Chemometrics and Intelligent Laboratory Systems 138 (2014) 97–109

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Soft sensor model development in multiphase/multimode processes based on Gaussian mixture regression

Xiaofeng Yuan, Zhiqiang Ge *, Zhihuan Song

State Key Laboratory of Industrial Control Technology, Institute of Industrial Process Control, Department of Control Science and Engineering, Zhejiang University, Hangzhou, 310027 Zhejiang, P. R. China

ARTICLE INFO

Article history:
Received 7 May 2014
Received in revised form 7 July 2014
Accepted 17 July 2014
Available online 25 July 2014

Keywords:
Soft sensor
Gaussian mixture regression
Multiphase process
Multimode process
Quality prediction

ABSTRACT

For complex industrial plants with multiphase/multimode data characteristic, Gaussian mixture model (GMM) has been used for soft sensor modeling. However, almost all GMM-based soft sensor modeling methods only employ GMM for identification of different operating modes, which means additional regression algorithms like PLS should be incorporated for quality prediction in different localized modes. In this paper, the Gaussian mixture regression (GMR) model is introduced for multiphase/multimode soft sensor modeling. In GMR, operating mode identification and variable regression are integrated into one model; thus, there is no need to switch prediction models when the operating mode changes from one to another. To improve the GMR model fitting performance, a heuristic algorithm is adopted for parameter initialization and component number optimization. Feasibility and efficiency of GMR based soft sensor are validated through a numerical example and two benchmark processes.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

In modern industrial plants, process operating condition is frequently monitored and controlled in order to improve process efficiency and produce high-quality products. However, these techniques are highly dependent on accurate model identification and reliable measurements. Especially, online analysis of key process variables is critical for process monitoring and control. In some cases, due to reasons of high investment cost, technical difficulty, measurement delay, and so on, the process often encounters the great challenge of lacking accurate real-time measurements of key process variables. Over the past decades, soft sensors have been widely used to tackle these problems, which provide frequent estimations of key process variables through those that are easy to measure online [1–5]. By far, the majority of soft sensors are based on data-driven modeling methods, which construct inferential models by using abundant process measurement data. As a category of data-driven soft sensors, the multivariate statistical techniques such as principal component analysis (PCA) [6], partial least squares (PLS) [7], and independent component analysis (ICA) [8] are extensively researched and employed in diverse processes. In addition, machine learning based nonlinear methods like artificial neural network (ANN) [9] and support vector machine (SVM) [10] have also been also applied for soft sensor modeling. Apart from these, many soft sensors have also been developed with the usage of fuzzy systems like Takagi–Sugeno fuzzy model [11–14].

Though different types of soft sensor modeling techniques have been applied for quality prediction, most of them are based on the assumption that process data are generated from a single operating region and follow a unimodal Gaussian distribution. For complex multiphase/multimode processes that are running at multiple operating conditions, the basic assumption of multivariate Gaussian distribution does not met because of the mean shifts or covariance changes. Consequently, data distribution may be complicated with arbitrary non-Gaussian patterns. Meanwhile, another problem in multiphase/multimode processes is that a single global soft sensor model is no longer well suited in predicting the output of key process variables. As mixture models can represent arbitrarily complex probability density functions, they are ideal tools to model complex multi-class data distribution.

By taking sufficient linear combinations of single multivariate Gaussian distributions, the Gaussian mixture model (GMM) can smoothly approximate almost any continuous density to arbitrary accuracy [15]. Thus the GMM technique has shown strong ability in dealing with non-Gaussian data and can be used for classification or cluster problems in various fields. Speech recognition [16], image segmentation [17], and robotic learning [18] are some typical applications. With respect to the process industry, GMM is extensively utilized for multiphase/multimode process monitoring and soft sensor applications [19–22]. In these researches, the main purpose of GMM is to identify and localize operating mode of data. For example, a novel multimode process monitoring approach based on finite Gaussian mixture models and Bayesian inference strategy is proposed in reference [20].

* Corresponding author. Tel.: +86 87951442.
E-mail address: gezhiqiang@zju.edu.cn (Z. Ge).

http://dx.doi.org/10.1016/j.chemolab.2014.07.013
0169-7439/© 2014 Elsevier B.V. All rights reserved.

As have been mentioned, GMM can classify those data similar to each other into the same class, and then certain data modes can be obtained in multiphase/multimode processes. In order to carry out soft sensor estimation of process variables, localized soft sensor algorithm should be constructed to model the relationship between output and input variables in each mode, respectively. In other words, two steps are carried out in sequence when building soft sensors for multiphase/multimode processes: mode localization step and localized regression step. It is cockamamie and time consuming. First of all, both classification and regression approaches should be incorporated. Second, multiple regression models should be trained, the number of which is equivalent to the number of operating modes. An improved method is to give a weighted output by assigning membership degrees to specific regions. For example, in Lughofer's research work [13,14], a certain number of piecewise linear predictors are build in different regions, then the output of the query sample is achieved by a weighted sum of all the piecewise linear predictors. Such a technique can obtain good prediction accuracy and address the non-linear problem of processes.

In a simpler and more direct way, we introduce a new soft sensing method for multiphase/multimode processes, which is called Gaussian mixture regression (GMR) [23,24]. GMR, first proposed in [23], is an extension of Gaussian mixture model and can be exploited for regression problems. The regression procedure is based on the Gaussian conditioning and linear combination properties of Gaussian distributions. By separating the data point into input and output parts, the joint probability distribution of input and output of data point is modeled in a GMM. Then the conditional probability distribution of output on input is estimated with parameters obtained in GMM. After the training step, output can be predicted when an input data comes.

Compared to deterministic regression methods like the fuzzy systems [13,14], GMR tries to find the relationship between the input and output under a probabilistic maximum likelihood framework. Besides, as mentioned in reference [24], GMR is easy to implement, and satisfies robustness and smoothness criterions that are common to a variety of fields. Although the theoretical considerations of GMR were presented two decades ago, it has come out with only few applications [25–28]. By far, GMR has mainly been utilized in area of robot programming by demonstration (PbD) for imitation learning of multiple tasks [28,29]. To our best knowledge, this method has not been applied in other fields yet. Particularly, no literature about GMR has been found for soft sensor application in chemical processes up to now. Therefore, the advantage of GMR in soft sensor modeling has not yet been explored.

Two practical but fundamental issues to be addressed in employing GMR are as follows: (1) how to determine the number of Gaussian components, (2) how to calculate the parameters in the mixture model. Too small number of components will result in under-fitting problem, while a large number suffers from computational burden and data overfitting; thus, a proper number of Gaussian components are critical to adequately describe the data in GMR. There are several techniques such as Schwarz's Bayesian inference criterion (BIC) [30], the minimum message length (MML) [31], and the minimum description length (MDL) [32] that can be used. For the second issue, the expectation maximization (EM) algorithm [33] is a classical method for learning the parameters of mixture models. By iteratively running E-step and M-step, the parameters will converge to optimum values. Nevertheless, this algorithm has some drawbacks like the need for a predetermined component number and critical dependence on initialization. In general, by executing certain different runs of EM algorithm with different initializations and different component numbers, and assessing each estimation with some criteria like BIC, the optimal number of components can be obtained by maximizing or minimizing the criterion. However, this is computational burden and time consuming.

In this paper, a simple absolute increment of log-likelihood (AIL) criterion is used to determine the optimal components, which will show great effectiveness in later case studies. To overcome the disadvantage of the EM algorithm, we adopt a heuristic procedure. We first start with an adequately number of components to have a fine-scale representation of data. Then by sequentially pruning the least probabilistic component, merging the least probabilistic one with the closest component and obtaining a nearly good initialization for the next run of EM, new mixture models with reduced components will be obtained successively. Detailed description of the schedule will be explained in Section 3.

The remainder of this paper is organized as follows. In Section 2, the definition of GMM and the EM algorithm for parameter estimation are briefly revisited. Then the GMR based soft sensor with optimal components selection is introduced in Section 3. A numerical example and two application examples are provided in Section 4. Finally, conclusions are made.

2. Preliminaries

2.1. Gaussian mixture model

Let $z \in R^d$ be a data point of d-dimensional variable. If z comes from a unimodal multivariate Gaussian distribution, then the probability density function is given by

$$f(z|\theta) = \frac{1}{\sqrt{(2\pi)^d|\sum|}} \exp\left\{-1/2(z-\mu)^T \sum{}^{-1}(z-\mu)\right\} \quad (1)$$

where μ is the mean vector and Σ is the covariance matrix, and $\theta=\{\mu,\Sigma\}$ are the parameters to determine a Gaussian distribution. When the data point z is from a mixture Gaussian model, then its probability density function can be formulated as follows [15]

$$p(z|\Omega) = \sum_{k=1}^{K} \omega_k f(z|\theta_k) \quad (2)$$

where K is the number of Gaussian components in GMM, ω_k is the probabilistic weight of the kth Gaussian component and subjects to condition of $\omega_k \geq 0$, $\sum_{k=1}^{K} \omega_k = 1$, $\theta_k = \{\mu_k,\Sigma_k\}$ represent the parameters in the kth Gaussian component (mean vector μ_k and covariance matrix Σ_k), and $\Theta = \{\theta_1, \theta_2, \cdots \theta_K\} = \{\mu_1, \sum_1, \mu_2, \sum_2, \cdots \mu_K, \sum_K\}$ denote all parameters defining each Gaussian component, respectively. Then the total parameters in the complete GMM with K components can be defined as $\Omega = \{\{\omega_1, \mu_1, \sum_1\}, \{\omega_2, \mu_2, \sum_2\}, \cdots \{\omega_K, \mu_K, \sum_K\}\}$, which involve both the Gaussian model parameters θ_k and the mixing probabilities ω_k ($1 \leq k \leq K$). Given a set of N independent and identically distributed training samples $Z = [z_1, z_2, \ldots, z_N]$, the likelihood and log-likelihood function of Z can be written as

$$L(Z,\Omega) = \prod_{n=1}^{N} \sum_{k=1}^{K} \omega_k f(z_n|\theta_k) \quad (3)$$

and

$$\log L(Z,\Omega) = \sum_{n=1}^{N} \log\left(\sum_{k=1}^{K} \omega_k f(z_n|\theta_k)\right) \quad (4)$$

2.2. EM algorithm for GMM

To build a GMM, the unknown parameter set Ω need to be estimated firstly. This problem is equal to find parameters that maximize the log-likelihood function formulated as

$$\hat{\Omega} = \arg\max_{\Omega}(\log L(Z,\Omega)) \quad (5)$$

Download English Version:

https://daneshyari.com/en/article/1180737

Download Persian Version:

https://daneshyari.com/article/1180737

Daneshyari.com