Statistics 291: Lecture 11 (February 27, 2024)

Bounds on $F_N(\beta)$ in the High-Temperature Case with External Field

Instructor: Mark Sellke

Scribe: Peter Luo

1 Introduction

Today we will introduce a new technique for studying models on the sphere: restricting H_N on subspherical bands. In particular, if one considers a band in S_N, the restriction of H_N will also be a spin glass, in one lower dimension.

In this class, we will concern ourselves with the high temperature case, with an external field applied. Next time we will cover shattering for pure p-spin models.

2 Recap

Consider the general p-spin model

$$H_N(x) = \sum_{p=1}^{P} \gamma_p H_{N,p}(x),$$

where

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

In the case of high-temperature with no external field, we have $\gamma_1 = 0$ and

$$\frac{1}{N}\log \mathbb{E} Z_N(\beta) = \beta^2\xi(1)/2, \qquad \frac{1}{N}\log \mathbb{E} Z_N(\beta)^2 = \max_{-1<R<1} \beta^2(\xi(1)+\xi(R)) + \frac{1}{2}\log(1-R^2).$$

Setting $\Phi(R) = \beta^2(\xi(1)+\xi(R)) + \frac{1}{2}\log(1-R^2)$, we can draw a rough sketch to see that the maximum of $\Phi(R)$ is attained at $R = 0$. This relies on $\gamma_1 = 0 \iff \xi'(0) = 0$ and the assumption $\beta \le \beta_0(\xi)$. Thus

$$\frac{1}{N}\log \mathbb{E} Z_N(\beta)^2 \approx \beta^2\xi(1).$$

We've computed the first and second moments of the partition function, the first moment exactly and the second moment to leading exponential order. By Paley-Zygmund and the Concentration of $F_N(\beta)$, we also obtained the bounds

$$\beta^2\xi(1)/2 - o(1) \le \mathbb{E} F_N(\beta) = \frac{1}{N}\mathbb{E}\log Z_N(\beta) \le \frac{1}{N}\log \mathbb{E} Z_N(\beta) = \beta^2\xi(1)/2.$$

However, if $\gamma_1 > 0$, then the maximum of $\Phi(R)$ will be achieved at some positive R, thus making it harder to find the second moment of $Z_N(\beta)$.

Another argument that the annealed free energy ought to be incorrect is as follows. Recall from lecture 4 that Gaussian integration by parts gives:

$$\frac{d}{d\beta}F_N(\beta) = \beta \cdot \left(\xi(1) - \mathbb{E}_{x,x'\sim\mu_\beta}[R(x,x')]\right).$$

This suggests that

$$F_N(\beta) \approx \beta^2\xi(1)/2 \iff \mathbb{E}_{x,x'\sim\mu_\beta}[\xi(R(x,x'))] \approx 0.$$

We would expect that with an external field, $R(x,x')$ is usually positive, so we should not expect the latter to be true.

3 Lower bound for $F_N(\beta)$

One easy lower bound is given by conditioning on $\mathbf{G}_N^{(1)}$ and considering a $1/N$-width neighborhood about its orthogonal complement. Then, $(\mathbf{G}_N^{(1)})^\perp$ satisfies

$$F_N(\beta) = F_\beta(H_N; S_N) \ge F_\beta(H_N; (\mathbf{G}_N^{(1)})^\perp) \ge \beta^2\xi_0(1)/2 - o(1),$$

where

$$\xi_0(R) = \sum_{p=2}^{P} \gamma_p^2 R^p = \xi(R) - \xi'(0)R$$

and $\beta \le \beta_0(\xi)$. However this bound is suboptimal.

A second approach would be to find a subsphere correlated with $\mathbf{G}_N^{(1)}$.

Definition 3.1. The band centered at x is

$$\mathrm{Band}(x) = \{y \in S_N : \langle x,x\rangle = \langle y,x\rangle\}.$$

By fixing $\mathbf{G}_N^{(1)}$, we have $\mathrm{Band}(\sqrt{q}\mathbf{G}_N^{(1)}) = \{\sqrt{q}\tilde{\mathbf{G}}_N^{(1)} + \sqrt{1-q}z : z \in S_N\}$, where $\tilde{\mathbf{G}}_N^{(1)} = \frac{\sqrt{N}\mathbf{G}_N^{(1)}}{\left\|\mathbf{G}_N^{(1)}\right\|}$ has norm $\sqrt{N}$. It follows that

$$F_\beta(H_N; S_N) \ge F_\beta(H_N; \mathrm{Band}(\sqrt{q}\mathbf{G}_N^{(1)})) = \frac{\beta}{N}H_N(\sqrt{q}\tilde{\mathbf{G}}_N^{(1)}) + F_\beta(H_N - H_N(\sqrt{q}\tilde{\mathbf{G}}_N^{(1)}); \mathrm{Band}(\sqrt{q}\mathbf{G}_N^{(1)})),$$

due to the identity

$$F_N(\beta) = \frac{1}{N}\log\int e^{\beta H_N(x)}\,dx = \beta H_N(x_*(q))/N + \frac{1}{N}\log\int e^{\beta(H_N(x) - H_N(x_*(q))}\,dx$$

for any x_*. The first term on the RHS is $\beta\sqrt{q}N$, but the second term is intractable, even though $\tilde{\mathbf{G}}_N^{(1)}$ will cancel out from subtraction. Ultimately, the reason we still cannot solve the second term is that $\nabla_{\mathrm{sph}}(H_N(\sqrt{q}\tilde{\mathbf{G}}_N^{(1)})) \ne 0$.

Lemma 3.2. *For* $y = \sqrt{q}\tilde{\mathbf{G}}_N^{(1)} + \sqrt{1-q}z,\ y' = \sqrt{q}\tilde{\mathbf{G}}_N^{(1)} + \sqrt{1-q}z'$, *where* $z, z' \in S_N$ *and* $x, x' \perp \sqrt{q}\tilde{\mathbf{G}}_N^{(1)}$,

$$\mathbb{E}[(H_N(y) - H_N(\sqrt{q}\tilde{\mathbf{G}}_N^{(1)}))(H_N(y') - H_N(\sqrt{q}\tilde{\mathbf{G}}_N^{(1)}))] = N(\xi(R(y,y') - \xi(q)) = N(\xi_0(q + (1-q)R(z,z')) - \xi_0(q)).$$

Proof. We have

$$\mathbb{E}H_N(y)H_N(y') = N\xi_0(R(y,y'))$$

and

$$R\left(y, \sqrt{q}\tilde{\mathbf{G}}_N^{(1)}\right) = R\left(\sqrt{q}\tilde{\mathbf{G}}_N^{(1)} y'\right) = R\left(\sqrt{q}\tilde{\mathbf{G}}_N^{(1)}, \sqrt{q}\tilde{\mathbf{G}}_N^{(1)}\right) = q.$$

□

Note that we can view $\tilde{R}(y, y') = R(z, z')$ as the "effective covariance on the band", so that

$$\tilde{\xi}(\tilde{R}) = \xi_0(q + (1-q)\tilde{R}) - \xi_0(q) \implies \tilde{\xi}'(0) = \xi_0'(q)(1-q).$$

It seems like we require two conditions to hold in order to find a good band centered at x:

- $R(x, \tilde{\mathbf{G}}_N^{(1)}) > 0$
- $\nabla_{\text{sph}}(H_N(\sqrt{q}\tilde{\mathbf{G}}_N^{(1)})) = 0.$

Denote $x_*(q) = \arg\max_{x \in \sqrt{q} S_N} H_N(x)$, so that

$$F_\beta(H_N; S_N) \geq F_\beta(H_N; \text{Band}(x_*(q)) = \frac{\beta}{N} H_N(x_*(q)) + F_\beta(H_N - H_N(x_*(q)); \text{Band}(x_*(q)). \tag{1}$$

Recall that for $\beta \leq \beta_0(\xi)$, $q \leq q_0(\xi)$, if $\xi'(1) > \xi''(1)$, then topologically trivially, the ground state energy satisfies GS $\approx \sqrt{\xi'(1)}$. We need this to see $\tilde{\xi}_q(R) = \xi(qR)$. In addition,

$$\tilde{\xi}_q'(1) = q\xi'(q) \geq q\xi'(0) = q\gamma_1^2 \geq \Omega(q), \qquad \tilde{\xi}_q''(1) = a^2\xi''(q) \leq O(q^2),$$

for $\gamma_1 > 0$. Equipped with these results, we can observe that the first term in the RHS of equation (1) is

$$\frac{\beta}{N} H_N(x_*(q)) = \beta\sqrt{q\xi'(q)},$$

due to the fact that if q is small and $\gamma_1 > 0$, so H_N is topologically trivial on S_N. We are now able to modify Lemma 4.1, replacing each $\sqrt{q}\tilde{\mathbf{G}}_N^{(1)}$ with $x_*(q)$:

Lemma 3.3. *For $y = \sqrt{q}x_*(q) + \sqrt{1-q}z$, $y' = \sqrt{q}x_*(q) + \sqrt{1-q}z'$, where $z, z' \in S_N$ and $z, z' \perp \sqrt{q}x_*(q)$,*

$$\mathbb{E}[(H_N(y) - H_N(x_*(q)))(H_N(y') - H_N(x_*(q)))] = \tilde{\xi}(\tilde{R}) - \xi'(0)\tilde{R}.$$

Since $\xi_q(\tilde{R})$ has no external field, then the second moment method works.

Theorem 3.4. $\liminf_{N\to\infty} F_N(\beta) \geq \max_{q \leq q_0(\xi)} \beta\sqrt{q\xi'(q)} + \frac{\beta^2}{2}\xi_q(1) + \frac{1}{2}\log(1-q)$.

Proof. Here is a sketch. We appeal to the idea of the proof of Kac-Rice. Because q is small, then we have topological trivialization on $\sqrt{q}S_N$, which implies

$$\frac{1}{N}\log \mathbb{E}|\text{Crt}_{\sqrt{q}S_N}(H_N)| \leq o(1).$$

Then,

$$\frac{1}{N}\log \mathbb{E}|\text{Crt}_{\sqrt{q}S_N}(H_N; I_\epsilon)| \leq -\delta \leq 0,$$

where

$$I_\epsilon(x) = \left\{\left|F_\beta(H_N - H_N(x), \text{Band}(x)) - \beta^2\xi_q(1)/2\right| \geq \epsilon\right\}.$$

Thus

$$\mathbb{P}[I_\epsilon(x) | \nabla_{\text{sph}}(H_N(x) = 0] \leq e^{-\delta N}.$$

In other words,

$$\mathbb{E}[\text{number of critical points } x \in \sqrt{q}S_N \text{ where } F_\beta(H_N; \text{Band}(x)) \text{ is too small}] \leq e^{-\delta N},$$

so Band$(x_*(q))$ has the correct free energy $\beta^2\xi_q(1)/2$. □

The moral of the story is that at high tempeartures with an external field,

- μ_β is replica-symmetric on a well-chosen band.
- Also in certain regimes for tensor PCA, non-spherical models.

4 Upper bound for $F_N(\beta)$

We now briefly outline an upper bound for the free energy. This is the replica-symmetric case of the Parisi formula.

Theorem 4.1. $F_\beta(\xi) \le \min_q \left(\frac{\beta^2}{2}(\xi(1) - \xi(q)) + \frac{q}{2(1-q)} + \frac{\log(1-q)}{2} \right)$

This matches the lower bound at $\beta \le \beta_0(\xi)$.

Proof. We interpolate $H_N \mapsto$ linear $L_N \sim \xi_{\mathrm{Lin}}(R) = \xi'(q)R$. Denoting $H_{N,\theta}(x) = \cos(\theta)H_N(x) + \sin(\theta)L_N(x)$, we have

$$\frac{d}{d\theta}\mathbb{E}F_{N,\theta} = \beta \sin\theta \cos\theta \mathbb{E}_{x,x' \sim \mu_{\beta,\theta}}[\xi(R(x,x')) - \xi(1) + \xi'(q)(1 - R(x,x'))].$$

The inner expression on the RHS is minimized at $R(x,x') = q$, (at least for ξ which is convex on $[-1,1]$). Thus it is lower-bounded by $\xi(q) - \xi(1) + \xi'(q)(1-q)$. Thus

$$F_\beta(\xi) \le \frac{\beta^2}{2}(\xi(1) - \xi(q) - (1-q)\xi'(q)) + \mathbb{E}F_\beta(L_N).$$

By some algebra, one can show that

$$\min_q \left(\frac{\beta^2}{2}(\xi(1) - \xi(q) - (1-q)\xi'(q)) + \mathbb{E}F_\beta(L_N) \right) = \min_q \left(\frac{\beta^2}{2}(\xi(1) - \xi(q)) + \frac{q}{2(1-q)} + \frac{\log(1-q)}{2} \right),$$

which completes the proof. □