مجله‌ی ریاضی شریف
دوره‌ی سوم، شماره‌ی دوم (بهار ۱۴۰۳)

هندسه‌ی فضای اندازه‌های احتمال

امین طالبی*

چکیده. می‌توان فضای اندازه‌های احتمال بر روی یک فضای متریک فشرده را به متریکی به نام متریک وسرشتاین [۱] مجهز کرد. در این مقاله سعی داریم تا به توصیفی مختصر از هندسه‌ی این فضا بپردازیم. ابتدا به مطالعه هندسه و توپولوژی فضای اندازه‌های احتمالی که بر روی متناهی نقطه جمع شده‌اند می‌پردازیم، سپس نشان می‌دهیم همان گونه که به هر عدد حقیقی می‌توان به عنوان یک دنباله کوشی از اعداد گویا نگاه کرد، به هر اندازه‌ی احتمال می‌توان به چشم یک دنباله کوشی از اندازه‌های روی متناهی نقطه از فضا نگریست. انگیزه‌ی ما برای انجام این کار، کاربرد آن در سیستم‌های دینامیکی و نظریه‌ی ارگودیک است.

۱. سرآغاز

یکی از سوالات محوری در شاخه سیستم‌های دینامیکی مطالعه موجوداتی است که تحت دینامیک، ناوردا می‌مانند. اندازه‌های احتمال از جمله موجوداتی هستند که می‌توان ناوردا بودن آن‌ها را تحت اثر دینامیک بررسی کرد. در بسیاری از مسائل موجود در سیستم‌های دینامیکی جواب به این سوال که چه اندازه‌های احتمالی تحت اثر دینامیک ناوردا می‌مانند، و همچنین توصیف اندازه‌های ناوردا، از اهمیت ویژه‌ای برخوردار است. یکی از انواع اندازه‌های احتمال که به صورت طبیعی در سیستم‌های دینامیکی به آن بر می خوریم، اندازه‌های احتمال دیراکی هستند، یعنی اندازه‌هایی که بر روی تعدادی متناهی از نقاط فضا جمع شده‌اند. رفتار این نوع اندازه‌ها و دنباله‌هایی از آن‌ها و این که در فضای تمام اندازه‌های احتمال چگونه پخش می‌شوند، به موضوعاتی مهم و پایه‌ای در سیستم‌های دینامیکی و نظریه ارگودیک مربوط می‌شود. با این انگیزه، در این مقاله قصد داریم تا به مطالعه هندسه‌ی فضای اندازه‌ها بر روی یک فضای متریک فشرده بپردازیم و نشان می‌دهیم که چگونه این فضای بزرگ را می‌توان توسط اندازه‌های قابل فهمی همچون اندازه‌های دیراکی تقریب زد و حسی نسبت به آن یافت. در این مقاله صحبتی از ارتباط این مباحث با سیستم‌های دینامیکی نخواهیم کرد، و امید این است که در مقاله‌ای دیگر به توضیحی مفصل از این موضوع بپردازیم.

۲. نمادگذاری‌ها و تعاریف اولیه

فرض کنید X یک فضای متریک فشرده مجهز به فاصله d است. احتمالا اکثر افرادی که این مقاله را می خوانند با تعریف اندازه‌ی احتمال آشنا هستند، ولی بیایید دوباره این مفهوم را تعریف کنیم. برای این منظور ابتدا نیاز داریم تا ساختاری به نام سیگماـجبر را تعریف کنیم:

تعریف ۱.۲. یک **سیگماـجبر** بر روی X ، یک مجموعه‌ی ناتهی از زیرمجموعه‌های X مانند $\mathcal{B}$ است که تحت اعمال مکمل گیری، اشتراک شمارا و اجتماع شمارا بسته باشد.

تعداد سیگماـجبرهایی که می توان بر روی یک فضای متریک متصور بود بسیار زیاد است، اما ما از میان تمامی سیگماـجبرهای موجود، با سیگماـجبر بورل کار خواهیم کرد. ولی جا دارد ذکر کنیم که در مطالعه سیستم‌های دینامیکی، مطالعه اندازه‌های روی سیگماـجبرهای دیگری به جز بورل نیز به طور طبیعی ظاهر می‌شود که ما در این مقاله به آن‌ها نمی‌پردازیم.

تعریف ۲.۲. سیگماـجبر بورل، کوچکترین سیگماـجبری است که تمام زیر مجموعه‌های باز فضای متریک X را شامل می‌شود.

تعریف ۳.۲. یک اندازه‌ی احتمال بر روی X ، تابعی مانند μ است که به هر عضو سیگماـجبر بورل، عددی حقیقی و نامنفی نسبت می‌دهد به طوری که

- $\mu(\emptyset)=0$
- برای هر تعداد شمارا و دو به دو مجزا از اعضای سیگما-جبر مانند $A_1, A_2, \ldots$ داریم:

$$\mu(\cup_i A_i)=\Sigma_i \mu(A_i).$$

در اینجا باید ذکر کنیم که در تعریف اندازه‌ی احتمال می‌توانستیم هر سیگما-جبر دیگری را به جای سیگما-جبر بورل در نظر بگیریم، ولی در این نوشته ما خود را به اندازه‌های احتمال روی سیگما-جبر بورل محدود می‌کنیم. فضای تمام اندازه‌های احتمال را بر روی فضای X با $\mathcal{M}_1(X)$ نمایش خواهیم داد. گرچه ما خود را به اندازه‌های احتمالی که سیگما-جبر بورل دارند محدود کردیم، ولی این فضا به خودی خود فضای بسیار بزرگی محسوب می‌شود. در ادامه سعی می‌کنیم این فضا را از طریق مطالعه‌ی خانواده‌ای از اندازه‌های ساده که برای تعریف آن‌ها تنها به متناهی نقطه از فضای X نیاز است، بفهمیم. در واقع برای فهم بهتر فضای همه‌ی اندازه‌های احتمال، می‌خواهیم کاری شبیه به ساختن اعداد حقیقی با روش در نظر گرفتن دنباله‌های کوشی از اعداد گویا انجام دهیم. یعنی هر اندازه‌ی احتمال را به صورت دنباله‌ای کوشی از اندازه‌های روی متناهی نقطه در فضا ببینیم. فهم اعداد حقیقی به عنوان دنباله‌های کوشی از اعداد گویا چیزی نیست که معمولا افراد در ذهن خود دارند و خیلی راحت به خاطر قابل تصور بودن اعداد حقیقی، و داشتن شکلی از آن در ذهن با آن کنار می‌آیند. ولی فضای اندازه‌های احتمال فضای بزرگی است و معمولا افراد با خواصش آن را می‌شناسند، لذا نیاز به طی کردن چنین مسیری برای درک آن موجه می‌نماید.

برای فهم بهتر فضای همه‌ی اندازه‌های احتمال، می‌خواهیم کاری شبیه به ساختن اعداد حقیقی با روش در نظر گرفتن دنباله‌های کوشی از اعداد گویا انجام دهیم.

۳. اندازه‌های احتمال بر روی فضاهای متریک متناهی

بد نیست برای شروع فهم بهتر فضای اندازه‌های احتمال، از حالتی شروع کنیم که فضای متریک ما مجموعه‌ای متناهی است. فرض کنید مجموعه X ، یک مجموعه N عضوی باشد. می‌توانیم اعضای آن را با اعداد ۱ تا N نام گذاری کنیم. در حقیقت اگر مجموعه این اعداد را با $\mathbb{Z}_N$ نمایش دهیم، می‌توانیم فکر کنیم که فضای متریک X همان $\mathbb{Z}_N$ است که به متریکی مانند d مجهز شده است. یکی از موضوعات قابل تاملی که می‌توان در مورد آن فکر کرد، فضای همه‌ی متریک‌های ممکن بر روی مجموعه‌ی $\mathbb{Z}_N$ است. بیایید این فضا را با $\mathcal{D}_N$ نمایش دهیم. یک متریک با نسبت دادن اعداد مثبتی به زوج نقاط فضای $\mathbb{Z}_N$ مشخص می شود. در نگاه اول می‌گوییم، برای این که متریک را مشخص کنیم $\binom{N}{2}$ تا زوج داریم و به ازای هر زوج یک عدد حقیقی مثبت انتخاب می‌کنیم. ولی باید به این نکته توجه کنیم که انتخاب این اعداد مثبت باید طوری باشد که در نامساوی مثلث صدق کند. با این شرط فضای $\mathcal{D}_N$ دیگر همان فضای $\mathbb{R}_+^{\binom{N}{2}}$ نخواهد بود، بلکه زیر مجموعه‌ای از این فضا خواهد بود. می‌توان به خواص مختلف این فضا از جمله توپولوژی این فضا، هندسه‌ی این فضا و دیگر خواص آن فکر کرد، اما در این مقاله قصد پرداختن به این موضوعات را نداریم، و می‌خواهیم تنها فرض کنیم که متریک d عضوی از این فضا است و سعی کنیم فضای اندازه‌های احتمال بر روی $(\mathbb{Z}_N, d)$ را مطالعه کنیم.

برای ساده ترین حالت، یعنی $N=1$ ، بدیهی است که فضای اندازه‌های احتمال یک فضای تک عضوی است. برای حالت $N=2$ ، هر اندازه احتمال با مشخص کردن $\mu(1)$ مشخص می شود که باید عددی در بازه $[0,1]$ باشد. پس اندازه‌های احتمال بر روی $\mathbb{Z}_2$ که به متریکی مانند d مجهز شده را می‌توان با بازه $[0,1]$ معادل در نظر گرفت. در حالت کلی برای فضای متریک $(\mathbb{Z}_N, d)$ ، یک اندازه‌ی احتمال μ (که سیگما-جبر آن سیگما-جبر بورل است) با مقدار آن روی هر یک از اعضای $\mathbb{Z}_N$ مشخص می‌شود. لذا اگر مقدار $\mu(i)$ را با c_i نشان دهیم، برای مشخص کردن اندازه μ کافیست اعداد c_i را طوری انتخاب کنیم که بردار $(c_1, \ldots, c_N)$ در سادک $N-1$ بعدی Δ_N در $\mathbb{R}^N$ بیفتد که

$$\Delta_N := \{(c_1, \ldots, c_N) | \forall i \in \mathbb{Z}_N, 0 \le c_i \le 1 \quad \& \quad \Sigma_{i=1}^N c_i = 1\}.$$

مشاهده کنید که اگر μ و ν دو اندازه‌ی احتمال باشند، برای هر عدد $t \in [0,1]$ می‌توان از روی این دو اندازه، اندازه‌ی احتمالی جدید ساخت، با گرفتن ترکیب خطی‌ای از این دو اندازه:

$$t\mu + (1-t)\nu.$$

با توجه به این نکته، متوجه می‌شویم که فضای اندازه‌های احتمال بر روی فضای متریک $(\mathbb{Z}_n, d)$ یک فضای محدب است. از آنجایی که ترکیب محدب دو اندازه‌ی احتمال، معادل ترکیب محدب دو بردار در سادک Δ_N است، محدب بودن فضای اندازه‌های احتمال را با محدب بودن Δ_N نیز می‌توان مشاهده کرد. هر مجموعه‌ی محدب دارای نقاط گوشه‌ای است، یعنی نقاطی که آن‌ها را نمی‌توان به صورت ترکیب محدب دو اندازه‌ی نامساوی با خودش نوشت. می‌توان دید که نقاط گوشه‌ای فضای اندازه‌های احتمال روی $(\mathbb{Z}_N, d)$ اندازه‌های دیراک یک‌ـنقطه‌ای هستند:

تعریف ۱.۳. یک اندازه‌ی احتمال بر روی مجموعه‌ی X را اندازه‌ی دیراکی یک‌ـنقطه‌ای گویند، هرگاه نقطه‌ای مانند $x \in X$ موجود باشد که به مجموعه‌ی تک عضوی $\{x\}$ اندازه‌ی کامل نسبت دهد. این اندازه را با δ_x نمایش می‌دهیم. اندازه‌ی μ بر روی مجموعه‌ی X را اندازه‌ی دیراکی k ـنقطه‌ای گویند هرگاه ترکیبی محدب با ضرایب ناصفر از k اندازه دیراک یک‌ـنقطه‌ای بر روی k نقطه دو به دو متفاوت در فضا باشد.

این تعریف در حالتی که مشغول مطالعه‌ی اندازه‌های احتمال روی مجموعه‌های متناهی هستیم تعریفی بدیهی به نظر می‌رسد، زیرا تمام اندازه‌های احتمال دیراکی هستند. اما در قسمت‌های بعدی این مقاله نیز از این تعریف استفاده خواهیم کرد.

سوالی که پیش می‌آید این است که پس نقش متریک d در این میان چیست؟ به نظر می‌رسد که فضای تمام اندازه‌های احتمال بر روی $(\mathbb{Z}_N, d)$ مستقل از d باشد. جواب این است که در واقع هندسه‌ی فضای اندازه‌های احتمال است که با تغییر متریک d تغییر می‌کند. پس باید بگوییم که چه هندسه‌ای بر روی اندازه‌های احتمال در نظر می‌گیریم تا این پاسخ را توجیه کنیم.

۴. هندسه‌ی فضای اندازه‌های احتمال بر روی فضاهای متریک متناهی

منظورمان از هندسه‌ی یک فضا، متریکی است که بر روی آن در نظرش می‌گیریم. حال فضای $(\mathbb{Z}_N, d)$ را فیکس کنید. برای مشخص کردن هندسه‌ی $\mathcal{M}_1((\mathbb{Z}_N, d))$ باید بگوییم که فاصله بین دو اندازه‌ی احتمال مانند μ و ν را چه تعریف می‌کنیم. بیایید فرض کنیم این دو اندازه پخش شدن ۱ کیلوگرم خاک را بر روی هر یک از اعضای $\mathbb{Z}_N$ مشخص کرده باشند. یعنی روی i به ترتیب $\mu(i)$ و $\nu(i)$ کیلوگرم خاک ریخته شده. و ما می‌خواهیم به عنوان فاصله‌ی بین μ و ν کمترین انرژی لازم برای تبدیل یکی از آن‌ها به دیگری را در نظر بگیریم. برای این منظور باید ابتدا مفهوم طرح انتقال [1] را تعریف کنیم. یک طرح انتقال بین این دو اندازه، به ما می‌گوید که برای تبدیل اندازه‌ی μ به اندازه‌ی ν چه مقدار از خاک ریخته شده بر روی جایگاه i ام را به جایگاه j ام منتقل کنیم. واضح است که از طرح انتقالی که اندازه‌ی μ را به اندازه‌ی ν تبدیل می‌کند، می توان با برعکس نگاه کردن به ماجرا، به طرح انتقالی برای تبدیل اندازه‌ی ν به اندازه‌ی μ رسید. تعریف ریاضیاتی این مفهوم از این قرار است:

تعریف ۱.۴. یک طرح انتقال بین اندازه‌ی μ و ν یک ماتریس $n \times n$ با درایه‌های نامنفی p_{ij} است که

$$\Sigma_i p_{ij} = \mu(i), \quad \Sigma_j p_{ij} = \nu(j)$$

در این تعریف دیدن این نکته راحت است که اگر ماتریس $P_{n\times n}$ طرح انتقالی بین اندازه‌ی μ و اندازه‌ی ν باشد، آنگاه ترانهاده آن $P^T_{n\times n}$ طرح انتقالی بین اندازه‌ی ν و اندازه‌ی μ خواهد بود. حال باید بگوییم که به یک طرح انتقال چگونه یک انرژی، یا یک هزینه نسبت می‌دهیم. قبل از شروع به بیان این مطلب بهتر است این را بگوییم که راه‌های متنوعی می‌توان برای نسبت دادن انرژی به یک طرح انتقال متصور شد و ما تنها یکی از آن‌ها را که بیشتر با اهداف ما برای مطالعه سیستم‌های دینامیکی همخوانی دارد را توضیح خواهیم داد.

دید ما برای تخصیص انرژی به یک طرح انتقال، دیدی جزء نگرانه است. به این صورت که ابتدا انرژی لازم برای حمل p_{ij} کیلوگرم خاک از مکان i به مکان j را معرفی می‌کنیم، در نتیجه انرژی کل حاصل جمع زدن چنین انرژی‌هایی خواهد بود. انتظاری که ما از انرژی داریم، اولا این است که نسبت به p_{ij} خطی باشد، یعنی اگر r برابر این مقدار را از i به j منتقل کردیم، r برابر انرژی اولیه نیاز باشد. ثانیا انتظار داریم که این انرژی به فاصله‌ی $d(i,j)$ وابستگی مستقیم داشته باشد. یعنی با افزایش فاصله بین i و j این انرژی بیشتر شود. وابستگی‌های مستقیم متفاوتی را می‌توان در نظر گرفت، اما ما همان طور که اشاره شد، به

[1] transport plan

منظور اهداف دینامیکی، وابستگی خطی را انتخاب می‌کنیم. به طور خاص می‌توانیم انرژی لازم برای حمل p_{ij} کیلوگرم خاک از مکان i به مکان j را برابر با

$$d(i,j)p_{ij}$$

در نظر بگیریم. واضح است که این عدد در دو خاصیتی که در بالا به آن‌ها اشاره شد صدق می‌کند. حال می‌توان انرژی یک طرح انتقال را به صورت زیر به دست آورد:

$$\Sigma_{1\leq i,j\leq n} d(i,j)p_{i,j}.$$

حال که انرژی یک طرح انتقال را تعریف کردیم، می‌توانیم فاصله بین دو اندازه‌ی احتمال روی $(\mathbb{Z}_N, d)$ را تعریف کنیم.

تعریف ۲.۴. فضای تمام طرح‌های انتقال موجود بین اندازه‌ی μ و اندازه‌ی ν را با $\pi(\mu,\nu)$ نمایش می‌دهیم. طرح انتقال P را یک طرح انتقال بهینه می‌نامیم اگر کمترین انرژی را در بین اعضای $\pi(\mu,\nu)$ داشته باشد. فاصله‌ی بین دو اندازه‌ی احتمال μ و ν را برابر این مقدار مینیمم در نظر می‌گیریم و آن را با d_w نمایش می‌دهیم:

$$d_w(\mu,\nu) := \min_{P\in\pi(\mu,\nu)} \Sigma_{1\leq i,j\leq n} d(i,j)p_{i,j}$$

چک کردن این که d_w در خواص فاصله صدق می‌کند را به خواننده می‌سپاریم.

حال می‌توانیم ببینیم که بعد از فیکس کردن فضای متریک $(\mathbb{Z}_N, d)$ ، فضای $\mathcal{M}_1((\mathbb{Z}_N, d))$ با متر d_w نیز تبدیل به یک فضای متریک می‌شود. اگر دوست داشته باشید، می‌توانید این فضا را همان سادک Δ_N ببینید که به متریکی جدید، و نه آن متریکی که از فضای $\mathbb{R}^N$ به ارث می‌برد مجهز شده است. آیا می‌توانید تصور کنید با تغییر متریک d هندسه‌ی $\mathcal{M}_1((\mathbb{Z}_N, d))$ چه تغییری می‌کند؟ به طور مثال اگر $d(i,j)$ را به سمت صفر میل دهیم، فاصله‌ی بین دو اندازه‌ی δ_i و δ_j نیز به سمت صفر میل می‌کند. همچنین می‌توان دید حداکثر فاصله‌ای که دو اندازه‌ی احتمال روی $(\mathbb{Z}_N, d)$ می‌توانند کسب کنند، دقیقا برابر قطر $(\mathbb{Z}_N, d)$ است.

۵. هندسه‌ی فضای اندازه‌های احتمال بر روی یک فضای متریک فشرده

حال فرض کنید (X,d) یک فضای متریک نه لزوما متناهی، ولی فشرده باشد. می‌خواهیم هندسه‌ی فضای $\mathcal{M}_1((X,d))$ را بفهمیم. بگذارید همانند بخش قبل، از اندازه‌های دیراکی شروع کنیم. تمام اندازه‌های دیراکی حداکثر k- نقطه‌ای را با $\mathcal{M}_1^k((X,d))$ نمایش می‌دهیم:

$$\mathcal{M}_1^k((X,d)) := \{\Sigma_{i=1}^k c_i\delta_{x_i} | (c_1,...,c_k)\in\Delta_k, x_i\in X\}.$$

دقت کنید چون به تعداد دلخواه می‌توانیم c_i صفر داشته باشیم، تمام اندازه‌های دیراک s- نقطه‌ای برای $k\geq s$ در مجموعه‌ی $\mathcal{M}_1^k((X,d))$ حضور دارند.

$$\mathcal{M}_1^1((X,d)) \subset \mathcal{M}_1^2((X,d)) \subset \mathcal{M}_1^3((X,d)) \subset ...$$

مجموعه‌ی تمام اندازه‌های دیراکی روی (X,d) را با $\mathcal{M}_1^*((X,d))$ نمایش می‌دهیم:

$$\mathcal{M}_1^*((X,d)) := \cup_{k=1}^{\infty}\mathcal{M}_1^k((X,d)).$$

گزاره ۱.۵. $\mathcal{M}_1^*((X,d))$ یک مجموعه‌ی محدب است که نقاط گوشه‌ای آن اعضای $\mathcal{M}_1^1((X,d))$ یا همان اندازه‌های دیراک یک-نقطه‌ای هستند.

اثبات. برای اثبات محدب بودن کافیست دقت کنید که اگر $\mu\in\mathcal{M}_1^k((X,d))$ و $\nu\in\mathcal{M}_1^s((X,d))$، آنگاه برای هر $t\in[0,1]$ اندازه‌ی $t\mu+(1-t)\nu$ در $\mathcal{M}_1^{k+s}((X,d))$ زندگی می‌کند. از طرفی برای اثبات قسمت دوم حکم، واضح است که هر عضو $\mathcal{M}_1^*((X,d))$ را به شکل یک ترکیب خطی متناهی از اعضای $\mathcal{M}_1^1((X,d))$ می‌توان نوشت و اعضای $\mathcal{M}_1^1((X,d))$ را نمی‌توان به صورت ترکیب خطی اندازه‌های دیگری نوشت. □

همان طور که فاصله‌ی بین دو اندازه‌ی احتمال بر روی یک فضای متریک متناهی را در بخش قبل تعریف کردیم، می‌توانیم بین دو اندازه‌ی دیراکی دلخواه بر روی (X,d) نیز یک فاصله تعریف کنیم و $\mathcal{M}^*_1((X,d))$ را مجهز به یک متریک کنیم. برای این منظور کافی است مفهوم طرح انتقال را برای این حالت کلی تر تعمیم دهیم. فرض کنید $\mu=\Sigma_{i=1}^k c_i\delta_{x_i}$ یک اندازه‌ی دیراک k- نقطه‌ای و $\nu=\Sigma_{j=1}^s b_j\delta_{y_j}$ یک اندازه‌ی دیراک s- نقطه‌ای باشد. یادآوری می‌کنیم که با توجه به تعریف، تمام x_i ها و همچنین تمام y_j ها دو به دو مجزا هستند. اگر دوباره به این دو اندازه‌ی احتمال به چشم دو نحوه‌ی پخش شدن یک کیلوگرم خاک به ترتیب بر روی k نقطه از فضا و s نقطه از فضا نگاه کنیم، یک طرح انتقال بین آن‌ها به ما می‌گوید چه مقدار از خاک موجود در نقطه x_i را به نقطه y_j منتقل کنیم. تعریف ریاضیاتی این مفهوم در این حالت از این قرار است:

تعریف ۲.۵. یک طرح انتقال بین دو اندازه‌ی دیراک k- نقطه‌ای μ و s- نقطه‌ای ν یک ماتریس $P_{k\times s}$ است که

$$\Sigma_j p_{ij}=\mu(x_i),\quad \Sigma_i p_{ij}=\nu(y_j).$$

می‌توان به یک طرح انتقال، نه به عنوان یک ماتریس، بلکه به عنوان یک اندازه‌ی دیراکی روی فضای $X\times X$ نگاه کرد:

گزاره ۳.۵. طرح‌های انتقال بین دو اندازه‌ی دیراک k- نقطه‌ای μ بر روی نقاط $x_1,\dots,x_k$ و s- نقطه‌ای ν بر روی نقاط $y_1,\dots,y_s$ در تناظر یک‌به‌یک با اندازه‌های دیراکی حداکثر $k\times s$- نقطه‌ای روی فضای $X\times X$ با فرم $\Sigma_{i,j}c_{ij}\delta_{(x_i,y_j)}$ هستند با این شرط که

$$\Sigma_j c_{ij}=\mu(x_i),\quad \Sigma_i c_{ij}=\nu(y_j).$$

اثبات. کافی است درایه p_{ij} از یک طرح انتقال را همان ضریب c_{ij} در اندازه‌ی $\Sigma_{i,j}c_{ij}\delta_{(x_i,y_j)}$ در نظر بگیرید، و بالعکس. □

حال همانند قبل می‌توانیم به یک طرح انتقال بین μ و ν انرژی

$$\Sigma_{i,j}d(x_i,y_j)p_{ij}$$

را نسبت دهیم. برای مشخص کردن فاصله‌ی بین این دو اندازه، کمترین انرژی را در بین تمام طرح‌های انتقال بین آن‌ها در نظر می‌گیریم:

تعریف ۴.۵. فضای تمام طرح‌های انتقال موجود بین دو اندازه‌ی دیراک k- نقطه‌ای μ بر روی نقاط $x_1,\dots,x_k$ و s- نقطه‌ای ν بر روی نقاط $y_1,\dots,y_s$ را با $\pi(\mu,\nu)$ نمایش می‌دهیم. طرح انتقال P را یک طرح انتقال بهینه نامیم اگر کمترین انرژی را در بین اعضای $\pi(\mu,\nu)$ داشته باشد. فاصله‌ی بین دو اندازه‌ی احتمال μ و ν را برابر این مقدار مینیمم در نظر می‌گیریم و آن را با d_w نمایش می‌دهیم:

$$d_w(\mu,\nu):=\min_{P\in\pi(\mu,\nu)}\Sigma_{1\leq i\leq k,1\leq j\leq s}d(x_i,y_j)p_{ij}.$$

اثبات این که d_w در این حالت نیز در خواص متریک صدق می‌کند به خواننده واگذار می‌شود.

حال بیایید فضای $\mathcal{M}^*_1((X,d))$ مجهز به متریک d_w را بهتر بفهمیم، همانطور که در بالا دیدیم این فضا از اجتماع مجموعه‌های تو در توی $\mathcal{M}^k_1((X,d))$ تشکیل شده است. حال به بررسی خواص توپولوژیک و همچنین خواصی از این فضا می‌پردازیم که متریک d_w در آن‌ها دخیل است. ابتدا به بررسی $\mathcal{M}^1_1((X,d))$ می‌پردازیم:

گزاره ۵.۵. (آ) نگاشت $p_1:X\to\mathcal{M}^1_1((X,d))$ که هر نقطه x را به اندازه‌ی احتمال δ_x می‌فرستد یک ایزومتری بین (X,d) و $(\mathcal{M}^1_1((X,d)),d_w)$ است.

(ب) $diam(X,d)=diam(\mathcal{M}^1_1((X,d)),d_w)=diam(\mathcal{M}^*_1((X,d)),d_w)$

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. □

با توجه به گزاره‌ی بالا متوجه می‌شویم که $(\mathcal{M}_1^1((X,d)),d_w)$ یک کپی ایزومتریک از X است و تمام خواص آن با X یکسان خواهد بود. وقتی $k>1$ است، برای فهمیدن بهتر $(\mathcal{M}_1^k((X,d)),d_w)$ بیایید به پاسخ به این سوال‌ها بپردازیم: آیا می‌توانیم توپولوژی (سرتاسری) این فضا را برحسب توپولوژی X توصیف کنیم؟ اعضایی از $\mathcal{M}_1^k((X,d))$ که در $\mathcal{M}_1^{k-1}((X,d))$ نیستند کدامند؟

حاصل ضرب دکارتی X در خودش به تعداد k مرتبه را با X^k نشان می‌دهیم. یادآوری می‌کنیم که سادک k بعدی Δ_k را به صورت زیر تعریف می‌کردیم:

$$\Delta_k := \{(c_1,...,c_k)|\forall i\in\mathbb{Z}_k, 0\le c_i\le 1 \quad \& \quad \Sigma_{i=1}^k c_i=1\}.$$

به هر $2k$- تایی مرتب $(x_1,...,x_k,c_1,...,c_k)\in X^k\times\Delta_k$ می‌توانیم یک اندازه‌ی احتمال به صورت زیر نسبت دهیم:

$$(x_1,...,x_k,c_1,...,c_k)\mapsto \Sigma_{i=1}^k c_i\delta_{x_i}.$$

این به ما نگاشتی از $X^k\times\Delta_k$ به $\mathcal{M}_1^k((X,d))$ می‌دهد که آن را p_k می‌نامیم.

$$\begin{array}{c} X^k\times\Delta_k \\ {\scriptstyle p_k}\big\downarrow \\ \mathcal{M}_1^k((X,d)) \end{array}$$

اثبات این نکته که p_k پیوسته است، کار راحتی است.

اگر مانند حالت $k=1$ نگاشت p_k تناظری یک‌به‌یک برقرار می‌کرد، کار تمام بود. ولی نگاشت p_k به سه دلیل نگاشتی یک‌به‌یک نیست. اولا این که اعضای X^k دارای ترتیب هستند و نگاشت p_k به این ترتیب حساس نیست. دوما برای اندازه‌هایی که محمل کمتر از k نقطه دارند، چون در هر نمایش آن‌ها به صورت $(x_1,...,x_k,c_1,...,c_k)$ حداقل دو تا x_i و x_j مساوی وجود دارد، می‌توانیم به جای c_i و c_j مقادیر c_i-t و c_j+t را جایگذاری کنیم و دوباره همان اندازه را بگیریم. سوما حتی اگر تمام x_i ها دو به دو متفاوت باشند ولی یکی از c_i ها برابر صفر باشد به این معنی است که اندازه‌ی شما نقطه‌ی x_i مربوطه را در محمل خود نمی‌بیند و نحوه‌های نمایش دیگری برای این اندازه وجود دارد. به طور مثال می‌توانید x_i را هر نقطه‌ی دلخواهی بگذارید. برای کم کردن میزان غیر یک‌به‌یک بودن نگاشت p ، از فضای خارج قسمتی

$$Q_k(X) := (X^k\times\Delta_k)/\sim$$

کمک می‌گیریم که در آن رابطه‌ی هم‌ارزی $\sim$ به صورت زیر تعریف شده است:

$$(x_1,...,x_k,c_1,...,c_k)\sim(x_{\sigma(1)},...,x_{\sigma(k)},c_{\sigma(1)},...,c_{\sigma(k)}) \quad \forall\sigma\in S_k.$$

در بالا S_k گروه جایگشت‌های روی $\mathbb{Z}_k$ را نشان می‌دهد. به هر کلاس هم‌ارزی $[(x_1,...,x_k,c_1,...,c_k)]\in Q_k(X)$ همانند بالا می‌توانیم یک اندازه‌ی احتمال به صورت زیر نسبت دهیم:

$$[(x_1,...,x_k,c_1,...,c_k)]\mapsto \Sigma_{i=1}^k c_i\delta_{x_i}.$$

این به ما نگاشتی از $Q_k(X)$ به $\mathcal{M}_1^k((X,d))$ می‌دهد که آن را q_k می‌نامیم. می‌توان چک کرد که q_k خوش تعریف است.

$$\begin{array}{c} Q_k(X) \\ {\scriptstyle q_k}\big\downarrow \\ \mathcal{M}_1^k((X,d)) \end{array}$$

دیدن این نکته سخت نیست که نگاشت q_k نیز پیوسته است.

اگر نگاشتی که هر عضو $(x_1,...,x_k,c_1,...,c_k)$ را به کلاس هم‌ارزی آن در $Q_k(X)$ می‌برد را با ι_k نمایش دهیم، از تعاریف نتیجه می‌شود که دیاگرام زیر جا به جایی است:

$$\begin{tikzcd} X^k \times \Delta_k \arrow[d, "\iota_k"'] \arrow[dr, "p_k"] & \\ Q_k(X) \arrow[r, "q_k"'] & \mathcal{M}_1^k((X,d)) \end{tikzcd}$$

همچنان نگاشت q_k به دلایل دوم و سوم ذکر شده در بالا، غیر یک‌به‌یک است. اما می‌توانیم زیر مجموعه‌ای خوب(!) از $Q_k(X)$ را بیابیم که q_k روی آن یک‌به‌یک باشد.

مجموعه‌ی تمام kـتایی های مرتب در X^k که دارای حداقل دو مولفه‌ی برابر هستند را با $D(X^k)$ نمایش می‌دهیم. همچنین تعریف کنید:

$$\Delta_k \supset \Delta_k^\circ := \{(c_1,...,c_k) | \forall i \in \mathbb{Z}_k, 0 < c_i \leq 1 \quad \& \quad \Sigma_{i=1}^k c_i = 1\}.$$

می‌توان به Δ_k° به چشم نقاط درونی Δ_k نگاه کرد. حال فضای خارج قسمتی زیر را در نظر بگیرید:

$$Q_k(X) \supset Q_k^\circ(X) := ((X^k \setminus D(X^k)) \times \Delta_k^\circ)/\sim .$$

می‌توان دید که $Q_k^\circ(X)$ زیرمجموعه‌ای باز و چگال از $Q_k(X)$ است. این همان مجموعه‌ی خوبی است که ما به دنبال آن می‌گشتیم، زیرا نه تنها q_k روی آن یک‌به‌یک است، بلکه یک همئومورفیسم به روی تصویرش نیز هست:

گزاره ۶.۵. نگاشت q_k مجموعه‌ی $Q_k^\circ(X)$ را به صورت همئومورفیسم به مجموعه‌ی اندازه های دیراکی دقیقا k تایی، یعنی $\mathcal{M}_1^k((X,d)) \setminus \mathcal{M}_1^{k-1}((X,d))$ می‌برد و مجموعه‌ی $Q^k(X) \setminus Q_k^\circ(X)$ را به صورت پوشا به مجموعه‌ی اندازه‌های دیراکی حداکثر $k-1$ نقطه‌ای ، یعنی $\mathcal{M}_1^k((X,d)) \supset \mathcal{M}_1^{k-1}((X,d))$ می‌برد.

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. □

نتیجه ۷.۵. فضای $\mathcal{M}_1^k((X,d))$ با متریک d_w فشرده است.

اثبات. $Q_k(X)$ فضایی فشرده است و q_k از این فضا به $\mathcal{M}_1^k((X,d))$ پوشا و پیوسته است، لذا حکم ثابت می‌شود. □

گزاره ۶.۵ به ما کمک می‌کند تا $\mathcal{M}_1^k((X,d))$ را به صورت استقرایی بهتر بفهمیم. در حقیقت به ما می‌گوید زیرمجموعه‌ی اعضایی از $\mathcal{M}_1^k((X,d))$ که در $\mathcal{M}_1^{k-1}((X,d))$ نیستند با $Q_k^\circ(X)$ همئومورف است. پس با افزایش k به اندازه‌ی یک واحد، می‌فهمیم که مجموعه‌ی نقاطی که به $\mathcal{M}_1^k((X,d))$ اضافه می‌شود از لحاظ توپولوژیک چه شکلی دارد. البته این برای فهمیدن شکل $\mathcal{M}_1^k((X,d))$ کافی نیست. زیرا ما باید بگوییم که تکه‌هایی که در هر مرحله به شکل قبلی اضافه می‌شوند، چطور به آن می‌چسبند. برای این منظور به این نکته دقت کنید که با توجه به گزاره‌ی بالا نگاشت q_k روی $Q^k(X) \setminus Q_k^\circ(X)$ را می‌توان نگاشتی به $\mathcal{M}_1^{k-1}((X,d))$ نیز در نظر گرفت. حال اگر فضای $Q_k(X)$ را از روی زیرمجموعه‌ی $Q^k(X) \setminus Q_k^\circ(X)$ با نگاشت q_k به فضای $\mathcal{M}_1^{k-1}((X,d))$ بچسبانیم، به عنوان چسباندن دو فضای توپولوژیک، فضای توپولوژیک حاصل با $\mathcal{M}_1^k((X,d))$ همئومورف خواهد شد. به این گونه، به صورت استقرایی با شروع از $\mathcal{M}_1^1((X,d))$ که با X هومئومورف است، و با چسباندن صحیح $Q_2(X)$ به آن می‌توانیم شکل $\mathcal{M}_1^2((X,d))$ را به دست آوریم و الی آخر.

توجه کنید که یکی از نتایج دیگری که از گزاره‌ی بالا می‌توان گرفت در مورد بعد $\mathcal{M}_1^k((X,d))$ است. به طور مثال اگر فضای X یک منیفلد m بعدی باشد، می‌توان دید که $Q^k(X) \setminus Q_k^\circ(X)$ یک منیفلد $km+k-1$ بعدی خواهد بود. این نکته نشان می‌دهد که در این حالت $\mathcal{M}_1^*((X,d))$ یک فضای متناهی البعد نخواهد بود. در حالت کلی می‌توان چک کرد که وقتی تعداد اعضای X نامتناهی باشد $\mathcal{M}_1^*((X,d))$ نیز یک فضای نامتناهی البعد خواهد شد.

حال به معرفی زیرمجموعه‌هایی از $\mathcal{M}_1^*((X,d))$ می‌پردازیم که شامل اعضایی هستند که در مطالعه‌ی رفتار آماری سیستم‌های دینامیکی به صورت طبیعی ظاهر می‌شوند.

فضای حاصل ضرب متقارن [1] kـام یک فضای توپولوژیک X ، به صورت خارج قسمت X^k تحت عمل گروه جایگشت‌های k تایی که مولفه‌های اعضای X^k را جابه جا می‌کند، تعریف می‌شود و با نماد $Sym_k(X)$ نمایش داده می‌شود. می‌توان دید که یک کپی همئومورف با $Sym_k(X)$ در $Q_k(X)$ زندگی می‌کند. اگر مقطعی از (کلاف تاری) $X^k \times \Delta_k$ را در نظر بگیرید که در آن تمام ضرایب c_i برابر $\frac{1}{k}$ هستند:

$$E_k := \{(x_1, ..., x_k, c_1, ..., c_k) | \forall i \in \mathbb{Z}_k \quad c_i = \frac{1}{k}\}.$$

می‌توان دید که $\iota_k(E_k)$ با $Sym_k(X)$ همئومورف است. حال بیایید تمام اندازه‌هایی که تحت نگاشت q_k از اعضای $\iota_k(E_k)$ به دست می‌آیند را در نظر بگیریم:

$$\mathcal{E}_1^k((X,d)) := q_k \circ \iota_k(E_k).$$

گزاره ۸.۵. نگاشت q_k یک همئومورفیسم از $\iota_k(E_k)$ به $\mathcal{E}_1^k((X,d))$ است.

اثبات. برای اثبات حکم بالا سخت‌ترین قسمت کار چک کردن یک‌به‌یک بودن است. برای این منظور دقت کنید که اگر

$$\Sigma_{i=1}^k \frac{1}{k}\delta_{x_i} = \Sigma_{i=1}^k \frac{1}{k}\delta_{y_i},$$

آنگاه از مساوی بودن محمل این دو اندازه، مساوی بودن دو مجموعه‌ی $\{x_1, ..., x_k\}$ و $\{y_1, ..., y_k\}$ را نتیجه می‌گیریم و از مساوی بودن ضرایب نیز نتیجه می‌گیریم که هر عضو از این مجموعه به تعدادی یکسان در بین x_i ها و y_i ها ظاهر می‌شود. لذا می‌توان با یک جایگشت، x_i ها را به y_i تبدیل کرد. پس دو کلاس هم‌ارزی زیر یکسان هستند:

$$[(x_1, ..., x_k, \frac{1}{k}, ..., \frac{1}{k})] = [(y_1, ..., y_k, \frac{1}{k}, ..., \frac{1}{k})].$$

و این یک‌به‌یک بودن را به ما می‌دهد. □

از آنجایی که $\iota_k(E_k)$ و $Sym_k(X)$ همئومورف هستند، گزاره‌ی بالا به ما می‌گوید که $\mathcal{E}_1^k((X,d))$ یک کپی همئومورف از $Sym_k(X)$ است. پس فهمیدن توپولوژی آن راحت‌تر از فهمیدن توپولوژی $\mathcal{M}_1^k((X,d))$ است.

بر خلاف مجموعه‌های $\mathcal{M}_1^k((X,d))$ مجموعه‌های $\mathcal{E}_1^k((X,d))$ تودرتو نیستند، و این که چه اشتراکی با هم دارند کمی پیچیده‌تر از قبل است:

گزاره ۹.۵. برای هر دو عدد طبیعی $k, s \in \mathbb{N}$ اگر بزرگترین مقسوم علیه مشترک آن دو را با (k,s) نمایش دهیم، آنگاه:

$$\mathcal{E}_1^k((X,d)) \cap \mathcal{E}_1^s((X,d)) = \mathcal{E}_1^{(k,s)}((X,d)).$$

اثبات. ابتدا ثابت کنید که اگر $l|k$ آنگاه $\mathcal{E}_1^l((X,d)) \subset \mathcal{E}_1^k((X,d))$. این یک طرف تساوی را نتیجه می‌دهد. برای اثبات طرف دیگر به این نکته دقت کنید که اگر عضوی از $\mathcal{E}_1^k((X,d))$ باشد که دارای محملی l عضوی باشد، برای این که تمام c_i ها برابر $\frac{1}{k}$ باشند راهی به جز این که $l|k$ باشد باقی نمی‌گذارد. □

ذکر این نکته مفید است که حرف‌هایی که تا به حال در مورد توپولوژی $\mathcal{M}_1^k((X,d))$ زده شد، تنها به توپولوژی فضای X وابسته است، و با تغییر متریک d، طوری که توپولوژی وابسته به آن تغییری نکند، فضاهایی که در بالا معرفی کردیم در حد همئومورفیسم یکسان باقی می‌مانند. بحث در مورد توپولوژی این فضاها به خودی خود مبحث جالبی است و حرف‌های بیشتری می‌توان در این مورد زد. خواننده‌ی علاقه‌مند را تشویق می‌کنیم که به این موضوع بیشتر فکر کند. به خصوص وقتی فضای X را مشخص کنید، این مسئله می‌تواند جالب باشد. یکی از ساده‌ترین حالات این مساله وقتی است که فضای X را دایره یا پاره‌خط بگیرید. حتی در این حالت فهمیدن توپولوژی $\mathcal{M}_1^k((X,d))$ کار ساده‌ای نیست. ما در اینجا به همین مقدار توضیح در مورد توپولوژی $\mathcal{M}_1^k((X,d))$ بسنده می‌کنیم و در ادامه به خواص هندسی این فضا که به متریک d_w مجهز شده می‌پردازیم.

[1] symmetric product

دقت کنید که در تعریف متریک d_w متریک روی فضای X دخیل است و خواهیم دید که خواص هندسی $\mathcal{M}_1^k((X,d))$ چگونه به آن وابسته است.

گزاره ۱۰.۵. برای هر $k \in \mathbb{N}$ مجموعه‌ی $\mathcal{M}_1^k((X,d))$ با متریک d_w فضای متریکی کامل است.

از آنجایی که با توجه به نتیجه‌ی ۷.۵ فضای $\mathcal{M}_1^k((X,d))$ فشرده است، و هر فضای متریک فشرده کامل است، حکم بالا به راحتی نتیجه می‌شود. ولی ما در اینجا به عنوان نقطه‌ای شروع برای فهم هندسه‌ی این فضا، اثباتی مستقیم برای آن می‌آوریم.

اثبات. فرض کنید $\mu_1, \mu_2, \ldots$ دنباله‌ای کوشی از اندازه‌های احتمال متعلق به $\mathcal{M}_1^k((X,d))$ باشد. محمل هر یک از این اندازه‌ها یک زیرمجموعه‌ی حداکثر k نقطه‌ای از X است. می‌توان زیردنباله‌ای مانند $\mu_{i_1}, \mu_{i_2}, \ldots$ انتخاب کرد به طوری که محمل μ_{i_n} ها در توپولوژی هاسدورف [1] به مجموعه‌ای متناهی شامل s عضو مانند $\{x_1, x_2, \ldots, x_s\}$ میل کند. می‌توان چک کرد که حتما $s \leq k$. حال ϵ را آنقدر کوچک بگیرید که گویی شعاع ϵ حول هر یک از x_i ها شامل هیچ x_j دیگری نباشد. اثبات این که حد زیر موجود است کار راحتی است:

$$\lim_n \mu_{i_n}(B_\epsilon(x_j)).$$

اگر این حد را c_j بنامیم می‌توان چک کرد که برای اندازه‌ی

$$\mu_\infty := \Sigma_{j=1}^s c_j \delta_{x_j}$$

فاصله‌ی d_w اندازه‌های μ_{i_n} از اندازه‌ی μ_∞ به صفر میل می‌کند. از آنجایی که $\{\mu_i\}_i \in \mathbb{N}$ یک دنباله کوشی بود، میل کردن یک زیردنباله‌ی آن به μ_∞ به معنی میل کردن کل دنباله به این اندازه است. برای اتمام اثبات کافیست توجه کنید که چون $s \leq k$ پس

$$\mu_\infty \in \mathcal{M}_1^k((X,d)).$$

□

حال بیایید به این فکر کنیم که اعضایی از $\mathcal{M}_1^k((X,d))$ که در $\mathcal{M}_1^{k-1}((X,d))$ نیستند حداکثر چقدر از این فضا فاصله می‌گیرند. فرض کنید $\mu = \Sigma_{i=1}^k c_i \delta_{x_i}$ چنین عضوی باشد. ما به دنبال یافتن کران بالایی برای

$$d_w(\mu, \mathcal{M}_1^{k-1}((X,d)))$$

هستیم. با توجه به مفروضاتمان، x_i ها دو به دو متفاوت هستند و تمام c_i ها ناصفرند. پس حداقل عددی مانند $1 \leq s \leq k$ داریم که $c_s \leq \frac{1}{k}$. حال اندازه‌ی

$$\mu' := \mu - c_s \delta_{x_s}$$

را در نظر بگیرید. محمل این اندازه شامل $k-1$ نقطه است، ولی این یک اندازه‌ی احتمال نیست زیرا اندازه‌ی کل فضا برابر $1 - c_s$ است. برای اندازه‌ی احتمال شدن، باید آن را نرمال کنیم:

$$\nu = \frac{1}{1 - c_s}\mu.$$

برای این که تخمینی از $d_w(\mu, \nu)$ به دست آوریم یک طرح انتقال بین این دو اندازه پیشنهاد می‌دهیم. این دو اندازه به مقدار خوبی باهم اشتراک دارند. برای هر $i \neq s$ مقدار خاک موجود بر روی x_i را دست نمی‌زنیم. مقدار خاک موجود بر روی x_s را نیز به نسبت وزن‌های دیگر x_i ها بین آن‌ها تقسیم می‌کنیم. یعنی برای $i \neq s$ از x_s به x_i به مقدار زیر خاک می‌دهیم:

$$\frac{c_s c_i}{1 - c_s}.$$

[1] توپولوژی هاسدورف توپولوژی‌ای است که بر روی فضای تمام زیرمجموعه‌های فشرده‌ی یک فضای متریک فشرده گذاشته می‌شود. با این توپولوژی، فضای تمام زیرمجموعه‌های فشرده‌ی X فضایی فشرده خواهد بود و لذا هر دنباله، زیردنباله‌ای همگرا دارد.

می‌توانید چک کنید که این طرح انتقال اندازه‌ی μ را به اندازه‌ی ν تبدیل می‌کند، و انرژی آن برابر

$$\Sigma i \neq s \frac{c_s c_i}{۱ - c_s} d(x_s, x_i)$$

خواهد بود. از آنجایی که $d(x_s, x_i) \leq diam(X, d)$ و این که ممکن است طرح انتقال‌های دیگری با انرژی کمتری بین این دو اندازه وجود داشته باشد، نتیجه می‌گیریم

$$d_w(\mu, \nu) \leq \frac{diam(X, d)}{k}. \qquad (۱.۵)$$

می‌بینیم که $\mathcal{M}_۱^k((X, d))$ خیلی چاق‌تر از $\mathcal{M}_۱^{k-۱}((X, d))$ نیست، مخصوصا اگر k عددی بزرگ باشد. سوالی که پیش می‌آید این است که برای $l \in \mathbb{N}$ که خیلی از k بزرگتر است، آیا $\mathcal{M}_۱^l((X, d))$ می‌تواند خیلی از $\mathcal{M}_۱^k((X, d))$ چاق‌تر باشد؟ گزاره‌ی زیر جوابی منفی به این سوال است:

گزاره ۱۱.۵. برای هر $k \in \mathbb{N}$ تعریف کنید:

$$\epsilon_k := \sup_{\mu \in \mathcal{M}_۱^*((X,d)))} d_w(\mu, \mathcal{M}_۱^k((X, d)).$$

آنگاه دنباله‌ی ϵ_k نزولی و همگرا به صفر است.

با توجه به این گزاره، اگر بخواهیم فضای $\mathcal{M}_۱^*((x, d))$ را به چیزی تشبیه کنیم، پیاز می‌تواند گزینه‌ی خوبی باشد. فضایی لایه‌لایه و تودرتو، که ضخامت لایه‌های بیرونی آن، کم و کمتر می‌شود. البته یک فرق اساسی این است که بعد لایه‌های فضای اندازه‌های دیراکی ثابت نیست، و افزایش می‌یابد. اثبات این گزاره در راستای هدفی که در این مقاله به دنبال آن هستیم، یعنی فهم بهتر از فضای اندازه‌های احتمال مهم است.

اثبات. برای فضای متریک (X, d)، عدد $\epsilon_k(X, d)$ را برابر مینیمم مقداری در نظر بگیرید که به ازای آن یک مجموعه‌ی k عضوی $\epsilon_k(X, d)$-چگال در X وجود داشته‌باشد. واضح است که شرط فشرده بودن (X, d) نتیجه می‌دهد که $\lim_{k\to\infty} \epsilon_k(X, d) =$ ۰. حال عدد k را فیکس کنید و مجموعه‌ی $\{x_۱, ..., x_k\}$ را طوری در نظر بگیرید که یک مجموعه‌ی $\epsilon_k(X, d)$-چگال باشد. اندازه‌ی دلخواه $\mu \in \mathcal{M}_۱^k((X, d))$ را در نظر بگیرید. از روی این اندازه، اندازه‌ای با محمل $\{x_۱, ..., x_k\}$ می‌سازیم که فاصله‌ی آن‌ها با هم کمتر از $\epsilon_k(X, d)$ باشد. بستار گوی به شعاع r حول نقطه‌ی x را با $\overline{B_r(x)}$ نمایش می‌دهیم. تمام جرم توزیع شده توسط μ در مجموعه‌ی $\overline{B_{\epsilon_k(X,d)}(x_۱)}$ را به نقطه‌ی $x_۱$ منتقل می‌کنیم. سپس تمام جرم توزیع شده توسط μ در مجموعه‌ی $\overline{B_{\epsilon_k(X,d)}(x_۲)} \setminus \overline{B_{\epsilon_k(X,d)}(x_۱)}$ را به نقطه‌ی $x_۲$ منتقل می‌کنیم. به همین شکل به صورت استقرایی جلو رفته و در مرحله‌ی s، جرم پخش شده توسط μ را در $\overline{B_{\epsilon_k(X,d)}(x_s)} \setminus \cup_{i=۱}^{s-۱} \overline{B_{\epsilon_k(X,d)}(x_i)}$ به نقطه‌ی x_s منتقل می‌کنیم. پس از k مرحله به این اندازه می‌رسیم:

$$\nu := \Sigma_{s=۱}^k c_s \delta_{x_s},$$

که در آن

$$c_s = \Sigma_{s=۱}^k \mu(\overline{B_{\epsilon_k(X,d)}(x_s)} \setminus \cup_{i=۱}^{s-۱} \overline{B_{\epsilon_k(X,d)}(x_i)}).$$

با توجه به نحوه‌ی به دست آوردن ν، که در حقیقت خود یک طرح انتقال بین μ و ν است، نتیجه می‌گیریم که

$$d_w(\mu, \nu) \leq \epsilon_k(X, d).$$

پس برای اثبات حکم کافی است قرار دهیم $\epsilon_k = \epsilon_k(X, d)$. □

اگر X نامتناهی عضو داشته باشد، پیدا کردن دنباله‌ای کوشی از اندازه‌های دیراکی که به هیچ یک از اعضای $\mathcal{M}_۱^*((X, d))$ میل نکند کار سختی نیست. به طور مثال فرض کنید $x_۱, x_۲, ...$ اعضای دوبه‌دو مجزا از X باشند. اندازه‌های احتمال زیر، دنباله‌ای کوشی می‌دهند که به هیچ یک از اعضای $\mathcal{M}_۱^*((X, d))$ همگرا نیست:

$$\nu_n := \frac{۲^n}{۲^n - ۱} \Sigma_{j=۱}^n \frac{۱}{۲^j} \delta_{x_j}.$$

برای کوشی بودن این دنباله می‌توان از روش به دست آوردن نامساوی ۱.۵ استفاده کرد و بدست آورد که

$$d_w(\nu_{n-۱},\nu_n) \leq \frac{diam(X,d)}{۲^n}.$$

برای X نامتناهی، فضای متریک $(\mathcal{M}_۱^*((X,d)),d_w)$ کامل نیست.

۶. کامل‌سازی فضای $(\mathcal{M}_۱^*((X,d)),d_w)$

همان طور که قبلا اشاره کردیم، هدف نهایی ما فهم فضای تمام اندازه‌های احتمال، توسط اندازه‌های احتمال دیراکی، همانند ساختن اعداد حقیقی با در نظر گرفتن دنباله های کوشی از اعداد گویا است. در اینجا مجموعه‌ی $\mathcal{M}_۱^*((X,d))$ حکم اعداد گویا را برای ما را دارد و دنباله‌های کوشی آن قرار است اندازه‌های احتمال دلخواه را بسازند. در ادامه درباره‌ی کامل‌سازی این فضا و خواص آن صحبت خواهیم کرد. فضای تمام دنباله‌های کوشی از اندازه‌های دیراکی را در نظر بگیرید:

$$\mathcal{C}_۱[(X,d)] := \{\underline{\mu} : (\mu_n)_{n\in\mathbb{N}} |\ \text{یک دنباله کوشی در } \mathcal{M}_۱^*((X,d)) \text{ است } (\mu_n)_{n\in\mathbb{N}}\}.$$

رابطه‌ی هم‌ارزی $\sim'$ را روی این فضا این‌گونه تعریف کنید که دو دنباله‌ی کوشی $(\mu_n)_{n\in\mathbb{N}}$ و $(\nu_n)_{n\in\mathbb{N}}$ معادلند هرگاه

$$\lim_{n\in\mathbb{N}} d_w(\mu_n,\nu_n) = ۰.$$

اگر فضای متریک ما کامل بود، این تعریف با این که بگوییم هر دو دنباله به یک نقطه میل کنند یکسان می‌شد. ولی ممکن است مانند مثالی که در انتهای فصل قبل آورده شد، دنباله‌ای کوشی باشد، ولی حد نداشته باشد. حال خارج قسمت $\mathcal{C}_۱[(x,d)]$ نسبت به این رابطه‌ی هم‌ارزی را در نظر بگیرید:

$$\overline{\mathcal{M}}_۱(X,d) := \mathcal{C}_۱[(x,d)]/\sim'.$$

اگر یک اندازه‌ی دیراکی داشته باشیم، می‌توانیم دنباله‌ی کوشی ثابت را در نظر بگیریم که تمام اعضای دنباله همان اندازه‌ی دیراکی اولیه هستند. پس می‌بینیم که فضای $\overline{\mathcal{M}}_۱(X,d)$ یک کپی از $\mathcal{M}_۱^*((X,d))$ را در خود دارد. و قبلا دیدیم که اگر X نامتناهی باشد، حتما اعضای بیشتری نیز دارد. حال می‌خواهیم فضای $\overline{\mathcal{M}}_۱(X,d)$ را مجهز به متریکی کنیم که تحدید آن به کپی موجود از $\mathcal{M}_۱^*((X,d))$ در این فضا، همان متریک d_w باشد. برای دو کلاس هم‌ارزی $[(\mu_n)_{n\in\mathbb{N}}]$ و $[(\nu_n)_{n\in\mathbb{N}}]$ تعریف کنید

$$\overline{d}_w([(\mu_n)_{n\in\mathbb{N}}],[(\nu_n)_{n\in\mathbb{N}}]) := \lim_{n\to\infty} d_w(\mu_n,\nu_n).$$

می‌توان چک کرد که در این تعریف این که چه نماینده‌هایی از این دو کلاس هم‌ارزی انتخاب شود فرقی ایجاد نمی‌کند. به راحتی می‌توان چک کرد که $\overline{d}_w$ در سه خاصیت متریک صدق می‌کند و همچنین $\overline{\mathcal{M}}_۱(X,d)$ با این متریک کامل است.

فضای متریک $(\overline{\mathcal{M}}_۱(X,d),\overline{d}_w)$ کامل است.

از حالا به بعد برای نمایش کپی‌هایی از $\mathcal{M}_۱^k((X,d))$ ها و $\mathcal{M}_۱^*((X,d))$ که در $\overline{\mathcal{M}}_۱(X,d)$ وجود دارند، با سوء استفاده از نمادگذاری، از خود این اسامی استفاده می‌کنیم:

$$\mathcal{M}_۱^k((X,d)) \subset \mathcal{M}_۱^*((X,d)) \subset \overline{\mathcal{M}}_۱(X,d).$$

گزاره ۱.۶. فضای متریک $(\overline{\mathcal{M}}_۱(X,d),\overline{d}_w)$ فشرده است.

اثبات. برای اثبات به این نکته دقت کنید که اگر $(\underline{\mu}_n)_n$ دنباله‌ای از دنباله‌های کوشی باشد، آنگاه می‌توان نزدیک‌ترین نقطه روی $\mathcal{M}_۱^k(X,d)$ به عضو k ام از $\underline{\mu}_n$ ها را در نظر گرفت، و از فشردگی هر یک از $\mathcal{M}_۱^k(X,d)$ ها به همراه گزاره‌ی ۱۱.۵ استفاده کرد و دنباله‌ای کوشی ساخت که زیردنباله‌ای از $\underline{\mu}_n$ ها به آن میل می‌کند. □

۷. تناظر بین $\mathcal{M}_1(X,d)$ و $\overline{\mathcal{M}}_1(X,d)$

در بخش قبل به فضای $(\overline{\mathcal{M}}_1(X,d),\overline{d}_w)$ رسیدیم که یک فضای متریک کامل و شامل (یک کپی از) تمام اندازه‌های دیراکی بود. اگر تعریفی از اندازه‌ی احتمال که در ابتدای این نوشته ارائه شد را فراموش کنیم، می‌توانیم اعضای این فضا را به عنوان اندازه‌های احتمال بر روی فضای (X,d) در نظر بگیریم. در این صورت دیگر نیازی به درگیر شدن با تعاریفی از قبیل سیگما_جبر، مجموعه‌های بورل و غیره که به صورت استاندارد در تعریف اندازه‌ی احتمال نقش ایفا می‌کنند نداریم، چون هیچ یک از این مفاهیم در تعریف جدیدی که از اندازه‌ی احتمال پیشنهاد دادیم ظاهر نمی‌شوند. اندازه‌های احتمال جدید ما یا اندازه‌ی احتمال دیراکی هستند، یا یک دنباله‌ی کوشی از چنین اندازه‌هایی. البته باید دقیق‌تر باشیم و بگوییم یک کلاس هم‌ارزی از دنباله‌های کوشی. بیایید چنین تعریف و دیدگاهی را با تعریف مرسوم از اندازه‌ی احتمال مقایسه کنیم و ببینیم آیا این تعریف نیز می‌تواند جایگزینی برای تعریف مرسوم باشد؟ آیا در کاربرد، و اثبات قضایای مربوط به اندازه‌های احتمال مواردی وجود دارد که استفاده از این تعریف کار را نسبت به استفاده از تعریف مرسوم راحت‌تر کند؟ یا حداقل آیا خواصی از اندازه‌های احتمال را که برای تعریف مرسوم به راحتی می‌توان چک کرد، برای این تعریف نیز می‌توان به راحتی چک کرد؟ و از همه مهم‌تر، موجوداتی که به عنوان اندازه‌ی احتمال معرفی کردیم، چه نسبتی با اندازه‌های احتمال مرسوم دارند؟ آیا تناظری یک‌به‌یک و طبیعی بین این موجودات و اندازه‌های احتمال مرسوم وجود دارد؟

به طور مثال از تعریف مرسوم به راحتی نتیجه می‌شود که ترکیب محدب دو اندازه‌ی احتمال، نیز از جنس اندازه‌ی احتمال است. چک کردن این نکته در $\overline{\mathcal{M}}_1(X,d)$ نیز کار راحتی است، و می‌توان دو کلاس هم‌ارزی را ترکیب محدب گرفت.

اما یکی از اولین انتظاراتی که ما از یک اندازه‌ی احتمال داریم، این است که اندازه‌ی یک زیرمجموعه‌ی اندازه‌پذیر را به ما بدهد. اصلا تعریف مرسوم از چنین جایی شروع می‌شود. ولی فعلا برای ما معلوم نیست که یک کلاس هم‌ارزی از دنباله‌های کوشی دیراکی به زیرمجموعه‌های فضا چه عددی باید نسبت دهد. پس به نظر در اولین قدم یک چنین تعریفی دچار مشکل می‌شود. اما بیایید سعی کنیم ببینیم چطور می‌توان از روی یک کلاس هم‌ارزی در $\overline{\mathcal{M}}_1(X,d)$ به زیرمجموعه‌های فضا، اندازه نسبت داد.

فرض کنید $\underline{\mu}=(\mu_n)_n$ یک دنباله‌ی کوشی از اندازه‌های دیراکی، و $[\underline{\mu}]$ کلاس هم‌ارزی آن باشد. اولین ایده‌ای که به ذهن می‌رسد این است که برای اندازه‌ی زیرمجموعه‌ای مثل $U\subset X$ چنین مقداری را در نظر بگیریم:

$$\lim_{n\to\infty}\mu_n(U).$$

اما به سرعت می‌توان دید که به چندین دلیل این مقدار گزینه‌ی درستی نیست. اول این که ممکن است حد بالا وجود نداشته باشد. دوم این که در صورت وجود ممکن است به این که ما چه نماینده‌ای از کلاس هم‌ارزی $[\underline{\mu}]$ انتخاب کنیم حساس باشد. به طور مثال اگر X را بازه‌ی $[0,1]$ و U را بازه‌ی $(0,1)$ در نظر بگیریم و تعریف کنیم

$$\mu_n=\delta_{\frac{1}{n}},$$

که دنباله‌ای کوشی است، می‌بینیم حد بالا برای این دنبله برابر ۱ است. اما می‌توانیم دنباله‌ای کوشی هم‌ارز این دنباله را در نظر بگیریم :

$$\mu_n=\delta_0.$$

و با انتخاب یکی در میان از این دو دنباله، دنباله‌ای بسازیم که حد برای آن موجود نباشد. اما از طرفی دقت کنید که هر نماینده‌ای که در این کلاس هم‌ارزی بگیریم، از جایی به بعد اعضای آن، بیشتر جرمشان در نزدیکی صفر است و حتی اگر این جرم‌ها داخل مجموعه‌ی U باشند، با انرژی خیلی کمی می‌توان آن‌ها را به خارج U هل داد. لذا به نظر خوب است که در بالا به جای $\lim$ از $\liminf$ استفاده کنیم. اینگونه مشکل وجود نداشتن حد نیز حل می‌شود. اما همچنان وابستگی به دنباله‌ی انتخاب شده از کلاس هم‌ارزی وجود دارد. برای رفع این مشکل هم می‌توان روی تمام نماینده‌های کلاس هم‌ارزی، اینفیمم گرفت و در نهایت به چنین مقداری برای اندازه‌ی U می‌رسیم:

$$\inf_{(\mu_n)_n\in[\underline{\mu}]}\liminf_{n\to\infty}\mu_n(U).$$

به نظر می‌رسد که این تعریف گزینه‌ی مناسبی برای اندازه‌ی حاصل از کلاس هم‌ارزی یک دنباله‌ی کوشی باشد. اما نه لزوما برای هر نوع مجموعه‌ای! به طور مثال اگر U مجموعه‌ی تک نقطه‌ای x باشد این مقدار صفر می‌شود، ولی انتظار ما این است که این کلاس هم‌ارزی همان اندازه‌ی دیراک روی صفر باشد. یا مثلا اگر U را مجموعه‌ی اعداد گنگ در $[۰, ۱]$ بگیرید و دنباله‌ی کوشی را به صورت زیر تعریف کنید:

$$\mu_n = \frac{۱}{n}\Sigma_{i=۱}^{n}\delta_{\frac{i}{n}},$$

این دنباله‌ی کوشی اگر بخواهد معادل یک اندازه باشد، آن اندازه باید اندازه‌ی لبگ باشد، ولی می‌توان دید که $\mu^*(U)$ برابر صفر خواهد شد، ولی اندازه‌ی مجموعه‌ی اعداد گنگ برای اندازه‌ی لبگ برابر یک است.

در دو مثال بالا مجموعه‌ی U مجموعه‌ای باز نیست. درادامه می‌بینیم که تعریف بالا روی خانواده‌ی زیرمجموعه‌های باز X همان خواصی را دارد که یک اندازه‌ی احتمال باید داشته باشد.

مجموعه‌ی تمام زیرمجموعه‌های باز X را با $\mathcal{U}$ نمایش می‌دهیم. دیدیم که به هر کلاس هم‌ارزی $[\underline{\mu}]$ می‌توان تابعی از $\mathcal{U}$ به $\mathbb{R}^{\geq ۰}$ نسبت داد. نام این نگاشت را ϕ^* می‌نامیم:

$$\phi^* : \overline{\mathcal{M}}_۱(X,d) \to \{\mathcal{U}\text{توابع حقیقی مقدار روی}\}$$

$$[\underline{\mu}] \mapsto \inf_{(\mu_n)_n\in[\underline{\mu}]} \liminf_{n\to\infty} \mu_n(.).$$

گزاره ۱.۷. برای هر $[\underline{\mu}] \in \overline{\mathcal{M}}_۱(X,d)$ نگاشت $\phi^*([\underline{\mu}])$ بر روی مجموعه‌های باز، دارای خاصیت سیگما_جمع پذیری است، به این معنی که اگر $U_۱, U_۲, ...$ تعدادی شمارا (یا متناهی) از مجموعه‌های باز دوبه دو مجزا باشند، آنگاه

$$\phi^*([\underline{\mu}])(\cup_{i=۱}^{\infty} U_i) = \Sigma_{i=۱}^{\infty}\phi^*([\underline{\mu}])(U_i).$$

اثبات. برای اثبات به این نکته دقت کنید که می‌توان در یک کلاس هم‌ارزی، دنباله‌ای کوشی یافت که اینفیمم را به طور همزمان، هم برای $\cup_{i=۱}^{\infty} U_i$ بدهد، و هم برای تک‌تک U_i ها. این‌گونه که با شروع از یک دنباله‌ی کوشی دلخواه، برای هر عضو μ_n از دنباله، جرمی از آن اندازه را که در درون و در فاصله‌ی نزدیکی از مرز U_i ها قرار دارد دقیقا به روی مرز آن U_i منتقل می‌کنیم. این‌گونه جرم درون یک U_i را داخل هیچ U_j دیگری نریختیم. حال هرچقدر اندیس μ_n بزرگتر شود، فاصله کمتری از مرز U_i ها برای انتقال جرم آن به روی مرز در نظر می‌گیریم. می‌توان دید طی این فرآیند به دنباله‌ای جدید از اندازه‌ها می‌رسیم که کوشی است، و با دنباله‌ی اولیه هم‌ارز است. دیدن این نکته سخت نیست که این دنباله مقدار هر دو طرف تساوی را محقق می‌کند. □

حال می‌خواهیم از روی $\phi^*([\underline{\mu}])$ یک اندازه‌ی احتمال با سیگما_جبر بورل بسازیم. برای افرادی که تجربه‌ی احتمالاتی یا نظریه‌ی اندازه‌ای دارند، خاصیت $\phi^*([\underline{\mu}])$ که در گزاره‌ی ۱.۷ به آن اشاره شد آشناست و شبیه خواص پیش_اندازه [1] است. یک پیش اندازه بر روی یک جبر از زیرمجموعه‌ها [2] تعریف می‌شود، که نسبت به خاصیت مکمل‌گیری، اجتماع و اشتراک متناهی بسته است. ثابت می‌شود [3] که هر پیش اندازه‌ی متناهی به صورت یکتا به یک اندازه گسترش داده می‌شود. اگر ما بتوانیم از این قضیه استفاده کنیم کار تمام است، ولی مجموعه‌ی زیرمجموعه‌های باز نسبت به مکمل‌گیری بسته نیست. اما حل این مشکل نیز کار سختی نیست و می‌توانیم در نهایت از روی یک دنباله‌ی کوشی یک اندازه بر روی سیگما_جبر بورل بسازیم:

گزاره ۲.۷. برای هر کلاس هم‌ارزی $[\underline{\mu}] \in \overline{\mathcal{M}}_۱(X,d)$ اندازه‌ی احتمال یکتای $\phi([\underline{\mu}])$ چنان موجود است که بر روی زیرمجموعه‌های باز با $\phi^*([\underline{\mu}])$ برابر است:

$$\forall U \in \mathcal{U}, \phi([\underline{\mu}])(U) = \phi^*([\underline{\mu}])(U).$$

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. □

---
[1] pre-measure
[2] algebra of subsets
[3] Carathéodory's extension theorem

می‌توانیم در گزاره‌ی بالا ϕ را به عنوان نگاشتی از $\overline{\mathcal{M}}_1(X,d)$ به $\mathcal{M}_1(X,d)$ در نظر بگیریم:

$$\phi : \overline{\mathcal{M}}_1(X,d) \to \mathcal{M}_1(X,d).$$

گزاره ۳.۷. نگاشت $\phi : \overline{\mathcal{M}}_1(X,d) \to \mathcal{M}_1(X,d)$

- یک‌به‌یک است،
- بر روی $\mathcal{M}_1^*(X,d)$ به صورت همانی عمل می‌کند:

$$\forall \mu \in \mathcal{M}_1^*(X,d), \quad \phi([(\mu_n = \mu)_n]) = \mu.$$

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. □

سوالی که مطرح می‌شود این است که آیا ϕ یک تناظر یک‌به‌یک است؟ یعنی باید ببینیم آیا ϕ پوشاست یا خیر. برای این کار باید ببینیم که چطور می‌توان در جهت عکس حرکت کرد و به یک اندازه‌ی احتمال دلخواه، دنباله‌ای کوشی از اندازه‌های دیراکی نسبت داد.

فرض کنید $\nu \in \mathcal{M}_1(X,d)$ اندازه‌ی احتمالی دلخواه و نه لزوما دیراکی باشد. ابتدا به صورت استقرایی، دنباله‌ای از افرازهای فضای X را می‌سازیم که قطر اعضای آن‌ها کم و کمتر می‌شود. برای نقاط $x_1^k, ..., x_k^k$ در X عدد مثبت ϵ_k را طوری درنظر بگیرید که در خواص زیر صدق کنند:

- مجموعه‌ی $\{x_1^k, ..., x_k^k\}$ در X یک مجموعه‌ی ϵ_k-چگال باشد،
- $\lim_{k\to\infty} \epsilon_k = 0$

برای هر k مجموعه‌ی تمام گوی‌های بسته به شعاع ϵ_k حول x_j^k ها و مکمل آن‌ها را داخل مجموعه‌ی F_k بریزید و قرار دهید

$$G_k := \cup_{s=1}^k F_k.$$

G_k شامل تعدادی متناهی عضو است. تمام اشتراکات ممکن این متناهی مجموعه نیز متناهی تا مجموعه خواهند بود. از بین تمام این اشتراکات، اعضای مینیمال را که به صورت سره شامل هیچ اشتراک دیگری نیستند در نظر بگیرید و داخل مجموعه‌ی P_k بریزید. از آنجایی که هر عضو X در یک اشتراک مینیمال خواهد افتاد، و اشتراکات مینیمال با هم اشتراک ندارند، P_k یک افراز از X به ما می‌دهد که قطر اعضای آن حداکثر ϵ_k است. حال بیایید دنباله‌ی دیراکی ν_k را به این صورت بسازید که داخل هر یک از اعضای افراز P_k یک عضو انتخاب کنید. اگر تعداد اعضای P_k برابر N_k باشد، اعضای افراز را با $A_1^k, A_2^k, ..., A_{N_k}^k$ و نقاط انتخاب شده از آن‌ها را با $y_1^k, ..., y_{N_k}^k$ نمایش می‌دهیم. تعریف کنید :

$$\nu_k := \Sigma_{i=1}^{N_k} \nu(A_i^k)\delta_{y_i^k}.$$

یعنی اندازه‌ی ν_k اندازه‌ای دیراکی بر روی نقاط انتخاب شده از هر عضو افراز است که اندازه‌ی ν به ما می‌گوید به هر نقطه، اندازه‌ی عضوی از افراز که شامل آن است را نسبت دهیم.

لم ۴.۷. دنباله‌ی $(\nu_k)_k$ کوشی است و کلاس هم‌ارزی این دنباله‌ی کوشی به انتخاب x_i^k ها و ϵ_k ها بستگی ندارد.

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. □

با توجه به لم بالا می‌توان نگاشت

$$\psi : \mathcal{M}_1(X,d) \to \overline{\mathcal{M}}_1(X,d)$$

را به این صورت تعریف کرد که اندازه‌ی ν را به کلاس هم‌ارزی دنباله‌ی $(\nu_k)_k$ که در فرآیند بالا ساخته شد ببرد.

گزاره ۵.۷. نگاشت $\psi : \mathcal{M}_1(X,d) \to \overline{\mathcal{M}}_1(X,d)$ معکوس نگاشت $\phi : \overline{\mathcal{M}}_1(X,d) \to \mathcal{M}_1(X,d)$ است.

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. □

پس بالاخره دیدیم که تناظری یک‌به‌یک بین دنباله‌های کوشی اندازه‌های دیراکی و اندازه‌های احتمال برقرار است.

۸. انتقال ساختار متریک از $\overline{\mathcal{M}}_1(X,d)$ به $\mathcal{M}_1(X,d)$

با توجه به تناظر یک‌به‌یکی که در بخش قبل به دست آمد می‌توانیم متریک $\overline{d}_w$ را به فضای اندازه‌های احتمال منتقل کنیم. از آنجایی که تناظر ما بر روی زیرمجموعه‌ی $\mathcal{M}_1^*(X,d)$ همانی بود، متریک روی این زیرمجموعه از $\mathcal{M}_1(X,d)$، همان متریک d_w خواهد بود که در ابتدا آن را تعریف کردیم. در حالت کلی نیز، متریک جدید بین دو اندازه‌ی احتمال دلخواه μ و ν را با $d_w(\mu,\nu)$ نمایش می‌دهیم

$$d_w(\mu,\nu) := \overline{d}_w(\psi(\mu),\psi(\nu)).$$

با توجه به پوشا بودن ψ نتیجه می‌گیریم که این نگاشت یک ایزومتری بین دو فضای $\overline{\mathcal{M}}_1(X,d)$ و $\mathcal{M}_1(X,d)$ است، لذا هر خاصیتی که تا کنون برای $(\overline{\mathcal{M}}_1(X,d),\overline{d}_w)$ اثبات کردیم، برای $(\mathcal{M}_1(X,d),d_w)$ نیز برقرار خواهد بود، به طور خاص:

نتیجه ۱.۸. فضای $(\mathcal{M}_1(X,d),d_w)$ فشرده است.

۹. بحث و نتیجه گیری

ما با تعریف یک متریک بر روی فضای اندازه‌های احتمال دیراکی شروع کردیم. سپس این فضا را کامل کردیم و به فضای متریک جدیدی رسیدیم که ثابت کردیم در تناظری یک‌به‌یک با فضای اندازه‌های احتمال قرار دارد. از روی این تناظر، متریکی بر روی فضای اندازه‌های احتمال به دست آوردیم که تحدیدش به اندازه‌های احتمال دیراکی، همان متریکی بود که از آن شروع کرده بودیم. لذا این‌گونه توانستیم بین هر دو اندازه‌ی احتمال دلخواه یک فاصله تعریف کنیم. خوب است به این نکته اشاره کنیم که مسیری که طی کردیم، می‌تواند به عنوان جایگزینی برای تعریف اندازه‌ی احتمال بر روی فضاهای متریک در نظر گرفته شود. یعنی به جای این که یک اندازه‌ی احتمال را تابعی تعریف کنیم که به اعضای سیگما_جبر بورل اعدادی نسبت می‌دهد که دارای خواص ذکر شده در ابتدای مقاله باشد، هر اندازه را به عنوان کلاس هم‌ارزی یک دنباله‌ی کوشی از اندازه‌های دیراکی ببینیم. البته در چنین دیدگاهی برای تعریف اندازه‌های احتمال، فقط اندازه‌های احتمال بر روی سیگما_جبر بورل را می‌توانیم بدست آوریم. این که با لفظ "تعریف جدید" از کاری که انجام دادیم یاد می‌کنیم از این منظر نیست که این تعریف را جایگزین تعریف مرسوم از اندازه‌های احتمال در نظر بگیریم، بلکه این است که می‌توان به عنوان دو تعریف معادل به این دو نگاه کرد و با توجه به موقعیت و کاربردی که مد نظرمان است، از هرکدام که کارمان را راحت‌تر به ثمر می‌رساند استفاده کنیم.

ذکر این نکته نیز واجب است که متریک d_w بر روی فضای $\mathcal{M}_1(X,d)$ متریک مشهوری است که سال‌ها پیش معرفی شده و مورد مطالعه قرار گرفته است. اما تعریف مرسوم این متریک به صورت مستقیم بین دو اندازه‌ی احتمال دلخواه انجام شده و این‌گونه نیست که ابتدا بر روی اندازه‌های دیراکی تعریف شده باشد. در ادبیات مربوطه، فاصله d_w اسامی متفاوتی دارد، که می توان به فاصله‌ی واسرشتاین [۱] ، فاصله‌ی کانتروویچ_روبینشتاین [۲] و یا فاصله‌ی بولدوزر [۳] اشاره کرد. در تعریف مرسوم، یک طرح انتقال به صورت کلی بین دو اندازه‌ی احتمال نه لزوما دیراکی تعریف می‌شود، و سپس انرژی طرح انتقال تعریف می‌شود و در نهایت فاصله را به عنوان اینفیمم انرژی‌های تمام طرح‌های انتقال ممکن در نظر گرفته می‌شود. ما نیز از تعریف مرسوم بر روی اندازه‌های دیراکی شروع کردیم، ولی از روشی دیگر آن را بین دو اندازه‌ی دلخواه گسترش دادیم. احکام زیادی مرتبط با هندسه‌ی این فضا با استفاده از تعریف مرسوم به دست آمده که ممکن است با استفاده از تعریف جدیدی که در اینجا ارائه دادیم با روشی متفاوت اثبات شوند. به طور مثال می توانید اثبات فشردگی این فضا را در [۱] ببینید و با اثباتی که در اینجا آورده شده مقایسه کنید.

به عنوان حسن ختام مقاله، بگذارید کمی در مورد ماهیت هندسی این طرز نگاه به فضای اندازه‌های احتمال صحبت کنیم. فضای حاصل ضرب‌های متقارن، در بعضی از شاخه‌های ریاضی بسیار مورد توجه قرار می‌گیرند و مطالعه می‌شوند. به طور مثال وقتی X یک رویه ریمانی دو بعدی است، حاصل ضرب‌های متقارن آن، منیفلدهایی مختلط خواهند شد که به ساختارهای بیشتری نیز مجهز هستند و احکام زیادی در مورد آن‌ها یافت شده است. این نکته بسیار جالب است که فضای $\mathcal{M}_1^*((X,d))$ تمام حاصل ضرب‌های متقارن فضای X را در خود دارد. یا حتی اگر از طرف دیگر به ماجرا نگاه کنیم، این که فضاهای حاصل

---
[1] Wasserestein distance

[2] Kantorovich–Rubinstein distance

[3] earth mover distance

ضرب‌های متقارن به طور همزمان در یک فضای بزرگتر در کنار یکدیگر نشسته‌اند نیز بسیار جالب است. اگر تعریف کنیم

$$\mathcal{E}_1^*(X,d) := \cup_{k=1}^{\infty} \mathcal{E}_1^k(X,d),$$

آنگاه می‌توانیم فرآیند کامل‌سازی‌ای که برای $\mathcal{M}_1^*(X,d)$ انجام دادیم و فضای اندازه‌های احتمال را گرفتیم، برای $\mathcal{E}_1^*(X,d)$ نیز انجام دهیم و دوباره در این حالت نیز به فضای اندازه‌های احتمال می‌رسیم. این نکته به ما می‌گوید که $\mathcal{E}_1^k(X,d)$ برای k بزرگ، در فضای اندازه‌های احتمال می‌تواند به دلخواه چگال باشد. داشتن فضایی تقریبا چگال از اندازه‌های احتمال که بسیار ساختارمند باشد و اعمال هندسی زیادی را بتوان بر روی آن انجام داد، این قابلیت را دارد که مسیرهای جدیدی برای فکر در مورد فضای اندازه‌های احتمال و مسائل مرتبط با آن را باز کند.

مراجع

[1] Figalli, A., & Glaudo, F. (2023). *An invitation to optimal transport, Wasserstein distances, and gradient flows* (2nd edition). Berlin: European Mathematical Society (EMS).

* **دانشگاه صنعتی شریف**
رایانامه: amin.s.talebi@gmail.com