MATH 585 – TOPICS IN MATHEMATICAL PHYSICS – FALL 2006
MATHEMATICS OF MEAN FIELD SPIN GLASSES AND THE REPLICA METHOD
LECTURE 11: THE POSITIVE INTEGER MOMENTS OF THE RANDOM PARTITION FUNCTION

S. STARR
MATHEMATICS DEPARTMENT, UNIVERSITY OF ROCHESTER

CONTENTS

1. MOTIVATION

Recall that, for $N > 0$, the Sherrington-Kirkpatrick Hamiltonian can be considered as a Gaussian family $(\mathsf{H}_N(\sigma) \,:\, \sigma \in \Omega^N)$ with covariance given by

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

where $q_N(\sigma,\sigma') = \frac{1}{N}\sum_{i=1}^N \sigma_i\sigma_i'$ is the spin-spin overlap. Also, the random partition function is defined as

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

where x stands in for βh, and

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

So, $w_N(\sigma;x)$ is the part of the Gibbs weight due to a pure external magnetic field. The random pressure is

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

and the quenched pressure is

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

Date: December 4, 2006.

We have now reviewed two important theorems about the random pressure and the quenched pressure. First, we reviewed a theorem of Guerra and Toninelli which proved that the thermodynamic limit for the quenched pressure exists,

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

Second, we reviewed a concentration-of-measure inequality, proved by Talagrand, who also used it to prove that

$$\mathbb{P}\{|\mathsf{p}_N(\beta, x) - p_N(\beta, x)| \;>\; t\} \;\leq\; 2e^{-2Nt^2/\beta^2}\,.$$

This bound is very strong. One implication of it is that

$$\mathscr{D} - \lim_{N \to \infty} \mathsf{p}_N(\beta, x) \;=\; p(\beta, x)\,.$$

I.e., the random variables, $\mathsf{p}_N(\beta, x)$, converge, in distribution, to a nonrandom limit, $p_N(\beta, x)$.

Given these two results, one would say that the next direction to investigate is an explicit formula for $p(\beta, x)$, the thermodynamic limit of the quenched pressure.

2. Positive Integer Moments

It turns out to be rather difficult to calculate

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

directly. On the other hand, a much easier question is to calculate

$$m_N(\beta, x; n) \;:=\; \mathbb{E}\left[\big(\mathsf{Z}_N(\beta, x)\big)^n\right]\,,$$

for $n = 0, 1, 2, \dots$. In this lecture we will state what the answer is. In the next lecture we will give the physicists' method of extrapolating $p(\beta, x)$ from the answer, although it is wrong. (Then we'll state another physicist's "fix" for the problem, which turns out to be right, although proving it is very difficult.)

Lemma 2.1 *For each $n = 0, 1, 2, \dots,$*

$$\begin{aligned} m_N(\beta, x; n) \;&:=\; \mathbb{E}\left[\big(\mathsf{Z}_N(\beta, x)\big)^n\right] \\ &= \sum_{\sigma^{(1)}, \dots, \sigma^{(n)} \in \Omega^N} \left[\prod_{a=1}^n w_N(\sigma^{(a)}; x)\right] \exp\left(\frac{\beta^2 N}{4} \sum_{a,b=1}^n [q_N(\sigma^{(a)}, \sigma^{(b)})]^2\right)\,. \end{aligned}$$

Proof. Write the nth power of $\mathsf{Z}_N(\beta, x)$ as an n-fold sum over Ω:

$$\begin{aligned} \big(\mathsf{Z}_N(\beta, x)\big)^n \;&=\; \left(\sum_{\sigma \in \Omega^N} w_N(\sigma; x)\, e^{-\beta \mathsf{H}_N(\sigma)}\right)^n \\ &= \sum_{\sigma^{(1)}, \dots, \sigma^{(n)}} \left[\prod_{a=1}^n w_N(\sigma^{(a)}; x)\right] e^{-\beta[\mathsf{H}_N(\sigma^{(1)}) + \dots + \mathsf{H}_N(\sigma^{(n)})]}\,. \end{aligned}$$

Since the random variable $\mathsf{H}_N(\sigma^{(1)}) + \cdots + \mathsf{H}_N(\sigma^{(n)})$ is a Gaussian, by the usual calculation

$$\mathbb{E}\left[e^{-\beta[\mathsf{H}_N(\sigma^{(1)})+\cdots+\mathsf{H}_N(\sigma^{(n)})]}\right] = \exp\left(\frac{\beta^2}{2}\mathrm{Var}\big(\mathsf{H}_N(\sigma^{(1)}) + \cdots + \mathsf{H}_N(\sigma^{(n)})\big)\right).$$

But the variance is exactly

$$\mathrm{Var}\big(\mathsf{H}_N(\sigma^{(1)})+\cdots+\mathsf{H}_N(\sigma^{(n)})\big) = \sum_{a,b=1}^{n} \mathbb{E}\left[\mathsf{H}_N(\sigma^{(a)})\mathsf{H}_N(\sigma^{(b)})\right] = \frac{N}{2}\sum_{a,b=1}^{n}[q_N(\sigma^{(a)},\sigma^{(b)})]^2 .$$

□

The formula in the lemma is not very explicit, because we still have to do the summation that is left. On the other hand, it is explicit enough to allow us to calculate the $N \to \infty$ limit in some sense.

3. Large Deviation Techniques

Let us define

$$\mathcal{P}_N(\beta, x; n) = \frac{1}{N}\log\big(m_N(\beta, x; n)\big).$$

Then the we can calculate the $N \to \infty$ limit of this.

Definition 3.1 *For each $n = 1, 2, \ldots$, define the set of $n \times n$ "overlap kernels" to be the set of real, $n \times n$ matrices, q, such that $q_{a,a} = 1$ for all $1 \le a \le n$, and such that q is positive semidefinite.*

Lemma 3.2 *For each $n = 0, 1, 2, \ldots$,*

$$\begin{aligned}\mathcal{P}(\beta, x; n) &:= \lim_{N\to\infty} \mathcal{P}_N(\beta, x; n)\\ &= \sup_{q\in\mathrm{Ov}(n)}\ \inf_{\lambda\in M_n(\mathbb{R})}\left[\sum_{a,b=1}^{n}\left(\frac{\beta^2}{4}q_{a,b}^2 - q_{a,b}\lambda_{a,b}\right) + \log\left(\sum_{\sigma\in\Omega^n} e^{x\sum_{a=1}^n \sigma_a + \sum_{a,b=1}^n \lambda_{a,b}\sigma_a\sigma_b}\right)\right].\end{aligned}$$

Proof. This is an elementary application of the large deviation techniques we learned: namely Varadhan's lemma and the discrete version of Cramér's theorem or Sanov's theorem. We want to calculate the limit of

$$\frac{1}{N}\log\left(\sum_{\sigma^{(1)},\ldots,\sigma^{(n)}\in\Omega^N}\left[\prod_{a=1}^{n} w_N(\sigma^{(a)}; x)\right]\exp\left(\frac{\beta^2 N}{4}\sum_{a,b=1}^{n}[q_N(\sigma^{(a)},\sigma^{(b)})]^2\right)\right).$$

If we take out a normalization, we get

$$\begin{aligned}&\frac{1}{N}\log\left(\sum_{\sigma^{(1)},\ldots,\sigma^{(n)}\in\Omega^N}\left[\prod_{a=1}^{n} w_N(\sigma^{(a)}; x)\right]\right)\\ &\qquad + \frac{1}{N}\log\left(\sum_{\sigma^{(1)},\ldots,\sigma^{(n)}\in\Omega^N}\left[\prod_{a=1}^{n} \widehat{w}_N(\sigma^{(a)}; x)\right]\exp\left(\frac{\beta^2 N}{4}\sum_{a,b=1}^{n}[q_N(\sigma^{(a)},\sigma^{(b)})]^2\right)\right).\end{aligned}$$

where

$$\widehat{w}_N(\sigma;x) \;=\; \frac{\widehat{w}_N(\sigma;x)}{\sum_{\sigma\in\Omega^N}\widehat{w}_N(\sigma;x)} \;=\; \prod_{i=1}^N \left[\frac{e^{x\sigma_i}}{2\cosh(x)}\right].$$

This is still a product measure. Moreover, the second sum can now be rewritten as

$$\frac{1}{N}\log\left(\mathbf{E}^{P_N}\left[\exp\left(N\cdot\frac{\beta^2}{4}\sum_{a,b=1}^n [q_{a,b}]^2\right)\right]\right),$$

where P_N is the induced measure on $M_n(\mathbb{R})$ obtained by taking the distribution of $(q_N(\sigma^{(a)},\sigma^{(b)}) : 1 \le a,b \le N)$ with the $\prod_{a=1}^n \widehat{w}_N(\sigma^{(a)};x)$ measure for $(\sigma^{(1)},\dots,\sigma^{(N)})$. By Varadhan's lemma,

$$\lim_{N\to\infty}\frac{1}{N}\log\left(\mathbf{E}^{P_N}\left[\exp\left(N\cdot\frac{\beta^2}{4}\sum_{a,b=1}^n [q_{a,b}]^2\right)\right]\right) \;=\; \sup_{q\in M_n(\mathbb{R})}\left[\frac{\beta^2}{4}\sum_{a,b=1}^n [q_{a,b}]^2 - I(q)\right],$$

where $I(q)$ is the large deviation rate function for the limits of the measures $P_1, P_2, \dots$.

Technically, Varadhan's lemma only applies when the range for q is compact. That is not an issue here, though, because $I(q)$ will be infinite unless q is in the support of some of the P_N's. The support of every P_N is a subset of $\mathrm{Ov}(n)$. The reason is that, no matter what random $(\sigma^{(1)},\dots,\sigma^{(N)})$ we have, we know that $(q_N(\sigma^{(a)},\sigma^{(b)}) : 1 \le a,b \le N) \in \mathrm{Ov}(n)$. Indeed, since $q_N(\sigma^{(a)},\sigma^{(b)})$ is just the dot-product, normalized by $1/N$, we obtain a "Gram matrix" which is always positive semidefinite. The diagonals are 1, just because $q_N(\sigma,\sigma) = 1$ for all $\sigma \in \Omega^N$. By Cauchy-Schwarz, we also know $|q_{a,b}| \le |q_{a,a}|^{1/2}|q_{b,b}|^{1/2} = 1$. Therefore, all matrix entries are bounded between -1 and 1. So $\mathrm{Ov}(n)$ is compact.

All that remains is to calculate $I(q)$. But

$$q_N(\sigma^{(a)},\sigma^{(b)}) \;=\; \frac{1}{N}\sum_{i=1}^N (\sigma_i^{(a)}\sigma_i^{(b)}).$$

So this can be viewed as the average of random variables $\sigma_i^{(a)}\sigma_i^{(b)}$ for i ranging from 1 to N. Moreover, the measure $\prod_{a=1}^n \widehat{w}_N(\sigma^{(a)};x)$ is a product measure in i. So it is actually an average of i.i.d. random variables. Therefore, by Cramér's theorem, we know

$$I(q) \;=\; \inf_{\lambda\in M_n(\mathbb{R})}\left[\sum_{a,b=1}^n \lambda_{a,b}q_{a,b} - \Lambda(\lambda)\right],$$

where Λ is the logarithmic moment generating function of P_1. But now, putting the constant normalization back in, we see that

$$\begin{aligned}\Lambda(\lambda) + n\log(2\cosh(x)) \;&=\; \mathbf{E}^{P_1}\left[e^{\sum_{a,b=1}^n \lambda_{a,b}q_{a,b}}\right] + n\log(2\cosh(x))\\ &=\; \log\left(\sum_{\sigma\in\Omega^n} e^{x\sum_{a=1}^n \sigma_a + \sum_{a,b=1}^n \lambda_{a,b}\sigma_a\sigma_b}\right).\end{aligned}$$

The normalization we added back in is precisely the normalization we took out originally to go from $w_N(\sigma;x)$ to $\widehat{w}_N(\sigma;x)$. So we are done. □

4. Duality for the Legendre-Fenchel Transform

One can simplify the formula from the last lemma. To get the simpler formula, we use the following elementary result about Legendre-Fenchel transforms. Let us simply quote the desired result from Dembo and Zeitouni's book [1]. (See Lemma 4.5.8 on pg. 152 of their book.)

Lemma 4.1 (Duality lemma) *Let $\mathcal{X}$ be a locally convex Hausdorff topological vector space. Let $f : \mathcal{X} \to (-\infty, \infty]$ be a lower semicontinuous, convex function, and define*

$$g(\lambda) \;=\; \sup_{x \in \mathcal{X}} \{\langle \lambda, x\rangle - f(x)\}\,.$$

Then $f(\cdot)$ is the Fenchel-Legendre transform of $g(\cdot)$, i.e.,

$$f(x) \;=\; \sup_{\lambda \in \mathcal{X}^*} \{\langle \lambda, x\rangle - g(\lambda)\}\,.$$

We will not prove the lemma here. Please refer to Dembo and Zeitouni's textbook. Instead, we will use the lemma to prove the final formula for $\mathcal{P}(\beta, x; n)$.

Lemma 4.2 *For each $n = 0, 1, 2, \ldots,$*

$$\mathcal{P}(\beta, x; n) \;=\; \sup_{q \in \mathrm{Ov}(n)} \left[-\frac{\beta^2}{4} \sum_{a,b=1}^{n} q_{a,b}^2 + \log\left(\sum_{\sigma \in \Omega^n} e^{x \sum_{a=1}^{n} \sigma_a + \frac{\beta^2}{2} \sum_{a,b=1}^{n} q_{a,b} \sigma_a \sigma_b}\right)\right]\,.$$

Proof. First of all, by making a transformation $q \to \frac{\sqrt{2}}{\beta} q$ and $\lambda \to \frac{\beta}{\sqrt{2}}\lambda$, the old formula for $\mathcal{P}(\beta, x; n)$ is

$$\mathcal{P}(\beta, x; n) \;=\; \sup_{q \in \frac{\beta}{2}\,\mathrm{Ov}(n)}\; \inf_{\lambda \in M_n(\mathbb{R})} \left[\frac{q^2}{2} - \langle q, \lambda\rangle + f(\lambda)\right]\,,$$

where $f(\lambda) = \Lambda(\frac{\beta}{\sqrt{2}}\lambda) + n \log(2\cosh(x))$, which is clearly convex (because all logarithmic moment generating functions are) and lower semicontinuous. The inner-product is the natural one, $\langle q, \lambda\rangle = \sum_{a,b=1}^{n} q_{a,b}\lambda_{a,b}$. Also, we write q^2 as a short notation for $\langle q, q\rangle$.

Now, choosing $\lambda = q$, we certainly see that

$$\inf_{\lambda \in M_n(\mathbb{R})} \left[\frac{q^2}{2} - \langle q, \lambda\rangle + f(\lambda)\right] \;\le\; -\frac{q^2}{2} + f(q)\,.$$

Therefore,

$$\mathcal{P}(\beta, x; n) \;\le\; \sup_{q \in \frac{\beta}{\sqrt{2}}\,\mathrm{Ov}(n)} \left[-\frac{q^2}{2} + f(q)\right]\,.$$

But on the other hand, for any $r \in M_n(\mathbb{R})$,

$$\frac{q^2}{2} \;\ge\; -\frac{r^2}{2} + \langle q, r\rangle\,.$$

Also, instead of restricting q to $\beta^{-1}\,\mathrm{Ov}(n)$, we could allow it to range over all of $M_n(\mathbb{R})$. It just means that for the new values of q, the infimum in λ will give $-\infty$, which does not affect the supremum in q. Doing this, we have

$$\begin{aligned}
\mathcal{P}(\beta,x;n) &= \sup_{q\in M_n(\mathbb{R})}\inf_{\lambda\in M_n(\mathbb{R})}\left[\frac{q^2}{2}-\langle q,\lambda\rangle+f(\lambda)\right]\\
&\geq -\frac{r^2}{2}+\sup_{q\in M_n(\mathbb{R})}\inf_{\lambda\in M_n(\mathbb{R})}\left[\langle q,r\rangle-\langle q,\lambda\rangle+f(\lambda)\right]\\
&= -\frac{r^2}{2}+\sup_{q\in M_n(\mathbb{R})}\left(\langle q,r\rangle-\sup_{\lambda\in M_n(\mathbb{R})}\left[\langle q,\lambda\rangle-f(\lambda)\right]\right)\\
&= -\frac{r^2}{2}+f(r)\,,
\end{aligned}$$

using the duality lemma. Since this is true for every $r\in M_n(\mathbb{R})$, we have

$$\mathcal{P}(\beta,x;n)\geq \sup_{r\in M_n(\mathbb{R})}\left[-\frac{r^2}{2}+f(r)\right].$$

But we also have upper bounds for the ostensibly smaller, but actually equal, supremum over $q\in\frac{\beta}{\sqrt{2}}\operatorname{Ov}(n)$. Therefore, we conclude

$$\mathcal{P}(\beta,x;n)=\sup_{q\in\frac{\beta}{\sqrt{2}}\operatorname{Ov}(n)}\left[-\frac{1}{2}q^2+f(q)\right],$$

which is also the maximum over all $q\in M_N(\mathbb{R})$. Making the substitution back to $q\to\frac{\beta}{\sqrt{2}}q$, we obtain the desired formula. □

Remark 4.3 This calculation was essentially done by Sherrington and Kirkpatrick in their original paper [2]. However, they did not use exactly the same argument presented here. Specifically, while both arguments involve large deviations, their calculation used large deviations for Gaussians rather than for the overlap itself. Specifically, they used a trick called the Hubbard-Stratonovich transformation, and they did not use the duality lemma for the Legendre-Fenchel transform. We will eventually make use of the Hubbard-Stratonovich transformation, too, later.

REFERENCES

[1] A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications*, Applications of Mathematics v. 38, Springer Verlag, New York, Inc., 1998.
[2] D. Sherrington and S. Kirkpatrick. Solvable model of a spin-glass. *Phys. Rev. Lett.* **35**, 1792–1796 (1975).

MATHEMATICS DEPARTMENT, UNIVERSITY OF ROCHESTER, ROCHESTER, NY 14627
E-mail address: sstarr@math.rochester.edu