

基于核主元分析与核密度估计的非线性过程故障监测与识别^①



郑天标, 肖应旺

(广东技术师范大学 自动化学院, 广州 510665)
通信作者: 郑天标, E-mail: 819906961@qq.com

摘要: 在针对将核主元分析 (kernel principal components analysis, KPCA) 与基于高斯分布的控制限 (control limits, CLS) 相结合会降低其性能的问题, 提出了一种基于核主元分析与核密度估计 (kernel principal components analysis-kernel density estimation, KPCA-KDE) 相结合的非线性过程故障监测与识别方法. 该方法采用核密度估计 (kernel density estimation, KDE) 技术来估计基于 KPCA 的非线性过程监控的 CLS. 通过研究 KPCA 和 KPCA-KDE 所有 20 个故障的检出率发现, 与相应的基于高斯分布的方法进行比较, KDE 具有较高的故障检出率; 此外, 基于 KDE 的检测延迟等于或低于其他方法. 通过改变带宽和保留的主元数量进行故障检测, KPCA 记录的 FAR 值较高, 相反, KPCA-KDE 方法仍然没有记录任何假报警. 在田纳西伊斯曼过程 (Tennessee Eastman, TE) 上的应用表明, KPCA-KDE 比基于高斯假设的 CLS 的 KPCA 在灵敏度和检测时间上都具有更好的监控性能.

关键词: 故障检测与识别; 过程监控; 非线性系统; 多变量统计; 核密度估计; 机器学习; 故障诊断; 核主元分析

引用格式: 郑天标, 肖应旺. 基于核主元分析与核密度估计的非线性过程故障监测与识别. 计算机系统应用, 2022, 31(10): 329-334. <http://www.c-s-a.org.cn/1003-3254/8732.html>

Nonlinear Process Fault Identification and Detection Based on KPCA-KDE

ZHENG Tian-Biao, XIAO Ying-Wang

(School of Automation, Guangdong Polytechnic Normal University, Guangzhou 510665, China)

Abstract: The combination of kernel principal components analysis (KPCA) and control limits (CLS) based on Gaussian distribution will undermine the performance. The fault detection and identification method for nonlinear process based on kernel principal components analysis-kernel density estimation (KPCA-KDE) is proposed. kernel density estimation (KDE) technology is adopted to estimate the CLS based on KPCA for nonlinear process monitoring. According to the detection rate of all 20 faults in KPCA and KPCA-KDE, KDE has a higher fault detection rate than the corresponding method based on Gaussian distribution. In addition, KDE-based detection delay is equal to or lower than other methods. By changing the bandwidth and the number of reserved pivots during the fault detection, KPCA records a larger FAR while the KPCA-KDE does not record any false alarms. The application on the Tennessee Eastman (TE) process shows that KPCA-KDE has better monitoring performance in sensitivity and detection time than KPCA based on Gaussian CLS.

Key words: fault detection and identification; process monitoring; nonlinear systems; multivariate statistics; kernel density estimation (KDE); machine learning; fault diagnosis; kernel principal components analysis (KPCA)

^① 基金项目: 国家自然科学基金 (61174123); 广东省自然科学基金 (S2013010015007)

收稿时间: 2022-01-13; 修改时间: 2022-02-17; 采用时间: 2022-02-22; csa 在线出版时间: 2022-07-25

近年来,关于多元统计过程的监控策略可谓百花齐放,早期由于主成分分析方法 (principal components analysis, PCA) 拥有良好的降维能力被普遍应用,但由于 PCA 不适用于非线性系统, Schölkopf 等人^[1]提出了核主元分析法 (kernel principal components analysis, KPCA), KPCA 通过非线性映射函数将原始输入空间映射到高维特征空间,然后再利用特征空间中映射数据点的内积就可解决非线性问题. 吴天昊等人^[2]将 KPCA 引入核电厂设备在线监控领域中,使监测手段提供更为早期的预报警. 吕宁等人^[3]提出一种改进的多向核主元分析故障诊断模型,使得非线性主元能够在高维特征空间中被快速提取. 吴洪艳等人^[4]提出了基于小波的 KPCA 故障监测方法,降低了 KPCA 计算的复杂性,缩短了计算时间. 霍特林的 T^2 统计量和 Q 统计量 (也称平方预测误差, squared prediction error, SPE) 是 PCA 与 KPCA 在过程监控中常用的两个统计量指标. T^2 统计量用于监控模型空间的变化,而 Q 统计量用于监控剩余空间的变化. 然而, T^2 和 Q 统计量仅适用于高斯分布过程的监控. 借鉴文献 [5-8] 提出的算法,本文采用核密度估计 (kernel density estimation, KDE) 方法推导核主元分析的控制限;然后,比较了基于高斯分布统计量控制限的 KPCA 和 KPCA-KDE 在田纳西伊斯曼 (TE) 过程中的故障检测性能. 结果表明,基于 KPCA-KDE 的监测方法比基于高斯假设的 KPCA 具有更好的性能.

1 基于 KPCA-KDE 的过程监控

1.1 KPCA 算法

给定 m 个训练样本 $X_k \in \mathbb{R}^n, k = 1, 2, \dots, m$, 由非线性映射 $X_k \in \mathbb{R}^n \rightarrow \varphi(X_k \in \mathbb{R}^F)$. 可以映射到高维特征空间. 特征空间中的协方差矩阵为:

$$C_F = \frac{1}{m} \sum_{j=1}^m \langle \varphi(X_j), \varphi(X_j) \rangle \quad (1)$$

在 $\varphi(X_j)$ 中, $j = 1, \dots, m$ 假设均值和单位方差为零. 为了对角化协方差矩阵,在特征空间中求解特征值问题:

$$\lambda a = C_F a \quad (2)$$

其中, λ 是 C_F 的特征值, 满足 $\lambda \geq 0$, $a \in \mathbb{R}^F$ 是对应的特征向量 ($a \neq 0$). 特征向量可表示为映射数据点的线性组合如下:

$$a = \sum_{i=1}^m \alpha_i \varphi(X_i) \quad (3)$$

用 $\varphi(X)$ 两边同时左乘式 (2):

$$\lambda \langle \varphi(X_k), a \rangle = \langle \varphi(X_k), C_F a \rangle \quad (4)$$

将式 (1)、式 (3) 代入式 (4), 得:

$$\lambda \sum_{i=1}^m \alpha_i \langle \varphi(X_k), \varphi(X_i) \rangle = \frac{1}{m} \sum_{i=1}^m \alpha_i \left\langle \varphi(X_k), \sum_{j=1}^m \varphi(X_j) \right\rangle \langle \varphi(X_j), \varphi(X_i) \rangle \quad (5)$$

不直接对式 (1) 中 C_F 的进行特征值分解并找到特征值和主成分, 而采用核技巧, 定义一个 $m \times m$ 的核矩阵, 如式 (6):

$$[K]_{ij} = K_{ij} = \langle \varphi(X_i), \varphi(X_j) \rangle = k(X_i, X_j) \quad (6)$$

对于所有 $i, j = 1, \dots, m$ 引入核函数的形式 $k(x, y) = \langle \varphi(x), \varphi(y) \rangle$ 在式 (5) 使内积的计算 $\langle \varphi(x_i), \varphi(x_i) \rangle$ 在特征空间中作为输入数据的函数. 在特征空间中不需要进行非线性映射和计算内积. 对于核矩阵, 将式 (5) 重写为:

$$\lambda \sum_{i=1}^m \alpha_i K_{ki} = \frac{1}{m} \sum_{i=1}^m \alpha_i \sum_{j=1}^m K_{kj} K_{ji} \quad (7)$$

注意, $k = 1, \dots, m$, 因此, 式 (7) 可以表示为:

$$\lambda m K \alpha = K^2 \alpha \quad (8)$$

式 (8) 等价于特征值问题:

$$m \lambda \alpha = K \alpha \quad (9)$$

更进一步, 核矩阵可以表示为式 (10):

$$K_{ctr} = K - UK - KU + UKU \quad (10)$$

其中, U 是一个 $m \times m$ 矩阵, 其中每个元素都等于 $1/m$. K_{ctr} 的特征值分解相当于 \mathbb{R}^F 中的 PCA. 本质上相当于解决式 (9) 中的特征值问题, 它产生特征向量为 $\alpha_1, \alpha_2, \dots, \alpha_m$, 相应的特征值为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$.

由于核矩阵 K_{ctr} 是对称的, 故导出的主元是标准正交的, 即:

$$\langle \alpha_i, \alpha_j \rangle = \delta_{i,j} \quad (i, j = 1, 2, \dots, m) \quad (11)$$

然后, 将 $\varphi(X_j)$ 投影到由特征向量张成的 PC 空间上, 就可提取出主元训练观测值 $X_j, j = 1, \dots, m$ 的非线性映射的得分向量.

$$Z_{K,j} = \langle \alpha_k, (K_{ctr}) \rangle = \sum_{i=1}^m \alpha_{k,i} \langle \varphi(X_i), \varphi(X_j) \rangle \quad (12)$$

应用内核技巧, 这可以表示为:

$$Z_{K,j} = \sum_{i=1}^m \alpha_{K,i} [K_{ctr}]_{i,j} \quad (13)$$

1.2 故障检测指标

特征空间中第 j 个样本的霍特林的 T^2 , 其 KPCA 故障检测表达式为:

$$T_j^2 = [Z_{1,j}, \dots, Z_{q,j}] \Omega^{-1} [Z_{1,j}, \dots, Z_{q,j}]^T \quad (14)$$

其中, $Z_{i,j}, i = 1, \dots, q$ 表示第 j 个样本的主元得分, q 为保留主元的数量, Ω^{-1} 表示保留主元对应的特征值矩阵的逆. 从 T^2 的分布可估计其控制极限. 如果所有采样都是高斯分布的, 那么 F 分布的控制限对应于一个显著性水平 α, T^2 , 可以表示为:

$$T_\alpha^2 \sim \frac{q(m-1)}{m-q} F_{q,m-q,\alpha} \quad (15)$$

其中, $F_{q,m-q,\alpha}$ 为显著性水平 α 对应的自由度为条件下的 F 分布临界值. Choi 等人^[9] 还提出了对 Q 统计量的简化计算. 对于第 j 个样本:

$$Q_j = \left\| \phi(X_j) - \hat{\phi}_q(X_j) \right\|^2 = \sum_{i=1}^m Z_{i,j}^2 - \sum_{i=1}^q Z_{i,j}^2 \quad (16)$$

如果所有分数都是正态分布, Q 统计量在 $100(1-\alpha)\%$ 置信水平的控制极限可以推导如下:

$$Q_\alpha = \theta_1 \left[\frac{C_\alpha h_0 \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0} \quad (17)$$

其中, $\theta_1 = \sum_{j=q+1}^n \lambda_j^i$, $h_0 = 1 - 2\theta_1\theta_3/3\theta_2^2$, λ_i 为特征值, C_α 为 $100(1-\alpha)\%$ 正态百分位数.

1.3 核密度估计

KDE 是从一组随机样本中使用合适的平滑概率密度函数对数据集进行拟合的过程. 它被广泛应用于概率密度函数 (probability density functions, PDFs) 的估计, 特别是单变量的随机数据^[10-13]. 尽管这些统计特征的过程是多变量的, 因为两者都是单变量的, 所以 KDE 可应用于 T^2 和 Q 统计. 给定一个随机变量 y , 从其 m 个样本中获得它的概率密度函数 $g(y), y_j, j = 1, \dots, m$ 如下:

$$g(y) = \frac{1}{mh} \sum_{j=1}^m k\left(\frac{y-y_j}{h}\right) \quad (18)$$

其中, k 为核函数, h 为带宽或平滑参数. 基于 Cheng 等人^[14] 提出的选择带宽的最佳方法, 本文在连续范围内对密度函数积分得到概率. 因此, 假设 PDF 在指定的显著水平上, y 小于 c 的概率为:

$$p(y < c) = \int_{-\infty}^c g(y) dy = \alpha \quad (19)$$

因此, 监测统计量 (T^2 和 Q) 的控制限可由各自的

概率密度函数估算值计算:

$$\int_{-\infty}^{T^2} g(T^2) dT^2 = \alpha \quad (20)$$

$$\int_{-\infty}^{Q_\alpha} g(Q) dQ = \alpha \quad (21)$$

1.4 在线监测

对于一个以均值为中心的观测样本, 使用训练例子计算相应的核向量 $X_j, j = 1, \dots, m$ 如下:

$$[K_{tt}]_j = k(X_j, X_{tt}) \quad (22)$$

然后将测试核向量标准化, 如式 (23) 所示:

$$K_{ctt} = K_{tt} - Ku_1 - Uk_{tt} + UKu_1 \quad (23)$$

$u_1 = 1/m[1, \dots, 1]^T \in \mathfrak{R}^m$. 计算相应的得分向量 (核主元) Z_{tt} :

$$Z_{tt,k} = \langle \alpha_k, (K_{ctt}) \rangle = \sum_{i=1}^m \alpha_{k,i} \langle \phi(X_i), \phi(X_{tt}) \rangle \quad (24)$$

也可写成:

$$Z_{tt,k} = \sum_{i=1}^m \alpha_{k,i} [K_{ctt}]_i \quad (25)$$

用向量形式表示为:

$$Z_{tt} = AK_{ctt} \quad (26)$$

其中, $A = [\alpha_1, \dots, \alpha_m]$.

2 KPCA-KDE 故障监测算法过程

2.1 离线监控模型的建立

TR1. 获取正常运行条件下的数据 (NOC), 使用数据集中代表不同变量的列的均值和标准差对数据进行规范化;

TR2. 决定内核函数的类型并确定内核参数;

TR3. 构造 NOC 数据的核矩阵并将其中心化;

TR4. 获得特征值及其对应的特征向量, 并将其按降序排列;

TR5. 使用式 (11) 对特征向量进行正交化;

TR6. 使用式 (13) 得到非线性得分向量;

TR7. 根据核化 NOC 数据, 利用式 (14) 和式 (16)

计算监控性能指标 (T^2 和 Q);

TR8. 用式 (20)、式 (21) 确定 T^2 和 Q 的控制限值.

2.2 在线监控模型的建立

TT1. 获取测试样本 X_{tt} 并使用离线步骤 1 中的均值

和标准差进行规范化;

TT2. 用式 (22) 获得样本的核向量;

TT3. 据式 (23) 对核向量进行中心化;

TT4. 据式 (25) 求出测试样本的主元值;

TT5. 将测试样本的 T^2 和 Q 与模型建立阶段得到的各自控制极限进行比较;

TT6. 如果 T^2 和 Q 都小于它们的监控统计数据, 则过程处于控制之中. 如果 T^2 或 Q 超过控制限度, 则过程失控, 进行故障识别, 识别故障的来源. 基于核主元分析与核密度估计的非线性过程故障监测与识别流程图见图 1.

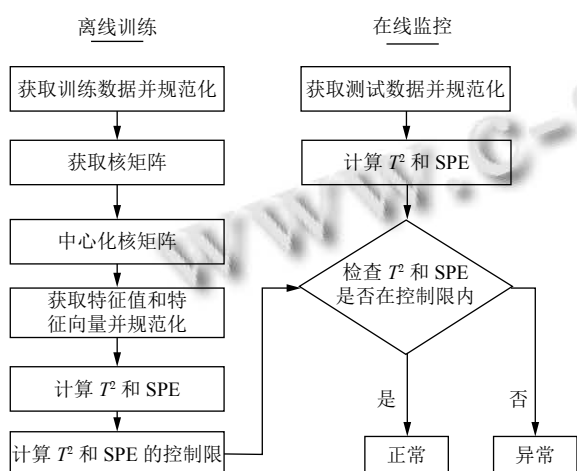


图 1 基于 KPCA-KDE 的故障监测流程图

2.3 故障变量识别

在检测到故障之后, 重要的是要识别与故障关联最紧密的变量, 以便于确定故障发生的原因.

贡献图是一种常用的故障识别方法, 它显示了各变量对故障区域的高统计指标值的贡献. 然而, 由于转换后的过程变量与原始过程变量之间存在非线性关系, 基于非线性 PCA 的故障识别不像线性主成分分析那样简单.

在本文中, 使用 Petzold 等人^[15]提出的敏感性分析原理识别故障变量. 该方法基于 Deng 等人^[16]研究的方法, 计算因参数变化而引起系统输出变化的变化率. 给定一个测试数据向量 $X_i \in \mathcal{R}^n$, 第 i 个变量对监测指标的贡献定义为:

$$T_{i,\text{con}}^2 = x_i a_i, Q_{i,\text{con}} = x_i b_i \quad (27)$$

其中, $a_i = \partial T^2 / \partial x_i$, $b_i = \partial Q / \partial x_i$.

在使用多元统计方法的故障识别中, 一种获得变量贡献的有效方法, 即通过对定义 T^2 和 Q 的函数在参考故障瞬间进行阶跃微分得到偏导数的方法来识别故障^[17].

3 应用程序

3.1 田纳西伊斯曼过程

TE 过程是基于真实工业过程的模拟^[18], 具有较强的非线性和动态特性, Chiang 等人^[19]把它用作评价和比较过程监控的基准过程. 该过程由分离器、压缩机、反应器、汽提器和冷凝器 5 个关键部件和编码为 A 到 H 的 8 个部件组成, 共采样 960 个, 变量 53 个, 其中连续变量 22 个, 3 个成分分析器采样的成分测量值 19 个, TE 过程中的操纵变量为 12 个. 采样每隔 3 分钟进行一次, 而每个故障在采样数 160 处引入. McAvoy 等人^[20]阐述了干扰和基线操作的信息工艺条件.

由于 T^2 和 Q 数据具有互补性, 因此联合使用它们进行故障检测. 由于可检测到的过程变化并不总是同时发生在模型空间和剩余空间中, 故当任一监控统计数据检测到故障时, 即检测到故障.

3.2 计算监控性能指标

由于从化学过程获得的测量通常是有噪声的, 统计量可能会随机地超过其控制限. 这相当于在没有实际发生故障的情况下统计量可能超过控制限, 即虚假警报. 换句话说, 一个统计量可能会超过它的控制限一次, 但是如果没有出现故障, 在随后的监控中, 统计量可能不会总是保持在它的控制限之上. 相反, 如果在连续的测量中统计量保持在其控制限以上, 则可能发生故障. van Sprang 等人^[21]使用故障检测规则解决了假警报的问题. 检测规则也为比较不同的监测方法提供了统一的依据. 在本文中, 当一个统计量在至少两个连续监控中超过其控制限时, 即认为发生了故障. 所有实验在使用基于此准则的数据进行测试时, 都记录了一个零的误报率的数据来用于评估不同技术的监视性能. 性能监控基于 3 个指标: 故障检出率 (fault detection rate, FDR)、误报率 (false alarm rate, FAR) 和检测延迟 (detection delay, DD). 故障检出率是正确识别的故障样本的百分比:

$$FDR = \frac{n_{fc}}{n_{tf}} \times 100\% \quad (28)$$

其中, n_{fc} 为正确识别的故障样本数, n_{tf} 为故障样本总数. FAR 是指在工厂正常运行期间, 正常样本中被确定为故障 (或异常) 的百分比:

$$FAR = \frac{n_{nf}}{n_{tf}} \times 100\% \quad (29)$$

其中, n_{nf} 表示被识别为故障的正常样本的数量, n_{tf} 为正常样本的总数量. 检测延迟为检测到引入故障之前

经过的时间。

3.3 应用结果

使用 TE 流程的故障 11 显示了基于 KPCA 的故障检测。故障 11 是反应堆冷却水入口温度的随机变化, 图 2 显示了故障的监控图。实线代表监测指标, 红线和绿线分别代表基于高斯分布和基于 KDE 的 99% 置信水平的控制限值。在 T^2 控制图中, 基于 KDE 的控制限值都低于基于高斯分布的控制限值。即与基于高斯分布的控制限相比, 监测指标超出基于 KDE 的控制限的程度更大。意味着使用基于 KDE 的 KPCA 技术的控制限比使用基于高斯分布的控制限具有更高的监控性能。

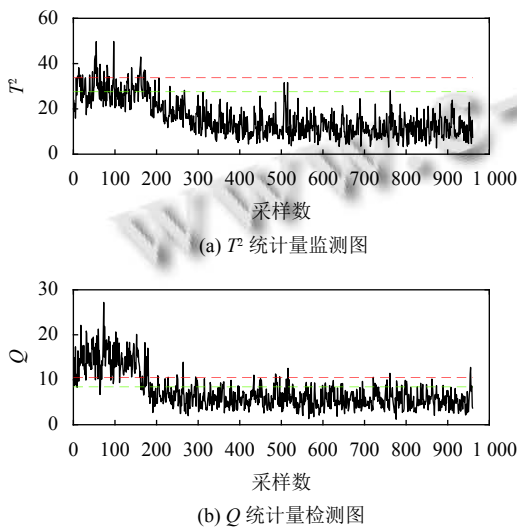


图 2 故障 11 的监视图

图 3 显示了故障 14 在 $\omega = 40$ 时 KPCA 和 KPCA-KDE 的监视图表。该故障表现为反应堆冷却水阀卡死, 大多数统计过程监测方法都能很容易地检测出该故障。在 $\omega = 40$ 时, KPCA 和 KPCA-KDE 结果一样, 在 $\omega = 10$ 时, KPCA 记录的误报率为 8.13%, 而 KPCA-KDE 的仍为零。此外, 当保留 25 个主元个数时, KPCA 记录的误报率也同样高。相反, KPCA-KDE 方法仍然没有记录任何假警报。通过研究 KPCA 和 KPCA-KDE 所有 20 个故障的检出率, 结果表明, 与相应的基于高斯分布的方法进行比较, KDE 具有较高的故障检出率; 此外, 基于 KDE 的检测延迟等于或低于其他方法。这意味着基于 KDE 控制限的方法比基于高斯分布的同类方法更早地检测出故障。因此, 相对于使用基于高斯假设的控制限, 将基于 KDE 的控制限方法与基于 KPCA 的故障检测方法相结合, 可提供更好的监控效果。

以故障 11 为例, 显示了基于 KPCA-KDE 的故障识别。故障 11 的出现引起反应堆冷却水流量的变化,

从而导致反应堆温度波动。在图 4 中所示的样本 300 处, 基于 T^2 和 Q 的贡献图均正确识别了两个故障变量。变量 9 是反应堆温度, 变量 32 对应于反应堆冷却水流速。尽管控制回路有可能在经过较长时间后补偿反应堆温度的变化, 但是通过贡献图可以正确识别出在引入故障后早期受影响的两个变量的波动。

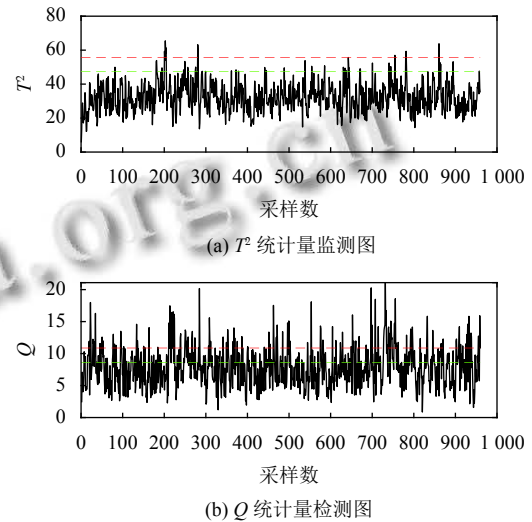


图 3 $\omega = 40$ 故障 14 基于 KPCA 的监视图

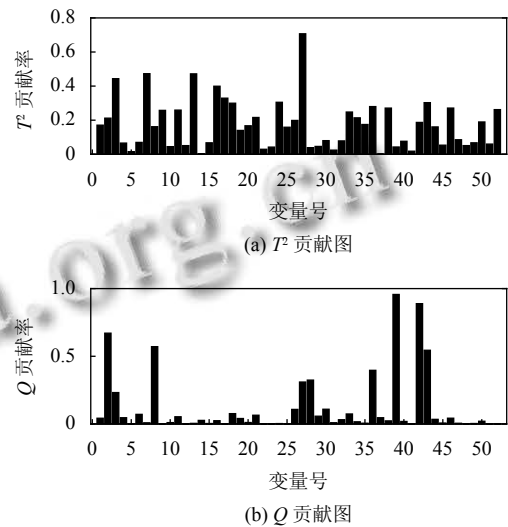


图 4 故障 11 的贡献图

为了测试 KPCA-KDE 技术的性能, 通过改变两个参数进行故障检测: 带宽和保留的主元数量。表 1 显示, 当保留 25 个主元时, KPCA 记录的 FAR 值也同样高。相反, KPCA-KDE 方法仍然没有记录任何假报警。因此, 除了提供更高的 FDR 和更早的检测外, KPCA-KDE 比基于高斯假设的 KPCA 方法监控性能更好, 当故障及早发现时, 操作人员将有更多的时间找出故障的根本原因, 以便采取补救措施。其次, 虽然有方法可

获得开发过程监控模型的最优设计参数,但不能保证最优值一直使用.造成这种情况的原因可能是人员缺乏经验,以及对流程本身缺乏了解.因此,监控方法性能越好,就更利于过程监控.

表1 在保留不同数量主元时的监测结果

主元 数量	KPCA			KPCA-KDE		
	FDR (%)	FAR (%)	DD (min)	FDR (%)	FAR (%)	DD (min)
10	99.88	0	3	99.88	0	3
15	99.75	0	6	99.75	0	6
20	96.75	0	6	99.88	0	3
25	99.75	8.13	3	99.88	0	6

4 结论与展望

本文研究了基于KPCA-KDE技术的非线性过程故障检测与识别.在这种方法中,用于构建控制图的控制限是直接从监测指标的概率密度函数得到的,而不是基于高斯分布的控制限.将该方法应用于田纳西伊斯曼过程,并与基于高斯假设的核主成分分析方法进行了性能比较.结果表明,基于KPCA-KDE比基于高斯分布的KPCA检测到的故障更多、更早.研究还表明,基于KDE的控制限比基于高斯假设的控制限更符合监测统计量的实际分布,因此具有更好的监控性能.接下来的研究工作是设计更优化的方法,既在复杂的环境中更能迅速准确的检测出故障,此外,将KPCA-KDE结合自适应算法也是一个有价值的研究方向.

参考文献

- Schölkopf B, Smola A, Müller KR. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 1998, 10(5): 1299–1319. [doi: 10.1162/089976698300017467]
- 吴天昊, 刘韬, 施海宁, 等. 基于核主元分析法的核电厂设备状态监测技术研究. *核动力工程*, 2020, 41(5): 132–137.
- 吕宁, 颜鲁齐, 白光远. 分段多向核主元分析的啤酒发酵过程故障检测. *计算机科学*, 2016, 43(S1): 25–27, 33.
- 吴洪艳, 黄道平. 基于特征向量提取的核主元分析法. *计算机科学*, 2009, 36(7): 185–187, 255. [doi: 10.3969/j.issn.1002-137X.2009.07.044]
- 杨大勇, 杨建华, 卢伟. 基于动态阈值的核密度估计前景检测算法. *计算机应用*, 2015, 35(7): 2033–2038. [doi: 10.11772/j.issn.1001-9081.2015.07.2033]
- 郭金玉, 张安宝, 李元. 基于故障检测的KPCA中核主元的选取方法. *计算机应用与软件*, 2021, 38(9): 60–66, 85. [doi: 10.3969/j.issn.1000-386x.2021.09.010]
- 郑丹, 陈路, 童楚东. 基于KPCA非线性故障监测过程优化. *数据通信*, 2021, (3): 23–25, 28. [doi: 10.3969/j.issn.1002-

5057.2021.03.005]

- Wang H, Peng MJ, Yu Y, *et al.* Fault identification and diagnosis based on KPCA and similarity clustering for nuclear power plants. *Annals of Nuclear Energy*, 2021, 150: 107786. [doi: 10.1016/j.anucene.2020.107786]
- Choi SW, Lee IB. Nonlinear dynamic process monitoring based on dynamic kernel PCA. *Chemical Engineering Science*, 2004, 59(24): 5897–5908. [doi: 10.1016/j.ces.2004.07.019]
- Bowman AW, Azzalini A. *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-plus Illustrations*. Oxford: Clarendon Press, 1977.
- 王萌萌, 梁沪丹, 寇俊克. 基于MATLAB的核密度估计研究. *科技视界*, 2021, (4): 45–47.
- 王国良, 梁德群, 王新年, 等. 基于分箱核密度估计的非参数多模态背景模型. *计算机应用*, 2007, 27(5): 1121–1122.
- Xiong L, Liang J, Qian JX. Multivariate statistical process monitoring of an industrial polypropylene catalyzer reactor with component analysis and kernel density estimation. *Chinese Journal of Chemical Engineering*, 2007, 15(4): 524–532.
- Cheng C, Chiu MS. Nonlinear process monitoring using JITL-PCA. *Chemometrics and Intelligent Laboratory Systems*, 2005, 76(1): 1–13. [doi: 10.1016/j.chemolab.2004.08.003]
- Petzold L, Li ST, Cao Y, *et al.* Sensitivity analysis of differential-algebraic equations and partial differential equations. *Computers & Chemical Engineering*, 2006, 30(10–12): 1553–1559.
- Deng XG, Tian XM, Chen S. Modified kernel principal component analysis based on local structure analysis and its application to nonlinear process fault diagnosis. *Chemometrics and Intelligent Laboratory Systems*, 2013, 127: 195–209. [doi: 10.1016/j.chemolab.2013.07.001]
- Martins JRRA, Sturdza P, Alonso JJ. The complex-step derivative approximation. *ACM Transactions on Mathematical Software*, 2003, 29(3): 245–262. [doi: 10.1145/838250.838251]
- Downs JJ, Vogel EF. A plant-wide industrial process control problem. *Computers & Chemical Engineering*, 1993, 17(3): 245–255.
- Chiang LH, Russell E, Braatz RD. *Fault Detection and Diagnosis in Industrial Systems*. London: Springer, 2001.
- McAvoy TJ, Ye N. Base control for the Tennessee Eastman problem. *Computers & Chemical Engineering*, 1994, 18(5): 383–413.
- van Sprang ENM, Ramaker HJ, Westerhuis JA, *et al.* Critical evaluation of approaches for on-line batch process monitoring. *Chemical Engineering Science*, 2002, 57(18): 3979–3991. [doi: 10.1016/S0009-2509(02)00338-X]

(校对责编: 牛欣悦)