Statistical Physics of Computation 2025 - Exercises

Vittorio Erba, Emanuele Troiani

Week 6

6.1 Phase diagrams of spiked-Wigner model

In this exercise we will take the result of the replica computation we previously did for the low-rank denoising problem and study the associated state equations in a couple of simple cases. Our starting point is the state equation on m:

$$m = \int Dz \mathrm{d}x_0 P_0(x_0) \frac{\int P_0(x)\, x\, x_0 \,\mathrm{d}x \;\exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m}xz + \lambda m x x_0\right\}}{\int P_0(x)\,\mathrm{d}x \;\exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m}xz + \lambda m x x_0\right\}} \tag{1}$$

which is the extremisation equation of the free entropy ϕ

$$\begin{aligned}\phi &= \max_m \phi(m) \\ &= \max_m \left[-\frac{\lambda}{4}m^2 + \int Dz P_X(x_0)\,\mathrm{d}x_0 \log\left(\int P_X(x)\,\mathrm{d}x\; e^{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m}xz + \lambda m x x_0}\right)\right]\end{aligned} \tag{2}$$

where Dz denotes integration against a standard Gaussian variable.

Gaussian prior. We start with a simple case, i.e. Gaussian prior on the components of x:

$$x, x_0 \sim \mathcal{N}(0,1)\,. \tag{3}$$

1. Show that the state equation becomes

$$m = \frac{\lambda m}{\lambda m + 1}\,. \tag{4}$$

We start with the integral

$$\int P_0(x)\,\mathrm{d}x\;\exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m}xz + \lambda m x x_0\right\} \tag{5}$$

which is a Gaussian integral

$$\begin{aligned}&\int P_0(x)\,\mathrm{d}x\;\exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m}xz + \lambda m x x_0\right\} \\ &= \frac{1}{\sqrt{2\pi}}\int \mathrm{d}x\;\exp\left\{-\frac{\lambda m + 1}{2}x^2 + \sqrt{\lambda m}xz + \lambda m x x_0\right\} \\ &= \frac{e^{\frac{(\lambda m x_0 + z\sqrt{\lambda m})^2}{2\lambda m + 2}}}{\sqrt{\lambda m + 1}}\,.\end{aligned}$$

Similarly we have

$$\int P_0\left(x\right) x x_0 \,\mathrm{d}x \;\; \exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m} x z + \lambda m x x_0\right\}$$
$$= \frac{e^{\frac{\left(\lambda m x_0 + z\sqrt{\lambda m}\right)^2}{2\lambda m+2}} \left(\lambda m x_0^2 + x_0 z \sqrt{\lambda m}\right)}{\left(\lambda m+1\right)^{3/2}} .$$

We can take the ratio to have

$$\frac{\int P_0\left(x\right) x\, x_0 \,\mathrm{d}x \;\; \exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m} x z + \lambda m x x_0\right\}}{\int P_0\left(x\right) \mathrm{d}x \;\; \exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m} x z + \lambda m x x_0\right\}} = \frac{\lambda m x_0^2}{\lambda m+1} + \frac{x_0 z\sqrt{\lambda m}}{\lambda m+1} .$$

The last two steps are just taking the expectation over x_0, z, which are both standard Gaussian variables, giving

$$m = \frac{\lambda m}{\lambda m+1} . \tag{6}$$

2. Compute the variational free entropy $\phi(m)$ for this prior.

Using one of the integral computed above, we have

$$\phi(m) = -\frac{\lambda}{4}q^2 + \int Dz \int P_0\left(x_0\right) \mathrm{d}x_0 \log\left(\int P_0\left(x\right) \mathrm{d}x \;\; \exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m} x z + \lambda m x x_0\right\}\right)$$
$$= -\frac{\lambda}{4}m^2 + \frac{\lambda m}{2} - \frac{1}{2}\log\left(1+\lambda m\right)$$

3. Solve the state equation. In principle it should have two solutions. How should one choose the right one? Use this criterion to find the correct $m(\lambda)$ for each λ.

The first step is to simply solve the state equation. We can write it as a second degree polynomial

$$m^2\lambda + m(1-\lambda) = 0 \tag{7}$$

It has a trivial solution $m = 0$ and a non-trivial one

$$m = 1 - \frac{1}{\lambda} \tag{8}$$

The dominant solution is the one with largest free entropy. One can compute that the free entropy of the trivial solution is zero, while the free entropy of the non-trivial solution is positive only for $\lambda > 1$. Thus, the trivial solution dominates for $0 < \lambda < 1$, and the non-trivial one dominates for $\lambda > 1$.

4. Is there a phase transition? Of which order? Justify.

Yes, there is a phase transition, as there is an abrupt change in the behavior of the system at $\lambda = 1$. In particular, the value of m changes continuously as one varies λ, but its derivative has a jump. This is the signature of a second order phase transition.

Sparse binary prior. A more interesting case is the sparse binary prior. The elements of x have a probability $0 < \rho < 1$ of being one and $1 - \rho$ of being zero

$$P_0(x_i) = \rho\delta(x_i - 1) + (1 - \rho)\delta(x_i) \quad \forall i\,, \tag{9}$$

where δ is the Dirac's delta.

5. Show that the state equation becomes

$$m = \int Dz \frac{\rho^2}{\rho + (1-\rho)e^{-\frac{\lambda m}{2} + z\sqrt{\lambda m}}} \tag{10}$$

We start with the integral

$$\int P_0\left(x\right) \mathrm{d}x \;\; \exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m} x z + \lambda m x x_0\right\} \tag{11}$$

The prior is

$$P_0\left(x\right) = (1-\rho)\delta(x) + \rho\delta(x-1) \tag{12}$$

which means the integral is just a simple sum.

$$\int P_0\left(x\right) \mathrm{d}x \;\; \exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m} x z + \lambda m x x_0\right\}$$
$$= (1-\rho) + \rho e^{-\frac{\lambda m}{2} + \lambda m x_0 + z\sqrt{\lambda m}}$$

Similarly we have

$$\int P_0\left(x\right) x x_0 \,\mathrm{d}x \;\; \exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m} x z + \lambda m x x_0\right\}$$
$$= \rho x_0 e^{-\frac{\lambda m}{2} + \lambda m x_0 + z\sqrt{\lambda m}}$$

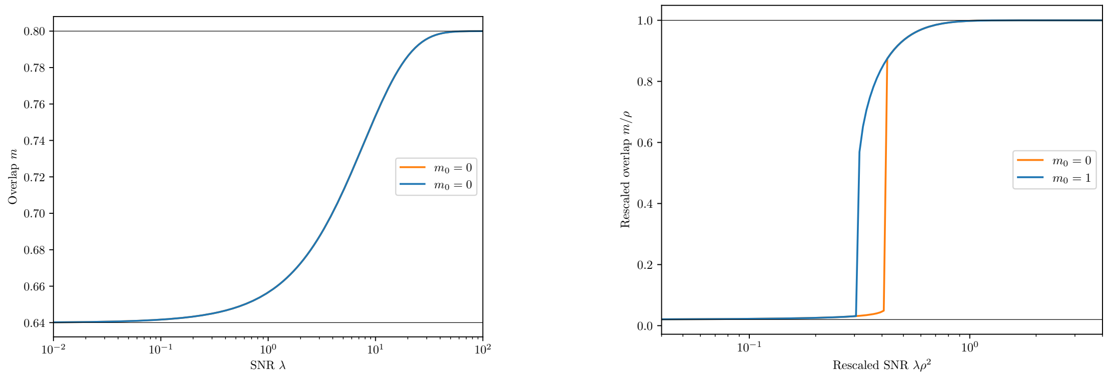
We can take the ratio to have

$$\frac{\int P_0\left(x\right) x\, x_0 \,\mathrm{d}x \;\; \exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m} x z + \lambda m x x_0\right\}}{\int P_0\left(x\right) \mathrm{d}x \;\; \exp\left\{-\frac{\lambda m}{2}x^2 + \sqrt{\lambda m} x z + \lambda m x x_0\right\}}$$
$$= \frac{\rho x_0}{\rho + (1-\rho)e^{\frac{\lambda m}{2} - \lambda m x_0 - z\sqrt{\lambda m}}}$$

Now we have to integrate over x_0. This is again a sum. The case $x_0 = 0$ will give zero contribution, thus we have

$$\int Dz P_0\left(x_0\right) \frac{\rho x_0}{\rho + (1-\rho)e^{\frac{\lambda m}{2} - \lambda m x_0 - z\sqrt{\lambda m}}}$$
$$= \int Dz \frac{\rho^2}{\rho + (1-\rho)e^{\frac{\lambda m}{2} - \lambda m - z\sqrt{\lambda m}}}$$
$$= \int Dz \frac{\rho^2}{\rho + (1-\rho)e^{-\frac{\lambda m}{2} + z\sqrt{\lambda m}}}$$

The last integral must be done numerically.



(a) Case with low sparsity $\rho = 0.8$. We notice how there is no phase transition and both initialisations converge to the same value.

(b) Case with high sparsity $\rho = 0.02$. Here we can see a first order phase transition. Notice how the axis are rescaled to make the plot clear

Figure 1: Overlaps as a function of the SNR ratio λ for the binary prior.

6. We would like to solve the state equation, but it would only be possible to do so numerically. We do that by a fixed point iteration scheme

$$m^{t+1} = \int Dz \frac{\rho^2}{\rho + (1-\rho)e^{-\frac{\lambda m^t}{2} + z\sqrt{\lambda m^t}}} \,. \tag{13}$$

Iterate numerically (in the language of your choice) the state evolution until convergence for a large value of ρ (for example $\rho = 0.8$) for several value of λ. Use $m^{t=0} = 0.01$ and $m^{t=0} = 0.99$ as initializations. What do you observe? Is the behavior of m at convergence dependent on the initialization? Are there any phase transitions?
Hint: you will need to integrate (possibly by quadrature) over z. We advise to choose a suitable integration interval (something like $[-5, 5]$ would do).

You can see the result in Figure 1a. There is no phase transition (no non-analyticity of the free entropy) for such a high ρ, and both informed $m^{t=0} = 0.99$ and non-informed $m^{t=0} = 0.01$ initializations converge to the same value. Notice how the overlap ranges from ρ^2 to ρ, as you can expect from looking at the state equation in the limits of large and small λ.

7. We want to repeat the same numerical experiment for small ρ. First, it is useful to estimate the magnitude of the SNR and of the magnetization in that limit. Show that

$$Q = Q_* = \mathbb{E}_{x \sim P_0} x^2 = \rho \,. \tag{14}$$

Argue that this implies $m < \rho$, so that it is useful to plot m/ρ as our order parameter for small ρ. Argue that the strength of the signal scales linearly with ρ, hence the effective signal to noise ratio of the model is $\lambda\rho^2$.

Q_* can be computed by simply computing the averaged Euclidean norm of the prior, giving $Q = Q_* = \rho$. By the Cauchy–Schwarz inequality we have $m = q \leq Q$, so $m \leq \rho$. As for λ, it's reasonable to assume that the intensity of the signal scales with the number of non-zero elements in x, which on average is ρ. We thus want to reabsorb this dependency on ρ in the SNR, so we define the rescaled SNR $\sqrt{\lambda'} = \sqrt{\lambda}\rho$, leading to $\lambda' = \lambda\rho^2$

8. Show numerically that for $\rho = 0.02$ small enough there is a first order phase transition by plotting m at convergence as a function of λ. Hint: In order to see anything from your plot you should use the scaling in m and λ from the previous point. Remember to initialize the state equation both in $m^{t=0} = 0.01$ and $m^{t=0} = 0.99$.

   You can see the result in Figure 1b. Notice in particular how the informed and uninformed initializations lead to different solutions in a range of SNRs. This is the signature of a first order phase transition, as the two solutions correspond to two different maxima of the free entropy, which as λ varies exchange as the dominating maximum of the free entropy in a discontinuous manner.

As discussed in class, to find the precise point at which the first order phase transition happens one needs to compare the value of the free entropy at the two solutions involved in the transition. For this prior, this is not so simple to do, as computing numerically the free entropy for small ρ leads to numerical instabilities. The solution is to expand the free entropy for small ρ after rescaling the magnetization and the SNR to highlight their natural dependence on ρ, in order to obtain a numerically more manageable approximate expression. This whole procedure starts being outside of the scope of these lectures, so we do not cover it here.