به نام خدا

آنالیز احتمالاتی در ابعاد بالا (۱-۲۵۸۴۱)

تمرین سری سوم

ترم پاییز ۰۴-۱۴۰۳

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

استاد: دکتر محمّدحسین یاسائی میبدی

مهلت تحویل: سه‌شنبه ۲۷ آذر ساعت ۲۳:۵۹

(*) مسائلی که با ستاره مشخّص شده‌اند امتیازی هستند و حل‌کردن آن‌ها نمره‌ی امتیازی خواهد داشت!

۱ چینینگ فرآیندهای با دم دلخواه

روش Chaining تنها به فرآیندهای زیرگوسی محدود نمی‌شود. فرض کنید $\{X_t\}_{t\in\mathcal{T}}$ فرآیندی تصادفی با میانگین صفر بوده و برای هر $\lambda \geq 0$ و $s,t\in\mathcal{T}$ داشته باشیم:

$$\log\,\mathbb{E}\left[e^{\lambda\frac{X_t-X_s}{d(t,s)}}\right]\leq\psi(\lambda),$$

که در آن ψ تابعی محدّب است و $\psi(0)=\psi'(0)=0$. نشان دهید:

$$\mathbb{E}\left[\sup_{t\in\mathcal{T}}X_t\right]\leq C\int_0^\infty\psi^{*-1}\left(2\log\mathcal{N}(\mathcal{T},d,\epsilon)\right)d\epsilon,$$

که در آن $\psi^*(x)=\sup_{\lambda\geq 0}\{\lambda x-\psi(\lambda)\}$.

۲ ماتریس تصادفی گوسی

فرض کنید درایه‌های ماتریس تصادفی $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ گوسی استاندارد و مستقل از هم و $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ دو بردار دلخواه داده‌شده باشند.

۱. با استفاده از صورت‌های معادل متغیّرهای تصادفی گوسی ثابت کنید:

$$\mathbb{P}\left[\mathbf{u}^\top\boldsymbol{A}^\top\boldsymbol{A}\mathbf{v}>a\right]\leq\frac{e^{-sa}}{(1-2s\mathbf{u}^\top\mathbf{v})^{m/2}},\quad s>0,$$
$$\mathbb{P}\left[\mathbf{u}^\top\boldsymbol{A}^\top\boldsymbol{A}\mathbf{v}\leq a\right]\leq\frac{e^{sa}}{(1+2s\mathbf{u}^\top\mathbf{v})^{m/2}},\quad s>0.$$

۲. ثابت کنید:

$$\mathbb{P}\left[\left|\frac{1}{m}\mathbf{u}^\top\boldsymbol{A}^\top\boldsymbol{A}\mathbf{v}-\mathbf{u}^\top\mathbf{v}\right|>\epsilon\mathbf{u}^\top\mathbf{v}\right]\leq 2e^{-m\left(\frac{\epsilon^2}{4}-\frac{\epsilon^3}{6}\right)}.$$

راهنمایی: می‌توانید از رابطه‌ی $\det\left\{\mathbf{I}-\mathbf{u}\mathbf{v}^\top\right\}=1-\mathbf{u}^\top\mathbf{v}$ استفاده نمایید.

۳ نابهینه بودن کران دودلی

در این تمرین قصد داریم نشان دهیم که کران دودلی در یک مسئله‌ی خاص، tight نیست. فرض کنید $\mathbf{e}_۱, \mathbf{e}_۲, \dots, \mathbf{e}_n$ پایه‌ی استاندارد (کانونی) فضای $\mathbb{R}^n$ باشد. مجموعه‌ی $\mathcal{T}$ را به شکل زیر تعریف می‌کنیم:

$$\mathcal{T} = \left\{ \frac{\mathbf{e}_k}{\sqrt{۱ + \log(k)}},\ k = ۱, \dots, n \right\}.$$

۱. عرض گوسی مجموعه‌ی $\mathcal{T}$ به صورت زیر تعریف می‌شود:

$$w(\mathcal{T}) = \mathbb{E}\left[\sup_{\mathbf{t}\in\mathcal{T}} \left\{\mathbf{t}^\top \boldsymbol{X}\right\}\right],$$

که $\boldsymbol{X} \sim \mathcal{N}(\circ, \mathbf{I}_n)$. نشان دهید عرض گوسی مجموعه‌ی $\mathcal{T}$ محدود است.

۲. نشان دهید انتگرال دودلی واگراست، یعنی:

$$\int_{\circ}^{+\infty} \sqrt{\log \mathcal{N}(\mathcal{T}, d, \epsilon)} d\epsilon \to \infty.$$

۴ نسخه‌ی با احتمال بالای کران دودلی

فرض کنید $\{X_t\}_{t\in\mathcal{T}}$ یک فرآیند تصادفی روی فضای متریک $(\mathcal{T}, d)$ باشد. مشابه با کران دودلی، فرض کنید $X_t - X_s$ فرآیندی زیرگوسی با پارامتر $d(s,t)$ است. نشان دهید که برای هر $u \geq \circ$، نامساویِ:

$$\sup_{s,t\in\mathcal{T}} |X_t - X_s| \leq C\left[\int_{\circ}^{+\infty} \sqrt{\log \mathcal{N}(\mathcal{T}, d, \epsilon)} d\epsilon + u\ \mathrm{diam}(\mathcal{T})\right]$$

با احتمال حدّاقل $۱ - ۷\exp(\frac{-u^۲}{۲})$ برقرار است.

۵ اصلاح کران دودلی برای فرآیندهای لیپشیتز

فرض کنید $\{X_t\}_{t\in\mathcal{T}}$ یک فرآیند تصادفی لیپشیتز و زیر گوسی نسبت به فضای متریک $(\mathcal{T}, d)$ باشد، یعنی به ازای هر $s, t \in \mathcal{T}$ و هر $\lambda \in \mathbb{R}$ داریم:

$$\mathbb{E}\left[e^{\lambda(X_t - X_s)}\right] \leq \exp\left(\frac{\lambda^۲ d^۲(s,t)}{۲}\right),$$

و همچنین متغیّری تصادفی مانند C وجود دارد که به ازای هر $s, t \in \mathcal{T}$ داریم:

$$|X_t - X_s| \leq C\dot{d}(s,t).$$

نشان دهید:

$$\mathbb{E}\left[\sup_{t\in\mathcal{T}}\{X_t\}\right] \leq \int_{\delta>\circ} \left\{۲\delta\,\mathbb{E}[C] + ۱۲\int_{\delta}^{+\infty} \sqrt{\log\left(\mathcal{N}(\mathcal{T}, d, \epsilon)\right)} d\epsilon\right\}.$$

۶ $(*)$ فاصله‌ی واسراشتاین

بردارهای تصادفی $\{\boldsymbol{X}_۱, \dots, \boldsymbol{X}_n\}$ به صورت i.i.d. از توزیع P روی مربّع واحد $[۰,۱]^۲$ نمونه‌برداری می‌شوند. توزیع تجربی متناظر با این نمونه‌ها را با P_n نشان می‌دهیم. فاصله‌ی واسراشتاین این دو توزیع به صورت زیر تعریف می‌شود:

$$\mathsf{W}_۱(P_n, P) = \sup_{f:۱-\mathrm{Lip}} \{\mathbb{E}_{P_n}[f(\boldsymbol{X})] - \mathbb{E}_P[f(\boldsymbol{X})]\}.$$

برای سادگی، فرض کنید لیپشیتز بودن نسبت به نرم ∞ سنجیده می‌شود. یعنی f تابعی ۱–لیپشیتز است، اگر و تنها اگر به ازای هر $\mathbf{x}, \mathbf{y}$ داشته باشیم:

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq \|\mathbf{x} - \mathbf{y}\|_\infty.$$

تعریف می‌کنیم: $\mathcal{F}_۰ = \{f : ۱-\mathrm{Lip}, f(۰) = ۰\}$.

۱. نشان دهید:

$$\mathcal{N}(\mathcal{F}_۰, \|\cdot\|_\infty, \epsilon) \leq \exp\left(\frac{c}{\epsilon^۲}\right).$$

۲. نشان دهید که انتگرال دودلی در این مسئله واگراست و کمکی به محاسبه‌ی کرانی برای فاصله‌ی واسراشتاین نمی‌کند.

۳. با استفاده از پرسش قبل، نشان دهید:

$$\mathbb{E}[\mathsf{W}_۱(P_n, P)] \leq C\frac{\log(n)}{\sqrt{n}}.$$

۷ روشی عجیب برای شبیه‌سازی ابر کره

فرض کنید m نقطه در فضای اقلیدسی $\mathbb{R}^n$ به صورت i.i.d از توزیع $\mathcal{N}(۰, \mathbf{I}_n)$ نمونه‌برداری شده‌اند. همچنین فرض کنید $m \geq \exp(Cn)$ که C یک ضریب ثابت به اندازه‌ی کافی بزرگ است. نشان دهید با احتمال بالا، پوش محدّب این نقاط تقریباً برابر یک گوی اقلیدسی nـبعدی با شعاع $\log n$ است. به طور دقیق‌تر، اگر تعریف کنیم $\mathcal{B}_۲^n(r) = \{\mathbf{x} \in \mathbb{R}^n : \|x\|_۲ \leq r\}$، و نقاط تصادفی را در مجموعه‌ی $\mathcal{A} = \{\boldsymbol{X}_۱, \dots, \boldsymbol{X}_m\}$ گردآوری کنیم، نشان دهید مقدار مثبت کوچکی مانند ϵ وجود دارد، به قسمی که:

$$(۱-\epsilon)\mathcal{B}_۲^n(\log n) \subseteq \mathrm{convhull}(\mathcal{A}) \subseteq (۱+\epsilon)\mathcal{B}_۲^n(\log n).$$

همچنین تعریف پوش محدب یا convex hull به این صورت است:

$$\mathrm{convhull}(\mathcal{A}) = \left\{\mathbf{y} \in \mathbb{R}^n : \exists a_۱, \dots, a_m \in [۰,۱] \text{ s.t. } \sum_{i=۱}^m a_i = ۱, \mathbf{y} = \sum_{i=۱}^m a_i \boldsymbol{X}_i\right\}.$$

۸ کمی ماتریس تصادفی

فرض کنید $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ یک ماتریس تصادفی با درایه‌های مستقل زیرگاوسی با پارامتر σ و میانگین صفر باشد. آنگاه برای هر $t > ۰$ داریم:

$$\mathbb{P}\left[\|\boldsymbol{A}\|_{\mathsf{op}} \geq C\sigma(\sqrt{p} + \sqrt{q} + t)\right] \leq e^{-t^۲},$$

برای یک ثابت مطلق C. بنابراین:

$$\mathbb{E}[\|\boldsymbol{A}\|_{\mathsf{op}}] \lesssim \sigma(\sqrt{p} + \sqrt{q}).$$

۹ $(*)$ آنتروپی کلاس توابع

کلاس توابع زیر را در نظر بگیرید:

$$\mathcal{F} := \{f : [۰,۱] \to [۰,۱] \mid \|f\|_{\mathrm{Lip}} \leq ۱\}.$$

نشان دهید برای هر $\epsilon > ۰$:

$$\mathcal{N}(\mathcal{F}, \|\cdot\|_\infty, \epsilon) \leq \left(\frac{۲}{\epsilon}\right)^{\frac{۲}{\epsilon}}.$$