Concatenation of Markov processes for Monte Carlo Integration

Sascha Holl
Max Planck Institute for Informatics
`sholl@mpi-inf.mpg.de`

Coauthor(s): Gurprit Singh

Markov Chain Monte Carlo (MCMC) is a sophisticated technique employed to sample from a probability distribution by generating an invariant time-homogeneous Markov process. When sampling by conventional methods is infeasible or computationally intractable, it is the only viable method available. MCMC has found widespread application in diverse fields such as statistics, machine learning, physics, and finance.

Even though MCMC is a powerful tool, it is also hard to control and tune the evolution of the underlying Markov process in practice. Simultaneously achieving rapid *local exploration* of the state space and *global discovery* of the target distribution is a challenging task.

The arguably most popular MCMC technique is the Metropolis-Hastings algorithm. Despite its popularity, which not least stems from its practical ease of use, it does not perform outstanding well in neither of the both aforementioned goals.

Its inherent reversibility leads to diffusive exploration behavior and backtracking, slowing down the local exploration and consequentially the convergence to the equilibrium distribution. Moreover, to ensure global discovery, practical implementations often need to rely on uninformed large scale perturbations of the current state.

Wang et al. [2] introduced a novel MCMC approach, which is based on the concatenation of Markov processes [1]. It allows the usage of an essentially arbitrary Markov process for local exploration. That way, the process can be chosen to satisfy a desired exploration behavior suitable for the state space at hand without worrying about invariance at this point. The process is executed up to a certain finite lifetime. After this time has elapsed, the process is *killed* and started afresh at a spawn location drawn from a *regeneration* distribution. The lifetime is chosen in a way ensuring that the overall process is invariant with respect to a given target distribution.

We generalize this idea and introduce it with appropriate rigor. We show how the validity of the method can be established for a more general class of Markov processes. We also allow the usage of a whole family of Markov processes for local exploration with possibly varying exploration characteristic. We establish a transfer mechanism between consecutive processes, which allows the user to specify the initial state of the newly spawned process to depend on the exit point of the previous one. Not least, we derive a Rao-Blackwellization technique which guarantees variance reduction in practice.

We showcase the potential of the framework in a practical rendering experiment. We compare the method proposed in [2] with existing methods based on Metropolis-Hastings algorithms with Random-Walk, Langevin and Hamiltonian proposals, respectively.

[1] Sharpe, Michael (1998). *General Theory of Markov Processes.* Pure and Applied Mathematics. Academic Press.

[2] Wang, Andi Q. and Pollock, Murray and Roberts, Gareth O. and Steinsaltz, David. Regeneration-enriched Markov processes with application to Monte Carlo. The Annals of Applied Probability. Institute of Mathematical Statistics.