

CONTRIBUTIONS TO COMPUTATIONAL BAYESIAN STATISTICS

LEAH SOUTH 

(Received 13 December 2019; first published online 5 February 2020)

2010 Mathematics subject classification: primary 65C05; secondary 65C60.

Keywords and phrases: Bayesian statistics, approximate Bayesian computation, sequential Monte Carlo, nested sampling, marginal likelihood, control variates, cell biology, ordinary differential equations.

Statisticians and applied practitioners often desire the ability to work with complex statistical models. Such models can lead to a more complete understanding of processes in comparison to simpler models that are computationally easy to fit. However, standard methods for performing inference on such models can be inefficient or infeasible. The aim of this thesis is to improve practitioners' ability to work in this setting through novel methodology and open-source software.

All methodology in this thesis is focused on Bayesian statistics, where the key quantities of interest are posterior expectations and the normalising constant of the posterior. A common way to estimate these quantities is through Monte Carlo methods which use pseudo-random sampling to help estimate the underlying integrals. Markov chain Monte Carlo (MCMC) is the most popular Monte Carlo method when the likelihood function is tractable or can be unbiasedly estimated. Standard methods like MCMC are infeasible when the likelihood function is intractable and can be inefficient when the posterior distribution is complex or multimodal. This thesis makes contributions to three areas of Monte Carlo inference: advancing methods for when the likelihood function is intractable, improving the efficiency and flexibility of sequential Monte Carlo (SMC) methods and extending *post-hoc* methods for variance reduction of estimators. These contributions are illustrated through challenging examples from biology, ecology and econometrics.

The first contribution of this thesis is related to likelihood-free methods such as approximation Bayesian computation (ABC). The thesis develops a novel method, Bayesian synthetic likelihood (BSL), which helps to reduce the limitations associated with ABC. ABC methods generally involve reducing the data to a set of summary

Thesis submitted to Queensland University of Technology in January 2019; degree approved on 28 August 2019; principal supervisor Christopher Drovandi, associate supervisor Anthony Pettit.

© 2020 Australian Mathematical Publishing Association Inc.

statistics. They essentially amount to choosing parameters that produce simulated summary statistics that are sufficiently ‘close’ to the observed summary statistics using nonparametric methods. Although ABC has been shown to be useful in applications, it has many drawbacks. ABC often involves many model simulations, significant tuning and difficulty in handling high-dimensional summary statistics. BSL helps to reduce these difficulties for a wide range of applications through the use of parametric assumptions.

Two contributions are made to SMC methods for static Bayesian models. SMC methods are naturally adaptable, embarrassingly parallelisable and capable of handling complicated and multimodal posterior distributions. However, their use for static models is comparatively infrequent relative to MCMC methods. One contribution of this work is to make better use of all information produced in the SMC process. Additionally, a new SMC method is developed to provide methods with improved theoretical support when there is a sharp peak in the posterior (a phase transition).

Finally, the fourth contribution is in variance reduction of Monte Carlo estimators of posterior expectations and the normalising constant of the posterior. Zero-variance control variates (ZV-CV) and control functionals (CF) are becoming increasingly popular for their ability to reduce the variance of estimators using derivatives of the log posterior density. The thesis develops a new method, regularised ZV-CV, that offers more efficient Bayesian inference for higher dimensional examples than those that are typically tackled by ZV-CV or CF.

Papers from this thesis are available at [1–4].

References

- [1] L. F. Price (née South), C. C. Drovandi, A. Lee and D. J. Nott, ‘Bayesian synthetic likelihood’, *J. Comput. Graph. Statist.* **27**(1) (2018), 11 pages.
- [2] R. Salomone, L. F. South, C. C. Drovandi and D. P. Kroese, ‘Unbiased and consistent nested sampling via sequential Monte Carlo’, Preprint, 2018, arXiv:1805.03924.
- [3] L. F. South, C. J. Oates, A. Mira and C. C. Drovandi, ‘Regularised zero-variance control variates for high-dimensional variance reduction’, Preprint, 2018, arXiv:1811.05073.
- [4] L. F. South, A. N. Pettitt and C. C. Drovandi, ‘Sequential Monte Carlo samplers with independent MCMC proposals’, *Bayesian Anal.* **14**(3) (2019), 753–776.

LEAH SOUTH, School of Mathematical Sciences,
Queensland University of Technology, 2 George Street,
Brisbane City, Queensland 4000, Australia
e-mail: leah.south@hdr.qut.edu.au