Machine Learning for Physical Scientists

Instructor: Thiparat Chotibut

Homework 3

Due: March 25, 2021 at noon

This homework encourages you to adopt theoretical techniques from disorder systems to analyze problems of statistical physics and of statistical inference. Please refer to lecture notes or consult reading materials from week 5-7 before attempting to solve the problems.

1 3-spin Model

In this problem, we'll solve a generalization of Ising model to include cubic spin interaction. The 3-spin ferromagnetic model has the Hamiltonian that reads

$$\mathcal{H}\left(\{S_i\}_{i=1}^N\right) = -\frac{1}{N^2}\sum_{i,j,k} S_i S_j S_k - h\sum_i S_i = -N\left[\left(\frac{\sum_i S_i}{N}\right)^3 + h\frac{\sum_i S_i}{N}\right], \tag{1}$$

where N is the total number of spins and the spin S_i is either ± 1.

(a) Let m be the magnetization, argue that the number of configurations with magnetization m, $\mathcal{N}(m)$, is given by

$$\mathcal{N}(m) = \left(\begin{array}{c} N \\ \frac{N+Nm}{2} \end{array}\right) = \frac{N!}{\left(\frac{1+m}{2}N\right)!\left(\frac{1-m}{2}N\right)!}.$$

Using the Stirling formula, show that the free energy per spin of the 3-spin model $f(\beta,h) = -\frac{1}{N\beta}\lim_{N\to\infty}\log Z_N(\beta,h)$ can be written in the asymptotically large N limit as

$$f(\beta,h) = \min_{m\in[-1,1]} \mathcal{F}_m(\beta,h,m),$$

where the free energy at fixed magnetization $\mathcal{F}_m(\beta,h,m)$ is given by

$$\mathcal{F}_m(\beta,h,m) = -m^3 - hm + \frac{1}{\beta}\left[\frac{1+m}{2}\log\left(\frac{1+m}{2}\right) + \frac{1-m}{2}\log\left(\frac{1-m}{2}\right)\right],$$

whose minimzer m^* is given by

$$\left\langle \frac{\sum_i S_i}{N}\right\rangle = m^*.$$

(b) From now we'll assume the external field is absent, i.e. $h = 0$. Plot $\mathcal{F}_m(\beta, h = 0, m)$ as a function of m for different values of the inverse temperature β. Compute the value of m^* for many temperatures to show that it has a *first order transition*, that is there's a discontinuity in $m*$ as one changes the temperature.

(c) Show (numerically) that there are three regions: for $\beta < \beta_s$ there is a unique minima in $m = 0$. For $\beta_s < \beta < \beta_c$ a second local minima appears. For $\beta_c < \beta$ the second minima is now the global one. Find the values of β_s and β_c.

2 Statistical Inference for Binary Estimation

We have studied in class the following problem: given a vector $\mathbf{x}^*$, where each component x_i^* is taken from $P_X(x)$, one is given a noisy symmetric $N \times N$ matrix $\mathbf{Y}$ such that

$$Y_{ij} = \sqrt{\frac{\lambda}{N}} x_i^* x_j^* + \omega_{ij}$$

where λ is the signal to noise ratio, and where $\omega_{ij} = \omega_{ji}$ is a noise taken from a Gaussian distribution $\mathcal{N}(0,1)$. We have seen that the free entropy reads $\Phi(\lambda) = \min_m \Phi_{\mathrm{RS}}(m;\lambda)$ where $m \geq 0$ and

$$\Phi_{\mathrm{RS}}(m;\lambda) = -\frac{\lambda m^2}{4} + \int P_X\left(x^*\right) \mathrm{d}x^* \int \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}} \mathrm{d}z \log \left\{ \int P_X(x) \mathrm{d}x \exp\left(-\frac{\lambda m}{2} x^2 + \lambda m x^* x + \sqrt{\lambda m} z x \right) \right\}$$

and the maximizer $m^*(\lambda)$ is the magnetization $m = E_{\mathbf{x}^*,\omega}\left[\left\langle \frac{1}{N} \sum_i x_i^* x_i \right\rangle_{\mathbf{x}^*,\omega} \right]$. We remind that the Minimum Mean Square Error is given by

$$\mathrm{MMSE}(\lambda) = E_{x^*}\left[\left(x^*\right)^2 \right] - m^*(\lambda).$$

(a) Assume now that each $x_i^* = \pm 1$ are distributed with equal probabilities, i.e. $P_X(x) = \frac{1}{2}[\delta(x-1) + \delta(x+1)]$, show that one can drastically simplify the expression as

$$\Phi_{\mathrm{RS}}(m,\lambda) = -\frac{\lambda m^2}{4} - \frac{\lambda m}{2} + \frac{1}{2} E_z \left[\log \left\{ \frac{1}{2} (\cosh(2\lambda m) + \cosh(2z\sqrt{\lambda m})) \right\} \right].$$

(b) Numerically compute the MMSE as a function of λ and show that there is a phase transition at $\lambda = 1$. Interpret your results in terms of learnability of statistical inference.