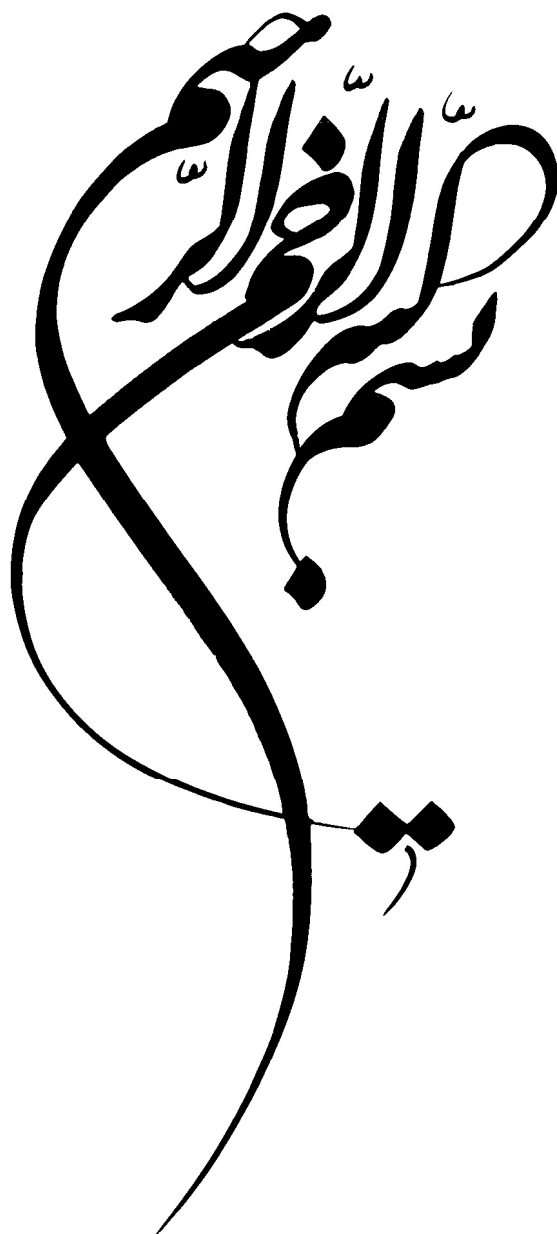






دانشگاه صنعتی شریف

دانشکده علوم ریاضی

پایان نامه کارشناسی ارشد

گرایش ریاضی محض

عنوان

انتخاب نقاط تصادفی از اشکال محدب در ابعاد بالا

نگارش

علی خزلی

استاد راهنما

دکتر کسری علیشاهی

ماه و سال

شهریور ۸۹

تصویب نامه

به نام خدا

دانشگاه صنعتی شریف

دانشکده علوم ریاضی

رساله کارشناسی ارشد

عنوان: انتخاب نقاط تصادفی از اشکال محدب در ابعاد بالا

نگارش: علی خزلی

کمیته ممتحنین:

امضاء استاد راهنما: دکتر کسری علیشاهی

امضاء ممتحن داخلی: دکتر سید رضا مقدسی

امضاء ممتحن خارجی: دکتر علی فروش باستانی

تاریخ

تقدیر

لازم است از استاد بزرگوارم دکتر کسری علیشاهی تشکر کنم که با صبر فراوان مرا در این پایان نامه راهنمایی کردند. همچنین از اساتید گرامی دکتر سید رضا مقدسی و دکتر علی فروش باستانی تشکر می کنم که زحمت داوری این پایان نامه را به عهده گرفتند.

چکیده

«چگونه می توانیم یک نقطه تصادفی با توزیع یک نواخت در یک شکل محدب $K \subseteq \mathbb{R}^n$ انتخاب

کنیم؟»

با توجه به کاربرد های این مسئله، مهم است که یک پاسخ برای این مسئله در ابعاد بالا هم کارا باشد. در این جا به الگوریتم هایی علاقه مندیم که مرتبه آن ها بر حسب بعد، چند جمله ای باشد. روش هایی که برای این مسئله در ابعاد بالا معرفی شده اند، همگی بر اساس روش مونت کارلو هستند. یعنی یک قدم زدن تصادفی در K ساخته می شود که توزیع پایای آن همان توزیع یک نواخت روی K باشد. در این صورت با شبیه سازی گام های این قدم زدن تصادفی «به تعداد کافی»، توزیع حالت نهایی «تقریباً» برابر با توزیع مطلوب است. مسئله مهم در روش مونت کارلو، تحلیل «زمان آمیختگی» یک قدم زدن تصادفی است؛ یعنی چند گام از آن باید انجام شود تا توزیع نهایی به مقدار قابل قبولی به توزیع پایا نزدیک شود.

در قسمت عمده این پایان نامه به تحلیل زمان آمیختگی قدم زدن تصادفی Hit-and-run - که سریع ترین قدم زدن تصادفی معرفی شده تا کنون است - می پردازیم. علاوه بر آن، در مورد یکی از مهم ترین کاربرد های این مسئله - که باعث گسترش بررسی آن شده است - توضیح می دهیم: تخمین حجم اشکال محدب. این موضوع از آن جا جلب توجه می کند که ثابت می شود هیچ الگوریتم غیر تصادفی نمی تواند در زمان چند جمله ای حجم اشکال محدب را با خطای نسبی محدود محاسبه کند.

کلمات کلیدی: اشکال محدب، تخمین حجم، فرآیند مارکوف، الگوریتم تصادفی، زمان آمیختگی،

رسانش، فاصله مجموع تغییرات

فهرست

تقدیر	۵
چکیده	۵
فهرست	۵
فهرست علایم و اختصارات	ح
۱. مقدمه	۱
۱.۱. مسئله محاسبه حجم	۱
۲.۱. تولید نقطه تصادفی	۵
۲. مدل محاسباتی	۱۱
۱.۲. اواراکل ها	۱۱
۲.۲. اجرای قدم زدن تصادفی Hit-and-run	۱۳
۳.۲. الگوریتم های تصادفی و غیر تصادفی	۱۵
۳.۲.۱. نتایج منفی برای الگوریتم های غیر تصادفی	۱۵
۳.۲.۲. الگوریتم های تصادفی	۱۹
۳. هسته های مارکوف و فرآیندهای مارکوف	۲۱
۱.۳. تعاریف اولیه	۲۱
۲.۳. آمیختگی در فرآیندهای مارکوف	۲۶
۲.۳.۱. فاصله مجموع تغییرات	۲۶
۲.۳.۲. جفت سازی و فاصله مجموع تغییرات	۳۲
۲.۳.۳. آمیختگی در فرآیندهای مارکوف	۳۴
۲.۳.۴. رسانش زنجیرهای مارکوف	۳۵

۴۳	۴. تولید یک نقطه تصادفی در یک شکل محدب.....
۴۳	۱.۴. مراحل تحلیل زمان آمیختگی قدم زدن تصادفی Hit-and-run.....
۴۵	۲.۴. قضایای اصلی.....
۴۸	۳.۴. فاصله هندسی و فاصله احتمالاتی.....
۵۷	۴.۴. یک نابرابری محیطی.....
۶۳	۵.۴. اندازه گام.....
۶۹	۶.۴. کران پایین برای رسانش.....
۷۲	۷.۴. کران برای زمان آمیختگی.....
۷۴	۸.۴. تولید یک دنباله از نقاط تصادفی.....
۸۱	۵. الگوریتم حجم.....
۹۱	۶. کارهای دیگر.....
۹۱	۱.۶. دقیق بودن مرتبه کران.....
۹۴	۲.۶. گرد کردن شکل.....
۹۸	۳.۶. نمونه گیری از توزیع های لوگ-مقعر.....
۱۰۰	۴.۶. روش Simulated Annealing.....
۱۰۳	ضمیمه أ. لم موضعی سازی.....
۱۱۷	ضمیمه ب. مقدمات هندسی.....
۱۱۷	۱. قضیه ای از هندسه مسطحه.....
۱۱۷	۲. گوی واحد.....
۱۲۰	۳. قضیه برون-مینکوفسکی.....
۱۲۹	مراجع.....
۱۳۱	فهرست کلمات.....
۱۳۲	واژه نامه (فارسی-انگلیسی).....
۱۳۴	واژه نامه (انگلیسی-فارسی).....

فهرست علائم و اختصارات

مجموعه اعداد حقیقی نامنفی	$\mathbb{R}^{\geq 0}$
حجم مجموعه A	$vol(A)$
حجم d بعدی مجموعه A	$vol_d(A)$
گوی به شعاع r حول x	$\mathbb{B}_r(x)$
گوی واحد n بعدی	$\mathbb{B}^n$
کره واحد n بعدی؛ یا همان $\partial\mathbb{B}^{n+1}$	$\mathbb{S}^n$
حجم گوی واحد n بعدی	π_n
جمع مینکوفسکی دو مجموعه A و B	$A + B$
مجموعه $\{tu : t \in \mathbb{R}\}$	$\mathbb{R} \cdot u$
مرز مجموعه A	∂A
معیاری برای فاصله دو نقطه p و q در شکل محدب K (ص ۴۸)	$d_K(p, q)$
توزیع احتمال یک نواخت روی K	$\mathbb{P}_K$
متوسط تابع f روی K	$\mathbb{E}_K(f)$
هشتک طول گام قدم زدن تصادفی Hit-and-run (ص ۴۹)	$F(x)$
معیاری برای فاصله نقطه x از مرز شکل محدب K (ص ۶۳)	$s(x)$

متقارن سازی اشتاینر	S_θ
بیش ترین فاصلهٔ نقاط مجموعهٔ C از مبدأ مختصات	$d(C)$
فاصلهٔ مجموع تغییرات دو توزیع μ و ν	$\ \mu - \nu\ _{TV}$
تابع چگالی توزیع μ نسبت به ν	$d\mu/d\nu$
تابع مشخصهٔ مجموعهٔ A	χ_A
تحدید توزیع μ به مجموعهٔ A	$\mu _A$
رسانش یک فرآیند مارکوف بین دو مجموعهٔ A و B	$\Phi(A, B)$
رسانش یک فرآیند مارکوف	Φ
نماد مرتبه که در آن عوامل $\log^c n$ و $\log R/r$ و عواملی که بر حسب	O^*
دقت هستند حذف شده‌اند (ص ۱۹)	
نماد مرتبهٔ دقیق	Θ

۱. مقدمه

۱.۱. مسئله محاسبه حجم

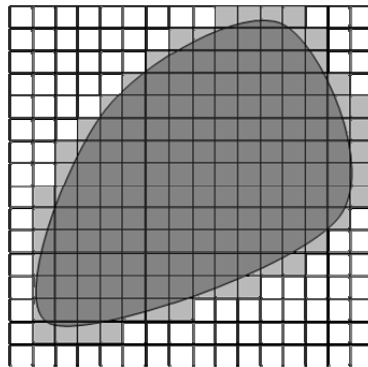
چگونه می‌توانیم حجم یک مجموعه محدب داده‌شده $K \subseteq \mathbb{R}^n$ را محاسبه کنیم؟

پاسخ این مسئله در بعد سه خیلی سخت نیست. اگر فرض کنیم K در یک مکعب مستطیل

$\prod_{i=1}^3 [a_i, b_i]$ قرار دارد، کافی است بازه‌ها را به تعداد زیادی زیربازه تقسیم کنیم و ببینیم چه تعداد از

مکعب‌های مستطیل کوچک ایجاد شده با K اشتراک دارند. به این ترتیب با هر دقت نسبی دلخواه

$\epsilon > 0$ می‌توان حجم K را محاسبه کرد.



شکل ۱

اگر چند جمله‌ای بودن مرتبه الگوریتم نسبت به $\frac{1}{\epsilon}$ مهم باشد، آن‌گاه روش بالا را می‌توان در هر بعد

ثابت به عنوان یک روش خوب در نظر گرفت. اما اگر مرتبه آن را نسبت به n در نظر بگیریم، با افزایش

n این روش عملاً بی‌ارزش است.

در این جا ما به الگوریتم‌هایی علاقه‌مندیم که مرتبه آن‌ها نسبت به n چند جمله‌ای باشد. بنابراین از روش افراز $\mathbb{R}^n$ به مکعب‌های ریز نمی‌توانیم استفاده کنیم. به علاوه، حتی در این روش ساده، برخی سؤالات جدید و مشکلات اساسی مطرح می‌شوند که پاسخ آن‌ها باید مشخص شود. مثلاً:

سؤال ۱: منظور ما از این که K یک «شکل محدب داده شده» است چیست؟ به چه شکلی داده شده است؟ به علاوه، چه طور می‌توانیم بفهمیم که یک مکعب مستطیل مشخص با K اشتراک دارد یا خیر؟

سؤال ۲: چه وقت یک الگوریتم را مناسب می‌دانیم؟

در این جا ما اشکال محدبی را در نظر می‌گیریم که به وسیله یک *اوراکل*^۱ داده شده‌اند. یعنی ما مستقیماً در مورد خود شکل اطلاعاتی نداریم اما از یک جعبه سیاه می‌توانیم سؤالات مشخصی را در مورد آن بپرسیم و پاسخ دریافت کنیم (تعریف ۲.۱.۲). هم‌چنین به الگوریتم‌هایی علاقه‌مندیم که در آن‌ها تعداد بارهای استفاده از اوراکل بر حسب بعد از مرتبه چند جمله‌ای باشد. می‌توانیم به دنبال یافتن مقدار دقیق حجم باشیم و یا به دنبال تخمینی از آن با خطای نسبی $\epsilon > 0$. در حالت دوم معمولاً می‌خواهیم

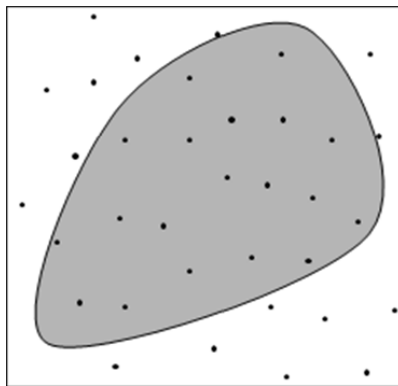
مرتبه الگوریتم نسبت به $\frac{1}{\epsilon}$ هم چند جمله‌ای (و در صورت امکان لگاریتمی) باشد.

فصل ۲ به توضیح در مورد مدل محاسباتی اختصاص دارد. در این فصل پس از توضیح در مورد نحوه داده شدن شکل و نحوه اجرای الگوریتم‌ها، به بیان برخی نتایج منفی در رابطه با محاسبه حجم می‌پردازیم. نشان می‌دهیم مسئله یافتن مقدار دقیق حجم یک شکل محدب در زمان چند جمله‌ای قابل حل نیست. اما تخمین زدن آن یکی از جاهایی از ریاضیات است که در آن قدرت روش‌های تصادفی به اثبات می‌رسد. نشان می‌دهیم در هر الگوریتم غیر تصادفی برای تخمین حجم، تعداد بارهای استفاده از

¹ Oracle

اوراکل از مرتبه حداقل نمایی است. اما با روش‌های تصادفی می‌توان الگوریتمی از مرتبه چندجمله‌ای برای تخمین حجم داد که این موضوع در فصل ۵ بررسی می‌شود.

ساده‌ترین روش تصادفی برای تخمین حجم K استفاده از روش مونت کارلو^۱ است. یعنی یک شکل Q شامل آن بیابیم (به عنوان مثال یک مکعب مستطیل) و تعدادی نقطه تصادفی با توزیع یکنواخت در Q تولید کنیم و ببینیم چه کسری از این نقاط داخل K قرار می‌گیرند. با توجه به این که احتمال وقوع یک نقطه در K برابر با $vol(K)/vol(Q)$ است، عدد حاصل تخمینی از $vol(K)/vol(Q)$ است. مشکل این روش در ابعاد بالا این است که حتی در حالت ساده‌ای که K یک گوی در $\mathbb{R}^n$ باشد، احتمال وقوع یک نقطه در K به صورت نمایی کاهش می‌یابد. بنابراین برای این که تخمین حاصل دارای انحراف معیار محدودی باشد باید تعداد نقاط حداقل به صورت نمایی رشد کند.



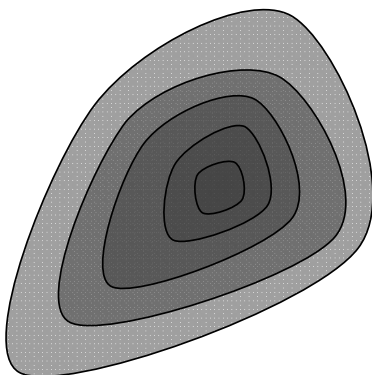
شکل ۲

یک راه برای اصلاح روش مونت کارلو، استفاده از روش مونت کارلو چندفازی^۲ است. یعنی مجموعه‌های محدبی مثل $K_0 \subseteq K_1 \subseteq \dots \subseteq K_m = K$ بیابیم که نسبت حجم K_{i+1} به حجم K_i از مرتبه چندجمله‌ای باشد و تعداد آن‌ها هم بر حسب n چندجمله‌ای باشد. سپس به روش فوق

¹ Monte Carlo

² Multi-phase Monte Carlo

حجم K را محاسبه کنیم. $vol(K_i)/vol(K_{i+1})$ را تقریب بزنیم و اگر حجم K_0 به راحتی قابل محاسبه باشد، به این روش



شکل ۲

در این روش باید به دو سؤال مهم پاسخ دهیم:

(۱) چگونه یک نقطه تصادفی با توزیع یک نواخت در یک شکل محدب تولید کنیم؟ بهترین روش‌های موجود برای تخمین حجم، همگی بر اساس تولید نقطه تصادفی هستند. به همین دلیل تأکید این پایان‌نامه روی این مبحث است (فصل ۴). در بخش بعد مقدمه‌ای در این مورد می‌آوریم.

(۲) چگونه مجموعه‌های محدبی مثل $K_1 \subseteq K_2 \subseteq \dots \subseteq K_m = K$ بیابیم که نسبت حجم K_{i+1} به حجم K_i از مرتبه چندجمله‌ای باشد و تعداد آن‌ها هم بر حسب n چندجمله‌ای باشد؟ معمول‌ترین روش این است که K_i را اشتراک K با یک گوی مناسب حول مبدأ بگیریم به طوری که حجم K_{i+1} حداکثر دو برابر حجم K_i باشد. در این مورد در فصل ۵ بیش‌تر توضیح می‌دهیم.

۲.۱. تولید نقطه تصادفی

همان طور که در بخش قبل گفتیم، تولید یک نقطه تصادفی در یک مجموعه محدب نقش کلیدی ای در الگوریتم های محاسبه حجم ایفا می کند. علاوه بر این کاربرد - که مهم ترین کاربرد آن است - در زمینه های دیگری مانند بهینه سازی تصادفی و تخمین انتگرال توابع هم از تولید نقاط تصادفی استفاده می شود. با توجه به کاربردهای مسئله، مهم است که یک راه حل برای این مسئله در ابعاد بالا هم کارا باشد. به عبارت دیگر می خواهیم مرتبه الگوریتم بر حسب بعد چند جمله ای باشد.

روش های مرسوم برای تولید نقاط تصادفی با توزیع داده شده روی یک ناحیه کران دار $S \subseteq \mathbb{R}^n$ ، محدود به روش های تبدیل، ترکیبی، رد و قبول و مونت کارلو هستند. در روش تبدیل، یک مجموعه B ، یک توزیع مناسب روی B و یک تابع $f: B \rightarrow S$ انتخاب می شوند طوری که اگر یک نقطه تصادفی $x \in B$ با توزیع مذکور تولید کنیم، $f(x)$ توزیع دلخواه ما را داشته باشد. مشکل روش تبدیل این است که تابع تبدیل در اکثر موارد به صورت صریح داده نمی شود و در نتیجه این روش عملاً به جز برای برخی توزیع های خاص قابل استفاده نیست. برای مثال در مورد توزیع یک نواخت روی سادک^۱، کره و بیضی گون می توان از این روش استفاده کرد. اما متأسفانه این روش را نمی توان به چندوجهی ها تعمیم داد. البته می توانیم چندوجهی را به سادک های مجزا افراز کنیم و با استفاده از روش ترکیبی از این سادک ها نمونه گیری کنیم. یعنی ابتدا یکی از سادک ها را به طور تصادفی انتخاب کنیم (با توزیعی که احتمال انتخاب شدن هر سادک متناسب با حجم آن است) و سپس یک نقطه تصادفی با توزیع یک نواخت در آن سادک تولید کنیم. اما تعداد سادک ها و سختی تعیین آن ها عملاً این روش را حتی برای بعدها نه چندان زیاد غیر قابل استفاده می کند.

¹ Simplex

در روش ردوقبول، یک مجموعه مناسب شامل S مانند S' می‌یابیم که بتوانیم از توزیع یک نواخت روی آن نمونه‌گیری کنیم. سپس عمل نمونه‌گیری در S' را به صورت مستقل آن قدر تکرار می‌کنیم که نقطه حاصل عضوی از S باشد. حاصل این الگوریتم، نقطه‌ای تصادفی در S است که توزیع آن یک نواخت خواهد بود. تعداد بارهایی که عمل نمونه‌گیری در S' انجام می‌شود، یک متغیر هندسی است که میانگین آن $vol(S')/vol(S)$ است. این روش در ابعاد بالا دارای یک مشکل جدی است و آن این است که برای تولید یک نقطه تعداد دفعاتی که باید الگوریتم تکرار شود با افزایش بعد به سرعت افزایش می‌یابد. برای مثال اگر بخواهیم برای تولید یک نقطه تصادفی با توزیع یک نواخت در کره واحد n -بعدی از روش ردوقبول و تولید نقاط تصادفی در مکعب $[-1, 1]^n$ استفاده کنیم، متوسط زمان لازم از مرتبه حداقل نمایی است زیرا نسبت حجم کره به حجم مکعب مذکور به صورت نمایی کاهش می‌یابد.

الگوریتم‌هایی که در این پایان‌نامه بررسی می‌کنیم، بر اساس روش مونت کارلو طراحی شده‌اند. روش مونت کارلو روشی قدرت‌مندی برای نمونه‌گیری «تقریبی» از یک توزیع داده شده است. این روش را متروپولیس^۱ به هنگام برخورد با مشکلات محاسباتی در فیزیک آماری مطرح کرد. ایده اصلی آن این است که یک فرآیند مارکوف ساخته شود که توزیع حدی آن همان توزیع داده شده باشد. در این صورت با شبیه‌سازی گام‌های این فرآیند «به تعداد کافی»، توزیع حالت نهایی «تقریباً» برابر با توزیع مطلوب است. معمولاً ساختن چنین فرآیند مارکوفی کار سختی نیست؛ معمول‌ترین ایده، استفاده از رابطه تعادل (تعریف ۵.۱.۳) است. مسئله سخت‌تر این است که چند گام لازم است تا به مقدار قابل قبولی به توزیع پایا نزدیک شویم و البته این که فرآیند را چه طور طراحی کنیم که سرعت هم‌گرایی زیاد باشد. تعداد گام‌های لازم در یک فرآیند مارکوف برای این که توزیع آن به مقدار مشخصی به توزیع حدی نزدیک شود را زمان آمیختگی این فرآیند می‌نامند. مطالعه زمان آمیختگی مبحث جدید و مهمی در نظریه احتمال است که در مسئله ما نیز نقشی کلیدی را ایفا می‌کند (بخش ۲.۳).

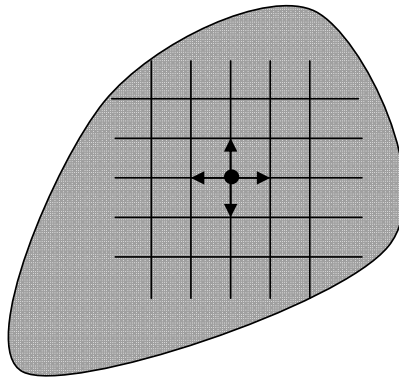
¹ Metropolis

در ادامه این مقدمه تنها به معرفی سه قدم زدن تصادفی می پردازیم. سپس در فصل ۳ مقدمات و نتایج مورد نیاز در زمینه فرآیندهای مارکوف و زمان آمیختگی را بیان می کنیم. پس از آن، فصل ۴ به تحلیل مهم ترین قدم زدن تصادفی برای تولید نقاط تصادفی، یعنی الگوریتم ۳.۲.۱ اختصاص دارد. در فصل ۵ به تحلیل الگوریتم حجم می پردازیم و در نهایت در فصل ۶ برخی کارهای دیگر مرتبط به این مباحث را معرفی می کنیم.

یک شرط که معمولاً روی الگوریتم ها گذاشته می شود، شرط تنبل بودن است. یعنی هر بار با احتمال $\frac{1}{4}$ اصلاً حرکت نمی کنیم و با احتمال $\frac{1}{4}$ مطابق یک قدم زدن تصادفی حرکت می کنیم. شرط تنبل بودن در حالت کلی به ما اجازه تحلیل های بهتری می دهد. هم چنین سرعت الگوریتم به طور متوسط نصف می شود و مرتبه الگوریتم تغییر نمی کند. در این مورد در بخش ۲.۳ ج. بیش تر توضیح می دهیم. بنا بر این در هر سه قدم زدن تصادفی زیر، شرط تنبل بودن را در نظر می گیریم و از تکرار آن خودداری می کنیم.

۱.۲.۱. الگوریتم: قدم زدن تصادفی روی نقاط شبکه ای داخل شکل

در این روش نقاطی از $\mathbb{R}^n$ را در نظر می گیریم که همه مؤلفه های آن ها مضرب صحیحی از یک عدد کوچک δ (که باید به اندازه مناسبی انتخاب شود) هستند. از یک نقطه دلخواه در K شروع می کنیم و در هر گام یکی از $2n$ همسایه نقطه فعلی را انتخاب می کنیم (با توزیع یک نواخت). اگر نقطه جدید عضوی از K بود، به آن نقطه حرکت می کنیم و در غیر این صورت اصلاً حرکت نمی کنیم.



شکل ۴

دایر^۱، فریز^۲ و کنان^۳ با استفاده از این روش اولین الگوریتم محاسبه حجم اشکال محدب با مرتبه چندجمله‌ای را ارائه کردند [۱]، هرچند مرتبه بالای آن (حدود n^{26}) اجازه استفاده عملی از آن را نمی‌داد. بعد از آن تغییرات متعددی در این الگوریتم ایجاد شد و مرتبه آن تا n^8 کاهش یافت.

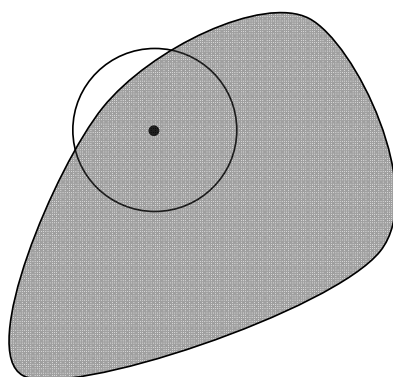
۲.۲.۱. الگوریتم: قدم زدن تصادفی با گام‌های کروی

در این روش در هر گام، یک نقطه تصادفی در یک گوی با شعاع δ حول نقطه فعلی انتخاب می‌کنیم (با توزیع یک‌نواخت). اگر نقطه جدید عضوی از K بود، به آن نقطه حرکت می‌کنیم و در غیر این صورت اصلاً حرکت نمی‌کنیم.

¹ Dyer

² Frieze

³ Kannan

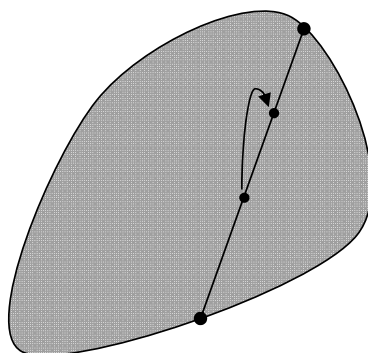


شکل ۵

این روش اولین قدم زدن تصادفی پیوسته‌ای بود که در مورد مسئله محاسبه حجم مورد تحلیل قرار گرفت.

۳.۲.۱. قدم زدن تصادفی Hit-and-run

در این روش در هر گام یک خط تصادفی گذرنده از نقطه فعلی انتخاب می‌کنیم (با استفاده از تولید یک نقطه تصادفی روی کره واحد با توزیع یک نواخت). سپس نقطه بعدی را روی اشتراک این خط با K به صورت تصادفی (با توزیع یک نواخت و به طور مستقل) انتخاب می‌کنیم.



شکل ۶

قدم‌زدن تصادفی Hit-and-run در سال ۱۹۸۱ توسط اسمیت^۱ معرفی شد [۲]. اسمیت نشان داد که در هر بعد ثابت توزیع نقاط تولید شده توسط این الگوریتم به طور نمایی به توزیع یکنواخت روی K میل می‌کند [۲]. اما کرانی که او برای نرخ هم‌گرایی آن به دست آورد بر حسب بعد نمایی است. این الگوریتم از همان ابتدا به عنوان الگوریتمی که به سرعت به توزیع یک‌نواخت هم‌گرا می‌شود شناخته شده بود اما سال‌ها طول کشید که یک اثبات دقیق برای این موضوع یافت شود. اولین اثبات در سال ۱۹۹۳ توسط لواژ^۲ ارائه شد [۳]. او همچنین نشان داد که مرتبه کران او بهترین مرتبه برای این الگوریتم است که این موضوع را در بخش ۱.۶ اثبات می‌کنیم.

این الگوریتم در عمل سریع‌ترین قدم‌زدن تصادفی ارائه شده تا کنون برای تولید نقاط تصادفی در اشکال محدب است، هرچند که با فرض شروع از یک توزیع مناسب، زمان آمیختگی قدم‌زدن با گام‌های کروی هم‌مرتبه با آن است. یک مزیت قدم‌زدن تصادفی Hit-nad-run نسبت به قدم‌زدن‌های تصادفی دیگر این است که توزیع اولیه تأثیری جدی روی زمان آمیختگی آن ندارد. این موضوع را در بخش ۲.۴ توضیح می‌دهیم.

¹ Smith

² Lovasz

۲. مدل محاسباتی

۱.۲. اوراکل‌ها

در این بخش می‌خواهیم در مورد نحوه داده شدن شکل صحبت کنیم. در همه جای این پایان‌نامه K یک شکل محدب و کران‌دار در $\mathbb{R}^n$ است.

۱.۱.۲. تعریف. یک شکل محدب K خوش‌تضمین^۱ با شعاع‌های r و R است هرگاه بدانیم که شامل یک گوی $\mathbb{B}_r(o)$ ^۲ است و همچنین کاملاً در یک گوی $\mathbb{B}_R(o)$ قرار دارد.

یک دلیل برای گذاشتن شرط خوش‌تضمینی این است که قرار است سؤالاتی از نوع «آیا x عضو K

است؟» بپرسیم. بدون شرط خوش‌تضمینی هیچ تضمینی وجود ندارد که با پرسیدن هر تعداد سؤال تنها

جواب «خیر» بگیریم. در این صورت اگر ندانیم K کجای فضا است، هیچ اطلاعاتی در مورد حجم آن

نمی‌توانیم به دست آوریم و حتی با فرض این که $K \subseteq \mathbb{B}_R(o)$ ، اگر کران پایینی برای حجم K نداشته باشیم، نمی‌توانیم آن را با خطای نسبی محدود تخمین بزنیم.

در کاربردهای مختلف اشکال محدب می‌توانند به روش‌های مختلفی داده شوند. برای مثال ممکن

است به عنوان پوش محدب تعدادی نقطه، توسط یک دستگاه نامساوی‌های خطی، و یا توسط سطح تراز

یک تابع محدب^۳ داده شوند. برای در نظر گرفتن یک حالت نسبتاً کلی، از مفهوم *اوراکل* استفاده

می‌کنیم. یک اوراکل مشابه یک جعبه سیاه است که می‌توانیم از آن سؤالات خاصی بپرسیم و جواب‌های

^۱ Well-guaranteed

^۲ منظور از $\mathbb{B}_r(x)$ گوی به شعاع r حول x است.

^۳ در اینجا منظور از سطح تراز مجموعه‌ای به شکل $\{x \in \mathbb{R}^n : f(x) \leq c\}$ است.

خاصی دریافت کنیم. با استفاده از اوراکل‌ها می‌توانیم نتایج را در حالت کلی و مستقل از نحوه داده‌شدن شکل بیان کنیم. بسته به نوع سؤالات و پاسخ‌ها چند نوع اوراکل تعریف می‌شوند. در این پایان‌نامه تنها از اوراکل عضویت قوی استفاده می‌کنیم که در زیر تعریف می‌شود.

۲.۱.۲. تعریف. یک شکل K توسط یک *اوراکل عضویت قوی*^۱ داده شده است در صورتی که مستقیماً اطلاعاتی در مورد آن نداشته باشیم به جز این که بدانیم خوش تضمین با شعاع‌های r و R است و همچنین بتوانیم سؤالاتی از نوع «آیا x عضو K است؟» بپرسیم و پاسخ بله یا خیر دریافت کنیم. به علاوه، فرض می‌کنیم پاسخ همواره درست است.

عبارت «قوی» به این معنی است که پاسخ اوراکل همیشه درست است. برای کنترل خطاهایی که توسط گرد کردن اعداد ایجاد می‌شوند، اوراکل‌های ضعیف نیز تعریف می‌شوند. برای مثال، در یک اوراکل عضویت ضعیف، اگر x به مقدار مشخصی نزدیک به مرز K باشد، پاسخ ممکن است درست یا غلط باشد. همچنین با توجه به برخی کاربردها، *اوراکل‌های جداکننده* نیز تعریف می‌شوند که در آن‌ها هنگامی که پاسخ «خیر» است، یک ابرصفحه که x را از K جدا می‌کند نیز داده می‌شود.

۳.۱.۲. نکته. برای در نظر گرفتن مرتبه یک الگوریتم تصادفی، در حالت کلی شمردن چیزهای مختلفی می‌تواند اهمیت داشته باشد. برای مثال تعداد دفعات استفاده از اوراکل، تعداد اعمال محاسباتی و یا تعداد اعداد تصادفی تولید شده. در این جا فقط تعداد دفعات استفاده از اوراکل را می‌شماریم. زیرا معمولاً استفاده از اوراکل پرهزینه‌ترین عمل در یک گام الگوریتم است و با کران دادن برای تعداد دفعات استفاده از اوراکل، بسته به نوع اوراکل می‌توان برای تعداد اعمال محاسباتی نیز کران داد. در الگوریتم‌های حجم معمولاً مرتبه تعداد اعمال محاسباتی به اندازه n^2 از مرتبه تعداد دفعات استفاده از اوراکل بیش‌تر است.

¹ Strong membership oracle

۲.۲. اجرای قدم‌زدن تصادفی Hit-and-run

در بخش قبل گفتیم که فرض می‌کنیم K توسط یک اوراکل عضویت قوی داده شده است. با این فرض باید در الگوریتم‌های مذکور در بخش ۲.۱ بیش‌تر دقت کنیم. اجرای قدم‌زدن تصادفی روی نقاط شبکه‌ای و هم‌چنین قدم‌زدن تصادفی با گام‌های کروی ساده است زیرا کافی است یک نقطه تصادفی با توزیع مربوطه تولید کنیم و ببینیم که آیا در K هست یا خیر. اما اجرای قدم‌زدن تصادفی Hit-and-run نیاز به دقت بیشتری دارد زیرا باید از یک وتر که ابتدا و انتهای آن را دقیقاً نمی‌دانیم یک نقطه با توزیع یک‌نواخت تولید کنیم. این کار با استفاده از روش ردو قبول ممکن است، زیرا با توجه به این که می‌دانیم قطر K حداکثر $2R$ است، می‌توانیم یک بازه به طول $4R$ شامل وتر مذکور ارائه کنیم. اگر طول وتر برابر با l باشد، متوسط تعداد گام‌های لازم در روش ردو قبول $4R/l$ است. اما می‌توان مرتبه این روش را به شیوه زیر به‌تر کرد.

۲.۲.۱. الگوریتم: قدم‌زدن تصادفی Hit-and-run برای اشکال محدب داده شده توسط اوراکل

عضویت: با شروع از نقطه x ، نقطه بعدی را به روش زیر تولید می‌کنیم.

(۱) متغیرهای نرمال استاندارد مستقل $Z_1, \dots, Z_n$ را تولید کنید و قرار دهید:

$$Z := (Z_1, \dots, Z_n),$$

$$u := \frac{Z}{|Z|}.$$

(۲) قرار دهید $a := x + 2Ru$.

(۳) تا هنگامی که $\frac{x+a}{2} \notin K$ ، قرار دهید $\frac{x+a}{2}$ ، $a := \frac{x+a}{2}$ و این گام را تکرار کنید.

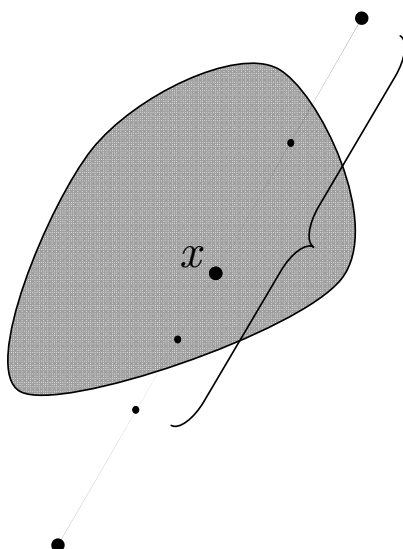
(۴) قرار دهید $b := x - 2Ru$.

(۵) تا هنگامی که $\frac{x+b}{2} \notin K$ ، قرار دهید $b := \frac{x+b}{2}$ و این گام را تکرار کنید.

(۶) یک متغیر تصادفی U با توزیع یک نواخت روی بازه $[0, 1]$ تولید کنید و قرار دهید

$y := U.a + (1-U).b$. در صورتی که $y \notin K$ ، این گام را تکرار کنید و در غیر این صورت y را

به عنوان نقطه بعدی در قدم زدن تصادفی Hit-and-run معرفی کنید.



شکل ۷

۲.۲.۲. تعریف: متغیرهای تصادفی $Z_1, \dots, Z_n$ و U تنها متغیرهای تصادفی استفاده شده در

الگوریتم بالا هستند. این متغیرها را منبع تصادف^۱ این الگوریتم می نامیم.

با داشتن یک دنباله مستقل از منابع تصادف، می توان چند گام از الگوریتم فوق را به صورت متوالی

انجام داد. هم چنین این متغیرها به نقطه شروع بستگی ندارند. بنابراین می توان الگوریتم بالا را با شروع از

دو نقطه متفاوت و با منبع تصادف یکسان انجام داد.^۲

^۱ Source of randomness

^۲ این در واقع یک جفت سازی برای یک قدم زدن تصادفی با نقاط شروع متفاوت است. به بخش ۲.۳ ب مراجعه کنید.

در صورتی که طول وتر $(x + \mathbb{R}.u) \cap K$ برابر با l باشد، برای گام‌های ۲ تا ۵ از مرتبه $\log R/l$ بار از اوراکل استفاده می‌شود. می‌توان نشان داد که با تکرار این الگوریتم به تعداد به اندازه کافی زیاد، در گام‌های ۲ تا ۵ به طور متوسط از مرتبه $\log R/r$ بار از اوراکل استفاده می‌شود. در گام ششم نیز با توجه به این که طول پاره خط واصل a و b حداکثر $2l$ است، به طور متوسط ۲ بار از اوراکل استفاده می‌شود. در نتیجه برای کران دادن برای تعداد دفعات استفاده از اوراکل، کافی است برای تعداد گام‌های مورد نیاز برای تکرار الگوریتم Hit-and-run کران ارائه دهیم. این کار پس از بیان مقدمات آن در فصل ۳، در فصل ۴ انجام می‌شود.

۳.۲. الگوریتم‌های تصادفی و غیر تصادفی

۳.۲.۱. نتایج منفی برای الگوریتم‌های غیر تصادفی

اولین سؤال در مورد محاسبه حجم این است که آیا می‌توان حجم یک شکل محدب را در زمان چندجمله‌ای دقیقاً محاسبه کرد یا خیر. مثلاً فرض کنید شکل توسط نابرابری‌های خطی با ضرایب گویا داده شده باشد که صورت‌ها و مخرج‌ها در آن‌ها کران دار هستند. پاسخ این سؤال خیر است. دایر و فریز و هم‌چنین خاچیان^۱ قضیه زیر را ثابت کردند.

۳.۲.۲. قضیه: مسئله محاسبه حجم یک چندوجهی که با یک دستگاه نابرابری‌های خطی با ضرایب

گویا داده شده باشد یک مسئله $\#P$ -سخت است.^۲

اثبات: برای این منظور نشان می‌دهیم محاسبه حجم یک مسئله $\#P$ -سخت را حل می‌کند. فرض

کنید P یک ترتیب جزئی روی $\{a_1, \dots, a_n\}$ باشد. چندوجهی‌ای که با روابط زیر داده می‌شود را در نظر بگیرید:

^۱ Khachiyan

^۲ منظور از کلاس $\#P$ ، مجموعه مسائل شمارشی‌ای است که متناظر با مسائل تصمیم‌گیری NP هستند.

$$0 \leq x_i \leq 1 \quad (i = 1, \dots, n)$$

$$x_i \leq x_j \text{ در صورتی که } a_i \leq a_j.$$

هر ترتیب کامل، متناظر با یک سادک با حجم $\frac{1}{n!}$ می شود و سادک های متناظر با ترتیب های مختلف مجزا هستند. بنابراین تعداد توسیع های خطی P^1 ، $n!$ برابر حجم چندوجهی بالا می شود. برای P^2 و وینکلر^۳ ثابت کردند که محاسبه تعداد توسیع های خطی مسئله ای $\#P$ -سخت است. بنابراین اگر $P \neq \#P$ ، آن گاه محاسبه حجم یک چندوجهی که با نابرابری های خطی داده شده باشد در زمان چندجمله ای ممکن نیست. $\square$

دیدیم محاسبه دقیق حجم ممکن نیست. بنابراین از این به بعد فقط روی مسئله تخمین حجم تمرکز می کنیم. قضیه زیر که توسط الکز^۴ ثابت شده است نشان می دهد که با محدود کردن خود به الگوریتم های غیر تصادفی نمی توانیم به نتایج خوبی امیدوار باشیم.

۳.۳.۲. قضیه. برای تخمین حجم یک شکل محدب در $\mathbb{R}^n$ که توسط یک اوراکل عضویت قوی داده شده باشد، خطای نسبی هر الگوریتم غیر تصادفی که q بار از اوراکل استفاده کند حداقل $\sqrt{2^n/(q+1)}$ است.

ایده اثبات این قضیه این است که برای هر الگوریتم غیر تصادفی، دو شکل محدب ساخته می شوند که الگوریتم برای آن ها یک نتیجه را بدهد و نسبت حجم آن ها لااقل $2^n/(q+1)$ باشد. ابتدا گزاره زیر را ثابت می کنیم.

۴.۳.۲. گزاره: فرض کنید $K \subseteq \mathbb{B}^n$ یک چندوجهی محدب با m رأس باشد.^۱ در این صورت

^۱ یعنی تعداد راه هایی که می توان P را به یک ترتیب کامل گسترش داد.

^۲ Brightwell

^۳ Winkler

^۴ Elekes

$$\frac{vol(K)}{vol(\mathbb{B}^n)} \leq \frac{m}{2^n}.$$

اثبات: فرض کنید K پوش محدب نقاط $x_1, x_2, \dots, x_m$ باشد. نشان می دهیم که اگر B_i را گوی به قطر $[0, x_i]$ بگیریم، آن گاه

$$K \subseteq \bigcup_{i \leq m} B_i.$$

در این صورت با توجه به این که شعاع هر B_i حداکثر $\frac{1}{2}$ است حکم ثابت می شود. فرض کنید

$$y \in K \quad \text{پس می توان نوشت} \quad y = \sum_{i=1}^q c_i x_i \quad \text{که} \quad c_j \geq 0 \quad \text{و} \quad \sum_{i=1}^q c_i = 1.$$

$$\text{بنابراین} \quad y \cdot y = \sum_{i=1}^q c_i y \cdot x_i \quad \text{پس} \quad \sum_{i=1}^q c_i y \cdot (y - x_i) = 0 \quad \text{پس} \quad i \leq q \quad \text{موجود است که}$$

$$y \cdot (y - x_i) \leq 0. \quad \text{این به این معنی است که زاویه} \angle Oyx_i \text{ منفرجه است (که } O \text{ مبدأ مختصات}$$

است) و بنابراین $y \in B_j$. $\square$

اثبات قضیه ۳.۳.۲: فرض کنید یک الگوریتم غیر تصادفی داده شده باشد. این که الگوریتم غیر تصادفی

است، به این معنی است که هر نقطه ای که به اوراکل می دهد فقط تابعی از نقاط داده شده قبلی و

جواب های گرفته شده است و همچنین خروجی آن نیز فقط تابع نقاط و جواب ها است. الگوریتم را برای

$K_1 = \mathbb{B}^n$ اجرا کنید و فرض کنید نقاط $x_1, \dots, x_q$ را به اوراکل بدهد. فرض کنید از بین این نقاط،

$x_1, \dots, x_{i_m}$ ($m \leq q$) داخل K_1 و بقیه خارج آن باشند. K_2 را پوش محدب $x_1, \dots, x_{i_m}$ بگیرید.

با توجه به این که هریک از x_i ها یا متعلق به هردوی K_1 و K_2 هستند و یا متعلق به هیچ کدام

آن ها، به صورت استقرایی ثابت می شود که اگر الگوریتم را برای K_2 استفاده کنیم، همان نقاط استفاده

^۱ منظور از $\mathbb{B}^n$ گوی واحد n -بعدی است.

می‌شوند و همان تخمین برای حجم K_p به دست می‌آید. طبق لم ۴.۳.۲ نسبت حجم K_1 به K_p حداقل $2^n/q$ است. پس خطای نسبی برای یکی از آن‌ها حداقل $\sqrt{2^n/q}$ است. البته ممکن است K_p دارای شرط خوش‌تضمینی نباشد. برای اصلاح این موضوع، K'_p را پوش محدب $K_p \cup \mathbb{B}_1(0)$ می‌گیریم. با اثباتی مشابه لم ۴.۳.۲، به دست می‌آوریم که K'_p داخل $q+1$ گوی به شعاع $\frac{1}{p}$ قرار دارد و به طور مشابه حکم ثابت می‌شود. $\square$

۵.۳.۲. نتیجه. هر الگوریتم غیر تصادفی که بخواهد حجم اشکال محدب داده شده توسط اوراکل عضویت قوی را با خطای نسبی محدود محاسبه کند، باید از مرتبه حداقل 2^n بار از اوراکل استفاده کند. $\square$

۶.۳.۲. نکته: قضیه فوق، یک قضیه در نظریه پیچیدگی محاسبه نیست. در واقع این قضیه بدون توجه به تعداد اعمال محاسباتی نشان می‌دهد که تعداد چندجمله‌ای نقطه اطلاعات کافی برای تخمین حجم را دربر ندارند. بنابراین این قضیه را می‌توان به عنوان نتیجه‌ای در نظریه اطلاعات قلمداد کرد. بارنی^۱ و فوردی^۲ قضیه دیگری ثابت کردند که در این‌جا فقط صورت آن را می‌آوریم.

۷.۳.۲. قضیه: هیچ الگوریتم غیر تصادفی وجود ندارد که در زمان چندجمله‌ای یک کران پایین $\underline{vol}(K)$ و یک کران بالای $\overline{vol}(K)$ را برای هر شکل محدب $K \subseteq \mathbb{R}^n$ (داده شده توسط یک اوراکل عضویت قوی) محاسبه کند به طوری که

$$\frac{\overline{vol}(K)}{\underline{vol}(K)} \leq (c \frac{n}{\log n})^n.$$

¹ Barany

² Furedi

۸.۳.۲. نکته: یک نتیجه این قضیه‌ها این است که الگوریتم‌های تصادفی ما برای محاسبه حجم باید از یک منبع واقعی بیت‌های تصادفی (یا اعداد تصادفی در $[0,1]$) استفاده کنند و نمی‌توانند از دنباله‌های شبه تصادفی استفاده کنند تا تبدیل به یک الگوریتم غیر تصادفی شوند. هر چند که در عمل شبیه‌سازی‌های کامپیوتری به نظر خوب کار می‌کنند.

۳.۲. الگوریتم‌های تصادفی

با توجه به آن چه گفتیم، برای تخمین حجم در زمان چندجمله‌ای ناچاریم از الگوریتم‌های تصادفی استفاده کنیم. در الگوریتم‌های تصادفی همیشه احتمال خطا وجود دارد. بنابراین علاوه بر دقت محاسباتی، یک دقت احتمالاتی نیز باید مشخص شود. ما به دنبال الگوریتمی هستیم که برای هر شکل محدب $K \subseteq \mathbb{R}^n$ (که توسط یک اوراکل عضویت قوی داده شده است)، با مشخص کردن هر $\epsilon, \eta > 0$ ، عدد حقیقی ζ را به دست آورد که به احتمال حداقل $1 - \eta$ داشته باشیم

$$(1 - \epsilon)\zeta < \text{vol}(K) < (1 + \epsilon)\zeta.$$

احتمال خطای η تأثیر زیادی در سرعت الگوریتم‌ها ندارد. اگر در تعدادی گام بتوانیم با احتمال

حداقل $\frac{3}{4}$ به خطای نسبی ϵ دست یابیم، آن گاه می‌توانیم همین روش را T بار به طور مستقل تکرار

کنیم و میانگین نتایج حاصل را محاسبه کنیم. کارپ^۱ و لوبی^۲ و هم‌چنین جروم^۳، والیان^۴ و وزیرانی^۵ نشان دادند که خطای احتمالی عدد حاصل، بر حسب T به طور نمایی کاهش می‌یابد. بنابراین با گذشت از

$\eta = \frac{1}{4}$ به یک η کوچک، تعداد گام‌ها تنها در یک عامل $\log \frac{1}{\eta}$ ضرب می‌شود. می‌توانیم به

وابستگی به ϵ نیز علاقه‌مند باشیم. این وابستگی معمولاً به شکل یک عامل $\frac{1}{\epsilon} \log \frac{1}{\epsilon}$ است.

^۱ Karp

^۲ Luby

^۳ Jerrum

^۴ Valiant

^۵ Vazirani

بهترین الگوریتمی که تا کنون ارائه شده، از مرتبه $O^*(n^4)$ است [۴]. منظور ما از نماد O^* این است که عامل های $\log^c(n)$ و همچنین ثابت هایی که به ϵ ، η و $\log(R/r)$ بستگی دارند را حذف می کنیم. در فصل ۵ یک الگوریتم حجم از مرتبه $O^*(n^5)$ ارائه می کنیم (الگوریتم ۱.۱.۵). همچنین در بخش ۴.۶ در مورد الگوریتم مرتبه چهار مذکور توضیح می دهیم.

۳. هسته‌های مارکوف و فرآیندهای مارکوف

در این فصل مقدمات و نتایج لازم در مورد فرآیندهای مارکوف را بیان می‌کنیم که یک مدل ریاضی کلی از الگوریتم‌های تصادفی ارائه می‌کنند. در بخش ۱.۳ تعاریف اولیه را می‌آوریم. سپس در بخش ۲.۳ در مورد آمیختگی در فرآیندهای مارکوف بحث می‌کنیم. هدف اصلی این فصل قضیه ۱۷.۲.۳ است که کرانی برای زمان آمیختگی یک فرآیند مارکوف ارائه می‌کند. از این کران در فصل بعد برای تخمین زمان آمیختگی قدم‌زدن تصادفی Hit-and-run استفاده می‌شود.

۱.۳. تعاریف اولیه

منظور از یک شیء تصادفی، نقطه‌ای تصادفی است که از یک مجموعه مشخص به نام فضای حالت انتخاب می‌شود. در صورتی که فضای حالت زیرمجموعه‌ای از $\mathbb{R}^n$ باشد، آن را یک نقطه تصادفی نیز می‌نامیم. یک فرآیند تصادفی، دنباله‌ای از اشیاء تصادفی $(X_t)_{t=1}^\infty$ است که روی یک فضا تعریف شده باشند. یک فرآیند تصادفی در واقع یک شیء تصادفی (روی فضایی دیگر) نیز هست، اما اگر t را به عنوان زمان در نظر بگیریم، یک فرآیند تصادفی را می‌توان به عنوان یک شیء متحرک تصادفی در نظر گرفت. یعنی مقدار آن در لحظه اول به صورت تصادفی با توزیع X_1 انتخاب می‌شود و پس از مشخص شدن مقدار آن تا زمان t ، در لحظه بعد با توزیع شرطی X_{t+1} نسبت به $X_1, \dots, X_t$ انتخاب می‌شود. این توزیع را به صورت زیر نمایش می‌دهیم

$$\mathbb{P}(X_{t+1} \in A | X_1, \dots, X_t)$$

که A زیر مجموعه‌ای از فضای نمونه است.^۱ این مقدار البته تابعی از t و $X_1, \dots, X_t$ و A است.

اگر این توزیع به زمان‌های گذشته (یعنی $X_1, \dots, X_{t-1}$) بستگی نداشته باشد، این فرآیند تصادفی را یک فرآیند مارکوف یا زنجیر مارکوف می‌نامیم. به عبارت دیگر

$$\mathbb{P}(X_{t+1} \in A | X_1, \dots, X_t) = P_{t, X_t}(A).$$

این معادل است با

$$\mathbb{P}(X_{t+1} \in A | X_1, \dots, X_t) = \mathbb{P}(X_{t+1} \in A | X_t).$$

به علاوه، اگر توزیع فوق به زمان نیز بستگی نداشته باشد، این فرآیند تصادفی را یک فرآیند مارکوف مستقل از زمان می‌نامیم.

۱.۱.۳. نکته: در یک فرآیند مارکوف، برای هر تابع اندازه‌پذیر و کران دار f روی فضای نمونه، داریم

$$\mathbb{E}(f(X_{t+1}) | X_1, \dots, X_t) = \mathbb{E}(f(X_{t+1}) | X_t).$$

این موضوع برای توابع مشخصه مجموعه‌ها و ترکیب‌های خطی آن‌ها واضح است. حالت کلی نیز با حد گرفتن به دست می‌آید.

در ادبیات معمول امروزی معمولاً منظور از یک فرآیند مارکوف، یک فرآیند مارکوف مستقل از زمان است. در این پایان‌نامه نیز همه فرآیندها مستقل از زمان هستند مگر این که خلاف آن تصریح شود.

یک فرآیند مارکوف را با دانستن احتمال‌های $\mathbb{P}(X_{t+1} \in A | X_t)$ می‌توان شناخت. برای مثال، Ω را یک مجموعه دلخواه (مجهز به یک σ -جبر) بگیرید و فرض کنید برای هر $u \in \Omega$ ، اندازه احتمال P_u روی Ω داده شده باشد به طوری که برای هر $A \subseteq \Omega$ (در آن σ -جبر) مقدار $P_u(A)$ تابعی اندازه‌پذیر از u باشد. خانواده $\mathcal{M} = \{P_u : u \in \Omega\}$ را یک هسته مارکوف می‌نامیم. یک هسته مارکوف به همراه یک توزیع اولیه Q روی Ω ، یک فرآیند مارکوف (مستقل از زمان) را تعریف می‌کند؛

^۱ این در واقع امید ریاضی شرطی تابع مشخصه پیشامد $\{X_{t+1} \in A\}$ نسبت به $X_1, \dots, X_t$ است.

یعنی دنباله $X_0, X_1, X_2, \dots$ از متغیرهای تصادفی در Ω که X_0 از توزیع Q_0 انتخاب می‌شود و پس از انتخاب $X_0, \dots, X_i$ ($i \geq 0$)، X_{i+1} از توزیع P_{X_i} انتخاب می‌شود.

۲.۱.۳. نکته: اگر Q_i توزیع X_i باشد، آن‌گاه

$$Q_{i+1}(A) = \int_{\Omega} P_u(A) dQ_i(u).$$

به‌عنوان مثال، فرض کنید $(\Omega, \mathcal{A}, \mu)$ یک فضای اندازه و $f: \Omega \times \Omega \rightarrow \mathbb{R}$ یک تابع انتگرال پذیر نسبت به اندازه حاصل ضربی $\mu \times \mu$ باشد. اگر به‌علاوه برای هر $u \in \Omega$ داشته باشیم

$$\int_{\Omega} f(u, v) d\mu(v) = 1, \quad \text{آنگاه } f \text{ یک هستهٔ مارکوف به صورت زیر تعریف می‌کند:}$$

$$P_u(A) = \int_A f(u, v) d\mu(v).$$

۳.۱.۳. تعریف: تابع f را چگالی/احتمال گذر این هستهٔ مارکوف نسبت به اندازه μ می‌نامیم.

$f(u, v)$ را با $f_u(v)$ نیز نشان می‌دهیم. این چگالی احتمال گذر را متقارن می‌نامیم هرگاه

$$f(u, v) = f(v, u).$$

توجه کنید که برای یک هستهٔ مارکوف دلخواه، لزوماً یک تابع احتمال گذر وجود ندارد. مثلاً ممکن

است P_u نسبت به μ اتم داشته باشد.

۴.۱.۳. تعریف: اندازهٔ احتمال Q روی Ω را یک توزیع پایا برای هستهٔ مارکوف $\mathcal{M}$ گوئیم، هرگاه

اگر X_0 را از این توزیع انتخاب کنیم، X_1 نیز همین توزیع را داشته باشد (پس هر X_i نیز همین توزیع را خواهد داشت). معادلاً برای هر $A \subseteq \Omega$ داشته باشیم

$$\int_{\Omega} P_u(A) dQ(u) = Q(A).$$

در این پایان نامه همیشه یک توزیع پایای ثابت Q را برای هسته مارکوف M در نظر می گیریم. در این جا به وجود و یکتایی توزیع پایا کاری نداریم. در کاربردهایی که مد نظر ما است، توزیع Q به صورت صریح داده می شود و یکتا خواهد بود. این موضوع از قضیه ۱۷.۲.۳ نتیجه می شود.

۵.۱.۳. **تعریف:** یک فرآیند مارکوف را برگشت پذیر^۱ می نامند هرگاه دو زوج (X_0, X_1) و (X_1, X_0)

هم توزیع باشند. به عبارت دیگر، برای هر دو زیرمجموعه A و B از Ω داشته باشیم

$$\mathbb{P}(X_0 \in A, X_1 \in B) = \mathbb{P}(X_0 \in B, X_1 \in A)$$

و یا معادلاً

$$\int_B P_u(A) dQ(u) = \int_A P_u(B) dQ(u).$$

که در آن Q توزیع اولیه فرآیند است. این رابطه، رابطه تعادل^۲ نام دارد.

۶.۱.۳. **گزاره:** اگر فرآیند مارکوف M با توزیع اولیه Q برگشت پذیر باشد، آنگاه Q یک توزیع پایا برای M است.

اثبات: در رابطه تعادل قرار دهید $B := \Omega$. $\square$

بررسی رابطه تعادل، معمولاً ساده ترین راه برای اثبات پایا بودن یک توزیع برای یک هسته مارکوف است. البته یک توزیع پایا لزوماً در رابطه تعادل صدق نمی کند.

در یک فرآیند مارکوف برگشت پذیر، طبق خاصیت مارکوف می توانیم برای هر n نتیجه بگیریم

$$\mathbb{P}(X_0 \in A_0, \dots, X_n \in A_n) = \mathbb{P}(X_n \in A_0, \dots, X_0 \in A_n).$$

همچنین اگر زنجیر مارکوف برای زمان های منفی نیز تعریف شده باشد^۳، آنگاه

$$\mathbb{P}(X_0 \in A_0, \dots, X_n \in A_n) = \mathbb{P}(X_0 \in A_0, \dots, X_{-n} \in A_n).$$

¹ Time-reversible

² Detailed balanced equation

³ طبق قضیه توسیع کولموگوروف، برای هر توزیع پایا در لحظه صفر، می توان فرآیند را برای زمان های صحیح نیز تعریف کرد.

به عبارت دیگر قانون فرآیند تحت قرینه کردن زمان ناوردا است و این علت انتخاب نام برگشت پذیری است.

۷.۱.۳. نکته: اگر M نسبت به Q چگالی احتمال گذر داشته باشد، برگشت پذیر بودن معادل با تقارن تابع چگالی احتمال گذر است. به طور کلی تر اگر f تابع چگالی احتمال گذر M نسبت به یک اندازه μ باشد و g چگالی احتمال Q نسبت به μ باشد، آن گاه برگشت پذیر بودن معادل با این است که

$$g(x)f(x,y) = g(y)f(y,x).$$

۸.۱.۳. تعریف: رابطه فوق ابزار مناسبی برای طراحی یک فرآیند مارکوف است که یک توزیع داده شده Q را به عنوان توزیع پایا داشته باشد. در صورتی که توزیع آن فرآیند به توزیع پایا میل کند، با شبیه سازی گام های فرآیند به تعداد کافی می توان یک نمونه تصادفی با توزیعی تقریباً برابر با Q تولید کرد. این روش روش مونت کارلو^۱ نام دارد که ابزار مهمی در شبیه سازی تصادفی است.

به عنوان مثال، الگوریتم های ۱.۲.۱ تا ۳.۲.۱ بر اساس روش مونت کارلو ساخته شده اند. یک ویژگی مهم روش مونت کارلو این است که کافی است تابع چگالی توزیع پایا، یعنی تابع g را با تقریب ضرب در یک عدد ثابت بدانیم. برای مثال، در الگوریتم های ۱.۲.۱ تا ۳.۲.۱، توزیع پایا توزیع یک نواخت روی یک مجموعه محدب K است که تابع چگالی آن نسبت به اندازه لبگ برابر با $\frac{1}{vol(K)}$ است. بنابراین برای استفاده از روش مونت کارلو نیازی نیست که حجم K را بدانیم. اتفاقاً در این پایان نامه از روش مونت کارلو برای تخمین حجم K استفاده می کنیم.

¹ Markov Chain Monte Carlo (MCMC)

۲.۳. آمیختگی در فرآیندهای مارکوف

یک قضیه اساسی در نظریه فرآیندهای مارکوف این است که در یک فضای حالت متناهی، هر فرآیند مارکوف تحویل‌ناپذیر و غیر متناوب یک توزیع پایای یکتا دارد و توزیع فرآیند در لحظه t ، مستقل از توزیع اولیه، به توزیع پایا میل می‌کند. در حالت کلی قضیه هریس حکم مشابهی را بیان می‌کند. منظور از زمان آمیختگی، هریک از چندین فرمول‌بندی ایده زیر است: چند گام از فرآیند لازم است تا توزیع آن به مقدار مشخصی نزدیک به توزیع پایا شود.

در بخش ۲.۳.۱ مفهوم نزدیک بودن توزیع‌ها را توسط یک متریک فرمول‌بندی می‌کنیم و به بیان خواص مقدماتی آن می‌پردازیم. سپس در بخش ۲.۳.۲ ب در مورد ارتباط این متریک و مفهومی به نام جفت‌سازی صحبت می‌کنیم که از آن تنها در فصل ۵ استفاده می‌شود. در بخش ۲.۳.۳ توضیحاتی در مورد منظور ما از آمیختگی می‌دهیم و در نهایت در بخش ۲.۳.۴ به هدف اصلی این فصل می‌پردازیم: قضیه‌ای برای دادن کران بالا برای زمان آمیختگی.

۲.۳.۱. فاصله مجموع تغییرات^۱

می‌توان گفت فاصله مجموع تغییرات طبیعی‌ترین فاصله روی توزیع‌ها است. اگر دو توزیع را به عنوان دو تابع روی زیرمجموعه‌های Ω در نظر بگیریم، فاصله مجموع تغییرات همان بیش‌ترین مقدار تفاوت این دو تابع است.

۲.۳.۲. تعریف: فاصله مجموع تغییرات بین دو توزیع احتمال μ و ν روی Ω (مجهز به یک σ -

میدان داده شده) به صورت زیر تعریف می‌شود

$$\|\mu - \nu\|_{TV} := \sup_{A \subseteq \Omega} |\mu(A) - \nu(A)|.$$

¹ Total variation distance

طبق نامساوی مثلث برای اعداد حقیقی، واضح است که فاصلهٔ مجموع تغییرات در نامساوی مثلث صدق می‌کند. در واقع مجموعهٔ همهٔ توزیع‌های احتمال روی Ω ، زیرمجموعهٔ مجموعهٔ همهٔ اندازه‌های علامت‌دار متناهی روی Ω است. این فضا یک فضای خطی است که مجهز به نرم $\|\mu\|_{TV} := \sup_A |\mu(A)|$ است. فاصلهٔ مجموع تغییرات در واقع همان متر القاشده توسط این نرم است.

۳.۲.۳. لم: نسبت به هریک از μ و ν تابعی محدب است.

اثبات: فرض کنید $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$ که در آن $0 \leq \alpha \leq 1$. در این صورت برای هر

$A \subseteq \Omega$ ، با استفاده از تحدب تابع قدر مطلق روی اعداد حقیقی، به دست می‌آوریم

$$\begin{aligned} |\mu(A) - \nu(A)| &\leq \alpha |\mu_1(A) - \nu(A)| + (1-\alpha) |\mu_2(A) - \nu(A)| \\ &\leq \alpha \|\mu_1 - \nu\|_{TV} + (1-\alpha) \|\mu_2 - \nu\|_{TV} \end{aligned}$$

و حکم ثابت می‌شود. $\square$

در حالتی که دو توزیع دارای تابع چگالی (نسبت به یک توزیع دیگر) باشند، مثلاً در حالتی که Ω متناهی است، گزارهٔ زیر یک فرمول صریح برای فاصلهٔ مجموع تغییرات ارائه می‌کند.

۴.۲.۳. گزاره: اگر f و g به ترتیب توابع چگالی احتمال μ و ν نسبت به اندازهٔ دلخواه φ روی $(\Omega, \mathcal{A})$ باشند، آن‌گاه

$$\begin{aligned} \|\mu - \nu\|_{TV} &= \int_{f(x) > g(x)} (f(x) - g(x)) d\varphi(x) \\ &= \frac{1}{2} \int_{\Omega} |f(x) - g(x)| d\varphi(x). \end{aligned}$$

اثبات: قرار دهید $B := \{x \in \Omega : f(x) \geq g(x)\}$. برای هر مجموعهٔ $A \subseteq \Omega$ داریم

$$\mu(A) - \nu(A) \leq \mu(A \cap B) - \nu(A \cap B) \leq \mu(B) - \nu(B).$$

نامساوی اول به این علت برقرار است که روی $A \setminus B$ داریم $f < g$ و نامساوی دوم به این علت درست است که روی $B \setminus A$ داریم $f \geq g$.

با استدلالی مشابه به دست می آوریم که

$$\nu(A) - \mu(A) \leq \nu(B^c) - \mu(B^c).$$

طرف های راست این دو نامساوی با هم برابر هستند (این موضوع با در نظر گرفتن تفاضل آن ها دیده می شود). همچنین در هر دو نامساوی تساوی اتفاق می افتد. بنابراین به دست می آوریم

$$\begin{aligned} \|\mu - \nu\|_{TV} &= \mu(B) - \nu(B) = \frac{1}{2}[\mu(B) - \nu(B) + \nu(B^c) - \mu(B^c)] \\ &= \frac{1}{2} \int_{\Omega} |f(x) - g(x)| d\varphi(x) \end{aligned}$$

و حکم ثابت می شود. $\square$

تعریف معادل دیگری نیز برای فاصله مجموع تغییرات وجود دارد که استفاده از آن در بسیاری از جاها مفید است:

۵.۲.۳. گزاره: فرض کنید μ فاقد/تم^۱ باشد. در این صورت

$$\|\mu - \nu\|_{TV} = \sup\left\{\int_{\Omega} f d\mu - \int_{\Omega} f d\nu\right\}$$

که در آن سوپریمم روی تمام توابع اندازه پذیر $f: \Omega \rightarrow [0, 1]$ گرفته می شود.

با توجه به این که مجموعه توزیع های علامت دار متناهی همان دوگان فضای توابع اندازه پذیر و کران دار است، عبارت طرف راست در واقع فاصله μ و ν به عنوان دو تابع است.

در واقع این گزاره برای دو توزیع احتمال دلخواه نیز درست است. اما در این جا فقط به این حالت احتیاج داریم که اثبات آن کوتاه تر است.

^۱ یک اتم مجموعه ای با اندازه ناصفر است به طوری که هر طور که آن را به دو زیرمجموعه افراز کنیم، اندازه یکی از آن دو صفر باشد.

برای اثبات لم بالا به یک تعریف دیگر نیاز داریم که در آینده نیز از آن استفاده خواهیم کرد.

۶.۲.۳. تعریف: برای دو توزیع μ و ν روی یک فضای اندازه $(\Omega, \mathcal{A})$ ، تابع $h_{\mu, \nu}(x)$ به ازای

$x \in [0, 1]$ به صورت زیر تعریف می شود

$$h_{\mu, \nu}(x) = \sup_f \int_{\Omega} f d\nu - x = \sup_f \int_{\Omega} f d\nu - \int_{\Omega} f d\mu$$

که در آن سوپریمم روی تمام توابع اندازه پذیر $f : \Omega \rightarrow [0, 1]$ که $\int_{\Omega} f d\mu = x$ گرفته می شود.

با ثابت گرفتن x ، $h_{\mu, \nu}(x)$ معیاری برای فاصله μ و ν است. همچنین

$$\|\mu - \nu\|_{TV} = \sup_x h_{\mu, \nu}(x).$$

برای اختصار $h_{\mu, \nu}(x)$ را با $h(x)$ نشان می دهیم.

۷.۲.۳. نکته: به سادگی دیده می شود که $h(x)$ تابعی مقعر بر حسب x است و

$$0 \leq h(x) \leq 1 - x. \text{ همچنین برای هر مجموعه } A \subseteq \Omega \text{ که } \mu(A) = x \text{ داریم}$$

$$-h(1-x) \leq \nu(A) - \mu(A) \leq h(x).$$

۸.۲.۳. لم: فرض کنید μ فاقد اتم باشد. در این صورت برای هر $0 \leq x \leq 1$ ، سوپریمم موجود در

تعریف $h(x)$ توسط تابع مشخصه یک مجموعه اتخاذ می شود.

اثبات: باید مجموعه ای بیابیم که اندازه آن نسبت به μ کوچک و اندازه آن نسبت به ν بزرگ باشد.

به صورت شهودی باید این مجموعه را بیش تر از جاهایی انتخاب کنیم که چگالی ν نسبت به μ زیاد

است. این ایده را به صورت زیر دنبال می کنیم.

مجموعه اندازه پذیری مثل U موجود است که $\mu(U) = 0$ و $\nu(U)$ ماکزیمم است. μ' و ν' را
تحدید μ و ν به $U \setminus \Omega$ بگیرید. حال ν' نسبت به μ' پیوسته مطلق است. پس طبق قضیه رادون-
نیکودایم^۱، مشتق $\phi = d\nu'/d\mu'$ وجود دارد.

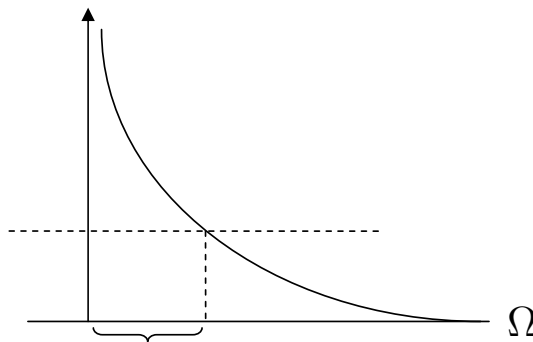
به ازای $t \geq 0$ ، تعریف کنید

$$A_t := U \cup \{u \in \Omega \setminus U : \phi(u) \geq t\},$$

$$A'_t := U \cup \{u \in \Omega \setminus U : \phi(u) > t\},$$

$$s := \inf\{t : \mu(A_t) \leq x\}.$$

البته برای $x = 0$ قرار دهید $s = \infty$ و $A_s = A'_s = U$. بنابراین $A'_s \subseteq A_s$ و برای هر
 $u \in A_s \setminus A'_s$ داریم $\phi(u) = s$. توجه کنید که $A_s = \cap\{A_t : t < s\}$ و
 $A'_s = \cup\{A_t : t > s\}$. بنابراین $\mu(A'_s) \leq x \leq \mu(A_s)$. با توجه به این که μ اتم ندارد، زیر
مجموعه $A'_s \subseteq V \subseteq A_s$ یافت می شود که $\mu(V) = x$ ^۲. نشان می دهیم V همان مجموعه
مطلوب است.



شکل ۸

چگالی ν نسبت به μ

^۱ Radon-Nikodym

^۲ این موضوع توسط سرپینسکی (Sierpiński) بیان شده است و به وسیله لم زرن ثابت می شود.

F را تابع مشخصه V بگیرید. بنابراین $\int_{\Omega} F d\mu = x$ و $\int_{\Omega} F d\nu = \nu(V)$. طبق نحوه انتخاب V ، روی مجموعه $V \setminus U$ داریم $\phi \geq s$ و روی $\Omega \setminus V$ داریم $\phi \leq s$. بنابراین برای هر تابع اندازه پذیر $f : \Omega \rightarrow [0, 1]$ به دست می آوریم

$$\begin{aligned} \int_{\Omega} f d\nu &= \int_U f d\nu + \int_{V \setminus U} f d\nu + \int_{\Omega \setminus V} f d\nu \\ &= \int_U f d\nu + \int_{V \setminus U} \phi d\mu - \int_{V \setminus U} (1 - f)\phi d\mu + \int_{\Omega \setminus V} f\phi d\mu \\ &\leq \nu(U) + \nu(V \setminus U) - s \int_{V \setminus U} (1 - f) d\mu + s \int_{\Omega \setminus V} f d\mu \\ &= \nu(V) - s\mu(V) + s \int_{\Omega} f d\mu = \nu(V) + s[x - \mu(V)] = \nu(V). \end{aligned}$$

بنابراین ماکزیمم موجود در تعریف $h(x)$ توسط F ماکزیمم می شود. $\square$

اثبات گزاره ۵.۲.۳: واضح است که $\|\mu - \nu\|_{TV} \leq \sup\{\int_{\Omega} f d\mu - \int_{\Omega} f d\nu\}$ (توابع مشخصه مجموعه ها را در نظر بگیرید). داریم $\sup\{\int_{\Omega} f d\mu - \int_{\Omega} f d\nu\} = \max h(x)$ که طبق تعریف h ، این ماکزیمم در یک نقطه اتخاذ می شود. بنابراین طبق لم قبل مجموعه A وجود دارد که $\nu(A) - \mu(A) = \sup\{\int_{\Omega} f d\mu - \int_{\Omega} f d\nu\}$ و به این ترتیب حکم ثابت می شود. $\square$

۹.۲.۳. لم: فرض کنید μ_1 و ν_1 به ترتیب توزیع های حاصل پس از یک گام از یک فرآیند مارکوف با

شروع از دو توزیع μ_0 و ν_0 باشند. در این صورت

$$\|\mu_1 - \nu_1\|_{TV} \leq \|\mu_0 - \nu_0\|_{TV}.$$

اثبات: برای هر مجموعه $A \subseteq \Omega$ داریم

$$\mu_1(A) - \nu_1(A) = \int_{\Omega} P_u(A) d\mu_0 - \int_{\Omega} P_u(A) d\nu_0.$$

$P_u(A)$ به عنوان تابعی از u اندازه پذیر است و $0 \leq P_u(A) \leq 1$. بنابراین طبق گزاره ۵.۲.۳

مقدار سمت راست حداکثر برابر با $\|\mu_0 - \nu_0\|_{TV}$ است. $\square$

۲.۳.ب. جفت سازی و فاصلهٔ مجموع تغییرات^۱

یک جفت سازی^۲ دو توزیع احتمال μ و ν ، یک جفت متغیر تصادفی (X, Y) است که توزیع های حاشیه ای X و Y ، به ترتیب μ و ν باشند. به عبارت دیگر $\mathbb{P}(X \in A) = \mu(A)$ و $\mathbb{P}(Y \in B) = \nu(B)$. توجه کنید که X و Y باید روی یک فضای احتمال تعریف شده باشند.

جفت سازی یک تکنیک عمومی است که در جاهای مختلف مورد استفاده قرار می گیرد. به عنوان مثال در فرآیندهای مارکوف با فضای حالت متناهی، می توان با استفاده از جفت سازی برای زمان آمیختگی کران بالا ارائه کرد [۵]. در این جا ما تنها به بیان ارتباط نزدیک بین جفت سازی و فاصلهٔ مجموع تغییرات می پردازیم.

۱.۰.۲.۳. مثال: اگر $\mu = \nu$ باشد، لااقل دو جفت سازی زیر برای آن ها وجود دارند:

X و Y را دو متغیر تصادفی مستقل با توزیع μ بگیرید.

X را یک متغیر تصادفی با توزیع μ بگیرید و قرار دهید $Y = X$.

برای هر دو توزیع احتمال دلخواه μ و ν ، همواره یک جفت سازی مستقل وجود دارد. اما اگر دو توزیع مساوی نباشند، آن گاه X و Y نمی توانند همه جا با هم مساوی باشند. گزارهٔ زیر می گوید که X و Y حداکثر چقدر می توانند با هم مساوی باشند.

۱.۱.۲.۳. گزاره:

$$\|\mu - \nu\|_{TV} = \inf\{\mathbb{P}(X \neq Y)\}$$

که در آن اینفیمم روی همهٔ جفت سازی های (X, Y) از μ و ν اتخاذ می شود. به علاوه، اینفیمم توسط یک جفت سازی اتخاذ می شود که آن را جفت سازی بهینه می نامیم.

^۱ از این زیر بخش فقط در فصل ۵ استفاده می شود.

^۲ Coupling

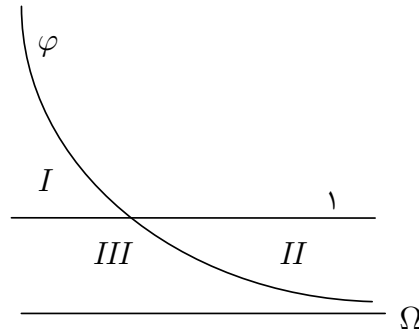
اثبات: برای هر جفت‌سازی مذکور و هر مجموعه $A \subseteq \Omega$ ، داریم

$$\mu(A) - \nu(A) = \mathbb{P}(X \in A) - \mathbb{P}(Y \in A) \leq \mathbb{P}(X \in A, Y \notin A) \leq \mathbb{P}(X \neq Y).$$

بنابراین $\|\mu - \nu\|_{TV} \leq \mathbb{P}(X \neq Y)$. حال کافی است یک جفت‌سازی بسازیم که

$$\|\mu - \nu\|_{TV} = \mathbb{P}(X \neq Y).$$

یافت که $\mu(U) = 0$ و روی مکمل U ، ν نسبت به μ چگالی احتمالی مثل φ داشته باشد.



شکل ۹

به صورت شهودی در شکل ۹ سه ناحیه ایجاد شده توسط نمودارهای تابع ثابت ۱ و φ را در نظر بگیرید. روش ما این طور است که یک نقطه تصادفی در اجتماع این سه ناحیه انتخاب می‌کنیم (با توزیع حاصل ضربی μ و اندازه لبگ). اگر این نقطه در ناحیه III بود، X و Y را مساوی با مؤلفه اول این نقطه می‌گیریم و در صورتی که در یکی از ناحیه‌های I یا II بود، طوری X و Y را انتخاب می‌کنیم که به ترتیب در تصویرنواحی I و II روی مؤلفه اول باشند (در این صورت X و Y لزوماً نابرابر خواهند بود). به طور دقیق‌تر، قرار دهید $A := \{u \in \Omega : \varphi(u) \leq 1\}$ و

$$p := \int_{\Omega} \min\{1, \varphi(u)\} du.$$

مشابه گزاره ۴.۲.۳، به دست می‌آوریم

$$p = 1 - \|\mu - \nu\|_{TV}.$$

یک سکه پرتاب کنید طوری که احتمال شیر آمدن آن برابر با p باشد. در صورتی که شیر آمد، X و

Y را برابر با هم از توزیع $\frac{1}{p} \min\{\mu, \nu\}$ انتخاب کنید و در صورتی که خط آمد، X را از توزیع

$\frac{1}{1-p}(\mu - \nu)|_A$ و Y را مستقلاً از توزیع $\frac{1}{1-p}(\nu - \mu)|_{\Omega \setminus A}$ انتخاب کنید. به سادگی دیده

می‌شود که (X, Y) یک جفت‌سازی برای μ و ν است و $P(X \neq Y) = \|\mu - \nu\|_{TV}$. $\square$

۲.۳. آمیختگی در فرآیند های مارکوف

در یک فرآیند مارکوف $\mathcal{M} = (\Omega, \mathcal{A}, P)$ ، P_μ^t را توزیع فرآیند در لحظه t با شروع از توزیع μ

می‌نامیم. همان‌طور که در ابتدای بخش ۲.۳ گفتیم، در یک فضای حالت متناهی برای هر فرآیند مارکوف

غیرمتناوب تحویل‌ناپذیر، یک توزیع پایای یکتای Q موجود است و برای هر توزیع اولیه μ ،

$\|P_\mu^t - Q\|_{TV}$ به صورت نمایی به صفر میل می‌کند. در یک فضای حالت کلی قضیه هریس حکم

مشابهی را (با فرض‌های متفاوت) بیان می‌کند. در این‌جا به حالت کلی قضیه هریس نیازی نداریم و تنها

به بیان یک نتیجه از آن اکتفا می‌کنیم.

۱۲.۲.۳. قضیه: اگر تابع چگالی احتمال گذر یک فرآیند مارکوف نسبت به یک توزیع پایای Q ،

همه‌جا از $\delta > 0$ بزرگتر باشد، آن‌گاه برای هر توزیع اولیه دلخواه μ داریم

$$\|P_\mu^t - Q\|_{TV} \leq (1 - \delta)^t.$$

معمولاً زمان آمیختگی^۱ را به شکل کمترین زمان t تعریف می‌کنند که برای هر توزیع اولیه μ

داشته باشیم $\|P_\mu^t - Q\|_{TV} < \epsilon$. با این تعریف، در صورت برقراری شرایط قضیه هریس می‌توان ثابت

کرد که بستگی زمان آمیختگی به ϵ لگاریتمی است. اما همان‌طور که در بخش ۲.۴ توضیح می‌دهیم، در

¹ Mixing time

مسئله مورد نظر ما $\|P_\mu^t - Q\|_{TV}$ بستگی زیادی به μ دارد. بنابراین زمان آمیختگی را بر حسب

توزیع اولیه و برای رده خاصی از آن‌ها در نظر خواهیم گرفت. به قضیه های ۲.۲.۴ و ۳.۲.۴ توجه کنید.

۱۳.۲.۳. **تعریف:** یک فرآیند مارکوف را تنبل^1 می‌نامیم هرگاه برای هر $x \in \Omega$ داشته باشیم

$$P_x(\{x\}) \geq \frac{1}{\mathfrak{p}}$$

ساخت که در هر گام با انداختن یک سکه به طور مستقل، با احتمال $\frac{1}{\mathfrak{p}}$ اصلاً حرکت نکنیم. به عبارت

$$\text{دیگر } P'_x := \frac{P_x + \delta_x}{\mathfrak{p}} \text{ که در آن } \delta_x \text{ توزیع اتمی روی } x \text{ است.}$$

در نسخه تنبل شده یک فرآیند مارکوف به طور متوسط نصف گام‌ها هدر می‌روند. بنابراین تنبل

کردن یک فرآیند مارکوف به نظر بی‌فایده می‌آید. اما در واقع این‌طور نیست. برای مثال در نسخه

غیر تنبل قدم‌زدن تصادفی روی یک شبکه (الگوریتم ۱.۲.۱) مشکل زوجیت مانع هم‌گرایی به توزیع پایا

می‌شود. اما این تنها فایده تنبل کردن یک فرآیند مارکوف نیست. در حالتی که فضای حالت متناهی

باشد، این کار باعث می‌شود که ماتریس احتمال گذر مثبت نیمه‌معین شود و این موضوع به تحلیل زمان

آمیختگی کمک می‌کند. در واقع وجود مقدار ویژه‌ای نزدیک به ۱ - باعث کند شدن هم‌گرایی می‌شود

[۵]. برای نسخه‌های تنبل شده فرآیندهای مارکوف می‌توان از قضیه‌های کلی که برای فرآیندهای تنبل

وجود دارند (برای مثال قضیه ۱۷.۲.۳) استفاده کرد. بنابراین در این پایان‌نامه از نسخه تنبل شده قدم‌زدن

تصادفی Hit-and-run استفاده می‌کنیم.

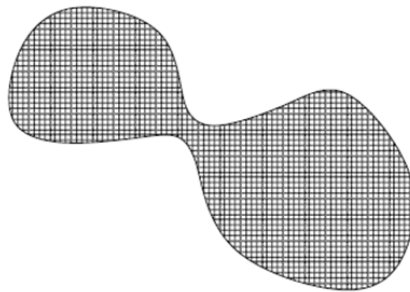
۲.۳. د. رسانش زنجیرهای مارکوف

در این بخش روشی برای کران دادن برای زمان آمیختگی بیان می‌کنیم. روش‌های متفاوتی برای این کار

وجود دارند اما روش‌هایی که در تمام الگوریتم‌هایی که در این پایان‌نامه ذکر می‌کنیم استفاده می‌شوند،

¹ Lazy

همگی بر اساس یک ایده استوارند: استفاده از *رسانش*. این روش را اولین بار سینکلر^۱ و جروم^۲ برای محاسبهٔ پرمونت^۳ یک ماتریس مطرح کردند. ایدهٔ اصلی آن، استفاده از مفهوم *تنگناها*^۴ است. تنگناها در یک فرآیند مارکوف ویژگی‌هایی هندسی هستند که زمان آمیختگی را کنترل می‌کنند. یک تنگنا دسترسی به قسمت‌هایی از Ω از برخی نقاط را سخت می‌کند و سرعت هم‌گرایی را محدود می‌کند. به‌عکس، اگر یک فرآیند مارکوف تنگنا نداشته باشد، می‌توان برای زمان آمیختگی آن کران بالا یافت. شکل زیر یک گراف به همراه یک تنگنای واضح را نشان می‌دهد. در زیر تنگناها را به صورت کمی بررسی می‌کنیم و از آن برای کران دادن برای زمان آمیختگی استفاده می‌کنیم.



شکل ۱۰

۱۴.۲.۳. تعریف: *شار/ارگودیک*^۵ هستهٔ مارکوف $\mathcal{M}$ با توزیع پایای Q را به صورت زیر تعریف

می‌کنیم

$$\Phi(A, B) := \int_A P_u(B) dQ(u)$$

و تعریف می‌کنیم

$$\Phi(A) := \Phi(A, A^c).$$

¹ Sinclair

² Jerrum

³ Permanent

⁴ Bottleneck

⁵ Ergodic flow

(برای هر $A, B \in \mathcal{A}$) $\Phi(A, B)$ احتمال این است که اگر از توزیع پایای Q شروع کنیم، داشته

باشیم $w \in A$ و $w \in B$.

به سادگی می توان دید که

$$\Phi(A, B) + \Phi(A, B^c) = Q(A),$$

$$\Phi(A, B) + \Phi(A^c, B) = Q(B)$$

و از این می توان نتیجه گرفت $\Phi(A) = \Phi(A^c)$. اثبات عکس این تساوی هم معتبر است؛ یعنی از

تساوی $\Phi(A) = \Phi(A^c)$ برای هر A ، پایا بودن توزیع Q نتیجه می شود.

اگر A یا A^c «کوچک» باشند، آنگاه $\Phi(A)$ نیز کوچک است زیرا

$\Phi(A) \leq \min\{Q(A), Q(A^c)\}$. اما کم بودن $\Phi(A)$ برای مجموعه های کوچک به معنای وجود

تنگنا نیست. در صورتی یک تنگنا ایجاد می شود که دو مجموعه A و A^c «بزرگ» باشند و $\Phi(A)$

کوچک باشد. این موضوع ایده تعریف زیر به عنوان معیاری برای سنجش تنگناها است.

۱۵.۲.۳. تعریف: رسانش^۱ هسته مارکوف $\mathcal{M}$ با توزیع پایای Q را به صورت زیر تعریف می کنیم

$$\Phi := \inf_{0 < Q(A) < \frac{1}{4}} \frac{\Phi(A)}{Q(A)}.$$

۱۶.۲.۳. نکته: همواره داریم $0 \leq \Phi \leq 1$. همچنین اگر فرآیند تنبل باشد، آنگاه $\Phi \leq \frac{1}{4}$. در واقع

رسانش نسخه تنبل شده یک فرآیند مارکوف نصف رسانش خود آن است. کوچک بودن Φ به معنای

وجود تنگنا است و بزرگ بودن آن به معنای عدم وجود تنگنا است.

هدف از این بخش اثبات قضیه زیر است.

¹ Conductance

۱۷.۲.۳. قضیه: اگر Q اتم نداشته باشد و فرآیند تنبل باشد، آن گاه به ازای

$$M := \sup_A Q_*(A)/Q(A)$$
 داریم

$$\|Q_n - Q\|_{T.V.} \leq \sqrt{\frac{M}{\gamma}} \left(1 - \frac{1}{\gamma} \Phi^2\right)^n.$$

۱۸.۲.۳. تعریف: طبق تعریف ۶.۲.۳، تعریف کنید

$$h_n(x) = h_{Q, Q_n}(x).$$

۱۹.۲.۳. لم: فرض کنید Q اتم نداشته باشد و $n \geq 0$. اگر $\frac{1}{\gamma} \leq x \leq 1$ آن گاه

$$h_{n+1}(x) \leq \frac{1}{\gamma} [h_n(x - \gamma x \Phi) + h_n(x + \gamma x \Phi)]$$

و اگر $0 \leq x \leq \frac{1}{\gamma}$ ، آن گاه

$$h_{n+1}(x) \leq \frac{1}{\gamma} [h_n(x - \gamma(1-x)\Phi) + h_n(x + \gamma(1-x)\Phi)].$$

اثبات: نامساوی اول را ثابت می کنیم. دومی هم به طور مشابه ثابت می شود. با توجه به این که Q اتم

ندارد، طبق لم ۸.۲.۳ مجموعه A موجود است که $Q(A) = x$ و

$h_{n+1}(x) = Q_{n+1}(A) - Q(A)$ به ازای $u \in \Omega$ تعریف کنید

$$f_1(u) := \begin{cases} \gamma P_u(A) - 1, & u \in A \\ 0, & u \notin A \end{cases}$$

و

$$f_\gamma(u) := \begin{cases} 1, & u \in A \\ \gamma P_u(A), & u \notin A \end{cases}.$$

همچنین قرار دهید $x_i := \int_{\Omega} f_i dQ$. به علت تنبیل بودن قدم زدن تصادفی، داریم $0 \leq f_i \leq 1$ و

$0 \leq x_i \leq 1$. همچنین چون Q پایا است، به دست می آوریم

$$x_{\gamma} + x_{\gamma} = \int_{\Omega} (g_{\gamma} + g_{\gamma}) dQ = 2 \int_{\Omega} P_u(A) dQ = 2Q(A) = 2x.$$

حال

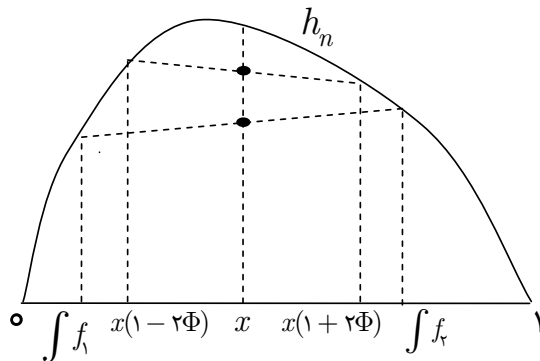
$$Q_{n+1}(A) = \int_{\Omega} P_u(A) dQ_n = \frac{1}{\gamma} \int_{\Omega} f_{\gamma} dQ_n + \frac{1}{\gamma} \int_{\Omega} f_{\gamma} dQ_n.$$

بنابراین

$$\begin{aligned} h_{n+1}(x) &= Q_{n+1}(A) - Q(A) = \frac{1}{\gamma} \left(\int_{\Omega} f_{\gamma} dQ_n - x_{\gamma} \right) + \frac{1}{\gamma} \left(\int_{\Omega} f_{\gamma} dQ_n - x_{\gamma} \right) \\ &\leq \frac{1}{\gamma} h_n(x_{\gamma}) + \frac{1}{\gamma} h_n(x_{\gamma}). \end{aligned}$$

از طرفی داریم

$$\begin{aligned} x - x_{\gamma} &= x_{\gamma} - x = 2 \int_{\Omega \setminus A} P_u(A) dQ + Q(A) - x \\ &= 2\Phi(A) \geq 2x\Phi. \end{aligned}$$



شکل ۱۱

پس طبق تقعر h_n به دست می آوریم

$$h_{n+1}(x) \leq \frac{1}{\varphi} h_n(x_1) + \frac{1}{\varphi} h_n(x_{\varphi}) \leq \frac{1}{\varphi} h_n(x - \varphi x \Phi) + \frac{1}{\varphi} h_n(x + \varphi x \Phi)$$

و لم ثابت می شود. $\square$

۲.۲.۳. لم: فرض کنید عدد حقیقی c موجود باشد که

$$h_{\circ}(x) \leq c \cdot \min\{\sqrt{x}, \sqrt{1-x}\}.$$

در این صورت برای هر $n \geq \circ$ داریم

$$h_n(x) \leq c \cdot \min\{\sqrt{x}, \sqrt{1-x}\} \left(1 - \frac{1}{\varphi} \Phi^{\varphi}\right)^n.$$

اثبات: از استقرا روی n استفاده می کنیم. حالت $n = \circ$ همان فرض لم است. فرض کنید حکم برای

n درست باشد. از لم قبل استفاده می کنیم. ابتدا فرض کنید $x \leq \frac{1}{\varphi}$. پس

$$\begin{aligned} h_{n+1}(x) &\leq \frac{1}{\varphi} [h_n(x - \varphi x \Phi) + h_n(x + \varphi x \Phi)] \\ &\leq \frac{c}{\varphi} [\sqrt{x - \varphi x \Phi} \left(1 - \frac{1}{\varphi} \Phi^{\varphi}\right)^n + \sqrt{x + \varphi x \Phi} \left(1 - \frac{1}{\varphi} \Phi^{\varphi}\right)^n] \\ &\leq \frac{c}{\varphi} \sqrt{x} (\sqrt{1 - \varphi \Phi} + \sqrt{1 + \varphi \Phi}) \left(1 - \frac{1}{\varphi} \Phi^{\varphi}\right)^n \\ &\leq c \sqrt{x} \left(1 - \frac{1}{\varphi} \Phi^{\varphi}\right)^{n+1}. \end{aligned}$$

نامساوی آخر به سادگی با به توان دو رساندن به دست می آید. برای $x \geq \frac{1}{\varphi}$ هم حکم به طور مشابه

ثابت می شود. $\square$

اثبات قضیه ۱۷.۲.۳: از تعریف M نتیجه می شود که $h_{\circ}(x) \leq Mx$. از طرفی $h_{\circ}(x) \leq 1 - x$.

بنابراین $h_{\circ}(x) \leq \sqrt{M} \min\{\sqrt{x}, \sqrt{1-x}\}$ (توجه کنید که $M \geq 1$). پس طبق لم قبل

و با توجه به این که
$$h_n(x) \leq \sqrt{M} \min\{\sqrt{x}, \sqrt{1-x}\} \left(1 - \frac{1}{r} \Phi^r\right)^n$$

□ حکم ثابت می شود. $\|Q_n - Q\|_{TV} = \max h(x)$

۴. تولید یک نقطه تصادفی در یک شکل

محدب

همان طور که در مقدمه گفته شد، برای تولید یک نقطه تصادفی در یک شکل محدب از روش مونت کارلو استفاده می‌کنیم. در این پایان‌نامه از بین سه قدم‌زدن تصادفی که در بخش ۲.۱ معرفی کردیم، تنها از نسخه تنبل‌شده قدم‌زدن تصادفی Hit-and-run استفاده می‌کنیم. برای ارائه یک الگوریتم، باید برای زمان آمیختگی این قدم‌زدن تصادفی یک کران بالا ارائه دهیم. در بخش ۱.۴ مراحل و ایده‌های اصلی این کار را به صورت خلاصه توضیح می‌دهیم. قضیه‌های اصلی که هدف این فصل هستند در بخش ۲.۴ معرفی می‌شوند. این قضایا را در چند گام در بخش‌های ۳.۴ تا ۷.۴ ثابت خواهیم کرد. در نهایت در بخش ۸.۴ در مورد تولید یک دنباله از نقاط تصادفی در یک شکل محدب صحبت می‌کنیم.

۱.۴. مراحل تحلیل زمان آمیختگی قدم‌زدن تصادفی Hit-

and-run

برای این کار برای رسانش این قدم‌زدن تصادفی یک کران پایین ارائه می‌کنیم و از قضیه ۱۷.۲.۳ استفاده می‌کنیم. فرض کنید K را به دو بخش S_1 و S_2 افراز کرده باشیم. S'_i را مجموعه نقاط «عمیقاً درون S_i » بگیرید؛ یعنی نقاطی که با شروع از آن‌ها، احتمال قدم‌زدن به مجموعه دیگر کم است^۱:

$$S'_i = \{x \in S_i : P_x(S_i^c) < \frac{1}{6400}\}, \quad i = 1, 2. \quad (1.4)$$

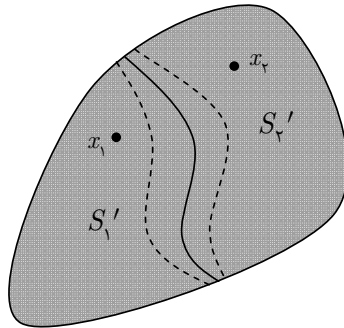
^۱ ثابتی که در این تعریف استفاده می‌شود تأثیری در مرتبه کران ما ندارد.

توجه کنید که اگر $x_i \in S'_i$ ($i = 1, 2$)، آن گاه داریم

$$P_{x_1}(S_1) - P_{x_2}(S_1) > 1 - \frac{1}{6400} - \frac{1}{6400} = 1 - \frac{1}{3200}.$$

$$\|P_{x_1} - P_{x_2}\|_{TV} > 1 - \frac{1}{3200}.$$

بنابراین



شکل ۱۲

یک گام از اثبات قضیه، ارائه گزاره‌ای مناسب به شکل زیر است.

۱.۱.۴. گزاره (طرح کلی): اگر دو نقطه x_1 و x_2 «به هم نزدیک باشند»، آن گاه P_{x_1} و P_{x_2} نیز

«به هم نزدیک هستند».

در بخش ۳.۴ تعبیری از «نزدیک بودن» را بیان می‌کنیم و چنین گزاره‌ای را ثابت می‌کنیم

(گزاره ۴.۳.۴). بنابراین نتیجه می‌گیریم که نقاط S'_1 و S'_2 از هم دور هستند. گام دیگر اثبات، ارائه

گزاره‌ای مناسب به شکل زیر است که در بخش ۴.۴ انجام می‌شود (قضیه ۳.۴.۴).

۲.۱.۴. گزاره (طرح کلی): اگر دو زیرمجموعه از K از هم دور باشند، آن گاه حجم آن‌ها نمی‌تواند

«زیاد» باشد.

در نتیجه حجم مجموعه نقاطی که احتمال قدم زدن از آن‌ها به بخش دیگر از $\frac{1}{6400}$ کمتر است، نمی‌تواند به معنایی که بیان خواهیم کرد زیاد باشد و به این ترتیب در بخش ۶.۴ برای $\Phi(S_1, S_2)$ کران پایینی به دست می‌آوریم.

۲.۴. قضایای اصلی

اسمیت در [۲] نشان داد که برای هر شکل کران دار داده شده، توزیع قدم زدن تصادفی Hit-and-run به صورت نمایی به توزیع یک‌نواخت میل می‌کند (حتی برای اشکال بازی که لزوماً محدب نیستند). این موضوع پس از محاسبه تابع چگالی احتمال گذر قدم زدن تصادفی Hit-and-run در لم ۵.۳.۴، از نکته ۷.۱.۳ و لم‌های ۶.۱.۳ و ۱۲.۲.۳ نتیجه می‌شود. اما کرانی که با این روش برای زمان آمیختگی به دست می‌آید بر حسب بعد نمایی است. هدف این فصل ارائه یک کران چند جمله‌ای برای زمان آمیختگی این قدم زدن تصادفی است.

یک مشکل که در همه قدم‌زدن‌های تصادفی ما وجود دارد این است که ممکن است نزدیک یک گوشه K «گیر کنیم». برای مثال در قدم‌زدن با گام‌های کروی، اگر از نزدیک یک گوشه مکعب n -بعدی شروع کنیم، احتمال جابه‌جا شدن در یک گام نزدیک به 2^{-n} است. بنابراین بیش‌تر گام‌ها هدر می‌روند و این موضوع باعث می‌شود که برای آمیختگی به زمان نمایی نیاز باشد. به علاوه، حتی با شروع از یک نقطه نه‌چندان مرزی هم زمان آمیختگی می‌تواند نمایی باشد. تنها راه شناخته شده برای اجتناب از این مشکل، شروع از یک توزیع گرم است:

۱.۲.۴. تعریف: فرض کنید Q توزیع پایای یک فرآیند مارکوف M باشد. توزیع آغازی Q_0 برای M را یک توزیع M -گرم^۱ می‌نامند هرگاه برای هر مجموعه A داشته باشیم

$$Q_0(A) \leq M.Q(A)$$

طبق قضیه رادون-نیکودایم، این موضوع معادل با این است که Q_0 نسبت به Q پیوسته باشد و

$$dQ_0/dQ \leq M$$

با این که شروع از یک توزیع گرم به نظر محدودیت زیادی است، اما غالباً این شرط در الگوریتم‌های خاص برآورده می‌شود؛ برای مثال هنگام تولید تعداد زیادی نقطه تصادفی و یا هنگام استفاده از روش مونت کارلو چندفازی (به بخش ۸.۴ و فصل ۵ مراجعه کنید).

در قدم‌زدن Hit-and-run هم طبق نتیجه ب.۱۸، طول وتری که از یک نقطه نزدیک گوشه مکعب خارج می‌شود به احتمال زیاد کوچک است. بنابراین مقدار جابجایی کم خواهد بود. اما همین جابجایی کوچک باعث می‌شود که در زمان چند جمله‌ای از گوشه خارج شویم. کرانی که برای زمان آمیختگی به دست خواهیم آورد، بر حسب فاصله نقطه شروع تا مرز شکل، لگاریتمی است. همچنین برای شروع از رده‌های خاصی از توزیع‌ها (برای مثال توزیع‌های گرم و یا توزیع‌های با نرم L^2 محدود) نیز کران‌هایی ارائه شده است که در کاربردهای مختلف استفاده شده‌اند.

مشکل دیگری که باعث کندی قدم‌زدن‌های تصادفی معرفی شده در بخش ۲.۱ می‌شود این است که ممکن است شکل مورد نظر دراز و باریک باشد. برای مثال مکعب مستطیل $[0, \epsilon]^{n-1} \times [0, 1]$ را در نظر بگیرید. اگر ϵ کوچک باشد، برای این که در قدم‌زدن تصادفی Hit-and-run از نقطه‌ای در قسمت $[0, \epsilon]^{n-1} \times [0, \frac{1}{p}]$ به نیمه دیگر گام برداریم، باید جهت انتخاب شده تقریباً موازی محور اول مختصات باشد که احتمال این اتفاق می‌تواند به دلخواه کم باشد. برای کنترل این موضوع باید میزان «گرد بودن»

¹ M -wam distribution

شکل را در نظر بگیریم. مطابق تعریف ۱.۱.۲، فرض می‌کنیم K شامل گوی به شعاع r حول مبدأ است و هم‌چنین کاملاً در گوی به شعاع R حول مبدأ قرار گرفته است. کرانی که در این جا ارائه می‌دهیم به

$$\frac{R}{r} \text{ نیز بستگی دارد.}$$

همان طور که گفتیم، زمان آمیختگی قدم‌زدن Hit-and-run بستگی زیادی به توزیع اولیه دارد. ابتدا قضیه زیر را ثابت خواهیم کرد که شرطی مشابه گرم‌بودن برای توزیع اولیه فرض می‌کند. از این قضیه در بخش ۸.۴ و فصل ۵ استفاده می‌شود.

۲.۲.۴. قضیه: فرض کنید $K \subseteq \mathbb{R}^n$ شکل محدبی باشد که شامل یک گوی به شعاع r حول مبدأ

است و در یک گوی به شعاع R حول مبدأ قرار گرفته است. فرض کنید σ توزیع اولیه و σ^m توزیع قدم‌زدن Hit-and-run تنبل‌شده در K بعد از m گام باشد. $\epsilon > 0$ و M را اعدادی

دلخواه بگیرید و فرض کنید به‌جز در یک مجموعه $S \subseteq K$ که $\sigma(S) \leq \frac{\epsilon}{4}$ داشته باشیم

$$d\sigma/d\pi \leq M \text{ (که در آن } \pi \text{ توزیع یک‌نواخت روی } K \text{ است). در این صورت به ازای}$$

$$m > 4^{\frac{n}{2}} \frac{n^{\frac{1}{2}} R^{\frac{n}{2}}}{r^{\frac{n}{2}}} \ln \frac{\sqrt{2} M}{\epsilon}$$

$$\|\sigma^m - \pi\|_{TV} < \epsilon. \text{ خواهیم داشت}$$

این قضیه نسبت به قضیه‌هایی که قبل از آن ثابت شده بودند این مزیت را دارد که بستگی کران داده شده به M و ϵ را از چندجمله‌ای (که برای قدم‌زدن با گام‌های کروی اجتناب‌ناپذیر است) به

لگاریتمی کاهش می‌دهد در حالی که برحسب n ، R و r بهترین کران موجود را حفظ می‌کند.

برای کنترل زمان آمیختگی با شروع از یک نقطه ثابت دلخواه، ابتدا یک گام از قدم‌زدن تصادفی را

برمی‌داریم و برای توزیع به دست آمده از قضیه قبل استفاده می‌کنیم.

۳.۲.۴. قضیه: با فرض های قضیه قبل، اگر از یک نقطه ثابت به فاصله d از مرز K شروع کنیم،

آن گاه به ازای

$$m > 2^{5\gamma} \frac{n^{\gamma} R^{\gamma}}{r^{\gamma}} \ln \frac{2R}{d\epsilon}$$

خواهیم داشت $\|\sigma^m - \pi\|_{TV} < \epsilon$.

۳.۴. فاصله هندسی و فاصله احتمالاتی

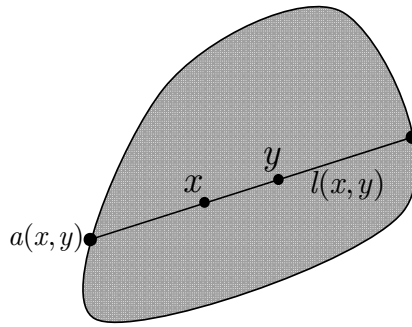
هدف این بخش اثبات گزاره ای به شکل ۱.۱.۴ است. یعنی اگر دو نقطه به اندازه مشخصی نزدیک هم باشند، آن گاه توزیع هایی که با یک بار قدم زدن با شروع از آنها به دست می آیند نیز به هم نزدیک هستند. در این جا نزدیک بودن دو توزیع را با معیار فاصله مجموع تغییرات می سنجیم.

همان طور که در بخش ۲.۴ اشاره کردیم، اگر دو نقطه نزدیک به یک گوشه شکل باشند، آن گاه طول گامی که از آن ها برداشته می شود به احتمال زیاد کوچک است. این در حالی است که برای نقاط دور از مرز چنین نیست. بنابراین فاصله مجموع تغییرات توزیع هایی که با یک بار قدم زدن از دو نقطه حاصل می شوند، علاوه بر فاصله اقلیدسی آن ها به فاصله آن ها از مرز و گوشه های شکل نیز بستگی زیادی دارد. این موضوع را در تعریف های زیر فرمول بندی می کنیم.

۱.۳.۴. تعریف: برای دو نقطه $x, y \in K$ را وتر گذرنده از این دو نقطه و $l(x, y)$ را طول

این وتر تعریف کنید. $a(x, y)$ را یکی از نقاط انتهایی $L(x, y)$ بگیرید که به x نزدیک تر است.

بنابراین انتهای دیگر $L(x, y)$ همان $a(y, x)$ است.



شکل ۱۳

۲.۳.۴. تعریف: برای چهار نقطه هم خط a, b, c, d ، نسبت ناهمساز^۱ آنها را به صورت زیر تعریف

می کنیم

$$(a, b; c, d) := \frac{(a - c) \cdot (b - d)}{(a - d) \cdot (b - c)}.$$

توجه کنید که چون نقاط هم خط هستند، به جای ضرب داخلی می توانیم فقط اندازه و علامت این بردارها نسبت به هم را در نظر بگیریم. برای دو نقطه p و q در یک شکل محدب K ، اگر

$r := a(p, q)$ و $s := a(q, p)$ ، تعریف می کنیم

$$d_K(p, q) := -(p, s; q, r) = \frac{|p - q| \cdot |r - s|}{|p - r| \cdot |q - s|}.$$

به وضوح داریم $d_K(p, q) = d_K(q, p)$. همچنین اگر p به q میل کند، $d_K(p, q)$ به صفر میل

می کند و اگر یکی از p و q به مرز K نزدیک شود، $d_K(p, q)$ به بینهایت میل می کند. به علاوه

چون $|p - r| + |q - s| \leq |r - s|$ ، به دست می آوریم که

$$d_K(p, q) \geq \frac{4|p - q|}{|r - s|} \geq \frac{4|p - q|}{\text{diam}(K)}.$$

^۱ Cross ratio

در واقع d_K یک متریک نیست زیرا در نامساوی مثلث صدق نمی کند. برای به دست آوردن متریک می توانیم متریک هیلبرت یعنی $\log(1 + d_K)$ را در نظر بگیریم.^۱ اما کار با خود d_K کافی است.

۳.۳.۴. **تعریف:** برای یک نقطه $x \in K$ ، $F(x)$ را برابر با عددی تعریف کنید که

$$P(|X_1 - X_*| \leq F(x) | X_* = x) = \frac{1}{\lambda}.$$

$F(x)$ در واقع هشتک^۲ طول یک گام با شروع از x را نشان می دهد.

حال آماده ایم که گزاره اصلی این بخش را بیان کنیم.

۴.۳.۴. **گزاره:** فرض کنید $u, v \in K$ طوری باشند که

$$|u - v| < \frac{2}{\sqrt{n}} \max\{F(u), F(v)\} \text{ و } d_K(u, v) < \frac{1}{16}$$

در این صورت

$$\|P_u - P_v\|_{TV} < 1 - \frac{1}{3200}.$$

برای اثبات این گزاره، به لم زیر احتیاج داریم.

۵.۳.۴. **لم:** تابع چگالی احتمال گذر الگوریتم Hit-and-run نسبت به اندازه لبگ با شروع از نقطه

$u \in K$ برابر است با

$$f_u(x) = \frac{2}{n\pi_n} \cdot \frac{1}{l(u, x)|x - u|^{n-1}}$$

که در آن π_n حجم کره واحد است.

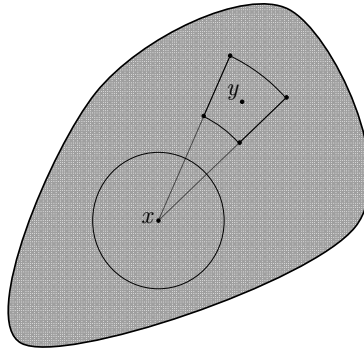
^۱ این متریک مشابه متریک هذلولی روی گوی واحد در مدل کلاین است

^۲ Octile

اثبات (طرح کلی): A را یک مجموعه بسیار کوچک^۱ در مختصات کروی با مرکز u بگیرید طوری که $l(u, x)$ در A تقریباً مقدار ثابت l باشد. به عبارت دیگر

$$A := \{x : |x - u| \in [t, t + dt], \frac{x - u}{|x - u|} \in dS\}$$

که در آن dS یک ناحیه بسیار کوچک روی کره واحد است.



شکل ۱۴

حال احتمال اینکه جهت تصادفی انتخاب شده در dS واقع شود برابر با $\frac{\gamma dS}{n\pi_n}$ است و پس از انتخاب

جهت، احتمال اینکه فاصله تا u در بازه $[t, t + dt]$ بیافتد تقریباً برابر با $\frac{dt}{l}$ است. پس

$$P_u(A) \approx \frac{\gamma dS dt}{nl\pi_n} \approx \frac{\gamma}{n\pi_n} \cdot \frac{vol(A)}{l \cdot t^{n-1}}$$

و حکم ثابت می شود. $\square$

اثبات گزاره ۴.۳.۴: فرض کنید $F(u) \geq F(v)$. برای یک زیرمجموعه دلخواه $A \subseteq K$ باید

ثابت کنیم

¹ Infinitesimal

$$P_u(A) - P_v(A) < 1 - \frac{1}{3200}.$$

طبق لم ۵.۳.۴، تابع چگالی P_v برابر است با

$$P_v(dx) = \frac{2}{n\pi_n} \cdot \frac{1}{l(v, x)|x - v|^{n-1}} dx.$$

برای این که این مقدار را به تابع چگالی P_u مربوط کنیم، باید برای $l(v, x)$ و $|x - v|$ کران‌هایی بر حسب $l(u, x)$ و $|x - u|$ ارائه کنیم. اما این کار به جز برای x هایی که به اندازه کافی از u, v دور باشند امکان پذیر نیست. به همین خاطر زیرمجموعه‌های زیر از A را در نظر می‌گیریم.

بخش‌هایی از A که به u نزدیک‌اند:

$$A_1 := \{x \in A : |x - u| < F(u)\},$$

$$A_2 := \{x \in A : |x - u| < \frac{1}{3} |u - a(u, x)|\},$$

$$A_3 := \{x \in A : |x - u| < \frac{1}{5} |u - a(x, u)|\}.$$

بخش دوم و سوم نقاطی از A هستند که نسبت به وتر $L(u, x)$ به u نزدیک‌اند.

بخش چهارم نقاطی هستند که پاره خط ux از تعامد با خط uv فاصله دارد:

$$A_4 := \{x \in A : |(x - u) \cdot (u - v)| > \frac{2}{\sqrt{n}} |x - u| \cdot |u - v|\}.$$

یعنی قدر مطلق کسینوس زاویه $\hat{x}uw$ بیش‌تر از $\frac{2}{\sqrt{n}}$ باشد. توجه کنید که نقاط نزدیک به v در

این بخش قرار می‌گیرند. در نهایت بخش پنجم $S := A \setminus (A_1 \cup A_2 \cup A_3 \cup A_4)$ است.

نشان می‌دهیم

$$P_u(A_1 \cup A_2 \cup A_3 \cup A_4) \leq \frac{4}{5},$$

$$P_v(S) \geq \frac{1}{6} P_u(S).$$

در این صورت خواهیم داشت

$$\begin{aligned} P_u(A) - P_v(A) &\leq P_u(A) - P_v(S) < P_u(A) - \frac{1}{6} P_u(S) \\ &\leq P_u(A) - \frac{1}{6} (P_u(A) - \frac{4}{5}) \leq 1 - \frac{1}{32}. \end{aligned}$$

و حکم ثابت خواهد شد.

x را گام تصادفی بعدی با شروع از u بگیرید. طبق تعریف $F(u)$ احتمال این که $x \in A_1$ برابر با

$\frac{1}{8}$ است. برای هر خط دلخواه l گذرنده از u ، اگر طول دو قطعه ایجاد شده روی آن خط را a و b

بنامیم، آن گاه $l \cap (A_1 \cup A_2)$ یک پاره خط با طول

$$\min\{a, \max\{\frac{a}{5}, \frac{b}{3}\}\} + \min\{b, \max\{\frac{b}{5}, \frac{a}{3}\}\}$$

است که با یک حالت گیری ساده دیده می شود که این مقدار حداکثر برابر با $\frac{2}{5}(a+b)$ است. پس

احتمال این که $x \in (A_1 \cup A_2)$ حداکثر برابر با $\frac{2}{5}$ است.

هم چنین برای این که $x \in A_3$ ، لازم و کافی است که خط تصادفی انتخاب شده گذرنده از u طوری

باشد که کسینوس زاویه آن با خط uv بیش تر از $\frac{2}{\sqrt{n}}$ باشد و طبق لم ب.۱۸ در ضمیمه ب. این مقدار

کمتر از $\frac{11}{4}$ است. در نتیجه

$$\mathbb{P}(x \in (A_{\lambda} \cup A_{\mu} \cup A_{\nu} \cup A_{\epsilon})) < \frac{1}{\lambda} + \frac{2}{\delta} + \frac{11}{\epsilon} = \frac{4}{\delta}$$

برای اثبات نامساوی دوم، فرض کنید $x \in S$. چون $x \notin A_{\lambda} \cup A_{\mu} \cup A_{\nu}$ ، می‌دانیم

$$|x - u| \geq F(u) \geq \frac{\sqrt{n}}{2} |u - v|,$$

$$|(x - u) \cdot (u - v)| \leq \frac{2}{\sqrt{n}} |x - u| \cdot |u - v|,$$

$$|x - u| \geq \frac{1}{2} |x - a(u, x)|,$$

$$|x - u| \geq \frac{1}{2} |x - a(x, u)|.$$

با استفاده از این روابط به دست می‌آوریم

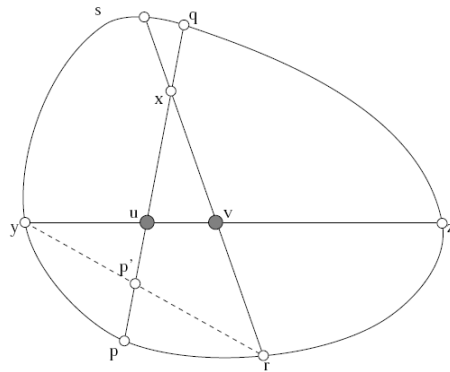
$$\begin{aligned} |x - v|^2 &= |x - u|^2 + |u - v|^2 + 2(x - u) \cdot (u - v) \\ &\leq |x - u|^2 + |u - v|^2 + \frac{4}{\sqrt{n}} |x - u| \cdot |u - v| \\ &\leq |x - u|^2 + \frac{4}{n} |x - u|^2 + \frac{\lambda}{n} |x - u|^2 \\ &= \left(1 + \frac{12}{n}\right) |x - u|^2. \end{aligned}$$

پس

$$|x - v| \leq \left(1 + \frac{6}{n}\right) |x - u|. \quad (2.4)$$

برای تقریب $l(v, x)$ ، مطابق شکل ۱۵ از نمادگذاری‌های زیر استفاده می‌کنیم. قرار دهید

$$\begin{aligned} r &:= a(v, x) & q &:= a(x, u) & p &:= a(u, x) & z &:= a(v, u) & y &:= a(u, v) \\ s &:= a(x, v) \end{aligned}$$



شکل ۱۵

طبق تحدب K و این که همگی این نقاط در یک صفحهٔ دو-بعدی قرار دارند، پاره خط‌های up و

ry یک‌دیگر را در نقطه‌ای مثل $p' \in K$ قطع می‌کنند. طبق قضیهٔ منلائوس (قضیهٔ ب.۱)

$$\frac{|x-r|}{|v-r|} = \frac{|u-y|}{|v-y|} \cdot \frac{|x-p'|}{|u-p'|}.$$

اما داریم

$$\frac{|u-y|}{|v-y|} = 1 - \frac{|u-v|}{|v-y|} > 1 - d_K(u, v).$$

از این دو رابطه به دست می‌آوریم

$$\begin{aligned} \frac{|x-v|}{|v-r|} &= \frac{|x-r|}{|v-r|} - 1 \geq (1 - d_K(u, v)) \frac{|x-p'|}{|u-p'|} - 1 \\ &= \frac{|x-u|}{|u-p'|} (1 - d_K(u, v)) \frac{|x-p'|}{|x-u|} \\ &> \frac{3}{4} \frac{|x-u|}{|u-p'|}. \end{aligned}$$

در نامساوی آخر از این که $|x-p'| \leq |x-p| < 4|x-u|$ و این که $d_K(u, v) \leq \frac{1}{16}$

استفاده کردیم. بنابراین

$$|v - r| < \frac{4}{3} \cdot \frac{|x - v|}{|x - u|} \cdot |u - p'| \leq \frac{4}{3} \cdot \frac{|x - v|}{|x - u|} \cdot |u - p|.$$

به طور مشابه با در نظر گرفتن تقاطع پاره خط‌های zs و uq به دست می‌آوریم:

$$|v - s| < \frac{4}{3} \cdot \frac{|x - v|}{|x - u|} \cdot |u - q|.$$

بنابراین

$$l(x, v) < \frac{4}{3} \cdot \frac{|x - v|}{|x - u|} \cdot l(x, u).$$

با استفاده از این کران و کران (۲.۴) می‌توانیم $P_v(S)$ را به نحو زیر تقریب بزنیم

$$\begin{aligned} P_v(S) &= \frac{2}{n\pi_n} \int_S \frac{1}{l(x, v)|x - v|^{n-1}} dx \\ &\geq \frac{3}{2n\pi_n} \int_S \frac{|x - u|}{l(x, u)|x - v|^n} dx \\ &\geq \frac{3}{2n\pi_n} \left(1 + \frac{\epsilon}{n}\right)^{-n} \int_S \frac{1}{l(u, x)|x - u|^{n-1}} dx \\ &\geq \frac{3}{4e^\epsilon} P_u(S) > \frac{1}{600} P_u(S) \end{aligned}$$

و گزاره ثابت می‌شود. $\square$

۴.۴. یک نابرابری محیطی

هدف این بخش، اثبات گزاره‌ای مناسب به شکل ۲.۱.۴ است. یعنی اگر دو زیرمجموعه از K از هم دور باشند، آن‌گاه حجم آن‌ها نمی‌تواند زیاد باشد. به چنین گزاره‌هایی، نابرابری‌های محیطی می‌گویند. مسئله کلاسیک نابرابری محیطی یافتن رویه‌ای با کمترین مساحت است که ناحیه‌ای با حجم ثابت را احاطه می‌کند. نوع دیگر مسائل نابرابری محیطی را اولین بار دایر، فریز و کنان برای تخمین رسانش قدم‌زدن تصادفی روی یک شبکه گسسته (الگوریتم ۱.۲.۱) مطرح کردند [۱]. هدف آن‌ها یافتن کران بالایی برای مساحت رویه‌هایی که K را به دو قسمت تقسیم می‌کنند بر حسب حجم آن دو قسمت بود. به صورت دقیق‌تر، یافتن عددی مثل ψ که برای هر مجموعه اندازه‌پذیر $S \subset K$ که $\partial S \cap K$ اندازه مینکوفسکی $(n-1)$ -بعدی داشته باشد، داشته باشیم

$$vol_{n-1}(\partial S \cap K) \geq \psi \cdot \min\{vol(S_1), vol(S_2)\}.$$

آن‌ها حدس زدند که $\psi \geq \frac{1}{D}$ که D قطر K است. این موضوع توسط لواژ و سیمونوویتز ثابت

شد [۶]. بهترین کران ممکن برای مسئله دوم بر حسب قطر $\frac{2}{D}$ است که توسط دایر و فریز ثابت شد [۷].

بعد از [۱] نیز همه روش‌هایی که برای نمونه‌گیری تصادفی روی یک شکل محدب و محاسبه حجم بیان شده‌اند، از نسخه‌های مختلفی از نابرابری‌های محیطی استفاده کرده‌اند ([۸] و [۶] و [۳] و [۹]). فرم کلی بیش‌تر نابرابری‌های محیطی استفاده شده به صورت زیر است:

۱.۴.۴. مسئله نابرابری محیطی: فرض کنید $d : K \times K \rightarrow \mathbb{R}^{\geq 0}$ یک تابع متقارن و $\epsilon > 0$

عددی حقیقی باشد. عددی مثل ψ بیابید که هرگاه K را به سه زیرمجموعه S_1, S_2 و S_3 افراز

کنیم به طوری که $\epsilon > d(S_1, S_2) := \inf\{d(x_1, x_2) : x_1 \in S_1, x_2 \in S_2\}$ آن گاه

$$vol(S_3) \geq \psi \frac{vol(S_1) \cdot vol(S_2)}{vol(K)}.$$

این شرط غالباً با شرط زیر جایگزین می شود

$$vol(S_3) \geq \psi \cdot \min\{vol(S_1), vol(S_2)\}.$$

با توجه به این که مقادیر سمت راست این دو نامساوی با یکدیگر نسبت حداکثر ۲ دارند، این دو

نامساوی برای ما تفاوت چندانی ندارند.

۲.۴.۴. نکته: دو نامساوی بالا را می توان به صورت زیر نوشت

$$\pi(S_3) \geq \psi \pi(S_1) \pi(S_2),$$

$$\pi(S_3) \geq \psi \cdot \min\{\pi(S_1), \pi(S_2)\}$$

که در آن π توزیع یک نواخت روی K است.

به عنوان مثال، مسئله ای که در بالا گفتیم حالت خاصی از این مسئله است. کافی است d را همان

متریک اقلیدسی و S_3 را ϵ -همسایگی رویه مورد نظر بگیرد و آن را از S_1 و S_2 کم کنید. سپس ϵ را

به صفر میل دهید.

در گزاره ۴.۳.۴، منظور ما از «نزدیک بودن» نقاط مشخص شده است. بنابراین در نابرابری محیطی

مورد نیاز ما باید از این تعبیر نزدیک بودن استفاده کنیم.

۳.۴.۴. قضیه: فرض کنید K یک شکل محدب در $\mathbb{R}^n$ و $h: K \rightarrow \mathbb{R}^+$ تابعی دلخواه باشد.

فرض کنید $S_\gamma, S_\gamma, S_\gamma$ افزای از K باشد به طوری که برای هر دو نقطه $u \in S_\gamma$ و $v \in S_\gamma$ و

هر نقطه x روی وتر K که از u و v می‌گذرد، داشته باشیم

$$h(x) \leq \min\left(\frac{1}{6}, \frac{1}{3} d_K(u, v)\right).$$

در این صورت

$$\text{vol}(S_\gamma) \cdot \text{vol}(K) \geq \int_K h(x) dx \min\{\text{vol}(S_\gamma), \text{vol}(S_\gamma)\}.$$

تابع h را در فصل ۶.۴ در اثبات قضیه ۱.۶.۴ معرفی می‌کنیم و نشان می‌دهیم که اگر S_γ و S_γ را

مشابه رابطه (۱.۴) در بخش ۱.۴ بگیریم، در فرض این قضیه صدق می‌کنند.

اثبات این قضیه بر اساس یک لم موضعی سازی است. لم‌های موضعی سازی را لواژ و سیمونوویتز^۱ مطرح

کردند [۸] که ایده هوشمندانه‌ای برای تبدیل مسائل انتگرالی چندبعدی به مسائلی یک‌بعدی است.

به خاطر این که این بخش طولانی نشود، معرفی و اثبات لم‌های موضعی سازی را به ضمیمه ۸. موکول

می‌کنیم.

اثبات: می‌توانیم فرض کنیم $\text{vol}(S_\gamma) \leq \text{vol}(S_\gamma)$. فرض کنید حکم درست نباشد. در این صورت

$$A \leq \frac{1}{2} \text{ وجود دارد که}$$

$$\text{vol}(K) = \frac{1}{A} \text{vol}(S_\gamma)$$

۹

$$\text{vol}(S_\gamma) < A \int_K h(x) dx.$$

¹ Simonovits

می‌توانیم فرض کنیم K باز است و S_1 و S_2 در K بسته‌اند. حال دو تابع $\chi_K - \frac{1}{A}\chi_{S_1}$ و $\chi_{S_2} - Ah$ از پایین پیوسته‌اند.^۱ با این که تابع اول پیوسته نیست، اما مشابه اثبات لم ۲.۱، برای هر $\epsilon > 0$ دو نقطه $a, b \in K$ و تابع خطی $l : [0, 1] \rightarrow \mathbb{R}^+$ با این خاصیت یافت می‌شوند که اگر

تعریف کنیم

$$G(t) := h(ta + (1-t)b) \text{ و } F(t) := l(t)^{n-1}$$

و

$$J_i := \{t \in [0, 1] : ta + (1-t)b \in S_i\}. \quad (i = 1, 2, 3)$$

آن گاه

$$\left| \int_0^1 F(t)dt - \frac{1}{A} \int_{J_1} F(t)dt \right| < \epsilon \int_0^1 F(t)dt \quad (3.4)$$

و

$$\int_{J_2} F(t)dt < A \int_0^1 G(t)F(t)dt. \quad (4.4)$$

قرار دهید $T := \int_0^1 F(t)dt$ و فرض کنید M کوچک‌ترین کران بالای h روی وتر گذرنده از a و b باشد. طبق رابطه (۴.۴) به دست می‌آوریم

$$\int_{J_2} F(t)dt < AMT. \quad (5.4)$$

هم‌چنین از رابطه (۳.۴) به دست می‌آوریم $T < \frac{1}{A(1-\epsilon)} \int_{J_1} F(t)dt$ با ضرب این دو رابطه

جدید به دست می‌آوریم

^۱ منظور از χ ، تابع مشخصه یک مجموعه است.

$$\int_{J_{\tau}} F(t)dt < \frac{M}{1-\epsilon} \int_{J_1} F(t)dt. \quad (6.4)$$

برای هر $0 \leq u < v \leq 1$ داریم

$$d_{[\cdot, 1]}(u, v) = \frac{v-u}{u(1-v)} \geq d_K(ua + (1-u)b, va + (1-v)b).$$

بنابراین

$$d_{[\cdot, 1]}(J_1, J_{\tau}) \geq d_K(S_1, S_{\tau}) \geq 3M.$$

حال برای افراز J_1, J_{τ}, J_{τ} از $[0, 1]$ ، از یک نابرابری محیطی دیگر استفاده می‌کنیم [۱۰]. این

نابرابری می‌گوید که برای هر افراز S_1, S_{τ}, S_{τ} از یک شکل محدب $K \subseteq \mathbb{R}^n$ و هر تابع لوگ-مقعر

$$f: K \rightarrow \mathbb{R}^+$$

$$\int_K f(x)dx \int_{S_{\tau}} f(x)dx \geq d_K(S_1, S_{\tau}) \int_{S_1} f(x)dx \int_{S_{\tau}} f(x)dx.$$

حالت یک‌بعدی این نامساوی را در قضیه ۳.۴ در ضمیمه ۳.۴ ثابت می‌کنیم. حال طبق حالت یک‌بعدی

این نامساوی به دست می‌آوریم

$$T \int_{J_{\tau}} F(t)dt \geq 3M \int_{J_1} F(t)dt \int_{J_{\tau}} F(t)dt.$$

با مقایسه این رابطه و رابطه (۶.۴) به دست می‌آوریم

$$T > \frac{3}{1-\epsilon} \int_{J_{\tau}} F(t)dt.$$

هم‌چنین از رابطه (۳.۴) به دست می‌آوریم

$$\int_{J_1} F(t)dt < A(1+\epsilon)T \leq \frac{1+\epsilon}{2}T.$$

با استفاده از دو رابطه آخر به دست می‌آوریم

$$\int_{J_{\tau}} F(t)dt = \int_{\circ}^1 F(t)dt - \int_{J_{\lambda}} F(t)dt - \int_{J_{\tau}} F(t)dt$$

$$> (1 - \frac{1+\epsilon}{2} - \frac{1-\epsilon}{3})T = \frac{1-\epsilon}{6}T.$$

حال رابطه (۵.۴) ایجاب می کند که $M > \frac{1-\epsilon}{3}$ و این با فرض $M \leq \frac{1}{6}$ متناقض است. $\square$

۵.۴. اندازه گام

برای این که بتوانیم طبق توضیحاتی که در بخش ۱.۴ دادیم، بین گزاره ۴.۳.۴ و قضیه ۳.۴.۴ ارتباط برقرار

کنیم، باید تابع مناسبی مانند h را طوری بیابیم که برای هر $u, v \in K$ ، اگر $d_K(u, v) \geq \frac{1}{\lambda}$ و یا

$$|u - v| \geq \frac{2}{\sqrt{n}} \max\{F(u), F(v)\}$$

روی وتر گذرنده از u و v آن گاه برای هر x روی وتر گذرنده از u و v

داشته باشیم $h(x) \leq \min(\frac{1}{\epsilon}, \frac{1}{3} d_K(u, v))$. برای این کار، در این بخش برای F یک کران پایین

مناسب می یابیم.

همان طور که در بخش ۲.۴ اشاره کردیم، متوسط طول یک گام با شروع از یک نقطه نزدیک به گوشه

کوچک است. در مورد قدم زدن با گام های کروی، لواژ و سیمونویتز در [۸] برای کنترل نقاط گوشه ای از

تعریف زیر استفاده کردند.

۱.۵.۴. تعریف: (رسانش موضعی قدم زدن با گام های کروی)

$$\lambda(x, t) := \frac{\text{vol}(K \cap \mathbb{B}_t(x))}{\text{vol}(\mathbb{B}_t(x))}.$$

رسانش موضعی در واقع احتمال جابجا شدن در یک گام در قدم زدن با گام های کروی با شروع از x

است. آن ها نشان دادند که اگر رسانش موضعی در همه نقاط به اندازه کافی بزرگ باشد، آن گاه قدم زدن

با گام های کروی سریع است.

«گیر کردن نزدیک یک گوشه»، می تواند به دو صورت باشد: هنگام شروع که از وضعیت نقطه شروع

اطلاعی نداریم و یا در گام های میانی. به صورت شهودی می توان گفت که مشکل در گام های میانی

خیلی جدی نیست زیرا احتمال رفتن به یک نقطه گوشه‌ای کوچک است. با توجه به این موضوع، کنان، لواژ و سیمونویتر در [۱۱] تنها با استفاده از متوسط رسانش موضعی برای زمان آمیختگی قدم‌زدن با گام‌های کروی کران دادند.

۲.۵.۴. تعریف: (متوسط رسانش موضعی)

$$\lambda(t) := \frac{1}{\text{vol}(K)} \int_K \lambda(x, t) dx.$$

از این دو تعریف در این بخش برای ارائه کرانی برای F استفاده می‌کنیم.

۳.۵.۴. نکته: از تحدب K ، به سادگی نتیجه می‌شود که $\lambda(x, t)$ بر حسب t نزولی است.

۴.۵.۴. تعریف: برای یک $\gamma \in [0, 1]$ دلخواه تعریف کنید

$$s(x) := \max\{t \in \mathbb{R}^+ : \lambda(x, t) \geq \gamma\}.$$

$s(x)$ میزان نزدیکی x به مرز را می‌سنجد.

۵.۵.۴. لم: برای هر $\gamma > 0$ ، $s(x)$ یک تابع مقعر است.

اثبات: فرض کنید $x_1, x_2 \in K$. برای $i = 1, 2$ قرار دهید $r_i := s(x_i)$ و

$A_i := (x_i + r_i \mathbb{B}^n) \cap K$. فرض کنید $x = \alpha x_1 + (1 - \alpha)x_2$ که $\alpha \in [0, 1]$ و تعریف کنید

$r := \alpha r_1 + (1 - \alpha)r_2$ و $A := \alpha A_1 + (1 - \alpha)A_2$. طبق تحدب K داریم $A \subseteq K$ و

به علاوه

$$\begin{aligned} A &\subseteq \alpha x_1 + (1 - \alpha)x_2 + \alpha r_1 \mathbb{B}^n + (1 - \alpha)r_2 \mathbb{B}^n \\ &= x + (\alpha r_1 + (1 - \alpha)r_2) \mathbb{B}^n = x + r \mathbb{B}^n. \end{aligned}$$

هم‌چنین طبق نامساوی برون-مینکوفسکی (نتیجه ب. ۱۴ در ضمیمه ب.) داریم:

$$\begin{aligned}
\text{vol}(A)^{\frac{1}{n}} &\geq \alpha \cdot \text{vol}(A_1)^{\frac{1}{n}} + (1 - \alpha) \cdot \text{vol}(A_2)^{\frac{1}{n}} \\
&\geq (\gamma \cdot \text{vol}(\mathbb{B}^n))^{\frac{1}{n}} (\alpha r_1 + (1 - \alpha) r_2) \\
&= \gamma^{\frac{1}{n}} \cdot \text{vol}(r \mathbb{B}^n)^{\frac{1}{n}}.
\end{aligned}$$

بنابراین

$$\text{vol}((x + r \mathbb{B}^n) \cap K) \geq \text{vol}(A) \geq \gamma \cdot \text{vol}(r \mathbb{B}^n)$$

و طبق نزولی بودن $\lambda(x, t)$ بر حسب t نتیجه می‌شود که $s(x) \geq r$. $\square$

۶.۵.۴. اگر $\gamma \geq \frac{63}{64}$ ، آن‌گاه برای هر $x \in K$ داریم

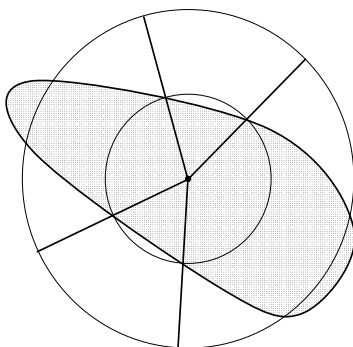
$$F(x) \geq \frac{s(x)}{32}.$$

اثبات: قرار دهید $s := s(x)$. A را مجموعهٔ نقاطی مثل u از $\partial \mathbb{B}^n := \mathbb{S}^{n-1}$ بگیرید که

$x + (\frac{s}{4})u \notin K$ و فرض کنید نسبت مساحت A به مساحت $\mathbb{S}^{n-1}$ برابر با p باشد. در نتیجه

$$(x + [\frac{s}{4}, s]A) \cap K = \emptyset \quad \text{پس}$$

$$\text{vol}((x + s \mathbb{B}^n) \setminus K) \geq p(\text{vol}(s \mathbb{B}^n) - \text{vol}(\frac{s}{4} \mathbb{B}^n)).$$



شکل ۱۶

از طرفی، طبق تعریف s داریم

$$\text{vol}((x + s\mathbb{B}^n) \setminus K) \leq (1 - \gamma)\text{vol}(s\mathbb{B}^n) \leq \frac{1}{64}\text{vol}(s\mathbb{B}^n).$$

پس

$$p \leq \frac{1}{64} / (1 - 2^{-n}) \leq \frac{1}{32}.$$

یک خط تصادفی l گذرنده از x در نظر بگیرد. با احتمال حداقل $1 - 2p$ خواهیم داشت

$$l \cap (x + \frac{s}{2}\mathbb{B}^n) \subset K$$

یکنواخت روی $l \cap K$ داریم

$$\mathbb{P}\left(\left\{|y - x| \leq \frac{s}{32}\right\} \middle| l\right) \leq \frac{1}{16}.$$

در نتیجه

$$s \mathbb{P}\left(|y - x| \leq \frac{s}{32}\right) \leq \frac{1}{16} + \frac{15}{16} \cdot \frac{1}{16} < \frac{1}{8}$$

و لم ثابت می شود. $\square$

تابعی که در نابرابری محیطی (قضیه ۳.۴.۴) استفاده خواهیم کرد، ضربی از $s(x)$ است (مراجعه کنید به اثبات قضیه ۷.۴). بنابراین باید برای متوسط s یک کران پایین ارائه دهیم. برای این کار از متوسط λ استفاده می کنیم. ابتدا به لم زیر توجه کنید.

۷.۵.۴. لم: فرض کنید $L \subseteq \partial K$ زیرمجموعه ای اندازه پذیر باشد و S مجموعه زوج های (x, y)

باشد که $x \in K$ و $y \notin K$ و $|x - y| < t$ و همچنین پاره خط xy با L متقاطع باشد.

در این صورت حجم $(2n)$ -بعدی S حداکثر برابر است با

$$\frac{t^{n+1}\pi_{n-1}}{(n+1)}\text{vol}_{n-1}(L)$$

که در آن π_{n-1} حجم کره واحد $(n-1)$ -بعدی است.

اثبات (طرح کلی): کافی است حکم را برای وقتی که L یک جزء بسیار کوچک^۱ سطح است ثابت کنیم. در این حالت بیشترین مقدار حجم S وقتی است که مرکز K در t -همسایگی L در یک ابرصفحه قرار بگیرد و K شامل نصف گوی به شعاع t به مرکز L باشد. حجم S در این حالت مستقل از K است و با یک محاسبه ساده حکم به دست می آید. $\square$

$$۸.۵.۴. \text{ نتیجه: اگر } K \text{ شامل یک گوی به شعاع } r \text{ باشد، آنگاه } \lambda(t) \geq 1 - \frac{t\sqrt{n}}{2r}.$$

اثبات: در لم قبل قرار دهید $L = \partial K$. بنابراین

$$\begin{aligned} \lambda(t) &= \frac{1}{\text{vol}(K) \cdot \text{vol}(t\mathbb{B}^n)} \int_K \text{vol}((x + t\mathbb{B}^n) \cap K) dx \\ &= \frac{1}{\text{vol}(K) \cdot \text{vol}(t\mathbb{B}^n)} \int_K \int_{y \in (x + t\mathbb{B}^n) \cap K} dy dx \\ &= 1 - \frac{1}{\text{vol}(K) \cdot \text{vol}(t\mathbb{B}^n)} \int_K \int_{y \notin (x + t\mathbb{B}^n) \cap K} dy dx \\ &= 1 - \frac{\text{vol}(S)}{\text{vol}(K) \cdot \text{vol}(t\mathbb{B}^n)} \geq 1 - \frac{t}{2\sqrt{n}} \frac{\text{vol}_{n-1}(\partial K)}{\text{vol}(K)}. \end{aligned}$$

در نامساوی آخر از لم قبل و نتیجه ب.۳ در ضمیمه ب. استفاده کردیم. همچنین طبق لم ب.۴ در

همان ضمیمه، اگر K شامل کره ای به شعاع r باشد، آن گاه $\text{vol}(K) \geq \frac{r}{n} \text{vol}_{n-1}(\partial K)$ و لم

ثابت می شود. $\square$

۹.۵.۴. لم: اگر K شامل یک گوی به شعاع r باشد، آن گاه

¹ Infinitesimal

$$\mathbb{E}_K(s) \geq \frac{r(1-\gamma)}{\sqrt{n}}.$$

اثبات: قرار دهید $p(t) := \mathbb{P}_K(\lambda(x, t) \geq \gamma)$. بنابراین با جدا کردن انتگرال روی این پیشامد و

مکمل آن به دست می‌آوریم $\lambda(t) \leq p(t) \times 1 + (1 - p(t)) \times \gamma$. پس

$$p(t) \geq \frac{\lambda(t) - \gamma}{1 - \gamma} \geq 1 - \frac{t\sqrt{(n)}}{2(1 - \gamma)}.$$

با استفاده از این رابطه به دست می‌آوریم

$$\begin{aligned} \mathbb{E}_K(s) &= \int_0^\infty P(s \geq t) dt \geq \int_0^\infty p(t) dt \\ &\geq \int_0^{\frac{2(1-\gamma)}{\sqrt{n}}} \left(1 - \frac{t\sqrt{n}}{2(1-\gamma)}\right) dt = \frac{1-\gamma}{\sqrt{n}} \end{aligned}$$

و لم ثابت می‌شود. $\square$

۶.۴. کران پایین برای رسانش

۱.۶.۴. قضیه: فرض کنید $K \subseteq \mathbb{R}^n$ مجموعه‌ای محدب با قطر D باشد که شامل کره‌ای به شعاع

$\frac{r}{2^{2\gamma}nD}$ است. در این صورت رسانش قدم‌زدن تصادفی Hit-and-run در K حداقل برابر با $\frac{r}{2^{2\gamma}nD}$ است.

اثبات: فرض کنید $K = S_1 \cup S_2$ افزایی از K به دو مجموعه اندازه‌پذیر باشد. طبق تعریف رسانش (۱۵.۲.۳) باید ثابت کنیم

$$\int_{S_1} P_x(S_2) dx \geq \frac{r}{2^{2\gamma}nD} \min\{vol(S_1), vol(S_2)\}. \quad (۷.۴)$$

نقاطی را در نظر بگیرید که عمیقاً داخل این دو مجموعه هستند؛ یعنی احتمال پریدن آن‌ها به مجموعه دیگر کم است:

$$S'_1 := \{x \in S_1 : P_x(S_2) < \frac{1}{64^{0.0}}\},$$

$$S'_2 := \{x \in S_2 : P_x(S_1) < \frac{1}{64^{0.0}}\}.$$

$$S'_2 := K \setminus (S'_1 \cup S'_2).$$

ابتدا فرض کنید $vol(S'_1) < vol(S_1)/2$ در این صورت

$$\int_{S_1} P_x(S_2) dx \geq \frac{1}{64^{0.0}} vol(S_1 \setminus S'_1) \geq \frac{1}{128^{0.0}} vol(S_1)$$

و رابطه (۷.۴) ثابت می‌شود. پس می‌توانیم فرض کنیم $vol(S'_1) \geq \frac{vol(S_1)}{2}$ و به‌طور مشابه

$$vol(S'_r) \geq \frac{vol(S_r)}{2}$$

برای هر $u \in S'_1$ و $v \in S'_r$ داریم

$$\|P_u - P_v\|_{TV} \geq P_u(S_1) - P_v(S_1) > 1 - \frac{1}{32}.$$

پس طبق لم ۴.۳.۴، یا

$$d_K(u, v) \geq \frac{1}{16} \quad (۸.۴)$$

و یا

$$|u - v| \geq \frac{2}{\sqrt{n}} \max\{F(u), F(v)\} \quad (۹.۴)$$

که در آن $s(x)$ در تعریف ۴.۵.۴ با ثابت $\gamma := \frac{63}{64}$ تعریف می‌شود. می‌خواهیم قضیه ۳.۴.۴ را برای

افراز S'_1, S'_r, S'_p و تابع $h(x) = s(x)/(48D\sqrt{n})$ اعمال کنیم. برای بررسی کردن فرض قضیه،

فرض کنید $u \in S'_1$ و $v \in S'_r$ و x نقطه‌ای روی وتر pq باشد که $p := a(u, v)$ و

$q := a(v, u)$ واضح است که $h(x) \leq \frac{1}{6}$. اگر رابطه (۸.۴) برقرار باشد، آن‌گاه به وضوح

$h(x) \leq d_K(u, v)/3$ پس فرض کنید رابطه (۹.۴) برقرار باشد. می‌توانیم فرض کنیم x بین u و

q است. حال طبق تقعر s (لم ۵.۵.۴)، داریم

$$s(x) \leq \frac{|x - p|}{|u - p|} s(u) \leq 32 \frac{|q - p|}{|u - p|} F(u) \quad (\text{طبق لم ۶.۵.۴})$$

$$\leq \mathfrak{I}\mathfrak{E} \left| \frac{q-p}{u-p} \right| \sqrt{n} |u-v| \quad (\text{طبق رابطه (۹.۴)})$$

$$= \mathfrak{I}\mathfrak{E} d_K(u, v) \sqrt{n} |q-v| \leq \mathfrak{I}\mathfrak{E} d_K(u, v) D \sqrt{n}.$$

و باز هم نتیجه می شود $h(x) \leq d_K(u, v) / \mathfrak{I}$ بنابراین طبق قضیه ۳.۴.۴

$$\begin{aligned} \text{vol}(S'_\mathfrak{r}) &\geq \frac{\mathfrak{I}}{\mathfrak{E}\mathfrak{I}D\sqrt{n}} \mathbb{E}_K(s) \min\{\text{vol}(S'_\mathfrak{I}), \text{vol}(S'_\mathfrak{r})\} \\ &\geq \frac{r}{\mathfrak{I}^{\mathfrak{I}\mathfrak{I}}nD} \min\{\text{vol}(S'_\mathfrak{I}), \text{vol}(S'_\mathfrak{r})\}. \end{aligned}$$

در این جا از لم ۹.۵.۴ استفاده کردیم. بنابراین

$$\begin{aligned} \int_{S_\mathfrak{I}} P_x(S_\mathfrak{r}) dx &= \frac{\mathfrak{I}}{\mathfrak{r}} \left(\int_{S_\mathfrak{I}} P_x(S_\mathfrak{r}) dx + \int_{S_\mathfrak{r}} P_x(S_\mathfrak{I}) dx \right) \\ &\geq \frac{\mathfrak{I}}{\mathfrak{r}} \cdot \frac{\mathfrak{I}}{\mathfrak{E}\mathfrak{E}\mathfrak{I}} \text{vol}(S'_\mathfrak{r}) \geq \frac{r}{\mathfrak{I}^{\mathfrak{I}\mathfrak{I}}nD} \min\{\text{vol}(S'_\mathfrak{I}), \text{vol}(S'_\mathfrak{r})\} \\ &\geq \frac{r}{\mathfrak{I}^{\mathfrak{I}\mathfrak{I}}nD} \min\{\text{vol}(S_\mathfrak{I}), \text{vol}(S_\mathfrak{r})\} \end{aligned}$$

و قضیه ثابت می شود. $\square$

۷.۴. کران برای زمان آمیختگی

اثبات قضیه ۲.۲.۴: طبق قضیه ۱.۶.۴ و نکته ۱۶.۲.۳، رسانش نسخه تنبل شده قدم زدن تصادفی Hit-

and-run در رابطه زیر صدق می کند

$$\Phi \geq \frac{r}{2^{2\wedge} nR}.$$

ابتدا فرض کنید σ توزیعی M -گرم باشد. یعنی همه جا داشته باشیم $d\sigma/d\pi \leq M$.

در این صورت طبق قضیه ۱۷.۲.۳ داریم

$$\|\sigma^m - \pi\|_{TV} \leq \sqrt{\frac{M}{2}} \left(1 - \frac{\Phi}{2}\right)^m.$$

بنابراین با توجه به این که $\ln(1+x) \leq x$ نتیجه می گیریم که برای

$$m > 2^{2\wedge} \frac{n^2 R^2}{r^2} \ln \frac{\sqrt{M/2}}{\epsilon}$$

خواهیم داشت $\|\sigma^m - \pi\|_{TV} < \epsilon$.

حال فرض کنید تنها خارج از یک مجموعه S که $\sigma(S) = \delta < \epsilon/2$ داشته باشیم

$d\sigma/d\pi \leq M$. قرار دهید $\sigma_1 := \frac{1}{\delta} \sigma|_S$ و $\sigma_2 := \frac{1}{1-\delta} \sigma|_{K \setminus S}$. σ_1 و σ_2 اندازه های احتمال

روی K هستند و $\sigma = \delta\sigma_1 + (1-\delta)\sigma_2$. بنابراین $\sigma^m = \delta\sigma_1^m + (1-\delta)\sigma_2^m$. σ_2 توزیعی

M -گرم است. بنابراین طبق لم ۳.۲.۳ و قضیه ۱۷.۲.۳ به دست می آوریم

$$\|\sigma^m - \pi\|_{TV} \leq \delta + (1-\delta) \sqrt{\frac{M}{2(1-\delta)}} \left(1 - \frac{\Phi}{2}\right)^m.$$

پس برای

$$m > 2^{\Delta \vee} \frac{n^2 R^2}{r^2} \ln \frac{\sqrt{2} M}{\epsilon}$$

خواهیم داشت $\|\sigma^m - \pi\|_{TV} < \epsilon$ و حکم ثابت می شود. $\square$

اثبات قضیه ۳.۲.۴: فرض کنید x نقطه ای به فاصله d از مرز K باشد و σ توزیع حاصل پس از یک گام از قدم زدن Hit-and-run با شروع از x باشد. در قضیه ۲.۲.۴، S را گوی $\mathbb{B}_{\frac{\epsilon d}{2}}(x)$ بگیرید.

پس $\sigma(S) \leq \frac{\epsilon}{2}$. همچنین طبق لم ۵.۳.۴، خارج از S تابع چگالی σ نسبت به اندازه لبگ حداکثر

$$\frac{1}{n\pi_n \left(\frac{\epsilon}{2}\right)^{n-1} d^n} \text{ است. پس چون } vol(K) \leq R^n \pi_n, \text{ خارج از } S \text{ داریم}$$

$$\frac{d\sigma}{d\pi} \leq \frac{R^n}{n\left(\frac{\epsilon}{2}\right)^{n-1} d^n}.$$

مقدار طرف راست را M بنامید. فرض کنید σ^i توزی حاصل از قدم زدن تصادفی Hit-and-run تنبل شده با شروع از σ پس از i گام باشد. در نتیجه توزیع آن پس از m گام با شروع از x برابر است

با $\sum_{i=1}^{\infty} 2^{-i} \sigma^{m-i}$ (این موضوع با در نظر گرفتن اولین گامی که از $\{x\}$ خارج می شویم دیده می شود).

پس طبق لم ۳.۲.۳ و مطالب مذکور در اثبات قضیه ۲.۲.۴، فاصله مجموع تغییرات این توزیع و π حداکثر برابر است با

$$\sum_{i=1}^{\infty} 2^{-i} \left[\frac{\epsilon}{2} + \sqrt{\frac{M}{2} \left(1 - \frac{\epsilon}{2}\right) \left(1 - \frac{\Phi^2}{2}\right)^{m-i}} \right].$$

و با یک محاسبه ساده حکم ثابت می شود. $\square$

۸.۴. تولید یک دنباله از نقاط تصادفی

در عمل معمولاً به تولید یک دنباله مستقل و هم توزیع از نقاط تصادفی $X_1, \dots, X_k$ نیاز داریم. در مسئله مورد نظر ما، می توانیم الگوریتم را چند بار مستقلاً تکرار کنیم. اما ایده بهتر این است که بعد از تولید هر نقطه، از آن به عنوان یک نقطه شروع جدید استفاده کنیم. در این صورت طبق قضیه ۲.۲.۴، توزیع نقاط در هر یک از گام های بعدی تقریباً یک نواخت خواهد بود. یک مشکل این روش این است که استقلال نقاط تولید شده از بین می رود. اما اگر الگوریتم را در مرتبه دوم و بعد «به تعداد کافی» گام انجام دهیم، نقاط حاصل «تقریباً مستقل» می شوند. همان طور که توزیع یک نواخت تقریبی در مسائل کاربردی کفایت می کند، استقلال تقریبی هم در عمل کافی است.

۱.۸.۴. تعریف: می گوئیم دو شیء تصادفی X و Y ، μ -مستقل هستند (که $0 < \mu < 1$)،

در صورتی که برای هر دو مجموعه A و B در برد آن ها داشته باشیم

$$|\mathbb{P}(X \in A, Y \in B) - \mathbb{P}(X \in A)\mathbb{P}(Y \in B)| \leq \mu.$$

این رابطه را می توان به این صورت نوشت

$$|\mathbb{P}(Y \in B | X \in A) - \mathbb{P}(Y \in B)| \leq \frac{\mu}{\mathbb{P}(X \in A)}.$$

بنابراین اگر دو نقطه تصادفی تقریباً مستقل باشند، توزیع شرطی آن ها نسبت به یک دیگر باید با توزیع خود آن ها تفاوت چندانی نداشته باشد. بنابراین برای استفاده از نقطه حاصل از الگوریتم Hit-and-run به عنوان نقطه شروع بعدی، در تکرار الگوریتم برای بار دوم هم چنان باید این قدر گام برداریم که آمیختگی تقریبی اتفاق بیفتد. اما با توجه به این که نقطه اول یک نقطه تصادفی است، تعداد گام های لازم بسیار کم تر خواهد بود (به مخرج کسر در رابطه فوق دقت کنید).

۲.۸.۴. نکته:

$$\begin{aligned} & \left| \mathbb{P}(X \in A, Y \in B) - \mathbb{P}(X \in A)\mathbb{P}(Y \in B) \right| \\ &= \left| \mathbb{P}(X \in A^c, Y \in B) - \mathbb{P}(X \in A^c)\mathbb{P}(Y \in B) \right|. \end{aligned}$$

بنابراین برای بررسی μ -مستقل بودن، کافی است تنها مجموعه‌هایی را در نظر بگیریم که

$$\mathbb{P}(X \in A) \geq \frac{1}{4} \text{ و } \mathbb{P}(Y \in B) \geq \frac{1}{4}.$$

هدف این بخش اثبات قضیه زیر است:

۳.۸.۴. قضیه: فرض کنید $K \subseteq \mathbb{R}^n$ شکل محدبی باشد که شامل یک گوی به شعاع r حول مبدأ

است و در یک گوی به شعاع R حول مبدأ قرار دارد. N را یک عدد طبیعی و $\epsilon > 0$ را عددی

دلخواه بگیرید. با این شرایط می‌توان توسط قدم‌زدن تصادفی Hit-and-run تنبل‌شده دنباله

$X_1, \dots, X_N$ از نقاط تصادفی در K را تولید کرد به طوری که شرایط زیر برقرار باشند

(۱) توزیع هریک از آن‌ها تقریباً یک‌نواخت باشد به این معنی که فاصله مجموع تغییرات آن از توزیع

یک‌نواخت روی K حداکثر ϵ باشد.

(۲) دوه دو ϵ -مستقل باشند.

(۳) تعداد گام‌های استفاده‌شده از مرتبه $\left(\frac{R^2}{r^2}\right)(n^2 + Nn^2)$ O^* باشد.

۴.۸.۴. نکته: طبق این قضیه، اگر حداقل n نقطه لازم داشته باشیم، تولید هر کدام به طور متوسط از

مرتبه $O^*\left(\frac{n^2 R^2}{r^2}\right)$ گام لازم دارد.

برای اثبات قضیه فوق ابتدا لم زیر را ثابت می‌کنیم.

۵.۸.۴. لم: فرض کنید $K \subseteq \mathbb{R}^n$ شکل محدبی با فرض‌های قضیه ۳.۸.۴ باشد. فرض کنید σ

توزیع اولیه و σ^m توزیع قدم‌زدن تصادفی Hit-and-run تنبل‌شده در K بعد از m گام باشد.

فرض کنید $\|\sigma - \pi\|_{TV} < \frac{\epsilon}{16}$. در این صورت به ازای $m = O^*\left(\frac{n^\gamma R^\gamma}{r^\gamma}\right)$ و $\sigma, \sigma^m, -\epsilon$

مستقل اند.

اثبات: برای هر دو مجموعه A و B داریم

$$\begin{aligned} & \left| \mathbb{P}(X_\bullet \in A, X_m \in B) - \mathbb{P}(X_\bullet \in A)\mathbb{P}(X_m \in B) \right| \\ &= \sigma(A) \cdot \left| \mathbb{P}(X_m \in B | X_\bullet \in A) - \sigma^m(B) \right|. \end{aligned} \quad (۱۰.۴)$$

طبق نکته ۲.۸.۴ کافی است فرض کنیم $\sigma(A) \geq \frac{1}{4}$ را شرطی شده σ به A و σ_A^m را توزیع

قدم‌زدن تصادفی Hit-and-run تنبل‌شده پس از m گام با شروع از σ_A بگیرید. بنابراین

$$\begin{aligned} & \left| \mathbb{P}(X_m \in B | X_\bullet \in A) - \sigma^m(B) \right| = \left| \sigma_A^m(B) - \sigma^m(B) \right| \\ & \leq \left\| \sigma^m - \pi \right\|_{TV} + \left\| \sigma_A^m - \pi \right\|_{TV}. \end{aligned} \quad (۱۱.۴)$$

طبق گزاره ۹.۲.۳ داریم

$$\left\| \sigma^m - \pi \right\|_{TV} \leq \left\| \sigma - \pi \right\|_{TV} < \frac{\epsilon}{16}.$$

یک مجموعه U را در نظر بگیرید که $\pi(U) = 0$ و $\sigma(U)$ ماکزیمم باشد. در این صورت خارج از

U ، σ نسبت به π پیوسته است. قرار دهید $S := U \cup \{x \in K \setminus U : d\sigma/d\pi > 2\}$ با توجه

به فرض $\left\| \sigma - \pi \right\|_{TV} \leq \frac{\epsilon}{16}$ و این که $\sigma(S) \geq 2\pi(S)$ ، به دست می‌آوریم $\sigma(S) \leq \frac{\epsilon}{8}$. پس

بنابراین طبق قضیه ۲.۲.۴، بعد از $m = O^*\left(\frac{n^\gamma R^\gamma}{r^\gamma}\right)$ $\sigma_A(S) \leq \sigma(S)/\sigma(A) \leq \frac{\epsilon}{4}$

خواهیم داشت

$$\|\sigma_A^m - \pi\|_{TV} < \frac{\epsilon}{2}$$

و طبق رابطه های (۱۰.۴) و (۱۱.۴) حکم ثابت می شود. $\square$

اثبات قضیه ۳.۸.۴: طبق قضیه ۳.۲.۴، در $O^*\left(\frac{n^\gamma R^\gamma}{r^2}\right)$ گام می توان نقطه X_1 را تولید کرد

به طوری که فاصله مجموع تغییرات توزیع آن از توزیع یک نواخت، حداکثر $\frac{\epsilon}{16}$ باشد. حال با چند بار

استفاده از لم ۵.۸.۴، می توان با ادامه دادن قدم زدن تصادفی Hit-and-run تنبل شده، نقاط

$X_2, \dots, X_N$ را در $O^*\left(\frac{Nn^\gamma R^\gamma}{r^2}\right)$ گام تولید کرد به طوری که فاصله مجموع تغییرات توزیع هریک از

آن ها از توزیع یک نواخت، حداکثر $\frac{\epsilon}{16}$ باشد و X_i و X_{i-1} ، ϵ -مستقل باشند. در واقع چون

$X_1, \dots, X_N$ در یک نمونه از یک قدم زدن تصادفی تولید شده اند، باز هم طبق لم ۵.۸.۴، دو به دو

ϵ -مستقل اند. $\square$

در ادامه این بخش برخی نتایج اولیه در مورد استقلال تقریبی می آوریم که در فصل ۵ مورد استفاده

قرار می گیرند. اگر بخواهید، می توانید آن ها را هنگام نیاز مطالعه کنید.

۶.۸.۴. لم: اگر X و Y ، μ -مستقل باشند و f و g دو تابع اندازه پذیر باشند، آن گاه $f(X)$ و

$g(Y)$ هم μ -مستقل اند.

اثبات: داریم

$$\begin{aligned} & \left| \mathbb{P}(f(X) \in A, g(Y) \in B) - \mathbb{P}(f(X) \in A)\mathbb{P}(g(Y) \in B) \right| \\ &= \left| \mathbb{P}(X \in f^{-1}(A), Y \in g^{-1}(B)) - \mathbb{P}(X \in f^{-1}(A))\mathbb{P}(Y \in g^{-1}(B)) \right| \\ &\leq \mu \end{aligned}$$

و لم ثابت می شود. $\square$

۷.۸.۴. لم: فرض کنید X و Y دو متغیر تصادفی μ -مستقل باشند که $0 \leq X \leq a$ و

$0 \leq Y \leq b$ در این صورت

$$|\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)| \leq \mu ab.$$

اثبات: داریم

$$\begin{aligned}\mathbb{E}(XY) &= \int_0^a \int_0^b \mathbb{P}(X \geq x, Y \geq y) dx dy \\ \mathbb{E}(X)\mathbb{E}(Y) &= \int_0^a \int_0^b \mathbb{P}(X \geq x) \mathbb{P}(Y \geq y) dx dy\end{aligned}$$

و حکم با در نظر گرفتن تفاضل این دو عبارت به دست می‌آید. $\square$

۸.۸.۴. لم: فرض کنید $X_0, X_1, \dots, X_i$ یک زنجیر مارکوف باشد (که می‌تواند به زمان وابسته باشد)

و X_i و X_{i-1} ، μ -مستقل باشند. در این صورت X_i و $(X_0, \dots, X_{i-1})$ نیز μ -مستقل اند.

اثبات: برای هر دو مجموعه A و B (در فضاهای مناسب) باید ثابت کنیم

$$\left| \mathbb{P}((X_0, \dots, X_i) \in A, X_i \in B) - \mathbb{P}((X_0, \dots, X_{i-1}) \in A) \mathbb{P}(X_i \in B) \right| \leq \mu.$$

قرار دهید $f(r) := \mathbb{P}((X_0, \dots, X_{i-1}, r) \in A)$ و g را تابع مشخصه B بگیرید. در

این صورت

$$\begin{aligned}\mathbb{P}((X_0, \dots, X_{i-1}) \in A) &= \mathbb{E}(f(X_{i-1})) \\ \mathbb{P}(X_i \in B) &= \mathbb{E}(g(X_i)).\end{aligned}$$

طبق خاصیت مارکوف، داریم

$$\begin{aligned}&\mathbb{P}((X_0, \dots, X_{i-1}, r) \in A, X_i \in B) \\ &= \mathbb{P}((X_0, \dots, X_{i-1}, r) \in A) \mathbb{P}(X_i \in B | X_{i-1} = r) \\ &= f(r) \mathbb{E}(g(X_i) | X_{i-1} = r) = \mathbb{E}(f(r) g(X_i) | X_{i-1} = r) \\ &= \mathbb{E}(f(X_{i-1}) g(X_i) | X_{i-1} = r).\end{aligned}$$

بنابراین

$$\mathbb{P}((X_1, \dots, X_{i-1}) \in A, X_i \in B) = \mathbb{E}(f(X_{i-1})g(X_i))$$

و حکم از لم ۷.۸.۴ نتیجه می شود. $\square$

۹.۸.۴. لم: فرض کنید X, Y, X' و Y' متغیرهای تصادفی باشند و هم‌چنین زوج‌های (X, Y) و (X', Y') مستقل باشند، X و Y, μ -مستقل باشند و X' و Y' نیز μ' -مستقل باشند. در این صورت زوج‌های (X, X') و (Y, Y') $(\mu + \mu')$ -مستقل هستند.

اثبات: می‌خواهیم ثابت کنیم برای هر دو مجموعه A و B (در فضا‌های مناسب) داریم

$$\left| \mathbb{P}((X, X') \in A, (Y, Y') \in B) - \mathbb{P}((X, X') \in A) \mathbb{P}((Y, Y') \in B) \right| \leq \mu + \mu'. \quad (۱۲.۴)$$

قرار دهید

$$A_r := \{r' : (r, r') \in A\}, \quad B_s := \{s' : (s, s') \in B\},$$

$$f(r) = \mathbb{P}(X' \in A_r), \quad g(s) := \mathbb{P}(Y' \in B_s),$$

$$h(r, s) = \mathbb{P}(X' \in A_r, Y' \in B_s).$$

بنابراین طبق استقلال دو زوج (X, Y) و (X', Y') به دست می‌آوریم

$$\mathbb{P}((X, X') \in A) = \mathbb{E}(f(X)),$$

$$\mathbb{P}((Y, Y') \in B) = \mathbb{E}(g(Y)),$$

$$\mathbb{P}((X, X') \in A, (Y, Y') \in B) = \mathbb{E}(h(X, Y)).$$

حال عبارت داخل قدر مطلق در رابطه (۱۲.۴) را می‌توانیم به صورت زیر بنویسیم

$$\begin{aligned} & \mathbb{E}(h(X, Y)) - \mathbb{E}(f(X))\mathbb{E}(g(Y)) \\ &= [\mathbb{E}(h(X, Y) - f(X)g(Y))] + [\mathbb{E}(f(X)g(Y)) - \mathbb{E}(f(X))\mathbb{E}(g(Y))]. \end{aligned} \quad (۱۳.۴)$$

طبق فرض، داریم

$$\left| h(r, s) - f(r)g(s) \right| = \left| \mathbb{P}(X' \in A_r, Y' \in B_s) - \mathbb{P}(X' \in A_r)\mathbb{P}(Y' \in B_s) \right| \leq \mu'.$$

بنابراین جمله اول سمت راست عبارت (۱۳.۴) حداکثر μ' است. جمله دوم نیز طبق لم ۷.۸.۴

حداکثر μ است و این حکم را ثابت می کند. $\square$

۵. الگوریتم حجم

در این فصل الگوریتم تخمین حجم را با جزئیات بیش تر می آوریم و به تحلیل آن می پردازیم.

۱.۱.۵ الگوریتم: فرض کنید $K \subseteq \mathbb{R}^n$ یک شکل داده شده توسط یک اوراکل عضویت قوی و با

شرط خوش تضمینی $\mathbb{B}_r(\circ) \subseteq K \subseteq \mathbb{B}_R(\circ)$ باشد. قرار دهید $m = \lceil n \log R/r \rceil$ و

$K_i := K \cap \mathbb{B}_{r \cdot 2^{i/n}}(\circ)$ ($i = 0, 1, \dots, m$) فرض کنید k عددی طبیعی و $\delta > 0$ عددی

حقیقی و دلخواه باشد (که مقدار آن ها بعداً معرفی می شود). نقاط تصادفی

$X_i^j \in K_i$ ($0 \leq i \leq m$ و $1 \leq j \leq k$) را به صورت زیر می سازیم به طوری که فاصلهٔ مجموع

تغییرات توزیع X_i^j از توزیع یک نواخت روی K_i حداکثر δ باشد.

$X_0^1, \dots, X_0^k$ را نقاط تصادفی مستقل با توزیع یک نواخت در $K_0 = \mathbb{B}_r(\circ)$ بگیرید.

فرض کنید $X_{i-1}^1, \dots, X_{i-1}^k$ ساخته شده باشند.

قدم زدن تصادفی Hit-and-run تنبل شده را در شکل K_i ، k بار مستقلاً با نقاط شروع

$X_{i-1}^1, \dots, X_{i-1}^k$ انجام دهید تا نقاط $X_i^1, \dots, X_i^k$ به دست آیند (تعداد گام ها را بعداً مشخص

می کنیم).

توسط این نقاط مقدار زیر را به عنوان تخمینی از $vol(K_{i-1})/vol(K_i)$ حساب کنید

$$W_i = \frac{1}{k} \sum_{j=1}^k \chi_{K_{i-1}}(X_i^j).$$

خروجی الگوریتم، مقدار $Z = r^n \pi_n / (W_1 \dots W_m)$ به عنوان تخمینی از حجم K است.

هدف این فصل این است که نشان دهیم برای هر $\epsilon > 0$ ، با انتخاب مقادیر مناسب k و تعداد گام‌های قدم‌زدن تصادفی Hit-and-run تنبل‌شده، خروجی این الگوریتم با احتمال بالا حجم K را با خطای نسبی حداکثر ϵ محاسبه می‌کند. در این بند مراحل اثبات را به طور خلاصه توضیح می‌دهیم و سپس به بیان جزئیات می‌پردازیم. ابتدا با استفاده از قضیه ۲.۲.۴ تعداد گام‌های قدم‌زدن تصادفی Hit-and-run تنبل‌شده را طوری انتخاب می‌کنیم که فاصله مجموع تغییرات توزیع X_i^j از توزیع یک‌نواخت روی K_i حداکثر δ باشد. سپس با استفاده از گزاره ۱۱.۲.۳، X_i^j را با یک نقطه تصادفی $\bar{X}_i^j$ در K_i تقریب می‌زنیم که توزیع آن دقیقاً توزیع یک‌نواخت روی K_i باشد و به وسیله آن‌ها $\bar{W}_1, \dots, \bar{W}_m$ را تعریف می‌کنیم. در لم ۴.۱.۵ نشان می‌دهیم که با احتمال زیاد برای هر i و j داریم $X_i^j = \bar{X}_i^j$. در لم ۵.۱.۵ استقلال تقریبی $\bar{X}_i^j$ ‌ها را بررسی می‌کنیم و سپس در لم ۶.۱.۵ خطای نسبی $\bar{W}_1 \dots \bar{W}_m$ را مورد مطالعه قرار می‌دهیم. در نهایت در قضیه ۸.۱.۵ با مشخص کردن پارامترها حکم را ثابت می‌کنیم.

با توجه به این که $K_i \subseteq K_{i-1}$ ، نتیجه می‌شود $\text{vol}(K_i) \leq \text{vol}(K_{i-1})$. μ_i را توزیع یک‌نواخت روی K_i بگیرد. بنابراین $\mu_{i-1} \leq 2\mu_i$. در قدم‌زدن تصادفی Hit-and-run موجود در الگوریتم فوق، با استفاده از قضیه ۲.۲.۴ تعداد گام‌ها را به اندازه‌ای در نظر بگیرید که با شروع از هر توزیعی که حداکثر ۴ برابر μ_i باشد، فاصله مجموع تغییرات توزیع نهایی از μ_i حداکثر δ/m باشد. توزیعی که با این تعداد گام با شروع از μ_{i-1} به دست می‌آید را ν_i بنامید.

۲.۱.۵. لم: فاصله مجموع تغییرات توزیع X_i^j از μ_i حداکثر δ است.

اثبات: اگر توزیع یک متغیر تصادفی X را با μ_X نشان دهیم، طبق لم ۹.۲.۳ داریم

$$\|\mu_{X_i^j} - \nu_i\|_{TV} \leq \|\mu_{X_{i-1}^j} - \mu_{i-1}\|_{TV}.$$

بنابراین

$$\|\mu_{X_i^j} - \mu_i\|_{TV} \leq \|\mu_{X_{i-1}^j} - \mu_{i-1}\|_{TV} + \sigma/m$$

و به طور استقرایی به دست می آوریم

$$\|\mu_{X_i^j} - \mu_i\|_{TV} \leq i\sigma/m \leq \sigma$$

و حکم ثابت می شود. $\square$

۳.۱.۵. **تعریف:** هریک از دنباله های $(X_0^j, \dots, X_m^j)$ را یک رشته می نامیم. توجه کنید که

رشته های مختلف از هم مستقل هستند. همچنین مراحل ساخته شدن X_i^j ها از روی X_{i-1}^j ها را فاز i -ام می نامیم.

توزیع X_i^j «تقریباً» برابر با μ_i است. برای بررسی به تر، X_i^j را با نقطه تصادفی $\bar{X}_i^j$ تقریب

می زنیم که توزیع آن دقیقاً μ_i باشد. j را ثابت بگیرید. قرار دهید $\bar{X}_0^j := X_0^j$. فرض کنید $\bar{X}_i^j$

تعریف شده باشد. به جای شروع از X_i^j ، الگوریتم Hit-and-run را در K_{i+1} با نقطه شروع $\bar{X}_i^j$ اجرا

کنید (با همان تعداد گام و با همان منبع تصادف) تا نقطه Y حاصل شود. بنابراین فاصله مجموع

تغییرات توزیع Y از μ_{i+1} حداکثر δ است. حال طبق گزاره ۱۱.۲.۳، $\bar{X}_{i+1}^j$ را نقطه ای تصادفی با

توزیع μ_{i+1} بگیرید که $\mathbb{P}(\bar{X}_{i+1}^j = Y) \geq 1 - \delta$ (در گزاره ۱۱.۲.۳ نحوه ساختن $\bar{X}_{i+1}^j$ را مستقل

از متغیرهای دیگر انجام دهید). این کار را برای رشته های مختلف به صورت مستقل تکرار می کنیم.

بنابراین رشته های جدید $(\bar{X}_0^j, \dots, \bar{X}_m^j)$ نیز از هم مستقل اند.

بالاخره تعریف کنید

$$\bar{W}_i := \frac{1}{k} \sum_{j=1}^k \chi_{K_{i-1}}(\bar{X}_i^j).$$

۵.۴.۱.۵. لم:

$$\mathbb{P}(X_i^j = \bar{X}_i^j, \forall i, j) \geq 1 - km\delta.$$

اثبات: در صورتی که $X_i^j = \bar{X}_i^j$ باشد، آن گاه $Y = X_{i+1}^j$. بنابراین به احتمال حداقل $1 - \delta$

خواهیم داشت $X_{i+1}^j = \bar{X}_{i+1}^j$. به این صورت با استقرا به دست می آوریم که

$$\mathbb{P}(X_l^j = \bar{X}_l^j, \forall l \leq i) \geq 1 - i\delta$$

و با استفاده از این موضوع حکم نتیجه می شود. $\square$

اگر می توانستیم تقریب زدن را طوری انجام دهیم که اعضای هر رشته هم مستقل باشند استدلال ساده تر می شد. اما برای این کار باید در هر فاز از یک شروع سرد شروع کنیم که به زمان زیادی احتیاج دارد. در روش ما باید استقلال تقریبی اعضای رشته ها را به دقت بررسی کنیم.

۵.۴.۱.۵. لم: الف) نقاط $\bar{X}_i^j$ و $\bar{X}_{i+1}^j$ ، (δ/m) -مستقل اند.

ب) دنباله $(\bar{X}_i^j, \dots, \bar{X}_{i+1}^j)$ و $\bar{X}_{i+1}^j$ ، (δ/m) -مستقل اند.

ج) متغیرهای $\bar{W}_1 \dots \bar{W}_i$ و $\bar{W}_{i+1}$ ، $(k\delta/m)$ -مستقل اند.

اثبات: الف) طبق طریقه ساختن $\bar{X}_{i+1}^j$ ، Y را نقطه تصادفی حاصل از قدم زدن تصادفی Hit-and-

run تنبل شده با شروع از $\bar{X}_i^j$ بگیرید. فرض کنید $A, B \subseteq K_{i+1}$ دو مجموعه دلخواه باشند. طبق

نکته ۲.۸.۴ کافی است فرض کنیم $\mu_i(A) \geq \frac{1}{\gamma}$. μ_i' را شرطی شده μ_i به A بگیرید. پس

$$\mu_i' \leq 2\mu_i \leq 4\mu_{i+1}$$

$$\left| \mathbb{P}(Y \in B \mid \bar{X}_i^j \in A) - \mathbb{P}(Y \in B) \right| \leq \delta/m.$$

بنابراین

$$\left| \mathbb{P}(Y \in B, \bar{X}_i^j \in A) - \mathbb{P}(Y \in B) \mathbb{P}(\bar{X}_i^j \in A) \right| \leq \delta/m.$$

پس Y و $\bar{X}_i^j$ ، (δ/m) -مستقل‌اند. با توجه به این که ساختن $\bar{X}_{i+1}^j$ از Y به طور مستقل از بقیه نقاط انجام شده بود، سه تایی $(\bar{X}_i^j, Y, \bar{X}_{i+1}^j)$ یک فرآیند مارکوف (وابسته به زمان) است. پس $(\bar{X}_{i+1}^j, Y, \bar{X}_i^j)$ نیز همین طور است و طبق لم‌های ۸.۸.۴ و ۶.۸.۴، $\bar{X}_i^j$ و $\bar{X}_{i+1}^j$ نیز (δ/m) -مستقل‌اند.

(ب) این موضوع از لم ۸.۸.۴ نتیجه می‌شود.

(ج) با چند بار استفاده از قسمت (ب) و لم ۹.۸.۴ نتیجه می‌شود که دو دنباله $(\bar{X}_{i+1}^1, \dots, \bar{X}_{i+1}^k)$

$\{\bar{X}_l^j : 0 \leq l \leq i, 1 \leq j \leq k\}$ ، $(k\delta/m)$ -مستقل‌اند و حکم از لم ۶.۸.۴ نتیجه می‌شود. $\square$

قرار دهید $R_i := \text{vol}(K_{i-1}) / \text{vol}(K_i)$. بنابراین $R_i \geq \frac{1}{\gamma}$ و $\mathbb{E}(\bar{W}_i) = R_i$

۶.۱.۵. لم (طرح کلی): به احتمال زیاد خطای نسبی $\bar{W}_1 \dots \bar{W}_m$ به عنوان تقریبی از $R_1 \dots R_m$ کم است.

معنی صورت این لم این است که برای خطای نسبی و خطای احتمالاتی مذکور کران ارائه می‌کنیم. برای این کار برای متوسط و واریانس $\bar{W}_1 \dots \bar{W}_m$ کران ارائه می‌کنیم. در این صورت می‌توان از قضیه چبیشف استفاده کرد. طبق قسمت (ج) در لم ۵.۱.۵، $\bar{W}_i$ و $\bar{W}_1 \dots \bar{W}_{i-1}$ ، $(k\delta/m)$ -مستقل‌اند. با توجه به این که تنها می‌دانیم $0 \leq \bar{W}_i \leq 1$ ، اگر بخواهیم مستقیماً از لم ۷.۸.۴ استفاده کنیم، به دست می‌آوریم

$$\left| \mathbb{E}(\bar{W}_1 \dots \bar{W}_i) - \mathbb{E}(\bar{W}_1 \dots \bar{W}_{i-1}) \mathbb{E}(\bar{W}_i) \right| \leq k\delta/m.$$

اما $k\delta/m$ می تواند نسبت به $R_1 \dots R_i$ بزرگ باشد. بنابراین $\bar{W}_i \dots \bar{W}_i$ را به صورت زیر تغییر

می دهیم. به صورت استقرایی تعریف کنید

$$U_i := \min\{U_{i-1}\bar{W}_i, \alpha R_1 \dots R_i\}.$$

بعداً نشان می دهیم که به احتمال زیاد $U_m = \bar{W}_1 \dots \bar{W}_m$. بنابراین برای متوسط و واریانس U_m

کران می یابیم. برای این کار به لم زیر احتیاج داریم.

۷.۱.۵. لم: فرض کنید $X \geq 0$ یک متغیر تصادفی و $a > 0$ عددی حقیقی باشد. در این صورت

$$\mathbb{E}(\min\{X, a\}) \geq \mathbb{E}(X) - \frac{\mathbb{E}(X^2)}{4a}.$$

اثبات: قرار دهید $X' := \min\{X, a\}$ و $X'' := X - X'$. توجه کنید که $X'X'' = aX''$.

بنابراین

$$\mathbb{E}(X^2) = \mathbb{E}((X' + X'')^2) \geq 4\mathbb{E}(X'X'') = 4\mathbb{E}(aX'') = 4a(\mathbb{E}(X) - \mathbb{E}(X'))$$

و حکم ثابت می شود. $\square$

اثبات لم ۶.۱.۵: طبق لم ۶.۸.۴، U_i و $\bar{W}_{i+1}$ ، $k\delta/m$ -مستقل اند. پس طبق لم ۷.۸.۴ خواهیم داشت

$$|\mathbb{E}(U_i \bar{W}_{i+1}) - R_{i+1} \mathbb{E}(U_i)| \leq \frac{k\delta}{m} \cdot \alpha R_1 \dots R_i \leq \frac{2k\delta}{m} \cdot \alpha R_1 \dots R_{i+1}. \quad (۱.۵)$$

بنابراین

$$\mathbb{E}(U_{i+1}) \leq \mathbb{E}(U_i \bar{W}_{i+1}) \leq R_{i+1} \mathbb{E}(U_i) + \frac{2k\delta}{m} \cdot \alpha R_1 \dots R_{i+1}.$$

حال به صورت استقرایی به دست می آوریم

$$\mathbb{E}(U_i) \leq (1 + \frac{2ik\delta}{m} \alpha) R_1 \dots R_i. \quad (۲.۵)$$

هم چنین با توجه به این که $\mathbb{E}(\bar{W}_i^2) \geq R_i^2$ ، به طور مشابه به دست می آوریم

$$\mathbb{E}(U_i^r) \leq (1 + \frac{\mathfrak{f}ik\delta}{m}\alpha^r)\mathbb{E}(\bar{W}_1^r)\dots\mathbb{E}(\bar{W}_i^r) \quad (3.5)$$

و باز هم طبق لم ۷.۸.۴

$$\begin{aligned} \mathbb{E}(U_i^r \bar{W}_{i+1}^r) &\leq \mathbb{E}(U_i^r)\mathbb{E}(\bar{W}_{i+1}^r) + \frac{k\delta}{m}\alpha^r R_1^r \dots R_i^r \\ &\leq (1 + \frac{\mathfrak{f}(i+1)k\delta}{m}\alpha^r)\mathbb{E}(\bar{W}_1^r)\dots\mathbb{E}(\bar{W}_{i+1}^r) \\ &\leq (1 + \mathfrak{f}k\delta\alpha^r)\mathbb{E}(\bar{W}_1^r)\dots\mathbb{E}(\bar{W}_{i+1}^r). \end{aligned} \quad (4.5)$$

حال برای $\mathbb{E}(U_i)$ یک کران پایین می یابیم. طبق لم ۷.۱.۵ و رابطه (۴.۵) به دست می آوریم

$$\begin{aligned} \mathbb{E}(U_{i+1}) &\geq \mathbb{E}(U_i \bar{W}_{i+1}) - \frac{\mathbb{E}(U_i^r \bar{W}_{i+1}^r)}{\mathfrak{f}\alpha R_1 \dots R_{i+1}} \\ &\geq \mathbb{E}(U_i \bar{W}_{i+1}) - (1 + \mathfrak{f}k\delta\alpha^r) \frac{\mathbb{E}(\bar{W}_1^r)\dots\mathbb{E}(\bar{W}_{i+1}^r)}{\mathfrak{f}\alpha R_1 \dots R_{i+1}}. \end{aligned}$$

از طرفی داریم

$$\mathbb{E}(\bar{W}_i^r) = \mathbb{E}(\frac{1}{k} \sum_{j=1}^k (\bar{X}_i^j)^r) = \frac{R_i}{k} + \frac{k-1}{k} R_i^r < (1 + \frac{r}{k}) R_i^r. \quad (5.5)$$

بنابراین

$$\mathbb{E}(U_{i+1}) \geq \mathbb{E}(U_i \bar{W}_{i+1}) - \frac{1}{\mathfrak{f}\alpha} (1 + \mathfrak{f}k\delta\alpha^r) (1 + \frac{r}{k})^{i+1} R_1 \dots R_{i+1}.$$

هم چنین طبق رابطه (۱.۵) داریم

$$\mathbb{E}(U_i \bar{W}_{i+1}) \geq \mathbb{E}(U_i) R_{i+1} - \frac{\mathfrak{f}k\delta}{m} \alpha R_1 \dots R_{i+1}.$$

بنابراین

$$\mathbb{E}(U_{i+1}) \geq \mathbb{E}(U_i) R_{i+1} - (\frac{\mathfrak{f}k\delta}{m} \alpha + \frac{1}{\mathfrak{f}\alpha} (1 + \mathfrak{f}k\delta\alpha^r) (1 + \frac{r}{k})^{i+1}) R_1 \dots R_{i+1}$$

و به صورت استقرایی به دست می آوریم

$$\mathbb{E}(U_i) \geq (1 - \frac{\mathfrak{r}ik\delta}{m}\alpha - \frac{k}{\lambda\alpha})(1 + \mathfrak{r}k\delta\alpha^{\mathfrak{r}})(1 + \frac{\mathfrak{r}}{k})^{i+1})R_1 \dots R_i. \quad (6.5)$$

از طرفی طبق روابط (3.5) و (5.5) به دست می آوریم

$$\mathbb{E}(U_m^{\mathfrak{r}}) \leq (1 + \mathfrak{r}k\delta\alpha^{\mathfrak{r}})(1 + \frac{\mathfrak{r}}{k})^m R_1^{\mathfrak{r}} \dots R_m^{\mathfrak{r}}. \quad (7.5)$$

به وسیله دو رابطه آخر می توان برای واریانس U_m کران داد. این کار را در اثبات قضیه بعد انجام

می دهیم.

هم چنین نشان می دهیم که به احتمال زیاد $U_m = \bar{W}_1 \dots \bar{W}_m$ داریم

$$\begin{aligned} \mathbb{P}(U_{i+1} \neq U_i \bar{W}_{i+1}) &= \mathbb{P}(U_i \bar{W}_{i+1} > \alpha R_1 \dots R_{i+1}) \\ &\leq \frac{\mathbb{E}(U_i \bar{W}_{i+1})}{\alpha R_1 \dots R_{i+1}} \leq \frac{1}{\alpha} + \mathfrak{r}k\delta. \end{aligned}$$

بنابراین

$$\mathbb{P}(U_m = \bar{W}_1 \dots \bar{W}_m) \geq 1 - m(\frac{1}{\alpha} + \mathfrak{r}k\delta). \quad (8.5)$$

از این اطلاعات در اثبات قضیه زیر استفاده می کنیم. $\square$

8.1.5. **قضیه:** برای هر $\epsilon > 0$ ، می توان پارامترهای الگوریتم 1.1.5 را طوری مشخص کرد که خطای

نسبی خروجی الگوریتم به احتمال $1 - o(1)$ کمتر از ϵ باشد و تعداد دفعات استفاده از اوراکل از

$$\text{مرتبه } O^*\left(\frac{n^{\mathfrak{r}} R^{\mathfrak{r}}}{r^{\mathfrak{r}}}\right) \text{ باشد.}$$

اثبات: برای سادگی، فقط به ذکر مرتبه پارامترها اکتفا می کنیم و از بیان مقدار دقیق آن ها خودداری

می کنیم. در صورتی که خطای نسبی $W_1 \dots W_m$ کم تر از $\epsilon/2$ باشد، خطای نسبی خروجی الگوریتم

کم تر از ϵ می شود. طبق توضیحات موجود در لم قبل داریم

$$1 - \alpha_1 \leq \frac{\mathbb{E}(U_m)}{R_1 \dots R_m} \leq 1 + \alpha_r,$$

$$\frac{\mathbb{E}(U_m^r)}{R_1^r \dots R_m^r} \leq 1 + \beta$$

که α_1 ، α_r و β توسط روابط (۶.۵)، (۲.۵) و (۷.۵) مشخص می‌شوند. می‌خواهیم از قضیهٔ چبیشف استفاده کنیم. فرض کنید

$$\alpha_1, \alpha_r \leq \epsilon/4. \quad (9.5)$$

در این صورت

$$\begin{aligned} \mathbb{P}\left(\left|\frac{U_m}{R_1 \dots R_m} - 1\right| > \frac{\epsilon}{4}\right) &\leq \mathbb{P}(|U_m - \mathbb{E}(U_m)| > \frac{\epsilon}{4} R_1 \dots R_m) \\ &\leq \frac{16 \text{Var}(U_m)}{\epsilon^r R_1^r \dots R_m^r} \leq \frac{16(\beta + 2\alpha_1)}{\epsilon^r}. \end{aligned}$$

با استفاده از این رابطه و لم ۴.۱.۵ و رابطهٔ (۸.۵)، نتیجه می‌گیریم که احتمال این که خطای نسبی

$W_1 \dots W_m$ به عنوان تقریبی از $R_1 \dots R_m$ بیش‌تر از $\epsilon/2$ باشد، حداکثر برابر است با

$$km\delta + m\left(\frac{1}{\alpha} + 2k\delta\right) + \frac{16(\beta + 2\alpha_1)}{\epsilon^r}.$$

بنابراین کافی است جملات $km\delta$ ، $\frac{m}{\alpha}$ ، $\frac{\beta}{\epsilon^r}$ و $\frac{\alpha_1}{\epsilon^r}$ از مرتبهٔ $o(1)$ باشند. داریم

$m = o(n \log R/r)$. به علت وجود جملهٔ $(1 + \frac{2}{k})^m$ در β و α_1 ، باید داشته باشیم

$k = o(m/\epsilon^r) = o(n \log(R/r)/\epsilon^r)$. به دلیل وجود جملهٔ k/α در α_r ، باید داشته باشیم

$\alpha = o(n \log(R/r)/\epsilon^r)$. در نهایت با توجه به شرط (۹.۵)، $\delta = o(\frac{\epsilon^\delta}{n^r \log^r R/r})$. به راحتی

دیده می‌شود که با این شرایط خواص مطلوب برقرار می‌شوند.

طبق توضیحات بعد از الگوریتم ۱.۱.۵ و قضیه ۲.۲.۴، تعداد گام‌های لازم از قدم‌زدن تصادفی Hit-

and-run تنبل‌شده برای تولید هر X_i^j ، از مرتبه $O^*\left(\frac{n^2 R^2}{r^2}\right)$ است. بنابراین تعداد دفعات استفاده از

اوراکل از مرتبه $O^*\left(\frac{n^5 R^2}{r^2}\right)$ است. توجه کنید که بستگی آن به ϵ نیز به شکل $\frac{1}{\epsilon^2} \log^{\gamma} \frac{1}{\epsilon}$ است. $\square$

۶. کارهای دیگر

۱.۶. دقیق بودن مرتبه کران

لواژ در [۳] پس از ارائه اولین کران چندجمله‌ای برای قدم‌زدن تصادفی Hit-and-run نشان داد که لاقل اگر ϵ را ثابت بگیریم، مرتبه کران موجود در قضیه ۲.۲.۴ دقیق است. به این منظور او استوانه $C_p := \mathbb{B}^{n-1} \times [-p, p]$ را در نظر گرفت. به ازای $K = C_p$ داریم

$$\sigma_{R/r = \sqrt{(1+p^2)}} \text{ را توزیع یک‌نواخت روی } \mathbb{B}^{n-1} \times \left[-\frac{p}{2}, \frac{p}{2}\right] \text{ بگیرید.}$$

۱.۱.۶. قضیه: با اجرای قدم‌زدن تصادفی Hit-and-run در C_p با توزیع اولیه σ و توقف در لحظه

$$T = \frac{n^2 p^2}{100}, \text{ اگر } n \text{ به اندازه کافی بزرگ باشد آن گاه فاصله مجموع تغییرات توزیع نقطه نهایی از}$$

$$\text{توزیع یک‌نواخت روی } C_p \text{ حد اقل } \frac{1}{12} \text{ است.}$$

اثبات شهودی این قضیه آسان است. مشابه نتیجه ب.۱۸، جهت انتخاب شده در هر لحظه تقریباً عمود بر صفحه دوبعدی گذرنده از محور n -ام مختصات و نقطه جاری است. بنابراین تا هنگامی که نزدیک به دو قاعده C_p نشویم، خط انتخاب شده به احتمال زیاد این دو قاعده را قطع نمی‌کند. در این صورت $n-1$ مؤلفه اول نقطه جاری توزیع یک‌نواخت روی $\mathbb{B}^{n-1}$ دارند و به احتمال زیاد نقطه جاری به فاصله

$$1 - \Theta\left(\frac{1}{n}\right) \text{ از محور } n\text{-ام مختصات است. بنابراین طول گام به احتمال زیاد } \Theta(\sqrt{n}) \text{ است و تصویر}$$

آن روی محور n -ام به اندازه $\Theta(1/n)$ است. پس اگر مؤلفه n -ام نقاط را نگاه کنیم، تقریباً یک

قدم زدن تصادفی با طول گام $1/n$ انجام می شود. بنابراین برای پوشاندن فاصله p با احتمال مناسب، به $\Theta(n^2 p^2)$ گام احتیاج است.

اما اثبات دقیق به جزئیات بیش تری نیاز دارد. فرض کنید $(v_t)_{t=1}^\infty$ قدم زدن تصادفی فوق باشد و y_t مؤلفه n -ام v_t باشد. ابتدا قدم زدن تصادفی $(w_t = (b_t, z_t))_{t=0}^\infty$ در استوانه $C^* = \mathbb{B}^{n-1} \times \mathbb{R}$ با همان توزیع اولیه را در نظر می گیریم که در آن $b_t \in \mathbb{B}^{n-1}$ و $z_t \in \mathbb{R}$ (توجه کنید که احتمال این که جهت انتخاب شده موازی محور n -ام مختصات باشد صفر است). اگر برای این دو قدم زدن تصادفی از یک منبع تصادف استفاده کنیم^۱، تا هنگامی که خط انتخاب شده یکی از دو قاعده C_p را قطع کند v_t با w_t برابر است. ابتدا میزان حرکت عمودی، یعنی z_t را بررسی می کنیم.

به وضوح z_t یک مارتینگل^۲ است، یعنی $\mathbb{E}(z_{t+1} - z_t | w_t) = 0$. در مقاله مذکور برای واریانس تغییرات z_t نیز کران زیر ارائه شده است.

$$\mathbb{E}((z_{t+1} - z_t)^2 | w_t) \leq \frac{3}{n(n+1)} + \frac{1}{n}(1 - |b_t|^2).$$

با توجه به این که سمت چپ این نامساوی همان $\mathbb{E}(z_{t+1}^2 - z_t^2 | w_t)$ است، می توان به این وسیله برای واریانس z_T کران ارائه کرد. اما برای این که بین z_T و y_T ارتباط برقرار کنیم، T' را مینیمم

T و اولین زمانی می گیریم که $|z_t|^2 > 12p^2/13$. T' یک زمان توقف برای قدم زدن تصادفی w_t

است. همچنین از رابطه فوق نتیجه می شود که فرآیند $z_t^2 - \frac{3t}{n(n-1)} - \frac{1}{n} \sum_{i=0}^{t-1} (1 - |b_i|^2)$ یک

زیرمارتینگل است. به این ترتیب طبق قضیه دوب^۳ به دست می آوریم که

^۱ این در واقع یک جفت سازی از این دو قدم زدن تصادفی است.

^۲ Martingale

^۳ Doob

$$\mathbb{E}\left((z_{T'})^2\right) \leq \frac{p^2}{\epsilon} \quad (۱.۶)$$

(برای این کار باید از این که $T' < T = n^2 p^2 / ۱۰۰$ و این که تا قبل از زمان T' توزیع b_t روی

$\mathbb{B}^{n-1}$ یک نواخت است استفاده کنیم).

حال می توانیم برای واریانس y_T کران دهیم. امید ریاضی $y_T - z_{T'}$ را به شرط سه پیشامد $\{v_T = w_{T'}\}$ ، $\{T' \neq T\}$ و $\{T' = T, v_T \neq w_T\}$ در نظر می گیریم. امید شرطی اول به وضوح صفر است. در صورت وقوع پیشامد دوم داریم $z_{T'} > ۱۲ p^2 / ۱۳$ در حالی که $y_T \leq p^2$. پس امید شرطی دوم حداکثر $p^2 / ۱۳$ است. حال ادعا می کنیم که احتمال وقوع پیشامد سوم از مرتبه $o(p^2)$ است. در صورت وقوع این پیشامد باید قبل از زمان T' خطی انتخاب شده باشد که یکی از دو قاعده C_p را قطع کند. با توجه به این که قبل از زمان T' داریم $|z_T^2| < ۱۲ p^2 / ۱۳$ ، می توان برای احتمال این پیشامد کران داد و به این ترتیب ادعا را ثابت کرد. به این ترتیب به دست می آوریم که

$$\mathbb{E}(y_T - z_{T'}) \leq p^2 / ۱۳ + o(p^2). \quad (۲.۶)$$

اما اگر فاصله مجموع تغییرات توزیع y_T از توزیع یک نواخت روی $[-p, p]$ کمتر از $۱/۱۲$ باشد،

به دست می آوریم

$$\mathbb{E}(y_T) > \frac{p^2}{۴}$$

که این موضوع با روابط (۱.۶) و (۲.۶) متناقض است. $\square$

۲.۶. گرد کردن شکل

همان طور که در فصل های ۴ و ۵ دیدیم، مرتبه الگوریتم ها بستگی زیادی به R/r دارند. یک روش برای کوچک کردن این نسبت، این است که یک تبدیل آفین مناسب روی K اعمال کنیم. با این کار توزیع یک نواخت روی K تبدیل به توزیع یک نواخت روی مجموعه محدب جدید می شود و حجم K نیز در دترمینان این تبدیل ضرب می شود. به این کار *ساندویچ کردن* K و به نسبت R/r نسبت *ساندویچی* می گویند.

طبق قضیه ای از لاونر^۱ و جان^۲، همواره می توان با اعمال یک تبدیل آفین مناسب به نسبت ساندویچی $R/r \leq n$ دست یافت [۱۲]. اما تا کنون هیچ الگوریتمی برای یافتن این تبدیل و یا حتی تخمین آن یافت نشده است. در روش های تصادفی از *ساندویچ کردن تصادفی* استفاده می شود. یعنی با احتمال به دلخواه زیاد، نسبت ساندویچی کوچک باشد. به این ترتیب دایر، فریز و کنان در [۱] به نسبت ساندویچی $R/r = O(n^{3/2})$ دست یافتند. سپس لواژ و سیمونیتز در [۸] این نسبت را به $O(n)$ کاهش دادند. بعد از آن کنان، لواژ و سیمونیتز در [۱۱] از *ساندویچ کردن تقریبی تصادفی* استفاده کردند. یعنی تبدیل آفینی بیابیم که به احتمال به دلخواه زیاد، با حذف قسمت ناچیزی از K نسبت ساندویچی کوچک باشد (توجه کنید که با حذف قسمت ناچیزی از K ، حجم آن و همچنین توزیع یک نواخت روی آن به مقدار ناچیزی تغییر می کنند). به این ترتیب می توان به نسبت ساندویچی $R/r = O(\sqrt{n})$ دست یافت. در این جا مراحل کار انجام شده در [۱۱] را به اختصار توضیح می دهیم.

¹ Lowner

² John

معیار دیگری که برای گرد بودن شکل استفاده می‌شود، مفهومی به نام موقعیت آیزوتروپیک است.

فرض کنید $X = [X_1, \dots, X_n]^t$ یک بردار تصادفی در $\mathbb{R}^n$ باشد. می‌گوییم X در موقعیت آیزوتروپیک^۱ قرار دارد، هرگاه داشته باشیم

$$\begin{aligned}\mathbb{E}(X) &= 0, \\ \mathbb{E}(XX^t) &= I\end{aligned}$$

که در آن I ماتریس واحد است. این روابط معادل اند با

$$\begin{aligned}\mathbb{E}(X_i) &= 0, \\ \mathbb{E}(X_i X_j) &= \delta_{ij}\end{aligned}$$

هم‌چنین رابطه دوم معادل است با این که برای هر بردار $v \in \mathbb{R}^n$ داشته باشیم

$$\mathbb{E}((v^t X)^2) = |v|^2.$$

می‌گوییم یک شکل S در موقعیت آیزوتروپیک است هرگاه متغیر تصادفی با توزیع یک‌نواخت روی

آن در موقعیت آیزوتروپیک باشد. به عبارت دیگر

$$\int_K x dx = 0,$$

$$\frac{1}{\text{vol}(K)} \int_K (v^t x)^2 dx = |v|^2.$$

گزاره زیر از [۱۱] ارتباط بین موقعیت آیزوتروپیک و نسبت ساندریچی را بیان می‌کند.

۱.۲.۶. گزاره: اگر شکل محدب K در موقعیت آیزوتروپیک باشد، آن‌گاه

$$\mathbb{B}_{\sqrt{(n+2)/n}}(0) \subseteq K \subseteq \mathbb{B}_{\sqrt{n(n+2)}}(0).$$

همواره یک تبدیل آفین وجود دارد که K را به موقعیت آیزوتروپیک می‌برد. کافی است قرار دهید

$b := \mathbb{E}(X)$ و $A := \mathbb{E}(XX^t)$ که X متغیر تصادفی با توزیع یک‌نواخت روی K است. در این

¹ Isotropic position

صورت $A^{-\frac{1}{2}}(K - b)$ در موقعیت آیزوتروپیک قرار دارد. اما با توجه به این که K توسط یک اوراکل داده شده است، نمی‌توانیم این تبدیل را دقیقاً محاسبه کنیم. اما می‌توانیم K را «تقریباً آیزوتروپیک» کنیم.

۲.۲.۶. **تعریف:** می‌گوییم شکل $K \subseteq \mathbb{R}^n$ ، θ -آیزوتروپیک است، هرگاه برای هر بردار v داشته باشیم

$$\left| \frac{1}{\text{vol}(K)} \int_K x dx \right| < \theta,$$

$$(1 - \theta) |v|^2 < \frac{1}{\text{vol}(K)} \int_K (v^t x)^2 dx < (1 + \theta) |v|^2.$$

اگر K در موقعیت θ -آیزوتروپیک قرار داشته باشد، آن‌گاه

$$\frac{1}{\text{vol}(K)} \int_K |x|^2 dx < n(1 + \theta).$$

بنابراین حجم قسمتی از K که خارج از گوی به شعاع $\sqrt{\frac{n(1 + \theta)}{\epsilon}}$ حول مبدأ قرار دارد حداکثر $\epsilon \text{vol}(K)$ است.^۱ همچنین گزاره‌ای شبیه به گزاره ۱.۲.۶ برای موقعیت تقریباً آیزوتروپیک برقرار است. بنابراین اگر بتوانیم برای یک θ ثابت به موقعیت θ -آیزوتروپیک دست یابیم، با حذف کسر ϵ از حجم K به نسبت ساندویچی از مرتبه $\sqrt{n}$ دست می‌یابیم.

اما چگونه باید K را به موقعیت تقریباً آیزوتروپیک تبدیل کنیم؟ با توجه به این که K توسط یک اوراکل داده شده است، تنها راه استفاده از نمونه‌گیری است: تعدادی نقطه از K به طور مستقل با توزیع تقریباً یک‌نواخت انتخاب می‌کنیم و مجموعه متناهی حاصل را با یک تبدیل آفین به موقعیت (دقیقاً) آیزوتروپیک می‌بریم. در این صورت این تبدیل آفین K را با احتمال زیاد به موقعیت تقریباً آیزوتروپیک

^۱ با استفاده از قضیه‌ای از بورل [19] می‌توان مرتبه این شعاع را نسبت به ϵ لگاریتمی کرد.

می‌برد. در [۱۱] نشان داده شده است که تعداد $O^*(n^2)$ از نقاط کافی است. این تعداد توسط بورگین^۱ [۱۳] و رودلسون^۲ [۱۴] به $O(n \log^2 n)$ کاهش یافته است.

با این که این کار به نظر دوری می‌آید (برای به‌تر کردن نمونه‌گیری، از نمونه‌گیری استفاده می‌کنیم)، اما می‌توانیم از تکنیک Bootstrap استفاده کنیم و به طور استقرایی، در هر گام زیرمجموعه بزرگ‌تری از K را به موقعیت تقریباً آیزوتروپیک ببریم. مشابه الگوریتم حجم، قرار دهید $K_i := K \cap \mathbb{B}_{r, 2^i/n}(\circ)$. در این صورت K_i یک گوی است و با یک تجانس به موقعیت آیزوتروپیک برده می‌شود. با فرض این که K_i را به موقعیت ۲-آیزوتروپیک برده باشیم، می‌توان نشان داد که K_{i+1} ، ۶-آیزوتروپیک است. بنابراین می‌توان به صورت بهینه (بدون مشکل تیز بودن شکل) از K_{i+1} نمونه‌گیری کرد. حال با استفاده از نمونه‌گیری، K_{i+1} را به موقعیت ۲-آیزوتروپیک تبدیل می‌کنیم و این کار را تکرار می‌کنیم. تعداد فازها $O(n \log R/r)$ است و در هر فاز به $O^*(n)$ نقطه نیاز است. بنابراین با توجه به این که نمونه‌گیری‌ها را در اشکال ۶-آیزوتروپیک انجام می‌دهیم، طبق قضیه ۳.۸.۴ مرتبه الگوریتم $O^*(n^5)$ خواهد بود.

پس از این که به روش فوق به نسبت ساندویچی $O^*(\sqrt{n})$ دست یافتیم، مرتبه الگوریتم حجمی که در فصل ۵ آوردیم، $O^*(n^5)$ می‌شود (بستگی آن به R/r نیز لگاریتمی است). لواژ و ومپالا^۳ در [۴] به یک الگوریتم حجم از مرتبه $O^*(n^4)$ دست یافتند که سریع‌ترین الگوریتم حجم یافته شده تا کنون است. در این مورد در بخش ۴.۶ توضیح می‌دهیم.

¹ Bourgain

² Rudelson

³ Vempala

۳.۶. نمونه‌گیری از توزیع‌های لوگ-مقعر

روش مونت کارلو برای نمونه‌گیری از یک توزیع دلخواه نیز می‌تواند به کار رود. هریک از قدم‌زدن‌های تصادفی مذکور در بخش ۲.۱ را می‌توان طوری اصلاح کرد که توزیع حدی آن یک توزیع با تابع چگالی دلخواه باشد. روش مورد استفاده، روش متروپولیس-هستینگز است که توسط نیکلاس متروپولیس^۱ در سال ۱۹۵۳ ابداع شد. شکل ساده‌ی این روش به صورت زیر است.

۱.۳.۶. الگوریتم متروپولیس-هستینگز^۲: فرض کنید $\mathcal{M}$ یک فرآیند مارکوف برگشت‌پذیر با تابع

چگالی احتمال گذر متقارن $f_u(v)$ نسبت به یک توزیع پایای Q باشد. هم‌چنین فرض کنید Q' توزیع دیگری با تابع چگالی g نسبت به Q باشد. $\mathcal{M}$ را به شکل زیر تغییر می‌دهیم تا Q' یک توزیع پایا برای آن باشد. برای نقطه شروع x ، ابتدا نقطه y را با توزیع P_x (توزیع احتمال گذر $\mathcal{M}$) انتخاب کنید. حال به احتمال $\min\{1, \frac{g(y)}{g(x)}\}$ به نقطه y بروید و در غیر این صورت در x

بمانید. به عبارت دیگر تابع چگالی احتمال گذر جدید (نسبت به π) را به شکل زیر تعریف می‌کنیم:

$$f'_x(y) = f_x(y) \min\{1, \frac{g(y)}{g(x)}\}. \quad (y \neq x)$$

توجه کنید که این تابع چگالی روی $\Omega \setminus \{x\}$ تعریف شده است. احتمال ثابت ماندن در x هم برابر

با $1 - \int_{\Omega \setminus x} f'_x(y) dQ(y)$ است. طبق نکته ۷.۱.۳ به راحتی دیده می‌شود که فرآیند مارکوف جدید

نسبت به توزیع Q' برگشت‌پذیر است. بنابر این طبق گزاره ۶.۱.۳، Q' یک توزیع پایا برای آن است.

توجه کنید که مشابه روش مونت کارلو، باز هم کافی است g را با تقریب ضرب در یک ثابت بدانیم.

¹ Nicholas Metropolis

² Metropolis-Hastings

زمان آمیختگی این قدم زدن تصادفی جدید بستگی زیادی به Q' دارد. در موردی که M یکی از سه الگوریتم مذکور در بخش ۲.۱ باشد، این مسئله برای توزیع‌های با تابع چگالی لوگ-مقعر تحلیل شده است. بسیاری از توزیع‌های مهم، مانند توزیع گاوسی تحدید شده به K ، لوگ-مقعر هستند. نمونه‌گیری از توزیع‌های لوگ-مقعر از همان اوایل کار، بعد از مقاله دایر، فریز و کنان مورد توجه بوده است [۱۵]. در ابتدا هدف این بود که تابع مشخصه K (که نمونه‌گیری از توزیع با تابع چگالی متناسب با آن انجام می‌شود) با یک تابع هموار تقریب زده شود، زیرا برای توابع چگالی هموار نقاط گوشه‌ای راحت‌تر کنترل می‌شوند. اما به مرور زمان با پیدا شدن کاربردهایی در بهینه‌سازی، خود این توزیع‌ها هم اهمیت پیدا کردند.

خواص توزیع‌های لوگ-مقعر به طور گسترده مورد مطالعه قرار گرفته‌اند. برای مثال لواژ و ومپالا در [۱۶] به بررسی خواص هندسی این توزیع‌ها می‌پردازند. از جمله ویژگی‌های مهمی که برای تحلیل الگوریتم‌ها لازم بوده است، نسخه‌هایی از نابرابری‌های محیطی برای توزیع‌های لوگ-مقعر است (رجوع کنید به نکته ۲.۴.۴). یک نمونه از این نابرابری‌ها قضیه ۳.۰ است که برای اثبات قضیه ۳.۴.۴ مورد استفاده قرار گرفت. با ارائه یک نابرابری محیطی مناسب، در [۱۷] کران زیر برای زمان آمیختگی این قدم زدن تصادفی ارائه شده است که هم مرتبه با کران موجود در قضیه ۲.۲.۴ است.

۲.۳.۶. قضیه: فرض کنید f یک تابع چگالی لوگ-مقعر و θ -آیزوتروپیک (تعریف ۲.۲.۶) در $\mathbb{R}^n$

باشد. π_f را توزیع با تابع چگالی متناسب با f نسبت به اندازه لبگ بگیرید. فرض کنید چگالی

توزیع اولیه نسبت به π_f حداکثر M باشد. در این صورت به ازای

$$m > 10^4 \theta^2 M^2 \frac{n^3}{\epsilon^2} \log \frac{1}{\epsilon}$$

فاصله مجموع تغییرات توزیع قدم زدن تصادفی Hit-and-run اصلاح شده از π_f حداکثر ϵ است.

۴.۶. روش Simulated Annealing

در این بخش، خلاصه‌ای از ایده‌های الگوریتم حجم معرفی شده در [۴] توسط لواژ و ومپالا را بیان می‌کنیم. این الگوریتم از مرتبه $O^*(n^4)$ است که سریع‌ترین الگوریتم حجم معرفی شده تاکنون است. همان‌طور که دیدیم، تعداد فازها در الگوریتم حجم ۱.۱.۵ از مرتبه $m = O(n \log R/r)$ است. باید سعی کنیم تعداد فازها را تا حد امکان کوچک انتخاب کنیم. زیرا هرچه تعداد فازها بیش‌تر باشد، علاوه بر این که تعداد نسبت‌های لازم برای محاسبه بیش‌تر می‌شود، باید برای هریک از نسبت‌ها خطای نسبی کوچک‌تری در نظر بگیریم. همان‌طور که در اثبات قضیه ۸.۱.۵ دیدیم، در مرتبه الگوریتم حجم عامل m^2 ظاهر می‌شود. اما می‌توان نشان داد که با روش انتخاب اشکال محدب تودرتو، به تعداد فازهای به‌تر از $\Omega(n)$ نمی‌توان دست یافت.

از سوی دیگر، محاسبه حجم حالت خاصی از محاسبه انتگرال توابع است. در [۴] به جای در نظر گرفتن دنباله‌ای از اشکال تودرتو، دنباله‌ای از توابع $f_1 \leq f_2 \leq \dots \leq f_m$ روی K ساخته می‌شود که f_m تابع مشخصه K است. نسبت انتگرال‌های $\int f_{i-1} / \int f_i$ توسط نمونه‌گیری تخمین زده می‌شود. برای این کار کافی است تعدادی نقطه تصادفی انتخاب کنیم که چگالی توزیع آن‌ها متناسب با f_i باشد و سپس میانگین تابع f_{i-1}/f_i روی این نمونه‌ها را حساب کنیم. مقدار $\int f_i$ نیز تخمین زده می‌شود و به این ترتیب تخمینی برای $\int f_m = \text{vol}(K)$ حاصل می‌شود.

در صورتی که f_i را تابع مشخصه K_i بگیریم، این روش تبدیل به همان روش مونت کارلو چندفازی می‌شود که در الگوریتم ۱.۱.۵ استفاده شده است. در مقاله مذکور از توابع به شکل $f(x) = e^{-x/T}$ استفاده می‌شود که در آن x مؤلفه اول x است. اگر T زیاد باشد، توزیع القاشده توسط f نزدیک به توزیع یک‌نواخت است و اگر T کم باشد، این توزیع متمرکز روی نقاطی از K می‌شود که x کم‌تری

دارند (البته قبل از این کار، تغییری در K ایجاد می‌شود که در دمای پایین مقدار $\int f$ قابل تخمین باشد. بعداً در این مورد توضیح می‌دهیم). همان‌طور که در بخش ۳.۶ دیدیم، می‌توان قدم‌زدن تصادفی Hit-and-run را توسط روش متروپولیس-هستینگز طوری اصلاح کرد که چگالی توزیع حدی آن متناسب با $f(x) = e^{-x/T}$ باشد. در [۱۷] و سپس در [۹] کران‌هایی برای زمان آمیختگی این قدم‌زدن تصادفی ارائه شده است. در این کران از معیار دیگری برای «گرد بودن» شکل استفاده شده است که در این جا به آن اشاره نمی‌کنیم. به این ترتیب به روشی مشابه الگوریتم ۱.۱.۵ می‌توان یک الگوریتم حجم ارائه کرد.

روش فوق بر اساس روش Simulated annealing است که توسط کرکپاتریک^۱ معرفی شد. Simulated annealing یک روش تصادفی عمومی برای بهینه‌سازی است. در این روش یک قدم‌زدن تصادفی در فضای جواب‌های ممکن انجام می‌شود که تابع احتمال گذر آن در چند مرحله توسط پارامتری به نام دما تغییر می‌کند. در یک دمای بالا، توزیع قدم‌زدن تصادفی به سرعت به توزیع یک‌نواخت روی تمام فضا میل می‌کند و هنگامی که دما پایین می‌آید، توزیع حدی آن به سمت جواب بهینه متمایل می‌شود. الگوریتم‌های مبتنی بر روش Simulated annealing معمولاً در عمل خوب کار می‌کنند، اما ارائه اثبات تئوری برای عمل کردن آن‌ها معمولاً مشکل است. در بالا ما عکس این کار را انجام دادیم. یعنی از یک دمای پایین (توزیع متمرکز) شروع کردیم و دما را افزایش دادیم.

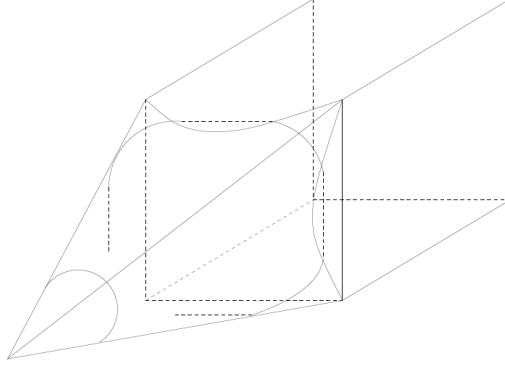
همان‌طور که گفتیم، باید در K تغییری ایجاد کنیم که مقدار $\int f$ قابل تخمین باشد. دیدیم که f متمرکز روی نقاطی از K است که x کوچک‌تری دارند. برای این کار مخروط زیر را در $\mathbb{R}^{n+1}$ در نظر بگیرید

$$C_a := \{(t, x) : 0 \leq t \leq a, x \in \mathbb{R}^n, |x| \leq t\}$$

¹ Kirkpatrick

و قرار دهید

$$K' := C_{\sqrt{R}} \cap [0, \sqrt{R}] \times K.$$



شکل ۱۷

بنابراین $C_r \subseteq K' \subseteq C_{\sqrt{R}}$. حال به جای K از K' استفاده کنید. با توجه به تمرکز f روی C_r در دماهای پایین، با محاسبه انتگرال f روی C_r و $C_{\sqrt{R}}$ تخمینی برای انتگرال آن روی K به دست می‌آید. پس از تخمین حجم K' با توجه به این که $[R, \sqrt{R}] \times K \subseteq K'$ و حجم آن لااقل نصف حجم K' است، با روش مونت کارلو می‌توان تخمینی برای حجم $[R, \sqrt{R}] \times K$ و بنابراین حجم K به دست آورد.

برای تحلیل این روش در مقاله مذکور از همان مراحل تحلیل الگوریتم ۱.۱.۵ استفاده شده است که البته جزئیات آن پیچیده تر است. در نهایت با انتخاب مناسب دماها تعداد فازها از مرتبه $O^*(\sqrt{n})$ شده است. هم‌چنین تغییری در نحوه گرد کردن شکل ایجاد شده است که مرتبه آن را به $O^*(n^{\frac{1}{4}})$ کاهش می‌دهد. به این ترتیب با کم شدن تعداد فازها، مرتبه الگوریتم حجم $O^*(n^{\frac{1}{4}})$ می‌شود.

ضمیمه أ. لم موضعی سازی

لم های موضعی سازی را اولین بار لواژ و سیمونویتز مطرح کردند. این لم ها ابزارهایی قوی برای تبدیل مسائل انتگرالی چندبعدی به مسائل یک بعدی هستند. برای مثال در [۱۸] چند مسئله مربوط به نابرابری های محیطی (بخش ۴.۴) و هم چنین موقعیت آیزوتروپیک (بخش ۲.۶) توسط لم های موضعی سازی اثبات شده است. در این ضمیمه دو نسخه از لم های موضعی سازی را می آوریم (لم های ۱.أ و ۲.أ) و سپس توسط آن ها یک نابرابری محیطی را ثابت می کنیم (قضیه ۳.أ). این نتایج در اثبات قضیه ۳.۴.۴ استفاده شده اند.

۱.أ. لم: فرض کنید g و h توابعی از پایین پیوسته و انتگرال پذیر روی $\mathbb{R}^n$ باشند به طوری که

$$\int_{\mathbb{R}^n} g(x) dx > 0 \text{ و } \int_{\mathbb{R}^n} h(x) dx > 0.$$

در این صورت دو نقطه $a, b \in \mathbb{R}^n$ و تابع خطی $l : [0, 1] \rightarrow \mathbb{R}^{\geq 0}$ موجوداند به طوری که

$$\int_0^1 l(t)^{n-1} g((1-t)a + tb) dt > 0 \text{ و } \int_0^1 l(t)^{n-1} h((1-t)a + tb) dt > 0.$$

می توانیم حکم را به صورت شهودی به این طریق بیان کنیم که اگر g و h انتگرال های مثبت داشته باشند، آن گاه یک مخروط ناقص با قاعده های بسیار کوچک یافت می شود که تحدید g و h به آن نیز انتگرال های مثبت داشته باشند. می توان نشان داد که به جای مخروط ناقص نمی توان اشکال خاص تری مانند استوانه (یعنی l یک تابع ثابت باشد) یا مخروط (یعنی l در یکی از دو انتهایش صفر باشد) گذاشت. هم چنین حکم در مورد سه انتگرال درست نیست. درواقع اگر بخواهیم علامت k انتگرال حفظ شود، باید از اشکال محدب $k-1$ بعدی استفاده کنیم.

اثبات ۱.۱: ادعا می‌کنیم کافی است حکم را برای توابع پیوسته و انتگرال پذیر ثابت کنیم. در حالت

کلی h و g را می‌توان به صورت حد صعودی توابع پیوسته و انتگرال پذیر $g_k \rightarrow g$ و $h_k \rightarrow h$

نوشت. حال طبق لم فانو^۱ خواهیم داشت

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} g_k = \int_{\mathbb{R}^n} g > 0 \text{ و } \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} g_k = \int_{\mathbb{R}^n} g > 0.$$

پس برای یک k به اندازه کافی بزرگ خواهیم داشت $\int_{\mathbb{R}^n} g_k > 0$ و $\int_{\mathbb{R}^n} h_k > 0$. اگر درستی

حکم را برای توابع پیوسته بدانیم، دو نقطه $a, b \in \mathbb{R}^n$ و یک تابع خطی نامنفی l می‌یابیم که

$$\int_0^1 l(t)^{n-1} h_k((1-t)a + tb) dt > 0 \text{ و } \int_0^1 l(t)^{n-1} g_k((1-t)a + tb) dt > 0.$$

و چون $g \geq g_k$ و $h \geq h_k$ حکم ثابت می‌شود.

هم‌چنین کافی است ثابت کنیم انتگرال‌های موجود در حکم بزرگتر یا مساوی صفر هستند. زیرا

در این صورت، اگر این حکم را در مورد $g - a$ و $h - a$ به کار ببریم که a یک تابع همه‌جا مثبت با

انتگرال کوچک است، حکم اصلی نتیجه می‌شود.

ادعا: دنباله‌ای از مجموعه‌های محدب مانند $K_0 \supseteq K_1 \supseteq K_2 \supseteq \dots$ موجود است که

$$\int_{K_i} h(x) dx > 0 \text{ و } \int_{K_i} g(x) dx > 0. \quad (1.7)$$

و هم‌چنین $K := \bigcap_i K_i$ یک نقطه یا یک پاره خط است.

برای اثبات این ادعا، K_0 را می‌توانیم یک گوی به اندازه کافی بزرگ انتخاب کنیم. پس از انتخاب

K_i ، یک نیم‌فضای H را در نظر بگیرید به طوری که

$$\int_{K_i \cap H} g(x) dx = \frac{1}{2} \int_{K_i} g(x) dx.$$

¹ Fano's lemma

مرز چنین نیم فضایی را یک ابرصفحه نیم ساز می نامیم. اگر در صورت لزوم H را با نیم فضای مکملش جایگزین کنیم، می توانیم فرض کنیم

$$\int_{K_i \cap H} h(x) dx > 0.$$

بنابراین $K_{i+1} := K_i \cap H$ شرط (۱.۷) را ارضا می کند. حال باید ثابت کنیم می توانیم در هر مرحله H را طوری انتخاب کنیم که K_i ها به یک نقطه یا یک پاره خط نزول کنند. برای این منظور از این نکته استفاده می کنیم که به ازای هر زیر فضای آفین $(n-2)$ -بعدی دلخواه مانند A ، لااقل یک ابرصفحه نیم ساز شامل آن موجود است. این موضوع به سادگی از قضیه مقدار میانی با چرخاندن یک ابرصفحه شامل A نتیجه می شود.

$A_0, A_1, \dots$ را دنباله تمام زیر فضاهای آفین $(n-2)$ -بعدی گویا بگیرید (یعنی زیر فضاهایی که با دو معادله خطی با ضرایب گویا تعریف می شوند). با قطع کردن K_i توسط یک ابرصفحه نیم ساز P_i که شامل A_i است و سپس انتخاب نیم فضای مناسب K_{i+1} را تعریف کنید. نشان می دهیم $K := \bigcap_i K_i$ حداکثر یک بعدی است. فرض کنید این طور نباشد. پس چون K محدب است، حداقل دوبعدی است. پس تصویر آن روی یکی از صفحه هایی که با دو تا از بردارهای یکه استاندارد تولید می شوند (مثلاً دو بردار اول) نیز هم چنان دوبعدی است و بنابراین شامل نقطه درونی گویایی مثل (r_1, r_2) است. صفحه $(n-2)$ -بعدی ای که با معادلات $x_1 = r_1$ و $x_2 = r_2$ تعریف می شود یکی از A_i ها است. اما چون K_{i+1} کاملاً یک طرف P_i است، K هم باید کاملاً یک طرف P_i باشد و این با انتخاب (r_1, r_2) متناقض است. پس ادعا ثابت می شود.

اگر K از تنها یک نقطه a تشکیل شده باشد، از پیوستگی g و h نتیجه می شود که $g(a) \geq 0$ و $h(a) \geq 0$. بنابراین حکم به ازای $a := b$ و $1 \equiv l$ برقرار است. پس فرض کنید K یک پاره خط با طول ناصفر است و a و b دو سر K هستند.

ادعا: تابع مقعر $\psi : [0, 1] \rightarrow \mathbb{R}^{\geq 0}$ که متحد با صفر نیست موجود است به طوری که

$$\int_0^1 \psi(t)^{n-1} g((1-t)a + tb) dt \geq 0, \quad (2.7)$$

$$\int_0^1 \psi(t)^{n-1} g((1-t)a + tb) dt \geq 0.$$

بدون کم شدن از کلیت می توانیم فرض کنیم $a = 0$ و $b = e_1$ که اولین بردار یکه استاندارد

است. تعریف کنید

$$Z_t := \{x \in \mathbb{R}^n : x_1 = t\}$$

و

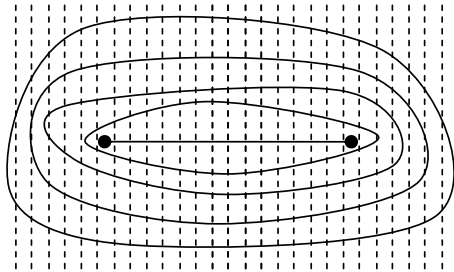
$$\psi_i(t) := \left(\frac{\text{vol}_{n-1}(K_i \cap Z_t)}{\text{vol}(K_i)} \right)^{\frac{1}{n-1}}$$

که در آن x_i مؤلفه i -ام x در مختصات استاندارد $\mathbb{R}^n$ است. همچنین فرض کنید α_i و β_i

به ترتیب مینیمم و ماکزیمم مؤلفه اول نقاط K_i باشند. به وضوح $\alpha_i \leq 0$ و $\beta_i \geq 1$ و همین طور با

افزایش i ، α_i به صفر و β_i به یک میل می کند. به علاوه، طبق قضیه برون-مینکوفسکی (قضیه ب.۵)،

ψ_i تابعی مقعر روی بازه $[\alpha_i, \beta_i]$ است.



شکل ۱۸

ادعا می کنیم که می توانیم زیر دنباله ای از توابع $\psi_i(t)$ را بیابیم که برای هر $0 \leq t \leq 1$ به تابعی

مثل $\psi(t)$ میل کنند. توجه کنید که چون ψ_i مقعر است و $\int_0^1 \psi_i(t)^{n-1} dt = 1$ ، پس ψ_i ها به طور

یک نواخت کران دار هستند. فرض کنید M کران بالایی برای آن ها باشد. طبق تقعر ψ_i ، برای هر

$$0 < \epsilon < s < t < 1 - \epsilon$$

$$-\frac{M}{\epsilon} \leq -\frac{\psi_i(1-\epsilon)}{1-\epsilon} \leq \frac{\psi_i(t) - \psi_i(s)}{t-s} \leq \frac{\psi_i(\epsilon)}{\epsilon} \leq \frac{M}{\epsilon}.$$

پس ψ_i ها روی هر بازه $[\epsilon, 1-\epsilon]$ به طور یک نواخت هم پیوسته هستند و وجود یک زیر دنباله

هم گرا از آن ها به سادگی با استفاده از قضیه آرزلا-آسکولی^۱ نتیجه می شود.

به وضوح تابع حدی ψ نامنفی و مقعر است و $\int_0^1 \psi(t)^{n-1} dt = 1$. حال اگر نقطه $x \in \mathbb{R}^n$ را با

(t, y) نشان دهیم که $t \in \mathbb{R}$ و $y \in \mathbb{R}^{n-1}$ خواهیم داشت

$$\begin{aligned} \int_{K_i} g(x) dx &= \int_{\alpha_i}^{\beta_i} \int_{K_i \cap Z_t} g(t, y) dy dt = \\ \int_{\alpha_i}^{\beta_i} \left[\frac{1}{\text{vol}_{n-1}(K_i \cap Z_t)} \int_{K_i \cap Z_t} g(t, y) dy \right] \text{vol}(K_i) \psi_i(t)^{n-1} dt. \end{aligned}$$

بنابراین

$$\frac{1}{\text{vol}(K_i)} \int_{K_i} g(x) dx = \int_{\alpha_i}^{\beta_i} \left[\frac{1}{\text{vol}_{n-1}(K_i \cap Z_t)} \int_{K_i \cap Z_t} g(t, y) dy \right] \psi_i(t)^{n-1} dt.$$

طرف چپ نامنفی است در حالی که طرف راست به $\int_0^1 g(t, \circ) \psi(t)^{n-1} dt$ میل می کند. پس این

انتگرال نامنفی است. این استدلال برای h هم کار می کند و ادعا ثابت می شود.

a ، b و ψ را طوری انتخاب کنید که دو شرط زیر برقرار باشند.

¹ Arzela-Ascoli

(الف) a و b را دو نقطه از K با کمترین فاصله بگیرید که نامساوی‌های (۲.۷) به ازای یک تابع مقعر غیر متحد با صفر $\psi \geq 0$ برقرار باشد. دوباره می‌توانیم فرض کنیم $a = 0$ و $b = e_1$. همچنین می‌توانیم فرض کنیم ψ پیوسته است (تنها در 0 و 1 ممکن است ناپیوسته باشد). تابع ψ و $1 \leq \alpha \leq \beta \leq 0$ را طوری انتخاب کنید که

(ب) ψ روی بازه $[\alpha, \beta]$ خطی باشد و $\beta - \alpha$ ماکزیمم باشد

با استفاده از فشردگی به سادگی امکان این انتخاب‌ها ثابت می‌شود. اگر ψ خطی باشد کار تمام است. در ادامه، با فرض این که ψ خطی نیست به تناقض می‌رسیم.

از این به بعد فقط با مقادیر g و h روی پاره خط واصل a و b سروکار داریم. تعریف کنید

$$\hat{h}(x) := h(x_1) \text{ و } \hat{g}(x) := g(x_1)$$

همچنین برای هر مجموعه محدب L ، تعریف کنید

$$\psi_L(t) := \text{vol}_{n-1}(L \cap Z_t)^{\frac{1}{n-1}}.$$

مجموعه K' را به شکل زیر تعریف کنید

$$0 \leq x_1 \leq 1, x_2, \dots, x_n \geq 0,$$

$$x_2 + \dots + x_n \leq \psi(x_1)$$

به عبارت دیگر $K \cap Z_t$ یک سادک $(n-1)$ -بعدی با رئوس $te_1, te_1 + \psi(t)e_2, \dots, te_1 + \psi(t)e_n$ و

$te_1 + \psi(t)e_n$ است. طبق تقعر ψ ، K' یک مجموعه محدب است. حال

$$\int_{K'} \hat{g}(x) dx = \frac{1}{n!} \int_0^1 \psi(t)^{n-1} g(te_1) dt \geq 0.$$

و به طور مشابه

$$\int_{K'} \hat{h}(x) dx \geq 0.$$

طبق شرط (الف) یکی از این انتگرال‌ها باید صفر باشد. پس بدون کم شدن از کلیت مسئله می‌توانیم

فرض کنیم

$$\int_{K'} \hat{g}(x) dx = 0.$$

دو عدد حقیقی $0 < \sigma < 1$ و $\tau > 0$ و زیرفضای آفین $(n-2)$ -بعدی A که با روابط

$x_1 = \sigma$ و $x_1 + \dots + x_n = \tau$ تعریف می‌شود را در نظر بگیرید. فرض کنید A با درون K'

تقاطع داشته باشد؛ یعنی $\psi(\sigma) < \tau$. ابرصفحه H شامل A موجود است که K' را به دو مجموعه

محدب L_A و L'_A تقسیم می‌کند که

$$\int_{L_A} \hat{g}(x) dx = \int_{L'_A} \hat{g}(x) dx = 0.$$

می‌توانیم نمادها را طوری انتخاب کنیم که

$$\int_{L_A} \hat{h}(x) dx \geq 0.$$

L_A باید با هر دوی Z_0 و Z_1 اشتراک داشته باشد زیرا در غیر این صورت ψ_{L_A} در شرط (الف) در

بازه کوچک‌تری صدق می‌کند و این با (الف) متناقض است.

$$\begin{aligned} L_A \cap Z_0 &\neq \emptyset, \\ L'_A \cap Z_0 &\neq \emptyset. \end{aligned} \tag{3.7}$$

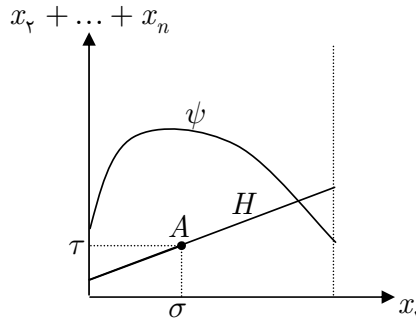
اکنون دو حالت داریم: یا L_A «روبه پایین» است (یعنی شامل نقطه σe_1 است) و یا «روبه بالا» است.

طبق شرط (3.7)، H نمی‌تواند بر محور اول مختصات عمود باشد و چون شامل A است، توسط

معادله‌ای به شکل $x_1 + \dots + x_n = l(x_1)$ داده می‌شود که l یک تابع خطی است. می‌خواهیم با

انتخاب مناسب A با یکی از (الف) و (ب) به تناقض برسیم. تصویر این اشکال روی صفحه دوبعدی

تولید شده توسط دو بردار e_1 و $e_1 + \dots + e_n$ به شکل زیر است.



شکل ۱۹

توجه کنید که اگر L_A «روبه پایین» باشد، آن گاه طبق (۳.۷) باید داشته باشیم $l(1), l(0) \geq 0$ و

نتیجه می گیریم

$$\psi_{L_A} = \min\{\psi(t), l(t)\}. \quad (4.7)$$

حالت اول: $\psi(0) = \psi(1) = 0$. قرار دهید $\sigma := \frac{1}{\psi}$. طبق (۳.۷)، L_A نمی تواند «روبه بالا» باشد.

پس از $0 \leq l(1), l(0)$ و $l(\frac{1}{\psi}) = \tau$ نتیجه می گیریم که اگر τ را به صفر میل دهیم، $l(t)$ روی بازه

$[0, 1]$ به صورت یک نواخت به صفر میل می کند. حال طبق (۴.۷)، ψ_{L_A} روی یک بازه خطی است که

طول این بازه به یک میل می کند و این با شرط (ب) متناقض است.

حالت دوم: $\psi(0) = 0$ و $\psi(1) > 0$. قرار دهید $\tau := \psi(1)\sigma$. اگر L_A «روبه بالا» باشد، آن گاه

طبق (۳.۷) باید داشته باشیم $l(t) = \psi(1)t$. حال با در نظر گرفتن ψ_{L_A} طبق حالت اول به تناقض

می رسیم. بنابراین فرض کنید L_A «روبه پایین» است. حال (۳.۷) ایجاب می کند که

$0 \leq l(1) \leq \psi(1)$ و طبق (۴.۷)، ψ_{L_A} روی بازه $[\sigma, 1]$ خطی است که با کوچک انتخاب کردن σ

متناقض با شرط (ب) است.

حالت سوم: $\psi(\circ), \psi(1) > \circ$. تمام توابع خطی $\eta: [\circ, 1] \rightarrow \mathbb{R}^{\geq \circ}$ را در نظر بگیرید به طوری که

$$\psi(\circ) \leq \eta(\circ) \leq \psi(1) \text{ و } \circ \leq \eta(1) \leq \psi(1) \text{ و همچنین روی شکل محدب } K_\eta \text{ که با روابط}$$

$$\circ \leq x_1 \leq 1, \quad x_1, \dots, x_n \geq \circ \\ \eta(x_1) \leq x_1 + \dots + x_n \leq \psi(x_1)$$

تعریف می شود داشته باشیم

$$\int_{K_\eta} \hat{h}(x) dx = \circ \text{ و } \int_{K_\eta} \hat{g}(x) dx = \circ$$

η ای را در نظر بگیرید که $\eta(\circ) + \eta(1)$ ماکزیمال باشد. اگر $\eta(\circ) = \psi(\circ)$ یا $\eta(1) = \psi(1)$

آن گاه با توجه به رابطه $\psi_{K_\eta} = \psi - \eta$ طبق حالت اول یا دوم به تناقض می رسیم. پس فرض کنید

$$\eta(\circ) < \psi(\circ) \text{ و } \eta(1) < \psi(1)$$

(σ, τ) را تقاطع پاره خط واصل دو نقطه $(\circ, \eta(\circ))$ و $(1, \psi(1))$ (در $\mathbb{R}^2$) با پاره خط واصل

$(\circ, \psi(\circ))$ و $(1, \eta(1))$ بگیرید. همچنین یک ابرصفحه H شامل A را طوری در نظر بگیرید که K_η

را به دو بخش M و M' تقسیم کند که

$$\int_M \hat{g}(x) dx = \int_{M'} \hat{g}(x) dx = \circ$$

فرض کنید M «روبه بالا» و M' «روبه پایین» باشد. توجه کنید که طبق انتخاب (σ, τ) ، یا

هردوی M, M' با هردوی Z_+ و Z_- اشتراک دارند و یا هیچ کدام از آن ها با هردوی Z_+ و Z_-

اشتراک ندارند. حالت دوم طبق (۳.۷) قابل قبول نیست. بنابراین K_l و طبق ماکزیمال بودن η باید

داشته باشیم

$$\int_M \hat{h}(x) dx < \circ$$

اما $M' \setminus K'$ یک مخروط ناقص است و داریم

$$\int_{K' \setminus M} g(x) dx = \int_{K'} g(x) dx - \int_M g(x) dx = 0.$$

و هم چنین

$$\int_{K' \setminus M} h(x) dx = \int_{K'} h(x) dx - \int_M h(x) dx > 0.$$

که با شرط (ب) متناقض است. $\square$

۲.۱. لم: فرض کنید K یک زیرمجموعه محدب و کران دار در $\mathbb{R}^n$ ، g یک تابع پیوسته روی K و h

یک تابع کران دار و از پایین پیوسته روی K باشند به طوری که

$$\int_T g(x) dx = 0 \text{ و } \int_T h(x) dx = 0.$$

در این صورت دو نقطه $a, b \in K$ و تابع خطی نامنفی $\mathbb{R} \rightarrow [0, 1]: l$ یافت می شوند که

$$\int_0^1 l(t)^{n-1} h((1-t)a + tb) dt = 0 \text{ و } \int_0^1 l(t)^{n-1} g((1-t)a + tb) dt = 0.$$

اثبات: می توانیم از اثباتی مشابه اثبات لم قبل استفاده کنیم. اما در این جا این مسئله را با استفاده از

لم قبل ثابت می کنیم. برای راحتی از یک نمادگذاری استفاده می کنیم. منظور از یک سوزن مثل N ، دو

نقطه $\mathbb{R}^n$ به همراه تابع خطی نامنفی و غیر متحد با صفر $\mathbb{R} \rightarrow [0, 1]: l$ است. برای هر تابع

$f: \mathbb{R}^n \rightarrow \mathbb{R}$ ، انتگرال f روی N را به این شکل تعریف می کنیم

$$\int_N f := \int_0^1 l(t)^{n-1} f((1-t)a + tb) dt.$$

مشابه اثبات لم قبل، کافی است فرض کنیم h پیوسته است. $\delta > 0$ را طوری انتخاب کنید که

$$\int_K (h - \delta) > 0. \text{ و } \epsilon > 0 \text{ را هم یک عدد حقیقی دلخواه بگیرید (که بعداً آن را به صفر میل}$$

می دهیم). داریم

$$\int_K (\epsilon^2 - g) > 0 \text{ و } \int_K (h - \delta + \frac{1}{\epsilon} g) > 0.$$

با ضرب در تابع مشخصه K ، این انتگرال‌ها را به عنوان انتگرال‌هایی روی $\mathbb{R}^n$ در نظر می‌گیریم و از لم قبل استفاده می‌کنیم. بنابراین سوزن N_ϵ موجود است که

$$\int_{N_\epsilon} (\epsilon^2 - g) > 0 \text{ و } \int_{N_\epsilon} (h - \delta + \frac{1}{\epsilon}g) > 0.$$

بنابراین

$$\int_{N_\epsilon} (h - \delta + \epsilon) > 0.$$

اگر M یک کران بالا برای h روی K باشد، آنگاه

$$\int_{N_\epsilon} g < \epsilon^2 \int_{N_\epsilon} 1 \text{ و } \int_{N_\epsilon} g > \int_{N_\epsilon} (-\epsilon h) \geq -M\epsilon \int_{N_\epsilon} 1$$

این سوزن‌ها را به‌ازای $\epsilon = \frac{1}{k}$ ($k = 1, 2, \dots$) در نظر بگیرید. با ضرب در یک ثابت مناسب،

می‌توانیم فرض کنیم که ماکزیمم هریک از توابع خطی متناظر با این سوزن‌ها برابر با یک است. حال یک زیر دنباله از این سوزن‌ها هم‌گرا می‌شود (به این معنی که نقاط انتهایی پاره‌خط‌های متناظر با آن‌ها هم‌گرا هستند و توابع خطی متناظر با آن‌ها هم‌گرا هستند). به سادگی با استفاده از پیوستگی g و h نتیجه می‌شود که سوزن حدی حاصل در حکم لم صدق می‌کند. $\square$

۳.۱. قضیه: فرض کنید $K := [0, 1] \subseteq \mathbb{R}$ و $f : K \rightarrow \mathbb{R}^+$ تابعی لوگ-مقعر باشد. در این صورت

برای هر افراز J_1, J_2, J_3 از K داریم

$$\int_K f(x)dx \int_{J_3} f(x)dx \geq d_K(J_1, J_2) \int_{J_1} f(x)dx \int_{J_2} f(x)dx.$$

این گزاره، نسخه‌ای از مسئله نابرابری محیطی برای توزیع‌های لوگ-مقعر است (رجوع کنید به

نکته ۲.۴.۴). با استفاده از لم موضعی سازی، دقیقاً مشابه اثبات قضیه ۳.۴.۴، می‌توان حالتی که

$K \subseteq \mathbb{R}^n$ یک شکل محدب دلخواه باشد را به این قضیه تقلیل داد. اما با توجه به این که ما تنها به حالت یک بعدی آن نیاز داریم (در اثبات قضیه ۳.۴.۴)، تنها این حالت را ثابت می کنیم.

اثبات: فرض کنید $d_K(J_1, J_2) = \epsilon$. می توانیم فرض کنیم J_1 و J_2 بسته هستند. همچنین می توانیم فرض کنیم برای هر بازه ماکسیمال در J_3 ، یکی از دو سر آن عضوی از J_1 و دیگری عضوی از J_2 است (در غیر این صورت این بازه را از J_3 حذف کنید و به یکی از J_1 و J_2 اضافه کنید). بنابراین چون $d_K(J_1, J_2) > 0$ و J_1 نیز اجتماعی از بازه های مجزا هستند.

ادعا: کافی است حکم را برای حالتی که $J_1 = [0, u]$ و $J_2 = [v, 1]$ ثابت کنیم.

فرض کنید حکم را برای این حالت ثابت کرده باشیم. پس طبق فرض های فوق، برای هر بازه ماکسیمال به شکل (a, b) در J_3 داریم

$$\int_a^1 f(x)dx \int_0^b f(x)dx \geq \epsilon \int_0^a f(x)dx \int_b^1 f(x)dx.$$

با جمع کردن روابط حاصل به دست می آوریم

$$\int_0^1 f(x)dx \int_{J_3} f(x)dx \geq \epsilon \sum \int_0^1 f(x)dx \int_b^1 f(x)dx.$$

که در آن مجموع روی تمام بازه های ماکسیمال (a, b) در J_3 گرفته می شود. فرض کنید $\{(a_i, b_i) : i = 1, 2, \dots\}$ تمام بازه های ماکسیمال در هریک از J_1 و J_2 و J_3 باشند و $c_i := \int_{a_i}^{b_i} f(x)dx$. حال مجموع فوق را می توان به شکل مجموعی از جملاتی مانند $c_i c_j$ نوشت. همچنین حاصل $\int_{J_1} f(x)dx \int_{J_2} f(x)dx$ را نیز می توان به شکل مجموع چنین جملاتی نوشت. فرض کنید $(a_i, b_i) \subseteq J_2$ و $(a_j, b_j) \subseteq J_1$. چون $d_K(J_1, J_2) > 0$ ، پس بین این دو بازه، بازه ای

از J_3 موجود است. بنابراین $c_i c_j$ در مجموع موجود در رابطه فوق ظاهر می‌شود. بنابراین

$$\int_{J_1} f(x) dx \int_{J_2} f(x) dx \text{ از مجموع موجود در این رابطه بیش تر نیست و ادعا ثابت می‌شود.}$$

حال حکم را برای حالتی که $J_1 = [0, u]$ و $J_2 = [v, 1]$ ثابت می‌کنیم. توجه کنید که اعداد c_i و

λ موجوداند که $f(u) = c_1 e^{\lambda u}$ و $f(v) = c_2 e^{\lambda v}$. پس چون f لوگ-مقعر است، در بازه $[u, v]$

داریم $f(x) \geq c_1 e^{\lambda x}$ و در بقیه نقاط به عکس. قرار دهید $x := \int_u^v f(t) dt$ ، $y := \int_u^v f(t) dt$ و

$$z := \int_v^1 f(t) dt. \text{ حکم این است که } \frac{y(x+y+z)}{xz} \geq \frac{v-u}{u(1-v)}. \text{ طرف چپ نسبت به } y$$

صعودی و نسبت به x و z نزولی است. پس کافی است f را با تابع $c_1 e^{\lambda x}$ جایگزین کنیم.

همچنین می‌توانیم فرض کنیم $c_1 = 1$. پس تنها کار باقی‌مانده این است که برای

$$0 \leq u < v \leq 1$$

$$\frac{(e^{\lambda v} - e^{\lambda u})(e^{\lambda} - 1)}{(e^{\lambda u} - 1)(e^{\lambda} - e^{\lambda v})} \geq \frac{v-u}{u(1-v)}.$$

توجه کنید که تابع $\frac{e^{\lambda t} - 1}{t}$ لوگ-محدب است زیرا

$$\frac{d^2}{dt^2} \ln\left(\frac{e^{\lambda t} - 1}{t}\right) = \frac{e^{\lambda t} + e^{-\lambda t} - 2 - \lambda^2 t^2}{t^2 e^{-t}(e^{\lambda t} - 1)^2}$$

و با در نظر گرفتن بسط تیلور صورت، دیده می‌شود که این مقدار نامنفی است. بنابراین برای هر

$$0 \leq a < b \text{ و } c \leq d \text{ که } a + d = b + c \text{ داریم}$$

$$\frac{e^{\lambda a} - 1}{a} \cdot \frac{e^{\lambda d} - 1}{d} \geq \frac{e^{\lambda b} - 1}{b} \cdot \frac{e^{\lambda c} - 1}{c}.$$

در حالت خاص

$$\frac{e^{\lambda(v-u)} - 1}{v-u} \cdot \frac{e^{\lambda} - 1}{1} \geq \frac{e^{\lambda v} - 1}{v} \cdot \frac{e^{\lambda(1-u)} - 1}{u}.$$

معادلاً

$$\frac{(e^{\lambda v} - 1)(e^{\lambda(1-u)} - 1)}{(e^{\lambda(v-u)} - 1)(e^{\lambda} - 1)} \leq \frac{v(1-u)}{v-u}.$$

با کم کردن عدد ۱ از هر دو طرف به دست می آوریم

$$\begin{aligned} \frac{u(1-v)}{v-u} &\leq \frac{(e^{\lambda v} - 1)(e^{\lambda(1-u)} - 1) - (e^{\lambda(v-u)} - 1)(e^{\lambda} - 1)}{(e^{\lambda(v-u)} - 1)(e^{\lambda} - 1)} \\ &= \frac{(e^{\lambda v} - e^{\lambda(v-u)})(e^{\lambda(1-v)} - 1)}{(e^{\lambda(v-u)} - 1)(e^{\lambda} - 1)} \end{aligned}$$

که معادل با حکم است. $\square$

ضمیمه ب. مقدمات هندسی

۱. قضیه‌ای از هندسهٔ مسطحه

این ضمیمه را با قضیه‌ای از هندسهٔ مسطحه آغاز می‌عکنیم.

ب.۱. قضیهٔ منلائوس^۱: فرض کنید A' ، B' و C' به ترتیب نقاطی روی اضلاع BC ، CA و AB از مثلث ABC باشند. اگر این سه نقطه هم خط باشند، آن گاه

$$\frac{BA'}{A'C} \cdot \frac{CB'}{B'A} \cdot \frac{AC'}{C'B} = ۱.$$

اثبات: $\frac{BA'}{A'C}$ برابر با نسبت فاصله‌های A و B از خط $A'B'C'$ است. عبارت مشابهی برای دو کسر

دیگر برقرار است و با ضرب کردن این سه عبارت حکم ثابت می‌شود. $\square$

در واقع اگر سه کسر فوق را به صورت علامت‌دار در نظر بگیریم، آن گاه هم خط بودن A' ، B' و

C' معادل با این است که حاصل ضرب این سه کسر برابر با -۱ شود.

۲. گوی واحد

ب.۲. قضیه: حجم گوی واحد n -بعدی برابر است با

$$\pi_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + ۱)}$$

^۱ Menelaus

که در آن $\Gamma(a) := \int_0^\infty r^{a-1} e^{-r} dr$ تابع گاما است.

اثبات: ابتدا توجه کنید که حجم گوی واحد $\frac{1}{n}$ برابر مساحت کره واحد است؛ زیرا می‌توانیم گوی را به

شکل اجتماع تعداد زیادی مخروط به ارتفاع ۱ در نظر بگیریم که حجم هر کدام $\frac{1}{n}$ مساحت قاعده آن‌ها

است. بنابراین کافی است مساحت کره واحد را حساب کنیم. برای این منظور انتگرال تابع زیر را به دو

روش حساب می‌کنیم

$$f(x) := e^{-\frac{1}{2}|x|^2}.$$

در روش اول این انتگرال را به حاصلضرب انتگرال‌های گاوسی یک‌بعدی تجزیه می‌کنیم:

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}^n} \prod_{i=1}^n e^{-\frac{1}{2}x_i^2} dx = \prod_{i=1}^n \int_{\mathbb{R}} e^{-\frac{1}{2}x_i^2} dx_i = (\sqrt{2\pi})^n.$$

در روش دوم این انتگرال را در مختصات قطبی بازنویسی می‌کنیم:

$$\begin{aligned} \int_{\mathbb{R}^n} f(x) dx &= \int_{r=0}^\infty \int_{\mathbb{S}^{n-1}} f(r\sigma) \cdot r^{n-1} d\sigma dr \\ &= \int_{r=0}^\infty n\pi_n r^{n-1} e^{-\frac{1}{2}r^2} dr = \frac{n}{2} \pi_n \Gamma\left(\frac{n}{2} + 1\right) \end{aligned}$$

که در آن $\mathbb{S}^{n-1}$ کره واحد n -بعدی و σ جزء سطح آن است. با مقایسه این دو رابطه حکم ثابت

می‌شود. $\square$

ب.۳. نتیجه: برای $n \geq 3$ داریم

$$\frac{\pi_n}{\pi_{n-1}} > \frac{2}{\sqrt{n}}.$$

اثبات: با توجه به این که $\Gamma(x+1) = x\Gamma(x)$ به دست می آوریم $\frac{\pi_n}{\pi_{n-1}} = \frac{n-1}{n} \cdot \frac{\pi_{n-2}}{\pi_{n-3}}$ و

حکم با استقرا ثابت می شود. $\square$

در واقع به وسیله فرمول استرلینگ می توان نشان داد که برای n های بزرگ، $\pi_n \approx \left(\sqrt{\frac{2\pi e}{n}}\right)^n$ و

$$\frac{\pi_n}{\pi_{n-1}} \approx \sqrt{\frac{2\pi}{n}}$$

ب. ۴. لم: اگر $K \subseteq \mathbb{R}^n$ شکل محدبی شامل یک کره به شعاع r باشد، آن گاه نسبت حجم K به

مساحت (حجم $(n-1)$ -بعدی) سطح آن حداقل برابر با $\frac{r}{n}$ است.

اثبات: کافی است فرض کنیم مرز K مشتق پذیر است و مشتق پیوسته دارد. در حالت کلی می توان حکم را با حد گرفتن به دست آورد. همچنین می توانیم فرض کنیم مرکز گوی مذکور همان مبدأ مختصات است. یک جزء کوچک سطح با مساحت dS حول نقطه $x \in \partial K$ در نظر بگیرید. در این صورت حجم جسم مخروط شکلی که از وصل کردن آن به مبدأ مختصات حاصل می شود تقریباً

برابر با $\frac{h(x)}{n} dS$ است که $h(x)$ فاصله مبدأ از ابرصفحه مماس بر K در نقطه x است. اما چون K

شکلی محدب و شامل $\mathbb{B}_r(0)$ است، همواره داریم $h(x) \geq r$ و با انتگرال گرفتن حکم ثابت می شود.

$\square$

لم فوق در واقع عکس مسئله کلاسیک نابرابری محیطی است که برای مساحت یک شکل در $\mathbb{R}^n$ با

حجم ثابت یک کران پایین ارائه می کند و تنها حالت تساوی آن وقتی است که K یک گوی باشد. لم

فوق با فرض این که یک شکل محدب خیلی «تیز» نباشد، کران بالایی برای مساحت ارائه می کند.

می توان دید تساوی در آن به ازای پوش محدب $\mathbb{B}_r(0)$ و یک نقطه حاصل می شود.

۳. قضیه برون-مینکوفسکی

ب.۵. قضیه برون^۱: فرض کنید $K \subseteq \mathbb{R}^n$ شکلی محدب باشد. به ازای یک بردار یکه u تعریف کنید

$$H_t := \{x \in \mathbb{R}^n : \langle x, u \rangle = t\}$$

که در آن $\langle x, u \rangle$ ضرب داخلی x و u است. در این صورت تابع

$$f(t) := (vol_{n-1}(K \cap H_t))^{\frac{1}{n-1}}$$

در جاهایی که ناصفر است تابعی مقعر است.

به ازای $n = 2$ حکم ساده است. پس فرض کنید $n > 2$. اثبات‌ها و تعمیم‌های زیادی از این

قضیه موجوداند. اثباتی که در این جا می‌آوریم بر اساس *متقارن سازی/اشتاینر* است. ابتدا تعریف و برخی

از خواص مقدماتی *متقارن سازی اشتاینر* را می‌آوریم و سپس به اثبات قضیه فوق می‌پردازیم.

فرض کنید که $C \subseteq \mathbb{R}^n$ شکلی فشرده باشد که مرکز ثقل آن بر مبدأ مختصات واقع است. θ را

یک بردار یکه دلخواه بگیرید. ابرصفحه عمود بر θ و گذرنده از مبدأ را با $\theta^\perp$ نشان می‌دهیم.

در صورتی که C محدب باشد، از قضیه برون در حالت دوبعدی نتیجه می‌شود که تابع زیر در جاهایی

که ناصفر است تابعی مقعر است

$$\begin{cases} h : \theta^\perp \rightarrow \mathbb{R} \\ h(x) := vol_1(C \cap (x + \mathbb{R} \cdot \theta)) \end{cases}$$

بنابراین در این صورت شکل زیر یک شکل محدب است

$$S_\theta(C) := \{x + r\theta : x \in \theta^\perp, |r| \leq \frac{h(x)}{2}\}.$$

¹ Brunn

این عمل را متقارن سازی /شتاینر^۱ در راستای θ می نامیم. شکل حاصل فشرده است، نسبت به $\theta^\perp$ متقارن است، و مرکز ثقل آن نیز بر مبدأ مختصات واقع است. همچنین هر خط موازی با θ ، C و $S_\theta(C)$ را در دو پاره خط هم طول قطع می کند. بنابراین حجم C و $S_\theta(C)$ برابر است.

ایده اثبات قضیه برون این است که عمل متقارن سازی را در راستای تعداد زیادی از جهت های عمود بر u روی K انجام دهیم به نحوی که هر $K \cap H_t$ تقریباً به شکل گوی شود.

ب.۶. تعریف: برای یک زیرمجموعه فشرده $C \subseteq \mathbb{R}^n$ تعریف کنید

$$d(C) := \max\{|x| : x \in C\}$$

و

$$A(C) := \{x \in C : |x| = d(C)\}.$$

ب.۷. نکته: واضح است که $d(S_\theta(C)) \leq d(C)$ زیرا برای هر $x \in \theta^\perp$ داریم

$$d(S_\theta(C) \cap (x + \mathbb{R} \cdot \theta)) \leq d(C \cap (x + \mathbb{R} \cdot \theta)).$$

ب.۸. نکته: فرض کنید θ و θ' دو بردار یکه متعامد باشند و C شکلی باشد که $S_\theta(C) = C$.

$$\square \quad S_\theta(S_{\theta'}(C)) = S_{\theta'}(C)$$

ب.۹. نکته: به سادگی دیده می شود که $S_\theta(C) \cap \partial \mathbb{B}_{d(C)}(\circ)$ متشکل از نقاطی از $\partial \mathbb{B}_{d(C)}(\circ)$ مانند

x است که اشتراک خط $x + \mathbb{R} \cdot \theta$ با C و $\mathbb{B}_{d(C)}(\circ)$ یکسان باشد. بنابراین

$d(S_\theta(C)) = d(C)$ اگر و تنها اگر خطی موازی با θ یافت شود که اشتراک آن با C و $\mathbb{B}_{d(C)}(\circ)$ یکسان باشد.

¹ Steiner symmetrization

ب.۱۰. **تعریف:** برای دو زیرمجموعه کران دار $A, B \subseteq \mathbb{R}^n$ ، فاصله هاسدورف^۱ آن ها کمترین عدد حقیقی مثل $d \geq 0$ است که $A \subseteq B + \mathbb{B}_d(0)$ و $B \subseteq A + \mathbb{B}_d(0)$. این عدد را با $d_H(A, B)$ نشان می دهیم.

به راحتی می توان دید که فاصله هاسدورف یک متریک است.

ب.۱۱. **گزاره:** الف) d نسبت به متریک هاسدورف پیوسته و vol نیم پیوسته از بالا است.

ب) S_θ نسبت به متریک هاسدورف نیم پیوسته است به این معنی که اگر $C_i \rightarrow C$ و

$$D \subseteq S_\theta(C), S_\theta(C_i) \rightarrow D$$

اثبات: پیوستگی d از این موضوع نتیجه می شود که $d(C + \mathbb{B}_\epsilon(0)) = d(C) + \epsilon$

نیم پیوستگی vol نیز از این که $\bigcap_{\epsilon > 0} (C + \mathbb{B}_\epsilon(0)) = C$ نتیجه می شود. برای اثبات نیم پیوستگی

S_θ نیز کافی است ثابت کنیم $\bigcap_{\epsilon > 0} S_\theta(C + \mathbb{B}_\epsilon(0)) = S_\theta(C)$. این موضوع نتیجه مستقیمی از

رابطه زیر است

$$\bigcap_{\epsilon > 0} ((C + \mathbb{B}_\epsilon(0)) \cap (x + \mathbb{R} \cdot \theta)) = C \cap (x + \mathbb{R} \cdot \theta)$$

که درستی آن به سادگی قابل مشاهده است. $\square$

ب.۱۲. **گزاره:** فرض کنید $C_1, C_2, \dots$ خانواده ای به طور یک نواخت کران دار از مجموعه های فشرده و

ناتهی در $\mathbb{R}^n$ باشند. در این صورت این دنباله، یک زیردنباله هم گرا دارد. به عبارت دیگر فضای مجموعه های فشرده با متر هاسدورف موضعاً فشرده است.

اثبات: این که به طور یک نواخت کران دار هستند، یعنی همگی آنها در یک گوی ثابت قرار دارند. پس

بدون کم شدن از کلیت، می توانیم فرض کنیم همه آن ها در مکعب واحد قرار دارند. برای عدد طبیعی i ،

¹ Hausdorff distance

مکعب واحد را به مکعب‌های ریزتری با طول ضلع 2^{-i} افراز کنید. برای هر مجموعه فشرده C ، $W_i(C)$ را اجتماع مکعب‌های کوچکی بگیرید که با C اشتراک دارند. با توجه به این که تعداد این مکعب‌ها متناهی است، بی‌نهایت تا از $W_i(C_j)$ ها (که i ثابت است) با هم برابر هستند. حال با یک استدلال قطری زیر دنباله‌ای مثل $j_1, j_2, \dots$ یافت می‌شود که برای هر i داشته باشیم

$$W_i(C_{j_i}) = W_i(C_{j_{i+1}}) = W_i(C_{j_{i+2}}) = \dots$$

با حذف بقیه اعضای دنباله، فرض می‌کنیم $j_i = i$.

با توجه به این که قطر یک مکعب به ضلع 2^{-i} برابر با $2^{-i}\sqrt{n}$ است، به سادگی نتیجه می‌شود که اگر داشته باشیم $W_i(C_1) = W_i(C_r)$ ، آن‌گاه $d_H(C_1, C_r) \leq 2^{-i}\sqrt{n}$. بنابراین $d_H(C_i, C_j) \leq 2^{-\min(i,j)}\sqrt{n}$ و نتیجه می‌شود که این دنباله یک دنباله کوشی است. ادعا می‌کنیم هم‌گرا نیز هست. قرار دهید

$$D_i := \overline{C_i \cup C_{i+1} \cup C_{i+2} \cup \dots},$$

$$D := D_1 \cap D_2 \cap D_3 \cap \dots$$

D_i ها مجموعه‌های فشرده و تودرتو هستند. بنابراین D فشرده و ناتهی است. ادعا می‌کنیم C_i ها به D هم‌گرا هستند. برای هر $j \geq i$ داریم $d_H(C_i, C_j) \leq 2^{-i}\sqrt{n}$. بنابراین $C_j \subseteq C_i + \mathbb{B}_{2^{-i}\sqrt{n}}(\circ)$ پس $D_i \subseteq C_i + \mathbb{B}_{2^{-i}\sqrt{n}}(\circ)$ و به دست می‌آوریم $D \subseteq C_i + \mathbb{B}_{2^{-i}\sqrt{n}}(\circ)$. به عکس، $\epsilon > 0$ را دلخواه بگیرید. قرار دهید

$$D'_i := \overline{D_i \setminus (D + \mathbb{B}_\epsilon(\circ))}. D'_i \text{ ها فشرده و تودرتو اند و } \bigcap_{i=1}^{\infty} D'_i = \emptyset. \text{ پس از جایی به بعد}$$

$D'_i = \emptyset$ ؛ یعنی $D_i \subseteq D + \mathbb{B}_\epsilon(\circ)$ و بنابراین $C_i \subseteq D + \mathbb{B}_\epsilon(\circ)$ ها به D

هم‌گرا هستند و حکم ثابت می‌شود. $\square$

ب.۱۳. لم: برای یک زیرمجموعه فشردۀ $C \subseteq \mathbb{R}^n$ ، X را مجموعه تمام اشکال به شکل $S_{\theta_1} o S_{\theta_2} o \dots o S_{\theta_k}(C)$ بگیرید که k عددی طبیعی و $\theta_1, \dots, \theta_k$ بردارهای یکه‌اند. در این صورت دنباله‌ای از اعضای X به یک گوی هم‌حجم با C میل می‌کند.

اثبات: فرض کنید $e_1, \dots, e_n$ بردارهای یکه استاندارد باشند. Y را مجموعه همه اشکال به شکل $S_{e_n} o \dots o S_{e_1} o S_{\theta_k} o \dots o S_{\theta_1}(C)$ بگیرید. طبق نکته ب.۸، Y این خاصیت اضافه را دارد که برای هر $C' \in Y$ و هر $i \leq n$ داریم $S_{e_i}(C') = C'$ کافی است حکم را برای Y ثابت کنیم.

طبق نکته ب.۷ همگی اعضای Y داخل یک گوی ثابت قرار دارند. بنابراین طبق گزاره ب.۱۱، $\bar{Y}$ نیز چنین است. پس طبق گزاره ب.۱۲ عضو $\bar{Y}$ مانند D وجود دارد که $d(D) = \min\{d(C') : C' \in Y\}$ طبق تعریف Y و گزاره‌های ب.۱۱ و ب.۱۲، برای هر انتخاب بردارهای یکه $\theta_1, \dots, \theta_k$ ، زیرمجموعه‌ای از $S_{\theta_1} o \dots o S_{\theta_k} o S_{e_n} o \dots o S_{e_1}(D)$ در $\bar{X}$ قرار دارد. بنابراین طبق نکته ب.۷ و تعریف D به دست می‌آوریم

$$d(S_{e_n} o \dots o S_{e_1} o S_{\theta_k} o \dots o S_{\theta_1}(D)) = d(D)$$

و نتیجه می‌شود که

$$d(S_{\theta_k} o \dots o S_{\theta_1}(D)) = d(D).$$

نشان می‌دهیم هر مجموعه فشردۀ با این خاصیت یک گوی است. فرض کنید D گوی نباشد. پس $A(D) \neq \mathbb{B}_{d(D)}(\circ)$ را نقطه‌ای از $\partial \mathbb{B}_{d(D)}(\circ)$ بگیرید که $\mathbb{B}_\epsilon(x) \cap D = \emptyset$. سپس نقاط $x_1, \dots, x_k \in \partial \mathbb{B}_{d(D)}(\circ)$ را طوری بگیرید که $\partial \mathbb{B}_{d(D)}(\circ) \subseteq \bigcup_{i=1}^k \mathbb{B}_\epsilon(x_i)$ و قرار دهید

$$\theta_i := \frac{x - x_i}{|x - x_i|}$$

$$(\partial \mathbb{B}_{d(D)}(\circ) \cap \mathbb{B}_\epsilon(x_i)) \cap S_{\theta_k} o \dots o S_{\theta_1}(D) = \emptyset.$$

بنابراین $\partial \mathbb{B}_{d(D)}(\circ) \cap S_{\theta_k} o \dots o S_{\theta_1}(D) = \emptyset$ پس $d(S_{\theta_k} o \dots o S_{\theta_1}(D)) < d(D)$ و این با فرض متناقض است.

حال ثابت می‌کنیم حجم D با حجم C برابر است و لم ثابت می‌شود. r را برابر با شعاع D قرار دهید. می‌دانیم دنباله $C_1, C_2, \dots$ از اشکال هم‌حجم با C موجود است که به D میل می‌کنند. به‌علاوه، برای هر $i \leq n$ و هر j داریم $S_{e_i}(C_j) = C_j$. از گزارهٔ ب.۱۱ نتیجه می‌شود که $vol(D) \geq vol(C)$. از طرف دیگر برای هر $\epsilon > 0$ ، از جایی به بعد داریم $d_H(D, C_j) < \epsilon$. نشان می‌دهیم در این صورت $C_j \supseteq \mathbb{B}_{r-\epsilon\sqrt{n}}(\circ)$ و حکم ثابت می‌شود. فرض کنید $x = (x_1, \dots, x_n) \in \mathbb{B}_{r-\epsilon\sqrt{n}}(\circ) \setminus C$. می‌توانیم فرض کنیم برای هر $i \leq n$ ، $x_i \geq 0$. (حالت‌های دیگر مشابه است). حال عضوی مثل $y \in \partial D$ موجود است که برای هر $i \leq n$ ، $y_i \geq x_i + \epsilon$. چون $D \subseteq C_j + \mathbb{B}_\epsilon(\circ)$ ، عضوی مانند $z \in C_j$ وجود دارد که $|y - z| \leq \epsilon$. حال برای هر $i \leq n$ داریم $z_i \geq x_i$ و از این‌که برای هر $i \leq n$ داریم $S_{e_i}(C_j) = C_j$ ، نتیجه می‌گیریم که $x \in C_j$ و این تناقض است. $\square$

اثبات قضیهٔ برون: بدون کم‌شدن از کلیت قضیه، می‌توانیم فرض کنیم که K فشرده است و مرکز ثقل آن بر مبدأ مختصات واقع است. فرض کنید $t_\lambda = (1-\lambda)t_1 + \lambda t_\lambda$ که $0 < \lambda < 1$ و $f(t_1), f(t_\lambda) \neq 0$. قرار دهید $K_i := K \cap H_{t_i}$. پس طبق تحدب K داریم $K_\lambda \supseteq (1-\lambda)K_1 + \lambda K_\lambda$. مشابه اثبات لم ب.۱۳، می‌توان بردارهای یکه $\theta_1, \dots, \theta_k$ عمود بر u را طوری در نظر گرفت که $K'_i := S_{\theta_k} o \dots o S_{\theta_1}(K_i)$ نزدیک به یک گوی $(n-1)$ -بعدی هم‌حجم با K_i باشد (به ازای $1 \leq i \leq 3$). با توجه به این‌که $K'_i = S_{\theta_k} o \dots o S_{\theta_1}(K) \cap H_{t_i}$ و

این که متقارن سازی اشتاینر تحدب را حفظ می کند، به دست می آوریم $K'_\# \supseteq (1-\lambda)K'_1 + \lambda K'_2$.

حال با میل کردن K'_i به یک گوی هم حجم با K_i حکم نتیجه می شود. $\square$

ب. ۱۴. نتیجه: برای دو شکل محدب $A, B \subseteq \mathbb{R}^n$ ، تابع $f(\lambda) := vol((1-\lambda)A + \lambda B)^{\frac{1}{n}}$ بازه $0 \leq \lambda \leq 1$ تابعی مقعر است.

اثبات: قرار دهید

$$K := \bigcup_{0 \leq \lambda \leq 1} ((1-\lambda)A + \lambda B) \times \{\lambda\} \subseteq \mathbb{R}^{n+1}.$$

به سادگی دیده می شود که K محدب است و حکم از قضیه برون نتیجه می شود. $\square$

ب. ۱۵. نکته: در قضیه برون می توان به جای تحدب K این شرط ضعیف تر را گذاشت که به ازای

$$f(t_\#), f(t_\#) \neq 0 \text{ و } 0 < \lambda < 1 \text{ که } t_\# = (1-\lambda)t_1 + \lambda t_2$$

$$K \cap H_{t_\#} \supseteq (1-\lambda)(K \cap H_{t_1}) + \lambda(K \cap H_{t_2}).$$

در اثبات فقط باید این نکته را جایگزین کنیم که اگر θ یک بردار یکه عمود بر u باشد، آن گاه $S_\theta(K)$ نیز خاصیت فوق را دارد. به سادگی دیده می شود که این موضوع از حالت دوبعدی همین قضیه نتیجه می شود. برای اثبات حالت دوبعدی، باید ثابت کنیم

$$vol((1-\lambda)A + \lambda B) \geq (1-\lambda)vol(A) + \lambda vol(B)$$

که در آن A و B دو زیرمجموعه فشرده از $\mathbb{R}$ هستند. با انتقال دادن هریک از A و B می توانیم فرض کنیم $\max(A) = \min(B) = 0$. در این حالت $(1-\lambda)A$ و λB دو زیرمجموعه مجزای $(1-\lambda)A + \lambda B$ هستند و حکم ثابت می شود. $\square$

ب. ۱۶. نتیجه (قضیه برون-مینکوفسکی)^۱: اگر $A, B \subseteq \mathbb{R}^n$ دو زیرمجموعه فشرده باشند، آنگاه تابع

¹ Brunn-Minkowski

$$f(\lambda) := \text{vol}((1-\lambda)A + \lambda B)^{\frac{1}{n}}$$

روی بازه $0 \leq \lambda \leq 1$ تابعی مقعر است. $\square$

ب. ۱۷. گزاره: فرض کنید A_1 و A_2 دو زیرمجموعه از کره واحد باشند که فاصله هر دو نقطه از آن‌ها حداقل ϵ باشد. در این صورت

$$\text{vol}_{n-1}(A_1) \cdot \text{vol}_{n-1}(A_2) \leq (1 - \epsilon^2)^n n^2 \pi_n.$$

اثبات: قرار دهید $A'_i := \{tx : x \in A_i, 0 \leq t \leq 1\}$ و $B := \frac{1}{2}(A'_1 + A'_2)$. به سادگی

دید می‌شود که فاصله هر عضو B از مبدأ حداکثر $\sqrt{1 - \epsilon^2}$ است. بنابراین

$$\text{vol}(B) \leq (1 - \epsilon^2)^{\frac{n}{2}} \pi_n.$$

از طرف دیگر طبق قضیه برون-مینکوفسکی داریم

$$\text{vol}(B)^{\frac{1}{n}} \geq \frac{1}{2}(\text{vol}(A'_1)^{\frac{1}{n}} + \text{vol}(A'_2)^{\frac{1}{n}}) \geq (\text{vol}(A'_1) \cdot \text{vol}(A'_2))^{\frac{1}{2n}}.$$

داریم $\text{vol}(A'_i) = \frac{1}{n} \text{vol}_{n-1}(A_i)$. بنابراین

$$\text{vol}(B) \geq \frac{1}{n} \sqrt{\text{vol}_{n-1}(A_1) \cdot \text{vol}_{n-1}(A_2)}.$$

با مقایسه این دو نامساوی حکم ثابت می‌شود. $\square$

ب. ۱۸. نتیجه: قرار دهید

$$A := \{x \in \mathbb{S}^{n-1} : |x| > \frac{\epsilon}{\sqrt{n}}\}.$$

در این صورت مساحت A حداکثر برابر است با

$$2n\pi_n e^{-\frac{\epsilon^2}{2}}.$$

اثبات: قرار دهید

$$A_1 := \{x \in \mathbb{S}^{n-1} : x_1 > \frac{\epsilon}{\sqrt{n}}\}$$

و

$$A_2 := \{x \in \mathbb{S}^{n-1} : x_1 < -\frac{\epsilon}{\sqrt{n}}\}.$$

حال فاصله هر دو عضو A_1 و A_2 حداقل برابر با $\frac{\epsilon}{\sqrt{n}}$ است. پس طبق گزاره قبل

$$vol_{n-1}(A_1) \cdot vol_{n-1}(A_2) \leq \left(1 - \frac{\epsilon^2}{n}\right)^n n^2 \pi_n^2 \leq e^{-\epsilon^2} n^2 \pi_n^2.$$

و با توجه به این که $vol_{n-1}(A_i) = \frac{1}{2} vol_{n-1}(A)$ حکم نتیجه می شود. $\square$

مراجع

1. *A random polynomial time algorithm for approximating the volume of convex bodies.* **M. Dyer, A. Frieze, R. Kannan.** 1989, Proc. of the 21'st Annual ACM Symposium on Theory of Computing, pp. 375-381.
2. *Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions.* **R. L. Smith.** 1984, Operations Res., Vol. 32, pp. 1296-1308.
3. *Hit-and-run mixes fast.* **L. Lovasz.** 1998, Math. Prog., pp. 443-461.
4. *Simulated annealing in convex bodies and an $O^*(n^4)$ volume algorithm.* **L. Lovasz, S. Vempala.** Los Alamitos, CA : s.n., 2003, IEEE Comput. Soc. Press, pp. 650- 659.
5. **D. A. Levin, Y. Peres, E. L. Wilmer.** *Markov Chains and Mixing Times.* s.l. : American Mathematical Society, 2008.
6. *Mixing rate of Markov chains, an isoperimetric inequality, and computing the volume.* **L. Lovasz, M. Simonovits.** s.l. : Proc. 31st Annual Symp. on Found. of Computer Science, IEEE Computer Soc., 1990, pp. 346-355.
7. *Computing the volume of convex bodies: a case where randomness provably helps.* **M. Dyer, A. Frieze.** s.l. : Proc. Symp. Appl. Math., 1991, Vol. 44, pp. 123-170.
8. *Random walks in a convex body and an improved volume algorithm.* **L. Lovasz, M. Simonovits.** 1993, Random Structures and Alg., pp. 4, 359-412.
9. *Hit-and-run from a corner.* **L. Lovasz, S. Vempala.** s.l. : SIAM Journal on Computing, 2006, Vol. 35(4), pp. 985-1005.
10. *Hit-and-Run is Fast and Fun.* **L. Lovasz, S. Vempala.** 2003.
11. *Random walks and an $O^*(n^5)$ volume algorithm for convex bodies.* **R. Kannan, L. Lovasz, M. Simonovits.** 11, 1997, Random Structures Algorithms, pp. 1-50.
12. *Extremum problems with inequalities as subsidiary conditions.* **F. John.** New York : Interscience Publishers, 1948, Studies and Essays Presented to R. Courant on his 60th Birthday, pp. 187- 204.
13. *Random points in isotropic convex sets.* **J. Bourgain.** Berkely, CA : Math. Sci. Inst. Publ., 1996, Convex geometric analysis, pp. 53-58.
14. *Random vectors in the isotropic position.* **M. Rudelson.** 164, 1999, J. Funct. Anal, pp. 60-72.
15. *Sampling and integration of near log-concave functions.* **D. Applegate, R. Kannan.** 1990.

16. *The geometry of logconcave functions and an $O^*(n^3)$ sampling algorithm.* **L. Lovasz, S. Vempala.** s.l. : Microsoft Research Tech. Rep.
17. *Where to Start a Geometric Random Walk?* **L. Lovasz, S. Vempala.** s.l. : Technical Report MSR-TR, 2003.
18. *Isoperimetric problems for convex bodies and a localization lemma.* **R. Kannan, L. Lovasz, M. Simonovits.** 13, 1995, J. Discr. Comput. Geom, pp. 541-559.
19. *The Brunn-Minkowski inequality in Gauss spaces.* **C. Borell.** 30, 1975, Invent. Math., pp. 207-216.
20. **A. Prekopa.** *Stochastic Programming.* Akademiai Kiado, Budapest and Kluwer : Dordrecht, 1995.
21. *How to compute the volume in high dimension.* **M. Simonovits.** 97, 2003, Math. Program. Ser. B, pp. 337- 374.

فهرست کلمات

- اتم, ۲۸
استقلال تقریبی, ۷۳, ۷۶
الگوریتم حجم, ۸۰
آمیختگی, ۳۴
آمیختگی تقریبی, ۷۳
اوراکل, ۲, ۱۱
اوراکل عضویت قوی, ۱۲
برگشت پذیری, ۲۴
توزیع پایا, ۲۳
توزیع گرم, ۴۴
توزیع لوگ-مقعر, ۹۸
جفت سازی, ۳۲
چگالی احتمال گذر, ۲۳
خطای احتمالاتی, ۱۹
خوش تضمینی, ۱۱
رابطه تعادل, ۲۴
رسانش, ۳۷
رسانش موضعی, ۶۲, ۶۳
رشته, ۸۲
روش Simulated Annealing, ۹۹
روش تبدیل, ۵
روش ترکیبی, ۵
روش رد و قبول, ۶
روش متروپولیس-هستینگز, ۹۷
روش مونت کارلو, ۳, ۲۵
روش مونت کارلو چندفازی, ۳
زمان آمیختگی, ۳۴
ساندویچ کردن, ۹۳
شار ارگودیک, ۳۶
شروع گرم, ۴۴
شیء تصادفی, ۲۱
فاز, ۸۲, ۹۹
فاصله مجموع تغییرات, ۲۶, ۳۲
فاصله هاسدورف, ۱۲۱
فرآیند تصادفی, ۲۱
فرآیند تنبل, ۳۵
فرآیند مارکوف, ۲۲
قدم زدن تصادفی Hit-and-run, ۹, ۱۳, ۴۶, ۴۷, ۷۴, ۹۰
قدم زدن تصادفی با گام های کروی, ۸, ۴۴, ۶۲
قدم زدن تصادفی روی نقاط شبکه ای, ۷
قضیه برون, ۱۱۹
قضیه برون-مینکوفسکی, ۱۲۵
قضیه منلائوس, ۱۱۶
قضیه هریس, ۳۴
لم موضعی سازی, ۱۰۲, ۱۱۱
مقارن سازی اشتاینر, ۱۱۹
منبع تصادف, ۱۴, ۱۹
موقعیت آیزوتروپیک, ۹۴
تقریباً آیزوتروپیک, ۹۵
نابرابری محیطی, ۵۶, ۵۷, ۱۱۲
نسبت ساندویچی, ۹۳
نسبت ناهمساز, ۴۸
هسته مارکوف, ۲۲

واژه‌نامه (فارسی-انگلیسی)

atom	اتم
oracle	اوراکل
mixing	آمیختگی
time-reversibility	برگشت پذیری
stationary distribution	توزیع پایا
warm distribution	توزیع گرم
coupling	جفت سازی
probability transition function	چگالی احتمال گذر
multiphase	چند فازی
well-guaranteed	خوش تضمین
detailed balance equation	رابطه تعادل
conductance	رسانش
local conductance	رسانش موضعی
thread	رشته
transformational method	روش تبدیل
composition method	روش ترکیبی
acceptance-rejection method	روش رد و قبول
mixing time	زمان آمیختگی
sandwiching	ساندویچ کردن
ergodic flow	شار ارگودیک
warm start	شروع گرم
random object	شیء تصادفی
phase	فاز
total variation distance	فاصله مجموع تغییرات
Hausdorff distance	فاصله هاسدورف
stochastic process	فرآیند تصادفی
lazy process	فرآیند تنبل
Markov process	فرآیند مارکوف

random walk	قدم زدن تصادفی
ball walk	قدم زدن تصادفی با گام های کروی
localization lemma	لم موضعی سازی
log-concave	لوگ-مقعر
symmetrization	متقارن سازی
source of randomness	منبع تصادف
isoperimetric inequality	ناابرابری محیطی
sandwiching ration	نسبت ساندویچی
cross ratio	نسبت ناهمساز
Markov kernel	هسته مارکوف

واژه‌نامه (انگلیسی-فارسی)

acceptance-rejection method	روش رد و قبول
Atom	اتم
ball walk	قدم زدن تصادفی با گام های کروی
composition method	روش ترکیبی
Conductance	رسانش
Coupling	جفت سازی
cross ratio	نسبت ناهمساز
detailed balance equation	رابطه تعادل
ergodic flow	شار ارگودیک
Hausdorff distance	فاصله هاسدورف
isoperimetric inequality	نابرابری محیطی
lazy process	فرآیند تنبل
local conductance	رسانش موضعی
localization lemma	لم موضعی سازی
log-concave	لوگ-مقعر
Markov kernel	هسته مارکوف
Markov process	فرآیند مارکوف
mixing	آمیختگی
mixing time	زمان آمیختگی
multiphase	چند فازی
oracle	اوراکل
phase	فاز
probability transition function	چگالی احتمال گذر
random object	شیء تصادفی
random walk	قدم زدن تصادفی
sandwiching	ساندویچ کردن
sandwiching ration	نسبت ساندویچی
source of randomness	منبع تصادف

stationary distribution	توزیع پایا
stochastic process	فرآیند تصادفی
symmetrization	متقارن سازی
thread	رشته
time-reversibility	برگشت پذیری
total variation distance	فاصلهٔ مجموع تغییرات
transformational method	روش تبدیل
warm distribution	توزیع گرم
warm start	شروع گرم
well-guarantied	خوش تضمین

Abstract

“How can we generate a random point with uniform distribution over a convex body $K \subseteq \mathbb{R}^n$?”

According to its applications, it's important for a solution to this problem to be applicable in high dimensions. Here, we are interested in algorithms with polynomial order with respect to the dimension. All existing methods for dealing with this problem are based on the Markov chain Monte Carlo method, i.e. a random walk is constructed in K such that its stationary distribution is the uniform distribution over K . Then, after simulating “enough” steps of this random walk, the distribution of the resulting point is “approximately” uniform. The real problem in Monte Carlo method is analyzing its “mixing time”, i.e. the number of steps needed to get close enough to the stationary distribution.

This thesis focuses mostly on analyzing the mixing time of the Hit-and-run walk which is the fastest existing random walk for the mentioned problem. Moreover, we explain one of the most important applications of this problem: approximating the volume of a convex body. A fascinating point about this application is that no deterministic polynomial-time algorithm can approximate the volume of a convex body with bounded relative error.

Keywords: convex body, approximating the volume, Markov chain, stochastic algorithm, mixing time, conductance, total variation distance



Sharif University of Technology

Department of Mathematical Science

M.S. Thesis

in Pure Mathematics

Title

Generating Random Points in a Convex Body in High Dimensions

Author

Ali Khezeli

Supervisor

Prof. Kasra Alishahi

August 2009