Mechanical Systems and Signal Processing 32 (2012) 153–169

Contents lists available at SciVerse ScienceDirect

Mechanical Systems and Signal Processing

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

Parameter estimation and model selection for a class of hysteretic systems using Bayesian inference

K. Worden [a,*], J.J. Hensman [b]

[a] Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, United Kingdom
[b] The Sheffield Institute for Translational Neuroscience, The University of Sheffield, 385A Glossop Road, Sheffield, S10 2HQ, United Kingdom

ARTICLE INFO

Article history:
Received 1 June 2011
Received in revised form
14 March 2012
Accepted 27 March 2012
Available online 13 April 2012

Keywords:
Nonlinear system identification
Bayesian inference
Markov Chain Monte Carlo (MCMC)
Deviance Information Criterion (DIC)
Duffing oscillator
Bouc-Wen hysteresis

ABSTRACT

The aim of this paper is to provide an overview of the possible advantages of adopting a Bayesian approach to nonlinear system identification in structural dynamics. In contrast to identification schemes which estimate maximum likelihood values (or other point estimates) for parameters, the Bayesian scheme discussed here provides information about the complete probability density functions of parameter estimates without adopting restrictive assumptions about their nature. Among other advantages of the Bayesian viewpoint are the abilities to make informed decisions about model selection and also to effectively make predictions over entire classes of models, with each individual model weighted according to its ability to explain the observed data.

The approach is illustrated using data from simulated systems, first a Duffing oscillator and then a new application to hysteretic system of the Bouc–Wen type. The modelling and identification of the latter type of system has long presented problems due to the fact that commonly used model structures like the Bouc–Wen model are nonlinear in the parameters, or have unmeasured states, etc. These issues have been dealt with in the past by adopting an optimisation-based approach to the problem; in particular, the differential evolution algorithm has proved very effective. An objective of the current paper is to illustrate how the Bayesian approach provides the same information and more as the optimisation approach; it yields parameter estimates and their associated confidence intervals, but can also provide confidence bounds on model predictions and evidence measures which can be used to select the most appropriate model from a candidate set. A new model selection criterion in this context – the Deviance Information Criterion (DIC) – is presented here.

1. Introduction

The subject of engineering dynamics has evolved for much of its history under the assumption that deterministic models are largely appropriate for system modelling and prediction. However, (fairly) recent developments in the subject have begun to accommodate the fact that uncertainty plays a more major role in structural system analysis than hitherto recognised. For example, the modelling of biomechanical systems faces the problem that the mechanical properties of tissue vary considerably from individual to individual and even within a single individual. Because of the issue of uncertainty, probabilistic reasoning is becoming much more common in the consideration of dynamic problems.

* Corresponding author. Tel.: +44 114 2227758.
E-mail address: k.worden@sheffield.ac.uk (K. Worden).

0888-3270/$ - see front matter
http://dx.doi.org/10.1016/j.ymssp.2012.03.019

In some areas, uncertainty has been (at least partially) accommodated in theory and practice for a long time; the field of *system identification* is a case in point. If one wishes to identify a parametric model of a system from measured data, one has to allow for the fact that noise may be present in any measurements, in order that the identified or estimated parameters for the model are completely meaningful. In general, the inclusion of *noise models* in linear [1,2] and nonlinear [3,4] approaches has often been considered sufficient. The main objective of noise models in 'classical' system identification would appear to be to ensure that there is no systematic bias in estimated parameters. It is probably true to say though that the probabilistic reasoning that underlies much of linear and nonlinear system identification (SI) in structural dynamics is often hidden. For example, although it is well-known that the least-squares estimators commonly used for SI are actually maximum-likelihood estimators under certain assumptions for the estimation residuals, the SI practitioner will generally implement the algorithms as an exercise in linear algebra and will usually treat the resulting set of crisp parameters as constituting 'the model'. Even if a covariance matrix is extracted, the user will usually use this only to provide confidence intervals or 'error bars' on the parameters; predictions will still be made using the crisp parameters produced by the estimation process. While such approaches are undeniably powerful and have allowed the identification of systems in many demanding engineering problems, they do not fully accommodate the fact that a given set of measured data may be consistent with a number of different parametric models. It is now becoming clear – largely as a result of the pioneering work of James Beck and colleagues and more recently from guidance from the machine learning community – that a more robust approach to parameter estimation, and also model selection, can be formulated on the basis of Bayesian principles for probability and statistics. Among the potential advantages offered by a Bayesian formulation are: the estimation procedure will return parameter distributions rather than parameters; predictions can be made by integrating over all parameters consistent with the data, weighted by their probabilities; evidence for a given model structure can be computed, leading to a principled means of model selection.

For the purpose of further discussion, it is useful to divide predictive models into two classes: white and black-box models. A *white-box* model will be taken here to mean one in which the equations of motion have been derived completely from the underlying physics of the problem of interest and in which the model parameters have direct physical meanings. Finite element models constitute one sub-class of such models. In contrast, a *black-box* model is usually formed by adopting a parametrised class of models with some universal approximation property and learning the parameters from measured data; in such a model, like a neural network, the parameters will not generally carry any physical meaning. Note that system identification, or learning from data, is essential to a black-box approach; in the case of the white-box model, one may also learn the parameters from data, or the parameters as well as the model structure may be determined by physical laws. Recent developments in system identification and machine learning theory have provided Bayesian approaches for the estimation of parameters in both white and black-box models.

In terms of the identification of black-box models, Bayesian methods and structures have recently begun to occupy a central position within the work of the machine learning community [5,6]. Bayesian methods for training multi-layer perceptron (MLP) neural networks are a good example of this trend [7]. The Gaussian process model is also achieving wide popularity [8]. In the context of white-box models, and in particular within the structural dynamics nonlinear system identification community, the use of Bayesian methods has not been so widespread; however, their pedigree is as long. One can find references to Bayesian methods in a monograph on parameter estimation from 1974 [17], and dating from the same year is perhaps the first paper on Bayesian methods for structural dynamic system identification [9]. By far the most systematic and extensive development of Bayesian system identification is the result of the work of James Beck and his various colleagues. Beck's early work on statistical system identification is summarised in [10]. The transition to the Bayesian framework is documented in [11]. This paper discusses the idea of marginalising over a group of models with different parameter estimates. It is argued that when a finite set of optimally parametrised models dominate, predictions can be effectively computed from a weighted sum of the optimal models with the density functions of individual models derived from a Laplace approximation. Least-squares or maximum likelihood parameter estimates are identified as a good first step towards a fully Bayesian approach. The use of the Laplace approximation – as it does in the Bayesian MLP [7] – removes the need to evaluate intractable high-dimensional integrals. In a later paper, Beck and Au [12] introduce a *Markov Chain Monte Carlo* (MCMC) method as a more general means of computing response quantities of interest represented by high-dimensional integrals. A neat approach to annealing the length scales of the MCMC method is introduced to speed up convergence of the algorithm. Bayesian methods of model selection are discussed in [13], with the evidence for a given model (based on observed data) represented by an 'Ockham factor'. The paper discusses the possibility of marginalising over different model *classes*. One of the examples discussed in the paper is a bilinear hysteretic oscillator representing a different class of hysteresis model than the one discussed later in the current paper. A very recent contribution by Muto and Beck [14] discusses identification and model selection for a further type of hysteretic system model—the Masing model. Bayesian methods for the system identification of differential equations have also been the subject of recent interest in the context of *systems biology* [15,16] and show considerable promise in the context of structural dynamics.

The structural dynamics community has recently considered new approaches to Bayesian system identification in the context of the state estimation problem and some powerful algorithms have been applied. The paper [18] discusses the use of the particle filter—a Bayesian generalisation of the Kalman filter, and an interesting recent reference is [19]. The Bayesian 'bootstrap' filter was used to identify hysteretic systems in [20]. Finally, a Bayesian approach based on an unscented Kalman filter has recently been applied to the Bouc–Wen hysteresis model in [21].

A very recent book [22] provides an interesting overview of the use of Bayesian methods in structural dynamics (largely in the context of civil engineering) and has a quite comprehensive set of references which the curious reader might wish to consult.