

12-12-2019

GPR Modeling Method Based on Kernel Slow Feature Analysis and Time Delay Estimation

Huilai Peng

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

Weili Xiong

1. College of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China; ;2. Key Laboratory of Advanced Process Control for Light Industry (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.

GPR Modeling Method Based on Kernel Slow Feature Analysis and Time Delay Estimation

Abstract

Abstract: A Gaussian process regression modeling method based on kernel slow feature analysis is proposed to deal with the time delay and nonlinear characteristics in industrial processes. *The time delay in the industrial data is effectively extracted by fuzzy curve analysis and the optimal time-delay in the data is obtained. The model data is reconstructed based on the optimal delay. The method of kernel slow feature analysis is used to extract the nonlinear features of the reconstructed data. The Gaussian process regression model is established based on the extracted features.* The effectiveness and performance of the method are verified by the simulation experiment of the soft measurement of butane concentration at the bottom of the de-butane tower.

Keywords

time delay, kernel slow feature analysis, fuzzy curve analysis, Gaussian process regression

Recommended Citation

Peng Huilai, Xiong Weili. GPR Modeling Method Based on Kernel Slow Feature Analysis and Time Delay Estimation[J]. Journal of System Simulation, 2019, 31(8): 1562-1571.

基于核慢特征分析和时滞估计的 GPR 建模

彭慧来¹, 熊伟丽^{1,2}

(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: 针对工业过程中存在的时滞和非线性特性, 提出了一种基于核慢特征分析的高斯过程回归建模方法。该方法通过模糊曲线分析充分挖掘工业数据中的时滞信息, 求出数据中的最优时滞, 并进行建模数据的重构; 通过核慢特征分析方法对重构数据进行非线性的特征提取; 基于提取后的特征建立高斯过程回归(GPR)模型。通过对脱丁烷塔塔底丁烷浓度软测量的仿真实验, 验证了方法的有效性与性能。

关键词: 时延; 核慢特征分析; 模糊曲线分析; 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) 08-1562-10

DOI: 10.16182/j.issn1004731x.joss.17-0389

GPR Modeling Method Based on Kernel Slow Feature Analysis and Time Delay Estimation

Peng Huilai¹, Xiong Weili^{1,2}

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

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

Abstract: A Gaussian process regression modeling method based on kernel slow feature analysis is proposed to deal with the time delay and nonlinear characteristics in industrial processes. *The time delay in the industrial data is effectively extracted by fuzzy curve analysis and the optimal time-delay in the data is obtained. The model data is reconstructed based on the optimal delay. The method of kernel slow feature analysis is used to extract the nonlinear features of the reconstructed data. The Gaussian process regression model is established based on the extracted features.* The effectiveness and performance of the method are verified by the simulation experiment of the soft measurement of butane concentration at the bottom of the de-butane tower.

Key words: time delay; kernel slow feature analysis; fuzzy curve analysis; Gaussian process regression

引言

在实际工业过程中, 一些关键变量的测量对于生产高质量的产品至关重要。但是在现有技术条件和经济代价等问题的约束下, 使得直接获取关键变

量变得十分困难。因此, 在这样的背景下, 软测量技术应运而生^[1-2]。常见的软测量建模方法如偏最小二乘^[3]、神经网络^[4]、最小二乘支持向量机^[5]等都可以取得良好的预测结果。高斯过程回归(Gaussian process regression, GPR)^[6-7]由于其具有预测精度高, 超参数容易确定, 输出具有概率意义等优点, 近几年被广泛应用于软测量建模。因此选取高斯过程回归作为本文的建模方法。

建模过程中, 一般都会对过程数据预先进行特征提取。常用的特征提取方法有: 线性的主成



收稿日期: 2017-08-09 修回日期: 2017-11-04;
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: 彭慧来(1991-), 男, 江西上饶, 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, 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>

• 1562 •

分分析^[8]和独立主元分析^[9], 以及非线性的核主元分析^[10], 但它们仅从变量的角度出发去构造新的成分, 而并未考虑样本之间的变化情况。慢特征分析^[11](slow feature analysis, SFA)作为一种新的特征提取方法, 能够从样本的角度出发, 分析样本的变化情况, 提取出一部分最缓慢的特征代表过程的本质特征。近年来 SFA 得到了各个研究领域的关注, 并且取得了广泛的应用, 例如, 盲信号分离^[12]、软测量建模^[13-14]、行为识别^[15]、故障诊断^[16]以及遥感^[17]。然而现有的慢特征分析方法多是基于多项式扩展, 其处理非线性的能力较弱。当数据的维度较大时, 多项式扩展会引起维数灾难, 也有可能造成数据的过拟合。针对工业过程中的强非线性和高维特征, 通过引入核函数扩展的核慢特征分析^[18-19](kernel slow feature analysis, KSFA)代替多项式扩展的慢特征分析, 加强了对非线性数据的特征提取能力, 从而实现对工业过程中的非线性特征信息的有效提取。

实际的生产过程中输入输出数据会存在显著的时间滞后, 如物料流动引起的传输时延、信号传输带来的纯时延等。若对这些时滞置之不理, 会导致输入和输出变量之间的因果关系发生变化, 从而导致建模精度的下降。因此, 为了确保建立的软测量模型能够在较长时间内实现输出变量的精确预测, 有必要引入过程中的时滞信息。

如何准确估计输入和输出之间的时延变得尤为重要。因为实际工业过程工艺复杂, 所以根据过程工艺进行时延估计^[20]是十分困难的。Zhang 等^[21]通过构建输入变量和输出变量的相关系数来估计时延, 由于只考虑了变量之间的线性关系, 对于非线性工业过程建模可能得不到理想的预测精度。阮宏镁等^[22]通过差分估计(difference estimation, DE)算法优化过程变量间的联合互信息来确定时延, 但是智能优化算法容易陷入局部最优, 且联合互信息分析相关性时计算复杂度较高, 运行时间较长。模糊曲线分析(fuzzy curve analysis, FCA)方法^[23]最早用于确定对输出变量重要的输入变量。本文通过引

入不同的时滞信息扩展输入变量, 通过 FCA 构造输入变量和输出变量之间的模糊逻辑, 从而确定对输出变量最重要的输入变量所对应的时滞, 得到各输入变量对应的最优时滞信息。这种方法的优点在于计算复杂度较低, 能够直观有效准确的确定出各输入变量的最优时滞信息。

综合上述分析, 同时考虑实际工业过程中的时延和非线性特性, 提出了基于核慢特征分析和时滞估计的高斯过程回归建模方法。该方法通过 FCA 挖掘出各辅助变量的最优时滞, 进行数据重构; 并且对重构的数据进行核慢特征分析, 基于提取出的非线性特征进行高斯过程回归软测量建模。最后, 通过脱丁烷塔数据建立的软测量模型对本文方法进行了有效性验证。

1 基于核慢特征分析的 GPR 建模

1.1 慢特征分析(SFA)描述

SFA 算法^[11]的目标是从大量的时序输入数据中, 提取出一部分变化最缓慢的成分来表征过程变量中蕴含的固有属性, 其数学描述如下:

给定一个 p 维的时序输入信号 $\mathbf{X}(t)=\{x_1(t), x_2(t), \dots, x_p(t)\}$, SFA 算法旨在找到一个 q 维的变换函数 $\mathbf{G}(\mathbf{x})=\{g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_q(\mathbf{x})\}$, 使得 q 维的输出信号 $\mathbf{S}(t)=\{S_1(t), S_2(t), \dots, S_q(t)\}$ 变化尽可能缓慢。其中 $S_j(t)=g_j(\mathbf{X}(t))$, $j \in [1, 2, \dots, q]$ 。SFA 算法最终转化为一个优化问题:

$$\min \Delta(S_j(t)) := \langle \dot{S}_j^2 \rangle_t \quad (1)$$

且满足约束条件:

$$\langle S_j \rangle_t = 0 \quad (2)$$

$$\langle (S_j)^2 \rangle_t = 1, \quad (3)$$

$$\langle S_i S_j \rangle_t = 0, \forall i \neq j, \quad (4)$$

式中: \dot{S}_j 为 S_j 的一阶导数; $\langle \cdot \rangle_t$ 为时间平均。

式(1)为优化目标函数, 使得输出信号变化最为缓慢; 式(2)表示的零均值约束, 仅仅是为了简化问题的求解; 单位方差约束(3)不但对输出信号的尺度进行了归一化, 还避免了常值解的出现, 使

输出信号必须包含一定的信息；式(4)保证了输出信号的各个分量之间是不相关的，避免了冗余信号的出现。除此之外，慢特征根据其变化快慢进行排序，若按升序排序，即 $S_1(t)$ 是最缓慢的特征， $S_2(t)$ 是次缓慢的特征，以此类推。

当变换函数 $g_j(\bullet)$ 为线性时，此时每个慢特征 S_i 都是输入变量的线性组合，即：

$$S_j = g_j(\mathbf{X}) = \mathbf{X}(t)\mathbf{w}_j \quad (5)$$

由于 $\mathbf{X}(t)$ 已经预先进行了均值化处理，因此满足式(2)的约束条件。式(1)优化目标和约束条件(3)可以重写为：

$$\begin{aligned} \min \Delta(S_j(t)) &= \min \langle \dot{S}_j^2 \rangle_t = \\ &= \min \langle (\dot{\mathbf{X}}(t) \cdot \mathbf{w}_j)^2 \rangle_t = \min \mathbf{w}_j^T \mathbf{A} \mathbf{w}_j \quad (6) \end{aligned}$$

$$s.t. \langle (S_j)^2 \rangle_t = \langle (\mathbf{X}(t) \cdot \mathbf{w}_j)^2 \rangle_t = \mathbf{w}_j^T \mathbf{B} \mathbf{w}_j = 1$$

式中： $\mathbf{A} = \langle \dot{\mathbf{X}}(t)^T \dot{\mathbf{X}}(t) \rangle_t$ ， $\mathbf{B} = \langle \mathbf{X}(t)^T \mathbf{X}(t) \rangle_t$ ，进而公式(6)的优化求解问题等价于如下的广义特征值分解问题：

$$\mathbf{A} \mathbf{W} = \mathbf{B} \mathbf{W} \mathbf{\Lambda} \quad (7)$$

式中： $\mathbf{\Lambda} = (\lambda_1, \dots, \lambda_q)$ 为广义特征值构成的对角矩阵且 $\lambda_1 < \lambda_2 < \dots < \lambda_q$ ， $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_p)$ 为对应的广义特征向量矩阵。式(1)的优化目标正是对角矩阵 $\mathbf{\Lambda}$ 中的主对角元素。

$$\Delta(S_j(t)) = \frac{\mathbf{w}_j^T \mathbf{A} \mathbf{w}_j}{\mathbf{w}_j^T \mathbf{B} \mathbf{w}_j} = \frac{\mathbf{w}_j^T \lambda_j \mathbf{B} \mathbf{w}_j}{\mathbf{w}_j^T \mathbf{B} \mathbf{w}_j} = \lambda_j \quad (8)$$

当变换函数 $g_j(\bullet)$ 为非线性时，此时的变换函数可以等效为线性变换函数在非线性空间的扩展，多项式扩展为常用的扩展方式，以二阶多项式扩展为例^[11]，定义如下：

$$\tilde{\mathbf{V}}(\mathbf{x}) = [x_1, \dots, x_p, x_1 x_1, x_1 x_2, \dots, x_p x_p] \quad (9)$$

经过二阶扩展后， $\tilde{\mathbf{V}}(\mathbf{x})$ 的维度由 p 维变成了 L 维，其中 $L=p+p(p+1)/2$ 。将 $\tilde{\mathbf{V}}(\mathbf{x})$ 进行均值化处理后得 $\mathbf{V}(\mathbf{x})$ ，将 $\mathbf{V}(\mathbf{x})$ 替换式(5)中的 $\mathbf{X}(t)$ ，就将非线性情况转化成了线性情况。

通过对式(7)求取广义特征值，获得相应的特征向量矩阵 \mathbf{W} ，即可求取相应的慢特征 $\mathbf{S}=\mathbf{X}(t) \cdot \mathbf{W}$ ，为了方便计算，式(7)的广义特征值分解通常会被

转化为两次简单的特征值分解(SVD)问题，推导流程如图 1 所示。

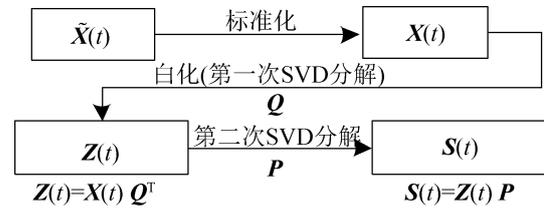


图 1 SFA 流程图
Fig. 1 Flow chart of SFA

1.2 基于核慢特征分析(KSFA)的特征提取

由于实际工业过程存在着非线性特征，为了加强对非线性过程的处理能力，慢特征分析一般会采用多项式扩展的方法，但当数据的维度较大时，多项式扩展往往会导致维数灾难，严重影响算法的速度和实际建模的精度。同时，采用多项式扩展的方法并不能显著提高算法的精度，特别是对于一些高维度且存在较强非线性特征的工业过程。

核慢特征分析通过采用核函数扩展代替多项式扩展，从而加强对非线性数据的处理。通过核函数变换将原始的低维数据映射到高维空间 φ ： $\mathbf{x} \rightarrow \varphi(\mathbf{x})$ ，然后在高维空间里进行数据处理。核函数的选取有多种选择，本文选择泛化和平滑估计能力强的高斯核函数。

$$K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j) = \varphi(i)^T \varphi(j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{c}\right) \quad (10)$$

式中： $\mathbf{x}_i, \mathbf{x}_j$ 为任意的样本； $\varphi(i), \varphi(j)$ 为 $\mathbf{x}_i, \mathbf{x}_j$ 样本在高维空间的映射。

核函数确定后，将核函数扩展方法代替多项式扩展从而加强对非线性数据的处理，从而实现核慢特征分析的非线性特征提取。主要改进在于将经过核函数扩展后的数据 \mathbf{K} 代替二阶多项式扩展的 $\mathbf{V}(t)$ ，优化目标就转换为：

$$\begin{aligned} \min \Delta(S_j(t)) &= \min \langle (\dot{\mathbf{K}} \cdot \mathbf{w}_j)^2 \rangle_t = \\ &= \min \mathbf{w}_j^T \langle \dot{\mathbf{K}}^T \dot{\mathbf{K}} \rangle_t \mathbf{w}_j \quad (11) \end{aligned}$$

同时优化目标求解问题等价于如下的广义特征值分解问题：

$$A_k W = B_k W A \quad (12)$$

式中: $A_k = \langle \dot{K}^T \dot{K} \rangle_t$ 表示 \dot{K} 的协方差阵;
 $B_k = \langle K^T K \rangle_t$ 表示 K 的协方差阵。

通过对式(12)求取广义特征值, 可求出相应的特征向量矩阵 W , 进而求取相应的核慢特征 $S_k = K \cdot W$, 求取的核慢特征 S_k 即可代表从工业过程数据中提取的非线性特征。式(12)的求解, 也可以简化为两次 SVD 分解。

1.3 GPR 软测量模型的建立

传统的核慢特征回归(KSFR)方法^[24], 可以理解为首先通过 KSFA 进行非线性的核慢特征的求取, 然后基于求出的核慢特征进行最小二乘回归建模。然而在存在强非线性的实际工业过程中, 即使采用了 KSFA 进行了非线性的特征提取, 但是采用线性的最小二乘回归建模方法还是难于建立较为准确的模型, 依然会导致建立的模型精度不高。为了避免这种线性建模方法带来的不足, 进一步提升模型的精度, 本文选择 GPR 作为建模方法。

GPR 是基于贝叶斯理论和统计学习理论发展起来的一种机器学习方法, 不仅能够得到预测输出, 而且同时给出预测值的估计方差, 适于处理高维数、小样本和非线性等复杂回归问题。因此本文选取 KSFA 进行非线性特征提取, 然后基于提取的非线性特征进行 GPR 建模, 能够实现更高的建模精度, 更加符合实际的工业过程特点。GPR 算法的描述如下:

给定一个训练样本集 $X \in R^{D \times N}$ 和 $y \in R^N$, 其中 D 代表样本的维数, N 代表样本的个数。对于一个新来的测试样本 x^* 可以由式(13)–(14)分别得到预测输出的均值和方差。

$$y^*(x^*) = k^* C^{-1} y \quad (13)$$

$$\sigma_{y^*}^2(x^*) = k - k^{*T} C^{-1} k^* \quad (14)$$

式中: $k^* = [k(x^*, x_1), \dots, k(x^*, x_N)]^T$ 是训练样本和测试输入之间的 $N \times 1$ 维协方差矩阵; C 是训练样本之间的 $N \times N$ 维协方差矩阵; k 是测试输入和自身的协方差。

本文采用的是平方指数协方差函数^[6]:

$$k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{1}{2}(x_i - x_j)^T M^{-1}(x_i - x_j)\right) + \delta_{ij} \sigma_n^2 \quad (15)$$

式中: 当 $i=j$ 时, $\delta_{ij}=1$, 否则为 0, $M = \text{diag}(l^2)$; l 为方差尺度; σ_f 为信号方差。

协方差函数的确定, 需对其中的未知参数 $\theta = [M, \sigma_f^2, \sigma_n^2]$ 进行估计, 一般可以采用极大似然估计求解:

$$L(\theta) = -\frac{1}{2} \log(\det(C)) - \frac{1}{2} y^T C^{-1} y - \frac{N}{2} \log(2\pi) \quad (16)$$

首先将参数 θ 设置为一个合理范围的随机值, 然后通过共轭梯度法进行寻优, 以获取参数 θ 的最优值。

2 基于模糊曲线分析的变量时滞估计

2.1 模糊曲线分析算法

模糊曲线法(FCA)最早由 Lin 等^[23]提出, 用于选择重要的输入变量。该方法通过构建输入和输出变量之间的模糊逻辑, 从而确定对输出变量重要的输入变量。假设有 m 个输入变量 $\{x_i, i=1, 2, \dots, m\}$ 和输出变量 y , 采集到的样本个数为 n , 令 x_{ik} 为第 k 个样本的第 i 个变量, y_k 为对应的输出变量。对于每个输入变量 $x_i, i=1, 2, \dots, m$ 在第 k 个样本处的模糊隶属度函数 $\phi_{ik}(x_i)$ 定义为:

$$\phi_{ik}(x_i) = \exp\left(-\left(\frac{x_{ik} - x_i}{b}\right)^2\right), k=1, 2, 3, \dots, n \quad (17)$$

式中: ϕ_{ik} 为变量 x_i 所对应的第 k 个样本点的输入变量模糊隶属函数, ϕ_{ik} 这里选取为高斯函数; b 选取为输入变量 x_i 范围的 20%。因此对于 n 个训练样本对应的每个输入变量都有 n 条模糊规则, 其中第 k 个规则描述为: R^k : if x_i is $\phi_{ik}(x_i)$, then y is y_k 。

通过式(18)对输入变量 x_i 进行重心去模糊化, 计算得到其对应的模糊曲线为 c_i 。进一步通过式(19), 可以计算出模糊曲线 c_i 的变化范围 Rc_i 。若得到的 Rc_i 越大, 表明此时输入变量对于输出变量越重要。

$$c_i(\mathbf{x}_i) = \frac{\sum_{k=1}^m \phi_{ik}(\mathbf{x}_i) \cdot y_k}{\sum_{k=1}^m \phi_{ik}(\mathbf{x}_i)} \quad (18)$$

$$Rc_i = \max(c_i) - \min(c_i) \quad (19)$$

由于共有 m 个输入变量，故能得到 m 个 Rc_i ，将 Rc_i 大小按降序排列，可以得到各输入变量的重要性。

2.2 模糊曲线法估计变量时滞

在软测量建模过程中，针对有时延的数据，通过模糊曲线法计算其最优时滞，其关键的一点在于将滞后时间的确定转化为求取重要输入变量的选择^[25]。对于时延过程，通过引入时滞信息，原有的输入变量 \mathbf{x}_i 转变为 $T_{\max}+1$ 维： $\{\mathbf{x}_i(t), \mathbf{x}_i(t-1), \dots, \mathbf{x}_i(t-d), \dots, \mathbf{x}_i(t-T_{\max})\}$ ，其中 t 为输出变量的采样时刻， d 为待确定的时滞变量， T_{\max} 为最大的变量时延值。根据模糊曲线法原理，将 $\{\mathbf{x}_i(t), \mathbf{x}_i(t-1), \dots, \mathbf{x}_i(t-d), \dots, \mathbf{x}_i(t-T_{\max})\}$ 转化为 $T_{\max}+1$ 个待选

输入变量，从中选取一个最重要的变量 $\mathbf{x}_i(t-d)$ ，则 d 即为该输入变量所对应的最优时滞。假设求取出各输入变量的最优时滞分别为 d_1, d_2, \dots, d_m ，则重构后的样本数据如图 2 所示，其中 m, n 分别表示样本维度和样本个数。

$x_1(T_{\max}+1-d_1)$	$x_2(T_{\max}+1-d_2)$...	$x_m(T_{\max}+1-d_m)$
$x_1(T_{\max}+1-d_1+1)$	$x_2(T_{\max}+1-d_2+1)$...	$x_m(T_{\max}+1-d_m+1)$
\vdots	\vdots		\vdots
$x_1(n-d_1)$	$x_2(n-d_2)$...	$x_m(n-d_m)$
$y(T_{\max}+1)$	$y(T_{\max}+1+1)$...	$y(n)$

图 2 重构后的数据
Fig. 2 Reconstructed data

3 基于核慢特征分析和时滞估计的 GPR 建模的具体建模步骤

本文的建模流程如图 3 所示。

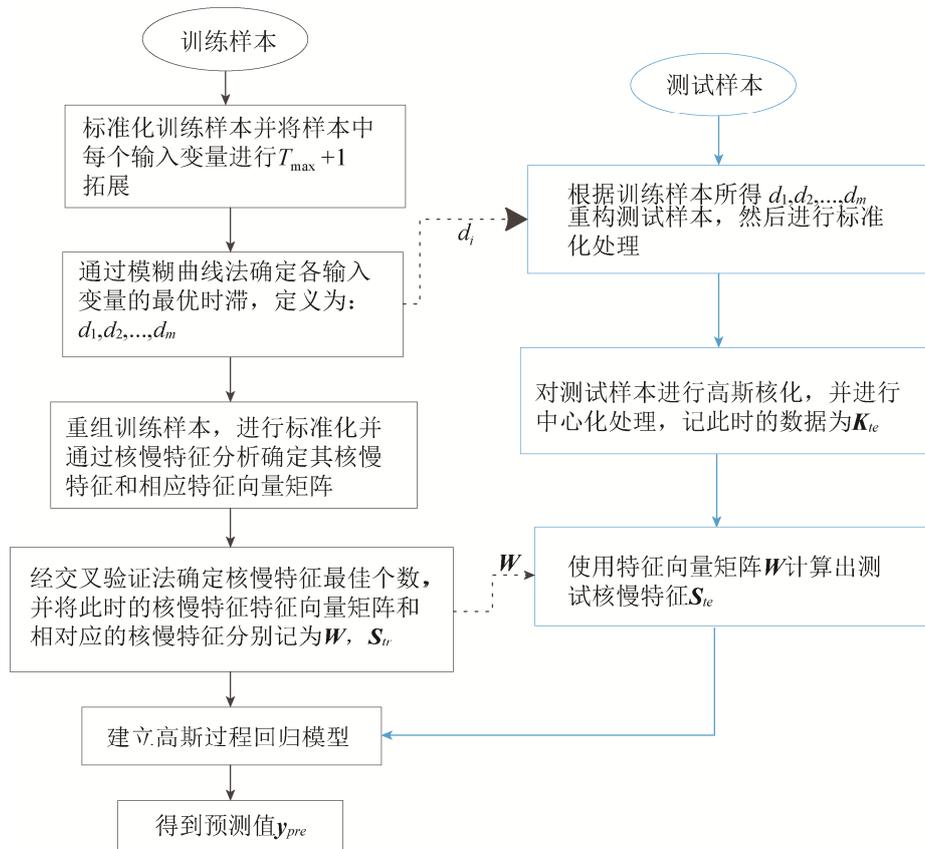


图 3 基于核慢特征分析和时滞估计的 GPR 建模流程图
Fig. 3 Modeling flow chart of FCA-KSFA-GPR method

具体的建模说明如下:

1) 训练部分

step 1: 获取训练样本数据集 $\{\tilde{\mathbf{X}}(t), \mathbf{y}(t)\}$, $t=1,2,3,\dots,N$, 对该数据集进行标准化后, 根据过程机理及经验, 确定出最大时滞估计 T_{\max} 。

step 2: 对于每一个输入变量 x_i , $i=1,2,\dots,m$, 进行 $T_{\max}+1$ 维扩展, 得 $\{x_i(t), x_i(t-1), \dots, x_i(t-d), \dots, x_i(t-T_{\max})\}$, 根据模糊曲线分析法, 由式(19)可确定出最优的时滞变量 $x_i(t-d)$, 将求出的各输入变量的最优时滞记为 $d_1, d_2, d_3, \dots, d_m$, 并按照图 2 所示将训练样本数据进行重构。

step 3: 将重构的训练样本标准化, 并进行高斯函数的核化处理, 将此时数据记为 $\tilde{\mathbf{K}}_{tr}$, 并对其进行中心化处理得 \mathbf{K}_{tr} 。

step 4: 对 \mathbf{K}_{tr} 作白化处理, 即得到 $\mathbf{Z}_k = \mathbf{K}_{tr} \mathbf{Q}_k^T$, 其中 $\mathbf{Q}_k = \mathbf{A}^{-1/2} \mathbf{U}^T$ 为白化矩阵, 可通过对 $\mathbf{B}_k = \langle \mathbf{K}_{tr}^T \mathbf{K}_{tr} \rangle_t$ 进行 SVD 分解得到, 即 $\mathbf{B}_k = \mathbf{U} \mathbf{A} \mathbf{U}^T$, 值得注意的是里有可能存在特征值极小的情况, 本文设置了阈值 $\varepsilon = 1 \times 10^{-6}$ 。当 $\lambda_{si} / \max(\lambda_{si})$ 小于阈值 ε , 则舍去该特征值。

step 5: 定义 $\mathbf{A}_k = \langle \mathbf{Z}_k^T \mathbf{Z}_k \rangle_t$, 对其进行 SVD 分解得 $\mathbf{A}_k = \mathbf{W}_k^T \mathbf{\Omega} \mathbf{W}_k$, 即可得到特征向量 $\bar{\mathbf{W}} = \mathbf{Q}_k^T \mathbf{W}_k$ 和相应的核慢特征 $\bar{\mathbf{S}}_{tr} = \mathbf{K}_{tr} \bar{\mathbf{W}}$ 。

step 6: 采用交叉验证法来确定最佳的核慢特征个数并将此时的特征向量和相应的核慢特征分别记为 \mathbf{W} , \mathbf{S}_{tr} 。

step 7: 通过 $\{\mathbf{S}_{tr}, \mathbf{y}\}$ 建立高斯过程回归模型。

2) 测试部分

step 1: 对于测试样本, 根据训练样本所得的 $d_1, d_2, d_3, \dots, d_m$ 重构测试样本, 并进行标准化, 记此时的数据为 $\tilde{\mathbf{K}}_{te}$ 。

step 2: 对 $\tilde{\mathbf{K}}_{te}$ 作中心化处理, 得 \mathbf{K}_{te} , 利用训练样本所得特征向量 \mathbf{W} , 即可求出测试样本的核慢特征 $\mathbf{S}_{te} = \mathbf{K}_{te} \mathbf{W}$ 。

step 3: 将 \mathbf{S}_{te} 代入训练样本建立好的高斯过程回归模型中, 即可实现对测试样本的预测。

4 仿真研究

脱丁烷塔过程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^[26], 主要任务是减少脱丁烷塔底部的丁烷浓度。该过程共有 7 个辅助变量分别为: x_1 顶层温度; x_2 顶层压力; x_3 回流流量; x_4 塔顶产品流出量; x_5 第 6 塔板温度; x_6 底层温度 1; x_7 底层温度 2 和 1 个主导变量脱丁烷浓度。整个样本集共有 2394 个样本, 采样时间为 12 min, 本文选取脱丁烷塔前 1000 组样本作为建模数据。脱丁烷塔 C_4 浓度分析仪的安装位置, 使得 C_4 浓度大约存在 30~75 min 的滞后, 因此辅助变量和主导变量存在大约 45~90 min 的滞后^[20], 则时延参数 d_i 的范围大约为 4~8, 仿照参考文献[22]的做法, 设置辅助变量的最大时延为 19, 共得 140 个候选变量。

本文选取了 FCA 对输入变量确定最佳时滞, 由 FCA 可得 7 个辅助变量所对应的最优时滞为 [7,8,6,16,6,6,0]。以脱丁烷塔的前 6 个辅助变量为例, 图 4 给出了如式(19)所示不同时延所对应的 R_c 的值。由图可知对前 6 个辅助变量, 当时滞分别为 [7,8,6,16,6,6,0] 时, 对应的 R_c 最大, 表明进行时滞估计后的变量与原变量相比, 对于主导变量的贡献性更大, 同时和主导变量的因果性更强。

图 5 中的(a)和(b)分别展示了训练样本和测试样本经过 KSFA 提取后的 4 个变化最缓慢的核慢特征, 可以发现 4 个核慢特征变化情况都不同。第 1 个核慢特征时变化最缓慢, 且越往后的慢特征变化越快。通过比较训练样本和测试样本的核慢特征, 可以发现两者的变化趋势基本保持一致, 进一步表明 KSFA 可以有效地提取变化缓慢的信息。

为了进一步研究本文方法的有效性, 将几种不同的建模方法进行了仿真结果对比分析。方法 1 为传统 GPR 方法; 方法 2 采用考虑时滞信息的 GPR 方法(简记为 F-GPR); 方法 3 采用考虑时滞信息, 并通过二阶多项式扩展的 SFA 特征提取的 GPR 方法(简记为 FS-GPR); 本文方法为考虑时滞信息,

通过 KSFA 特征提取的 GPR 方法(简记为 FKS-GPR)。本文方法的高斯核宽度, 通过多次试验设置为 5.1。为了更好的对预测结果进行评价, 本文选取均方根误差 (RMSE) 和跟踪性能^[7] (TP) 作为性能指标, 其定义为:

$$RMSE = \sqrt{\sum_i^N (y_i - \hat{y}_i)^2 / N} \quad (20)$$

$$TP = 1 - \sigma_{error}^2 / \sigma_{property}^2 \quad (21)$$

式中: y_i 表示真实值; \hat{y}_i 表示预测值; N 为测试样本的个数; σ_{error}^2 表示真实值和预测值之间的误差方差; $\sigma_{property}^2$ 为真实值的方差。

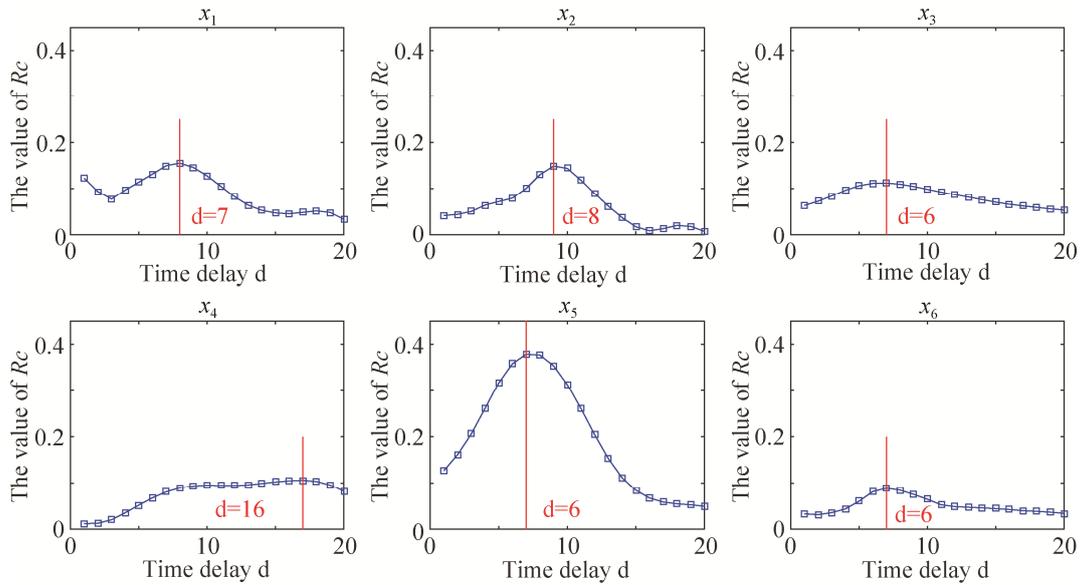
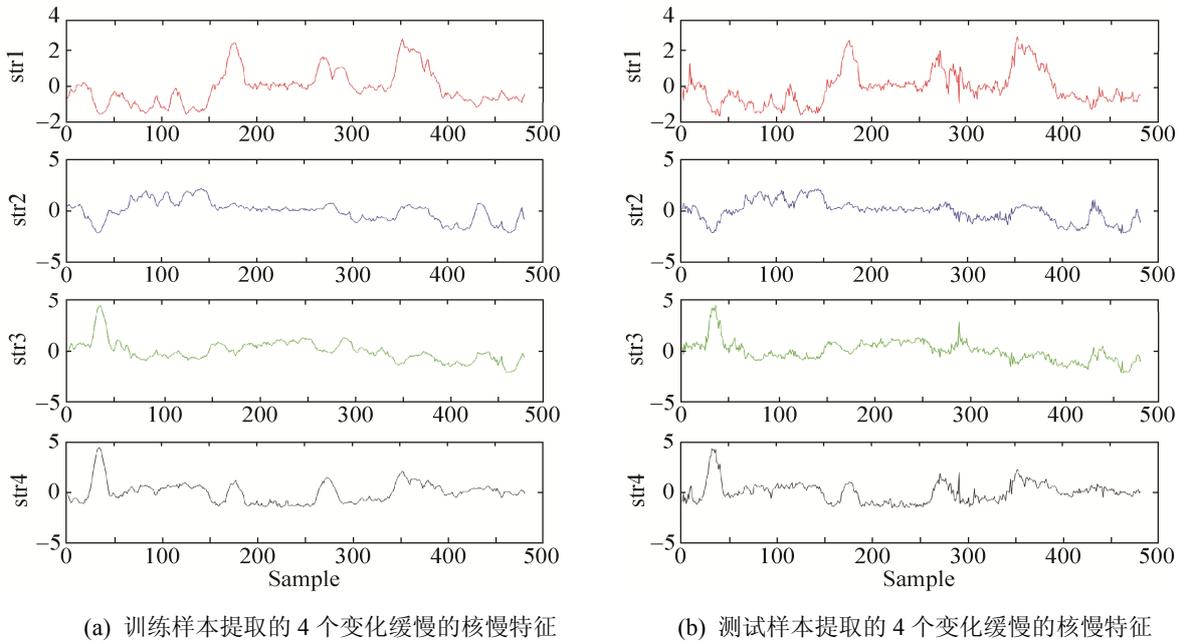


图 4 前 6 个输入变量不同时延所对应的 Rc 值
Fig. 4 Rc value of different time delay in the first six input variables



(a) 训练样本提取的 4 个变化缓慢的核慢特征
(b) 测试样本提取的 4 个变化缓慢的核慢特征
图 5 训练样本和测试样本提取的 4 个变化缓慢的核慢特征
Fig. 5 Four slowest kernel slow features extracted from training data and testing data

表 1 给出了 4 种建模方法的预测结果比较, 根据均方根误差比较可知, 本文方法的 RMSE 是最小的, 而传统 GPR 方法最大。F-GPR 由于考虑了过程数据的时滞信息, 致使重构后的数据对主导变量的贡献率增强, 且更加符合实际生产过程中因果关系的, 因此相对于 GPR 方法预测精度更高, 跟踪系数更高。FS-GPR 方法相对于 F-GPR 方法, 考虑了将时滞后重构数据进行了特征提取, 因此具有更小的 RMSE 和更大的 TP, 这是因为慢特征分析作为一种无监督的特征提取方法, 能挖掘数据中的部分先验知识, 以提高软测量模型的精度, 达到良好的预测效果。本文方法与 FS-GPR 方法不同的在于引入核函数扩展来替代二阶多项式扩展, 对于实际工业过程, 由于过程中往往会存在着比较强的非线性, 虽然 FS-GPR 方法也采用了二阶多项式扩展来实现对非线性数据的处理能力, 但是这种方法的弊端在处

理强非线性数据和高维数据的时候, 表现的比较明显, 而本文方法采用 KSFA 来提取非线性特征, 就避免了这种问题, 因此得到了更好的预测效果。

表 1 4 种建模方法的预测结果比较

Tab. 1 Comparison of prediction results of four modeling methods

建模方法	RMSE	TP
GPR	0.043 9	0.761 5
F-GPR	0.037 4	0.826 6
FS-GPR	0.032 3	0.870 6
FKS-GPR	0.028 5	0.899 8

图 6 给出了 4 种建模方法更详细的预测结果, 从图中可清晰地看出, 本文方法相对于其它 3 种方法真值和预测值之间的偏差更小, 因而能够更好地跟踪实际测量的 C_4 浓度, 实现更高的预测精度。

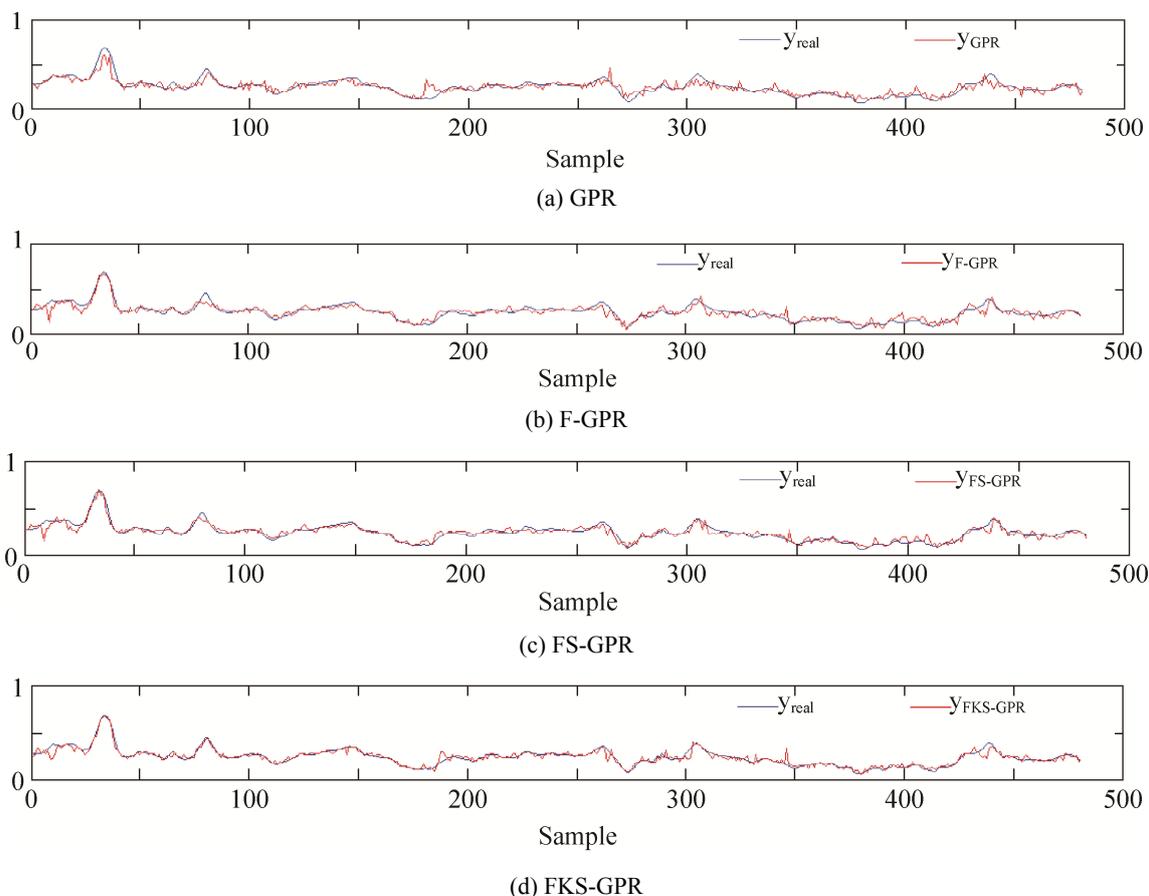


图 6 4 种建模方法的预测结果图

Fig. 6 Prediction results of the four modeling methods

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

• 1569 •

为了更加直观地展现所得效果,图7给出了4种建模方法的散点图,散点越接近于中间45°的黑色线,表明预测值越接近于真实值,软测量建模精度越高。由图7可清晰地看出本文方法建立模型得到的脱丁烷塔C₄浓度散点更加接近于黑色线,因此本文所提出的软测量建模方法能够有效地提高预测精度。

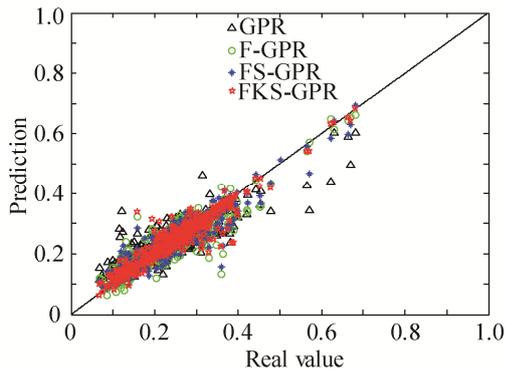


图7 4种建模方法的散点图

Fig. 7 Scatter plot of four modeling methods

5 结论

本文提出了一种基于核慢特征分析和时滞估计的高斯过程回归建模方法。针对实际工业过程的时滞特性,采用模糊曲线法进行分析,从而得到每个输入变量的最优时滞。同时由于实际工业数据存在着非线性特性,通过核函数对慢特征分析进行扩展,以更好地挖掘出工业数据中的非线性特征,反映过程中的数据变化特性,最后基于提取后的特征进行高斯过程回归建模。通过对脱丁烷塔数据的仿真实验,表明本文方法的优越性。

参考文献:

- [1] Yuan X, Ge Z, Song Z. Locally weighted kernel principal component regression model for soft sensing of nonlinear time-variant processes[J]. *Industrial & Engineering Chemistry Research* (S0888-5885), 2014, 53(35): 13736-13749.
- [2] Petr K, Ratko G, Bogdan G. Review of adaptation mechanisms for data-driven soft sensors [J]. *Computers and Chemical Engineering* (S0098-1354), 2011, 35(1): 1-24.
- [3] Facco P, Doplicher F, Bezzo F, et al. Moving average PLS soft sensor for online product quality estimation in an industrial batch polymerization process[J]. *Journal of Process Control* (S0959-1524), 2009, 19(3): 520-529.
- [4] Sun K, Liu J, Kang J L, et al. Development of a variable selection method for soft sensor using artificial neural network and nonnegative garrote[J]. *Journal of Process Control* (S0959-1524), 2014, 24(7): 1068-1075.
- [5] 刘毅, 王海清. 采用最小二乘支持向量机的青霉素发酵过程建模研究[J]. *生物工程学报*, 2006, 22(1): 144-149.
Liu Yi, Wang Haiqing. Modeling optimization of SVM for a fermentation soft sensor[J]. *Chinese Journal of Biotechnology*, 2006, 22(1): 144-149.
- [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] Zhang W, Li Y, Xiong W, et al. Adaptive soft sensor for online prediction based on enhanced moving window GPR[C]//*Control, Automation and Information Sciences (ICCAIS), 2015 International Conference on. Changshu China: IEEE, 2015: 291-296.*
- [8] 范雪莉, 冯海泓, 原猛. 基于互信息的主成分分析特征选择算法[J]. *控制与决策*, 2013, 28(6): 915-919.
Fan Xueli, Feng Haihong, Yuan Meng, et al. PCA based on mutual information for feature selection[J]. *Control and Decision*, 2013, 28(6): 915-919.
- [9] Lee J M, Yoo C K, Lee I B. Statistical process monitoring with independent component analysis[J]. *Journal of Process Control* (S0959-1524), 2004, 14(5): 467-485.
- [10] Zhang Y. Enhanced statistical analysis of nonlinear processes using KPCA, KICA and SVM[J]. *Chemical Engineering Science* (S0009-2509), 2009, 64(5): 801-811.
- [11] Wiskott L, Sejnowski T J. Slow feature analysis: Unsupervised learning of invariances[J]. *Neural computation* (S0899-7667), 2002, 14(4): 715-770.
- [12] Blaschke T, Zito T, Wiskott L. Independent slow feature analysis and nonlinear blind source separation[J]. *Neural computation* (S0899-7667), 2007, 19(4): 994-1021.
- [13] Shang C, Yang F, Gao X, et al. Extracting latent dynamics from process data for quality prediction and performance assessment via slow feature regression[C]//*American Control Conference (ACC), 2015. Chicago, IL, USA: IEEE, 2015: 912-917.*

- [14] Shang C, Huang B, Yang F, et al. Probabilistic slow feature analysis - based representation learning from massive process data for soft sensor modeling[J]. *AICHe Journal* (S0001-1541), 2015, 61(12): 4126-4139.
- [15] Zhang Z, Tao D. Slow feature analysis for human action recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (S0162-8828), 2012, 34(3): 436-450.
- [16] Shang C, Huang B, Yang F, et al. Slow feature analysis for monitoring and diagnosis of control performance[J]. *Journal of Process Control* (S0959-1524), 2016, 39: 21-34.
- [17] Zhang L, Wu C, Du B. Automatic radiometric normalization for multitemporal remote sensing imagery with iterative slow feature analysis[J]. *IEEE Transactions on Geoscience and Remote Sensing* (S0196-2892), 2014, 52(10): 6141-6155.
- [18] Böhmer W, Grünwälder S, Nickisch H, et al. Regularized sparse kernel slow feature analysis[J]. *Machine Learning and Knowledge Discovery in Databases* (S0302-9743), 2011: 235-248.
- [19] Böhmer W, Grünwälder S, Nickisch H, et al. Generating feature spaces for linear algorithms with regularized sparse kernel slow feature analysis[J]. *Machine Learning* (S0885-6125), 2012, 89(1/2): 67-86.
- [20] Fortuna L, Graziani S, Rizzo A, et al. *Soft sensors for monitoring and control of industrial processes*[M]. London: Springer Science & Business Media, 2007.
- [21] Zhang J, Jin Q, Xu Y. Inferential estimation of polymer melt index using sequentially trained bootstrap aggregated neural networks[J]. *Chemical engineering & technology* (S0930-7516), 2006, 29(4): 442-448.
- [22] 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]. *CIESC Journal*, 2014, 65(11): 4497-4502.
- [23] Lin Y, Cunningham G A. A new approach to fuzzy-neural system modeling[J]. *IEEE Transactions on Fuzzy systems* (S1063-6706), 1995, 3(2): 190-198.
- [24] 蒋昕祎, 杜红彬, 李绍军. 基于核慢特征回归与互信息的常压塔软测量建模[J]. *化工学报*, 2017, 68(5): 1977-1986.
Jiang Xinyi, Du Hongbin, Li Shaojun. Atmospheric tower soft sensor based on kernel slow feature regression and mutual information[J]. *CIESC Journal*, 2017, 68(5): 1977-1986.
- [25] 刘瑞兰. 软测量技术若干问题的研究及工业应用[D]. 杭州: 浙江大学, 2004.
Liu Ruilan. *Some Studies on Soft Sensor Technology and Their Applications to Industry Process*[D]. Hangzhou: Zhejiang University, 2004.
- [26] Fortuna L, Graziani S, Xibilia M G. Soft sensors for product quality monitoring in debutanizer distillation columns[J]. *Control Engineering Practice* (S0967-0661), 2005, 13(4): 499-508.