Pre-Fetching: Parallelisation of MCMC

Supervisor: Murray Pollock (Warwick)*

Interested students should schedule a meeting to discuss this project prior to selection
***Find updated project listing / availability at https://warwick.ac.uk/mpollock/projects ***

Overview

In many practical statistical settings it is simply not possible to analytically compute expectations of many quantities we are interested in with respect to the probability distributions induced by our chosen models. One popular (so called *Monte Carlo*) approach to tackle this problem is to simply construct an estimator of the quantity of interest by sampling randomness which is representative of the model. In particular, letting φ be a function of interest, π be our probability distribution, and supposing we have access to $X_1, X_2, \ldots, X_N \overset{\text{iid}}{\sim} \pi$, then,

$$\mathbb{E}_\pi \left[\varphi(X)\right] := \int_{\mathcal{S}} \varphi(x) \cdot \pi(x) \, \mathrm{d}x \approx \frac{1}{N} \sum_{i=1}^{N} \varphi\left(X_i\right). \tag{1}$$

Although seemingly naïve, the Monte Carlo approach is extremely powerful as it transpires that under weak assumptions the estimator is consistent, unbiased, and the rate at which the estimator converges to the expectation of interest is well understood.

Of course, critical to the Monte Carlo method is being able to sample $X_1, X_2, \ldots, X_N \overset{\text{iid}}{\sim} \pi$, and in general this is not possible. The *Markov chain Monte Carlo* method is perhaps the most powerful and popular Monte Carlo approach when direct access to samples of π are not possible. The innovation in this approach is to construct a *Markov chain* $\{X_j : j = 1, 2, \ldots\}$ with dynamics such that the limiting distribution of the Markov chain is π. Sampling π can then be achieved by simply sampling from the dynamics (which are typically simpler) of the constructed Markov chain.

Unfortunately, sampling from Markov chains is intrinsically sequential in nature, and so is poorly suited to any attempt to exploit modern multi- and many- core computing settings to parallelise. One approach which appears in the literature is *pre-fetching* [Brockwell, 2006, Strid, 2010, Angelino et al., 2014]. In essence, this approach uses all available computer resources to simultaneously consider many possible future states and steps of the Markov chain, and when all resources are exhausted chooses the 'best' future. This project will involve fully exploring existing literature on pre-fetching MCMC, attempting to implement such parallel schemes, and determining the extent to which they offer an advantage over single-core approaches (or indeed, other parallel schemes).

References

[Angelino et al., 2014] Angelino, E., Kohler, E., Waterland, A., Seltzer, M., and Adams, R. P. (2014). Accelerating mcmc via parallel predictive prefetching. *arXiv preprint arXiv:1403.7265*.

*Email: *m.pollock@warwick.ac.uk*

[Brockwell, 2006] Brockwell, A. E. (2006). Parallel markov chain monte carlo simulation by pre-fetching. *Journal of Computational and Graphical Statistics*, 15(1):246–261.

[Strid, 2010] Strid, I. (2010). Efficient parallelisation of metropolis–hastings algorithms using a prefetching approach. *Computational Statistics & Data Analysis*, 54(11):2814–2835.