STAT 206A: Gibbs Measures | Invited Speaker: Andrea Montanari

Lecture 5: Random Energy Model

Lecture date: September 12 | *Scribe: Sebastien Roch*

This is a guest lecture by Andrea Montanari (ENS Paris and Stanford) on the Random Energy Model (REM) in physics.

1 Disordered Models

In this lecture, we consider so-called *disordered models* in statistical physics. These are particle systems where the energy function is random. Therefore, we have two levels of randomness. We use the notation $\mathbf{E}, \mathbf{P}$ to denote averaging with respect to the energy and we use the notation $\langle \cdot \rangle$ to denote averaging with respect to the (random) Boltzmann distribution.

The state space is $\{0,1\}^N$ which we sometimes denote equivalently $\{1,\ldots,2^N\}$. We denote the (random) energy function $E : \{0,1\}^N \to \mathbb{R}$. Thus, for $\beta \in [0,+\infty]$, the Boltzmann distribution is

$$p(\boldsymbol{x}) = \frac{1}{Z(\beta)} \exp\left(-\beta E(\boldsymbol{x})\right),$$

where the so-called partition function is

$$Z(\beta) = \sum_{\boldsymbol{x}} \exp\left(-\beta E(\boldsymbol{x})\right).$$

We are interested in the typical properties of p under $\mathbf{P}$.

Example: Random k-SAT. Let $F : \{0,1\}^N \to \{0,1\}$ be a Boolean function in k-CNF form. For example, $F(\boldsymbol{x}) = (x_1 \vee x_2 \vee x_3) \wedge (\bar{x}_2 \vee x_3 \vee \bar{x}_4)$. Let

$$E_F(\boldsymbol{x}) = \#\{\text{clauses violated by } \boldsymbol{x}\}, \tag{1}$$

which we think of as an energy function. Imagine that we pick F uniformly at random over all k-SAT formulas with N variables and $M = \alpha N$ clauses. In particular, with the energy in (1), if $\beta = +\infty$, the Boltzmann distribution is uniform over all satisfying assignments in F. The first two moments of E under $\mathbf{P}$ are easily computed:

$$\mathbf{E}[E(\boldsymbol{x})] = \frac{\alpha N}{2^k},$$

and

$$\begin{aligned}\mathbf{Cov}[E(\boldsymbol{x}), E(\boldsymbol{y})] &= \alpha N(\mathbf{P}[\boldsymbol{x} \;\&\; \boldsymbol{y} \text{ violate an arbitrary clause}] \\ &\quad -\mathbf{P}[\boldsymbol{x} \text{ violates an arbitrary clause}]\mathbf{P}[\boldsymbol{y} \text{ violates an arbitrary clause}]).\end{aligned}$$

Note that the correlation between two assignments $\boldsymbol{x}$ and $\boldsymbol{y}$ depends only on their Hamming distance. Note also that both assignments can violate a clause simultaneously only if they agree on all variables involved. Therefore,

$$\mathbf{Cov}[E(\boldsymbol{x}), E(\boldsymbol{y})] = N\alpha\left(\frac{1}{2^k}(1-\delta)^k - \frac{1}{2^{2k}}\right) \equiv Nf(\delta),$$

where the Hamming distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ is $N\delta$. For large k, the function f is non-negligible from 0 up to $O(k^{-1})$.

2 Random Energy Model

The REM was introduced by Derrida [1]. It is defined as follows. Each energy level is an independent Gaussian with mean 0 and variance $N/2$, i.e. $\{E(\boldsymbol{x})\}_{\boldsymbol{x}\in\{0,1\}^N}$ are i.i.d. $\mathcal{N}(0, N/2)$ and

$$\mathbf{Cov}[E(\boldsymbol{x}), E(\boldsymbol{y})] = \frac{N}{2}\mathbf{1}\{\boldsymbol{x} = \boldsymbol{y}\}.$$

We follow Derrida's treatment of the REM [1]. The main quantities we want to compute are the so-called *free energy*

$$F_N(\beta) = -\frac{1}{\beta}\log Z_N(\beta),$$

the *internal energy*

$$U_N(\beta) = \langle E(\boldsymbol{x})\rangle_\beta = \frac{\partial}{\partial\beta}[\beta F_N(\beta)],$$

and the *canonical entropy*

$$S_N(\beta) = \mathcal{H}(p_\beta) = \beta^2\frac{\partial}{\partial\beta}F_N(\beta),$$

where $\mathcal{H}$ is the Shannon entropy.

3 Thermodynamic Properties

Using the equivalent state space $\{1, \ldots, 2^N\}$, we compute the number of energy levels in the interval $[N\epsilon, N(\epsilon+\delta)]$

$$A(\epsilon, \epsilon+\delta) = \#\{i \,:\, E_i \in [N\epsilon, N(\epsilon+\delta)]\},$$

which up to subexponential factors is, in expectation,

$$\mathbf{E}[A(\epsilon,\epsilon+\delta)] = 2^N \int_{\epsilon}^{\epsilon+\delta} \sqrt{\frac{N}{\pi}} e^{-Nx^2} \mathrm{d}x \doteq e^{N \max_{x\in[\epsilon,\epsilon+\delta]} s_Q(x)},$$

where

$$s_Q(x) = \log 2 - x^2.$$

Let $\epsilon_* = \sqrt{\log 2}$. Note that s is positive on $(-\epsilon_*, \epsilon_*)$. There are two cases of interest:

1. When $[\epsilon,\epsilon+\delta] \cap [-\epsilon_*,\epsilon_*] \neq \emptyset$, $\mathbf{E}[A(\epsilon,\epsilon+\delta)]$ is exponentially large and the random variable $A(\epsilon,\epsilon+\delta)$ is concentrated.
2. When $[\epsilon,\epsilon+\delta] \cap [-\epsilon_*,\epsilon_*] = \emptyset$, $\mathbf{E}[A(\epsilon,\epsilon+\delta)]$ is exponentially small and the random variable $A(\epsilon,\epsilon+\delta)$ is almost surely 0.

Next, we compute the partition function. Let

$$s(x) = \begin{cases} s_Q(x), & x \in [-\epsilon_*, \epsilon_*], \\ 0, & \text{o.w.} \end{cases}$$

We have

$$Z_N(\beta) = \sum_{i=1}^{2^N} e^{-\beta E_i} \doteq \int_{-\infty}^{+\infty} e^{N(s(x)-\beta x)} \mathrm{d}x \doteq e^{N \max_x [s(x)-\beta x]}.$$

To summarize, we can prove the following.

Proposition 1 *Let*

$$\varphi(\beta) = \max_x [s(x) - \beta x].$$

Then we have

$$\lim_{N\to+\infty} \frac{1}{N} \log Z_N(\beta) = \varphi(\beta).$$

Also,

$$\mathbf{P}[|\log Z_N - N\varphi| \geq N\xi] \leq e^{-N\xi^2/2}.$$

4 The Phase Transition

There is a simple graphical way to compute $\varphi(\beta)$: find the point on the curve of $s(x)$ with slope β. See Figure 5.2 in [2]. Therefore, it is easy to show that the *free energy density*

$$f(\beta) = \lim_{N\to+\infty} \frac{1}{N} F_N(\beta) = -\frac{\varphi(\beta)}{\beta} = \begin{cases} -\frac{\beta}{4} - \frac{\log 2}{\beta}, & \beta \leq \beta_c, \\ -\sqrt{\log 2}, & \beta > \beta_c, \end{cases}$$

where $\beta_c = 2\sqrt{\log 2}$. Also, the *entropy density* is

$$s(\beta) = \lim_{N\to+\infty} \frac{1}{N} S_N(\beta) = \begin{cases} -\frac{\beta^2}{4} + \log 2, & \beta \le \beta_c, \\ 0, & \beta > \beta_c, \end{cases}$$

and the *energy density* is

$$u(\beta) = \lim_{N\to+\infty} \frac{1}{N} U_N(\beta) = \begin{cases} -\frac{\beta}{2}, & \beta \le \beta_c, \\ -\sqrt{\log 2}, & \beta > \beta_c. \end{cases}$$

There is a phase transition at the point β_c which is seen by the discontinuity in the second derivative of the free energy density. It is a so-called second order phase transition. The typical behavior in the two phases are:

1. At high temperature, i.e. $\beta \le \beta_c$, there are exponentially many states with energy density $-\beta/2$ and the Boltzmann distribution is roughly uniform over them.
2. At low temperature, i.e. $\beta > \beta_c$, the Boltzmann distribution is concentrated on a subexponential number of states of lowest energy density $-\sqrt{\log 2}$.

5 Low Temperature Regime

In this Section, we consider some more properties of the low temperature regime.

5.1 Ruelle's Reformulation

A different characterization of the REM in terms of a point process was given by Ruelle [3]. Let $\bar{\omega}_1 \ge \bar{\omega}_2 \ge \cdots$ be a Poisson point process on $\mathbb{R}_+$ with intensity $m\omega^{-1-m}$ for $0 \le m < 1$. Consider the random variables

$$\bar{p}_i = \frac{\bar{\omega}_i}{\sum_j \bar{\omega}_j}.$$

Then the values $\{\bar{p}_i\}_i$ behave like the large N limit of the Bolztmann distribution of the REM.

To see this, consider the regime $\beta > \beta_c$. As we pointed out before, the Boltzmann distribution is concentrated on a small number of states with energy density roughly $-\sqrt{\log 2}$. Thus, consider the following rescaling of the energy

$$E_i = -N\sqrt{\log 2} + z_i.$$

By plugging this expression for E_i into the density of a $\mathcal{N}(0, N/2)$, it is easy to see that the z_i's have a density roughly proportional to $e^{-\beta_c z}$. Now, consider again the original probabilities of the REM

$$p_i = \frac{\omega_i}{\sum_j \omega_j},$$

where $\omega_i \sim e^{-\beta_c z_i}$. Then one can show that in the large N limit, the ω_i's form a Poisson point process with intensity proportional to $\omega^{-1-\beta_c/\beta}$. This corresponds to the Ruelle reformulation with $m = \beta_c/\beta$.

Note that this Poisson process has an accumulation point at 0. Also, notice that the larger β is (i.e. the lower the temperature is), the "fatter" the tail of the process is, indicating that the Boltzmann distribution is dominated by a few large values.

5.2 Condensation

To quantify further the *condensation* phenomenon at low temperature, consider the variable

$$Q_{\boldsymbol{x},\boldsymbol{y}} = \begin{cases} 1, & \text{if } \boldsymbol{x} = \boldsymbol{y}, \\ 0, & \text{o.w.}, \end{cases}$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are two states. The inverse of the quantity

$$\mathbf{P}[Q_{\boldsymbol{x},\boldsymbol{y}} = 1] = \mathbf{E}\left[\sum_i p_i^2\right],$$

measures the "number of states" on which the Boltzmann distribution is concentrated. Note that

$$\mathbf{E}\left[\sum_i p_i^2\right] = \mathbf{E}\left[\frac{\sum_i e^{-2\beta E_i}}{Z(\beta)^2}\right] = \mathbf{E}\left[\frac{Z(2\beta)}{Z(\beta)^2}\right].$$

From our previous calculations for Z, one can derive

$$\lim_{N\to+\infty} \mathbf{P}[Q_{\boldsymbol{x},\boldsymbol{y}} = 1] = \begin{cases} 0, & \beta \le \beta_c, \\ 1 - \frac{\beta_c}{\beta}, & \beta > \beta_c. \end{cases}$$

References

[1] B. Derrida, Random-Energy Model: Limit of a Family of Disordered Models, PRL, 45(2):79–82, 1980.

[2] M. Mezard and A. Montanari, Constraint Satisfaction Networks in Physics and Computation, Oxford University Press, In Press.

[3] D. Ruelle, A mathematical reformulation of Derrida's REM and GREM, CMP, 108(2):225–239, 1987.