



Spectral likelihood expansions for Bayesian inference



Joseph B. Nagel^{*}, Bruno Sudret^{**}

ETH Zürich, Institute of Structural Engineering, Chair of Risk, Safety & Uncertainty Quantification, Stefano-Franscini-Platz 5, CH-8093 Zürich, Switzerland

ARTICLE INFO

Article history:

Received 19 June 2015

Received in revised form 9 October 2015

Accepted 22 December 2015

Available online 4 January 2016

Keywords:

Uncertainty quantification

Bayesian inference

Inverse problem

Surrogate modeling

Polynomial chaos

Posterior expansion

ABSTRACT

A spectral approach to Bayesian inference is presented. It pursues the emulation of the posterior probability density. The starting point is a series expansion of the likelihood function in terms of orthogonal polynomials. From this spectral likelihood expansion all statistical quantities of interest can be calculated semi-analytically. The posterior is formally represented as the product of a reference density and a linear combination of polynomial basis functions. Both the model evidence and the posterior moments are related to the expansion coefficients. This formulation avoids Markov chain Monte Carlo simulation and allows one to make use of linear least squares instead. The pros and cons of spectral Bayesian inference are discussed and demonstrated on the basis of simple applications from classical statistics and inverse modeling.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

In view of inverse modeling [1,2] and uncertainty quantification [3,4], Bayesian inference establishes a convenient framework for the data analysis of engineering systems [5–7]. It adopts probability theory in order to represent, propagate and update epistemic parameter uncertainties. The prior distribution captures the uncertainty of unknown model parameters before the data are analyzed. A posterior is then constructed as an updated distribution that captures the remaining uncertainty after the data have been processed. The computation of this posterior is the primary task in Bayesian inference.

Some simplified statistical models admit closed-form expressions of the posterior density. Beyond these so-called conjugate cases, computational approaches either aim at evaluating expectation values under the posterior or drawing samples from it [8]. This is usually accomplished with stochastic methods such as Markov chain Monte Carlo [9,10]. Nowadays this class of techniques constitutes the mainstay of Bayesian computations. The posterior is explored by realizing an appropriate Markov chain over the prior support that exhibits the posterior as its long-run distribution. In turn, the obtained sample is used to empirically approximate the statistical quantities of interest. These include the characteristics of the posterior and the predictive distribution. This way of proceeding suffers from some inherent deficiencies. The presence of sample autocorrelation and the absence of a convergence criterion cause severe practical problems. Moreover, Markov chain Monte Carlo typically requires a large number of serial forward model runs. Since in engineering applications even a single model run can be computationally taxing, this may be prohibitive. In the recent past, numerous enhancements have been proposed in order to accelerate Markov chain Monte Carlo for Bayesian inverse problems. This includes the implementation of more efficient sampling algorithms, e.g. transitional Markov chain Monte Carlo [11,12] or Hamiltonian Monte Carlo [13–15], and the

^{*} Corresponding author.

^{**} Principal corresponding author.

E-mail addresses: nagel@ibk.baug.ethz.ch (J.B. Nagel), sudret@ibk.baug.ethz.ch (B. Sudret).

substitution of the forward model with an inexpensive metamodel, e.g. based on Gaussian process models [16,17] or polynomial chaos expansions [18–20]. Although these approaches promise significant speedups, they still inherit all principle shortcomings of sample-based posterior representations.

Unfortunately there are only few fundamental alternatives to stochastic sampling. Variational Bayesian inference establishes such an alternative where the posterior is sought through deterministic optimization [21–23]. In particular, a member from a simple parametric family of probability densities is selected such that some distance to the posterior is minimized. In this regard, the Kullback–Leibler divergence is often chosen as a relative measure of the dissimilarity of two probability densities. The procedure commonly rests upon some simplifying independence assumptions. Variational methods are regarded as less computing intensive than Markov chain Monte Carlo, yet they are only approximate. They are prominently used in machine learning and computer science [24,25], and since recently such methods are applied to inverse problems [26,27], too. A particularly interesting implementation of variational Bayesian inference has been proposed in [28]. The posterior is parametrized as a transformation of the prior density and can be computed based on the corresponding back-transformation. More specifically, a random variable transformation is sought in polynomial form such that the Kullback–Leibler divergence between the prior density and the back-transformed posterior density is minimized. This formulation is supported by arguments from optimal transport theory which also allows for a practical regularization of the problem. Finally, samples from the posterior distribution are obtained by independently sampling from the prior and applying the polynomial map. Another approach to certain Bayesian inverse problems has been recently devised in [29,30]. Based on monomial Taylor expansions of the forward model and of the posterior density, the computation of expectation values under the posterior is tackled by sparse numerical quadrature.

In this paper we propose a novel approach to surrogate the posterior probability density in itself. The main idea is to decompose the likelihood function into a series of polynomials that are orthogonal with respect to the prior density. It is shown that all statistical quantities of interest can then be easily extracted from this spectral likelihood expansion. Emulators of the joint posterior density and its marginals are derived as the product of the prior, that functions as the reference density, and a linear combination of polynomials, that acts as an adjustment. In doing so, the model evidence simply emerges as the coefficient of the constant term in the expansion. Moreover, closed-form expressions for the first posterior moments in terms of the low-order expansion coefficients are given. The propagation of the posterior uncertainty through physical models can be easily accomplished based on a further postprocessing of the expansion coefficients. In this sense, spectral Bayesian inference is semi-analytic. While the corrections required for an expansion of the posterior with respect to the prior as the reference density may be large, they can be small for an expansion around a properly chosen auxiliary density. A change of the reference density is therefore suggested in order to increase the efficiency of computing a posterior surrogate. The devised formulation entirely avoids Markov chain Monte Carlo. Instead it draws on the machinery of spectral methods [31–33] and approximation theory [34–36]. It is proposed to compute the expansion coefficients via linear least squares [37,38]. This allows one to make use of the wealth of statistical learning methods [39,40] that are designed for this type of problems. The approach features a natural convergence criterion and it is amenable to parallel computing.

The scope of applicability of the proposed approach covers problems from Bayesian inference for which the likelihood function can be evaluated and for which polynomials can be constructed that are orthogonal with respect to the prior, possibly after a carefully chosen variable transformation. This excludes statistical models that involve intractable likelihoods [41,42], i.e. the likelihood cannot be exactly evaluated. It also excludes improper prior distributions [43,44], i.e. the prior does not integrate to one or any finite value, and models with pathologic priors such as the Cauchy distribution for which the moments are not defined [45,46]. Many hierarchical Bayesian models [47,48] are not covered by the devised problem formulation. They are either based on conditional priors, which does not allow for orthogonal polynomials, or on integrated likelihoods, which can only be evaluated subject to noise.

Spectral likelihood expansions complement the existing array of Bayesian methods with a way of surrogating the posterior density directly. They have the potential to remedy at least some of the shortcomings of Markov chain Monte Carlo. Yet, their practical implementation poses challenges. Hence, the goal of this paper is to discuss and investigate the possibilities and limitations of the approach. The method of spectral likelihood expansions is therefore applied to well-known calibration problems from classical statistics and inverse heat conduction. We restrict the analysis to low-dimensional problems. The final results are compared with corresponding results from Markov chain Monte Carlo simulations.

The manuscript is structured as follows. The principles of Bayesian inference are summarized in Section 2. Surrogate forward modeling with polynomial chaos expansions is reviewed in Section 3. After that, spectral likelihood expansions are introduced as an alternative approach to Bayesian inference in Section 4. Two well-known Gaussian fitting examples and the identification of thermal properties of a composite material serve as numerical demonstrations in Section 5. Finally, it is summarized and concluded in Section 6.

2. Bayesian inference

Let $\mathbf{x} = (x_1, \dots, x_M)^\top \in \mathcal{D}_{\mathbf{x}}$ with $\mathcal{D}_{\mathbf{x}} = \mathcal{D}_{x_1} \times \dots \times \mathcal{D}_{x_M} \subset \mathbb{R}^M$ be a vector of unknown parameters. The goal of statistical inference is to deduce these unknowns from the observed data $\mathbf{y} = (y_1, \dots, y_N)^\top \in \mathbb{R}^N$. In Bayesian statistics one adapts probabilistic models for representing uncertainty. Hence, let $(\Omega, \mathcal{F}, \mathcal{P})$ be a suitable probability space with a sample space Ω , a σ -field \mathcal{F} and a probability measure \mathcal{P} . On this space one can define a prior model of the unknowns and an

Download English Version:

<https://daneshyari.com/en/article/6930533>

Download Persian Version:

<https://daneshyari.com/article/6930533>

[Daneshyari.com](https://daneshyari.com)