STAT 291 **Spring 2024**

Problem Set 3

Instructions: many points are available in the problems below. 100 **points will count as a perfect score.** Subproblems are equally weighted unless stated otherwise, and you can earn credit for later subproblems without solving the previous ones. Points will be cumulative throughout the semester, so you will get credit for earning more than 100 points on this problem set if you choose to do so. You may also submit solutions to *extra problems* (listed in a separate file on the course website) at the end of any problem set. **Solutions may be handwritten or Latexed** and should be submitted in PDF format via Canvas or email. The official due date for this assignment is **March 22**, though you are recommended to complete it before Spring break.

Collaboration with your classmates is encouraged. **Please identify everyone you worked with at the beginning of your solution PDF** (e.g. *Collaborators: Alice, Bob, and GPT4*). Your solutions should be written entirely by you, even if you collaborated to solve the problems.

The first person to report each mistake in this problem set (by emailing me and Yufan) will receive up to 5 extra points, depending on the mistake.

Problem 1. Spiked Random Matrices (120 points)

In this problem, we let G_N be a $GOE(N)$ random matrix, and consider the quadratic spin glass Hamiltonian

$$H_N(\boldsymbol{x}) = \langle \boldsymbol{x}, G_N \boldsymbol{x} \rangle + h\langle \mathbf{g}, \boldsymbol{x} \rangle$$

where $\mathbf{g} \sim \text{Normal}(0, I_N)$ and $h > 0$ is a scalar. We also define the spiked random matrix Hamiltonian:

$$\widehat{H}_N(\boldsymbol{x}) = \left\langle \boldsymbol{x}, \left(G_N + \frac{\alpha \mathbf{g}^{\otimes 2}}{N} \right) \boldsymbol{x} \right\rangle = \langle \boldsymbol{x}, G_N \boldsymbol{x} \rangle + \frac{\alpha}{N} \langle \mathbf{g}, \boldsymbol{x} \rangle^2.$$

We will be interested in their maximum values. In the latter case, this amounts to understanding the top eigenvector of $A_N = G_N + \frac{\alpha \mathbf{g}^{\otimes 2}}{N}$.

1. Show that H_N satisfies the topological trivialization criterion from class for all $h > 0$. Conclude that it has ground state energy $\lim_{N\to\infty} \sup_{\boldsymbol{x}\in\mathcal{S}_N} H_N(\boldsymbol{x})/N = \sqrt{h^2+4}$.

2. By varying the value of h, find the limiting maximum value of $H_N(\cdot)/N$ on $\{\boldsymbol{x} \in \mathcal{S}_N \ : \ R(\boldsymbol{x}, \mathbf{g}) = q\}$ for each $q \in [0, 1]$. (Hint: this is similar to the later part of Homework 1 Problem 1, but with ground states instead of free energies.)

3. Using the previous part, show that the asymptotic maximum eigenvalue of A_N is:

$$\operatorname*{p-lim}_{N\to\infty} \lambda_{\max}(A_N) = \begin{cases} \alpha + \frac{1}{\alpha}, & \alpha > 1, \\ 2, & \alpha \in [0, 1]. \end{cases}$$

4. Show that for any α, the second eigenvalue of A_N is $2 \pm o(1)$ with high probability. Thus the top eigenvalue is an outlier when $\alpha > 1$. (Cauchy's eigenvalue interlacing theorem may be helpful.)

We will next see how to explicitly construct the outlier eigenvector when $\alpha > 1$, given access to G_N and $\mathbf{g}$ separately. This will give an alternate proof of the lower bound. Below, to break the $\pm$ symmetry we take the top eigenvector of A_N to have positive overlap with $\mathbf{g}$, and normalize it to have length $\sqrt{N}$ (i.e. lie on $\mathcal{S}_N$).

5. For $\alpha > 1$, find the limiting overlap R_* between $\mathbf{g}$ and the maximizer of H_N. (Possibly using the formulas from the previous parts.)

6. Consider $H_N(\boldsymbol{x})$ restricted to the band

$$Band_{R_*}(\mathbf{g}) \equiv \{\boldsymbol{x} \in \mathcal{S}_N \ : \ R(\boldsymbol{x}, \widetilde{\mathbf{g}}) = R_*\},$$

where $\widetilde{\mathbf{g}} = \sqrt{N}\mathbf{g}/\|\mathbf{g}\|\mathcal{S}_N$ is a slightly normalized version of $\mathbf{g}$. This band is another sphere (of dimension $N-2$), and the restriction of $H_N(\cdot)$ is still a quadratic function. We recenter this restriction of H_N by defining, for $\boldsymbol{\sigma} = R_*\widetilde{\mathbf{g}} + \sqrt{1-R_*^2}\boldsymbol{\rho} \in Band_{R_*}(\mathbf{g})$:

$$\widetilde{H}_N(\boldsymbol{\rho}) = H_N(\boldsymbol{\sigma}) - H_N(R_*\widetilde{\mathbf{g}}).$$

Show that conditionally on $\mathbf{g}$, the function $\widetilde{H}_N$ is another spin glass Hamiltonian on $Band_{R_*}(\mathbf{g})$, and find its effective parameters in the sense of lecture 11. (They should be essentially identical to H_N modulo scaling.)

7. Using the similarity, identify a sub-band $Band_{R_*}^{(2)} \subseteq Band_{R_*}(\mathbf{g})$, defined by overlaps with both $\mathbf{g}$ and $\nabla H_N(R_*\boldsymbol{x})$, such that the maximizer of H_N is within distance $o(\sqrt{N})$ of $Band_{R_*}^{(2)}$ with high probability.

8. Fix $\delta > 0$. By repeating the above construction a large constant $K = K(\alpha, \delta)$ number of times, find a point $\boldsymbol{x}_* \in \mathcal{S}_N$ such that

$$H_N(\boldsymbol{x}_*) \geq \sup_{\boldsymbol{x} \in \mathcal{S}_N} H_N(\boldsymbol{x}) - \delta N \tag{1}$$

holds with high probability.

9. Adapt this algorithm to find a point $\widehat{\boldsymbol{x}}_*$ which approximately maximizes $\widehat{H}_N$ in the same sense.

10. Show that for $\delta > 0$ small enough depending on $\varepsilon > 0$, with high probability as $N \to \infty$, the top eigenvector of A_N is within $\varepsilon\sqrt{N}$ of $\widehat{\boldsymbol{x}}_*$. (Hint: recall Part 4 above.)

11. (**Harder, counts as two parts**) Give an analogous recursive algorithm for topologically trivial mixed p-spin models. That is, for $\xi'(1) > \xi''(1)$ with Hamiltonian H_N, using $K(\delta)$ gradient computations of H_N, show how to compute a δ-approximate maximizer in the sense of (1).

Problem 2. Langevin Dynamics and Shattering (80 points)

1. Let M_t be an $\mathbb{R}^N$-valued continuous time martingale with $M_0 = 0$ taking the form

$$\mathrm{d}M_t = a_t \mathrm{d}B_t$$

where the symmetric $N \times N$ matrices a_t are progressively measurable and have all eigenvalues in $[0,1]$ almost surely. **Though it is not strictly necessary, you may also assume that** $\sup_{t\in[0,\delta]} \|M_t\|$ **is almost surely bounded by a constant depending on** (N, δ)**.**

   (a) Using Ito's formula or otherwise, show that for any unit vector v, the value $e^{\alpha\langle M_t, v\rangle - \alpha^2\|v\|^2 t/2}$ is a super-martingale.

   (b) Show that for fixed $\delta > 0$, one has $\sup_{t\in[0,\delta]} \|M_t\| \leq 1000\sqrt{\delta N}$ with probability $1 - e^{-N}$.

2. Given C-bounded H_N, prove the following holds for Langevin dynamics started from any $x_0 \in \mathcal{S}_N$. For any β, ε there exists $\delta > 0$ such that for large N,

$$\mathbb{P}[\sup_{0 \le t \le \delta} \|x_t - x_0\| \le \varepsilon\sqrt{N}] \ge 1 - e^{-N}.$$

(Hint: it might help to decompose $x_t - x_0$ into its martingale and finite-variation parts and bound them separately. If you assumed $\|M_t\|$ is almost surely bounded in the previous part, make sure to check the condition applies here.)

3. In class, we proved the concentration of $H_N(\boldsymbol{x}_t)/N$ for any fixed $t = O(1)$, where $\boldsymbol{x}_t$ is given by Langevin dynamics for the spherical spin glass Hamiltonian H_N. **Assuming** the limit

$$E(t) = \operatorname*{p-lim}_{N \to \infty} H_N(\boldsymbol{x}_t)/N$$

exists for all $t \ge 0$, show $E(t)$ is a continuous function.

   - For 50% extra credit, show there exists a subsequence $N_1 < N_2 < \dots$ along which the random function $t \mapsto H_{N_i}(\boldsymbol{x}_t)/N_i$ converges in distribution within the space $C([0,T];\mathbb{R})$ for any $T \ge 0$.

4. Suppose $H_N : \mathcal{S}_N \to \mathbb{R}$ is C-bounded and admits a shattering decomposition at inverse temperature β, with cluster diameter $r\sqrt{N} \le \sqrt{N}/100$ and separation $s\sqrt{N} \ge 10r\sqrt{N}$. Initialize Langevin dynamics from its stationary measure $\boldsymbol{x}_0 \sim \mu_\beta(H_N)$. Prove that for some $c > 0$, for N sufficiently large:

$$\mathbb{P}\left[\sup_{0 \le t \le e^{cN}} \|\boldsymbol{x}_t - \boldsymbol{x}_0\| \le 2r\sqrt{N}\right] \ge 1 - e^{-cN}.$$

Conclude that the mixing time of Langevin dynamics is exponentially large in the shattered phase. (Hint: note that $\boldsymbol{x}_t \sim \mu_\beta(H_N)$ holds for any fixed $t \ge 0$. Show that if $\boldsymbol{x}_\tau$ is far from every cluster at a stopping time τ, this separation is likely to persist on $t \in [\tau, \tau + \delta]$.)

Problem 3. Survey (4 points)

Rate both of the problems above that you worked on from 1 to 5 based on:

- Interestingness (1 for boring, 5 for exciting)
- Difficulty (1 for too easy, 5 for too hard)
- Learning (1 for almost nothing, 5 if you learned a lot).

Optionally, you are encouraged to elaborate on your ratings, and to share any other comments you have regarding this problem set or the recent lectures. Suggestions on potential improvements for future weeks or years of the course are especially appreciated.