهیه شده است. این مقاله را بیش از ۳۵ سال پیش نوشته‌ام و اکنون نسبت به برخی مطالب آن نظری متفاوت دارم. با این حال بهتر دیدم فقط به چند تصحیح کوچک، اکثراً چاپی، اکتفا کنم (نگارنده).
^۱ داستایوفسکی در رمان برادران کارامازوف، کتاب دوازدهم، بخش نهم.

ارتباط مستقیم تمایل‌های فلسفی در هر دوره از تاریخ ریاضیات با چهره و واقعیت ریاضیات آن دوره است. خواهیم دید که گرایش‌های متداول یک دوره، هرچند که از دید معاصران طبیعی و بدیهی به نظر برسد، ممکن است عکس‌العمل یا طغیانی بیش نبوده باشد. به عکس، آنچه که امروز در نظر ما نوعی جزمی‌گرایی در بعضی دوره‌های تاریخ ریاضیات به نظر می‌رسد، ممکن است در زمان خود، نقشی مهم در حفظ یکپارچگی و بقای این علم دارا بوده باشد.

به‌طور کلی، فلسفه ریاضیات مشتمل بر مقوله‌های وجود^۱ (هستی‌شناسی)، معرفت^۲ (شناخت‌شناسی) و روش‌شناسی^۳ است. مقوله «وجود» در فلسفه ریاضی، به بررسی اشیاء (مفاهیم) ریاضی و نحوه وجود آنها اختصاص دارد. در مقوله‌های معرفت و روش‌شناسی، مسائلی از قبیل ماهیت اثبات ریاضی، منشأ اعتقاد ما به قضیه‌های ریاضی و فرآیند رشد نظریه‌های ریاضی مورد گفتگو قرار می‌گیرد. البته این مقوله‌ها از یکدیگر تفکیک شده نیستند و یک مکتب فلسفی ریاضی باید موضع منسجمی در برابر سؤالاتی که از این مقوله‌ها بر می‌خیزند، اتخاذ کند. بد نیست یکی دو نمونه از سؤال‌هایی را همیشه در فلسفه ریاضی مطرح بوده است و در این سخنرانی نیز مرتباً به آنها اشاره خواهد شد، بیان کنم. یکی از اساسی‌ترین مسائل فلسفه ریاضی، تشریح منشأ یقین^۴ در ریاضیات است. همه ما می‌دانیم که گزاره‌های ریاضی از درجه یقینی بالاتر از درجه یقین احکام سایر علوم برخوردارند. در واقع بسیاری از اوقات، از ریاضیات به‌عنوان تنها رشته معرفت بشری که در آن، «حقیقت مطلق» معنا دارد، یاد می‌شود. منشأ این درجه یقین چیست؟ آیا احکام ریاضی به درستی بیانگر حقیقت مطلق هستند و یا اینکه در این احکام نیز نوعی نسبیت وجود دارد؟ آیا درستی احکام ریاضی صرفاً وابسته به ماهیت صوری آنها است و یا اینکه این احکام دارای محتوای خبری هستند؟ مسئله دیگری که از قدیم‌الایام، هم در میان فلاسفه و هم در میان ریاضیدانان جنجال‌برانگیز بوده است، مسئله دوگانگی «پیوسته» و «گسسته» است. این سؤال که در ریاضیات، کمیت‌ها یا مفاهیم متصل، تقدم دارند و یا کمیت‌ها یا مفاهیم منفصل، احتمالاً قدمتی همدیاف با قدمت خود فلسفه دارد. خواهیم دید که جواب‌های رایج به این سؤال، همواره ارتباط نزدیکی با گرایش‌های ریاضی زمانه و نیز دیدگاه‌های متداول فلسفه ریاضی داشته‌اند.

در سخنرانی‌های گذشته و به‌ویژه بحث‌های متعاقب آنها، پدیده‌ای به چشم می‌خورد که لازم است آن را نیز به‌عنوان یک نکته مقدماتی در میان بگذارم؛ چه به گمان من، این مطلب رابطه‌ای اساسی با دو واقعیت عینی زیر دارد: یکی اینکه اکثر ریاضیدانان از بحث‌های علنی درباره فلسفه ریاضی گریزانند؛ اگرچه کمتر ریاضیدانی هست که به‌طور خصوصی تمایل و سلیقه ویژه در این زمینه نداشته باشد. دیگر اینکه موضوع فلسفه ریاضی در دو - سه دهه اخیر تحول عمده‌ای به خود ندیده و در مقایسه با بیشتر مقولات فلسفی، رشته‌ای عقیم و عقب‌مانده است. مطلبی که در ذهن دارم این است که علی‌رغم عنوان سمینار

^۱ontology ^۲epistemology ^۳methodology ^۴certainty

فلسفه ریاضی، در بحث‌های جلسات گذشته، صحبت‌های بسیاری از مکانیک کوانتومی و نظریه نسبیت (که از نظریه‌های فیزیک نوین هستند) به میان آمد در حالی که عملاً اشاره‌ای به نظریه‌های ریاضی معاصر نشد؛ البته سواى آنچه دربارهٔ نظریه مجموعه‌ها و منطق ریاضی (به اصطلاح، «مبانی ریاضی») بیان گردید. بی‌شک عده‌ای ایراد مرا نابجا دانسته، دلایلی از این قبیل در رد آن اظهار خواهند کرد:

۱- تمام ریاضیات امروز (به استثنای بعضی گوشه‌های نظریه رسته‌ها) بر پایهٔ نظریه مجموعه‌ها بنا شده است. بنابراین کافی است توجه فلسفه ریاضی را به بررسی فلسفی نظریه مجموعه‌ها و منطق ریاضی (که در ارتباط نزدیک با نظریه مجموعه‌ها است) معطوف کنیم.

۲- بیشتر تحقیقات ریاضی امروز، در جهت رشد درونی ریاضی است و در نتیجه، اثری بر مقولات فلسفی ندارد.

۳- نظریه‌های مکانیک کوانتومی و نسبیت، تأثیری ژرف بر بینش فلسفی انسان از جهان نهاده‌اند؛ در حالی که تنها اکتشاف ریاضی معاصر که بتوان نقشی مشابه برایش قائل شد، قضیهٔ گودل و شاید ظهور بی‌نهایت‌های بالفعل در نظریه مجموعه‌ها باشد که این‌ها نیز به مبحث نظریه مجموعه‌ها تعلق دارند.

در مورد ادعای اول، باید بگویم که حتی اگر بنا ساختن ریاضیات بر پایهٔ نظریه مجموعه‌ها را بپذیریم، تأثیر این نظریه بر بیشتر پژوهش‌های ریاضی، امری است کاملاً سطحی و از نقش یک زبان تجاوز نمی‌کند. مثلاً حتی نمی‌توان نقش نظریه مجموعه‌ها را در ریاضیات با نقش ریاضیات در فیزیک، مقایسه کرد، زیرا از آغاز فیزیک به عنوان یک علم در قرن هفدهم میلادی تا به امروز، نهادهای ریاضی نقشی عمیق و حیاتی در بیان قوانین فیزیک داشته‌اند؛ در حالی که بیشتر مطالب ریاضیات امروز، دنبالهٔ طبیعی اکتشافات پیش از ظهور نظریه مجموعه‌ها است. کمتر کشف مهم ریاضی در قرن بیستم است که نتوان بدون جدا کردن محتوای اصلی، به زبان ریمان و پوانکاره ترجمه کرد. غور در مبانی نظریه مجموعه‌ها و منطق ریاضی حداکثر همان قدر به فلسفه و شناخت ریاضی کمک می‌کند که مطالعهٔ دستور زبان برای بررسی ادبیات. در واقع سعی خواهیم کرد ثابت کنم که نقش نظریه مجموعه‌ها از این نیز ناچیزتر است و اتکای متداول امروزی به آن، خود زائیدهٔ دیدگاه‌های خاص در فلسفه ریاضی است.

در مورد ادعای دوم، مرز مشخصی میان مسائل درونی و مسائلی که می‌توانند واکنش فلسفی ایجاد کنند وجود ندارد، زیرا سیر پژوهش‌های واقعاً بدیع و اثرات احتمالی آنها، قابل پیش‌بینی نیست. نباید فراموش کرد که نظریه مجموعه‌های کانتور، تا حدود زیادی، زادهٔ پژوهش‌های او در بعضی مسائل کاملاً درونی مربوط به سری‌های مثلثاتی بود. در هر صورت، اگر فلسفه یک علم، تحولات آن علم را نادیده بگیرد، به سرعت تبدیل به مجموعه‌ای بی‌ارتباط با موضوع اصلی خود می‌شود. مدعی هستم که این دقیقاً وضعیتی است که در فلسفه ریاضی پدید آمده است. در این سخنرانی، سعی خواهم کرد نشان دهم که سه

مکتب غالب فلسفه ریاضی، یعنی منطق‌گرایی^۱، شهودگرایی^۲ و صورتگرایی^۳، هر یک با روی گرداندن از تجربه عینی ریاضی‌کاران، به مقوله‌ای مستقل از ریاضیات تبدیل شده است که جوابگوی کنجکاوی‌های فلسفی ریاضیدانان حرفه‌ای نیست.

سرانجام، دربارهٔ نکتهٔ سوم، شکی نیست که نظریه‌های فیزیک نوین، دگرگونی‌های عمیقی در دیدگاه‌های فلسفی ایجاد کرده‌اند، ولی کافی است یادآوری کنم صحبت اصلی ما در اینجا فلسفه ریاضی است نه فلسفه به مفهوم عام یا بررسی عواملی که بر دیدگاه‌های فلسفی اثر گذاشته‌اند.

۲. دوران باستان

برای بررسی تاریخ فلسفه ریاضی، به نظر من باید سه دوره را در نظر گرفت. دورهٔ اول، از ادوار باستانی تا اواخر قرن هفدهم میلادی؛ دورهٔ دوم که با ابداع حساب دیفرانسیل و انتگرال توسط نیوتن و لایب‌نیتس آغاز می‌شود و حدوداً دو قرن به درازا می‌کشد؛ و دورهٔ سوم، از زمان دقیق‌سازی مبانی آنالیز ریاضی توسط وایرستراس و دیگران تا به امروز.

در دورهٔ اول، مبحث هستی‌شناسی در ریاضی، از بررسی «وجود» در سایر علوم جدا نیست. ریاضیات علم کمیت‌ها و شکل‌ها است و هدف آن بررسی این گونه خواص اشیاء تجربی است. پایهٔ تجربی ریاضی در این دوره، به‌روشنی مشهود است. هندسه از اندازه‌گیری شروع می‌شود و حساب از تدوین جدول‌های نجومی و دفترداری امور تجاری و اداری سرچشمه می‌گیرد. کشف عدد ثابت پی (یعنی اینکه نسبت محیط دایره به قطر آن، عددی مستقل از اندازهٔ دایره است)، یکی از نخستین کشف‌های علم تجربی است. یونانیان باستان، برای آنچه ما امروز قضیه می‌نامیم دو لغت متمایز Theorema و Porisma را به‌کار می‌بردند. کلمهٔ Theorema به معنی بینش یا شهود، به آن دسته از قضیه‌های ریاضی اطلاق می‌شود که صحتشان قبل از اثبات دقیق ریاضی، تجربه شده است و نوعی واقعیت حائز اهمیت یا مستقل از ذهن را بیان می‌کنند. در حالی که Porisma به معنای «بادآورده»، واقعیتی است که نخست از طریق استنتاج ریاضی ظاهر می‌شود. این تمایز نشان می‌دهد که جنبهٔ صوری ریاضی در نظر یونانیان، نقشی فرعی دارد و محتوای عینی نتایج ریاضی است که در درجهٔ اول، مورد نظر آنها است. باید توجه داشت که حتی توسل به مُثُل افلاطونی در این اعتقاد که قضیه‌های ریاضی علی‌الاصول خواص اشیاء طبیعی را بیان می‌کنند، خللی وارد نمی‌کند. به‌عکس، از آنجا که در دیدگاه افلاطونی، جهان ایده‌ها از عینیت پایداری برخوردار است و مفاهیم ریاضی، والاترین وسیلهٔ ارتباط انسان با آن جهان می‌باشند، اصولاً عینیت‌های ریاضی مورد سؤال نیست. نیز روش اصل موضوعی اقلیدس در این دوره، اثری بر دیدگاه هستی‌شناسی ریاضی ندارد،

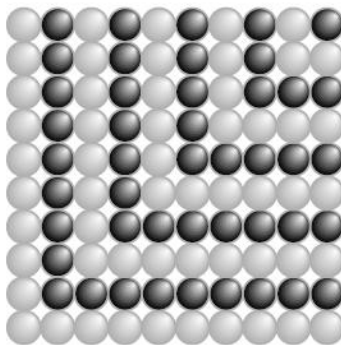
^۱logicism ^۲intuitionism ^۳formalism

زیرا اصول اقلیدس، از نظر یونانیان بیانگر ابتدایی‌ترین خواص هندسی فضای واقعی محسوب می‌شود و صحت آنها به تجربه، مورد پذیرش عموم است.

در مبحث معرفت و روش‌شناسی در این دوره، آنچه ریاضیات را از سایر علوم متمایز می‌سازد، درجه یقین ریاضی است و یقین در ریاضی، مبتنی بر «برهان» است و برهان نیز وسیله‌ای برای «دیدن» یا «اکتشاف» یک واقعیت است با کنار هم قرار دادن روابطی که صدقشان مورد تردید نیست. مثلاً برای اثبات اینکه مجموع هر چند عدد فرد متوالی که از یک آغاز شوند، مربع کامل است، یعنی

$$1 + 3 + 5 + \dots + (2n + 1) = k^2,$$

فیثاغورسیان شکل زیر را رسم می‌کنند:



حتی در عصر دقت و وسواس کنونی، کمتر کسی با اندک ذوق ریاضی است که با چند لحظه نگاه کردن به این شکل، به حکم مورد نظر یقین نیاورد. البته این حکم را می‌توان به روش استقرا، از اصول پئانو^۱ ثابت نمود. اینکه کدام اثبات ارجح است، به دیدگاه فلسفی ما بستگی دارد. به گمان من، اکثر ریاضی‌کاران، اثبات فیثاغورسی را روشن‌کننده‌تر و زیباتر می‌یابند. البته فراموش نکنیم که روش استقرا، یک وسیله کشف واقعیت نیست، بلکه حربه‌ای برای تأیید حدس‌هایی است که به روش‌های دیگر به آنها رسیده‌ایم.

در اینجا لازم است چند کلمه‌ای هم درباره جدال «پیوسته-گسسته» بیان کنم. شواهد تاریخی روشنی هست که حساب و هندسه پیدایش مستقل از یکدیگر دارند [۲۵]. حساب که با اعداد طبیعی سروکار دارد، تجلی ریاضی مفاهیم گسسته است. در حالی که موضوع هندسه بر پایه پیوستار^۲ صفحه و فضا می‌باشد. اولین برخورد این دو علم، در اندازه‌گیری طول‌های هندسی است. در اینجا حساب در هندسه کاربرد می‌یابد؛ به‌ویژه با توسعه مفهوم عدد به اعداد کسری، به نظر می‌رسد که بتوان همه طول‌های هندسی را با علائم حساب نمایش داد. کشف اولین عدد ناگویا (طول قطر مربع واحد) توسط فیثاغورسیان، بحرانی

^۱G. Peano ^۲continuum

در ریاضی پدید آورد که نتیجه آن، تفوق هندسه بر حساب است و توسعه مفهوم عدد به نحوی که بتوان به هر طول، عددی نسبت داد. تا زمان کشف (یا اختراع) حساب دیفرانسیل و انتگرال، هندسه اقلیدسی، مظهر یقین علمی و محک سنجش یقین برای سایر مقولات است. حتی ابداع دستگاه مختصات و هندسه تحلیلی توسط دکارت و فرما، که نتیجه آن جبری سازی هندسه است، از تقدّم فلسفی هندسه نمی‌کاهد، زیرا علم جبر (دنباله طبیعی حساب)، خود بر پایه پیوستار اعداد حقیقی، یعنی خط اقلیدسی بنا می‌شود.

۳. دوره دوم

دو قرنی که دوره دوم ریاضیات خوانده‌ایم، از یک سو دوران فتوحات درخشان و شگرف ریاضی است و از سوی دیگر، دوران بحران فلسفی در مبانی ریاضیات است. آنالیز نیوتن و لایبنیتس متکی بر مفهوم بی‌نهایت کوچک^۱ است که جای مناسبی در پیوستار اقلیدسی ندارد و فرض وجود آن خالی از تناقض و معما نیست. بینش اتمی لایبنیتس با فرض تقدّم پیوستار، تضاد دارد. لایبنیتس که به خوبی از نادقیق بودن مبانی حساب دیفرانسیل و انتگرال آگاه بود، هرگز موفق نشد به نظری قاطع و نهایی درباره ماهیت و درجه یقین این علم دست یابد. او در یک نوشته اظهار می‌کند که صحت نتایج حساب دیفرانسیل و انتگرال، تقریبی است ولی درجه خطا ماهیتاً چنان کوچک است که قابل اندازه‌گیری نیست. در جایی دیگر، اذعان می‌کند که تنها دلیل پذیرش بی‌نهایت کوچک‌ها، کارایی آنها است و اینکه در عمل به نتایج درستی منجر می‌شوند [۲۳]. علی‌رغم این معضلات فلسفی، حساب دیفرانسیل و انتگرال حربه چنان نیرومندی در اکتشافات ریاضی است که مشکلات مربوط به مبانی آن، همگی به حاشیه رانده می‌شوند. وضعیت مشابهی همزمان در علم فیزیک پدید می‌آید. قدرت مکانیک نیوتنی در پیش‌بینی پدیده‌های نجومی چنان است که تردیدهای شروع فلسفی نسبت به آن، حل‌نشده به خاموشی سپرده می‌شوند. ریاضیدانان این دوره، به خوبی از تزلزل و عدم یقینی که در مبانی آنالیز وجود دارد آگاه هستند و از این رو، بحث و مجادله درباره صحت قضیه‌های ریاضی، بسیار رایج است. با وجود این، تا مدت مدیدی، همت موفقیت‌آمیزی برای دقیق ساختن تعریف‌ها گمارده نمی‌شود. دوم اینکه علی‌رغم روشن نبودن مفاهیم اولیه، اشتباه فاحشی در ریاضیات رخ نمی‌دهد که موجب آشفتگی یا رکود تحقیق در ریاضی شود. در اینجا ممکن است ایراد گرفته شود که ادعای کوشی مبنی بر اینکه «حد یک دنباله از تابع‌های پیوسته، همواره پیوسته است»، نمونه‌ای از اشتباه‌های فاحش در این دوره است (مثال‌های دیگر از این نوع، فراوان است). البته چنین عبارتی در چارچوب دیدگاه‌های فلسفی ریاضی امروزی، یک اشتباه نابخشودنی محسوب می‌شود، ولی باید از دیدگاه دیگری به مسئله نگاه کرد. تعبیر دقیق کوشی از این ادعا هنوز روشن نیست و بحث‌های فراوانی در این زمینه شده است [۱۳].

^۱ infinitesimal به تعبیر لایبنیتس و fluxion به تعبیر نیوتن.

آنچه مسلم است این است که کوشی و ریاضیدانان زمان او، از مثال‌های ناقض این ادعا (به مفهوم امروزی تعاریف)، مانند سری فوریه‌ای که به یک تابع ناپیوسته میل می‌کند، کاملاً آگاه بوده‌اند. مثلاً آبل در جایی می‌نویسد که «قضیه کوشی استثنائاتی دارد» ولی ادعای کوشی را به نحوی، اساساً درست می‌داند. اگر امروزه این طرز بیان در ریاضی خیلی غریب به نظر می‌رسد به این دلیل است که دیدگاه هستی‌شناسی ریاضی در یک قرن اخیر، دگرگونی عظیمی به خود دیده است. در نظر ما، «تابع پیوسته» و «همگرایی»، اشیاء اصیل و اساسی ریاضی هستند که نهایت دقت در به‌کار بستن آنها رعایت می‌شود. در حالی که حتی در اوایل قرن نوزدهم میلادی، دیدگاه در مورد مقوله وجود، اساساً همان دیدگاه دوره اول است. اشیاء ریاضی یا عینیت فیزیکی دارند و یا مفهوم‌های بسیار مأنوس ریاضی هستند که در اثر تجربه متمادی، عملاً عینیت یافته‌اند (مثلاً عددهای صحیح، شکل‌های هندسی و معادلات جبری). «اشتباه» در ریاضی وقتی رخ می‌دهد که نتیجه‌ای درباره این مفاهیم ملموس به‌دست آید که نادرستی آن قابل تحقیق باشد. در زمینه معرفت نیز اثبات ریاضی هنوز یک حربه اکتشاف است. قضیه‌ها و نظریه‌های کلی ریاضی، وسیله‌هایی برای روشن ساختن پدیده‌های خاص هستند. به قول کانت: «اثبات، یک طریق رسیدن به حقیقت است مانند دیدن و شنیدن.» چون کشف واقعیت‌های جالب، هدف نهایی ریاضیدانان است، برخی مانند اوایلر از هیچ وسیله‌ای برای دست یافتن به واقعیت‌های ریاضی فروگذار نمی‌کنند. مثالی را که جورج پولیا گفته است [۱۷] برای کسانی که ندیده‌اند، نقل می‌کنم. اوایلر می‌خواهد مجموع سری $\sum_{n=1}^{\infty} \frac{1}{n^2}$ را محاسبه کند. او می‌داند که اگر ریشه‌های یک چندجمله‌ای، عددهای $\alpha_1, \alpha_2, \dots, \alpha_n$ باشند، آن چندجمله‌ای به‌صورت $c(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$ است که در آن، c عددی ثابت است. تابع $\sin x$ یک چندجمله‌ای نیست و معادله $\sin x = 0$ بی‌نهایت جواب دارد که عبارت‌اند از:

$$0, \pm\pi, \pm2\pi, \dots$$

با وجود این، اوایلر به آزمایشی ماجراجویانه دست می‌زند. او حاصل ضرب بی‌نهایت عامل خطی را که از این ریشه‌ها به‌دست می‌آیند، در نظر می‌گیرد:

$$c(x^2 - \pi^2)(x^2 - 4\pi^2) \cdots$$

آیا این حاصل ضرب معنا دارد؟ آیا می‌توان با انتخاب مقدار مناسبی برای c ، این حاصل ضرب را مساوی با $\sin x$ دانست؟ البته اوایلر کاملاً با سری تیلور آشنا است. او می‌داند که

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

ضریب x در این عبارت برابر با ۱ است و در عبارت قبلی برابر با

$$c(-\pi^2)(-4\pi^2) \cdots$$

به همین دلیل، اویلر c را به‌طور صوری برابر با

$$\left(\frac{-1}{\pi^2}\right)\left(\frac{-1}{4\pi^2}\right)\dots$$

می‌گیرد و با توزیع عوامل آن در حاصل ضرب فوق، به فرمول زیر دست می‌یابد:

$$\sin x = x \left(1 - \frac{x^2}{\pi^2}\right) \left(1 - \frac{x^2}{4\pi^2}\right) \dots$$

البته هنوز چیزی «ثابت نشده است» ولی اویلر می‌داند که ضریب جمله درجه اول، درست است. جمله ناصفر بعدی، جمله درجه سوم است. اگر این فرمول درست باشد، مقایسه ضریب جمله درجه سوم آن با ضریب جمله درجه سه در بسط تیلور تابع $\sin x$ ، نتیجه می‌دهد

$$-\frac{1}{6} = -\frac{1}{\pi^2} - \frac{1}{4\pi^2} - \frac{1}{9\pi^2} + \dots$$

بنابراین

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots = \frac{\pi^2}{6}.$$

اوایلر آگاه است که «اثبات» او یقین اثبات‌های اقلیدسی را ندارد، پس بی‌درنگ به تحقیق درباره این تساوی می‌پردازد و با محاسبه دو طرف تا تعداد زیادی رقم بعد از اعشار، یقین حاصل می‌کند که به واقعیت دست یافته است!

امروزه فیزیکدانان همچنان به چنین ماجراجویی‌هایی می‌پردازند و ریاضیدانان هم در خفای دفتر کارشان، به این گونه آمایش‌ها دست می‌زنند ولی افشای چنین اعمالی در جامعه ریاضی گناه کبیره محسوب می‌شود. یادآوری می‌کنم که مقصود من از آوردن این مثال‌ها، تأکید بر این مطلب است که در دوره دوم تاریخ ریاضی نیز، مانند دوره اول، اشیاء ریاضی عینیتی مشابه سایر علوم دارند و اثبات ریاضی تنها، حربه‌ای است برای دست یافتن به روابط این اشیاء یا استحکام بخشیدن به روابطی که صحت آنها به طریق تجربی مشاهده شده است. حتی وایرشتراس که بعداً به عنوان نقطه عطف دیدگاه فلسفی ریاضی از او یاد خواهد شد، در نامه‌ای به یک دوست ریاضیدان خود می‌نویسد: «اینکه برای یک ریاضیدان به عنوان مکتشف، استفاده از هر روشی مشروع است، امری است بدیهی...»^۱ تفاوت عمده فلسفه‌های ریاضی در دوره‌های اول و دوم در این است که ریاضیات دوره دوم (عمدتاً حساب دیفرانسیل و انتگرال)، بر پایه متزلزل «بی‌نهایت کوچک‌ها» بنا شده است و یقین ریاضی که قرن‌ها در هندسه اقلیدسی والاترین تجلی خود را یافته بود، دستخوش بحرانی نگران‌کننده می‌شود تا جایی که بزرگترین ریاضیدانان این دوره، به وجود «استثنا» برای قضیه‌های خود، تن در می‌دهند.

^۱ نقل از لاکاتش [۱۳].

پدیده دیگری از دوره دوم که شایسته است به آن اشاره شود، پیدایش هندسه‌های نااقلیدسی است؛ گرچه به نظر من درباره اثر واقعی آن بر فلسفه ریاضی دوره دوم، غلو شده است. لباچفسکی و بویویی^۱ مستقل از یکدیگر درصدد بودند که از طریق برهان خلف، اصل توازی اقلیدس را از سایر اصول هندسه اقلیدسی نتیجه بگیرند. تلاش‌های این دو، مانند کوشش‌های متقدمانشان، به نتیجه‌ای نمی‌رسد ولی هر یک با نقض اصل پنجم، به مجموعه‌ای چشمگیر از قضایای یک نوع هندسه نااقلیدسی دست می‌یابد که در آن، تناقضی به چشم نمی‌خورد. هر دو عملاً قانع می‌شوند که نمی‌توان اصل توازی را از اصل‌های بدیهی‌تر اقلیدس نتیجه گرفت و انکار آن هم تناقضی ایجاد نمی‌کند. بر خلاف آنچه برخی تاریخ‌نویسان دوره بعد تاریخ ریاضی ادعا می‌کنند، عدم توجه مراکز عمده ریاضی به «کشفیات» لباچفسکی و بویویی به سبب عدم شهرت این دو یا عدم تعلق آنها به این مراکز نیست. باید به دیدگاه فلسفی در آن زمان توجه کرد. در اوایل قرن نوزدهم، احتمالاً همه ریاضیدانان بزرگ، به نحوی به استقلال اصل توازی یقین یافته‌اند و نتایج لباچفسکی و بویویی، تنها شواهدی بیشتر در این زمینه هستند. هیچ‌یک از این دو ریاضیدان، مدلی از یک فضای نااقلیدسی ارائه نمی‌کند.^۲ اصولاً آنچه در آن زمان مطرح است، سازگاری منطقی نظریه‌های ریاضی نیست، بلکه قدرت این نظریه‌ها در بیان روابط هندسی قابل ادراک است. کانت که مدعی است اثبات‌های ریاضی را نمی‌توان به‌طور کامل به استنتاج‌های منطقی محض تبدیل کرد، به شهودهای پیشینی^۳ از فضا و زمان معتقد است و اصول اقلیدس را از جمله شهودهای پیشینی می‌داند.

با این زمینه فلسفی و آنچه قبلاً درباره دیدگاه‌های فلسفه ریاضی در دوره دوم گفتیم، روشن است که هندسه نااقلیدسی تنها در صورتی در انظار ریاضیدانان این دوره دارای اهمیت است که نااقلیدسی بودن فضای قابل ادراک به‌طور جدی مطرح شود و چنین سؤالی در آن دوره مطرح نبود. باید این نکته را هم افزود که در بررسی هندسه‌های نااقلیدسی، گاوس^۴ به مراتب از لباچفسکی و بویویی جلوتر است. گاوس که با روش‌های تحلیلی، هندسه دیفرانسیل را پایه‌گذاری می‌کند، موفق می‌شود مدل‌هایی دوبله از هندسه‌های نااقلیدسی ارائه کند که به‌صورت رویه‌هایی در فضای سه بُعدی اقلیدسی ظاهر می‌شوند. به زبان امروزی، گاوس ثابت می‌کند که اگر هندسه اقلیدسی سازگار باشد، نظریه‌های نااقلیدسی نیز سازگارند. ولی جو فلسفه ریاضی در آن زمان برای افشای این نتایج مساعد نبود و گاوس هرگز مبادرت به نشر این قضیه‌ها نکرد.

^۲ مدلی که گاهی به نام «هندسه لباچفسکی» از آن یاد می‌شود (مثلاً در کتاب هندسه لباچفسکی ترجمه احمد بیرشک، انتشارات دانشگاه صنعتی شریف، ۱۳۵۹)، در واقع ساخته پوانکاره است.

^۴ گاوس خود هرگز مطلب را به این صورت بیان نمی‌کند. بعدها بلترامی نشان داد که رویه‌هایی که خمیدگی ثابت منفی دارند، مدل‌هایی برای هندسه لباچفسکی و بویویی هستند.

اکنون می‌پردازیم به کشمکش میان «گسسته و پیوسته» در دوره دوم. پیوستار اعداد حقیقی و توسعه آن به پیوستار اعداد مختلط که در واقع زمینه نیروی عظیمی است که آنالیز به ریاضیات این دوره بخشیده است، خود ماهیتی معماگونه پیدا می‌کند که سرچشمه اصلی بحران فلسفه ریاضی در این دوره است. اعتقاد به شهود پیشینی از پیوستار اعداد حقیقی (خط راست) با وجود بی‌نهایت کوچک‌ها، تناقض مستقیم دارد. اگر نظریه اتمی لایبنیتس درست باشد، وجود پیوستار، قهراً نقض می‌شود (مقصود از «وجود»، به تعبیر این دوره است). ریاضیدانان دوره دوم به توافقی در این زمینه نمی‌رسند. از یک سو، آنها ناگزیر به استفاده از بی‌نهایت کوچک‌ها هستند و از سوی دیگر، موفق نمی‌شوند نظریه‌ای مدون و منطقی در توجیه این مخلوقات ارائه دهند. در برخی تعبیرهای خط حقیقی (در آنالیز) حضور بی‌نهایت کوچک‌ها واجب به نظر می‌رسد و در بعضی تعبیرهای دیگر (هندسه محض)، بی‌نهایت کوچک‌ها از صحنه محو می‌شوند. خط اقلیدسی، مستحکم‌ترین سنگر یقین در علم، به معماگون‌ترین مفهوم‌ها تبدیل می‌شود.

۴. دوره سوم

برای رفع بحران بنیادی ریاضی، لازم است که ریاضیات بر پایه‌ای مستحکم بازسازی شود و مانند گذشته، یقین، به ریاضیات بازگردد. اگر پیوستار اقلیدسی به عنوان شهود پیشینی قابل اعتماد نیست، باید ریاضیات و به ویژه آنالیز را بر مبنای دیگری بنا نهاد. نامزد اصلی شهود پیشینی خالی از ابهام، مجموعه اعداد طبیعی است. کرونکر اعلام می‌کند که «خداوند اعداد طبیعی را آفرید و بقیه کار انسان است» و ددکیند برای اولین بار پیشنهاد دقیقی برای تعریف پیوستار یک بُعدی اقلیدسی ارائه می‌کند. این کوشش‌ها وقتی باثمر خواهند بود که بتوان آنالیز را بر مبنایی مستحکم بنا کرد. وایرستراس وارد صحنه می‌شود. او از یک سو مانند ددکیند، تعریفی دقیق از اعداد حقیقی پیشنهاد می‌کند (به زبان امروزی، اعداد گویا رده‌های هم‌ارزی زوج‌های مرتب اعداد صحیح هستند و اعداد حقیقی رده‌های هم‌ارزی دنباله‌های کوشی از اعداد گویا) و از سوی دیگر، با دقیق ساختن مفهوم حد (به صورت کنونی)، بی‌نهایت کوچک‌ها را از آنالیز می‌راند. به این ترتیب ظاهراً وایرستراس موفق می‌شود که با حسابی ساختن^۱ آنالیز، همه قضیه‌های آنالیز حقیقی و مختلط را بر پایه‌ای استوار سازد که بار دیگر، یقین در ریاضیات احیا شود. این نقطه عطفی است که آن را آغاز دوره سوم فلسفه ریاضیات می‌دانیم.

با اینکه نظریه اتمی لایبنیتس و بی‌نهایت کوچک‌های او جایی در نظام وایرستراس ندارند، سازمانی که وایرستراس برای پیوستار حقیقی قائل می‌شود، خود دربرگیرنده نوعی اتم‌گرایی است. به این صورت که شهود پیوستگی خط حقیقی، تقدم خود را از دست می‌دهد و سازمان خاص اجزای خط حقیقی است که منشأ ادراک پیوستگی محسوب می‌شود. این که حتی «سازمان» و «ترتیب» هم خود تشریح اتمی پیدا

^۱arithmetization

می‌کنند، ثمره نظریه مجموعه‌های کانتور است که بعداً به آن خواهیم پرداخت. لازم است در اینجا بیفزاییم که تعریف‌های دقیق دوره و ایرشتراس، خود منشأ ایجاب مشروعیت برای مخلوقات ریاضی غیرشهودی و حتی ضدشهودی می‌شوند. از این قبیل است «تابع‌های همه‌جا پیوسته و هیچ‌جا مشتق‌پذیر»، «منحنی‌های فضا پُرکن» و انواع گوناگون تابع‌هایی که در عملکرد روزمره ریاضی جایی ندارند. اگرچه این «عفریت‌آفرینی» نقشی مثبت در دقیق ساختن اثبات‌های ریاضی دارد و به یقین متزلزل شده ریاضی حیاتی نو می‌بخشد، پذیرش این راه‌حل، به قیمت از دست رفتن حیثیت «وجود» در ریاضیات تمام می‌شود. دیگر نه تنها هندسه نااقلیدسی، قبح نیست، موجودات عجیب و غریبی مانند «منحنی فضا پُرکن» در کنار مفاهیم ریشه دارد و زمان آزموده ریاضی قد علم می‌کنند و ادعای وجود دارند. این وضعیت با پیدایش نظریه مجموعه‌های کانتور و برنامه‌های بعدی فرگه، راسل و هیلبرت به افراطی‌گری‌های خطرناکی کشیده می‌شود ولی در زمان و ایرشتراس، معدود عفریت‌های موجود، زنگ خطری به صدا در نیاوردند و تنها مایه کنجکاوی ریاضیدانان شدند.

کارهای و ایرشتراس، بولتزانو، ددکیند و معاصرانشان در مبانی آنالیز، روشن می‌سازد که مفهوم «مجموعه‌ای از اعداد حقیقی» کارایی فوق‌العاده‌ای در روشن ساختن سازمان اعداد حقیقی دارد و قوام و دقت تازه‌ای به قضیه‌های آنالیز می‌بخشد. تعمیم مفهوم انتگرال توسط ریمان و موفقیت‌هایی که در بررسی سری‌های فوریه نصیب او می‌شود، مؤید این دیدگاه هستند. کانتور که مجموعه نقاط همگرایی سری‌های فوریه را بررسی می‌کند و در درک ساختمان مجموعه‌های اعداد حقیقی از دیگران پیشی گرفته است، دو گام به جلو برمی‌دارد: یکی در نظر گرفتن مجموعه‌ها به عنوان مفهوم عام و دیگری، آفریدن اعداد اصلی و ترتیبی نامتناهی به عنوان یکی از ویژگی‌های مجموعه‌ها. به این ترتیب، «اتم‌گرایی» و ایرشتراس توسط کانتور به بیشتر بخش‌های ریاضیات سرایت می‌کند. بناکردن ریاضیات بر پایه نظریه مجموعه‌ها را نوعی «اتم‌گرایی» خوانده‌ایم به این سبب که شرط لازم و کافی برای تساوی دو مجموعه، یکی بودن عضوهای آنها است و لذا ماهیت یک مجموعه، از ترکیب اعضای آن، کاملاً مشخص می‌شود.

کانتور حتی «پیوستار» را برای همیشه به عنوان یک شهود پیشینی از میان برمی‌دارد. او پیوستار به معنای عام را یک مجموعه همبند و کامل^۱ تعریف می‌کند.^۲ مفهوم ترتیب نیز با در نظر گرفتن زیرمجموعه‌های حاصل ضرب دکارتی یک مجموعه در خودش، تعریفی در چارچوب نظریه مجموعه‌ها می‌یابد. سرانجام، تسرملو با اثبات قضیه خوش‌ترتیبی (معادل با اصل انتخاب)، استقلال کامل ترتیب از ترکیب یک مجموعه را اعلام می‌کند.

^۲ این تعریف کانتور برای زیرمجموعه‌های فضای اقلیدسی است ولی می‌توان آن را در هر فضای توپولوژیک به کار بست.

^۱ perfect

آنچه که بحران جدیدی در فلسفه ریاضی به وجود می‌آورد، یکی به کار بستن جسورانه بی‌نهایت بالفعل توسط کانتور است و دیگری، ظهور تناقض‌های گوناگون در مورد مجموعه‌های بینهایت بزرگ. ریاضیدانان که از دوران پیشین به ظهور تناقض‌های ناشی از به کار بردن بی‌نهایت بزرگ و بی‌نهایت کوچک، آشنا بوده و پس از سالیان دراز کوشش، موفق شده‌اند به کمک تعریف‌های وایرستراس، همه بی‌نهایت‌ها را به بینهایت‌های بالقوه کاهش دهند، این بار، مقاومت زیادی در برابر نظریه کانتور از خود نشان می‌دهند. آنچه واقعاً نگران‌کننده است، این است که مفهوم ظاهراً بی‌آزاری مانند «مجموعه» می‌تواند منشأ پارادکس‌هایی مانند پارادکس بورالی-فورتی^۱ و پارادکس راسل باشد. جهان ریاضی که مدت کوتاهی است از یقین باز یافته این علم، سرمست شده است، ناگهان دستخوش تزلزل تازه‌ای می‌شود. ریاضیدانان بزرگ آن زمان، نظریات گوناگونی در این زمینه دارند. بعضی مانند پوانکاره بر این عقیده‌اند که نظریه مجموعه‌ها «مرضی است که سرانجام ریاضیات از آن رهایی خواهد یافت» و برخی دیگر مانند هیلبرت معتقدند که «کسی هرگز ما را از بهشتی که کانتور برایمان آفریده است، بیرون نخواهد راند». در هر صورت، موقعیت نظریه مجموعه‌ها و بحران بی‌نهایت نمی‌تواند مورد توجه مکتب‌های فلسفه ریاضی که در همان زمان پایه‌گذاری می‌شدند، قرار نگیرد.

۵. مکتب‌های فلسفه ریاضی

با توجه به صحبت‌هایی که دربارهٔ جوّ ریاضی حاکم در قرن نوزدهم کردیم، زمینه پیدایش سه مکتب منطق‌گرایی، شهودگرایی و صورتگرایی باید روشن باشد. این سه مکتب، هر یک به نحوی خواستار حل دو مسئله «وجود» و «یقین» هستند و کوشش‌هایی که در راه حسابی‌ساختن ریاضیات توسط وایرستراس و دیگران انجام گرفته است، اثری عمیق بر دیدگاه آنها دارد. منطق‌گرایی که اعلام ایجاد آن را باید کتاب مبانی حساب^۲ فرگه به سال ۱۸۸۴ دانست، ریاضیات را جزئی از منطق و قضیه‌های ریاضی را گزاره‌هایی در منطق می‌داند. فرگه در کتاب خود، اشتقاق اعداد طبیعی از مفاهیم منطقی را توصیف می‌کند. پس از فرگه، راسل نیز در اوایل قرن بیستم، با انتشار کتاب اصول ریاضی^۳ [۲۰]، تر منطق‌گرایی و چگونگی بازسازی ریاضیات بر مبنای منطق را پیشنهاد کرده و در اثر عظیم خود با وایتهد [۲۱]، به پیاده کردن این برنامه همت می‌گمارد. کار راسل و وایتهد، نظریه مجموعه‌های کانتور را نیز دربر می‌گیرد و این دو با ابداع نظریه انواع^۴، بر تضادهای نظریه مجموعه‌ها فائق می‌آیند. به این ترتیب، راه‌حلی که منطق‌گرایان برای مسائل وجود و یقین پیشنهاد می‌کنند، این است که «وجود» در ریاضی، همان وجود در منطق است و منشأ «یقین» نیز یقین در منطق. بعداً که به توصیف صورتگرایی می‌پردازیم، خواهیم دید که این دیدگاه با مشکلات بزرگی روبه‌رو است و در واقع، نه تنها کلیه ایرادهایی که به صورتگرایی وارد است، به منطق‌گرایی

^۱Burali-Forti paradox ^۲Die Grundlagen der Arithmetik ^۳type theory

نیز منتقل می‌شود، بلکه چنانچه ریاضیات را جزئی از منطق بپنداریم، این مسائل در سطحی گسترده‌تر، همهٔ منطق را دچار تزلزل خواهند کرد. فرگه و راسل در اواخر کار، کوشش‌های خود را شکست خورده محسوب می‌کردند و علی‌رغم کمک بزرگی که پیدایش این مکتب به توسعهٔ منطق ریاضی کرد، امروز می‌توان گفت که منطق‌گرایی به‌عنوان فلسفهٔ ریاضی، طرفداران زیادی ندارد.

در دههٔ سوم قرن بیستم، هیلبرت، برنامهٔ فرگه و راسل را از دیدگاه واقع‌بینانه‌تری از سر می‌گیرد: دیدگاهی که به صورت‌گرایی معروف است. هیلبرت، نخست مسئلهٔ وجود را با انکار کامل آن، حل می‌کند به این شکل که از دیدگاه هیلبرت، اشیاء ریاضی، نمادهای وضع‌شده‌ای بیش نیستند که حدود کارکرد و ویژگی‌های آنها توسط اصول موضوعه مشخص می‌شود. روش رسیدن به گزاره‌های درست ریاضی، استنتاج متناهی‌وار^۱ است. این در واقع تعمیم کاری است که پئانو در مورد علم حساب شروع کرده بود. تفاوت اصلی آن با دستگاه اقلیدسی این است که تعبیر معنایی نمادهای هیلبرت منحصر به فرد نیست؛ در حالی که در دستگاه اقلیدسی، نمادها، نشان‌دهندهٔ اشیاء خارجی مشخصی هستند. البته روش استنتاج منطقی نیز در برنامهٔ هیلبرت، تعریف دقیق دارد و با پیشرفت‌هایی که در اثر رواج منطق‌گرایی نصیب منطق ریاضی شده است، اکنون می‌توان مسائل ظریف‌تری را در نظر قرار داد. شرایطی که هیلبرت برای قابل قبول بودن یک نظریهٔ ریاضی قائل است، این است که اصول موضوعه منجر به تناقض نشوند (اصل سازگاری) و گزاره‌های درست نظریه همه در درون آن نظریه، اثبات‌پذیر باشند (اصل تمامیت). بنابراین برنامهٔ هیلبرت، اصل موضوعی ساختن همهٔ نظریه‌های ریاضی و اثبات سازگاری و تمامیت آنها است با این هدف که برای همیشه، پارادکس‌هایی که استقرار یقین در ریاضیات را به مخاطره انداخته‌اند، از صحنه پاک شوند. اینکه آرزوی هیلبرت با کشف غیرمنتظرهٔ گودل دربارهٔ ناتمامیت حساب نقش بر آب می‌شود، داستان آشنایی است. گودل ثابت می‌کند که تمامیت و سازگاری اصول موضوعهٔ پئانو یا هر مجموعهٔ متناهی دیگری که شامل آن شود، نمی‌توانند توأماً برقرار باشند. به عبارت دیگر، اگر نظریهٔ اعداد از نظر منطقی سازگار باشد، همواره در آن گزاره‌های درست ولی غیر قابل اثبات (به مفهوم برنامهٔ هیلبرت، یعنی در چارچوب اصول و روش استنتاج متناهی) وجود خواهند داشت. در اینکه قضیهٔ گودل، ادعای صورت‌گرایی و منطق‌گرایی را به‌عنوان فلسفه‌های جامع علم ریاضی، برای همیشه مردود می‌سازد، شکی نیست. در آینده سعی خواهم کرد نشان دهم که این اشکال درونی دو مکتب نامبرده، ممکن است تنها یک اشکال «موضعی» باشد و در واقع ناتوانی‌های اساسی‌تری متوجه این دو مکتب است.

قبل از پرداختن به بررسی انتقادی، لازم است مختصری نیز از مکتب عمدهٔ دیگر این دوره، یعنی شهودگرایی بگوئیم. این دیدگاه که به ابتکار براوئر پایه‌گذاری می‌شود، تفاوت‌هایی اساسی با دو مکتب صورت‌گرایی و منطق‌گرایی دارد. مبنای فلسفهٔ ریاضی براوئر را می‌توان در دو اصل زیر خلاصه کرد:

۱- اشیاء ریاضی، ساختمان‌های ذهنی هستند؛

^۱ finitary deduction

۲- این اشیاء بر مبنای شهود پیشینی از اعداد طبیعی (به عنوان اشیاء اولیه) و به طریق منطق شهودی ساخته می شوند. منطق شهودی مانند منطق غیرصوری، عادی است با این تفاوت عمده که در آن، اصل طرد شق ثالث^۱ انکار می شود.

کنار گذاشتن اصل طرد شق ثالث، نقش اساسی در دیدگاه فلسفه شهودگرایی نسبت به «وجود» دارد. از آنجا که اشیاء ریاضی در نظر براوتر ساخته ذهنی انسان هستند، تنها به وجود آن اشیائی می توان حکم کرد که بشود آنها را در تعداد متناهی گام ساخت. استدلال های وجودی که مبتنی بر برهان خلف هستند، طریقه ای برای دست یافتن به شیء ریاضی مورد نظر ارائه نمی کنند، بلکه تنها وجود را به عنوان یک استلزام منطقی (مبتنی بر اصل طرد شق ثالث) حکم می نمایند. هر گونه اثبات وجود که روشی برای ساختن، دست کم به طور نظری، همراه نداشته باشد، از دیدگاه شهودگرایی مردود است. از همین قبیل است اصل انتخاب و همه مطالب نظریه مجموعه ها که به بی نهایت بالفعل مربوط می شوند. از این دیدگاه، تنها بی نهایت های بالقوه در ریاضیات معنا دارند و اعداد اصلی و ترتیبی بی نهایت بزرگ، در چارچوب وجودی این فلسفه نمی گنجند. نمونه ای از بی نهایت بالقوه، تعداد اعداد طبیعی است، بدین معنی که می توان هر عدد طبیعی را ساخت و فرآیند ساختن این اعداد، هیچ جا ختم نمی شود. بنابراین انکار اصل طرد شق ثالث، وسیله ای است که به کمک آن، براوتر استقلال اشیاء ریاضی از ذهن سازنده را سلب می کند.

به همین روال، یک گزاره ریاضی تنها در صورتی صادق است که بتوان صدق آن را با استدلال شهودگرایانه (بدون استفاده از اصل طرد شق ثالث) اثبات کرد و در صورتی کاذب است که بتوان به همین روش، از آن تناقضی نتیجه گرفت. یک گزاره ریاضی تا وقتی که صدق یا کذب آن به این طریق ثابت نشده باشد، بی معنی است. برای مثال، از دید شهودگرایان، گزاره «سیاوش پسر کیکاووس، خالی بر شانه دارد» بی معنی است، زیرا «سیاوش پسر کیکاووس» ساخته ذهن فردوسی (یا حماسه سرایان پیش از او) است و تنها آن احکامی درباره وی صادق یا کاذب اند که بتوان آنها را از اساطیر مربوط، استنتاج کرد یا رد نمود. چون ظاهراً هیچ جا در افسانه های باستانی ایران، اشاره ای به «خال روی شانه» سیاوش نشده است، این گزاره نه صادق است و نه کاذب، بلکه بی معنی است. به همین ترتیب، گزاره های درست و غیر قابل اثبات گودل از دیدگاه شهودگرایی بی معنی هستند.

با این دیدگاه براوتر، پس از او افراد دیگری مانند هیتینگ^۲ و بیشاپ^۳ به بازسازی ریاضیات می پردازند. مقدار قابل ملاحظه ای از ریاضیات سنتی (عملاً همه چیز درباره اعداد حقیقی و مختلط) را می توان به این طریق بنا کرد ولی آنچه نیاز به اصل انتخاب یا استفاده از بی نهایت های بالفعل دارد، از ریاضیات طرد می شود. در این کوشش، بخش بزرگی از منطق شهودگرایی به دست می آید که استدلال های منطقی قابل استفاده در این مکتب را توصیف می کنند. مقایسه روش ساختن اعداد حقیقی به طریق

^۱principle of excluded middle ^۲A. Heyting ^۳E. Bishop

شهودگرایانه و صوری-منطقی، هر ریاضیدان را قانع می‌سازد که در این زمینه، اشیاء مورد نظر عملاً یکی هستند و تفاوت، تنها در دو چیز است: یکی تعبیر فلسفی و دیگری، اینکه بعضی استدلال‌های شهودگرایانه، به دلیل محدود بودن زمینه وجود، طولانی‌تر از اثبات‌های سنتی هستند.

۶. نقد مکاتب سه‌گانه

نقد سه مکتب معروف فلسفه ریاضی را از شهودگرایی آغاز می‌کنیم. نخستین انتقادی که به شهودگرایی وارد است، تکیه آن بر اعداد طبیعی به عنوان شهود پیشینی است. در واقع شهودگرایی و در عصر حاضر، ساخت‌گرایی^۱ بیشاپ دنباله طبیعی کوشش‌های حسابی‌سازی ریاضیات هستند. در توجیه این گرایش، گفته می‌شود که از شهود پیشینی فضا و زمان که کانت به آن اعتقاد داشت، شهود فضا با پیدایش هندسه‌های نااقلیدسی، ثبات خود را از دست می‌دهد ولی شهود زمان، یا به‌طور دقیق‌تر، شهود توالی زمان، سرچشمه اعداد طبیعی است. در این زمینه می‌توان ایراد گرفت که همچنان که هندسه‌های نااقلیدسی، شهود پیشینی فضا را متزلزل ساخته است، نظریه نسبیت نیز مطلق بودن زمان را برای همیشه برانداخته است.^۲ در دیدگاه فیزیک امروزی از جهان طبیعی، زمان و فضا اساساً در هم آمیخته‌اند و قابل تفکیک نیستند. به علاوه، واقعاً بعید به نظر می‌رسد که از نظر تاریخی یا روانی، شهود زمان، منشأ اعداد طبیعی باشد. پژوهش‌های پیازه درباره درک مفاهیم ریاضی توسط کودکان، مؤید این است که کودکان، مفاهیم مجردی همچون اعداد طبیعی را در بازی کردن و تجربه با اشیاء و شرکت فعال در محیط، کسب می‌کنند نه از طریق انفعالی ادراک گذشت زمان (مثلاً [۱۰] را ببینید). البته در اینجا انتقاد ما به حسابی‌سازی ریاضیات نیست، بلکه به توجیه حسابی‌سازی بر اساس شهود زمان است. در زمینه استفاده از شهود فضا و زمان، به نظر ما رنه توم^۳ دیدگاه واقع‌بینانه‌تری دارد [۱۰، ۲۴]. به نظر توم، آنچه ما به عنوان شهود پیشینی از ادراک فضا و زمان کسب می‌کنیم، پیوستگی موضعی است. این فرض، با نظریه نسبیت و نااقلیدسی بودن فضا مغایرتی ندارد، بلکه تنها پیشنهادی است برای اتخاذ پیوستار موضعی به عنوان شهود پیشینی. جالب خواهد بود اگر بتوان یک دستگاه شهودی بر مبنای این شهود پیشینی، بنا نهاد. چنین مجموعه‌ای انعکاس دقیق‌تر دیدگاه کانت خواهد بود تا شهودگرایی براور که در واقع تحت تأثیر تز تقدم گسسته کرونکر است. حتی اگر زمانی، کوانتمی بودن فضا و زمان به مثابه یک اصل فیزیکی مورد قبول قرار گیرد، شکی نیست که ادراک ما از پیوستار، تقریب بسیار قابل اعتمادی از جهان طبیعی است که واقعیت عملی خود را در قلمرو ادراک و کارکرد انسان‌ها از دست نخواهد داد.

^۲ به نظر می‌رسد این ایراد را اول بار K. Popper در [۱۸] مطرح کرده است.

^۱constructivism ^۳R. Thom

شاید عمده‌ترین اشکال شهودگرایی این باشد که به سبب محدودیت‌هایی که در مقوله «وجود» قائل می‌شود، نمی‌توان آن را در جایگاه یک فلسفه جامع ریاضی پذیرفت. امروزه، بی‌نهایت بالفعل، جای خود را در ریاضیات باز کرده است و بی‌می واقعی درباره ناسازگار بودن آن احساس نمی‌شود. اصل طرد شق ثالث همچنان که از دوره یونان باستان مرسوم بوده است، هنوز مورد قبول و استفاده ریاضیدانان است و اشیاء ریاضی در دیدگاه ریاضی‌کاران، استقلال مشابه استقلال اشیاء فیزیکی دارند. رشد چشمگیر منطق ریاضی آن را به شاخه‌ای غنی در ریاضیات تبدیل کرده است و یکی از ثمرات آن ریاضیات ناستانده^۱ است. آنالیز ناستانده به ما نشان می‌دهد که بی‌نهایت کوچک‌های لایب‌نیتس را طوری می‌توان توجیه کرد که دقت مورد انتظار ریاضی را داشته باشند. شاید مهیج‌تر از این، دیدگاه‌های کاملاً نوینی است که حساب ناستانده در این رشته گشوده است. تجربه ریاضی نیم قرن اخیر، نه تنها مفاهیم ریاضی را آن‌طور که براوتر طالب بود، در بند نکشیده است، بلکه روش‌هایی که براوتر کاملاً غیر قابل قبول می‌دانست، مانند گذشته، به غنی ساختن ریاضی کمک شایانی می‌کنند. امروز شاید بیش از گذشته انتقاد هیلبرت بجا به نظر می‌رسد که شهودگرایان را به نوعی کوتاه‌فکری مشابه تمایلات عده‌ای که روزگاری مخالف استفاده از اعداد «موهومی» بودند، متهم می‌کرد. البته این انتقاد تا وقتی وارد است که شهودگرایی ادعای دربرگرفتن همه ریاضیات را داشته باشد؛ ولی نمی‌توان منکر شد که شهودگرایی به عنوان یک رشته از ریاضی ارزش فراوانی دارد. همچنان که پژوهش‌های بیشاپ نشان می‌دهد، بررسی «ساخت‌پذیری» مفاهیم ریاضی، خود مقوله‌ای جذاب است که ممکن است حتی در خارج ریاضیات هم به کار گرفته شود.

اکنون می‌پردازیم به صورتگرایی. همان‌طور که قبلاً گفتیم، آنچه به عنوان انتقاد از صورتگرایی گفته خواهد شد، به منطق‌گرایی نیز وارد است. به اشکال درونی صورتگرایی که با کشف قضیه ناتمامیت گودل برملا شد، اشاره کردیم. بی‌شک قضیه گودل نارسا بودن دستگاه‌های اصل موضوعی و مکانیسم اثبات در آن را به نحوی که صورتگرایان پیشنهاد کرده‌اند، به وجهی قاطع نشان داده است. در چند دهه اخیر، صورتگرایی چنان غلبه‌ای بر چهره ریاضیات داشته است که در نتیجه آن، بعضی افراد، قضیه گودل را یک نقصان درونی ریاضیات تلقی کرده‌اند. ولی نباید اثبات به معنای عرفی را با اثبات به طریق صورتگرایانه، یکی گرفت. چه بسا استدلال‌هایی که در چارچوب صوری یک نظریه اصل موضوعی، نمی‌گنجند ولی از نظر هر ریاضیدانی به عنوان اثبات (استدلال قانع‌کننده) مورد قبول هستند. مثلاً درستی گزاره‌هایی که گودل اثبات‌ناپذیر بودن آنها را ثابت می‌کند، به راحتی مورد پذیرش است. مثالی با همین ماهیت را که برای من جالب است، نقل می‌کنم.^۲ فرض کنید روزی ثابت شود که حدس گلدباخ^۳ (اینکه هر عدد زوج بزرگتر از ۲ مجموع دو عدد اول است) در چارچوب اصول پنانو اثبات‌پذیر نیست. در نتیجه نمی‌توان مثالی نقض

^۲ مرجع [۷] بسیار روشن‌کننده بوده است.

^۱ non-standard mathematics ^۳ Goldbach conjecture

در مجموعهٔ استاندارد اعداد طبیعی برای آن یافت و مطلب اخیر نشان می‌دهد که حدس گلدباخ برای اعداد طبیعی درست است. کدام ریاضیدان است که این استدلال را مردود بداند؟ اگر $G(2n)$ حدس گلدباخ باشد که می‌گوید عدد زوج $2n$ را می‌توان به صورت مجموع دو عدد اول نوشت و PA مجموعهٔ اصول پثانو باشد، ما نشان داده‌ایم که

$$\forall n \in \mathbb{N} : PA \vdash G(2n)$$

در حالی که اثبات قابل قبول از دیدگاه صورتگرایی باید به شکل زیر باشد:

$$PA \vdash (\forall n \in \mathbb{N} : G(2n)).$$

در دیدگاه صورتگرایی، اثبات ریاضی دیگر آن وسیلهٔ کشف مانند دیدن و شنیدن نیست، بلکه یک فرآیند ماشینی است که یک رایانهٔ به اندازهٔ کافی نیرومند نیز قادر به انجام آن است. اگر توجه کنیم که گزاره‌های جالب ریاضی، عموماً در تعداد نامتناهی موردِ خاص صادق هستند، آیا معقول است انتظار داشته باشیم که بتوان با روش‌های متناهی ماشینی صورتگرایی به آنها دست یافت؟ اینکه اثبات ریاضی، به مفهوم فرعی، ما را قادر می‌سازد در تعداد متناهی گام، به واقعیتی که در تعداد نامتناهی موردِ صادق است، دست بیابیم، اعجازی است که باید ریشهٔ آن را در روانشناسی اثبات و مدل‌سازی ریاضی از تجربه^۱ جستجو کرد. رنه توم به تشابهی میان صوری‌سازی نظریه‌های ریاضی و وضع کردن دستگاه مختصات اشاره می‌کند. همان‌طور که معمولاً نمی‌توان یک خمینه را با دستگاه مختصات واحدی توصیف کرد و استفاده از دستگاه مختصات تنها یک حربهٔ موضعی است، نباید انتظار داشت که بتوان همهٔ ویژگی‌های یک دسته اشیاء ریاضی را با وضع یک دستگاه صوری واحد، بیان کرد. البته کارایی موضعی دستگاه‌های صوری را قرن‌ها تجربهٔ ریاضی تأیید می‌کند. بی‌دلیل نیست که علی‌رغم قضیهٔ گودل، صورتگرایی بیش از دو مکتب دیگر، مورد قبول ریاضیدانان حرفه‌ای واقع شده است.

در همین زمینه، به رویدادی در سال‌های اخیر اشاره می‌کنم که تا حد زیادی گویای واقعیت است. چند سال پیش اعلام شد که حدس چهار رنگ، بالاخره ثابت شده است. آنچه تازگی داشت استفادهٔ غیرعادی (ده‌ها ساعت) از رایانه در اثبات این حدس بود. به‌طور خلاصه می‌توان گفت اثبات این حدس منجر به تحقیق تعداد زیادی از موارد خاص شده و کاربرد رایانه، در تحقیق صحت حدس در این موارد بود. حجم عمل رایانه در تحقیق این حالت‌های خاص چنان عظیم بود که تنها راه تأیید صحت آن، به جریان انداختن مجدد برنامهٔ رایانه‌ای یا نوشتن برنامه‌ای دیگر با همین منظور بود. جالب اینجاست که با وجود قدمت مسئله، حل آن هیجان‌چندانی در محافل ریاضی ایجاد نکرد. علت سردی استقبال را باید در این یافت که هنوز هم از نظر یک ریاضیدان حرفه‌ای، مهم‌ترین نقش اثبات روشن ساختن روابط است نه انباشتن

^۱ در اینجا مقصود، تجربه به معنای عام است که شامل تجربه با اشیاء ریاضی هم می‌شود.

پی در پی و بدون تبعیض حقایق منطقی. آنچه در مسئله چهار رنگ واقعاً مطلوب است درک کامل مسئله رنگ کردن است (که هنوز حل نشده است) نه اینکه چه عدد به خصوصی برای این کار کفایت می‌کند. آزمایش‌های ماشینی با مفاهیم ریاضی می‌تواند نقش مهمی در ریاضی ایفا کند البته نه در زمینه اثبات، بلکه در زمینه تولید مثال و شاهد برای دنبال کردن مطالب.

دیدگاه آزادمنشانه صورت‌گرایان و منطق‌گرایان درباره «وجود» در ریاضیات («هر چه منجر به تناقض نشود، وجود دارد»)، اثر مختلطی بر تجربه ریاضی داشته است که بی‌شبهت به اثر آزاد شدن سبک‌های هنری در عصر حاضر نیست. از یک سو این دیدگاه، رهایی‌آفرین است و ریاضیدان را از قید الزاماتی که در قرن‌های پیشین، پیشرفت موضوع‌هایی مانند هندسه ناقلیدسی را دشوار ساخته بود، آزاد می‌سازد و از سوی دیگر، همچنان که قبلاً اشاره کردیم، این آزادی، منجر به پیدایش موجودات عجیب و غریبی در ریاضیات شده است که بدون حاکمیت یک دستگاه ارزشی در ریاضی، علم ما را به هرج و مرج می‌کشاند. در هر حال، اگر «وجود» در ریاضی مطابق نظر هیلبرت و راسل مستقر شود، دیگر مقوله «وجود» اهمیت بیش از اندازه‌ای در کار ریاضی نخواهد داشت. در واقع آنچه به جای «وجود»، اهمیت عملی و فلسفی در ریاضیات پیدا خواهد کرد، ارزشمندی موجودات ریاضی است. تحولات دو دهه اخیر در ریاضی، در جهتی است که می‌توان تبلور نوعی اجماع نظر در این زمینه را احساس کرد. تعریف‌ها و مباحثی در ریاضی قابل توجه‌اند که مبتنی بر تجربه (به مفهوم عام) باشند و نظریه‌های نو وقتی ارزش پیگیری علمی دارند که مشکل‌گشای مسائل ریشه‌دار باشند.

خلاصه آنچه در باب نقد مکتب‌های سه‌گانه فلسفه ریاضی گفتیم این است که این مکتب‌ها به علت اشتغال بیش از اندازه به بازگرداندن یقین به ریاضیات، از تجربه عینی این علم روی گردانیدند و کوشش کردند ریاضیات را بر پایه‌های غیرتاریخی و حتی ضد تاریخی بازسازی کنند. در مورد شهودگرایی، موفقیت این کوشش به قیمت طرد بخش بزرگی از علم ریاضی تمام می‌شود و در منطق‌گرایی و صورت‌گرایی، قضیه گودل، این کوشش‌ها را از درون به شکست محکوم می‌کند. این مکتب‌ها را از این رو غیرتاریخی و ضد تاریخی خوانده‌ایم که هر کدام به نحوی، تحت تأثیر نهضت حسابی‌سازی ریاضیات در قرن نوزدهم، مفاهیم اولیه و متعارف ریاضی را طوری نقض کردند که تعریف‌ها دیگر جنبه معرفی ندارند و حتی اگر به لحاظ منطقی، بازگویی کلیه ویژگی‌های شناخته شده مفاهیم‌ها باشند، از نظر عملی بر شهود ریاضیدانان منطبق نیستند. کدام ریاضیدان است که در تفکر ریاضی خود، زوج مرتب (a, b) را به صورت $\{\{a\}, \{a, b\}\}$ و یا عدد حقیقی را یک رده هم‌ارزی از دنباله‌های کوشی رده‌های هم‌ارزی زوج مرتب‌های اعداد صحیح تجسم کند؟! گمانم بر این است که به اندازه کافی علت‌های عدم استقبال ریاضیدانان از مکتب‌های فلسفی سه‌گانه را توصیف کرده باشم.

۷. ریاضیات به مثابه یک علم طبیعی

معمولاً گفته می‌شود که ریاضیدانان حرفه‌ای اکثراً افلاطون‌گرا^۱ هستند. این گفته اگر به این معنی تلقی شود که امروزه نوعی مکتب مدوّن متأثر از افلاطون‌گرایی در ریاضیات وجود دارد، چنین ادعایی درست نیست. امروز مکتب افلاطونی مشخص و شناخته‌شده‌ای در ریاضیات وجود ندارد و شاید اکثر ریاضیدانان، مثل افلاطونی را در زمره خرافات ماقبل علم بیندارند.^۲ آنچه می‌توان تمایل افلاطون‌گرایی در میان ریاضیدانان نامید، این احساس ذهنی هر شخص ریاضی‌کار است که مفاهیم ریاضی، نوعی عینیت مستقل از ذهن ریاضیدان دارند و پژوهش ریاضی، مانند پژوهش در علوم طبیعی، نوعی اکتشاف است. مفهومی‌های عمده ریاضی، هیچ‌یک قرارداد دلخواه یک ریاضیدان نیستند، بلکه همه در طول زمان، پس از آزمایش با پیشنهادهای گوناگون، تحقق یافته‌اند. دامنه جهان ریاضی در طول قرن‌ها پژوهش چنان گستره‌ای پیدا کرده است که ریشه بسیاری از مفاهیم برای تازه‌کارها، روشن نیست و پایه‌های تجربی این علم نیز از نظر بسیاری افراد مکتوم است. با اینکه بعضی مفهومی‌های ریاضی با چند درجه تجرید از جهان فیزیکی ملموس فاصله دارند، تجربه و انس ریاضی‌کاران با آنها، احساسی عینیتی ایجاد می‌کند که عملاً همدریف عینیتی است که فیزیکدان نظری با آن سروکار دارد. فراموش نشود که در فیزیک هم مفهومی‌هایی مانند انرژی پتانسیل، نمودار فاینمن و حتی الکترون، مجردات و مدل‌هایی هستند که برای هماهنگ ساختن مشاهدات فیزیکی ابداع شده‌اند و نمی‌توان ادعا کرد هیچ‌یک از آنها عینیتی بیش از عینیت عدد حقیقی، مثلث، گروه یا انتگرال دارد. اگر امکان ترسیم یا مشاهده یک مثلث کاملاً دقیق در جهان وجود ندارد، هیچ دستگاه مکانیکی بدون اصطکاک نیز هرگز یافت نمی‌شود. چنانچه این واقعیت را که هم از نظر ریشه تاریخی و هم از نظر عملکرد روزمره، ریاضیات یک علم طبیعی است (به مفهوم عام، یعنی با شامل دانستن برخی مفاهیم مأنوس ذهنی در میان اشیاء طبیعی) در دیدگاه یک مکتب فلسفی ریاضی منظور نشود، آن دیدگاه اعتبار چندانی در دید ریاضی‌کاران نخواهد داشت. در واقع آنچه ریاضیدانان حرفه‌ای در سال‌های اخیر در زمینه فلسفه ریاضی نوشته‌اند، در جهت احیای این دیدگاه معتبر تاریخی است و می‌توان پیدایش تدریجی جنبش‌هایی در این مسیر را کم‌کم احساس کرد.

برای اینکه صحبت خود را با نوای امیدبخشی به پایان برسانم، اشاره‌ای کوتاه به کارهای ایمره لاکاتش می‌کنم که شاید سرآغاز یک جنبش ثمربخش در فلسفه ریاضی باشد. اولین و مشهورترین اثر لاکاتش، کتاب اثبات و ابطال [۱۲] است که مبتنی بر رساله سال ۱۹۶۱ وی در دانشگاه کمبریج است. در این اثر و در کارهای متعاقب آن^۳، لاکاتش به بررسی رشد نظریه‌های ریاضی از دیدگاه شناخت‌شناسی پرداخته است.^۴ تنها استثنای جدی در این مورد، شاید ریاضیدان روسی، شافارویچ، باشد. به گفتاری از او که در [۴] نقل شده است، مراجعه کنید.^۵ رجوع کنید به [۱۳] و فهرست مراجع آن.

^۱Platonist

و تصویری رسم می‌کند که با تجربه تاریخی ریاضی بسیار سازگارتر از سه مکتب یادشده است. او رشد نظریه‌های ریاضی را محصول فعل و انفعالی میان حدس‌های معقول و کوشش برای اثبات (غیرضروری) این حدس‌ها می‌داند. به نظر لاکاتش، قضیه‌ها و حتی تعریف‌های ریاضی، تاریخچه تحولی دارند که در طی آن، در اثر برخورد با مثال‌های ناقض موضعی و کلی^۱، شکل می‌گیرند. مقصود از مثال ناقض موضعی، مثالی است که یک گام از استدلال اکتشافی^۲ را نقض کند و مثال ناقض کلی، مثالی است که حکم اصلی را باطل سازد. نمونه‌ای که لاکاتش در کتاب اثبات و ابطال به سبک مناظره‌های افلاطونی می‌پروراند، تحول مفهوم شاخص اویلر^۳ است که با حدس اولیه دکارت - اویلر در مورد چندوجهی‌ها شروع می‌شود و پس از کوشش‌های متمادی و اثبات‌های ناقص، به تدریج به صورت دقیقی درمی‌آید که پوانکاره عرضه کرده است. لاکاتش معتقد است که نظریه‌های زنده و متحرک ریاضی هیچ‌کدام، اقلیدسی نیستند، یعنی جریان انتقال «درستی»، از اصول به نتایج نیست، بلکه این نظریات، «شبه‌تجربی»^۴ هستند، یعنی نادرستی نتایج موجب ترمیم و تحول اصول می‌شود. این دیدگاه طبعاً تحت تأثیر نظرات کارل پوپر در زمینه فلسفه علوم طبیعی است و نیز تأکیدهای پولیا درباره نقش اساسی استدلال‌های اکتشافی را به یاد می‌آورد.

در مقوله «وجود»، لاکاتش چیزی از خود بجا نگذاشته است. این ممکن است به سبب مرگ نابهنگام وی در سال ۱۹۷۴ باشد یا به این دلیل که او اساساً اهمیتی برای مسئله وجود قائل نبوده است (سواى آنچه درباره تحول تعریف‌ها در ریاضی اظهار می‌کند). در هر صورت، چون قرار است یک سخنرانی کامل از این سمینار به بحث درباره آثار لاکاتش اختصاص داده شود، در اینجا سخن را کوتاه می‌کنم. باشد که ریاضی‌کاران و فلاسفه ریاضی بیش از گذشته به یکدیگر نزدیک شوند تا فلسفه ریاضی سیاقی دوباره بیابد.

مراجع

- [1] Ayer, A. J., *Language, Truth and Logic*, Dover, New York, 1952.
- [2] Benacerraf, P., and Putnam, H. (editors), *Philosophy of Mathematics: Selected Readings*, Prentice-Hall, Inc., 1964.
- [3] Boyer, C. B., *The History of Calculus and Its Conceptual Development*, Dover, New York, 1959.
- [4] Davis, P. J., and Hersch, R., *The Mathematical Experience*, Birkhäuser, Boston, 1981.
- [5] Dawson, J., The Godel incompleteness theorem from length-of-proof perspective, *Amer. Math. Monthly*, **86** (1979), 740-747.
- [6] Dummett, M., *Elements of Intuitionism*, Oxford University Press, Oxford, New York, 1977.
- [7] Goodman, N. D., Mathematics as an objective science, *Amer. Math. Monthly*, **86** (1979), 540-551.

^۱local and global counterexamples ^۲heuristic argument ^۳Euler characteristic ^۴quasi-empirical

- [8] Hersch, R., Some proposals for reviving the philosophy of mathematics, *Adv. in Math.*, **31** (1979), 31-50.
- [9] Heyting A., *Intuitionism: An Introduction*, North-Holand, Amsterdam, 1956.
- [10] Howson, A. G. (editor), *Developments in Mathematical Education*, Cambridge University Press, Cambridge, 1973.
- [11] Kant, I., *Critique of Pure Reason*, St. Martin's Press, New York, 1965.
- [12] Lakatos, I., *Proofs and Reputations: The Logic of Mathematical Discovery*, Cambridge University Press, Cambridge, 1976.
- [13] Lakatos, I., *Mathematics, Science and Epistemology*, Cambridge University Press, Cambridge, 1978.
- [14] Manin, Yu., *A Course in mathematical Logic*, Springer-Verlag, New York, 1977.
- [15] Manin, Yu., *Mathematics and Physics*, Birkhäuser, Boston, 1981.
- [16] Poincare, H., *Science and Hypothesis*, Dover, New York, 1938.
- [17] Polya, G., *Patterns of Plausible Reasoning* (2 Vols.), Princeton University Press, 1954.
- [18] Popper, K., *Objective Knowledge: An Evolutionary Approach*, Oxford University Press, 1972.
- [19] Renyi, A., *Dialogues on Mathematics*, Holden-Day, 1967.
- [20] Russell, B., *The Principles of Mathematics*, Cambridge University Press, Cambridge, 1903.
- [21] Russell, B., Whitehead A. N., *Principia Mathematica*, Cambridge University Press, Cambridge, 1910.
- [22] Snapper, E., Are mathematical theorems analytic or synthetic? *Math. Intelligencer*, **3** (1981), 85-88.
- [23] Struik, D. J., *A Source Book in Mathematics: 1200-1800*, Harvard University Press, 1969.
- [24] Thom, R., Modern mathematics: an educational and philosophical error, *Am. Sci.*, **59** (1971), 695-699.
- [25] van der Waerden, L., *Science Awakening*, Noordhoff, Netherlands, 1975.
- [26] Weyl, H., *Philosophy of Mathematics and Natural Science*, Princeton University Press, 1949.
- [27] Wilder, R., *The Foundations of Mathematics*, John-Wiley & Sons, New York, 1965.

میراث فوریه*

ژان-پیر کاهان

برگردان: سعید مقصودی

۱. غرض از مقاله حاضر

میراث فوریه وجوه بسیاری دارد. فوریه در وهله اول یک فیزیکدان و یک ریاضیدان است. نام فوریه برای ریاضیدانان، فیزیکدانان، مهندسان و کلاً دانشمندان علوم طبیعی آشناست. اصطلاحاتی مانند معادله فوریه یا همان معادله گرما، سری فوریه، ضرایب فوریه، انتگرال‌های فوریه، تبدیل‌های فوریه، آنالیز فوریه، تبدیل‌های سریع فوریه جزء اصطلاحات علمی روزمره‌اند. مقاله «نظریه تحلیلی گرما» رویدادی دوران‌ساز در علم به شمار می‌آید. اما فوریه در حوزه مصرشناسی نیز معروف است. او مقدمه جامعی بر مجموعه کتاب‌های «مصر»^۱ نگاشته است و زمانی که سنگ رشید^۲ کشف شد، در مصر حضور داشت و باعث آشنایی ژان فرانسوا شامپولین^۳، رمزگشای کتیبه‌های هیروگلیف، با این موضوع شد. فوریه اهل سیاست و مدیریت هم بود. در انقلاب فرانسه شرکت داشت (آراگو^۴ گفته است که فوریه ثمره ناب انقلاب فرانسه بود، زیرا در ابتدا بنا بود به سلک کشیشان درآید) و در مقام «دبیر دائمی مؤسسه مطالعات مصر»^۵ بناپارت و مونژ^۶ را در مصر همراهی کرد. پس از آن، بناپارت او را به سمت فرماندار گرنویل^۷ انتخاب کرد و وی در آنجا اقدامات مهمی در زمینه آموزش و بهداشت صورت داد. فوریه پس از سقوط ناپلئون هم

عبارات و کلمات کلیدی. سری فوریه؛ سری مثلثاتی؛ قضیه‌های همگرایی سری فوریه؛ قضیه یکتایی کانتور؛ آنالیز ریاضی در قرن نوزدهم.

* نام و نشان مقاله اصلی از این قرار است:

Kahan, Jean-Pierre, *The Heritage of Fourier, Perspectives in Analysis: Essays in Honor of Lennart Carleson's 75th Birthday*, Springer-Verlag, New York, 2005, 83-95.

^۱Description de l'Egypte ^۲Rosetta stone ^۳Jean-François Champollion ^۴Arago ^۵Institut d'Egypte

^۶Monge ^۷Grenoble

به پاریس آمد و همان‌جا ماند و به عضویت دو فرهنگستان علوم^۱ و فرهنگستان فرانسه^۲ درآمد؛ به مقام دبیری دائم فرهنگستان علوم برگزیده شد و در رسمیت بخشیدن به علم آمار در فرانسه، نقش داشت.

«نظریه تحلیلی گرما» و ابداع ابزارهای مربوط به آن، تنها اثر علمی فوریه نیست. او به معادله‌های جبری نیز علاقه‌مند بود و مطالعاتش درباره تعیین محل ریشه‌های این معادله‌ها، عین تحولی است که در آثار استورم^۳ نسبت به آثار دکارت^۴ در این زمینه شاهد هستیم. متأسفانه فوریه اعتنایی به تحقیقات گالوا در این زمینه نکرد و به تحقیقات خود او درباره نامعادله‌ها، یا به اصطلاح خودش، «تحلیل نامتعین‌ها»^۵ هم اعتنایی نمی‌شود. تصور داریو^۶ این بود که فوریه در اهمیت این موضوع راه اغراق رفته است و به همین سبب، در ویراستی که از مجموعه آثار علمی فوریه تهیه کرد، مقاله‌های مربوط به آن موضوع را کنار گذاشت. اگر آن مقاله‌ها منتشر می‌شد، برنامه‌ریزی خطی و آنالیز محدب نیز جزء میراث فوریه محسوب می‌شد.

فوریه انسانی فرهیخته و فیلسوفی با معیارهای قرن هجدهم بود. او از جهتی، یکی از آخرین نمایندگان مکتب عصر روشنگری است. از سوی دیگر، آثارش منبع اطلاعاتی مهمی درباره اوگوست کنت^۷ است. کنت نقطه آغاز مکتب «پوزیتیویسم»^۸ فرانسه در قرن نوزدهم بود.

در اینجا به بخشی کوچک اما مهم از میراث علمی او خواهیم پرداخت که همان موضوع بسط یک تابع برحسب سری مثلثاتی و دستور محاسبه ضرایب مربوط به آن است. این موضوع یکی از راه‌هایی است که می‌توان تأمل درباره فوریه و ارتباط او با فیزیک و فلسفه طبیعی را از آن آغاز کرد و استمرار تحقیقاتش در حوزه ریاضیات توسط دیگران را نیز مورد بررسی قرار داد.

درباره طرز نگاه‌ها به فوریه، تصوراتی من این است که برای مدت طولانی در فرانسه و حتی شاید در آلمان، اهمیت چندانی به او داده نشد و تنها در همین سال‌های اخیر است که از او بسیار نام برده می‌شود. این امر با بیان چند مطلب و نقل قول در ادامه، روشن خواهد شد. برای توضیح تداوم تحقیقات فوریه درباره سری‌های مثلثاتی توسط دیگران، به موضوعات بسیاری بر اساس اهمیت تاریخی و اهمیت حال حاضر آنها خواهیم پرداخت. به هر حال، بخش زیادی از همایش با موضوع «چشم‌اندازهایی در آنالیز»^۹ را می‌شود شرح میراث فوریه محسوب کرد.

بخش عمده این مقاله تشکیل شده است از اظهارنظرهای مختلف و نقل قول‌ها (۱.۲) ویکتور هوگو، ۲.۲ یاکوبی، ۳.۲ فوریه، ۴.۲ دیریکله و ریمان). در بخش ۳، به نظریه ریمان درباره سری‌های مثلثاتی و در بخش ۴، به مسئله همگرایی (۱.۴ قضیه کارلسون، ۲.۴ شکل‌های متفاوت همگرایی) پرداخته شده است. در پایان، در بخش ۵، به موضوع بازگشت دوباره فوریه خواهیم پرداخت.

^۱ Académie des sciences ^۲ Académie Française ^۳ Jacques C. F. Sturm ^۴ Descartes ^۵ Analyse indéterminée ^۶ Darboux ^۷ Auguste Comte ^۸ positivism ^۹ Perspectives in Analysis

۲. چند نقل قول

۱۰۲. ویکتور هوگو. کار را با ویکتور هوگو^۱ آغاز می‌کنیم. هوگو در سال ۱۸۶۲ در جزیرهٔ گرنزی^۲ در تبعید به‌سر می‌برد و در همان‌جا بخش زیادی از کتابش را به نگارش درآورد. این همان سالی است که رمان او با نام «بینوایان» منتشر شد. این رمان دربرگیرندهٔ مطالب زیادی راجع به روزگار فرانسه در آغاز قرن نوزدهم است. عنوان یکی از فصل‌های آن، «به سال ۱۸۱۷» است. در این فصل، ژوزف فوریه طی عبارتی کوتاه در داستان ظاهر می‌شود:

«در فرهنگستان علوم فرانسه یک فوریهٔ مشهور بود که اخلاف فراموشش کردند، و در همان موقع نمی‌دانم در کدام کلبهٔ محقر، فوریهٔ گمنامی وجود داشت که زمان آینده به‌یادش خواهد آورد.» (بینوایان، ویکتور هوگو، ۱۸۶۲)

منظور مورد اول، ژوزف فوریه و دومی، شارل فوریهٔ آرمان‌گرا است. مسلماً ویکتور هوگو، ژوزف فوریه را «افتخار ملی» به‌شمار نمی‌آورد. فوریه با فرانسوا آراگو که بعداً در سِمَت «دبیر دائم فرهنگستان علوم» جانشین فوریه شد، رابطهٔ دوستی داشت. پس از درگذشت فوریه در سال ۱۸۳۰، آراگو به‌عنوان یک دوست و ویکتور کوژن^۳ به‌عنوان عضوی از فرهنگستان فرانسه، سوگنامه‌هایی نگاشتند. هر دو سوگنامه به‌لحاظ مطالب مربوط به زندگی ژوزف فوریه بسیار ارزشمند و جالب‌اند، لیکن در هیچ‌یک از آنها به تحقیقات وی در ریاضیات اعتنایی نشده است و یا ارزش کمی برای آن قائل شده‌اند. آراگو که در جوانی و خیلی زودتر از فوریه به عضویت فرهنگستان علوم برگزیده شد، احتمالاً بی‌رغبتی زیاد لاگرانژ را به موضوع تجزیهٔ یک تابع برحسب سری‌های مثلثاتی به‌خاطر داشته و خبر داشته است که در اعطای جایزهٔ فرهنگستان به فوریه برای تحقیقاتش دربارهٔ نظریهٔ تحلیلی گرما در سال ۱۸۱۱، لاگرانژ شک و تردیدهایی را دربارهٔ آن نظریه هم از حیث کلیت و هم به‌لحاظ دقت، مطرح کرده بود. فوریه رقیبانی نظیر کوشی و پواسن^۴ داشت. از این مطالب روشن می‌شود که چرا سوگنامهٔ آراگو، چیزی شبیه به یک مراسم خاکسپاری مجلل به نظر می‌رسد. واقع امر این است که فرانسویان تا همین اواخر بر جایگاه فوریه واقف نبودند. در پاریس «خیابان شارل فوریه» وجود دارد، اما خیابان ژوزف فوریه وجود ندارد. در اولین ویراست دانشنامهٔ انویرسال^۵، که در فرانسه هم‌تراز دانشنامهٔ بریتانیکا است، هیچ مقاله‌ای دربارهٔ ژوزف فوریه نیامده است؛ در ویراست ششم آن به سال ۱۹۷۴ هم وضع به‌همین منوال بود. پیش از این گفته شد که داریو تنها بخشی از آثار ریاضی فوریه، و در واقع نظریهٔ تحلیلی گرما، را منتشر کرد و هیچ‌گاه مجموعه آثار فوریه انتشار نیافت. تا همین سال‌های اخیر هم در فرانسه، حداقل بجز گرنوبل، توجه زیادی به زندگی و آثار ژوزف فوریه نمی‌شد؛ اما در شهر گرنوبل مدت‌ها پیش، مؤسسهٔ ریاضی آنجا به اسم مؤسسهٔ ژوزف فوریه نامگذاری شد و در سال ۱۹۸۷ دانشگاهی را که این مؤسسه وابسته به آن بود، به اسم دانشگاه ژوزف فوریه نامگذاری کردند. در

^۱Victor Hugo ^۲Guernsey ^۳Victor Cousin ^۴Poisson ^۵Universalis

سال ۱۹۹۸ کتاب فوق‌العاده‌ای منتشر شد: «فوریه، خالق ریاضیات فیزیک»^۱ نوشته ژان دُنیر^۲ ریاضیدان و ژان-برنارد روبر^۳ فیزیکدان. نشانه‌هایی وجود دارد گویای این که نام ژوزف فوریه دیگر در فرانسه از یادها نخواهد رفت.

۲.۲. یاکوبی. مهم‌تر از این موارد، این است که کارل گوستاف یاکوبی^۴ که در آن زمان ۲۶ سال داشت، چند هفته‌ای پس از درگذشت فوریه به او ادای احترام کرد و در نامه‌ای به لژاندر^۵ نوشت:

«شایسته بود که آقای پواسُن داوری خود را درباره آقای فوریه فقید در چنین زمان نامناسبی منتشر نمی‌کرد و من و آبل را به دلیل عدم توجه کافی به موضوع حرکت گرما مورد سرزنش قرار نمی‌داد. در حقیقت، آقای فوریه هدف غایی در ریاضیات را علاوه بر تبیین پدیده‌های طبیعی، نفع عامه مردم می‌دانست. لیکن در جایگاه یک فیلسوف، طبعاً آگاه بود که یگانه غایت علم همان سروری عقل آدمی است و از این حیث، پرسش از اعداد، به اندازه پرسش از جهان هستی ارزشمند است.»

عبارت کلیدی «سروری عقل آدمی» برخلاف دیدگاه فوریه، بدل به شعار ریاضیات محض می‌شود. به‌ویژه می‌بینیم که ژان دیودونه^۶ این شعار را عنوان یکی از کتاب‌های مشهورش قرار داده است.

۳.۲. فوریه. طی چند جمله در مقدمه کلی مقاله فوریه در نظریه تحلیلی گرما، می‌توان به‌روشنی به دیدگاه او پی برد. در اینجا [ترجمه] گزیده‌ای از این جمله‌ها را می‌آوریم.

«بررسی دقیق طبیعت پربارترین منبع اکتشافات ریاضیات است.» «معادله‌های تحلیلی را می‌توان برای همه پدیده‌های عمومی به‌کار برد. زبانی ساده‌تر، فراگیرتر، کم‌خطاتر و تواناتر از این زبان برای بیان روابط پایدار بین اجسام طبیعی وجود ندارد.» «از این منظر، آنالیز ریاضی وسعتی به اندازه خود طبیعت دارد. ویژگی اصلی‌اش بیان روشن آن است. در آن، نشانی از مفاهیم مبهم نیست. پدیده‌های بسیار متفاوتی را به یکدیگر مربوط می‌سازد و شباهت‌های پنهان آنها را آشکار و بدین وسیله آنها را قابل فهم و اندازه‌گیری می‌کند. گویا آنالیز ریاضی، قابلیت از ذهن آدمی است تا جبرانی باشد برای کوتاهی عمر و نقص حواس انسان.»

این تعریفی باشکوه از آنالیز ریاضی است. ضمناً باید افزود که مسائل مربوط به زندگی انسان‌ها و صنعت نیز در فوریه تعلق خاطر ایجاد کرده بودند و این امر به موازات «بررسی دقیق طبیعت» یکی دیگر از منابع مهم تحقیقات و اکتشافات او بود. در اینجا ترجمه گزیده [دیگری] را به تقریب می‌آوریم:

^۱Fourier, créateur de la physique mathématique ^۲Jean Dhombres ^۳Jean- Bernard Robert ^۴Carl Gustav Jacobi ^۵Legendre ^۶Jean Dieudonné

«معادله گرما، همانند معادله‌های مربوط به تارهای مرتعش یا حرکت مایعات، در مبحث کاملاً جدیدی از آنالیز ریاضی قرار می‌گیرد [یعنی معادله‌های دیفرانسیل با مشتقات جزئی]. پس از تشکیل معادله، باید به جستجوی جواب‌ها پرداخت. بدین معنی که از یک بیان کلی، به جواب خصوصی مقید به شرایط از پیش تعیین‌شده، دست یافت. انجام چنین تحقیقاتی، دشوار و نیازمند نوع جدیدی از آنالیز بر پایه قضایای جدید است. . . . این روش هیچ ابهامی در جواب‌ها باقی نمی‌گذارد. این مطالعه و بررسی، همچون هر مطالعه و بررسی مفید دیگر، در آخر به کاربردهای عددی می‌رسد.»

نگرش فوری به‌خوبی در مقدمه کلی گفته‌شده در بالا، بیان شده است: کار از پدیده‌های طبیعی آغاز می‌شود و با نتیجه‌گیری‌های عددی پایان می‌یابد. در این بین، «نوع جدیدی از آنالیز» مورد نیاز است که امروزه آنالیز فوری نامیده می‌شود. اکنون این رهیافت، نسبت به پنجاه سال پیش، با وجود روش‌های مدل‌سازی رایانه‌ای، جلوه‌ای مدرن‌تر یافته است و دلیل رواج دوباره دیدگاه‌های فوری را روشن می‌کند.

۴.۲. دیریکله و ریمان. دیریکله و ریمان دنباله‌روان اصلی فوری بودند. دیریکله در مدتی که بین سال‌های ۱۸۲۲ و ۱۸۲۵ در پاریس اقامت داشت، فوریه را ملاقات کرد. دیریکله در آن زمان، بیست سال هم نداشت. او در سال ۱۸۲۹ اولین حکم کلی و بدون غلط را درباره همگرایی سری‌های فوری منتشر کرد. آن مقاله با ادای احترام به فوریه (هندسه‌دان نامی که راهی نو در کاربرد آنالیز گشوده است) و انتقادی از رویکرد کوشی به مسئله همگرایی آغاز می‌شود. نقطه آغاز بحث دیریکله در این موضوع همانند فوریه، تابعی است که به طریقی مشخص شده است. پس از آن، دستورهای برحسب انتگرال داده می‌شوند که ضرایب مربوط به سری فوریه را به‌دست می‌دهند. مسئله این بود که نشان دهیم این سری فوریه به آن تابع همگرا است. راه‌حل دیریکله برای این مسئله، شاهکاری در آنالیز بود و به شرایط مشهور دیریکله در باب این مسئله منتهی شد. نکته مهمی هم به دنبال داشت: برای اینکه دستورهای انتگرالی مربوط، معنی‌دار باشند، لازم است تابع داده شده در برخی شرایط صدق کند. دیریکله طی بحثی کوتاه در این باره، مثال معروفش از تابعی را که روی نقاط گویا یک مقدار و روی نقاط گنگ مقدار دیگری اختیار می‌کند می‌آورد؛ چنین تابعی نه با تعریف وی و نه با تعریفی که بعدها ریمان برای انتگرال بیان کرد، روی هیچ بازه‌ای انتگرال‌پذیر نیست. از زمان دیریکله به‌بعد، مسئله همگرایی سری فوریه (یعنی، سری مثلثاتی‌ای که ضرایب آن از دستورهای انتگرالی فوریه به‌دست آمده باشند) به دو سؤال بنیادی پیوند می‌خورد: منظور از تابع چیست؟ منظور از انتگرال چیست؟ واقعیت این است که اصطلاح سری فوریه^۱ نخستین بار در رساله ریمان در باب سری‌های مثلثاتی، نگاشته‌شده در سال ۱۸۵۴، آمده است. رساله ریمان با تاریخچه‌ای

^۱Fouriersche Reihe

انتقادی از موضوع مورد بحث از زمان آغاز مناقشه راجع به تارهای مرتعش در قرن هجدهم تا زمان انتشار مقاله یادشده از دیریکله، شروع می‌شود. مقاله دیریکله، از یک سو، درباره ادامه تحلیلی و هارمونیک توابع و از سوی دیگر، درباره همگرایی معمولی و مطلق سری‌های عددی، ملهم از اشتباهات کوشی، بود. چند صفحه‌ای از رساله هم به انتگرال‌های عادی و تعمیم‌یافته اختصاص داشت. انتگرال ریمان تابع کراندار روی بازه کراندار طی چند سطر تعریف شده است و (پیش از لبگ!) ریمان شرایط لازم و کافی صریحی برای انتگرال‌پذیری تابع به‌دست می‌دهد.

یکی از مضمون‌های سراسری رساله ریمان، تصدیق نقش فوریه در این موضوع است. پس از کشمکش‌های راجع به تارهای مرتعش بین دالامبر^۱، اویلر^۲، دانیل برنولی^۳ و لاگرانژ^۴، «موضوع نمایش تابع دلخواه برحسب عبارتی تحلیلی، تقریباً به مدت پنجاه سال هیچ پیشرفت اساسی‌ای نکرده بود تا اینکه فوریه نکته‌ای بیان کرد و این مسئله ظاهر جدیدی به خود گرفت و در این بخش از ریاضیات، دورانی نو آغاز گردید. دیری نپایید که اهمیت فوق‌العاده این موضوع برای پیشرفت فیزیک ریاضیاتی معلوم شد. فوریه این نکته را بیان می‌کند که برای سری مثلثاتی مفروض

$$f(x) = a_1 \sin x + a_2 \sin 2x + \dots + \frac{1}{p} b_0 + b_1 \cos x + b_2 \cos 2x + \dots,$$

ضرایب که با فرمول‌های

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx$$

معین می‌شوند، خوش‌تعریف‌اند. او اظهار می‌کند که تعریف این سری برای تابع کاملاً دلخواه $f(x)$ معنی‌دار است. فوریه $f(x)$ را تابعی به‌اصطلاح ناپیوسته می‌گیرد (نمودار یک خط شکسته بالای محور x ها) و سری‌ای به‌دست می‌آورد که در واقع همواره مقدار آن تابع را به‌دست می‌دهد. «ریمان پس از بحث درباره دلیل بی‌میلی لاگرانژ و خصومت پواسن، اظهار می‌کند که «در واقع فوریه بود که ماهیت سری‌های مثلثاتی را به‌طور کامل و به‌درستی دریافت. از آن پس، این سری‌ها بارها در فیزیک ریاضیاتی به‌منظور نمایش توابع دلخواه به‌کار گرفته شده‌اند و به آسانی می‌توان متقاعد شد که در هر مورد خاص، سری فوریه مورد بحث، واقعاً به مقدار تابع همگرا است.» برای این کار، یک برهان کلی مورد نیاز بود. این برهان را، پس از کارهای کوشی در این باره، دیریکله برای رده وسیعی از توابع ارائه داد. این رده (به گفته ریمان) تمامی نیازهای احتمالی فیزیک را در بر می‌گیرد. با وجود این، «کاربرد سری‌های فوریه محدود به تحقیقات فیزیک نیست؛ امروزه آنها را با موفقیت در حوزه‌ای از ریاضیات محض، به نام نظریه اعداد، به‌کار می‌گیرند و ظاهر امر این است که مهم‌ترین توابع در آن نظریه، جزء رده تابعی نیستند که دیریکله بررسی کرده است.»

^۱d'Alembert ^۲Euler ^۳Daniel Bernoulli ^۴Lagrange

۳. نظریه سری‌های مثلثاتی ریمان

با این حال، ریمان در مهم‌ترین و بدیع‌ترین بخش رساله‌اش مسیرهایی را دنبال نمی‌کند که فوریه و دیریکله ترسیم کردند. ریمان به جای اینکه از یک تابع آغاز کند و ویژگی‌های سری فوریه (و به‌یژه مسئله همگرایی) آن را بررسی کند، کار را با یک سری مثلثاتی همه‌جا همگرا شروع می‌کند و به بررسی ویژگی‌های تابع مجموع می‌پردازد. ابزار و روش‌های ابداعی ریمان در این کار، به نظریه سری‌های مثلثاتی ریمان مشهور است. این روش، مستلزم انتگرال‌گیری دوگانه صوری از سری مفروض و عمل ضرب صوری در سری‌های (یا توابع) به قدر کافی منظم است. از دیدگاه امروزی، این مطلب تمهیدی بر رهیافت نظریه توزیع‌های شوارتس است. آشکارتر اینکه ریمان در آنجا مفاهیم «توابع هموار» زیگموند^۱ و «اندازه‌نماها»^۲ و «تابع‌نماها»^۳ سیلم^۴ را به کار می‌برد که بعدها معرفی شدند.

نظریه سری‌های مثلثاتی ریمان نقش بزرگی در تاریخ ریاضیات ایفا کرد. دلیل این امر، طرح دو مسئله‌ای است که ریمان آنها را بدون پاسخ گذاشت و به دست گئورگ کانتور حل شدند: یکی اینکه ثابت کنید اگر یک سری مثلثاتی، همه‌جا همگرا باشد، ضرایب آن به صفر می‌گرایند و مسئله دشوارتر دوم اینکه ثابت کنید چنانچه سری مثلثاتی، همه‌جا به صفر همگرا باشد، آن سری، سری صفر است. حکم اخیر همان «قضیه یکتایی» کانتور است. کانتور سپس در پی «گسترشی» از این حکم برآمد: آیا اگر فرض را به این صورت ضعیف‌تر کنیم که سری مورد بحث، خارج از یک مجموعه مفروض همگرا به صفر باشد، این نتیجه به قوت خود باقی می‌ماند؟ کانتور ثابت کرد برای مجموعه‌های «تحویل‌پذیر»^۵ (مجموعه با بستار شمارا) پاسخ مثبت است. این فرصت اولین فرصتی بود که کانتور برای ابداع نظریه‌اش درباره اعداد حقیقی و مجموعه‌های اعداد حقیقی می‌یافت. لذا نظریه سری‌های مثلثاتی به لحاظ تاریخی سرمنشأ نظریه مجموعه‌ها است. مجموعه‌های یکتایی (که پاسخ پرسش گفته شده برای آنها مثبت است) و مجموعه‌های چندگانگی^۶ (نقطه مقابل مجموعه‌های یکتایی) به صورت یک مکان ایده‌آل برای به کار بستن روش‌های آنالیز تابعی، نظریه احتمال، نظریه اعداد و منطق درآمدند. ریاضیدانانی که آثار مهمی در این زمینه دارند عبارت‌اند از: لبگ، مارتسل ریس^۷، یانگ، منشف^۸، رایچمن^۹، باری^{۱۰}، زیگموند، مارتسینکیه‌ویچ^{۱۱}، سیلم و پس از آنها آر. کوفمن^{۱۲}، تی. کورنر^{۱۳}، جی. پی. کاهان، بی. مندلبرو^{۱۴}، وای. کتس نلسون^{۱۵}، بی. کانس^{۱۶} و متأخرتر از آنها ای. اس. کخریس^{۱۷}، و ای. لوو^{۱۸}، جی. دبز^{۱۹} و جی. سن رمون^{۲۰}، جی. بورگن^{۲۱}، ام. آس^{۲۲} و جی. وانگ^{۲۳}. تاریخچه مختصری از این موضوع را می‌توان در [۴] یافت. قضیه یکتایی کانتور بدین معنی است که چنانچه مجموع یک سری مثلثاتی همه‌جا همگرا، معلوم باشد، ضرایب آن خوش‌تعریف‌اند. برای

^۱Zygmund ^۲pseudomeasures ^۳pseudofunctions ^۴Salem ^۵reducible ^۶multiplicity ^۷Marcel Riesz ^۸Menchoff ^۹Rajchman ^{۱۰}Bari ^{۱۱}Marcinkiewicz ^{۱۲}R. Kaufman ^{۱۳}T. Körner ^{۱۴}B. Mandelbrot ^{۱۵}Y. Katznelson ^{۱۶}B. Connes ^{۱۷}A. S. Kechris ^{۱۸}A. Louveau ^{۱۹}G. Debs ^{۲۰}J. Saint-Raymond ^{۲۱}J. Bourgain ^{۲۲}M. Ash ^{۲۳}G. Wang

محاسبه این ضرایب، نوع جدیدی از دستورات فوریه به انضمام معنای جدیدی از انتگرال مورد نیاز است. مقصود نظریه «کلی سازی»^۱ مرتبه دوم ای. دانژوا^۲ همین امر است. کتاب «درس‌هایی درباره حساب ضرایب سری‌های مثلثاتی» وی (به سال ۱۹۴۱، ۱۹۴۹) در چهار جلد، تقریر کامل همین موضوع است.

۴. مسئله همگرایی پس از دیریکله

نظریه ریمان و قضیه یکتایی کانتور جزء میراث فوریه‌اند، منتها قسمت اصلی آن نیستند. قسمت اصلی را خود ریمان به خوبی بیان کرده که همان مسئله همگرایی است که دیریکله مورد بررسی قرار داده است. تصور دیریکله، علی‌رغم اینکه از عهده اثبات آن برنیامد، این بود که سری فوریه تابع پیوسته باید نقطه‌وار به خود تابع همگرا باشد. نادرستی این مطلب را پل دوبوا-ریمون^۳ بعدها در سال ۱۸۷۳ ثابت کرد و تابعی پیوسته ساخت که سری فوریه‌اش در نقطه مفروضی واگرا است. ساختار آن مثال، صورتی از «متراکم سازی تکنیکی‌ها» است. لبگ و همچنین فیر^۴ مثال‌های دیگری به دست داده‌اند. امروزه این مثال به عنوان کاربردی متداول از قضیه باناخ-اشتاینهاوس ذکر می‌شود و در اثبات آن از این مطلب که نرم L^1 هسته دیریکله (یا، به اصطلاح، ثابت لبگ) کردار نیست، استفاده می‌شود. آیا می‌شود تابع پیوسته‌ای ساخت که سری فوریه آن همه‌جا باشد؟ این سؤال نیز تا سال ۱۹۶۵، پیش از اثبات قضیه کارلسون، بی‌پاسخ مانده بود و کتس نلسون و نگارنده مقاله حاضر ثابت کردند یا چنین تابعی را می‌توان ساخت و یا سری فوریه هر تابع پیوسته، تقریباً همه‌جا به خود تابع همگرا است. نکته در این بود که تابعی پیوسته بسازیم که سری فوریه آن روی مجموعه پوچ مفروضی واگرا باشد.

۱.۴. قضیه کارلسون. قضیه سال ۱۹۶۶ کارلسون^۵ پرسش گفته شده در بالا را پاسخ داد: سری فوریه تابع پیوسته، تقریباً همه‌جا به خود تابع همگرا است و برای توابع پیوسته چیزی بهتر از این نمی‌توان گفت [۱]. ولی چنانچه رده توابع مورد بحث را بزرگتر بگیریم، گفتنی‌های بسیاری داریم. کارلسون ثابت کرد این نتیجه برای L^2 برقرار است (۱۹۶۶)، و آر. هانت^۶ آن را برای L^p که $1 < p$ ثابت کرد (۱۹۶۷). از سوی دیگر، این نتیجه برای L^1 برقرار نیست، زیرا کلموگورف تابعی در L^1 ساخت که سری فوریه‌اش تقریباً همه‌جا (۱۹۲۲) و حتی همه‌جا و اگر است (۱۹۲۶). وضعیت در نزدیکی L^1 در این اواخر مورد بررسی قرار گرفته است و همان‌طور که خواهیم دید، باعث مطرح شدن پرسش‌های پیچیده‌ای شده است. اکنون برای اینکه اهمیت نتیجه سال ۶۷-۱۹۶۶ کارلسون-هانت را درک کنیم، به ویراست و چاپ‌های پیاپی کتاب «سری‌های مثلثاتی» [۱۱] زیگموند رجوع می‌کنیم. بهترین نتیجه شناخته شده پیش از ۱۹۶۶ برای توابع عضو L^2 در حالت کلی متعلق به کلموگورف بود (۱۹۲۲): برای یک دنباله شکافدار^۷ آدامار مانند (n_k) ، مثلاً $n_k = 2^k$ ، مجموعه‌های جزئی مرتبه n_k تقریباً همه‌جا به خود تابع میل می‌کند. برای

^۱totalization ^۲A. Denjoy ^۳Paul du Bois-Raymond ^۴Fejer ^۵Carleson ^۶R. Hunt ^۷situation near

^۸lacunary

توابع عضو L^p ، $p > 1$ ، در حالت کلی همین حکم برقرار است، منتها اثبات آن متکی به یک ابزار پیچیده موسوم به نظریه لیتلوود-پیلی^۱ است (۱۹۳۸). دو فصل از ویراست دوم کتاب زیگموند به نظریه لیتلوود-پیلی و همگرایی تقریباً همه جایی مجموع جزئی مرتبه n_k اختصاص دارد که کاربرد اصلی آن نظریه محسوب می شود. کشف کارلسون، تغییری اساسی در چشم انداز موضوع مورد بحث ایجاد کرد. نخست اینکه نتیجه کلموگورف دیگر بهترین نتیجه نبود. دوم اینکه به سبب گسترشی که هانت از قضیه کارلسون به همه L^p ها، $p > 1$ ، به دست داد، قضیه لیتلوود-پیلی تبدیل به وسیله ای منسوخ در بررسی مسئله همگرایی شد. این موضوع چیز غیرمنتظره ای برای زیگموند بود. در زمان حیاتش چندین «چاپ» جدید از ویراست دوم کتاب «سری های مثلثاتی» به بازار آمد، ولی هیچ «ویراست سومی» عرضه نشد. «ویراست سوم» آن کتاب می بایست یک بازنویسی کامل می بود. برای این کار، قضیه کارلسون-هانت می بایست آورده می شد و بسیاری از نتایج جزئی و روش های مربوط، کنار گذاشته می شد؛ نگاه داشتن نظریه لیتلوود-پیلی هم دیگر هیچ دلیلی نداشت. خوشبختانه زیگموند بازنویسی آن را بدین صورت دنبال نکرد و کتاب با همان شرح زیبایش از نظریه لیتلوود-پیلی، بدون قضیه کارلسون-هانت، باقی مانده است.

وضعیت در نزدیکی L^1 ارتباط نزدیکی با رفتار مجانبی مجموع های جزئی $S_n(f, x)$ دارد وقتی که $f \in L^1$. نخستین نتیجه مهم در این باره از آن هاردی است (۱۹۱۳): اگر $f \in L^1$ ، آن گاه تقریباً همه جا $S_n(f, x) = o(\log n)$. هاردی حدس زد که این بهترین نتیجه ممکن است. در جهت عکس این حکم، مثال سال ۱۹۲۶ کلموگورف وجود تابعی مانند $f \in L^1$ را ثابت کرد که همه جا $\lim_{n \rightarrow \infty} |S_n(f, x)| = \infty$. برای چه دنباله های صعودی $l(n)$ تابع $f \in L^1$ موجود است به طوری که همه جا $\lim_{n \rightarrow \infty} (|S_n(f, x)|/l(n)) = \infty$ ؟ بهترین نتیجه در این باره، تا چند سال پیش این بود که شرط $l(n) = o(\log \log n)$ حکم مورد نظر را به دست می دهد (چن^۲ ۱۹۶۲). در سال ۱۹۹۹، کونیگین^۳ تا $l(n) = o((\log n / \log \log n)^{1/2})$ پیش رفت، و بچکارف^۴ در سال ۲۰۰۳ ثابت کرد که اگر به جای دایره، گروه کانتور را بگذاریم، شرط $l(n) = o((\log n)^{1/2})$ کفایت می کند. اکنون مسئله دشوار، بهتر کردن نتیجه هاردی یا کونیگین است. یک مورد آزمودنی عبارت است از $l(n) = (\log n)^p$ که $1/2 < p < 1$.

۲.۴. شکل های متفاوت همگرایی. رواج دوباره سری های فوریه در قرن بیستم عمدتاً به دلیل مطرح شدن روش های جدید در مطالعه مسئله همگرایی بوده است. در وهله نخست، می توان به جای همگرایی عادی، فرایندهای مجموع پذیری را وارد کار کرد. این راه را فیِر در سال ۱۹۰۰ گشود و به معرفی مفاهیم هسته های مثبت، همانی های تقریبی، ضربگرها و پیش منتهی شد. مهم ترین این مفاهیم، یعنی پیش، خیلی دیر صورت بندی شد و عملاً هیچ تعریف صوری هم حوزه واقعی این مفهوم را پوشش نمی دهد.

^۱Littlewood-Paley ^۲Chen ^۳Konyagin ^۴Bochkarev

جبرهای پیچشی وینر^۱ باعث شد درک بهتری از مفهوم پیچش حاصل شود. جبرهای پیچشی وینر یکی از سرچشمه‌های مفهوم حلقه‌های نرم‌دار (جبرهای باناخ) گلفاند^۲ بودند. از سوی دیگر، همگرایی را می‌توان در فضای توابع و همچنین برای مفاهیم مجموع‌پذیری مورد بررسی قرار داد. این مسیر، پیوندی محکم بین سری‌های فوریه و سرچشمه‌های آنالیز تابعی ایجاد کرد. انتگرال لبگ (۱۹۰۱) و کاربردهای آن در سری‌های فوریه (۱۹۰۶)، تکان‌های اولیه این پیوند به حساب می‌آیند و بعد از آنها قضیه ریس-فیشر (۱۹۰۷) قرار می‌گیرد. حکم ساده L^p کامل است» (که ابتدا نیازمند تعریف فضاهای L^p و فضای متریک کامل است) جزء میراث فوریه است. سری‌های فوریه نخستین نمونه از سری‌های متعامد است و این‌گونه سری‌ها در همه بخش‌های آنالیز ظاهر می‌شوند. در L^2 -همگرایی آنها شکی نیست، ولی مسائل همگرایی تقریباً همه‌جایی آنها، نظیر حالت سری‌های فوریه کلاسیک، واقعاً جالب‌اند. هزاران مقاله و صدها کتاب راجع به این موضوعات در حال نگارش است و من در اینجا دیگر بر آنها تأکید نمی‌کنم.

۵. فوریه دوباره باز می‌گردد

اجازه دهید برگردیم به سراغ فوریه. فوریه سری‌های مثلثاتی را اولین بار و نوعاً در حل مسئله‌ای درباره توزیع دما در یک جسم صلب مفروض به شکل استوانه‌ای با قاعده نیم‌نوار $-\pi/2 \leq x \leq \pi/2$ ، $-\infty < z < \infty$ و $0 \leq y < \infty$ به کار برد. دما در قاعده افقی ($y = 0$) برابر با ۱ و در ضلع‌های قائم ($x = \pm\pi/2$) برابر با ۰ است. در حالت تعادل گرمایی، مقدار دما، $u(x, y)$ ، در معادله

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

صدق می‌کند که $u(\pm\pi/2, y) = 0$ و $u(x, 0) = 1$ ($-\pi/2 < x < \pi/2$). جوابی صوری برای این معادله به صورت

$$u(x, y) = a_1 e^{-y} \cos x + a_3 e^{-3y} \cos 3x + a_5 e^{-5y} \cos 5x + \dots \quad (1.5)$$

است و شرط آخر مسئله با قید

$$1 = a_1 \cos x + a_3 \cos 3x + a_5 \cos 5x + \dots \quad \left(-\frac{\pi}{2} < x < \frac{\pi}{2}\right) \quad (2.5)$$

بیان می‌شود. فوریه پیش از آوردن دستورهای انتگرالی ضرایب، آنها را به شیوه‌ای عجیب و غریب با استفاده از مشتق‌گیری، برش، و حل یک دستگاه خطی محاسبه می‌کند. داریو اظهار می‌کند که این روش، طبیعی است و حق با داریو است. کاری که فوریه انجام داد این بود که چند جمله‌ای مثلثاتی $1 - \sum a_n \cos nx$ (فرد) از درجه مفروض را که در ۰ تا حد ممکن تخت^۳ است مورد توجه قرار داد و سپس حد a_n را هنگامی

^۱Wiener ^۲Gelfand ^۳flat

که درجه به بینهایت میل می‌کند، یعنی $4/\pi n^{n+1}(-1)$ ، محاسبه کرد. مشتق این چندجمله‌ای‌های مثلثاتی تا حد ممکن در 0 تخت هستند، فرد نیز هستند و انتگرال آنها روی $(-\pi/2, 0)$ برابر با 1 است. به زبان نظریه توزیع‌ها، آنها به $2(\delta_{-\pi/2} - \delta_{\pi/2})$ همگرا هستند و این مطلب، برای اثبات (۱.۵) وقتی $y > 0$ کفایت می‌کند. برهان (۲.۵) دشوارتر است.

البته فوریه چیزی از نظریه توزیع‌ها نمی‌دانست، ولی توانایی زیادی در درک محاسبات داشت. پس عاقلانه است که پیش از مطرود شمردن روش‌ها و احکام وی، سعی کنیم از آنها سر در بیاوریم. فوریه پس از به‌دست دادن دستورهای انتگرالی ملاحظه کرد که به دلیل وجود عامل‌های نمایی، همگرایی سری (۱.۵) به‌ازای $y > 0$ بسیار سریع است. نظرش درباره سری (۲.۵) این بود که همگرایی آن را می‌توان ثابت کرد و عملاً هم بعداً در آن فصل اثبات درستی برای آن آورد. او سپس این فرآیند را به توابع دیگر گسترش و دستورهای فوریه را برای توابع متناوب کلی با دوره تناوب 2π به‌دست داد. پس از آن، ادعا کرد که این روش را می‌توان برای توابع دلخواه به‌کار گرفت و یک سری مثلثاتی به‌دست آورد که همواره به آن تابع همگرا باشد. از نگاه دقیق و رسمی، این مطلب نادرست است و ما همین الان برخی از تحقیقات مهم برای یافتن مفاهیم و احکام صحیح در این باره را مورد بررسی قرار دادیم. فوریه به سبب عدم دقت در بیان برخی از احکام، به‌ویژه از طرف داربو، در معرض انتقاد بوده است و این امر تا حدودی علت عدم محبوبیت طولانی مدت او در بین فرانسویان بود. ولی فوریه سزاوار احترام است نه به این دلیل که چند قضیه اثبات کرد و احکام بی‌نقصی بیان کرد، بلکه از این نظر که برنامه‌ای بلندمدت را در ریاضیات به راه انداخت. در هر بخشی از نظریه تحلیلی گرما می‌توان نمونه‌هایی از عباراتی یافت که نامعقول انگاشته می‌شدند و شاید به‌نظر، پیشگویانه نیز برسند. یک نمونه ذکر می‌کنیم ([۲] فصل III ص. ۲۵). فوریه از دستورهای انتگرالی ضرایب، دستور

$$F(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} F(\alpha) d\alpha \left(\frac{1}{2} + \sum_i \cos i(x - \alpha) \right)$$

را نتیجه می‌گیرد. داربو در حاشیه این مطلب می‌نویسد که پرانتزها هیچ معنایی ندارند. اما توضیحات فوریه از این قرار است:

«عبارت (...) تابعی از x و α را نمایش می‌دهد به‌طوری که اگر در تابعی مانند $F(\alpha)$ ضرب شود و اگر بعد از نوشتن $d\alpha$ ، بین حدود $\alpha = -\pi$ و $\alpha = \pi$ از آن انتگرال گرفته شود، تابع داده شده $F(\alpha)$ به تابع مشابهی از x ، یعنی ضرب‌شده آن در نصف محیط دایره، تغییر خواهد یافت.»

به نظر من، آشکار است که فوریه فهمی شهودی از اندازه دیراک و طریقه به‌کار بستن آن داشته است (به‌ویژه، در بخش‌های جلوتر رساله، در قسمت مشتق و نمایش به‌صورت انتگرال فوریه). گفتنی است که اولین

نوشته لوران شوارتس^۱ دربارهٔ نظریهٔ توزیع‌ها، مقاله‌ای بود (۱۹۴۵) با عنوان «تعمیم مفاهیم تابع، مشتق، تبدیل فوریه، و کاربردهای ریاضیات در فیزیک» [۱۰]. بنابراین از یک دیدگاه می‌توان گفت که نظریهٔ توزیع‌ها جزء میراث فوریه است.

در باب مشتقات جدید و جالب‌تر سری‌های فوریه به اختصار سخن خواهیم گفت. تبدیل سریع فوریه که کولی^۲ و تاکی^۳ (۱۹۶۵) آن را «الگوریتمی برای محاسبهٔ ماشینی سری‌های فوریهٔ مختلط» معرفی کرده بودند، کاربردهای ارزشمندی در همهٔ بخش‌های علوم از اختر فیزیک گرفته تا زیست‌شناسی یافته است. موجک‌های ایو میر^۴ از کارهای فیزیکدانان و مهندسين سرچشمه گرفت و بلافاصله بستر مشترکی برای متخصصان شاخه‌های مختلف علوم و صنایع ایجاد کرد. این موضوع کاملاً شناخته‌شده و همچنان در حال پیشرفت است^۵ و بازگو کردن آن بی‌مورد است. نکتهٔ اصلی، تا جایی که به میراث فوریه ربط دارد، این است که FFT [تبدیل سریع فوریه] و موجک‌ها گواه آن هستند که فلسفهٔ فوریه، که در بالا آمد، دوباره رواجی چشمگیر خواهد یافت.

در ابتدا اشاره کردیم که در اینجا خود را به میراث ریاضی فوریه و، به بیان دقیق‌تر، سری‌های فوریه محدود خواهیم کرد. عملاً هم تعداد زیادی از موارد جذاب در تاریخ سری‌های فوریه را انتخاب کردیم و ذکری از هیچ مسئلهٔ حل‌نشده‌ای به میان نیاوردیم. ترتیب سخنرانی من دقیقاً این نبود. لیکن اکنون خوش‌تر دارم با تأمل در باب گذشته، و نه با ارائهٔ فهرستی شخصی از مسائل حل‌نشده، سهمی در «چشم‌اندازهایی در آنالیز» داشته باشم. گذشتهٔ ریاضیات، پرمایه است و مسیری که گذشته از طریق آن در جدیدترین تحقیقات ریاضی طنین‌انداز می‌شود، جزء «چشم‌اندازهایی در آنالیز» به‌شمار می‌آید.

مراجع

- [1] Carleson, L., On convergence and growth of partial sums of Fourier series, *Acta Math.* **116** (1966), 135-157.
- [2] Fourier J., *Théorie Analytique de la Chaleur*, Firmin Didot Père et Fils, Paris, 1822.
- [3] Jaffard, S., Meyer, Y., Ryan, R. D., *Wavelets*, SIAM, Philadelphia, 2001.
- [4] Kahane, J.-P., Sets of uniqueness and sets of multiplicity, *Banach Center Publications*, **56** (2002), 55-68.
- [5] Kahane, J.-P., Lemarié-Rieusset, P.G., *Séries de Fourier et Ondelettes*, Cassini, Paris, 1998.
- [6] Katznelson, Y., *An Introduction to Harmonic Analysis*, 3rd. edn., Cambridge University Press, Cambridge, 2004.
- [7] Kolmogorov, A. N., *Selected Works*, vol. I, Kluwer, Netherlands, 1991.

^۵ برای نمونه [۳] را نگاه کنید.

- [8] Konyagin, S. V., On divergence of trigonometric Fourier series everywhere, *C. R. Acad. Sci. Paris, Ser. I, Math.*, **329** (1999), 693-697.
- [9] Riemann, B., *Über die Darstellbarkeit Einer Function Durch Eine Trigonometrische Reihe*, Habilitationsschrift, Universität Göttingen, 1854.
- [10] Schwartz, L., Généralisation de la notion de fonction, de dérivation, de transformation de Fourier, et applications mathématique et physiques, *Annales Univ. Grenoble*, **21** (1945), 57-74.
- [11] Zygmund, A., *Trigonometric Series*, 3rd. edn., Cambridge University Press, Cambridge, 2002.

سعید مقصودی: زنجان-دانشگاه زنجان-گروه ریاضی

رایانامه: s_maghsodi@znu.ac.ir

مقایسه‌ای از رویکردهای نیمن - پیرسون، بیز و کم- بیشینه در آزمون فرضیه‌های فازی

عباس پرچمی، سید محمود طاهری، و ناصر رضا ارقامی

چکیده

در این مقاله، سه رویکرد نیمن-پیرسون، بیز و کم-بیشینه را به مسئله آزمون فرضیه‌های فازی (در حالتی که داده‌ها معمولی هستند) مطرح و مقایسه می‌کنیم. در هر سه رویکرد، اگر فرضیه‌ها دقیق در نظر گرفته شوند، آزمون مربوط معادل با آزمون‌های نیمن-پیرسون، بیز و کم-بیشینه برای فرضیه‌های دقیق خواهد بود. یک مثال کاربردی نیز در حوزه کشاورزی ارائه می‌شود و به وسیله این سه رویکرد مورد بررسی قرار می‌گیرد.

۱. فرضیه‌های فازی: انگیزه‌ها و کاربردها

در رویکرد کلاسیک به آمار، فرضیه‌های آماری به وسیله مجموعه‌های معمولی (دقیق) بیان می‌شوند. اما در عمل، به دلایل مختلف، با مواردی مواجه می‌شویم که صورت‌بندی فرضیه‌ها با استفاده از مجموعه‌های فازی مناسب‌تر است. به مثال‌های زیر توجه کنید.

مثال ۱.۱. یک آزمون رایج در بسیاری از مسائل کاربردی، آزمون مقایسه میانگین‌های دو جامعه مستقل است. مثلاً در حوزه علوم تربیتی: مقایسه دو روش آموزش از لحاظ میزان یادگیری، در حوزه کشاورزی: مقایسه رشد یک نوع گیاه بر اساس دو روش کشت، در حوزه مهندسی سازه: مقایسه مقاومت دو روش فناوری بتن. در اغلب این مسائل، برابری دقیق میانگین‌ها مورد نظر نیست، بلکه کافی است بررسی کنیم که آیا میانگین‌های دو جامعه «تقریباً برابر» هستند یا خیر؟ لذا مناسب‌تر و طبیعی‌تر این است که

عبارات و کلمات کلیدی. فرضیه فازی؛ آزمون بیز؛ آزمون کم-بیشینه؛ لم نیمن-پیرسون؛ احتمال پیشامد فازی.

فرضیه‌های مورد آزمون به این صورت تنظیم شوند:

$$\begin{cases} \mu_1 \text{ و } \mu_2 \text{ تقریباً برابر هستند : } \tilde{H}_0 \\ \mu_1 \text{ و } \mu_2 \text{ تفاوت بسیار دارند : } \tilde{H}_1 \end{cases}$$

در اینجا می‌توانیم مفاهیم «تقریباً برابر» و «تفاوت بسیار» را بر اساس مجموعه‌های فازی مدل‌سازی و در تنظیم فرضیه‌ها از آنها استفاده کنیم.

مثال ۲.۱. گاهی مناسب‌تر و واقعی‌تر است که برای صورت‌بندی فرضیه‌های مورد آزمون، از واژه‌های کلامی به جای ارائه مقادیر عددی دقیق استفاده کنیم. برای مثال، فرض کنید ادعایی از طرف نهاد سازنده یک موشک زمین به زمین، مبنی بر کم خطا بودن موشک‌های تولیدی در اصابت به هدف، مطرح شده است. در این مورد، اگر میانگین خطا در برخورد به هدف را با متر اندازه‌گیری کنیم، فرضیه‌ها مثلاً این گونه بیان می‌شوند:

$$\begin{cases} H_0 : \mu \leq 100 \\ H_1 : \mu > 100 \end{cases}$$

که در آن، μ متوسط فاصله محل اصابت موشک‌های تولیدی تا هدف است. اما روشن است که هرچه موشک نزدیک‌تر به هدف اصابت کند، تأثیر و شدت تخریب آن بیشتر است و هرچه دورتر اصابت کند، تأثیر کمتری دارد. لذا مناسب‌تر است که فرضیه‌های مورد آزمون، به کمک مجموعه‌های فازی به صورت زیر طراحی شوند که در آنها به جای عدد دقیق ۱۰۰، از واژه‌های کوچک و بزرگ استفاده شده است:

$$\begin{cases} \mu \text{ کوچک است : } \tilde{H}_0 \\ \mu \text{ بزرگ است : } \tilde{H}_1 \end{cases}$$

مثال ۳.۱. گاهی کافی است مطمئن شویم که در یک جامعه یا سازوکار تحت مطالعه، وضعیت جدید از وضعیت قبلی به اندازه‌ای قابل قبول بهتر شده است. در این موارد فرضیه‌های فازی، فرضیه‌هایی معقول‌تر و طبیعی‌تر هستند. مثلاً جامعه‌ای را در نظر بگیرید که نرخ شیوع یک بیماری خاص در آن حدوداً ۲۰٪ بوده است. پس از انجام یک واکسیناسیون همگانی، می‌خواهیم بررسی کنیم که آیا این نسبت «بسیار کم» شده است یا خیر. در این مورد، طبیعی است که نخست واژه‌های کلامی «حدوداً ۲۰٪» و «بسیار کم» را با دو مجموعه فازی مدل‌سازی و آن‌گاه فرضیه‌های فازی زیر را آزمون کنیم:

$$\begin{cases} \theta \text{ حدوداً } 20\% \text{ است : } \tilde{H}_0 \\ \theta \text{ بسیار کوچک است : } \tilde{H}_1 \end{cases}$$

مثال ۴.۱. گاهی متغیر مورد بررسی و مشاهدات مربوط به آن، مبهم است. در این موارد نیز بسا بهتر باشد که فرضیه‌های مورد آزمون را نادقیق در نظر بگیریم. مثلاً اگر بخواهیم در مورد «رضایت شغلی» در میان افراد یک جامعه بررسی انجام دهیم، داده‌هایی (مشاهده‌هایی) به صورت زیر خواهیم داشت: «رضایت متوسط»، «رضایت نسبتاً زیاد»، «رضایت کم»، «رضایت خیلی کم»، در این مورد، طبیعی‌تر است که فرضیه‌های مورد آزمون را نیز فازی در نظر بگیریم. مثلاً به جای آزمون فرضیه

متوسط رضایت شغلی بیشتر از ۸۲/۶ است : H

این فرضیه فازی را آزمون کنیم که

متوسط رضایت شغلی نسبتاً زیاد است : $\tilde{H}$

مثال‌های بالا نمونه‌هایی از مسائلی هستند که در آنها تنظیم فرضیه‌ها به صورت مجموعه‌های فازی، موجه یا مطلوب است. از این دست مسائل بسیار است. به منظور طولانی نشدن مقاله، از آوردن مثال‌های دیگر چشم‌پوشی می‌کنیم. در این مقاله، نخست توضیح می‌دهیم که چگونه می‌توان فرضیه‌های نادقیق را در قالب مجموعه‌های فازی صورت‌بندی کرد و سپس سه رویکرد به آزمون فرضیه‌های فازی را بررسی و مقایسه می‌کنیم.

۲. فرضیه‌های فازی: مفاهیم و تعاریف‌ها

در این بخش، چند تعریف و رابطه مقدماتی را درباره فرضیه‌های فازی مرور می‌کنیم. در ادامه فرض می‌کنیم که Θ فضای پارامتر مورد بحث است.

تعریف ۱.۲. هر گزاره فازی به شکل « θ به صورت تابع H است : $\tilde{H}$ » است که در آن، $H : \Theta \rightarrow [0, 1]$ یک فرضیه فازی نامیده می‌شود. عبارت « θ به صورت تابع H است» دلالت بر این دارد که هر مقدار خاص مانند θ ، با درجه $H(\theta)$ متعلق به فضای پارامتر است.

به دیگر سخن، هر ادعا درباره پارامتر مجهول جامعه که به وسیله یک مجموعه فازی بیان گردد، یک فرضیه (پارامتری) فازی است [۱]. توجه کنید که با جایگذاری تابع نشانگر به جای تابع عضویت، تعریف فوق تبدیل به تعریف فرضیه معمولی در آمار کلاسیک می‌شود و بنابراین فرضیه‌های معمولی (غیرفازی)، حالت خاصی از فرضیه‌های فازی هستند. برای مثال، فرضیه غیرفازی $\theta = \theta_0 : H$ را می‌توان یک فرضیه فازی با تابع عضویت $I(\theta = \theta_0)$ در نظر گرفت.

مثال ۲.۲. فرض کنید θ نسبت (مجهول) افراد بزرگسال مرد ساکن در استان گلستان باشد که مبتلا به سرطان روده هستند. هر یک از فرضیه‌های زیر که تابع عضویت مربوط به آنها در ادامه درج شده است، یک فرضیه فازی در مورد مقدار θ است (شکل ۱ را ببینید):

$\tilde{H}_1$: کوچک است (نسبت مبتلایان کم است)

$\tilde{H}_2$: بسیار کوچک است (نسبت مبتلایان بسیار کم است)

$\tilde{H}_3$: θ تقریباً ۰/۱۵ است

$$H_1(\theta) = \begin{cases} \frac{0.3-\theta}{0.3} & 0 \leq \theta < 0.3 \\ 0 & 0.3 \leq \theta \leq 1 \end{cases}$$

$$H_2(\theta) = \begin{cases} \left(\frac{0.3-\theta}{0.3}\right)^4 & 0 \leq \theta < 0.3 \\ 0 & 0.3 \leq \theta \leq 1 \end{cases}$$

$$H_3(\theta) = \begin{cases} 0 & 0 \leq \theta < 0.10 \\ \frac{\theta-0.10}{0.05} & 0.10 \leq \theta < 0.15 \\ \frac{0.20-\theta}{0.05} & 0.15 \leq \theta < 0.20 \\ 0 & 0.20 \leq \theta \end{cases}$$

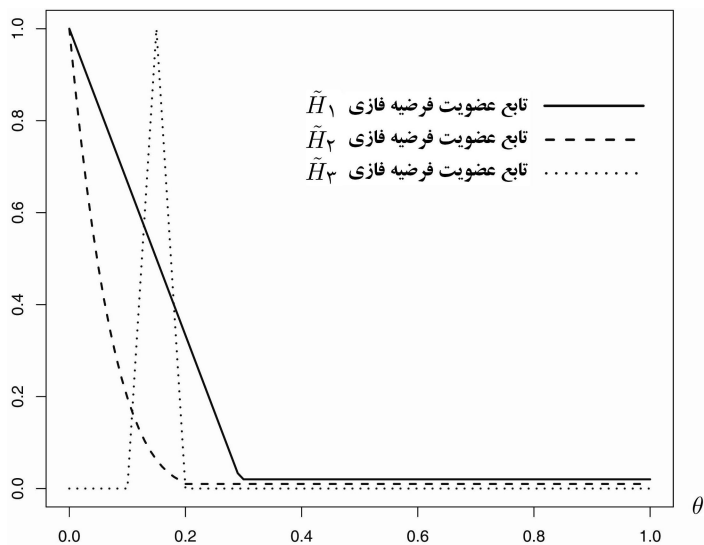
در همین مثال، فرضیه‌های زیر را در نظر بگیرید:

$$\begin{cases} \theta \text{ کوچکتر یا مساوی } 0.20 \text{ است: } H_0 \\ \theta \text{ بزرگتر از } 0.20 \text{ است: } H_1 \end{cases}$$

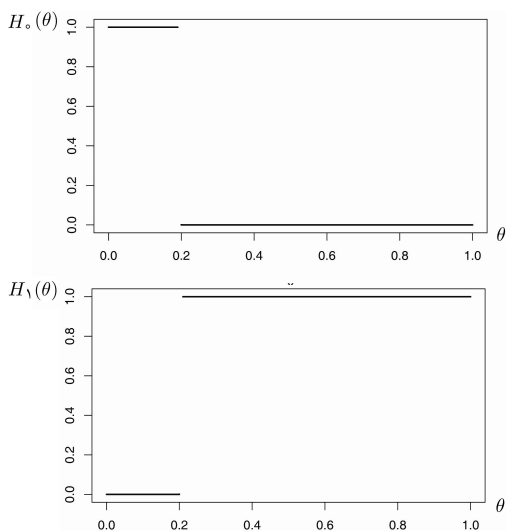
هر دو فرضیه بالا، فرضیه‌های دقیق (غیرفازی) هستند و البته می‌توان این فرضیه‌ها را حالت‌های خاص از فرضیه‌های فازی با توابع عضویت زیر در نظر گرفت (شکل ۲ را ببینید):

$$H_0(\theta) = \begin{cases} 1 & 0 \leq \theta \leq 0.20 \\ 0 & 0.20 < \theta \leq 1 \end{cases} \quad H_1(\theta) = \begin{cases} 0 & 0 \leq \theta \leq 0.20 \\ 1 & 0.20 < \theta \leq 1 \end{cases}$$

مثال ۳.۲. میزان جذب عنصر سمی کادمیوم توسط ریشه گیاه تربچه از یک خاک آلوده به $CdNO_3$ به عنوان متغیر تصادفی در یک تحقیق کشاورزی در نظر گرفته شده است. بر اساس تحقیقات پایس و بنتون [۱۳]، اگر میزان جذب کادمیوم توسط گیاه (برحسب میلی‌گرم در هر کیلوگرم ماده خشک) در بازه $[0.05, 0.2]$ باشد، آن‌گاه جذب ایده‌آلی توسط گیاه صورت گرفته است. همچنین حداکثر میزان جذب



شکل ۱. نمودار توابع عضویت سه فرضیه فازی در مثال ۲.۲



شکل ۲. نمودار توابع عضویت فرضیه‌های دقیق H_0 (نمودار بالایی) و H_1 (نمودار پایینی) در مثال ۲.۲

قابل قبول کادمیوم برای گیاه تربچه برابر ۳ میلی‌گرم در هر کیلوگرم ماده خشک است. فرض کنید ادعایی

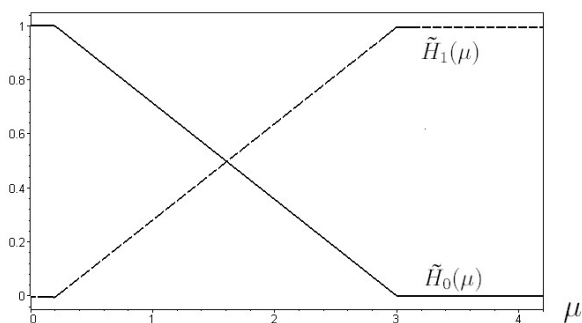
دربارهٔ تریچه‌های کاشته‌شده در یک تحقیق گلخانه‌ای مطرح شده است مبنی بر اینکه میانگین جذب کادمیوم (μ) منطبق بر جذب مناسب معرفی‌شده توسط پایس و بنتون [۱۳] است. برای بررسی صحت و سقم این ادعا، فرضیه‌ها به شکل زیر صورت‌بندی شده‌اند:

$$(۱.۲) \quad \begin{cases} \tilde{H}_0: \text{میانگین جذب کادمیوم کم است} \\ \tilde{H}_1: \text{میانگین جذب کادمیوم زیاد است} \end{cases}$$

در اینجا مفاهیم نادقیق «کم» و «زیاد» به ترتیب به وسیلهٔ

$$H_0(\mu) = \begin{cases} 1 & 0 \leq \mu < 0.2 \\ \frac{3-\mu}{2.8} & 0.2 \leq \mu < 3 \\ 0 & \mu \geq 3 \end{cases}$$

و $H_1(\mu) = 1 - H_0(\mu)$ مشخص شده‌اند (شکل ۳). از آنجا که کادمیوم عنصری سمی است، می‌توان



شکل ۳. توابع عضویت فرضیه‌های فازی در مثال ۳.۲

ادعا کرد که جذب کمتر آن برای گیاه مناسب‌تر بوده و لذا در معرفی $H_0(\mu)$ ، میزان عضویت اعداد $[0, 0.2]$ به مجموعهٔ فازی جذب مناسب کادمیوم، برابر ۱ در نظر گرفته شده است (شکل ۳ را ببینید). توجه داشته باشید که اگر به منظور صورت‌بندی این ادعا، از فرضیه‌های معمولی و غیرفازی

$$(۲.۲) \quad \begin{cases} H_0: \mu \leq 0.2 + 2/8t \\ H_1: \mu > 0.2 + 2/8t \end{cases}$$

استفاده کنیم، آنگاه به ازای مقدارهای مختلف $0 \leq t \leq 1$ می‌توان فرضیه‌های H_0 و H_1 زیادی را مطرح ساخت که لزوماً آزمون آنها منتهی به جواب یکسان نمی‌شود و امکان متناقض شدن نتایج آزمون فرضیه‌های دقیق (۲.۲) به ازای مقدارهای مختلف $t \in [0, 1]$ وجود دارد. دلیل این امر، ابهام موجود در تعیین میزان

مناسب مقدار جذب توسط پایس و بنتون است و لذا ادعای مبهم آنها نیز قابلیت صورت‌بندی دقیق در قالب فرضیه‌های دقیق را ندارد. بنابراین برای رفع این مشکل، می‌توان به کمک مجموعه‌های فازی، مسئله را به وسیله فرضیه‌های فازی (۱.۲) صورت‌بندی کرد. این موضوع از مزیت‌های به کارگیری فرضیه‌های فازی در برخی از مسائل کاربردی و واقعی به شمار می‌آید.

تعریف ۴.۲. ([۲۰]) فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال $f(x; \theta)$ و انتگرال تابع عضویت فرضیه $\tilde{H}$ متناهی باشد. «تابع چگالی احتمال وزنی X تحت فرضیه فازی $\tilde{H}$ » را با نماد $f(x; \tilde{H})$ نشان می‌دهیم و به صورت

$$f(x; \tilde{H}) = \int_{\theta} H^*(\theta) f(x; \theta) d\theta$$

تعریف می‌کنیم که در آن، H^* تابع عضویت نرمال‌شده فرضیه $\tilde{H}$ ، عبارت است از

$$H^*(\theta) = \frac{H(\theta)}{\int_{\theta} H(\theta) d\theta}.$$

در حالت گسسته به جای انتگرال از مجموع‌یابی استفاده می‌شود.

توجه داشته باشید که در حالت پیوسته، تابع عضویت نرمال‌شده لزوماً یک تابع عضویت نیست، زیرا اگر $\int_{\theta} H(\theta) d\theta < 1$ ، آن‌گاه ممکن است به‌ازای برخی از θ ها تابع عضویت نرمال‌شده، $H^*(\theta)$ ، مقادارهایی بزرگتر از یک اختیار کند.

تابع $f(x; \tilde{H})$ یک تابع چگالی احتمال است، زیرا اولاً $f(x; \tilde{H})$ نامنفی است و ثانیاً

$$\begin{aligned} \int_x f(x; \tilde{H}) dx &= \int_x \int_{\theta} H^*(\theta) f(x; \theta) d\theta dx \\ &= \int_{\theta} H^*(\theta) \left(\int_x f(x; \theta) dx \right) d\theta \\ &= \int_{\theta} H^*(\theta) d\theta = 1. \end{aligned}$$

به‌علاوه اگر H فرضیه معمولی (غیرفازی) « $H : \theta = \theta_0$ » باشد، آن‌گاه $f(x; H) = f(x; \theta_0)$.

تعریف ۵.۲. فرض‌های تعریف ۴.۲ را در نظر بگیرید. در این صورت احتمال پیشامد A تحت فرضیه فازی $\tilde{H}$ با

$$P_{\tilde{H}}(A) = \int_{\theta} H^*(\theta) P_{\theta}(A) d\theta$$

مقایسه‌ای از رویکردهای نیمن - پیرسون، بیز و کم- بیشینه در آزمون فرضیه‌های فازی _____ ۵۰

تعریف می‌شود. در حالت گسسته، به جای انتگرال مجموعیابی قرار می‌گیرد.

دقت کنید که با فرض امکان جابه‌جایی ترتیب انتگرال‌گیری، داریم

$$\begin{aligned}\int_{\theta} H^*(\theta) P_{\theta}(A) d\theta &= \int_{\theta} H^*(\theta) \int_A f(x; \theta) dx d\theta \\ &= \int_A \int_{\theta} H^*(\theta) f(x; \theta) d\theta dx \\ &= \int_A f(x; \tilde{H}) dx.\end{aligned}$$

اگر H فرضیه مرکب و غیرفازی « $H : \theta \in \Theta'$ » و Θ' متناهی باشد، آنگاه می‌توان $P_H(A)$ را متوسط $P_{\theta}(A)$ روی $\theta \in \Theta'$ دانست، زیرا بنابر تعریف ۵.۲ در این حالت داریم

$$\begin{aligned}P_H(A) &= \frac{\int_{\theta} I_{\Theta'}(\theta) P_{\theta}(A) d\theta}{\int_{\Theta'} d\theta} \\ &= \frac{\int_{\Theta'} P_{\theta}(A) d\theta}{\int_{\Theta'} d\theta} \\ &= E_{\Theta'} P_{\theta}(A)\end{aligned}$$

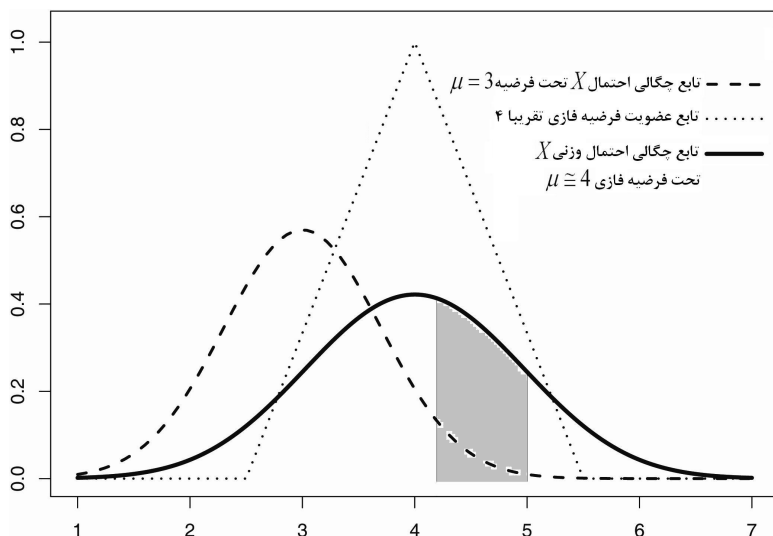
که در اینجا، منظور از $E_{\Theta'}$ امید ریاضی نسبت به توزیع یکنواخت روی مجموعه Θ' است. همچنین در حالت خاص‌تر اگر H فرضیه ساده غیرفازی « $H : \theta = \theta_0$ » باشد، آنگاه $P_H(A) = P_{\theta_0}(A)$.

مثال ۶.۲. فرض کنید X دارای توزیع نرمال با میانگین مجهول μ و واریانس $1/5$ باشد. همچنین فرض کنید مفهوم «تقریباً ۴» به وسیله تابع عضویت زیر ارائه شده باشد:

$$H(\mu) = \begin{cases} \frac{\mu - 2/5}{1/5} & 2/5 < \mu \leq 4 \\ \frac{5/5 - \mu}{1/5} & 4 < \mu \leq 5/5 \\ 0 & \text{اگر نه} \end{cases}$$

طبق تعریف ۴.۲، تابع چگالی احتمال وزنی X تحت فرضیه فازی « $\tilde{H} : \mu \cong ۴$ » به صورت زیر است (شکل ۴ را ببینید):

$$\begin{aligned} f(x; \mu \cong ۴) &= \int_{\mu} H^*(\mu) f(x; \mu) d\mu \\ &= \frac{۲}{۳} \int_{\mu} H(\mu) \frac{۱}{\sqrt{۱/۴}\pi} e^{-\frac{(x-\mu)^2}{1/4}} d\mu \\ &= \frac{۲}{۳\sqrt{1/4}\pi} \int_{2/5}^4 \frac{\mu - 2/5}{1/5} e^{-\frac{(x-\mu)^2}{1/4}} d\mu \\ &\quad + \frac{۲}{۳\sqrt{1/4}\pi} \int_4^{5/5} \frac{5/5 - \mu}{1/5} e^{-\frac{(x-\mu)^2}{1/4}} d\mu, \quad x \in \mathbb{R}. \end{aligned}$$



شکل ۴. تابع چگالی احتمال وزنی نرمال تحت فرضیه فازی $\mu \cong ۴$ و تابع چگالی احتمال نرمال تحت فرضیه $\mu = ۳$ در مثال ۶.۲

حال بر اساس تعریف ۵.۲ و به کمک تابع چگالی احتمال وزنی X ، مثلاً می توان احتمال این پیشامد را که متغیر تصادفی X بین اعداد $۴/۲$ و ۵ باشد تحت فرضیه فازی « $\tilde{H} : \mu \cong ۴$ » به صورت زیر محاسبه

کرد:

$$\begin{aligned} P_{\tilde{H}}(4/2 < X < 5) &= \int_{4/2}^5 f(x; \tilde{H}) dx \\ &= \frac{2}{3\sqrt{1/4}\pi} \int_{4/2}^5 \int_{2/5}^4 \frac{\mu - 2/5}{1/5} e^{-\frac{(x-\mu)^2}{1/4}} d\mu dx \\ &\quad + \frac{2}{3\sqrt{1/4}\pi} \int_{4/2}^5 \int_4^{5/5} \frac{5/5 - \mu}{1/5} e^{-\frac{(x-\mu)^2}{1/4}} d\mu dx \\ &= 0/255 \end{aligned}$$

که برابر با سطح ناحیه خاکستری در شکل ۴ است.

۳. آزمون نیمن-پیرسون

در ابتدای این بخش، برخی از تعریف‌های مربوط به آزمون فرضیه‌های فازی را بر اساس مراجع [۱۸، ۶] مرور می‌کنیم. فرض کنید $X = (X_1, \dots, X_n)$ نمونه‌ای تصادفی از جامعه‌ای با تابع چگالی احتمال یا تابع جرم احتمال $f(x_i; \theta)$ و $x = (x_1, \dots, x_n)$ بردار مقدارهای مشاهده‌شده باشد که در آن، $\theta \in \Theta$ پارامتر مجهول است. می‌خواهیم فرضیه‌های فازی زیر را آزمون کنیم:

$$\begin{cases} \tilde{H}_0: \theta \text{ به صورت تابع } H_0 \text{ است;} \\ \tilde{H}_1: \theta \text{ به صورت تابع } H_1 \text{ است;} \end{cases} \quad (1.3)$$

تعریف ۱.۳. در مسئله آزمون فرضیه‌های فازی (۱.۳) هر تابع Φ (تابعی از مجموعه مقدارهای ممکن نمونه تصادفی به بازه $[0, 1]$) به طوری که $\Phi(x)$ احتمال رد $\tilde{H}_0$ به ازای مقدار مشاهده‌شده x را نشان دهد، یک تابع آزمون نامیده می‌شود.

از این پس، آزمونی که تابع آزمون آن $\Phi(X)$ باشد را به اختصار «آزمون Φ » می‌نامیم.

تعریف ۲.۳. تابع توان آزمون Φ به صورت

$$\beta_{\Phi}(\theta) = P_{\theta} \{d(X) = a_1\} = E_{\theta} [\Phi(X)]$$

تعریف می‌شود که در آن، $d(X) = a_1$ به معنی رد H_0 است.

تعریف ۳.۳. فرض کنید $\Phi(X)$ یک تابع آزمون برای آزمون فرضیه‌های فازی (۱.۳) باشد و همچنین $\int_{\Theta} H_0(\theta) d\theta < \infty$ و $\int_{\Theta} H_1(\theta) d\theta < \infty$. احتمال ارتکاب خطای نوع اول آزمون Φ عبارت

است از

$$\begin{aligned}\alpha_{\Phi} &= P_{\tilde{H}_0} \{d(X) = a_1\} \\ &= \int_{\Theta} H_0^*(\theta) P_{\theta} \{d(X) = a_1\} d\theta \\ &= \int_{\Theta} H_0^*(\theta) E_{\theta} \Phi(X) d\theta\end{aligned}$$

و احتمال ارتکاب خطای نوع دوم، عبارت است از

$$\begin{aligned}\beta_{\Phi} &= P_{\tilde{H}_1} \{d(X) = a_0\} \\ &= \int_{\Theta} H_1^*(\theta) P_{\theta} \{d(X) = a_0\} d\theta \\ &= \int_{\Theta} H_1^*(\theta) [1 - E_{\theta} \Phi(X)] d\theta.\end{aligned}$$

(در حالت گسسته، از مجموع یابی به جای انتگرال استفاده می‌شود).

اگر به تسامح، تابع عضویت فرضیه فازی $\tilde{H}_0$ را به چشم تابع چگالی احتمال θ بنگریم، آنگاه α_{Φ} را می‌توان میانگین وزنی $E_{\theta} \Phi(X)$ تلقی نمود. تفسیری مشابه را می‌توان برای β_{Φ} بیان کرد. در حالت آزمون فرضیه ساده در مقابل ساده (مانند آنچه در لم نیمن-پیرسون معمولی وجود دارد)، α_{Φ} و β_{Φ} به احتمال ارتکاب خطاهای نوع اول و دوم معمولی تبدیل می‌شوند.

تعریف ۴.۳. آزمون Φ را یک آزمون در سطح α (معناداری) می‌نامیم اگر $\alpha_{\Phi} \leq \alpha \in [0, 1]$. همچنین α_{Φ} را اندازه آزمون Φ می‌گوییم.

تعریف ۵.۳. آزمون در سطح α ی Φ را بهترین آزمون در سطح α می‌نامیم اگر به ازای هر آزمون در سطح α ی Φ^* ، داشته باشیم $\beta_{\Phi} \leq \beta_{\Phi^*}$.

قضیه ۶.۳. (لم نیمن-پیرسون تعمیم یافته [۱۸]). فرض کنید $X = (X_1, \dots, X_n)$ نمونه‌ای تصادفی از جامعه‌ای با تابع چگالی احتمال یا تابع جرم احتمال $f(x_i; \theta)$ و $x = (x_1, \dots, x_n)$ بردار مقدارهای مشاهده شده باشد که در آن، $\theta \in \Theta$ پارامتر مجهول است. برای آزمون فرضیه‌های فازی

$$\begin{cases} \tilde{H}_0: \theta \text{ به صورت تابع } H_0 \text{ است;} \\ \tilde{H}_1: \theta \text{ به صورت تابع } H_1 \text{ است;} \end{cases}$$

الف) هر آزمون به صورت

$$\Phi(x) = \begin{cases} 1 & \frac{\int_{\Theta} H_1(\theta)f(x;\theta)d\theta}{\int_{\Theta} H_0(\theta)f(x;\theta)d\theta} > k \\ 0 & \frac{\int_{\Theta} H_1(\theta)f(x;\theta)d\theta}{\int_{\Theta} H_0(\theta)f(x;\theta)d\theta} < k \end{cases} \quad (2.3)$$

برای هر $k \geq 0$ ، بهترین آزمون در اندازه خود است. به علاوه به ازای $k = \infty$ ، آزمون

$$\Phi(x) = \begin{cases} 1 & \int_{\Theta} H_0(\theta)f(x;\theta)d\theta = 0 \\ 0 & \int_{\Theta} H_0(\theta)f(x;\theta)d\theta > 0 \end{cases} \quad (3.3)$$

بهترین آزمون در اندازه صفر است.

ب) برای هر $1 \leq \alpha \leq 0$ آزمونی به شکل (۲.۳) یا (۳.۳) وجود دارد به طوری که در آن، $\alpha_{\Phi} = \alpha$.

۴. آزمون کم-بیشینه

در ابتدای این بخش، مفاهیم مقدماتی نظریه تصمیم را به کوتاهی مرور می‌کنیم [۱۵، ۱۶]. فرض کنیم $\mathbb{R} \rightarrow \Theta \times \mathcal{A} : \mathcal{L}(\theta, a)$ تابع زیان باشد که در آن، $\mathcal{A}$ فضای اعمال است. تابع $\mathcal{L}(\theta, a)$ نشان‌دهنده مقدار زیان است وقتی مقدار پارامتر θ است و ما تصمیم a را اتخاذ نموده‌ایم. فرض کنیم $\mathcal{D}$ رده توابع تصمیم (توابعی از $\mathbb{R}^n$ به $\mathcal{A}$) باشد. تابع $\{0\} \cup \mathbb{R}^+ \rightarrow \Theta \times \mathcal{D} : \mathcal{R}$ را که به صورت $\mathcal{R}(\theta, d) = E_{\theta}\mathcal{L}(\theta, d(X))$ تعریف می‌شود، تابع ریسک می‌نامیم. بر اساس این تعریف‌ها، رویکرد کم-بیشینه، تصمیم $d_m \in \mathcal{D}$ را به گونه‌ای انتخاب می‌کند که

$$\max_{\theta} \mathcal{R}(\theta, d_m) \leq \max_{\theta} \mathcal{R}(\theta, d).$$

چنین تصمیمی، در صورت وجود، تصمیم کم-بیشینه (مینیمکس) نامیده می‌شود. مسئله آزمون فرضیه « $H_0 : \theta = \theta_0$ » در مقابل فرضیه « $H_1 : \theta = \theta_1$ » را می‌توان حالتی خاص از بحث فوق در نظر گرفت که در آن، فضای اعمال متشکل از دو عمل a_0 (پذیرش H_0) و a_1 (پذیرش H_1) است. بنابراین در این حالت، $\mathcal{A} = \{a_0, a_1\}$ و $\Theta = \{\theta_0, \theta_1\}$. تابع زیان نیز در مسئله آزمون فرضیه معمولاً به صورت

$$\begin{cases} \mathcal{L}(H_1, a) = C_1 & a = a_0, \quad C_1 > 0 \\ \mathcal{L}(H_0, a) = C_0 & a = a_1, \quad C_0 > 0 \\ \mathcal{L}(H_0, a) = 0 & a = a_0 \\ \mathcal{L}(H_1, a) = 0 & a = a_1 \end{cases} \quad (1.4)$$

در نظر گرفته می‌شود که در آن، C_0 زیان ناشی از ارتکاب خطای نوع اول و C_1 زیان ناشی از ارتکاب خطای نوع دوم است.

در مسئله آزمون فرضیه‌های فازی نیز حالتی را در نظر می‌گیریم که فضای اعمال متشکل از تنها دو نقطه a_0 و a_1 است که در آن، به ازای $j = 0, 1$ ، عمل a_j نشان‌دهنده پذیرش فرضیه فازی « θ به صورت تابع H_j است: $\tilde{H}_j$ » (و رد فرضیه دیگر) است.

تعریف ۱.۴. در آزمون فرضیه‌های فازی

$$\begin{cases} \tilde{H}_0: \theta \text{ به صورت تابع } H_0 \text{ است} \\ \tilde{H}_1: \theta \text{ به صورت تابع } H_1 \text{ است} \end{cases}$$

تصمیم d_m را تصمیم کم-بیشینه و یا به طور معادل، Φ_m را آزمون کم-بیشینه نامیم اگر به ازای هر $d \in \mathcal{D}$ داشته باشیم

$$\max \left\{ \mathcal{R}(\tilde{H}_0, d_m), \mathcal{R}(\tilde{H}_1, d_m) \right\} \leq \max \left\{ \mathcal{R}(\tilde{H}_0, d), \mathcal{R}(\tilde{H}_1, d) \right\}.$$

قضیه ۲.۴ ([۱۴]). فرض کنید $X = (X_1, \dots, X_n)$ نمونه‌ای تصادفی از جامعه‌ای با تابع چگالی احتمال $f(x_i; \theta)$ و $x = (x_1, \dots, x_n)$ بردار مقدارهای مشاهده شده باشد که در آن، $\theta \in \Theta$ پارامتر مجهول است. در آزمون فرضیه‌های فازی

$$\begin{cases} \tilde{H}_0: \theta \text{ به صورت تابع } H_0 \text{ است} \\ \tilde{H}_1: \theta \text{ به صورت تابع } H_1 \text{ است} \end{cases}$$

با در نظر داشتن تابع زیان

$$\begin{cases} \mathcal{L}(\tilde{H}_1, a) = C_1 & a = a_0, \quad C_1 > 0 \\ \mathcal{L}(\tilde{H}_0, a) = C_0 & a = a_1, \quad C_0 > 0 \\ \mathcal{L}(\tilde{H}_0, a) = 0 & a = a_0 \\ \mathcal{L}(\tilde{H}_1, a) = 0 & a = a_1 \end{cases} \quad (۲.۴)$$

آزمون کم-بیشینه، فرضیه فازی $\tilde{H}_0$ را رد می‌کند اگر و تنها اگر $\lambda(x) \geq k$ که در آن،

$$\lambda(x) = \frac{\int_{\Theta} H_1(\theta) f(x; \theta) d\theta}{\int_{\Theta} H_0(\theta) f(x; \theta) d\theta} \quad (۳.۴)$$

و مقدار ثابت k به گونه‌ای انتخاب می‌شود که

$$C_1 P_{\tilde{H}_1} \left\{ \frac{\int_{\Theta} H_1(\theta) f(X; \theta) d\theta}{\int_{\Theta} H_0(\theta) f(X; \theta) d\theta} < k \right\} = C_0 P_{\tilde{H}_0} \left\{ \frac{\int_{\Theta} H_1(\theta) f(X; \theta) d\theta}{\int_{\Theta} H_0(\theta) f(X; \theta) d\theta} \geq k \right\}. \quad (۴.۴)$$

نتیجه ۳.۴. تحت زبان $1-0$ ، یعنی حالتی که $C_0 = C_1 = 1$ ، آزمون کم- بیشینه، فرضیه فازی $\tilde{H}_0$ را در مقابل $\tilde{H}_1$ رد می‌کند اگر و تنها اگر $\lambda(x) \geq k$ که در آن، ثابت k از رابطه

$$P_{\tilde{H}_1} \left\{ \frac{\int H_1(\theta) f(X; \theta) d\theta}{\int H_0(\theta) f(X; \theta) d\theta} < k \right\} = P_{\tilde{H}_0} \left\{ \frac{\int H_1(\theta) f(X; \theta) d\theta}{\int H_0(\theta) f(X; \theta) d\theta} \geq k \right\},$$

یعنی رابطه $\alpha_{\Phi} = \beta_{\Phi}$ به دست می‌آید.

به کمک قضیه ۶.۳ می‌توان احتمال ارتکاب خطای نوع دوم ($P_{\tilde{H}_1} \{ \tilde{H}_0 \}$) در آزمون فرضیه‌های فازی را با فرض ثابت بودن احتمال ارتکاب خطای نوع اول ($P_{\tilde{H}_1} \{ \tilde{H}_0 \}$) کمینه کرد. اما در قضیه ۲.۴، سطح ثابت α بی برای $P_{\tilde{H}_0} \{ d(X) = a_1 \}$ در آزمون فرضیه‌های فازی در نظر گرفته نمی‌شود و در عوض، آزمون کم- بیشینه با آزاد اختیار کردن احتمالات خطای نوع اول و دوم، مقدار

$$\max \{ \mathcal{R}(\tilde{H}_0, d), \mathcal{R}(\tilde{H}_1, d) \}$$

را کمینه می‌سازد.

ایده اصلی قضیه ۲.۴ جانشینی توابع عضویت فرضیه‌های $\tilde{H}_1$ و $\tilde{H}_0$ به جای تابع‌های نشانگر

$$H_1(\theta) = \begin{cases} 1 & \theta = \theta_1 \\ 0 & \theta = \theta_0 \end{cases} \quad \text{و} \quad H_0(\theta) = \begin{cases} 1 & \theta = \theta_0 \\ 0 & \theta = \theta_1 \end{cases}$$

است. در این حالت داریم

$$\int_{\Theta} H_0(\theta) f(x; \theta) d\theta = f(x; \theta_0), \quad \int_{\Theta} H_1(\theta) f(x; \theta) d\theta = f(x; \theta_1)$$

و در نتیجه با توجه به قضیه ۲.۴، آزمون کم- بیشینه، فرضیه H_0 را رد می‌کند اگر و تنها اگر $\frac{f(x; \theta_1)}{f(x; \theta_0)} \geq k$ که در آن، ثابت k از رابطه زیر به دست می‌آید

$$C_1 P_{\theta_1} \left\{ \frac{f(X; \theta_1)}{f(X; \theta_0)} < k \right\} = C_0 P_{\theta_0} \left\{ \frac{f(X; \theta_1)}{f(X; \theta_0)} \geq k \right\}.$$

بنابراین آزمون کم- بیشینه برای فرضیه ساده در مقابل ساده که در صفحه ۵۲۱ از [۱۶] مطرح شده است، حالتی خاص از قضیه ۲.۴ است.

فرض کنیم $X = (X_1, \dots, X_n)$ نمونه‌ای تصادفی از جامعه‌ای با تابع جرم احتمال $f(x_i; \theta)$ ، و $x = (x_1, \dots, x_n)$ بردار مقدارهای مشاهده شده باشد. در این حالت اگر خود را محدود به آزمون‌های غیرتصادفی کنیم، آن‌گاه در بسیاری از حالت‌ها k بی نمی‌توان یافت که در رابطه (۴.۴) صدق کند. در این

موارد، آزمون کم-بیشینه برای فرضیه‌های فازی در رده آزمون‌های غیرتصادفی وجود ندارد و باید آن را در رده آزمون‌های تصادفی که به شکل

$$\Phi(x) = \begin{cases} 1 & \frac{\int_{\Theta} H_1(\theta) f(x; \theta) d\theta}{\int_{\Theta} H_0(\theta) f(x; \theta) d\theta} > k \\ \delta & \frac{\int_{\Theta} H_1(\theta) f(x; \theta) d\theta}{\int_{\Theta} H_0(\theta) f(x; \theta) d\theta} = k \\ 0 & \frac{\int_{\Theta} H_1(\theta) f(x; \theta) d\theta}{\int_{\Theta} H_0(\theta) f(x; \theta) d\theta} < k \end{cases} \quad (5.4)$$

هستند، جستجو کرد. لذا با توجه به قضیه ۲.۴، در این حالت باید k به گونه‌ای یافت شود که در رابطه

$$\mathcal{R}_{Rand}(\tilde{H}_0, k) = \mathcal{R}_{Rand}(\tilde{H}_1, k)$$

صدق کند. توجه داشته باشید که در رابطه اخیر، ریسک آزمون تصادفی (۵.۴) را می‌توان برحسب ریسک آزمون غیرتصادفی با ناحیه بحرانی $\frac{\int_{\Theta} H_1(\theta) f(x; \theta) d\theta}{\int_{\Theta} H_0(\theta) f(x; \theta) d\theta} \geq k$ به صورت زیر محاسبه کرد:

$$\begin{aligned} \mathcal{R}_{Rand}(\tilde{H}, k) &= E_{\theta} \mathcal{L}(\tilde{H}, d_m(X)) \\ &= \mathcal{L}(\tilde{H}, a_0) P_{\tilde{H}_1} \{d_m(X) = a_0\} + \mathcal{L}(\tilde{H}, a_1) P_{\tilde{H}_0} \{d_m(X) = a_1\} \\ &= \begin{cases} \mathcal{R}(\tilde{H}_1, k) + (1 - \delta) \left(\mathcal{R}(\tilde{H}_1, \bar{k}) - \mathcal{R}(\tilde{H}_1, k) \right) & \text{تحت فرضیه } \tilde{H}_1 \\ \mathcal{R}(\tilde{H}_0, \bar{k}) + \delta \left(\mathcal{R}(\tilde{H}_0, k) - \mathcal{R}(\tilde{H}_0, \bar{k}) \right) & \text{تحت فرضیه } \tilde{H}_0 \end{cases} \end{aligned}$$

که در آن، $\bar{k}$ اولین (کوچکترین) عضو دنباله تغییرات $\frac{\int_{\Theta} H_1(\theta) f(x; \theta) d\theta}{\int_{\Theta} H_0(\theta) f(x; \theta) d\theta}$ بعد از k است.

۵. آزمون بیز

فرض کنیم $X = (X_1, \dots, X_n)$ نمونه‌ای تصادفی از جامعه‌ای با تابع چگالی احتمال یا تابع جرم احتمال $f(x_i; \theta)$ با بردار مشاهدات $x = (x_1, \dots, x_n)$ و $\theta \in \Theta$ دارای تابع چگالی پیشین $\pi(\theta)$ باشد. طاهری و بهبودیان [۱۹] ناحیه بحرانی آزمون بیز را برای فرضیه‌های فازی

$$\begin{cases} \tilde{H}_0: \theta \text{ به صورت تابع } H_0 \text{ است} \\ \tilde{H}_1: \theta \text{ به صورت تابع } H_1 \text{ است} \end{cases}$$

تحت تابع زیان $1-0$ به صورت

$$\int_{\Theta} H_0(\theta) \pi(\theta|x) d\theta \leq \int_{\Theta} H_1(\theta) \pi(\theta|x) d\theta$$

معرفی نمودند که در آن، $\pi(\theta|x)$ توزیع پسین θ به شرط مشاهدات می‌باشد. این آزمون بیز تحت تابع زیان (۲.۴) نیز قابل تعمیم است:

تعریف ۱.۵. یک آزمون بیز تحت تابع زیان (۲.۴) فرضیه فازی $\tilde{H}_0$ را در مقابل فرضیه فازی $\tilde{H}_1$ رد می‌کند اگر و تنها اگر

$$C_0 \int_{\Theta} H_0(\theta) \pi(\theta|x) d\theta \leq C_1 \int_{\Theta} H_1(\theta) \pi(\theta|x) d\theta. \quad (۱.۵)$$

بنابر تعریف ۱.۵، آزمون بیز فرضیه فازی $\tilde{H}_0$ را در مقابل فرضیه فازی $\tilde{H}_1$ رد می‌کند اگر و تنها اگر $\frac{\int_{\Theta} H_1(\theta) \pi(\theta|x) d\theta}{\int_{\Theta} H_0(\theta) \pi(\theta|x) d\theta}$ بیشتر از نسبت زیان ارتکاب خطای نوع اول به زیان ارتکاب خطای نوع دوم، یعنی $\frac{C_1}{C_0}$ باشد. بنابراین در حالت خاص $C_0 = C_1 = ۱$ ، تعریف ۱.۵ معادل با تعریف آزمون بیز در حالت فرضیه‌های فازی تحت تابع زیان $۱-۰$ می‌شود ([۱۹]، صفحه ۴۰).

در آزمون فرضیه « $H_0: \theta \in \Theta_0$ » در مقابل « $H_1: \theta \in \Theta_1$ »، توابع عضویت فرضیه‌ها به صورت

$$H_0(\theta) = \begin{cases} ۱ & \theta \in \Theta_0 \\ ۰ & \text{اگر نه} \end{cases} \quad \text{و} \quad H_1(\theta) = \begin{cases} ۱ & \theta \in \Theta_1 \\ ۰ & \text{اگر نه} \end{cases}$$

تعریف می‌شوند و بنابراین می‌توان نتیجه گرفت

$$\int_{\Theta} H_0(\theta) \pi(\theta|x) d\theta = \int_{\Theta_0} \pi(\theta|x) d\theta = P(\theta \in \Theta_0|x)$$

و

$$\int_{\Theta} H_1(\theta) \pi(\theta|x) d\theta = \int_{\Theta_1} \pi(\theta|x) d\theta = P(\theta \in \Theta_1|x)$$

که در آن، $\Theta = \Theta_0 \cup \Theta_1$ و $\Theta_0 \cap \Theta_1 = \emptyset$. در نتیجه همانند آزمون فرضیه‌های معمولی به روش بیز، ناحیه رد این آزمون بیز به صورت

$$C_0 P(\theta \in \Theta_0|x) \leq C_1 P(\theta \in \Theta_1|x) \quad (۲.۵)$$

است (صفحه ۱۶۴ از [۸] را ببینید).

۶. مثال کاربردی در کشاورزی

فرض کنید یک محقق کشاورزی با توجه به امکان گسترش یک آفت، می‌خواهد نسبت به سمپاشی یا عدم سمپاشی یک مزرعه تصمیم‌گیری کند. در حالت معمولی، نسبت گیاهانی که دچار آفت می‌شوند حدود ۱۰٪ است، در حالی که این احتمال وجود دارد که به دلیل مسائل محیطی، این نسبت به حدود ۳۰٪

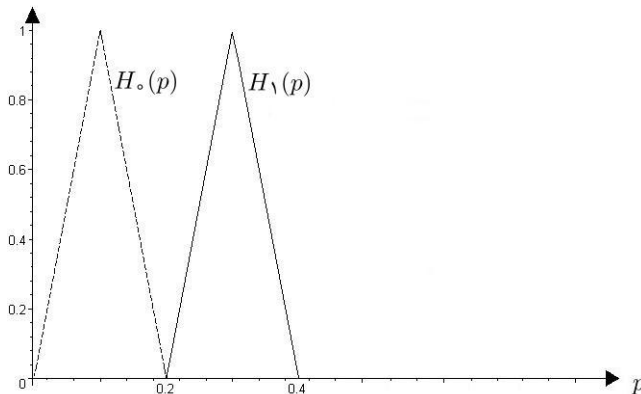
افزایش یافته باشد. بنابراین تمایل او آزمون فرضیه‌های فازی

$$\begin{cases} \tilde{H}_0 : p \cong \circ/1^\circ \text{ (تقریباً } \circ/1^\circ \text{ است)} \\ \tilde{H}_1 : p \cong \circ/3^\circ \text{ (تقریباً } \circ/3^\circ \text{ است)} \end{cases}$$

است که مفاهیم فازی تقریباً $\circ/1^\circ$ و تقریباً $\circ/3^\circ$ به ترتیب به وسیله توابع عضویت زیر تعریف می‌شوند (شکل ۵):

$$H_0(p) = \begin{cases} 1^\circ p & \circ \leq p < \circ/1 \\ 1^\circ - 1^\circ p & \circ/1 \leq p < \circ/2 \\ \circ & \text{اگر نه} \end{cases}, H_1(p) = \begin{cases} 1^\circ p - 2^\circ & \circ/2 \leq p < \circ/3 \\ 4^\circ - 1^\circ p & \circ/3 \leq p < \circ/4 \\ \circ & \text{اگر نه} \end{cases}$$

وی نمونه‌ای شامل ۲۲ گیاه را بررسی می‌کند و متوجه می‌شود که ۴ گیاه آفت دارند. اگر فرض کنیم متغیرهای تصادفی مستقل از هم X_i به ازای $i = 1, \dots, 22$ در نمونه تصادفی دارای توزیع برنولی با احتمال موفقیت p است، آن‌گاه آماره $S = \sum_{i=1}^{22} X_i$ دارای توزیع دوجمله‌ای با پارامترهای $n = 22$ و p خواهد بود و مقدار مشاهده شده آن برابر است با $\sum_{i=1}^{22} x_i = 4$.



شکل ۵. توابع عضویت فرضیه‌های فازی $\tilde{H}_0$ و $\tilde{H}_1$ در مثال کاربردی

الف) آزمون نیمین-پیرسون: با توجه به گسسته بودن توزیع آماره آزمون، طبق قضیه ۶.۳، آزمون

$$\Phi(x) = \begin{cases} 1 & \frac{\int H_1(p)f(x;p)dp}{\int H_0(p)f(x;p)dp} > k \\ \delta & \frac{\int H_1(p)f(x;p)dp}{\int H_0(p)f(x;p)dp} = k \\ \circ & \frac{\int H_1(p)f(x;p)dp}{\int H_0(p)f(x;p)dp} < k \end{cases} \quad (1.6)$$

مقایسه‌ای از رویکردهای نیمن - پیرسون، بیز و کم- بیشینه در آزمون فرضیه‌های فازی _____ ۶۰

پرتوان‌ترین آزمون در سطح $\alpha = 0/10$ است که در آن، مقدارهای ثابت k و $\delta(x)$ از رابطه $\alpha_\Phi = 0/10$ به دست می‌آیند. طبق نتایج جدول ۱، تابعی اکیداً صعودی از آماره $\sum_{i=1}^{22} x_i$ است و لذا می‌توان پرتوان‌ترین آزمون را به صورت زیر بازنویسی کرد:

$$\Phi(x) = \begin{cases} 1 & \sum_{i=1}^{22} x_i > k \\ \delta & \sum_{i=1}^{22} x_i = k \\ 0 & \sum_{i=1}^{22} x_i < k. \end{cases} \quad (2.6)$$

از طرفی، با توجه به تعریف ۳.۳، α_Φ برابر است با

$$\begin{aligned} & P_{p \cong 0/10} \{d(X) = a_1\} \\ &= \int H_0^*(p) P_p(d(X) = a_1) dp \\ &= 10 \int_0^1 H_0(p) \left\{ P_p \left(\sum_{i=1}^{22} X_i > k \right) + \delta P_p \left(\sum_{i=1}^{22} X_i = k \right) \right\} dp \\ &= 10 \int_0^{0/10} 10p \left\{ \left(\sum_{s=k+1}^{22} \binom{22}{s} p^s (1-p)^{22-s} \right) + \delta \binom{22}{k} p^k (1-p)^{22-k} \right\} dp \\ &+ 10 \int_{0/10}^{0/20} (2 - 10p) \left\{ \left(\sum_{s=k+1}^{22} \binom{22}{s} p^s (1-p)^{22-s} \right) + \delta \binom{22}{k} p^k (1-p)^{22-k} \right\} dp. \end{aligned}$$

مقادیر α_Φ برای همه مقدارهای محتمل k به ازای $\delta = 0$ محاسبه شده و در جدول ۴ با نماد $\alpha_\Phi|_{\delta=0}$ نشان داده شده است. با توجه به ستون‌های اول و دوم این جدول، می‌توان نتیجه گرفت که $k = 4$ و در نتیجه عدد ثابت δ از رابطه تک‌مجهولی

$$\begin{aligned} 0/10 &= 0/95 + 10\delta \left\{ \int_0^{0/10} 10p \binom{22}{4} p^4 (1-p)^{18} dp \right. \\ &\quad \left. + \int_{0/10}^{0/20} (2 - 10p) \binom{22}{4} p^4 (1-p)^{18} dp \right\} \end{aligned}$$

برابر با ۰/۰۴۶ محاسبه می‌گردد. پس در این مثال کاربردی، آزمون تصادفی

$$\Phi(x) = \begin{cases} 1 & \sum_{i=1}^{22} x_i > 4 \\ 0/046 & \sum_{i=1}^{22} x_i = 4 \\ 0 & \sum_{i=1}^{22} x_i < 4 \end{cases}$$

پرتوان‌ترین آزمون در سطح $0/10$ است. در نتیجه بر اساس مشاهده $\sum_{i=1}^{22} x_i = 4$ ، پرتوان‌ترین آزمون، فرضیه فازی $\tilde{H}_0$ را با احتمال ۰/۰۴۶ رد و با احتمال ۰/۹۵۴ می‌پذیرد.

جدول ۱. مقدارهای عددی $\frac{\int H_1(p)f(x;p)dp}{\int H_0(p)f(x;p)dp}$ برای یافتن پرتوان‌ترین آزمون و همچنین مقدارهای عددی $\frac{\int H_1(p)\pi(p|x)dp}{\int H_0(p)\pi(p|x)dp}$ برای یافتن آزمون بیز به‌ازای مقدارهای مختلف $\sum_{i=1}^{22} x_i$ در مثال کاربردی:

$\frac{\int_0^1 H_1(p)\pi(p x)dp}{\int_0^1 H_0(p)\pi(p x)dp}$	$\frac{\int_0^1 H_1(p)f(x;p)dp}{\int_0^1 H_0(p)f(x;p)dp}$	$\sum_{i=1}^{22} x_i$
۰/۰۰۷	۰/۰۰۵	۰
۰/۰۲۸	۰/۰۲۵	۱
۰/۰۹۶	۰/۰۹۷	۲
۰/۲۸	۰/۳۲۳	۳
۰/۸۰	۰/۹۳۲	۴
۲/۰۹	۲/۴۹	۵
۵/۲۶	۶/۴۸	۶
۱۲/۹	۱۶/۷	۷
۳۱/۰	۴۳/۹	۸
۷۴/۳	۱۱۶	۹
۱۷۷	۳۱۶	۱۰
۴۲۳	۸۷۶	۱۱
$۱/۰۱ \times ۱۰^۳$	۲۴۶۶	۱۲
$۲/۴۶ \times ۱۰^۳$	۷۰۴۹	۱۳
$۱/۳۱ \times ۱۰^۴$	$۲/۰۴ \times ۱۰^۴$	۱۴
$۳/۸۵ \times ۱۰^۴$	$۵/۹۶ \times ۱۰^۴$	۱۵
$۱/۱۵ \times ۱۰^۵$	$۴۱/۷۶ \times ۱۰^۵$	۱۶
$۳/۵۱ \times ۱۰^۵$	$۵/۲۳ \times ۱۰^۵$	۱۷
$۱/۰۹ \times ۱۰^۶$	$۱/۵۶ \times ۱۰^۶$	۱۸
$۳/۴۴ \times ۱۰^۶$	$۴/۶۸ \times ۱۰^۶$	۱۹
$۱/۰۷ \times ۱۰^۷$	$۱/۴۱ \times ۱۰^۷$	۲۰
$۲/۶۲ \times ۱۰^۷$	$۴/۲۵ \times ۱۰^۷$	۲۱
$۷/۹۲ \times ۱۰^۷$	$۱/۲۸ \times ۱۰^۸$	۲۲

مقایسه‌ای از رویکردهای نیمن - پیرسون، بیز و کم- بیشینه در آزمون فرضیه‌های فازی _____ ۶۲

ب) آزمون کم-بیشینه: علاوه بر فرض‌های فوق، فرض کنید هزینه سمپاشی برابر 10° واحد پولی و هزینه زیان ناشی از عدم سمپاشی در حالی که آفت شیوع پیدا کرده باشد، معادل 60° واحد پولی باشد. بنابراین تابع زیان مالی را می‌توان همانند جدول ۲ برحسب واحد پولی در نظر گرفت.

جدول ۲. تابع زیان مالی برای سمپاشی و یا عدم سمپاشی برحسب واحد پولی
 $(\mathcal{L}^*(\theta, a))$:

عمل	حالت طبیعت	
	آفت حدود 30°	آفت حدود 10°
عدم سمپاشی	60°	0°
سمپاشی	10°	10°

اگر نسبت شیوع آفت حدودا 30° باشد، بهترین حالت آن است که مزرعه سمپاشی شود. ولی حالت دوم آن است که با وجود حدودا 30° آفت، مزرعه سمپاشی نشود که در این صورت، زیان کشاورز برابر 60° واحد خواهد بود. از مقایسه این دو حالت درمی‌یابیم که در حالت دوم، زیان کشاورز 50° واحد بیشتر از حالت اول است و تنها افسوس و تأسف 50° واحد پولی را خواهد خورد. با توجه به آنچه مطرح شد، تابع زیان جدول ۲ نیاز به تصحیح داشته و مناسب‌تر است که به‌ازای هر حالت طبیعت، تمامی زیان‌ها نسبت به بهترین (کمترین) زیان مقایسه گردند. لذا با توجه به جدول ۲، به‌وسیله فرمول

$$\mathcal{L}(\theta, a) = \mathcal{L}^*(\theta, a) - \inf_{a \in \mathcal{A}} \mathcal{L}^*(\theta, a)$$

تابع زیان تصحیح‌شده در جدول ۳ درج شده است (برای توضیحات بیشتر به صفحه ۱۱۳ از [۱۵] و یا صفحه ۳۷۶ از [۸] مراجعه کنید). بنابراین برای تصمیم‌گیری کم-بیشینه، می‌توان زیان مطرح شده در جدول ۳ را به‌صورت زیر بازنویسی کرد:

$$\begin{aligned} \mathcal{L}(\tilde{H}_0, a_0) &= 0, & \mathcal{L}(\tilde{H}_0, a_1) &= 10, \\ \mathcal{L}(\tilde{H}_1, a_0) &= 50, & \mathcal{L}(\tilde{H}_1, a_1) &= 0. \end{aligned} \quad (3.6)$$

جدول ۳. تابع زیان تصحیح شده مالی برای سمپاشی و یا عدم سمپاشی برحسب واحد پولی:

حالت طبیعت		تصمیم
آفت حدود $0/10$ ($\tilde{H}_0$)	آفت حدود $0/30$ ($\tilde{H}_1$)	
۰	۵۰	عدم سمپاشی (a_0)
۱۰	۰	سمپاشی (a_1)

توجه داشته باشید که اگرچه ناحیه بحرانی آزمون کم-بیشینه به صورت (۵.۴) مطرح شده است، با توجه به اکیداً صعودی بودن $\frac{\int H_1(p)f(x;p)dp}{\int H_0(p)f(x;p)dp}$ نسبت به $\sum_{i=1}^{22} x_i$ ، این آزمون معادل آزمون

$$\Phi_m(x) = \begin{cases} 1 & \sum_{i=1}^{22} x_i > k \\ \delta & \sum_{i=1}^{22} x_i = k \\ 0 & \sum_{i=1}^{22} x_i < k \end{cases} \quad (4.6)$$

است. از طرفی، دامنه تغییرات متغیر تصادفی گسسته $\sum_{i=1}^{22} X_i$ مجموعه $\{0, 1, \dots, 22\}$ است و لذا همان طور که در جدول ۴ مشهود است نمی توان k ی در این دامنه یافت که در $\mathcal{R}(\tilde{H}_0, k) = \mathcal{R}(\tilde{H}_1, k)$ صدق کند. جدول ۴ حاوی مقدارهای ریسک آزمون غیرتصادفی تحت فرضیه های فازی $\tilde{H}_1$ و $\tilde{H}_0$ است که با در نظر داشتن مقدارهای $0/10 = C_0$ و $0/30 = C_1$ به صورت زیر برحسب k محاسبه می شوند:

$$\begin{aligned} \mathcal{R}(\tilde{H}_0, k) &= C_0 P_{p \cong 0/10} \left(\sum_{i=1}^{22} X_i \geq k \right) \\ &= 0/10 C_0 \int_0^1 H_0(p) P_p \left(\sum_{i=1}^{22} X_i \geq k \right) dp \\ &= 0/10^2 \int_0^{0/10} 0/10 p \left(\sum_{s=k}^{22} \binom{22}{s} p^s (1-p)^{22-s} \right) dp \\ &\quad + 0/10^2 \int_{0/10}^{0/20} (2 - 0/10 p) \left(\sum_{s=k}^{22} \binom{22}{s} p^s (1-p)^{22-s} \right) dp \end{aligned}$$

$$\begin{aligned}
 \mathcal{R}(\tilde{H}_1, k) &= C_1 P_{p \cong \circ / \circ} \left(\sum_{i=1}^{22} X_i < k \right) \\
 &= 1 \circ C_1 \int_{\circ}^1 H_1(p) P_p \left(\sum_{i=1}^{22} X_i < k \right) dp \\
 &= 5 \circ \circ \int_{\circ / \circ}^{\circ / \circ} (1 \circ p - 2) \left(\sum_{s=\circ}^{k-1} \binom{22}{s} p^s (1-p)^{22-s} \right) dp \\
 &\quad + 5 \circ \circ \int_{\circ / \circ}^{\circ / \circ} (4 - 1 \circ p) \left(\sum_{s=\circ}^{k-1} \binom{22}{s} p^s (1-p)^{22-s} \right) dp.
 \end{aligned}$$

به‌طور مشابه با توجه به مقدارهای $1 \circ$ و $5 \circ$ ، C_1 می‌توان ریسک آزمون تصادفی را نیز تحت فرضیه‌های فازی $\tilde{H}_1$ و $\tilde{H}_\circ$ برحسب k به‌صورت زیر محاسبه کرد:

$$\begin{aligned}
 \mathcal{R}_{Rand}(\tilde{H}_\circ, k) &= C_\circ \left\{ P_{p \cong \circ / 1 \circ} \left(\sum_{i=1}^{22} X_i > k \right) + \delta P_{p \cong \circ / 1 \circ} \left(\sum_{i=1}^{22} X_i = k \right) \right\} \\
 &= 1 \circ^2 \int_{\circ}^{\circ / 1 \circ} 1 \circ p \left\{ \left(\sum_{s=k+1}^{22} \binom{22}{s} p^s (1-p)^{22-s} \right) + \delta \binom{22}{k} p^k (1-p)^{22-k} \right\} dp \\
 &\quad + 1 \circ^2 \int_{\circ / 1 \circ}^{\circ / \circ} (2 - 1 \circ p) \left\{ \left(\sum_{s=k+1}^{22} \binom{22}{s} p^s (1-p)^{22-s} \right) + \delta \binom{22}{k} p^k (1-p)^{22-k} \right\} dp
 \end{aligned}$$

و

$$\begin{aligned}
 \mathcal{R}_{Rand}(\tilde{H}_1, k) &= C_1 \left\{ P_{p \cong \circ / \circ} \left(\sum_{i=1}^{22} X_i < k \right) + (1 - \delta) P_{p \cong \circ / \circ} \left(\sum_{i=1}^{22} X_i = k \right) \right\} \\
 &= 5 \circ \circ \int_{\circ / \circ}^{\circ / \circ} (1 \circ p - 2) \left\{ \left(\sum_{s=\circ}^{k-1} \binom{22}{s} p^s (1-p)^{22-s} \right) + (1 - \delta) \binom{22}{k} p^k (1-p)^{22-k} \right\} dp \\
 &\quad + 5 \circ \circ \int_{\circ / \circ}^{\circ / \circ} (4 - 1 \circ p) \left\{ \left(\sum_{s=\circ}^{k-1} \binom{22}{s} p^s (1-p)^{22-s} \right) + (1 - \delta) \binom{22}{k} p^k (1-p)^{22-k} \right\} dp.
 \end{aligned}$$

اکنون باید مقدار k را به گونه‌ای از دامنه تغییرات آن به‌دست آورد که در دو رابطه

$$\mathcal{R}(\tilde{H}_\circ, k) > \mathcal{R}(\tilde{H}_1, k) \quad \text{و} \quad \mathcal{R}(\tilde{H}_\circ, k+1) < \mathcal{R}(\tilde{H}_1, k+1)$$

صدق کند. بر اساس اطلاعات جدول ۴، ابتدا $k = ۳$ به دست می آید و سپس نتیجه می شود که

$$\Phi_m(x) = \begin{cases} ۱ & \sum_{i=1}^{۲۲} x_i > ۳ \\ \delta & \sum_{i=1}^{۲۲} x_i = ۳ \\ ۰ & \sum_{i=1}^{۲۲} x_i < ۳ \end{cases}$$

تابع آزمون تصادفی کم-بیشینه است و در آن، عدد ثابت δ از رابطه

$$\mathcal{R}(\tilde{H}_0, ۴) + \delta \left(\mathcal{R}(\tilde{H}_0, ۳) - \mathcal{R}(\tilde{H}_0, ۴) \right) = \mathcal{R}(\tilde{H}_1, ۳) + (1 - \delta) \left(\mathcal{R}(\tilde{H}_1, ۴) - \mathcal{R}(\tilde{H}_1, ۳) \right)$$

به دست می آید. رابطه اخیر را می توان به کمک اطلاعات جدول ۴ به صورت

$$۲/۰۴ + \delta (۳/۸۰ - ۲/۰۴) = ۱/۴۶ + (1 - \delta) (۴/۲۵ - ۱/۴۶)$$

بازنویسی کرد و در نتیجه $\delta = ۰/۴۸۶$. بنابراین با در نظر داشتن تابع زیان (۳.۶)، آزمون تصادفی کم-بیشینه به صورت

$$\Phi_m(x) = \begin{cases} ۱ & \sum_{i=1}^{۲۲} x_i > ۳ \\ ۰/۴۸۶ & \sum_{i=1}^{۲۲} x_i = ۳ \\ ۰ & \sum_{i=1}^{۲۲} x_i < ۳ \end{cases}$$

است و لذا بر اساس مشاهده $\sum_{i=1}^{۲۲} x_i = ۴$ ، تابع زیان (۳.۶)، فرضیه فازی $\tilde{H}_0$ در مقابل فرضیه فازی $\tilde{H}_1$ در این مثال کاربردی رد می گردد.

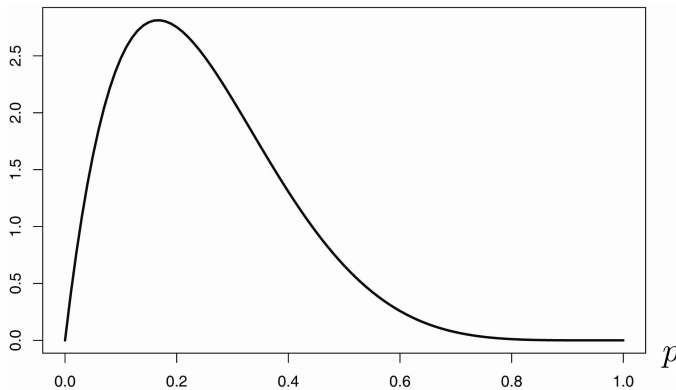
(ج) آزمون بیز: اکنون علاوه بر فرض های بالا و تابع زیان بند (ب)، فرض کنید بتوانیم طبق شواهد موجود و نظرات کارشناسان، توزیع پیشین $Beta(۲, ۶)$ را برای p در نظر بگیریم، یعنی به ازای $۰ \leq p \leq ۱$ داریم $\pi(p) = ۴۲p(1-p)^۵$ (شکل ۶ را ببینید). برای محاسبه توزیع پسین، می نویسیم

$$\begin{aligned} \pi(p|x) &\propto \pi(p) f(x|p) \\ &= (۴۲p(1-p)^۵) \left(p^{\sum_{i=1}^{۲۲} x_i} (1-p)^{۲۲-\sum_{i=1}^{۲۲} x_i} \right) \\ &\propto p^{۱+\sum_{i=1}^{۲۲} x_i} (1-p)^{۲۷-\sum_{i=1}^{۲۲} x_i} \end{aligned}$$

و بنابراین $p|x \sim Beta(۲ + \sum_{i=1}^{۲۲} x_i, ۲۸ - \sum_{i=1}^{۲۲} x_i)$

جدول ۴. مقادیر α_Φ به‌ازای $\delta = 0$ برای یافتن پرتوان‌ترین آزمون تصادفی و همچنین مقدارهای $\mathcal{R}(\tilde{H}_0, k)$ و $\mathcal{R}(\tilde{H}_1, k)$ برای یافتن آزمون تصادفی کم-بیشینه برای همه مقدارهای محتمل k :

$\mathcal{R}(\tilde{H}_1, k)$	$\mathcal{R}(\tilde{H}_0, k)$	$\alpha_\Phi _{\delta=0}$	k
0	10	1	0
0/039	8/46	0/846	1
0/339	6/12	0/612	2
1/46	3/80	0/380	3
4/25	2/04	0/204	4
9/30	0/95	0/095	5
16/5	0/39	0/039	6
24/9	0/13	0/013	7
33/0	$4/36 \times 10^{-2}$	$4/36 \times 10^{-3}$	8
39/7	$1/19 \times 10^{-2}$	$1/19 \times 10^{-3}$	9
44/5	$2/84 \times 10^{-3}$	$2/84 \times 10^{-4}$	10
47/4	$4/12 \times 10^{-3}$	$4/12 \times 10^{-4}$	11
48/9	$6/42 \times 10^{-5}$	$6/42 \times 10^{-6}$	12
49/6	$8/47 \times 10^{-6}$	$8/47 \times 10^{-7}$	13
49/8	$9/39 \times 10^{-7}$	$9/39 \times 10^{-8}$	14
49/9	$8/66 \times 10^{-8}$	$8/66 \times 10^{-9}$	15
$\simeq 50$	$6/57 \times 10^{-9}$	$6/57 \times 10^{-10}$	16
$\simeq 50$	$4/03 \times 10^{-10}$	$4/03 \times 10^{-11}$	17
$\simeq 50$	$1/95 \times 10^{-11}$	$1/95 \times 10^{-12}$	18
$\simeq 50$	$7/17 \times 10^{-13}$	$7/17 \times 10^{-14}$	19
$\simeq 50$	$1/88 \times 10^{-14}$	$1/88 \times 10^{-15}$	20
$\simeq 50$	$3/70 \times 10^{-16}$	$3/70 \times 10^{-17}$	21
$\simeq 50$	$\simeq 0$	$\simeq 0$	22



شکل ۶. تابع چگالی احتمال $Beta(2, 6)$ به عنوان پیشین p در مثال کاربردی

بر اساس تعریف ۱.۵ و با توجه به مقدار $\frac{C_0}{C_1} = \frac{1}{6} = 0.167$ ، آزمون بیز به صورت زیر به دست می آید:

$$\Phi_B(x) = \begin{cases} 1 & \frac{\int H_1(p)\pi(p|x)dp}{\int H_0(p)\pi(p|x)dp} \geq 0.167 \\ 0 & \frac{\int H_1(p)\pi(p|x)dp}{\int H_0(p)\pi(p|x)dp} < 0.167. \end{cases}$$

بر اساس نتایج جدول ۱، تابعی اکیداً صعودی از آماره $\sum_{i=1}^{22} x_i$ است و در نتیجه آزمون فوق معادل آزمون

$$\Phi_B(x) = \begin{cases} 1 & \sum_{i=1}^{22} x_i \geq 3 \\ 0 & \sum_{i=1}^{22} x_i < 3 \end{cases}$$

خواهد بود. بنابراین با به کار بردن تابع زیان (3.6) و چگالی پیشین $Beta(2, 6)$ و همچنین مقدار مشاهده شده $\sum_{i=1}^{22} x_i = 4$ در این مثال کاربردی، آزمون بیز فرضیه فازی $\tilde{H}_0$ را در مقابل فرضیه فازی $\tilde{H}_1$ رد می کند.

۷. نتیجه گیری و کارهای آینده

سه رویکرد نیمن-پیرسون، کم-بیشینه و بیز که در این مقاله برای آزمون فرضیه های فازی مبتنی بر داده های دقیق مطرح شد، قابل مقایسه با برخی از تحقیقات دیگر در این زمینه است. آرنولد [۷] پرتوان ترین آزمون برای فرضیه های فازی (با داده های دقیق) را به گونه ای متفاوت با رویکرد نیمن-پیرسونی بخش ۳،

معرفی کرده است. وی احتمال خطاهای نوع اول و دوم را با

$$\alpha_{\Phi} = \sup_{\{\theta \in \Theta: H_0(\theta) > H_1(\theta)\}} \{(H_0(\theta) - H_1(\theta)) \Phi(t, \theta)\}$$

و

$$\beta_{\Phi} = \sup_{\{\theta \in \Theta: H_0(\theta) < H_1(\theta)\}} \{(H_1(\theta) - H_0(\theta)) (1 - \Phi(t, \theta))\}$$

تعریف می‌کند که در آن، $\Phi(t, \theta)$ تابع توان بر اساس آماره آزمون T است. لذا تعمیم او برای احتمال خطاهای نوع اول و دوم مبتنی بر سوپریمم اختلاف توابع عضویت فرضیه‌های فازی است، در حالی که تعمیم مطرح شده در این مقاله مبتنی بر احتمال وزنی پیشامدهای فازی است. رویکرد مطرح شده در بخش ۴ از این مقاله برای تعمیم آزمون کم-بیشینه با کار کازالس و همکاران [۹] متفاوت است، زیرا آنها آزمون کم-بیشینه را برای فرضیه‌های غیرفازی مبتنی بر داده‌های فازی معرفی کردند ولی ما در این مقاله مسئله آزمون کم-بیشینه را برای فرضیه‌های فازی مبتنی بر داده‌های غیرفازی مورد بررسی قرار داده‌ایم. دلگادو و همکاران [۱۰] نیز روش بیزی متفاوتی را نسبت به آنچه در بخش ۵ برای آزمون فرضیه‌های فازی مطرح شد، مبتنی بر تجزیه مسئله فازی به خانواده‌ای از مسائل غیرفازی، ارائه کردند. در آزمون فرضیه‌های فازی بر اساس داده‌های معمولی، که در مقاله حاضر بررسی شد، ناحیه‌های بحرانی آزمون‌های کم-بیشینه و نیمن-پیرسون (در سطح معناداری α) به صورت

$$\frac{\int_{\Theta} H_1(\theta) f(x; \theta) d\theta}{\int_{\Theta} H_0(\theta) f(x; \theta) d\theta} \geq k$$

است که در آن، مقدار ثابت k به ترتیب از رابطه‌های

$$C_1 P_{\tilde{H}_1} \{\tilde{H}_0\} = C_0 P_{\tilde{H}_0} \{\tilde{H}_0\}$$

و $\alpha = P_{\tilde{H}_0} \{\tilde{H}_0\}$ محاسبه می‌شوند. از سوی دیگر، ناحیه بحرانی آزمون بیز، به صورت

$$\frac{\int_{\Theta} H_1(\theta) \pi(\theta) f(x; \theta) d\theta}{\int_{\Theta} H_0(\theta) \pi(\theta) f(x; \theta) d\theta} \geq k$$

است که $k = \frac{C_0}{C_1}$. در این مقاله، آزمون‌های کم-بیشینه و بیز مبتنی بر تابع زیان (۲.۴) محاسبه و معرفی شده‌اند. از آنجا که معمولاً مسائل واقعی و کاربردی در علوم مختلف آمیخته با ابهام هستند، لذا یک مثال کاربردی در بخش ۶ به منظور اشاره بیشتر به قابلیت‌های کاربردی این آزمون‌ها مطرح و حل شد.

گفتنی است که موضوع آزمون فرضیه‌های فازی بر اساس رویکردهای مختلف توسط دیگر محققان دنبال و بررسی شده است که برای مرور برخی از پژوهش‌های اخیر می‌توانید به [۱، ۲، ۳، ۱۱، ۱۲، ۱۷، ۲۰]

مراجعة نمایید. همچنین حقیقی‌نژاد و طاهری در [۴، ۵] به بررسی برآوردگرهای بیز و کم-بیشینه بر پایه تابع زیان فازی پرداخته‌اند. لذا ایده به‌کارگیری تابع زیان فازی را می‌توان برای آزمون فرضیه‌های فازی در روش‌های بیز و کم-بیشینه به‌عنوان کارهای آینده مورد بررسی قرار داد. همچنین در راستای این تحقیق می‌توان به تعمیم ایده پیشنهادی برای آزمون‌های کم-بیشینه و بیز در حالتی که تابع زیان ثابت نبوده، بلکه به مقدار پارامتر θ بستگی داشته باشد، پرداخت. از دیگر موضوعات بالقوه، می‌توان به آزمون فرضیه در چارچوب نظریه تصمیم مبتنی بر متغیر تصادفی فازی و آزمون فرضیه برای ارزیابی مدل‌های رگرسیون فازی اشاره کرد.

مراجع

- [۱] پرچمی، ع.، ماشین‌چی، م.، آزمون فرضیه در محیط فازی بر پایه p -مقدار، سری سیستم‌های فازی و محاسبات نرم، جلد ۲ (مباحثی در آمار و احتمال فازی)، انتشارات انجمن سیستم‌های فازی ایران، تهران ۱۳۹۰، ۱۸-۱.
- [۲] ترابی، ح.، شمشیری، ر.، تعمیمی از آزمون نسبت درستی برای فرض‌های فازی، اندیشه آماری، سال سیزدهم، شماره یکم (۱۳۸۷)، ۲۶-۲۰.
- [۳] چاچی، ج.، روش‌های آماری بر اساس اطلاعات نادقیق، رساله دکتری آمار، دانشکده علوم ریاضی، دانشگاه صنعتی اصفهان، اصفهان ۱۳۹۱.
- [۴] حقیقی‌نژاد، ا.، طاهری، س.م.، برآوردگرهای بیز و مینیمکس بر پایه تابع زیان فازی، گزارش هفتمین کنفرانس آمار ایران، دانشگاه علامه طباطبائی، تهران ۱۳۸۳، ۵۸۰-۵۶۷.
- [۵] حقیقی‌نژاد، ا.، طاهری، س.م.، برآوردگرهای بیز و مینیمکس بر پایه داده‌های فازی، گزارش چهارمین همایش مجموعه‌های فازی و کاربردهای آن، دانشگاه مازندران، بابل، ۱۳۸۲، ۵۴-۴۲.
- [۶] طاهری، س.م.، ماشین‌چی، م.، مقدمه‌ای بر احتمال و آمار فازی، انتشارات دانشگاه شهید باهنر کرمان، کرمان ۱۳۸۷.
- [7] Arnold, B. F., An approach to fuzzy hypothesis testing, *Metrika*, **44** (1996), 119-126.
- [8] Berger, J. O., *Statistical Decision Theory and Bayesian Analysis*, 2nd. edn., Springer-Verlag, New York, 1985.
- [9] Casals, M. R., Gil, M. A., Gil, P., The fuzzy decision problem: An approach to the problem of testing statistical hypotheses with fuzzy information, *European Journal of Operational Research*, **27** (1986), 371-382.
- [10] Delgado, M., Verdegay, J. L., Vila, M. A., Testing fuzzy hypotheses, a Bayesian approach, in: M. M. Gupta et al. (editors), *Approximate Reasoning in Expert Systems*, North-Holland, Amsterdam, 307-316, 1985.
- [11] Hesamian, G., Chachi, J., Two-sample Kolmogorov-Smirnov fuzzy test for fuzzy random variables, *Statistical Papers*, **56** (2015), 61-82.
- [12] Hesamian, G., Shams, M., Parametric testing statistical hypotheses for fuzzy random variables, *Soft Computing*, **20** (2016), 1537-1548.

- [13] Pais, I., Benton, J. J., *The Handbook of Trace Elements*, St. Lucie Press, Florida, 1997.
- [14] Parchami, A., Mashinchi, M., Testing fuzzy hypotheses with crisp data: The minimax approach, *2nd. Joint Congress on Fuzzy and Intelligent Systems*, Maleke-Ashtar University, Tehran, 2008.
- [15] Parmigiani, G., Inoue, L. Y. T., Lopes, H. F., *Decision Theory Principles and Approaches*, John Wiley & Sons, New Jersey, 2009.
- [16] Rohatgi, V. K., Ehsanes Saleh, A. K., *An Introduction to Probability and Statistics*, 2nd. edn., John Wiley & Sons, New York, 2001.
- [17] Taheri, S. M., Arefi, M., Testing fuzzy hypotheses based on fuzzy test statistics, *Soft Computing*, **13** (2009), 617-625.
- [18] Taheri, S. M., Behboodian, J., Neyman-Pearson Lemma for fuzzy hypotheses testing, *Metrika*, **49** (1999), 3-17.
- [19] Taheri, S. M., Behboodian, J., A Bayesian approach to fuzzy hypotheses testing, *Fuzzy Sets and Systems*, **123** (2001), 39-48.
- [20] Torabi, H., Behboodian, J., Sequential probability ratio test for fuzzy hypotheses testing with vague data, *Austrian Journal of Statistics*, **34** (2005), 25-38.

عباس پرچمی: گروه آمار، دانشکده ریاضی و کامپیوتر، دانشگاه شهید باهنر کرمان
رایانامه: parchami@uk.ac.ir

سید محمود طاهری: دانشکده فنی، دانشگاه تهران
رایانامه: sm_taheri@ut.ac.ir

ناصررضا ارقامی: گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد
رایانامه: arghami_nr@yahoo.com

تجزیه نامنفی ماتریسی: روشی برای تحلیل داده‌های نامنفی

مهسا یوسفی و منصور رزقی

چکیده

امروزه پیشرفت فناوری رایانه‌ای منجر به افزایش حجم داده‌ها و نیز به وجود آمدن پایگاه‌های بزرگ داده‌ها شده است. در نتیجه روش‌های مختلفی برای کشف دانش از آنها، معرفی شده است و یا در حال معرفی هستند. در این راستا، یکی از شاخه‌های نسبتاً جدید علمی موسوم به داده‌کاوی مورد توجه زیادی قرار گرفته است. ماتریس‌های داده در کاربردهای داده‌کاوی، غالباً نامنفی هستند که این ویژگی محدودیت‌هایی را در استفاده از روش‌های ماتریسی کلاسیک به همراه دارد. اگرچه به کارگیری این روش‌ها سبب کاهش بُعد داده‌های بزرگ می‌شود، اما تعبیری صحیح از داده‌های نامنفی از آنها به دست نمی‌آید. اخیراً روش جدیدی با نام تجزیه نامنفی ماتریسی برای نمایش خطی داده‌های نامنفی پیشنهاد شده است که علاوه بر کاهش بُعد داده‌ها، محدودیت روش‌های کلاسیک را ندارد. در این روش، ماتریس بزرگ متناظر با داده‌های نامنفی به دو ماتریس نامنفی کوچک تجزیه می‌شود. در این مقاله، ابتدا روش‌های کلاسیک را مرور می‌کنیم. سپس تجزیه نامنفی ماتریسی با نسخه‌های مختلف آن معرفی و مسائل مهم داده‌کاوی مانند رده‌بندی و خوشه‌بندی برای این روش بررسی می‌شود.

۱. مقدمه

رشد روش‌های تولید داده‌ها بسیار سریع‌تر از توانایی ما در درک و استفاده از آنها است. برای درک این سرعت رشد، کافی است برای چند لحظه به تعداد افراد آنلاینی فکر کنیم که تجربیاتشان را با استفاده از متن، فیلم، عکس و ... با یکدیگر به اشتراک می‌گذارند. آن وقت می‌توان به سرعت تولید داده‌ها و افزایش

عبارات و کلمات کلیدی. تجزیه نامنفی ماتریسی؛ کاهش بُعد؛ کمترین مربعات متناوب با قید نامنفی؛ خوشه‌بندی؛ رده‌بندی.

حجم آنها در هر ثانیه پی‌برد. اما داده‌ها فقط زمانی مفید هستند که مورد پردازش قرار گرفته باشند. در واقع، برای ما مهم این است که از حجم بسیار بالای داده‌های خام اولیه، به دانش پنهان شده و الگوی مفید در آنها دست یابیم. این الگو باید ساده، معتبر و قابل فهم برای توصیف ارتباط میان داده‌ها باشد. در این راستا، شاخه نسبتاً جدید علمی موسوم به داده‌کاوی^۱ در حال گسترش است که می‌توان انگیزه شکل‌گیری آن را انجام تحقیقات در زمینه‌هایی نظیر آمار، یادگیری ماشینی و مدیریت پایگاه داده‌ها دانست. با داده‌کاوی می‌توان اطلاعات باارزشی را از مجموعه داده‌های بزرگ استخراج و بر اساس آنها تصمیمات مهمی اتخاذ کرد. در یک تقسیم‌بندی کلی می‌توان یادگیری ماشینی را به دو دسته اصلی: یادگیری بدون نظارت^۲ مانند خوشه‌بندی^۳ و یادگیری با نظارت^۴ مانند رده‌بندی^۵ تقسیم کرد. در داده‌کاوی، از خوشه‌بندی و رده‌بندی به‌سان ابزارهایی جهت توصیف داده‌ها و مدل کردن پیش‌بینی آنها استفاده می‌شود. اساس مشترک رهیافت‌های مختلف در تحلیل داده‌ها، پیدا کردن یک مدل مناسب برای نمایش آنها است تا بینش و تصویری صحیح از آنها به‌وجود آید. هدف ما، توصیف داده‌کاوی به‌عنوان شاخه جدید علمی نیست، بلکه بر آن هستیم که به اهمیت و کاربرد مدل(های) مناسب در آن بپردازیم. رهیافت بسیار معمولی که آن را مدل کاهشی^۶ می‌نامیم، مدلی است که سعی می‌کند از پیچیدگی داده‌های اولیه و با بُعد بالا بکاهد و ضمن حفظ ملزومات مسئله همراه با داده‌ها، ساختار پنهان آنها را آشکار کند. در واقع، این مدل ضمن کاهش داده‌ها، در سطحی نزدیک‌تر (واقعی‌تر) به سیستم اولیه قرار دارد. مدل‌های کاهشی می‌توانند خطی و یا غیرخطی باشند که در این مقاله، مدل‌های خطی را مطالعه می‌کنیم. در این مقاله، سه مدل کاهشی خطی به نام‌های تجزیه مقدار تکین (SVD)^۷، تحلیل مؤلفه اصلی (PCA)^۸ و تجزیه نامنفی ماتریسی (NMF)^۹ بیان می‌شوند و روی مدل اخیر بیشتر تمرکز خواهیم کرد. مدل NMF برای تقریب داده‌های نامنفی ذخیره‌شده در ماتریس نامنفی $A \in \mathbb{R}^{m \times n}$ ، به‌دنبال تولید دو ماتریس نامنفی دیگر مانند $W \in \mathbb{R}^{m \times r}$ و $H \in \mathbb{R}^{r \times n}$ با شرط $r \ll \min\{m, n\}$ است به‌طوری که معادله تقریبی $A \approx WH$ برقرار شود. در این معادله تقریبی، رتبه ماتریس WH از رتبه ماتریس A کمتر است. در بخش دوم، تجزیه ماتریسی را توضیح خواهیم داد. روش‌های کلاسیک مشهور در تجزیه ماتریسی را که در داده‌کاوی مورد استفاده قرار می‌گیرند، مرور می‌کنیم و محدودیت آنها را در توصیف داده‌ها نشان می‌دهیم. سپس روش جدید NMF را معرفی و نسخه‌های مختلف موجود از آن را به تفصیل بررسی می‌کنیم. بخش سوم، به دو کاربرد منتخب از روش NMF برای مسائل خوشه‌بندی و رده‌بندی به همراه نتایج عددی اختصاص دارد. در بخش پایانی، به جمع‌بندی و نتیجه‌گیری مطالب این مقاله خواهیم پرداخت.

^۱data mining ^۲unsupervised learning ^۳clustering ^۴supervised learning ^۵classification ^۶reduction model ^۷singular value decomposition ^۸principle component analysis ^۹nonnegative matrix factorization

۲. تجزیه ماتریسی

به طور کلی، تجزیه ماتریسی ابزاری برای تحلیل داده‌ها است. هر تجزیه تعبیرهای مختلفی را از ساختار ضمنی داده‌ها آشکار می‌سازد که البته این تعبیرها از نظر ریاضی هم‌ارز هستند. برخی از این تعبیرها در قالب مثال‌هایی در قسمت ۱.۳.۲ توضیح داده می‌شوند. یک پایگاه داده را می‌توان ماتریسی با m سطر و n ستون در نظر گرفت. برای مثال، در ماتریس متناظر با یک پایگاه داده تصویری، هر ستون می‌تواند معرف یک تصویر باشد (شکل ۱ را ببینید). تجزیه ماتریسی بازنمایش ویژه و ساده‌ای از ماتریس داده اولیه از طریق تولید ماتریس‌های جدید است. لذا به کارگیری آن، به این دلیل که داده‌ها اغلب پیچیده هستند و منجر به ناکارآمدی روش‌های موجود در داده‌کاوی می‌شوند، معقول و مناسب خواهد بود. حتی می‌توان از آن برای پاکسازی داده‌ای^۱ در شرایطی که داده‌ها آلوده به نوفه باشند نیز استفاده کرد. با مطالعه مرجع [۲۰] می‌توان جزئیات بیشتری در این باره به دست آورد. تجزیه مقدار تکین، تحلیل مؤلفه اصلی و تجزیه نامنفی ماتریسی از روش‌های شناخته شده در مبحث تجزیه ماتریسی هستند. در این مقاله، سعی بر این است که چگونگی محاسبه و آشکارسازی ساختار ضمنی ماتریس داده توسط NMF به تفصیل ارائه گردد. اما قبل از آن، دو تجزیه دیگر را مرور می‌کنیم.

۱.۲. تجزیه مقدار تکین. هر ماتریس $A \in \mathbb{R}^{m \times n}$ با $m \geq n$ را می‌توان به صورت

$$A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T \quad (۱.۲)$$

تجزیه کرد که در آن، $\Sigma \in \mathbb{R}^{n \times n}$ ماتریس قطری با نمایش

$$\Sigma = \text{diag}(\sigma_1, \dots, \sigma_n), \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0.$$

است و ماتریس‌های $U \in \mathbb{R}^{m \times m}$ و $V \in \mathbb{R}^{n \times n}$ متعامد هستند. رابطه (۱.۲) را تجزیه مقدار تکین (SVD) ماتریس A می‌نامیم. این تجزیه را می‌توان به صورت بسط زیر نیز نمایش داد:

$$A = \sum_{i=1}^n \sigma_i u_i v_i^T. \quad (۲.۲)$$

رتبه ماتریس A با تعداد مقادیر تکین ناصفر آن برابر است. در عمل، به علت وجود خطاهای مختلف، امکان دارد استقلال خطی ستون‌های ماتریس تغییر کند. به تعبیر ریاضی، این ناسازگاری به سبب وجود یک یا چند مقدار تکین بسیار کوچک رخ می‌دهد. این رویداد منجر به افزایش تأثیر خطا می‌شود که البته می‌توان با نادیده گرفتن این مقادیر تکین کوچک، از آن جلوگیری کرد. در نتیجه تعداد مقادیر تکین

^۱data cleaning

بزرگ به رتبه عددی^۱ ماتریس اشاره دارد که به روشنی از رتبه اصلی ماتریس کوچکتر است ($r < n$). در نمایش زیر که به روش تجزیه مقدار تکین برشی (TSVD)^۲ معروف است، ماتریس A_r جایگزین ماتریس A می‌شود:

$$A \approx \sum_{i=1}^r \sigma_i u_i v_i^T =: A_r. \quad (3.2)$$

از این رو می‌توان گفت که هر یک از داده‌های اولیه به صورت

$$a_j \approx \sum_{i=1}^r (\sigma_i v_{ij}) u_i, \quad j = 1, \dots, n \quad (4.2)$$

توسط u_i ها بازسازی می‌شود به طوری که سهم (وزن) هر بردار u_i برابر با $\sigma_i v_{ij}$ است. به این ترتیب، ماتریس متعامد U ماتریس پایه‌ای برای A خواهد بود. مراجع [۶] و [۱۲] را برای مطالعه جزئیات بیشتر در مورد SVD و کاربردهای آن ببینید.

۲.۲. تحلیل مؤلفه اصلی. در تحلیل مؤلفه اصلی [۲۳] نیز به دنبال یک تبدیل خطی متعامد هستیم تا داده‌ها توسط این تبدیل به دستگاه مختصات جدید انتقال یابند و بر اساس آن، آسان‌تر توصیف شوند. برای این منظور، محورهای مختصات جدید، پی در پی در جهتی قرار می‌گیرند که داده‌ها دارای بیشترین پراکندگی (واریانس) باشند. در این روش، پس از تعیین ماتریس داده نرمال شده A و ماتریس کوواریانس C متناظر با آن، بردارهای ویژه ماتریس C محاسبه و بر اساس کاهش مقادیر ویژه متناظرشان، مرتب می‌شوند. با توجه به مقارن بودن ماتریس C ، بردارهای ویژه چون متعامدند، تشکیل پایه می‌دهند. برای کاهش بُعد داده‌ها و حفظ اطلاعات مهم، r تا از بردارهای ویژه نرمال شده متناظر با مقادیر ویژه بزرگتر انتخاب خواهند شد. این r بردار پایه‌ای، تبدیل خطی $Q = [q_1 q_2 \dots q_r]$ (ماتریس پایه) را برای تبدیل داده‌ها از فضای چند بُعدی به فضایی با بُعد کمتر تشکیل می‌دهند. در روش تحلیل مؤلفه اصلی، داده‌های اولیه روی زیرفضای تولیدشده توسط ماتریس پایه C تصویر می‌شوند که به این ترتیب، ماتریس متناظر $H = [h_1 h_2 \dots h_r] \in \mathbb{R}^{r \times n}$ (ماتریس اوزان) به دست می‌آید [۲۳]:

$$A \approx QH = \sum_{j=1}^r q_j h_j^T =: A_r. \quad (5.2)$$

از این رو می‌توان گفت که هر یک از داده‌های اولیه، توسط r مؤلفه اصلی به صورت زیر بازسازی می‌شوند:

$$a_j \approx Qh_j = \sum_{i=1}^r q_i h_{ij}, \quad j = 1, \dots, n. \quad (6.2)$$

خیلی وقت‌ها به دلیل وجود شباهت‌هایی بین روش‌های SVD و PCA، این دو به جای هم در نظر گرفته می‌شوند. در این باره، مطالعه [۲۴] پیشنهاد می‌شود.

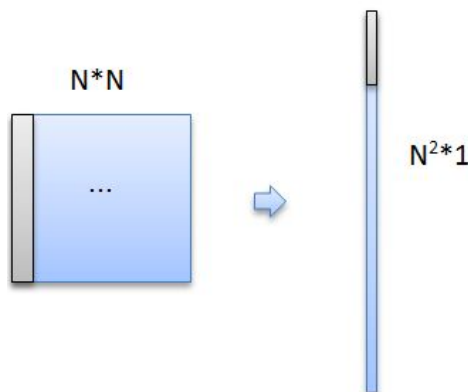
۳.۲. تجزیه نامنفی ماتریسی.

۱.۳.۲. نظریه. بسیاری از داده‌های دنیای واقعی پیرامون ما از جمله شدت نور (پیکسل‌های مربوط به تصویر دیجیتال) و غلظت شیمیایی در بیوانفورماتیک (ژن‌ها)، داده‌هایی با مقادیر نامنفی هستند. روشن است که بردارهای پایه‌ای در SVD (u_i ها) و PCA (q_i ها) دارای خاصیت تعامد هستند. این خاصیت سبب ایجاد مؤلفه‌هایی با علامت‌های دلخواه در بردارهای پایه‌ای و اوزان می‌گردد. پس این روش‌ها نمی‌توانند شرط نامنفی بودن داده‌ها را تضمین کنند. این محدودیت، انگیزه کافی برای تمرکز بر ابزاری بهتر و قوی‌تر مانند NMF را به وجود می‌آورد. این تجزیه، روشی سودمند برای نمایش خطی داده‌های نامنفی است به طوری که واقعیت فیزیکی (اطلاعات) آنها حفظ شود. NMF عوامل ماتریسی $W \in \mathbb{R}^{m \times r}$ و $H \in \mathbb{R}^{r \times n}$ را برای تقریب ماتریس نامنفی $A \in \mathbb{R}^{m \times n}$ با شرط $r \ll \min\{m, n\}$ چنان می‌یابد که رابطه تقریبی $A \approx WH$ برقرار باشد. پیدا کردن چنین تقریبی، نیازمند تابع هزینه‌ای است که کیفیت تقریب را به خوبی نشان دهد. یکی از این تابع‌ها می‌تواند اندازه فاصله دو ماتریس نامنفی از یکدیگر باشد. معمولاً این معادله تقریبی به صورت مسئله بهینه‌سازی

$$\min_{\substack{W \geq 0 \\ H \geq 0}} f(W, H) \equiv \frac{1}{2} \|A - WH\|_F^2 \quad (۷.۲)$$

صورت‌بندی می‌شود. در اینجا معنای نمادهای $H \geq 0$ و $W \geq 0$ این است که کلیه درآیه‌های این ماتریس‌ها نامنفی هستند. در سراسر این مقاله، منظور از NMF رابطه (۷.۲) خواهد بود. همان‌طور که قبلاً اشاره شد، تعبیرهای مختلفی از ساختار ضمنی داده‌ها توسط تجزیه‌های ماتریسی آشکار می‌شوند که از نظر ریاضی هم‌ارز هستند [۲۱].

مثال ۱.۲ (تعبیر عاملی). هر تصویر دیجیتال، یک آرایه مستطیلی $N \times N$ از پیکسل‌ها است و هر پیکسل توسط شدت نورش نمایش داده می‌شود. چون شدت نور توسط مقادیری نامنفی معین می‌شود، می‌توان هر تصویر را یک ماتریس نامنفی در نظر گرفت. همانند شکل ۱، از پشت سر هم قرار دادن ستون‌های ماتریس $N \times N$ ، بردار N^2 بُعدی متناظر با آن حاصل خواهد شد و مختصات این بردار، معرف پیکسل‌های تصویر هستند. به این ترتیب، ماتریس داده n ستونی متناظر با n داده تصویری تولید می‌شود. توجه به شکل برداری معادله تقریبی $A \approx WH$ با نمایش $a \approx Wh$ ، اهمیت ویژه‌ای دارد. در این نمایش، h و a معرف یک ستون از H و A هستند. این معادله تقریبی نشان می‌دهد که هر ستون از داده‌های ماتریس A را می‌توان به صورت ترکیبی خطی از r ستون ماتریس W نوشت (رابطه‌های (۴.۲)



شکل ۱. تبدیل داده از قالب ماتریس به بردار

و (۶.۲) را نیز ببینید). روشن است که ضرایب این ترکیب خطی مؤلفه‌های بردار h هستند. لذا به W ماتریس پایه و به ستون‌های آن، بردارهای پایه‌ای اطلاق می‌شود و همچنین H ماتریس اوزان خواهد بود. با توجه به نمایش $a \approx Wh$ ، می‌توان گفت که نمونه a توسط عامل ماتریسی W بازسازی می‌شود.

مثال ۲.۲ (تعبیر هندسی). متنی متشکل از پنج جمله و دو موضوع متفاوت در اختیار است که جمله‌های آن را در قالب ۵ سند به ترتیب زیر مشاهده می‌کنید به‌طوری که موضوع سند آخر (football) مستقل از موضوع سایر اسناد (google) است [۱۲]:

The **google matrix** P is a model of the **Internet**.

P_{ij} is nonzero if there is a **link** from **web page** j to i .

The **google matrix** is used to **rank** all **web pages**.

The **ranking** is done by solving a **matrix eigenvalue** problem.

England dropped out of the top 10 in **FIFA ranking**.

ماتریس متن-سند متناظر با این متن، به صورت ماتریس داده

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

تشکیل می‌گردد که در آن، اعداد ۱ و ۰ نشان‌دهندهٔ حضور و عدم حضور کلمات کلیدی در هر سند هستند. با این دانش، می‌توان میانگین بردارهای ستونی اول تا چهارم را نمایندهٔ خوشه‌ای با موضوع google و طبیعتاً بردار ستونی آخر را به‌عنوان نمایندهٔ خوشه‌ای با موضوع football دانست. هر نماینده را می‌توان یک بردار پایه‌ای برای ماتریس پایهٔ $W \in \mathbb{R}^{10 \times 2}$ در نظر گرفت و مختصات ستون‌های ماتریس A را برحسب این بردارهای پایه‌ای به‌صورت زیر محاسبه کرد:

$$\min_{H \geq 0} \|A - WH\|. \quad (۸.۲)$$

ستون‌های ماتریس H همان مختصات جدید داده‌های اولیه هستند که با توجه به بیشینهٔ مقادیر هر یک از این ستون‌ها، می‌توان خوشهٔ متناظر با هر یک از داده‌ها را تعیین کرد (خوشه‌بندی). اگر در (۸.۲) قید مثبت بودن حذف شود، این احتمال وجود خواهد داشت که در H مؤلفه‌های منفی ظاهر شود. این اتفاق بیان می‌کند که ماتریس نامنفی A با ماتریس رتبهٔ پائین‌تر اما با درآیه‌های منفی، تقریب زده شده است که این در واقعیت، غیرطبیعی به‌نظر می‌رسد. چون در این شرایط، حضور یا عدم حضور کلمات در اسناد، با مقادیر منفی بیان خواهد شد که تعبیر صحیحی نیست. در تعبیر هندسی، وجود پایه برای تصویر کردن داده‌ها ضروری است، اما نکتهٔ قابل توجه این است که مفهوم پایه در NMF برای ستون‌های ماتریس W ، برخلاف دو تجزیهٔ دیگر، با تعریف پایه در جبرخطی، فرق دارد. به‌عبارت دیگر، ماتریس W تنها با نام ماتریس پایه، نامگذاری و شناخته می‌شود.

۲.۳.۲. روش‌های محاسبه. تابع هزینه در مسئلهٔ NMF (۷.۲) همزمان نسبت به دو عامل ماتریسی H و W نامحدب است. لذا نمی‌توان در پی یافتن کمینهٔ سراسری بود. از این‌رو همگرایی به نقطهٔ ایستا، هدفی است که در همهٔ الگوریتم‌های NMF دنبال می‌شود. بر اساس قضیه‌ای در [۳]، اگر (W, H) نقطهٔ کمینهٔ موضعی مسئلهٔ (۷.۲) باشد، آنگاه شرایط KKT^۱ به‌صورت

$$H \geq 0, W \geq 0$$

$$\nabla_H f(W, H) = W^T W H - W^T A \geq 0,$$

$$\nabla_W f(W, H) = W H H^T - A H^T \geq 0,$$

$$H. * \nabla_H f(W, H) = 0, W. * \nabla_W f(W, H) = 0$$

برقرارند. نمادهای $*$ و ∇ به‌ترتیب، نشان‌دهندهٔ ضرب مؤلفه‌ای عوامل و گرادیان تابع هستند. پس (W, H) یک نقطهٔ ایستا برای مسئلهٔ (۷.۲) است اگر و تنها اگر در شرایط KKT صدق کند.

^۱Karush–Kuhn–Tucker conditions

در سال ۱۹۹۹ مطالعات جدیدی درباره NMF توسط لی و سانگ^۱ [۱۴] آغاز گردید. البته پیش از این، در مورد تجزیه مثبت ماتریسی^۲ [۱۹] چندین اثر از پاترو^۳ در سال‌های ۱۹۹۴، ۱۹۹۷ و ۱۹۹۹ منتشر شده بود. با وجود این، اثر لی و سانگ به‌عنوان اساس و NMF استاندارد در نظر گرفته می‌شود. روش‌های تناوبی مختلفی به‌منظور سرعت بخشیدن به الگوریتم استاندارد NMF پیشنهاد شده‌اند. در این باره می‌توان به پژوهش‌های لین^۴ در سال ۲۰۰۵ [۱۶، ۱۷]، بری^۵ و همکاران وی در سال ۲۰۰۷ [۲] و سپس کیم و پارک^۶ در سال‌های ۲۰۰۷ و ۲۰۰۸ [۸، ۹] اشاره کرد. کلی‌ترین دسته‌بندی برای الگوریتم‌های NMF عبارت است از: قواعد بهنگام ضربی^۷، روش کمترین مربعات متناوب^۸ و روش‌های گرادیان^۹.

۳.۳.۲. قواعد بهنگام ضربی. پرکاربردترین رهیافت برای مسئله (۷.۲) را می‌توان قواعد بهنگام ضربی دانست که توسط لی و سانگ در سال ۲۰۰۱ پیشنهاد گردید [۱۵]. گفتیم که تابع هزینه مسئله NMF نامحدب است و لذا نمی‌توان نقطه کمینه سراسری را پیدا کرد. اما روش‌هایی در بهینه‌سازی عددی وجود دارند که برای پیدا کردن کمینه موضعی به‌کار می‌روند. شاید روش‌های کاهش گرادیان [۳] از ساده‌ترین آنها باشند. به‌طور کلی، الگوریتم لی و سانگ را می‌توان یک روش کاهش گرادیان دانست که در قالب یک فرم متناوب برای بهنگام کردن مؤلفه‌های عوامل ماتریسی از قواعد

$$W_{ia} \leftarrow W_{ia} - \eta_{ia} (\nabla_W f(W, H))_{ia}, H_{bj} \leftarrow H_{bj} - \mu_{bj} (\nabla_H f(W, H))_{bj} \quad (۹.۲)$$

در خلاف جهت گرادیان استفاده می‌کند. در اینجا $\eta_{ia} = \frac{W_{ia}}{(WHH^T)_{ia}}$ و $\mu_{bj} = \frac{H_{bj}}{(W^TWH)_{bj}}$ به‌عنوان طول گام در نظر گرفته می‌شوند. توجه می‌کنیم که همگرایی روش‌های مبتنی بر گرادیان، به انتخاب طول گام حساس است. لی و سانگ برای کمینه کردن نرم فروبنیوس $\|A - WH\|_F^2$ ، قواعد بهنگام ضربی

$$W_{ia} \leftarrow W_{ia} \frac{(AH^T)}{(WHH^T)_{ia}}, H_{bj} \leftarrow H_{bj} \frac{(W^TA)_{bj}}{(W^TWH)_{bj}} \quad (۱۰.۲)$$

را پیشنهاد کردند. روشن است که وقتی $A = WH$ ، هر یک از ضرایب این قواعد برابر با یک خواهند شد؛ به این معنی که $\nabla_W f(W, H) = 0$ و $\nabla_H f(W, H) = 0$. لذا این روش یک نوع روش نقطه ثابت است. تحت این شرایط، لی و سانگ ثابت کردند که نرم فروبنیوس $\|A - WH\|_F^2$ تحت قواعد (۱۰.۲) ناافزایشی است. البته نسخه اصلی این قضیه در [۱۵] شامل قسمت دومی نیز هست که بعداً معلوم شد لزوماً درست نیست [۲، ۷، ۱۶]. لی و سانگ در قسمت دوم نسخه اصلی قضیه ادعا کردند که این نرم، تحت قواعد (۱۰.۲)، در یک نقطه ایستا تغییر نمی‌کند به این دلیل که الگوریتم لی و سانگ، یک روش نقطه ثابت است و لذا اگر با شروع از $H > 0$ به نقطه ایستای $H_* > 0$ همگرا شود،

^۱Daniel Lee and Sebastian Seung ^۲positive matrix factorization ^۳Paatero Pentti ^۴Chih-Jen Lin

^۵Michael Berry ^۶Hyunsoo Kim and Haesun Park ^۷multiplicative update rules ^۸alternating least

squares method ^۹gradient methods

آن‌گاه $\nabla_H f(W_*, H_*) = 0$. اما علی‌رغم مثبت بودن عامل ماتریسی در همه تکرارها، ممکن است برخی از مؤلفه‌های آن به صفر میل کنند، زیرا ما با دستگاه‌هایی سروکار داریم که اعداد در آنها دارای دقت محدود هستند و در نتیجه اعداد بسیار کوچک، به صفر گرد می‌شوند. به عبارت دیگر، مؤلفه‌های صفر بهنگام نمی‌شوند. حال اگر در یکی از تکرارها، مؤلفه‌ای مانند H_{bj} صفر شود، آن‌گاه $H_{*bj} = 0$. اما با توجه به شرط $\nabla_H f(W_*, H_*)_{bj} \geq 0$ ، معلوم نیست قاعده بهنگام چگونه باید عمل کند، زیرا وقتی $\nabla_H f(W_*, H_*) \neq 0$ ، روش باید نقطه فعلی را بهنگام کند که در این صورت حتماً بهینه بودن جواب فعلی نقض خواهد شد. حتی گاهی ممکن است با صفر شدن مؤلفه‌ای، مقدار گرادیان منفی شود. کلیه استدلال‌های بیان‌شده، برای قاعده دوم نیز برقرار است. با دلایل ذکر شده، الگوریتم لی و سانگ دارای همگرایی قوی نیست. معمولاً برای جلوگیری از به‌وجود آمدن مشکلات عددی، عدد کوچک مثبتی مانند $\epsilon = 10^{-9}$ به مخرج کسر هر یک از قواعد اضافه می‌کنند. الگوریتم لی و سانگ برای تابع هزینه نرم‌دار، به‌صورت زیر معرفی می‌شود:

مقداردهی اولیه: $W^k \geq 0$ و $H^k \geq 0$ وقتی $k = 1$

تکرار مراحل زیر تا برقراری شرط توقف:

(۱) اعمال قواعد (۱۰.۲) با افزودن مقدار ϵ به مخرج کسرها؛

(۲) $k = k + 1$

۴.۳.۲. **روش کمترین مربعات متناوب.** با توجه به نامحدب بودن تابع هزینه در مسئله NMF، تبدیل این مسئله به دو زیرمسئله محدب و در نتیجه استفاده از ویژگی‌های بهینه‌سازی محدب منطقی به‌نظر می‌رسد. در اینجا با ثابت نگه‌داشتن یکی از عوامل ماتریسی، می‌توان عامل ماتریسی دیگر را از یک مسئله کمترین مربعات به‌دست آورد. به این ترتیب در روش کمترین مربعات متناوب با قید نامنفی (ANLS)^۱، مسئله (۷.۲) به دو مسئله کمترین مربعات با قید نامنفی به‌صورت

$$\min_{H \geq 0} \frac{1}{2} \|A - WH\|_F^2 \quad (11.2)$$

با عامل ثابت W و

$$\min_{W \geq 0} \frac{1}{2} \|A^T - H^T W^T\|_F^2 \quad (12.2)$$

با عامل ثابت $H(H^T)$ بازنویسی می‌شود و تا برقراری شرط همگرایی، به‌طور تناوبی به‌دنبال هم می‌آیند. از آنجا که عامل سمت راست در زیرمسئله‌های (۱۱.۲) و (۱۲.۲) یک ماتریس است، در ادامه از نام NLS^۲ ماتریسی برای آنها استفاده می‌شود.

^۱alternating non-negativity-constrained least squares ^۲non-negative least squares

در سال ۲۰۰۷، بری و همکاران وی [۲] الگوریتم پایه ALS^۱ را برای NMF بیان کردند. در این الگوریتم، هر یک از زیرمسئله‌های (۱۱.۲) و (۱۲.۲) به صورت LS نامقید حل می‌شود و در صورت به وجود آمدن درآیه‌های منفی در عامل ماتریسی، مقدار صفر جایگزین آن درآیه‌ها خواهد شد. این عمل، ضمن برقراری شرط نامنفی بودن عوامل ماتریسی تولیدشده، به تنگی^۲ آنها نیز کمک می‌کند.

مقداردهی اولیه: $W^k \geq 0$ وقتی $k = 1$

تکرار مراحل زیر تا برقراری شرط توقف:

(۱) حل مسئله (۱۱.۲) به صورت نامقید برای عامل ماتریسی H ؛

(۲) تعویض درآیه‌های منفی ماتریس H با مقدار صفر (عمل تصویر کردن)؛

(۳) حل مسئله (۱۲.۲) به صورت نامقید برای عامل ماتریسی H ؛

(۴) تعویض درآیه‌های منفی ماتریس W با مقدار صفر (عمل تصویر کردن)؛

(۵) $k = k + 1$.

اگرچه این الگوریتم موجب تسریع محاسبات می‌شود، اما تحلیل همگرایی برای آن سخت خواهد بود، زیرا قیدهای نامنفی در دو زیرمسئله محذب دقیقاً برقرار نمی‌شوند. در حالت کلی، این الگوریتم به نقطه ایستا همگرا نیست و از آن به طور مستقیم استفاده نمی‌شود. معمولاً این الگوریتم به عنوان یک راه انداز برای الگوریتم‌های دیگر به کار می‌رود. هر NLS ماتریسی به ترتیب با ثابت در نظر گرفتن W و H ، با مسائل

$$\min_{H_{:j} \geq 0} \frac{1}{2} \sum_j \|A_{:j} - WH_{:j}\|_2^2, \quad j = 1, 2, \dots, n \quad (13.2)$$

و

$$\min_{W_{:j}^T \geq 0} \frac{1}{2} \sum_j \|A_{:j}^T - H^T W_{:j}^T\|_2^2, \quad j = 1, 2, \dots, m \quad (14.2)$$

معادل است. از این رو زیرمسئله‌های (۱۱.۲) و (۱۲.۲) به ترتیب به n و m مسئله NLS برداری تبدیل می‌شود. نماد $X_{:j}$ را برای نشان دادن ستون j ام ماتریس X به کار می‌بریم. برای حل هر NLS برداری، می‌توان از الگوریتم NNLS [۱۳] با نام کاربردی مجموعه فعال^۳ استفاده کرد. این الگوریتم در نرم افزار متلب با نام تابع lsqnonneg در دسترس است.

می‌دانیم که ناحیه پذیرفتنی یک LS مقید با قید نامنفی، همان قید مفروض است. با این فرض، هر جواب مسئله NLS برداری باید در ناحیه پذیرفتنی باشد و هر جواب پذیرفتنی دارای دو رده از متغیرها است که توسط آنها شناسایی می‌شود. رده اول، شامل متغیرهایی است که در درون ناحیه پذیرفتنی و رده دوم، شامل متغیرهایی است که بر روی مرز ناحیه پذیرفتنی قرار دارند. مجموعه غیرفعال، توسط اندیس متناظر با متغیرهای رده اول و مجموعه فعال توسط اندیس متناظر با متغیرهای رده دوم شناخته می‌شوند.

^۱alternating least squares ^۲sparcity ^۳active set

سازوکار الگوریتم مجموعه فعال در حل مسئله NLS برداری این گونه است که برای هر جواب پذیرفتنی آغازین، متغیرها به طور تکراری بین مجموعه فعال و غیرفعال جابه‌جا می‌شوند. در چنین حالتی، تابع هزینه به طور پیوسته و یکنواخت کاهش پیدا می‌کند در حالی که پذیرفتنی بودن جواب، محفوظ می‌ماند. برای درک و دریافت جزئیات بیشتر این الگوریتم، مطالعه مرجع [۲۰] پیشنهاد می‌شود. روشن است که استفاده از الگوریتم مجموعه فعال برای همه NLS‌های برداری، حتی با انجام محاسبات خیلی سریع، منجر به اجرایی‌کُند و غیر قابل قبول می‌شود. به منظور ارائه جایگزینی که برای حل زیرمسائل (۱۳.۲) و (۱۴.۲) مناسب باشد، می‌توان به اقدامات کیم و پارک [۸] در استفاده از روش مجموعه فعال ترکیبیاتی (FC-NNLS)^۱ و نیز اقدامات لین [۱۷] در استفاده از روش تصویر گرادیان^۲ اشاره کرد.

الگوریتم FC-NNLS توسط بنتم و کینن^۳ برای حل NLS ماتریسی پیشنهاد گردید. در حقیقت، این الگوریتم استفاده مستقیم از الگوریتم مجموعه فعال است. اشاره کردیم که هر NLS ماتریسی را می‌توان به چندین NLS برداری تبدیل کرد. تعداد NLS‌های برداری به تعداد ستون‌های ماتریس ضرایب زیرمسائل بستگی دارد. لذا می‌توان به جای اجرای الگوریتم تکراری مجموعه فعال برای هر NLS برداری، فقط یک تکرار از این الگوریتم را برای یک NLS ماتریسی به کار برد. در این استراتژی ترکیبیاتی، ستون‌هایی از ماتریس ضرایب که دارای مجموعه‌های غیرفعال یکسان‌اند با هم در نظر گرفته می‌شوند و مسئله به‌ازای متغیرهای متناظر با مجموعه غیرفعال، حل می‌شود. برای مطالعه دقیق ساختار الگوریتم FC-NNLS همراه با شبه‌کد متلب آن به نام fcnnls، به [۱] رجوع کنید. اگرچه سرعت رهیافت پیشنهادی توسط کیم و پارک مطلوب است، نسخه‌هایی سریع‌تر از آن در مراجع [۱۰، ۱۱] ارائه شده است.

مقداردهی اولیه: $W \geq 0$

تکرار مراحل زیر تا برقراری شرط توقف:

(۱) حل مسئله (۱۱.۲) با استفاده از fcnnls؛

(۲) حل مسئله (۱۲.۲) با استفاده از fcnnls.

۵.۳.۲. روش‌های گرادیان. برای تعیین عوامل ماتریسی در NMF، همانند قواعد بهنگام ضربی، می‌توان از قواعد بهنگام مبتنی بر روش گرادیان به صورت زیر استفاده کرد:

$$H = H - \alpha_H \nabla_H f(W, H), \quad W = W - \alpha_W \nabla_W f(W, H). \quad (۱۵.۲)$$

پارامترهای α_H و α_W به عنوان طول گام، وابسته به الگوریتم بوده و در خلاف جهت گرادیان انتخاب می‌شوند. در (۱۵.۲) برای تعیین هر یک از عوامل ماتریسی، عامل دیگر، ثابت در نظر گرفته می‌شود. اما نکته قابل توجه، چگونگی انتخاب طول گام‌ها است. می‌توان در ابتدا از مقدار یک برای آنها استفاده

^۱fast combinatorial non-negative least squares ^۲projected gradient ^۳Mark H. Van Benthem and Michael R. Keenan

کرد و در تکرارهای بعدی این مقدار را نصف نمود. با وجود این، تضمینی در منفی نشدن عوامل وجود ندارد. از طرفی، اگر بخواهیم عمل تصویر کردن را انجام دهیم، به این معنی که هر مقدار منفی را با صفر جایگزین کنیم، تحلیل همگرایی سخت‌تر خواهد شد. قبلاً اشاره کردیم که می‌توان الگوریتم NMF/MUR را یک روش کاهش گرادیان در نظر گرفت. در آنجا (روابط ۹.۲) مشاهده شد که همگرایی حتی با انتخاب دقیق‌تر طول گام‌ها، آرام است. لذا در حالت کلی می‌توان نتیجه گرفت که رهیافت‌های مبتنی بر روش گرادیان، همگرایی مطلوبی ندارند. محاسبه عوامل ماتریسی NMF را با روش‌های گرادیان می‌توان به سه مرحله: محاسبه گرادیان، انتخاب طول گام و تصویر کردن عامل ماتریسی جدید در فضایی نامنفی، برای هر تکرار تقسیم کرد. گاهی می‌توان مراحل انتخاب طول گام و عمل تصویر کردن عوامل ماتریسی جدید را در یک تکرار و با هم در نظر گرفت تا کاهش کافی در مقدار تابع هزینه حاصل شود. به این ترتیب در هر تکرار، الگوریتم‌های مبتنی بر روش گرادیان از یک حلقه داخلی به‌ازای هر قاعده بهنگام برخورداری خواهند شد. می‌توان به‌طور فرضی، کران‌های بالای بسیار بزرگی را برای قیود زیرمسئله‌های (۱۱.۲) و (۱۲.۲) در نظر گرفت. آن وقت هر یک از آنها یک مسئله مقید کراندار خواهد بود که لین برای حل این زیرمسئله‌ها، از رهیافتی مبتنی بر روش تصویر گرادیان [۳] استفاده کرد. راه‌های مختلفی برای انتخاب طول گام در این روش وجود دارد که در میان آنها، روش آرمیثو [۳] الگوریتمی ساده و مؤثر است. در این راستا، لین از یک سو کاهش هزینه را با توجه به درجه دو بودن تابع هدف و از سوی دیگر، انعطاف بیشتری را برای تعیین طول گام مورد توجه قرار داد. بر این اساس، طول گامی به‌عنوان حدس اولیه در نظر گرفته می‌شود و سپس در جریان الگوریتم افزایش یا کاهش می‌یابد. رهیافت لین این امکان را می‌دهد که طول گام‌های بزرگتر از یک نیز در جستجو، امتحان شوند. جزئیات بیشتر در مورد استفاده لین از این رهیافت به‌همراه شبه‌کد متلب آن با نام `alspsgrad`، در [۱۷] موجود است.

مقداردهی اولیه: $W \geq 0$ و $H \geq 0$

تکرار مراحل زیر تا برقراری شرط توقف [۱۷]:

(۱) حل مسئله (۱۱.۲) با استفاده از `alspsgrad`؛

(۲) حل مسئله (۱۲.۲) با استفاده از `alspsgrad`.

۶.۳.۲. *مقداردهی اولیه و شرط توقف.* برای NMF دو مبحث مقداردهی اولیه و شرط توقف، در دستیابی به عوامل ماتریسی بهین و همگرایی نقش به‌سزایی دارند. مشکلی که اغلب الگوریتم‌های NMF با آن روبه‌رو هستند، عدم تضمین همگرایی به نقطه کمینه سراسری است. مقداردهی اولیه ضعیف (مانند روش تصادفی) اغلب همگرایی آرام و گاهی جواب‌های بی‌ربط و غلطی را نتیجه می‌دهد. اگر یک مقدار اولیه خوب نتواند به قدر کافی همگرایی به نقطه ایستا را ضمانت کند، حتماً به اندازه کافی از تعداد تکرارها خواهد کاست. کارایی بسیاری از الگوریتم‌های NMF تحت تأثیر انتخاب ماتریس‌های اولیه است و لذا

مهم است که روش‌های سازگار و کارا برای مقداردهی اولیه به عوامل ماتریسی در دست باشند. مقداردهی اولیه مبتنی بر روش NMF [۵] و مقداردهی مبتنی بر روش SVD [۴] از این گونه الگوریتم‌ها هستند. در این مقاله، از مقداردهی مبتنی بر NMF استفاده می‌شود که در آن، ابتدا c زوج ماتریس اولیه (W, H) به صورت تصادفی و یا خروجی از یک الگوریتم ساده و سریع مانند NMF/ALS تولید می‌شوند. معمولاً ۱۰ تا ۲۰ زوج ماتریس کافی است. سپس عواملی مانند $(W^{i_{\min}}, H^{i_{\min}})$ که در آن،

$$i_{\min} = \operatorname{argmin}_{1 \leq i \leq c} \|A - W^i H^i\|_F^2$$

انتخاب می‌شوند. به این معنی که کمترین مقدار برای تابع هزینه مسئله NMF به‌ازای این عوامل به دست می‌آید. در نتیجه از آنها می‌توان به عنوان عوامل ماتریسی اولیه استفاده کرد. برای هر الگوریتم تکراری، از جمله الگوریتم‌های تکراری NMF باید شرط خاتمه‌ای وجود داشته باشد. علاوه بر معیارهای توقفی که در مرجع [۵] برای NMF آمده‌اند، معیارهای قوی‌تر دیگری توسط لین برای رهیافت مبتنی بر روش تصویر گردایان و توسط کیم و پارک برای رهیافت مبتنی بر روش مجموعه فعال معرفی شدند. البته شرط توقفی که لین [۱۷] معرفی کرد، متناسب با همان رهیافت مربوطه است؛ در حالی که شرط توقف پیشنهادی کیم و پارک، برای همه روش‌های محاسبه NMF از عمومیت لازم برخوردار است. ما نیز در این مقاله از معیار توقف کیم و پارک استفاده می‌کنیم. شرایط KKT را می‌توان به صورت فشرده

$$\min(W, \nabla_W f(W, H)) = 0, \quad \min(H, \nabla_H f(W, H)) = 0 \quad (۱۶.۲)$$

بیان کرد. به عبارت‌های سمت چپ علامت تساوی، مانده KKT اطلاق می‌شود. عبارت

$$\Delta = \frac{\Delta_0}{\sigma_W + \sigma_H} \quad (۱۷.۲)$$

مانده KKT نرمال شده است که در آن، σ_W و σ_H تعداد مؤلفه‌های W و H هستند که در حال حاضر، مقدار مانده KKT به‌ازای آنها صفر نشده است و

$$\begin{aligned} \Delta_0 = & \sum_{i=1}^m \sum_{a=1}^r |\min(W_{ia}, (\nabla_W f(W, H))_{ia})| \\ & + \sum_{b=1}^r \sum_{j=1}^n |\min(H_{bj}, (\nabla_H f(W, H))_{bj})|. \end{aligned} \quad (۱۸.۲)$$

بنابراین اگر Δ_1 مقدار Δ بعد از یک تکرار و ϵ یک تلورانس معین باشد، معیار همگرایی KKT که کیم و پارک ارائه کردند، به صورت زیر است:

$$\Delta \leq \epsilon \Delta_1. \quad (۱۹.۲)$$

۳. کاربردها و نتایج عددی

کاربردهای NMF از همان دهه ۹۰ میلادی با اقدامات پاترو با نام تجزیه مثبت ماتریسی آغاز شدند و سپس با اقدامات لی و سانگ مورد توجه ویژه‌ای قرار گرفتند. در حال حاضر، طیف گسترده‌ای از این کاربردها در مقاله‌های گوناگون مورد بحث هستند. NMF یک روش کاهش بُعد است که علاوه بر داشتن این توانایی، می‌توان از آن به عنوان یک خوشه‌بندی یا رده‌بندی نیز استفاده کرد. خوشه‌بندی یا رده‌بندی داده‌ها از روی مختصات جدید آنها شدنی خواهد بود. به عبارت دیگر، پس از اعمال NMF بر روی ماتریس داده‌ها و کاهش بُعد آن، انجام عمل خوشه‌بندی یا رده‌بندی در مختصات جدید ساده‌تر می‌شود. در این مقاله، به این دو کاربرد می‌پردازیم. در کاربردهای زیر، پارامتر r مقدار معینی دارد که توسط کاربر تعیین می‌شود. این مقدار می‌تواند با توجه به نوع کاربردها، تعداد خوشه‌ها و یا رده‌ها باشد.

۱.۳. خوشه‌بندی. مسئله خوشه‌بندی همان گروه‌بندی بدون نظارت داده‌ها در چندین گروه (خوشه) با ویژگی‌هایی مشابه است. اگر بحث تفسیر داده‌ها خصوصاً داده‌های نامفی مطرح نباشد، k-means معمول‌ترین روش خوشه‌بندی است [۱۲]. پارامتر k در k-means معادل با پارامتر r در NMF است. در عمل خوشه‌بندی با NMF، حداکثر تعداد درآیه در ستون ماتریس H شناسایی شده و داده متناظر با این ستون به خوشه متناظر با محل این عنصر، تخصیص می‌یابد. برای خوشه‌بندی داده‌ها با روش k-means، نمونه به عنوان افزایش اولیه تعیین می‌شوند به طوری که این نمونه‌ها، دارای کمترین فاصله اقلیدسی از میانگین کل باشند. در روش NMF نیز از همین افزایش برای مقداردهی اولیه عامل ماتریسی متناظر استفاده می‌گردد. اما برای شرط توقف روش k-means، وابستگی بهین بین داده‌ها کمتر از تلورانس δ در نظر گرفته [۱۲] و برای روش NMF از شرط توقف KKT استفاده می‌شود.

۱.۱.۳. خوشه‌بندی/اسناد. یکی از مسائل مورد بحث متن‌کاوی^۱، خوشه‌بندی داده‌های متنی با ویژگی‌های پنهان (مانند موضوع) می‌باشد. با فراخوانی ماتریس متن-سند در مثال ۲ و توضیحات آن، روشن است که تعداد خوشه‌ها $k = r = 2$ ، زیرا به وجود دو موضوع در اسناد آگاهی داریم. با این فرض، تعداد اسناد شناسایی شده برای موضوع google و موضوع football توسط هر دو روش k-means (با $\delta = 10^{-10}$) و NMF (با $\epsilon = 10^{-3}$) به ترتیب برابر با ۳ و ۱ به دست آمد. لذا هر دو روش در تشخیص موضوع تنها یک سند دچار ضعف بوده‌اند.

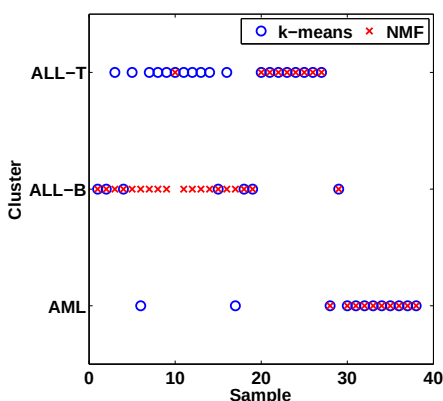
۲.۱.۳. تحلیل ژنی در ریزآرایه. داده‌های ژنی ALLAML^۲ متشکل از سرطان حاد لنفوم (ALL) با ۱۹ نمونه سلول از نوع B و ۸ نمونه سلول از نوع T و سرطان حاد لنفوم (AML) با ۱۱ نمونه سلول است. این داده‌ها شامل ۵۰۰۰ ژن هستند. برای ماتریس داده ریزآرایه $A \in \mathbb{R}^{5000 \times 38}$ نتایج خوشه‌بندی دو روش k-means با $\delta = 10^{-10}$ و NMF با $\epsilon = 10^{-6}$ و $k = r = 3$ در جدول ۱ ارائه شده

^۱text mining ^۲<http://www.broadinstitute.org/cgi-bin/cancer/datasets.cgi>

است. شکل ۲ نحوه خوشه‌بندی نمونه‌ها را در ۳ خوشه توسط این دو روش نمایش می‌دهد. روشن است که روش NMF در خوشه‌بندی داده‌های ژنی، عملکرد بهتری نسبت به روش k-means داشته است. مشاهده می‌شود که این روش تنها در تشخیص خوشه مربوط به نمونه‌های ۱۰ و ۲۹ ناتوان است.

جدول ۱. نتایج خوشه‌بندی داده‌های ژنی توسط k-means و NMF

نمونه ۱۱ AML	نمونه ۸ ALL-T	نمونه ۱۹ ALL-B	
۱۰	۸	۶	روش k-means
۱۰	۸	۱۸	روش NMF



شکل ۲. نحوه تفکیک داده‌های ژنی به سه خوشه توسط k-means و NMF

۲.۳. رده‌بندی. مسئله رده‌بندی داده‌ها عبارت است از تخصیص داده مفروض به یکی از چندین رده از پیش تعیین شده. رده‌بندی را می‌توان برحسب هر دو شیوه یادگیری با نظارت و بدون نظارت توصیف کرد که در این مقاله، از NMF به صورت یادگیری با نظارت استفاده می‌شود. بخش مهمی از روند شناسایی الگو، مسئله رده‌بندی در حوزه تصویر است که از آن می‌توان به دو مسئله استاندارد شناسایی چهره و شناسایی ارقام دست‌نویس اشاره کرد.

۱.۲.۳. شناسایی چهره. در شناسایی تصاویر چهره توسط NMF [۲۵]، همه داده‌های تصویری آموزشی و آزمایشی روی فضای تولیدشده توسط بردارهای پایه‌ای (ستون‌های ماتریس W) تصویر شده و بردارهای ویژگی جدید (ستون‌های ماتریس H) تولید می‌شوند. سپس عمل مقایسه به منظور رده‌بندی

چهره‌ها توسط این بردارهای ویژگی انجام می‌گیرد. معیارهای مختلفی برای این مقایسه وجود دارد که معمولاً از ساده‌ترین شکل آن یعنی فاصله اقلیدسی استفاده می‌شود. پایگاه داده (ORL)^۱ شامل ۴۰۰



شکل ۳. سه رده از پایگاه داده ORL

تصویر سیاه و سفید از چهره ۴۰ فرد متفاوت (۴۰ رده و هر یک شامل ۱۰ تصویر) در اندازه 92×112 با نورپردازی و حالات چهره مختلف مفروض است. در این مقاله، ۲۰ رده از این پایگاه انتخاب و از هر رده، ۴ تصویر آموزشی و ۶ تصویر آزمایشی با تبدیل اندازه به 16×16 در نظر گرفته شده است. به این ترتیب، بر اساس مثال ۱، ماتریس داده $A \in \mathbb{R}^{256 \times 80}$ تشکیل می‌گردد. ماتریس داده A از ۸۰ تصویر آموزشی تشکیل شده است که عوامل ماتریسی آن به صورت $A \approx W_1 H_1$ تعیین می‌شوند. همچنین ماتریس داده A' متشکل از ۱۲۰ چهره آزمایشی بر W_1 به صورت $A' \approx W_1 H_2$ تصویر می‌شود. اشاره شد که همه داده‌های تصویری برای تولید بردارهای ویژگی جدید، باید روی فضای تولیدشده توسط بردارهای پایه‌ای تصویر شوند. لذا H_1 و H_2 به ترتیب شامل بردارهای ویژگی جدید در فضای تولیدشده توسط W_1 برای بازسازی چهره‌های آموزشی و آزمایشی است. از آنجا که داده‌های آموزشی شامل ۴ تصویر از هر فرد است، متناظراً میانگین ۴ بردار ویژگی جدید (h_{m_i}) محاسبه شده و این عمل برای هر ۲۰ رده آموزشی تکرار می‌شود. سرانجام رده مربوط به چهره نامعلوم z ام به صورت زیر تعیین می‌گردد:

$$\min_{m_i} \|(H_2)_{:j} - h_{m_i}\|, \quad i = 1, \dots, 20, \quad j = 1, \dots, 120. \quad (1.3)$$

در جدول ۲ تعداد تشخیص درست چهره‌های آزمایشی ORL به روش NMF با رتبه $r = 20$ ارائه شده است. با توجه به این نتایج، میزان موفقیت روش NMF در رده‌بندی چهره‌ها ۸۴/۱۶ درصد است. با استفاده از روش PCA، این نتیجه ۹۴/۱۶ درصد برآورد شده است. به این ترتیب، روش NMF علاوه بر رفع مشکل تفسیر و نمایش مبتنی بر اجزاء، می‌تواند قابل رقابت با روش PCA نیز باشد. حتی این نتیجه را می‌توان برای NMF با معیارهای مقایسه‌ای بهتری در (۱.۳) بهبود داد [۲۵].

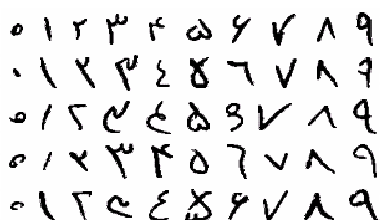
۲.۲.۳. شناسایی ارقام دست‌نویس. رده‌بندی تصاویر ارقام دست‌نویس [۱۲] با توجه به ماتریس پایه (شاخص) هر رده انجام می‌پذیرد. برخلاف مسئله قبل که روش NMF بر روی همه داده‌ها اعمال

^۱<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

جدول ۲. نتایج رده‌بندی چهره‌ها با روش NMF

۱۰	۹	۸	۷	۶	۵	۴	۳	۲	۱	شماره متناظر با ردهٔ چهره
۵	۶	۶	۶	۶	۵	۵	۶	۶	۵	تعداد تشخیص درست
۲۰	۱۹	۱۸	۱۷	۱۶	۱۵	۱۴	۱۳	۱۲	۱۱	شماره متناظر با ردهٔ چهره
۶	۵	۴	۲	۴	۲	۶	۴	۶	۶	تعداد تشخیص درست

شد، در اینجا برای داده‌های هر رده جداگانه به‌کار می‌رود. مجموعه ارقام هُدی، اولین مجموعهٔ بزرگ ارقام دست‌نویس فارسی است. این مجموعه متشکل از ۱۰ رده با ارقام سیاه و سفید ۰-۹ است که تعداد ارقام آموزشی و آزمایشی موجود برای هر رده از آن به‌ترتیب، ۶۰۰۰ و ۲۰۰۰ نمونه است.



شکل ۴. نمونه ارقام دست‌نویس هُدی

برای بررسی عددی، مجموعه ارقام آموزشی با ۱۰۰۰ نمونه و مجموعه ارقام آزمایشی با ۲۰۰ نمونه (۲۰ نمونه از هر رده) و با تبدیل اندازه به ۱۶×۱۶ انتخاب شده‌اند. ماتریس A_i متناظر با ردهٔ ارقام هم‌نوع تشکیل و مسئلهٔ $\rho_i = \min_x \|A_i x - d\|_2^2$ برای شناسایی رقم آزمایشی نامعلوم $d \in \mathbb{R}^{256}$ حل می‌شود. به این ترتیب، رقم d به ردهٔ i ام تخصیص خواهد یافت. اما می‌توان از روش NMF برای تجزیهٔ ماتریس A_i استفاده کرد. لذا برای رده‌بندی رقم آزمایشی d کافی است مسئلهٔ کوچک‌تر زیر حل شود:

$$\rho_i = \min_y \|W_i y - d\|_2^2, \quad y = H_i x, \quad i = 1, \dots, 10. \quad (2.3)$$

نتایج حاصل از رده‌بندی ارقام با استفاده از روش NMF با رتبهٔ ماتریسی $r = 10$ در جدول ۳ ارائه شده که بر اساس آن، میزان موفقیت این روش ۹۳ درصد است. اما میزان تشخیص درست با استفاده از روش SVD در رده‌بندی ارقام [۱۸]، برابر با ۹۳/۵ درصد است. با توجه به این نتیجه، شناسایی ارقام توسط روش NMF علاوه بر تفسیر صحیح، دارای دقت عملکرد قابل مقایسه‌ای با روش SVD است.

جدول ۳. نتایج رده‌بندی ارقام با روش NMF

ردۀ ارقام	۰	۱	۲	۳	۴	۵	۶	۷	۸	۹
تعداد تشخیص درست	۱۹	۲۰	۱۹	۲۰	۱۸	۱۹	۱۷	۱۹	۲۰	۱۵

بحث و نتیجه‌گیری

ارائه یک مدل کاهشی مناسب برای نمایش داده‌های نامنفی، برای درک و تفسیر آنها اهمیت دارد. NMF برخلاف روش‌هایی چون SVD و PCA، بردارهای وزن و پایه‌ای را همراه با قیدهای نامنفی به‌دست می‌آورد. در تقریب ماتریسی، این رویداد برای سایر روش‌ها به‌دلیل متعامد بودن بردارهای پایه‌ای و در نتیجه احتمال حضور درآیه‌های منفی اتفاق نمی‌افتد. روش NMF با توجه به قیدهای نامنفی می‌تواند نمایشی مبتنی بر اجزاء را به‌خوبی معنا بخشد. لذا این روش نسبت به سایر روش‌های نامبرده، به تغییرات جزئی داده‌ها حساس است. هرچند دستگاه‌های رایانه‌ای، تفسیر داده‌های با بُعد بالا را تسهیل می‌کنند، اما به هیچ وجه، انعطاف‌پذیری و شهود انسان را ندارند. به این معنی که ویژگی(های) مجموعه داده‌ها موجب می‌شوند که عملکرد روش‌ها برای آنها متفاوت باشد. مثلاً مشاهده شد که دقت روش NMF بیشتر از دقت روش‌های دیگر در کاربردهای مذکور نیست. از جمله بارزترین دلایل این رخداد، به‌کارگیری معیاری مانند فاصله اقلیدسی در این کاربردها است. هرچند نمی‌توان هیچ ملاک خاصی را در برتری معیارها نسبت به یکدیگر بیان کرد، اما می‌توان معیارهای جایگزین مناسب‌تری را نسبت به فاصله اقلیدسی در نظر گرفت که برای مطالعه این جایگزین‌ها، مطالعه مرجع [۲۵] پیشنهاد می‌شود. بارزترین مشکل NMF، عدم تضمین همگرایی به نقطه کمینه سراسری است. چون مسئله NMF محدب نیست، انتظار می‌رود چندین نقطه کمینه موضعی موجود باشد و لذا به‌دلیل عدم یکتایی جواب، نرمال کردن ستون‌های W در الگوریتم‌ها مطلوب به‌نظر می‌رسد. در الگوریتم‌های تکراری NMF چندین و چند تکرار لازم است تا همگرایی به نقطه بهین اتفاق افتد که این امر موجب می‌شود تا NMF در مقایسه با PCA زمان بیشتری را صرف کند. اما چون $r \ll \min\{m, n\}$ ، این امیدواری وجود دارد که این روش نیازمند فضای ذخیره‌سازی کمتری باشد.

چگونگی انتخاب مقدار r در PCA بر اساس بزرگی مقادیر ویژه قابل تعیین است. اما در حالت کلی، هیچ معیار مشخصی برای تعیین مقدار بهین آن برای NMF در دسترس نیست. غالباً این مقدار در کاربردهایی نظیر کاربردهای بیان شده در بخش قبل به‌عنوان تعداد خوشه و یا تعداد رده، از قبل مشخص است. اما در حالت کلی، نحوه انتخاب بهین این مقدار به‌صورت یک مسئله باز در حال مطالعه است. در مرجع [۲۵] به یکی از رهیافت‌های مورد بررسی در این زمینه پرداخته شده است.

قدردانی

نویسندگان بر خود لازم می‌دانند از داوران گرامی در ارائه دیدگاهشان، سردبیر محترم و سرکار خانم صمدیان برای پیگیری امور چاپ مقاله قدردانی نمایند.

مراجع

- [1] Benthem, M. H. V., Keenan, M. R., Fast algorithm for the solution of large-scale nonnegativity-constrained least squares problems, *J. Chemometrics*, **18** (2004), 441-450.
- [2] Berry, M. W., Browne, M., Langville, A. N., Pauca, V. P., Plemmons, R. J., Algorithms and applications for approximate nonnegative matrix factorization, *Computational Statistics and Data Analysis*, **1** (2007), 155-173.
- [3] Bertsekas, D. P., *Nonlinear Programming*, 2nd. edn., Athena Scientific, Belmont, Massachusetts, 1999.
- [4] Boutsidis C., Gallopoulos E., SVD-based initialization: A head start for nonnegative matrix factorization, *Pattern Recognition*, **41** (2008), 1350-1362.
- [5] Cichocki, A., Zdunek, R., Phan, A. H., Amari, S., *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, John Wiley & Sons, New York, 2009.
- [6] Golub, G. H., Van Loan, C. F., *Matrix Computations*, 3rd. edn., Johns Hopkins University Press, Baltimore and London, 1996.
- [7] Gonzales, E. F., Zhang, Y., *Accelerating the Lee-Seung Algorithm for Nonnegative Matrix Factorization*, Rice University, 2005.
- [8] Kim, H., Park, H., Nonnegative matrix factorization based on alternating non-negativity-constrained least squares and the active set method, *SIAM J. Matrix Anal. Appl.*, **30(2)** (2008), 713-730.
- [9] Kim, H., Park, H., Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis, *Bioinformatics*, **23** (2007), 1495-1502.
- [10] Kim, J., Park H., Toward faster nonnegative matrix factorization: A new algorithm and comparisons, *Proc. 8th IEEE Int. Conf. Data Mining*, 2008, 353-362.
- [11] Kim, J., Park H., Fast nonnegative matrix factorization: An active-set-like method and comparisons, *SIAM Journal on Scientific Computing (SISC)*, **33** (2011), 3261-3281.
- [12] Elden, L., *Matrix Methods in Data Mining and Pattern Recognition*, Fundamentals of Algorithms, Society for Industrial and Applied Mathematics, 2007.
- [13] Lawson, C. L., Hanson, R. J., *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, NJ, 1974.

- [14] Lee, D. D., Seung, H. S., Learning the parts of objects by non-negative matrix factorization, *Nature*, **401** (1999), 788-791.
- [15] Lee, D. D., Seung, H. S., Algorithms for non-negative matrix factorization, *Advances in Neural Information Processing*, **13**, 2001.
- [16] Lin, C. J., On the convergence of multiplicative update algorithms for non-negative matrix factorization, *IEEE Transactions on Neural Networks*, **18(6)** (2005), 1589-1596.
- [17] Lin, C. J., Projected gradient methods for nonnegative matrix factorization, *Neural Computation*, **19(10)** (2005), 2756-2779.
- [18] Mazack, M. J. M., *Non-negative Matrix Factorization with Applications to Handwritten Digit Recognition*, Working paper, University of Minnesota, 2009.
- [19] Paatero P., Tapper U., Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, *J. Environmetrics*, **5** (1994), 111-126.
- [20] Rasmus, B., Sijmen, D. J., A fast non-negativity-constrained least squares algorithm, *Journal of Chemometrics*, **11(5)** (1997), 393-401.
- [21] Skillicorn, D. B., *Understanding Complex Datasets: Data Mining with Matrix Decompositions*, Chapman & Hall/CRC, 2007.
- [22] Tan, P. N., Steinbach, M., Kumar, V., *Introduction to Data Mining*, Addison-Wesley, Boston, 2005.
- [23] Turk, M., Pentland, A., Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, **3(1)** (1991), 71-86.
- [24] Wall, M. E., Rechtsteiner, A., Rocha, L. M., Singular Value Decomposition and Principal Component Analysis, in: *A Practical Approach to Microarray Data Analysis*, Berrar, D. P., Dubitzky, W. and Granzow, M. (editors), 2003, 91-109.
- [25] Yun, X., *Non-negative Matrix Factorization for Face Recognition*, Hong Kong Baptist University, Ph.D. Thesis, 2007.

مهسا یوسفی: دانشگاه صنعتی سهند تبریز، گروه ریاضی کاربردی

رایانامه: mahsa.yousefi1987@gmail.com , m_yousefi@sut.ac.ir

منصور رزقی: دانشگاه تربیت مدرس، گروه علوم کامپیوتر

رایانامه: rezghi@modares.ac.ir

حجم گوی های واحد تعمیم یافته*

زیانفو وانگ

برگردان: رضا کهکشانی و رسول کاظمی نجف آبادی

۱. سرآغاز

هر دانشجویی با شکل های هندسی لوزی، استوانه، مربع، ستاره و گوی آشنایی دارد. اما از دیدگاه ما، این ها همگی گوی های تعمیم یافته هستند. به وسیله تبدیل های خطی و غیر خطی می توان گوی اقلیدسی استاندارد را به انواع گوی های عجیب تغییر شکل داد. هدف از این مقاله، ارائه دستوری واحد برای محاسبه حجم گوی های یکه تعمیم یافته در فضا های n - بعدی است. گوی واحد تعمیم یافته در $\mathbb{R}^n$ عبارت است از مجموعه

$$\mathbb{B}_{p_1 p_2 \dots p_n} = \{ \mathbf{x} = (x_1, \dots, x_n) : |x_1|^{p_1} + \dots + |x_n|^{p_n} \leq 1 \} \quad (1.1)$$

که در آن،

$$p_j > 0 \quad j = 1, 2, \dots, n.$$

اگر اعداد $p_1, p_2, \dots, p_n$ همگی بزرگتر یا مساوی ۱ باشند، گوی واحد $\mathbb{B}_{p_1 p_2 \dots p_n}$ محدب است. چون به ازای $1 < p < \infty$ تابع $x \mapsto |x|^p$ روی $[-1, 1]$ محدب نیست، $\mathbb{B}_{p_1 p_2 \dots p_n}$ نیز وقتی $n > 1$ ممکن است محدب نباشد. اگر قرار دهیم $1 \leq p_1 = p_2 = \dots = p_n = p$ ، گوی واحد فضای $\ell_p(\mathbb{R}^n)$ به دست می آید. گوی واحد فضای $\ell_2(\mathbb{R}^n)$ را با $\mathbb{B}$ نشان می دهیم. همان طور که در شکل ۱

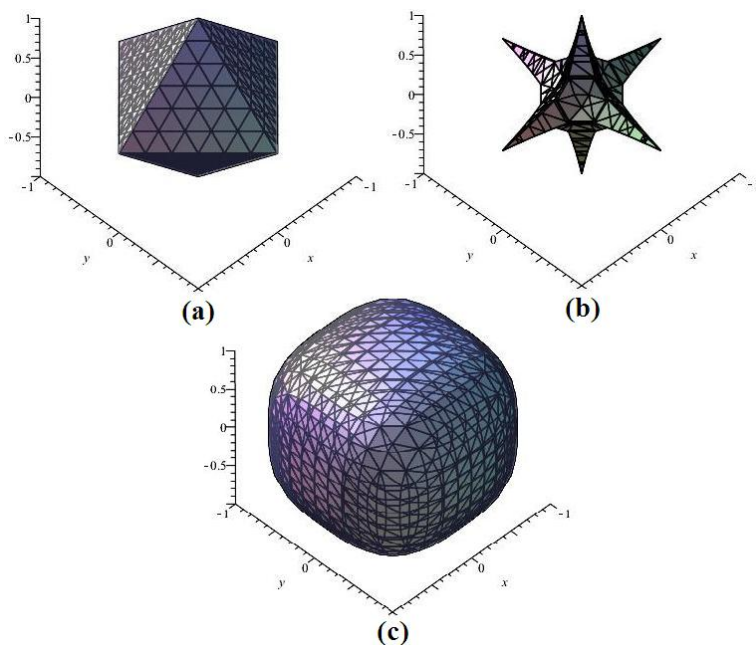
عبارات و کلمات کلیدی: فضای اقلیدسی؛ دستور تعویض متغیر؛ حجم گوی واحد؛ گوی واحد تعمیم یافته.
* نام و نشان مقاله اصلی از این قرار است:

Wang, Xianfu, Volumes of generalized unit balls, *mathematics magazine*, 78 (2005), 390-395.

با ارائه مثال‌هایی در $\mathbb{R}^3$ نشان داده شده است، با انتخاب مقادیر مختلف برای p_i ها، می‌توان در شکل گوی‌های تعمیم‌یافته تغییر زیادی ایجاد کرد.

۲. دستور حجم گوی واحد تعمیم‌یافته

مقاله‌ای از فولند^۱ [۵] انگیزه‌ای گردید تا دستوری واحد برای محاسبه حجم این گوی‌ها به دست آوریم. اگرچه دستورهای حجم گوی اقلیدسی استاندارد $\mathbb{B}$ و سادک^۲ مدت زیادی است که شناخته شده‌اند ([۴]، صفحه‌های ۲۰۸ و ۲۲۰)، این دستور واحد (نسبتاً) جدید است. حیرت‌آور است که گوی‌ها هر قدر هم عجیب به نظر برسند، حجم آنها را می‌توان با دستوری واحد به شرح زیر محاسبه نمود:



شکل ۱. (a) $|x_1| + |x_2| + |x_3| \leq 1$ ؛ (b) $|x_1|^{1/2} + |x_2|^{1/2} + |x_3|^{1/2} \leq 1$ و (c) $|x_1|^3 + |x_2|^{1/2} + |x_3|^3 \leq 1$

^۱Folland

^۲یک سادک n - بعدی در یک فضای اقلیدسی، مجموعه ترکیب‌های محدب $(n+1)$ بردار مستقل خطی مانند $p_1, p_2, \dots, p_n$ است؛ یعنی مجموعه همه نقاط $x = \lambda_0 p_0 + \dots + \lambda_n p_n$ به‌طوری که $\lambda_0 + \lambda_1 + \dots + \lambda_n = 1$ و برای هر i ، $\lambda_i \geq 0$. مثلاً یک مثلث همراه با درونش یک سادک دو بعدی است - م.

قضیه ۱.۲. فرض کنیم

$$p_j > 0 \quad j = 1, 2, \dots, n.$$

حجم گوی واحد تعمیم یافته $\mathbb{B}_{p_1 p_2 \dots p_n}$ در $\mathbb{R}^n$ برابر است با

$$2^n \frac{\Gamma(1 + \frac{1}{p_1}) \cdots \Gamma(1 + \frac{1}{p_n})}{\Gamma(\frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_n} + 1)}. \quad (1.2)$$

حجم آتبریک - هشتم مثبت، یعنی جایی که همه مؤلفه‌های x مثبت‌اند، با حذف عامل 2^n از دستور بالا به دست می‌آید.

دستور بالا شامل تابع گاما است که تعریف آن را برای خوانندگانی که ممکن است آشنا نباشند، یادآوری می‌کنیم. به ازای هر $0 < t < \infty$ تعریف می‌کنیم

$$\Gamma(t) := \int_0^\infty s^{t-1} e^{-s} ds.$$

این انتگرال برای هر $t > 0$ همگرا است و در مورد آن، از دو نتیجه استفاده می‌کنیم. یکی اینکه برای $u > 0$ و $v > 0$ داریم

$$\Gamma(u+1) = u\Gamma(u) \quad (2.2)$$

و دیگری این است که

$$\int_0^1 s^{u-1} (1-s)^{v-1} ds = \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)}. \quad (3.2)$$

اگرچه انتگرال مذکور در تعریف $\Gamma(t)$ به ازای $t \leq 0$ بی‌نهایت می‌شود، تساوی (۲.۲) یک دستور تحلیلی برای ادامه تعریف $\Gamma(t)$ روی $(-\infty, 0)$ فراهم می‌آورد. آن وقت تابع Γ تنها در نقاط

$$t = 0, -1, -2, \dots$$

ناپایسته است. جزئیات بیشتر را می‌توان در [۶] صفحه‌های ۳۴۴-۳۴۶ یافت.

اثبات. گام اول: با این واقعیت که

$$V(\mathbb{B}_{p_1 \dots p_n}) = \int_{\mathbb{B}_{p_1 p_2 \dots p_n}} 1 dx$$

و به‌کارگیری تعویض متغیری که گوی تعمیم‌یافته را به گوی استاندارد $\mathbb{B}$ تغییر شکل می‌دهد، شروع می‌کنیم.

گیریم $y_1 = x_1^{p_1/2}, y_2 = x_2^{p_2/2}, \dots, y_n = x_n^{p_n/2}$. برای تابع

$$\phi(\mathbf{y}) := (y_1^{\frac{2}{p_1}}, \dots, y_n^{\frac{2}{p_n}})$$

درمیان ماتریس ژاکوبی برابر است با

$$J\phi(\mathbf{y}) = \frac{2}{p_1} \cdots \frac{2}{p_n} y_1^{\frac{2}{p_1}-1} \cdots y_n^{\frac{2}{p_n}-1}.$$

خواننده می‌تواند برای ملاحظه جزئیات اثبات دستور تعویض متغیر که در گام بعدی برهان مورد نیاز است،

[۶] صفحه ۴۳۲ را مطالعه کند. با استفاده از این دستور، به دست می‌آوریم

$$\int_{\mathbb{B}_{p_1 \cdots p_n}} 1 d\mathbf{x} = \int_{\mathbb{B}} |J\phi(\mathbf{y})| d\mathbf{y} = \frac{2^n}{p_1 \cdots p_n} \int_{\mathbb{B}} |y_1|^{\frac{2}{p_1}-1} \cdots |y_n|^{\frac{2}{p_n}-1} d\mathbf{y}.$$

گام دوم: فرض کنیم

$$\alpha_j > -1 \quad j = 1, 2, \dots, n.$$

ادعا می‌کنیم

$$\int_{\mathbb{B}} |x_1|^{\alpha_1} \cdots |x_n|^{\alpha_n} d\mathbf{x} = \frac{\Gamma(\beta_1) \cdots \Gamma(\beta_n)}{\Gamma(\beta_1 + \cdots + \beta_n + 1)} \quad (4.2)$$

که در آن، برای هر $i = 1, \dots, n$ ، $\beta_i := (\alpha_i + 1)/2$. برای بررسی صحت این ادعا، یک دستور

بازگشتی به دست می‌آوریم. انتگرال موجود در (۴.۲) را با $I(\alpha_1, \dots, \alpha_n)$ نشان می‌دهیم. سپس آن را

به صورت یک انتگرال مکرر که با x_1 به عنوان بیرونی‌ترین متغیر شروع می‌شود، محاسبه می‌کنیم:

$$I(\alpha_1, \dots, \alpha_n) = \int_{-1}^1 |x_1|^{\alpha_1} \int_{x_1^2 + \cdots + x_n^2 \leq 1 - x_1^2} |x_2|^{\alpha_2} \cdots |x_n|^{\alpha_n} dx_2 \cdots dx_n dx_1.$$

انتگرال داخلی روی گوی به شعاع $r = \sqrt{1 - x_1^2}$ محاسبه می‌شود. مجدداً با تعویض متغیر

$$(x_2, \dots, x_n) = r(y_2, \dots, y_n),$$

به دست می‌آوریم

$$\begin{aligned} & \int_{x_1^2 + \cdots + x_n^2 \leq r^2} |x_2|^{\alpha_2} \cdots |x_n|^{\alpha_n} dx_2 \cdots dx_n \\ &= \int_{y_1^2 + \cdots + y_n^2 \leq 1} r^{(n-1) + \alpha_2 + \cdots + \alpha_n} |y_2|^{\alpha_2} \cdots |y_n|^{\alpha_n} dy_2 \cdots dy_n \end{aligned}$$

که نتیجه می‌دهد $I(\alpha_1, \dots, \alpha_n)$ برابر است با

$$\begin{aligned} & \int_{-1}^1 |x_1|^{\alpha_1} (1 - x_1^2)^{\frac{(n-1)}{2} + \frac{\alpha_2 + \dots + \alpha_n}{2}} \int_{y_2^2 + \dots + y_n^2 \leq 1} |y_2|^{\alpha_2} \dots |y_n|^{\alpha_n} dy_2 \dots dy_n dx_1 \\ &= 2 \int_0^1 x_1^{\alpha_1} (1 - x_1^2)^{\frac{(n-1)}{2} + \frac{\alpha_2 + \dots + \alpha_n}{2}} dx_1 \cdot \int_{y_2^2 + \dots + y_n^2 \leq 1} |y_2|^{\alpha_2} \dots |y_n|^{\alpha_n} dy_2 \dots dy_n \\ &= \int_0^1 (x_1^2)^{\frac{\alpha_1-1}{2}} (1 - x_1^2)^{\frac{(\alpha_2 + \dots + \alpha_n + n+1)}{2} - 1} d(x_1^2) \int_{y_2^2 + \dots + y_n^2 \leq 1} |y_2|^{\alpha_2} \dots |y_n|^{\alpha_n} dy_2 \dots dy_n. \end{aligned}$$

لذا بنابر (۳.۲)،

$$I(\alpha_1, \dots, \alpha_n) = \frac{\Gamma((\alpha_1 + 1)/2) \Gamma((\alpha_2 + \dots + \alpha_n + n + 1)/2)}{\Gamma((\alpha_1 + \dots + \alpha_n + n + 2)/2)} \cdot I(\alpha_2, \dots, \alpha_n)$$

این رابطه، یک دستور بازگشتی میان $I(\alpha_1, \dots, \alpha_n)$ و $I(\alpha_2, \dots, \alpha_n)$ به دست می‌دهد. با $n - 1$ بار به کار بردن این دستور بازگشتی و پس از ساده سازی‌های لازم، به دست می‌آوریم

$$I(\alpha_1 \dots \alpha_n) = \frac{\Gamma(\frac{\alpha_1+1}{2}) \dots \Gamma(\frac{\alpha_{n-1}+1}{2}) \frac{\alpha_n+1}{2} \Gamma(\frac{\alpha_n+1}{2})}{\Gamma(\frac{\alpha_1+\dots+\alpha_n+n}{2} + 1)} \cdot I(\alpha_n). \quad (5.2)$$

اما

$$I(\alpha_n) = \int_{x^2 \leq 1} |x|^{\alpha_n} dx = 2 \int_0^1 x^{\alpha_n} dx = \frac{2}{\alpha_n + 1}$$

که با قرار دادن این مقدار در (۵.۲)، تساوی (۴.۲) نتیجه می‌شود.

گام سوم: اگر برای $\alpha_i = 2/p_i - 1, i = 1, \dots, n$ ، نتیجه می‌دهد که

$$I(2/p_1 - 1, \dots, 2/p_n - 1) = \frac{\Gamma(1/p_1) \dots \Gamma(1/p_n)}{\Gamma(1/p_1 + \dots + 1/p_n + 1)}.$$

لذا

$$\begin{aligned} V(\mathbb{B}_{p_1 \dots p_n}) &= 2^n \frac{1}{p_1} \dots \frac{1}{p_n} I(2/p_1 - 1, \dots, 2/p_n - 1) \\ &= 2^n \frac{\Gamma(1 + 1/p_1) \dots \Gamma(1 + 1/p_n)}{\Gamma(1/p_1 + \dots + 1/p_n + 1)}. \end{aligned}$$

حجم قسمت مثبت دستگاه مختصات در $\mathbb{R}^n$ از این دستور حاصل می‌شود، زیرا 2^n تا ابر یک - هشتم در

□

$\mathbb{R}^n$ وجود دارد.

۳. حالت‌های خاص

ممکن است استدلال کنید که مقدار p_i در (۱.۱) نمی‌تواند بی‌نهایت باشد، اما خواننده می‌تواند حالت حدی را در نظر بگیرد. می‌نویسیم

$$x^\infty = \begin{cases} 0 & 0 \leq x < 1 \\ 1 & x = 1 \end{cases}$$

و با جدا کردن برخی حالت‌های خاص ادامه می‌دهیم (برای دانشجویان حسابان):
(الف) $p_i = +\infty$: چون Γ بر $(0, +\infty)$ پیوسته است، $V(\mathbb{B}_{p_1 \dots p_n})$ برابر است با

$$2 \cdot 2^{n-1} \frac{\Gamma(1 + 1/p_1) \cdots \Gamma(1 + 1/p_{i-1}) \Gamma(1 + 1/p_{i+1}) \cdots \Gamma(1 + 1/p_n)}{\Gamma(1/p_1 + \cdots + 1/p_{i-1} + 1/p_{i+1} + \cdots + 1/p_n + 1)}.$$

به‌ویژه وقتی $p_1 = \cdots = p_n = +\infty$ ، حجم گوی برابر با 2^n است و شکل آن یک اَبرمکعب n -بعدی است (به‌استثنای بخشی از مرکز که آنجا حداقل دوتا از x_i ‌ها همزمان برابر ۱ هستند). وقتی $p_1 = p_2 = 2$ و $p_3 = \infty$ ، گوی تعمیم‌یافته، یک استوانهٔ مستدیر در $\mathbb{R}^3$ است.
(ب) وقتی $0 < p_1 = p_2 = \cdots = p_n = p$ داریم

$$V(\mathbb{B}_{p_1 \dots p_n}) = 2^n \frac{(\Gamma(1 + 1/p))^n}{\Gamma(n/p + 1)} = \frac{(2/p)^n (\Gamma(1/p))^n}{(n/p) \Gamma(n/p)}.$$

به یاد آورید که $\Gamma(1) = 1$ ، $\Gamma(1/2) = \pi^{1/2}$ و $\Gamma(n) = (n-1)!$. به‌ازای $p = 2$ ، گوی تعمیم‌یافته همان گوی اقلیدسی استاندارد با حجم $\frac{2^n \pi^{n/2}}{n \Gamma(n/2)}$ است. به‌ازای $p = 1$ ، گوی تعمیم‌یافته

$$\{(x_1, \dots, x_n) : |x_1| + \cdots + |x_n| \leq 1\}$$

یک لوزی n -بعدی با حجم $\frac{2^n}{n!}$ است. به‌ازای $p = \frac{1}{2}$ ، حجم گوی تعمیم‌یافته برابر با $\frac{2^{2n}}{(2n)!}$ است و به شکل یک ستاره n -بعدی است. این‌ها دو تا از گوی‌های نشان داده شده در شکل ۱ هستند. شگفت این‌که به‌ازای $0 < p < \infty$ درمی‌یابیم که وقتی n بزرگتر می‌شود، حجم گوی n -بعدی کمتر می‌شود و

$$\lim_{n \rightarrow \infty} V(\mathbb{B}_{p,p,\dots,p}) = \lim_{n \rightarrow \infty} \frac{(2/p)^n (\Gamma(1/p))^n}{(n/p) \Gamma(n/p)} = 0.$$

در اینجا دستور استرلینگ را به‌کار برده‌ایم:

$$\Gamma(x) \sim \sqrt{2\pi} x^{x-1/2} e^{-x}$$

که در آن، نماد $\sim$ بدین معنی است که وقتی $x \rightarrow \infty$ ، نسبت مقادیرهای چپ و راست به ۱ میل می‌کند ([۶] صفحه ۳۵۳).

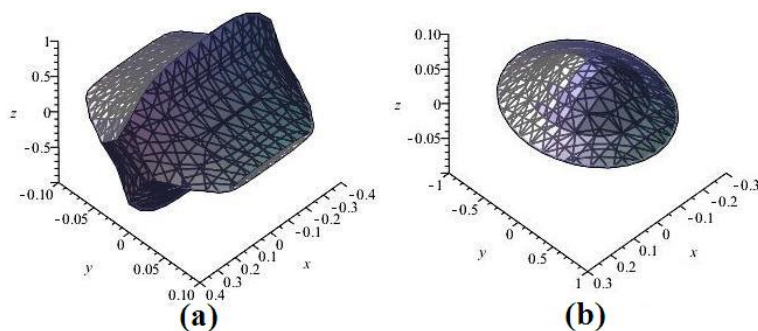
(پ) برای بیضی‌گون

$$\left\{ (x_1, \dots, x_n) : \frac{|x_1|^{p_1}}{a_1^{p_1}} + \dots + \frac{|x_n|^{p_n}}{a_n^{p_n}} \leq 1 \right\}$$

با $a_i > 0$ ، یک تبدیل خطی ساده آن را به صورت (۱) تبدیل می‌کند و بر مبنای قضیه، حجم آن از دستور

$$a_1 \dots a_n \times \frac{\Gamma(1 + 1/p_1) \dots \Gamma(1 + 1/p_n)}{\Gamma(1/p_1 + \dots + 1/p_n + 1)}$$

به‌دست می‌آید. برای اینکه ایده‌ای از شکل بیضی‌گون در $\mathbb{R}^3$ به‌دست آورید، در شکل ۲، دو بیضی‌گون را در $\mathbb{R}^3$ نشان داده‌ایم.



شکل ۲. (a) $16|x_1|^4 + 3|x_2|^{1/2} + |x_3| \leq 1$ و (b) $16|x_1|^2 + |x_2|^2 + 4|x_3|^{1/2} \leq 1$

پس از اینکه این نتیجه را به‌دست آوردم، دکتر ج. م. بوروین از دانشگاه سیمون فریزر به اینجانب اطلاع داد که دیریکله، ریاضیدان آلمانی قرن نوزدهم، با استفاده از روشی متفاوت، نتیجه‌ای مشابه به‌دست آورده است ([۳] صفحه‌های ۱۵۳-۱۵۹). یک برهان فاقد استقرا برای دستور حجم گوی واحد در $\ell_p(\mathbb{R}^n)$ ، از طریق تبدیل لاپلاس، توسط بوروین و بایلی در ([۲] صفحه‌های ۱۹۵-۱۹۷) ارائه شده است. به‌طور مشابه، با استفاده از تبدیل لاپلاس می‌توان برهانی فاقد استقرا برای دستور حجم گوی‌های تعمیم‌یافته (۱.۲) استخراج نمود. در خاتمه، باید ذکر کنیم که ویژگی‌های بیشتری درباره‌ی تابع گاما و حجم گوی‌های اقلیدسی را می‌توان در استرومبرگ ([۷] صفحه‌های ۳۹۴-۳۹۵) یافت.

مراجع

- [1] Baker, J. A., Integration over spheres and the divergence theorem for balls, *Amer. Math. Monthly*, **104** (1997), 36-47.
- [2] Borwein, J. M., Bailey, D. H., *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, A K Peters, Natick Mass., 2003.
- [3] Edwards, J., *A Treatise on the Integral Calculus*, Vol. II, Chelsea Publishing Company, New York, 1992.
- [4] Fleming, W., *Functions of Several Variables*, Springer-Verlag, New York, 1997.
- [5] Folland, G. B., How to integrate a polynomial over a sphere, *Amer. Math. Monthly*, **108** (2001), 446-448.
- [6] ———, *Advanced Calculus*, Prentice Hall, Upper Saddle River, NJ, 2002.
- [7] Stromberg, K. R., *An Introduction to Classical Real Analysis*, Wadsworth International Mathematics Series, 1981.

رضا کهکشانی: دانشکده علوم ریاضی، دانشگاه کاشان

رایانامه: kahkeshanireza@kashanu.ac.ir

رسول کاظمی نجف‌آبادی: دانشکده علوم ریاضی، دانشگاه کاشان

رایانامه: r.kazemi@kashanu.ac.ir

رابطه‌ای معادل با رابطه فیثاغورس و سه‌تایی فیثاغورسی وارون

جواد بهبودیان، صدیقه بلادی، و مهرداد سیمکانی

چکیده

در این مقاله، یک رابطه کم و بیش معروف معادل با رابطه فیثاغورس ارائه می‌گردد. معادل بودن این رابطه با رابطه فیثاغورس با روش‌های هندسی، جبری و مثلثاتی ثابت می‌شود. سپس تعریف سه‌تایی فیثاغورسی اولیه بیان خواهد شد و در پایان در برابر هر سه‌تایی فیثاغورسی اولیه، یک سه‌تایی فیثاغورسی وارون معرفی می‌شود.

۱. مقدمه

این مقاله را به یاد و پاس خدمات گران‌بهای استاد پرویز شهریاری می‌نویسیم که در تاریخ ۹۱/۲/۲۲ چشم از جهان فرو بست. شهریاری در طول ۸۶ سال زندگی خود، با نوشتن کتاب‌ها و مقاله‌های گوناگون و سخنرانی‌های متنوع در سمینارهای ریاضی، آموزش ریاضی را در کشور گسترش داد. بسیاری از دانش‌آموختگان پنجاه سال اخیر در ایران، در دوران تحصیل خود، کم و بیش از آثار شهریاری بهره‌مند شده‌اند. شهریاری، این معلم توانا، به ریاضی سنتی و دبیرستانی از جمله هندسه، جبر، مثلثات و حساب استدلالی علاقه و توجه ویژه‌ای داشت.

در این مقاله، مثلث $\triangle ABC$ را با دو زاویه حاده $\angle B$ و $\angle C$ و اضلاعی به طول a ، b و c در نظر می‌گیریم. فرض می‌کنیم h طول ارتفاع AH باشد. اگر $\angle A = 90^\circ$ ، در $\triangle ABC$ رابطه زیبا و تاریخی فیثاغورس، یعنی

$$a^2 = b^2 + c^2 \quad (1.1)$$

عبارات و کلمات کلیدی. رابطه فیثاغورس؛ سه‌تایی فیثاغورسی اولیه؛ سه‌تایی فیثاغورسی وارون.

رابطه‌ای معادل با رابطه فیثاغورس و سه‌تایی فیثاغورسی وارون. _____ ۱۰۰

برقرار است. این رابطه، اثبات‌های گوناگون و کاربردهای فراوان دارد (به مرجع‌های [۱] و [۳] نگاه کنید). در ضمن، می‌دانیم که اگر در مثلثی (۱.۱) برقرار باشد، آن مثلث قائم‌الزاویه است. در هر مثلث قائم‌الزاویه رابطه کم و بیش معروف

$$\frac{1}{h^2} = \frac{1}{b^2} + \frac{1}{c^2} \quad (2.1)$$

نیز برقرار است. این رابطه به کمک مساحت مثلث قائم‌الزاویه، S ، به شرح زیر به آسانی ثابت می‌شود:

$$2S = bc = ah,$$

$$b^2 c^2 = a^2 h^2,$$

$$b^2 c^2 = (b^2 + c^2) h^2. \quad (3.1)$$

حال دو طرف (۳.۱) را بر $b^2 c^2 h^2$ تقسیم می‌کنیم تا (۲.۱) به دست آید.

در این مقاله، معادل بودن رابطه‌های (۱.۱) و (۲.۱) را بررسی می‌کنیم. به بیان دیگر، نشان می‌دهیم رابطه (۲.۱) شرط لازم و کافی برای رابطه (۱.۱) است. در بند قبل نشان دادیم رابطه (۲.۱) شرط لازم برای رابطه (۱.۱) است. در بخش ۲ کافی بودن رابطه (۲.۱) را با سه روش اثبات می‌کنیم. سرانجام در بخش سه، به کمک رابطه (۲.۱) در برابر هر سه‌تایی فیثاغورسی اولیه، یک سه‌تایی فیثاغورسی وارون معرفی خواهیم کرد.

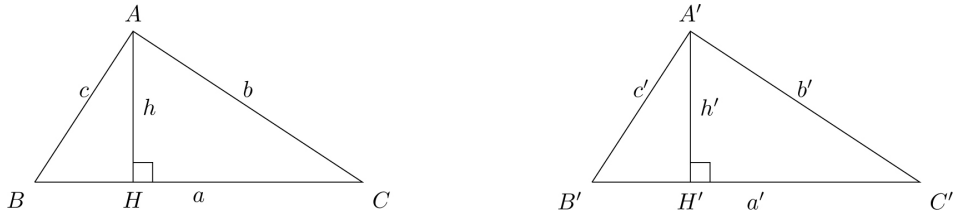
۲. اثبات کفایت

در این بخش، با سه روش هندسی، جبری و مثلثاتی ثابت می‌کنیم که رابطه (۲.۱) شرط کافی برای رابطه (۱.۱) است تا نشان دهیم در رابطه (۲.۱) کاربرد هندسه، جبر و مثلثات نهفته است.

۱.۲. روش هندسی. با روش هندسی نشان می‌دهیم که اگر رابطه (۲.۱) در $\triangle ABC$ برقرار باشد، این مثلث قائم‌الزاویه است. برای اثبات، مثلث قائم‌الزاویه $A'B'C'$ را طوری رسم می‌کنیم که $\angle A' = 90^\circ$ ، $A'B' = c' = c$ ، $A'C' = b' = b$ و $A'H' = h'$ ارتفاع وارد بر $B'C'$ باشد (شکل ۱).

در مثلث قائم‌الزاویه $A'B'C'$ بنابر شرط لازم، داریم

$$\frac{1}{h'^2} = \frac{1}{b^2} + \frac{1}{c^2}. \quad (1.2)$$



شکل ۱

از مقایسه (۲.۱) با (۱.۲)، $h = h'$ نتیجه می‌شود. از طرف دیگر، می‌دانیم هرگاه وتر و یک ضلع از مثلث قائم‌الزاویه‌ای با وتر و یک ضلع از مثلث قائم‌الزاویه دیگر برابر باشند، دو مثلث هم‌نهشت هستند. بنابراین $\triangle ABH \cong \triangle A'B'H'$ و در نتیجه $\angle BAH = \angle B'A'H'$. همچنین

$$\triangle ACH \cong \triangle A'C'H'$$

و از این رو $\angle CAH = \angle C'A'H'$. پس

$$\angle A = \angle A' = 90^\circ$$

و لذا $\triangle ABC$ قائم‌الزاویه است.

۲.۲. روش جبری. پیش از اینکه با استفاده از روش جبری نشان دهیم رابطه (۲.۱) شرط کافی برای رابطه فیثاغورس است، یک لم ساده را بیان می‌کنیم.

لم ۱.۲. شرط لازم و کافی برای قائم‌الزاویه بودن $\triangle ABC$ (شکل ۱) این است که

$$AH^2 = BH \times CH.$$

□

اثبات. با استفاده از تشابه دو مثلث ABH و ACH ، اثبات می‌شود.

اینک (۲.۱) را به صورت

$$0 = b^2 c^2 - b^2 h^2 - c^2 h^2$$

می‌نویسیم. به دو طرف تساوی بالا h^4 را می‌افزاییم تا رابطه

$$h^4 = h^4 + b^2 c^2 - b^2 h^2 - c^2 h^2$$

و به دنبال آن با فاکتورگیری از رابطه به دست آمده

$$h^4 = (b^2 - h^2)(c^2 - h^2)$$

رابطه‌ای معادل با رابطه فیثاغورس و سه‌تایی فیثاغورسی وارون. _____ ۱۰۲

نتیجه شود. از این‌رو با استفاده از دو مثلث قائم‌الزاویه ABH و ACH در شکل ۱، می‌نویسیم

$$AH^2 = BH^2 \times CH^2$$

یا

$$AH^2 = BH \times CH.$$

در نتیجه بنابر لم بالا، $\angle A = 90^\circ$.

۳.۲. روش مثلثاتی. با توجه به راه‌حلی که در ادامه ارائه خواهیم داد، می‌بینیم که روش مثلثاتی ساده‌ترین روش است. نخست (۲.۱) را به صورت

$$\frac{h^2}{b^2} + \frac{h^2}{c^2} = 1$$

می‌نویسیم. اینک با استفاده از دو مثلث قائم‌الزاویه $\triangle ABH$ و $\triangle ACH$ در شکل ۱، چون

$$\sin^2 C + \sin^2 B = 1,$$

داریم

$$\sin^2 C + \sin^2 B = \cos^2 B + \sin^2 B$$

و در نتیجه $\sin^2 C = \cos^2 B$. از این‌رو $\sin C = \cos B$ و لذا $\angle C = 90^\circ - \angle B$. بنابراین $\angle A = 90^\circ$.

۳. سه‌تایی فیثاغورسی اولیه و سه‌تایی فیثاغورسی وارون

در این بخش، اندکی حساب استدلالی به‌کار می‌بریم و دو تعریف مهم سه‌تایی فیثاغورسی اولیه و سه‌تایی فیثاغورسی وارون را بیان می‌کنیم.

تعریف ۱.۳. سه‌تایی فیثاغورسی، یک مجموعه سه عضوی از اعداد صحیح مانند a, b, c است به‌طوری که $a^2 = b^2 + c^2$. این سه‌تایی، سه‌تایی فیثاغورسی اولیه خوانده می‌شود هرگاه بزرگترین مقسوم‌علیه مشترک a, b, c و برابر با یک باشد.

گیریم a, b, c یک سه‌تایی فیثاغورس دلخواه باشد و $d = \gcd(a, b, c)$. اگر قرار دهیم $a = da', b = db', c = dc'$ داریم

$$b'^2 + c'^2 = \frac{b^2 + c^2}{d^2} = \frac{a^2}{d^2} = a'^2$$

با این ویژگی که $\gcd(a', b', c') = 1$. پس سه‌تایی a' ، b' و c' یک سه‌تایی فیثاغورسی اولیه است. بنابراین هدفمان را یافتن کلیه سه‌تایی‌های فیثاغورسی اولیه قرار می‌دهیم، زیرا هر سه‌تایی فیثاغورسی را می‌توان از یک سه‌تایی فیثاغورسی اولیه به دست آورد. همچنین بدون کم شدن از کلیت می‌توان سه‌تایی فیثاغورسی اولیه a ، b و c را مثبت در نظر گرفت.

در هر سه‌تایی فیثاغورسی اولیه، هر دوتایی نسبت به هم اول هستند؛ مثلاً $\gcd(a, b) = 1$ ، زیرا اگر a و b بر عددی صحیح بخش پذیر باشند، با توجه به $a^2 = b^2 + c^2$ ، c هم باید بر آن عدد بخش پذیر باشد که چنین نیست.

قضیه ۲.۳ که برهان جالب آن را برتون در [۲] نوشته است، تمام سه‌تایی‌های فیثاغورسی اولیه را تعیین می‌کند.

قضیه ۲.۳. تمام سه‌تایی‌های فیثاغورسی اولیه a ، b و c عبارت‌اند از

$$a = s^2 + t^2, \quad b = 2st, \quad c = s^2 - t^2$$

که در آن، برای اعداد صحیح s و t داریم $s > t > 0$ و $\gcd(s, t) = 1$. به علاوه، یکی از دو عدد s و t فرد و دیگری زوج است.

پیش از آن که سه‌تایی فیثاغورسی وارون را تعریف کنیم، این نکته را خاطرنشان می‌کنیم که در مثلث قائم‌الزاویه ABC (شکل ۱) با سه‌تایی فیثاغورسی اولیه a ، b و c ، ارتفاع $h = AH$ نمی‌تواند عددی صحیح باشد، زیرا

$$\gcd(a, b) = 1, \quad \gcd(a, c) = 1$$

و

$$h = \frac{bc}{a} = \frac{2st(s^2 - t^2)}{s^2 + t^2}.$$

اینک سه‌تایی فیثاغورسی اولیه a ، b و c را با سه‌تایی فیثاغورسی

$$a' = (s^2 + t^2)^2, \quad b' = 2st(s^2 + t^2), \quad c' = (s^2 - t^2)(s^2 + t^2)$$

متناظر می‌کنیم. در نتیجه

$$h' = \frac{b'c'}{a'} = 2st(s^2 - t^2)$$

عدد صحیح خواهد شد.

رابطه‌ای معادل با رابطه فیثاغورس و سه‌تایی فیثاغورسی وارون. _____ ۱۰۴

تعریف ۳.۳. سه‌تایی فیثاغورسی وارون، یک مجموعه سه عضوی از اعداد صحیح

$$h' = 2st(s^2 - t^2), \quad b' = 2st(s^2 + t^2), \quad c' = (s^2 - t^2)(s^2 + t^2)$$

است که در رابطه

$$\frac{1}{h'^2} = \frac{1}{b'^2} + \frac{1}{c'^2}$$

صدق کنند و در آن، اعداد صحیح $s > t > 0$ چنان باشند که $\gcd(s, t) = 1$ و یکی از دو عدد s و t فرد و دیگری زوج باشد. می‌توان نشان داد که $\gcd(h', b', c') = 1$.

مثال ۴.۳. فرض می‌کنیم $s = 2$ و $t = 1$. بنابراین سه‌تایی فیثاغورسی اولیه، ۵، ۴ و ۳ خواهد بود. اگر این سه‌تایی را در $s^2 + t^2 = 5$ ضرب کنیم، سه‌تایی فیثاغورسی ۲۵، ۲۰ و ۱۵ متناظر با ۵، ۴ و ۳ به دست می‌آید. پس با این سه‌تایی، $h' = \frac{2 \times 15}{5} = 12$ و در نتیجه سه‌تایی فیثاغورسی وارون، ۱۲، ۲۰ و ۱۵ می‌شود. برای این سه‌تایی داریم

$$\frac{1}{12^2} = \frac{1}{20^2} + \frac{1}{15^2}.$$

مثال ۵.۳. فرض می‌کنیم $s = 3$ و $t = 2$. بنابراین سه‌تایی فیثاغورسی اولیه، ۱۳، ۱۲ و ۵ $b = 12, a = 13, c = 5$ خواهد بود. اینک اگر این سه‌تایی را در $s^2 + t^2 = 13$ ضرب کنیم، سه‌تایی فیثاغورسی $a' = 169, b' = 156, c' = 65$ و سه‌تایی فیثاغورسی وارون $h' = 60, b' = 156, c' = 65$ حاصل می‌شود.

در پایان، به عنوان نتیجه به نکته‌ای جالب اشاره می‌کنیم. فرض کنیم در $\triangle ABC$ ، زاویه B منفرجه باشد. بدیهی است که ارتفاع AH بیرون از مثلث قرار می‌گیرد. با یکی از روش‌های بالا، مثلاً روش مثلثاتی، به آسانی می‌توان ثابت کرد که رابطه (۲.۱) شرط لازم و کافی برای $\angle B - \angle C = 90^\circ$ است. بنابراین به طور کلی رابطه (۲.۱) در مثلث ABC شرط لازم و کافی برای $\angle B + \angle C = 90^\circ$ (یا $\angle C$ و $\angle B$ هر دو حاده) یا $\angle B - \angle C = 90^\circ$ ($\angle B$ منفرجه) است.

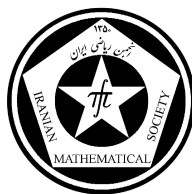
مراجع

- [۱] بهبودیان، جواد، نگاهی به قضیه فیثاغورس، رشد آموزش ریاضی، شماره ۱۶ (۱۳۶۶)، صص. ۲۰-۲۶.
- [2] Burton, D. M., *Elementary Number Theory*, 5th edn., McGraw Hill, 2002.
- [3] Loomis, E. S., *The Pythagorean Proposition*. The National Council of Teachers of Mathematics, 1968.

جواد بهبودیان: دانشگاه شیراز، دانشکده علوم ریاضی
رایانامه: behboodian@susc.ac.ir

صدیقه بلادی: آموزش و پرورش ناحیه ۱ شیراز، دبیر ریاضی
رایانامه: Sedigheh.Beldadi@yahoo.com

مهرداد سیمکانی: دانشگاه میشیگان فلینت، بخش ریاضی
رایانامه: simkani@umich.edu



FARHANG va ANDISHE-ye RIYĀZI

An Expository Journal of the
Iranian Mathematical Society

ISSN 1022-6443

Vol. 35, No. 2, Fall 2016 & Winter 2017

Editor-in-Chief

Ruhollah Jahanipur, University of Kashan
jahanipu@kashanu.ac.ir

Editorial Board

Mohammad Gholamzadeh-Mahmoudi, Sharif University of Technology, Tehran
mmahmoudi@sharif.ir

Behnam Hashemi, Shiraz University of Technology, Shiraz
hoseynhashemi@gmail.com

Alireza Jamali, Kharazmi University, Tehran
jamali@khu.ac.ir

Ruhollah Jahanipur, University of Kashan, Kashan
jahanipu@kashanu.ac.ir

Saeed Maghsoudi, University of Zanjan, Zanjan
s_maghsodi@znu.ac.ir

Mojgan Mahmoudi, Shahid Beheshti University, Tehran
m-mahmoudi@sbu.ac.ir

Rostam Mohammadian, Shahid Chamran University of Ahvaz, Ahvaz
mohamadian_r@scu.ac.ir

Alimohammad Nazari, Arak University, Arak
a-nazari@araku.ac.ir

Abolfazl Rafiepour, Shahid Bahonar University of Kerman, Kerman
drafiepour@gmail.com

Editorial Office Manager

F. Samadian, Iranian Mathematical Society
iranmath@ims.ir

P. O. Box 13145-418

Tehran - Iran

Tel: 88808855, 88807795, 88807775

e-mail: iranmath@ims.ir

web: mct.iranjournals.ir

<http://www.ims.ir>