دانشگاه صنعتی شریف

دانشکده علوم ریاضی

پایان‌نامه کارشناسی ارشد

ریاضی گرایش محض

عنوان

گذار فاز در مساله‌های بهینه‌سازی محدب با داده‌های تصادفی

نگارنده

دلبر فقیه

استاد راهنما

دکتر کسری علیشاهی

بهمن ۹۹

بسم الله الرحمن الرحیم

دانشگاه صنعتی شریف

دانشکده علوم ریاضی

پایان‌نامه

برای دریافت درجه کارشناسی‌ارشد در رشته

ریاضی گرایش محض

عنوان: گذار فاز در مساله‌های بهینه‌سازی محدب با داده‌های تصادفی

نگارش: دلبر فقیه

اعضای هیأت داوران:

دکتر کسری علیشاهی (استاد راهنما) امضاء: . . . .

دکتر امضاء: . . . .

دکتر امضاء: . . . .

تاریخ: ۱۳۹۶/۰۶/۱۵

مقدمه

دسته ای روش های حل مسائل بهینه سازی محدب path-following methods نام دارند که ایده ی اصلی آنها در نظر گرفتن پارامتری از مسئله می باشد که سختی سوال حاصل از آن است و با شروع از یک جواب بدیهی به جواب اصلی نزدیک می شوند.مشکل آنجایی رخ می دهد که در جواب سوال عدم پیوستگی ای نسبت به آن پارامتر دیده شود به این صورت که درسایر نقاط در طول مسیر جواب سوال بهینه سازی نسبت به آن پارامتر مورد نظر از جواب ها، بی نهایت بار مشتق پذیر است و بنابراین به روزرسانی جواب کار راحتی می باشد اما در آن نقاط یکی از مشتق ها مثلا مشتق اول ناپیوسته است و این کار را دشوار می کند.این تغییرات ناگهانی را گذار فاز گویند.در تحلیل داده چنین گذارهایی زمانی رخ می دهند که پیچیدگی مدل یا تعداد داده های دور افتاده در انتخاب مدل خطی برازش داده **robust** یا compresd sensing reconstructions فراتر از آستانه برود.در هندسه ی ترکیبیاتی این گذارها تغییرات ناگهانی در ویژگی های شمارش صورت پلی توپ های محدب با ابعاد متفاوت می باشند. [۲]

نظریه ی حجم های ذاتی مخروط تاریخچه ای غنی دارد و امروزه با پیدا کردن کاربردهای گوناگون در مسائل بهینه سازی و بخصوص مسائل شرح داده شده در پاراگراف فوق مورد توجه قرار گرفته است. سامرویل با بررسی ارتباط بین زاویه و حجم پلی توپ های محدب در فضای n -بعدی پایه های این نظریه را ایجاد کرد. [۳]

چکیده

در رفتار بسیاری از مساله‌های بهینه‌سازی محدب با قیدهای تصادفی در ابعاد بالا، بر حسب تعداد قیدها تغییرات ناگهانی یا گذار فاز مشاهده شده است. از نمونه‌های شناخته‌شده‌ی این موضوع مساله‌ی بازسازی یک بردار تُنُک یا یک ماتریس با رتبه‌ی کم بر مبنای تعدادی مشاهده‌ی خطی تصادفی است. در هر دو مثال روش‌هایی مبتنی بر بهینه‌سازی محدب ابداع و مشاهده و اثبات شده است که وقتی تعداد مشاهدات از یک آستانه‌ی مشخص بیش‌تر (کم‌تر) شود جواب مساله با احتمال نزدیک به یک (صفر) به درستی بردار یا ماتریس اولیه را بازسازی می‌کند. به تازگی نتایجی به دست آمده است که توضیح می‌دهد چرا این پدیده‌ی گذار فاز در مساله‌های بهینه‌سازی محدب با داده‌های تصادفی فراگیر است. این نتایج به مفاهیم هندسه تصادفی در ابعاد بالا و قضایایی درباره‌ی تجمع حجم‌های ذاتی متکی است که به تعریفی از بُعد آماری برای کُنج‌های محدب و مسائل بهینه‌سازی محدب منجر شده و ارتباطاتی بین موضوعات از قبل شناخته شده به وجود آورده است. هدف این پایان‌نامه معرفی و بررسی این نتایج جدید است.

واژه‌های کلیدی: بهینه‌سازی محدب، گذار فاز، هندسه تصادفی، بردارهای تُنُک، ماتریس‌های کم‌رتبه

فهرست مطالب

فصل ۱
معرفی مساله

فصل ۲
پیش نیاز

گوییم مجموعه C محدب است هرگاه به ازای هر دو نقطه درون C شامل پاره خط واصل آن دو نقطه باشد. مخروط محدب $C \subset \mathbb{R}^n$ مجموعه ای است محدب که برای هر $\lambda > 0$ داشته باشیم $\lambda C = C$.
وجه سره [۱] C مجموعه ای است به فرم $F = C \cap H$ که H یک ابرصفحه ی پشتیبان [۲] است. F وجه C نامیده می شود اگر و فقط اگر F وجهی سره از C باشد یا خود آن.مجموعه وجوه C را با $\mathcal{F}(C)$ نشان می دهیم. گستره خطی [۳] مخروط C (line(C)) کوچکترین زیرفضا خطی شامل C می باشد و $line(C) = C \bigoplus (-C)$ که $\bigoplus$ جمع مینکوسکی دو مجموعه است.بعد وجه F برابر است با بعد گستره ی آن و درون نسبی [۴] relent(F) درون F است در Line(F) . مخروط را pointed گوییم هرگاه شامل هیچ زیرفضا خطی با بعد بیشتر از ۰ نباشد.اگر C مخروطی باشد که pointed نباشد آنگاه شامل زیرفضایی غیرخطی و نابدیهی از بعد حداکثر $k > 0$، $L = C \cap (-C)$ است که مشمول در هر ابرصفحه ی پشتیبان C (و بنابراین در هر وجه آن) خواهد بود.تصویر اقلیدسی مخروط محدب از نقطه x نزدیک ترین نقطه مخروط به آن تعریف می شود(یعنی $\Pi_C(x) = argmin\{\| x - y \| : y \in C\}$).قطبی یک مخروط عبارت است از مجموعه ی نرمال های بیرونی آن(یعنی $C^o = \{u \in \mathbb{R}^n \mid < u, x > \leq 0 \forall x \in C\}$) قطبی یک مخروط همواره بسته و محدب است.یک مخروط را خوددوگان [۵] گویند هرگاه $C^o = C$. یک تابع محدب را سره [۶] گوییم هر گاه حداقل یک مقدار متناهی را اخذ کند و هرگز $-\infty$ را نگیرد.

۰-۲.۱ تصاویر اقلیدسی

تعریف ۲.۱. فرض کنید $E \subset \mathbb{R}^n$ یک مجموعه محدب بسته باشد.تصویر اقلیدسی نگاشت زیر می باشد

$$\pi_E : \mathbb{R}^n \to E, \pi_E := argmin\{\| x - y \| : y \in E\}$$

[1] Proper face
[2] Supporting hyperplane
[3] Line span
[4] relative interior
[5] Self-dual
[6] proper convex function

از آنجایی که نرم اکیدا محدب است تصویر اقلیدسی خوش تعریف می باشد. خواص:

$$\| \pi_E(x) - \pi_E(y) \| \leq \| x - y \|, \ \| (I - \pi_E)(x) - (I - \pi_E)(y) \| \leq \| x - y \| \ for all x, y \in \mathbb{R}^n$$

$$|dist(x, E) - dist(y, E)| \leq \| x - y \| \ \ for \ all x, y \in \mathbb{R}^n$$

مربع فاصله همه جا مشتق پذیر است و داریم

$$\nabla dist^۲(x, E) = ۲(x - \pi_E(x)) \ for \ all x \in \mathbb{R}^n$$

فرض کنید $C \subset \mathbb{R}^n$ یک مخروط محدب بسته باشد.آنگاه

$$\pi_C(\tau x) = \tau \pi_C(x) \ \tau \geq ۰$$

مخروط C یک تجزیه ی متعامد به $\mathbb{R}^n$ القا می کند.

$$x = \pi_C(x) + \pi_(C^o)(x), \ < \pi_C(x), \pi_(C^o)(x) >= ۰ \ for all x \in \mathbb{R}^n$$

این تجزیه معادله فیثاغورسی زیر را نتیجه می دهد

$$\| x \|^۲ = \| \pi_C(x) \|^۲ + \| \pi_(C^o)(x) \|^۲ \ \ for \ all x \in \mathbb{R}^n$$

و در پی آن نتیجه ی زیر را خواهیم داشت

$$dist(x, C) = \| x - \pi_C(x) \| = \| \pi_(C^o)(x) \| \ for all \ x \in \mathbb{R}^n$$

بنابراین

$$\nabla \| \pi_C(x) \|^۲ = \nabla dist^۲(x, C^o) = ۲(x - \pi_(C^o)(x)) = ۲(\pi_C(x) + \pi_(C^o)(x) - \pi_(C^o)(x)) = ۲\pi_C(x) \ for all \ x \in \mathbb{R}^n$$

$$\Rightarrow \nabla \| \pi_C(x) \|^۲ = ۲\pi_C(x) for all \ x \in \mathbb{R}^n$$

برای دو مخروط بسته محدب $C_۱, C_۲ \subset \mathbb{R}^{n_۱}, \mathbb{R}^{n_۲}$ داریم

$$\pi_{C_۱ \times C_۲}(x_۱, x_۲) = (\pi_{C_۱}(x_۱), \pi_{C_۲}(x_۲)) for all \ x_۱ \in \mathbb{R}^{n_۱}, x_۲ \in \mathbb{R}^{n_۲}$$

فصل ۳

مقدمات

۱-۳ حجم ذاتی

[1]

تعریف ۱.۳. برای مخروط چند وجهی $C \subseteq \mathbb{R}^n$ و دو وجه $F, G \in \mathcal{F}(C)$ تعریف می کنیم

$$v_F(G) = \mathbb{P}\{\Pi_G(g) \in relint F\}$$

که $g \sim \mathcal{N}(\mathbb{R}^n)$ بردار گوسی استاندارد در $\mathbb{R}^n$ است. k -امین حجم ذاتی C برای $0 \leq k \leq n$ بصورت زیر تعریف می شود

$$v_k(C) = \sum_{F \in \mathcal{F}_k(C)} v_F(C)$$

مثال ۲.۳ فرض کنید L_j زیرفضا j -بعدی $\mathbb{R}^n$ باشد.آنگاه L_j مخروطی چند وجهی است که دقیقا یک وجه دارد.پس Π_{L_j} هر نقطه را به این وجه j -بعدی تصویر می کند.بنابراین برای $0 \leq k \leq n$

$$v_k(L_j) = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}$$

▲

مثال ۳.۳ کنج نامنفی $\mathbb{R}^n_+$ یک مخروط چند وجهی است. $\Pi_{\mathbb{R}^n_+}(g)$ در درون نسبی یک وجه k -بعدی قرار دارد اگر و تنها اگر k -امین مختص g مثبت باشد.مختص ها مستقل اند و با احتمال یک دوم مثبت و یا منفی اند در نتیجه

$$v_k(\mathbb{R}^n_+) = \binom{n}{k} 1/2^n$$

[1] intrinsic volume

▲

خواص حجم ذاتی

۱) ناوردایی متعامد.برای تبدیل متعامد $Q \in O(n)$

$$v_k(QC) = v_k(C)$$

۲) قطبیت

$$v_k(C) = v_(n-k)(C^o)$$

۳) قاعده ضرب

$$v_k(C \times D) = \sum_{i+j=k} v_i(C)v_j(D)$$

تعریف ۴.۳. فرض کنید C یک مخروط محدب بسته در $\mathbb{R}^n$ باشد و $\{C_i : i = ۱,۲,۳,\cdots\} \subseteq \mathbb{R}^n$ مخروط های چند وجهی که در متر هاسدورف مخروطی [۲] به C میل می کنند.برای هر *k=0,1,...,n*

$$v_k(C) := \lim_{i\to\infty} v_k(C_i)$$

خواص:

۱) توزیع.حجم های ذاتی یک توزیع احتمال روی {0,1,...,n} تعریف می کنند

$$\sum_{k=۰}^{n} v_k(C) = ۱,\ v_k(C) \geq ۰$$

۲) فرمول Gauss–Bonnet. زمانی که C زیرفضا نیست

$$\sum_{k=۰,keven}^{n} v_k(C) = \sum_{k=۰,kodd}^{n} v_k(C) = ۱/۲$$

حجم ذاتی یک مخروط محدب بسته حول بعد آماری مخروط در مقیاس تعیین شده توسط بعد آماری متمرکز است.

قضیه ۵.۳. فرض کنید C یک مخروط محدب بسته باشد.پهنا گذار [۳] بصورت زیر تعریف می شود

$$\omega(C) := \sqrt{min\{\delta(C), \delta(C^O)\}}$$

تابع زیر را در نظر بگیرید

$$p_C(\lambda) := ۴exp(\frac{-\lambda^۲ ۸}{\omega^۲(C)+\lambda}) for \lambda \geq ۰$$

[2]conic Hausdorff metric
[3]transition width

آنگاه

$$k_- \leq \delta(C) - \lambda + 1 \Rightarrow t_{k_-}(C) \geq 1 - p_C(\lambda)$$

$$k_+ \geq \delta(C) + \lambda \Rightarrow t_{k_+}(C) \leq p_C(\lambda)$$

نامساوی ضعیف تر زیر وجود دارد

$$p_C(\lambda) \leq \begin{cases} 4e^{-\lambda^2/16\omega^2(C)} & 0 \leq \lambda \leq \omega^2(C) \\ 4e^{-\lambda/16} & \lambda \geq \omega^2(C) \end{cases}$$

برهان. ■

۲-۳ فرمول مخروطی اشتاینر

اندازه ی d -بعدی ϵ -همسایگی یک ساختار محدب (محدب و فشرده)

$K \subseteq \mathbb{R}^n$ را می توان بصورت چند جمله ای از ϵ و با درجه حداکثر d و ضرایب حجم های ذاتی بیان کرد

$$vol(K + \epsilon B^n) = \sum_{i=0}^{n} v_i(K)\omega_{n-i}\epsilon^{n-i}$$

که B^n گوی واحد است و $\omega_{n-i} = vol(B^{n-i}) = \frac{2\pi^{(n-i)/2}}{\Gamma((n-i)/2+1))}$ و $v_i(K)$ حجم های ذاتی اقلیدسی می باشند. [3]

مثال ۶.۳

$$vol(K + \epsilon B^2) = \text{مساحت} + \epsilon \cdot \text{محیط} + \pi \cdot \epsilon^2$$

▲

برای بیان نتیجه مشابه در مخروطهای محدب یا مجموعه های محدب کروی (اشتراک مخروط های محدب با کره واحد) ، باید در مورد مفهوم فاصله توافق کنیم.یک انتخاب طبیعی $arccos(\| \Pi_C(x) \| / \| x \|)$ است که همواره کمتر مساوی پی دوم می باشد و تساوی در صورتی رخ می دهد که x عصو قطبی C باشد.برای x با نرم یک و $\alpha \leq \pi/2$ داریم $arccos(\| \Pi_C(x) \| / \| x \|) \leq \pi/2$ اگر و فقط اگر $\| \Pi_C(x) \|^2 \geq cos^2(\alpha)$.

تعریف ۷.۳. فرض کنید L_k زیر فضا k -بعدی $\mathbb{R}^n$ باشد.تعریف می کنیم

$$I_k^n(\epsilon) = \mathbb{P}\{\| \Pi_{L_k}(\theta) \|^2 \geq \epsilon\}, \epsilon \in [0, 1]$$

که θ توزیع یکنواخت روی کره واحد در $\mathbb{R}^n$ است.

استدلال هندسی پایه نشان می دهد که I_k^n نسبت نقاطی در کره S^{n-1} است که درون زاویه $arccos(\sqrt{\epsilon})$ از زیر فضای L_k قرار دارد. [۴]

۲-۳.۱ فرمول کروی اشتاینر

فرض کنید C یک مخروط محدب بسته در $\mathbb{R}^n$ باشد.برای هر $\epsilon \in [0, 1]$

$$\mathbb{P}\{\| \Pi_{L_k}(\theta) \|^2 \geq \epsilon\} = \sum_{i=0}^{n} v_i(K) I_k^n(\epsilon)$$

۳-۳ بعد آماری

تعریف ۳.۸. فرض کنید C یک مخروط محدب در $\mathbb{R}^n$ باشد.بعد آماری [4] $\delta(C)$ بصورت زیر تعریف می شود

$$\delta(\bar{C}) = \sum_{k=0}^{n} k v_k(\bar{C})$$

که $\bar{C}$ بستار C است.

خواص

۱) برای مخروط های محدب تو در تو $C \subseteq K \subseteq \mathbb{R}^n$ داریم

$$0 \leq \delta(C) \leq \delta(K) \leq n$$

۲)بعد آماری زیر فضا خطی با بعد آن برابر است. ۳)برای هر ماتریس متعامد $U \in \mathbb{R}^{n\times n}$ داریم $\delta(UC) = \delta(C)$. ۴) مجموع بعد آماری یک مخروط محدب بسته و بعد آماری قطبی آن برابر است با بعد فضا. ۵) عد آماری حاصل ضرب دو مخروط محدب بسته برابر است با مجموع ابعاد آماری آن دو. (بعد آماری تحت نشاندن ناوردا است.)

قضیه ۳.۹. فرض کنید C یک مخروط محدب بسته در $\mathbb{R}^n$ باشد.داریم

۱.(نمایش گاوسی) $\delta(C) = \mathbb{E}[\| \Pi_C(g) \|^2], g \sim \mathcal{N}(0, 1_n)$

۲.(نمایش کروی) $\delta(C) = n\mathbb{E}[\| \Pi_C(\theta) \|^2], \theta \sim U(S^{n-1})$

۳.(نمایش قطبی) $\delta(C) = \mathbb{E}[dist^2(g, C^o)]$

۴.(نمایش Mean-squared-width) $\delta(C) = \mathbb{E}[(\sup_{y\in C\cap B^n} < y, g >)^2], g \sim \mathcal{N}(0, 1_n)$

برهان. ۱)نمایش گاوسی و کروی مطابق یکدیگر اند پس

$$\mathbb{E}[\| \Pi_C(g) \|^2] = n\mathbb{E}[\| \Pi_C(\theta) \|^2] = n \int_0^1 \mathbb{P}[\| \Pi_C(\theta) \|^2 \geq \epsilon] \, d\epsilon$$

[4] Statistical dimension

طبق فرمول کروی اشتاینر و تعریف تابع tropic داریم

$$\mathbb{E}[\| \Pi_C(g) \|^2] = n\sum_{k=0}^{n} v_k(C)\int_0^1 \mathbb{P}[\| \Pi_(L_k)(\theta) \|^2 \geq \epsilon]\, d\epsilon = \sum_{k=0}^{n} v_k(C)(n\mathbb{E}[\| \Pi_(L_k)(\theta) \|^2 \geq \epsilon])$$

$$= \sum_{k=0}^{n} kv_k(C)$$

که L_k یک زیرفضا k-بعدی دلخواه می باشد. ■

قضیه ۱۰.۳. فرض کنید C یک مخروط خوددوگان در $\mathbb{R}^n$ باشد آنگاه $\delta(C) = n/2$.

برهان. طبق خواص بعد آماری داریم

$$\delta(C) + \delta(C^o) = \delta(C) + \delta(C) = (\mathbb{R}^n) = dim(\mathbb{R}^n) = n \rightarrow \delta(C) = n/2$$

■

تعریف ۱۱.۳. مخروط مدور $circ_n(\alpha)$ در $\mathbb{R}^n$ با زاویه $0 \leq \alpha \leq /2$ بصورت زیر تعریف می شود

$$circ_n(\alpha) := \{x \in \mathbb{R}^n \ :| x_1 |\geq\| x \| cos(\alpha)\}$$

قضیه ۱۲.۳. بعد آماری مخروط های مدور در $\delta(circ_n(\alpha)) = nsin^2(\alpha) + O(1)$ صدق می کند.
ترم ارور تقریبا با $cos(2\alpha)$ برابر است.

برهان. زاویه ی $\beta = \beta(\theta) = arccos(\theta_1)$ را بین θ و بردار $(1, 0, \cdots, 0)$ تعریف می کنیم پس

$$F(\beta) :=\| \Pi_{circ_n(\alpha)}(\theta) \|^2= \begin{cases} 1 & 0 \leq \beta < \alpha \\ cos^2(\beta - \alpha) & \alpha \leq \beta < \pi/2 + \alpha \\ 0 & \pi/2 + \alpha \leq \beta \leq \pi \end{cases}$$

حال به کمک نمایش کروی بعد آماری را محاسبه می کنیم.

$$\delta(circ_n(\alpha)) = n\mathbb{E}[\| \Pi_{circ_n(\alpha)}(\theta) \|^2] = n\frac{\Gamma(n/2)}{\sqrt{\pi}\Gamma((n-1)/2))}\int_0^\pi sin^{n-2}(\beta)F(\beta)\, d\beta$$

می توان به کمک لاپلاس انتگرال را تخمین زد [5]

$$\int_0 sin^{n-2}(\beta)F(\beta)\, d\beta = \sqrt{\frac{2\pi}{n}}F(\frac{\pi}{2}) + O(n^{-3/2})$$

به کمک Gautschi's Inequality می توان نوشت

$$\sqrt{n-2} < \frac{\sqrt{2}\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2}} < \sqrt{n}$$

در نتیجه

$$\frac{n\sqrt{n-2}}{\sqrt{2\pi}}(\sqrt{\frac{2\pi}{n}}F(\frac{\pi}{2})+O(n^{-3/2})) < n\frac{\Gamma(\frac{n}{2})}{\sqrt{\pi}\Gamma(\frac{n-1}{2})}(\sqrt{\frac{2\pi}{n}}F(\frac{\pi}{2})+O(n^{-3/2})) < \frac{n\sqrt{n}}{\sqrt{2\pi}}(\sqrt{\frac{2\pi}{n}}F(\frac{\pi}{2})+O(n^{-3/2})$$

$$ncos^2(\frac{\pi}{2}-\alpha)+\frac{n\sqrt{n-2}}{\sqrt{2\pi}}O(n^{-3/2}) < \delta(circ_n(\alpha)) < ncos^2(\frac{\pi}{2}-\alpha)+\frac{n\sqrt{n}}{\sqrt{2\pi}}(O(n^{-3/2}))$$

$$nsin^2(\alpha)+\frac{n\sqrt{n-2}}{\sqrt{2\pi}}O(n^{-3/2}) < \delta(circ_n(\alpha)) < nsin^2(\alpha)+\frac{n\sqrt{n}}{\sqrt{2\pi}}O(n^{-3/2})$$

پس $\delta(circ_n(\alpha)) = nsin^2(\alpha)+O(1)$. برای تخمین بهتر ترم ارور با $cos(2\alpha)$ از این واقعیت استفاده می شود که حجم ذاتی مخروط مدور در تساوی زیر صدق می کند.

$$v_k(circ_n(\alpha)) = \frac{1}{2}\binom{\frac{n-2}{2}}{\frac{k-1}{2}}sin^{k-1}(\alpha)cos^{n-k-1}(\alpha)fork = 1, 2, \cdots, n$$

برهان. اضافه شود. ■

■

تعریف ۱۳.۳. *(signed) permutahedron* تولید شده توسط بردار $x \in \mathbb{R}^n$ پوش محدب تمام جایگشت های مختصات آن بردار می باشد

$$\mathcal{P}(x) := conv\{\sigma(x) : \sigma a coordinate permutation\}$$

مخروط نرمال $\mathcal{N}(E,x)$ از مجموعه محدب $E \subset \mathbb{R}^n$ و $x \in E$ بصورت زیر تعریف می شود

$$\mathcal{N}(E,x) = \{s \in \mathbb{R}^n :< s, y-x >\leq 0 \ \forall y \in E\}$$

قضیه ۱۴.۳. فرض کنید بردار $x \in \mathbb{R}^n$ با درایه هایی متمایز باشد.بعد آماری مخروط نرمال در راس *(signed) permutahedron* تولید شده توسط X در معادله ی زیر صدق می کند

$$\delta(\mathcal{N}(\mathcal{P}(x),x)) = \sum_{i=1}^{n} i^{-1},\ \delta(\mathcal{N}(\mathcal{P}_{\pm}(x),x)) = \frac{1}{2}\sum_{i=1}^{n} i^{-1}$$

برهان. برای اثبات از لم زیر استفاده می کنیم. ■

تعریف ۱۵.۳. یگروه متقارن متناهی یک زیرگروه متناهی از گروه متعامد(گروه ماتریس هایی که از تبدیلات متعامد یک فضای اقلیدسی پدید می آید) است که با تقارن حول ابر صفحه ها تولید می شود.

تعریف ۱۶.۳. هر گروه متقارن متناهی $\mathbb{R}^n$ را به مجموعه ای از مخروط های محدب افراز می کند که اتاق [5] نامیده می شوند.

[5] chambers

اتاق خانواده نامتناهی $A_{n-۱}$ و BC_n از گروه های متقارن متناهی تحویل ناپذیر با مخروط های زیر ایزومتریک می باشند

$$C_A = \{x \in \mathbb{R}^n : x_۱ \le x_۲ \le \cdots \le x_n\},\ C_{BC} = \{x \in \mathbb{R}^n : ۰ \le x_۱ \le x_۲ \le \cdots \le x_n\}$$

لم ۱۷.۳. $\mathcal{N}(\mathcal{P}(x), x)$ با C_A و $\mathcal{N}(\mathcal{P}_\pm(x), x)$ با C_{BC} ایزومتری می باشد.

برهان. برهان.(و اثبات رو کامل کن) ∎

۳-۳.۱ بعد آماری مخروط های کاهشی

تعریف ۱۸.۳. مخروط کاهشی $\mathcal{D}(f,x)$ از تابع محدب سره $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ در نقطه $x \in \mathbb{R}^n$ عبارت است از پوش محدب اختلال هایی است که f نزدیک x افزایش نمی یابد.

$$\mathcal{D}(f,x) = \bigcup_{\tau > ۰} \{y \in \mathbb{R}^n : f(x+\tau y) \le f(x)\}$$

این مخروط ها همواره محدب اند و لزوما بسته نمی باشند.مخروط کاهشی یک تابع محدب هموار؟ زیرفضا می باشد پس مطالعه این مخروط ها زمانی حائز اهمیت است که تابع هموار نباشد.

قضیه ۱۹.۳. (بعد آماری مخروط کاهشی نرم بینهایت).فرض کنید برای بردار $x \in \mathbb{R}^n$، $۱ \le \#\{i : |x_i| = \| x \|_\infty\} = s \le n$ آنگاه

$$\delta(D(\| \,.\, \|_\infty, x)) = n - s/۲$$

برهان. کافی است نشان دهیم $\delta(D(\| \,.\, \|_\infty, x)) = \mathbb{R}^s_+ \times \mathbb{R}^{n-s}$ در نتیجه خواهیم داشت

$$\delta(D(\| \,.\, \|_\infty, x)) = \delta(\mathbb{R}^s_+) + \delta(\mathbb{R}^{n-s}) = s/۲ + n - s = n - s/۲$$

فرض کنید $y \in \mathbb{R}^s_+ \times \mathbb{R}^{n-s}$ پس $\lambda > ۰$ وجود دارد که

$$۰ \le \lambda y_i \le ۱ fori = ۱, \cdots, s-۱, s$$

$$\lambda |y_j| \le ۱ - |x_j| forj = s+۱, \cdots, n-۱, n$$

از آنجایی که برای $i = ۱, \cdots, s$ داریم $y_i \ge ۰$ پس برای $\lambda > ۰$ داریم $\lambda y_i \ge ۰$ و زمانی که $\lambda^{-۱} \ge max_{i=۱,\cdots,s} y_i$ آنگاه $\lambda y_i \le ۱$.زمانی که $\lambda^{-۱} \ge min_{j=s,\cdots,n} |y_i|/(۱ - |x_j|)$ را داشته باشیم نامساوی دوم رخ می دهد.بنابراین به ازای $\lambda^{-۱}$ به اندازه ی کافی بزرگ شرایط فوق صادق اند. پس

$$|x_i + \lambda y_i| = |\lambda y_i - ۱| \le ۱\ fori = ۱, \cdots, s$$

$$|x_j + \lambda y_j| = |x_j| + \lambda |y_j| \leq 1 for j = s, \cdots, n$$

با گرفتن ماکسیمم روی مختصه ها داریم $\| x \|_\infty = 1 \geq \| x + \lambda y \|_\infty$ و طبق تعریف $y \in D(\| \,.\, \|_\infty, x)$ و لذا $D(\| \,.\, \|_\infty, x) \supset \mathbb{R}^s_+ \times \mathbb{R}^{n-s}$. برای طرف عکس فرض کنید $y \notin \mathbb{R}^s_+ \times \mathbb{R}^{n-s}$ یا معادلا وجود داشته باشد $i_* \in \{1, \cdots, s\}$ که $y_{i_*} < 0$ آنگاه برای هر $\lambda > 0$ داریم $|x_{i_*} + \lambda y_{i_*}| = 1 - \lambda y_{i_*} > 1$ پس $\| x \|_\infty = 1 \geq \| x + \lambda y \|_\infty$ و لذا $y \notin D(\| \,.\, \|_\infty, x)$ در نتیجه $D(\| \,.\, \|_\infty, x) \subset \mathbb{R}^s_+ \times \mathbb{R}^{n-s}$ و اثبات کامل است. [۶] ■

تعریف ۲۰.۳. زیردیفرانسیل [۶] یک تابع محدب سره $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ در نقطه $x \in \mathbb{R}^n$ مجموعه محدب بسته زیر تعریف می شود

$$\partial f(x) = \{s \in \mathbb{R}^n : f(y) \geq f(x) + < s, y - x > \forall y \in \mathbb{R}^n\}$$

اگر $\partial f(x) \neq$ گوییم f زیرمشتق پذیر [۷] است.

مثال ۲۱.۳ ▲

بین مخروط های کاهشی و زیردیفرانسیل یک دوگانی وجود دارد که می توان برای محاسبه ی بعد آماری مخروط های کاهشی از آن استفاده کرد.

قضیه ۲۲.۳. فرض کنید زیردیفرانسیل $\partial f(x)$ ناتهی فشرده و شامل مبدا نباشد

$$\mathcal{D}(f, x)^o = cone(\partial f(x)) := \bigcup_{\tau > 0} \tau . \partial f(x)$$

برهان. به کمک نتیجه زیر ■

قضیه ۲۳.۳. فرض کنید f یک تابع محدب سره باشد و x نقطه ای باشد که f در آن زیرمشتق پذیر بوده اما مینیمم خود را در آن اخذ نمی کند.آنگاه مخروط نرمال $C = \{z | f(z) \leq f(x)\}$ در x بستار مخروط محدب تولید شده توسط $\partial f(x)$ است.

برهان. djvu ■

نتیجه ۲۴.۳. فرض کنید f یک تابع محدب سره باشد و x نقطه درونی $dom\ f$ باشد که f(x) مینیمم f نباشد.بردار x^* نرمال C در x است اگر و فقط اگر $\lambda \geq 0$ وجود داشته باشد که $x^* \in \lambda \partial f(x)$.

برهان. ■

[6] subdifferential
[7] subdifferentiable

قضیه ۲۵.۳. (کرانی برای بعد آماری مخروط های کاهشی).فرض کنید $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ یک تابع محدب سره و $x \in \mathbb{R}^n$ باشد و زیردیفرانسیل $\partial f(x)$ ناتهی فشرده و شامل مبدا نباشد.تابع زیر را تعریف می کنیم $J(\tau) := J(\tau; \partial f(x)) := \mathbb{E}[dist^2(g, \tau \cdot \partial f(x))] for \tau \geq 0$ که $g \sim \mathcal{N}(0, I)$ خواهیم داشت $\mathcal{D}(f,x) \leq inf_(\tau \geq 0) J(\tau)$.

برهان. بعد آماری مخروط کاهشی با بعد آماری بستار آن برابر است.برای اثبات از نمایش قطبی استفاده می‌کنیم و از این واقعیت که قطبی مخروط کاهشی مخروط تولید شده توسط زیردیفرانسیل می باشد؟ استفاده می شود.

$$\delta(\mathcal{D}(f,x)) = \delta(\mathcal{D}(f,x)^{oo}) = \mathbb{E}[dist^2(g, \mathcal{D}(f,x)^o)]$$

$$= \mathbb{E}[dist^2(g, \bigcup_{\tau > 0} \tau \cdot \partial f(x))] = \mathbb{E}[inf_{\tau \geq 0} dist^2(g, \tau \cdot \partial f(x))]$$

■

لم ۲۶.۳. فرض کنید S زیرمجموعه ای ناتهی؟ فشرده و محدب از $\mathbb{R}^n$ باشد که شامل مبدا نیست.به خصوص اعدادی وجود دارند که برای هر $s \in S$ داریم $b \leq \| s \| \leq B$. $u \in \mathbb{R}^n$ را ثابت درنظر گرفته و تعریف می کنیم

$$J_u : \tau \to dist^2(u, \tau S) for \tau \geq 0$$

خواص زیر برقرار خواهند بود:

۱)تابع محدب است. ۲) $(\tau) \geq (\tau \| \ \|)^2 \tau \geq \| \ \| /$ بخصوص مینیمم خود را در بازه ی فشرده $[\underline{0,2\|\|}]$ اخذ می کند. ۳)تابع J_u بطور پیوسته مشتق پذیر است و مشتق آن به شکل زیر خواهد بود

$$J'_u(\tau) = -\frac{2}{\tau} < u - \pi_{\tau S}(u), \pi_{\tau S}(u) > for \ \tau > 0$$

مشتق راست $J'_u(0)$ وجود دارد و $J'_u(0) = lim_{\tau \downarrow 0} J'_u(\tau)$. ۴) $|J'_u(\tau)| \leq 2B(\| u \| + \tau B) for \ \tau \geq 0$ ۵)نگاشت $u \mapsto J'_u(\tau)$ برای $\tau \geq 0$ لیپشیتز است:

$$|J'_u() - J'_y(\tau)| \leq 2B \| u - y \| \ for \ all u, y \in \mathbb{R}^n$$

برهان. تحدب.برای $\tau > 0$

$$J_u(\tau) = (inf_{s \in S} \| u - \tau s \|)^2 = (\tau \cdot inf_{s \in S} \| u/\tau - s \|)^2 = (\tau \cdot dist(u/\tau, S))^2$$

فاصله تا یک مجموعه بسته ی محدب؟ تابعی محدب است.تبدیل پرسپکتیو یک تابع محدب؟ محدب است و مربع یک تابع نامنفی محدب نیز محدب است پس $J_u(\tau)$ محدب است. پیوستگی.

■

قضیه ۲۷.۳. فرض کنید x برداری با درایه های ناصفر در $\mathbb{R}^n$ باشد.آنگاه بعد آماری مخروط کاهشی نرم یک در x در کران های زیر صدق می کند.

$$\psi(sn) - \frac{۲}{\sqrt{sn}} \le \frac{\delta(\mathcal{D}(\| \cdot \|_۱, x)}{n} \le \psi(sn)$$

تابع $\psi : [۰, ۱] \to [۰, ۱]$ بصورت زیر تعریف می شود

$$\psi(\rho) := inf_{\tau \ge ۰}\{\rho(۱ + \tau^۲) + (۱ - \rho) \int_\tau^\infty (u - \tau)^۲ \cdot \phi(u)\, du\}$$

که $\phi(u) := \sqrt{۲/\pi} e^{-u^۲۲}$ یک چگالی احتمال روی $[۰, \infty)$ است.اینفیمم برای یک τ یکتا مثبت با حل معادله ثابت زیر بدست می اید

$$\int_\tau^\infty (\frac{u}{\tau} - ۱) \cdot \phi(u)\, du = \frac{\rho}{۱ - \rho}$$

برهان. ■

۳-۴ فرمول مخروطی kinematic

تعریف ۲۸.۳. فرض کنید C یک مخروط محدب در $\mathbb{R}^n$ باشد.برای هر $k = ۰, ۱, \cdots, n$ k -امین *tail functional* بصورت زیر تعریف می شود

$$t_k(C) := v_k(C) + v_{k+۱}(C) + \cdots + v_n(C)$$

و k -امین *half-tail functional* بصورت $h_k(C) := \sum_{j=k, j-keven}^n v_j(C)$ تعریف می شود.

قضیه ۲۹.۳. *(Interlacing)* برای هر مخروط محدب بسته در $\mathbb{R}^n$ که زیر فضای خطی نیست داریم

$$۲h_k(C) \ge t_k(C) \ge ۲h_{k+۱}(C) \forall k = ۰, ۱, \cdots, n - ۱$$

لم ۳۰.۳. *(Conic kinematic formula)* فرض کنید C و K مخروط هایی محدب و بسته در $\mathbb{R}^n$ باشد و C زیرفضا نباشد.پس

$$\mathbb{P}\{C \cap QK \ne ۰\} = ۲h_{n+۱}(C \times k)$$

برای زیرفضا خطی n-m بعدی L_{n-m} از $\mathbb{R}^n$ شرح فوق به فرمول *Crofton* کاهش می یابد

$$\mathbb{P}\{C \cap QL_{n-m} \ne ۰\} = ۲h_{m+۱}(C)$$

برهان. ■

برهان. فرض کنید L_{n-k+1} زیرفضا خطی با بعد $n-k+1$ باشد و L_{n-k} زیر فضا خطی با بعد $n-k$ در L_{n-k+1} باشد.فرمول Crofton نشان می دهد half-tail functionals در حال کاهش اند

$$2h_{k+1}(C) = \mathbb{P}C \cap QL_{n-k} \neq 0 \leq \mathbb{P}C \cap QL_{n-k+1} \neq 0 = 2h_k(C)$$

از طرفی $\frac{1}{2}t_k(C) = \frac{1}{2}[h_k(C) + h_{k+1}(C)]$ در نتیجه $2h_k(C) \geq t_k(C) \geq 2h_{k+1}(C)$. ■

تبصره ۳۱.۳. (توسیع فرمول *kinematic*).با استقرا فرمول *kinematic* برای خانواده $\{C, k_1, \cdots, k_r\}$ از مخروط های محدب در $\mathbb{R}^n$ تعمیم می یابد.اگر C زیرفضا نباشد آنگاه

$$\mathbb{P}\{C \cap Q_1 k_1 \cap \cdots \cap Q_r k_r \neq 0\} = 2h_{rn+1}(C \times k_1 \times \cdots \times k_r)$$

هر ماتریس $Q_i \in \mathbb{R}^{n \times n}$ پایه متعامد تصادفی است که مستقل از بقیه انتحاب می شود.

این نتیجه در تجزیه و تحلیل مسائل demixing [8] با بیش از دو متغیر اصلی استفاده می شود. [۷]

۴-۳.۱ تخمین فرمول kinematic

قضیه ۳۲.۳. فرض کنید $\lambda \geq 0$ و C یک مخروط محدب در $\mathbb{R}^n$ باشد برای یک زیرفضا $n-m$ بعدی l_{n-m} داریم

$$m \geq \delta(C) + \lambda \Rightarrow \mathbb{P}\{C \cap Ql_{n-m} \neq \{0\}\} \leq p_C(\lambda)$$

$$m \leq \delta(C) - \lambda \Rightarrow \mathbb{P}\{C \cap Ql_{n-m} \neq \{0\}\} \geq 1 - p_C(\lambda)$$

برای هر مخروط محدبK در $\mathbb{R}^n$ داریم

$$\delta(C) + \delta(K) \leq n - 2\lambda \Rightarrow \mathbb{P}\{C \cap QK \neq \{0\}\} \leq p_C(\lambda) + p_K(\lambda)$$

$$\delta(C) + \delta(K) \geq n + 2\lambda \Rightarrow \mathbb{P}\{C \cap QK \neq \{0\}\} \geq 1 - (p_C(\lambda) + p_K(\lambda))$$

که $p_C(\lambda)$ و $p_K(\lambda)$ در $----$ تعریف شده اند.

برهان. از آنجایی که $\mathbb{P}\{C \cap QK \neq \{0\}\} = \mathbb{P}\{\bar{C} \cap Q\bar{K} \neq \{0\}\}$ پس می توان فرض کرد C و K بسته اند.اگر C زیرفضا باشد شماره ۱ به وضوح برقرار است.فرض کنید C زیرفضا نباشد طبق فرمول Crofton داریم

$$\mathbb{P}\{C \cap Ql_{n-m} \neq \{0\}\} = 2h_{m+1}(C) \leq t_m(C)$$

8 nois

طبق قضیه $---$ زمانی که $m \geq \delta(C)+\lambda$ داریم $t_m(C) \leq p_C(\lambda)$. خط دوم بطور موازی نشان داده خواهد شد و بدین ترتیب بخش اول ثابت می شود. حال $\delta(C)+\delta(K)n-۲\lambda$ را در نظر بگیرید.طبق فرمول kinematic داریم

$$\mathbb{P}\{C \cap Ql_{n-m} \neq \{۰\}\} = ۲h_{n+۱}(C \times K) \leq t_n(CK)$$

لم ۳۳.۳. فرض کنید C و K مخروط هایی محدب و بسته باشند.آنگاه

$$t_{\lceil \delta(C)+\delta(K)+۲\lambda \rceil}(C \times K) \leq t_{\lceil \delta(C)+\lambda \rceil}(C) + t_{\lceil \delta(K)+\lambda \rceil}(K)$$

برهان. فرض کنید $C, K \subset \mathbb{R}^n$ مخروط هایی محدب و بسته باشند.متغیرهای مستقل تصادفی X وy را تعریف می کنیم بطوریکه

$$\mathbb{P}\{X = k\} = v_k(C), \mathbb{P}\{Y = k\} = v_k(K), fork = ۰, ۱, \cdots, n$$

طبق خواص حجم ذاتی

$$\mathbb{P}\{X + Y = k\} = v_k(C \times K) fork = ۰, ۱, \cdots, ۲n$$

پس

$$\mathbb{P}\{X + Y \geq \delta(C) + \delta(K) + ۲\lambda\} \leq \mathbb{P}\{X \geq \delta(C) + \lambda\} + \mathbb{P}\{Y \geq \delta(K) + \lambda\}$$

و در نتیجه

$$t_{\lceil \delta(C)+\delta(K)+۲\lambda \rceil}(C \times K) \leq t_{\lceil \delta(C)+\lambda \rceil}(C) + t_{\lceil \delta(K)+\lambda \rceil}(K)$$

■

از آنجایی که tail functionals ها ضعیفا نزولی می باشند به کمک لم فوق فرض $\delta(C)+\delta(K) \leq n-۲\lambda$ نتیجه می دهد

$$t_n(C \times K) \leq t_{\lceil \delta(C)+\delta(K)+۲\lambda \rceil}(C \times K) \leq t_{\lceil \delta(C)+\lambda \rceil}(C) + t_{\lceil \delta(K)+\lambda \rceil}(K)$$

طبق قضیه $---$ داریم

$$t_{\lceil \delta(C)+\lambda \rceil}(C) + t_{\lceil \delta(K)+\lambda \rceil}(K) \leq p_C(\lambda) + p_K(\lambda)$$

پس $\mathbb{P}\{C \cap QK \neq \{۰\}\} \leq p_C(\lambda) + p_K(\lambda)$ و اثبات کامل است.دومین خط نیز بطور موازی نشان داده میشود.

■

قضیه ۳۴.۳. $\eta \in (0,1)$ را ثابت در نظر بگیرید.فرض کنید C و K مخروط هایی محدب در $\mathbb{R}^n$ باشند و Q یک پایه ی تصادفی متعامد در $\mathbb{R}^{n\times n}$.آنگاه

$$\delta(C)+\delta(K) \leq n - a_\eta\sqrt{n} \Rightarrow \mathbb{P}\{C\cap QK \neq \{0\}\} \leq \eta$$

$$\delta(C)+\delta(K) \geq n + a_\eta\sqrt{n} \Rightarrow \mathbb{P}\{C\cap QK \neq \{0\}\} \geq 1-\eta$$

$a_\eta = \sqrt{(8log(4\eta))}$.

برهان. مشابه قضیه ی قبل ابتدا فرض می کنیم $\delta(C)+\delta(K) \leq n - 2\lambda$.طبق قضیه ی $---$

$$\mathbb{P}\{C\cap QK \neq \{0\}\} = 2h_{n+1}(C\times K) \leq t_n(C\times K)$$

از طرفی داریم $\delta(C)+\delta(K) = \delta(C\times K)$ و طبق قضیه $---$

$$t_n(C\times K) \leq p_{C\times K}(\lambda) \leq 4exp(\frac{-\lambda^2 8}{\delta(C\times K)+\lambda} \leq 4e^{-\lambda^2 8n}$$

پس

$$\mathbb{P}\{C\cap QK \neq \{0\}\} \leq 4e^{-\lambda^2 8n}$$

حال با تغییر متغیر λ به $a_\eta\sqrt{n}$ که $a_\eta = \sqrt{8log(4\eta)}$ خواهیم داشت

$$\mathbb{P}\{C\cap QK \neq \{0\}\} \leq \eta$$

. ■

تبصره ۳۵.۳. (قضیه ی فرایندهای گاوسی) اگر پایه ی متعامد تصادفی را با ماتریس نرمال استاندارد عوض کنیم شاهد مسئله ی متفاوتی در هندسه ی تصادفی خواهیم بود.رودلسن و ورشاینین از نتایج این قضیه در جهت مطالعه ی روش مینیمم سازی نرم l_1 برای فشرده سازی استفاده کردند.مدل متعامد تصادفی برخلاف اینکه بسیار طبیعی تر از مدل گاوسی است در خیلی از کاربردها اعمال نمی شود.

۵-۳ بعد آماری متعارف است.

حال به بررسی برخی پیامدهای قابل توجه بعد آماری در هندسه ی انتگرالی می پردازیم. فرض کنید ϕ_n خانواده ای از مخروط های محدب در $\mathbb{R}^n$ باشد که با متر هاوسدورف مخروطی برای تشکیل یک فضای متریک فشرده مجهز شده است.یک تابعک هندسی $v:\phi_n \to \mathbb{R}$ تابع ارزیاب پیوسته ناوردا نسبت به دوران [9] گفته می شود هرگاه در ویژگی های زیر صدق کند: ۱)ارزه(۱) .برای مخروط بدیهی $v(\{0\}) = 0$

[9] a continuous, rotationinvariant valuation

. ۲)ارزه(۲) .اگر $C, K, C \cup K \in \phi_n$ آنگاه $v(C \cup K) + v(C \cap K) = v(C) + v(K)$. ۳)ناوردایی نسبت به دوران.برای هر $C \in \phi_n$ داریم $v(UC) = v(C)$ برای هر ماتریس متعامد $U \in \mathbb{R}^{n \times n}$. ۴)پیوستگی.اگر $C_i \mapsto C$ در ϕ_n آنگاه $lim_{i \to \infty} vC_i \to v(C)$. توابع ارزیاب پیوسته ناوردا نسبت به دوران؟ اندازه های هندسی طبیعی هستند که روی مخروط های محدب تعریف شده اند.یک مورد خاص این توابع حجم ذاتی است.بنابر تعریف بعد آماری خواهیم دید که این ویژگی توسط بعد آماری نیز به ارث برده می شود.هر تابع ارزیاب پیوسته ناوردا نسبت به دوران؟ توسط مقادیری که روی زیرفضا های خطی می گیرد مشخص می شود در نتیجه بعد آماری یک تابع ارزیاب پیوسته ناوردا نسبت به دوران یکتا است که برای هر زیرفضا مقدار آن با بعد زیرفضا برابر می باشد.به عبارت دیگر بعد آماری بطور طبیعی بعد خطی را به کلاس مخروط های محدب بسته توسیع داد.

فصل ۴

بررسی

۴-۱ مسائل معکوس خطی منظم با یک مدل تصادفی

فرض کنید $x_\circ \in \mathbb{R}^n$ برداری با s درایه ی ناصفر باشد که دقیق آن را نمی شناسیم و A یک ماتریس تصادفی $m \times n$ با درایه های متغیر نرمال استاندارد باشد و $z_\circ = Ax_\circ \in \mathbb{R}^m$ برداری باشد که ما به آن دسترسی داریم.مسئله ی فشرده سازی شناسایی $x_\circ$ از روی A و $z_\circ$ است.زمانی که m بزرگتر از n باشد حل این مسئله در صورتی راحت است که اطلاعاتی در مورد $x_\circ$ داشته باشیم برای مثال تنک بودن آن.

۴-۱.۱ حل مسائل معکوس خطی به کمک بهینه سازی محدب

فرض کنید $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ یک تابع محدب سره باشد.می توان $x_\circ$ را با حل مسئله بهینه سازی محدب زیر شناسایی کرد.

$$minimize\ f(x)\ subject\ to\ z_\circ = Ax$$

تابع f تنظیم کننده [1] نامیده می شود.

مثال ۴.۱ (بردار تنک) زمانی که بدانیم $x_\circ$ برداری تنک می باشد می توان از مینیمم سازی l_1 نرم برای یافتن $x_\circ$ استفاده کرد.

$$minimize \parallel x \parallel_1\ \ subjecttoz_\circ = Ax$$

که $\parallel x \parallel_1 = \sum_{i=1}^{n} |x_i|$. ▲

مثال ۴.۲ (ماتریس با رتبه کم) اگر رتبه ی کم باشد می توان از Schatten 1-norm استفاده کرد.

$$minimize \parallel x \parallel_{S_1}\ \ subjecttoz_\circ = Ax$$

[1] regularizer

که $\| x \|_{S_1} = tr(\sqrt{x^*x})$. ▲

زمانی موفق به شناسایی x می شویم که مسئله ی بهینه سازی یک مینیمم ساز یکتا $\hat{x}$ داشته باشد که $\hat{x} = x_0$. برای بهبود شرایط مینیمم سازی از مخروط های کاهشی استفاده می شود.

قضیه ۳.۴. x_0 جواب بهینه یکتا مسئله خواهد بود اگر و فقط اگر $\mathcal{D}(f, x_0) \cap null(A) = \{0\}$.

برهان. [۸] ■

۱-۴.۲ مسئله ی معکوس خطی با داده های تصادفی

یک تکنیک مدل سازی طبیعی ترسیم A بطور تصادفی از توزیع نرمال استاندارد در $\mathbb{R}^n$ است.در این حالت هسته ماتریس A یک زیرفضا جهت دار تصادفی خواهد بود که با توجه به قضیه ی قبل ما نیازمند محاسبه ی احتمال اشتراک این زیرفضای تصادفی با مخروط کاهشی خواهیم بود.

فرمول kinematic عبارت دقیقی را برای احتمال موفقیت مسئله ی بهینه سازی با A تصادفی را به ما می دهد و با استناد به تخمین این فرمول نتیجه ی ساده ای بدست می اید که منجر به شناسایی گذار شدید می شود.

۱-۴.۳ گذار فاز در مسائل معکوس خطی با اندازه گیری های تصادفی

قضیه ۴.۴. $\eta \in (0, 1)$ را ثابت در نظر بگیرید و فرض کنید $x_0 \in \mathbb{R}^n$ برداری ثابت و $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ یک تابع محدب سره باشد. $A \in \mathbb{R}^{m\times n}$ را ماتریسی با درایه های مستقل نرمال استاندارد درنظر بگیرید و $z_0 = Ax_0$.آنگاه

$m \leq \delta(\mathcal{D}(f, x_0)) - a_\eta\sqrt{n} \Rightarrow$ مسئله ی بهینه سازی با احتمال کمتر از η موفقیت آمیز خواهد بود

$m \geq \delta(\mathcal{D}(f, x_0)) + a_\eta\sqrt{n} \Rightarrow$ مسئله ی بهینه سازی با احتمال بیشتر از $1 - \eta$ موفقیت آمیز خواهد بود

$a_\eta := \sqrt{(8 log(4\eta))}$.

برهان. توزیع نرمال استاندارد روی $\mathbb{R}^{n\times m}$ تحت دوران ناوردا ست. بنابراین فضای پوچ $null(A)$ یک زیرفضا تصادفی یکنواخت $d - m$ بعدی از $\mathbb{R}^n$ می باشد. پس بعد آماری آن $d - m$ است.طبق قضیه ی داریم

$$n - m + \delta(\mathcal{D}(f, x_0)) \leq n - a_\eta\sqrt{n} \Rightarrow \mathbb{P}\{null(A) \cap Q\mathcal{D}(f, x_0) \neq \{0\}\} \leq \eta$$

$$n - m + \delta(\mathcal{D}(f, x_0)) \geq n + a_\eta\sqrt{n} \Rightarrow \mathbb{P}\{null(A) \cap Q\mathcal{D}(f, x_0) \neq \{0\}\} \geq 1 - \eta$$

و طبق قضیه ی داریم

$$\delta(\mathcal{D}(f, x_0)) + a_\eta\sqrt{n} \leq m \Rightarrow \mathbb{P}\{null(A) \cap Q\mathbb{D}(f, x_0) = \{0\}\} \geq 1 - \eta$$

$$\delta(\mathcal{D}(f,x_\circ)) - a_\eta\sqrt{n} \geq m \Rightarrow \mathbb{P}\{null(A) \cap Q\mathcal{D}(f,x_\circ) \neq \{\circ\}\} \leq \eta$$

■

این قضیه نشان می دهد که ما همیشه با یک گذار فاز برای حل مسئله ی معکوس خطی با داده های تصادفی روبرو هستیم.گذار زمانی رخ می دهد که $\delta(\mathcal{D}(f,x_\circ)) = m$.

از آنجایی که فقط m معادله داریم نمی توان $x_\circ \in \mathbb{R}^n$ را با حل $Ax = z_\circ$ و مشاهده ی $z_\circ \in \mathbb{R}^m$ شناسایی کرد.مسئله بهینه سازی محدب با تابع تنظیم کننده $n - \delta(\mathcal{D}(f,x_\circ))$ معادله موثر به دستگاه اضافه می کند و در نتیجه زمانی که $m \geq \delta(\mathcal{D}(f,x_\circ))$ می توانیم $x_\circ$ را بازیابی کنیم. بعلاوه این قضیه امکان مقایسه ی عملکرد تنظیم کننده های مختلف را نیز به ما می دهد.

۲-۴ مسائل demixing

در این مسائل ما به دنبال استخراج ساختارها از یک ترکیب آن ها هستیم.فرض کنید $z_\circ \in \mathbb{R}^n$ را به فرم $z_\circ = x_\circ + Uy_\circ$ داریم که $x_\circ, y_\circ \in \mathbb{R}^n$ را نمی شناسیم اما $U \in \mathbb{R}^{n\times n}$ را که یک ماتریس متعامد می باشد می شناسیم.برای شناسایی زوج $(x_\circ, y_\circ)$ باید فرض کنیم هر مولفه برای کاهش درجه آزادی ساخته شده است. علاوه بر این اگر دو نوع ساختار منسجم باشند (یعنی با یکدیگر هم تراز باشند) ممکن است جدا کردن آنها غیرممکن باشد ، بنابراین برای مدل سازی جهت گیری نسبی دو سیگنال سازنده استفاده از ماتریس U به مناسب خواهد بود.

۲-۴.۱ حل مسائل demixing به کمک بهینه سازی محدب

فرض کنید f و g توابعی محدب و سره روی $\mathbb{R}^n$ باشند.می توان نوشت

$$minimize f(x) \ subject \ to g(y) \leq g(y_\circ) \ and z_\circ = x + Uy$$

در این رویکرد به اطلاعات جانبی $g(y_\circ)$ نیازمندیم. از اینرو گاهی در عمل استفاده از لاگرانژین طبیعی تر می باشد.

مثال ۴.۵ (تنک + تنک)فرض کنید سیگنال اول $x_\circ$ در پایه ی استاندارد و سیگنال دوم $Uy_\circ$ در پایه ی مشخص U تنک باشند.در این مورد می توان از نرم یک استفاده کرد

$$minimize \ \ \| x \|_1 \ \ subject \ to \ \ \| y \|_1 \leq \| y_\circ \|_1 \ \ and z_\circ = x + Uy$$

این رویکرد ترکیب سیگنال های پراکنده انالیز مولفه های مورفولوژیکی [2] نامیده می شود.
▲

[2]morphological component analysis

مثال ۶.۴ (کم رتبه + تنک)فرض کنید $Z_{\circ} = X_{\circ} + U(Y_{\circ})$ مشاهده شده باشد. $X_{\circ}$ یک ماتریس کم رتبه است و $Y_{\circ}$ تنک. U یک تبدیل متعامد شناخته شده روی فضای ماتریس ها می باشد.می توان نوشت

$$minimize \ \| x \|_{S_1} \ subject \ to \| Y \|_1 \leq \| Y_{\circ} \|_1 \ and \ Z_{\circ} = X_{\circ} + U(Y)$$

این مسئله demixing تجزیه ی $rank - sparsity$ نامیده می شود. ▲

قضیه ۷.۴. فرض کنید f و g توابعی محدب سره ای باشند.زوج $(x_{\circ}, y_{\circ})$ جواب بهینه یکتا مسئله بهینه سازی است اگر و تنها اگر $\mathcal{D}(f, x_{\circ}) \cap (-U\mathcal{D}(g, y_{\circ})) = \{\circ\}$.

قضیه ۸.۴. $\eta \in (\circ, 1)$ را ثابت در نظر بگیرید و فرض کنید $x_{\circ}, y_{\circ} \in \mathbb{R}^n$ برداری ثابت و $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ و $g : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ توابعی محدب سره باشند.پایه ی متعامد تصادفی $U \in \mathbb{R}^{n \times n}$ را درنظر بگیرید و فرض کنید $z_{\circ} = x_{\circ} + Uy_{\circ}$.آنگاه

$$\delta(\mathcal{D}(f, x_{\circ})) + \delta(\mathcal{D}(g, y_{\circ})) \geq n + a_{\eta}\sqrt{n} \Rightarrow$$

مسئله ی بهینه سازی با احتمال کمتر از اتا موفقیت آمیز خواهد بود

$$\delta(\mathcal{D}(f, x_{\circ})) + \delta(\mathcal{D}(g, y_{\circ})) \leq n - a_{\eta}\sqrt{n} \Rightarrow$$

مسئله ی بهینه سازی با احتمال بیشتر ازیک منهای اتا موفقیت آمیز خواهد بود
$a_{\eta} = \sqrt{8 log(4\eta)}$.

برهان. طبق قضیه داریم

$$\delta(\mathcal{D}(f, x_{\circ})) + \delta(\mathcal{D}(g, y_{\circ})) \leq n - a_{\eta}\sqrt{n} \Rightarrow \mathbb{P}\{\mathcal{D}(f, x_{\circ}) \cap (-U\mathcal{D}(g, y_{\circ})) \neq \{\circ\}\} \leq \eta$$

$$\delta(\mathcal{D}(f, x_{\circ})) + \delta(\mathcal{D}(g, y_{\circ})) \geq n + a_{\eta}\sqrt{n} \Rightarrow \mathbb{P}\{\mathcal{D}(f, x_{\circ}) \cap (-U\mathcal{D}(g, y_{\circ})) \neq \{\circ\}\} \geq 1 - \eta$$

و طبق قضیه ی داریم

$$\delta(\mathcal{D}(f, x_{\circ})) + \delta(\mathcal{D}(g, y_{\circ})) \geq n + a_{\eta}\sqrt{n} \Rightarrow \mathbb{P}\{\mathcal{D}(f, x_{\circ}) \cap (-U\mathcal{D}(g, y_{\circ}))\} \geq 1 - \eta$$

$$\delta(\mathcal{D}(f, x_{\circ})) + \delta(\mathcal{D}(g, y_{\circ})) \leq n - a_{\eta}\sqrt{n} \Rightarrow \mathbb{P}\{\mathcal{D}(f, x_{\circ}) \cap (-U\mathcal{D}(g, y_{\circ}))\} \leq \eta$$

■

بهینه سازی زمانی موثر است که بعد آماری کلی دو مخروط های کاهشی کوچکتر از بعد فضا n باشد.

فصل ۵
ارجاعات مهم

فصل ۶

نتیجه گیری

مراجع

[1] Amelunxen, D. and Lotz, M. Intrinsic Volumes of Polyhedral Cones: A Combinatorial Perspective , 10.1007/s00454-017-9904-9.

[2] Donoho, D. L. and Tanner, J. Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. . *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 367, 4273–4293.

ت

[3] Sommerville, D. M. Y. The Relations Connecting the Angle-Sums and Volume of a Polytope in Space of n Dimensions *Proc. R. Soc. Lond. A 1927 115, 103-119*, 10.1098/rspa.1927.0078.

ت، ۶

[4] Amelunxen, D. and Lotz, M. Living on the edge: phase transitions in convex programs with random data *A Journal of the IMA (2014) 3, 224–294*,10.1093/imaiai/iau005.

۷

[5] Ablowits, Mark.J Complex Variables: Introduction and Applications Second Edition (Cambridge Texts in Applied Mathematics) 2nd *A Journal of the IMA (2014) 3, 224–294*,10.1093/imaiai/iau005.

۸

[6] McCoy, M B. ageometric analysis of convex demixing 0,0.

۱۱

[7] McCoy, M B. and Joel, A T The achievable performance of convex demixing 0,0.

۱۴

[8] Chandrasekaranm, V. Rechtw, B. Parrilom, PA. and Willskym,A S. The Convex Geometry of Linear Inverse Problems Foundations of Computational Mathematics, Vol. 12, No. 6, pp. 805-849, 2012,10.1007/S10208-012-9135-7.

۱۹

Abstract

Keywords: Convex optimization, phase transition, stochastic geometry, sparse vectors, low-rank matrices.

Sharif University of Technology

Department of Mathematical Sciences

Thesis

Submitted in partial fulfillment of the requirements for the degree of

Master in pure mathematics

Phase Transition in Convex Optimization Problems with Random Data

By

Delbar Faghih

Supervisor

Prof. K. Alishahi

September 2017