

Advanced Probabilistic Machine Learning and Applications

Alessandro Lonardi and Caterina De Bacco

November 23, 2021

Contents

1	Tutorial 5: TAP approach	2
1.1	Exercise 1: Sampling from the SK model	2
1.2	Exercise 2: planted SK model	3
2	Errata	4

1 Tutorial 5: TAP approach

1.1 Exercise 1: Sampling from the SK model

Consider the Hamiltonian of the SK model with $h = 0$:

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

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) algorithm as we learned for the Curie-Weiss model in a previous tutorial.

Perform the following tasks.

- (a) Write a code to perform the MCMC dynamics:
 - (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.) Start by configurations of spins extracted uniformly at random. Run your dynamics for a long enough time (say, with $t_{\max} = 10^2 N$ 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.

For numerical reasons implement the fixed point iterations as follows:

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

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

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

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

$$= - \underbrace{\sum_{\substack{i > j \\ i \neq k, j \neq k}} J_{ij} s_i s_j}_{\text{do not contain } s_k} - s_k \underbrace{\sum_{i \neq k} J_{ki} s_i}_{\text{contains } s_k} \quad (2)$$

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

$$H_{\text{now}} - H_{\text{flip}} = H(\mathbf{s}_{\text{old}}) - H(\mathbf{s}_{\text{new}}) \quad (3)$$

$$= \left[-s_k^{\text{old}} \sum_{i \neq k} J_{ki} s_i^{\text{old}} \right] - \left[-s_k^{\text{new}} \sum_{i \neq k} J_{ki} s_i^{\text{old}} \right] \quad (4)$$

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

$$= -2s_k^{\text{old}} \sum_{i \neq k} J_{ki} s_i^{\text{old}} \quad (6)$$

1.2 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 $P(\mathbf{s}, \mathbf{J})$:

$$s_i \sim \frac{1}{2} \delta(s_i - 1) + \frac{1}{2} \delta(s_i + 1), \quad \forall i = 1, \dots, N \quad (7)$$

$$J_{ij} | s_i, s_j \sim \mathcal{N}\left(\frac{s_i s_j}{\sqrt{N}}, \sigma^2\right), \quad \forall (i, j) \in N^2 \quad (8)$$

Perform the following tasks.

- Use Bayes’ theorem to write the posterior distribution $P(\mathbf{s} | \mathbf{J})$.
- Rewrite it as a Boltzmann distribution similar to the SK model with $\beta := \frac{1}{\sigma^2}$.
- 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 a function to *sample* an instance of (\mathbf{s}, \mathbf{J}) .
- 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 $N \rightarrow \infty$.

For numerical reasons implement the fixed point iterations as follows:

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

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

- Run some experiments ($N_{\text{real}} \in [10, 100]$ re-samplings of J, s at your choice) for $N = 10, 100, 1000$ and varying $\sigma^2 \in [0.1, 2]$. Plot the overlap $\text{ov} := |\sum_i s_i m_i|/N$ (where s_i are the sampled spins, and m_i the magnetization at convergence) for TAP and MF as a function of the noise σ^2 for various N . Comment on what you observe.

(a) The posterior distribution is:

$$P(\mathbf{s}|\mathbf{J}) = \frac{P(\mathbf{J}, \mathbf{s})}{P(\mathbf{J})} \quad (9)$$

$$\propto \prod_{ij} \mathcal{N}\left(\frac{s_i s_j}{\sqrt{N}}, \sigma^2\right) \prod_i p_0(s_i) \quad (10)$$

$$p_0(s_i) = \frac{1}{2} \delta(s_i - 1) + \frac{1}{2} \delta(s_i + 1) \quad (11)$$

(b) It can be rewritten as:

$$\begin{aligned} P(\mathbf{s}|\mathbf{J}) &\propto \prod_{ij} \exp\left(-\frac{1}{2\sigma^2} \left(J_{ij} - \frac{s_i s_j}{\sqrt{N}}\right)^2\right) \\ &\propto \prod_{ij} \exp\left(-\frac{1}{2\sigma^2} \left(J_{ij}^2 - 2J_{ij} \frac{s_i s_j}{\sqrt{N}} + \frac{1}{N}\right)\right) \\ &\propto \exp\left(\frac{\beta}{\sqrt{N}} \sum_{ij} J_{ij} s_i s_j\right) \end{aligned}$$

2 Errata

Two notes about what has been said in class:

- the computations 1.2b are indeed right. I added all the steps for clarity.
- when computing the overlap with the MF method, in 1.2e, the standard deviations are not zero, but low.