probability theory → Markov chains ★★★★★

Fast MCMC for the Ising Model

Simple MCMC algorithms for drawing samples from the Ising model, such as Metropolis or the Gibbs sampler, are extremely inefficient for modeling the Ising model, particularly near the critical temperature. In this problem, we will explore 2 algorithms for fast MCMC for the Ising model. The key idea with both methods will be the introduction of an enlarged state space, and using this to flip clusters of spins at a time, instead of single spins.

Consider the graph $G = (V, E)$; the state space $\Omega_0 = \{\pm 1\}^V$ for the classic simulations. However, we will enlarge this to the following state space: $\Omega = \{\pm 1\}^V \cup \{0, 1\}^E$. Now, elements of the state space consist of (s_i, b_{ij}), with s_i the spins at each site and b_{ij} auxiliary bond variables which we will introduce shortly, defined as being either 0 or 1 on each edge. In this problem, you should take the Ising model to have the equilibrium distribution

$$\mathrm{p}_{\mathrm{eq}}(s_i) \sim \exp\left[\frac{1}{2T}\sum_{ij\in E} s_i s_j\right].$$

We begin by analyzing the following algorithm, called the **Swendsen-Wang algorithm**. Now, we are going to introduce a Markov chain with a stationary distribution $\mathrm{p}_{\mathrm{eq}}(s_i, b_{ij})$ with the marginal $\mathrm{p}_{\mathrm{eq}}(s_i)$ the same as for the Ising model. Suppose we defined a Markov chain with equilibrium distribution

$$\mathrm{p}_{\mathrm{eq}}(s_i, b_{ij}) = \prod_{\substack{ij\in E \\ s_i \neq s_j}} \mathbf{1}(b_{ij} = 0) \times \prod_{\substack{ij\in E \\ s_i = s_j}} \left[\mathbf{1}(b_{ij} = 0) + \alpha \mathbf{1}(b_{ij} = 1)\right].$$

(a) Show that, for a special choice of α, the marginal distribution $\mathrm{p}_{\mathrm{eq}}(s_i)$ is the equilibrium distribution of the Ising model. What choice of α gives a specific T?

(b) Given some (s_0, b_0), the Swendsen-Wang algorithm proceeds as follows. We first choose $(b_{ij})_1$ according to the distribution $\mathrm{p}_{\mathrm{eq}}(b_{ij}|(s_i)_0)$; then we choose $(s_i)_1$ according to the distribution $\mathrm{p}_{\mathrm{eq}}(s_i|(b_{ij})_1)$. What are these conditional probability distributions?

(c) Why is the Swendsen-Wang algorithm a Markov chain with the specified equilibrium distribution – is it the unique equilibrium distribution?

(d) Code up the Swendsen-Wang algorithm for an arbitrary graph, and present your code.

(e) Now, choose the graph to be the discrete torus $\mathbb{Z}_L \times \mathbb{Z}_L$, with edges to the 4 nearest neighbors. Try to determine how long this algorithm will take to reach equilibrium, using a method of your choice. Plot some curves of the mixing time τ for the Markov chain vs. L, at varying T. Comment on what happens, and compare to the results for the single-step algorithms.

Now, we consider a slight variant of the above algorithm – the **Wolff algorithm**. The Wolff algorithm proceeds as follows. We pick a random spin on the graph, and create a cluster recursively in the following way: for each edge from the chosen vertex, if the pair of spins on the nodes is aligned, add the new spin to the cluster with probability $1 - \mathrm{e}^{-2/T}$. Once a cluster has finished growing in this way, we flip the spin of the entire cluster.

(f) Show that the Ising stationary distribution is reversible for a transition matrix given by the above algorithm.

(g) Implement this algorithm.

(h) Repeat similar numerical experiments to the previous part, and compare the performance of this algorithm to the Swendsen-Wang algorithm.

(i) What advantage do you think this algorithm has over the Swendsen-Wang algorithm (it was proposed 2 years after)?