1 Random matrix ensembles

We will consider these three types of random matrices.

1. I.i.d. matrix ensembles. The coefficients $\xi_{i,j}$ are i.i.d. (real or complex) random variables. We will often assume that $\mathbf{E}\xi = 0$ and $\mathrm{Var}(\xi) = 1$.

2. Symmetric Wigner matrix ensembles. The upper triangular coefficients $\xi_{i,j}$, where $j \geq i$, are independent and real, but the lower triangular coefficients $\xi_{i,j}$, where $j < i$, equal their transpose $\xi_{j,i}$. The distribution of the diagonal coefficients and that of the strictly upper triangular coefficients may be different. An important example is the Gaussian Orthogonal Ensemble (GOE), in which the upper triangular entries have distribution $N(0,1)_{\mathbb{R}}$ while the diagonal entries have distribution $N(0,2)_{\mathbb{R}}$.

3. Hermitian Wigner matrix ensembles. The upper triangular coefficients $\xi_{i,j}$, where $j \geq i$, are independent and complex, but the lower triangular coefficients $\xi_{i,j}$, where $j < i$, equal their adjoint $\overline{\xi_{j,i}}$. An important example is the Gaussian Unitary Ensemble (GUE), in which the upper triangular entries have distribution $N(0,1)_{\mathbb{C}}$ while the diagonal entries have distribution $N(0,1)_{\mathbb{R}}$.

2 The operator norm of random matrices

We will be interested in the distribution of the eigenvalues or their related quantities of a $n \times n$ matrix ensemble $M = (\xi_{i,j})$. We will first focus on the operator norm

$$\|M\| := \sup_{\|x\|_2=1} \|Mx\|_2.$$

Note that $\|M\|$ is the square root of the largest eigenvalue of M^*M. How large should $\|M\|$ be if the entries are random? Let say all the entries $\xi_{i,j}$ satisfy $|\xi_{i,j}| \leq 1$. Then because of the fact that all the norms on a finite dimensional vector space are equivalent, by computing the Hilbert–Schmidt norm, we may guess that $\|M\| \sim n$. However, this argument does not make use of independence. (In fact, if you think carefully, the argument does not quite make sense. Although the norms are equivalent on $\mathbb{C}^{n^2}$, the constants in the "norm inequality" do depend on n. Anyway, the point is we will not be able to get a good bound if we don't use the independence structure.) We will see in fact that $\|M\| \sim \sqrt{n}$.

2.1 A concentration inequality

It is usually difficult to derive a limit law for a sequence of random variables (X_n). Instead proving a concentration inequality would be easier and would help us understanding the asymptotic behavior of X_n more. To be more precise, if we have a bound on $\mathbf{P}(X_n > x)$, then we will understand the random variable X_n better.

In our case, we will be interested in the following quantity:

$$\mathbf{P}(\|M\| > \lambda), \quad \lambda \geq 0.$$

By definition, we have

$$\mathbf{P}(\|M\| > \lambda) = \mathbf{P}\left(\text{there exists } x \in \mathbb{C}^n \text{ with } \|x\|_2 = 1 \text{ such that } \|Mx\|_2 > \lambda\right).$$

Let's forget about "there exists $x \in \mathbb{C}^n$ with $\|x\|_2 = 1$ such that" and focus on $\mathbf{P}(\|Mx\|_2 > \lambda)$ first. We have the following.

Lemma 2.1. *Suppose that $\xi_{i,j}$ are independent, $\mathbf{E}\xi_{i,j} = 0$ and $|\xi_{i,j}| \leq 1$. Let $x \in \mathbb{C}^n$ be a unit vector. Then for sufficiently large A, we have*

$$\mathbf{P}(\|Mx\|_2 \geq A\sqrt{n}) \leq C\exp(-cAn)$$

for some absolute constants $C, c > 0$.

Proof. Write $x = (x_1, \ldots, x_n)$. Then each row of Mx is $\sum_{j=1}^n \xi_{i,j}x_j$. By Hoeffding's inequality (we will state the inequality below), we have

$$\mathbf{P}\left(\left|\sum_{j=1}^n \xi_{i,j}x_j\right| \geq \lambda\right) \leq 4\exp(-\lambda^2/4).$$

Note that

$$\begin{aligned}
\mathbf{E}\exp\left(c\left|\sum_{j=1}^n \xi_{i,j}x_j\right|^2\right) &= \mathbf{E}\int_0^{\exp\left(c\left|\sum_{j=1}^n \xi_{i,j}x_j\right|^2\right)} 1 \,\mathrm{d}t \\
&= \mathbf{E}\int_0^\infty \mathbf{1}_{\left\{\exp\left(c\left|\sum_{j=1}^n \xi_{i,j}x_j\right|^2\right) \geq t\right\}}\mathrm{d}t \\
&= \int_0^\infty \mathbf{P}\left(\exp\left(c\left|\sum_{j=1}^n \xi_{i,j}x_j\right|^2\right) \geq t\right)\mathrm{d}t \\
&= 1 + \int_1^\infty \mathbf{P}\left(\exp\left(c\left|\sum_{j=1}^n \xi_{i,j}x_j\right|^2\right) \geq t\right)\mathrm{d}t \\
&= 1 + \int_1^\infty \mathbf{P}\left(\left|\sum_{j=1}^n \xi_{i,j}x_j\right| \geq \sqrt{\frac{\log t}{c}}\right)\mathrm{d}t \\
&\leq 1 + \int_1^\infty 4t^{-1/4c}\mathrm{d}t =: C < \infty
\end{aligned}$$

when $c > 0$ is sufficient small. So by independence,

$$\mathbf{E}\exp(c\|Mx\|_2^2) = \prod_{i=1}^n \mathbf{E}\exp\left(c\left|\sum_{j=1}^n \xi_{i,j}x_j\right|^2\right) \leq C^n.$$

Finally, by Markov's inequality,

$$\mathbf{P}(\|Mx\|_2 \geq A\sqrt{n}) \leq C^n\exp(-cA^2n) = \exp((-cA^2 + \log C)n).$$

Hence the inequality holds when A is sufficiently large. $\square$

Here is the statement of Hoeffding's inequality.

Theorem 2.2 (Hoeffding's inequality)**.** *Let $X_1, \dots, X_n$ be independent real-valued random variables such that $X_i \in [a_i, b_i]$ almost surely for all $i \le n$. Then for every $t \ge 0$,*

$$\mathbf{P}\left(\left|\sum_{i=1}^{n}(X_i - \mathbf{E}X_i)\right| \ge t\right) \le 2\exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

Recall that we were trying to bound

$$\mathbf{P}\left(\text{there exists } x \text{ with } \|x\|_2 = 1 \text{ such that } \|Mx\|_2 > \lambda\right) = \mathbf{P}\left(\bigcup_{x:\|x\|_2=1}\{\|Mx\|_2 > \lambda\}\right).$$

We cannot apply subadditivity on the right hand side since it is an uncountable union. However, this can be "reduced" to a finite union because the map $x \mapsto \|Mx\|_2$ is (Lipschitz) continuous.

Definition 2.3. *A set $\Sigma \subseteq S$ is called an ε-net ($\varepsilon > 0$) of S if for any $x, y \in \Sigma$, $x \ne y$, we have $\|x - y\|_2 \ge \varepsilon$.*

Lemma 2.4. *Let $\varepsilon \in (0, 1)$, and let Σ be an ε-net of $\{x \in \mathbb{C}^n : \|x\|_2 = 1\}$. Then there exists $C > 0$ such that $\#\Sigma \le (C/\varepsilon)^n$.*

Proof. Consider the balls of radius $\varepsilon/2$ centered around each point in Σ. These balls are disjoint, and obviously covered by a ball of radius $3/2$ centered at the origin. This proves the lemma. □

Lemma 2.5. *Let Σ be a maximal $1/2$-net of $\{x \in \mathbb{C}^n : \|x\|_2 = 1\}$. Then for any $\lambda > 0$, we have*

$$\mathbf{P}(\|M\| > \lambda) \le \mathbf{P}\left(\bigcup_{y\in\Sigma}\{\|My\|_2 > \lambda/2\}\right).$$

Proof. Let $x \in \mathbb{C}^n$ with $\|x\|_2 = 1$ such that $\|M\| = \|Mx\|_2$. By maximality, there exists $y \in \Sigma$ such that $\|x - y\|_2 < 1/2$. Therefore,

$$\|Mx\|_2 - \|My\|_2 \le \|M(x - y)\|_2 \le \|M\|\|x - y\|_2 < \|M\|/2.$$

Rearranging, we have $\|My\|_2 > \|M\|/2$. If $\|M\| > \lambda$, then $\|My\|_2 > \lambda/2$ for some $y \in \Sigma$, and this completes the proof. □

We will continue next time.

References

[1] Tao, Terence. Topics in random matrix theory. Graduate Studies in Mathematics, 132. American Mathematical Society, Providence, RI, 2012. x+282 pp. ISBN: 978-0-8218-7430-1