Journal of Computational Physics 308 (2016) 81–101

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Emulation of higher-order tensors in manifold Monte Carlo methods for Bayesian Inverse Problems

Shiwei Lan [a,*], Tan Bui-Thanh [b], Mike Christie [c], Mark Girolami [a,d,*]

[a] Department of Statistics, University of Warwick, Coventry CV4 7AL, UK
[b] Department of Aerospace Engineering and Engineering Mechanics, Institute for Computational Engineering & Sciences, The University of Texas, Austin, TX 78712, USA
[c] Institute of Petroleum Engineering, Heriot-Watt University, Edinburgh, EH14 4AS, UK
[d] The Alan Turing Institute for Data Science, British Library, 96 Euston Road, London, NW1 2DB, UK

ARTICLE INFO

Article history:
Received 22 July 2015
Received in revised form 17 November 2015
Accepted 14 December 2015
Available online 17 December 2015

Keywords:
Markov Chain Monte Carlo
Hamiltonian Monte Carlo
Gaussian Process emulation
Bayesian Inverse Problems
Uncertainty quantification

ABSTRACT

The Bayesian approach to Inverse Problems relies predominantly on Markov Chain Monte Carlo methods for posterior inference. The typical nonlinear concentration of posterior measure observed in many such Inverse Problems presents severe challenges to existing simulation based inference methods. Motivated by these challenges the exploitation of local geometric information in the form of covariant gradients, metric tensors, Levi-Civita connections, and local geodesic flows have been introduced to more effectively locally explore the configuration space of the posterior measure. However, obtaining such geometric quantities usually requires extensive computational effort and despite their effectiveness affects the applicability of these geometrically-based Monte Carlo methods. In this paper we explore one way to address this issue by the construction of an emulator of the model from which all geometric objects can be obtained in a much more computationally feasible manner. The main concept is to approximate the geometric quantities using a Gaussian Process emulator which is conditioned on a carefully chosen design set of configuration points, which also determines the quality of the emulator. To this end we propose the use of statistical experiment design methods to refine a potentially arbitrarily initialized design online without destroying the convergence of the resulting Markov chain to the desired invariant measure. The practical examples considered in this paper provide a demonstration of the significant improvement possible in terms of computational loading suggesting this is a promising avenue of further development.

1. Introduction

In Bayesian Inverse Problems one needs to draw samples from a typically high dimensional and complicated intractable probability measure [1]. Samples are needed to estimate integrals for e.g. point estimates or interval estimates for uncertainty quantification. Random Walk Metropolis (RWM) is hampered with its random walk nature, and Hybrid Monte Carlo (HMC) [2–9] with its exploitation of local gradients and approximate Hamiltonian flows in an expanded phase space can substantially improve over RWM. Riemannian Manifold Hamiltonian Monte Carlo [5] further takes advantage of local metric

* Corresponding authors.
E-mail addresses: S.Lan@warwick.ac.uk (S. Lan), tanbui@ices.utexas.edu (T. Bui-Thanh), mike.christie@pet.hw.ac.uk (M. Christie), M.Girolami@warwick.ac.uk (M. Girolami).

http://dx.doi.org/10.1016/j.jcp.2015.12.032
0021-9991/

tensors to adapt the transition kernel of the Markov chain to the local structure of the probability measure, and indeed the proposal mechanism is provided by the local geodesic flows on the manifold of probability measures [5]. This has been demonstrated to allow Markov Chain Monte Carlo (MCMC) to effectively explore the types of challenging posterior measures observed in many Inverse Problems, see e.g. [10] and the example in Fig. 1 in this paper.

The challenge here is that these geometric objects including gradients, metrics, connection components are typically expensive to compute, hindering their application in practice. This is due to the requirement of a single forward solve of the model in evaluating the likelihood, and this increases with the choice of metric tensor and associated connections (second and third order tensors), see [10] for detailed developments which exploit adjoint solver codes.

In this contribution we investigate the feasibility of emulating these expensive geometric quantities using a Gaussian Process model [11]. The remainder of the paper has the following structure. Section 2 briefly reviews Hamiltonian Monte Carlo methods, Sections 3 and 4 detail the Gaussian Process emulation of potential energies, gradients, second order metric tensors and third order tensor metric connections. Since it is impossible to emulate the expected Fisher metric [5] based on the Gaussian Process assumption, we propose to emulate the empirical Fisher information in this work. The accuracy of the GP emulator to approximate these geometric quantities depends on the design set, or configurations, which should be well spread over the distribution to capture its geometry. Given a well chosen design set that is fixed a priori, the emulated MCMC algorithms can scale up with dimensions reasonably well due to emulation of higher oder tensors also being a linear prediction problem. However it is unreasonable to assume that such a good design set is available initially or indeed generally. Therefore Section 5 introduces *regeneration* [12–14] as a general adaptation framework and experimental design algorithm *Mutual Information for Computer Experiments (MICE)* [15] to refine the design set. It is a well-known challenging problem to obtain good design sets in high dimensions in general. Although the proposed method by exploiting MCMC samples may be not ready for applications of thousand dimensions, it is a novel attempt and worth further development. We illustrate the advantage of emulation for geometric Monte Carlo algorithms over their full versions with examples in Section 6. Finally in Section 7, we summarize the contribution and discuss some future directions of investigation.

2. Review of dynamics and geometry inspired simulation methods

2.1. HMC

Hybrid Monte Carlo (HMC) [2,3] is a Metropolis style algorithm that reduces its random walk behavior by making distant proposals guided by Hamiltonian flows. These distant proposals are found by numerically simulating Hamiltonian dynamics, whose state space consists of its *position* vector, $\boldsymbol{\theta} \in \mathbb{R}^D$, and its momentum vector, $\mathbf{p} \in \mathbb{R}^D$. In application to statistical models, $\boldsymbol{\theta}$ consists of the model parameters (and perhaps latent variables), and $\mathbf{p}$ are auxiliary variables. The objective is to sample from the posterior distribution $\pi(\boldsymbol{\theta}|\mathcal{D}) \propto \pi(\boldsymbol{\theta})L(\boldsymbol{\theta}|\mathcal{D})$, where $\pi(\boldsymbol{\theta})$ is the prior and $L(\boldsymbol{\theta}|\mathcal{D})$ is the likelihood function. We define the *potential energy* as $U(\boldsymbol{\theta}) := -\log \pi(\boldsymbol{\theta}|\mathcal{D})$, and the *kinetic energy*, $K(\mathbf{p})$, similarly as the minus log of the density of $\mathbf{p}$, which is usually assumed $\mathbf{p} \sim \mathcal{N}(\mathbf{0}, \mathbf{M})$. Then the total energy, *Hamiltonian* function is defined as their sum:

$$H(\boldsymbol{\theta}, \mathbf{p}) = U(\boldsymbol{\theta}) + K(\mathbf{p}) = -\log \pi(\boldsymbol{\theta}|\mathcal{D}) + \frac{1}{2}\mathbf{p}^{\mathsf{T}}\mathbf{M}^{-1}\mathbf{p} \tag{1}$$

Therefore the joint density of $\boldsymbol{\theta}$ and $\mathbf{p}$ is $\pi(\boldsymbol{\theta}, \mathbf{p}) \propto \exp(-H(\boldsymbol{\theta}, \mathbf{p}))$. Note, the covariance matrix $\mathbf{M}$ is also referred as the constant *mass matrix*.

Given the current state $\boldsymbol{\theta}$, we sample the momentum $\mathbf{p} \sim \mathcal{N}(\mathbf{0}, \mathbf{M})$, and evolve the joint state $\mathbf{z} := (\boldsymbol{\theta}, \mathbf{p})$ according to *Hamilton's equations*:

$$\begin{aligned} \dot{\boldsymbol{\theta}} &= \frac{\partial H}{\partial \mathbf{p}} = \mathbf{M}^{-1}\mathbf{p} \\ \dot{\mathbf{p}} &= -\frac{\partial H}{\partial \boldsymbol{\theta}} = -\nabla_{\boldsymbol{\theta}} U(\boldsymbol{\theta}) \end{aligned} \tag{2}$$

The resulting Hamiltonian dynamics are 1) time reversible, and 2) volume preserving. In practice, however, it is difficult to solve Hamiltonian's equations analytically, so numerical methods such as, *leapfrog* (or Störmer–Verlet) [16,17], to approximate these equations by discretizing time with small step size ε. In the standard HMC algorithm, L, of these leapfrog steps, with some step size, ε, are used to propose a new state, which is either accepted or rejected according to the Metropolis acceptance probability [One can refer to [3], for more details].

2.2. RHMC

While HMC explores the parameter space more efficiently than *Random Walk Metropolis (RWM)*, it does not fully exploit the geometric properties of the parameter space. In some complex scenarios, e.g. the concentrated nonlinear distribution in Fig. 1, HMC does not explore the parameter space as 'straightforwardly' as RHMC does. To take advantage of the Riemannian geometry of statistical models, Girolami and Calderhead [5] propose *Riemannian Manifold HMC (RHMC)* to improve the efficiency of the standard HMC by automatically adapting to the local structure of the parameter space. Following the