Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 199 (2017) 1056–1061

Procedia Engineering

www.elsevier.com/locate/procedia

X International Conference on Structural Dynamics, EURODYN 2017

Approximate Bayesian Computation by Subset Simulation for model selection in dynamical systems

Majid K. Vakilzadeh[a,*], James L. Beck[b], Thomas Abrahamsson[a]

[a] *Department of Applied Mechanics, Chalmers University of Technology, Gothenburg, Sweden*
[b] *Division of Engineering and Applied Science, California Institute of Technology, CA, USA*

Abstract

Approximate Bayesian Computation (ABC) methods are originally conceived to expand the horizon of Bayesian inference methods to the range of models for which only forward simulation is available. However, there are well-known limitations of the ABC approach to the Bayesian model selection problem, mainly due to lack of a sufficient summary statistics that work across models. In this paper, we show that formulating the standard ABC posterior distribution as the exact posterior PDF for a hierarchical state-space model class allows us to independently estimate the evidence for each alternative candidate model. We also show that the model evidence is a simple by-product of the ABC-SubSim algorithm. The validity of the proposed approach to ABC model selection is illustrated using simulated data from a three-story shear building with Masing hysteresis.

Keywords: Approximate Bayesian Computation; Subset Simulation; Bayesian model selection; Masing hysteretic models

1. Introduction

Due to exclusive foundation of Bayesian statistics on the probability logic axioms, it provides a rigorous framework for model updating and model selection. In this approach, a key idea is to construct a *stochastic model class* $\mathcal{M}$ consisting of the following fundamental probability distributions [1]: a set of parameterized input-output probability models $p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{u}, \mathcal{M})$ for predicting the system behavior of interest $\boldsymbol{y}$ for given input $\boldsymbol{u}$ and a *prior* probability density function (PDF) $p(\boldsymbol{\theta}|\mathcal{M})$ over the parameter space $\boldsymbol{\Theta} \in \mathbb{R}^{N_p}$ of $\mathcal{M}$ that reflects the relative degree of plausibility of each input-output model in the set. When data $\mathcal{D}$ consisting of the measured system input $\hat{\boldsymbol{u}}$ and output $\hat{\boldsymbol{z}}$ are available, the prior PDF $p(\boldsymbol{\theta}|\mathcal{M})$ can be updated through Bayes' Theorem to obtain the *posterior* PDF for the uncertain model parameters $\boldsymbol{\theta}$ as:

$$p(\boldsymbol{\theta}|\mathcal{D}, \mathcal{M}) \propto p(\hat{\boldsymbol{z}}|\boldsymbol{\theta}, \hat{\boldsymbol{u}}, \mathcal{M})p(\boldsymbol{\theta}|\mathcal{M}) \tag{1}$$

where $p(\hat{\boldsymbol{z}}|\boldsymbol{\theta}, \hat{\boldsymbol{u}}, \mathcal{M})$ denotes the *likelihood function* of $\boldsymbol{\theta}$ which gives the probability of getting the data based on the input-output probability model $p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{u}, \mathcal{M})$.

* Corresponding author. Tel.: +46-031-772-1510;
E-mail address: khorsand@chalmers.se

1877-7058 © 2017 The Authors. Published by Elsevier Ltd.
Peer-review under responsibility of the organizing committee of EURODYN 2017.
10.1016/j.proeng.2017.09.291

There are some model classes, e.g., hidden Markov models, for which the likelihood function is difficult or even impossible to compute, but one might still be interested to perform Bayesian parameter inference or model selection. ABC methods were originally conceived to circumvent the need for computation of the likelihood by simulating samples from the corresponding input-output probability model $p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{u}, \mathcal{M})$.

The basic idea behind ABC is to avoid evaluation of the likelihood function in the posterior PDF $p(\boldsymbol{\theta}|\mathcal{D}, \mathcal{M}) \propto p(\hat{\boldsymbol{z}}|\boldsymbol{\theta}, \hat{\boldsymbol{u}}, \mathcal{M})\, p(\boldsymbol{\theta}|\mathcal{M})$ over the parameter space $\boldsymbol{\theta}$ by using an augmented posterior PDF:

$$p(\boldsymbol{\theta}, \boldsymbol{y}|\mathcal{D}, \mathcal{M}) \propto P(\hat{\boldsymbol{z}}|\boldsymbol{y}, \boldsymbol{\theta})\, p(\boldsymbol{y}|\boldsymbol{\theta}, \hat{\boldsymbol{u}}, \mathcal{M})\, p(\boldsymbol{\theta}|\mathcal{M}) \tag{2}$$

over the joint space of the model parameters $\boldsymbol{\theta}$ and the model output $\boldsymbol{y}$ that is simulated using the distribution $p(\boldsymbol{y}|\boldsymbol{\theta}, \hat{\boldsymbol{u}}, \mathcal{M})$. The interesting point of this formulation is the degree of freedom brought by the choice of function $P(\hat{\boldsymbol{z}}|\boldsymbol{y}, \boldsymbol{\theta})$. The original ABC algorithm defines $P(\hat{\boldsymbol{z}}|\boldsymbol{y}, \boldsymbol{\theta}) = \delta_{\hat{\boldsymbol{z}}}(\boldsymbol{y})$, where $\delta_{\hat{\boldsymbol{z}}}(\boldsymbol{y})$ is equal to 1 when $\hat{\boldsymbol{z}} = \boldsymbol{y}$ and equal to 0 otherwise, to retrieve the target posterior distribution when $\boldsymbol{y}$ exactly matches $\hat{\boldsymbol{z}}$. However, the probability of generating exactly $\hat{\boldsymbol{z}} = \boldsymbol{y}$ is zero for continuous stochastic variables. Pitchard et al. [2] broadened the realm of the applications for which ABC algorithm can be used by replacing the point mass at the observed output data $\hat{\boldsymbol{z}}$ with an indicator function $\mathbb{I}_{S(\epsilon)}(\boldsymbol{y})$, where $\mathbb{I}_{S(\epsilon)}(\boldsymbol{y})$ gives 1 over the set $S(\epsilon) = \{\boldsymbol{y} : \rho(\boldsymbol{\eta}(\hat{\boldsymbol{z}}) - \boldsymbol{\eta}(\boldsymbol{y})) \leq \epsilon\}$ and 0 elsewhere, for some chosen metric ρ and low-dimensional summary statistic $\boldsymbol{\eta}$. The ABC algorithm based on this formulation thus gives samples from the true posterior distribution when the tolerance parameter ϵ is sufficiently small and the summary statistics $\boldsymbol{\eta}(.)$ are sufficient. These conditions pose some difficulties for computer implementation of this algorithm which renders it far from a routine use for parameter inference and model selection. Firstly, a sufficiently small tolerance parameter ϵ means that only predicted model outputs $\boldsymbol{y}$ lying in a small local neighborhood centered on the observed data vector $\hat{\boldsymbol{z}}$ are accepted. This leads to the problem of rare-event simulation. To circumvent this problem, Chiachio et al. [3] developed a new algorithm, called ABC-SubSim, by incorporating the Subset Simulation algorithm [4] for rare-event simulation into the ABC algorithm. Secondly, the lack of a reasonable vector of summary statistics that works across models hinders the use of an ABC algorithm for model selection [5].

In this study, we show that formulating a dynamical system as a general hierarchical state-space model enables us to solve the inherent difficulty of the ABC technique to model selection. Using this formulation, one can independently estimate the model evidence for each model class. We also show that the model evidence can be estimated as a simple by-product of the recently proposed multi-level MCMC algorithm, called ABC-SubSim. The effectiveness of the ABC-SubSim algorithm for Bayesian model class selection with simulated data is illustrated using a three-story shear building with Masing hysteresis [6].

2. Formulation

In this section, we review the formulation of a Bayesian hierarchical model class for dynamical systems and then we address the Bayesian model updating and model selection approach for this class of models.

2.1. Formulation of hierarchical stochastic model class

In this section, we present the formulation for a hierarchical stochastic state-space model class $\mathcal{M}$ to predict the uncertain input-output behavior of a system. We start with the general case of a discrete-time finite-dimensional stochastic state-space model of a real dynamic system:

$$\begin{aligned} \forall n \in \mathbb{Z}^+,\ \boldsymbol{x}_n &= \boldsymbol{f}_n(\boldsymbol{x}_{n-1}, \boldsymbol{u}_{n-1}, \boldsymbol{\theta}_s) + \boldsymbol{w}_n \quad &&\text{(State evolution)} \\ \boldsymbol{y}_n &= \boldsymbol{g}_n(\boldsymbol{x}_n, \boldsymbol{u}_n, \boldsymbol{\theta}_s) + \boldsymbol{v}_n &&\text{(Output)} \end{aligned} \tag{3}$$

where $\boldsymbol{u}_n \in \mathbb{R}^{N_I}$, $\boldsymbol{x}_n \in \mathbb{R}^{N_s}$ and $\boldsymbol{y}_n \in \mathbb{R}^{N_o}$ denote the (external) input, dynamic state and output vector at time t_n, and $\boldsymbol{\theta}_s \in \mathbb{R}^{N_p}$ is a vector of uncertain-valued model parameters. In (3), the uncertain state and output prediction errors $\boldsymbol{w}_n$ and $\boldsymbol{v}_n$ are introduced to account for the model being always an approximation of the real system behavior. The prior distributions, $\mathcal{N}(\boldsymbol{w}_n|\boldsymbol{0}, \boldsymbol{Q}_n(\boldsymbol{\theta}_w))$ and $\mathcal{N}(\boldsymbol{v}_n|\boldsymbol{0}, \boldsymbol{R}_n(\boldsymbol{\theta}_v))$, $\forall n \in \mathbb{Z}^+$, are