Mechanical Systems and Signal Processing 66-67 (2016) 31–42

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Fundamental two-stage formulation for Bayesian system identification, Part I: General theory

Siu-Kui Au [b], Feng-Liang Zhang [a,*]

[a] Research Institute of Structural Engineering and Disaster Reduction, Tongji University, China
[b] Center for Engineering Dynamics and Institute for Risk and Uncertainty, University of Liverpool, UK

ARTICLE INFO

Article history:
Received 24 August 2014
Received in revised form
21 March 2015
Accepted 1 April 2015
Available online 6 June 2015

Keywords:
Bayesian
Double counting principle
Modal identification
Prediction error
System identification
Two-stage approach

ABSTRACT

Structural system identification is concerned with the determination of structural model parameters (e.g., stiffness, mass) based on measured response data collected from the subject structure. For linear structures, one popular strategy is to adopt a 'two-stage' approach. That is, modal identification (e.g., frequency, mode shape) is performed in Stage I, whose information is used for inferring the structural parameters in Stage II. Different variants of Bayesian two-stage formulations have been proposed in the past. A prediction error model is commonly introduced to build a link between Stages I and II, treating the most probable values of the natural frequencies and mode shapes identified in Stage I as 'data' for Stage II. This type of formulation, which casts a prediction error model through descriptive statistics, involves heuristics that distort the fundamental nature of the Bayesian approach, although it has appeared to be inevitable. In this paper, a fundamental theory is developed for the Bayesian two-stage problem. The posterior distribution of structural parameters is derived rigorously in terms of the information available in the problem, namely the prior distribution of structural parameters, the posterior distribution of modal parameters in Stage I and the distribution of modal parameters conditional on the structural parameters that connects Stages I and II. The theory reveals a fundamental principle that ensures no double-counting of prior information in the two-stage identification process. Mathematical statements are also derived that provide insights into the role of the structural modeling error. Beyond the original structural model identification problem that motivated the work, the developed theory can be applied in more general settings. In the companion paper, examples with synthetic and real experimental data are provided to illustrate the proposed theory.

1. Introduction

System identification involves making inference about the parameters of a mathematical model based on observed measurements of the real system. Driven by the increasing demand for understanding and using mathematical models of nature and engineered systems consistent with observations, it has become one of the most important problems in modern science and engineering. The Bayesian approach provides a fundamental means for system identification, resolving uncertainties due to the lack of information in the context of probability logic [1–3]. The parameters are viewed as uncertain

* Corresponding author. Tel.: +86 21 6598 7352; fax: +86 21 6598 2668.
E-mail addresses: fengliangzhang@hotmail.com, fengliangzhang@tongji.edu.cn (F.-L. Zhang).

http://dx.doi.org/10.1016/j.ymssp.2015.04.025
0888-3270/

variables and the identification results are cast in terms of their probability distribution after incorporating information from the observed data.

Let $\boldsymbol{\theta}$ be a set of model parameters to be identified from available data D. Bayesian system identification aims at determining the 'posterior distribution' (i.e., given data), $p(\boldsymbol{\theta}|D)$. Using Bayes' Theorem,

$$p(\boldsymbol{\theta}|D) = p(D)^{-1} p(\boldsymbol{\theta}) p(D|\boldsymbol{\theta}) \tag{1}$$

where $p(D)^{-1}$ is a normalizing constant; $p(\boldsymbol{\theta})$ is the 'prior distribution' of $\boldsymbol{\theta}$ (i.e., in the absence of data); and $p(D|\boldsymbol{\theta})$ is the 'likelihood function' that gives the distribution of D for a given $\boldsymbol{\theta}$. If the relationship between $\boldsymbol{\theta}$ and D is complicated, the identification problem can be very challenging. For example, it can be difficult in the first place to formulate the likelihood function $p(D|\boldsymbol{\theta})$ in an explicit form conducive to analytics or computations. The problem may not be 'globally identifiable', i.e., there is more than one or even an infinite number of most probable values, reflecting the fact that the available data is not sufficient for delineating their plausibility. In this case it is also difficult to extract information (e.g., descriptive statistics) about the posterior distribution of $\boldsymbol{\theta}$ [4].

In view of the difficulty of identifying $\boldsymbol{\theta}$ directly from the data D, a 'two-stage' approach has been suggested to convert the original problem into two sub-problems which are more intuitive. This has been motivated by structural system identification problems, where the target is to identify the structural parameters (e.g., stiffness, mass) from vibration data (e.g., acceleration) measured from the subject structure [5–7]. In Stage I the modal properties, i.e., natural frequencies, damping ratios, mode shapes, etc., are first identified. Their identification result is then used for identifying the structural parameters in Stage II. Although intuitive, formulating the two-stage problem in strict accordance with Bayes' rule is mathematically non-trivial. One needs to express the posterior distribution of $\boldsymbol{\theta}$ in Stage II in terms of the posterior distribution of the modal parameters in Stage I, in an explicit manner consistent with the relationship between the modal properties and the data (Stage I), the relationship between the modal parameters and the structural parameters and the prior information on the structural parameters.

Two-stage formulations with different variants have been proposed, e.g., [8–12]. Applications can be found in, e.g., [13–15]. The existing formations, however, involve heuristics in the formulation of the likelihood function $p(D|\boldsymbol{\theta})$ in Stage II. In one popular formulation, in order to link Stages I and II, the most probable value (MPV, a descriptive statistic) of the modal parameters in Stage I is taken as 'data' and modeled to consist of the structural model prediction (which depends on the structural parameters) and a prediction error. The statistical properties of the latter is determined from either ensemble statistics of identification results in Stage I in early developments, or from the posterior statistics in more recent developments [16]. Casting a prediction error model on the most probable modal parameters, which is merely a descriptive statistic characterizing the posterior distribution, has philosophical issues and distorts the fundamental nature of a Bayesian approach. Due to the non-trivial nature of the two-stage problem, however, the heuristic treatment has so far appeared to be unavoidable.

In this work, we develop a general fundamental theory for the Bayesian two-stage problem and apply it to structural system identification based on ambient vibration data. It is presented in two companion papers. In this paper, we derive the equation that fundamentally expresses the posterior distribution of $\boldsymbol{\theta}$ in terms of the posterior distribution of the parameters identified in Stage I. Theoretical issues associated with the formulation are investigated in detail. In the companion paper [17], the general theory is applied to the case of structural model identification using ambient vibration data. Illustrative examples with synthetic and experimental data are presented to verify the method and investigate its applications.

2. Problem context

For clarity we first present the context of the two-stage identification problem. Recall from the introduction that the target is to determine the posterior distribution of $\boldsymbol{\theta}$ from the measured data D. For discussion purposes we refer $\boldsymbol{\theta}$ as the 'structural parameters'. This terminology is motivated from the structural system identification problem. Suppose there is a 'data prediction model' whose set of parameters $\boldsymbol{\alpha}$ can be readily identified from D, in the sense that the likelihood function $p(D|\boldsymbol{\alpha})$ is available in explicit form and conducive to computations. Although not mathematically required, $\boldsymbol{\alpha}$ is often globally identifiable from the data and its posterior statistics (e.g., most probable value and covariance matrix) can be determined efficiently. Both $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$ are related to the data D but intuitively the relationship between $\boldsymbol{\alpha}$ and D is more direct and characteristic. Without loss of generality, divide the parameters in $\boldsymbol{\alpha}$ into two groups,

$$\boldsymbol{\alpha} = [\varpi, \boldsymbol{\upsilon}] \tag{2}$$

where ϖ is related to $\boldsymbol{\theta}$ but $\boldsymbol{\upsilon}$ is not (and is possibly null). The relationship between ϖ and $\boldsymbol{\theta}$ is described through a 'structural prediction model' that gives a prediction of ϖ for a given $\boldsymbol{\theta}$ at least probabilistically, in terms of the conditional distribution $p(\varpi|\boldsymbol{\theta})$.

The idea of a two-stage approach is to first identify ϖ from D (Stage I) and then use the identification result to identify $\boldsymbol{\theta}$ (Stage II). The potential advantage is that the result in Stage I is often demanded and hence calculated anyway; and is relatively easy to obtain and check intuitively, providing a quick consolidation of the information in the data that is useful for making inference about $\boldsymbol{\theta}$. Of course, Stage II can still be difficult but it is likely to be easier than the direct approach and