36-709: Advanced Statistical Theory I **Spring 2020**

Lecture 9: February 18

Lecturer: Siva Balakrishnan *Scribe: Oscar Li*

9.1 Background

In this class, we will talk about matrix concentration inequalities. Our basic set up is, given a collection of independent symmetric random matrices $Q_1, \ldots, Q_n \in \mathbb{S}^{d\times d}$, with mean $\mathbf{0}_{d\times d}$, we would like to bound the maximum eigenvalue of their average $\overline{Q} = \frac{1}{n}\sum_{i=1}^n Q_i$:

$$\mathbb{P}(\gamma_{\max}(\overline{Q}) \geq t).$$

A standard way to approach this is through a Chernoff argument, for any $s > 0$:

$$\begin{aligned}
&\mathbb{P}(\gamma_{\max}(\overline{Q}) \geq t)\\
&= \mathbb{P}(s\gamma_{\max}(\overline{Q}) \geq st)\\
&= \mathbb{P}(\gamma_{\max}(s\,\overline{Q}) \geq st)\\
&= \mathbb{P}[\exp(\gamma_{\max}(s\,\overline{Q})) \geq \exp(st)]\\
&\overset{\text{(i)}}{=} \mathbb{P}[\gamma_{\max}(\exp(s\,\overline{Q})) \geq \exp(st)]\\
&\overset{\text{(ii)}}{\leq} \exp(-st)\cdot\mathbb{E}[\gamma_{\max}(\exp(s\,\overline{Q}))]\\
&\overset{\text{(iii)}}{\leq} \exp(-st)\cdot\mathbb{E}[\mathrm{tr}(\exp(s\,\overline{Q}))]\\
&\overset{\text{(iv)}}{=} \exp(-st)\cdot\mathrm{tr}(\mathbb{E}[\exp(s\,\overline{Q})]).
\end{aligned} \tag{9.1}$$

In step (i) we use the fact that the matrix exponential preserves the relative ordering of the symmetric matrix's eigenvalues. Step (ii) is standard Markov inequality. Step (iii) uses the fact that $\exp(s\,\overline{Q})$ is a positive definite matrix. Therefore its trace, which is a sum of its eigenvalues, is greater than its maximum eigenvalue. Step (iv) trace is a linear operator so it can commute with expectation.

Taking inf over all of $s > 0$ in the above argument would complete the Chernoff argument. Now it remains for us to bound the term $\mathrm{tr}(\mathbb{E}[\exp(s\,\overline{Q})])$. In the standard scalar case, the exponential of the average could be written as the product of individual exponentials $\exp(\frac{1}{n}\sum_{i=1}^n X_i) = \prod_{i=1}^n \exp(\frac{X_i}{n})$. Then we can proceed through properties of individual random variable's moment generating function directly. However in matrix exponential, we need $AB = BA$ to have $\exp(A)\exp(B) = \exp(A+B)$. As a result, we cannot directly do a factorization on arbitrary realization of $Q_1, \ldots, Q_n$. We will see later in this lecture how to handle this problem to eventually relate back to individual random matrix's moment generating function.

9.2 Sub-Gaussian and sub-exponential matrices

Like the real-valued random variable case, we characterize a class of random matrix variables through their moment generating function.

Definition 9.1 (Sub-Gaussian random matrices) *A zero-mean symmetric random matrix $Q \in \mathbb{S}^{d\times d}$ is sub-Gaussian with matrix parameter V^2, $V \in \mathbb{S}_+^{d\times d}$, if*

$$\mathbb{E}[\exp(tQ)] \preceq \exp\left(\frac{t^2V^2}{2}\right), \text{ for all } t \in \mathbb{R}\,.$$

Definition 9.2 (Sub-Exponential random matrices) *A zero-mean symmetric random matrix $Q \in \mathbb{S}^{d\times d}$ is sub-exponential with parameters (V^2, b), $V \in \mathbb{S}_+^{d\times d}$, $b \geq 0$ if*

$$\mathbb{E}[\exp(tQ)] \preceq \exp\left(\frac{t^2V^2}{2}\right), \text{ for all } |t| \leq \frac{1}{b}.$$

Remark 9.3 *Sub-Gaussian random matrix with parameter V^2 is sub-exponential with parameter $(V^2, 0)$.*

Example 9.4 *For any fixed symmetric matrix A, define the random matrix variable $Q = \epsilon A$, where $\epsilon \in \{+1, -1\}$ is the Rademacher random variable. Then Q is a sub-Gaussian random matrix with $V^2 = A^2$. To see this, for any $t \in \mathbb{R}$,*

$$\begin{aligned}
\mathbb{E}[\exp(tQ)] &= \mathbb{E}[\sum_{k=0}^{\infty} \frac{t^k \epsilon^k A^k}{k!}] \\
&= \mathbb{E}[\sum_{k=0}^{\infty} \frac{t^{2k} \epsilon^{2k} A^{2k}}{(2k)!}] + \mathbb{E}[\sum_{k=0}^{\infty} \frac{t^{2k+1} \epsilon^{2k+1} A^{2k+1}}{(2k+1)!}] \\
&= \sum_{k=0}^{\infty} \frac{t^{2k} A^{2k}}{(2k)!} \\
&\preceq \sum_{k=0}^{\infty} \frac{(t^2)^k (A^2)^k}{2^k \cdot k!} \\
&= \sum_{k=0}^{\infty} \frac{(\frac{tA^2}{2})^k}{k!} \\
&= \exp\left(\frac{tA^2}{2}\right).
\end{aligned}$$

A similar argument can show that if we replace the ϵ Rademacher random variable with a zero mean sub-Gaussian real-valued random variable g with parameter σ^2, the random matrix $Q = gA$ is sub-Gaussian with $V^2 = \sigma^2 A^2$.

Remark 9.5 *We know that bounded real-valued random variable is sub-Gaussian. So we might naturally wonder whether a "bounded" random matrix is guaranteed to be sub-Gaussian: if random matrix Q always satisfies $Q^2 \preceq V^2$, can we bound its moment generating function $\mathbb{E}[\exp(tQ)] \preceq \exp\left(\frac{\lambda t^2 V^2}{2}\right)$ with some constant λ for any t? The answer to this question is No. However, we (might) see in homework that even though we cannot directly bound this type of random matrices' moment generating function, through a symmetrization argument we can still get concentration bound of their averages.*

9.3 Matrix Concentration inequalities

With the appropriate definitions in place, we are ready to state the first matrix concentration inequality.

Theorem 9.6 (Sub-Gaussian bound) *Let $\{Q_i\}_{i=1}^n$ be a sequence of zero-mean independent symmetric random matrices $Q_i \in \mathbb{S}^{d\times d}$ that are sub-Gaussian with parameters $\{V_i^2\}_{i=1}^n$. Denote their average by $\overline{Q} = \frac{1}{n}\sum_{i=1}^n Q_i$. Then for all $t > 0$, we have the bound*

$$\mathbb{P}(\gamma_{\max}(\overline{Q}) \geq t) \leq d\exp(-\frac{nt^2}{2\sigma^2}).$$

where $\sigma^2 = \|\frac{1}{n}\sum_{i=1}^n V_i^2\|_{\text{op}}$.

Before we prove this theorem, we make a few comments.

Remark 9.7 *Comparing this result with the real-valued analog, we notice that there is an extra d in the upper bound. This additional term d is in some sense unavoidable. Consider the following example:*

Let $n = d$, $Q_i = \epsilon_i diag(\boldsymbol{e}_i)$, $\epsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$, where $diag(\boldsymbol{e}_i)$ denotes the diagonal matrix whose only nonzero entry is at (i,i) with value 1. In this case $\gamma_{\max}(\overline{Q}) = \max_{i=1}^n \frac{\epsilon_i}{n}$. Notice that Q_i is a sub-Gaussian random matrix with parameter $diag(\boldsymbol{e}_i)$. Thus $\sigma^2 = \|\frac{1}{n}I_{n\times n}\|_{\text{op}} = \frac{1}{n}$. Then the probability

$$\mathbb{P}(\gamma_{\max}(\overline{Q}) \geq t) = \mathbb{P}(\max_{i=1}^n \epsilon_i \geq nt) = 1 - \prod_{i=1}^n (1 - \mathbb{P}(\epsilon_i \geq nt)) \approx d\exp\left(-\frac{n^2t^2}{2}\right).$$

The last step can also also be seen roughly from a union bound over the independent Gaussian random variables. This shows that the bound in the theorem is tight with the factor d. There is another argument in Martin's book that reasons about this by comparing the Big O order of the random variable implied by the tail bound and the Big O order of the gaussian max divided by n. See p.176 of High Dimensional Statistics: a non-assymptotic viewpont.

To prove this result, we need two deep facts:

(1) **(**log **is operator monotone)** for positive definite matrices A and B,

$$A \preceq B \implies \log(A) \preceq \log(B).$$

(2) **(Lieb's theorem)** For any symmetric matrix $H \in \mathbb{S}^{d\times d}$, the function $f\colon \mathbb{S}_+^{d\times d} \to \mathbb{R}$, $f(A) = \mathrm{tr}(\exp(H + \log(A))$ is concave.

Now we are ready to prove Theorem 9.6.

Proof: [Theorem 9.6]

Based on the Chernoff argument (9.1), we now proceed to bound $\mathrm{tr}(\mathbb{E}[\exp(s\,\overline{Q})])$.

$$\begin{aligned}
&\mathrm{tr}(\mathbb{E}[\exp(s\,\overline{Q})])\\
=&\mathbb{E}_{Q_1,\dots Q_{n-1}}[\mathbb{E}_{Q_n}[\mathrm{tr}(\exp(\frac{s}{n}\sum_{i=1}^{n-1} Q_i + \frac{s}{n}Q_n))]]\\
=&\mathbb{E}_{Q_1,\dots Q_{n-1}}[\mathbb{E}_{Q_n}[\mathrm{tr}(\exp(\frac{s}{n}\sum_{i=1}^{n-1} Q_i + \log\exp(\frac{s}{n}Q_n)))]]
\end{aligned}$$

The last step uses the fact that matrix log composed with matrix exponential is identity over $\mathbb{S}^{d\times d}$. For fixed value of $Q_1, \dots, Q_{n-1}$, we see that by Lieb's theorem the function

$$f(X) = \mathrm{tr}(\exp(\frac{s}{n}\sum_{i=1}^{n-1} Q_i + \log X))$$

is concave. By Jensen's inequality, we have $\mathbb{E}[f(X)] \le f(\mathbb{E}[X])$. In our case, the random variable X is $\exp(\frac{s}{n}Q_n)$. As a result, we have

$$\begin{aligned}
&\mathbb{E}_{Q_1,\dots Q_{n-1}}[\mathbb{E}_{Q_n}[\mathrm{tr}(\exp(\frac{s}{n}\sum_{i=1}^{n-1} Q_i + \log\exp(\frac{s}{n}Q_n)))]]\\
\le&\mathbb{E}_{Q_1,\dots Q_{n-1}}[\mathrm{tr}(\exp(\frac{s}{n}\sum_{i=1}^{n-1} Q_i + \log(\mathbb{E}_{Q_n}[\exp(\frac{s}{n}Q_n)])))].
\end{aligned}$$

Repeating this trick $n-1$ times, we get that

$$\mathrm{tr}(\mathbb{E}[\exp(s\,\overline{Q})]) \le \mathrm{tr}(\exp(\sum_{i=1}^{n} \log(\mathbb{E}_{Q_i}[\exp(\frac{s}{n}Q_i)]))).$$

Recall our comment in the first section about the inability to directly factorize the exponential of matrix sum, we see now this problem is avoided here through the application of

Lieb's theorem. We now can deal with individual random matrix Q_i's moment generating function.

By theorem assumption that Q_i is V_i^2 sub-Gaussian, we have

$$\mathbb{E}[\exp(\frac{s}{n}Q_i)] \preceq \exp\left(\frac{s^2 V_i^2}{2n^2}\right).$$

By operator monotonicity of matrix log, we have $\log(\mathbb{E}[\exp(\frac{s}{n}Q_i)]) \preceq \frac{s^2 V_i^2}{2n^2}$. Summing these matrix inequalities over i, we have

$$\sum_{i=1}^{n} \log(\mathbb{E}_{Q_i}[\exp(\frac{s}{n}Q_i)] \preceq \frac{s^2 \sum_{i=1}^n V_i^2}{2n^2}.$$

A fact about the function $X \mapsto \text{tr}(\exp(X))$ is that for any pair of symmetric matrices $Q \preceq R$, we have $\text{tr}(\exp(Q)) \leq \text{tr}(\exp(R))$. Applying this fact, we have

$$\text{tr}(\exp(\sum_{i=1}^{n} \log(\mathbb{E}_{Q_i}[\exp(\frac{s}{n}Q_i)]))) \leq \text{tr}(\exp(\frac{s^2 \sum_{i=1}^n V_i^2}{2n^2})).$$

Because the trace of a $d \times d$ positive definite matrix is upper bouned by d times its maximum eigenvalue (operator norm), we have

$$\begin{aligned}
\text{tr}(\exp(\frac{s^2 \sum_{i=1}^n V_i^2}{2n^2})) \leq & d \cdot \left\| \exp\left(\frac{s^2 \sum_{i=1}^n V_i^2}{2n^2}\right) \right\|_{\text{op}} \\
\leq & d \cdot \exp\left(\left\| \frac{s^2 \sum_{i=1}^n V_i^2}{2n^2} \right\|_{\text{op}} \right) \\
\leq & d \cdot \exp\left(\frac{s^2}{2n} \left\| \frac{\sum_{i=1}^n V_i^2}{n} \right\|_{\text{op}} \right) \\
\leq & d \cdot \exp\left(\frac{s^2 \sigma^2}{2n}\right).
\end{aligned}$$

Thus we have shown that

$$\text{tr}(\mathbb{E}[\exp(s\,\overline{Q})]) \leq d \cdot \exp\left(\frac{s^2\sigma^2}{2n}\right).$$

Plugging this into (9.1), we get

$$\mathbb{P}(\gamma_{\max}(\overline{Q}) \geq t) \leq \inf_{s>0} d \cdot \exp\left(\frac{s^2\sigma^2}{2n} - st\right).$$

Let $s = \frac{nt}{\sigma^2}$, we have $\mathbb{P}(\gamma_{\max}(\overline{Q}) \geq t) \leq d\exp(\frac{-nt^2}{2\sigma^2})$. The proof is complete. ■

Remark 9.8 *With the same assumptions made in Theorem 9.6, we can also get an* ***operator norm*** *bound of the symmetric average matrix* $\overline{Q}$*:*

$$\mathbb{P}(\|\overline{Q}\|_{\mathrm{op}} \geq t) \leq 2d \exp(\frac{-nt^2}{2\sigma^2}).$$

To see this, we notice that $\|\overline{Q}\|_{\mathrm{op}} = \max\{\gamma_{\max}(\overline{Q}), -\gamma_{\min}(\overline{Q})\}$. *By union bound, we then have*

$$\begin{aligned}\mathbb{P}(\|\overline{Q}\|_{\mathrm{op}} \geq t) \leq & \mathbb{P}(\gamma_{\max}(\overline{Q}) \geq t) + \mathbb{P}(-\gamma_{\min}(\overline{Q}) \geq t) \\ = & \mathbb{P}(\gamma_{\max}(\overline{Q}) \geq t) + \mathbb{P}(\gamma_{\max}(-\overline{Q}) \geq t).\end{aligned}$$

When Q_i *is* V_i^2 *sub-Gaussian, so is* $-Q_i$. *Using Theorem 9.6 on the collection of random matrices* $\{-Q_i\}_{i=1}^n$*, we can bound the term* $\mathbb{P}(\gamma_{\max}(-\overline{Q}) \geq t)$. *Thus we have* $\mathbb{P}(\|\overline{Q}\|_{\mathrm{op}} \geq t) \leq 2d\exp(\frac{-nt^2}{2\sigma^2})$.

Theorem 9.6 can be seen as the matrix analog for concentration of sub-Gaussian random variables. There are also concentration inequalities for bounded random matrices (recall from an earlier comment in this lecture that they are not sub-Gaussian random matrices). For bouned random matrices, if we know their "variance", we can also incorporate this information into the bound:

Definition 9.9 (Variance of random matrix) *For random matrix* $Q \in \mathbb{S}^{d\times d}$*, define its variance as* $\mathbb{V}[Q] = \mathbb{E}[Q^2] - (E[Q])^2$.

Remark 9.10 *For any random matrix* Q, $\mathbb{V}[Q] \succeq 0_{d\times d}$.

Now we state the Bernstein inequality for symmetric random matrices.

Theorem 9.11 (Bernstein bound for random matrices) *Let* $\{Q_i\}_{i=1}^s$ *be a sequence of independent zero-mean symmetric random matrices with bounded operator norm: for some* $b > 0$, $\|Q_i\|_{\mathrm{op}} \leq b$ *for all* i. *Then for all* $t \geq 0$*, we have*

$$\mathbb{P}(\gamma_{\max}(\frac{1}{n}\sum_{i=1}^n Q_i) \geq t) \leq d\exp\left(-\frac{nt^2}{2(\sigma^2 + bt)}\right),$$

where $\sigma^2 = \left\|\frac{1}{n}\sum_{i=1}^n \mathbb{V}[Q_i]\right\|_{\mathrm{op}}$.

So far we have proved concentration bounds for symmetric (which are necessarily square) matrices. However, the bounds can be extended to non-symmetric and/or nonsquare random matrices by forming a *self-adjoint dilation*. Given a random matrix $Q_i \in \mathbb{R}^{d_1 \times d_2}$, we form the following matrix $\widetilde{Q_i} \in \mathbb{S}^{(d_1+d_2)\times(d_1+d_2)}$:

$$\widetilde{Q_i} = \begin{bmatrix} \mathbf{0}_{d_1 \times d_1} & Q_i \\ Q_i^T & \mathbf{0}_{d_2 \times d_2} \end{bmatrix}.$$

We can easily see that the new matrix $\widetilde{Q_i}$ is symmetric. In addition, one can show that $\|\widetilde{Q_i}\|_{\text{op}} = \|Q_i\|_{\text{op}}$. This technique can be used to prove a Bernstein bound (c.f. Theorem 9.11) for non-symmetric and/or non-square matrices. There, we define σ^2 as

$$\sigma^2 = \max\left\{ \left\| \frac{1}{n}\sum_{i=1}^{n} Q_i Q_i^T \right\|_{\text{op}}, \left\| \frac{1}{n}\sum_{i=1}^{n} Q_i^T Q_i \right\|_{\text{op}} \right\}.$$

9.4 Conclusion

In this class, we have defined sub-Gaussian and sub-exponential random matrices and introduced concentration bounds for both cases (Theorem 9.6 sub-Gaussian, Theorem 9.11 sub-Exponential). A nice paper with comprehensive results and examples on this topic is Joel Tropp's paper "User-friendly tail bounds for sums of random matrices" [Tro11]. In the paper Tropp proposed the application of Lieb's Theorem to bound the mgf of the sum in terms of individual mgfs, which provides a general framework for this type of matrix concentration analysis.

[Tro11] Tropp, Joel A. "User-friendly tail bounds for sums of random matrices." Foundations of computational mathematics 12.4 (2012): 389-434.