جلسه ۲۴

در این جلسه به ادامه اثبات خواص ذکر شده برای تابع آنتروپی کوانتومی می‌پردازیم.

۱ آنتروپی اندازه‌گیری

قضیه ۱ فرض کنید که سیستمی در حالت ρ داریم. یک اندازه‌گیری در «پایه متعامد یکه» دلخواه $\{|y_j\rangle\}$ روی این سیستم انجام می دهیم (با استفاده از $M_j = |y_j\rangle\langle y_j|$). احتمال اینکه حاصل اندازه‌گیری j باشد، برابر است با $q_j = \langle y_j|\rho|y_j\rangle$.
در این صورت آنتروپی احتمالات مربوط به حاصل اندازه‌گیری بزرگ‌تر مساوی آنتروپی ρ است:

$$H(\rho) \leq H(\{q_j\}).$$

نکته ۲ آنتروپی فون نیومن $H(\rho)$ به این معناست که ماتریس چگالی را قطری کرده و آنتروپی شانون مقادیر ویژه آن را محاسبه کنیم. $H(\{q_j\})$ نیز به این معناست که ماتریس چگالی را در پایه $|y_j\rangle$ نوشته و آنتروپی شانون عناصر روی قطر اصلی را بدست آوریم. پس قضیه فوق را می‌توان به این شکل تفسیر کرد: آنتروپی فون نیومن یک ماتریس چگالی همواره بزرگتر یا مساوی آنتروپی عناصر روی قطر اصلی نمایش ماتریس چگالی در هر پایه دلخواه است.

در اینجا دو اثبات برای قضیه فوق می آوریم.

۱.۱ اثبات اول قضیه ۱

ابتدا پایه‌ی متعامد یکه‌ای که ماتریس چگالی در آن می‌شود را درنظر می‌گیریم:

$$\rho = \sum_i p(x_i)\ |x_i\rangle\langle x_i|$$

که در اینجا $|x_i\rangle$ برای مقادیر مختلف x_i تنها یک برچسب گذاری روی بردارهای ویژه ρ است. همچنین چون جمع مقادیر ویژه ρ برابر یک است، آنها را می‌توان با یک توزیع احتمال نشان داد که در اینجا از نمادگذاری $p(x_i)$ استفاده کرده‌ایم.

حال پایه متعامد یکه با اعضای $|y_j\rangle$ را در نظر بگیرید. اگر ماتریس ρ را در این پایه بنویسیم، در حالت کلی ماتریس حاصل قطری نخواهد بود. می‌دانیم عناصر روی قطر اصلی ماتریس در پایه جدید به صورت زیر تعریف می‌شوند :

$$\begin{aligned} q_j &= \langle y_j|\rho|y_j\rangle \\ &= \langle y_j|\big(\sum_i p(x_i)|x_i\rangle\langle x_i|\big)|y_j\rangle \\ &= \sum_i p(x_i)|\langle y_j|x_i\rangle|^2. \end{aligned}$$

ادعا می‌کنیم که می‌توان توزیع شرطی‌ای به شکل زیر تعریف کرد:

$$p(Y = y_j|X = x_i) = |\langle y_j|x_i\rangle|^2.$$

این تعریف از احتمال شرطی مشروع است زیرا به ازای هر x_i

$$|\langle y_j|x_i\rangle|^2 \geq 0, \qquad \sum_j |\langle y_j|x_i\rangle|^2 = 1.$$

با استفاده از این تعریف می‌توان توزیع احتمال حاشیه‌ای Y را یافت. با درنظر گرفتن معادلات فوق و قضیه احتمال کل داریم $q_j = p(y_j)$.

هدف ما این است که نشان دهیم آنتروپی شانون توزیع $\{q_j\}$ از آنتروپی شانون مقدار ویژه‌های ρ یا $\{p(x_i)\}$ بیشتر است. می‌توان به توزیع Y به عنوان متغیر تصادفی حاصل از عبور دادن X از کانال $p(y|x)$ نگاه کرد. پس برای اثبات قضیه، کافیست نشان دهیم عبور دادن یک متغیر تصادفی کلاسیک از کانال $p(y|x)$ باعث افزایش آنتروپی آن می‌شود.

بدین منظور، تعریف می کنیم:

$$\sigma = \sum_j p(y_j)|y_j\rangle\langle y_j|$$

ثابت می کنیم که

$$D(\sigma||\rho) = H(\sigma) - H(\rho)$$

که با استفاده از نامنفی بودن آنتروپی نسبی نتیجه می‌گیریم $H(\sigma) \geq H(\rho)$ و با توجه به تعریف σ ، قضیه اثبات می شود. می‌دانیم که

$$D(\sigma||\rho) = \text{tr}(\rho \log \rho - \rho \log \sigma) = -H(\rho) - \text{tr}(\rho \log \sigma)$$

ثابت می‌کنیم :

$$\text{tr}(\rho \log \sigma) = -H(\sigma).$$

توجه کنید که

$$\log \sigma = \sum_j \log(p(y_j))|y_j\rangle\langle y_j|.$$

داریم:

$$\begin{aligned}
\mathrm{tr}(\rho\log\sigma) &= \mathrm{tr}\big((\sum_i p(x_i)|x_i\rangle\langle x_i|)(\sum_j \log p(y_j)|y_j\rangle\langle y_j|)\big)\\
&= \mathrm{tr}(\sum_{ij} p(x_i)\log p(y_j)|x_i\rangle\langle x_i|y_j\rangle\langle y_j|)\\
&= \mathrm{tr}(\sum_{ij} p(x_i)\log p(y_j)\langle x_j|y_j\rangle|x_i\rangle\langle y_j|)\\
&= \mathrm{tr}(\sum_{ij} p(x_i)\log p(y_j)\langle x_j|y_j\rangle\langle y_j|x_i\rangle)\\
&= \sum_{ij} p(x_i)\log(p(y_j))\ |\langle x_i|y_j\rangle|^2\\
&= \sum_{ij} p(x_i)p(y_j|x_i)\log p(y_j)\\
&= \sum_{j} p(y_j)\log p(y_j)\\
&= -H(\sigma)
\end{aligned}$$

نکته ۳ شکل فشرده تری از اثبات بالا را می‌توانید در اثبات قضیه ۱۱.۹ صفحه ۵۱۵ از کتاب درسی بیابید.

۲.۱ اثبات دوم قضیه ۱

در اینجا، می‌خواهیم با استفاده از یکی از خواص جالب دنباله‌ها، قضیه اصلی را ثابت کنیم. برای این منظور ابتدا به بررسی خاصیتی از دنباله‌ها بنام غلبه کردن[1] می‌پردازیم.

۳.۱ غلبه کردن دنباله‌ها

برای وضوح نمادگذاری در ادامه از حروف پررنگ برای نشان دادن دنباله‌ها و از حروف معمولی برای نشان دادن مولفه‌های دنباله‌ها استفاده می‌کنیم. برای شروع نیاز به یک تعریف داریم:

تعریف ۴ فرض کنید که $\mathbf{a}$ دنباله دلخواهی باشد. در این صورت $\mathbf{a}^{\downarrow}$ برابر دنباله $\mathbf{a}$ است وقتی اعضای آن بصورت نزولی مرتب شده باشند. مثلا اگر $\mathbf{a} = (0,3,2)$ آنگاه $\mathbf{a}^{\downarrow} = (3,2,0)$. همچنین $a_i^{\downarrow}$ بیانگر مولفه i-ام دنباله $\mathbf{a}^{\downarrow}$ است. پس برای هر بردار d بعدی $\mathbf{a}$ داریم: $a_1^{\downarrow} = \max_i a_i$، $a_d^{\downarrow} = \min_i a_i$.

تعریف ۵ فرض کنید دو دنباله $\mathbf{a}$ و $\mathbf{b}$ را در فضای بردارهای d بعدی داریم. می‌گوییم دنباله $\mathbf{a}$ بر دنباله $\mathbf{b}$ غلبه می‌کند اگر و فقط اگر جمع k مولفه‌ی بزرگتر $\mathbf{a}$، بیشتر مساوی k مولفه‌ی بزرگتر $\mathbf{b}$ باشد و بعلاوه جمع مولفه‌های دو بردار با هم

[1] Majorization

مساوی باشند. به عبارت دیگر

$$\sum_{i=1}^{k} a_i^{\downarrow} \geq \sum_{i=1}^{k} b_i^{\downarrow}, \qquad k = 1, \ldots, d-1,$$

$$\sum_{i=1}^{d} a_i^{\downarrow} = \sum_{i=1}^{d} b_i^{\downarrow}.$$

غلبه دنباله $\mathbf{a}$ بر دنباله $\mathbf{b}$ را با نماد $\mathbf{a} \succ \mathbf{b}$ نشان می‌دهیم.

مثال ۶ ترتیب اعضای دو دنباله تاثیری بر غلبه کردن آنها بر هم ندارد. مثلا $(0,3) \prec (1,2)$ معادل $(3,0) \prec (2,1)$ می‌باشند. مثال دیگری از غلبه کردن این است که

$$(1,2,3) \prec (0,3,3) \prec (0,0,6).$$

همچنین برای هر مقدار دلخواه n داریم:

$$\left(\frac{1}{n}, \ldots, \frac{1}{n}\right) \prec \left(\frac{1}{n-1}, \ldots, \frac{1}{n-1}, 0\right) \prec \cdots \prec \left(\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0\right) \prec (1, 0, \ldots, 0).$$

تمرین ۷ ثابت کنید دنباله هر توزیع احتمال دلخواه بر روی الفبای n-تایی بر دنباله توزیع یکنواخت (بر روی الفبای n-تایی) غلبه می‌کند.

تمرین ۸ اگر جمع مولفه‌های دو بردار با هم مساوی نباشند، دو بردار قابل مقایسه با هم نیستند. حال دو بردار بسازید که جمع مولفه‌هایشان با هم برابر باشد ولی قابل مقایسه با یکدیگر نباشند (هیچ کدام بر دیگری غلبه نکند).

تمرین ۹ دنباله $\mathbf{a} = (p_1, p_2)$ را در نظر بگیرید. نشان دهید $\mathbf{a}$ بر دنباله‌ی دو تایی $\mathbf{b}$ غلبه می‌کند اگر و فقط اگر λ در بازه $[0,1]$ وجود داشته باشد به طوری که

$$\mathbf{b} = \lambda(p_1, p_2) + (1-\lambda)(p_2, p_1).$$

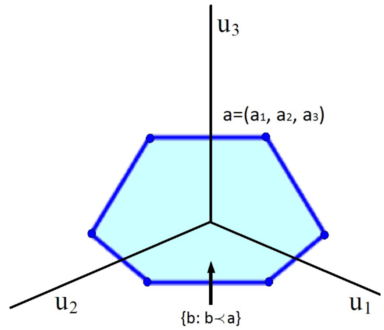
نتیجه بگیرید که $\mathbf{a} \succ \mathbf{b}$ خواهد بود، اگر و فقط اگر دنباله $\mathbf{b}$ در پوش محدب (p_1, p_2) و (p_2, p_1) قرار گیرد.

تمرین ۱۰ نشان دهید اگر $\mathbf{a} \succ \mathbf{b}$ و $\mathbf{a} \succ \mathbf{c}$ آنگاه

$$\mathbf{a} \succ \frac{1}{2}\mathbf{b} + \frac{1}{2}\mathbf{c}$$

که منظور از $\frac{1}{2}\mathbf{b} + \frac{1}{2}\mathbf{c}$ میانگین‌گیری مولفه‌های متناظر دو بردار است. نتیجه بگیرید که مجموعه بردارهایی که $\mathbf{a}$ بر آنها غلبه می‌کند محدب است.

در حالت کلی قضیه زیر را داریم:



شکل ۱: نمایش هندسی قضیه ۱۱ در بعد 3. تمام دنباله‌هایی که دنباله $\mathbf{a}$ بر آنها غلبه می‌کند با رنگ آبی مشخص شده‌اند. این دنباله‌ها ترکیب خطی از نقاط مرزی محدوده آبی رنگ هستند که همان جایگشت‌های اعضای دنباله $\mathbf{a}$ هستند.

قضیه ۱۱ (تعبیر هندسی غلبه کردن) برای دو دنباله $\mathbf{a}$ و $\mathbf{b}$، رابطه $\mathbf{a} \succ \mathbf{b}$ برقرار است اگر و فقط اگر دنباله $\mathbf{b}$ در پوش محدب دنباله‌هایی که از جایگشت‌های مختلف اعضای $\mathbf{a}$ بدست می‌آیند قرار گیرد.

اثبات قضیه بالا را در اینجا نمی‌آوریم، اما این قضیه بصورت تصویری در شکل ۳.۱ نشان داده شده است.

حال به اثبات قضیه ۱ باز می‌گردیم. این اثبات دو مرحله دارد: (۱) ابتدا ثابت می‌کنیم که دنباله مقادیر ویژه ρ بر دنباله $\{q_j\}$ غلبه می‌کند. (۲) اگر دنباله احتمالات $\mathbf{a}$ بر دنباله احتمالات $\mathbf{b}$ غلبه کند، آنگاه $H(\mathbf{a}) \leq H(\mathbf{b})$.

بخش اول اثبات: در مباحث جبر خطی در ابتدای درس دیدیم که برای هر بردار بطول واحد $|y\rangle$ مقدار $\langle y|\rho|y\rangle$ همواره کمتر مساوی بزرگترین مقدار ویژه ماتریس مثبت نیمه معین ρ است. در واقع

$$\lambda_{\max} = \max_{|e\rangle:\ \||e\rangle\|=1} \langle e|\rho|e\rangle$$

پس

$$\lambda_{\max} \geq \max_j q_j.$$

اگر بخواهیم نامساوی‌های باقیمانده را اثبات کنیم نیاز به تعمیمی از قضیه‌ای که در بالا استفاده کردیم داریم. ابتدا آن را بازنویسی می‌کنیم

$$\begin{aligned}\lambda_{\max} &= \max_{|e\rangle:\ \||e\rangle\|=1} \langle e|\rho|e\rangle \\ &= \max_{|e\rangle:\ \langle e|e\rangle=1} \mathrm{tr}(\langle e|\rho|e\rangle) \\ &= \max_{|e\rangle:\ \mathrm{tr}(|e\rangle\langle e|)=1} \mathrm{tr}(\rho|e\rangle\langle e|)\end{aligned}$$

در نتیجه، عبارت فوق را می توان با استفاده از عملگر تصویر $\Pi \geq 0$ این گونه نوشت:

$$p(x_1) = \max_{\Pi:\ \mathrm{rank}(\Pi)=1} \mathrm{tr}(\rho\Pi)$$

تعمیمی از این قضیه، به عنوان اصل ماکزیمم کی-فان [۲] برقرار است.

قضیه ۱۲ برای هر ماتریس مثبت معین دلخواه ρ جمع بیشترین k مقدار ویژه ρ برابر است با

$$\max_{\Pi:\, rank(\Pi)=k} tr(\rho\Pi)$$

که در آن ماکزیمم گیری روی عملگرهای تصویری $\Pi \geq 0$ است.

با استفاده از قضیه فوق به راحتی می‌توان دید که جمع بزرگترین k مقدار ویژه ρ بیشتر یا مساوی جمع k عضو بزرگ دنباله $\{q_j\}$ است. کافیست در اصل کی-فان، عملگر تصویر متناظر با زیرفضای k-بعدی تولید شده تعدادی از $\langle y_j|$-ها را در نظر گرفت.

بخش دوم اثبات: از قضیه زیر استفاده می‌کنیم.

قضیه ۱۳ اگر برای دو دنباله احتمال $\mathbf{a}$ و $\mathbf{b}$ داشته باشیم $\mathbf{a} \succ \mathbf{b}$ آنگاه:

$$H(b) \geq H(a)$$

اثبات: اثبات این قضیه برای دنباله‌های به طول 2، با استفاده از تعبیر هندسی که بیان کردیم واضح است. برای اثبات حالت کلی نیز از تعبیر هندسی استفاده می‌کنیم. می‌دانیم که می‌توان دنباله $\mathbf{b}$ که دنباله $\mathbf{a}$ بر آن غلبه کند را به صورت ترکیب محدبی از دنباله‌هایی که از جایگشت‌های مختلف مولفه‌های $\mathbf{a}$ بدست می‌آیند، نوشت. یعنی

$$\mathbf{b} = \sum_i \lambda_i \mathbf{a}_{\pi_i}$$

برای یک دنباله از وزن‌های نامنفی λ_i به طوری که $\sum_i \lambda_i = 1$. در اینجا هر یک از $\mathbf{a}_{\pi_i}$-ها یک جایگشت از دنباله $\mathbf{a}$ است. بنابر خاصیت تحدب تابع آنتروپی داریم:

$$\begin{aligned} H(\mathbf{b}) &= H(\sum_i \lambda_i \mathbf{a}_{\pi_i}) \\ &\leq \sum_i \lambda_i H(\mathbf{a}_{\pi_i}) \\ &= \sum_i \lambda_i H(\mathbf{a}) \\ &= H(\mathbf{a}) \end{aligned}$$

که تساوی $H(\mathbf{a}_{\pi_i}) = H(\mathbf{a})$ به این دلیل برقرار است که اعضای هر $\mathbf{a}_{\pi_i}$ جایگشتی از اعضای $\mathbf{a}$ است و درنتیجه آنتروپی تمام دنباله‌های مذبور برابر با آنتروپی همان دنباله اصلی $\mathbf{a}$ است.

[۲] Ky Fan's Maximum Principle