

1. Pseudo-Marginal Markov Chain Monte Carlo

Proposer: Geoff Nicholls

Brief Description: The Metropolis Hastings algorithm is a procedure for simulating an update of a Markov chain. It can be adjusted so that the transition matrix it simulates is stationary with respect to any particular probability distribution the user wishes to provide. The realised sequence of Markov Chain states can then be used to estimate expectations in that target distribution, via the ergodic theorem for irreducible and aperiodic Markov chains. One of the attractions of this algorithm is that it is sufficient to know the target distribution up to an overall constant. However, there are applications in which the target distribution cannot even be evaluated up to a constant. This happens, for example, in some Bayesian inference, where the target distribution is a posterior probability distribution proportional to the product of prior and likelihood, and the likelihood is given as the solution of an equation which we cannot solve. Such problems are called "doubly intractable".

Following an idea due to Mark Beaumont, the authors Gareth Roberts and Christophe Andrieu gave an extension to the Metropolis Hastings algorithm which deals with these doubly intractable problems. The idea is to replace the target distribution with an unbiased estimate of the target distribution. Importance sampling is used to estimate the target distribution, and this estimate plugged into the standard Metropolis Hastings algorithm for the target. The authors above showed that the resulting chain would be ergodic, with equilibrium equal the chosen target distribution.

In this project the student will examine the new algorithm, and develop example applications to illustrate its strengths and weaknesses. The project can be developed at three levels: we can make simple applications to cases where we know the distribution exactly (so pseudo marginal is not needed); we can make applications to problems from spatial statistics, where the distributions are less familiar, but the estimation is still based on simple importance sampling; we can replace the importance sampler with a particle filter, which is a more elaborate estimation scheme, and work on simple time-structured examples from simulated financial time-series.

Andrieu,C. and Roberts,G. O., (2009), 'The pseudo-marginal approach for efficient Monte Carlo computations', *Annals of Statistics* vol 37, (2), pp 697-725
Beaumont,M. A., (2003), 'Estimation of population growth or decline in genetically monitored populations', *Genetics* vol 164 (3), pp 1139-1160}
There are some nice examples in Darren Wilkinson's blog: <https://darrenjw.wordpress.com/tag/pseudo-marginal/>

Prerequisite courses: BS3a or Part A *Simulation*

Type of project: Theoretical/Simulation study

Computing required? Yes.

Data available? Not applicable