Markov-chain Monte Carlo: A modern primer

Lecture 3: Advanced subjects
Part 3.1: Meta algorithms

Werner Krauth
Département de physique, Ecole normale supérieure
Paris, France

CANES Masterclass
Faculty of Natural & Mathematical Sciences, King's College London
London, Great Britain

3 March 2022

References

- D. A. Levin, Y. Peres, E. L. Wilmer, **Markov Chains and Mixing Times**, (American Mathematical Society, 2008),
- D. Frenkel, B Smit, **Understanding Molecular Simulation: From Algorithms to Applications**, (Elsevier, 2001)
- B. Hajek, **Cooling Schedule for optimal annealing**, Mathematics of Operations Research (1988)
- W. Krauth **Statistical Mechanics: Algorithms and Computations**, (Oxford University Press, 2006)

Characteristic times in MCMC 1/3

- Correlation time.
- Mixing time.
- Cover time.

Characteristic times in MCMC 2/3

- Correlation time: Time to move from one i ($\sim \pi_i$) to an independent j ($\sim \pi_j$).
- Mixing time: Time to reach a $j \sim \pi_j$ starting from $i \sim \pi^{\{0\}}$ with worst $\pi^{\{0\}}$.
- Cover time: Time to have seen all samples, starting from the worst initial sample x: $t_{\text{cov}} = \max_{x \in \Omega} \mathbb{E}\left[\tau_{\text{cov}}(x)\right]$ (with $\tau_{\text{cov}}(x)$ the time to have seen all $i \in \Omega$).

Characteristic times in MCMC 3/3

Example (SSEP of N hard spheres on path graph P_{2N}):

- Correlation time: $\propto N^3$.
- Mixing time: $\propto N^3 \log N$.
- Cover time: N^N.

Consequences:

1. Difficult to know normalization of π:
   - ... What is $Z = \sum_x \pi_x$? (Thermodynamic integration)
2. Difficult to know Ω:
   - ... What is $\min_x \pi_x$?
   - ... What is $\max_x \pi_x$? (Simulated annealing)
   - ... What is conductance?
3. Difficult to explore Ω:
   - Is $\Omega = \emptyset$?
   - Have we seen all of Ω? (Multicanonical MC)

Missing element in P_N: Combinatorial explosion of sample space.

Thermodynamic integration, simulated annealing, etc.

- Normal MCMC (algorithm development): keep π, change P, keep Ω.
- Lifted MCMC: keep π, keep P, change Ω.
- Thermodynamic integration, sim annealing, multicanonical MC: change π, keep P, keep Ω.

Debate:

- Sampling algorithms development *vs.* Metaheuristics.

Thermodynamic integration

- All of MCMC: concerned with π_i/π_j, norm of π_i (usually) irrelevant.
  - Metropolis filter: $\mathcal{P}(i \to j) = \min(1, \pi_j/\pi_i)$.
  - NB: Flow: $\mathcal{F}_{ij} = \pi_i \mathcal{P}_{ij}$ (usually) unknown.
- All of physics: concerned with $Z = \sum_{i \in \Omega} \pi_i$
  ($\pi_i = \exp(-E_i/kT)$)
- All of physics: Partitition function known analytically in some limits:
  - High-temperature limit: $T \to \infty \Leftrightarrow \beta \to 0$
  - Ideal-gas limit: density $\rho \to 0$, interactions $\to 0$.
  - Ideal-solid limit: density $\rho \to \rho_{\max}$, interactions $\to$ harmonic.
  - Keep Ω (usually), change π.
- Creating a path from a known limit to the relevant $\{\Omega, \pi\}$ is called “Thermodynamic integration”.
- Path must (normally) be smooth (avoid phase transitions).

Thermodynamic integration (example)

- V-shaped: $\pi_i = \text{const} |\frac{n+1}{2} - i| \; \forall i \in \Omega$.
- const $= \frac{4}{n^2}$ is unknown.
- $\tilde{V}$ -shaped: $\tilde{\pi}_i = |\frac{n+1}{2} - i| \; \forall i \in \Omega$.
- Partition function $Z = \sum_{i \in \Omega} \tilde{\pi}_i$.
- Consider $\tilde{\pi}^\alpha$ (“pi to the power alpha”).
- $\alpha \to 0$: high-temperature limit $\pi_i^0 = 1 \; \forall i$.
- “Riemann integration” path $\alpha(t)$ with $\alpha(0) = 0$ and $\alpha(1) = 1$.

Thermodynamic integration

- Partition function $Z(\alpha) = \sum_{i \in \Omega} \tilde{\pi}_i^{\alpha}$
  (NB: π: normalized, $\tilde{\pi}$: non-normalized).
- Fundamental expression 1

$$Z(\alpha') = \sum_{i \in \Omega} \tilde{\pi}_i^{\alpha'} = \sum_{i \in \Omega} \tilde{\pi}_i^{\alpha} \frac{\tilde{\pi}_i^{\alpha'}}{\tilde{\pi}_i^{\alpha}}$$

- Fundamental expression 2:

$$\frac{Z(\alpha')}{Z(\alpha)} = \frac{1}{Z(\alpha)} \sum_{i \in \Omega} \tilde{\pi}_i^{\alpha'} = \sum_{i \in \Omega} \pi_i^{\alpha} \frac{\tilde{\pi}_i^{\alpha'}}{\tilde{\pi}_i^{\alpha}} = \mathbb{E}\left(\frac{\tilde{\pi}_i^{\alpha'}}{\tilde{\pi}_i^{\alpha}} \right)_{\alpha}$$

- Fundamental expression 3:

$$Z(1) = \left[\frac{Z(1)}{Z(0.75)} \right] \left[\frac{Z(0.75)}{Z(0.5)} \right] \left[\frac{Z(0.5)}{Z(0.25)} \right] \left[\frac{Z(0.25)}{Z(0)} \right] Z(0)$$

- Only $Z(0)$ is known.

From thermodynamic integration to simulated annealing

Thermodynamic integration:

- MCMCs at $\alpha(t)$ with $\alpha(0) = 0$ and $\alpha(1) = 1$
- Samplings at different α independent.

Simulated annealing:

- The sampling at $\alpha = 1$ may be difficult in MCMC.
- Use the same sequence of α, but starting sample at α' equals final sample at α. Simulated annealing:
- Samplings at different α dependent.
- Most studied (basic convergence theorem: Hajek (1988))
- Meta-MCMC algorithm: Work on "cooling schedule" α_t rather than on $P^{\{\alpha=1\}}$.

Simulated annealing

Simulated annealing can be tested in:

- V-shaped stationary distribution
- Simulated annealing: Sequence of samplings that feed on each other (feed-forward), single sampling “active” at any time.
- Simulated tempering (feed-back, many samplings simultaneously active).