HW4 - Jacob Shkrob

Jacob Shkrob

November 29, 2022

1 Exercise 39

Our goals in the exercise are to use the Gibbs sampler to generate samples from the Ising model, show the distribution of the (mean) magnetization

$$f(\sigma) = \sum_{i \in \mathbb{Z}_L^2} \sigma_i,$$

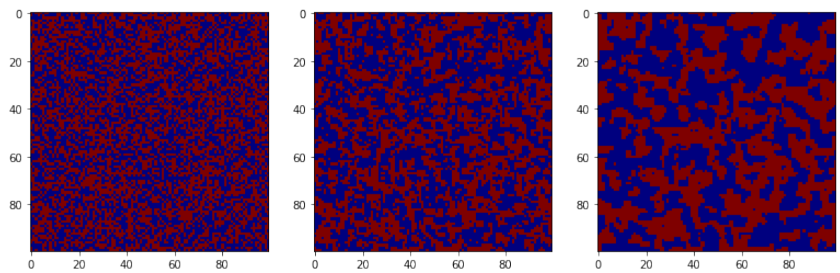
compute the integrated autocorrelation times (IAT) for $f(\sigma)$ for both the deterministic (sweeping) Gibbs sampler and randomized Gibbs sampler, and compare results between different regimes (temperature, lattice size L, etc). First, a brief review of the Ising model in 2-d. Consider a measure $\pi(\sigma)$ on the space $\{-1, 1\}^{\mathbb{Z}_L^2}$:

$$\pi(\sigma) = \frac{e^{\beta \sum_{i \leftrightarrow j} \sigma_i \sigma_j}}{\mathcal{Z}}.$$

In this model, β represents the inverse temperature $k_B T^{-1}$. In order to use Gibbs sampling, we need to calculate the conditional distributions in each coordinate. Thankfully, this process is simple and the resulting distribution is given by the simple formula

$$\pi(\sigma_i | \sigma_{[i]}) = \left(\frac{e^{\beta} \sum_{j \leftrightarrow i} \sigma_j}{e^{\beta \sum_{j \leftrightarrow i} \sigma_j} + e^{-\beta \sum_{j \leftrightarrow i} \sigma_j}} \right) \delta_1(\sigma_i) + \left(\frac{e^{\beta \sum_{j \leftrightarrow i} \sigma_j}}{e^{-\beta \sum_{j \leftrightarrow i} \sigma_j} + e^{-\beta \sum_{j \leftrightarrow i} \sigma_j}} \right) \delta_{-1}(\sigma_i).$$

Here, the notation $[i]$ means all the lattice points in $\mathbb{Z}_L^2 - \{i\}$. Hence, our Gibbs sampling scheme will be based on two different regimes regarding the selection of indices i to condition on. We propose two different regimes: **deterministic** and **random**. In the deterministic regime, we select our index i row-by-row, until we hit the number of interations. In the random regime, we select an index i at random, until we hit the number of interactions. Each method has its benefits and pitfalls. Below in Figure 1 are example outputs of the Ising model at temeperature parameter $\beta = 1$ after 1000 iterations of the Gibbs process.



(a) $\beta = 1$ initial configuration (b) $\beta = 1$ after 10000 steps (c) $\beta = 1$ after 20000 steps

Figure 1: Example output of the 2d Ising model on lattice size $L = 100$

Before continuing, we also include a figure (Figure 2) which shows an emperical histogram of the magnetizations $f(\sigma)$ using the endpoints of $n = 1000$ Markov chains that were based on Gibbs sampling. We fixed the temperature parameter $\beta = 1.0$ and lattice size $L = 50$, halting each chain after 3000 steps.

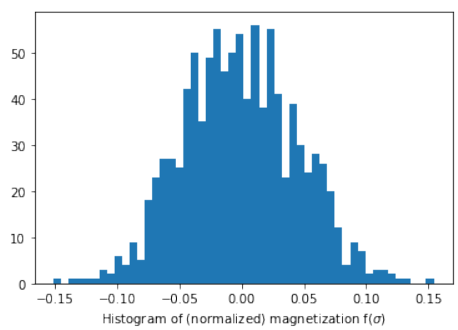


Figure 2: Histogram of mean magnitization of 2d-Ising model using Gibbs sampling (n=1000, $\beta = 1$, $L = 50$, iterations = 3000)

In our analysis of Gibbs sampling, and future algorithms like Metroplolis and Jarzynski, we will be computing the integrated autocorrelation time (IAT) for the chain $(X_n)_{n\geq 1}$ for the function f is defined to be

$$\tau_f = 1 + 2\sum_{j=1}^{\infty} \mathbf{Cor}\left[f(X_j), f(X_0)\right].$$

Note that this τ_f appears in the Central Limit Theorem result for the Markov chain $\{f(X_k)\}_{k\geq 1}$:

$$\lim_{N\to\infty} \sqrt{N}(\bar{f}_N - \pi[f]) = \mathcal{N}(0, \tau_f \sigma_f^2),$$

where $\sigma_f^2 = \mathbf{Var}\left(X_1\right), X_1 \sim \pi$. Therefore, the lower τ_f is, the sharper our estimator of $\pi[f]$ will be. It is therefore desirable to estimate τ_f for various different MCMC schemes. To actually estimate τ_f, say $\hat{\tau}_f$, we write,

$$\hat{\tau}_f = 1 + 2\sum_{j=1}^{N} \hat{\rho}_f(j),$$

where $\hat{\rho}_f(j)$ is the normalized autoccorrelation function of $\{f(X_n)\}_{n\geq 1}$, which can be written as $\hat{\rho}_f(j) = \frac{\hat{c}_f(j)}{\hat{c}_f(0)}$ where $\hat{c}_f(j)$ are sample autocovariances of the finite chain $\{f(X_j)\}_{j=1}^{N}$. As stated in (Goodman and Weare, 2010), the fast Fourier transform computationally efficiently copute these coefficients, and we use this method throughout the rest of the writeup.

Figure 3 below illustrates the difference between estimates of τ_f for each of the two methods of selecting i. We note that random selection of i appears to triumph, so we will continue using random selection from now on. Additionally, we are interested in the effect of (1) lattice number L on the IAT estimates and (2) temperature parameter β on the IAT estimates. Below, Figure ?? shows how these two quantities change.

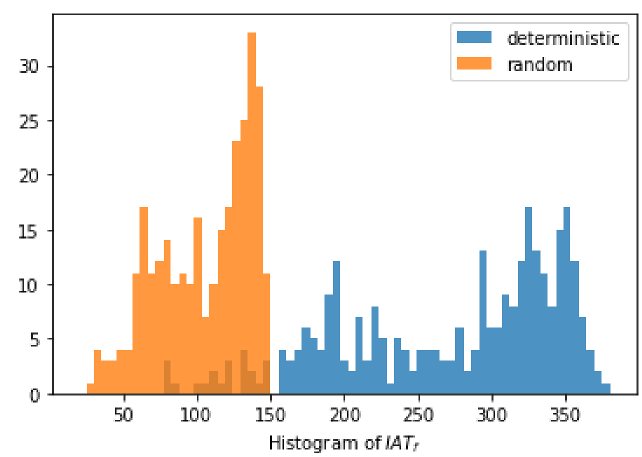
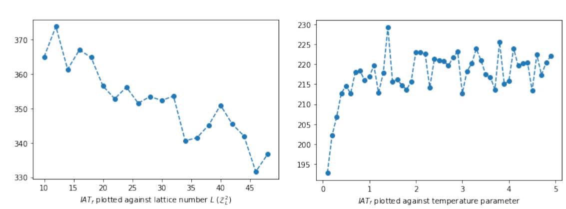


Figure 3: IAT estimates for deterministic and random Gibbs sampling



(a) IAT_f as lattice number L changes

(b) IAT_f as temperature β changes

2 Exercise 42

In this exercise, we use the Metropolis-Hastings algorithm to sample from the Ising model instead. The metropolis algorithm for the ising model is easily constructed: at each iteration of the chain k, select a lattice entry i at random and flip the sign of σ_i. Since the probability of of the newly flipped configuration Y_{k+1} given X_k, the current configuration, is reflexive, we only need to care about the ratio $\frac{\pi(\tilde{\sigma}_i)}{\tilde{\sigma}_i}$. Thankfully, this calculation can be easily found to be

$$\frac{\pi(Y_{k+1})}{\pi(X_k)} = e^{-4\beta X_{i,k} \sum_{j \leftrightarrow i} X_{j,k}}.$$

Therefore, we can easily formulat the acceptance probability that we need to use for the Metropolis-based scheme. Below, Figure 5 shows a histogram comparing the integrated autocorrelation time for the scheme and the integrated autocorrelation time for the randomized-version of Gibbs sampling.

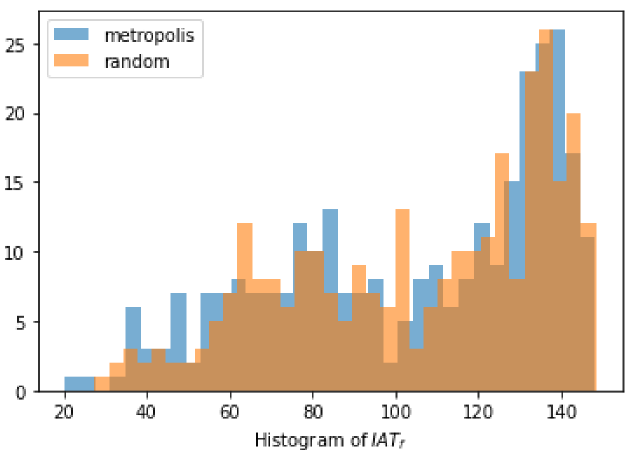


Figure 5: IAT estimates of magnetization f for Metropolis-Hasting sampling and random Gibbs sampling, $L = 50$, $\beta = 1$, and we used $n = 300$ estimates

As we can see, the MH scheme slightly outperforms the random Gibbs scheme (on average, the time estimated is lower), where $\hat{\tau}_f(random) \approx 105$ and $\hat{\tau}_f(MH) \approx 104$. Therefore, Metropolis would be a more advised selection when it comes to performance.

3 Exercise 43

The goal of the exercise is to use Jarzynski's method (without resampling) to generate sampled from the 2d-Ising model. The first necessary part of Jarzynski's method is to choose an appropriate sequence $\{\pi_k\}$ of measures for which $\pi_k = \pi^{\frac{k}{n}}$

and choose transition operators $\mathcal{T}_k$ for which $\pi_{k-1}\mathcal{T}_k = \pi_{k-1}$. Based on the previous exercises, for the Metropolis algorithm, this amounts to replacing π with $\pi^{\frac{k}{N}}$, which leads to scaling the factor in the exponential by $(\frac{k}{N})$. In the Gibbs algorithm, we just change the probabilities from $e^{\beta\sum_{j\leftrightarrow i}\sigma_j}$ to $e^{\beta\frac{k}{N}\sum_{j\leftrightarrow i}\sigma_j}$ at each step when we choose i at random. We note that the choise of n when creating the sequence $\{\pi^{\frac{k}{n}}\}$ is important and will be carfuly selected to minimize the variance of the magnetization estimates. Using normalized importance weights, we have the following important recursion: let the weight at time step k be denoted as W_k, so that

$$W_0 = 1,$$
$$W_k = W_{k-1}\frac{\pi_k(X_k)}{\pi_{k-1}(X_k)} = W_{k-1}\pi(X_k)^{\frac{1}{n}}.$$

Since we do not want to compute the normalizing constant for the Ising model, we run multiple different chains and compute the weight for each chain $\{W_m\}_{m=1}^{M}$ at time step k via the formula

$$W_{m,k} = \frac{W_{m,k-1}w_{k-1}(X_{m,k-1})}{\sum_{j=1}^{M}W_{j,k-1}w_{k-1}(X_{j,k-1})}.$$