Assignment 4

Random Matrix Theory in Data Science and Statistics

(EN.553.796, Fall 2025)

Assigned: November 7, 2025   Due: 11:59pm EST, December 1, 2025

Solve all problems. Each problem is worth an equal amount towards your grade.

Submit solutions in LaTeX. Write in complete sentences. Include and justify all steps of your arguments, but avoid writing excessive discussion that is not contributing to your solution. You are welcome to include images if you think that will help explain your solutions.

Problem 1 (Concentration inequalities)**.** This problem will give you some intuition for how Poincaré and modified log-Sobolev (MLS) inequalities behave for scalar random variables. Recall that, for $X \in \mathbb{R}$ a scalar random variable, the Poincaré inequality is the condition that, for some $C > 0$, for all continuously differentiable f where both sides are finite,

$$\mathsf{Var}[f(X)] \leq C \cdot \mathbb{E}(f'(X))^2, \tag{1}$$

and the MLS inequality is instead the condition that, for the same class of f,

$$\mathsf{Ent}[\exp(f(X))] \leq C \cdot \mathbb{E}\left[(f'(X))^2 \exp(f(X))\right]. \tag{2}$$

For this problem, you may make further mild assumptions on the class of f the inequalities quantify over, as needed to ensure that all expectations you encounter are guaranteed to converge (the actual definition of the correct "domain" of these inequalities can get quite technical).

1. Suppose that X satisfies the Poincaré inequality (1) for some constant C. Show that there exists some $K > 0$ depending only on the law of X such that the absolute moments of X are bounded as
$$\mathbb{E}|X|^n \leq (Kn)^n \tag{3}$$
for all $n \geq 1$. Using this, give an example of a random variable all of whose moments are finite but which does not satisfy a Poincaré inequality.

   (**Hint:** Consider $f(x) = x^m$ in the Poincaré inequality.)

2. Show that a random variable $X \sim \mathsf{Exp}(1)$ (with density $\mathbb{1}\{x \geq 0\} \exp(-x)\, dx$) satisfies a Poincaré inequality for some $C > 0$. Compute its moments and show that they are at least $\mathbb{E}|X|^n \geq (cn)^n$ for some $c > 0$, so the bound in (3) is roughly tight.

   (**Hint:** Show that you may assume $f(0) = 0$ in the Poincaré inequality without loss of generality. Bound $\mathsf{Var}[f(X)] \leq \mathbb{E}[f(X)^2]$, and explore how integration by parts works in the integrals associated to such expectations.)

3. Show that $X \sim \mathsf{Exp}(1)$ does *not* satisfy an MLS inequality for any $C > 0$.

   (**Hint:** Consider $f(X) = \lambda X$ as $\lambda \to 1$ from below.)

4. Show that if X satisfies the MLS inequality with constant C, then it also satisfies the Poincaré inequality with constant $2C$. Together with Part 3, this implies that, up to the specific values of the constants, the MLS inequality is a strictly stronger condition than the Poincaré inequality.

   (**Hint:** For a given $f(x)$ to test in the Poincaré inequality, consider $\widetilde{f}(x) = \lambda f(x)$ in the MLS inequality in the limit $\lambda \to 0$.)

Problem 2. This problem is a continuation of Homework 3, Problem 3, Part 3. Recall that statement: if $\boldsymbol{g} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ and $\boldsymbol{h} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda})$ are arbitrary N-dimensional centered Gaussian vectors such that $\mathbb{E}(g_i - g_j)^2 \leq \mathbb{E}(h_i - h_j)^2$ for all $i, j \in [N]$, then $\mathbb{E} \max_i g_i \leq \mathbb{E} \max_i h_i$.

1. The *Gaussian width* of a compact set $\mathcal{X} \subset \mathbb{R}^d$ is

$$\omega(\mathcal{X}) := \mathop{\mathbb{E}}_{\boldsymbol{g} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_d)} \max_{\boldsymbol{x} \in \mathcal{X}} \langle \boldsymbol{g}, \boldsymbol{x} \rangle.$$

   By the orthogonal invariance of $\boldsymbol{g}$, you may view this as measuring the width of $\mathcal{X}$ in the direction of a random ray through the origin in $\mathbb{R}^d$; hence the name. Let $\boldsymbol{W} \sim \mathsf{GOE}(d, 1)$ (recall this means $W_{ij} = W_{ji} \sim \mathcal{N}(0, 1 + \mathbb{1}\{i = j\})$ independently for all $i \leq j$). Show that, for any compact $\mathcal{X} \subseteq \mathbb{S}^{d-1}(1)$,

$$\mathop{\mathbb{E}}_{\boldsymbol{W} \sim \mathsf{GOE}(d)} \max_{\boldsymbol{x} \in \mathcal{X}} \boldsymbol{x}^\top \boldsymbol{W} \boldsymbol{x} \leq 2\omega(\mathcal{X}).$$

   (**Hint:** Use ϵ-nets to reduce to the case of finite $\mathcal{X}$ and use the inequality from Homework 3 cited above. You will find it useful to prove the linear-algebraic inequality $\|\boldsymbol{w}\boldsymbol{x}^\top - \boldsymbol{y}\boldsymbol{z}^\top\|_F^2 \leq \|\boldsymbol{w} - \boldsymbol{y}\|^2 + \|\boldsymbol{x} - \boldsymbol{z}\|^2$ for all vectors $\boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ of compatible sizes and each of unit norm.)

2. Specializing the above result, prove the following three bounds. Note that there are no further error terms in any of these statements; they hold exactly as written and

non-asymptotically for all d:

$$\mathbb{E}[\lambda_1(\boldsymbol{W})] = \mathbb{E}\left[\max_{\boldsymbol{x}\in\mathbb{S}^{d-1}(1)} \boldsymbol{x}^\top \boldsymbol{W}\boldsymbol{x}\right] \leq 2\cdot\sqrt{d},$$

$$\mathbb{E}\left[\max_{\boldsymbol{x}\in\{\pm 1/\sqrt{d}\}^d} \boldsymbol{x}^\top \boldsymbol{W}\boldsymbol{x}\right] \leq 2\sqrt{\frac{2}{\pi}}\cdot\sqrt{d},$$

$$\mathbb{E}\left[\max_{\substack{\boldsymbol{x}\in\mathbb{S}^{d-1}(1) \\ x_i\geq 0 \text{ for all } i\in[d]}} \boldsymbol{x}^\top \boldsymbol{W}\boldsymbol{x}\right] \leq \sqrt{2}\cdot\sqrt{d}.$$

Explain why the first result is sensible in relation to but does not follow from the semicircle limit theorem.

3. Let $\boldsymbol{G} \sim \mathcal{N}(0,1)^{\otimes d\times m}$, i.e., a random rectangular $d\times m$ matrix with i.i.d. standard Gaussian entries. Adapt your argument from above to show that

$$\mathbb{E}[\|\boldsymbol{G}\|] \leq \sqrt{d} + \sqrt{m}.$$

Explain why this result is sensible in relation to but does not follow from the Marchenko-Pastur limit theorem.

Problem 3. This problem, using a technique in a somewhat similar spirit to Homework 3, Problem 3 used above, explores the concentration of functions of Gaussian random variables without using MLS inequalities. Let $f:\mathbb{R}^n\to\mathbb{R}$ be a smooth function and let $\boldsymbol{g}\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{I}_n)$ be a standard n-dimensional Gaussian vector.

1. Let $\boldsymbol{g},\boldsymbol{h}\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{I}_n)$ independently, so that $\boldsymbol{h}$ is an independent copy of $\boldsymbol{g}$. Write $\boldsymbol{g}(t) = \sin(t)\boldsymbol{g} + \cos(t)\boldsymbol{h}$ for $t\in[0,\pi/2]$. This is a kind of interpolation in the style of Homework 3, Problem 3, but between two independent copies of the *same* Gaussian random vector. Show that

$$f(\boldsymbol{g}) - f(\boldsymbol{h}) = \int_0^{\pi/2} \langle \nabla f(\boldsymbol{g}(t)), \boldsymbol{g}'(t)\rangle\, dt,$$

and that $\mathsf{Law}((\boldsymbol{g}(t),\boldsymbol{g}'(t))) = \mathsf{Law}(\boldsymbol{g},\boldsymbol{h}) = \mathcal{N}(\boldsymbol{0},\boldsymbol{I}_{2n})$ for all $t\in[0,\pi/2]$, whereby the integrand above has the same law at each $t\in[0,\pi/2]$.

2. Let $\Psi:\mathbb{R}\to\mathbb{R}$ be a convex function. Show that

$$\mathop{\mathbb{E}}_{\boldsymbol{g},\boldsymbol{h}} \Psi(f(\boldsymbol{g}) - f(\boldsymbol{h})) \leq \mathop{\mathbb{E}}_{\boldsymbol{g},\boldsymbol{h}} \Psi\left(\frac{\pi}{2}\langle\nabla f(\boldsymbol{g}),\boldsymbol{h}\rangle\right),$$

and therefore also that

$$\mathop{\mathbb{E}}_{\boldsymbol{g}} \Psi(f(\boldsymbol{g}) - \mathbb{E}f(\boldsymbol{g})) \leq \mathop{\mathbb{E}}_{\boldsymbol{g},\boldsymbol{h}} \Psi\left(\frac{\pi}{2}\langle\nabla f(\boldsymbol{g}),\boldsymbol{h}\rangle\right).$$

(**Hint:** Jensen's inequality, repeatedly.)

3. Suppose now that f is also *L-Lipschitz*, i.e. that $\|\nabla f(\boldsymbol{x})\| \le L$ for all $\boldsymbol{x} \in \mathbb{R}^n$ (for a smooth function this is equivalent to the usual definition). Prove that $f(\boldsymbol{g}) - \mathbb{E}f(\boldsymbol{g})$ is $(\frac{\pi}{2}L)^2$-subgaussian, and therefore that

$$\mathbb{P}\left[|f(\boldsymbol{g}) - \mathbb{E}f(\boldsymbol{g})| > t\right] \le 2\exp\left(-\frac{2}{\pi^2}\frac{t^2}{L^2}\right).$$

   This is a slightly weaker (in the constant in the exponential) version of the Gaussian Lipschitz concentration inequality we discussed in class. In fact, relatively simple further arguments show that the same holds even if f is not smooth, and can be Lipschitz in the weaker sense that $|f(\boldsymbol{x}) - f(\boldsymbol{y})| \le L\|\boldsymbol{x} - \boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. You may use this without proof below.

   (**Hint:** Choose the Ψ to use in Part 2 that makes sense to show subgaussianity.)

4. Let $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ for any $n \times n$ covariance matrix $\boldsymbol{\Sigma} \succeq \boldsymbol{0}$. Write $\sigma^2 := \max_{i=1}^n \mathsf{Var}[x_i] = \max_{i=1}^n \Sigma_{ii}$. Let $M := \max_{i=1}^n x_i$. Prove that there is an absolute constant $c > 0$ such that

$$\mathbb{P}[|M - \mathbb{E}M| > t] \le 2\exp\left(-c\frac{t^2}{\sigma^2}\right).$$

   (**Hint:** Write $\boldsymbol{x} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{g}$ for $\boldsymbol{g} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$ and use Part 3.)

5. Prove that the map $\boldsymbol{\lambda} : \mathbb{R}^{d\times d}_{\mathsf{sym}} \to \mathbb{R}^d$ mapping a matrix to its eigenvalues in descending order is 1-Lipschitz when matrices are given the Frobenius norm and vectors the ℓ^2 norm. Using that, prove that there is an absolute constant $c > 0$ such that, for $\boldsymbol{W} \sim \mathsf{GOE}(d)$, $\widehat{\boldsymbol{W}} := \frac{1}{\sqrt{d}}\boldsymbol{W}$, any L-Lipschitz function $f : \mathbb{R} \to \mathbb{R}$, and any $i \in [d]$,

$$\mathbb{P}\left[\left|\frac{1}{d}\sum_{i=1}^{d} f(\lambda_i(\widehat{\boldsymbol{W}})) - \mathbb{E}\frac{1}{d}\sum_{i=1}^{d} f(\lambda_i(\widehat{\boldsymbol{W}}))\right| > t\right] \le 2\exp\left(-c\frac{t^2}{L^2}d^2\right),$$
$$\mathbb{P}\left[\left|f(\lambda_i(\widehat{\boldsymbol{W}})) - \mathbb{E}f(\lambda_i(\widehat{\boldsymbol{W}}))\right| > t\right] \le 2\exp\left(-c\frac{t^2}{L^2}d\right).$$

   Note that the first result makes it easy to upgrade weak convergence in expectation of $\mathsf{esd}(\widehat{\boldsymbol{W}})$ to the semicircle law to various stronger kinds of weak convergence (weak convergence in probability, L^2 or any L^p, almost surely, and so forth).

   (**Hint:** For the part about the $\boldsymbol{\lambda}$ map, look through your old homework problems.)