HDP - HW5 	400101294 محمدرضا ... 	تمرین سری 5 آمار و احتمالات ابعاد بالا

سوال 1

$G = G^H$, $B \in S_n^{d\times d}$, $Q = GB$, $G \in \mathbb{R} \sim \text{Subg}(\sigma^2)$

(I) $\mathbb{E}_G[e^{tGB}] = \mathbb{E}_G\left[\sum_i e^{tG\lambda_i} u_i u_i^H\right] = \sum_i \mathbb{E}[e^{tG\lambda_i}] u_i u_i^H \le \exp\left(\frac{\sigma^2 B^2 t^2}{2}\right)$

Take out the certainty

$\le \sum_{i=1}^n e^{\frac{\sigma^2 \lambda_i^2 t^2}{2}} u_i u_i^H$ (*)

(*) Since $\sum_{i=1}^n u_i u_i^H = I = UU^H$.

$\to$ Thus, $Q \sim \text{Subg}(\sigma^2 B^2)$ (a.s.) Q.E.D

(II) $\mathbb{E}_{B,G}(\cdot) = \mathbb{E}_B \mathbb{E}_G(\exp(tGB)) \le \mathbb{E}_B\left[\exp\left(\frac{\sigma^2 B^2 t^2}{2}\right)\right] = \mathbb{E}_B\left(\sum_i e^{\frac{\sigma^2 \lambda_i^2 t^2}{2}} u_i u_i^H\right)$

$\le \mathbb{E}_B\left[\sum_i e^{\frac{\sigma^2 t^2 \lambda^2_{max}(B)}{2}} u_i u_i^H\right] \le \mathbb{E}_B\left[\sum_i e^{\frac{\sigma^2 t^2 b^2}{2}} u_i u_i^H\right] = \mathbb{E}_B\left[e^{\frac{\sigma^2 t^2 b^2}{2}} I\right]$

$\lambda_i \le \lambda_{max} \le b$

$\Rightarrow \mathbb{E}_{B,G}[\exp(tGB)] \le \exp\left(\frac{\sigma^2 t^2 b^2}{2}\right) I = \exp\left(\frac{t^2\sigma^2 b^2}{2} I\right) \sim \sigma^2 b^2 I$ subg

Q.E.D

سوال 2

(I) $\mathbb{E}\left[\lambda_{max}\left(\frac{1}{n}\sum_{i=1}^n Q_i\right)\right] \le \mathbb{E}\left[\frac{1}{\beta}\log\left(\sum_i e^{\beta\lambda_i}\right)\right] = \frac{1}{\beta}\mathbb{E}\left[\log\left(tr\{\exp(\beta S_n)\}\right)\right]$

$\le \frac{1}{\beta}\log \mathbb{E}_{Q_1\dots Q_n}\left[tr\left\{\exp\left(\frac{\beta}{n}\sum_{i=1}^n Q_i\right)\right\}\right] = \frac{1}{\beta}\log\left\{\mathbb{E}_{Q_{1:n-1}}\mathbb{E}_{Q_n}\left\{tr\left(\exp\left(\frac{\beta}{n}\sum_{i=1}^{n-1}Q_i + \frac{\beta}{n}Q_n\right)\right)\right\}\right\}$

Jensen

$Tr\{\exp(H + \log(X))\}$ is concave

$\le \frac{1}{\beta}\log\left(\mathbb{E}_{Q_{1:n-1}} Tr\left\{\exp\left(\frac{\beta}{n}\sum_{i=1}^{n-1}Q_i + \log\left(\mathbb{E}_{Q_n}\left[\exp\left(\frac{\beta}{n}Q_n\right)\right]\right)\right)\right\}\right)$

applying this n-2 times

$\le \frac{1}{\beta}\log\left(tr\left\{\exp\left(\sum_{i=1}^n \log \mathbb{E}_{Q_i}\left[\exp\left(\frac{\beta}{n}Q_i\right)\right]\right)\right\}\right)$

$A \le B$, $Tr\{e^A\} \le Tr\{e^B\}$, $\log(A) \le \log(B)$

$\le \frac{1}{\beta}\log\left(Tr\left\{\exp\left(\sum_{i=1}^n \log\left(\exp\left\{\frac{1}{2}\frac{\beta^2}{n^2}V_i\right\}\right)\right)\right\}\right) = \frac{1}{\beta}\log\left(tr\left\{\exp\left(\frac{\beta^2}{2n}\sum_{i=1}^n \frac{V_i}{n}\right)\right\}\right)$

$= \frac{1}{\beta}\log\left(\sum_{j=1}^n e^{\frac{\beta^2}{2n}\lambda_j\left(\frac{1}{n}\sum_{i=1}^n V_i\right)}\right) \le \frac{1}{\beta}\log\left(\sum_{j=1}^n e^{\frac{\beta^2}{2n}\left\|\frac{1}{n}\sum_{i=1}^n V_i\right\|_{op}}\right)$

$= \frac{\log(d)}{\beta} + \frac{\beta}{2n}\left\|\frac{1}{n}\sum_{i=1}^n V_i\right\|_{op} \Rightarrow \mathbb{E}\left[\lambda_{max}\left(\frac{1}{n}\sum_{i=1}^n Q_i\right)\right] \le \inf_\beta \underbrace{\left\{\frac{\log d}{\beta} + \frac{\beta}{2n}\|\dots\|_{op}\right\}}_{f(\beta)}$

$\frac{\partial f(\beta)}{\partial\beta} = 0 \to \frac{+\log d}{\beta^2} = \frac{1}{2n}\left\|\frac{1}{n}\sum_{i=1}^n V_i\right\|_{op} \to \beta^* = \sqrt{\frac{2n\log d}{\frac{1}{n}\|\Sigma V_i\|_{op}}}$ $\longrightarrow$

$$\Rightarrow \mathbb{E}\left[\lambda_{max}^{\left(\frac{1}{n}\sum_{i=1}^{n}\Phi_i\right)}\right] \leq f(\beta^*) = \sqrt{\frac{2\left\|\frac{1}{n}\sum_{i=1}^{n}V_i\right\|_{op}\cdot \log d}{n}}$$ Q.E.D.

(II) consider the matrix $A_i = \begin{bmatrix} 0_{d\times d} & \Phi_i^T \\ \Phi_i & 0_{d\times d}\end{bmatrix} = A_i^T$, $Z_i = \begin{pmatrix} V_i & 0 \\ 0 & V_i\end{pmatrix}$

$$\rightarrow \lambda_{max}^{(A_i)} = \sup \frac{u^T A_i u}{\|u\|} \underset{u=(x\ y)^T}{=} \sup_{x,y} \frac{(x\ y)\begin{bmatrix}0 & \Phi_i \\ \Phi_i & 0\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}}{\|x\|_2^2+\|y\|_2^2} = \frac{2x^T\Phi_i y}{\|x\|_2^2+\|y\|_2^2}$$

$$= \sup_y \sup_x \{-\} = \sup_{y\in\mathbb{R}^d} \frac{\|\Phi_i y\|}{\|y\|} = \|\Phi_i\|_{op} \xrightarrow{\text{By Induction}} \lambda_{max}\left(\frac{1}{n}\sum_{i=1}^{n}A_i\right) = \left\|\frac{1}{n}\sum_{i=1}^{n}\Phi_i\right\|_{op}$$

- since $A_i^2 = \begin{bmatrix}\Phi_i^2 & 0 \\ 0 & \Phi_i^2\end{bmatrix} \rightarrow A_i \sim Z_i$ subG

$A_i = \begin{bmatrix}1 & 0 \\ 0 & 0\end{bmatrix}\otimes\Phi_i + \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\otimes\Phi_i \rightarrow$ disjoint spaces!

$$\Rightarrow \mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^{n}\Phi_i\right\|_{op}\right] = \mathbb{E}\left[\lambda_{max}\left(\frac{1}{n}\sum_{i=1}^{n}A_i\right)\right] \leq \sqrt{\frac{2\left\|\frac{1}{n}\sum_{i=1}^{n}Z_i\right\|_{op}\cdot\log(2d)}{n}}$$

Since $Z_i = I_{2\times 2}\otimes V_i \Rightarrow \frac{1}{n}\sum Z_i = I_{2\times 2}\otimes\left(\frac{1}{n}\sum_{i=1}^{n}V_i\right) \Rightarrow \left\|\frac{1}{n}\sum_{i=1}^{n}Z_i\right\|_{op} = \left\|\frac{1}{n}\sum_{i=1}^{n}V_i\right\|_{op}$

which finally suggests
$$\boxed{\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^{n}\Phi_i\right\|_{op}\right] \leq \sqrt{\frac{2\left\|\frac{1}{n}\sum_{i=1}^{n}V_i\right\|_{op}\cdot\log 2d}{n}}}$$ Q.E.D.

$A_i \in\mathbb{R}^{d_1\times d_2}$, $A_i = G_i B_i$, $\begin{cases} G_i\in\mathbb{R} \\ \mathbb{E}G_i = 0 \\ B_i \perp\!\!\!\perp G_i\end{cases}$ $\forall k\in\mathbb{N}-\{1\}: \mathbb{E}[G_i^k] \leq \frac{k!}{2} b_1^{k-2}\sigma^2$, $\|B_i\|_{op}\leq b_2$ $3\,\mathrm{O}'_2$

let $S_n = \frac{1}{n}\sum_{i=1}^{n}A_i \in\mathbb{R}^{d_1\times d_2}$, $T_n = \begin{pmatrix} 0_{d_1\times d_2} & S_n^T \\ S_n & 0_{d_1\times d_1}\end{pmatrix}$

In the previous parts we have seen that $\|S_n\|_{op} = \lambda_{max}^{(T_n)}$. let $C_i = \begin{pmatrix}0 & A_i^T \\ A_i & 0\end{pmatrix}$

$\Rightarrow T_n = \frac{1}{n}\sum_{i=1}^{n}C_i \xrightarrow{\text{thus}} \mathbb{P}\left[\|S_n\|_{op}\geq\delta\right] = \mathbb{P}\left[\lambda_{max}^{(T_n)}\geq\delta\right] = \mathbb{P}\left[\lambda_{max}\left(\frac{1}{n}\sum_{i=1}^{n}C_i\right)\geq\delta\right]$

$$\Rightarrow \mathbb{P}\left(\lambda_{max}^{(T_n)}\geq\delta\right) \underset{s,\delta>0}{=} \mathbb{P}\left(e^{s\lambda_{max}^{(T_n)}}\geq e^{s\delta}\right) \leq \frac{\mathbb{E}\left[e^{s\lambda_{max}^{(T_n)}}\right]}{e^{\delta s}} =$$

ادامه در صفحه بعد

ادامه تمرین ۳:

$$\Rightarrow \mathbb{P}\left[\left\|\frac{1}{n}\sum_{i=1}^{n} A_i\right\|_{op} \geq \delta\right] \leq e^{-s\delta}\cdot \operatorname{tr}\left\{\exp\left(\sum_{i=1}^{n}\log \mathbb{E}\left[e^{\frac{s A_i}{n}}\right]\right)\right\}$$

plug $\beta = \frac{1}{n}$ in last inequalities

$$\leq e^{-s\delta}\cdot \operatorname{tr}\left\{\exp\left(\frac{\sigma^2 b_2^2 s^2}{2n^2(1-b_1 b_2 s/n)}\right)\right\} = e^{-s\delta}\cdot \operatorname{tr}\left\{I\cdot \exp\left(\frac{s\sigma^2 b_2^2}{2(n^2-b_1 b_2 s n)}\right)\right\}$$

$$= (d_1+d_2)\cdot \exp\left(\frac{s^2\sigma^2 b_2^2}{2(n^2-b_1 b_2 s n)} - s\delta\right)$$

optimizing over s will lead to $s^* = \frac{n\delta}{\sigma^2 b_2^2 + b_1 b_2 \delta}$

$$\mathbb{P}\left[\left\|\frac{1}{n}\sum_{i=1}^{n} A_i\right\|_{op} \geq \delta\right] \leq \frac{s^2\sigma^2 b_2^2}{2n(n-b_1 b_2\delta)} - s\delta = \frac{-1}{2}\frac{n\delta^2}{2(\sigma^2 b_2^2 + b_1 b_2\delta)}$$

$$\longrightarrow \mathbb{P}\left[\left\|\frac{1}{n}\sum_{i=1}^{n} A_i\right\|_{op} \geq \delta\right] \leq (d_1+d_2)\cdot \exp\left(\frac{-n\delta^2}{4(\sigma^2 b_2^2 + b_1 b_2 \delta)}\right)$$

$2 \geq e^{-1/2}$

$$\leq 2(d_1+d_2)\cdot \exp\left(\frac{-n\delta^2}{2(\sigma^2 b_2^2 + b_1 b_2\delta)}\right)$$

Q.E.D. ∎

(I) $\{x_i\}_{i=1}^{n}$, $\mathbb{E}\{x_i\}=0$, $x_i \overset{iid}{\sim} P_\alpha$, $\mathbb{E}[x_i x_i^H] = C$, $\hat{C} = \operatorname{diag}\left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^H\right)$ تمرین ۴:

Both $C, \hat{C}$ are diagonal $\Rightarrow \hat{C}_i = \frac{1}{n}\sum_{j=1}^{n}(x_j)_{(i)}^2 = \frac{1}{n}\|x_i^T\|_2^2$

as we saw $\frac{\|x_i^T\|_2^2}{\sigma^2 nc} \sim \operatorname{subexp}\left(\frac{4}{n}, \frac{4}{n}\right)$.

$$\frac{\|x_i^T\|_2^2}{nc} \sim \operatorname{subexp}\left(\frac{4\sigma^2}{n}, \frac{4\sigma^2}{n}\right) \longrightarrow \mathbb{E}\left[\frac{\|x_i^T\|_2^2}{n}\right] = \frac{1}{n}\mathbb{E}\left[\sum_{j=1}^{n}(x_j^T)_{(i)}^2\right] = C_i$$

$$\mathbb{P}\left[\frac{|\hat{C}_i - C_i|}{c} \geq t\right] \leq 2\exp\left(\frac{-n}{4\sigma^2}\left(\frac{t^2}{\sigma^2}\wedge t^2\right)\right).$$

let $s = \frac{t}{c\sigma^2} \Rightarrow \mathbb{P}\left[\frac{|\hat{C}_i - C_i|}{\sigma^2} \geq s\right] \leq 2e^{-\frac{n}{4}(\delta\wedge\delta^2)c^2}$

$$\leq 2\exp\left(\frac{-n}{4}\min\left\{s - \frac{4\log(d)}{nc}, s^2 - \frac{4\log(d)}{nc}\right\}\right) \overset{\text{some algebra}}{\leq} 2\exp\left(-\frac{n^2}{4}(\delta\wedge\delta^2)\right)$$

Q.E.D. ∎

$$\mathbb{P}\left[\lambda_{max}^{(T_n)} \geq \delta\right] \leq e^{-s\delta}\mathbb{E}\left[e^{s\lambda_{max}^{(T_n)}}\right] \leq e^{-s\delta}\mathbb{E}\left[\sum_{i=1}^{n} e^{s\lambda_i(T_n)}\right] = e^{-s\delta}\mathbb{E}\left[\mathrm{tr}\{e^{sT_n}\}\right]$$

$$= e^{-s\delta}\mathrm{tr}\left[\mathbb{E}\left[e^{sT_n}\right]\right] = e^{-s\delta}\,\mathrm{tr}\left\{\mathbb{E}\left[\exp\left(\frac{s}{n}\sum_{i=1}^{n} C_i\right)\right]\right\} = \cancel{e^{-s\delta}\,\mathrm{tr}\{\mathbb{E}}$$

$\mathrm{tr}\{\mathbb{E}\} = \mathbb{E}\{\mathrm{tr}\}$

$\mathrm{tr}\{e^{H + \log X}\}$ is concave

$$\Rightarrow \mathbb{P}\left[\lambda_{max}^{(T_n)} \geq \delta\right] \leq e^{-s\delta}\,\mathrm{tr}\left\{\mathbb{E}\left[\exp\left(\sum_{i=1}^{n}\log\left(e^{\frac{sC_i}{n}}\right)\right)\right]\right\} \leq e^{-s\delta}\,\mathrm{tr}\left\{\exp\left(\sum_{i=1}^{n}\log \mathbb{E}\left[e^{\frac{sC_i}{n}}\right]\right)\right\}$$

(*) Now we wish to make a bound for $\mathbb{E}\left[e^{\frac{sC_i}{n}}\right]$:

$$\mathbb{E}\left[e^{\beta C_i}\right] = \mathbb{E}\left[\exp\left(\begin{bmatrix} 0 & \beta A_i^T \\ \beta A_i & 0\end{bmatrix}\right)\right] = \mathbb{E}_{B_i}\mathbb{E}_{G_i}\left[\exp\left(\beta G_i \begin{bmatrix} 0 & B_i^T \\ B_i & 0\end{bmatrix}\right)\right]$$

Taylor series

$$= \mathbb{E}_{B_i}\mathbb{E}_{G_i}\left[\sum_{k=0}^{\infty}\frac{1}{k!}\beta^k G^k D_i^k\right] = I + \mathbb{E}_{B_i}\mathbb{E}_{G_i}\left[\sum_{k=2}^{\infty}\frac{1}{k!}\beta^k G_i^k D_i^k\right]$$

$$= I + \mathbb{E}_{B_i}\mathbb{E}_{G_i}\left[\sum_{k=2}^{\infty}\sum_{\ell=1}^{d_1+d_2}\frac{1}{k!}\beta^k G_i^k \lambda_{\ell,i}^k u_{\ell,i}u_{\ell,i}^H\right] = I + \mathbb{E}_{B_i}\mathbb{E}_{G_i}\left[\dots\right]$$

$$\longrightarrow = I + \mathbb{E}_{B_i}\mathbb{E}_{G_i}\left[\sum_{\ell=1}^{d_1+d_2} u_{\ell,i}u_{\ell,i}^H \sum_{k=2}^{\infty}\frac{1}{k!}\beta^k G_i^k \lambda_{\ell,i}^k\right]$$

$$\preceq I + \mathbb{E}_{B_i}\mathbb{E}_{G_i}\left[\sum_{\ell=1}^{d_1+d_2} u_{\ell,i}u_{\ell,i}^H \sum_{k=2}^{\infty}\frac{1}{k!}\beta^k G_i^k \|D_i\|_{op}\right]$$

$\longrightarrow$ all are concave matrices

$$= I\left(1 + \mathbb{E}_{B_i}\mathbb{E}_{G_i}\left[\sum_{k=2}^{\infty}\frac{1}{k!}\beta^k G_i^k \|D_i\|_{op}^k\right]\right)$$

$$\Rightarrow \mathbb{E}\left[e^{\beta C_i}\right] \preceq I\left(1 + \mathbb{E}_{G_i}\left[\sum_{k=2}^{\infty}\frac{1}{k!}\beta^k G_i^k b_2^k\right]\right)$$

since $\|D_i\|_{op} \leq \|B_i\|_{op} \leq b_2$

$$= I\left(1 + \frac{\beta^2\sigma^2 b_2^2}{2}\sum_{k=0}^{\infty}(\beta b_1 b_2)^k\right) \preceq I\left(1 + \sum_{k=2}^{\infty}\frac{\beta^k}{k!}\frac{k!\, b_1^{k-2}\sigma^2 b_2^k}{2}\right) = I\left(1 + \frac{\beta^2\sigma^2 b_2^2}{2(1 - b_1 b_2\beta)}\right)$$

$$\Rightarrow \log\left\{\mathbb{E}\left[e^{\beta C_i}\right]\right\} \preceq \log\left(I\left(1 + \frac{\beta^2\sigma^2 b_2^2}{2(1-b_1b_2\beta)}\right)\right) \preceq \frac{\sigma^2 b_2^2\beta^2}{2(1-b_1b_2\beta)}I$$

$\log(I(1+x)) \preceq \log(I\cdot x)$

which suggests

$$\Rightarrow \mathbb{P}\left[\left\|\frac{1}{n}\sum_{i=1}^{n}A_i\right\|_{op} \geq \delta\right] \leq e^{-s\delta}\,\mathrm{tr}\left\{\exp\left(\sum_{i=1}^{n}\log\mathbb{E}\, e^{\frac{sC_i}{n}}\right)\right\}$$

بالاسم...، ٩٥

(IV) $\|\hat{C}-C\|_{op} = \max_i \{|\hat{C}_i - C_i|\} \geq \|\hat{C}-C\|_m$

$\Rightarrow \mathbb{P}\left[\|\hat{C}-C\|_{op} \geq t\right] = \mathbb{P}\left[\max_i \|\hat{C}_i - C\|^* \geq t\right] = \mathbb{P}\left[\max_i \|\hat{C}_i - C\|^m \geq t^m\right]$

$\leq \frac{1}{t^m} \mathbb{E}\left[\max_i |\hat{C}_i - C_i|^m\right] \leq t^{-m} \sum_{i=1}^{d} \mathbb{E}\left[|\hat{C}_i - C_i|^m\right] = t^{-m} \sum_{i=1}^{d} \|\hat{C}_i - C_i\|_m^m$

$\longrightarrow \hat{C}_i = \frac{1}{n}\sum_{j=1}^{n} (X_j)_{(i)}^2 = \frac{1}{n}\sum_{j=1}^{n}|X_{ji}|^2 \longrightarrow \hat{C}_i - C_i = \frac{1}{n}\sum_{j=1}^{n} \underbrace{(|X_{ji}|^2 - D_{ii})}_{Z_{ij}} = \sum_{j=1}^{D} Z_{ij}$

$\longrightarrow$ clearly $\mathbb{E} Z_{ij} = 0 \longrightarrow \|Z_{ij}\|_m^m = \mathbb{E}\left[\frac{1}{n^m}(|X_{ji}|^2 - D_{ii})^m\right] \leq \frac{\bar{C}}{n^m} < \infty$

we can see that the theorem still holds for Z_{ij}

$\|\hat{C}_i - C_i\|_m = \|\sum_{j=1}^{dn} Z_{ij}\|_m \leq k\left\{\left(\sum_{j=1}^{n} \mathbb{E}\left[|Z_{ij}|^m\right]\right)^{1/m} + \left(\sum_{j=1}^{m} |Z_{ij}|^2\right)^{1/2m}\right\}$

$\Rightarrow \mathbb{E}\left[|Z_{ij}|^m\right] = n^{-m}\mathbb{E}\left[(X_{ji} - D_{ii})^m\right] \leq n^{-m}\cdot\bar{C} \longrightarrow \left(\sum_{j=1}^{n} \mathbb{E}|Z_{ij}|^m\right)^{1/m} \leq (n\bar{C}\cdot n^{-m})^{1/m} \leq \frac{(n\bar{C})^{1/m}}{n}$

$\Rightarrow \mathbb{E}\left[|Z_{ij}|^2\right] = \mathbb{E}\left[(|Z_{ij}|^m)^{\frac{2}{m}}\right] \underset{\text{Jensen}}{\leq} \left(\mathbb{E}|Z_{ij}|^m\right)^{2/m} \leq (\bar{C}\cdot n^{-m})^{2/m} = \frac{\bar{C}^{2/m}}{n^2}$

thus $\|\hat{C}_i - C_i\|_m \leq k\left\{\frac{1}{n^{\frac{m-1}{m}}}\bar{C}^{1/m} + \frac{1}{\sqrt{n}}\bar{C}^{1/m}\right\} \leq \frac{2k}{n^{1-\frac{1}{m}}}\bar{C}^{1/m} = \frac{2k}{n}(n\bar{C})^{1/m}$

$\Rightarrow \mathbb{P}\left[\|\hat{C}-C\|_{op} \geq t\right] \leq t^{-m}\cdot d\left(\frac{2k}{n}(n\bar{C})^{1/m}\right)^m = d\cdot t^{-m}\cdot n\bar{C}\left(\frac{2k}{n}\right)^m$

$t = 4\delta\sqrt{\frac{d^{4/m}}{n}} \Rightarrow \mathbb{P}\left[\|\hat{C}-C\|_{op} \geq 4\delta\sqrt{\frac{d^{2/m}}{n}}\right] \leq \frac{\bar{C}k^m \cdot n^{1-\frac{m}{2}}}{2^m\delta^m} \leq \frac{\bar{C}\cdot k^m}{(2\delta)^m} = \bar{C}(2\delta)^{-m}$

$\Rightarrow \boxed{\mathbb{P}\left[\|\hat{C}-C\|_{op} \geq 4\delta\sqrt{\frac{d^{2/m}}{n}}\right] \leq \frac{C\cdot k^m}{(2\delta)^m}}$ Q.E.D. ■

$$\left\| \sum_{i=1}^n \varepsilon_i A_i \right\|_p \le C\sqrt{p+\log 2d}\ \left\| \sum_{i=1}^n A_i^2 \right\|_{op}^{1/2}$$

سوال ۵:

$$\begin{cases} S = \sum_{i=1}^n (x^T A_i y)^2 \\ Z_{xy} = \sum_{i=1}^n \varepsilon_i x^T A_i y \sim S \text{ subg} \end{cases}$$

$$\Rightarrow \left\| \sum_{i=1}^n \varepsilon_i A_i \right\|_{op} = \sup_{x,y\in S^d} \left\{ x^T\left(\sum_{i=1}^n \varepsilon_i A_i\right) y \right\} = \sup_{x,y\in S^d} \left\{ \underbrace{\sum_{i=1}^n \varepsilon_i x^T A_i y}_{Z_{xy}} \right\}$$

$$\Rightarrow \sum_{i=1}^n |x^T A_i y|^2 = \sum_{i=1}^n x^T A_i y y^T A_i x = x^T\left(\sum_{i=1}^n A_i y y^T A_i\right) x \le \|x\|_2^2 \cdot \lambda_{max}\left(\sum_{i=1}^n A_i y y^T A_i\right)$$

$$\le \sum_{i=1}^n \|A_i y\|_2^2 \cdot \|x\|_2^2 = \|x\|_2^2 \cdot \sum_{i=1}^n y^T A_i y = \|x\|_2^2 \cdot y^T\left(\sum A_i\right) y$$

$$\le \|x\|_2^2 \cdot \|y\|_2^2 \cdot \left\| \sum_{i=1}^n A_i^2 \right\|_{op} \xrightarrow{\text{Thus}} Z_{xy} \sim \left\| \sum_{i=1}^n A_i^2 \right\|_{op} \text{ subgaussian}$$

now we'll use the hoeffding tight version bound:

$$\mathbb{P}\left[\left\| \sum_{i=1}^n \varepsilon_i A_i \right\|_{op} \ge t \right] \le 2d \cdot \exp\left(-\frac{t^2}{2\sigma^2}\right), \quad \sigma^2 = \left\| \sum_{i=1}^n A_i^2 \right\|_{op}$$

$$\le \exp\left(\log 2d - \frac{t^2}{2\sigma^2}\right)$$

from Vershynin

$$\mathbb{E}\left[\|X\|^p\right] = p\int_0^\infty t^{p-1}\, \mathbb{P}(|X| \ge t)\, dt \xrightarrow{\text{let}} t = \sqrt{2\sigma^2 \log 2d + u^2} \le u + \sqrt{2\sigma^2 \log 2d}$$

$$\Rightarrow \mathbb{P}\left[\left\| \sum_{i=1}^n \varepsilon_i A_i \right\|_{op} \ge t \right] \le e^{-\frac{u^2}{2\sigma^2}}$$

$$\Rightarrow \mathbb{E}\left[\|X\|^p\right] = p\int_0^\infty t^{p-1} \cdot \mathbb{P}\left(\left\| \sum_{i=1}^n \varepsilon_i A_i \right\|_{op} \ge t \right) dt = p\int_0^{\sqrt{2\sigma^2 \log 2d}} t^{p-1}\, \mathbb{P}(\cdots \ge t)\, dt + p\int_{\sqrt{2\sigma^2\log 2d}}^\infty t^{p-1}\, \mathbb{P}(\cdots \ge t)\, dt$$

$$\le p\int_0^{\sqrt{2\sigma^2\log 2d}} t^{p-1}\, dp + p\int_{\sqrt{2\sigma^2\log 2d}}^\infty \left(u + \sqrt{2\sigma^2 \log 2d}\right)^{p-1} e^{-\frac{u^2}{2\sigma^2}}\, du$$

$$\le \left(2\sigma^2 \log 2d\right)^{p/2} + p\cdot\left(2\sqrt{2\sigma^2 \log 2d}\right)^p + 2^{p-1}\cdot p\int_0^\infty u^{p-1} e^{-\frac{u^2}{2\sigma^2}}\, du$$

$$\le \left(1 + 2^p \cdot p\right)\left(2\sigma^2 \log 2d\right)^{p/2} + \left(\frac{8\sigma^2 p}{2e}\right)^{p/2} \le 2^{p+1}\cdot p\left(2\sigma^2 \log 2d\right)^{p/2} + \left(2\sigma^2 p\right)^{p/2}$$

ادامه سوال 5:

$$\Rightarrow \mathbb{E}\left[\left|\sum_{i=1}^{n}\varepsilon_i A_i\right|^p\right]^{1/2} \le (\ldots)^{1/p} \le 4p^{1/p}\cdot(2\sigma^2\log 2d)^{1/2} + (2\sigma^2 p)^{1/2}$$

$$\le C\cdot\sigma\cdot p^{1/p}\sqrt{\log 2d + p} \le C\cdot\sigma\cdot e^{1/e}\sqrt{\log 2d + p}$$

let $\sigma^2 = \|\sum_{i=1}^{n} A_i^2\|_{op}$ $\qquad \forall p:\ p^{1/p} \le e^{1/e}$

با تعریف σ

$$\Rightarrow \boxed{\mathbb{E}\left[\left|\sum_{i=1}^{n}\varepsilon_i A_i\right|^p\right]^{1/p} \le C\cdot\sqrt{p+\log 2d}\cdot\left\|\sum_{i=1}^{n}A_i^2\right\|_{op}^{1/2}}$$ Q.E.D.

سوال 6:

(I) $A = U\Lambda U^H$, $C(V) \to$ column space of V

$M = \begin{bmatrix} I \\ 0 \end{bmatrix}$

$$\Rightarrow V = \tilde{V}_{1:d} = \tilde{V}M = \tilde{V}\begin{bmatrix} I \\ 0 \end{bmatrix} \to V^TAV = M^T\underbrace{\tilde{V}^TU}_{X}\Lambda U^H\tilde{V}M = M^TX\Lambda X^HM$$

$X^TX = U^H\underbrace{\tilde{V}\tilde{V}^T}_{I}U = I \to X\Lambda X^{-1}, A$ have the same spectrum (basis

$$\Rightarrow \lambda_{max}(V^TAV) = \max_{x\in\mathbb{R}^d}\left\{\frac{x V^TAVx}{x^Tx}\right\} = \max_{x\in\mathbb{R}^d}\left\{\frac{(Vx)^TA(Vx)}{x^TV^TVx}\right\} = \max_{x\in C(V)}\left\{\frac{x^TAx}{x^Tx}\right\}$$

if we do this again we can say that

$$\lambda_k = \max_{\substack{S\subseteq\mathbb{R}^n\\ \dim(S)=k}}\left\{\min_{\substack{x\in S\\ x\neq 0}}\left\{\frac{x^TAx}{x^Tx}\right\}\right\} = \min_{\substack{T\subseteq\mathbb{R}^n\\ \dim(T)=n-k+1}}\left\{\max_{\substack{x\in T\\ x\neq 0}}\left\{\frac{x^TAx}{x^Tx}\right\}\right\}$$

Q.E.D.

(II) $\mathbb{P}\left[\lambda_k(X) \geq t\right] = \mathbb{P}\left[\min_{V \in V^d_{d-k+1}} \lambda_{max}(V^T X V) \geq t\right] \leq e^{-\theta t}\, \mathbb{E}\left[\exp\left(\theta \cdot \min_{V \in V} \lambda_{max}(V^T X V)\right)\right]$

$$= \leq e^{-\theta t}\, \mathbb{E}\left[\min_{V \in V^d_{d-k+1}} \exp(\theta \lambda_{max}(V^T X V))\right] \leq e^{-\theta t} \min_{V \in V^d_{d-k+1}} \left\{\mathbb{E} \exp(\theta \lambda_{max}(V^T X V))\right\}$$

Concavity

$$\leq e^{-\theta t} \min_{V} \left\{\mathbb{E}\left[\max\, \exp(\theta \lambda_i(V^T X V))\right]\right\} \leq e^{-\theta t} \min_{V} \left\{\mathbb{E}\left[\sum_{i=1}^{n} \theta \lambda_i(V^T X V)\right]\right\}$$

$$\leq e^{-\theta t} \min_{V} \mathbb{E}\left[\mathrm{tr}\left\{\exp(\theta V^T X V)\right\}\right]$$

Q.E.D.

$$\Rightarrow \boxed{\mathbb{P}\left[\lambda_k(X) \geq t\right] \leq \inf_{\theta > 0} \left\{e^{-\theta t} \cdot \min_{V \in V^d_{d-k+1}} \mathbb{E}\left[\mathrm{tr}\left\{\exp(\theta V^T X V)\right\}\right]\right\}}$$

(III) $\mathbb{E}\left[\mathrm{tr}\left\{\exp\left(\sum_{j=1}^{n} V^T X_j V\right)\right\}\right] = \mathbb{E}_{X_{2:n}}\, \mathbb{E}_{X_1}\left[\mathrm{tr}\left\{\exp(\cdots)\right\}\right]$ → Jensen

$$= \leq \mathbb{E}_{X_{2:n}}\left[\mathrm{tr}\left\{\exp\left(\sum_{j=2}^{n} V^T X_j V + V^T \log\left(\mathbb{E}[e^{X_1}]\right) V\right)\right\}\right]$$

الموجة 6b.

do this n-2 times more

$$\leq \mathrm{tr}\left\{\exp\left(\sum_{j=1}^{n} V^T \cdot \log(\mathbb{E}[e^{X_i}]) \cdot V\right\} \leq \mathrm{tr}\left\{\exp\left(\sum_{j=1}^{n} V^T A_j V\right)\right.$$

which finally suggests that:

$$\boxed{\mathbb{E}\left[\mathrm{tr}\left\{\exp\left(\sum_{j=1}^{n} V^T X_j V\right)\right\}\right] \leq \mathrm{tr}\left\{\exp\left(\sum_{j=1}^{n} V^T A_j V\right)\right\}}$$

Q.E.D.

(IV) $\mathbb{P}\left[\lambda_k\left(\sum_{j=1}^{n} X_j\right) \geq t\right] \leq \inf_{\theta > 0} \left\{e^{-\theta t} \cdot \min_{V \in V^n_{d-k+1}} \cdot \mathbb{E}\left[\mathrm{tr}\left\{\exp\left(\theta V^T \sum_{j=1}^{n} X_j \theta\right)\right\}\right]\right.$

similar to part 3, we can get:

$$\mathbb{E}\left[\mathrm{tr}\left\{\exp\left(\theta \sum_{j=1}^{n} V^T X_j V\right)\right\}\right] \leq \mathrm{tr}\left\{\exp\left(g(\theta) \sum_{j=1}^{n} V^T A_j V\right)\right\}$$

thus we obtain that:

$$\mathbb{P}\left[\lambda_k\left(\sum_{j=1}^{n} X_j\right) \geq t\right] \leq \inf_{\theta > 0} \left\{e^{-\theta t} \cdot \min_{V \in V^d_{d-k+1}} \right\} \mathrm{tr}\left\{\exp\left(g(\theta) \sum_{j=1}^{n} V^T A_j V\right)\right\}$$

$$= \inf_{\theta > 0} \left\{\min_{V \in V^d_{d-k+1}} \left\{e^{-\theta t} \cdot \mathrm{tr}\left\{\exp\left(g(\theta) \sum_{j=1}^{n} V^T A_j V\right)\right\}\right\}\right.$$

Q.E.D.

Scanned by CamScanner

(V) Since e^x is convex, $\forall \lambda \in [0,1]: e^{\lambda\theta} = \exp(\lambda\theta + \bar{\lambda}(0))$

$\leq \bar{\lambda}(1) + \lambda e^{\theta} = 1 + \lambda(e^{\theta}-1)$

$\longrightarrow e^{\theta X} = \sum_{i=1}^{n} e^{\theta\lambda_i} u_i u_i^H \leq \sum_{i=1}^{n} \left(1+(e^{\theta}-1)\lambda_i\right) u_i u_i^H\Big) = \underbrace{\sum_{i=1}^{n} u_i u_i^H}_{I} + (e^{\theta}-1)\sum_{i=1}^{n}\lambda_i u_i u_i^H$

$\Rightarrow e^{\theta X} = I + (e^{\theta}-1)X \longrightarrow \boxed{\mathbb{E}[e^{\theta X}] = I + (e^{\theta}-1)\mathbb{E}[X]}$

$\mathbb{E}[e^{\theta X}] \leq I + (e^{\theta}-1)\mathbb{E}[X] = \sum_{i=1}^{n}\left(1+(e^{\theta}-1)\lambda_i\right) u_i u_i^H \leq \sum_{i=1}^{n} \exp((e^{\theta}-1)\lambda_i) u_i u_i^H$

$\longrightarrow \mathbb{E}[e^{\theta X}] \leq \exp((e^{\theta}-1)\mathbb{E}X) \longrightarrow \boxed{\mathbb{E}[e^{\theta X}] \leq \exp([e^{\theta}-1]\cdot \mathbb{E}X)}$ Q.E.D

$\forall j: \lambda_{max}(V^T X_j V) \leq \psi(V) \longleftrightarrow V^T X_j V \leq \psi(V) I$

$\longrightarrow \mathbb{E}\left[\exp\left(\beta \cdot \frac{V^T X_j V}{\psi(V)}\right)\right] \leq \exp\left((e^{\beta}-1)\,\mathbb{E}\left[\frac{V^T X_j V}{\psi(V)}\right]\right)$

thus $\forall \theta \geq 0: \mathbb{E}\left[\exp(\theta V^T X_j V)\right] \leq \exp\Big(\underbrace{\frac{e^{\theta\psi(V)}}{\psi(V)}}_{f(\theta)} \cdot \underbrace{V^T(\mathbb{E}X_j) V}_{A_j}\Big)$

$\Rightarrow \mathbb{P}\left[\lambda_k\left(\sum_{j=1}^{n} X_j\right) \geq (1+\delta)\mu_k\right] \leq \inf_{\theta>0}\left\{\min_{V\in V^d_{d-k+1}}\left\{e^{-\theta(1+\delta)\mu_k} \cdot \mathrm{tr}\left\{\exp\left(g(\theta)\sum_{j=1}^{n} V^T A_j V\right)\right\}\right\}\right\}$

$\leq \inf_{\theta>0}\left\{e^{-\theta(1+\delta)\mu_k} \cdot \mathrm{Tr}\left\{\exp\left(\frac{e^{\theta\psi(V_+)}-1}{\psi(V)} V_+^T(\mathbb{E}X_j)V_+\right)\right\}\right\}$

$\leq \inf_{\theta>0}\left\{e^{-\theta(1+\delta)\mu_k} \cdot (d-k+1)\,\lambda_{max}\left\{\exp\left(\frac{e^{\theta\psi(V_+)}-1}{\psi(V_+)} \cdot V_+^T(\mathbb{E}X_j)V_+\right)\right\}\right\}$

$= \inf_{\theta>0}\left\{e^{-\theta(1+\delta)\mu_k} \cdot (d-k+1) \cdot \exp\left(\frac{e^{\theta\psi(V_+)}-1}{\psi(V_+)} \cdot \mu_k\right)\right\}$

$= (d-k+1) \inf_{\theta>0}\left\{\exp\left(\frac{e^{\theta\psi(V_+)}-1}{\psi(V_+)} \cdot \mu_k\right)\right\} = (d-k+1)\exp\left(\inf_{\theta>0}\{---\}\right)$

Now in order to optimize the term over θ, we differentiate it with respect to θ:

$$\text{let } h(\theta) = \frac{e^{\theta\psi(V_+)}-1}{\psi(V_+)}$$

ادامه سوال ۵:

$$\frac{\partial h(\theta)}{\partial \theta} = 0 \rightarrow e^{\theta \psi(V_+)} = 1+\delta \rightarrow \theta^* = \frac{\log(1+\delta)}{\psi(V_+)}$$

$$\rightarrow \mathbb{P}\left[\lambda_k\left(\sum_{j=1}^{n} X_j\right) \geq (1+\delta)\mu_k\right] \leq (d-k+1)\; h(\theta^*)$$

$$\leq (d-k+1)\cdot\left[\exp\left(\delta-(1+\delta)\log(1+\delta)\right)\right]^{\frac{\mu_k}{\psi(V_+)}} = \left[\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right]^{\frac{\mu_k}{\psi(V_+)}}\cdot(d-k+1)$$

which proves that

$$\boxed{\mathbb{P}\left[\lambda_k\left(\sum X_j\right) \geq (1+\delta)\mu_k\right] \leq (d-k+1)\cdot\left[\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right]^{\frac{\mu_k}{\psi(V_+)}}}$$

Q.E.D.

$$\Rightarrow \lambda_k\left(\sum_{j=1}^{n} X_j\right) = -\lambda_{d-k+1}\left(-\sum_{j=1}^{n} X_j\right).$$

$$\mathbb{P}\left[\lambda_k\left(\sum_{j=1}^{n} X_j\right) \leq (1-\delta)\mu_k\right] = \mathbb{P}\left(\lambda_{d-k+1}\left(-\sum X_j\right) \geq -(1-\delta)\mu_k\right)$$

$$\leq \inf_{\theta>0}\left\{\min_{V\in V^d_{d-k+1}}\left\{e^{\theta(1-\delta)\mu_k}\cdot \mathrm{tr}\left\{\exp\left(g(\theta)\sum_{j=1}^{n} V^T A_j V\right)\right\}\right\}\right\}$$

$$\leq \inf_{\theta>0}\left\{e^{\theta(1-\delta)\mu_k}\cdot \mathrm{Tr}\left\{\exp\left(\frac{e^{-\theta\psi(V)}-1}{\psi(V)}\right)\cdot\left(V^T \mathbb{E}X_j V\right)\right\}\right\}$$

$$\leq \inf_{\theta>0}\left\{e^{\theta(1-\delta)\mu_k}\cdot k\cdot \lambda_{max}\left\{\exp\left(\frac{e^{-\theta\psi(V)}-1}{\psi(V)}\cdot\left(V^T \mathbb{E}X_j V\right)\right)\right\}\right\}$$

$$= \inf_{\theta>0}\left\{e^{\theta(1-\delta)\mu_k}\cdot k\cdot \exp\left(\frac{e^{-\theta\psi(V)}-1}{\psi(V)}\right)\cdot\lambda_{max}\left\{V^T\mathbb{E}X_j V\right\}\right\}$$

$$= \inf_{\theta>0}\left\{e^{\theta(1-\delta)\mu_k}\cdot k\cdot \exp\left(\frac{e^{-\theta\psi(V)}-1}{\psi(V)}\cdot(-\mu_k)\right)\right\}$$

$$\geq k\cdot \inf_{\theta>0}\left\{\exp\left(\frac{e^{-\theta\psi(V_-)}-1}{\psi(V_+)}\cdot\mu_k + \theta(1-\delta)\mu_k\right)\right\}$$

$\hookrightarrow$ optimizing over $\theta \rightarrow \frac{\partial}{\partial\theta} = 0 \rightarrow \theta^* = -\frac{\log(1-\delta)}{\psi(V_-)}$

which finally implies:

$$\mathbb{P}\left[\lambda_k\left(\sum_{j=1}^{n} X_j\right) \leq (1-\delta)\mu_k\right] \leq k\cdot\exp\left(-\frac{\delta\mu_k}{\psi(V_-)} - \frac{\mu_k}{\psi(V_-)}\cdot(1-\delta)\log(1+\delta)\right)$$

$$= k\cdot\left(e^{-\delta}/(1-\delta)^{1-\delta}\right)$$

Q.E.D.

Scanned by CamScanner