STAT C206A / MATH C223A : Stein's method and applications Fall 2007

Lecture 23

Lecture date: Oct. 19, 2007 *Scribe: Partha Dey*

Recall that, in the Sherrington Kirkpatrick model, the probability of a configuration $\boldsymbol{\sigma} = (\sigma_i)_{i=1}^N \in \{-1,+1\}^N$ is

$$\mathbf{P}(\boldsymbol{\sigma}) = Z_N^{-1} \exp\bigg(\frac{\beta}{\sqrt{N}} \sum_{i<j} g_{ij}\sigma_i\sigma_j + h\sum_{i=1}^N \sigma_i \bigg)$$

where $(g_{ij})_{1\le i<j\le N}$ are i.i.d. standard gaussian random variables and

$$Z_N = \sum_{\boldsymbol{\sigma}\in\{-1,+1\}^N} \exp\bigg(\frac{\beta}{\sqrt{N}} \sum_{i<j} g_{ij}\sigma_i\sigma_j + h\sum_{i=1}^N \sigma_i \bigg)$$

is the normalizing constant. Suppose $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2$ are i.i.d. configurations from this Gibbs measure given $\boldsymbol{g} = (g_{ij})_{i<j}$. The overlap between $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2$ is defined as

$$R_{12} = \frac{1}{N}\sum_{i=1}^N \sigma_i^1\sigma_i^2.$$

Suppose $\langle\cdot\rangle$ denotes conditional expectation w.r.t. the Gibbs measure given $\boldsymbol{g}$ and ν denotes unconditional expectation, *i.e.* $\nu(f) = \mathbf{E}\langle f\rangle$. Then we have the following result.

Theorem 1 $\exists\ \beta_0 > 0$ *such that for all* $\beta \in [0,\beta_0]$ *and for all* h

$$\frac{\log Z_N}{N} \to \log 2 + \mathbf{E}\log\cosh(\beta z\sqrt{q} + h) + \frac{\beta^2(1-q)^2}{4}$$

where q *satisfies* $q = \mathbf{E}\tanh^2(\beta z\sqrt{q}+h)$ *and* $z \sim N(0,1)$.

Idea of the proof: Choose any arbitary number $q\in[0,1]$. Consider the alternative Gibbs measure $\propto \exp(\sum_{i=1}^N (\beta z_i\sqrt{q} + h)\sigma_i)$ where $z_1, z_2, \ldots, z_N$ are i.i.d. N(0, 1) random variables independent of $\boldsymbol{g}$. Let ν_0 be the unconditional law of this Gibbs measure. Note that σ_i's are independent under this Gibbs measure (both conditionally and unconditionally) and this measure is easier to handle. Also

$$\frac{\beta}{\sqrt{N}} \sum_{1\le i<j\le N} g_{ij}\sigma_i\sigma_j + h\sum_{i=1}^N \sigma_i = \sum_{i=1}^N \bigg(\frac{\beta}{2} l_i + h\bigg)\sigma_i$$

where $l_i = \frac{1}{\sqrt{N}} \sum_{j=1, j\neq i}^N g_{ij}\sigma_j$. The main idea is to show that for β sufficiently small, with a proper choice of q one can compare ν_0 and ν "in some sense". In the last lecture we proved that $N^{-1}(\log Z_N - \mathbf{E}\log Z_N) \to 0$ in probability. Today we'll prove that

$$\mathbf{E}\left(\frac{\log Z_N}{N}\right) \leq \log 2 + \mathbf{E}\log\cosh(\beta z\sqrt{q} + h) + \frac{\beta^2(1-q)^2}{4} \text{ for all } q \in [0,1], \beta \geq 0, h \in \mathbb{R}.$$

Lemma 2 (Gaussian Interpolation) *Suppose* $\boldsymbol{X} = (X_1, \ldots, X_n)$ *and* $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ *are two centered gaussian random vectors independent of each other. Let* $F : \mathbb{R}^n \to \mathbb{R}$ *be a* C^2 *function and let*

$$\varphi(t) = \mathbf{E}F(\sqrt{t}\boldsymbol{X} + \sqrt{1-t}\boldsymbol{Y}).$$

Then we have

$$\varphi'(t) = \frac{1}{2}\sum_{i,j=1}^n (\mathbf{E}(X_iX_j) - \mathbf{E}(Y_iY_j)) \cdot \mathbf{E}\left(\frac{\partial^2 F}{\partial x_i \partial y_j}(\sqrt{t}\boldsymbol{X} + \sqrt{1-t}\boldsymbol{Y})\right).$$

In particular we have

$$\mathbf{E}F(\boldsymbol{X}) - \mathbf{E}F(\boldsymbol{Y}) = \int_0^1 \varphi'(t)dt.$$

Proof: Exercise. □

For each $\boldsymbol{\sigma} \in \{-1, +1\}^N$, let

$$u_{\boldsymbol{\sigma}} = \frac{\beta}{\sqrt{N}}\sum_{i<j} g_{ij}\sigma_i\sigma_j \text{ and } v_{\boldsymbol{\sigma}} = \beta\sqrt{q}\sum_{i=1}^N z_i\sigma_i.$$

Then the normalizing constants in the S-K model and in the alternative model are

$$Z_N = \sum_{\boldsymbol{\sigma}} \exp(u_{\boldsymbol{\sigma}} + h\sum \sigma_i) \quad \text{and} \quad Z_N^0 = \sum_{\boldsymbol{\sigma}} \exp(v_{\boldsymbol{\sigma}} + h\sum \sigma_i)$$

respectively. So if we define a function $Z : \mathbb{R}^{\{-1,+1\}^N} \to \mathbb{R}$ as

$$Z(\boldsymbol{x}) = \sum_{\boldsymbol{\sigma} \in \{-1,+1\}^N} w_{\boldsymbol{\sigma}} \exp(x_{\boldsymbol{\sigma}})$$

where $\boldsymbol{x} = (x_{\boldsymbol{\sigma}})_{\boldsymbol{\sigma} \in \{-1,+1\}^N}$ and $w_{\boldsymbol{\sigma}} = \exp(h\sum_{i=1}^N \sigma_i)$, we have $Z_N = Z(\boldsymbol{u})$, $Z_N^0 = Z(\boldsymbol{v})$ where $\boldsymbol{u} = \{u_{\boldsymbol{\sigma}}\}_{\boldsymbol{\sigma} \in \{-1,+1\}^N}$ and $\boldsymbol{v} = \{v_{\boldsymbol{\sigma}}\}_{\boldsymbol{\sigma} \in \{-1,+1\}^N}$. Let

$$F(\boldsymbol{x}) = \frac{\log Z(\boldsymbol{x})}{N} \text{ for } \boldsymbol{x} \in \mathbb{R}^{\{-1,+1\}^N}$$

and $\varphi(t) = \mathbf{E}F(\sqrt{t}\boldsymbol{u} + \sqrt{1-t}\boldsymbol{v})$. We are interested in

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)dt.$$

Clearly we have

$$\begin{aligned}
\frac{\partial F}{\partial \boldsymbol{x}_{\boldsymbol{\sigma}}} &= \frac{\partial}{\partial \boldsymbol{x}_{\boldsymbol{\sigma}}}\left(\frac{\log Z(\boldsymbol{x})}{N}\right) = \frac{1}{NZ(\boldsymbol{x})} w_{\boldsymbol{\sigma}} \exp(x_{\boldsymbol{\sigma}}) \\
\text{and } \frac{\partial^2 F}{\partial \boldsymbol{x}_{\boldsymbol{\tau}} \partial \boldsymbol{x}_{\boldsymbol{\sigma}}} &= -\frac{1}{N(Z(\boldsymbol{x}))^2} w_{\boldsymbol{\tau}} w_{\boldsymbol{\sigma}} \exp(x_{\boldsymbol{\sigma}} + x_{\boldsymbol{\tau}}) + \frac{1}{NZ(\boldsymbol{x})} w_{\boldsymbol{\sigma}} \exp(x_{\boldsymbol{\sigma}}) \cdot \mathbf{1}_{\{\boldsymbol{\sigma}=\boldsymbol{\tau}\}}.
\end{aligned}$$

Let $U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) = \frac{1}{2}\mathbf{E}(u_{\boldsymbol{\sigma}^1} u_{\boldsymbol{\sigma}^2} - v_{\boldsymbol{\sigma}^1} v_{\boldsymbol{\sigma}^2})$. Then

$$\begin{aligned}
\varphi'(t) &= \sum_{\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2} U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \mathbf{E}\left(\frac{\partial^2 F}{\partial \boldsymbol{x}_{\boldsymbol{\sigma}^1} \partial \boldsymbol{x}_{\boldsymbol{\sigma}^2}} (\sqrt{t}\boldsymbol{u} + \sqrt{1-t}\boldsymbol{v}) \right) \\
&= \frac{1}{N} \sum_{\boldsymbol{\sigma}} U(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \frac{w_{\boldsymbol{\sigma}} \exp(\sqrt{t} u_{\boldsymbol{\sigma}} + \sqrt{1-t} v_{\boldsymbol{\sigma}})}{Z_t} \\
&\quad - \frac{1}{N} \sum_{\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2} U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \frac{w_{\boldsymbol{\sigma}^1} w_{\boldsymbol{\sigma}^2} \exp(\sqrt{t}(u_{\boldsymbol{\sigma}^1} + u_{\boldsymbol{\sigma}^2}) + \sqrt{1-t}(v_{\boldsymbol{\sigma}^1} + v_{\boldsymbol{\sigma}^2}))}{Z_t^2}
\end{aligned}$$

where $Z_t = \sum_{\boldsymbol{\sigma}} w_{\boldsymbol{\sigma}} \exp(\sqrt{t} u_{\boldsymbol{\sigma}} + \sqrt{1-t} v_{\boldsymbol{\sigma}})$. For each $t \in [0, 1]$ we have a gibbs measure $\propto \exp(\sqrt{t} u_{\boldsymbol{\sigma}} + \sqrt{1-t} v_{\boldsymbol{\sigma}} + h \sum \sigma_i)$ where $u_{\boldsymbol{\sigma}} = \frac{\beta}{\sqrt{N}} \sum_{i<j} g_{ij} \sigma_i \sigma_j$ and $v_{\boldsymbol{\sigma}} = \beta \sqrt{q} \sum_{i=1}^N z_i \sigma_i$. Let $\langle \cdot \rangle_t$ denote the expectation w.r.t. this gibbs measure. Let ν_t denote the unconditional expectation. Then

$$\varphi'(t) = \frac{1}{N}(\mathbf{E}\langle U(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \rangle_t - \mathbf{E}\langle U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) \rangle_t).$$

Now

$$\begin{aligned}
U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) &= \frac{\beta^2}{2N} \mathbf{E}\left((\sum_{i<j} g_{ij} \sigma_i^1 \sigma_j^1)(\sum_{i<j} g_{ij} \sigma_i^2 \sigma_j^2) \right) - \frac{\beta^2 q}{2} \mathbf{E}\left((\sum_{i=1}^N Z_i \sigma_i^1)(\sum_{i=1}^N Z_i \sigma_i^2) \right) \\
&= \frac{\beta^2}{2N} \sum_{i<j} \sigma_i^1 \sigma_i^2 \sigma_j^1 \sigma_j^2 - \frac{\beta^2 q}{2} \sum_{i=1}^N \sigma_i^1 \sigma_i^2 \\
&= \frac{\beta^2}{4N} \left((\sum_{i=1}^N \sigma_i^1 \sigma_i^2)^2 - N \right) - \frac{\beta^2 q}{2} \sum_{i=1}^N \sigma_i^1 \sigma_i^2 \\
&= \frac{\beta^2 N}{4} \left(R_{12}^2 - \frac{1}{N} \right) - \frac{\beta^2 q N}{2} R_{12} \\
\implies \frac{1}{N} U(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) &= \frac{\beta^2}{4} \left(R_{12}^2 - \frac{1}{N} \right) - \frac{\beta^2 q}{2} R_{12}.
\end{aligned}$$

Note that

$$\frac{1}{N}U(\boldsymbol{\sigma},\boldsymbol{\sigma}) = \frac{\beta^2}{4}\left(1-\frac{1}{N}\right) - \frac{\beta^2 q}{2}.$$

Now plugging in the values of $U(\boldsymbol{\sigma},\boldsymbol{\sigma}), U(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2)$ we have

$$\varphi'(t) = \left(\frac{\beta^2}{4} - \frac{\beta^2 q}{2}\right) - \left(\frac{\beta^2}{4}\mathbf{E}\langle R_{12}^2\rangle_t + \frac{\beta^2 q}{2}\mathbf{E}\langle R_{12}\rangle_t\right) = -\frac{\beta^2}{4}\mathbf{E}\langle (R_{12}-q)^2\rangle_t + \frac{\beta^2}{4}(1-q)^2.$$

This gives, in particular,

$$\varphi(1) \le \varphi(0) + \frac{\beta^2(1-q)^2}{4} \quad \forall\ 0 \le q \le 1.$$

Now note that

$$\begin{aligned}
\varphi(0) &= \frac{1}{N}\mathbf{E}\log\Bigg(\sum_{\boldsymbol{\sigma}\in\{-1,+1\}^N}\prod_{i=1}^{N}\exp((\beta z_i\sqrt{q}+h)\sigma_i)\Bigg)\\
&= \frac{1}{N}\mathbf{E}\log\prod_{i=1}^{N}(\exp(\beta z_i\sqrt{q}+h)+\exp(-\beta z_i\sqrt{q}-h))\\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbf{E}\log\left(2\cosh(\beta z_i\sqrt{q}+h)\right) = \log 2 + \mathbf{E}\log\cosh(\beta z\sqrt{q}+h)
\end{aligned}$$

where $z \sim N(0,1)$. So for any $0 \le q \le 1$, we have

$$\mathbf{E}\left(\frac{\log Z_N}{N}\right) \le \log 2 + \mathbf{E}\log\cosh(\beta Z\sqrt{q}+h) + \frac{\beta^2(1-q)^2}{4}.$$

This inequality is called **Guerra's inequality** and this holds for all $\beta \ge 0, h \in \mathbb{R}$.

Exercise 3 *Prove that the R.H.S. of Guerra's inequality is minimized when*

$$q = \mathbf{E}\tanh^2(\beta z\sqrt{q}+h).$$