Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-18 (2016) 428–433

Parameter Estimation for Predictive Simulation of Oscillatory Systems with Model Discrepancy ⋆

Jerry A. McMahan, Jr. * **Ralph C. Smith** **

* *University of Dayton Research Institute, Dayton, OH 45469 USA (e-mail: jmcmahan1@udayton.edu).*
** *North Carolina State University, Raleigh, NC 27695 USA (e-mail: rsmith@ncsu.edu)*

Abstract: A Bayesian framework provides a methodology in which inferences from measurement data can be used to bound the uncertainties in the predictive simulation of a physical system. The accuracy of these bounds relies on the satisfaction of statistical assumptions on the measurement error. Discrepancies between the model and the true physics can invalidate these assumptions. We examine the effect of such model discrepancies in the context of an oscillating cantilever beam. First we illustrate the influence of discrepancies in a simplified model of purely periodic signals and then we observe how discrepancies affect the accuracy of prediction uncertainty bounds using Bayesian parameter inference on a Euler-Bernoulli beam model. Our study shows small changes in the inference setup can result in significant differences in prediction accuracy and calls attention to important considerations for the practical application of Bayesian parameter estimation.

Keywords: parameter estimation, Bayesian inversion, modeling errors, oscillation, prediction intervals

1. INTRODUCTION

Parameter estimation is an important step in developing practical models for control applications. Typically the relationship between parameters and model output is nonlinear, making estimation a non-trivial task. In classical approaches, parameter estimation is cast as a nonlinear programming problem where the difference in model output and measurement data is mapped into a non-negative cost functional. One of the most common forms of cost functional is the sum-of-squares residual, which computes the sum of the squared difference between model output and measurement data. Various standard optimization algorithms have been designed to find the set of parameters that minimizes this cost. The resulting optimal parameters are taken as the true values for the system model, which can then be used for control applications.

A shortcoming of this approach is that no mathematical model is ever perfect, inevitably leading to differences between the model and reality. One way to address this is to employ the Bayesian framework, as described in, e.g., Kennedy and O'Hagan (2001). The general idea is that if one knows the statistics of the model error, then the uncertainty in the parameter estimates can be statistically characterized as well. This uncertainty can be used to quantify the uncertainty in the model output, providing a better idea of the likely output of the true system and improving the usefulness of the model. For instance, Crews et al. (2013) and McMahan and Smith (2014) show how Bayesian uncertainty information can be used to design robust nonlinear controllers. More information on Bayesian parameter estimation in the context of uncertainty quantification can be found in Smith (2014).

As promising as this approach is, a crucial problem is the difficulty of accurately characterizing the statistics of the modeling error. A widely used assumption is that the measurement error is independent, identically distributed (i.i.d) Gaussian noise. The validity of this assumption realist on the system of interest being modeled well enough that the only model errors are completely random and independent. Unfortunately, model errors are frequently correlated due to unmodeled physics, making them effectively as unpredictable as random errors but not independent. We refer to these correlated errors as model discrepancies.

There are some approaches for estimating such discrepancies from data. One is described in Raol et al. (2004) which applies the variational method of Detchmendy and Sridhar (1966) to estimate unknown model dynamics from data (effectively a Kalman filter). It is not clear, however, how this method can be used for prediction beyond the fitting interval, where no data is available from which to infer the nature of the discrepancy. Other examples are found in Heino and Somersalo (2004); Heino et al. (2005), but these are concerned with problems such as medical

⋆ This research was supported in part by the Air Force Office of Scientific Research grant AFOSR FA9550-11-1-0152. It was also supported by the Consortium for Advanced Simulation of Light Water Reactors (http://www.casl.gov), an Energy Innovation Hub, (http://www.energy.gov/hubs) for Modeling and Simulation of Nuclear Reactors under U.S. Department of Energy Contract No. DE-AC05-00OR22725.

10.1016/j.ifacol.2016.10.203

imaging where parameter inference is the end-goal rather than a step in producing dynamic simulations to predict model behavior at some unknown time. Model discrepancy in the Bayesian context has been recently discussed in Brynjarsdóttir and O'Hagan (2014), which develops some simulated examples of Bayesian parameter estimation in a simple algebraic model with model discrepancy using Gaussian processes to model the discrepancy term.

In this paper, we focus on the effects of model discrepancy on Bayesian parameter estimation for damped oscillatory systems with the goal of predictive simulation. We first outline the Euler-Bernoulli model and its numerical approximation and then give a brief overview of the Bayesian framework for parameter estimation. To avoid the complexities of analyzing the beam model while still considering some of its key features, we examine in some detail some of the characteristics of parameter estimation and prediction for a simple model of periodic signals with model discrepancy terms. Finally, we return to calibration of the beam model and study an example of estimating its parameters from experimental data obtained from a vibrating cantilever beam. The example illustrates significant differences in prediction accuracy occurring due to changes in the time interval used for calibration.

2. MODEL DEVELOPMENT

We consider the problem of identifying the physical parameters of a model of an aluminum beam driven by a piezoelectric patch. In this section, we briefly outline the mathematical model for the dynamics of the beam whose parameters we will later estimate. Further details on the model and its numerical approximation can be found in Chapters 7 and 8 of Smith (2005) and the references therein.

2.1 Euler-Bernoulli Beam with Kelvin-Voigt Damping

The strong formulation of the Euler-Bernoulli beam with no applied force is

$$\rho\frac{\partial^2 w}{\partial^2 t} + \gamma\frac{\partial w}{\partial t} + \frac{\partial^2 M}{\partial x^2} = 0$$

where ρ is the density of the beam, γ is the viscous damping due to friction from air, and M is the bending moment of the beam. For brevity the boundary conditions corresponding to the cantilever mounting of the beam are omitted. The bending moment is made up of the three terms

$$\begin{aligned} M(x,t) &= M_e + M_d + M_p \\ &= EI\frac{\partial^2 w}{\partial x^2} + cI\frac{\partial^3 w}{\partial x^2 \partial t} - k_p V(t)\mathbb{1}_p(x), \end{aligned}$$

where M_e, M_d, M_p are the elastic, damping, and patch moments, respectively. Here $\mathbb{1}_p(x)$ is the characteristic function for the patch region.

The weak formulation is

$$\rho\int_0^L \frac{\partial^2 w}{\partial^2 t}\phi\, dx + \gamma\int_0^L \frac{\partial w}{\partial t}\phi\, dx + \int_0^L \frac{\partial^2 M}{\partial x^2}\phi\, dx = 0.$$

The space of test functions is defined to be $W = H_0^2(0,L) = \{\phi \in H^2(0,L) \mid \phi(0) = 0,\ \phi'(0) = 0\}$. This definition ensures that ϕ satisfies the boundary conditions for the cantilever beam at the fixed end and that ϕ is twice-differentiable with respect to x. Using integration by parts twice on the moment term yields

$$\rho\int_0^L \frac{\partial^2 w}{\partial^2 t}\phi\, dx + \gamma\int_0^L \frac{\partial w}{\partial t}\phi\, dx + \int_0^L M\frac{\partial^2 \phi}{\partial x^2}\, dx = 0.$$

Expanding the moment term gives

$$\begin{aligned} \rho\int_0^L \frac{\partial^2 w}{\partial^2 t}\phi\, dx + \gamma\int_0^L \frac{\partial w}{\partial t}\phi\, dx + EI\int_0^L \frac{\partial^2 w}{\partial x^2}\frac{\partial^2 \phi}{\partial x^2}\, dx \\ + cI\int_0^L \frac{\partial^3 w}{\partial x^2 \partial t}\frac{\partial^2 \phi}{\partial x^2}\, dx = k_p V(t)\int_{p_l}^{p_r} \frac{\partial^2 \phi}{\partial x^2}\, dx. \end{aligned}$$

The integration limits for the term on the right hand side come from $[p_l, p_r]$, the interval in space in which the patch lies.

2.2 Numerical Approximation

A semi-discrete ODE approximation of the PDE model suitable for computational implementation is derived from the weak approximation by projecting the infinite-dimensional function space W onto a finite-dimensional space W^N. Here $N+1$ is the dimension of W^N, the vector space of all linear combinations of a suitably chosen ϕ_j for $j = 1, \ldots, N+1$. In this paper, ϕ_j are standard cubic B-splines modified to ensure satisfaction of the essential boundary conditions (see Smith (2005)). The transverse displacement is approximated as

$$w(x,t) \approx w^N(x,t) = \sum_{i=1}^{N+1} w_i(t)\phi_j(x).$$

By substituting w^N for w in the weak formulation we obtain the vector-valued system

$$M\ddot{w}(t) + C\dot{w}(t) + Kw(t) = f(t),$$

where $w(t) = [w_1(t) \cdots w_{N+1}(t)]$ is the vector of coefficients and the dot notation indicates the derivative with respect to t. The M, C, and K terms are respectively the $(N+1)\times(N+1)$ mass, damping, and stiffness matrices with entries

$$\begin{aligned} [M]_{ij} &= \int_0^L \rho\phi_i\phi_j\, dx, \\ [C]_{ij} &= \int_0^L \left(\gamma\phi_i\phi_j + cI\phi_i''\phi_j''\right) dx, \\ [K]_{ij} &= \int_0^L EI\phi_i''\phi_j''\, dx. \end{aligned}$$

The prime notation indicates a derivative with respect to x. The f term is the $(N+1)\times 1$ vector discretization of the input given by

$$[f(t)]_i = k_p V(t)\int_{pl}^{pr} \phi_i''\, dx.$$

We rearrange to obtain the first-order system

$$\begin{aligned} \dot{z}(t) &= Az(t) + F(t) \\ y(t) &= D(x_{obs})z(t) \end{aligned}$$

where $z(t) = [w(t)\ \dot{w}(t)]^T$ and

$$A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix}, \quad F = \begin{bmatrix} 0 \\ M^{-1}f(t) \end{bmatrix},$$