

控制与决策

Control and Decision

基于广义主成分分析的重构故障子空间建模方法

杜柏阳, 孔祥玉, 冯晓伟

引用本文:

杜柏阳, 孔祥玉, 冯晓伟. 基于广义主成分分析的重构故障子空间建模方法[J]. *控制与决策*, 2021, 36(4): 808–814.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.0818>

您可能感兴趣的其他文章

Articles you may be interested in

[基于自适应混合核典型变量分析的工业过程质量相关故障检测](#)

Quality-related fault detection for industrial processes based on adaptive mixed kernel canonical variable analysis

控制与决策. 2021, 36(4): 801–807 <https://doi.org/10.13195/j.kzyjc.2020.0690>

[基于多块信息提取的AUV资源勘查系统故障检测](#)

Fault detection of AUV resource exploration system based on multi-block information extraction

控制与决策. 2021, 36(4): 790–800 <https://doi.org/10.13195/j.kzyjc.2019.0732>

[标签Petri网的路径信息在故障诊断中的应用](#)

Application of path information of labeled Petri nets in fault diagnosis

控制与决策. 2021, 36(2): 325–334 <https://doi.org/10.13195/j.kzyjc.2019.0698>

[基于马尔可夫过程的多部件系统劣化状态空间划分模型](#)

Multi-component system state space partition model based on Markov process

控制与决策. 2021, 36(2): 418–428 <https://doi.org/10.13195/j.kzyjc.2019.0480>

[基于不变网络模型和故障注入的分布式信息系统故障溯源方法](#)

Fault source location algorithm for distributed information system based on invariant network and fault injection

控制与决策. 2020, 35(11): 2723–2732 <https://doi.org/10.13195/j.kzyjc.2019.0214>

基于广义主成分分析的重构故障子空间建模方法

杜柏阳¹, 孔祥玉^{1†}, 冯晓伟²

(1. 火箭军工程大学 导弹工程学院, 西安 710025; 2. 火箭军工程大学 核工程学院, 西安 710025)

摘要: 在线故障诊断是工业过程中十分重要的问题. 相比传统贡献图而言, 基于重构的故障诊断受到特别关注. 传统的主元分析方法没有考虑故障数据中同时包含正常工况信息和故障信息, 因而提取出故障子空间对故障的描述准确性不足. 为提高故障子空间的准确性, 提出一种基于广义主成分分析的重构故障子空间建模方法. 首先, 同时考虑正常工况数据和故障数据, 分析数据关联, 提取出两个数据的广义主成分, 利用投影关系建立故障子空间模型; 然后, 构建主成分分析故障监测模型, 通过监测重构数据筛选广义主成分和故障方向数量, 得到正常运行和故障子空间最优组合. 该方法充分利用正常工况和故障工况的数据, 所提取的故障子空间能够更加充分地反映故障信息, 对后续提高故障诊断的准确性具有重要的作用. 最后, 通过 Matlab 数值仿真和 TE 工业过程验证所提出方法的有效性.

关键词: 广义主成分分析; 故障子空间; 重构故障诊断; TE 工业过程

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.0818

开放科学(资源服务)标识码(OSID):



引用格式: 杜柏阳, 孔祥玉, 冯晓伟. 基于广义主成分分析的重构故障子空间建模方法[J]. 控制与决策, 2021, 36(4): 808-814.

Reconstructed fault subspace modelling method based on generalized principal component analysis

DU Bo-yang¹, KONG Xiang-yu^{1†}, FENG Xiao-wei²

(1. College of Missile Engineering, The Rocket Force University of Engineering, Xi'an 710025, China; 2. College of Nuclear Engineering, The Rocket Force University of Engineering, Xi'an 710025, China)

Abstract: Fault diagnosis has been a crucial task for industrial processes. Compared with the traditional contribution plot method, the reconstruction-based fault diagnosis has drawn special attention, especially for its superior performance in the fault data processing. However, the conventional principal component analysis method pays little attention to the fact that the fault data contains both the fault information and the normal operating information. So the fault subspace modelled by the conventional principal component analysis is not accuracy enough. In order to improve the accuracy of the fault subspace, a reconstructed fault subspace modelling method is proposed based on generalized principal component analysis. First, we analyze the correlation between the normal operation data and fault data to obtain the generalized principal components, and model the fault subspace using projection relationship. Then, the principal component analysis monitoring system is constructed, and the optimal combination of normal operating and fault subspaces are obtained by monitoring the reconstruction data to select the appropriate number of generalized principal components and fault directions. The method makes full use of the normal operating data and faulty operating data, thus the extracted fault subspace can reflect more fault information than the previous methods, which plays an important role to help improve the fault diagnosis accuracy. Finally, both Matlab numerical simulation and Tennessee Eastman (TE) industrial process verify the effectiveness of the proposed method.

Keywords: generalized principal component analysis; fault subspace; reconstructed fault diagnosis; Tennessee Eastman industrial process

0 引言

随着现代工业的发展,大型自动化系统的结构越来越复杂,一处故障可能导致停机停产,甚至会造成

机毁人亡的灾难性事件,因而对工业产品的可靠性、安全性要求越来越高.基于此,在线故障诊断方法越来越受到人们的重视^[1-2].

收稿日期: 2019-06-10; 修回日期: 2019-11-14.

基金项目: 国家自然科学基金项目(61374120, 61673387, 61903375, 61833016).

责任编委: 吴立刚.

[†]通讯作者. E-mail: xiangyukong01@163.com.

实现在线故障诊断的方案较多. 传统的基于数据驱动的故障诊断方法, 如人工神经网络算法^[3]等, 通过大量历史数据和当前测量数据, 采用训练神经网络的方式得到故障监测和判别模型, 从而实现对故障的监测和定位. 该方法处理过程直观, 比较容易理解, 但是算法的计算复杂度较高, 处理的时间不易满足实际需求. 针对这种问题, 有的学者提出了基于降维的故障诊断方法, 例如偏最小二乘算法^[4]、主成分分析算法^[5]. 该类方法利用历史正常数据构建统计模型, 得到两种统计量 T^2 和SPE的控制限, 通过计算当前测量数据的 T^2 和SPE统计量实现对故障的实时监测. 该方法构建模型的计算复杂度较低, 对过程数据的异常值感应敏感、准确, 只是现有算法对故障定位和诊断的精度还有待于提高. 基于重构的贡献图方法在故障定位和诊断精度方面具有良好的性能, 在严重的故障拖尾效应下, 该方法依然可以得到正确的结果^[6]. 重构方法表明, 在故障工况过程数据中, 不仅包含故障过程因素, 还包含正常过程因素, 这对人们研究故障工况过程数据有很大的启发作用. Lu等^[7]通过改进传统的重构故障方法, 实现对非线性飞机传感器的故障重构, 扩展了人们改进重构故障方法的思路.

通过构建 T^2 和SPE统计模型, Zhao等^[8]给出了基于重构故障的多个故障诊断定位的方法, 该方法要实现对多个故障的准确定位和诊断, 需依赖准确的故障子空间. 当前, 故障子空间提取方法主要研究的对象是故障数据, 通过直接对故障数据的分析和分解可以得到故障子空间, 例如基于奇异值分解的故障子空间提取方法^[9]、阈值故障子空间提取方法^[10]和通过相对变化分析寻找故障变化方向的方法^[11]. 前两种算法都是在数据层面直接处理得到故障子空间. 实际上, 从故障重构的角度理解, 直接剔除部分数据的做法只是较大概率地去掉了包含正常过程信息的数据, 因而传统方法所提取的故障子空间只在一定意义上反映了故障的方向. Zhao等^[12]提出了故障相关主成分分析方法, 通过考察故障数据和正常工作数据的相对变化研究过程监测. 该方法在监测系统的两个方向下挖掘出故障相关监测方向和无关方向, 从而提高了故障监测的准确性. 在提取故障子空间方法的角度上, 故障工况的数据中, 正常的过程信息会污染数据的故障信息, 导致提取的故障子空间不准确. 相对而言, 阈值故障子空间提取方法比奇异值分解的方法得到的故障子空间要纯粹一些, 因为它一定程度上考虑了非故障数据信息对故障子空间的影响. 因此,

在故障工况的数据中, 将故障信息的数据分离出来, 然后进行故障子空间的提取, 所提取出的故障子空间能够更加精确地代表故障方向.

本文提出的广义主成分分析算法, 通过提取正常工况过程数据和故障工况过程数据的广义主成分, 得到故障工况过程数据中属于正常过程的信息空间, 从而分离出故障工况过程数据中故障信息的数据, 以此提取出故障子空间. 该方法从信息要素的角度, 对故障工况的过程数据进行有效分离, 可以充分利用已有数据中包含的有效信息, 提取出准确的故障子空间. 此外, 为进一步提高故障子空间的提取精度, 本文还利用PCA监测模型构造基于统计量的故障子空间优化方法. 本文提出的方法对于故障库的制备, 提高多个故障的定位和诊断效率, 以及对产品部件进行剩余寿命预测, 给装备维护提供指导和建议等方面, 具有重要的意义.

1 基于主成分分析的统计过程监控

考虑一个正常工况下的过程数据 X , 输入变量个数为 m , 采集的样本数量为 n . 假设 X 已经过标准化处理, 其均值为零, 方差为 1 ^[13]. 于是, 通过主成分分析可以构建监测模型^[14]

$$\begin{cases} T = XP^T, \\ E = X(I - PP^T). \end{cases} \quad (1)$$

其中: $P \in \mathbf{R}^{m \times j_t}$ 和 $T \in \mathbf{R}^{n \times j_t}$ 分别为过程数据 X 中主元子空间的负载矩阵和得分矩阵, $E \in \mathbf{R}^{m \times j_e}$ 为过程数据 X 中残差子空间的残差得分矩阵, $j_e + j_t = m$. 对应主元子空间和残差子空间, 通过统计量 T^2 和SPE可以实现对两个子空间下系统状态的实时监控. 其计算公式分别为

$$\begin{cases} T^2(i) = X(i)P^T \Lambda^{-1} (X(i)P^T)^T, \\ \text{SPE}(i) = E(i)E^T(i). \end{cases} \quad (2)$$

其中: $X(i)$ 是 X 的第 i 行; Λ 是通过历史数据建立主元模型的特征对角阵, 可表示为 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$. T^2 统计量主要表示当前输入数据与主元空间中标准数据状态的距离, 而SPE统计量主要是评估测量变量在残差空间的变异程度^[15].

通过变量的统计分布规律可以计算出对应两个统计量的控制限, 作为判断系统状态的依据. T^2 统计量服从F分布, 而SPE统计量服从 χ^2 分布^[16], 即

$$\begin{cases} T^2(i) \sim \frac{J_t(n^2 - 1)}{n(n - j_t)} F_{j_t, n - j_t, \alpha}, \\ \text{SPE}(i) \sim g\chi_{h^2, \alpha}. \end{cases} \quad (3)$$

其中: α 是两个分布的置信度, $g = \frac{\theta_1}{\theta_2}$, $h = \frac{2\theta_2^2}{\theta_1}$, θ_2 是建模时历史数据 SPE 的均值, θ_1 是该数据的方差. 容易理解: 当输入的测量数据处于正常工况时, 该数据对应的 T^2 统计量和 SPE 统计量就处在相应控制限以下; 而当输入的测量数据出现异常值时, 该数据对应的 T^2 统计量和 SPE 统计量将变化到相应控制限以上. 不同子空间下的数据高出控制限的情况不相同, 面向过程的数据异常主要导致主元子空间所对应的统计量 T^2 的值变化较大, 其他情况则主要导致残差子空间所对应的统计量 SPE 变化较大. 一般而言, 大多数异常值对两个统计量都会产生影响.

对样本故障信息的深入分析可以实现故障诊断. 故障重构^[9]是实现故障定位与诊断的一种十分有效的方法, 其基本目的是通过故障子空间重构故障, 使得重构后的状态数据可以消除故障信息. 如果能够得到一个故障子空间, 则重构后的数据可以表示为

$$X^*(i) = X(i) - \Sigma f(i). \quad (4)$$

其中: $\Sigma \in \mathbf{R}^{m \times r_f}$ 是由正交向量张成的故障子空间; $f(i) \in \mathbf{R}^{r_f \times 1}$ 是当前数据在故障方向的得分, r_f 表示沿着故障信息变化的主要故障方向.

对 T^2 统计量而言, 最大程度消除 X 故障信息的影响可以通过最小化差异距离获得^[5], 即

$$\begin{aligned} f_T(i) &= \arg \min \|\Lambda^{-\frac{1}{2}} X(i) P^T\| = \\ &= \arg \min \|\Lambda^{-\frac{1}{2}} (X(i) - \Sigma f(i)) P^T\| = \\ &= X(i) (\Sigma^T P P^T \Sigma)^{-1} \Sigma^T P P^T = \\ &= X(i) (\hat{\Sigma}^T \hat{\Sigma})^{-1} \hat{\Sigma}^T. \end{aligned} \quad (5)$$

其中: $\hat{\Sigma} = P P^T \Sigma$, 可以理解为故障子空间在主元子空间 P 中的表达; $P P^T = P (P^T P)^{-1} P^T$.

对 SPE 统计量而言, 对 X^* 的最优估计是寻找 X^* 到主元子空间的最小距离^[5], 即

$$\begin{aligned} f_S(i) &= \arg \min \|(I - P P^T) X^*(i)\| = \\ &= \arg \min \|X(i) - \Sigma f(i)\| = \\ &= X(i) ((I - P P^T) \Sigma^T \Sigma (I - P P^T))^{-1} (I - P P^T) \Sigma^T = \\ &= X(i) (\tilde{\Sigma}^T \tilde{\Sigma})^{-1} \tilde{\Sigma}^T. \end{aligned} \quad (6)$$

其中: $\tilde{\Sigma} = (I - P P^T) \Sigma$, 可以理解为故障子空间在残差子空间中的表达; $I - P P^T = I - P (P^T P)^{-1} P^T$.

得到重构数据后, 对重构后的数据再进行故障检测, 分别计算此时的 T^2 统计量和 SPE 统计量, 判断统计量是否存在于控制限以下, 便可判断出此时故障的影响是否完全消除.

2 基于广义主成分分析的故障方向提取

基于奇异值分解的故障子空间提取方法, 因为对数据特征分解不够充分, 难以得到精确的故障子空间. 为提高包含故障信息数据的质量, 阈值故障子空间通过人工设置阈值的方式, 筛选出能充分反映故障信息的数据以提取故障子空间, 相比于传统的奇异值提取方法, 提高了故障子空间的准确度. 但是该方法的准确性依赖于阈值的设置方式, 同时硬性地去除部分数据也会导致故障信息的流失, 造成不必要的浪费. 针对传统故障子空间提取方法的不足, 本文提出的基于广义主成分分析的提取方法则利用正常工况数据和故障工况数据的广义主子空间, 实现故障数据在数值分布上的提纯, 从而可以得到更加精确的故障子空间, 同时也有利于提高提取故障子空间的效率.

从另一个角度考虑重构数据公式(4), 可以认为 $\Sigma f(i)$ 是一个整体, 于是, 故障数据包含两个部分: 正常工况部分和故障工况部分. 从信息角度来看, 正常工况的部分为该数据在正常工况的子空间投影, 故障工况的部分为该数据在故障工况的子空间, 也就是故障子空间的投影. 从这个观点可以理解, 奇异值分解方法所提取的故障子空间只能建立在一定的可检测界限上. 阈值故障子空间法通过一定的筛选, 减少了故障数据中的正常工况的部分. 实际上, 直接提取故障子空间十分困难, 但是如果能够得到正常工况的子空间, 将原始数据在正常工况下投影, 并从原始数据中去掉该投影部分, 则剩余的数据就是原始数据在故障子空间中的投影. 进而, 通过主成分分析可以提取出对应的故障子空间.

上述步骤的核心环节是获取故障数据在正常工况下的子空间, 广义主成分分析方法是解决这个问题非常好的方案.

通过广义主成分得到的广义主元空间, 表达的是历史正常工况数据 Y 和当前故障工况数据 X 公共程度最高的子空间^[17]. 假设 Y 的协方差为 C_y , X 的协方差为 C_x , 广义主成分的基本表达式为

$$W C_x W = W C_y W \Lambda. \quad (7)$$

其中: 对角阵 Λ 由特征对 (C_x, C_y) 的广义特征值组成, 矩阵 W 是特征对 (C_x, C_y) 的广义主成分. 将特征值矩阵 Λ 的各个元素从大到小排列, 对应的 $W(1)$ 方向为数据 Y 和数据 X 线性关系最强的方向.

该算法的基本步骤如下:

step 1: 计算 $C_y = Y^T Y$, $C_x = X^T X$.

step 2: 通过广义主成分算法提取特征对 (C_x, C_y) 的广义主成分 W , 有

$$[W, \Lambda] = \text{eig}(C_x, C_y), \quad (8)$$

其中 eig 表示求广义特征的运算. 利用 W 的前 j 个列向量 \tilde{W}_j 重构故障工况数据, 得到故障数据部分 $X_f = X(I - \tilde{W}\tilde{W}^T)$.

step 3: 对矩阵 X_f^T 做奇异值分解, 有

$$X_f^T = U\Omega V^T, \quad (9)$$

其中 Ω 对角线上的奇异值按照降序排列, 筛选出前 i 个故障方向 $U(1:i)$, 此时的故障子空间 $\Sigma = U(1:i)$.

图1给出了广义主成分提取故障子空间的流程.

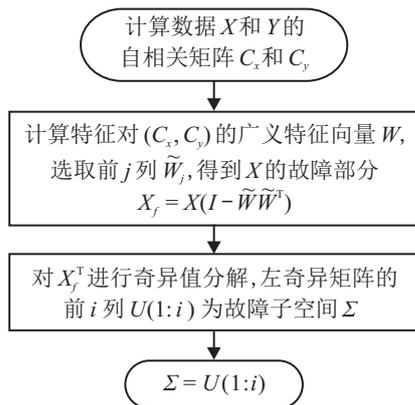


图1 广义主成分提取故障子空间流程

需要说明的是, 在该算法中, i 和 j 的值需要预先设定好. 对于任意给定的 i 和 j 而言, 都存在对应的故障子空间 Σ , 他们的差别在于描述故障信息能力不同. 能否得到描述故障信息能力最强、形式最为简单的故障子空间成为问题的关键. 这个步骤需要在基于重构误差的故障监测模型下通过筛选构建模型获得.

3 基于重构误差的故障子空间构建

为进一步提高算法提取出故障子空间的精确性, 可以利用重构误差检验所得故障子空间的性能. 定义重构超出量为重构数据的统计量超出控制限 (l_{T^2}, l_{SPE}) 的部分之和, 未超出的统计量则不计入. 有

$$\begin{cases} s_{T^2} = \sum_{p=1}^n (T^2(p) - l_{T^2}), T^2(p) \geq l_{T^2}; \\ s_{SPE} = \sum_{p=1}^n (SPE(p) - l_{SPE}), SPE(p) \geq l_{SPE}. \end{cases} \quad (10)$$

实际上, 满足 $s_{T^2} = s_{SPE} = 0$ 的故障子空间很多, 因此, 需要通过优化故障子空间得到最简单有效的组合. 优化故障子空间的具体步骤如图2所示.

step 1: 预处理正常工况数据, 得到均值 $\text{mean}(Y)$ 和方差 $\text{var}(Y)$, 从而得到正常工况数据, 利用 Y 构建 PCA 监测模型, 有

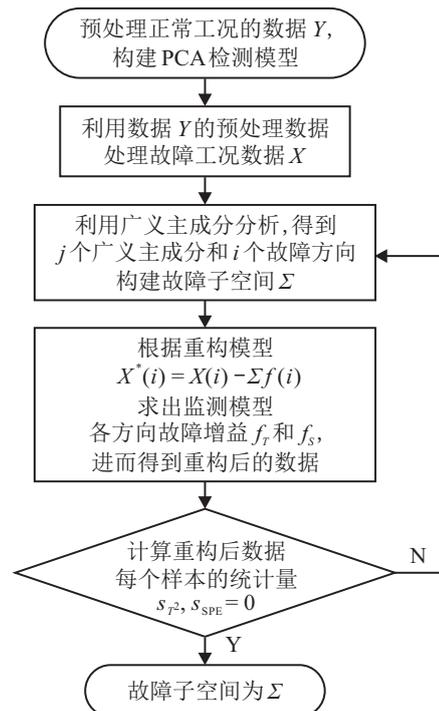


图2 重构误差的故障子空间优化流程

$$\begin{cases} T = YP^T, \\ E = Y(I - PP^T), \end{cases} \quad (11)$$

并计算出对角阵 $\text{diag}(Y)$. 计算 T^2 统计量和 SPE 统计量的控制限 l_{T^2} 和 l_{SPE} .

step 2: 利用 Y 的预处理结果均值 $\text{mean}(Y)$ 和方差 $\text{var}(Y)$, 将故障工况数据减掉 $\text{mean}(Y)$ 后, 再除以 $\text{var}(Y)$, 得到故障工况数据 X .

step 3: 基于广义主成分分析提取故障方向, 选择 j 个广义主元, 并选择 U 中前 i 个故障方向组成故障子空间 Σ , 初始的 i 和 j 设置为 1.

step 4: 利用故障子空间, 通过式 (5) 和 (6), 分别计算样本的 T^2 和 SPE 统计量在各个方向的故障增益 $f_T(i)$ 和 $f_S(i)$, 进而消除 X 在故障子空间的投影 $\Sigma f_T(i)$ 和 $\Sigma f_S(i)$, 得到重构数据 X^* .

step 5: 计算重构数据 X^* 每个样本的统计量 T^2 和 SPE, 进行故障监测. 当 $s_{T^2}, s_{SPE} > 0$ 时, 如果 $i = m$, 则令 $i = 1, j = j + 1$, 返回 step 3; 如果 $i < m$, 则令 $i = i + 1, j$ 不变, 返回 step 3; 当 $s_{T^2} = s_{SPE} = 0$ 时, 输出此时的故障子空间.

本文提出的基于监测统计的故障子空间提取方法, 通过广义主成分分析, 消除故障数据中主要包含正常工况信息的部分, 得到的故障子空间能够更加精确地描述与故障相关的变化规律. 这更新了学者对故障描述的认识, 避免了传统提取故障子空间算法中直接求得故障子空间的偏差, 从而保证测量数据中故障因子的完全提取, 提高了模型的子空间分解效率.

4 数值分析与仿真实验

本节通过两个算例验证本文所提出算法在提取故障子空间的有效性,第1个算例为典型的数值算例^[18],第2个为TE工业过程.

4.1 数值算例下故障子空间提取

数值仿真的模型如下:

$$X^*(i) = Gt(i) + e(i). \quad (12)$$

其中: $X \in \mathbf{R}^{5 \times 500}$ 是记录5个传感器的500个数据点; $G \in \mathbf{R}^{5 \times 2}$ 是建立过程状态与传感器测量的测量矩阵; $t \in \mathbf{R}^{2 \times 500}$ 代表两个独立的过程,且服从标准正态分布; $e \in \mathbf{R}^{5 \times 500}$ 代表5个独立的噪声,服从均值为0、方差为0.01的高斯分布. 参照文献[9]的研究,测量矩阵为

$$G = \begin{bmatrix} -0.1670 & -0.1352 \\ -0.5671 & -0.3695 \\ -0.1608 & -0.1019 \\ 0.7574 & -0.0563 \\ -0.2258 & -0.9119 \end{bmatrix}. \quad (13)$$

故障发生100采样点以后的第5个变量上,故障幅度为6,可得故障工况数据为

$$X(i) = X^*(i) - \Sigma f(i). \quad (14)$$

其中 $\Sigma = [0 \ 0 \ 0 \ 0 \ 1]^T$, $f = 6$.

图3为 T^2 统计量故障检测结果,图中细线部分表示该过程的 T^2 统计量控制限,没有重构的故障工况数据的 T^2 统计量,在100点以后明显超过控制限,在控制限上方震荡,这说明PCA监测模型监测到系统在100点以后出现系统故障. 同时,图3中控制限下方的数据表示利用本文算法所提取出故障子空间的重构数据,数据都在控制限以下,这说明重构后的数据已经不再包含故障信息了. 为更加清晰地展示重构后的数据,重构后的数据 T^2 统计量如图3中小框所示.

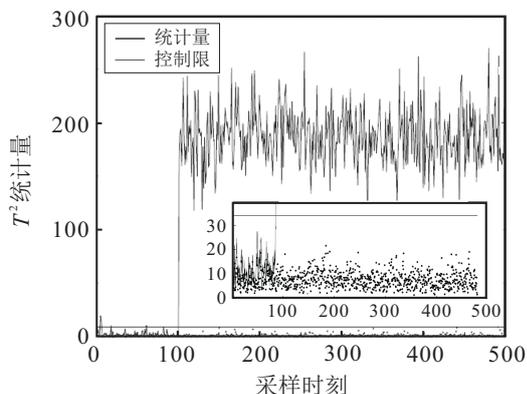


图3 仿真数据在重构前(实线)和重构后(点线) T^2 统计量的变化情况

SPE统计量的变化情况如图4所示. 未去除故障信息的数据变化情况与 T^2 统计量的变化情况相类似,也是在100点以后明显超过控制限,这说明在残差空间,PCA监测模型也检测到系统在100点以后出现故障. 图4中小框是将纵坐标40以下的部分放大后的内容,图中点线都在细线以下,说明在残差空间中,经过重构的数据,其SPE统计量都分布在控制限以下. 需要说明的是,经过故障子空间提取的重构算法,样本减去的数值只是该样本在故障子空间的投影. 因此,经过重构后,样本越充分地回归到控制限以下,说明提取出的故障子空间越准确.

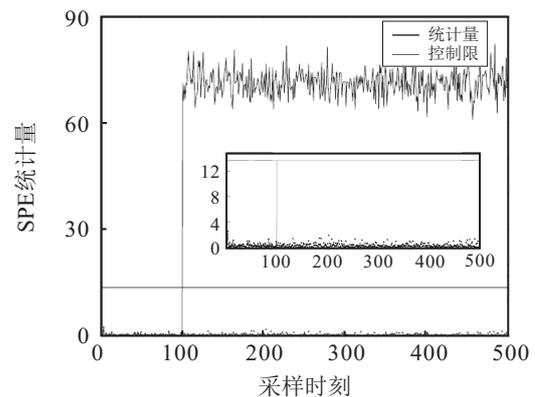


图4 仿真数据在重构前(实线)和重构后(点线)SPE统计量的变化情况

由于仿真数据的结构比较简单,故障子空间只需要一次优化就可得到重构超出量都等于0的结果. 传统的SVD方法^[10]提取故障子空间的结果为

$$\begin{bmatrix} -0.033 & 0.023 & 0.239 & -0.586 & 0.699 \\ -0.760 & -0.222 & 0.043 & 0.595 & 0.132 \end{bmatrix}^T,$$

该故障子空间的可解释性很差. 而本文算法提取故障子空间为 $[-0.002 \ -0.001 \ 0.000 \ -0.003 \ 1.000]^T$, 这个故障方向与实验数据的设置非常符合. 从该方向可以直接看出故障主要存在于第5个变量位置,直接解释了故障的原因. 需要说明的是,能够直接通过变量位置定位故障的情况并不多见,大多数故障出现时,并不是直接对应某一个变量,因此,通过已知的故障数据构建故障库,建立故障子空间与实际故障状况的逻辑关联,通过故障子空间准确描述系统的故障,是解决复杂故障或者多个故障识别的关键问题.

4.2 TE工业过程故障子空间提取

田纳西伊斯曼工业过程是由伊斯曼化工公司提出的基于实际化工生产过程的仿真实例^[19]. 此过程已经广泛作为连续过程的策略、监测、诊断的研究平台. 此过程有4种反应物(A、B、C、D),生成两种产物(G和H),并包含一种副产品. 整个过程包括5个主要的操作单元:反应器、冷凝器、气液分离器、循环压缩

机和汽提塔。

整个生产过程以3种不同的产物混合比例(G/H)构成6种不同的操作模式,本实验只考虑其中的一种基本工况模式. 整个过程包含总共52个观测变量,其中有11个操作变量和41个测试变量,测试变量含有22个连续变量和19个成分变量. TE工业过程内部的系统之间深度耦合,高度非线性,而且开环不稳定,是目前过程控制领域几个具有挑战性的控制问题之一. 在本实验中仿真数据包含一个正常运行工况和11个可操作的故障工况. 所有的工况都有480个训练样本和960个测试样本. 在故障的测试数据中,每3 min 采样一次,过程仿真时间是48 h,故障出现在第9个小时开始的时候,即从第161组数据开始引入故障.

本文选取故障02的数据作详细展示. 故障02表示的是进料口某种组分含量发生了变化,另外两种组分的进料流量比保持不变,这会导致化工系统出现一个阶跃的故障. 通过构建重构超出量的模型,采用PCA监测方法,通过故障02的数据验证本文所提出的基于广义主成分分析的故障子空间提取方法的有效性,故障02的数据展示的是两种统计量的结果.

图5显示的是采用故障子空间提取前后, T^2 统计量故障检测结果的效果对比,其中实线表示没有重构的故障工况数据的 T^2 统计量. 该统计量在160点以后明显超过了控制限,在控制限上方震荡,这说明PCA监测模型监测到TE化工过程在160点以后出现系统故障. 图5中的小框放大了纵坐标为40以下的图片内容,其中点线表示采用本文提出的故障子空间重构故障工况数据的 T^2 统计量,可以看出所有统计量都没超过控制限. 从而验证了本文算法的有效性.

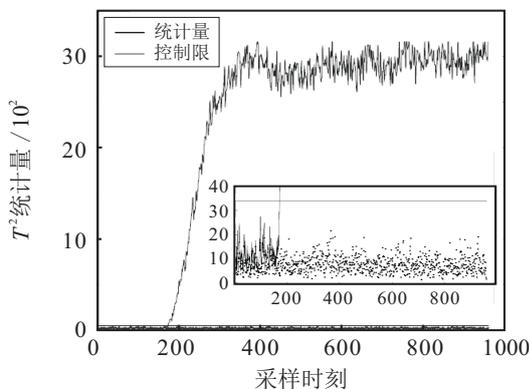


图5 故障02数据在重构前(实线)和重构后(点线) T^2 统计量的变化情况

SPE统计量的变化情况如图6所示,未重构的故障工况SPE统计量的变化情况与 T^2 统计量相似,在160点以后明显超过控制限,这说明在残差空间,PCA

监测模型也检测到系统在160点以后出现故障. 显然,从统计量的变化形式来看,TE工业过程数据的信息结构比仿真算例中的信息结构复杂. 图6中小框放大了纵坐标30以下的图片内容,点线表示重构后的SPE统计量,显然都分布在控制限以下. 这说明对于TE工业过程,本文提出的故障子空间提取算法能够准确地去除故障数据中的故障信息.

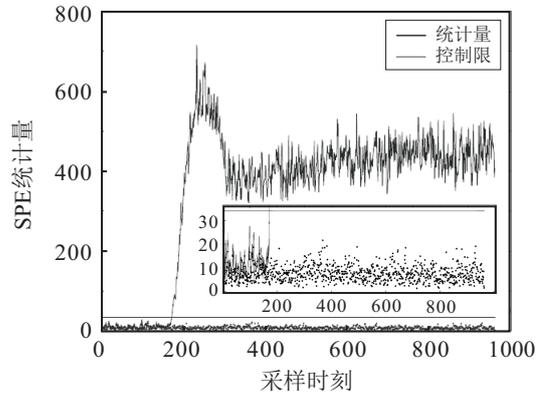


图6 故障02数据在重构前(实线)和重构后(点线)SPE统计量的变化情况

实际上,本文方法对于所有的故障过程都能够有效地提取故障子空间. 为进一步展示本文方法的有效性,将其与机遇重构的传统方法进行比较. 当故障数据完全返回到正常工况时,故障子空间模型中包含故障的方向数量是不相同的. 两种方法对TE工业过程的12个故障工况的数据实施故障重构,得到的结果如表1所示.

表1 TE工业过程不同故障经本文方法重构后的结果

故障序号	本文方法		传统方法	
	广义主成分个数	故障子空间方向数	主元空间的故障方向	残差空间的故障方向
1	1	1	9	7
2	2	4	9	6
3	1	1	24	5
4	5	1	17	11
5	3	8	4	3
6	5	7	9	1
7	4	8	15	9
8	5	9	24	3
9	7	3	26	3
10	4	4	15	11
11	5	2	8	4
12	3	7	21	6

由表1可以看出,采用本文方法提取的广义主成分和故障子空间,大部分故障工况下的维度都小于传统方法需要的维度. 这说明本文方法在构造故障子空间模型中具有较高的精度,在达到相同效果的前提

下需要较少的方向数量,这在实施故障诊断过程中能够减少计算压力。

5 结论

本文针对故障诊断过程中的故障子空间提取问题,提出了基于广义主成分分析的故障子空间提取算法。相比于传统的故障子空间提取方法,该方法同时考虑正常工况和故障工况的测量数据,针对其中分布和变化规律相同的数据信息提取出正常工况信息子空间,从而提纯故障数据,得到更加准确的故障子空间。该方法所提出的故障子空间对于多个故障诊断、故障库的制备、故障定位等问题具有重要的意义。最后,仿真算例和TE工业过程都验证了本文所提出算法的有效性。

参考文献(References)

- [1] Prakash O, Samantaray A K, Bhattacharyya R. Model-based multi-component adaptive prognosis for hybrid dynamical systems[J]. *Control Engineering Practice*, 2018, 72: 1-18.
- [2] Ge Z Q, Song Z H, Gao F R. Review of recent research on data-based process monitoring[J]. *Industrial & Engineering Chemistry Research*, 2013, 52(10): 3543-3562.
- [3] Liu L, Wang Z S, Zhang H G. Adaptive fault-tolerant tracking control for MIMO discrete-time systems via reinforcement learning algorithm with less learning parameters[J]. *IEEE Transactions on Automation Science and Engineering*, 2017, 14(1): 299-313.
- [4] Zhang Y W, Zhou H, Qin S J, et al. Decentralized fault diagnosis of large-scale processes using multiblock kernel partial least squares[J]. *IEEE Transactions on Industrial Informatics*, 2010, 6(1): 3-10.
- [5] Alcalá C F, Qin S J. Reconstruction-based contribution for process monitoring with kernel principal component analysis[J]. *Industrial & Engineering Chemistry Research*, 2010, 49(17): 7849-7857.
- [6] Alcalá C F, Qin S J. Reconstruction-based contribution for process monitoring[J]. *Automatica*, 2009, 45(7): 1593-1600.
- [7] Lu P, van Kampen E J, de Visser C, et al. Nonlinear aircraft sensor fault reconstruction in the presence of disturbances validated by real flight data[J]. *Control Engineering Practice*, 2016, 49: 112-128.
- [8] Zhao C H, Gao F R. Fault subspace selection approach combined with analysis of relative changes for reconstruction modeling and multifault diagnosis[J]. *IEEE Transactions on Control Systems Technology*, 2016, 24(3): 928-939.
- [9] Yue H H, Qin S J. Reconstruction-based fault identification using a combined index[J]. *Industrial & Engineering Chemistry Research*, 2001, 40(20): 4403-4414.
- [10] 宁超, 陈茂银, 周东华. 基于阈值故障子空间提取算法的多重故障重构[J]. *上海交通大学学报*, 2015, 49(6): 780-785.
(Ning C, Chen M Y, Zhou D H. Fault reconstruction for multiple failure modes based on threshold fault subspace extraction algorithm[J]. *Journal of Shanghai Jiao Tong University*, 2015, 49(6): 780-785.)
- [11] Zhao C H, Sun Y X. Subspace decomposition approach of fault deviations and its application to fault reconstruction[J]. *Control Engineering Practice*, 2013, 21(10): 1396-1409.
- [12] Zhao C H, Gao F R. Fault-relevant principal component analysis (FPCA) method for multivariate statistical modeling and processing monitoring[J]. *Chemometrics and Intelligent Laboratory System*, 2014, 133: 1-16.
- [13] Jackson J E. A user's guide to principal components[M]. New York: Wiley, 1991: 57-221.
- [14] El-Thalji I, Jantunen E. A summary of fault modelling and predictive health monitoring of rolling element bearings[J]. *Mechanical Systems and Signal Processing*, 2015, 60: 252-272.
- [15] Shang J, Chen M, Ji H, et al. Recursive transformed component statistical analysis for incipient fault detection[J]. *Automatica*, 2017, 80: 313-327.
- [16] Jackson J, Mudholkar G. Control procedures for residuals associated with principal component analysis[J]. *Technometrics*, 1979, 21(3): 341-349.
- [17] Feng X, Kong X, Ma H, et al. A novel unified and self-stabilizing algorithm for generalized eigenpairs extraction[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(12): 3032-3044.
- [18] 葛志强, 宋执环, 杨春节. 基于MCUSUM-ICA-PCA的微小故障检测[J]. *浙江大学学报: 工学版*, 2008, 42(3): 373-377.
(Ge Z Q, Song Z H, Yang C J. Small shift detection based on MCUSUM-ICA-PCA[J]. *Journal of Zhejiang University: Engineering Science*, 2008, 42(3): 373-377.)
- [19] Downs J J, Vogel E. A plant-wide industrial process control problem[J]. *Computers & Chemical Engineering*, 1993, 17(3): 245-255.

作者简介

杜柏阳(1990—), 男, 博士生, 从事复杂系统信号处理、流程工业过程故障检测和诊断的研究, E-mail: duboyang.epgc@163.com;

孔祥玉(1967—), 男, 教授, 博士生导师, 从事大型工业过程复杂系统建模与分析、数字信号处理、流程工业过程故障检测和诊断等研究, E-mail: xiangyukong01@163.com;

冯晓伟(1986—), 男, 讲师, 博士, 从事大型工业过程复杂系统建模与分析、数字信号处理、流程工业过程故障检测和诊断等研究, E-mail: xiaowei121416@163.com.

(责任编辑: 李君玲)