PHYS 142/242

Lecture 15: Ising Model and MCMC

Javier Duarte — February 12, 2024

2D Ising model

The Hamiltonian is

$$E(\sigma) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j$$

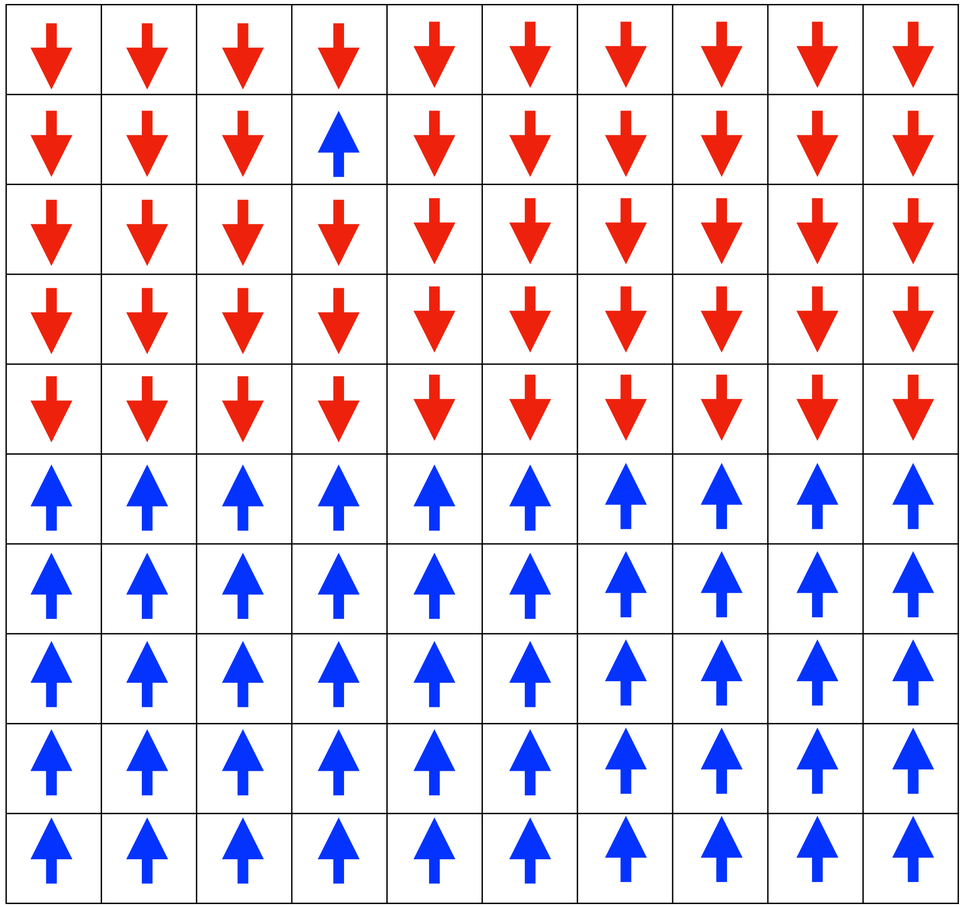
where $\sigma_k \in \{-1, +1\}$ represents the spin of each site. The probability of any given state is

$$p(\sigma) = \frac{1}{Z} \exp\left(-\frac{E(\sigma)}{k_{\mathrm{B}} T}\right)$$

where

$$Z = \sum_{\sigma} \exp\left(-\frac{E(\sigma)}{k_{\mathrm{B}} T}\right).$$

With 10×10 sites $2^{100} \approx 10^{30}$ possible states

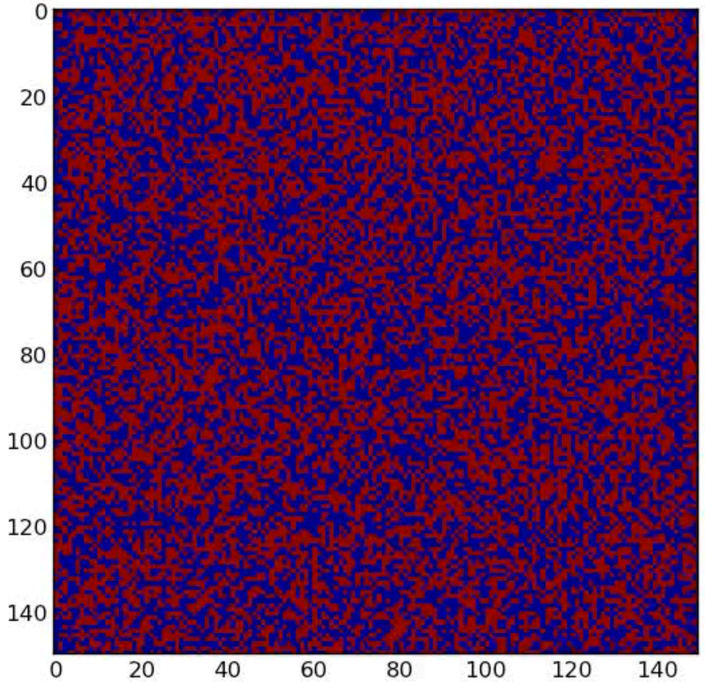


Ising model phase transition

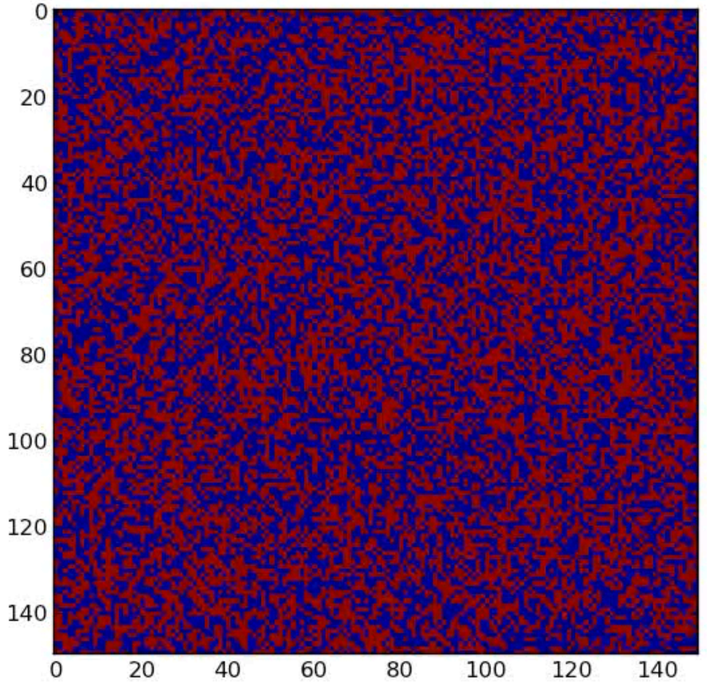
With 150×150 lattice, $2^{22500} \approx 10^{6773}$ possible states!

Phase transition occurs at a critical temperature: ordered phase at low T, disordered phase at high T. How can we compute this?

High T



Low T



Ising model phase transition

Often interested in quantities like the average magnetization, i.e. first we average over the lattice

$$\sigma_{\text{ave}} = \frac{1}{|\Lambda|} \sum_{k \in \Lambda} \sigma_k$$

then we take a thermodynamic average

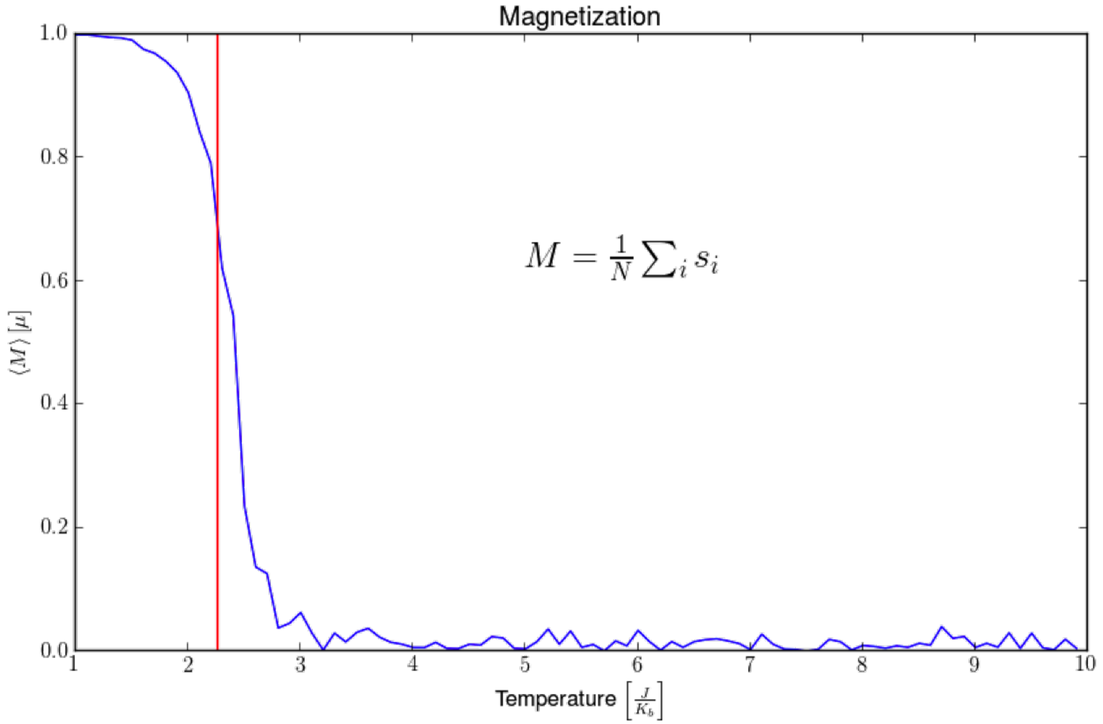
$$\langle \sigma_{\text{ave}} \rangle = \sum_{\sigma} \sigma_{\text{ave}} p(\sigma)$$

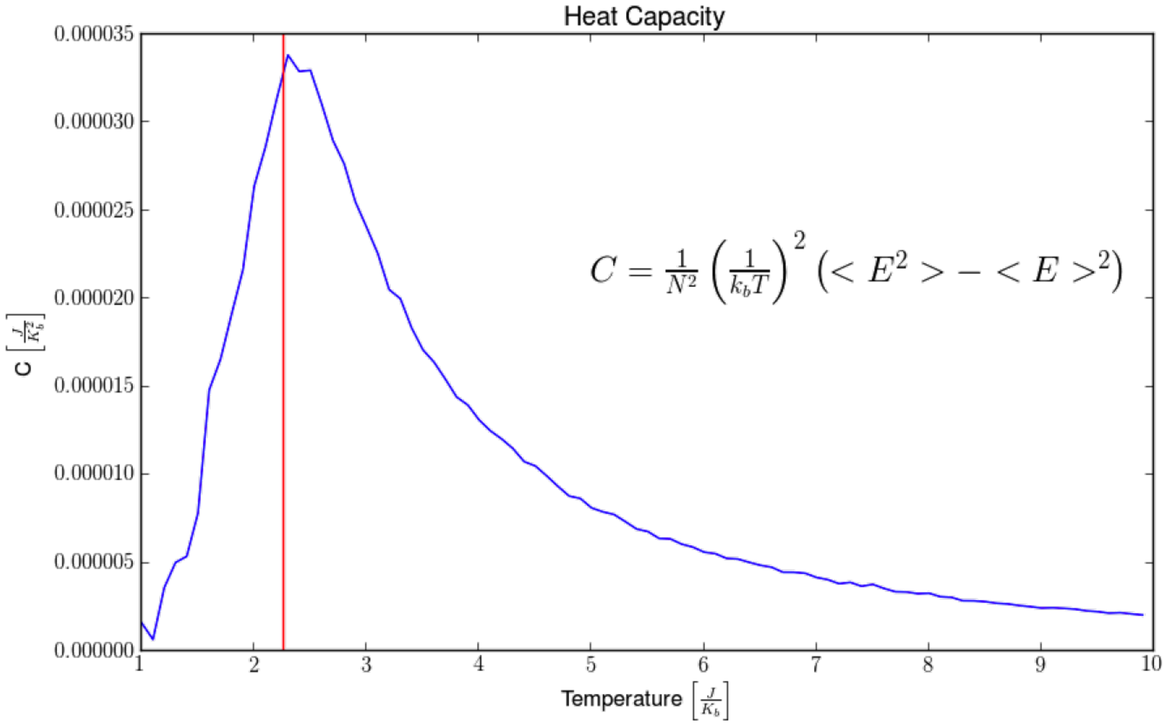
If we plot this versus temperature, we can see there is some critical temperature $T_{\text{c}}$where the system goes from being a ferromagnetic (spins are aligned) to a paramagnetic (spins are not aligned)

Ising model phase transition

In 2D, we can actually analytically calculate the critical temperature as solved in L. Onsager, Phys. Rev. 65, 117 (1944)

$$\frac{k_{\mathrm{B}}T_{\mathrm{c}}}{J} = \frac{2}{\ln(1+\sqrt{2})} \approx 2.269$$





Applying MCMC to the 2D Ising model

To apply the MCMC method to the Ising model, we design a Markov process using the Metropolis algorithm as follows

1. On step k, randomly choose one of the spins i and consider flipping it $\sigma_i \rightarrow -\sigma_i$
2. Calculate the change in energy that would result from flipping spin i, i.e. the quantity:

$$\Delta E = -\left[J \sum_{\langle ij \rangle} \sigma_j \right] \Delta\sigma_i,$$

where $\Delta\sigma_i = -2\sigma_i$ is the change in σ_i due to the spin flip

If $\Delta E \leq 0$, accept the spin flip

If $\Delta E > 0$, accept the spin flip with probability $\exp(-\Delta E/k_{\mathrm{B}}T)$. Otherwise, reject the flip.

3. Update the moving average of $\langle \sigma_{\mathrm{ave}} \rangle$ (or whatever quantity we are interested in).

4. Repeat.

Practical considerations for MCMC

- Several practical considerations
  - Acceptance rate: the rate at which we accept proposals is important. We want this to be relatively high (to reduce total running time), but not always 1 (which may mean we're not sufficiently sampling the full space). Typical values are 15% to 50%
  - Length of chain: Need long enough chain to "converge" to stationary distribution
  - Number of chains: Running multiple chains can help estimate uncertainty and whether each chain has converge
  - Burn in: Samples to remove from the beginning of each chain
  - Thinning: Removing every nth sample
  - Error and (integrated) autocorrelation time

Burn-in

"Burn-in": common to throw out the first few states of a Markov chain, maybe the first 100 or the first 1000

The idea is to get rid of "transient behavior" connected to an improbable intial configuration

Equivalent: pick an initial state near equilibrium (i.e. not unlikely)

Autoregressive series

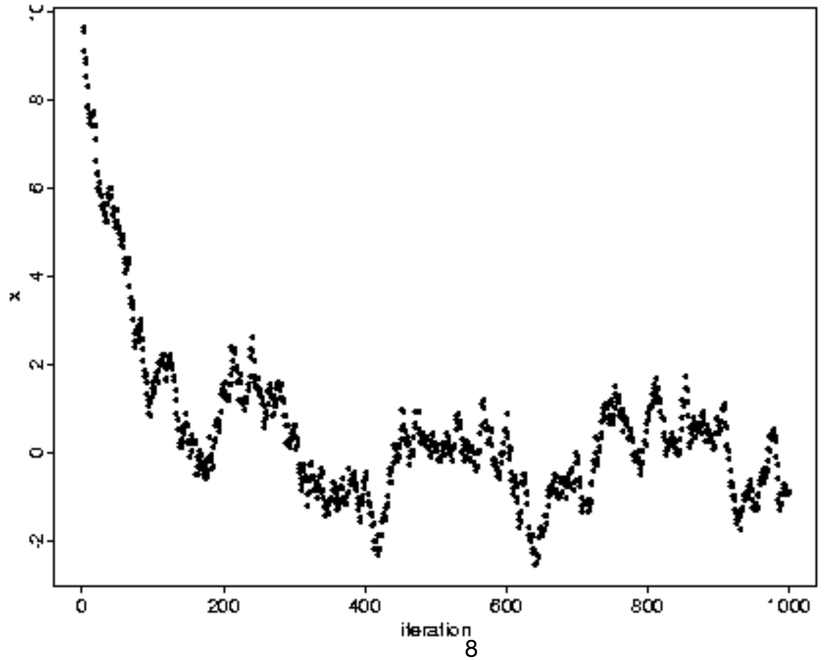
$$x_{n+1} = rx_n + e_n$$

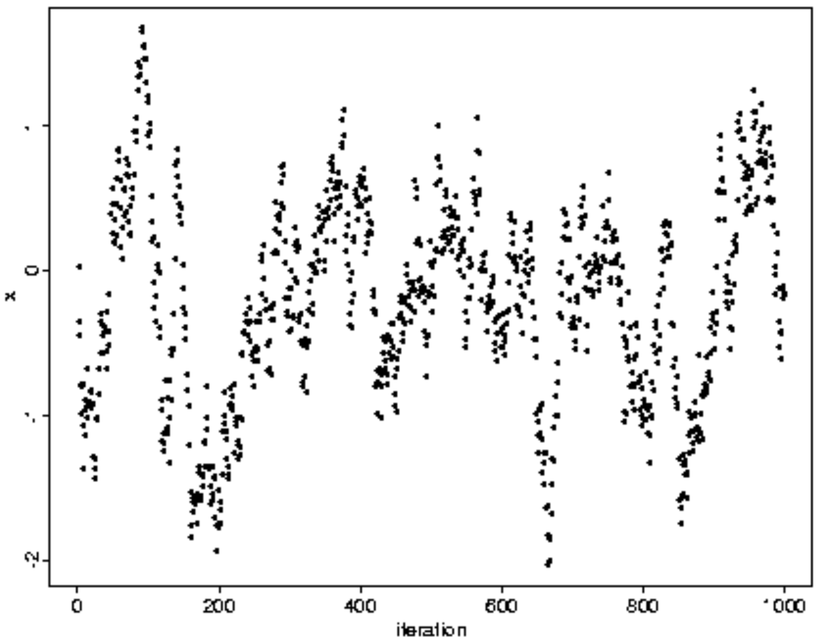
with $r = 0.98$ and

$$e_n \sim \mathcal{N}(0,1)$$

starting from $x_0 = 10$ (left)

and $x_0 = 0$ (right)





MC error

Our goal is to estimate the expectation value of some physical observable f.

Using traditional MC sampling, the estimate is given by the average over N samples

$$\bar{f} = \frac{1}{N}\sum_{i=1}^{N} f_i$$

If these samples are independent, then the variance on this estimator is

$$\sigma^2_{\bar{f}} = \frac{1}{N}\mathrm{Var}[f_i] = \frac{1}{N}\left(\langle f^2\rangle - \bar{f}^2\right)$$

and the error decreases as $1/\sqrt{N}$ as we generate more samples.

MCMC error

However, in MCMC, the samples are not independent! What is the variance?

$$\sigma_{\bar{f}}^2 = \frac{\tau_f}{N} \mathrm{Var}[f_i]$$

where τ_f is the integrated autocorrelation time for the chain.

N/τ_f is the effective number of samples and τ_f is the number of steps that are needed before the chain “forgets” where it started.

Autocorrelation function

Can estimate the autocorrelation function of the observable f as

$$c_f(t) = \langle (f_i - \bar{f})\,(f_{i+t} - \bar{f}) \rangle = \frac{1}{N-t} \sum_{i=1}^{N-t} (f_i - \bar{f})\,(f_{i+t} - \bar{f})$$

Note the variance is a special case: $c_f(0) = \mathrm{Var}[f_i]$

We can also define the normalized autocorrelation (like a Pearson correlation coefficient)

$$\rho_f(t) = c_f(t)/c_f(0)$$

It's probably more appropriate to define $c_f(t)$ as "autocovariance," but in practice you have to figure out what the literature means from context

Integrated autocorrelation time

The integrated autocorrelation time is then given by

$$\tau_f = \sum_{t=-\infty}^{\infty} \rho_f(t) = 1 + 2 \sum_{t=1}^{N} \rho_f(t)$$

Trying to approximate this directly is difficult. At longer lags, $\rho_f(t)$ starts to contain more noise than signal and summing all the way out to N will result in a very noisy estimate of τ_f. Instead we want to cut off the some at some $M \ll N$

$$\tau_f(M) = 1 + 2 \sum_{t=1}^{M} \rho_f(t)$$

One suggestion: using smallest M where $M \geq 5\tau_f(M)$

Estimating

Estimating integrated autocorrelation time: https://emcee.readthedocs.io/en/stable/tutorials/autocorr/