

12-12-2019

A Just-in-time Learning Soft Sensing Modeling Method Based on Bayesian Gaussian Mixture Model

Qi Cheng

1. College of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China; ;

Weili Xiong

2. Key Laboratory of Advanced Control for Light Industry Process of Ministry of Education, Jiangnan University, Wuxi 214122, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

A Just-in-time Learning Soft Sensing Modeling Method Based on Bayesian Gaussian Mixture Model

Abstract

Abstract: For some time-varying industrial processes with non-Gaussian properties, the model established by the general soft-sensing method is difficult to meet the accuracy requirement. To solve the above problems effectively, a *JITL soft sensor modeling method (BGMM)* is proposed based on *Bayesian Gaussian Mixture Model*. For the given training sample set, the number of components of the Gaussian mixture model is optimized by *Bayesian Information Criterion (BIC)*; For new test samples, *Gaussian Process Regression (GPR)* model is established by using the *BGMM similarity criterion* for the training samples to find out the most similar set; The model is used to predict the test samples. The effectiveness of the proposed method is verified by modeling and simulating the concentration of butane at the butane tower bottom .

Keywords

Gaussian mixture model, just-in-time learning, Bayesian information criterion, Gaussian process regression

Recommended Citation

Qi Cheng, Xiong Weili. A Just-in-time Learning Soft Sensing Modeling Method Based on Bayesian Gaussian Mixture Model[J]. *Journal of System Simulation*, 2019, 31(8): 1555-1561.

基于 BGMM 的即时学习软测量建模方法

祁成¹, 熊伟丽²

(1. 江南大学 物联网工程学院 自动化研究所, 无锡 214122; 2. 江南大学 轻工过程先进控制教育部重点实验室, 无锡 214122)

摘要: 对于具有非高斯特性的时变工业过程, 一般的软测量方法建立的模型很难满足精度要求。为有效解决上述问题, 提出一种基于贝叶斯高斯混合模型(BGMM)的即时学习软测量建模方法。对于给定的训练样本集, 利用贝叶斯信息准则对高斯混合模型的成分个数进行优化; 对于新的测试样本, 利用 BGMM 相似度准则从训练样本中找出与之最相似的一组样本建立高斯过程回归模型; 用该模型对测试样本进行预测。通过脱丁烷塔塔底丁烷浓度的软测量建模仿真, 验证了所提方法的有效性。

关键词: 高斯混合模型; 即时学习; 贝叶斯信息准则; 高斯过程回归

中图分类号: TP273 文献标识码: A 文章编号: 1004-731X (2019) 08-1555-07

DOI: 10.16182/j.issn1004731x.joss.17-0285

A Just-in-time Learning Soft Sensing Modeling Method Based on Bayesian Gaussian Mixture Model

Qi Cheng¹, Xiong Weili²

(1. College of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China;

2. Key Laboratory of Advanced Control for Light Industry Process of Ministry of Education, Jiangnan University, Wuxi 214122, China)

Abstract: For some time-varying industrial processes with non-Gaussian properties, the model established by the general soft-sensing method is difficult to meet the accuracy requirement. To solve the above problems effectively, a JITL soft sensor modeling method (BGMM) is proposed based on Bayesian Gaussian Mixture Model. For the given training sample set, the number of components of the Gaussian mixture model is optimized by Bayesian Information Criterion (BIC); For new test samples, Gaussian Process Regression (GPR) model is established by using the BGMM similarity criterion for the training samples to find out the most similar set; The model is used to predict the test samples. The effectiveness of the proposed method is verified by modeling and simulating the concentration of butane at the butane tower bottom.

Keywords: Gaussian mixture model; just-in-time learning; Bayesian information criterion; Gaussian process regression

引言

在一些复杂的工业过程中, 随着产品质量要求的不断提高, 需要对一些直接决定产品质量的过

程变量进行严格的监测和控制。但是由于某些测量仪器价格昂贵或者技术条件的制约, 使得这些变量无法用在线仪器测量得到。对于这些难以测得的变量, 可以通过建立软测量模型的方法进行估计和预测, 常用的软测量方法有偏最小二乘法^[1] (partial least squares, PLS)、人工神经网络^[2] (artificial neural networks, ANN)、支持向量机^[3] (support vector machine, SVM)等。PLS 可以很好的处理过程的线



收稿日期: 2017-06-14 修回日期: 2017-07-28;
基金项目: 国家自然科学基金(61773182), 国家重点研发计划子课题(2018YFC1603705-03);
作者简介: 祁成(1994-), 男, 安徽滁州, 硕士生, 研究方向为工业过程建模; 熊伟丽(1978-), 女, 江苏无锡, 博士生, 教授, 博导, 研究方向为复杂工业过程建模及优化、软测量技术。

<http://www.china-simulation.com>

• 1555 •

性问题,然而实际工业过程常常呈现非线性,因此线性方法不再适用。非线性建模方法如 ANN、SVM 等,虽然可以较好地处理过程的非线性,但存在优化参数较多等问题。近年来,高斯过程回归^[4](Gaussian process regression, GPR)受到了越来越多的关注,作为一种非参数概率模型^[5],它不仅能够得到模型的预测值,还能得到预测值对模型的信任值。与 ANN、SVM 等方法相比,GPR 需要优化的参数较少,在解决小样本、非线性问题中具有独特的优势^[6]。

在离线建立好的模型投入运行以后,由于生产环境的一些变化以及对产品质量要求的不断改变等原因,之前建立好的模型可能会出现不再适用于当前工况的情况,预测的结果不能满足精度要求。针对这一问题,常用的解决方法有基于滑动窗^[7](Moving window, MW)和基于即时学习^[8](Just-in-time learning, JITL)的方法,但是 MW 的窗口长度难以确定且不适用于过程的突变,JITL 方法根据相似输入产生相似输出的原理^[9],选择与测试样本最相似的一组训练样本来建立局部模型进行预测,可以较好地解决过程突变问题。

对于一些呈现非高斯性的时变工业过程,传统的 JITL 方法是基于欧氏距离或者和角度相结合的相似度准则来选择相似数据,无法充分考虑到过程数据的非高斯性。因此,本文提出一种基于贝叶斯高斯混合模型(Bayesian Gaussian mixture model, BGMM)的在线建模方法,它不仅考虑到了过程的时变性,而且在选择相似数据建立局部 GPR 模型时,充分考虑到数据的非高斯特性,更为合理的选择相似数据。首先,利用贝叶斯信息准则(Bayesian information criterion, BIC)确定最优的高斯成分个数;然后,当新的测试数据到来时,计算其隶属于每个高斯成分的后验概率,并求出其与训练数据之间的马氏距离,将两者融合作为相似度指标;最后,从原始的训练样本中选取相似度最大的一组数据来建立当前的 GPR 模型,并进行模型输出预测。通过对脱丁烷塔塔底丁烷浓度的建模仿真,表明

所提建模方法有较高的预测精度和泛化能力。

1 基于贝叶斯信息准则的高斯混合模型

1.1 高斯混合模型

GMM 在聚类方面有很好的效果,并且已经成功应用于很多领域^[10]。当高斯成分(满足单一高斯分布的一组数据就是一个高斯成分)足够多时,任意数据分布都可以用 GMM 进行拟合。已知包含 N 个训练样本的数据集 $\mathbf{X}\{\mathbf{x}_i \in R^m, i=1,2,\dots,N\}$, m 表示输入数据的维数,它的概率密度函数可以表示为:

$$p(\mathbf{X}|\Theta) = \sum_{k=1}^K \alpha_k p(\mathbf{X}|\theta_k) = \prod_{i=1}^N \sum_{k=1}^K \alpha_k p(\mathbf{x}_i|\theta_k) \quad (1)$$

式中: $\Theta = [\alpha_1, \mu_1, \Sigma_1; \alpha_2, \mu_2, \Sigma_2; \dots; \alpha_k, \mu_k, \Sigma_k]$ 是 GMM 的参数; K 是高斯成分的个数; θ_k 为第 k 个高斯成分的参数; $\theta_k = (\mu_k, \Sigma_k)$, μ_k 和 Σ_k 分别为第 k 个高斯成分的均值和协方差矩阵; α_k 为第 k 个高斯成分所占的比例,且 $\sum_{k=1}^K \alpha_k = 1$; $0 < \alpha_k < 1$,

其中第 k 个高斯成分的概率密度函数为:

$$p(\mathbf{x}_i|\theta_k) = \frac{1}{\sqrt{(2\pi)^m |\Sigma_k|}} \exp\left[-\frac{1}{2}(\mathbf{x}_i - \mu_k)\Sigma_k^{-1}(\mathbf{x}_i - \mu_k)^T\right] \quad (2)$$

一般通过期望最大化算法^[11]对 GMM 方法中的未知参数进行求解。具体求解过程分为 E 步和 M 步,其描述如下:

E 步:用当前第 l 次更新的参数 α_k^l 和 θ_k^l ,通过贝叶斯公式计算第 i 个训练样本属于第 k 个高斯成分的概率。其中 C_k 表示第 k 个高斯成分。

$$\beta_{ik}^l = p^l(C_k|\mathbf{x}_i) = \frac{\alpha_k^l p(\mathbf{x}_i|\theta_k^l)}{\sum_{k=1}^K \alpha_k^l p(\mathbf{x}_i|\theta_k^l)} \quad (3)$$

M 步:更新算法参数

$$\alpha_k^{l+1} = \frac{\sum_{i=1}^N \beta_{ik}^l}{N} \quad (4)$$

$$\mu_k^{l+1} = \frac{\sum_{i=1}^N \beta_{ik}^l \mathbf{x}_i}{\sum_{i=1}^N \beta_{ik}^l} \quad (5)$$

$$\Sigma_k^{l+1} = \frac{\sum_{i=1}^N \beta_{ik}^l (\mathbf{x}_i - \mu_k^l)^T (\mathbf{x}_i - \mu_k^l)}{\sum_{i=1}^N \beta_{ik}^l} \quad (6)$$

1.2 贝叶斯信息准则

利用 GMM 对数据进行聚类时, 需要对高斯成分个数进行确定, 如果高斯成分选取较少, 则模型易出现欠拟合的情况, 成分个数过多又会出现过拟合现象, 限制了算法的应用范围。针对这一问题, 一些学者提出了多种确定 GMM 成分个数的方法, 其中常用的有 Akaike 信息准则^[12] (Akaike information criterion, AIC) 和贝叶斯信息准则^[13] (BIC) 等。AIC 可由如下公式来描述:

$$\text{AIC} = -2 \log p(\mathbf{X} | \Theta) + 2d \quad (7)$$

$$\log p(\mathbf{X} | \Theta) = \log \prod_{i=1}^N \sum_{k=1}^K \alpha_k p(\mathbf{x}_i | \theta_k) = \sum_{i=1}^N \log \sum_{k=1}^K \alpha_k p(\mathbf{x}_i | \theta_k) \quad (8)$$

式中: $\log p(\mathbf{X} | \Theta)$ 表示训练样本的对数似然函数; d 表示 K 个高斯成分所具有的自由参数的个数。在文献[14]中 Mehrjou 等提出式(7)中第 2 项用 $2d$ 表示不够准确。因此, 用 BIC 来代替 AIC, BIC 的描述如下:

$$\text{BIC} = -2 \log p(\mathbf{X} | \Theta) + d \log(N) \quad (9)$$

式中: N 表示训练样本的个数。且 Mehrjou 的实验也表明 BIC 确定高斯成分数量始终优于 AIC^[14]。因此, 本文采用 BIC 来确定最优的 GMM 成分个数, 目的是使 BIC 的值最小, 此时高斯成分的个数即为最优值。

2 基于 JITL 的 GPR 建模

2.1 高斯过程回归

GPR 是一种非参数概率模型, 已经成为机器学习领域的一种重要的方法^[15], 其基本原理描述如下。

已知训练样本集 $\mathbf{X} \{\mathbf{x}_i \in R^m, i=1, 2 \dots N\}$ 和 $\mathbf{Y} \{y_i \in R, i=1, 2 \dots N\}$ 分别代表 m 维输入数据和 1

维输出数据。输入和输出之间的关系可以表示为:

$$y_i = f(\mathbf{x}_i) + \varepsilon \quad (10)$$

式中: f 表示一种未知的函数形式; ε 表示均值为 0, 方差为 σ_N^2 的白噪声。

对于新的测试样本 \mathbf{x}_q , 则它的输出预测值 y_q 也满足高斯分布, 其均值和方差可以分别表示为:

$$y_q(\mathbf{x}_q) = \mathbf{c}^T(\mathbf{x}_q) \mathbf{C}^{-1} \mathbf{Y} \quad (11)$$

$$\sigma_{y_q}^2(\mathbf{x}_q) = \mathbf{c}(\mathbf{x}_q, \mathbf{x}_q) - \mathbf{c}^T(\mathbf{x}_q) \mathbf{C}^{-1} \mathbf{c}(\mathbf{x}_q) \quad (12)$$

式中: $\mathbf{c}(\mathbf{x}_q) = [c(\mathbf{x}_q, \mathbf{x}_1), \dots, c(\mathbf{x}_q, \mathbf{x}_N)]^T$ 是测试输入数据与训练输入数据的协方差矩阵; $\mathbf{C} = \Sigma + \sigma_N^2 \mathbf{I}$ 为训练输入数据之间的协方差矩阵; $\mathbf{c}(\mathbf{x}_q, \mathbf{x}_q)$ 表示测试输入数据与本身的协方差值。

GPR 可以选择不同的协方差函数, 本文选择径向基协方差函数, 其函数描述如下:

$$c(\mathbf{x}_i, \mathbf{x}_j) = \nu \exp \left[-\frac{1}{2} \sum_{i=1}^m \omega_i (\mathbf{x}_{ii} - \mathbf{x}_{jj})^2 \right] + \sigma_N^2 \delta_{ij} \quad (13)$$

式中: ν 表示先验知识的总体度量; ω_i 表示每维数据相对应的权重; δ_{ij} 为 Kronecher 算子, 表示各辅助变量的相对重要程度。

一般用极大似然估计得到式(13)中的参数 $\theta = [\nu, \sigma_n^2, \omega_1, \dots, \omega_D]$, 其对数似然函数为:

$$L(\theta) = -\frac{1}{2} \log(\det(\mathbf{C})) - \frac{1}{2} \mathbf{Y}^T \mathbf{C}^{-1} \mathbf{Y} - \frac{N}{2} \log(2\pi) \quad (14)$$

先将参数 θ 设置为一个合理的初值, 然后用共轭梯度法得到优化的参数^[6]。参数确定后, 对于新的测试数据, 可由式(11)得到软测量模型输出。

2.2 JITL-GPR 建模方法

虽然 GPR 方法能对非线性工业过程进行比较准确的描述, 但对于一些工业过程具有时变特性, 会出现模型性能退化的情况, JITL 方法通过不断地更新局部模型可以有效的解决上述问题。

JITL 方法是根据相似输入产生相似输出的思想, 从训练样本中选择与当前到来的测试样本最相似的一组训练样本来建模。JITL 的核心是相似度准则的选取, 基于欧式距离和角度的相似度准则是一种常用的方法, 即:

$$SI = \gamma \sqrt{\exp(-d^2)} + (1-\gamma) \cos \theta \quad (15)$$

式中: 距离 d 表示当前到来的测试样本与训练样本之间的 2 范数; θ 表示这 2 个样本之间的夹角, γ 为一系数, 取值在 0 到 1 之间。

然而, 对于一些非高斯工业过程, GMM 可以较好地对过程的非高斯性进行描述, 相比于传统的相似度准则, 基于 BGMM 的相似度准则可以更好地选择相似样本来建立 GPR 模型。由第 2 节得到了最优的高斯成分个数 K 和各个成分的参数 Θ , 对应的相似度准则可以表示为:

$$SI(\mathbf{x}_q, \mathbf{x}_i) = \sum_{k=1}^K p(C_k | \mathbf{x}_q) \exp[-\sqrt{(\mathbf{x}_q - \mathbf{x}_i) \Sigma_k^{-1} (\mathbf{x}_q - \mathbf{x}_i)^T}] \quad (16)$$

式中: \mathbf{x}_q 表示新到来的样本; \mathbf{x}_i 表示第 i 个训练样本; $p(C_k | \mathbf{x}_q)$ 表示新到来的样本 \mathbf{x}_q 属于第 k 个高斯成分的后验概率, $(\mathbf{x}_q - \mathbf{x}_i) \Sigma_k^{-1} (\mathbf{x}_q - \mathbf{x}_i)^T$ 为 2 样本之间的马氏距离。针对当前到来的 \mathbf{x}_q , 利用上述相似度准则, 选择与 \mathbf{x}_q 最相似的一组数据建立当前的 GPR 模型。

3 基于 BGMM 的 JITL-GPR 的建模步骤

本文软测量建模方法的流程图如图 1 所示。

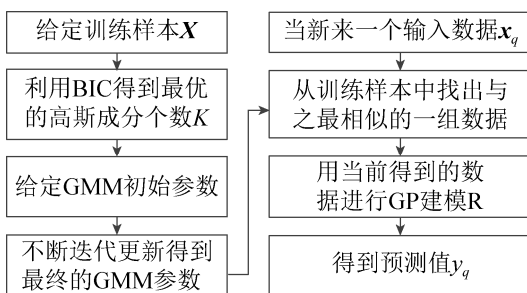


图 1 基于 BGMM 的 JITL-GPR 的流程图
Fig. 1 Flow chart of JITL-GPR based on BGMM

具体建模步骤如下:

(1) 已知训练样本 X , 利用 BIC 确定最优的高斯成分个数 K 。

(2) 给定 GMM 的初始参数, 并利用式(4)–(6) 不断迭代, 直到前后两次参数的差值小于设定好的

阈值, 得到 GMM 的参数。

(3) 当来到一个新的输入数据 \mathbf{x}_q , 利用公式(16)从训练样本中选择与之相似度最高的一组样本, 建立局部的 GPR 模型。

(4) 利用步骤(3)建立的局部模型对 \mathbf{x}_q 进行预测, 得到对应的预测输出 y_q 。

4 仿真实验

脱丁烷塔是石油炼制生产过程中脱硫和石脑油分离装置的重要组成部分^[16], 其中脱丁烷塔底的丁烷浓度对石油炼制的质量起到决定性作用, 因此对塔底丁烷浓度的实时监测和控制是非常有必要的, 但是由于塔底丁烷浓度不能直接通过在线仪器测量得到, 只能通过人工分析方法得到, 人工分析往往需要花费大量的时间, 软测量成为一种解决该问题的有效方法。脱丁烷塔的工艺过程如图 2 所示, 其中 $x_1 \sim x_7$ 为选取的辅助变量, 其描述如表 1 所示, 主导变量为脱丁烷塔塔底丁烷浓度。预测结果采用均方根误差(RMSE)作为衡量指标, 计算公式如下:

$$RMSE = \sqrt{\frac{\sum_{j=1}^M (y_j - y_j^*)^2}{M}} \quad (17)$$

式中: y_j 表示第 j 测试输出数据的真实值; y_j^* 表示第 j 测试输出数据的预测值; M 表示测试样本的个数。

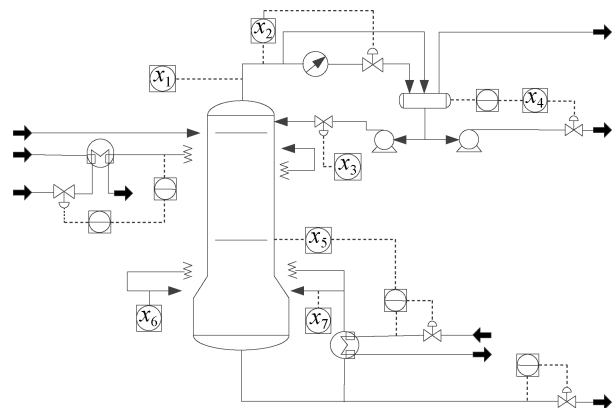


图 2 脱丁烷塔的工艺过程
Fig. 2 Process of deutanizer

表 1 脱丁烷塔的输入变量及其描述
Tab. 1 Input variables of the debutanizer and its description

输入变量	意义	输入变量	意义
x_1	塔顶温度	x_5	第 6 层塔板温度
x_2	塔顶压力	x_6	塔底温度 1
x_3	塔顶回流量	x_7	塔底温度 2
x_4	塔顶产品流出量		

从脱丁烷塔的数据库中选取了 1 400 组数据, 其中 700 组作为训练样本, 700 组作为测试样本。为了进一步分析所提算法的性能, 分别采用 SVM、ANN, GPR 3 种建模方法对脱丁烷塔塔底丁烷浓度进行建模预测, 3 种方法的预测效果如图 3 所示。从图中可以看出, 与 SVM、ANN 相比, GPR 建模方法的预测结果更逼近于真实值。3 种方法所对应的 RMSE 如表 2 所示, 从中也可以看出本文所用的 GPR 方法得到了更高的预测精度。

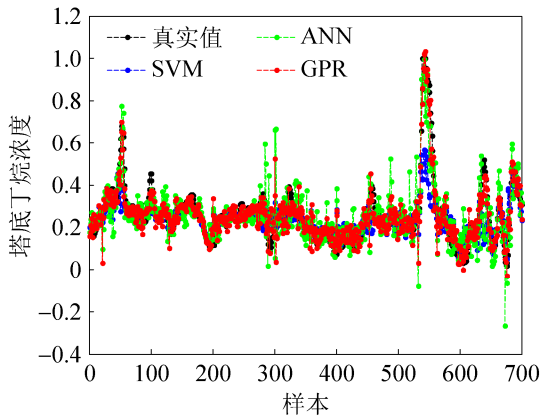


图 3 3 种方法的预测值
Fig. 3 Predicted values of three methods

表 2 3 种方法的 RMSE
Tab. 2 RMSE of three methods

方法	SVM	ANN	GPR
RMSE	0.1028	0.0909	0.0557

在确定用 GPR 建模后, 用 BIC 确定最优的高斯成分个数。BIC 的值与高斯成分个数的关系如图 4 所示。

由图 4 可知, 当高斯成分个数 K 为 4, BIC 的值最小, 说明 K 为 4 是最优的高斯成分个数。因此, 将训练数据划分为 4 类并得到其参数值。当新的测试样本到来时, 根据本文所提相似度准则选取

相似度最大的一部分数据建立局部的 GPR 模型, 并对新到来的样本进行预测, 得到模型的预测值。局部数据数量的选取对模型的预测精度起到至关重要的作用, 这里, 选取原始训练样本的 10%~100% (步长为 10%) 分别建立实时模型, 观察在不同数据比例下的预测效果。为了更好地对预测结果进行比较分析, 选择基于欧式距离和角度的 JITL 与所提方法进行比较。2 种方法在不同数据比例下的 RMSE 如图 5 所示。

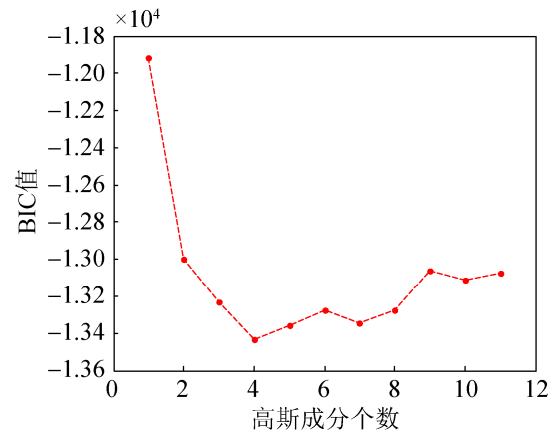


图 4 不同的高斯成分对应的 BIC 值
Fig. 4 BIC Values of Different Gauss Components

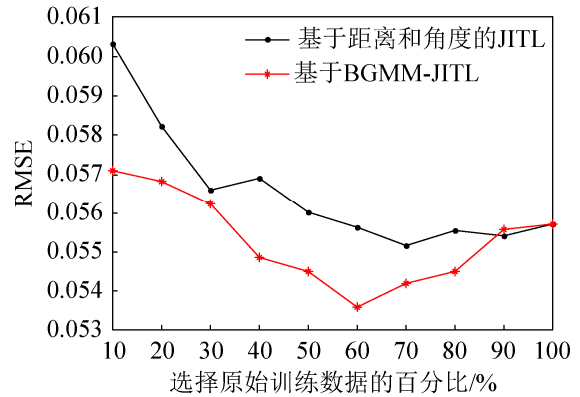


图 5 选取训练样本的不同比例建模的 RMSE
Fig. 5 RMSE of the training samples with different proportions

从图 5 中可以看出, 当选取的数据比例为 100% 时, 2 种方法的预测精度是一样的, 这是因为在这种情况下, 2 种方法都相当于普通的 GPR 建模。基于距离和角度的 JITL-GPR 相比于普通的 GPR 对预测精度的提升很有限, 而本文所提方法

对预测精度的提升比较明显。这是由于该方法充分考虑了预测样本的非高斯性,通过建立高斯混合模型,提取过程的非高斯特性,更加合理地对相似样本进行选取,从而提高了模型预测精度。

2 种方法在数据比例为 50%~80%下的 RMSE 如表 3 所示。

表 3 2 种方法在比例为 50%~80%下的 RMSE

Tab. 3 The RMSE of two methods at a ratio of 50%~80%

方法\比例	50%	60%	70%	80%
距离角度 JITL	0.056 0	0.055 6	0.055 2	0.055 5
BGMM- JITL	0.054 5	0.053 6	0.054 2	0.054 5

当选取样本的百分比为 60%时,取出 100 组数据比较两种方法的预测误差,预测误差如图 6 所示。

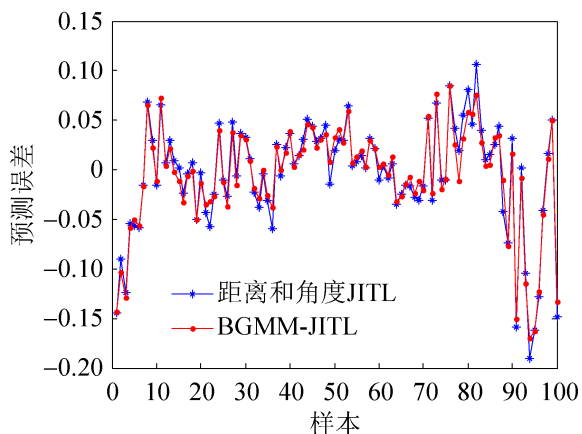


图 6 2 种方法的预测误差

Fig. 6 Prediction error of two methods

从图 6 中可以看出,相比于距离和角度的 JITL-GPR,本文所提方法的预测误差更小,精度更高。为了验证 BIC 的有效性,分别选取高斯成分个数为 2, 3, 4 和 5 建立局部的 GPR 模型;并选取训练样本的不同比例观察不同高斯成分个数对建模精度的影响。不同高斯成分个数在不同数据比例下的 RMSE 如图 7 所示。可以看出,当 K 为 4 时,且选取训练样本数目的 60%时,所得到的预测精度最高;从比例 50%~80%也可以看出 K 为 4 时,建模精度最高。

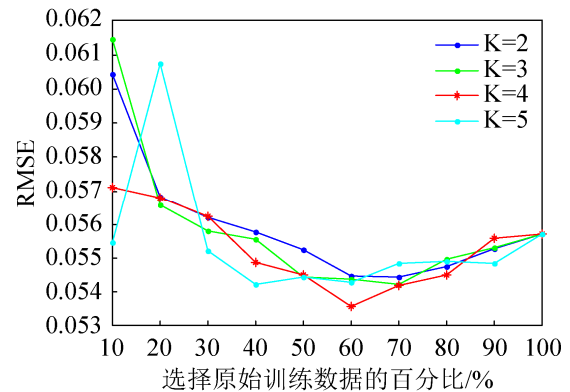


图 7 不同高斯成分在不同比例下的 RMSE

Fig. 7 RMSE of different Gaussian components and different proportions

综上所述,通过 BIC 优化的高斯混合模型对于确定合理的高斯成分个数是有效的,基于此建立的即时学习软测量模型相对于普通的 JITL-GPR 方法预测精度有了很大的提高。上述仿真结果充分验证了所提 BGMM-JITL-GPR 方法的有效性。

5 结论

从时变工业过程的非高斯性、非线性特征出发,提出一种基于 BGMM 的 JITL-GPR 建模方法,通过不断地更新模型,有效地克服了因工况变化而导致的模型退化的问题;同时,优化的 GMM 可以很好的表现过程的非高斯性,相比基于距离和角度的相似度准则,基于 BGMM 的相似度准则能更合理地选取相似数据进行建模。通过脱丁烷塔塔底丁烷浓度的软测量建模研究,验证了其相对于传统 GPR 和基于距离和角度的 JITL-GPR 有更高的精度,对有明显非高斯性、非线性工业过程的软测量建模问题提供了一种可行的思路。

参考文献:

- [1] Liu J, Chen D S, Shen J F. Development of self-validating soft sensors using fast moving window partial least squares[J]. Industrial & Engineering Chemistry Research (S0888-5885), 2010, 49(22): 11530-11546.
- [2] Choi D J, Park H. A hybrid artificial neural network as a software sensor for optimal control of a wastewater treatment process[J]. Water research (S0043-1354),

- 2001, 35(16): 3959-3967.
- [3] Desai K, Badhe Y, Tambe S S, et al. Soft-sensor development for fed-batch bioreactors using support vector regression[J]. *Biochemical Engineering Journal* (S1369-703X), 2006, 27(3): 225-239.
- [4] Ge Zhiqiang. Active probabilistic sample selection for intelligent soft sensing of industrial processes[J]. *Chemometrics & Intelligent Laboratory Systems* (S0169-7439), 2016, 151: 181-189.
- [5] Rasmussen C E. Gaussian processes for machine learning[J]. *Lecture Notes in Computer Science* (S0129-0657), 2004, 3176: 63-71.
- [6] 何志昆, 刘光斌, 赵曦晶, 等. 高斯过程回归方法综述[J]. *控制与决策*, 2013, 28(8): 1121-1129.
He Zhikun, Liu Guangbin, Zhao Xijing, et al. Overview of Gaussian process regression[J]. *Control and Decision*, 2013, 28(8): 1121-1129.
- [7] 王春鹏, 于佐军, 孟凡强. 折息移动窗递推 PLS 算法及其在聚丙烯生产过程中的应用[J]. *化工学报*, 2013, 64(12): 4592-4598.
Wang Chunpeng, Yu Zuojun, Meng Fanqiang. Discount moving window recursive PLS algorithm and its application to process of polypropylene production[J]. *Journal of Chemical Industry and Engineering*, 2013, 64(12): 4592-4598.
- [8] Ge Z, Song Z. A comparative study of just-in-time-learning based methods for online soft sensor modeling [J]. *Chemometrics and Intelligent Laboratory Systems* (S0169-7439), 2010, 104(2): 306-317.
- [9] 李庆良, 雷虎民. 一种基于即时学习的多模型在线建模方法[J]. *四川大学学报(工程科学版)*, 2010, 42(1): 196-200.
Li Qingliang, Lei Humin. An Online Multiple-model Modeling Method Based on Lazy Learning[J]. *Journal of Sichuan University(Engineering Science Edition)*, 2010, 42(1): 196-200.
- [10] Witten I H, Frank E, Hall M A, et al. *Data Mining: Practical machine learning tools and techniques*[J]. *BioMedical Engineering Online* (S1475-925X), 2006, 5(1).
- [11] 李昌利, 沈玉利. 期望最大算法及其应用(1)[J]. *计算机工程与应用*, 2008, 44(29): 61-64.
Li Changli, Shen Yuli. Tutorial of EM algorithm and its application(1)[J]. *Computer Engineering and Applications*, 2008, 44(29): 61-64.
- [12] Shimo-onoda K, Tanaka T, Furushima K, et al. Akaike's information criterion for a measure of linkage disequilibrium[J]. *Journal of human genetics* (S1434-5161), 2002, 47(12): 0649-0655.
- [13] Zanini A, Woodbury A D. Contaminant source reconstruction by empirical Bayes and Akaike's Bayesian Information Criterion[J]. *Journal of contaminant hydrology*, 2016, 185: 74-86.
- [14] Mehrjou A, Hosseini R, Araabi B N. Improved Bayesian information criterion for mixture model selection[J]. *Pattern Recognition Letters* (S0167-8655), 2016, 69: 22-27.
- [15] Han J, Zhang X P, Wang F. Gaussian Process Regression Stochastic Volatility Model for Financial Time Series[J]. *IEEE Journal of Selected Topics in Signal Processing* (S1941-0484), 2016, 10(6): 1015-1028.
- [16] 阮宏镁, 田学民, 王平. 基于联合互信息的动态软测量方法[J]. *化工学报*, 2014, 65(11): 4497-4502.
Ruan Hongmei, Tian Xuemin, Wang Ping. Dynamic soft sensor method based on joint mutual information[J]. *Journal of Chemical Industry and Engineering*, 2014, 65(11): 4497-4502.