Solution Sheet II

Introduction to Random Matrices, IST Austria
Published October 25, 2017. *Due* November 7, 2017.

LÁSZLÓ ERDŐS, DOMINIK SCHRÖDER

1 Bound on largest eigenvalue

The goal of this exercise is to obtain a direct proof of the fact that the operator norm of random matrices with independent entries of size $1/\sqrt{N}$ is bounded. Because the argument is based on simple concentration estimates it is only applicable to the case of uniformly subgaussian matrix entries.

Definition. A centered random variable X is called subgaussian if there exists $a > 0$ such that for all $t \in \mathbb{R}$ it holds that $\mathbf{E}\, e^{tX} \le e^{at^2/2}$. A family of random variables is called uniformly subgaussian if the same constant a can be chosen for all random variables.

Problem 1. *Prove that the following three assertions about a centered random variable X are equivalent:*

(i) X *is subgaussian,*

(ii) there exists $b > 0$ such that for all $\lambda > 0$, $\mathbf{P}(|X| \ge \lambda) \le 2e^{-b\lambda^2}$,

(iii) there exits $c > 0$ such that $\mathbf{E}\, e^{cX^2} \le 2$.

Solution.

(i)⇒(ii) It follows from Markov's inequality that

$$\mathbf{P}(X \ge \lambda) \le \frac{\mathbf{E}\, e^{tX}}{e^{t\lambda}} \le e^{at^2/2 - t\lambda} \le e^{-\lambda^2/2a}$$

and consequently $\mathbf{P}(|X| \ge \lambda) \le 2e^{-\lambda^2/2a}$.

(ii)⇒(iii) We find

$$\mathbf{E}\, e^{cX^2} = \int_0^\infty \mathbf{P}(e^{cX^2} \ge t)\, \mathrm{d}t = 1 + \int_0^\infty 2c\lambda e^{c\lambda^2} \mathbf{P}(|X| \ge \lambda)\, \mathrm{d}\lambda \le 1 + 4c \int_0^\infty \lambda e^{(c-b)\lambda^2}\, \mathrm{d}\lambda$$

and the claim follows from choosing c small enough.

(iii)⇒(i) Since X is centered it follows that

$$\begin{aligned}\mathbf{E}\, e^{tX} &= 1 + \mathbf{E} \int_0^1 (1-r)(tX)^2 e^{rtX}\, \mathrm{d}r \le 1 + t^2\, \mathbf{E}\, X^2 e^{|tX|} \le 1 + t^2 e^{t^2/2c}\, \mathbf{E}\, X^2 e^{cX^2/2} \\ &\le 1 + t^2 e^{t^2/2c}\, \mathbf{E}\, e^{cX^2} \le 1 + 2t^2 e^{t^2/2c}\end{aligned}$$

and the claim follows for some a small enough. We note that in the second inequality we used the often useful weighted Cauchy-Schwarz inequality $|tX| = \left|\frac{t}{\sqrt{c}}(\sqrt{c}X)\right| \le \frac{t^2}{2c} + \frac{cX^2}{2}$.

□

Let H be a real $N \times N$ random matrix with independent entries $(h_{ij})_{i,j\in[N]}$ of zero mean such that the normalized random variables $\sqrt{N}h_{ij}$ are uniformly subgaussian. The goal of this exercise is to prove that there exists a constant $c > 0$ such that

$$\mathbf{P}(\|H\| > C) \le e^{-cNC^2} \tag{1}$$

for all C large enough.

Problem 2. *Let x be a fixed vector of (Euclidean) length $\|x\| = 1$. Show that there exists a constant $c > 0$ such that*

$$\mathbf{P}(\|Hx\| > C) \le e^{-cNC^2} \tag{2}$$

for all C large enough.

Solution. We first observe that the random variables $\sqrt{N}(Hx)_i$ are uniformly subgaussian. Indeed,

$$\mathbf{E}\, e^{t\sqrt{N}(Hx)_i} = \prod_j \mathbf{E}\, e^{t\sqrt{N}h_{ij}x_j} \le \prod_j e^{at^2x_j^2/2} = e^{at^2/2}.$$

Thus it follows from Problem 1 that there exists a constant $c > 0$ such that $\mathbf{E}\, e^{c[\sqrt{N}(Hx)_i]^2} \le 2$ for all i and we can conclude from Markov's inequality that

$$\mathbf{P}(\|Hx\| \ge C) \le \frac{\mathbf{E}\, e^{cN\|Hx\|^2}}{e^{cC^2N}} = \frac{\prod_i \mathbf{E}\, e^{cN(Hx)_i^2}}{e^{cC^2N}} \le \frac{2^N}{e^{cC^2N}}$$

and the claim follows for some slightly smaller c.

□

Problem 3. *Let $P \subset S := \{x \in \mathbb{R}^N \mid \|x\| = 1\}$ be a maximal $1/2$-separated set, i.e., a set for which for any $x \neq y \in P$ we have that $\|x - y\| \ge 1/2$ and for each $z \in S \setminus P$ there exist $x \in P$ such that $\|x - z\| < 1/2$. Prove that*

$$\mathbf{P}\left(\|H\| > C\right) \le \mathbf{P}\left(\|Hx\| > C/2 \text{ for some } x \in P\right). \tag{3}$$

Solution. By compactness we find $x \in S$ such that $\|H\| = \|Hx\|$. If $x \in P$, then (3) follows immediately. Otherwise we find $y \in P$ such that $\|x - y\| < 1/2$ from maximality. It follows that

$$\|Hy\| = \|Hx + H(y - x)\| \ge \|H\| - \|H\|/2 = \|H\|/2$$

and we conclude that $\|Hy\| > C/2$ for some $y \in P$ whenever $\|H\| > C$. □

Problem 4. *Combine the statements of problems 2 and 3 to prove* (1).

Solution. Let P be any maximal $1/2$-separated set. Then the balls of radius $1/4$ around each point $x \in P$ are disjoint and at the same time contained in the ball of radius $5/4$ around the origin. It follows that $|P| \le K^N$ for some constant K. Thus it follows that

$$\mathbf{P}(\|H\| \ge C) \le \sum_{x \in P} \mathbf{P}(\|Hx\| \ge C/2) \le K^N e^{-cNC^2/4}$$

and we can choose c appropriately such that (1) holds for C large enough. □

2 Interlacing eigenvalues

Let A be a Hermitian $N \times N$ matrix and let B be the $(N-1) \times (N-1)$ principal submatrix, i.e.,

$$A = \begin{pmatrix} B & a \\ a^* & b \end{pmatrix}$$

for some vector $a \in \mathbb{C}^{N-1}$ and scalar $b \in \mathbb{R}$ and suppose that A and B have disjoint, simple spectra. Let $\lambda_1 < \lambda_2 < \cdots < \lambda_N$ and $\mu_1 < \cdots < \mu_{N-1}$ denote the ordered eigenvalues of A and B. The goal of this exercise is to prove that the eigenvalues of B interlace the eigenvalues of A, i.e., that

$$\lambda_1 < \mu_1 < \lambda_2 < \cdots < \lambda_{N-1} < \mu_{N-1} < \lambda_N. \tag{4}$$

We begin with an equivalent characterisation of interlacing.

Problem 5. *Let f, g be polynomials with distinct simple real roots and leading positive coefficients of degree N and $N-1$. Prove that the roots interlace if and only if for each $\lambda \in [0,1]$ the polynomial $\lambda f + (1-\lambda)g$ has only real roots.*

Solution. Denote the ordered union of roots of f and g by $\nu_1 < \cdots < \nu_{2N-1}$. Since both polynomials have leading positive coefficients it follows that both f and g are positive in the interval $I_0 = (\nu_{2N-1}, \infty)$. In the interval $I_1 = (\nu_{2N-2}, \nu_{2N-1})$ the polynomials f, g thus have opposite signs and continuing inductively we find that the polynomials f, g have equal signs in some interval $I_j = (\nu_{2N-1-j}, \nu_{2N-j})$ if and only if j is even.

Now suppose that $h_\lambda := \lambda f + (1-\lambda) g$ has real roots for all λ which can only lie in the intervals $I_1, I_3, \dots, I_{2N+1}$. The roots of h_λ are continuous functions of λ and thus cannot jump from one interval to another. By considering λ close to 0 and 1 it follows that for each odd interval the two interval endpoints have to be roots of f and g, respectively. In particular it follows that the roots interlace.

Now suppose that the roots interlace. Then $h_\lambda(\nu_{2N-1}) > 0$, $h_\lambda(\nu_{2N-3}) < 0$, $h_\lambda(\nu_{2N-5}) > 0$ and so on and we find N real roots from the intermediate value theorem. □

Problem 5 was wrong as stated. It turns out that the polynomials $f(x) = x^2 - 1$ and $g(x) = x - 2$ provide a counter example. Can you spot the mistake in the above argument? The correct version follows now.

Problem 5. *Let f, g be polynomials with distinct simple real roots and leading positive coefficients of degree N and $N-1$. Prove that the roots interlace if and only if for each $\lambda, \mu \in \mathbb{R}$ the polynomial $h = \lambda f + \mu g$ has only real roots.*

Solution. Denote the ordered roots of f by $\alpha_1 < \cdots < \alpha_N$. First assume that the roots are interlacing and fix $\mu, \lambda \in \mathbb{R}$. Then $\operatorname{sgn} h(\alpha_k) = \operatorname{sgn} \mu g(\alpha_k)$ is alternating since the same is true for g. Thus h has at least $N-1$ real roots by the intermediate value theorem and as complex roots come in pairs of complex conjugates and $\deg h = N$ it follows that all N roots of h are real.

Conversely, assume that h has only real roots for all λ, μ. Then the rational function $\psi(x) = \lambda/\mu + g(x)/f(x)$ has a derivative ψ' of constant sign. Indeed, if not then we could find a root of ψ' and by choosing λ/μ appropriately we could create a pair of complex roots of $h = \mu \psi f$, in contradiction with the assumption. Thus the sign of $g' f - f' g = f^2 \psi'$ is also constant and thus $f'(\alpha_i) g(\alpha_i)$ has the same sign for all i. As $f'(\alpha_i)$ alternates in sign so does $g(\alpha_i)$ and we can find rind a root of g between any two consecutive α_i. □

It turns out that Problem 5 provides a useful characterisation of interlacing for proving (4).

Problem 6. *Find a Hermitian matrix $C = C(\lambda)$ such that*

$$\det(x - C) = \det(x - A) + \frac{\mu}{\lambda} \det(x - B)$$

and conclude from Problem 5 that the eigenvalues of B interlace those of A, i.e., (4).

Solution. By linearity of the determinant

$$\det \begin{pmatrix} x - B & -a \\ -a^* & x - b + \frac{\mu}{\lambda} \end{pmatrix} = \det \begin{pmatrix} x - B & -a \\ -a^* & x - b \end{pmatrix} + \det \begin{pmatrix} x - B & -a \\ 0 & \frac{\mu}{\lambda} \end{pmatrix} = \det(x - A) + \frac{\mu}{\lambda} \det(x - B),$$

where the polynomial on the lhs. is real-rooted as the characteristic polynomial of a Hermitian matrix. By multiplying both sides by λ it follows from Problem 5 that the roots of $\det(x - B)$ interlace those of $\det(x - A)$, proving the claim. □

3 Resolvent identities

The goal of this exercise is to prove two identities relating resolvent elements to those of the resolvent of certain minors. Let H be a Hermitian $N \times N$ matrix. For $i \in [N]$ let $H^{(i)}$ denote the matrix with the i-th row and column set to zero, i.e., $H^{(i)}_{kl} = H_{kl} \mathbf{1}_{i \neq k} \mathbf{1}_{i \neq l}$. The entries of the resolvent $G^{(i)}(z) := (H^{(i)} - z)^{-1}$ of such a minor satisfy the *first resolvent decoupling identity* for $i, j \neq k$

$$G_{ij} = G^{(k)}_{ij} + \frac{G_{ik} G_{kj}}{G_{kk}} \tag{5a}$$

as well as the *second resolvent decoupling identity* for $i \neq j$

$$G_{ij} = -G_{ii} \sum_{k \neq i} h_{ik} G^{(i)}_{kj} = -G_{jj} \sum_{k \neq j} G^{(j)}_{ik} h_{kj}. \tag{5b}$$

Before proving (5a)–(5b) we recall the standard resolvent expansion formula.

Problem 7. *Show that for matrices A, B it holds that*

$$(A-B)^{-1} = A^{-1} + A^{-1}B(A-B)^{-1} = A^{-1} + (A-B)^{-1}BA^{-1}, \tag{6}$$

provided that all inverses exist.

Solution. We only prove the first equality as the proof for the second is identical. Multiplying both sides with $(A-B)$ from the right gives

$$1 = 1 - A^{-1}B + A^{-1}B = 1$$

and the claimed equality follows. $\square$

Problem 8. *Use* (6) *to first prove* (5b) *and then use* (6) *and* (5b) *to prove* (5a)*.*

Solution. We first prove (5b) and again only prove the first equality as the proof of the second one is identical. It follows from (6) that

$$G = (H-z)^{-1} = (H^{(i)} - z - [H^{(i)} - H])^{-1} = G^{(i)} + G(H^{(i)} - H)G^{(i)}$$

and consequently

$$G_{ij} = G_{ij}^{(i)} - \sum_k G_{ik}h_{ki}G_{ij}^{(i)} - \sum_k G_{ii}h_{ik}G_{kj}^{(i)} = -G_{ii}\sum_{k\neq i} h_{ik}G_{kj}^{(i)},$$

where the second equality followed from $G_{ij}^{(i)} = 0$.

We now proceed with the proof of (5a). It follows from (6) that

$$G = (H-z)^{-1} = (H^{(k)} - z - [H^{(k)} - H])^{-1} = G^{(k)} + G(H^{(k)} - H)G^{(k)}$$

and therefore it follows from (5b) that

$$G_{ij} = G_{ij}^{(k)} - \sum_{l\neq j} G_{ik}h_{kl}G_{lj}^{(k)} = G_{ij}^{(k)} + \frac{G_{ik}G_{kj}}{G_{kk}},$$

completing the proof. $\square$