Statistical Physics of Computation - Exercises

Emanuele Troiani, Vittorio Erba, Yizhou Xu

September 2024

Week 9

9.1 Replica computation for the spiked-Wigner model

We are given as observation a $N \times N$ symmetric matrix Y created as

$$\boldsymbol{Y} = \sqrt{\frac{\lambda}{N}} \underbrace{\boldsymbol{x}^* \boldsymbol{x}^{*\intercal}}_{N\times N \text{ rank-one matrix}} + \underbrace{\boldsymbol{\xi}}_{\text{symmetric iid noise}}$$

where $\boldsymbol{x}^* \in \mathbb{R}^N$ with $x_i^* \overset{\text{i.i.d.}}{\sim} P_X(x)$, $\xi_{ij} = \xi_{ji} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ for $i \leq j$. We keep the prior P_0 generic, as long as it is factorized over all the components of the vector $\boldsymbol{x}$.

This is called the spiked-Wigner model in statistics. The name "Wigner" refer to the fact that Y is a Wigner matrix (a symmetric random matrix with components sampled randomly from a Gaussian distribution) plus a "spike", that is a rank one matrix $\boldsymbol{x}^*\boldsymbol{x}^{*\intercal}$. We will use it as an example of recovering a signal with low rank structure corrupted by noise which we want to clean up.

Our task shall be to recover the vector $\boldsymbol{x}$ from the knowledge of $\boldsymbol{Y}$, the signal-to-noise ratio λ and the prior P_0. As we saw during the lecture, this can be achieved using the posterior estimation, i.e. by computing the posterior distribution and evaluating some statistics over it.

In this exercise, you will compute the normalization factor of the posterior distribution, i.e. the partition function of the problem, and derive the state equation for the overlap order parameter.

1. Show that the posterior distribution $P(\boldsymbol{x}|\boldsymbol{Y})$ for the problem can be written as

$$P(\boldsymbol{x}|\mathbf{Y}) = \frac{1}{Z(\mathbf{Y})} \left[\prod_{i=1}^{N} P_X(x_i)\right] \left[\prod_{i\leq j} \frac{e^{-\frac{\lambda}{2N}x_i^2x_j^2 + \sqrt{\frac{\lambda}{N}}x_ix_jy_{ij}}}{\sqrt{2\pi}}\right] \tag{1}$$

   for a specific $Z(\mathbf{Y})$. How is $Z(\mathbf{Y})$ defined for this measure?

We are interested in computing the averaged free entropy associated to the posterior distribution, i.e.

$$\lim_{N\to\infty} \mathbb{E}_{\boldsymbol{Y}} \left[\frac{1}{N} \log Z(\boldsymbol{Y})\right]$$

which we can compute using the replica method.

2. Show that the averaged replicated partition function equals

$$\mathbb{E}_{\mathbf{Y}}[Z^n] = \int \mathrm{d}\boldsymbol{Y} e^{-\frac{1}{2}\sum_{i\le j} y_{ij}^2} \prod_{\alpha=0}^{n} \int \mathrm{d}\boldsymbol{x}^{(\alpha)} \left(\prod_{i=1}^{N} P_X\left(x_i^{(\alpha)}\right)\right)\left(\prod_{i\le j} \frac{e^{-\frac{\lambda}{2N}\left(x_i^{(\alpha)}\right)^2\left(x_j^{(\alpha)}\right)^2+\sqrt{\frac{\lambda}{N}}x_i^{(\alpha)}x_j^{(\alpha)}y_{ij}}}{\sqrt{2\pi}}\right) \tag{2}$$

where you should notice that we are taking the product $n+1$ replicas.

3. Integrate over the disorder, i.e. the observation Y, to get at leading order in N

$$\mathbb{E}_{\mathbf{Y}}[Z^n] = \int \prod_{\alpha,i} P_X\left(x_i^{(\alpha)}\right)\mathrm{d}x_i^{(\alpha)}\ \exp\left(\frac{\lambda N}{2}\sum_{\alpha<\beta}\left(\sum_i \frac{x_i^{(\alpha)}x_i^{(\beta)}}{N}\right)^2\right) \tag{3}$$

4. Introduce the appropriate order parameters and obtain

$$\mathbb{E}_{\mathbf{Y}}[Z^n] = \int \prod_{\alpha<\beta} \mathrm{d}\hat{q}_{\alpha\beta}\,\mathrm{d}q_{\alpha\beta}\ \exp\left(NI_{\text{energy}}(q_{\alpha\beta},\hat{q}_{\alpha\beta}) + NI_{\text{entropy}}(\hat{q}_{\alpha\beta})\right) \tag{4}$$

where we defined

$$I_{\text{energy}}(q_{\alpha\beta},\hat{q}_{\alpha\beta}) = \frac{\lambda}{2}\sum_{\alpha<\beta} q_{\alpha\beta}^2 - \sum_{\alpha<\beta} q_{\alpha\beta}\hat{q}_{\alpha\beta} \tag{5}$$

and

$$I_{\text{entropy}}(\hat{q}_{\alpha\beta}) = \log\left(\int \prod_{\alpha} P_X\left(x_\alpha\right)\mathrm{d}x_\alpha\ \exp\left\{\sum_{\alpha<\beta}\hat{q}_{\alpha\beta}x_\alpha x_\beta\right\}\right) \tag{6}$$

where we stress that here the integral over dx_α runs over the real numbers for all $\alpha = 0,\dots,n$.

5. Which replica ansatz should you impose in this computation?

6. Show that in the ansatz you discussed in the previous point, the energetic term satisfies

$$\lim_{n\to 0}\frac{1}{n}I_{\text{energy}}(q_{\alpha\beta},\hat{q}_{\alpha\beta}) = \frac{\lambda}{4}q^2 - \frac{q\hat{q}}{2} \tag{7}$$

for q and $\hat{q}$ real numbers.

7. Show that in the ansatz you discussed in the previous point, the entropic term satisfies

$$I_{\text{entropy}}(\hat{q}_{\alpha\beta}) = \log\left(\int Dz\ \left[\left(\int P_X\left(x\right)\mathrm{d}x\ \exp\left\{-\frac{\hat{q}}{2}x^2 + \sqrt{\hat{q}}xz\right\}\right)^{n+1}\right]\right) \tag{8}$$

for $\hat{q}$ real numbers.

8. Show that in the small n limit, the entropic term satisfies

$$\lim_{n\to 0}\frac{1}{n}I_{\text{entropy}}(\hat{q}_{\alpha\beta}) = \int Dz\, I(\hat{q},z)\log\left(I(\hat{q},z)\right) \tag{9}$$

where we defined $I(\hat{q}, z)$ as

$$I(\hat{q}, z) = \int P_X(x)\, \mathrm{d}x \;\; \exp\left\{-\frac{\hat{q}}{2}x^2 + \sqrt{\hat{q}}xz\right\} \tag{10}$$

and that it can be equivalently rewritten as

$$\lim_{n\to 0} \frac{1}{n} I_{\text{entropy}}(\hat{q}_{\alpha\beta}) = \int Dz \int P_X(x_0)\, \mathrm{d}x_0 \log\left(\int P_X(x)\, \mathrm{d}x \;\; \exp\left\{-\frac{\hat{q}}{2}x^2 + \sqrt{\hat{q}}xz + \hat{q}xx_0\right\}\right) \tag{11}$$

9. Argue finally that the free entropy

$$\phi = \lim_{N\to\infty} \mathbb{E}_{\boldsymbol{Y}}\left[\frac{1}{N}\log Z(\boldsymbol{Y})\right] \tag{12}$$

can be expressed as

$$\phi = \text{extr}_{q,\hat{q}}\left[\frac{\lambda}{4}q^2 - \frac{q\hat{q}}{2} + \int Dz \int P_X(x_0)\, \mathrm{d}x_0 \log\left(\int P_X(x)\, \mathrm{d}x \;\; \exp\left\{-\frac{\hat{q}}{2}x^2 + \sqrt{\hat{q}}xz + \hat{q}xx_0\right\}\right)\right] \tag{13}$$

or equivalently as

$$\phi = \text{extr}_{q}\left[-\frac{\lambda}{4}q^2 + \int Dz P_X(x_0)\, \mathrm{d}x_0 \log\left(\int P_X(x)\, \mathrm{d}x \;\; \exp\left\{-\frac{\lambda q}{2}x^2 + \sqrt{\lambda q}xz + \lambda q x x_0\right\}\right)\right] \tag{14}$$