Advanced Probabilistic Machine Learning and Applications

Caterina De Bacco and Daniela Leite

Exercise 1: sampling from the SK model

Consider again the Hamiltonian of the SK model:

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

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 a previous tutorial.

(a) <u>Write a code</u> to perform the MCMC dynamics, and start by configurations extracted uniformly at random.
i) Sample a particular realization of $\mathbf{J} \sim \mathcal{N}(0, \sigma^2/\sqrt{N})$, for $\sigma^2 = 1$ and $\beta = 1.1$.
ii) Run your dynamics for a long enough time (say, with $t_{\max} = 10^3 N$ attempts to flips spins) and monitor the value of the magnetization $m = \sum_i s_i/N$ as a function of time. Make a plot for $N = 10, 100, 1000$ spins.

(b) Iterate MF and TAP equations and compare the values obtained from the ground truth MCMC sampler. To do so, draw a scatter plot of the single values m_i obtained from the MCMC sampler and the approximations (one plot per approximation, for a total of 2 plots).
Comment. Similarly to before, implement the consistency equations as

$$m_i^{(t+1)} = \tanh\left(\beta \sum_j J_{ij} \, m_j^{(t)} \right) \qquad \text{Mean Field}$$

$$m_i^{(t+1)} = \tanh\left(\beta \sum_j J_{ij} \, m_j^{(t)} - \beta^2 m_i^{(t-1)} \sum_j J_{ij}^2 (1 - (m_j^{(t)})^2) \right) \qquad \text{TAP}$$

[Solution]

We first need an expression for $H_{now} - H_{old}$ for a single-variable flip. Since we may try different β, N, to avoid confusion it is better to write them explicitly. For any $k \in 1, \ldots, N$, we can split the Hamiltonian as:

$$
\begin{aligned}
H(\mathbf{s}) \;&=\; -\sum_{i<j} J_{ij} s_i s_j & (2) \\[2mm]
&=\; -\underbrace{\sum_{\substack{i>j \\ i \neq k, j \neq k}} J_{ij} s_i s_j}_{\text{do not contain } s_k} - s_k \underbrace{\sum_{i \neq k} J_{ki} s_i}_{\text{contains } s_k} & (3)
\end{aligned}
$$

Suppose we flip the k-th spin; then the only difference is $s_k^{\text{new}} = -s_k^{\text{old}}$, yielding:

$$H_{\text{now}} - H_{\text{flip}} \quad = \quad H(\mathbf{s}_{old}) - H(\mathbf{s}_{new}) \tag{4}$$

$$= \quad \left[-s_k^{old} \sum_{i \neq k} J_{ki} s_i^{old} \right] - \left[-s_k^{new} \sum_{i \neq k} J_{ki} s_i^{old} \right] \tag{5}$$

$$= \quad \left[-s_k^{old} \sum_{i \neq k} J_{ki} s_i^{old} \right] - \left[s_k^{old} \sum_{i \neq k} J_{ki} s_i^{old} \right] \tag{6}$$

$$= \quad -2 s_k^{old} \sum_{i \neq k} J_{ki} s_i^{old} \tag{7}$$

Exercise 2: planted SK model

We consider a model similar to the SK model, with the difference that the J_{ij} are generated from a particular realization, called "planted". Specifically, we start from the following generative model for the $P(\mathbf{s}, \mathbf{J})$:

$$s_i \quad \sim \quad \frac{1}{2} \delta(s_i - 1) + \frac{1}{2} \delta(s_i + 1) \quad , \quad i = 1, \ldots, N \tag{8}$$

$$J_{ij} | s_i, s_j \quad \sim \quad \mathcal{N}\left(\frac{s_i s_j}{\sqrt{N}}, \sigma^2 \right) \quad , \quad (i, j) \in N^2 \tag{9}$$

Goal: estimate the planted $\mathbf{s}$ given the $\mathbf{J}$.

(a) Use Bayes' theorem to write the posterior distribution $P(\mathbf{s}|\mathbf{J})$.

(b) Rewrite it as a Boltzmann distribution similar to the SK model with $\beta := \frac{1}{\sigma^2}$.

(c) We would like to estimate the mean of $P(\mathbf{s}|\mathbf{J})$. However, this is not tractable analytically. We will instead use the approximation introduced in the class.
Write in a *jupyter* notebook a function to *sample* an instance of $(\mathbf{s}, \mathbf{J})$.

(d) In the same notebook, write a function that implements the TAP equation to approximate the mean $\hat{\mathbf{s}}$ of $P(\mathbf{s}|\mathbf{J})$. This is an iteration that, if it converges, gives a very good approximation for $\hat{\mathbf{s}}$ as $N \to \infty$.

For numerical reasons implement the fixed point iterations as follows:

$$m_i^{(t+1)} = \tanh\left(\frac{1}{\sigma^2 \sqrt{N}} \sum_j J_{ij} m_j^{(t)} \right) \qquad \text{Mean Field}$$

$$m_i^{(t+1)} = \tanh\left(\frac{1}{\sigma^2 \sqrt{N}} \sum_j J_{ij} m_j^{(t)} - m_i^{(t-1)} \frac{1}{N\sigma^4} \sum_j J_{ij}^2 \left(1 - (m_j^{(t)})^2\right) \right) \qquad \text{TAP}$$

(e) Run some experiments ($N_{real} \in [10, 100]$ re-samplings of J, s at your choice) for $N = 10, 100, 1000, 5000$ and fixed $\sigma^2 = 0.1$ and check that the overlap[1] with ground-truth improves with the iterations.

---

[1]The overlap is defined as: $overlap(\mathbf{m}, \mathbf{s}_0) := |\frac{\mathbf{m} \cdot \mathbf{s}_0}{N}|$

(f) i) Run sum experiments ($N_{real} \in [10, 100]$ re-samplings of J, s at your choice) for $N = 10, 100, 1000, 5000$ and varying $\sigma^2 \in [0.1, 2]$.
ii) Repeat the same experiments but using the MF approximation instead.
iii) Plot the performance metrics values at convergence for TAP and MF as a function of the noise σ^2 for various N.
Comment on what you observe.

[Solution]

(a) The posterior distribution is:

$$P(\mathbf{s}|\mathbf{J}) \quad = \quad \frac{P(\mathbf{J}, \mathbf{s})}{P(\mathbf{J})} \tag{10}$$

$$\propto \quad \prod_{ij} \mathcal{N}(s_i s_j, \sigma^2) \prod_i p_0(s_i) \tag{11}$$

$$p_0(s_i) \quad = \quad \frac{1}{2} \delta(s_i - 1) + \frac{1}{2} \delta(s_i + 1) \tag{12}$$

(b) It can be rewritten as:

$$P(\mathbf{s}|\mathbf{J}) \propto e^{\beta \sum_{ij} J_{ij} s_i s_j} \tag{13}$$

3