Rigorous REM solution and Concentration

1 Rigorous Calculation from REM

First recall the definition of the restricted energy model (REM). For $x \in \{\pm 1\}^n$ and an inverse temperature parameter $\beta > 0$, we have

$$E(x) \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left(0, \frac{n}{2}\right), \quad p_\beta(x) = \frac{1}{Z_\beta} \exp\left(\beta E(x)\right)$$

where the normalizing constant (also called partition function)

$$Z_\beta = \sum_{x \in \{\pm 1\}^n} \exp(\beta E(x))$$

makes p_β into a probability distribution. We previously computed (nonrigorously, using the replica trick) that as $n \to \infty$,

$$\frac{1}{n} \log Z_\beta \to \psi(\beta) = \begin{cases} \log 2 + \frac{\beta^2}{4} & \beta \le \beta_c \text{ (high temperature)} \\ \beta\sqrt{\log 2} & \beta > \beta_c \text{ (low temperature)} \end{cases}$$

where $\beta_c = 2\sqrt{\log 2}$ is the threshold between the two regimes. Now, we will rigorously prove that

$$\frac{1}{n} \log Z_\beta = \psi(\beta) + o(1)$$

by demonstrating lower and upper bounds.

1.1 Upper Bounds

We will first demonstrate that

$$\frac{1}{n} \mathbb{E}\left[\log Z_\beta\right] \le \psi(\beta) + o(1).$$

For the high temperature regime, we have that by Jensen

$$\mathbb{E}[\log Z_\beta] \le \log \mathbb{E}[Z_\beta] = \log\left(\sum_{x \in \{\pm 1\}^n} \mathbb{E}[\exp(\beta E(x))]\right) = \log\left(2^n \cdot \exp\left(\frac{n\beta^2}{4}\right)\right) = n\left(\log 2 + \frac{\beta^2}{4}\right).$$

And in the low temperature case, we have that taking a derivative yields that

$$\frac{\partial}{\partial \beta} \log Z_\beta = \frac{\sum_{x \in \{\pm 1\}^n} E(x) \exp(\beta E(x))}{\sum_{x \in \{\pm 1\}^n} \exp(\beta E(x))} = \mathbb{E}_{x \sim p_\beta}[E(x)] \le \max_{x \in \{\pm 1\}^n} E(x).$$

Taking an expectation (and interchanging derivative and integral) yields

$$\frac{\partial}{\partial \beta} \frac{1}{n} \mathbb{E}[\log Z_\beta] \le \frac{1}{n} \mathbb{E}\left[\max_{x \in \{\pm 1\}^n} E(x)\right] = \sqrt{\log 2} + o(1)$$

since the supremum of 2^n many independent standard Gaussians is on the order of $\sqrt{2 \log 2^n} = \sqrt{2n \log 2}$, and we have that each $E(x)$ is i.i.d. $\mathcal{N}(0, n/2)$. Combining the bounds in the two regimes we have that

$$\frac{1}{n} \mathbb{E}[\log Z_\beta] \le \psi(\beta) + o(1).$$

1.2 Lower Bounds

In the low temperature regime, consider that

$$\log Z_\beta = \log\left(\sum_{x\in\{\pm1\}^n} \exp(\beta E(x))\right) \geq \log\left(\max_{x\in\{\pm1\}^n} \exp(\beta E(x))\right) = \max_{x\in\{\pm1\}^n} \beta E(x)$$

and so in turn

$$\frac{1}{n}\mathbb{E}[\log Z_\beta] \geq \beta\sqrt{\log 2} + o(1).$$

The high temperature case is harder. We first note that

$$\mathbb{P}(E(x) \in [n\alpha, n(\alpha+\epsilon)]) = \frac{1}{\sqrt{2\pi}} \int_{n\alpha}^{n\alpha+1} \exp\left(-\frac{x^2}{n}\right) dx$$

and so, under Binomial-Poisson approximation,

$$|\{x \in \{\pm1\}^n \mid E(x) \in [n\alpha, n(\alpha+\epsilon)]\}| \approx \text{Poisson}\left(\frac{2^n}{\sqrt{2\pi}} \int_{n\alpha}^{n(\alpha+\epsilon)} \exp\left(-\frac{x^2}{n}\right) dx\right).$$

Now consider that by approximating $\int_{n\alpha}^{n(\alpha+\epsilon)} \exp(-x^2/n)dx \approx n\epsilon\exp(-n\alpha^2)$ we arrive at

$$\frac{1}{n}\log\left(\frac{2^n}{\sqrt{2\pi}} \int_{n\alpha}^{n(\alpha+\epsilon)} \exp\left(-\frac{x^2}{n}\right) dx\right) = \log 2 - \alpha^2 + o(1).$$

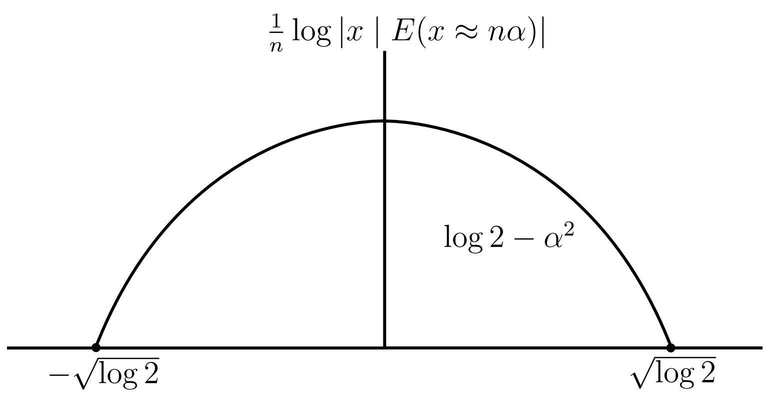

Figure 1: The number of $x \in \{\pm1\}^n$ such that $E(x) \approx n\alpha$ is (up to exponential accuracy) $\exp(n(\log 2 - \alpha^2))$.

Then,

$$Z_\beta = \sum_{x\in\{\pm1\}^n} \exp(\beta E(x)) \geq |\{x \in \{\pm1\}^n \mid E(x) \geq n\alpha\}| \cdot \exp(\beta n\alpha)$$

so

$$\frac{1}{n}\log Z_\beta \geq \beta\alpha + \log 2 - \alpha^2 + o(1)$$

whereby taking $\alpha = \beta/2$ yields

$$\frac{1}{n}\log Z_\beta \geq \log 2 + \frac{\beta^2}{4} + o(1)$$

as desired.

1.3 Concentration

Above, we have proved bounds on $\mathbb{E}[\log Z_\beta]$; to conclude we therefore need to show that $\log Z_\beta$ concentrates well, i.e. $\frac{1}{n}\log Z_\beta \to \frac{1}{n}\mathbb{E}[\log Z_\beta]$. To do this, we first introduce an inequality which we will justify later.

Theorem 1 (Poincaré's Inequality)**.** *For $Z \sim \mathcal{N}(0, \sigma^2 I_n)$ and $f : \mathbb{R}^n \to \mathbb{R}$ differentiable, we have*

$$\mathrm{Var}(f(Z)) \leq \sigma^2 \mathbb{E}[|\nabla f(Z)|_2^2].$$

Granting the above, we have that

$$\mathrm{Var}(\log Z_\beta) \leq \frac{n}{2}\mathbb{E}[|\nabla_E \log Z_\beta|_2^2] \leq \frac{n\beta^2}{2}$$

and so

$$\mathrm{Var}\left(\frac{1}{n}\log Z_\beta\right) \leq \frac{\beta^2}{2n} \implies \frac{1}{n}\log Z_\beta - \mathbb{E}\left[\frac{1}{n}\log Z_\beta\right] = O_{\mathbb{P}}\left(\frac{\beta}{\sqrt{2n}}\right).$$

This fact, combined with the bounds in the previous sections, establishes that

$$\lim_{n\to\infty} \frac{1}{n}\log Z_\beta = \psi(\beta)$$

as desired.

2 Hermite Polynomials

To prove Poincaré's Inequality, we first quickly develop some basic theory of the (probabilist's) Hermite polynomials.

Definition 1. We say that a collection of polynomials $\{p_n\}_{n=0}^\infty$, with $\deg(p_n) = n$, is an **orthogonal basis** of $L^2(\mu)$ if

$$\mathbb{E}_{X\sim\mu}[p_n(X)p_m(X)] = 0 \iff n \neq m$$

and

$$\mathrm{span}(p_0, p_1, p_2, \dots) = \mathrm{span}(1, x, x^2, \dots).$$

We say that it is **orthonormal** if for all $n \in \mathbb{N}$,

$$\mathbb{E}_{X\sim\mu}[p_n(X)^2] = 1.$$

We now introduce our premier example of such an orthogonal basis:

Definition 2. The **probabilist's Hermite polynomials** are defined by

$$\mathrm{He}_n(x) = (-1)^n e^{x^2/2}\left(\frac{d^n}{dx^n}e^{-x^2/2}\right).$$

The first few polynomials are:

$$\mathrm{He}_0(x) = 1$$
$$\mathrm{He}_1(x) = x$$
$$\mathrm{He}_2(x) = x^2 - 1$$

and so on. One useful identity about the Hermite polynomials will be their relation to the exponential generating function $e^{tx-t^2/2}$. Specifically, write the Taylor expansion

$$e^{-(x-t)^2/2} = \sum_{n=0}^{\infty} \frac{\partial^n}{\partial t^n} e^{-\frac{(x-t)^2}{2}} \bigg|_{t=0} \frac{t^n}{n!}$$

and note that by symmetry

$$\frac{\partial^n}{\partial t^n} e^{-\frac{(x-t)^2}{2}} = (-1)^n \frac{\partial^n}{\partial x^n} e^{-\frac{(x-t)^2}{2}}$$

so

$$\sum_{n=0}^{\infty} \frac{\partial^n}{\partial t^n} e^{-\frac{(x-t)^2}{2}} \bigg|_{t=0} \frac{t^n}{n!} = \sum_{n=0}^{\infty} \frac{\partial^n}{\partial x^n} e^{-\frac{x^2}{2}} \frac{t^n}{n!}$$

and, multiplying by e^{x^2}, we get that

$$e^{tx-t^2/2} = \sum_{n=0}^{\infty} \frac{\mathrm{He}_n(x) t^n}{n!}.$$

We may now check that the Hermite polynomials are orthogonal.

Lemma 1. *Let $Z \sim \mathcal{N}(0,1)$. Then*

$$\mathbb{E}[\mathrm{He}_n(Z)\mathrm{He}_m(Z)] = \begin{cases} n! & n = m \\ 0 & \textit{otherwise} \end{cases}.$$

Proof. We use the generating function identity. First, note that

$$\mathbb{E}\left[e^{tZ-t^2/2} e^{sZ-s^2/2}\right] = \mathbb{E}[e^{(s+t)Z}] e^{-\frac{t^2}{2}-\frac{s^2}{2}} = e^{st} = \sum_{n=0}^{\infty} \frac{(ts)^n}{n!}$$

since $\mathbb{E}[e^{(s+t)Z}] = e^{(s+t)^2/2}$ follows from the formula for the MGF of a standard Gaussian. On the other hand, we have that

$$\begin{aligned} \mathbb{E}\left[e^{tZ-t^2/2} e^{sZ-s^2/2}\right] &= \mathbb{E}\left[\left(\sum_{n=0}^{\infty} \frac{\mathrm{He}(Z) t^n}{n!}\right)\left(\sum_{n=0}^{\infty} \frac{\mathrm{He}(Z) s^n}{n!}\right)\right] \\ &= \sum_{n,m=0}^{\infty} \frac{\mathbb{E}\left[\mathrm{He}_n(Z)\mathrm{He}_m(Z)\right]}{n!m!} t^n s^m \end{aligned}$$

so we get what we want by matching terms in the power series. □

The above (and some more facts about the Hermite polynomials), combined with the following approximation fact which we will take for granted, will let us demonstrate Poincaré's identity.

Theorem 2. *Let $Z \sim N(0,1)$ and f be a function such that $\mathbb{E}[f(Z)^2] < \infty$; then for al $\epsilon > 0$, there is some polynomial p_ϵ such that*

$$\mathbb{E}[(f(Z) - p_\epsilon(Z))^2] \leq \epsilon.$$