Probabilistic Engineering Mechanics 46 (2016) 107–119

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

journal homepage: www.elsevier.com/locate/probengmech

Multi-model polynomial chaos surrogate dictionary for Bayesian inference in elasticity problems

Andres A. Contreras [a,*], Olivier P. Le Maître [a,b], Wilkins Aquino [a], Omar M. Knio [a,c]

[a] Duke University, Durham, NC, United States
[b] LIMSI-CNRS, Orsay, France
[c] King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

ARTICLE INFO

Article history:
Received 2 April 2016
Received in revised form
17 June 2016
Accepted 1 August 2016
Available online 19 September 2016

Keywords:
Bayesian inference
Model selection
Uncertainty quantification
Polynomial chaos
Elasticity imaging
Soft tissue

ABSTRACT

A method is presented for inferring the presence of an inclusion inside a domain; the proposed approach is suitable to be used in a diagnostic device with low computational power. Specifically, we use the Bayesian framework for the inference of stiff inclusions embedded in a soft matrix, mimicking tumors in soft tissues. We rely on a polynomial chaos (PC) surrogate to accelerate the inference process. The PC surrogate predicts the dependence of the displacements field with the random elastic moduli of the materials, and are computed by means of the stochastic Galerkin (SG) projection method. Moreover, the inclusion's geometry is assumed to be unknown, and this is addressed by using a dictionary consisting of several geometrical models with different configurations. A model selection approach based on the evidence provided by the data (Bayes factors) is used to discriminate among the different geometrical models and select the most suitable one. The idea of using a dictionary of pre-computed geometrical models helps to maintain the computational cost of the inference process very low, as most of the computational burden is carried out off-line for the resolution of the SG problems. Numerical tests are used to validate the methodology, assess its performance, and analyze the robustness to model errors.

1. Introduction

The nondestructive characterization of the parameters describing a physical system is a task of great importance and interest in various disciplines within science and engineering. Examples of such tasks include seismic imaging [1,2], health monitoring of infrastructure [3–5], and more recently elasticity imaging [6,7]. Elasticity imaging is a very promising branch of medical diagnosis which applies inverse problem techniques to compute the elasticity modulus given a set of measurements of a displacement or velocity field that is the result of some excitation force [8]. The idea is inspired by the palpation technique used by doctors to determine the presence of abnormal tissue through the sense of touch [9,10]. Palpation, however, is limited in detecting anomalies that lie deep in the body or which are too small [11]; moreover, it tends to be qualitative as opposed to quantitative. Elasticity imaging takes palpation to the next level by extending its range and effectiveness, all in a more quantitative manner. The general goal of this work is to use a collection of models within a Bayesian framework to estimate the contrast between the elastic properties of different regions in a given domain.

The elasticity imaging technique encompasses three basic steps: first, the body is deformed through an applied external load, then the deformation field is measured (e.g. using ultrasound techniques), and finally the elastic properties are estimated by solving an inverse problem. To approach this problem, Oberai et al. [8] assume that the displacements are governed by the equations of equilibrium of an incompressible, linear-elastic solid undergoing small, quasi-static deformation, and cast the problem as a non-linear optimization problem; the objective is to find a shear modulus field that minimizes the discrepancy between the measured and predicted displacement fields. Another optimization approach is based on minimizing the modified error in constitutive equation functional [12], which measures the discrepancy in the constitutive equations that connect kinematically admissible strains and dynamically admissible stresses in addition to measuring the discrepancy between the measured and predicted displacement fields. Other approaches include direct inversion methods [13–15], but these methods, although computationally less expensive, tend to be more sensitive to noise in measurement data. All these approaches are deterministic, and consequently result in a single estimate of the elastic modulus, which does not accommodate for the quantification of uncertainty.

Important insights emerge by approaching inverse problems

* Corresponding author.
E-mail address: andres.contreras@duke.edu (A.A. Contreras).

http://dx.doi.org/10.1016/j.probengmech.2016.08.004
0266-8920/

using a probabilistic framework. Some of the methods introduced to deal with this problem include the extended maximum likelihood method [16], the spectral stochastic method [17,18], the sparse grid collocation approach [19,20], stochastic reduced order models [21], and the Bayesian inference approach [22,23]. In the Bayesian formalism, one obtains additional insight by computing a probability distribution that summarizes all available information about the elastic moduli (e.g. we can estimate moments, marginal distributions, and quantiles), as opposed to the single value obtained in the deterministic setting. Specifically, in the context of elasticity imaging, Koutsourelakis [24] use a Bayesian framework to obtain probabilistic estimates of the material properties that account for various possible sources of uncertainty; this work deals with simplified geometries and large contrast ratio of the elastic properties. Another interesting approach is proposed by Iglesias et al. [25], where the Bayesian framework is applied under an infinite dimensional setting; this work, however, is limited to deterministic (known) elastic properties and it requires an appropriate prior model.

For complex forward models, extracting information from the posterior distribution can be very computationally expensive. Several techniques are applied to address this computational challenge, such as the use of a two-stage MCMC to increase the acceptance rate of the algorithm by using an inexpensive approximation of the posterior distribution [26–28]; the use of proper orthogonal decomposition (POD) to construct a reduced-order model for the direct simulations [29,30]; the use of adaptive hierarchical sparse grid collocation (ASGC) to obtain an approximate stochastic solution to the forward problem using piecewise linear interpolation [31]; and the use of polynomial chaos (PC) to approximate the solution of the stochastic forward model either through collocation [32,33] or through the stochastic Galerkin method [34]. A related application of PC representations in the context of inverse acoustic scattering problems is found in [35], where PC expansions are integrated with optimization methods for the probabilistic characterization of hidden obstacles and inclusions in acoustic media.

The objective of this work is to develop a method that can be used in a diagnostic device with a low computational power to quickly assess the presence of an inclusion in a given domain. To achieve this, the proposed approach breaks the process in two steps: (1) an offline or pre-processing step where surrogate models are constructed for different geometrical models and (2) an online step where a model selection and inference are performed on the basis of observations to assess the presence of an inclusion. This is advantageous, since the main computational cost is carried by the construction of the surrogate models, which is something that can be done offline with a dedicated computer. Thus, once the surrogate models have been constructed, the computational cost of the model selection and inference problem is relatively low and can be effectively handled by the diagnostic device with limited computational power. In more details, we extend the Bayesian approach proposed by Marzouk et al. [34] to the case of multiple geometrical models as follows. First, a dictionary of inclusion geometries is considered and for each of these geometries a suitable polynomial chaos expansion of the displacement field is computed, in terms of the unknown parameters (in our case Young's modulus and Poisson's ratio in soft matrix and inclusion) by means of the stochastic Galerkin (SG) method [36,37]. The SG allows for a fine control approximation error. When observations are made available, the PC surrogates can be used to derive corresponding approximations of the posterior distribution for the elastic properties given a geometry. Then, these posteriors can be compared by computing the evidences or Bayes factors of the geometrical models, in order to rank them and select the best one (or few best ones). The posterior distribution (s) of the elastic properties for the best model (or best ones) can then be used to reach a decision confirming or refuting the presence of an inclusion, analyzing for instance the ratios between the mean properties in the inclusion and soft matrix domains.

The outline of the paper is as follows. In Section 2, we introduce the mechanical model of the elasticity problem and derive the polynomial chaos expansions of the displacement field for a given geometry. In Section 3, we describe the use of the Bayesian framework to solve both inverse problem and the model selection problem. In Section 4 we present some numerical results showing the behavior of the approach when the exact geometry of the model is known. Then, in Section 5, we look at the case when the exact model geometry is unknown and construct a dictionary of surrogate models and rank them based on the evidence provided by the data; also we test the robustness of the approach with respect to errors in the mechanical model. Finally, in Section 6 we provide concluding remarks.

2. Physical model and polynomial chaos expansion

2.1. Physical model

2.1.1. Continuous problem

The strong form of the equilibrium equations of a linear-elastic solid undergoing static deformation due to boundary loads and displacements can be expressed as:

$$\nabla\cdot\boldsymbol{\sigma} = 0 \quad \text{in } \Omega \tag{1}$$

with boundary conditions:

$$\boldsymbol{\sigma}\cdot\boldsymbol{n} = \boldsymbol{\tau} \ \text{ on } \Gamma_\tau, \quad \boldsymbol{u} = \boldsymbol{u_0} \ \text{ on } \Gamma_u, \tag{2}$$

where $\boldsymbol{\sigma} = \mathbb{C}: \boldsymbol{\epsilon} \equiv \mathbb{C}_{ijkl}\,\epsilon_{kl}$ is the stress tensor; $\boldsymbol{\epsilon}(\boldsymbol{u}) = (\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^T)/2$ is the linearized strain tensor; $\boldsymbol{u}$ is the displacement field; $\boldsymbol{n}$ the unit normal to the boundary; $\boldsymbol{\tau}$ is the traction vector; Ω is the spatial domain; Γ_τ and Γ_u form a partition of the boundary Γ of Ω; $\boldsymbol{u_0}$ is the essential boundary condition; and $\mathbb{C}$ is the fourth-order constitutive tensor of linear elasticity. Under the assumption of an isotropic medium, the constitutive tensor has only two independent elastic constants and can be written as:

$$\mathbb{C}_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) \tag{3}$$

where δ_{ij} is the Kronecker delta and λ and μ are the Lamé constants [38]. This decomposition of $\mathbb{C}$ is very advantageous for the computation of the PC coefficients described in Section 2.2.1.

The forward problem consists in finding the displacement field $\boldsymbol{u}$ that satisfies (1) for a given constitutive tensor $\mathbb{C}$ (i.e. known material properties). The weak formulation of the forward problem is obtained after defining the space of trial functions, $\mathcal{S} = \left\{\boldsymbol{u} | u_i \in H^1(\Omega),\ \boldsymbol{u} = \boldsymbol{u_0} \text{ on } \Gamma_u\right\}$, and the space of test functions, $\mathcal{V} = \left\{\boldsymbol{v} | v_i \in H^1(\Omega),\ \boldsymbol{v} = \boldsymbol{0} \text{ on } \Gamma_u\right\}$. Multiplying (1) by an arbitrary $\boldsymbol{v} \in \mathcal{V}$, integrating over the spatial domain, using the divergence theorem, and the symmetry of $\mathbb{C}$ we get:

$$a\left(\boldsymbol{u}, \boldsymbol{v}\right) = \left(\boldsymbol{\tau}, \boldsymbol{v}\right), \quad \forall\ \ \boldsymbol{v} \in \mathcal{V}, \tag{4}$$

where

$$a\left(\boldsymbol{u}, \boldsymbol{v}\right) \equiv \int_\Omega \boldsymbol{\epsilon}(\boldsymbol{v}) : \mathbb{C} : \boldsymbol{\epsilon}(\boldsymbol{u})\,d\Omega, \quad \left(\boldsymbol{\tau}, \boldsymbol{v}\right) \equiv \int_{\Gamma_\tau} \boldsymbol{\tau}\cdot\boldsymbol{v}\,d\Gamma_\tau. \tag{5}$$

The function $\boldsymbol{u} \in \mathcal{S}$ that satisfies (4) is the equivalent weak solution of (1).

2.1.2. Finite element formulation

Using standard Voigt notation [39], the displacement fields, test