Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 323 (2017) 373–388
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Calculation of Hessian under constraints with applications to Bayesian system identification

Siu-Kui Au, Yan-Long Xie*

Institute for Risk and Uncertainty and Centre for Engineering Dynamics, University of Liverpool, United Kingdom

Received 21 January 2017; received in revised form 22 April 2017; accepted 15 May 2017
Available online 30 May 2017

Highlights

- Theory for efficient derivation and calculation of Hessian under constraints.
- Allows separate treatment of subject function and constraints.
- Theory applied to Bayesian operational modal analysis.
- Simpler expressions for posterior uncertainty and coherent treatment.

Abstract

In Bayesian system identification with globally identifiable models, the posterior (i.e., given data) probability density function (PDF) of model parameters can be approximated by a Gaussian PDF. The most probable value (MPV) of the parameters is equal to the mean of the Gaussian PDF. It maximises the posterior PDF, or equivalently, minimises the negative of logarithm (NL) of the posterior PDF. The covariance matrix of the Gaussian PDF is equal to the Hessian of the NL at the MPV. Model parameters can be subjected to constraints, which must be accounted for in the calculation of the posterior covariance matrix. In applications such as modal identification, existing strategies define a set of free parameters and map them to the model parameters so that the constraints are always satisfied. The Hessian of the NL with respect to the free parameters is obtained and then transformed to give the posterior covariance matrix of the model parameters where constraints are accounted for. Analytical expressions for this Hessian are complicated because of the composite actions of the NL and the mapping; and this creates significant burden in computer coding. In this work, a theoretical framework is developed for evaluating the Hessian of a function under constraints in a systematic manner. It is applied to obtain new analytical expressions for evaluating the posterior covariance matrix in Bayesian operational modal analysis. The resulting expressions are simpler than existing ones based on direct differentiation. They allow problems with similar mathematical structures to be computer-coded in a coherent manner. Numerical examples are presented to illustrate consistency and computational aspects.

Keywords: Bayesian system identification; Constraint; Hessian; Lagrange multiplier; Operational modal analysis

* Correspondence to: Harrison Hughes Building, Brownlow Hill, Liverpool, L69 3GH, United Kingdom.
E-mail address: yanlong.xie@liverpool.ac.uk (Y.L. Xie).

http://dx.doi.org/10.1016/j.cma.2017.05.021
0045-7825/

1. Introduction

System identification aims at identifying the parameters of a mathematical model from measured data. It is under intensive research in many disciplines, e.g., structural dynamics in mechanical and civil engineering [1–5]. Bayesian approach offers a fundamental means to address uncertainties in system identification [6–8]. In this approach, identification results are encapsulated in the 'posterior' (i.e., given data) probability distribution of parameters. Let $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_{n_\theta}]^T$ be a vector of parameters to be identified and D denotes the measured data. Without much loss of generality, $\boldsymbol{\theta}$ and D are assumed to be continuous-valued. According to Bayes' Theorem, the posterior probability density function (PDF) of $\boldsymbol{\theta}$ for given D is

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

where $p(\boldsymbol{\theta})$ is the 'prior distribution', reflecting the analyst's knowledge in the absence (or without using the information) of data; $p(D)$ is a normalising constant, immaterial to the distribution of $\boldsymbol{\theta}$; $p(D|\boldsymbol{\theta})$ is the 'likelihood function', which is the PDF of D for a given $\boldsymbol{\theta}$. For convenience in analysis and computation, the posterior PDF is often written as

$$p(\boldsymbol{\theta}|D) \propto e^{-L(\boldsymbol{\theta})} \tag{2}$$

where

$$L(\boldsymbol{\theta}) = -\ln[p(D|\boldsymbol{\theta})p(\boldsymbol{\theta})] \tag{3}$$

is called the 'negative logarithm function' (NL) in this work.

One major task in Bayesian system identification is the determination of the statistics associated with the posterior PDF. The computational strategy depends on the topology of the posterior PDF, which reflects whether the data has provided sufficient information for identifying the parameters [9,10]. For 'globally identifiable problems', which is typical in well-posed problems, the posterior PDF has a unique maximum in the interior of the parameter space. The location of the maximum is the posterior 'most probable value' (MPV), $\hat{\boldsymbol{\theta}}$, at which $L(\boldsymbol{\theta})$ is minimised. Approximating $L(\boldsymbol{\theta})$ by a second order Taylor series about $\hat{\boldsymbol{\theta}}$ leads to a Gaussian approximation of $p(\boldsymbol{\theta}|D)$ with a mean equal to $\hat{\boldsymbol{\theta}}$ and a covariance matrix $\hat{\mathbf{C}}$ equal to the inverse of the Hessian of $L(\boldsymbol{\theta})$ at the MPV, i.e.,

$$p(\boldsymbol{\theta}|D) \approx \frac{1}{(2\pi)^{n_\theta/2}\sqrt{|\hat{\mathbf{C}}|}} \exp\left[-\frac{1}{2}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})^T \hat{\mathbf{C}}^{-1}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})\right] \tag{4}$$

where

$$\hat{\mathbf{C}} = (\nabla^2 \hat{L})^{-1} \tag{5}$$

and the hat '^' denotes that the Hessian is evaluated at the MPV.

Depending on how the system identification problem is formulated, there can be constraints among some of the model parameters. For example, in modal identification, mode shapes are subjected to scaling constraints. The entries of the covariance matrix of process noise in a state-space model are subjected to symmetry constraints. While constraints can be handled by proper parameterisation or Lagrange multipliers in the determination of MPV, the problem is more non-trivial for the posterior covariance matrix. Simply taking the Hessian of the NL with respect to (w.r.t.) the original parameters (which are under constraints) at the MPV does not give the right answer.

One way to account for constraints is to define a set of 'free parameters' and map them to the model parameters so that the constraints are always satisfied. The Hessian w.r.t. the free parameters is obtained and then transformed to give the posterior covariance matrix of the model parameters where constraints are considered. This mapping approach has been adopted in the derivation of analytical expressions for evaluating the posterior covariance matrix in Bayesian operational modal analysis (OMA), e.g., [11] for single setup data (see also [12]) and [13] for multiple setup data. The case for single setup data is the conventional setting in OMA where all degrees of freedom (DOFs) are synchronously measured during the same time period. The case for multiple setup data is demanded in practice where it is desired to obtain the mode shape comprising more DOFs than the number of available synchronous data channels (limited by, e.g., the number the sensors). In both cases, the constraint arises from mode shape scaling. Due to the composite action of the NL and the mapping function, the expressions are generally complicated, which creates burden in the development of computer codes. This motivated the present work to develop a more systematic method