Advanced Probabilistic Machine Learning and Applications

Alessandro Lonardi and Caterina De Bacco

November 9, 2021

Contents

1 Tutorial 4: Mean field approach

1.1 Exercise 1: Curie-Weiss model

The Hamiltonian of the Curie-Weiss model can be written as:

$$H(\mathbf{s}) = -\frac{J}{2N}\sum_{i,j} s_i s_j - h\sum_i s_i. \tag{1}$$

From the lecture, we know that the Mean Field free variational energy is:

$$F[Q] = -\frac{J}{2}m^2 - hm + \frac{1}{\beta}\left[\frac{1+m}{2}\log\left(\frac{1+m}{2}\right) + \frac{1-m}{2}\log\left(\frac{1-m}{2}\right)\right]$$

where $m := \mathbb{E}_Q\left[\sum_i s_i/N\right]$ is the expected total magnetization.
Perform the following tasks.

1. Plot the function $F[Q]$ as a function of m in the two following cases:
   (i) at $h = 0$, for different values of J larger and lower than 1;
   (ii) at a value of J larger than 1 for different values (positive and negative) of h.
   Describe what you see in both cases.

1. The plot of function $F[Q]$ with $h = 0$ and $J = 0.9, 1, 1.2, 1.5$ are shown in Fig. (1).

   For small couplings ($J < 1$), the curve has only one minimum in zero. For high couplings ($J > 1$), the curve has two local symmetric minimums around zero. There is a phase transition at $J = 1$.

   The plot of function $F[Q]$ with $J = 1.2$ and $h = -0.1, -0.05, 0, 0.03, 0.07$ is shown in Fig. (2).

   Since $J > 1$, the minimum drifts away from zero. The number of local minima can be one or two depending on the value of $|h|$. If $h < 0$, the global minimum will be closer to -1, if $h > 0$, the global minimum will be closer to 1. When h moves past zero, the global minimum changes sharply to from a value to another with opposite sign.

Figure 1: Gibbs free energy Curie-Weiss for varying parameters.

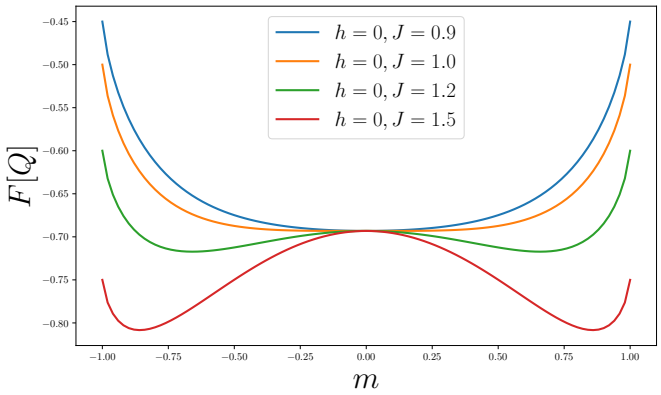
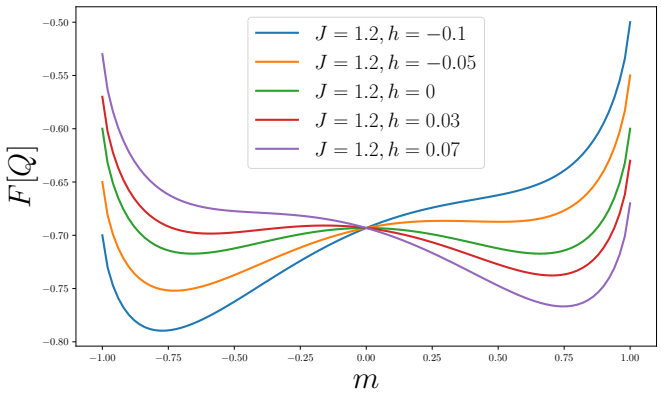


Figure 2: Gibbs free energy Curie-Weiss for varying external field.



2. The minimizer m^* is also the solution of the self-consitency equation:

$$m = \tanh\left[\frac{J}{2}m + h\right].$$

Compute the value of m^* in the three following cases:
(i) $h = 10^{-6}$ and J between 0 and 2;
(ii) $h = -10^{-6}$ and J between 0 and 2;
(iii) $J = 1.5$ and h between -1 and 1.

2. The plot of function $m^*(\beta, h)$ under $h = \pm 10^{-6}$ and $J \in (0, 2]$ is shown in Fig. (3).
The plot of function $m^*(\beta, h)$ under $J = 1.5$ and $h \in [-1, 1]$ is shown in Fig. (4).

3. Focusing on the case $J = 1.5$. With $h = 0.1$, 0.2, how many solutions to the self-consistency equations are there? Which one is the correct one? (Plot the function $F[Q]$ to answer to these questions).

3. The plot of function $F[Q]$ under $J = 1.5$ and $h = 0.1, 0.2$ is shown in Fig. (5).
For $h = 0.1$, there are two solutions, the right one has lower function value thus it is the correct one. For $h = 0.2$, there is only one solution, so it is the correct one.

Figure 3: Magnetization Curie-Weiss for varying couplings.

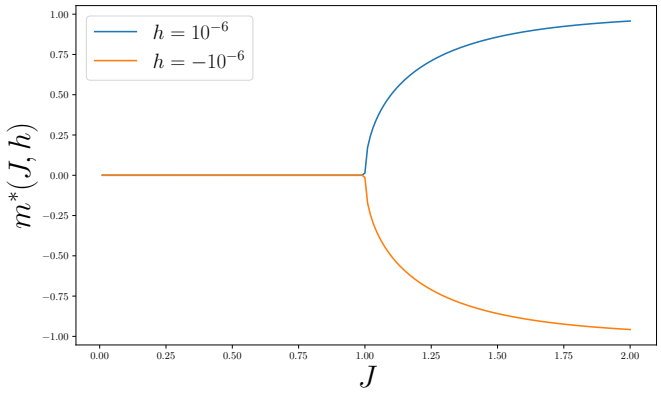


Figure 4: Magnetization Curie-Weiss for varying external field.

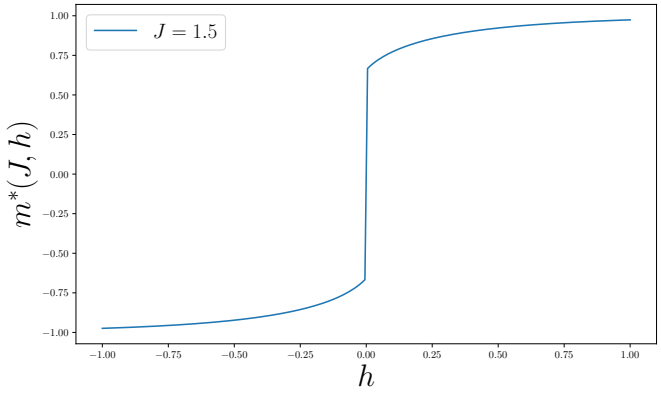
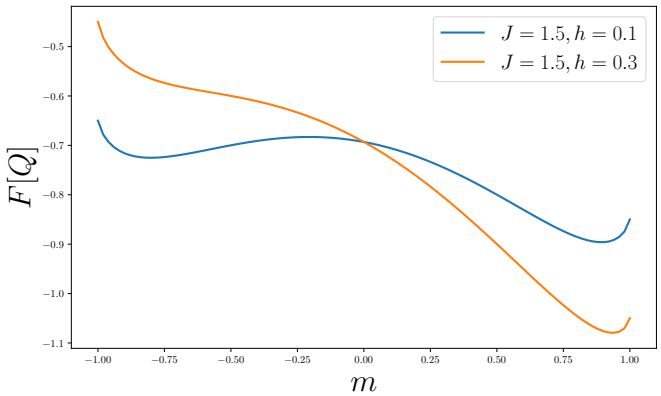


Figure 5: Gibbs free energy Curie-Weiss for varying external field.



1.2 Exercise 2: Sampling from the Curie-Weiss model (MCMC method)

Consider again the Curie-Weiss Hamiltonian in Equation (1). A practical way to sample configurations of N spins from the Gibbs probability distribution:

$$P(\mathbf{s}) = \frac{e^{-\beta H(\mathbf{s})}}{Z},$$

is the Monte-Carlo-Markov-Chain (MCMC) method, and in particular the Metropolis-Hastings algorithm. This works as follows:

1. Choose a starting configuration for the N spins values $s_i = \pm 1$, for $i = 1, \ldots, N$
2. Choose a spin i at random. Compute the current value of the energy H_{now} and the value of the energy H_{flip} if the spins i is flipped (that is if $s_i^{\text{new}} = -s_i^{\text{old}}$)
3. Sample a number $r \sim U(0, 1)$
4. **if** $r < \exp[\beta(H_{\text{now}} - H_{\text{flip}})]$ **then:** perform the flip (i.e. $s_i^{\text{new}} = -s_i^{\text{old}}$)
5. **if** $r \geq \exp[\beta(H_{\text{now}} - H_{\text{flip}})]$ **then:** leave s_i as it is
6. Go back to step 2

If one is running this program long enough, it is guaranteed that the final configuration of spins will have been chosen with the correct probability.

Perform the following tasks.

1. Write a code to implement the MCMC dynamics.
2. Run the code starting from a configuration where all spins are equal to 1. Take $h = 0$, $J = 1$, $\beta = 1.2$ and run your dynamics for a time that is long enough (say, with $t_{\max} = 100N$ attempts to flips spins). Monitor the value of the magnetization per spin $m = \sum_i s_i/N$ as a function of time. Make a plot for $N = 10, 50, 100, 200, 1000$ spins.
   Compare with the exact solution at $N = \infty$.
   Comment.

We first need an expression for $H_{\text{now}} - H_{\text{flip}}$ for a single-variable flip. For any $k \in 1, \ldots, N$, we can split the Hamiltonian as:

$$\begin{aligned} H(\mathbf{s}) &= -\frac{J}{N}\sum_{i<j} s_i s_j - h\sum_{i=1}^{N} s_i && (2)\\ &= \underbrace{-\frac{J}{N}\sum_{\substack{i<j\\ i\neq k, j\neq k}} s_i s_j - h\sum_{i\neq k} s_i}_{\text{do not contain } s_k} \underbrace{-\frac{J}{N} s_k \sum_{i\neq k} s_i - h\, s_k}_{\text{contains } s_k} && (3) \end{aligned}$$

Suppose we flip the k-th spin; then the only difference in the Hamiltonians H_{now} and H_{flip} is given by the flip $s_k^{\text{new}} = -s_k^{\text{old}}$, yielding:

$$\begin{aligned} H_{\text{now}} - H_{\text{flip}} &= H(\mathbf{s}_{old}) - H(\mathbf{s}_{new}) && (4)\\ &= \left[-\frac{J}{N} s_k^{\text{old}} \sum_{i\neq k} s_i^{\text{old}} - h s_k^{\text{old}}\right] - \left[-\frac{J}{N} s_k^{\text{new}} \sum_{i\neq k} s_i^{\text{new}} - h s_k^{\text{new}}\right] && (5)\\ &= \left[-\frac{J}{N} s_k^{\text{old}} \sum_{i\neq k} s_i^{\text{old}} - h s_k^{\text{old}}\right] - \left[\frac{J}{N} s_k^{\text{old}} \sum_{i\neq k} s_i^{\text{old}} + h s_k^{\text{old}}\right] && (6)\\ &= -2 s_k^{\text{old}} \left[\frac{J}{N}\sum_{i\neq k} s_i^{\text{old}} + h\right] && (7) \end{aligned}$$

If $N \to \infty$, the Boltzmann distribution is dominated by configurations whose magnetization equals m^*, the minimum of the Hamiltonian. So if we can run a simulation with infinite N, the

plot of m will start with a warm-up stage to reach m^*, once m^* is reached, the plot will get stuck there.

1. Since $h = 0$, there the two local minimizers have same free energy value so either one can achieve the global minimum. Here, we start our MCMC chain at all-one configuration, which is closer to the positive minimizer. However, it is possible that the random walk will reach the negative minimizer. In the long run, the chain will first walk to the positive minimizer and stay there for a while, then walk to the negative minimizer and stay there for a while. This will transition will keep happening back and forth. The larger N is, the longer the chain stay at the minimizers, and the less fluctuation the chain will deviated from the minimizer.

2. Start by a configuration where all spins are equal to 1 and take $h = -0.1$, $J = 1$, $\beta = 1.2$. Monitor again the value of the magnetization per spin $m = \sum_i s_i/N$ as a function of time. Make a plot for $N = 10, 50, 100, 200, 1000$ spins.
   Compare with the exact solution at $N = \infty$.
   Comment.

2. In this case we have similar behavior as for point 1. The only difference is that setting $h = -0.1$ breaks down the balance between the two local minimizers. As discussed in Point 2c. The negative minimizer is the correct one, which is associated to the global minima of the free energy. Again, the trace plot will start with a warm-up stage to reach m^*, once m^* is reached, the trace plot will get stuck at there. The larger N is, the less fluctuation the chain will deviated from m^*.

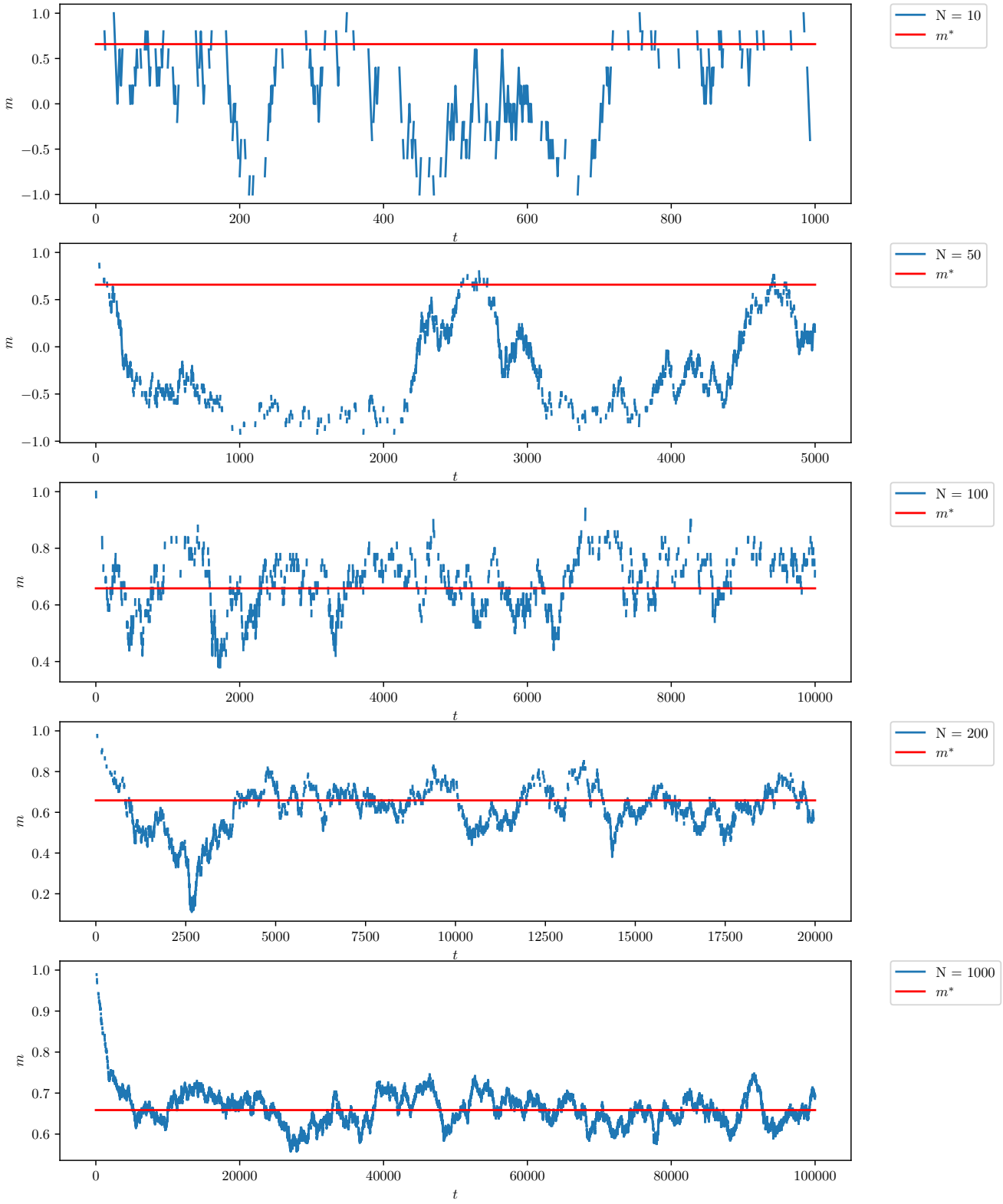


Figure 6: Instantaneous magnetization for $\beta = 1.2, h = 0$ and $N = 10, 50, 100, 200, 1000$

Figure 7: Instantaneous magnetization for $\beta = 1.2, h = -0.1$ and $N = 10, 50, 100, 200, 1000$