What is the optimal MCMC jumping step?

Jian Wang, UCLA Physics and Astronomy, January 2019

The question is as follows. In a Monte Carlo simulation, each updating cost average time t_u, each calculation cost average time t_c. We also know that the Monte Carlo has a correlated step τ. Then what is the the optimal jumping steps J between two calculations? The goal is to make the calculation a smaller error bar, within fixed total time T.

Listing 1: a typical MCMC code

```
while(true){
    for(int i=0;i<J;i++){
        updating();
    }
    calculating();
}
```

The two extreme limits are bad: (1) $J \gg \tau$ is wasting a lot of time to update. (2) $J = 1 \ll \tau$ is wasting time to calculate the "same" sample .

1

1.1 Theoretical solution

Effective sample size is introduced as:

$$\text{Effective Sample Size} = \frac{N}{1 + 2\tau_{\text{effective}}} \tag{1}$$

We should have $N = \frac{T}{t_u J + t_c}$ and $\tau_{\text{effective}} = \tau/J$, then

$$\text{ESS} = \frac{T}{t_u J + t_c} \frac{1}{1 + 2\frac{\tau}{J}} \tag{2}$$

In Eq. (2), the dimension of t_u and t_c are time, the dimension of τ and J are unit-less. To maximum ESS is to minimize its denominator:

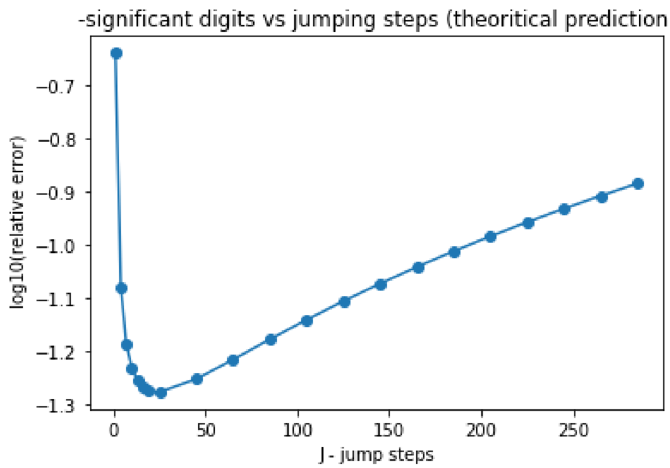
$$t_u J + t_c + 2t_u\tau + 2t_c\tau\frac{1}{J} \geq t_c + 2t_u\tau + 2\sqrt{t_u 2t_c\tau} \tag{3}$$

As a result of $a + b \geq 2\sqrt{ab}$, the optimal J takes:

$$J_{\text{optimal}} = \sqrt{\frac{2t_c\tau}{t_u}} \tag{4}$$

The log 10 of relative error is equivalent to negative significant digits.

$$-\text{Sig.Digits} = \log_{10}\left(\frac{\text{error}}{\text{mean}}\right) \propto \log_{10}\sqrt{\frac{1}{\text{ESS}}} \tag{5}$$



Compare $J_{\text{bad}} = 1$ and $J_{\text{optimal}} = 25$, the significant digit is increased from 0.7 to 1.3. This means the J_{bad} will spend $(10^{1.3-0.7})^2 \approx 15$ times extra CPU hours than the J_{optimal} in order to acheive the same accuracy.

1.2 test

Instead of looking for t_c, t_u, τ and admitting the assumption $\frac{N}{1+2\tau_{\text{effective}}}$. (why not $\frac{N}{1+3\tau_{\text{effective}}}$, $\frac{N}{1+4\tau_{\text{effective}}}$, ..) In practice, we generate a list of J values, test them and find the optimal J from the graph.

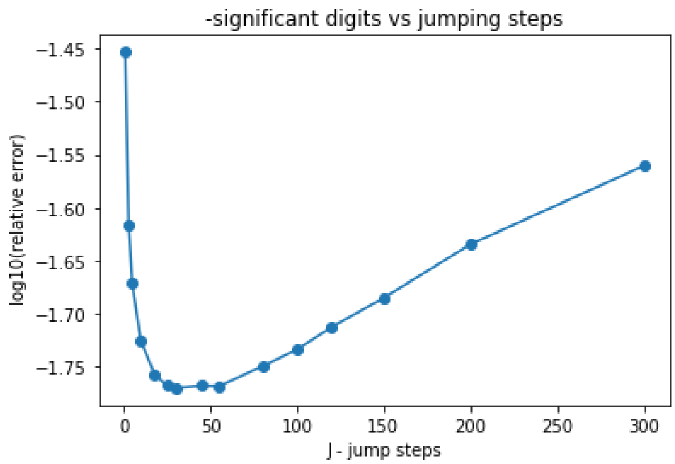


Figure 1: A real test example for 2D Ising model, we are using Wolff updating, and calculating the spin-spin correlation function. The optimal $J = 26$. The shape of the curve agrees with theory prediction.

1.3 Significant digits - CPU Time scaling relation

Running longer time MCMC will shift the curve down. The relation comes from central limit theory $\sigma_{\text{data mean}} = \frac{\sigma_{\text{data}}}{\sqrt{N}}$. The

$$\text{Sig.Digits}_T = \text{Sig.Digits}_{1\ \text{second}} + \log_{10}\sqrt{\frac{T}{1\ \text{second}}} \tag{6}$$