APMLA: assignment 1 Block II

Caterina De Bacco and Nicolò Ruggeri

Exercise 1: Inverse SK model

Consider the *inverse* SK model: given a set of M data samples $\mathbf{s}_1, \ldots, \mathbf{s}_M$, estimate the values of $\mathbf{J}$ and $\mathbf{h}$ that better explains the observed data, where each sample is drawn forma a Boltzmann distribution with Hamiltonian:

$$H(\mathbf{s}) = -\sum_{i \neq j} J_{ij} s_i s_j - \sum_i h_i s_i \quad . \tag{1}$$

We now want to *sample* configurations of N variables from the corresponding Boltzmann distribution for a particular realization of the couplings and comparing the empirical mean of the magnetizations with the one inferred using TAP and MF.
For sampling, we use the Monte-Carlo-Markov-Chain (MCMC) Metropolis-Hastings algorithm as we learned for the Curie Weiss model in tutorial 7.

(a) Derive equations for estimating the parameters h_i^{TAP} and J_{ij}^{TAP} analogous to what we have done in the lectures for MF. Consider only the case $i \neq j$.
The equation for the $\mathbf{J}$ should be equivalent to the following expression with matrices:

$$\mathbf{C} = \beta \mathbf{P}(1 + \mathbf{J}\,\mathbf{C} - \beta \mathbf{S}\,\mathbf{C} + 2\beta \mathbf{M}\,\mathbf{C})$$

which uses the matrix $\mathbf{M}$ and the diagonal matrices $\mathbf{P}$ and $\mathbf{S}$ with entries:

$$P_{ii} = 1 - m_i^2 \tag{2}$$

$$S_{ii} = \sum_k J_{ik}^2 (1 - m_k^2) \tag{3}$$

$$M_{ik} = m_i J_{ik}^2 m_k \tag{4}$$

Consider only one of the 2 solutions of the quadratic equation for J_{ij} (the one with the + sign).
Assume that $J_{ii} = 0\,, \forall i = 1, \ldots, N$, i.e. when writing sums like $\sum_{j=1}^N J_{ij} s_j$ this is equivalent to $\sum_{j \neq i} J_{ij} s_j$.
Hint: use the linear response theorem $C_{ij} := < s_i s_j >_D - < s_i >_D < s_j >_D = \frac{\partial < s_i >_D}{\partial h_j}$.

(b) Fill up the *jupyter* notebook uploaded on github with the skeleton of a code to test TAP and MF in this inference task. Throughout the following exercises we assume $\beta = 1$.
i) Extract a "ground-truth" set of parameters $h_i^{GT} \sim \mathcal{N}(0, 0.01)$ and $J_{ij}^{GT} \sim \mathcal{N}(0, J_0/N)$ with $J_0 = 1$.
ii) Generate $M = 10000$ samples of a system of $N = 10, 20, 50, 100$ (if you can try also larger N, but for $N \sim 1000$ the code runs real slow) random variables extracted using the above ground-truth. For the MCMC sampler, consider a burn-in period $T_{eq} = 100N$ i.e. number of MCMC steps performed before the first sample is returned (each step is a flip of one single variable). Select samples every $d_{sample} = 10N$, so that samples are not correlated.
iii) Infer $\mathbf{h}^{model}$,$\mathbf{J}^{model}$ for model being MF and TAP using the equations derived in the lectures and in (a).

(c) Repeat this for other values of $J_0 = 0.1, 2$, i.e. tuning the coupling strength.

(d) Generate scatter plots with J_{ij}^{model} vs J_{ij}^{GT}, similar for h_i^{model} vs h_i^{GT} (one plot for each of the 3 values of J_0). Plot all the different realizations in N on the same scatter plot, distinguishing them by marker type. At the end we have a total of 12 plots (one per model, one per parameter J or h, one per value of J_0.

(e) Calculate RMSE between inferred and ground-truth values of the parameters.

(f) Generate plots with RMSE as a function of N for the 3 different J_0 values (using different markers' type). We have a total of 2 plots, one per parameter J or h.

(g) Plot $\bar{J}$ and $\bar{h}$ as a function of N, where $\bar{J}$ is the average over the i, j. Similar for $\bar{h}$. In this plot there should be 3 different curves: one for the ground-truth, one for MF and one for the TAP. There is a total of 6 plots, two per J_0.

(h) Same plot but for the variances $\bar{h^2} - \bar{h}^2$ and $\bar{J^2} - \bar{J}^2$.

(i) Comment on the main results you observe.