Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 299 (2016) 215–244
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Sparse Variational Bayesian approximations for nonlinear inverse problems: Applications in nonlinear elastography

Isabell M. Franck, P.S. Koutsourelakis*

Professur für Kontinuumsmechanik, Technische Universität München, Boltzmannstrasse 15, 85747 Garching (b. München), Germany

Received 18 December 2014; received in revised form 15 October 2015; accepted 19 October 2015
Available online 12 November 2015

Abstract

This paper presents an efficient Bayesian framework for solving nonlinear, high-dimensional model calibration problems. It is based on a Variational Bayesian formulation that aims at approximating the exact posterior by means of solving an optimization problem over an appropriately selected family of distributions. The goal is two-fold. Firstly, to find lower-dimensional representations of the unknown parameter vector that capture as much as possible of the associated posterior density, and secondly to enable the computation of the approximate posterior density with as few forward calls as possible. We discuss how these objectives can be achieved by using a fully Bayesian argumentation and employing the marginal likelihood or evidence as the ultimate model validation metric for any proposed dimensionality reduction. We demonstrate the performance of the proposed methodology for problems in nonlinear elastography where the identification of the mechanical properties of biological materials can inform non-invasive, medical diagnosis. An Importance Sampling scheme is finally employed in order to validate the results and assess the efficacy of the approximations provided.

Keywords: Uncertainty quantification; Variational Bayesian; Inverse problem; Dimensionality reduction; Elastography; Dictionary learning

1. Introduction

The extensive use of large-scale computational models poses several challenges in model calibration as the accuracy of the predictions provided depends strongly on assigning proper values to the various model parameters. In mechanics of materials, accurate mechanical property identification can guide damage detection and an informed assessment of the system's reliability [1]. Identifying property-cross correlations can lead to the design of multi-functional materials [2]. Permeability estimation for soil transport processes can assist in detection of contaminants, oil exploration [3].

Deterministic optimization techniques which have been developed to address these problems [4], lead to point estimates for the unknowns without rigorously considering the statistical nature of the problem and without providing

* Corresponding author. Tel.: +49 89 289 16690.
E-mail addresses: franck@tum.de (I.M. Franck), p.s.koutsourelakis@tum.de (P.S. Koutsourelakis).
URL: http://www.contmech.mw.tum.de (P.S. Koutsourelakis).

http://dx.doi.org/10.1016/j.cma.2015.10.015

quantification of the uncertainty in the inverse solution. Statistical approaches based on the Bayesian paradigm [5] on the other hand, aim at computing a (posterior) probability distribution on the parameters of interest. Bayesian formulations offer several advantages as they provide a unified framework for dealing with the uncertainty introduced by the incomplete and noisy measurements. Significant successes have been noted in applications such as geological tomography [6], medical tomography [7], petroleum engineering [8], as well as a host of other physical, biological, or social systems [9,10]. Representations of the parametric fields in existing deterministic and Bayesian approaches (artificially) impose a minimum length scale of variability usually determined by the discretization size of the governing PDEs [11]. As a result they give rise to a very large vector of unknowns. Inference in high-dimensional spaces using standard Markov Chain Monte Carlo (MCMC) schemes is generally impractical as it requires an exuberant number of calls to the forward simulator in order to achieve convergence. Advanced schemes such as those employing Sequential Monte Carlo samplers [12,13], adaptive MCMC [14], accelerated MCMC methods [15] or spectral methods [16] can alleviate some of these difficulties particularly when the posterior is multi-modal but still pose significant challenges in terms of the computational cost [17].

This work is particularly concerned with the identification of the mechanical properties of biological materials, in the context non-invasive medical diagnosis. While in certain cases mechanical properties can also be measured directly by excising multiple tissue samples, non-invasive procedures offer obvious advantages in terms of ease, cost and reducing risk of complications to the patient. Rather than X-ray techniques which capture variations in density, the identification of stiffness or mechanical properties in general, can potentially lead to earlier and more accurate diagnosis [18,19], provide valuable insights that differentiate between modalities of the same pathology [20] and monitor the progress of treatments. In this paper we do not propose new imaging techniques but rather aim at developing rigorous statistical models and efficient computational tools that can make use of the data/observables (i.e. noisy displacements of deformed tissue) from existing imaging modalities (such as magnetic resonance [21], ultrasonic) in order to produce certifiable estimates of mechanical properties. The primary imaging modality considered in this project is ultrasound elasticity imaging (elastography [22,23]). It is based on ultrasound tracking of pre- and post-compression images to obtain a map of position changes and deformations of the specimen due to an external pressure/load. The pioneering work of Ophir and coworkers [24] followed by several clinical studies [25–27] have demonstrated that the resulting strain images typically improve the diagnostic accuracy over ultrasound alone.

Beyond a mere strain imaging there are two approaches for inferring the constitutive material parameters. In the *direct approach*, the equations of equilibrium are interpreted as equations for the material parameters of interest, where the inferred strains and their derivatives appear as coefficients [28]. While such an approach provides a computationally efficient strategy, it does not use the raw data (i.e. noisy displacements) but transformed versions i.e. strain fields (or even-worse, strain derivatives) which arise by applying sometimes ad hoc filtering and smoothing operators. As a result the informational content of the data is compromised and the quantification of the effect of observation noise is cumbersome. Furthermore, the smoothing employed can smear regions with sharply varying properties and hinder proper identification.

The alternative to direct methods, i.e. *indirect or iterative* procedures admit an inverse problem formulation where the discrepancy between observed and model-predicted displacements is minimized with respect to the material fields of interest [29–32]. While these approaches utilize directly the raw data, they generally imply an increased computational cost as the forward problem and potentially derivatives have to be solved/computed several times. This effort is amplified when stochastic/statistical formulations are employed as those arising in the Bayesian paradigm. Technological advances have led to the development of hand-carried ultrasound systems in the size of a smartphone [33]. Naturally their accuracy and resolution does not compare with the more expensive traditional ultrasound machines or even more so MRI systems. If however computational tools are available that can distill the informational content from noisy and incomplete data then this would constitute a major advance. Furthermore, significant progress is needed in improving the computational efficiency of these tools if they are to be made applicable on a patient-specific basis.

In this work we advocate a Variational Bayesian (VB) perspective [34,35]. Such methods have risen into prominence for probabilistic inference tasks in the machine learning community [36–38] but have recently been employed also in the context of inverse problems [39,40]. They provide *approximate* inference results by solving an optimization problem over a family of appropriately selected probability densities with the objective of minimizing the Kullback–Leibler divergence [41] with the exact posterior. The success of such an approach hinges upon the selection of appropriate densities that have the capacity of providing good approximations while enabling efficient