Mixing time and simulated annealing for the stochastic cellular automata, and beyond

Akira Sakai[1]

There are many real-life situations where we must choose one option among extremely many, as quickly and optimally as possible. Such combinatorial optimization problems are ubiquitous and possibly quite hard to solve them fast. One approach to find an optimal solution to a given problem is to translate it into an Ising Hamiltonian on a finite graph $G=(V,E)$ with no multi- or self-edges and find a ground state that corresponds to an optimal solution.

Given a system of spin-spin couplings $\boldsymbol{J}=\{J_{x,y}\}_{\{x,y\}\in E}$ (with $J_{x,y}=0$ if $\{x,y\}\notin E$) and external magnetic fields $\boldsymbol{h}=\{h_x\}_{x\in V}$, the Ising Hamiltonian of a spin configuration $\boldsymbol{\sigma}=\{\sigma_x\}_{x\in V}\in\Omega\equiv\{\pm1\}^V$ is defined as

$$H(\boldsymbol{\sigma})=-\frac{1}{2}\sum_{x,y\in V}J_{x,y}\sigma_x\sigma_y-\sum_{x\in V}h_x\sigma_x. \tag{1}$$

Let $\mathrm{GS}=\arg\min_{\boldsymbol{\sigma}\in\Omega}H(\boldsymbol{\sigma})$, which is the set of spin configurations where the Hamiltonian attains its minimum value. A standard method to find an element from GS is to use a Markov chain Monte Carlo (MCMC) to sample the Gibbs distribution $\pi^{\mathrm{G}}_\beta(\boldsymbol{\sigma})\propto e^{-\beta H(\boldsymbol{\sigma})}$ at the inverse temperature $\beta\ge0$, which attains its highest peaks on GS. There are several MCMCs that can generate the Gibbs distribution as the equilibrium distribution. One of them is the Glauber dynamics, which is defined by the transition probability

$$P^{\mathrm{G}}_\beta(\boldsymbol{\sigma},\boldsymbol{\sigma}^x)=\frac{1}{|V|}\frac{w^{\mathrm{G}}_\beta(\boldsymbol{\sigma}^x)}{w^{\mathrm{G}}_\beta(\boldsymbol{\sigma})+w^{\mathrm{G}}_\beta(\boldsymbol{\sigma}^x)}\overset{(1)}{=}\frac{1}{|V|}\frac{e^{-\beta\tilde{h}_x(\boldsymbol{\sigma})\sigma_x}}{2\cosh(\beta\tilde{h}_x(\boldsymbol{\sigma}))}, \tag{2}$$

where $(\boldsymbol{\sigma}^x)_y=(-1)^{\delta_{x,y}}\sigma_y$ and $\tilde{h}_x(\boldsymbol{\sigma})=\sum_{y\in V}J_{x,y}\sigma_y+h_x$; if the Hamming distance between $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$ is bigger than 1, $P^{\mathrm{G}}_\beta(\boldsymbol{\sigma},\boldsymbol{\tau})$ is defined to be zero. Since P^{G}_β is aperiodic, irreducible and reversible with respect to π^{G}_β, the Glauber dynamics has a unique equilibrium distribution π^{G}_β. However, since the number of spin-flips per update is at most one, it is potentially slow and may not be so useful in practice. It has been longed for a way to update many spins at once, independently of each other.

One such MCMC is in a particular class of probabilistic cellular automata [1, 3]. Since the abbreviation PCA has long been used for principal component analysis in statistics, we would rather call it the stochastic cellular automata (SCA). It is defined by the doubled Hamiltonian with the pinning parameters $\boldsymbol{q}=\{q_x\}_{x\in V}$ as

$$\begin{aligned}\tilde{H}(\boldsymbol{\sigma},\boldsymbol{\tau})&=-\frac{1}{2}\sum_{x,y\in V}J_{x,y}\sigma_x\tau_y-\frac{1}{2}\sum_{x\in V}h_x(\sigma_x+\tau_x)-\frac{1}{2}\sum_{x\in V}q_x\sigma_x\tau_x\\&=-\frac{1}{2}\sum_{x\in V}\big(\tilde{h}_x(\boldsymbol{\sigma})+q_x\sigma_x\big)\tau_x-\frac{1}{2}\sum_{x\in V}h_x\sigma_x.\end{aligned} \tag{3}$$

[1]Faculty of Science, Hokkaido University. https://orcid.org/0000-0003-0943-7842

Let

$$w^{\mathrm{SCA}}_{\beta,\boldsymbol{q}}(\boldsymbol{\sigma}) = \sum_{\boldsymbol{\tau}} e^{-\beta\tilde{H}(\boldsymbol{\sigma},\boldsymbol{\tau})} \overset{(3)}{=} \prod_{x\in V} 2e^{\frac{\beta}{2}h_x\sigma_x}\cosh\big(\tfrac{\beta}{2}(\tilde{h}_x(\boldsymbol{\sigma})+q_x\sigma_x)\big), \tag{4}$$

and define the SCA transition probability as

$$P^{\mathrm{SCA}}_{\beta,\boldsymbol{q}}(\boldsymbol{\sigma},\boldsymbol{\tau}) = \frac{e^{-\beta\tilde{H}(\boldsymbol{\sigma},\boldsymbol{\tau})}}{w^{\mathrm{SCA}}_{\beta,\boldsymbol{q}}(\boldsymbol{\sigma})} \overset{(4)}{=} \prod_{x\in V}\frac{e^{\frac{\beta}{2}(\tilde{h}_x(\boldsymbol{\sigma})+q_x\sigma_x)\tau_x}}{2\cosh(\frac{\beta}{2}(\tilde{h}_x(\boldsymbol{\sigma})+q_x\sigma_x))}. \tag{5}$$

Because of this product form, all spins are updated at once, independently of each other. Since $\tilde{H}$ is symmetric, $P^{\mathrm{SCA}}_{\beta,\boldsymbol{q}}$ is reversible with respect to $\pi^{\mathrm{SCA}}_{\beta,\boldsymbol{q}}(\boldsymbol{\sigma}) \propto w^{\mathrm{SCA}}_{\beta,\boldsymbol{q}}(\boldsymbol{\sigma})$, which is different from the Gibbs distribution, so we cannot naively use it to search for an element in GS. However, since $\tilde{H}(\boldsymbol{\sigma},\boldsymbol{\sigma}) = H(\boldsymbol{\sigma}) - \frac{1}{2}\sum_{x\in V} q_x$, the total-variation distance between $\pi^{\mathrm{SCA}}_{\beta,\boldsymbol{q}}$ and π^{G}_{β} tends to zero as $\min_{x\in V} q_x \uparrow \infty$. We take $\min_{x\in V} q_x$ large enough (e.g., $= \lambda/2$, where λ is a half of the largest eigenvalue of the matrix $[-J_{x,y}]_{V\times V}$) to ensure that $\tilde{H}$ attains its minimum values on the diagonal entries:

$$\min_{\boldsymbol{\sigma},\boldsymbol{\tau}\in\Omega} \tilde{H}(\boldsymbol{\sigma},\boldsymbol{\tau}) = \min_{\boldsymbol{\sigma}\in\Omega}\tilde{H}(\boldsymbol{\sigma},\boldsymbol{\sigma}), \qquad \arg\min_{\boldsymbol{\sigma}\in\Omega}\tilde{H}(\boldsymbol{\sigma},\boldsymbol{\sigma}) = \mathrm{GS}. \tag{6}$$

At the symposium, I will explain the following results from the joint paper [2]:

(i) If the temperature is sufficiently high (depending only on $\boldsymbol{J}$ and $\boldsymbol{q}$), the mixing time for the time-homogeneous SCA is at most of order $\log|V|$, which is much smaller than that for the Glauber dynamics under zero magnetic field.

(ii) If the temperature drops in time as $\beta_n \propto \log n$, then the time-inhomogeneous SCA weakly converges to the uniform distribution $\pi^{\mathrm{G}}_{\infty}$ on GS. The sequence $\{\beta_n\}_{n=0}^{\infty}$ is a standard cooling schedule in simulated annealing applied to single-spin dynamics.

I will also show some numerical results on the so-called ε-SCA, which is introduce to try to improve performance of the SCA, and explain the current status of the ongoing joint work with Fukushima-Kimura, Kamijima, Kawamoto, Kawamura and Noda.

References

[1] P. Dai Pra, B. Scoppola and E. Scoppola. Sampling from a Gibbs measure with pair interaction by means of PCA. *J. Stat. Phys.*, **149** (2012): 722–737.

[2] B.H. Fukushima-Kimura, S. Handa, K. Kamakura, Y. Kamijima and A. Sakai. Mixing time and simulated annealing for the stochastic cellular automata. Submitted. `arXiv:2007.11287`

[3] B. Scoppola and A. Troiani. Gaussian mean field lattice gas. *J. Stat. Phys.*, **170** (2018): 1161–1176.