Journal of Process Control 58 (2017) 11–22

Contents lists available at ScienceDirect

Journal of Process Control

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

Research paper

Multi-model multivariate Gaussian process modelling with correlated noises

Xiaodan Hong [a], Biao Huang [b,*], Yongsheng Ding [a,*], Fan Guo [a], Lei Chen [a], Lihong Ren [a]

[a] *Engineering Research Center of Digitized Textile & Apparel Technology, Ministry of Education, College of Information Science and Technology, Donghua University, Shanghai 201620, China*
[b] *University of Alberta, Edmonton, AB T6G 2R3, Canada*

ARTICLE INFO

Article history:
Received 10 April 2017
Received in revised form 19 July 2017
Accepted 8 August 2017

Keywords:
Multivariate GPR
Correlated Gaussian noises
Gaussian mixture model (GMM)
Multiple-model
Dependent multivariate GPR (DMGPR)
EM algorithm
PSO algorithm

ABSTRACT

A composite multiple-model approach based on multivariate Gaussian process regression (MGPR) with correlated noises is proposed in this paper. In complex industrial processes, observation noises of multiple response variables can be correlated with each other and process is nonlinear. In order to model the multivariate nonlinear processes with correlated noises, a dependent multivariate Gaussian process regression (DMGPR) model is developed in this paper. The covariance functions of this DMGPR model are formulated by considering the "between-data" correlation, the "between-output" correlation, and the correlation between noise variables. Further, owing to the complexity of nonlinear systems as well as possible multiple-mode operation of the industrial processes, to improve the performance of the proposed DMGPR model, this paper proposes a composite multiple-model DMGPR approach based on the Gaussian Mixture Model algorithm (GMM-DMGPR). The proposed modelling approach utilizes the weights of all the samples belonging to each sub-DMGPR model which are evaluated by utilizing the GMM algorithm when estimating model parameters through expectation and maximization (EM) algorithm. The effectiveness of the proposed GMM-DMGPR approach is demonstrated by two numerical examples and a three-level drawing process of Carbon fiber production.

1. Introduction

Data-driven approaches have received a great attention in recent years and have found many applications in various engineering fields, such as fault-tolerant control [1], process monitoring [2] and system identification [3–5]. Data-driven modelling has been widely utilized in system identification due to the complexity of the first-principle models [3–5]. The Gaussian process regression (GPR) algorithm is a data-driven modelling approach in the field of machine learning, and has a found good adaptability to high dimension, small sample set, and nonlinear problems [6–8]. The popularity of GPR is partly due to its solid theoretical basis in Bayesian Statistics, and partly because its hyper-parameters can be adaptively acquired [9,10]. In addition, it has a strong generalization ability and convenience of implementation [7,9,11]. Owing to these advantages, the GPR algorithm has been widely used in system identification [12], response surface modelling [13], dynamic process modelling [14,15], ensemble learning [16], and other application fields [4,17,18]. Various empirical modelling results have demonstrated superiority of the GPR model to other supervised regression models such as support vector machine, fuzzy model, and neural networks, due to its good prediction accuracy along with ability to provide probability distribution of the predictions [7,13,15,19].

Owing to the computational complexity in the GPR model, the Gaussian process is commonly considered under the univariate GPR modelling framework. However, most of the existing industrial processes are multivariate rather than univariate. To model multivariate Gaussian processes, some GPR modelling approaches have also been developed in the literatures [20–24]. A pragmatic and straightforward approach is to model multiple outputs independently without considering their correlation. Although this is not the most ideal approach, it has been used in some practical applications [13,19,20]. However, in order to improve prediction accuracy of the multivariate GPR model, one should consider not only the correlation between data points, but also the correlation between the output functions [21–24].

Furthermore, the existing multivariate GPR modelling approaches commonly assume the independent Gaussian white noises without considering the correlation between noises that

* Corresponding authors.
E-mail addresses: bhuang@ualberta.ca (B. Huang), ysding@dhu.edu.cn (Y. Ding).

http://dx.doi.org/10.1016/j.jprocont.2017.08.004
0959-1524/

affect different output variables [7,25]. However, in real industrial processes, the noises affecting multiple output variables can be correlated with each other due to reasons such as variation of the ambient conditions, correlation of the sensors etc. In this paper, we model multivariate Gaussian processes with correlated Gaussian noises; that is, different output variables are corrupted by different Gaussian noises, and there are correlation between these noises, so that a dependent multivariate Gaussian process regression (DMGPR) model is developed. A key to the proposed DMGPR model is the formulation of the covariance function, which describes the "between-data" covariance, the "between-output" covariance, as well as the covariance between correlated noises.

In addition, owing to the complexity of nonlinear systems and the nature of multi-mode operations in industrial processes, the single MGPR model may not be capable of achieving good prediction results in general. The multiple-model approach is an effective modelling approach which improves prediction accuracy and robustness of the model. In recent years, this approach has been widely studied in the field of system identification and intelligent control [26–28], and has been successfully applied in many industrial processes, such as maneuvering target tracking [29,30], prediction of silicon content in hot metal [31], and industrial process modelling in multiple operating environments [32] among others. In this paper, to improve prediction performance of the proposed DMGPR model, we will use the multiple-model approach to optimize the DMGPR model, and propose a GMM-DMGPR modelling approach which considers multiple DMGPR models by employing the GMM algorithm to classify the samples. The GMM algorithm is a clustering method which clusters data to multiple Gaussian components and calculates the probability of data belonging to each component [32,33]. Then, these probabilities are considered as the weights of the data in multiple sub-DMGPR models when estimating hyper-parameters of the composite GMM-DMGPR model through expectation and maximization (EM) algorithm. Hence, unlike the traditional multiple-model approach which models each sub-model with only the classified samples [26–28,32], the proposed GMM-DMGPR approach improves prediction accuracy by using all sampled data to model each sub-DMGPR model along with the weights of these samples belonging to each sub-model.

The main contributions of this paper are as follows: (i) a dependent multivariate Gaussian process regression (DMGPR) model with correlated noises is developed; (ii) a composite multiple-model approach is considered to optimize the DMGPR model; (iii) the GMM-DMGPR modelling approach is proposed by utilizing the GMM algorithm to cluster data along with their weights.

The rest of the paper is organized as follows: In Section 2, a traditional MGPR model with independent Gaussian white noise is revisited. In Section 3, a DMGPR model with correlated Gaussian noises is developed, and a GMM-DMGPR modelling approach is proposed. A verification of the performance of the GMM-DMGPR approach is performed in Section 4 by employing numerical examples and a three-level drawing of a Carbon fiber example. In Section 5, conclusions are presented.

2. Revisit of traditional MGPR model

The multivariate GPR model with independent Gaussian noise can be described by the following equation:

$$Y(X) = F(X) + E \tag{1}$$

where $X(\in \mathbb{R}^d)$ represents the input vector, $Y = [y_1, \cdots, y_q](\in \mathbb{R}^q)$ is the corresponding output vector at X, $F(\cdot)$ denotes an latent function, and $F(X) = \left[f_1(X), \cdots, f_q(X)\right](\in \mathbb{R}^q)$ is the latent function vector at X. Here, $F(X)$ is assumed to follow a q-variate joint Gaussian distribution with mean 0 and covariance matrix B_q, that is, $F(X) \sim G_q(0, B_q)$, where B_q is a $q \times q$ symmetric matrix which describes the correlation between outputs, and has the following equation:

$$B_q = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{q1} & b_{q2} & \cdots & b_{qq} \end{bmatrix} \tag{2}$$

where element b_{gh} in (2) represents the covariance between output variables $f_g(X)$ and $f_h(X)$, $g, h = 1, \cdots, q$. In addition, $E = \left[\varepsilon_1, \cdots, \varepsilon_q\right]$ represents the noise vector, and each element $\varepsilon_g (g = 1, \cdots, q)$ is assumed to be an independent Gaussian white noise with mean 0 and variance $\sigma_g^2, g = 1, \cdots, q$. That means the observation noise ε_g of each output variable y_g follows an independent Gaussian distribution as $\varepsilon_g \sim N(0, \sigma_g^2)$; consequently, E follows the joint Gaussian distribution as $E \sim G(0, S_q)$. Owing to the independence among q noises $\varepsilon_1, \cdots, \varepsilon_q$, the noise covariance matrix S_q can be described as a $q \times q$ diagonal matrix, i.e. $S_q = diag(\sigma_1^2, \cdots, \sigma_q^2)$.

We consider a sample set $\left\{X_i, Y_i\right\}_{i=1}^n$, where n denotes the number of given samples, $X_i(\in \mathbb{R}^d)$ and $Y_i = [y_{i1}, \cdots, y_{iq}](\in \mathbb{R}^d)$ are the input and output at sampling instant i, respectively. Under the assumption of the Gaussian process, for the output of the dimension $g(=1, \cdots, q)$, the n latent function values $f_g(X) = \left\{f_g(X_1), \cdots, f_g(X_n)\right\}$ follow a Gaussian process with mean 0 and a covariance function $k(X, X') = Cov(f_g(X), f_g(X'))$, that is, $f_g(X) \sim GP(0, k(X, X'))$. So, the Gaussian process prior of the n latent function values $f_g(X) = [f_g(X_1), \cdots, f_g(X_n)]^T$ can be written as $f_g(X) \sim N(0, K_n)$, where K_n represents an $n \times n$ covariance matrix which describes the correlation between different sample points and is formulated by employing the covariance function $k(X, X')$ as the following equation:

$$K_n = \begin{bmatrix} k(X_1, X_1) & k(X_1, X_2) & \cdots & k(X_1, X_n) \\ k(X_2, X_1) & k(X_2, X_2) & \cdots & k(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(X_n, X_1) & k(X_n, X_2) & \cdots & k(X_n, X_n) \end{bmatrix} \tag{3}$$

where element $k(X_i, X_j)$ in (3) represents the covariance between function values $f_g(X_i)$ and $f_g(X_j)$, $i, j = 1, \cdots, n$. To simplify the implementation, we consider a squared exponential function to describe the covariance function $k(X, X')$ as given below [7,11]:

$$k(X, X') = \sigma_f^2 \exp\left(-\sum_{h=1}^{d} \frac{(x_h - x_h{'})^2}{2l_h^2}\right) \tag{4}$$

where $h(=1, \cdots, d)$ is the input dimension index, x_h and x'_h represent the h-th dimension of inputs $X \in \mathbb{R}^d$ and $X' \in \mathbb{R}^d$, respectively. σ_f^2 denotes the magnitude or signal variance of the covariance function, and l_h is the characteristic length-scale corresponding to the h-th input dimension.

Owing to the Gaussian process assumption, $F(X) \sim G_q(0, B_q)$ and $f_g(X) \sim N(0, K_n)$, the $n \times q$ latent function vector $[F^T(X_1), \cdots, F^T(X_n)]^T$ follows a matrix Gaussian distribution $MN(0, K_n, B_q)$, considering the "between-data" correlation and the "between-output" correlation simultaneously. Consequently, the distribution of $[F^T(X_1), \cdots, F^T(X_n)]^T$ can be described as follows [21]:

$$[F^T(X_1), \cdots, F^T(X_n)]^T \sim G_{nq}\left(0, B_q \otimes K_n\right) \tag{5}$$