A very brief introduction to simulating the Ising model using an MCMC method

Radoslav Harman

December 30, 2015

1 Metropolis-Hastings algorithm

The aim of the Metropolis-Hastings algorithm is to generate samples from a "target" probability π on a finite set $\mathcal{X}$. The set $\mathcal{X}$ is often called the "space of states" and the elements are called "states", which is a terminology adopted from physics. For each state $x \in \mathcal{X}$, let $q(\cdot|x)$ be a "candidate", or "proposal" probability on $\mathcal{X}$ and let $X_1 \equiv x_1 \in X$ be an initial state.

Under very general assumptions, the following algorithm produces a sequence of $(X_i)_{i=1}^{\infty}$ of random variables on $\mathcal{X}$ which converges to the distribution π.

1. Set $t \leftarrow 1$.
2. Generate a "candidate" Y from the distribution $q(\cdot|X_t)$.
3. Calculate
$$\alpha \leftarrow \min\left(1, \frac{\pi(Y)q(X_t|Y)}{\pi(X_t)q(Y|X_t)}\right).$$
4. Generate random variable $U \sim U(0,1)$.
5. If $U < \alpha$ set $X_{t+1} \leftarrow Y$, otherwise set $X_{t+1} \leftarrow X_t$.
6. Set $t \leftarrow t+1$ and continue by step 2.

The algorithm is most often used with symmetric candidate probabilities satisfying $q(y|x) = q(x|y)$ for all $x, y \in \mathcal{X}$, in which case it is called just the Metropolis algorithm. (Sometimes $q(y|x)$ is the same for all $x, y \in \mathcal{X}$, which we call the "independence sampler".) In this case, the formula for α reduces to a simpler form and the behaviour of the algorithm is easier to interpret and understand: The algorithm will always make a transition to a state y with a higher probability than x, but it can also make a transition to a state y with a lower probability than x; the probability of such a transition is equal to $\pi(y)/\pi(x)$.

Note that for the Metropolis-Hastings algorithm it is not necessary to know the "normalising constant" of the distribution π. Metropolis-Hastings algorithm is often useful for a very large state space $\mathcal{X}$ and an complex probability π, or if we do not know the normalizing constant of π, which means that the direct sampling methods are infeasible.

Metropolis-Hastings algorithm is a special MCMC (Markov-chain Monte Carlo) method; clearly, the sequence $(X_i)_{i=1}^{\infty}$ of random states forms a Markov chain. If q is chosen such that this Markov chain is ergodic, we obtain the desired convergence to π. At the beginning of algorithm, we sometimes use the so-called "burn-in" period.

2 Ising model

The Ising model is a model of ferromagnetism in statistical physics. Consider a state space $\mathcal{X}$ formed by all matrices $B \in \{-1, +1\}^{N \times N}$. Note that the size of $\mathcal{X}$ is 2^{N^2}, which is an astronomic number even for moderate N. The individual elements of each state B, called "sites", represent magnetic dipole moments of atomic spins. We will assume that two different sites of the $N \times N$ lattice are "adjacent" if both their coordinates differ by at most 1.

The probability of a state B is a function of its "energy", which is in the case of zero external magnetic field and the same, unit interaction strength given by the formula

$$H(B) = \sum -B(i_1, j_1)B(i_2, j_2),$$

where the sum is taken with respect to all unordered couples (i_1, j_1), (i_2, j_2) of sites adjacent in the lattice. Note that two adjacent sites with the opposite spin contribute to the sum by one and two adjacent sites with the same spin reduce the sum by one. The probability of a state B is, using the Boltzmann law

$$\pi(B) = e^{-\beta H(B)}/Z_\beta,$$

where $\beta = (Tk_B)^{-1}$, T is the temperature of the system, k_B is the Boltzmann constant and Z_β is the normalization constant.

To generate a random state in accord with π, we need to construct a candidate distribution on $\mathcal{X}$. A simple symmetric candidate probability '$q(.|B_x)$ is the uniform distribution on all states By which differ from B_x at exactly one site. Now we can calculate the acceptance probability α, which is a minimum of 1 and

$$\frac{\pi(B_y)}{\pi(B_x)} = e^{-\beta(H(B_y)-H(B_x))} = \exp\left[\beta\left(\sum B_y(i_1,j_1)B_y(i_2,j_2) - B_x(i_1,j_1)B_x(i_2,j_2)\right)\right],$$

where again the sum is taken with respect to all unordered pairs (i_1, j_1), (i_2, j_2) of sites adjacent in the lattice. However, note that the two states differ in only one site, say (i,j). Therefore, the differences $B_y(i_1,j_1)B_y(i_2,j_2) - B_x(i_1,j_1)B_x(i_2,j_2)$ will be zero for all adjacent pairs of sites, except for those that one of the sites is equal to $(i.j)$. Therefore, the sum above is equal to

$$\sum B_y(i_1,j_1)B_y(i,j) - B_x(i_1,j_1)B_x(i,j),$$

where the sum is taken with respect to (i_1,j_1) adjacent to (i,j). But $B_y(i,j) = -B_x(i,j)$ and $B_y(i_1,j_1) = B_x(i_1,j_1)$ therefore the sum is equal to

$$-2B_x(i,j)\sum B_x(i_1,j_1).$$

Summarizing, we have

$$\frac{\pi(B_y)}{\pi(B_x)} = \exp\left[-2\beta B_x(i,j)\sum B_x(i_1,j_1)\right],$$

where the sum is taken with respect to (i_1,j_1) adjacent to (i,j).

During the lecture, we will demonstrate the Metropolis-Hastings algorithm applied to the simplest Ising model in the environment R.