Brief Notes on Monte Carlo for Spin Systems

Ethan Huecker

July 2024

In stat-mech, for a system in equilibrium with its environment at temperature T, the expectation value of some observable $\mathcal{O}$ is given by

$$\langle \mathcal{O} \rangle = \sum_c \mathcal{O}(c) P(c) = \frac{1}{\mathcal{Z}} \sum_c \mathcal{O}(c) e^{-\beta E(c)}, \tag{1}$$

where $\beta = 1/k_B T$ and $\mathcal{Z}$ is the canonical partition function. Here, $\{c\}$ is a set of different configurations for the system. In the case of the Ising model of N spins, there are 2^N configurations, and this sum is not tractable. In many cases, $P(c)$ for some given c is quite small and can be neglected, and so instead of summing over all configurations we can approximate the average over a subset of configurations sampled from the distribution $P(c)$ itself.

Ex: Say we average a function $f(x),\ x \in [a, b]$ over a distribution $p(x)$. The average is

$$\langle f \rangle = \int_a^b f(x) p(x) dx. \tag{2}$$

Let's approximate this by sampling N values from $p(x)$, x_i, and so $p(x)dx$ can be viewed as the probability of drawing $x_n \in [x, x + dx]$. All that we need to do is take an empirical average of $f(x)$ over these values,

$$\langle f \rangle \approx \frac{1}{N} \sum_{i=1}^{N} f(x_i), \quad x_i \sim p(x). \tag{3}$$

The right-hand side is known as the Monte Carlo estimator, which in the limit $N \to \infty$ is the true average. In particular, its average is precisely (2):

$$\left\langle \frac{1}{N} \sum_{i=1}^{N} f(x_i) \right\rangle = \frac{1}{N} \sum_{i=1}^{N} \langle f(x_i) \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle f(x) \rangle_p = \frac{1}{N} \sum_{i=1}^{N} \int_a^b f(x) p(x) dx = \int_a^b f(x) p(x) dx, \tag{4}$$

as it should be.

How do we generate the configurations/states for spins defined on a lattice? We construct a *Markov chain* whose elements' distribution approximates $\exp(-\beta E(c))/\mathcal{Z}$; the Markov chain's equilibrium distribution matches the target.[1] The more steps in the chain, the closer this match becomes. There are many methods in constructing Markov chains, but for spin-systems obeying a Boltzmann like distribution, the *Metropolis-Hastings* algorithm (MHA) is most common. The idea of MHA is to start with a random spin configuration, then propose a (local) change of a lattice site (say the flipping of a spin), and compute the energy difference ΔE between these configurations. If $\Delta E < 0$, we accept this configuration with probability 1, and if $\Delta E > 0$ we accept only with probability $\exp(-\beta \Delta E)$. By not totally disregarding unfavorable transitions $\Delta E > 0$ we are able to account for thermal fluctuations. This acceptance condition can be written as $\min(1, \exp(-\beta \Delta E))$, and satisfies the condition for detailed balance.

By "accepting" the proposal, I mean that under some $c \to c'$, if $\Delta E < 0$ we formally set $c \to c'$, then repeat. If $\Delta E > 0$, we first generate a random number $r \in [0, 1]$, then if $r < \exp(-\beta \Delta E)$ we formally set $c \to c'$. This process is known as the *Metropolis step*. Thermalization involves performing multiple sweeps of these Metropolis steps over each lattice site until the distribution of the configurations is well approximated by $\exp(-\beta E(c))/\mathcal{Z}$. Once the system is thermalized, we can continue performing Metropolis steps, but now each subsequent configuration is effectively taken from a Boltzmann distribution. For each of these we begin evaluating the desired observable $\mathcal{O}$, sum them all up, and divide by the number of steps n performed after thermalization:

$$\langle \mathcal{O} \rangle \approx \frac{1}{n} \sum_c \mathcal{O}(c), \tag{5}$$

[1] A Markov chain is a sequence of configurations where the probability of moving to the next state depends only on the current state, not the full history.

which is the equivalent of the example (3). Note that the amount of steps required for thermalization is not given, and depends on the system in consideration. Also, we don't have to start with a random spin configuration, as if we wish to evaluate an observable in the Ising model for low-T we can start with a lattice with most of the spins aligned.