Posterior Inference in Latent Gaussian Models Using Manifold MCMC Methods

Maurizio Filippone
Department of Computing Science
University of Glasgow
maurizio@dcs.gla.ac.uk

Mark Girolami
Department of Statistical Science
University College London
girolami@dcs.gla.ac.uk

1 Abstract

In this work, we study the inference problem in latent Gaussian models [5]. These models comprise a set of hyperparameters that parametrize the covariance function of a Gaussian Process from which latent functions are drawn; observations follow distributions (such as Bernoulli, Poisson, and so on) with statistics given by a transformed version of the latent function at their location. Efficiently sampling from the posterior distribution of the latent process and hyperparameters is complex because of their strong coupling [3, 4]. Gibbs style samplers, based on alternatively sampling one set of variables conditioned on the other and the observed data, are often employed in this kind of scenario. Although they are convenient from an implementation point of view, they are extremely inefficient, given that fixing the latent variables induces a sharply peaked posterior on the hyperparameters. The result is a very slow exploration of the posterior distribution that leads to a very poor Effective Sample Size (ESS) for the length-scale parameters of the covariance function [3]. This limitation can be overcome by whitening the prior over the latent variables. This technique simply amounts in replacing the sampling from the latent variables with their whitened version given by the covariance matrix. In other words, rather than sampling the latent process, we sample the (uncorrelated standard normal) variables that would be used to generate the latent variables.

We consider a set of recently proposed methods for efficiently sampling from highly correlated posterior distributions that make use of the natural geometry of the underlying statistical model [2]. The key quantity in information geometry is the Fisher Information (FI). It has already been shown that manifold methods applied to the unwhitened models are not efficient in terms of ESS per number of computations needed [1]. This is mainly due to their computationally intensive nature and to the block diagonal structure of the FI that does not yield a natural decoupling between the sets of variables. In this work, we apply manifold methods to the whitened models. We show that in this case the FI matrix is no longer block diagonal, and is therefore able to capture the correlations between the two sets of variables. We provide a set of experiments showing that manifold methods applied to these models yield a considerably better ESS than methods that do not exploit the geometry of the statistical model.

References

[1] M. Filippone. Discussion of the paper “Riemann Manifold Langevin and Hamiltonian Monte Carlo” by M. Girolami and B. Calderhead. *Journal Of The Royal Statistical Society Series B*.

[2] M. Girolami and B. Calderhead. Riemann Manifold Langevin and Hamiltonian Monte Carlo. *Journal Of The Royal Statistical Society Series B* (with discussion), to appear.

[3] I. Murray and R. P. Adams. Slice sampling covariance hyperparameters of latent Gaussian models. In *Advances in Neural Information Processing Systems 23 (NIPS 2010)* to appear.

[4] R. M. Neal. Regression and classification using gaussian process priors (with discussion). *Bayesian Statistics*, 6:475–501, 1999.

[5] C. E. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.