A Massively Parallel Evolutionary Markov Chain Monte Carlo Algorithm for Sampling Complicated Multimodal State Spaces

Wendy K. Tam Cho and Yan Y. Liu
University of Illinois at Urbana-Champaign

Summary

– Markov Chain Monte Carlo (MCMC) methods can sample from unknown distributions with asymptotic theoretical convergence
– For large problems with difficult to traverse state spaces, the convergence time for MCMC may be prohibitively long.
– MCMC is theoretically attractive but successful implementation for large and complex problems is challenging.
– Small transitions in the Markov chain result in small Metropolis-Hastings ratios and thus ineffective chains.
– Key to success is intelligent and efficient state space traversal
– Evolutionary Algorithms (EAs) are a class of optimization heuristics
  – EAs enable a diversified search
  – EAs lend themselves easily to parallelization
  – In an Evolutionary Markov Chain Monte Carlo (EMCMC), EA operators guide the movement of the Markov Chain, enabling large, diversified, and effective movement in the state space, which, along with parallelization to boost computation, hastens convergence of the Markov Chain.

Guiding Markov Chains with Optimization Heuristics

We wish to sample from a distribution $f(x) = \exp(-V(x)) / Z$.

In an EA, $H(x)$ is the fitness function and $x$ is a chromosome. Each chromosome has a weight that is proportional to its Boltzmann probability.

Our EA mutation operator (Liu and Cho, 2018), randomly chooses a chromosome $x_s$ and mutates it to $x_g$. The proposal is accepted with probability $\min(1, r_m)$ where

$$r_m = \frac{H(x_g)}{H(x_s)} = \exp(-V(x_g) - V(x_s)) / (1 + r_H)$$

is the Metropolis-Hastings (MH) ratio and $\theta_H$ is the transition probability. Since this mutation operator is a symmetric operator, the MH ratio can be computed from the fitness values of the chromosomes.

Our EA crossover operator randomly chooses two chromosomes, a source solution, $x_s$, and a different target solution, $x_t$. The operator is a spatial path relinking operator that converts a random allele from its value in $x_s$ to its value in $x_t$. If a feasible solution, $x_g$, is found on the path, chromosome $x_s$ is replaced in the population with $x_g$.

In the framework of Adaptive Direction Sampling (Roberts and Gilks, 1994), which is a Gibbs Sampler, $f^{(0)} = f^{(1)} = f^{(2)} = f^{(3)}$, where $f^{(0)} = f^{(1)}$, and $f^{(2)} = f^{(3)}$. A random walk between the source and target solutions, captured in $r$, determines the child solution.

The crossover operator guides larger moves that enable more efficiently and effective sampling within the MCMC framework.

Application to Computational Redistricting


Acknowledgements

Computational experiments in this research used the Blue Waters sustained-petaflop computing resources, which is supported by the NSF (Grants OCI-0725070 and ACI-1238931) and the state of Illinois under the project, “Massively Parallel Evolutionary Markov Chain Monte Carlo for Sampling Complicated Multimodal State Spaces.” Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. Research support was provided by the National Science Foundation, Grant SES-1725145/1728902, “Collaborative Research: High Performance Computational Standards for Redistricting.”

Contact

Wendy K. Tam Cho (wendycho@nmsu.edu), Departments of Political Science, Statistics, Mathematics, Asian American Studies, the College of Law, and NCSA at the University of Illinois at Urbana-Champaign. Yan Y. Liu (yliu@nmsu.edu), National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign.