به نام خدا

آنالیز احتمالاتی در ابعاد بالا (۲۵۸۴۱-۱)

تمرین سری پنجم

ترم پاییز ۱۴۰۳-۰۴

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

استاد: دکتر محمّدحسین یاسائی میبدی

مهلت تحویل: جمعه ۵ بهمن ساعت ۲۳:۵۹

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

۱ پارامتر ماتریس‌های زیرگوسی

ماتریس تصادفی $\mathbf{Q} = G\mathbf{B}$ را در نظر بگیرید که $G \in \mathbb{R}$ یک متغیر تصادفی زیرگوسی با پارامتر σ است.

۱. فرض کنید G یک توزیع متقارن حول صفر دارد و $\mathbf{B} \in \mathbb{S}^{d\times d}$ یک ماتریس غیرتصادفی است. نشان دهید $\mathbf{Q}$ یک ماتریس زیرگوسی با پارامتر $\mathbf{V} = \sigma^2 \mathbf{B}^2$ است.

۲. فرض کنید G یک توزیع متقارن حول صفر دارد و $\mathbf{B} \in \mathbb{S}^{d\times d}$ یک ماتریس تصادفی و مستقل از G است که $\|\mathbf{B}\|_2 \leq b$. نشان دهید $\mathbf{Q}$ یک ماتریس زیرگوسی با پارامتر $\mathbf{V} = \sigma^2 b^2 \mathbf{I}_d$ است.

۲ کران روی میانگین نرم

دنباله ماتریس‌های تصادفی متقارن d-بعدی با میانگین صفر $\{\mathbf{Q}_i\}_{i=1}^n$ را در نظر بگیرید که ماتریس $\mathbf{Q}_i$ زیرگوسی با پارامتر $\mathbf{V}_i$ است.

۱. ثابت کنید:

$$\mathbb{E}\left[\lambda_{\max}\left(\frac{1}{n}\sum_{i=1}^n \mathbf{Q}_i\right)\right] \leq \sqrt{\frac{2\left\|\frac{1}{n}\sum_{i=1}^n \mathbf{V}_i\right\|_{\mathsf{op}} \log d}{n}}.$$

۲. ثابت کنید:

$$\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^n \mathbf{Q}_i\right\|_{\mathsf{op}}\right] \leq \sqrt{\frac{2\left\|\frac{1}{n}\sum_{i=1}^n \mathbf{V}_i\right\|_{\mathsf{op}} \log(2d)}{n}}$$

۳ کران برنشتاین روی ماتریس‌ها با نرم بی‌کران

دنباله ماتریس‌های تصادفی $d_1 \times d_2$-بعدی $\{\mathbf{A}_i\}_{i=1}^n$ را درنظر بگیرید که داریم: $\mathbf{A}_i = G_i \mathbf{B}_i$ به طوری که $G_i \in \mathbb{R}$ یک متغیّر تصادفی با میانگین صفر است و $\mathbf{B}_i$ یک ماتریس تصادفی مستقل از G_iها است. فرض کنید برای $j = 2, 3, \ldots$ داریم:

$|\mathbb{E}[G_i^k]| \leq \frac{k!}{۲} b_۱^{k-۲} \sigma^۲$ و همچنین $\|\mathbf{B}_i\|_{\mathsf{op}} \leq b_۲$ است. نشان دهید برای هر $\delta > ۰$ داریم:

$$\mathbb{P}\left[\Big\|\frac{۱}{n}\sum_{i=۱}^{n}\mathbf{A}_i\Big\|_{\mathsf{op}} \geq \delta\right] \leq ۲(d_۱ + d_۲)\exp\left(\frac{-n\delta^۲}{۲(\sigma^۲ b_۲^۲ + b_۱ b_۲ \delta)}\right).$$

۴ تخمین ماتریس کواریانس قطری

دنباله بردارهای d-بعدی $\{\boldsymbol{X}_i\}_{i=۱}^{n}$ را درنظر بگیرید که به صورت مستقل و هم‌توزیع از یک توزیع با میانگین صفر و ماتریس کواریانس قطری $\mathbf{C}$ تولید شده‌اند. فرض کنید تخمینی که برای ماتریس کواریانس می‌سازیم به صورت زیر باشد:

$$\widehat{\mathbf{C}} = \operatorname{diag}\left(\frac{۱}{n}\sum_{i=۱}^{n}\boldsymbol{X}_i\boldsymbol{X}_i^{\top}\right).$$

۱. اگر هر کدام از بردارهای $\boldsymbol{X}_i$ زیرگوسی با پارامتر σ باشند، نشان دهید برای هر $\delta > ۰$ داریم:

$$\mathbb{P}\left[\frac{\|\widehat{\mathbf{C}} - \mathbf{C}\|_{\mathsf{op}}}{\sigma^۲} \geq c_۰\sqrt{\frac{\log d}{n}} + \delta\right] \leq c_۱ e^{-c_۲ \cdot n \cdot \min\{\delta, \delta^۲\}}.$$

۲. حال به جای فرض زیرگوسی بودن $\boldsymbol{X}_i$ها، فرض کنید برای یک $m \geq ۲$ که زوج است و برای هر $i = ۱, ۲, \ldots, n$ و $j = ۱, ۲, \ldots, d$ ثابت $c_۴$ وجود دارد که داریم:

$$\mathbb{E}\left[(X_{ij}^۲ - D_{jj})^m\right] \leq c_۴.$$

نشان دهید برای هر $\delta > ۰$ داریم:

$$\mathbb{P}\left[\|\widehat{\mathbf{C}} - \mathbf{C}\|_{\mathsf{op}} \geq ۴\delta\sqrt{\frac{d^{\frac{۲}{m}}}{n}}\right] \leq c_۵\left(\frac{۱}{۲\delta}\right)^m.$$

راهنمایی: برای متغیّرهای تصادفی اسکالر مستقل و با میانگین صفر Z_i که $\|Z_i\|_m < \infty$، ثابت k وجود دارد به طوری که:

$$\Big\|\sum_{i=۱}^{n} Z_i\Big\|_m \leq k\left\{\left(\sum_{i=۱}^{n}\mathbb{E}\left[|Z_i|^m\right]\right)^{۱/m} + \left(\sum_{i=۱}^{n}\mathbb{E}\left[Z_i^۲\right]\right)^{۱/۲}\right\}.$$

نیازی به اثبات قضیه‌ی گفته شده در راهنمایی نیست.

۵ نامساوی خینتچین ماتریسی

تغیّرهای تصادفی رادماخر مستقل $\varepsilon_۱, \ldots, \varepsilon_n$ را در نظر بگیرید. همچنین فرض کنید ماتریس‌های $\mathbf{A}_۱, \ldots, \mathbf{A}_n$ ماتریس‌های غیرتصادفی و متقارن با ابعاد $d \times d$ هستند. نشان دهید برای هر $p \in [۱, \infty)$ داریم:

$$\left(\mathbb{E}\left[\Big\|\sum_{i=۱}^{n}\varepsilon_i\mathbf{A}_i\Big\|_{\mathsf{op}}^{p}\right]\right)^{۱/p} \leq C\sqrt{p + \log d}\,\Big\|\sum_{i=۱}^{n}\mathbf{A}_i^۲\Big\|_{\mathsf{op}}^{۱/۲}.$$

۶ نامساوی تمرکز اندازه برای تمامی مقادیرویژه

در این سوال قصد داریم نامساوی تمرکز اندازه‌ای برای تمامی مقادیرویژه‌ها به دست آوریم.

۱. **قضیه‌ی کورانت-فیشر:** ماتریس متقارن $\mathbf{A}$ با مقادیر ویژه‌ی $\lambda_{max}(\mathbf{A}) = \lambda_۱(\mathbf{A}) \geq \ldots \geq \lambda_n(\mathbf{A}) = \lambda_{min}(\mathbf{A})$ را در نظر بگیرید. نشان دهید:

$$\lambda_k(\mathbf{A}) = \min_{\mathbf{V}\in\mathbb{V}^n_{n-k+۱}} \lambda_{\max}(\mathbf{V}^\top\mathbf{A}\mathbf{V})$$
$$\lambda_k(\mathbf{A}) = \max_{\mathbf{V}\in\mathbb{V}^n_k} \lambda_{\min}(\mathbf{V}^\top\mathbf{A}\mathbf{V}),$$

که در آن:

$$\mathbb{V}^n_d = \{\mathbf{V} \in \mathbb{R}^{n\times d} : \mathbf{V}^\top\mathbf{V} = \mathbf{I}\}.$$

۲. فرض کنید $\mathbf{X}$ یک ماتریس متقارن تصادفی با بعد d باشد. نشان دهید برای هر $k \leq d$ و هر $t \in \mathbb{R}$ داریم:

$$\mathbb{P}\left[\lambda_k(\mathbf{X}) \geq t\right] \leq \inf_{\theta>۰} \min_{\mathbf{V}\in\mathbb{V}^d_{d-k+۱}} \left\{ e^{-\theta t} \cdot \mathbb{E}\left[\mathsf{Tr}\left\{e^{\theta\mathbf{V}^\top\mathbf{X}\mathbf{V}}\right\}\right]\right\}.$$

۳. ماتریس‌های تصادفی مستقل و متقارن $\mathbf{X}_۱,\ldots,\mathbf{X}_n$ با ابعاد d را در نظر بگیرید. همچنین فرض کنید ماتریس‌های غیرتصادفی و متقارن $\mathbf{A}_۱,\ldots,\mathbf{A}_n$ با ابعاد d وجود دارند، به طوری که برای هر $j = ۱,\ldots,n$:

$$\mathbb{E}\left[e^{\mathbf{X}_j}\right] \preceq e^{\mathbf{A}_j}.$$

نشان دهید برای هر $\mathbf{V} \in \mathbb{V}^d_k$ که $k \leq d$ داریم:

$$\mathbb{E}\left[\mathsf{Tr}\left\{\exp\Big(\sum_{j=۱}^n \mathbf{V}^\top\mathbf{X}_j\mathbf{V}\Big)\right\}\right] \leq \mathsf{Tr}\left\{\exp\Big(\sum_{j=۱}^n \mathbf{V}^\top\mathbf{A}_j\mathbf{V}\Big)\right\}.$$

راهنمایی: از قضیه‌ی زیر استفاده کنید. نیاز به اثبات آن نیست.
اگر $\mathbf{H}$ یک ماتریس متقارن k بعدی و $\mathbf{V} \in \mathbb{V}^d_k$ که $k \leq d$ باشد. آنگاه نگاشت زیر برای ماتریس‌های مثبت معیّن یک نگاشت مقعّر است:

$$\mathbf{A} \mapsto \mathsf{Tr}\left\{\exp\Big(\mathbf{H} + \mathbf{V}^\top\log(\mathbf{A})\mathbf{V}\Big)\right\}.$$

۴. ماتریس‌های تصادفی مستقل و متقارن $\mathbf{X}_۱,\ldots,\mathbf{X}_n$ با ابعاد d را در نظر بگیرید. همچنین فرض کنید ماتریس‌های غیرتصادفی و متقارن $\mathbf{A}_۱,\ldots,\mathbf{A}_n$ با ابعاد d وجود دارد، به طوری که برای هر $\mathbf{V} \in \mathbb{V}^d_{d-k+۱}$ داریم:

$$\mathbb{E}\left[e^{\theta\mathbf{V}^\top\mathbf{X}_j\mathbf{V}}\right] \preceq e^{g(\theta)\mathbf{V}^\top\mathbf{A}_j\mathbf{V}},$$

که $g : (۰,\infty) \mapsto [۰,\infty)$. نشان دهید برای هر $t \in \mathbb{R}$ داریم:

$$\mathbb{P}\left[\lambda_k\Big(\sum_{j=۱}^n \mathbf{X}_j\Big) \geq t\right] \leq \inf_{\theta>۰} \min_{\mathbf{V}\in\mathbb{V}^d_{d-k+۱}} \left[e^{-\theta t}\cdot \mathsf{Tr}\left\{\exp\Big(g(\theta)\sum_{j=۱}^n \mathbf{V}^T\mathbf{A}_j\mathbf{V}\Big)\right\}\right].$$

۵. ماتریس‌های تصادفی مستقل و مثبت نیمه معین $\mathbf{X}_۱,\ldots,\mathbf{X}_n$ با ابعاد d را در نظر بگیرید. همچنین برای هر $k \leq d$ تعریف کنید:

$$\mu_k = \lambda_k\Big(\sum_{j=۱}^n \mathbb{E}\left[\mathbf{X}_j\right]\Big).$$

فرض کنید $\mathbf{V}_+ \in \mathbb{V}^d_{d-k+۱}$ و $\mathbf{V}_- \in \mathbb{V}^d_k$ در رابطه‌ی زیر صدق می‌کنند:

$$\mu_k = \lambda_{\max}\Big(\sum_{j=۱}^{n} \mathbf{V}_+^\top \mathbb{E}\left[\mathbf{X}_j\right] \mathbf{V}_+\Big) = \lambda_{\min}\Big(\sum_{j=۱}^{n} \mathbf{V}_-^\top \mathbb{E}\left[\mathbf{X}_j\right] \mathbf{V}_-\Big).$$

به ازای هر $\delta > ۰, \tilde{\delta} \in [۰, ۱)$ نشان دهید:

$$\mathbb{P}\left[\lambda_k\Big(\sum_{j=۱}^{n} \mathbf{X}_j\Big) \geq (۱+\delta)\mu_k\right] \leq (d-k+۱)\cdot\left[\frac{\mathrm{e}^{\delta}}{(۱+\delta)^{۱+\delta}}\right]^{\mu_k/\Psi(\mathbf{V}_+)},$$
$$\mathbb{P}\left[\lambda_k\Big(\sum_{j=۱}^{n} \mathbf{X}_j\Big) \leq (۱-\tilde{\delta})\mu_k\right] \leq k\cdot\left[\frac{\mathrm{e}^{-\tilde{\delta}}}{(۱-\tilde{\delta})^{۱-\tilde{\delta}}}\right]^{\mu_k/\Psi(\mathbf{V}_-)}.$$

که در آن تابع $\Psi : \bigcup_{۱\leq k\leq d} \mathbb{V}^d_k \to \mathbb{R}$ تابعی است که در شرایط زیر صدق می‌کند:

$$\max_j \lambda_{\max}\left(\mathbf{V}^T \mathbf{X}_j \mathbf{V}\right) \leq \Psi(\mathbf{V}) \qquad \text{almost surely for each } \mathbf{V} \in \bigcup_{۱\leq k\leq d} \mathbb{V}^d_k$$

راهنمایی: از لم زیر استفاده کنید و آن را اثبات کنید:
فرض کنید $\mathbf{X}$ یک ماتریس تصادفی مثبت نیمه معیّن است به طوری که $\lambda_{\max}(\mathbf{X}) \leq ۱$. آنگاه به ازای هر $\theta \in \mathbb{R}$ داریم:

$$\mathbb{E}\left[e^{\theta\mathbf{X}}\right] \preceq \exp\left((e^{\theta}-۱)\,\mathbb{E}\left[\mathbf{X}\right]\right).$$