Proceedings of the 24th UK Conference of the
Association for Computational Mechanics in Engineering
31 March - 01 April 2016, Cardiff University, Cardiff

BAYESIAN STATISTICAL INFERENCE ON THE MATERIAL PARAMETERS OF A HYPERELASTIC BODY

***Jack S. Hale[1], Patrick E. Farrell[2,4] and Stéphane P.A. Bordas[1,3,5].**

[1]Research Unit in Engineering Science, Faculté des Sciences, de la Technologie et de la Communication, Université du Luxembourg, Campus Kirchberg, 6, rue Coudenhove-Kalergi, L-1359, Luxembourg.
[2]Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK.
[3]Cardiff University Institute of Mechanics & Advanced Materials The Queen's Building, The Parade, Cardiff, CF24 4AG, UK.
[4]Center for Biomedical Computing, Simula Research Laboratory, Oslo, Norway.
[5]University of Western Australia Intelligent Systems for Medicine Laboratory 35 Stirling Highway, Crawley, WA 6009, Australia.

*jack.hale@uni.lu

ABSTRACT

We present a statistical method for recovering the material parameters of a heterogeneous hyperelastic body. Under the Bayesian methodology for statistical inverse problems, the posterior distribution encodes the probability of the material parameters given the available displacement observations and can be calculated by combining prior knowledge with a finite element model of the likelihood.

In this study we concentrate on a case study where the observations of the body are limited to the displacements on the surface of the domain. In this type of problem the Bayesian framework (in comparison with a classical PDE-constrained optimisation framework) can give not only a point estimate of the parameters but also quantify uncertainty on the parameter space induced by the limited observations and noisy measuring devices.

There are significant computational and mathematical challenges when solving a Bayesian inference problem in the case that the parameter is a field (i.e. exists infinite-dimensional Banach space) and evaluating the likelihood involves the solution of a large-scale system of non-linear PDEs. To overcome these problems we use dolfin-adjoint to automatically derive adjoint and higher-order adjoint systems for efficient evaluation of gradients and Hessians, develop scalable maximum aposteriori estimates, and use efficient low-rank update methods to approximate posterior covariance matrices.

Key Words: *Bayesian inference, dolfin-adjoint, posterior, FEniCS, hyperelasticity.*

1. Theoretical framework

Following the infinite-dimensional presentation of Stuart [1], we introduce the parameter-to-observable map $\mathcal{G} : \mathcal{M} \to \mathcal{Y}$ as a deterministic function that maps the *parameters* $m \in \mathcal{M}$ to the observables $y \in \mathcal{Y}$, where $\mathcal{M}, \mathcal{Y}$ are Banach spaces:

$$y = \mathcal{G}(m). \tag{1}$$

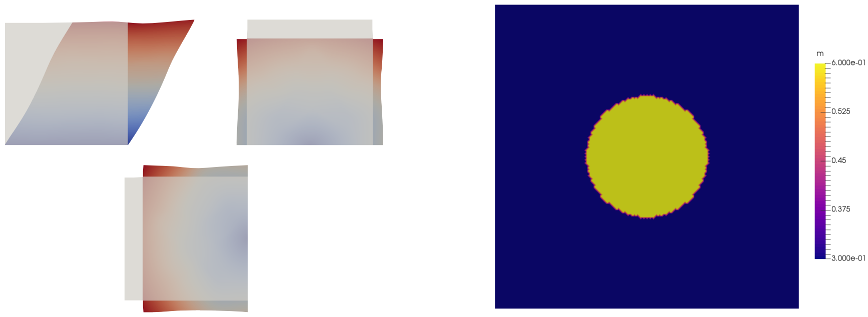
In our case, every evaluation of this map $\mathcal{G}$ will involve solving a PDE governing the behaviour of a geometrically non-linear hyperelastic solid.

The parameter m we wish to infer is the (spatially-varying) shear-like parameter of the following Neo-Hookean energy potential W:

$$W(X, \mathrm{I}_C, \mathrm{III}_C) := \frac{m}{2}(\mathrm{I}_C - 2) - m \ln J + \frac{\lambda}{2}(\ln J)^2 \tag{2}$$

where $\mathrm{I}_C = \mathrm{tr}(C)$ and $\mathrm{III}_C = \det(C)$ are the first and third invariants of the right Cauchy-Green tensor $C = F^T F$, F is the deformation gradient and $J = \det F = (\mathrm{III}_C)^{1/2}$. The displacement field $u^* \in \mathcal{V}$ at equilibrium can then be found through a standard minimisation problem of the following form:

$$u^* = \arg\min_{u \in \mathcal{V}} \left\{ \int_{\Omega_0} W(X, \mathrm{I}_C, \mathrm{III}_C)\, \mathrm{d}x_0 - \int_{\partial_N \Omega_0} t \cdot u \, \mathrm{d}s \right\}, \tag{3}$$



(a) Observed displacement fields y_{obs}.

(b) Shear-related parameter m.

Figure 1: (a) Three virtual experiments (shear and two friction-free compression tests) on a hyperelastic body (b) with a parameter field m given by a circular inclusion in a softer matrix. Grey shape is the square undeformed configuration of the body. Colour shows magnitude of displacement at equilibrium. Note the warped deformation on the boundary - if we only have these limited observations, how much can we tell about the parameters in the centre of the domain? This is the question we attempt to answer with this work.

where $\mathrm{d}s$ and $\mathrm{d}x_0$ are measures on the undeformed configuration domain Ω_0 and its boundary $\partial_N \Omega_0$, respectively and t are the external applied tractions on $\partial_N \Omega_0$. Virtual experiments and the exact parameter field used to create them are shown in Figure 1.

We state the Bayesian solution to the infinite-dimensional inference problem as follows: We describe the observer's prior beliefs about the parameter m through the prior probability measure μ_0. Given the likelihood model π_{like}, which gives the probability that we will observe y given the parameters m, the goal of the inference problem is to find the posterior probability measure μ^{post} (Stuart [1], Theorem 6.2) as:

$$\frac{\mathrm{d}\mu^{\text{post}}}{\mathrm{d}\mu_{\text{prior}}} \propto \pi_{\text{like}} \left(y^{\text{obs}} - \mathcal{G}(u) \right) \tag{4}$$

where the Radon-Nikodym derivative (Stuart [1], Theorem 6.29) $\frac{\mathrm{d}\mu^{\text{post}}}{\mathrm{d}\mu_{\text{prior}}}$ is the derivative of the posterior probability measure μ^{post} with respect to the prior probability measure μ_0.

We further assume that our prior knowledge can be expressed by a Gaussian distribution with mean m_0 and covariance operator C_0, or more compactly $\pi_{\text{prior}} \sim \mathcal{N}(m_0, C_0)$. Again, following the well-posedness result of Stuart [1], we solve a Helmholtz-like PDE problem to generate actions of our prior covariance on a vector.

Furthermore we assume that our noise model is white-noise Gaussian with mean zero and covariance operator Γ_{noise}. We can then re-write the posterior more concretely as:

$$\pi_{\text{post}}(m|y^{\text{obs}}) \propto \exp\left(-\frac{1}{2} ||y_{\text{obs}} - \mathcal{G}(m)||^2_{\Gamma^{-1}_{\text{noise}}} - \frac{1}{2} ||m - m_0||^2_{C_0^{-1}} \right). \tag{5}$$

Taking the logarithm of the above equation results in the following weighted least-squares functional:

$$\hat{J}(m) := -\ln \pi_{\text{post}}(m|y_{\text{obs}}) = \frac{1}{2} ||y_{\text{obs}} - \mathcal{G}(m)||^2_{\Gamma^{-1}_{\text{noise}}} + \frac{1}{2} ||m - m_0||^2_{C_0^{-1}}, \tag{6}$$

We characterise the Bayesian posterior via extraction of two pieces of information. The maximum a posteriori point m_{MAP} is characterised by the maximum of the above functional:

$$m_{\text{MAP}} := \arg\max_{m \in \mathcal{M}} \pi_{\text{post}} = \arg\min_{m \in \mathcal{M}} (-\ln \pi_{\text{post}}) \tag{7}$$

This is a classical point estimate similar to those found in the PDE-constrained optimisation literature, but in that case the norms used are usually somewhat arbitrary. By formulating out problem in the Bayesian seting, our problem has rigorous statistical meaning [1].

An effective simplification in the case that the parameter-to-observable map $\mathcal{G}$ is a linear operator $A : \mathcal{M} \to \mathcal{Y}$ we can write the following semi-analytical expressions for the MAP point and the posterior covariance:

$$m_{\text{MAP}} = C(A^* \Gamma_{\text{noise}}^{-1} y^{\text{obs}} + C_0^{-1} m_0), \tag{8}$$

$$C = (A^* \Gamma_{\text{noise}}^{-1} A + C_0^{-1})^{-1}. \tag{9}$$

where * denotes the usual adjoint operation. After some some algebraic manipulation we can show that the posterior distribution is infact Gaussian and can be written:

$$\pi_{\text{post}} \sim \mathcal{N}(m_{\text{MAP}}, \mathcal{H}^{-1}). \tag{10}$$

where the Hessian operator $\mathcal{H} = C^{-1}$ is the second Fréchét derivative of the weighted least-squares functional $\hat{J}(m)$ defined above. It is worth pointing out at this stage that our hyperelastic forward problem does not lead to a linear parameter-to-observable map $\mathcal{G}$. Thus, the above result does *not* for our problem hold because the map $\mathcal{G}$ induces non-Gaussianity into the posterior.

However, a useful approximation, as long as the posterior is not too non-Gaussian, is to evaluate the Hessian of the functional around the MAP point and use it as an approximation to the true second moment of the distribution about the MAP point.

2. Solution approach

We implement our solver within the dolfin-adjoint package [3], which is based on the finite element solver DOLFIN from the FEniCS Project [4]. We express the forward model in the high-level Unified Form Language (UFL) before automatically deriving finite element cell tensors for the adjoint and higher-order adjoint equations using symbolic manipulations.

We first solve the problem of finding the MAP point using a mesh-independent bound-constrained quasi-Newton optimisation algorithm that uses the gradients from dolfin-adjoint to efficiently drive the optimisation.

Then once we have found the MAP point, we evaluate likelihood Hessian actions from dolfin-adjoint within a Krylov-Schur type eigenvalue solver to extract information about the directions in parameter space that are most constrained by the observations, with respect to the prior. We use an efficient low-rank update procedure from Spantini et al. [2] to construct an approximation to the posterior covariance.

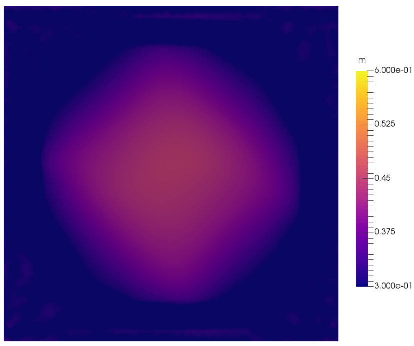
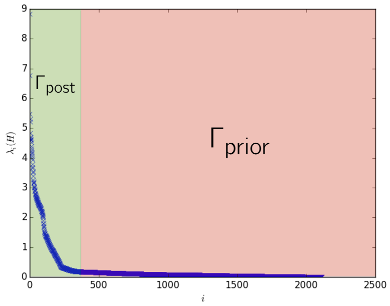
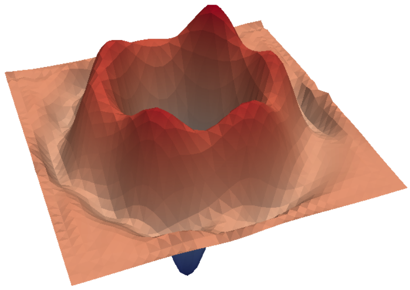


Figure 2: Maximum a posteriori (MAP) point of the Bayesian inference problem. We can detect the stiffness in the centre of the body, but the maximum value and precise radius are not recovered particularly well. This uncertainty is quantified by the information in the posterior covariance fig. 3.



(a) Spectrum of the inverse of the posterior covariance.



(b) Trailing eigenvector.

(c) Leading eigenvector.

Figure 3: (a) The leading eigenvalues $\lambda_i \gg 1$ (green) correspond to the directions in parameter space *most* informed by the observations via the likelihood. Conversely, the trailing eigenvalues $\lambda_i \ll 1$ (red) correspond to the directions in parameter space *least* informed, and thus correspond to the information originally contained in our prior. Plots of (b) trailing and (c) leading eigenvectors. The least constrained direction points towards the parameters in the centre. The most constrained direction points towards the parameters at the corner where in effect we have two independent 'sensors' touching one piece of material.

3. Results

We show our results along with full descriptions in Figures 2 and 3.

References

[1] A.M. Stuart. Inverse problems: a Bayesian perspective. Acta Numerica 19 (2010).

[2] Alessio Spantini, Antti Solonen, Tiangang Cui, James Martin, Luis Tenorio, Youssef Marzouk. Optimal low-rank approximations of Bayesian linear inverse problems. arXiv:1407.3463v2.

[3] Patrick E. Farrell, David A. Ham, Simon W. Funke and Marie E. Rognes. Automated derivation of the adjoint of high-level transient finite element programs, SIAM Journal on Scientific Computing 35.4, pp. C369-C393. doi:10.1137/120873558i (2013).

[4] M. S. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes and G. N. Wells, The FEniCS Project Version 1.5, Archive of Numerical Software, 3(100), doi:10.11588/ans.2015.100.20553i (2016).