Statistics 291: Lecture 8 (February 15th, 2024)

Kac-Rice IV: topological trivialization

Instructor: Mark Sellke

Scribe: Xiaodong Yang

1 Topological trivialization

Before everything, we recap the determinant bounds for GOE(N) matrices,

$$\frac{1}{N}\log|\det(\mathrm{GOE}(N)-tI_N)|\approx\psi(t):=\int\log|u-t|\mathrm{d}\nu_{sc}(t)$$
$$=\frac{t^2}{4}-\frac{1}{2}+\mathbb{1}\{|t|\geq 2\}\left(-\frac{t+\sqrt{t^2-4}}{4}+\log\left(\frac{\sqrt{t^2-4}+|t|}{2}\right)\right).$$

Natural Langevin dynamics. The concept of topological trivialization relates to the question of

"When are spin glasses actually glassy?"

To motivate this point, consider the natural Langevin dynamics, which initialize from $x_0\in S_N$ (uniform or independent of $H_{N,p}$) and proceed by

$$\mathrm{d}x_t=P_x^{\perp}\mathrm{d}B_t+\left(\beta\nabla_{\mathrm{sph}}H_{N,p}(x_t)+\begin{bmatrix}\textit{another drift term by Ito}\\ \textit{correction for curvature}\end{bmatrix}\right)\mathrm{d}t.$$

where B_t is N-dimensional Brownian motion. The stationary law of this continuous-time Markov chain is the Gibbs measure μ_β. But at very low temperature (large β), these dynamics can be very complex with exponential mixing time. An "aging" phenomenon is anticipated at low temperature, which can be informally stated in the context of $1\ll t\ll N$,

$$\lim_{t\to\infty}\lim_{N\to\infty}R(x_t,x_{\xi t})=f(\xi),\quad f(\alpha)\approx\alpha^{-c}. \tag{1}$$

Intuitively, if there exist "many critical points", the dynamics easily get stuck and tend to be complex.

Topological trivialization. This concept refers to the case of $O(1)$ critical points (2 critical points in particular). It is likely to happen when the Hamiltonian is added with an external field,

$$H_N=\left[\text{disordered term e.g. } H_{N,p}\right]+\left[\text{simple "signal" term}\right].$$

One specific example is tensor PCA, where for large SNR $\lambda>0$, the Hamiltonian takes the form of

$$H_N=H_{N,p}+N\lambda R(x,\sigma)^p.$$

It then suggests that given a warm start x_0 which is correlated with σ, it would be easy to find the MLE (maximizer of H_N).

2 Spherical p-spin model with random external field

Now we move to our model for today,

$$H_N(x) = H_{N,p}(x) + h\langle \vec{g}, x\rangle, \quad \vec{g} \sim \text{Normal}(0, I_N). \tag{2}$$

As shown later, this model has only 2 critical points with high probability if $h > \sqrt{p(p-2)}$. It would need a 2-dimensional variational problem to solve this optimization problem of this Hamiltonian. Similar conclusions also hold for another alternative setup where $\vec{g} \in S_N$.

The Hamiltonian $H_N(x)$ can be seen as a centered Gaussian process with covariance

$$\mathbb{E}H_N(x)H_N(y) = N\xi(R(x,y)), \quad \xi(R) = R^p + h^2 R. \tag{3}$$

Under this setup, we have

$$\xi(1) = h^2 + 1, \quad \xi'(1) = h^2 + p, \quad \xi''(1) = p(p-1).$$

Similar results also hold for a mixed p-spin model in which the Hamiltonian is taken as

$$H_N(x) = \sum_{p=1}^{P} \gamma_p^2 H_{N,p}(x), \quad \mathbb{E}H_N(x)H_N(y) = N\xi(R(x,y)), \quad \xi(R) = \sum_{p=1}^{P} \gamma_p^2 R^p.$$

Thus we will work in this generality during the rest of class. In general, it holds that

- If $\xi'(1) > \xi''(1)$, there are only 2 critical points.
- If $\xi'(1) < \xi''(1)$, it holds that $\mathbb{E}|\text{Crt}_{S_N}(H_N)| \geq e^{cN}$.

Our subsequent analysis starts with a joint Gaussian distribution for $H_N(x)$ and its spherical gradient and tangent Hessian.

Proposition 2.1 (Lemma 3.2 in [1]). *For fixed $x \in S_N$,*

(a) $\nabla_{\text{sph}} H_N(x)$ *is independent of* $\left(H_N(x), \nabla_{\text{rad}} H_N(x), \nabla^2_{\tan} H_N(x)\right)$, *with* $\nabla_{\text{sph}} H_N(x) \sim$ *Normal*$(0, \xi'(1) I_{N-1})$;

(b) $\left(\frac{H_N(x)}{N}, \nabla_{\text{rad}} H_N(x)\right)$ *is a centered Gaussian with covariance*

$$\begin{pmatrix} \xi(1) & \xi'(1) \\ \xi'(1) & \xi'(1) + \xi''(1) \end{pmatrix} = \sum_{p=1}^{P} \gamma_p^2 \begin{pmatrix} 1 & p \\ p & p^2 \end{pmatrix};$$

(c) $\nabla^2_{\tan} H_N(x)$ *is independent of* $(H_N(x), \nabla_{\text{rad}} H_N(x))$, *with* $\nabla^2_{\tan} H_N(x) \sim \sqrt{\xi''(1)}\,\text{GOE}(N-1)$.

According to Kac-Rice formula, the main term of expected number of critical points is

$$\det\left(\nabla^2_{\text{sph}} H_N(x)\right) = \det\left(\nabla^2_{\tan} H_N(x) - \nabla_{\text{rad}} H_N(x) I_{N-1}\right).$$

Specifically, $Z = \nabla_{\text{rad}} H_N(x)$ is the most important. Given Z, the conditional distribution of $E = \frac{H_N(x)}{N}$ is still Gaussian with

$$\mathbb{E}[E|Z] = \frac{\xi'(1)Z}{\xi'(1) + \xi''(1)}, \quad \text{Var}(E|Z) = \frac{1}{N}\left(\xi(1) - \frac{\xi'(1)^2}{\xi'(1) + \xi''(1)}\right). \tag{4}$$

It suffices to count the number of critical points with respect to each value of Z, since it follows that

$$\frac{1}{N}\log \mathbb{E}|\text{Crt}(H_N, E \approx a, Z \approx b)| \approx \frac{1}{N}\log \mathbb{E}|\text{Crt}(H_N, Z \approx b)| - \frac{(a - \mathbb{E}[E|Z=b])^2}{2N\text{Var}(E|Z)}.$$

When we fix Z, Kac-Rice formula still has the following ingredients:

— $\mathrm{Vol}_{N-1}(S_N) = (2\pi e)^{N/2}$;

— gradient density at 0, $\varphi_{\nabla_{\mathrm{sph}} H_N(x)}(0) \approx (2\pi\xi'(1))^{-N/2}$;

— density for Z, $\exp\left(-\frac{NZ^2}{2(\xi'(1)+\xi''(1))}\right)$;

— determinant

$$\mathbb{E}|\det(\sqrt{\xi''(1)}\,\mathrm{GOE}(N-1) - ZI_{N-1})| \approx \xi''(1)^{N/2}\exp\left(N\psi(Z/\sqrt{\xi''(1)})\right).$$

Therefore, in conclusion, we should have

$$\frac{1}{N}\log\mathbb{E}|\mathrm{Crt}(H_N, Z)| \approx \Phi_{\mathrm{rad}}(Z) := \frac{1}{2}\left(1+\log\frac{\xi''(1)}{\xi'(1)} - \frac{Z^2}{\xi''(1)+\xi'(1)}\right) + \psi\left(\frac{Z}{\sqrt{\xi''(1)}}\right). \tag{5}$$

Since $\psi(0) = -1/2$, we would have $\Phi_{\mathrm{rad}}(0) = \frac{1}{2}\log\frac{\xi''(1)}{\xi'(1)}$. Hence if $\xi''(1) > \xi'(1)$, we already have an exponentially large number of critical points in expectation from $Z \approx 0$, justifying the definition of the non-trivial phase. We focus on the latter case $\xi''(1) < \xi'(1)$ subsequently and show the number of total expected critical points is $e^{o(N)}$.

Besides the value at 0, since Φ_{rad} is even, we only need to investigate its property on the positive halfline. Due to

$$\psi'(t) = \mathrm{Re}\left(\frac{t-\sqrt{t^2-4}}{2}\right) = \begin{cases} t/2, & 0 \le t \le 2,\\ \dfrac{t-\sqrt{t^2-4}}{2}, & t \ge 2,\end{cases}$$

we can also obtain its derivative by

$$\Phi'_{\mathrm{rad}}(Z) = \begin{cases} Z\left(\dfrac{1}{2\xi''(1)} - \dfrac{1}{\xi'(1)+\xi''(1)}\right), & 0 \le Z \le 2\sqrt{\xi''(1)},\\ Z\left(\dfrac{1}{2\xi''(1)} - \dfrac{1}{\xi'(1)+\xi''(1)}\right) - \dfrac{\sqrt{Z^2-4\xi''(1)}}{2\xi''(1)}, & Z \ge 2\sqrt{\xi''(1)}.\end{cases}$$

Similarly, the second derivative is (now discontinuous at $2\sqrt{\xi''(1)}$ and):

$$\Phi''_{\mathrm{rad}}(Z) = \begin{cases} \left(\dfrac{1}{2\xi''(1)} - \dfrac{1}{\xi'(1)+\xi''(1)}\right), & 0 \le Z < 2\sqrt{\xi''(1)},\\ \left(\dfrac{1}{2\xi''(1)} - \dfrac{1}{\xi'(1)+\xi''(1)}\right) - \dfrac{Z}{2\xi''(1)\sqrt{Z^2-4\xi''(1)}}, & Z > 2\sqrt{\xi''(1)}.\end{cases}$$

Recall that $\Phi_{\mathrm{rad}}(Z) = \Phi_{\mathrm{rad}}(-Z)$. Since $\xi''(1) < \xi'(1)$, it is **convex** on $[-2\sqrt{\xi''(1)}, 2\sqrt{\xi''(1)}]$. Since $\sqrt{Z^2-4\xi''(1)} < |Z|$, it is **concave** outside $[-2\sqrt{\xi''(1)}, 2\sqrt{\xi''(1)}]$. Thus (see picture), if we find Z_* with

$$\Phi_{\mathrm{rad}}(Z_*) = \Phi'_{\mathrm{rad}}(Z_*) = 0$$

we will have shown that

$$\max_{Z\in\mathbb{R}} \Phi_{\mathrm{rad}}(Z) = 0.$$

Moreover the maximum will be achieved exactly at $\pm Z_*$, and we must have $Z_* > 2\sqrt{\xi''(1)}$. From this we will conclude (a weak form of) topological trivialization.

The solution turns out to be:

$$Z_* = \frac{\xi'(1)+\xi''(1)}{\sqrt{\xi'(1)}} = \sqrt{\xi''(1)}\cdot\left(a+\frac{1}{a}\right) > 2\sqrt{\xi''(1)}$$

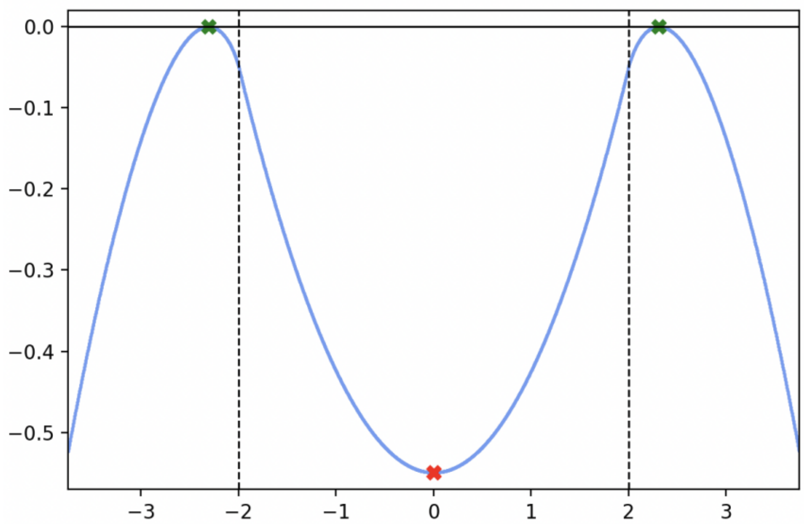

Figure 1: A typical diagram of Φ_{rad} , rescaled so $\xi''(1) = 1$.

for $a = \sqrt{\frac{\xi'(1)}{\xi''(1)}} > 1$. Note that

$$\sqrt{\left(a + \frac{1}{a}\right)^2 - 4} = a - \frac{1}{a} \quad \Longrightarrow \quad \sqrt{Z^2 - 4\xi''(1)} = \sqrt{\xi''(1)}\left(a - \frac{1}{a}\right) = \frac{\xi'(1) - \xi''(1)}{\sqrt{\xi'(1)}}. \tag{6}$$

Using this, it is easy to check that $\Phi'_{\text{rad}}(Z_*) = 0$.

To check $\Phi_{\text{rad}}(Z_*) = 0$ involves more miraculous cancellations (skipped in class). Let $Y_* = Z_*/\sqrt{\xi''(1)}$. Recalling (1), note that (6) gives

$$\log\left(\frac{\sqrt{Y_*^2 - 4} + |Y_*|}{2}\right) = \log\left(\sqrt{\frac{\xi'(1)}{\xi''(1)}}\right).$$

This cancels the log term in (5). The main remaining cancellation comes down to

$$\frac{(\xi'(1) - \xi''(1))^2}{4\xi'(1)\xi''(1)} - \frac{1}{2} - \frac{(\xi'(1) + \xi''(1)) \cdot (\xi'(1) - \xi''(1))}{4\xi'(1)\xi''(1)} = \frac{Z_*^2}{2(\xi'(1) + \xi''(1))} = \frac{\xi'(1) + \xi''(1)}{2\xi'(1)}.$$

In any case, we end up with the following theorem.

Theorem 2.2. *If $\xi''(1) < \xi'(1)$, then the following hold*

(a) $\mathbb{E}|\operatorname{Crt}_{S_N}(H_N)| = e^{o(N)}$;

(b) for any $\epsilon > 0$, expected number critical points with $Z \notin [-Z_* - \epsilon, -Z_* + \epsilon] \cup [Z_* - \epsilon, Z_* + \epsilon]$ is e^{-cN};

(c) for any $\epsilon > 0$, expected number critical points with $E \notin [-E_* - \epsilon, -E_* + \epsilon] \cup [E_* - \epsilon, E_* + \epsilon]$ is e^{-cN} where

$$E_* = \mathbb{E}[E|Z_*] = \sqrt{\xi'(1)}.$$

We end this lecture with following two corollaries.

Corollary 2.3. *If $\xi''(1) < \xi'(1)$, it holds that*

$$\operatorname*{p-lim}_{N\to\infty} \max_{x\in S_N} \frac{H_N(x)}{N} = \sqrt{\xi'(1)}.$$

Proof. The maximal value of H_N must occur at a critical point, so we must have with high probability:

$$\max_{x\in S_N} \frac{H_N(x)}{N} \approx \pm E_*.$$

We have the symmetry in distribution: $H_N \overset{d.}{=} -H_N$. This implies $\mathbb{E}\max_x \frac{H_N(x)}{N} \geq 0$. Consequently by concentration, the limiting maximum value must be E_*. □

Corollary 2.4. *If $\xi''(1) < \xi'(1)$, all critical points are local maxima or minima.*

Proof. By Theorem 2.2 and sign symmetry, it suffices to show the expected number of saddle points (i.e. critical points which are not local optima) with $Z \in [Z_* - \epsilon, Z_* + \epsilon]$ is at most $e^{-c(\xi,\epsilon)N}$, for small $\epsilon > 0$.

We use Proposition 3.3 from last lecture, with E replaced by Z. Analogously to (5), this gives the upper bound

$$\frac{1}{N}\log \mathbb{E}|\,\mathrm{Crt}^{\mathrm{saddle}}(H_N, Z \in [Z_*-\epsilon, Z_*+\epsilon])| \leq \sup_{Z\in[Z_*-\epsilon,Z_*+\epsilon]} \frac{1}{2}\left(1+\log\frac{\xi''(1)}{\xi'(1)} - \frac{Z^2}{\xi''(1)+\xi'(1)}\right) + \psi\left(\frac{Z}{\sqrt{\xi''(1)}}\right) - \frac{1}{2N}\log \mathbb{P}[I|Z].$$

Here I is the event of being a saddle point. Since we have shown trivialization, it suffices to show that

$$\inf_{Z\in[Z_*-\epsilon,Z_*+\epsilon]} \frac{1}{N}\log \mathbb{P}[I|Z] > c(\xi,\epsilon) > 0$$

for small enough ϵ depending on ξ. Indeed, since $Z_* > 2\sqrt{\xi''(1)}$, the event I holding requires the tangential Hessian

$$\nabla^2_{tan} H_N() \sim \sqrt{\xi''(1)}\, GOE(N-1)$$

to have an outlier eigenvalue, which has probability at most e^{-cN}. This concludes the proof. □

References

[1] David Belius, Jiri Cerny, Shuta Nakajima, and Marius A Schmidt. Triviality of the geometry of mixed p-spin spherical hamiltonians with external field. *Journal of Statistical Physics*, 186(1):12, 2022. 2