A flat histogram method for inference with probabilistic and deterministic constraints

Stefano Ermon, Carla Gomes, Ashish Sabharwal and Bart Selman
Department of Computer Science
Cornell University
Ithaca, NY 14850
{ermonste,gomes,sabhar,selman}@cs.cornell.edu

Reasoning in a context where both probabilistic and deterministic dependencies are present at the same time is a challenging task with many real-world applications. Markov Chain Monte Carlo (MCMC) methods provide a general framework for sampling and probabilistic inference from complex probability distributions, as captured, for example, in a graphical model representation. However, in the presence of a set of hard constraints (i.e., deterministic dependencies), it often becomes difficult to even reach states that satisfy all such dependencies in the the Markov Chain.

We will consider a novel MCMC sampling strategy, inspired by the Wang-Landau method ([1]), which is a so-called *flat histogram* sampling strategy from statistical physics. Given a combinatorial space and an energy function (for instance, that describes the log-likelihood of each configuration), a *flat histogram* method is a sampling strategy based on an Adaptive Markov Chain that converges to a steady state where it samples uniformly from sets of configurations with equal energy. This form of sampling will spend approximately the same amount of time in areas with low density configurations (usually, low energy states) and in high density areas of the search space. Such a sampling strategy generally leads to a much broader coverage of the state space compared to more traditional MCMC approaches, such as Metropolis-Hastings or Gibbs. In particular, it solves the problems caused by near-deterministic dependencies, that greatly slow down inference by creating low probability regions that are difficult to traverse (and eventually breaking down the ergodicity in the limit of deterministic dependencies). Another advantage of this method is that we obtain the full density of states distribution, where the density of states is defined as a function that for each energy level E gives the number of configurations with that energy. This represents a rich description of the state space and we will show that upon designing the right energy function, the density of states can be used to infer complex statistical properties of the probability distribution, such as marginals for all possible levels of softness of the constraints.

Building on [2], we will consider a modification of the Wang-Landau method that incorporates a random-walk style component to focus the Markov Chain more quickly on areas where all hard constraints are satisfied. By enforcing a detailed balance condition, we maintain uniform sampling across the different energy levels and the consistency of the method. We provide empirical data to show the practical effectiveness of our method by comparing it in a marginal computation task with general purpose techniques such as Gibbs sampling and specialized methods such as MC-SAT [3].

References

[1] F. Wang and DP Landau. Efficient, multiple-range random walk algorithm to calculate the density of states. *Physical Review Letters*, 86(10):2050–2053, 2001.

[2] S. Ermon, C. Gomes, and B. Selman. Computing the density of states of Boolean formulas. In *Proceedings of the 16th International Conference on Principles and Practice of Constraint Programming*, 2010.

[3] H. Poon and P. Domingos. Sound and efficient inference with probabilistic and deterministic dependencies. In *Proc. of AAAI-06*, volume 21, page 458, 2006.