Markov-Chain Monte-Carlo for the 2D Ising model

Jannik Daun

March 7, 2024

Contents

1 The 2D Ising Model **1**

2 Applying Metropolis-Hastings to the 2D Ising Model **2**

3 Metropolis-Hastings Algorithm **4**

1 The 2D Ising Model

The two dimensional Ising model is a model for magnetism in crystals. Atoms are placed at the sites of a lattice $\Lambda := \{1, \dots, N\}^2$. Each atom $i \in \Lambda$ has a spin $x(i) \in \{-1, 1\}$. The state space of the system is therefore $\{-1, 1\}^\Lambda$. The atoms at the edges of the lattice are connected together forming a torus so that every atom has four neighbours (periodic boundary conditions). The local energy H^i_{loc} of the atom $i \in \Lambda$ in the state $x \in \{-1, 1\}^\Lambda$ is defined by

$$H^i_{\text{loc}}(x) := -\sum_{j \in \Lambda : i \sim j} x(i)x(j),$$

where $i \sim j$ means that i and j are neighbours on the lattice. The energy of a lattice configuration $x \in \{-1, 1\}^\Lambda$ is then given by the Hamiltonian

$$H(x) = \frac{1}{2} \sum_{i \in \Lambda} H^i_{\text{loc}}(x).$$

To every link between neighbouring atoms we can associate the negative of the product of the spin values of these atoms. Then $H(x)$ is just the sum of the link values over all links in the state x.
The thermal equilibrium distribution $\pi : \{-1, 1\}^\Lambda \to (0, 1]$ of the crystal at the fixed inverse temperature $\beta > 0$ is given by

$$\pi(x) = \frac{1}{Z} e^{-\beta H(x)}, \tag{1}$$

where Z is a normalisation constant. The goal is to approximate the expectation value of the absolute value $|m|$ of the magnetisation m per lattice site, which is defined by

$$|m| = \mathbb{E}_\pi \left(x \mapsto \frac{1}{N^2} \Big| \sum_{i \in \Lambda} x(i) \Big| \right). \tag{2}$$

Monte-Carlo integration is used to this end: The idea is to generate a finite number of lattice states $x_1, \ldots, x_n \in \{-1,1\}^\Lambda$ that have (approximately) the distribution π and to then estimate

$$|m| \approx \frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{N^2}\left|\sum_{i\in\Lambda} x_j(i)\right|\right) \tag{3}$$

using the law of large numbers. To generate lattice states that have approximately the distribution π we will use the Metropolis-Hastings algorithm.

2 Applying Metropolis-Hastings to the 2D Ising Model

This section assumes knowledge on the Metropolis-Hastings algorithm. An explanation of the Metropolis-Hastings algorithm can be found in section 3.

The strict positivity condition placed on π to use the Metropolis-Hastings algorithm is satisfied. Next the reference Markov-Kernel q needs to be defined. To this end define x^i for $x \in \{-1,1\}^\Lambda$ by

$$x^i(j) = \begin{cases} x(j), \text{ if } i \neq j \\ -x(j), \text{ if } i = j. \end{cases}$$

That is x^i is simply the state x but with the spin flipped at the lattice site i. Now define the reference Markov kernel $q : \{-1,1\}^\Lambda \times \{-1,1\}^\Lambda \to [0,1]$ by

$$q(x,y) = \begin{cases} N^{-2}, \text{ if } y = x^i \text{ for some } i \in \Lambda, \\ 0, \text{ else.} \end{cases} \tag{4}$$

That is $q(x,-)$ is simply the uniform distribution on all lattice states that can be reached from the state x with a single spin flip.

It can be verified that q satisfies the conditions outlined in the preceding section to produce an ergodic Markov-Chain. To use the Metropolis-Hastings algorithm the following expression (see equation 7) needs to be calculated for $x, y \in \{-1,1\}^\Lambda$ with $x \neq y$ and $q(x,y) \neq 0$:

$$\min\left\{\frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}, 1\right\} = \min\left\{e^{-\beta(H(y)-H(x))}, 1\right\}.$$

Now y must be of the form $y = x^i$ for some $i \in \Lambda$ since $q(x,y) \neq 0$. Therefore

$$\begin{aligned} &H(y) - H(x) = H^i_{\text{loc}}(x^i) - H^i_{\text{loc}}(x) \\ &= -\sum_{j\in\Lambda : i\sim j} x^i(i)x(j) + \sum_{j\in\Lambda : i\sim j} x(i)x(j) = -2H^i_{\text{loc}}(x), \end{aligned}$$

since the only links that can change their values are those connected to i.

Let $(\Omega, \mathbb{P})$[1] be the underlying probability space. For all $t \in \mathbb{N}$ let $I_t : \Omega \to \Lambda$ and $U_t : \Omega \to [0,1]$ be (respectivly) i.i.d. uniformly distributed random variables (and independant of each other). Let $X_0 : \Omega \to \{-1,1\}^\Lambda$ be a random variable

[1] Ω could for example be the set of all possible seeds used for the random number generator

with arbitrary distribution. In the code the delta distribution for the state where all spins are -1 is used.

For an outcome (a seed of the RNG) $\omega \in \Omega$ and all $t \in \mathbb{N}$ define $x_t := X_t(\omega), i_t := I_t(\omega), u_t := U_t(\omega)$. Then the Metropolis-Hastings algorithm to generate the state x_{t+1} from the state x_t is given by algorithm 1[2]. Algorithm

Algorithm 1 Metropolis-Hastings Algorithm for the Ising model

calculate i_{t+1}
calculate u_{t+1}
if $u_{t+1} \leq \min\{\exp(2\beta H_{\text{loc}}^{i_{t+1}}(x_t)), 1\}$ **then**
  $x_{t+1} = x_t^{i_{t+1}}$
else
  $x_{t+1} = x_t$
end if

1 can easily be improved by noticing that u_{t+1} only needs to be generated if $H_{\text{loc}}^{i_{t+1}}(x_t) < 0$, because if $H^{i_{t+1}}(x_t) \geq 0$ then the if statement will allways be true. This optimisation leads to algorithm 2. Furthermore there are only two

Algorithm 2 Optimized Metropolis-Hastings Algorithm for the Ising model

calculate i_{t+1}
if $H^{i_{t+1}}(x_t) \geq 0$ **then**
  $x_{t+1} = x_t^{i_{t+1}}$
else
  calculate u_{t+1}
  if $u_{t+1} \leq \exp(2\beta H_{\text{loc}}^{i_{t+1}}(x_t))$ **then**
    $x_{t+1} = x_t^{i_{t+1}}$
  else
    $x_{t+1} = x_t$
  end if
end if

values $H^{i_{t+1}}(x_t)$ can take if it is negative: -2 and -4. This can be used to map the value of $H_{\text{loc}}^{i_{t+1}}(x_t)$ to the value of $\exp(2\beta H_{\text{loc}}^{i_{t+1}}(x_t))$ instead of calculating it over and over again.

The code implements algorithm 2 with the additional optimization mentioned after it. The lattice configuration x_t is stored in a 2D array A_t (=a matrix) with dimensions $(N+2) \times (N+2)$, where $A_t[i,j] = x_t(i,j)$ for all $i,j \in \{1,\ldots,N\}$. The first and last row/column of A are only used to enforce the periodic boundary condition:

$$A_t[N+1,j] = A_t[1,j] \text{ and } A_t[0,j] = A_t[N,j]$$

for all $j \in \{1,\ldots,N\}$ and

$$A_t[i,N+1] = A_t[i,1] \text{ and } A_t[i,0] = x_t[i,N] \tag{5}$$

[2]When comparing to section 3: Note that the random variable defined by $Y_t(\omega) = X_{t-1}(\omega)^{I_t(\omega)}$ has distribution $q(X_{t-1}(\omega), -)$ and that U_t together with the inequality check is used in place of the Bernoulli random variable

for all $j \in \{1, \dots, N\}$. This makes it easy to calculate $H_{\text{loc}}^{(i,j)}(x_t)$ for any lattice site $(i,j) \in \Lambda$ as

$$H_{\text{loc}}^{(i,j)}(x_t) = -A_t[i,j] \cdot (A_t[i+1,j] + A_t[i-1,j] + A_t[i,j+1] + A_t[i,j-1]). \quad (6)$$

This technique has the downside that the first/last row or column have to be edited, whenever one of the lattice sites at the boundary of the lattice is flipped.

To use Metropolis-Hastings to calculate m the state x_t has to be calculated for large t so that the Metropolis-Hastings Markov-Chain is close to the equilibrium distribution and after that the subsequent states can be used to sample $|m|$ as described in the preceding section. In the code the state is also advanced by a large amount (not just 1) between taking samples to break the correlation between subsequent samples.

3 Metropolis-Hastings Algorithm

This section explains the Metropolis-Hastings algorithm. Some knowledge on (discrete, finite state space) Markov-Chains and probability theory is assumed. Let E be a finite set (the state space) and $(\Omega, \mathbb{P})$ a probability space. The Metropolis-Hastings algorithm defines an ergodic Markov chain $(X_t : \Omega \to E)_{t \in \mathbb{N}}$ that has a given probability distribution $\pi : E \to (0,1]$ as its limiting distribution.

Let $q : E \times E \to [0,1]$ be a Markov kernel. q will be called the reference kernel. Let $X_0 : \Omega \to E$ be a random variable with arbitary distribution. First consider $t = 1$ in the following:

Let $Y_t : \Omega \to E$ be a random variable distributed according to $q(X_{t-1}, \cdot)$ and $B_t : \Omega \to \{0,1\}$ a Bernoulli distributed random variable with success probability

$$s_t = \min\left\{\frac{\pi(Y_t)q(Y_t, X_{t-1})}{\pi(X_{t-1})q(X_{t-1}, Y_t)}, 1\right\}. \quad (7)$$

Now set

$$X_{t+1}(\omega) = \begin{cases} Y_{t+1}(\omega), \text{ if } B_{t+1}(\omega) = 1, \\ X_t(\omega), \text{ else.} \end{cases} \quad (8)$$

In this way X_t, Y_t, B_t can be recursively defined for all $t \in \mathbb{N}$.

$(X_t)_{t \in \mathbb{N}}$ is a Markov chain since the next state X_{t+1} is a function of the current state X_t and the random influences B_{t+1}, Y_{t+1} whose distribution only depends on the current state X_t. Furthermore $(X_t)_{t \in \mathbb{N}}$ is also a time homogenous Markov chain with Markov kernel p, which can be seen by constructing it: Per definition $p_{t,t+1}(x,y) = \mathbb{P}\{X_{t+1} = y | X_t = x\}$ (the conditional probability that $X_{t+1} = y$ given that $X_t = x$) and therefore Since $p \coloneqq p_{t,t+1}$ is independant of t it follows (quite generally) that $\forall t, n \in \mathbb{N}$

$$\begin{aligned} p_{t,t+n}(x,y) &= \mathbb{P}\{X_{t+n} = y | X_t = x\} = \sum_{z \in E} \mathbb{P}\{X_{t+n} = y, X_{t+1} = z | X_t = x\} \\ &= \sum_{z \in E} \mathbb{P}\{X_{t+n} = y | X_{t+1} = z\} \mathbb{P}\{X_{t+1} = z | X_t = x\} \\ &= \sum_{z \in E} \mathbb{P}\{X_{t+n} = y | X_{t+1} = z\} p(x,z) = \cdots = p^n(x,y), \end{aligned}$$

where $p^n(x,y) = \sum_{z\in E} p^{n-1}(z,y)p(x,z)$ and $p^0 = 1$. This follows from the Markov property and repeated conditioning. Therefore $(X_t)_{t\in\mathbb{N}}$ is a time homogenous Markov chain with kernel p.
It is left to check that (X_t) is ergodic and that it has π as a stationary distribution. To show that π is a stationary distribution of the Markov chain it is enough to check the following:

$$\pi(x)p(x,y) = \pi(y)p(y,x) \quad \forall x,y \in E \tag{9}$$

since then $\forall y \in E$

$$(\pi p)(y) = \sum_{x\in E} p(x,y)\pi(x) = \pi(y)\sum_{x\in E} p(y,x) = \pi(y),$$

because p is a Markov kernel. Now to check Equation 9 for (X_t): Let $x,y \in E$, if $x \neq y$ and $q(x,y) \neq 0$ then

$$\pi(x)p(x,y) = \pi(x)q(x,y)\min\left\{\frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}, 1\right\} = \min\left\{\pi(y)q(y,x), \pi(x)q(x,y)\right\}$$

and if also $q(y,x) \neq 0$ then

$$\pi(y)p(y,x) = \pi(y)q(y,x)\min\left\{\frac{\pi(x)q(x,y)}{\pi(y)q(y,x)}, 1\right\} = \min\left\{\pi(y)q(y,x), \pi(x)q(x,y)\right\},$$

if instead $q(y,x) = 0$ then both expressions are 0. If $x \neq y$ and $q(x,y) = 0$ then

$$\pi(x)p(x,y) = 0 = \pi(y)p(y,x).$$

In the case that $x = y$ the equation is trivially satisfied. Therefore (X_t) has π as a stationary distribution. It is left to show that (X_t) is ergodic. Let the digraph associated to q be strongly connected and aperiodic. Furthermore let q satisfy $q(x,y) \neq 0 \Leftrightarrow q(y,x) \neq 0$ for all $x,y \in E$. Then the Digraph associated to p is also strongly connected and aperiodic since then $p(x,y) \neq 0 \Leftrightarrow q(x,y) \neq 0$ for all $x,y \in E$. So (X_t) is ergodic and therefore its distribution will converge to π independant of the starting distribution of X_0.