Automatica 78 (2017) 163–173

Contents lists available at ScienceDirect

Automatica

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

Online model regression for nonlinear time-varying manufacturing systems ☆

Jinwen Hu [a], Min Zhou [b], Xiang Li [c], Zhao Xu [a]

[a] Northwestern Polytechnical University, Xi'an, China

[b] National University of Singapore, Singapore

[c] Singapore Institute of Manufacturing Technology, Singapore

ARTICLE INFO

Article history:
Received 22 December 2014
Received in revised form
9 August 2015
Accepted 27 November 2016

Keywords:
Model regression
Gaussian process regression
Manufacturing systems

ABSTRACT

This paper addresses the online modeling for time-varying manufacturing systems with random unknown model variations between production batches. By modeling the system as a Gaussian process, we first apply the standard Gaussian process regression (GPR) method for estimating the system model, which provides the optimal model estimate with the minimum mean square error (MSE). Then, an iterative form of the method is derived which is more computation efficient but maintains the estimation optimality. However, such optimality is obtained by continuously updating the covariances between the estimated model values and the measurements, which would make the storage and computation unaffordable when the control input can vary within an infinite control space. Due to such a limitation, a suboptimal interactive GPR method is further proposed by trading off the computation efficiency and the estimation accuracy, where the trade-off can be tuned by a designed parameter. Finally, effectiveness and performance of the proposed methods are demonstrated via both simulation and case study by comparing to the conventional nonlinear modeling methods.

1. Introduction

In the manufacturing industry, model regression is always a tough issue which come across by the manufacturers when they want to optimize the production performance, improve the product quality or do failure prognosis for the machines (Buzacott & Shanthikumar, 1993). Nowadays, the manufacturing system is getting more and more complex due to the highly automated and integrated components. As a result, the physical models are hard to derive or even intractable (Hou & Jin, 2011). The method of design of experiments (DOE) or empirical knowledge can help build an initial reference system model for common usage. However, the machine conditions most often are changing from time to time in different environment or with different task allocation (Nassif, 2000). The machines may also suffer from sudden shifts due to component failure and continuous gradual drifts due to normal wear (Moyne, Castillo, & Hurwitz, 2010) in the daily batch productions. Therefore, manufacturers need an online model regression method to track the model variations of the systems so that they can always keep a correct view from input to output and thus better control their product quality.

Online model regression methods (also known as online learning or adaptive modeling methods) have been studied for many years. Some are focusing on the model that can reveal the overall system behavior including its transient and steady state performance, while the others are only focusing on the model of the steady state (Åström & Wittenmark, 2013; Benveniste, Métivier, & Priouret, 2012; Goodwin & Sin, 2013; Wang, Liang, Pan, & Wang, 2016; Wang, Song, Liang, & Pan, 2016). In general, the former is more difficult as it requires more information to build a transient response model which usually involves integration and differentiation. However, in many manufacturing processes, especially for the batch processes, knowing the steady state model is enough since they use fixed set points as inputs of the

☆ This research is funded by the State Key Laboratory of Geo-information Engineering under grant agreement No. SKLGIE2015-M-3-4 and is supported by National Science Foundation of China (Grant No. 61473230), National Science Foundation for Young Scholars of China (Grant No. 61603303) and Aviation Science Foundation (Grant No. 2014ZC53030). The material in this paper was partially presented at the 13th International Conference on Control, Automation, Robotics and Vision (ICARVC 2014), December 10–12, 2014, Marina Bay Sands, Singapore. This paper was recommended for publication in revised form by Associate Editor Jan Komenda under the direction of Editor Ian R. Petersen.

E-mail addresses: hujinwen@nwpu.edu.cn (J. Hu), zhoum@u.nus.edu (M. Zhou), xli@simtech.a-star.edu.sg (X. Li), zhaoxu@nwpu.edu.cn (Z. Xu).

http://dx.doi.org/10.1016/j.automatica.2016.12.012
0005-1098/

machines during the production of each batch, which is known as supervisory control. In such a case, the steady state input–output correlation can be built easily by using many current existing data mining methods such as the partial least squares (Geladi & Kowalski, 1986; Qin, 1998), neural networks (NN) (Chai, Hou, Lewis, Hussain, & Zhao, 2011; Mevawalla, May, & Kiehlbauch, 2011; Monostori & Prohaszka, 1993), fuzzy regression (Chang & Ayyub, 2001; Gandomi & Alavi, 2011; Kacprzyk, 1992), curve fitting (Liu, Pan, Dezert, & Martin, 2016; Liu, Pan, Dezert, & Mercier, 2015; Juliano, 2001; Silverman et al., 1985), local regression (Cleveland, Devlin, & Grosse, 1988; Cleveland & Loader, 1996), Gaussian process regression (GPR) (Nguyen-Tuong, Seeger, & Peters, 2009; Pillonetto & Nicolao, 2010; Quiñonero-Candela and Rasmussen, 2005; Rasmussen, 2004; Wang, Liang, Pan, Zhao, & Yang, 2014; Xu, Choi, Dass, & Maiti, 2013), etc. In many batch manufacturing processes such as vapor phase epitaxy, lithography, and chemical mechanical planarization in the semiconductor industry, the set points (or so-called recipes) are tuned between batches to improve output quality against unknown system shift and drift (Campbell, Firth, Toprac, & Edgar, 2002; Castillo, 1996; Firth, Campbell, Toprac, & Edgar, 2006; Moyne et al., 2010). The parameters of such time-varying system thus need to be estimated based on the batch production data.

So far in the literatures, most online model regression methods are focused on fast computation when data are coming in real-time, rather than on the adaptation of the methods to model variations. For example, the online NN or kernel-based learning methods in Fan and Song (2013), Gönen and Alpaydın (2011), Kivinen, Smola, and Williamson (2004), Liang, Huang, Saratchandran, and Sundararajan (2006), Liu, Wang, Yu, and Li (2010), Liu, Gao, Li, and Wang (2012), White (1989), and the online GPR methods in Chan, Liu, and Chen (2013), Lawrence, Seeger, and Herbrich (2003), Liu and Gao (2015), Nguyen-Tuong, Peters, and Seeger (2008), Quiñonero-Candela and Rasmussen (2005) and Rasmussen (2004) provide iterative calculations for sequential data processing so as to reduce the computation load. Specifically in Chan et al. (2013), a selective recursive Gaussian process modeling algorithm is proposed, which adaptively selects the data of a given size that are most efficient in reducing the estimation uncertainty so that the computation load is constrained when online streaming data are coming in. In most current works on batch-to-batch control, the conventional model regression methods are only used for obtaining an initial estimate of the system model using historical data during which the model variations are ignored (Good & Qin, 2006; Liu, Jia, Liu, & Huang, 2007; Moyne et al., 2010). The estimated model is then used as a benchmark to identify the model variation of each batch, where the models are usually assumed to be linear. However, such a way of regression modeling is not accurate and reliable since the data of historical batches also suffer from model variations which cannot be ignored. Therefore, it is required to design a new online model regression method which takes into account such variations in all batches of production. The key challenge in dealing with the model regression of a time-varying nonstationary process is how to adaptively fuse the data with time-varying nonlinear spatio-temporal correlations.

In this paper, we mainly study the online model regression for time-varying systems under the framework of GPR (Rasmussen, 2004). This topic is studied due to the following reasons. First, compared with other model regression methods, the GPR method is more general since it is applicable for not only deterministic systems but also stochastic systems and it tells not only what the estimated model is, but also how good the estimation is by its associated variance. Second, unlike other nonlinear model regression methods which only aim to find a good match between the inputs and the outputs, the GPR method explains in-depth data correlations and can be easily combined with Bayesian inference to do forecasting (Pillonetto, Chiuso, & Nicolao, 2011). Third, the standard GPR method is not applicable for online regression with an infinite number of set points in that the dimension of data matrix will go up to infinity if the real-time production data stream in. Finally, the existing modifications of GPR method are mainly designed for reducing computation load and have not considered the effect of model variations.

The main contribution of this paper is that an online model regression method is designed for nonlinear time-varying systems, which is computation efficient in estimating the output values and also provides a confidence evaluation on the model estimates. It can be applied in the modeling of complex nonlinear manufacturing processes with unknown random shifts and drifts, where the model estimate is continuously updated using online input and output data. First, we directly extend the standard GPR method for model regression of time-varying systems with batch data processing, which gives the optimal estimate with minimum mean square error (MSE). Second, the extended method with batch data processing is further modified into an iterative form so that it is more computation efficient for online regression. However, in order to keep the estimation optimality, the covariance between any two model outputs must be continuously updated along with the model estimate, which requires a large storage space. Third, regarding such a limitation, a new suboptimal iterative GPR method is proposed so that both computation and storage requirements are affordable at each iteration. In the meantime, the asymptotic estimation stability is guaranteed when model variation exists. Finally, the proposed methods are validated via simulation and case study as compared with other conventional model regression methods.

The paper is organized as follows. In Section 2, basic assumptions are given and problem is formulated. The method development is illustrated in Section 3. In Section 4, simulation and case study are implemented to demonstrate the effectiveness of the proposed methods. Section 5 draws the conclusions.

2. Basic assumptions and problem formulation

In this paper, we consider the nonlinear system model $f_t(u_t) : \mathbb{R}^q \mapsto \mathbb{R}$, where t is the discrete time index number corresponding to the time at which the quantities are evaluated, u_t is the control input vector selected from a predefined control space $\mathbb{U} \subset \mathbb{R}^q$ and f_t is some concerned quality characteristic. There can be more than one quality characteristics, but each of which is individually considered. Thus, without loss of generality f_t is assumed to be a scalar. At each evaluation time t during the manufacturing process, the quality characteristic is measured and the following measurement is obtained:

$$z_t = f_t(u_t) + v_t, \tag{1}$$

where v_t is the measurement noise and its variance is denoted by σ_v^2.

Note that in the manufacturing applications, t may represent the batch or run number of production so that z_t represents the sample mean of all the output/quality measurements taken for the tth batch (Good & Qin, 2006). In this paper, we do not consider the change of process setting or measurements within a production batch. The function f_t denotes the unknown nonlinear system model which is time-varying. The control settings and quality measurements of the process are usually stored in a database. To use such information, denote by U_t a matrix enclosing all the control inputs up to time t with the ith column being u_i ($i = 1, 2, \ldots, t$), and similarly denote by Z_t a vector with the ith entry being z_i.

In the real manufacturing processes, engineers can always get some prior model estimate by design of experiments or domain