کران‌های بالا و پایین

برای فرآیندهای تصادفی

آنالیز احتمالاتی ابعاد بالا

استاد: یاسایی میبدی

باتشکر ویژه از: آقایان مسیحا و منیری

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

تابستان ۱۴۰۰

بسم‌الله الرّحمن الرّحیم

فهرست مطالب

پیشگفتار

این گزارش نتیجه پژوهش و مطالعه کتاب کران‌های بالا و پایین برای فرآیندهای تصادفی آقای تالاگراند است. در این گزارش ضمن بررسی مفصل فصل‌های اول تا هفتم این کتاب، سعی شده تا شهودی که پشت قضایای کتاب پنهان بوده بیان گردد و بعضا با استفاده از تصاویر مناسب درک بهتری به مخاطب داده شود. در ضمن برخلاف رویکرد بسیار جبری کتاب، در این گزارش از ذکر مفصل محاسباتی این اثبات‌ها می‌گذریم و به خواننده مشتاق توصیه می‌کنیم برای دیدن اثبات دقیق آن‌ها به کتاب مراجعه نماید.

در لابه‌لای این پژوهش و در لایه‌های متفاوت کتاب؛ حس کردم به مبحثی جدید پاگذاشتم. غیر از تخمین‌هایی که مربوط به آنالیز و پیوستگی توپولوژیک است، و غیر از ماهیت ابعاد بالا و احتمالاتی قضایا؛ بنظرم تئوری پشت قضایا می‌تواند پلی جدید میان هندسه‌دیفرانسیل و توپولوژی جبری باشد. همانطور که در کتاب هم آمده با استفاده از اندازه‌های متفاوت و همچنین تابعک‌ها و تورهای درهم‌تنیده؛ به طرزی اسرارآمیز توانستیم به متری از میزان محدب بودن فضا دست پیدا کنیم. مفهومی که در هندسه کلاسیک بررسی می‌شود اما بدون نیاز به یک ضرب‌داخلی و یا زاویه که در بحث آن در هندسه به نظر غیرممکن می‌آید. رازی که در سرتاسر این گزارش سعی شده تا با استفاده از تصاویر گوناگون پرده از آن برداشته‌شود.

درنهایت لازم می‌دانم از آقای منیری که به عنوان دستیار پژوهش راهنمایی‌های لازم را به من کردند تشکر کنم. همچنین تشکر ویژه‌ای از آقای مسیحا که بدون استفاده از شهود و تسلط بالای ایشان به موضوعات طرح شده؛ این پژوهش ممکن نبود. همچنین از استاد درس، جناب آقای یاسایی، که با حمایت‌های شان همزمان با همه‌گیری کرونا، کار را برای ما دانشجویان بسیار ساده‌تر کردند هم کمال تشکر را دارم.

سیدابوالفضل رحیمی

تابستان ۱۴۰۰

چکیده

برای بدست آوردن کرانی چفت بروی فرآیندهای زیرگاوسی نیاز به تئوری و ابزار پیشرفته‌تری حس می‌شد. برای برطرف کردن فاصله‌ی میان کران سوداکف و دادلی نیاز به معرفی ابزارهای جدیدتری بود. تئوری‌ای که با معرفی تابعک‌ها و تابع لاندای تالاگراند برطرف شده‌اند. روشی که به عنوان زنجیره‌سازی عمومی شناخته می‌شود. از سوی دیگر و با تعمیم‌های طبیعی این تئوری که برای مثال به درخت‌ها منجر می‌شود؛ می‌توان دسته مسائل بیشتری را حل نمود. اینجاست که این توابع و تعمیم‌ها بنظر همچون ابزارهای هندسه‌دیفرانسیل نوعی محاسبه از ویژگی‌های محلی فرآیندهای تصادفی روی فضای اصلی می‌دهند که در نهایت و با کران‌های بالا و پایین به یک ویژگی کلی منجر می‌شود. اندازه‌گیری‌هایی که بنظر قطره‌ای از یک تئوری بسیار گسترده ترند. این ویژگی‌ها نه تنها برخلاف هندسه دیفرانسیل به زاویه یا خمیدگی محلی مرتبط نیستند؛ بلکه به طول و فاصله صرف فضای متریک نیز متکی نیستند و همانطور که بحث خواهد شد متر و معیاری برای اندازه‌گیری بزرگی فضا و همچنین محدب بودن آن خواهند بود.

۱- فرآیندهای گاوسی[۱] و زنجیرسازی عمومی[۲]

در این گزارش فرآیندهای (X_t) همگی به نوعی دارای شرط افزایش[۳] ۱.۱ زیر هستند:

$$\forall u > 0, P(|X_s - X_t| \geq u) \leq 2e^{-\frac{u^2}{2d(s,t)^2}} \qquad (۱.۱)$$

که درآن فاصله d ذکر می‌شود. همچنین تمامی این فرآیندها میانگین صفرند؛ یعنی: $E[X_t]=0$. همانطور که خواننده از قبل آشناست؛ برای اندازه‌گیری بزرگی[۴] فرآیند به مقدار $\sup_{t\in T} X_t$ توجه می‌کنیم. در این بخش ابتدا به زنجیرسازی می‌پردازیم و

[۱] Gaussian Process
[۲] Generic chaining
[۳] Increment condition
[۴] size

سپس به ایده اصلی این بخش که همان زنجیرسازی عمومی است می‌پردازیم و در حین بررسی آن به تابع گاما تالاگراند[۵] و تورهای تودرتو[۶] و روش ساخت آن‌ها بوسیله تابعک‌ها[۷] می‌پردازیم. درنهایت هم به یک مثال پراهمیت برای ذکر تفاوت میان دو کران دادلی و عمومی، یعنی بیضیگون‌ها[۸] در یک فضای هیلبرت می‌پردازیم.

۱-۱- زنجیرسازی عمومی

باتوجه به متمرکز بودن[۹] فرآیندها، به راحتی می‌توان دید که برای تخمین $\sup_{t\in T} X_t$ می‌توان درحالت کلی به بررسی مقدار $\sup_{t\in T} X_t - X_{t_0}$. حال و برای انجام محاسبات بیشتر و ابتدا به لم ساده زیر اشاره می‌کنیم:

لم۱.۱.۱. اگر شرط $\forall u>0, P(Y\geq u)\leq Ae^{-\frac{u^2}{B^2}}$ برای یک متغیر تصادفی Y برقرار باشد، آنگاه همواره داریم:

$$E[Y]\leq LB\sqrt{logA}$$

با استفاده از کران اجتماع[۱۰] به همراه شرط (۱.۱) به راحتی داریم:

$$P(\sup(X_t - X_{t_0})\geq u)\leq 2card(T)\exp(-\frac{u^2}{2\Delta(T)^2})$$

پس با استفاده از لم ۱.۱.۱ و شرط بالا می‌توان به اولین تخمین از معیار مورد علاقه‌مان رسید:

$$E\sup_{t\in T} X_t \leq L\Delta(T)\sqrt{log(card(T))} \qquad (۱.۲)$$

حال به زنجیره‌سازی باز می‌گردیم. ایده اصلی کران زدن بروی کاردینال مجموعه T است. کاری که با تقسیم کل فضا به قسمت‌های کوچکتر صورت می‌گیرد.

$$X_t - X_{t_0} = (X_t - X_{\pi_1(t)}) + (X_{\pi_1(t)} - X_{\pi_2(t)}) + \ldots + (X_{\pi_{n-1}(t)} - X_{\pi_n(t)}) + X_{\pi_n(t)} - X_{t_0}$$

---

[۵] Talagrand

[۶] Admissible Sequences

[۷] Functionals

[۸] Ellipsoid

[۹] $EX_t=0$

[۱۰] Union bound

که در اینجا هرکدام از π_iها تخمین‌های هر مرتبه دقیق‌تری از t هستند که در یک مجموعه T_i حضور دارند. پس برای یافتن یک کران مناسب برای معیار اندازه‌مان نیازمندیم تا یک کران مناسب روی اندازه این مجموعه‌های T_i بگذاریم. کران مناسب ما اعداد $N_n = 2^{2^n}$ هستند. این اعداد بگونه‌ای اندازه مجموعه را که با $\sqrt{log(N_n)} = 2^{\frac{n}{2}}$اندازه‌گیری می‌شوند هر مرتبه دو برابر می‌کنند. البته شهود اندازه مجموعه هم که با رادیکال لگاریتم اندازه گرفتیم اینست که این تابع به گونه‌ای معکوس تابع $\exp(-x^2)$ است که در (۱.۱) آمده است. پس به گونه‌ای به کاردینال مجموعه دست پیدا کردیم. حال برای فاصله نیز در هر مرتبه فاصله نقطه t مورد نظر را با هر یک از تخمین‌هایش با کران فاصله t از کل مجموعه T در آن تخمین، کران می‌زنیم. پس ما توانستیم با استفاده از زنجیرسازی به کران زیر از مساله اصلیمان برسیم:

$$E \sup_{t\in T} X_t \leq L \sup_{t\in T} \sum_n 2^{\frac{n}{2}} d(t, T_n) \qquad (۱.۳)$$

حال دو رویکرد متفاوت که منجر به کران شل[۱۱] دادلی و کران چفت[۱۲] عمومی می‌شود را می‌توان در پیش گرفت. ابتدا اگر سوپریمم را داخل جمع برد و سپس با تعریف انتروپی به صورت: $\begin{array}{l} e_n(T) := \inf_{T_n} \sup_t d(t, T_n) \\ s.t.\ card(T_n) \leq N_n \end{array}$ به کران دادلی زیر می‌رسیم:

$$E \sup_t X_t \leq L \sum_n 2^{\frac{n}{2}} e_n(T) \qquad (۱.۴)$$

$$E \sup_t X_t \leq L \int_0^\infty \sqrt{log(N(T, d, \epsilon))} d\epsilon \qquad (۱.۵)$$

که البته با تعریف اعداد گنجایشی[۱۳] به کران معادل ۱.۵ که فرم انتگرالی کران دادلی است؛ می‌رسیم. درواقع آنچه که رخداده؛ به گونه‌ای تثبیت کردن یک فاصله ϵ مناسب است و سپس کاردینالیتی مجموعه با اعداد گنجایشی کنترل شده است. ترتیبی که در حرکت از ۱.۳ به ۱.۴ نیز با استفاده از بگونه‌ای کران جمع به جای سوپریمم، شل بودن آن دیده‌می‌شود. حال به صورتی دیگر به مساله نگاه می‌کنیم.

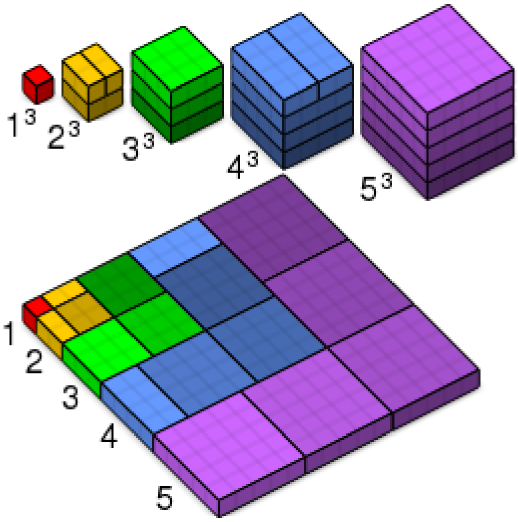
تعریف ۱.۶. منظور از یک **تور تودرتو** $(A_n)_{n=0}^\infty$ برای یک مجموعه T یک دنباله از مجموعه‌های تودرتو است که برای هر n داریم: $card(A_n) \leq N_n$. (تصویر ۱)

[۱۱] loose

[۱۲] sharp

[۱۳] Packing numbers

با استفاده از این تعریف، تعریف تابع گاما تالاگراند نیز به نظر طبیعی به نظر می‌رسد.



تصویر ۱. شمایی از یک تور تودرتو. هربار این تورها سعی می‌کنند به دقت بالاتری برسند و همچنین تعدادشان هم دوبرابر می‌شود. $N_n^2=N_{n+1}$

تعریف ۱.۷. برای یک فضای متریک (T, d) تعریف می‌کنیم:

$$\gamma_\alpha(T,d) := \inf_{A_n} \sup_t \sum_{n\geq 0} 2^{\frac{n}{\alpha}} \Delta(A_n(t))$$

در واقع این تعریف به طور طبیعی بزرگی فضای متریک را درنظر می‌گیرد. چون $A_0(t)=T$ پس همواره $\gamma_\alpha(T,d) \geq \Delta(T)$.

حال با ترکیب کران ۱.۳ و جایگزین کردن یک تور تودرتو به کران زنجیرسازی عمومی زیر می‌رسیم:

$$E \sup_t X_t \leq L\gamma_2(T,d) \qquad (۱.۸)$$

بدیهی است برای آنکه بتوانیم از این کران به خوبی استفاده کنیم؛ نیازمند بدست آوردن یک تور تودرتو هستیم. روشی عمومی برای بدست آوردن این تور استفاده از تابعک‌هاست که در بخش بعدی مورد بررسی قرار می‌گیرند.

پیش از ادامه دادن و بررسی تابعک‌ها، خوب است نامساوی‌هایی که مربوط به متغیرهای زیرنمایی و شبه برنشتاین و همچنین نامساوی‌های تمرکز اندازه حول این تابع گاما نیز بپردازیم.

قضیه ۱.۹. فرض کنید فضای متریک T با دو متر d_1 و d_2 داده شده است. و همچنین برای یک متغیر تصادفی X اندیس‌گزاری شده روی این فضا داشته باشیم:

$$\forall s,t \in T, \forall u \geq 0$$
$$P(|X_s - X_t| \geq u) \leq 2\exp\left(-\min\left(\frac{u^2}{d_2(s,t)^2}, \frac{u}{d_1(s,t)}\right)\right)$$

آنگاه به روشی کاملا مشابه حالت قبل خواهیم داشت:

$$E\sup_{t,s\in T} |X_s - X_t| \leq L(\gamma_2(T,d_2) + \gamma_1(T,d_1))$$

اثبات این قضیه چیزی نیست جز درنظر گرفتن یک تور تودرتو که با توجه به هر دو متر به اندازه کافی کوچک باشد.

حال و درنهایت خوب است به دو نامساوی تمرکز اندازه نیز با ابزار جدیدی که معرفی کردیم بپردازیم.

قضیه ۱.۱۰ برای یک متغیر تصادفی که در ۱.۱ صدق می‌کند؛ داریم:

$$P(\sup_{s,t\in T}|X_s - X_t| \geq L(\gamma_2(T,d) + u\Delta(T)) \leq L\exp(-u^2)$$

۱-۲- تابعک‌ها

همانطور که در بخش قبل بررسی شد؛ زنجیرسازی به طور طبیعی به تعاریفی منجر می‌شود که به کران‌هایی همچون ۱.۸ می‌رسیم. این کران‌ها اما، غیرقابل دستیابی‌اند مگر اینکه بتوان یکسری تور تودرتو برای یک فضای متریک ساخت. روشی که برای ساخت این تورها در نظر خواهیم گرفت؛ مبتنی بر مفهوم جدیدی به نام تابعک‌هاست که سعی می‌کنیم در این بخش با آن‌ها بیشتر آشنا شویم.

ابتدا برای چند پاراگراف مفهوم تورهای تودرتو را فراموش کنید. ابتدا به تعریف و ویژگی‌هایی از تابعک‌ها می‌پردازیم و سپس دوباره به آن‌ها باز خواهیم گشت.

تعریف ۱.۱۱ گوییم F یک تابعک روی فضای متری T است؛ هرگاه:

۱. F به هر زیرمجموعه H از T یک عدد مثبت F(H) نسبت دهد.

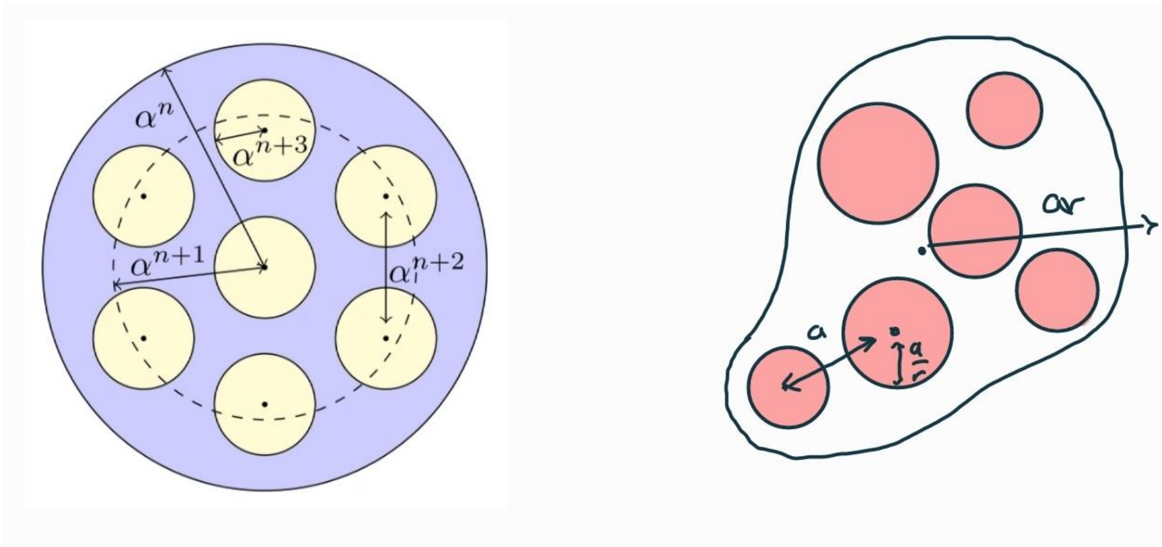
۲. F صعودی باشد. یعنی $H \subseteq H' \Rightarrow F(H) \leq F(H')$

درواقع می‌خواهیم هر تابعک به نوعی بزرگی یک مجموعه را اندازه‌گیری کند. البته ممکن است ذهن خواننده به توابع اندازه‌ای چون لبگ یا هآر[۱۴] سوق یابد؛ ولی بیشتر مفهوم بزرگی همان است که تا اینجا با آن سر و کار داشته‌ایم.

مثال ۱.۱۲ $F(H) = \gamma_\alpha(H,d)$ و یا $F(H) = \sup_{t\in H} X_t$ هردو مثال‌هایی برای یک تابعک هستند.

ویژگی دیگری که از یک تابعک انتظار داریم که اینست که بزرگی اجتماع مجموعه‌هایی که از یکدیگر دورند از بزرگی آن مجموعه‌ها به اندازه‌ای کافی بزرگ‌تر باشد.(تصویر ۲)

[۱۴] Haar measure



تصویر ۲. شمایی از جدابودن دو مجموعه(سمت راست) و تعمیم آن در کتاب ونرایت[۱۵](سمت چپ)

تعریف ۱.۱۳ فرض کنید عدد مثبت a و عدد صحیح $r > 3$ داده شده است. گوییم $H_1, H_2, ..., H_l \subseteq T$ یک مجموعه (a,r) جداست اگر: $\forall l \le m, H_l \subset B(t_l, \frac{a}{r})$ و همچنین نقاط $t_1, \ldots, t_l$ از یکدیگر به اندازه کافی دور باشند و برای یک s:

$$\forall l \le m, t_l \in B(s, ar);$$
$$\forall l, l' \le m, l \ne l' \Rightarrow d(t_l, t_{l'}) \ge a$$

حال به اندازه کافی بزرگ‌تر را تعریف می‌کنیم. درواقع این انتظار را در قالب شرط رشد[۱۶] بیان می‌کنیم. این شرط بیان می‌کند که بزرگی اجتماع حداقل به اندازه $\frac{a\sqrt{log(N)}}{c^*}$ از بزرگی کوچک‌ترین مجموعه، بزرگ‌تر است. البته تنها برای N هایی این شرط رشد برای ما مهم است که به اندازه کافی زیاد باشند. که همان $N = N_n$ ها برای ما کافی است.

برای خواننده‌ای که با توپولوژی آشناست؛ پس از در نظر گرفتن تورها[۱۷] به جای دنباله‌ها، طبیعی است که از دوگان آن‌ها که همان فیلترها[۱۸] هستند نیز استفاده کنیم. اینجاست که به جای درنظر گرفتن یک تابعک، دنباله‌ای نزولی از آن‌ها را تعریف می‌کنیم تا بتوان به مقدار کافی روی آن‌ها شرط گذاشت.

---

[۱۵] Ven wright

[۱۶] Growth condition

[۱۷] nets

[۱۸] filters

تعریف ۱.۱۴. گوییم دنباله (F_n) از تابعک‌ها در **شرط رشد با پارامترهای r, c*** صدق می‌کنند هرگاه برای هر عدد طبیعی n و هر a مثبت که $m = N_{n+1}$ باشیم: برای هر $H_1,\ldots,H_m$ که (a,r) جدا[۱۹] باشند؛ داشته باشیم:

$$F_n(\bigcup_{l\le m} H_l) \ge c^* a 2^{\frac{n}{2}} + \min_l F_{n+1}(H_l)$$

خوب است به محدودیت بزرگی که این شرط روی فضای متریک می‌گذارد توجه کنیم. درواقع این شرط به گونه‌ای بیان می‌کند که اگر $a > 2^{-n/2} F_0(T)/c^*$ باشد؛ آنگاه هر گوی $B(s,ar)$ را می‌توان با N_{n+1} گوی $B(t,a)$ پوشاند.

قضیه ۱.۱۵. هر فضای متریک دارای یک دنباله از تابعک‌هاست که $F_0 = \gamma_2(T,d)$ برای r=4 و c*=.5.

برهان. این قضیه اثبات آسانی دارد؛ که برای اثبات آن کافیست تابعک‌های زیر را درنظر بگیرید:

$$F_n(H) = \inf_A \sup_t \sum_{k\ge n} 2^{k/2} \Delta(A_k(t))$$

قضیه ۱.۱۶. تابعک‌های $F_n(H) = \gamma_2(H,d)$ شرط رشد را با پارامترهای r=8 و c*=.25. ارضا می‌کند.

برهان این قضیه نیز اثبات کلاسیک آنالیزی دارد و تنها به نتایج آن اینجا رضایت می‌دهیم.

تا اینجا فهمیدیم که توابع گاما تالاگراند می‌توانند به عنوان یک دنباله از تابعک‌ها روی یک فضای متریک تعریف شوند. حال و در قضیه اساسی زیر که برهان آن‌را در اینجا ذکر نمی‌کنیم؛ ثابت می‌شود که به‌گونه‌ای عکس قضیه ۱.۱۵ نیز درست است. به این معنا که کوچکترین تابعک با شرط رشد همان گاما تالاگراند است.

قضیه اساسی جدایی ۱.۱۷. فرض کنید در فضای متریک T یک دنباله از تابعک‌ها باشد که در شرط رشد صدق می‌کنند. در این صورت:

$$\gamma_2(T,d) \le \frac{Lr}{c^*} F_0(T) + Lr\Delta(T)$$

---

[۱۹] (a,r)-seperated

۱-۳- فرآیندهای گاوسی و فضای هیلبرت

در این بخش به فرآیندهای گاوسی و قدرتی که قضیه اساسی جدایی به ما داده می‌پردازیم.

قضیه ۱.۱۸. قضیه بزرگ‌کردن اندازه[۲۰]. برای فرآیند گاوسی X و فاصله $d(s,t)=(E(X_s-X_t)^2)^{1/2}$ **زنجیره سازی برای محاسبه بزرگی این فرآیند گاوسی کافی[۲۱] است.**

$$\frac{1}{L}\gamma_2(T,d)\le E\sup_t X_t \le L\gamma_2(T,d)$$

برهان. سمت راست این نامساوی این قضیه که بسادگی از ۱.۸ نتیجه می‌شود. برای سمت چپ اما، از قضیه اساسی جدایی و تابعک‌های زیر استفاده می‌کنیم:

$$F_n(H)=F(H)=\sup_{H^*\subset H, H^*\, finite} E\sup_{t\in H^*} X_t$$

اگر ثابت کنیم که شرط رشد را این تابعک‌ها ارضا می‌کنند و با گذاشتن کرانی بروی $\Delta(T)$ می‌توان قضیه را اثبات کرد.

اینجا طرحی از اثبات شرط رشد را بیان می‌کنیم.

لم ۱.۱۹. سوداکف[۲۲]. می‌دانیم که اگر نقاط $t_1,\ldots,t_m$ به گونه‌ای باشند که فاصله دوبه‌دوی آن‌ها از a بیشتر باشد آنگاه:

$$E\sup X_{t_p}\ge \frac{a}{L_1}\sqrt{\log m}$$

لم ۱.۲۰. تعمیم سوداکف. فرض کنید σ پارامتر زیرگواسی $\sup X_t$ باشد. H_1 را گوی حول t_1 در سوداکف به شعاع همین σ گیرید. حالا اگر $H=\cup H_l$ آنگاه داریم:

$$E\sup_H X_t \ge \frac{a}{L1}\sqrt{\log m}-L_2\sigma\sqrt{\log m}+\min_l \sup_{H_l} X_t$$

---

[۲۰] Majorizing measure theorem

[۲۱] Sufficient

[۲۲] Sudakov

حال با استفاده از این دو لم برهان ۱.۱۸ ساده است. تابعک‌های معرفی شده را بگیرید. این تابعک‌ها طبق ۱.۲۰ خاصیت رشد را با $r \geq 2L_1L_2, c^* = 1/L$ دارد. تنها برای استفاده از قضیه اساسی جدایی باید $\Delta(T)$ را کراندار کنیم که داریم:

$$E\max(X_t, X_s) = E\max(X_s - X_t, 0) = \frac{1}{\sqrt{2\pi}} d(s,t) \Rightarrow \Delta(T) \leq \sqrt{2\pi} E\sup_T X_t \bullet$$

حال و با استفاده از قضیه ۱.۱۸ می‌توان قضیه زیر را که به نوعی تعمیمی از اسلپیان[۲۳] است؛ ثابت کرد:

قضیه ۱.۲۱ فرض کنید که X_t ها گاوسی و Y_t ها هم خاصیت ۱.۱ را با متر $d(s,t) = E(X_s - X_t)^2$ ارضا کنند.

دراینصورت داریم: $E\sup_T |Y_s - Y_t| \leq LE\sup_T X_t$

برهان: $E\sup_T |Y_s - Y_t| \leq L\gamma_2(T,d) \leq L' E\sup_T X_t$

می‌دانیم که کران دادلی و سوداکف نمی‌توان بزرگی یا سایز فرآیند تصادفی را محاسبه کنند. به همین منظور چند مثال از چفت‌بودن قضایای بالا را می‌آوریم و سپس در بخش بعدی سعی می‌کنیم این حالت خاص را به بیضیگون‌ها در فضاهای هیلبرت ربط دهیم.

مثال ۱.۲۲ برای مجموعه اندیس‌گذار $\{1, 2, ..., N_s\}$ فرآیند $X_i = \frac{g_i}{\sqrt{logi}}$ را بگیرید که g گاوسی استاندارد است.

درینصورت داریم :

$$\sum_n e_n(T) 2^{n/2} \leq \frac{s-2}{L}$$
$$\sum_n d(p, T_n) 2^{n/2} \leq \sum_{n \leq m} L 2^{n/2} 2^{-m/2} \leq L$$

همانطور که دیده می‌شود در کران دادلی یک عبارت s که معادل $\sqrt{\log N}$ است اضافه شده است.

[۲۳] Slepian lemma

۱-۴- بیضیگون

این بخش ادامه مسیر قبل است تا بتوانیم در یک حالت کلی چفت نبودن کران دادلی را بررسی کنیم. آنچه اینجا بررسی می‌کنیم بیضیگون‌ها هستند که در عمل نیز در بسیاری از مسائل بدست می‌آیند.

منظور از یک بیضیگون در فضای هیلبرت زیر است و فرآیندهای زیر را در نظر می‌گیریم:

$$\epsilon = \{t \in l^2; \sum_{i\geq 1}^{\infty} \frac{t_i^2}{a_i^2} \leq 1\}$$
$$X_t = \sum_i t_i g_i$$

گزاره ۱.۲۳ داریم:

$$\frac{1}{L}(\sum_i a_i^2)^{1/2} \leq E\, sup_{\epsilon} X_t \leq (\sum_i a_i^2)^{1/2}$$

برهان. سمت راست به طور ساده از نامساوی کوشی شوارتز و نامساوی سمت چپ هم از ترکیب زیرگاوسی بودن با پارامتر همین جمع بدست می‌آید.

نتیجه ۱.۲۴ پس برای هر بیضیگون نتیجه گرفتیم: $\gamma_2(\epsilon) \leq L(\sum a_i^2)^{1/2}$

حال در ادامه سعی می‌کنیم اعداد انتروپی بیضیگون را حساب کنیم و نشان دهیم کران انتروپی دادلی برای محاسبه بزرگی بیضیگون مناسب نیست. بدین منظور فرض کنید ضرایب بیضیگون یعنی a_iها صعودی نیست. پس داریم:

$\frac{1}{2}\sum_{n\geq 1} 2^n\, a_{2^n}^2 \leq \sum a_i^2 \leq \sum_{n\geq 0} 2^n\, a_{2^n}^2$ پس در کل دراین حالت داریم:

$\frac{1}{L}(\sum_{n\geq 1} 2^n\, a_{2^n}^2)^{1/2} \leq E \sup_{\epsilon} X_t \leq (\sum_{n\geq 1} 2^n\, a_{2^n}^2)^{1/2}$ اما باتوجه به گزاره زیر می‌بینیم که دادلی اصلا چفت نیست:

گزاره ۱.۲۵ برای یک بیضیگون داریم:

$$\frac{1}{L}\sum 2^{n/2} a_{2^n} \leq \sum 2^n\, e_n(\epsilon) < L\sum 2^{n/2} a_{2^n}$$

۱-۵- جمع بندی

باتوجه به آنچه در اینجا نشان‌داده شد؛ بنظر می‌رسد برای یک فضای محدب احتمالا کران دادلی با یک نسبت رادیکال بعد هاوسدورف فضا از چفت بودن فاصله دارد. دلیل آن هم بالا رفتن اعداد انتروپی فضاست. در فصل بعدی اما سعی می‌کنیم بسیاری از فضاهایی که این دو کران یکسان عمل می‌کنند بپردازیم. یکی از سوالات مناسبی که پیش می‌آید رابطه تورهای تودرتو با محدب بودن است. البته باید مختصر درباره فضاهایی که زنجیرسازی عمومی در آن‌ها کران چفتی بدست نمی‌دهد هم سخن گفت.

۲- سری‌های تصادفی فوریه و مثلثاتی

هنگامی که فضا همگن است، کران دادلی چفت است! این نخستین گامیست که برمی‌داریم تا بیشتر با ابزاری که ساختیم آشنا شویم.

۲-۱- مترهای انتقال-ناوردا

فرض کنید فضای متریک T دارای این خاصیت است که T فشرده است و متر آن هم نسبت به انتقال[۲۴] ناوردا است. حال انداز هآر نرمال‌شده را روی این فضا در نظر بگیرید که یعنی: $\mu(T)=1$. حال قضیه زیر را در نظر بگیرید:

قضیه ۲.۱. یک متر انتقال-ناوردا پیوسته d را در نظر بگیرید. با تعریف زیر داریم:

$$\epsilon_n = \inf\{\epsilon \geq 0\,;\, \frac{\mu(T)=1}{\mu(B_d(0,\epsilon))} \leq N_n\}$$

$$\frac{1}{L}\sum \epsilon_n 2^{n/2} \leq \gamma_2(T,d) \leq L\sum \epsilon_n 2^{n/2}$$

[۲۴] Translation invariant

این قضیه درواقع براحتی با استفاده از لم زیر ثابت می‌شود و معادل بودن دو کران دادلی و زنجیرسازی عمومی را نشان می‌دهد.

لم ۲.۲. بین اعداد انتروپی و اپسیلون‌های تعریف بالا رابطه زیر درست است:

$$\epsilon_n \le e_n(T) \le 2\epsilon_n$$

پس با توجه به قضیه ۲.۱ به خاصیت دیگر بزرگی فضا پی بردیم. در ادامه به قضیه ساده ولی مهم زیر اشاره می‌کنیم که در برای قضیه پسیر[۲۵] اساسی است.

قضیه ۲.۳ فرض کنید خانواده ای از d_ω که مترهای مستقل از انتقال هستند در اختیار داریم. آنگاه با درنظر گرفتن $\Delta(T) = (E\Delta(T, d_\omega)^2)^{1/2}, d(s,t) = (Ed_\omega(s,t)^2)^{1/2}$ داریم:

$$(E\,\gamma_2(T, d_\omega)^2)^{1/2} \le L\,\gamma_2(T,d) + L\Delta(T)$$

به طور شهودی این قضیه بیان می‌کند که میانگین بزرگی تحت مترهای مختلف، حداکثر به اندازه بزرگی روی یک متر علاوه بر قطر فضاست.

۲-۲- قضیه مارکوس-پسیر[۲۶]

در این قسمت باتوجه به آنچه که در بخش قبل گفته شد؛ سعی می‌کنیم به صورت کوتاه به بررسی جمع‌های مثلثاتی بپردازیم و به یک نامساوی چفت برای جمع‌های گاوسی دست یابیم. بررسی این نکته که سمت راست تساوی‌های این بخش با سمت راست قضیه ۲.۳ بسیار یکسانند، سرنخی از روند اثبات این قضایا می‌دهد.

در این بخش متغیر تصادفی ξ_i متغیرهای مستقل و متقارن اند و χ_i هم یک گروه ضربی روی دایره واحد مختلط است. در این فصل تمام خانواده‌های اندیس‌گزار متناهی‌اند مگر آنکه خلاف آن گفته شود.

قضیه ۲.۴. با شرایط بالا و اگر ξ_i ممنت دوم[۲۷] هم داشته باشد آنگاه با تعریف زیر خواهیم داشت:

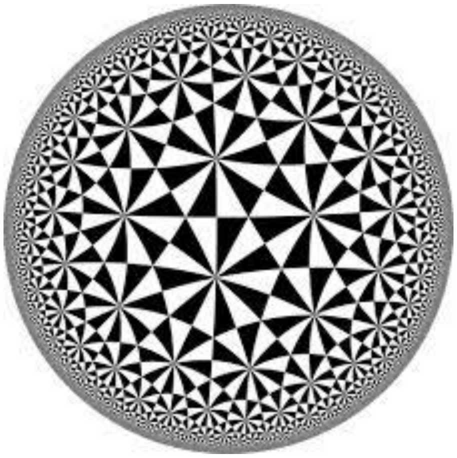
---

[۲۵] Pesier

[۲۶] Marcus-Pesier Theorem

$$d(s,t):=\sum E|\xi_i|^2|\chi_i(s)-\chi_i(t)|^2$$

$$(E\sup_T|\sum \xi_i a_i\chi_i(t)|^2)^{1/2}\le L(\gamma_2(T,d)+(\sum|a_i^2|E|\xi_i|^2)^{1/2s}$$



تصویر۳. از شماهایی که در آن کران زنجیرسازی عمومی از کران دادلی بهتر است. دلیل آن هم داشتن اندازه بیشتر(متراکم‌تر بودن) نقاط دورتر است.

در واقع قضیه بالا با در نظرگرفتن کسی‌ها به صورت گاوسی، قضیه محکمیست به اندازه قضایای فصل اول. ایده اصلی اثبات هم اضافه کردن متغیرهای برنولی تصادفی و همچنین استفاده از قضیه بنت برای ترکیب خطی برنولی هاست.

در ادامه و برای دستیابی به قضیه‌ای مشابه قضیه ۱۸.۱ یا بزرگ‌کردن اندازه بودن دو قضیه زیر نیز بررسی شده‌اند. تا بتوان جمع‌های مثلثاتی را نیز به راحتی بررسی کرد.

لم ۲.۵. $E\sup\sum a_i g_i\chi_i\ge\frac{1}{L}(\gamma_2(T,d)+\sum|a_i^2|^{1/2})$ این لم ساده با درنظر گرفتن بخش حقیقی و موهومی و با استفاده از قضیه۱۸.۱ یا بزرگ‌کردن اندازه بودن براحتی ثابت می‌شود.

لم ۲.۶. برای ϵ_i که برابر متغیر برنولی متقارن باشد؛ براحتی داریم:

$$E\sup\sum\xi_i x_i\ge E\sup\sum E|\xi_i|\epsilon_i x_i$$

درنتیجه برای متغیرها تصادفی گاوسی خواهیم داشت: $E\sup\epsilon_i x_i\le\sqrt{\frac{\pi}{2}}E\sup\sum g_i x_i$

قضیه مارکوس-پسیر زیر درنهایت ثابت می‌کند که در حالت خاصی که $x_i=a_i\chi_i$ باشند؛ نامساوی بالا برعکس هم می‌شود. دقت کنید که در فصل قبل دیدیم که سمت راست در حالت کلی به اندازه $\sqrt{\log m}$بیشتر از دست چپ است.

قضیه ۲.۷. مارکوس-پیسیر: برای متغیرهای تصادفی مستقل گاوسی g_i و متغیرها برنولی مستقل ϵ_i داریم:

$$E\sup\sum a_i g_i\chi_i\le LE\sup\sum a_i\epsilon_i\chi_i$$

پس در حالت کلی بزرگی جمع‌های مثلثاتی روی یک فضا را توانستیم با استفاده از زنجیره سازی بدست آوریم:

---

[۲۷] Second momentom

گزاره ۸.۲. با درنظر گرفتن دو متر به صورت:

$$d_1(s,t)=\sum |a_i|^2\,(E|\xi_i|)^2\,|\chi_i(s)-\chi_i(t)|$$
$$d_1(s,t)=\sum |a_i|^2\,(E\xi_i^{\,2})\,|\chi_i(s)-\chi_i(t)|$$

داریم:

$$\frac{1}{L}(\gamma_2(T,d_1)+(\sum |a_i|^2\,(E|\xi_i|)^2)^{1/2})\le E\sup\sum a_i\xi_i\chi_i\le L(\gamma_2(T,d_2)+(\sum |a_i|^2\,(E\xi_i^{\,2}))^{1/2})$$

۳- قضایای تطابق[۲۸]

بیضیگون بحث شده در فصل ۱.۴ را در نظر بگیرید. در این فصل سعی می‌کنیم نشان دهیم دلیل اصلی اینکه اعداد انتروپی تخمین بزرگ‌تر شده‌ای از بزرگی بیضیگون می‌دهند بدلیل محدب بودن بیضیگون است. سپس با استفاده از این قضیه به مسئله تطابق حمله می‌کنیم و نشان می‌دهیم بیضیگون برای جفت سازی کافی است.

۳-۱- قضیه بیضیگون

ابتدا به تعریف نرم محدب زیر توجه کنید.

تعریف ۱.۳ برای p>2 به یک نرم||.|| pمحدب[29] گوییم اگر برای یک η مثبت داشته باشیم:

$$\|x\|,\|y\|\le 1\Rightarrow \left\|\frac{x+y}{2}\right\|\le 1-\eta\|x-y\|^p$$

با توجه به این تعریف بررسی اینکه نرم زیر روی بیضیگون ۲محدب است با پارامتر ۰.۱۲۵.

---

[۲۸] Matching theorems

[۲۹] P-convex

$$\| x \|_{\epsilon} = (\sum \frac{t_i^2}{a_i^2})^{1/2}$$

در ادامه برای نزدیک‌شدن به مفهوم جفت سازی و همچنین ارتباط میان محدب بودن و بزرگی تعریف ۲.۳ زیر را در نظر بگیرید.

$$\gamma_{\alpha,\beta}(T,d) := (\inf \sup_t \sum (2^{n/\alpha} \Delta(A_n(t),d))^{\beta})^{\frac{1}{\beta}}$$

درواقع اینجا برای متغیرهایی است که با $\exp(-u^{\alpha})$ کم می‌شوند. این تعریف همچنین بگونه‌ای یک میزان اهمیت به بزرگی و کوچکی می‌دهد که با پارامتر β بالا کنترل می‌شود. با استفاده از این تعریف همان لگاریتم قبلی ظاهر خواهد شد!

لم ۳.۳ اگر فضای متریک T دارای حداکثر N_m عنصر باشد؛ آنگاه $\gamma_2(T,d) \le \sqrt{\log N_m}\gamma_{2,2}(T,d)$. برهان این لم براحتی و با استفاده از نامساوی کوشی–شوارتز بدست می‌آید اما لم قضیه زیر که صورت معادل قضیه بیضیگون است را نتیجه می‌دهد.

قضیه ۴.۳ برای هر فضای متریک داریم:

$$\gamma_{\alpha,p}(T) \le K(\alpha,p,\eta) \sup_n 2^{n/\alpha} e_n(T)$$
$$\gamma_{\alpha,p}(T) \le K(\alpha)(\sum (2^{n/\alpha} e_n(T))^p)^{1/p}$$
$$\sup_n 2^{n/\alpha} e_n(T) \le K(\alpha)\gamma_{\alpha,p}(T)$$

و قضیه بیضیگون زیر

قضیه ۵.۳. بیضیگون. برای هر α داریم:

$$\gamma_{\alpha,2}(\epsilon) \le K(\alpha) \sup_{\xi} \xi \, (card\{i; a_i \ge \xi\})^{1/\alpha}$$

مثال. فرض کنید F کلاس همه توابع از $[0,1] \to \mathbb{R}$ باشد که $f(0) = f(1) = 0$ و این توابع مشتق پذیر و $\sup |f'| \le 1$.

آنگاه
$$\gamma_{1,2}(F,d) = L$$
$$d(f,g) = (\int_{[0,1]} (f-g)^2)^{1/2}$$

۳-۲- تطابق

ستون فقرات این قسمت مسئله جذاب زیر است:

با دانستن اینکه نقاط Y_i به طور یکنواختی روی مربع واحد هستند؛ نقاط X_i که به صورت تصادفی انتخاب می‌شوند چقدر از این نقاط دورند. یا به عبارتی دیگر هزینه انتقال این نقاط به نقاط اولیه یکنواخت چقدر است؟

این هزینه را عموما به عنوان فاصله درنظر می‌گیرند. هزینه‌ای که ممکن است جمع فواصل باشد یا کمترین فاصله یا بیشترین فاصله.

ابتدا به بررسی نقاط Y_iها به صورت یکنواخت و غیر تصادفی روی مربع واحدند را بررسی می‌کنیم. به طور طبیعی نقاط به صورت $(k/\sqrt{(N)}, l/\sqrt{N})$ را می‌توان به جای این نقاط متصور بود. البته می‌توان تصور کرد که به تعداد N مستطیل هم مساحت به مساحت $1/N$ داریم که هر کدام از دیگری جدا بوده و شامل یکی از نقاط ایگرگ هستند.

مهمترین ابزار برای حمله به مسئله تطابق لم زیر است؛

لم ۶.۳. ماتریس $C=(c_{ij})_{i,j\le N}$ را در نظر بگیرید. حال تعریف کنید:

$$M(C)=\inf\sum_i c_{i\pi(i)}$$

حال تمام دنباله‌های $\omega_i, \omega_{i'}$ را بگیرید که در آن $\omega_i+\omega_j{}'\le c_{ij}$ در اینصورت:

$$M(C)=\sup\sum \omega_i+\omega_i{}'$$

اثبات این لم چیزی جز محاسبات جبری ندارد؛ اما مهمترین قسمت یافتن هزینه تطابق با کمک یک سری دنباله کمکی ωهاست.

حال با کمک این قضیه می‌توان به راحتی گزاره زیر را اثبات کرد.

گزاره ۷.۳. نقاط Y_i, X_i را درنظر بگیرید و همچنین F را خانواده توابع ۱لیپ‌شیتز گیرید؛ در اینصورت داریم:

$$\inf_\pi \sum d(X_i, Y_{\pi(i)}) = \sup_f \sum f(X_i) - f(Y_i)$$

برهان. داریم:

$$\sum f(X_i) - f(Y_i) = \sum f(X_i) - f(Y_{\pi(i)}) \le \sum d(X_i, Y_{\pi(i)})$$
$$\inf_\pi \sum d(X_i, Y_{\pi(i)}) = sup \sum \omega_i + \omega_i'$$
$$now\ consider\ f: f(x) = \min(-\omega_j' + d(x, Y_j)) \Rightarrow \sum \omega_i + \omega_i' \le \sum f(X_i) - f(Y_i)$$

حال بیایید به مساله اصلی که همان تطابق است بازگردیم. این مساله را می‌توان با یافتن کمینه برای کمیت $\sum \psi(X_i, Y_{\pi(i)})$ معادل دانست. این تابع به نوعی میزان فاصله را اندازه‌گیری می‌کند؛ اما یک متر نیست و لازم نیست هیچ‌گونه چیزی شبیه نامساوی مثلث را ارضا کند. البته طبیعی است که $\psi(x,x) = 0$ را داشته باشیم. حال و مشابه گزاره ۷.۳ می‌توان به نتیجه مشابه زیر رسید:

$$\inf \sum \psi(X_i, Y_{\pi(i)}) \le \sup \sum (f(X_i) + \omega_i')$$
$$where\ f(x) := min_j(-\omega_j' + \psi(x, Y_j))$$

که در ابتدایی سوپریمم روی تمام خانواده امگاست. از آنجایی که نقاط ابتدایی به طور یکنواخت و غیر تصادفی در فضا پراکنده‌اند می‌توان فرض کرد که $N\int hd\lambda = \sum h(Y_i)$ پس می‌توان دست راست بالا را به صورت زیر نوشت:

$$\sum (f(X_i) - \int fd\mu) + \sum (\omega_i' + f(Y_i))$$

حال اما باتوجه به نوع تعریف تابع f می‌دانیم که قسمت آخر منفی‌است. حال اگر کلاس کلاس توابع را بگونه‌ای درنظر بگیریم که قسمت دوم جمع‌بالا حداقل یک مقدار $-A$ بشود(که در بیشتر اوقات این‌چنین است؛ همانند آنچه در مثال بالا دیدیم.) می‌توان یافتن یک کران برای تطابق را به یافتن کران برای فاصله تجربی میانگین تابع f از مقدار واقعی توصیف کرد. یعنی مسأله به یافتن سوپریمم موجودات زیر تقلیل یافت:

$$Z_f = \sum_{i \le N} (f(X_i) - \int fd\mu)$$
$$\sup_{f \in F} Z_f$$

۳-۳- قضیه تطابق آجتای[۳۰]

اینجا و برای مثال به بررسی یک تطابق خاص درحالتی که تابع هزینه فاصله اقلیدسی است؛ می‌پردازیم. قضیه‌ای که همراه با روشی که در فصل بعد از آن کمک می‌گیریم به کرانی بسنده برای این حالت خاص تطابق می‌انجامد.

قضیه ۸.۳. برای فاصله اقلیدسی و با توجه به تعاریف پیش از این داریم:

$$E\inf_{\pi}\sum d(X_i,Y_{\pi(i)})\le L\sqrt{N\log N}$$

قضیه‌ای که به همراه روش فصل بعد دوطرفه می‌شود؛ یعنی:

$$\frac{1}{L}\sqrt{N\log N}\le E\inf_{\pi}\sum d(X_i,Y_{\pi(i)})\le L\sqrt{N\log N}$$

طبق آنچه که در فصل قبل به آن اشاره کردیم؛ دیدیم که کران مناسب برای این تطابق به یافتن کرانی مناسب برای توزیع تجربی می‌انجامد. البته پیش از این نیز قضیه زیر درمورد توابع ۱لیپ‌شیتز را فرض می‌کنیم.

قضیه ۹.۳. برای کلاس توابع ۱لیپ‌شیتز می‌دانیم:

$$E\sup_f|\sum_i(f(X_i)-\int f\,d\lambda)|\le L\sqrt{N\log N}$$

برهان قضیه ۸.۳. ابتدا با توجه به گزاره ۷.۳ داشتیم:

$$\inf_{\pi}\sum d(X_i,Y_{\pi(i)})=\sup_f\sum f(X_i)-f(Y_i)$$
$$\sum f(X_i)-f(Y_i)\le|\sum(f(X_i)-\int f\,d\lambda)|+|\sum(f(Y_i)-\int f\,d\lambda)|$$

اولین مولفه جمع در بالا که با توجه به قضیه۹.۳ از آنچه که خواسته شده کمتر است. تنها باید روی مقدار دوم یک کران مناسب پیدا کنیم که برای آن هم داریم:

$$|\sum(f(Y_i)-\int f\,d\lambda)|\le\sum_i|f(Y_i)-N\int_{R_i}f\,d\lambda|$$

---

[۳۰] Ajtai matching theorem

که در آن مستطیلی است که Y_i در آن قرار دارد. حال می‌دانیم که $Y_i \in R_i \subset B(Y_i, \frac{10}{\sqrt{N}})$ پس چون f ۱لیپ‌شیتز است؛ پس مقدار آن حداکثر به همین مقدار از $f(Y_i)$ فاصله دارد. پس این مقدار حداکثر $L\sqrt{N}$ است. پس در نهایت قضیه ۸.۳ ثابت شد.

خوب است کمی به برهان قضیه ۹.۳ نیز توجه کنیم. در این برهان تنها از قضیه های برنشتاین[۳۱] برای توابع ۱لیپ‌شیتز و فرم برنشتاین فصل یک نیاز است که استفاده شود. با استفاده از این دو به فرم زیر می‌رسیم و سپس با روشی جبری می‌توان توابع گاما تالاگراند سمت راست نامساوی ۱۰.۳ را کنترل کرد. البته که کنترل نرم ۲ بسیار سخت‌تر است.

گزاره ۱۰.۳ برای خانواده توابع ۱لیپ‌شیتز که دارای تابع تباهیده[۳۲] هستند داریم:

$$E \sup_f |\sum f(X_i) - \int f \, d\lambda| \leq L(\sqrt{N}\gamma_2(F, d_2) + \gamma_1(F, d_\infty))$$

۴- هنر کران‌های پایین و درخت‌ها

در ادامه به ذکر تعمیمی زیبا از ابزارهای قبلی می‌پردازیم. این‌بار اما، هدف به‌دست آوردن کران پایین برای بزرگی یک فرآیند تصادفی است. یک شروع خوب برای بدست آوردن یک کران پایین مناسب، پیدا کردن ساختاری است که به اندازه کافی در کل فضا کشیده شده‌باشد؛ از سوی دیگر این ساختار باید کوچک هم باشد. اگر ساختار فضا را به صورت یک گراف ببینیم؛ همانطور که خواننده آشنا با نظریه گراف‌ها بنظرش می‌رسد؛ در نظر گرفتن یک ساختار بلوکی بسیار برای این منظور مناسب است. البته می‌دانیم این بلوک‌ها یک درخت تشکیل می‌دهند. حال در این بخش هم ما درنظر داریم تا این شهود را دقیق کنیم تا به ابزاری برای بدست‌آوردن کران پایین برای بزرگی یک فضا دست پیدا کنیم.

۴-۱- درخت‌ها

[۳۱] Bernstein inequality

[۳۲] 0 is in F

در ادامه توضیحات قسمت قبل، ابتدایی ترین ویژگی یک درخت نداشتن دور یا یال اضافی است. با درنظر گرفتن گوی‌ها به عنوان راس‌ها، بنظر می‌آید که برای حذف یال‌های اضافی باید مجموعه‌ها یا اشتراک نداشته‌باشند؛ یا زیرمجموعه هم باشند. برای پوشا بودن هم راس ریشه را می‌توان کل فضا گرفت. پس به تعریف زیر رسیدیم:

تعریف ۴.۱ منظور از درخت τ مجموعه از زیرمجموعه‌های فضای T است به گونه‌ای که:

۱. $\forall A, B \in \tau, A \cap B \neq \varnothing \to A \subset B, or, B \subset A$

۲. درخت یک عضو ماکسیمال دارد.

با درنظر گرفتن این تعریف تعداد فرزندان c(A) یک مجموعه A در یک درخت را می‌توان تعدادB هایی گرفت که داریم:

$$\forall C : B \subset C \subset A \to A = C \vee B = C$$

این شرط اضافی را که هر عضوی که C آن صفر بود آنگاه تک نقطه است را هم برای اینکه به اندازه کافی درخت گسترده باشد اضافه می‌کنیم.

حال همانطور که در فصل یک دیدیم؛ نیاز به یک شرط همچون شرط رشد برای تخمین مناسب داریم. اینبار این شرط را بگونه‌ای دیگر مطرح می‌کنیم. برای پیدا کردن یک مفهوم مشابه ابتدا به تعریف خوشه نیاز داریم.

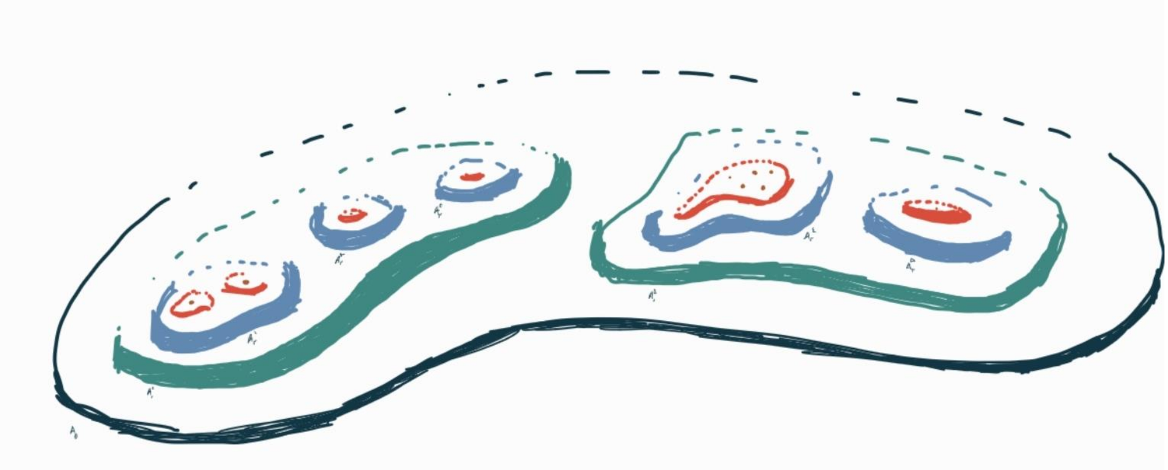
تعریف ۲.۴. منظور از یک خوشه در یک درخت دنباله $A_0, A_1, ..., A_k$ است که هر دو عضو متوالی پدر و فرزندند و اولین عضو نیز کل درخت و عضو انتهایی آن هم یک تک عضو است. پس می‌توان به هر عضو که در یکی از مجموعه‌های درخت حضور دارد یک و دقیقا یک خوشه[۳۳] متناظر کرد. حال تکیه‌گاه هر درخت را مجموعه کل این اعضا تعریف می‌کنیم و با S_τ نشان می‌دهیم.

[۳۳] branch

حال مفهوم دور بودن و یا پراکنده بودن درخت را با معرفی عدد تمرکز که با $s(A)$ که یک عدد صحیح است نشان می‌دهیم. انتظار داریم تمرکز[۳۴] پدر کمتر از فرزند باشد و همچنین مانند مفهوم (a,r)جدا که در فصل یک معرفی شد، خاصیت زیر را برای هر A که c(A) ناصفر است داشته باشیم:

$$B_1, B_2 \subset A \Rightarrow d(B_1, B_2) \geq 4^{-s(A)}$$

حال و با این تعاریف، آماده‌ایم تا مفهوم بزرگی را برای یک درخت معرفی کنیم.



تصویر ۴. یک مجموعه و که توسط یک درخت تقسیم بندی شده‌است. هر کدام از مجموعه‌های همرنگ فرزندان یک پدر هستند و همچنین خوشه‌های متفاوت آن به مجموعه‌های تک عضوی منتهی شده است. یک خوشه را می‌توان به صورت مجموعه‌هایی که ناهمرنگ‌اند در بالا دید.

تعریف۴.۳. منظور از عمق[۳۵] یک درخت عدد زیر است:

$$d(\tau) := \inf_t \sum_{t \in A} 4^{-s(A)} \sqrt{\log c(A)}$$

اگر به شباهت این تعریف با گاما دقت کنید درمی‌یابید که بگونه‌ای یک درخت به اندازه بزرگ‌ترین خوشه‌اش بزرگ است و انتظار داریم این مقدار از بزرگی فضای کلی که درخت در آن نشسته کمتر باشد. پس حال به مقدار زیر توجه کنید:

---

[۳۴] separation

[۳۵] depth

$$D(T)=\sup\{d(\tau);\ \tau \text{ is a seperated tree of } T\}$$

البته می‌توان مفهوم تمرکز را با تعریف زیر که به نوعی تعمیم آن است دقیق‌تر کرد. اینبار بجای $s(A)$ به هر کدام از مجموعه‌ها عدد $j(A)$ نسبت می‌دهیم و به آن عدد مرتب‌بودن مجموعه می‌گوییم.

تعریف ۴.۴. یک درخت را یک درخت مرتب[۳۶] گوییم هرگاه برای هر $A\in\tau$ که $c(A)\geq 1$ بتوان یک عدد صحیح $j(A)\in\mathbb{Z}$ و یک نقطه t و نقاط $t_1,\dots,t_{c(A)}$ نسبت داد که:

$$t_i\in B(t,4^{-j(A)})$$
$$\forall i,j: d(t_i,t_j)\geq 4^{-j-1}$$

و همچنین هر کدام از گوی های $B(t_i,4^{-j-2})$ شامل دقیقا یک فرزند A باشند. و برای این درخت عمق را به طور مشابه تعریف کرده:

$$d'(\tau)=inf_t\sum_{t\in A}4^{-j(A)}\sqrt{\log c(A)}$$
$$D'(T)=\sup_\tau d'(\tau)$$

با توجه به تعاریف بالا چندگزاره بدیهی را می‌توان براحتی محاسبه کرد.

گزاره ۵.۴. داریم:

$$\gamma_2(T)\leq LD'(T)$$
$$D'(T)\leq 16D(T)$$

پس با استفاده از این مفهوم جدید توانستیم دوباره تخمینی در همان اردر بزرگی قبلی بیابیم.

باتوجه به همین تعاریف و ایده محاسبه خوشه‌های متوالی از یک درخت می‌توان کران‌های پایینی برای مسائلی که پیش از این هم به آنها پرداخته شده بود پیدا کرد. غیر از مسئله فرآیند گاوسی، برای مسئله تطابق که در بخش قبل آن‌را بررسی کردیم هم می‌توان یک کران پایین مناسب یافت. کرانی که چفت بودن قضیه بخش قبل را هم اثبات می‌کند و ما تنها به ذکر آن بسنده می‌کنیم.

---

[۳۶] regular

قضیه ۸.۴. برای کلاس توابع ۱لیپ‌شیتز روی مربع واحد داریم:

$$E\sup_f |\sum_{i\le N}(f(X_i)-\int f d\lambda)|\ge \frac{1}{L}\sqrt{N\, logN}$$

حال یک اندازه احتمال μ روی فضای متریک با تکیه‌گاه شمارا بگیرید. می‌خواهیم نشان دهیم که کمیت زیر هم به نوعی بزرگی اندازه فضا را اندازه می‌گیرد.

$$I_\mu(t):=\int_0^\infty \sqrt{\log\frac{1}{\mu(B(t,\epsilon)}}d\epsilon$$

قضیه ۶.۴. با شروط مشابه قبل داریم:

$$\sup_t I_\mu(t)\le L\gamma_2(T,d)$$
$$d(\tau)\le L\sup_t I_\mu(t)$$

و به اندازه احتمالی که سمت چپ نامساوی قضیه ۶.۴ را کمینه می‌کند، اندازه احتمال مهم می‌گویند.

پس توانستیم به چهار کمیت از بزرگی فضاها با استفاده از تعاریف این بخش برسیم که در گزاره زیر گفته شده‌اند.

گزاره ۷.۴ با استفاده از مطالب این بخش چهار کمیت زیر از بزرگی یک فضای متریک معادلند.

$$D(\tau), D'(\tau), sup_t I_\mu(t), \gamma_2(T)$$

۵- جمع بندی

در این پژوهش سعی کردیم تا با زنجیرسازی عمومی به روشی مناسب برای بدست‌آوردن کران‌های بالا دست پیدا کنیم. روش هایی که نتایج آن‌را می‌توان در بخش‌های ۱ تا ۳ دید. یک موضوع مناسب برای ادامه این مطالب، پژوهش درباره زمانی‌است که کران زنجیرسازی عمومی هم چفت نیست.

در لابه‌لای مسائل و همچنین در فصل ۴ام نیز به بررسی روشی که منجر به به دست آوردن کران پایین برای یک فرآیند می‌شود پرداختیم و دیدیم که با معیارهای متفاوتی می‌توان بزرگی را تخمین زد. از مسیر پیش‌رو نیز می‌توان به لزوم تهیه و تدوین یک تئوری همه جانبه گفت که طی آن همه معیارهای مطروحه در فصل ۴ام را بتوان بدست آورد. این معیار

می‌تواند پلی باشد میان هندسه و آنالیز حقیقی زیرا با این تئوری همه‌جانبه قادر خواهیم بود مفهومی که از ترکیبیات و توپولوژی تا اندازه به آن دست یافتیم را زیر یک پرچم متحد کند.

مراجع

1. *Upper and Lower Bounds for Stochastic Processes, by Michel Talagrand, first Edition.*
2. *High-Dimensional Probability, An Introduction with Applications in Data Science, by Roman Vershynin.*