MATH 585 – TOPICS IN MATHEMATICAL PHYSICS – FALL 2006
MATHEMATICS OF MEAN FIELD SPIN GLASSES AND THE REPLICA METHOD
LECTURE 9: QUADRATIC REPLICA COUPLING II, GUERRA AND TONINELLI'S THEOREM ON SUPERADDITIVITY OF PRESSURE

S. STARR
MATHEMATICS DEPARTMENT, UNIVERSITY OF ROCHESTER

CONTENTS

All Gaussians are always assumed to be centered.

1. REVIEW OF THE SHERRINGTON-KIRKPATRICK SPIN GLASS

In this lecture we will discuss Guerra and Toninelli's proof of the existence of the thermodynamic limit of the pressure for the Sherrington-Kirkpatrick model and many other spin glass models. Let us start by reminding ourselves of the definition of the Hamiltonian. As usual, $\Omega = \{+1, -1\}$ and Ω^N is the set of all spin configurations $\sigma = (\sigma_1, \ldots, \sigma_N)$ with each $\sigma_i \in \Omega$. The random Hamiltonian is

$$\mathsf{H}_N(\sigma; h) \;=\; -\frac{1}{\sqrt{2N}} \sum_{i,j=1}^{N} \mathsf{J}_{ij} \sigma_i \sigma_j - h \sum_{i=1}^{N} \sigma_i \,.$$

where $h \in \mathbb{R}$ is a nonrandom number and $(\mathsf{J}_{ij} \,:\, i, j = 1, \ldots, N)$ are i.i.d. $N(0,1)$ random variables. Therefore, $\mathsf{H}_N(\sigma; h)$ is, itself a random variable. Let us define $\mathsf{H}_N(\sigma) = \mathsf{H}_N(s; 0)$ to be the Hamiltonian without external magnetic field. This can potentially be very confusing, but as we will show, we can more easily treat the magnetic field separately by some tricks. The advantage is that the Hamiltonian without magnetic field,

$$\mathsf{H}_N(\sigma) \;=\; -\frac{1}{\sqrt{2N}} \sum_{i,j=1}^{N} \mathsf{J}_{ij} \sigma_i \sigma_j \,,$$

is a pure Gaussian random variable, or rather it is a Gaussian centered random variable. So, the joint distribution of the family $\mathsf{H}_N = (\mathsf{H}_N(\sigma) \,:\, \sigma \in \Omega^N)$ is completely determined by

Date: November 9, 2006.

the covariance. Let us calculate this.

$$\begin{aligned}
\mathbb{E}[\mathsf{H}_N(s)\mathsf{H}_N(\sigma')] &= \frac{1}{2N} \sum_{i,j,k,\ell=1}^{N} \mathbb{E}[\mathsf{J}_{ij}\mathsf{J}_{k\ell}]\sigma_i\sigma_j\sigma'_k\sigma'_\ell \\
&= \frac{1}{2N} \sum_{i,j,k,\ell=1}^{N} \delta_{ik}\delta_{j\ell}\sigma_i\sigma_j\sigma'_k\sigma'_\ell \\
&= \frac{1}{2N} \sum_{i,j=1}^{N} \sigma_i\sigma_j\sigma'_i\sigma'_j \\
&= \frac{1}{2N} \left(\sum_{i=1}^{N} \sigma_i\sigma'_i\right)^2 .
\end{aligned}$$

Recall that for the Curie-Weiss model, we had several different versions of the Hamiltonian, which either included diagonal terms $\sigma_i\sigma_i$ or only included off–diagonal terms $\sigma_i\sigma_j$ for $i < j$. In the form of the Curie-Weiss model that included diagonal terms, we could rewrite

$$H_N^{\mathrm{CW}}(\sigma) = \frac{N}{2} m_N(\sigma)^2 \quad \text{where} \quad m_N(\sigma) = \frac{1}{N}\sum_{i=1}^{N} \sigma_i .$$

This was crucial to the large-deviations approach to the Curie-Weiss model. In the version of the Sherrington-Kirkpatrick model we just wrote down we also include diagonal terms. This does not mean we are going to necessarily use large deviations to analyze it. But it is true that, defining

$$q_N(\sigma,\sigma') := \frac{1}{N}\sum_{i=1}^{N} \sigma_i\sigma'_i ,$$

we have

$$\mathbb{E}[\mathsf{H}_N(\sigma)\mathsf{H}_N(\sigma')] = \frac{N}{2} q_N(\sigma,\sigma')^2 .$$

Note that, since the SK Hamiltonian $\mathsf{H}_N = (\mathsf{H}_N(\sigma) : \sigma \in \Omega^N)$ is Gaussian, it is natural to work with the covariance as much as possible. This is a trivial comment, but one which seems worth making.

The function $q_N : \Omega^N \times \Omega^N \to [-1, 1]$ is called the spin-spin overlap. It has some simple features which we would like to explicitly state now. First of all, and most importantly,

$$q_N(\sigma,\sigma') = \frac{1}{N}\vec{\sigma}\cdot\vec{\sigma}' ,$$

where we put arrows over the spin configurations to remind ourselves that $\Omega^N \subset \mathbb{R}^N$. We could define the scaled Euclidean distance

$$d_N(\sigma,\sigma') = \frac{1}{2N}\|\vec{\sigma} - \vec{\sigma}'\|_2 = 1 - q_N(\sigma,\sigma') .$$

Of course, it is also the case that $(q_N(\sigma,\sigma') : \sigma,\sigma' \in \Omega^N)$ is a positive semidefinite kernel, simply because it is the covariance of H_N. For later reference, let us observe one more simple fact. If $N = N_1 + N_2$ for $N_1, N_2 \geq 1$, then

$$q_N(\sigma, \tilde{\sigma}) \,=\, \frac{N_1}{N} q_{N_1}(\sigma^{(1)}, \tilde{\sigma}^{(1)}) + \frac{N_2}{N} q_{N_2}(\sigma^{(2)}, \tilde{\sigma}^{(2)})\,, \tag{1.1}$$

where

$$\sigma^{(1)} \,=\, (\sigma_1, \ldots, \sigma_{N_1}) \quad \text{and} \quad \sigma^{(2)} \,=\, (\sigma_{N_1+1}, \ldots, \sigma_{N_1+N_2})\,,$$

and similarly for $\tilde{\sigma}^{(1)}$ and $\tilde{\sigma}^{(2)}$, defined relative to $\tilde{\sigma}$. This is important to Guerra and Toninelli's convexity argument.

The real quantity of interest is the random pressure. First, define the random partition function

$$\mathsf{Z}_N(\beta) \,=\, \sum_{\sigma \in \Omega^N} e^{-\beta \mathsf{H}_N(\sigma)}\,.$$

If we want to include the external magnetic field, then we can. Let us use the parameter x in place of βh. Then we can define

$$\mathsf{Z}_N(\beta, x) \,=\, \sum_{\sigma \in \Omega^N} w_N(\sigma; x)\, e^{-\beta \mathsf{H}_N(\sigma)}\,,$$

where $(w_N(\sigma; x)\, : \sigma \in \Omega^N)$ is a weight-factor

$$w_N(\sigma; x) \,=\, e^{x \sum_{i=1}^N \sigma_i} \,=\, \prod_{i=1}^N e^{x\sigma_i}$$

Note that this factorizes. In particular, if $N = N_1 + N_2$, as we considered before,

$$w_N(\sigma; x) \,=\, w_{N_1}(\sigma^{(1)}; x) w_{N_2}(\sigma^{(2)}; x)\,.$$

The random pressure is

$$\mathsf{p}_N(\beta, x) \,=\, \frac{1}{N} \log\left(\mathsf{Z}_N(\beta, x)\right)\,.$$

We can define $\mathsf{p}_N(beta) \,=\, \mathsf{p}_N(\beta, 0)$. Also, we define the *"quenched pressure"* to be the expectation

$$p_N(\beta, x) \,=\, \mathbb{E}[\mathsf{p}_N(\beta, x)]\,,$$

which is nonrandom. Guerra and Toninelli proved that the following limit exists,

$$p(\beta, x) \,:=\, \lim_{N \to \infty} p_N(\beta, x)\,.$$

This is the thermodynamic limit of the quenched pressure. In the next lecture we will discuss what is called "self-averaging" of the random pressure, which proves that $|\mathsf{p}_N(\beta, x) - p_N(\beta, x)|$ converges to 0, in the limit $N \to \infty$, in distribution (and in L^2). Somewhat surprisingly, the "self-averaging" property was proved, rigorously in 1991, long before the existence of the thermodynamic limit of the quenched pressure. (Even then, it was a footnote on a paper proving a more complicated result, and the authors, Pastur and Shcherbina claimed to be surprised that the easy proof had not been found before.) The reason is that, using the martingale method it is easier to handle the fluctuations of $\mathsf{p}_N(\beta, h)$ than it is to hand the expectation of $\mathsf{p}_N(\beta, h)$, itself.

2. Guerra and Toninelli's Theorem for the SK Model

Let us state the main theorem from [1], specialized to the SK model.

Theorem 2.1 *For each $\beta \in [0,\infty)$ and $x \in \mathbb{R}$, there is superadditivity of $(N p_N(\beta,x) : N \in \mathbb{Z}_{>0})$. Namely, if $N_1, N_2 \geq 1$ then*

$$(N_1+N_2)\, p_{N_1+N_2}(\beta,x) \;\geq\; N_1\, p_{N_1}(\beta,x) + N_2\, p_{N_2}(\beta,x)\,.$$

Before proving this theorem, let us note the consequence. By Fekete's lemma, applied to the subadditive sequence, $-N\, p_N(\beta,x)$, we see that

$$p(\beta,x) \;=\; \lim_{N\to\infty} p_N(\beta,x)\,,$$

does exist, and it equals

$$p(\beta,x) \;=\; \sup_{N\geq 1} p_N(\beta,x) \in \mathbb{R}\cup\{+\infty\}\,.$$

In order to prove that $p(\beta,x) \neq +\infty$, let us recall, from Lecture 2, the definition of the "annealed pressure",

$$p_N^{\mathrm{A}}(\beta,x) \;:=\; \frac{1}{N}\log\left(\mathbb{E}[\mathsf{Z}_N(\beta,x)]\right)\,.$$

Recall that, by Jensen's inequality,

$$p_N(\beta,x) \;\leq\; p_N^{\mathrm{A}}(\beta,x)\,.$$

Also, recall that, by an explicit calculation (using the moment generating function for Gaussians), we have

$$p_N^{\mathrm{A}}(\beta,x) \;=\; \frac{\beta^2}{2} \;+\; \log(\cosh(2x))\,,$$

independent of N. (When we did the calculation before, there was a prefactor of $\frac{N}{N-1}$ because we considered the version of the SK model that did not include diagonal terms. But it is trivial to derive this identity from the previous one.) Therefore there is an upper bound for all $p_N(\beta,x)$ which is uniform in N, so the limit cannot be $+\infty$.

Proof. As in the statement of the lemma, let N_1 and N_2 be integers ≥ 1, and let $N = N_1+N_2$. Let $\mathbb{R}^{\Omega^N}$ be the set of all vectors $v = (v(\sigma) : \sigma \in \Omega^N)$. Define $\psi : \mathbb{R}^{\Omega^N} \to \mathbb{R}$ by

$$\psi(v;\beta,x) \;=\; \log\left(\sum_{\sigma\in\Omega^N} w_N(\sigma;x)\, e^{-\beta v(\sigma)}\right)\,.$$

Note that, defining $\mathsf{H}_N = (\mathsf{H}_N(\sigma) : \sigma \in \Omega^N)$, we have

$$\psi(\mathsf{H}_N;\beta,x) \;=\; N\,\mathsf{p}_N(\beta,x)\,.$$

Therefore,

$$p_N(\beta,x) \;=\; \mathbb{E}\left[\psi(\mathsf{H}_N;\beta,x)\right]\,.$$

This puts us in a position to apply the generalized Slepian's lemma for the function $\psi(\cdot;\beta,x)$ applied to the Gaussian family H_N, if we can find another Gaussian family, and verify all the hypotheses of the theorem. We do have the conditions for the sign of the mixed partial derivatives of $\psi(\cdot;\beta,x)$,

$$\frac{\partial^2}{\partial v(\sigma)\partial v(\sigma')}\psi(v;\beta,x) \,=\, -\beta^2\, \frac{w_N(\sigma)w_N(\sigma')e^{-\beta[v(\sigma)+v(\sigma')]}}{\left(\sum_{\sigma''\in\Omega^N} w_N(\sigma'')e^{-betav(\sigma'')}\right)^2}\,,$$

as long as $\sigma \neq \sigma'$. This is obviously nonpositive (which uses in part that all the $w_N(\sigma)$ are nonnegative). Suppose we find another Gaussian process $\mathsf{Y} = (\mathsf{Y}(\sigma)\,:\,\sigma\in\Omega_N)$ such that

$$\mathbb{E}[\mathsf{Y}(\sigma)^2] \,=\, \mathbb{E}[\mathsf{H}_N(\sigma)^2]$$

for all $\sigma \in \Omega_N$, and

$$\mathbb{E}[\mathsf{Y}(\sigma)\mathsf{Y}(\sigma')] \,\geq\, \mathbb{E}[\mathsf{H}_N(\sigma)\mathsf{H}_N(\sigma')]\,,$$

for all $\sigma \neq \sigma'$. Then, by the generalized Slepian's lemma, with the subset of the index set equal to $A = \{(\sigma,\sigma)\,:\,\sigma\in\Omega^N\}$, we will have

$$\mathbb{E}[\psi(\mathsf{H}_N;\beta,x)] \,\geq\, \mathbb{E}[\psi(\mathsf{Y};\beta,x)]\,.$$

In other words, $Np_N(\beta,x) \geq \mathbb{E}[\psi(\mathsf{Y};\beta,x)]$. With the same decomposition of $\sigma\in\Omega^N$ into $\sigma^{(1)}\in\Omega^{N_1}$ and $\sigma^{(2)}\in\Omega^{N_2}$ that we did before, define

$$\mathsf{Y}(\sigma) \,=\, \mathsf{H}'_{N_1}(\sigma^{(1)}) + \mathsf{H}''_{N_2}(\sigma^{(2)})\,,$$

where we think of H_N, H'_{N_1} and H''_{N_2} as all being independent of one another, but having the correct covariance for the SK Hamiltonian. Then

$$\mathbb{E}[\mathsf{Y}(\sigma)\mathsf{Y}(\tilde{\sigma})] \,=\, \frac{N_1}{2}q_{N_1}(\sigma^{(1)},\tilde{\sigma}^{(1)})^2 + \frac{N_2}{2}q_{N_2}(\sigma^{(2)},\tilde{\sigma}^{(2)})^2\,.$$

Using equation (1.1), we see that

$$\mathbb{E}[\mathsf{Y}(\sigma)\mathsf{Y}(\tilde{\sigma})] \,\geq\, \mathbb{E}[\mathsf{H}_N(\sigma)\mathsf{H}_N(\tilde{\sigma})]\,,$$

is true for all $\sigma,\sigma'\in\Omega^N$, either by completing-the-square, or else by just using convexity of the map $q\mapsto q^2$. Moreover, if $\sigma=\sigma'$, we have

$$q_N(\sigma,\sigma) \,=\, q_{N_1}(\sigma^{(1)},\sigma^{(1)}) \,=\, q_{N_2}(\sigma^{(2)},\sigma^{(2)}) \,=\, 1\,.$$

The self-overlap is always 1. So in this case, we do have

$$\mathbb{E}[\mathsf{Y}(\sigma)^2] \,=\, \mathbb{E}[\mathsf{H}_N(\sigma)^2]$$

Therefore, the general Slepian's lemma applies,

$$Np_N(\beta,x) \,\geq\, \mathbb{E}[\psi(\mathsf{Y};\beta,x)]\,.$$

But now we observe that, by independence, and factorization of $w_N(\sigma;x) = w_{N_1}(\sigma^{(1)};x) w_{N_2}(\sigma^{(2)};x)$, we have

$$\begin{aligned}
\psi(\mathsf{Y};\beta,x) &= \log\left[\sum_{\sigma\in\Omega^N} w_N(\sigma;x)\exp\left(-\beta\left[\mathsf{H}'_{N_1}(\sigma^{(1)}) + \mathsf{H}''_{N_2}(\sigma^{(2)})\right]\right)\right] \\
&= \log\left[\sum_{\sigma^{(1)}\in\Omega^{N_1}} w_{N_1}(\sigma^{(1)};x)\exp\left(\beta\mathsf{H}'_{N_1}(\sigma^{(1)})\right)\right] \\
&\quad + \log\left[\sum_{\sigma^{(2)}\in\Omega^{N_2}} w_{N_2}(\sigma^{(2)};x)\exp\left(\beta\mathsf{H}'_{N_2}(\sigma^{(2)})\right)\right] \\
&= N_1\,\mathsf{p}'_{N_1}(\beta,x) + N_2\,\mathsf{p}''_{N_2}(\beta,x)\,,
\end{aligned}$$

where $\mathsf{p}_N(\beta,x)$, $\mathsf{p}_{N_1}(\beta,x)$ and $\mathsf{p}_{N_2}(\beta,x)$ are independent random pressures. (Note that this would be of potential interest, if, for example N_1 and N_2 were equal so that some confusion could arise about the notation of $\mathsf{p}_{N_1}(\beta,x)$ and $\mathsf{p}_{N_2}(\beta,x)$ without the primes.) Therefore,

$$\begin{aligned}
Np_N(\beta,x) &\geq \mathbb{E}[\psi(\mathsf{Y};\beta,x)] \\
&= \mathbb{E}\left[N_1\,\mathsf{p}'_{N_1}(\beta,x) + N_2\,\mathsf{p}''_{N_2}(\beta,x)\right] \\
&= N_1, p_{N_1}(\beta,x) + N_2\,p_{N_2}(\beta,x)\,,
\end{aligned}$$

as claimed. □

Remark 2.2 Guerra and Toninelli's theorem applies to many other spin glass models as well. These include the p-spin models for even integers $p \geq 0$, and Derrida's Random Energy Model. Let us delay the discussion of the extension until after we have introduced those models, in a later lecture. Once the definitions have been made, it will be trivial to see how the generalized proof goes through, using just the argument from above, and in particular, convexity.

References

[1] F. Guerra and F. L. Toninelli. The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.* **230** (2002), 71–79.

Mathematics Department, University of Rochester, Rochester, NY 14627
E-mail address: sstarr@math.rochester.edu