Q1)

a) $Q = gB$ $\quad E[e^{\lambda Q}] = \sum_{k=0}^{\infty} \frac{\lambda^{2k} g^{2k} B^{2k}}{(2k)!} \leq \sum \frac{1}{k!}\left(\frac{\lambda^2 g^2 B^2}{2}\right)^k$

we know that $E[e^{\lambda g}] \leq e^{\frac{\sigma^2 \lambda^2}{2}}$

means that: $\sum \frac{\lambda^2 g^2}{(2k)!} \leq \sum \frac{1}{k!}\left(\frac{\lambda^2 g^2}{2}\right)^k \leq \sum \frac{1}{k!}\left(\frac{\lambda^2 \sigma^2}{2}\right)^k$ (1)

From (1) $\rightarrow$ $\sum \frac{1}{k!}\left(\frac{\lambda^2 g^2 B^2}{2}\right)^k \leq \sum \frac{1}{k!}\left(\frac{\lambda^2 \sigma^2 B^2}{2}\right)^k$

$\Longrightarrow E[e^{\lambda Q}] \leq \exp\left(\frac{\lambda^2 \sigma^2 B^2}{2}\right)$

$\Longrightarrow V = \sigma^2 B^2$ ✓ $\rightarrow$ wainwright

or

$B = U_g D U^T$ $\quad E[e^{\lambda Q}] = E[e^{\lambda U_g^T D U}] = U^T E[e^{\lambda g D}] U$

$\leq U^T e^{\frac{\lambda^2 \sigma^2 D^2}{2}} U = \frac{1}{2} e^{\frac{1}{2}\lambda^2 \sigma^2 B^2}$ ✓

b) $[e^{\lambda Q}] = E[e^{\lambda g B}] = \sum E\left[\frac{\lambda^{2k} g^{2k} B^{2k}}{(2k)!}\right] \Longrightarrow \sum \frac{1}{k!}\left(\frac{\lambda^2 \sigma^2 b^2 I}{2}\right)^k$

$\Longrightarrow B^{2k} \leq (bI)^2 \leq b^2 I$

$\Longrightarrow E[e^{\lambda Q}] \leq \exp\left(\frac{\lambda^2 \sigma^2 b^2 I}{2}\right) \Rightarrow$ $V = b^2 I$ ✓

or we can conclude from a that

$B \leq bI_d$ and so matrix exponential is monotone with diagonalisable matrices so the b hold! ●

Q2)

$$E\left[\lambda_{max}\left(\overbrace{\frac{1}{n}\sum Q_i}^{S_n}\right)\right] = E\left[\lambda_{max}(S_n)\right]$$

$$\exp\left(\lambda E\left[\lambda_{max}(S_n)\right]\right) \leq E\left[e^{\lambda \overbrace{\lambda_{max}(S_n)}^{Y_{max}(S_n)}}\right] \quad (1)$$

$$E\left[e^{\lambda Y_{max}(S_n)}\right] = E\left[Y_{max}\left(e^{\lambda S_n}\right)\right] \leq E\left[tr\left(e^{\lambda S_n}\right)\right]$$

$$= tr\left(E\left[e^{\lambda S_n}\right]\right) = tr\left(\Psi_{nS_n}(\lambda/n)\right)$$

lemma 6.13 $\rightarrow$ mainwright $\Rightarrow$ $tr\,\Psi_{S_n}(\lambda) \leq tr\left(e^{\sum_{i,j} \Psi_{Q_i}(\lambda)}\right)$ (1)

(1) $\rightarrow$ $tr\left(\Psi_{S_n}(\lambda/n)\right) \leq tr\left(\exp\left(\sum_{i,j} \Psi_{Q_i}\left(\frac{\lambda}{n}\right)\right)\right)$

by sub-g exponential $\rightarrow$ $\Psi_{Q_i}(\lambda/n) \leq \exp\left(\frac{\lambda^2}{2n^2} V_i\right)$ (2)

(2) $\rightarrow$ $tr\left(\exp\left(\sum_{i,j} \Psi_{Q_i}(\lambda/n)\right)\right) \leq tr\left(\exp\left(\frac{\lambda^2}{2n^2}\sum V_i\right)\right)$

now have $tr\left(\exp\left(\frac{\lambda^2}{2n^2}\sum V_i\right)\right)$

which is less than norm.

$$\sigma^2 = \left\| \frac{1}{n} \Sigma_i V_i \right\|_2$$

$$\Longrightarrow \quad \operatorname{tr}\left(\exp\left(\frac{\lambda^2}{2n^2} \left(\Sigma_i V_i \right) \right) \right) \leq \operatorname{tr} \exp\left(\frac{\lambda^2 \sigma^2}{2n} I \right)$$

$$\Longrightarrow \quad \leq d \exp\left(\frac{\lambda^2 \sigma^2}{2n} \right)$$

$$\Longrightarrow \exp\left(\lambda E\left[Y_{max}(S_n) \right] \right) \leq \exp\left(\frac{\lambda^2 \sigma^2}{2n} \lg d \right)$$

$$E\left[Y_{max}(S_n) \right] \leq \frac{\lambda \sigma^2}{2n} \lg d$$

$$\frac{\lambda \sigma^2}{2n} \lg d = \sqrt{\frac{2\sigma^2 \lg d}{n}} \Longrightarrow$$

$$\frac{\lambda^2 \cancel{\sigma^4} \cancel{\lg^2 d}}{4 \cancel{n^2} n} = \frac{2 \cancel{\sigma^2} \cancel{\lg d}}{\cancel{n}} \Longrightarrow \lambda^2 = \frac{8n}{\sigma^2}$$

$$\lambda = \frac{2}{\sigma} \sqrt{2n \lg d}$$

$$\Longrightarrow E\left[Y_{max}(S_n) \right] \leq \sqrt{\frac{2\sigma^2}{n} \lg d}$$

b) $E\left[\left\|\frac{1}{n}\sum Q_i\right\|_2\right] = E\left[\max\{Y_{max}(S_n), |Y_{min}(S_n)|\}\right]$

as we did for Part a it goes:

$$E\left[e^{\lambda|S_n|_2}\right] = E\left[e^{\lambda \max\{Y_{max}(S_n), |Y_{min}(S_n)|\}}\right]$$

by union bound we have $\Longrightarrow$

$$E\left[e^{\lambda|S_n|_2}\right] \leq E\left[e^{\lambda Y_{max}(S_n)} + e^{\lambda|Y_{min}(S_n)|}\right]$$

if $Y_{max} > |Y_{min}| \Longrightarrow$ second term not needed.

if $Y_{max} < |Y_{min}| \Rightarrow -Y_{min} > Y_{max} \Rightarrow$

$$\max\{Y_{max}, |Y_{min}|\} = \max\{Y_{max}, -Y_{min}\} \quad (1)$$

$$(1) \Longrightarrow E\left[e^{\lambda|S_n|}\right] \leq E\left[e^{\lambda Y_{max}} + e^{-\lambda Y_{min}}\right]$$

$$\Longrightarrow E\left[e^{\lambda|S_n|}\right] \leq d\, e^{\lambda^2\sigma^2/2n} + d\, e^{\lambda^2\sigma^2/2n}$$

$$\Longrightarrow E\left[e^{\lambda|S_n|}\right] \leq 2d\, e^{\lambda^2\sigma^2/2n}$$

$$\Longrightarrow \exp\left(\lambda E\left[\left\|\frac{1}{n}\sum Q_i\right\|_2\right]\right) \leq e^{\frac{\lambda^2\sigma^2}{2n}} 1.j2d$$

$$\Longrightarrow E\left[\left\|\frac{1}{n}\sum Q_i\right\|_2\right] \leq \frac{\lambda\sigma^2}{2n} 1.j2d$$

by doing same in a)

$$\frac{\lambda\sigma^2}{2n} \log 2d = \sqrt{\frac{2\sigma \log 2d}{n}} \Rightarrow$$

$$\lambda = \frac{2}{\sigma}\sqrt{n \log 2d}$$

$$\Rightarrow E\left[\left\|\frac{1}{n}\sum Q_i\right\|_2\right] \leq \sqrt{\frac{2\sigma \log 2d}{n}}$$ ✓

Q3) $Q_i = \begin{bmatrix} 0 & A_i \\ A_i^T & 0 \end{bmatrix} \Rightarrow Q^2 = \begin{bmatrix} A_i A_i^T & 0 \\ 0 & A_i^T A_i \end{bmatrix}$

$\det(Q_i^2 - \lambda I) \Rightarrow (A_i A_i^T - \lambda I)(A_i^T A_i - \lambda I) \Rightarrow$ eigvalues of Q_i^2 are either eigvalues of $A_i A_i^T$ or $A_i^T A_i \Rightarrow$ these are the same.

$$\|Q_i\|_2 = \sqrt{\|A_i A_i^T\|_2} = \sqrt{A_i^T A_i} = \|A_i\|_2$$

now Bernstein condition for $g_i Q_i \rightarrow$ are.

↳ theorem 6.17

$$E[Q_i^j] \leq \frac{1}{2} j!\, b^{j-2} \, var(Q) \rightarrow \text{condition}$$

$$E[(g_i Q_i)^j] \leq \frac{j!}{2} (b_1 b_2)^{j-2} \, var(g_i Q)$$

$$var(Q) \Rightarrow ? \quad E[(g_i Q_i)^2] \Rightarrow \| \sum_i v(Q_i) \|_2 =$$

$$\max \{ \| E[g_i A_i A_i^T g_i^T] \|_2 , \| E[g_i A_i^T A_i g_i] \|_2 \} = \sigma^2 b_2^2]$$

$$\Rightarrow E[(g_i Q_i)^j] \leq \frac{j!}{2} (b_1 b_2)^{j-2} \sigma^2 b_2^2]$$

note $\Rightarrow \sigma^2 = \max \{ \| \frac{1}{n} \sum A_i A_i^T \|_2 , \| \frac{1}{n} \sum [A_i^T A_i] \|_2 \}$

↳ from 6.43 book!

ok now we have Bernsteen condition $\Rightarrow$

$$E[(g_i Q_i)^j] \leq \frac{j!}{2} (b_1 b_2)^{j-2} \sigma^2 b_2^2]$$

now because eigvalues of Q_i are symmetric so

if λ is eigvalue of $g_i Q_i$ then $-\lambda$ is also eigvalue

$\Rightarrow$ we only need one-sided bound of Bernstein

$$\Rightarrow P[\| \frac{1}{n} \sum Q_i \| \geq \delta] \leq \text{rank}(\sum v_i) \, e^{\frac{-n\delta^2}{2\sigma^2}}$$

↳ $(\sigma^2 b_2^2 + b_1 b_2 \delta)$

$$d_{+1} \left\{ \begin{bmatrix} 0 & A_i \\ A_i^T & 0 \end{bmatrix} \right. \leftarrow d_1 + d_2$$

$$\Rightarrow P[\| \frac{1}{n} \sum Q_i \| \geq \delta] \leq (d_1 + d_2) \, e^{\frac{-n\delta^2}{(\sigma^2 b_2^2 + b_1 b_2 \delta)}}$$

✓ inspired from ex 6.11 book.

Q4)

a) inspired By section 6.5.1 with collary 6.24

first we define.

$$T_{\lambda_n}(M) = (\hat{M})_{ij} = \begin{cases} m_{ij} & |m_{ij}| > \lambda_n \\ 0 & \text{otherwise} \end{cases}$$

we use collary 6.24 here without proof for proof go to HDS (wainwright) collary 6.24 so why bother copying the proof from the book.

because real cov is D and Diagonal, almostsurly we can think that $\hat{\Sigma}_{ij}$ $i \neq j$ are very small so when we apply $T_{\lambda_n}(\hat{\Sigma})$ they become zero

now we use collary 6.24 the we know D has at most 1 non-zero entries per row so

$\Rightarrow$ with setting $\frac{\lambda_n}{\sigma^2} = 8\sqrt{\frac{\log d}{n}} + \delta$ means that

$\hat{\Sigma}_{ij} \leq \sigma^2 8 \sqrt{\frac{\log d}{n}} + \delta$ are zero

then we can have $\Rightarrow$

$$\hat{D} = \underbrace{T_{\lambda_n}(\hat{\Sigma})}_{\Downarrow} = \mathrm{Diag}(\hat{\Sigma})$$

this leaves all non-diag zero almost surly.

So we can say that $\mathrm{Diag}(\hat{\Sigma})$ is doing $T_{\lambda_n}(\hat{\Sigma})$ Job so

$$\Rightarrow P\left[\|T_{\lambda_n}(\hat{\Sigma}) - \Sigma\|_2 \geq 2\underset{\downarrow 1}{s}\lambda_n\right] \leq 8e^{\frac{-n}{16}\min(\delta,\delta^2)}$$

So we have $P\left[\|\hat{D} - D\|_2 \geq 2\left(8\sqrt{\frac{\log d}{n}} + \delta\right)\sigma^2\right] \leq 8e^{\frac{-n}{16}\min(\delta,\delta^2)}$

$$\Rightarrow P\left[\frac{\|\hat{D} - D\|}{\sigma^2} \geq \underset{\downarrow 2}{c_0}\sqrt{\frac{\log d}{n}} + \delta\right] \leq c_1 \exp\left(-c_2\, n \min(\delta, \delta^2)\right)$$

$\Rightarrow$ By using collary 6.24 then Proof is complete. ∎

for Proof of collary 6.24 or theorem 6.23 ∎

for more elaborate Proof we can have:

$x_{ij} \in \mathrm{sub}\,G(\sigma) \Rightarrow \sum_{i=1}^{n} x_{ij}^2$ is sub-Exponential with Param $(\underset{\downarrow 1}{2\sigma^2}, 4\sigma^2)$

$$\Longrightarrow \quad \frac{1}{n}\sum X_{ij}^2 \in \text{sub-E}\left(\frac{2\sigma^2}{\sqrt{n}}, \frac{4\sigma^2}{n}\right)$$

$$P\left[(\hat{D}_{ii} - D_{ii}) \geq t\right] \leq 2e^{-\frac{n}{8}\min(t,t^2)}$$

$\Longrightarrow$ By union bound we have $P\{\|\hat{D}-D\|_2 \geq t\} \leq 2d\, e^{-n/8 \min(t,t^2)}$

$$\Rightarrow P[\|\hat{D}-D\| \geq t+\delta] \leq 2de^{-n/8\min(t,t^2) - n/8\,\delta^2}$$

$C_0^2/8 = 1 \Rightarrow C_0^2 = 8$

$$\delta = C_0\sqrt{\frac{\log d}{n}}$$

$$\Rightarrow P[\|\hat{D}-D\|_2 > t+\delta] \leq 2d\, e^{-n/8(\min(t,t^2)) - \frac{n}{8} C_0^2 \frac{\log d}{n}}$$

$$P[\|\hat{D}-D\|_2 > t+\delta] \leq 2\cancel{d}\, e^{-n/8\min(t,t^2)} \frac{1}{\cancel{d}} = 2e^{-n/8\min(t,t^2)}$$

when $n < C_0^2 \log d$ we have $\Rightarrow$

$$P[\|\hat{D}-D\|_2 \geq C_0^2 \log d/n + t] \leq 2e^{-n/8\min(t,t^2)}$$

$\Longrightarrow$ Combining the results we have.

$$P\left[\|\hat{D}-D\|_2 \geq \min\left(C_0 \log d/n, \sqrt{C_0^2 \log n} + \delta\right)\right] \leq 2e^{-\frac{n}{8}\min(\delta^2,\delta)}$$

$C_0 = \sqrt{8}$

b) $\Rightarrow \underbrace{P\left[\|\hat{D}-D\|_2^m \geq (4\delta)^m \sqrt{\frac{d^2}{n^m}}\right]}_{P} \xrightarrow{\text{markov}}$

$$P \leq \frac{1}{(2\delta)^m} \times \frac{n^{m/2}}{d\, 2^m} E\left[\max_{i\in d} |\hat{D}_{ii} - D_{ii}|^m\right] \quad (1)$$

$$\frac{1}{d} E\left[\max |\hat{D}_{ii} - D_{ii}|^m\right] \leq \frac{1}{d} \sum_i E\left[|D_{ii} - D_{ii}|^m\right] = \|\hat{D}_{11} - D_{11}\|_m^m$$

union bound

$$\Rightarrow P\left(\|\hat{D} - D\| \geq 48\sqrt{\frac{d^{2/m}}{n}}\right) \leq \frac{1}{(28)^m} \frac{n^{m/2}}{2^m} \underbrace{\|\hat{D}_{11} - D_{11}\|_m^m}$$

$$\text{so } \|\hat{D}_{11} - D_{11}\|_m^m \leq \frac{k}{n^m}\left(\left[\sum E(x_{i1} - D_{11})^2\right]^{\frac{1}{2}} + \left[\sum |x_{i1}^2 - D_{11}|^m\right]^{\frac{1}{m}}\right)^m$$

$$\Rightarrow \leq \frac{k}{n^m}\left(\sqrt{n C_2} + (n C_3)^{\frac{1}{m}}\right)^{\frac{1}{m}} \leq k \frac{2^m}{n^{m/2}} \max(C_2, C_3)$$

combine

$$\Rightarrow P\left[\ \ldots \geq 48\sqrt{\frac{d^{2/m}}{n}}\right] \leq \frac{k \max(C_2, C_3)}{(28)^m} = k' \frac{1}{(28)^m}$$

Q5) this is inspired by section 6.2.2 from an Introduction to matrix concentration Inequalities

$$E[\bar{R}_m] = \frac{1}{m} \sum_{i=1}^{m} E[R_k] = \frac{1}{m} \times m \sum B_i = B$$

$$p_i \times \frac{1}{p_i} B_i$$

So $P[\|\bar{R}_m - \overbrace{E[R_m]}^{B}\|_2 \geq \delta]$?

assum $\{X_k\}$ independent, random matrices, with common dimension $d_1 \times d_2$

$E[X_k] = 0$ and $\|X_k\| \leq L$ for all k

$$S = \sum_k X_i$$

$$\mathrm{var}(S) = \max\{\|E[SS^T], E[S^TS]\}\} = \max\{\|\sum_k E[XX^T]\|, \|\sum E[S^TS]\|\}$$

Bernstein $\rightarrow$ $P[\|S\| \geq \delta] \leq (d_1+d_2)\exp\left(\frac{-\delta^2}{2(V(S)+L\delta)}\right)$

in wainwright this was σ^2

so this is Bernstein in other notation.

let's continue with Proof nnn with Bernstein we have:

$$P[|\bar{R}_n - B| \geq \delta] \leq (d_1+d_2)\exp\left(\frac{-m\delta^2}{2(V(R)+L\delta}\right)$$

$d_1 = d_2 = d \Rightarrow 2d$

$$V(R) = \left\|\frac{1}{m}\sum \mathrm{var}(R_i)\right\| = \left\|\sum \frac{1}{p_i}B_i^2 - B^2\right\| = \sigma^2$$

$* \ \sigma^2 = \left\| \sum \frac{1}{p_i} B_i^2 - B^2 \right\|$

$L = \| B_K - B \| \leq$ some bound $\Rightarrow \max \left\{ \left(\frac{1}{p_i} - 1\right) \|B_i\| + \sum_{i \neq j} B_j \right\} = b$

set $L = b =$

$\rightarrow$ this part was with help of mohammad mavshal pirgheibi

$$\Rightarrow P\{ \qquad \} \leq 2d \exp\left(\frac{-m \delta^2}{2(\sigma^2 + b\delta)} \right)$$

the actual bound with very specific detail

is $\leq 2d \exp\left(\frac{-\delta^2}{2\left(V(A) + \frac{L\delta}{3}\right)} \right)$

which tighter so our bound still works!! Yey...