Discussion of "Riemann manifold Langevin and Hamiltonian Monte Carlo methods" by M. Girolami and B. Calderhead

Anindya Bhadra

Department of Statistics, Texas A&M University, College Station, TX 77843-3143.

The authors are to be congratulated for a novel design of an efficient and automatic choice of the preconditioning matrix for MALA or mass matrix for HMC schemes. The clever use of local curvature information results in possible improvements in the relative speed of convergence to the high dimensional target distribution, as demonstrated by the authors using various illustrative examples.

The full MMALA and RMHMC schemes described by the authors require (a) evaluations of the partial derivatives up to the third order for the log-likelihood function and (b) inversion of the position specific metric tensor of the Riemann manifold formed by the parameter space. In the general case, considering the absence of nice analytical properties inducing sparsity in the covariance matrix etc., these two steps are computationally intensive (as pointed out by the authors) and in many of the examples the authors are forced to resort to a simplified version of the MMALA scheme.

One interesting application would be to use the authors' approach for an MCMC based stochastic optimization scheme like simulated annealing. At the early stages of the algorithm, with high temperature, the use of local curvature information in the MCMC proposal should result in high acceptance rate and the search should quickly reach a neighborhood of the global maxima (ref. Fig. 1). However, once this has been achieved, the use of local curvature information is unlikely to have much further benefit at the expense of the substantial computational burden imposed by steps (a) and (b) mentioned above and a vanilla MCMC (or global MALA or HMC) can be used once the temperature has cooled sufficiently to do a local search.

Numerical schemes for computing derivatives that are needed by the authors are often unstable. Ionides et al. (2006) proposed iterated filtering, a derivative free maximum likelihood based inference technique for partially observed Markovian state space models (an example of such a model is presented by the authors in section 8) that has been succesfully applied in many scientific applications (e.g., Laneri et al., 2010; Bretó et al., 2009; He et al., 2010). From the computational perspective, a favorable comparison of iterated filtering with Particle MCMC technique presented in Andrieu et al. (2010) is presented in Bhadra (2010) and iterated filtering presents a viable maximum likelihood alternative to Bayesian inference in many difficult situations.

References

Andrieu, C., Doucet, A., and Holenstein, R. (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 72:269–342.

Bhadra, A. (2010). Discussion of Particle Markov chain Monte Carlo methods by Andrieu et al. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 72:314–315.

Bretó, C., He, D., Ionides, E. L., and King, A. A. (2009). Time series analysis via mechanistic models. *Annals of Applied Statistics*, 3:319–348.

He, D., Ionides, E. L., and King, A. A. (2010). Plug-and-play inference for disease dynamics: Measles in large and small towns as a case study. *Journal of the Royal Society Interface*, 7:271–283.

Ionides, E. L., Bretó, C., and King, A. A. (2006). Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences of the USA*, 103:18438–18443.

Laneri, K., Bhadra, A., Ionides, E. L., Bouma, M., Dhiman, R. C., Yadav, R. S., and Pascual, M. (2010). Forcing versus feedback: Epidemic malaria and monsoon rains in NW India. *PLoS Computational Biology*, 6:e1000898.