Chemical Engineering Science 137 (2015) 796–806

Contents lists available at ScienceDirect

Chemical Engineering Science

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

Sequential Bayesian adaptive Monte Carlo model discrimination framework with application to chemical kinetics

Samira Masoumi *, Park M. Reilly, Thomas A. Duever

Department of Chemical Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1

HIGHLIGHTS

- SBMCMD framework is a combination of model selection and design of experiment (DOE).
- SBMCMD has the ability of handling nonlinear systems using MCMC methods.
- Adaptive MCMC sampling method has been used in the new implementations of SBMCMD.
- Applying DOE lets to discriminates models with the minimum number of experiments.
- The effect of the preliminary data and experimental error has been discussed.

ARTICLE INFO

Article history:
Received 29 December 2014
Received in revised form
31 May 2015
Accepted 29 June 2015
Available online 10 July 2015

Keywords:
Model selection
Sequential Bayesian
Design of experiment
Model uncertainty
Adaptive MCMC
Model discrimination

ABSTRACT

Sequential Bayesian Monte Carlo Model Discrimination (SBMCMD) framework has been previously proposed by the authors for the purpose of determining the underlying mechanisms of a system such as a chemical reaction (Masoumi et al., 2013). SBMCMD relies on sampling from the model parameters distribution using Markov Chain Monte Carlo, MCMC, methods. Effective tuning of MCMC methods, when applying to some nonlinear models, can be tedious and challenging. This limits using SBMCMD in many practical applications. The aim of this paper is to address this limitation and facilitate exploiting of the proposed framework with regards to nonlinear structured models. This is achieved by using adaptive random-walk Metropolis–Hasting method for sampling from the models parameter. This method is an adaptive MCMC algorithm that takes care of adjusting its parameters automatically.

Two implementations of the adaptive SBMCMD framework have been presented and applied to case studies. Results of two implementations have been compared, and the effect of preliminary data has been discussed.

1. Introduction

The objective of model discrimination techniques is to select the “best” among a set of proposed candidate models. This work deals with the problem of determining an underlying physical or chemical mechanism of a system; hence the considered candidate models are mechanistic in nature. In particular we are interested in determining chemical reaction mechanisms, which leads to nonlinear in parameter models with generally more complicated structures comparing with empirical models.

Sequential model discrimination refers to procedures in which a Design of Experiments (DOE) technique is used in conjunction with a model selection method. Gathering experimental information, for example in the case of reaction kinetics, could be expensive and time-consuming. The DOE step helps to find experimental conditions that contain the maximum information with respect to the model discrimination objective. Consequently, the “best” model could be discriminated in the minimum number of experiments.

To be clear, in this paper the term model selection refers to the mathematical process of selecting a model from a candidate set, given data. On the other hand, the term model discrimination indicates a sequential, iterative process, including experimental design and model selection. A recent elegant paper by Galagali and Marzouk (2014) presented a Bayesian model selection algorithm for kinetic models of chemical reactions.

SBMCMD which stands for Sequential Bayesian Monte Carlo Model Discrimination is a framework for model discrimination rather than only model selection, the novelty of which is that, it is a combination of a well-known model discrimination experimental

* Corresponding author at: Department of Chemical Engineering, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada N2L 3G1.
E-mail address: smasoumi@uwaterloo.ca (S. Masoumi).

http://dx.doi.org/10.1016/j.ces.2015.06.067
0009-2509/

design procedure with MCMC marginalization likelihood model selection methods, thus yielding a general Bayesian sequential framework of great value to practicing engineers and scientists.

Markov Chain Monte Carlo (MCMC) methods have been used in the SBMCMD to overcome the need for model linearization, which is required in most sequential model discrimination methods described in the literature, for example Buzzi Ferraris et al. (1984), Buzzi-Ferraris and Forzatti (1983, 1990), Stewart et al. (1996), Ucinski and Bogacka (2005) and Schwaab et al. (2006). The methods presented in this paper, are new implementations of the previously presented framework by the authors (Masoumi et al., 2013). In the new implementations, adaptive random walk Metropolis–Hasting, MH, method is used instead of acceptance-rejection method that has been used in the previous paper of authors. As a result, adaptive sampling fine tunes the multi-dimensional MH with no need for any information about the parameters distribution or expert judgment. In what follows we first establish the general idea of a Bayesian sequential model discrimination framework in Section 2. Sections 3 and 4 describe the model selection and experimental design criteria which have been implemented in SBMCMD. In Section 5 the details of the Adaptive Proposal (AP) Random Walk Metropolis Hasting algorithm, which is used in the execution of the MCMC marginal likelihood model selection methods, are presented. Then, in Section 6 two alternative implementations of the adaptive SBMCMD framework are introduced. Section 7 shows a reaction kinetics example, and finally we make some concluding remarks in Section 8.

2. A Bayesian sequential model discrimination framework

Suppose we have n observations y which are dependent upon a set of parameters θ. Applying Bayes' theorem (Box and Tiao, 1992) results in Eq. (1):

$$P(\theta|\mathbf{y}) \propto \pi(\theta) l(\theta|\mathbf{y}) \tag{1}$$

where $P(\theta|\mathbf{y})$ denotes the posterior probability of parameters given observed data $\mathbf{y}$, $\pi(\theta)$ the parameter prior probability which represents the already available information and $l(\theta|\mathbf{y})$, the likelihood of θ given the data. Considering K candidate models to represent the underlying model of a system, Eq. (1) can be cast into a suitable form for discriminating the "best" one. For this purpose, parameter θ should be replaced by the model index M_k, where represents the kth rival model. Hence, Bayes' theorem could be re-written as Eq. (2) for any of the K candidate models:

$$P(M_k|\mathbf{y}) \propto \pi(M_k) L(M_k|\mathbf{y}) \quad k = 1, \ldots, K \tag{2}$$

In the above equation, the likelihood, represented by L, is the likelihood of the model given the data which its estimation methods will be explained in the next section.

A non-sequential model selection method uses Eq. (2) just once. By contrast, a sequential model discrimination method uses Bayes' formula repeatedly as new data points become available until a stop criterion is satisfied. This process is represented by the following pseudo-code.

Inputs:
K: Number of rival models
$N_{\text{max experiment}}$: Maximum number of experiments
$P_{\text{selection criterion}}$: Stop criterion probability
$\mathbf{y}_0$: Preliminary information
x: Input condition

Algorithm:
Set rival model priors: $P_0(M_k|\mathbf{y}_0)$

for $t = 1$ to $N_{\text{max experiment}}$

- Design the next experiment inputs x_t
- Carry out the experiment to observe the new output data y_t
- for $k = 1$ to K
  - Estimate $L(M_k|y_t)$
  - Update posterior probability :
  - $P_t(M_k|\mathbf{y_t}) = P_{t-1}(M_k|\mathbf{y_{t-1}}) L(M_k|y_t) / \Sigma_k P_{t-1}(M_k|\mathbf{y_{t-1}}) L(M_k|y_t)$
- stop if $\exists\, k \in \{1, \ldots, K\} : P_t(M_k|\mathbf{y_t}) > P_{\text{selection criterion}}$

In the above $\mathbf{y}_t$ represents the vector of all observed data collected up to and including experiment t and y_t is the observed data at time t.

Please note that SBMCMD is a model discrimination method, therefore it is based on the assumption of having one model among the rival ones which can predict the system behavior better than the other candidates. Thus SBMCMD may pick the "best" model even if none of the candidate models can represent the system. Due to this, an adequacy check could be applied after model discrimination to make sure that the final model can actually predict the system behavior. For the adequacy check and when the purpose of model discrimination is obtaining a model to predict the system behavior, best parameters value in the selected model are needed. In the process of SBMCMD procedure, parameters distribution in each candidate model is also obtained. So when the SBMCMD stops, the selected model may have a single-modal parameter distribution. In this case the selected model along with the parameter with highest probability could be used to represent the system. Otherwise, a model regression method should be followed by the model discrimination to find the parameters of the model.

3. Model selection methods

Estimation of the models likelihood, $L(M_k|y_t)$ will be studied in this section since it is the main part of model selection methods. Two of the most popular types of model selection methods are marginalized likelihood methods and those which are based on information criteria.

Methods based on information criteria compare rival models according to their maximum likelihood. These methods penalize the number of parameters. Because without a penalty factor, the models with the larger number of parameters give a better fit to the data points, even though the model with the larger number of parameters does not necessarily predict the system behavior better. Some examples of information criteria model selection methods are Akaike's information theory, AIC, (Akaike, 1987), and Bayesian information criteria, BIC (Schwarz, 1978). Mallows's Cp (Mallows, 1973) is another criterion similar to the information criteria which also considers an approximation of the true error.

The marginalization likelihood model selection methods make up another category of model selection methods. The Marginal likelihood, which is also referred to as "evidence" of each model, is calculated by integrating out the model parameters from the likelihood equation of the model as shown in Eq. (3).

$$L(M_k|\mathbf{y}) = \int_\theta l(\theta|M_k, \mathbf{y}) \pi(\theta|M_k) d\theta \tag{3}$$

where $l(\theta|M_k, \mathbf{y})$ represents the likelihood of the parameters θ given the model M_k, $\pi(\theta|M_k)$ is the prior probability of parameters under model M_k, and $\mathbf{y}$ is the vector of all data points. Bayesian model selection methods, which exploit the marginal likelihood, are called marginalized methods. These methods have become more popular with the advance in computing facilities. Because, most of these methods need more extensive numerical