

研究论文

基于稀疏性非负矩阵分解的故障监测方法

王帆, 杨雅伟, 谭帅, 侍洪波

(华东理工大学化工过程先进控制和优化技术教育部重点实验室, 上海 200237)

摘要: 提出了基于稀疏性非负矩阵分解 (SNMF) 的故障监测方法。非负矩阵分解 (NMF) 是一种新的降维方法, 可以得到原始数据的低秩近似矩阵。与传统的多元统计过程监控方法如主成分分析 (PCA) 相比, NMF 对潜变量的性质没有假设, 除了非负性的要求。将稀疏编码和非负矩阵分解方法结合在一起, 因为施加了稀疏性的约束, 稀疏性非负矩阵分解方法可以得到对数据更稀疏的表示。在分解时对低秩近似矩阵进行正交化处理, 从而在降维时除去变量中的冗余信息, 将信息集中到更少的投影方向上。然后, 用 SNMF 方法来提取过程的潜变量, 并定义新的监测指标来进行故障监测。使用核密度估计 (KDE) 方法来计算新定义的监测指标的控制上限。最后, 将提出的基于 SNMF 的监测方法应用于 TE 过程来评估其监测性能, 并与基于传统 NMF 和 PCA 的方法进行比较。仿真实验结果表明了所提出新方法的可行性。

关键词: 故障监测; 非负矩阵分解; 主元分析; 稀疏编码; 统计过程监控

DOI: 10.11949/j.issn.0438-1157.20141660

中图分类号: TP 277

文献标志码: A

文章编号: 0438—1157 (2015) 05—1798—08

Fault detection method based on sparse non-negative matrix factorization

WANG Fan, YANG Yawei, TAN Shuai, SHI Hongbo

(Key Laboratory of Advanced Control and Optimization for Chemical Processes of Ministry of Education,
East China University of Science and Technology, Shanghai 200237, China)

Abstract: In this paper, a novel fault detection method based on sparse non-negative matrix factorization (SNMF) is proposed. NMF (non-negative matrix factorization) is a new dimension reduction technique that can find a low-rank matrix approximation from the original data. In contrast to the conventional multivariate statistical process monitoring methods, for example PCA, NMF has no assumption about the nature of latent variables, except for non-negativity. Combining linear sparse coding and NMF, SNMF can learn much sparser representation via imposing sparseness constraints. During factorization, low-rank matrix is orthogonalized to remove redundant information and concentrate information on fewer directions of projection. Then, SNMF is used to extract the latent variables that drive a process and new statistical metrics are defined for fault detection. Kernel density estimation (KDE) is adopted to calculate the confidence limits of defined statistical metrics. Afterwards, the proposed method is applied to the Tennessee Eastman process to evaluate the monitoring performance, comparing with conventional NMF and PCA. The results from the experiment show the feasibility of the new method.

Key words: fault detection; non-negative matrix factorization; PCA; sparse coding; statistical process monitoring

2014-11-05 收到初稿, 2015-01-14 收到修改稿。

联系人: 侍洪波。第一作者: 王帆 (1990—), 男, 硕士研究生。

基金项目: 国家自然科学基金项目 (61374140); 国家自然科学基金青年基金项目 (61403072)。

Received date: 2014-11-05.

Corresponding author: Prof. SHI Hongbo, hbshi@ecust.edu.cn

Foundation item: supported by the National Natural Science Foundation of China (61374140, 61403072).

引言

随着现代工业过程变得越来越复杂,许多工业过程,如化工过程,通常包含大量高维并且互相关联的数据。在这种情况下,多元统计过程监控(MSPM)方法已经被提出,用来从过程数据中提取有用信息,并将其用来进行系统的故障监测。传统的基于数据的多元统计过程监控方法主要包括主成分分析方法(PCA)、偏最小二乘方法(PLS)、独立成分分析方法(ICA)。这些方法已经被成功地用于化工过程的故障监测。

主成分分析是过程监测与故障诊断研究应用最广泛的一种方法,它通过将得到的样本数据方差进行最优线性降维处理,解决了变量之间的相关性^[1]。主成分分析方法提取的特征可以反映出原始数据的全局信息^[2]。基于PCA来进行故障监测的成果已有很多^[3-5]。

然而,基于PCA的故障监测方法,假定过程的潜变量服从多元高斯分布,这与实际的工业情况是不符的,实际过程的一部分变量并不服从高斯分布,这样用PCA对过程进行监控,势必会导致监测性能不好^[6]。而且经过PCA处理之后,潜变量是互相正交且不相关的。最近出现了一种新的降维方法,叫做非负矩阵分解(NMF)^[7],它没有对潜变量的性质进行假设,除了要求其非负。NMF通过从原始数据中找到一个低秩矩阵逼近,可以产生对原始数据的局部特征的稀疏表示。NMF在最近几年吸引了越来越多的关注,在许多应用中都有良好的表现,比如数据挖掘和模式识别^[8-9]。非负矩阵分解的方法要求非负性,这样在进行分解时就只允许加而非减的组合,使非负矩阵分解方法可以学习局部的基于部分的表示^[10]。每一个潜变量激活一个可见变量的子集,或者说是激活一个“部分”。所有的潜变量激活作用叠加在一起,产生一个整体^[2]。从这个角度看,NMF可以找到一个中间表示(基于部分的),介于数据的局部和全局结构之间^[11]。

在化工过程故障监测领域,非负矩阵分解方法的应用很少。Li等^[11]提出了应用传统的非负矩阵分解方法进行非高斯过程的故障监测,用仿真实验证明了非负矩阵分解方法可以用于化工过程的故障监测,而且比基于主成分分析的方法有更广泛的应用能力。随后,Li等^[2]又改进了PNMF(投影非负矩阵分解),提出了GNMF(广义非负矩阵投影),主要目的是减轻对原始数据的非负约束。但是,Li等

虽然在文中提到了NMF可以增强对数据的压缩性和可解释性是由于其基于部分的稀疏表示,也提到了稀疏性是数据压缩的关键属性,但并没有提出关于施加稀疏性的改进。因此,本文提出了采用SNMF(稀疏性非负矩阵分解)^[12]方法来进行化工过程故障监测。由于在非负矩阵分解基础上引入了稀疏编码(sparse coding)方法,SNMF可以得到对数据集更稀疏的表示。另外,对传统的非负矩阵分解得到的原始数据低秩近似矩阵,即系数矩阵 H ,进行正交化处理,从而在降维时除去变量中的冗余信息,得到更少的投影方向。因为施加了稀疏性和正交化,信息集中到了更少的投影方向上,所以相比于传统的非负矩阵分解方法具有优越性。同时,本文提出的稀疏性非负矩阵分解,其算法中的目标函数和更新策略不同于Li等所使用的方法。最后用仿真实验证明了在应用于故障监测时,基于稀疏性非负矩阵分解(SNMF)的方法具有可行性。

1 稀疏性非负矩阵分解

1.1 非负矩阵分解

非负矩阵分解是一种线性的、非负的近似数据表示。给定一个非负数据矩阵 $X \in R^{m \times n}$,NMF可以找到一个近似的分解,将 X 分解为两个非负矩阵的乘积,如式(1)所示,

$$X \approx WH \quad (1)$$

其中, $W \in R^{m \times k}$, $H \in R^{k \times n}$, k 通常选择为 $(m+n)k \leq mn$ ^[7]。 W 是由基向量 w_i 组成的基矩阵, H 是由系数向量 h_i 组成的系数矩阵。经过这种分解方法得到的基矩阵 W 和系数矩阵 H 都具有一定的稀疏性^[7]。

这样,非负矩阵分解问题就归结为一个非线性优化问题,用一个目标函数来测度低秩近似的效果。Lee等^[10]定义了两种简单的目标函数,如式(2)、式(3)所示

$$\arg \min_{W, H} E(W, H) = \frac{1}{2} \|X - WH\|^2 \quad (2)$$

$$\text{s.t. } \forall ij, W_{ij}, H_{ij} \geq 0$$

$$\arg \min_{W, H} F(W, H) = \sum_{i,j} (X_{ij} \lg \frac{X_{ij}}{(WH)_{ij}} - X_{ij} + (WH)_{ij}) \quad (3)$$

$$\text{s.t. } \forall ij, W_{ij}, H_{ij} \geq 0$$

Li等^[2,11]使用的是式(2)所示的目标函数,本文使用的是式(3)所示的目标函数。目标函数的不同,决定了相应的更新策略也是不同的。Lee等^[10]提出

了一种迭代方法来解决这个问题,如式(4)、式(5)所示

$$H_{aj} \leftarrow \frac{H_{aj} \sum_i X_{ij} \frac{W_{ia}}{\sum_l W_{il} H_{lj}}}{\sum_k W_{ka}} \quad (4)$$

$$W_{ia} \leftarrow \frac{W_{ia} \sum_j X_{ij} \frac{H_{aj}}{\sum_l W_{il} H_{lj}}}{\sum_v h_{av}} \quad (5)$$

这个方法叫做乘法更新(MU)算法,可以保证非负性,但是有可能会卡在非平稳点^[13]。使用乘法更新算法解决优化问题,就可以得到 W 和 H 。基矩阵 W 保留了原始数据的空间关系和数据结构,系数矩阵 H 可以被看作是原始数据矩阵的低秩近似。从统计过程监测方法的角度来看,需要观察矩阵 H 来反映工业过程的状态^[11]。因此,基于非负矩阵分解(NMF)的监测模型如式(6)所示

$$X = W\hat{H} + E \quad (6)$$

其中, $\hat{H} = (W^T W)^{-1} W^T X$, $\hat{H}$ 是低秩近似的重建, E 是表示预测误差的残差部分,这和基于PCA的方法比较相似^[11]。相应地,监测指标如式(7)、式(8)所示

$$N^2 = H^T \hat{H} \quad (7)$$

$$\text{SPE} = (X - \hat{X})^T (X - \hat{X}) \quad (8)$$

其中, $\hat{X} = W\hat{H}$ 。

1.2 稀疏编码

稀疏编码的概念指的是一种代表方法。在这种代表方法中,少数单位就可以有效地代表典型数据向量,这就意味着,在表示时,大部分单位的价值接近于零,只有少部分单位作用明显,价值是显著的非零值^[14]。经过非负矩阵分解后得到的矩阵本身具有稀疏性特点,因此,将稀疏编码的方法与非负矩阵分解结合是很合理的。对非负矩阵分解施加一定的稀疏性约束,可以得到更稀疏的表示。

线性稀疏编码模型的思想如下:假设观察到的数据集是一些基向量的线性组合,基于这些基向量的系数代表了为作进一步分析的新的特征向量。稀疏编码方法假定这些潜变量呈现出稀疏性,然后通过寻找一种特定形式的低熵编码来最大化稀疏性,这种低熵编码的概率密度的高峰出现在零周围并且有严重的厚尾性^[15]。

用 $I(x, y)$ 来表示一个图像,它可以表示为一些基函数的线性组合^[12],如式(9)所示

$$I(x, y) = \sum_i a_i f_i(x, y) \quad (9)$$

其中, $f_i(x, y)$ 是基函数, a_i 是系数。如上所述,线性稀疏编码可以表示为一个优化问题,通过最小化式(10)这个损失函数^[12]

$$E = -[\text{preserve information}] - \lambda[\text{sparseness of } a_i] \quad (10)$$

其中, λ 是正常数,决定了第2项相对于第1项的重要性。第1项测度了编码描述图像的好坏,由式(11)来表示^[15]

$$[\text{preserve information}] = -\sum_{x,y} \left[I(x, y) - \sum_i a_i y_i(x, y) \right]^2 \quad (11)$$

1.3 稀疏性非负矩阵分解

传统的非负矩阵分解方法,只对矩阵元素施加了非负约束,然而,对矩阵元素施加稀疏性约束,就可以得到更稀疏的表示,即降维后得到更少的投影方向。因此,把稀疏编码和非负矩阵分解结合在一起就可以得到稀疏性非负矩阵分解。进而,考虑到传统的非负矩阵分解得到的原始数据低秩近似,即系数矩阵 H ,不具有正交性,这就意味着变量之间存在冗余信息,从而考虑对其进行正交化处理。这样在降维时,可以将变量中的冗余信息除去,得到更少的投影方向。使信息集中到更少的投影方向上,与稀疏性的思想也是契合的。此时需要解决的优化问题如式(12)所示^[12]

$$\begin{aligned} \arg \min_{W, H} F(W, H) &= \sum_{i,j} \left(X_{ij} \lg \frac{X_{ij}}{(WH)_{ij}} - X_{ij} + (WH)_{ij} \right) + \\ &\quad \alpha \sum_{k,j} H_{kj} \\ \text{s.t. } W, H &\geq 0, \sum_i W_{ik} = 1, HH^T = I \end{aligned} \quad (12)$$

相应的乘法更新(MU)方法如式(13)~式(15)所示

$$H_{kj} \leftarrow \frac{H_{kj} \sum_i X_{ij} \frac{W_{ik}}{\sum_l W_{il} H_{lj}}}{1 + \alpha} \quad (13)$$

$$W_{ik} \leftarrow \frac{W_{ik} \sum_j X_{ij} \frac{H_{kj}}{\sum_l W_{il} H_{lj}}}{\sum_j H_{kj}} \quad (14)$$

$$W_{ik} \leftarrow \frac{W_{ik}}{\sum_u X_{ku} H_{iu}} \quad (15)$$

这种方法是通过最小化所有 H_{kj} 的和来进行稀疏性的限制,从而对传统的非负矩阵分解方法施加了稀疏性,将其改进成为稀疏性非负矩阵分解。对矩阵 H 的正交化处理体现在约束条件中的 $HH^T = I$,在更新算法求解过程中,体现在式(15)。没有对矩阵 H 施加正交处理时,式(15)原本应该是 $W_{ik} \leftarrow \frac{W_{ik}}{\sum_i W_{ik}}$,在迭代更新时以 HX^T 代替 W^T ,从

而得到式(15)所示的更新迭代公式。这样代替是因为 $X = WH$, 两边转置得到 $X^T = H^T W^T$, 两边左乘矩阵 H 得到 $HX^T = HH^T W^T$, 约束条件 $HH^T = I$, 所以在迭代更新时以 HX^T 代替 W^T , 这样可以保证其正交性。

然后就可以得到如式(6)所示的监测模型。相应地, 监测指标如式(7)和式(8)所示。

1.4 稀疏度比较

本文使用的稀疏性定量表示是基于 L1 范数和 L2 范数之间的关系^[14], 如式(16)所示

$$\text{spareness}(W_i) = \frac{\sqrt{m} - \|W_i\|_1 / \|W_i\|_2}{\sqrt{m} - 1} \quad (16)$$

其中, $W \subset R^{m \times k}$, W_i 是基矩阵 W 的列向量。稀疏度的值会有两种极端情况: 如果 W_i 中所有元素都相等且是非零值, 则稀疏度为零; 如果 W_i 中只有一个元素是非零值, 其余元素值都为零, 则稀疏度为 1。

对于基矩阵 W , 其稀疏度由式(17)计算

$$\text{spareness}(W) = \frac{\sum_{i=1}^k \text{spareness}(W_i)}{k} \quad (17)$$

使用随机生成的矩阵进行计算验证。用 MATLAB 随机产生 100×10 的矩阵进行计算。NMF 和 SNMF 两种方法条件相同, 都取 k 为 5, 循环次数为 1000 次。对矩阵 W 和 H 的初始化都采用随机初始化方法。在 SNMF 方法中, α 的值选取为 10。每个方法运行 10 次, 通过式(16)和式(17)来计算基矩阵 W 的稀疏度, 最后选用 10 次数据的平均结果。计算结果如表 1 所示。

表 1 NMF 和 SNMF 方法中 W 的稀疏度

Table 1 Sparseness of W in NMF and SNMF

Item	W_1	W_2	W_3	W_4	W_5	W
NMF	0.180	0.176	0.173	0.169	0.189	0.178
SNMF	0.203	0.257	0.231	0.193	0.211	0.219

如表 1 所示, 在随机计算中, 经过 SNMF 方法得到的基矩阵 W 的稀疏度大于传统 NMF 方法得到的稀疏度。通过稀疏度比较可以证明, 由于施加了稀疏性的约束和进行了正交化处理, SNMF 可以得到比传统 NMF 更加稀疏的表示, 从而具有了对数据更好的表示能力。

2 基于稀疏性非负矩阵分解的故障监测方法

基于上述提出的稀疏性非负矩阵分解, 这里提

出一种新的故障监测方法。

2.1 初始化

作为一种迭代算法, 对 W 和 H 的初始化就显得很重要。因为不同的初始化条件可能会导致不同的结果。正随机数初始化和奇异值分解 (SVD) 是比较常见的初始化方法, 其中 SVD 是初始化基矩阵 W 的较好方法^[16]。

本文采用 PCA 中的负荷矩阵 P 的绝对值作为基矩阵 W 的初始化, 这样做方便了两种算法的比较, 而且 PCA 选取主元的数目可以作为选择基矩阵 W 维数大小的参考。系数矩阵 H 的初始化采用正随机数的方法。

2.2 监测指标

如前文所示, 基于稀疏性非负矩阵分解 (SNMF) 的监测模型如式(6)所示。相应地, 监测指标如式(7)、式(8)所示。

2.3 控制限

核密度估计 (KDE) 被用来确定两个监测指标 N^2 和 SPE 的控制上限。核密度估计是一类数据驱动的技术, 用于密度函数的非参数估计。这是一个很有力的工具, 可以从给定的样本集中得到密度函数的经验分布^[17]。 N^2 和 SPE 两个统计量都是单变量, 所以很适合用 KDE 这个方法来进行计算。

样本集 $X = \{x_i, i = 1, 2, \dots, n\}$, 密度函数为 $P(x)$, 可以用式(18)表示

$$P(x < \theta) = \int_{-\infty}^{\theta} P(x) dx \quad (18)$$

对单变量核估计的 KDE 方程如式(19)所示^[18]

$$\hat{P}(x) = \frac{1}{nh} \sum_i K\left(\frac{x - x_i}{h}\right) \quad (19)$$

其中, $\hat{P}(x)$ 是概率密度函数的估计, n 是样本数, h 是带宽, $K(\cdot)$ 是核函数, 可以是多项式函数、高斯函数等, 这里使用高斯函数。对核函数的要求如式(20)所示

$$\int_{-\infty}^{+\infty} K(x) dx = 1, K(x) \geq 0 \quad (20)$$

最后, 给定 95% 的置信区间, 就可以通过 KDE 计算出两个监测指标 N^2 和 SPE 的控制上限。

2.4 故障监测流程

基于稀疏性非负矩阵分解 (SNMF) 方法的故障监测流程如下。

建模:

(1) 初始化, 正常样本数据 X 经过 PCA 后得到负荷矩阵 P , 用来初始化 W 。 H 用正随机数方法来进行初始化;

(2) 用稀疏性非负矩阵分解 (SNMF) 方法来计算得到 W 和 H ;

(3) 根据式(7)和式(8)分别计算统计量 N^2 和 SPE;

(4) 使用 KDE 计算出两个监测指标 N^2 和 SPE 的控制上限。

监测:

(1) 根据测试数据得到新的系数矩阵 H ;

(2) 根据式(7)和式(8)分别计算新的 N^2 和 SPE;

(3) 通过比较新的 N^2 和 SPE 与相应的控制上限, 判断其是否为故障。

3 仿真实验

3.1 TE 过程

TE 过程可以人为设定 21 种故障, 已经被广泛地用于评估和比较过程监测方法的有效性。TE 过程主要包括反应器、冷凝器、循环压缩机、解吸塔和汽液分离器 5 个单元。有 4 种反应物和 2 种产品,

此外, 还包含 1 种副产品和 1 种惰性气体。该过程共有 41 个测量变量和 12 个操作变量。41 个测量变量中, 22 个是连续的过程测量变量, 19 个是非连续的过程测量变量。Downs 等^[19]给出了关于 TE 过程的详细信息。TE 过程的工艺流程图如图 1 所示。

TE 过程可以仿真正常工作状态和 21 种故障情况。21 种故障类型如表 2 所示。

仿真数据的采样间隔为 3 min, 每一个样本数据集包括 52 个变量。正常数据作为训练数据用来建模。对每一个故障情况, 进行 960 次采样, 故障从第 161 次采样开始产生, 这样就得到了测试数据集。

这里分别将基于 PCA 的监测方法、基于 NMF 的监测方法和基于 SNMF 的监测方法, 应用于 TE 过程, 进行比较。在基于 PCA 的方法中, 选择方差贡献度为 85%, 此时, 对应的 k 值为 27。所以在基于 NMF 和基于 SNMF 的方法中, 相应地将 k 值选取为 27。另外, 这两种方法的循环次数都是 1000 次。在 SNMF 方法中, α 的值选取为 10。这是因

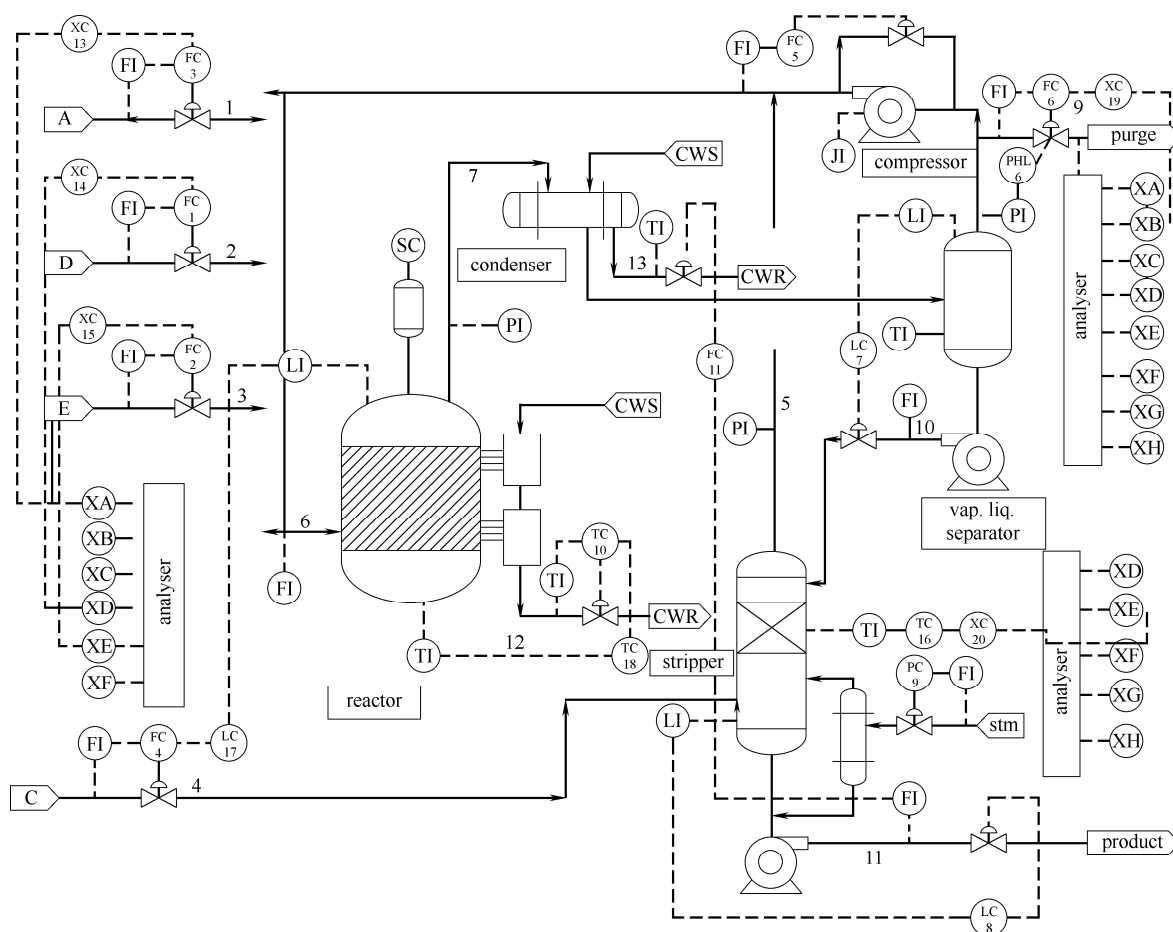


图 1 TE 过程的工艺流程

Fig.1 Process flow sheet of Tennessee Eastman process

表 2 TE 过程 21 种故障类型
Table 2 Process faults for TE process

No.	Description	Type
1	A/C feed ratio, B composition constant (stream 4)	step
2	B composition, A/C ratio constant (stream 4)	step
3	D feed temperature	step
4	reactor cooling water inlet temperature	step
5	condenser cooling water inlet temperature	step
6	A feed loss (stream 1)	step
7	C header pressure loss-reduced availability (stream 4)	step
8	A, B, C feed composition (stream 4)	random variation
9	D feed temperature (stream 2)	random variation
10	C feed temperature (stream 4)	random variation
11	reactor cooling water inlet temperature	random variation
12	condenser cooling water inlet temperature	random variation
13	reaction kinetics	slow drift
14	reactor cooling water valve	sticking
15	condenser cooling water valve	sticking
16	unknown	unknown
17	unknown	unknown
18	unknown	unknown
19	unknown	unknown
20	unknown	unknown
21	value for stream 4 was fixed at steady state position	constant position

表 3 PCA, NMF, SNMF 3 种方法在 TE 过程的故障监测率
Table 3 Fault detection rates of PCA, NMF and SNMF in TE process

Fault number	PCA		NMF		SNMF	
	$T^2/\%$	SPE/%	$N^2/\%$	SPE/%	$N^2/\%$	SPE/%
1	99.25	99.875	99.625	99.625	99.75	99.125
2	98.625	98.625	98.625	98.25	98.625	98.5
4	26.125	100	87.875	96.5	99.375	99.25
5	22.875	35.375	24.625	42.125	32.125	53
6	98.75	100	100	99.875	100	99.875
7	100	100	100	99.875	100	100
8	97.25	97.375	96.875	93.75	97.375	95.5
10	24.25	47.5	37.625	47.125	38.625	53.375
11	43.5	70.875	36.875	62.125	63	64.875
12	98.25	96.125	91	92.25	90.75	93
13	94.5	95	93.25	91.5	93.875	93.125
14	98.5	100	86.875	87.625	91.875	90.125
16	8.5	46.125	53.875	63.625	66.5	64.625
17	76.25	95.75	55.875	57.25	59	60.75
19	6.375	20.875	18.25	21.5	20.625	22.875
20	24.125	56.5	37.75	53.875	50.25	53

为, 由迭代更新公式(13)~式(15)可以看出, 当 α 的值选取为 0 时, SNMF 退化为传统的 NMF; 而当 α 的值选取过大时, 不利于迭代算法的进行。 α 的取值过小和过大都对结果有不利的影响, 所以经过尝

试, α 的值选取为 10。

3.2 结果讨论

故障监测率作为评价监测好坏的标准。3 种方法的监测结果如表 3 所示。

由于故障 3、9 和 15 的测试数据在均值、方差和高阶矩上没有可观察的变化, 所以统计方法的故障监测率都很低^[20], 这里没有进行分析讨论。对于故障 18, 由于 NMF 和 SNMF 这两种方法的误报率过高, 所以不予讨论。对于故障 21, 由于 PCA 的监测指标 SPE 的误报率过高, 所以也不予讨论。对于其余的故障情况, 基于 NMF 和 SNMF 的方法, 误报率始终在 5% 以内, 而基于 PCA 的方法的 SPE 指标, 在故障 8、12、17 和 19 时, 误报率都稍微超过了 5%。可以从表 3 中看出, 对于故障 1、2、6、7、8, 3 种方法得到的结果基本相同, 这是因为这些故障情况的故障程度很明显, 所以 3 种方法都能取得很好的故障监测率。

对于故障 4、5、10、16、19 和 20, 基于 SNMF 的方法都好于基于 PCA 的方法。有些情况下, SNMF 的结果不如 PCA, 这是因为 SNMF 降维后的投影方向相当于 PCA 的主元, 因为稀疏性约束和正交化处理, 基于 SNMF 的方法将信息集中到了较少的投影方向上, 可能会丢失一些与这些投影方向相关性差的信息。而 PCA 的主元数较多, 相比之下, 包含更多的信息。但同时也注意到, 在这些情况下, SNMF 的结果与 PCA 相差不多, 说明 SNMF 用很少的投影方向依然捕捉到了相关信息。

对于故障 4、5、10、11、13、14、16、17、19 和 20, 基于 SNMF 的方法都好于基于 NMF 的方法。故障 4、11、16 和 20 的 N^2 指标, SNMF 相比于 NMF 有很明显的提高 (超过 10%), 在故障 11 上提高了 26.125%。最值得注意的是在故障 4 上的表现, 基于 PCA 的方法, 监测指标 T^2 的故障监测率很低, 而基于 NMF 和 SNMF 的方法, 监测指标 N^2 有较高的故障监测率。在故障 4 的监测上, 也可以很明显地体现出基于 SNMF 的方法好于基本的 NMF 方法。3 种方法对故障 4 的监测情况如图 2 所示。

从整体上看, 基于 NMF 和 SNMF 的方法好于基于 PCA 的方法, 而且由于加入了稀疏性的约束和正交化的处理, 基于 SNMF 的方法可以得到更稀疏的表示, 因此其监测效果较基于传统的 NMF 的方法, 有了一定程度的提高。

4 结 论

相比于主成分分析 (PCA), 非负矩阵分解 (NMF) 是一种更普遍的方法, 因为这种方法对潜变量没有除了非负要求之外的任何假设。在非负性限制的基础上, 再加入稀疏性约束和正交化处理,

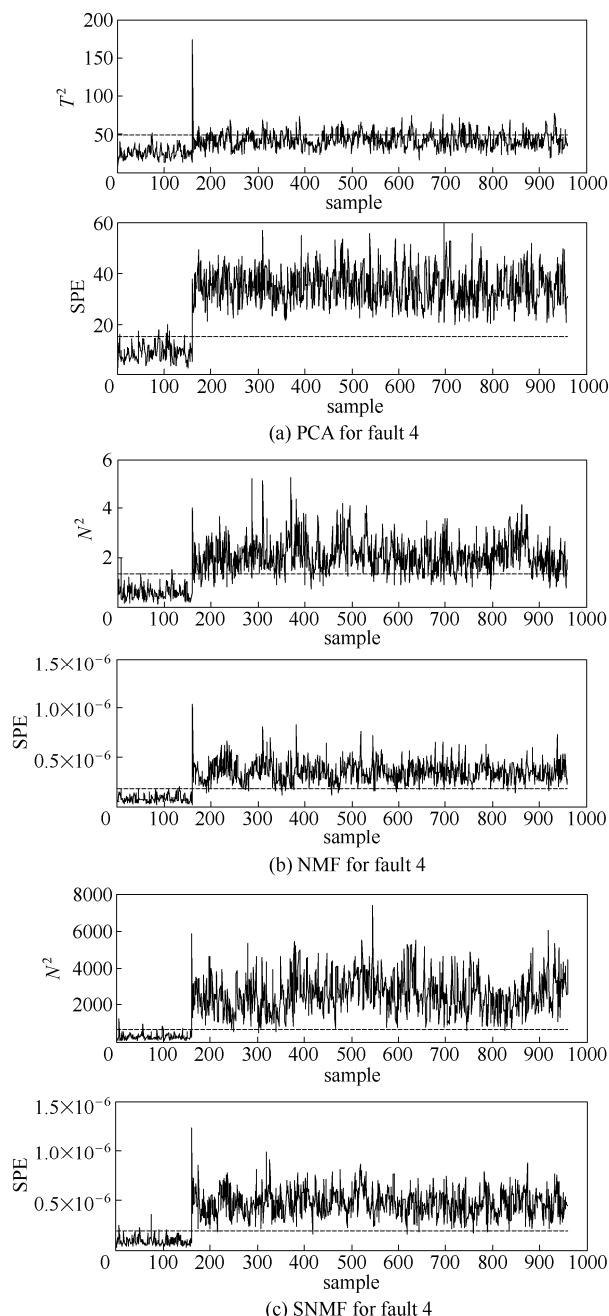


图 2 PCA, NMF 和 SNMF 对故障 4 的监测结果

Fig.2 Monitoring results of PCA, NMF and SNMF for fault 4

可以更好地提高这种方法对数据集的表示能力。本文提出了使用稀疏性非负矩阵分解 (SNMF) 的方法来进行化工过程的故障监测, 因为将稀疏编码的概念引入到传统的非负矩阵分解方法上, 所以可以得到对数据集更稀疏的表示能力, 另外, 又对低秩近似矩阵进行正交化处理使得降维时除去了变量的冗余信息, 将信息集中在更少的投影方向上, 进而在故障监测时得到更好的效果。仿真实验的结果表明, 在 TE 过程的监控上, 基于稀疏性非负矩阵分

解(SNMF)的故障监测方法好于传统的非负矩阵分解(NMF)方法和主成分分析(PCA)方法。

然而, 非负矩阵分解的方法在近几年还在不断的发展中, 以后会出现更多关于此方法的改进。另外, 关于该方法应用于化工过程的故障监测时, 还有一些值得考虑的问题, 比如初始化问题和收敛性问题。

References

- [1] Xia Luyue (夏陆岳), Pan Haitian (潘海天), Zhou Mengfei (周猛飞), Cai Yijun (蔡亦军), Sun Xiaofang (孙小方). Process monitoring and fault diagnosis of propylene polymerization based on improved multiscale principle component analysis [J]. *CIESC Journal* (化工学报), 2011, **62** (8): 2312-2317
- [2] Li Xiangbao, Yang Yupu, Zhang Weidong. Statistical process monitoring via generalized non-negative matrix projection [J]. *Chemometrics and Intelligent Laboratory Systems*, 2013, **121**: 15-25
- [3] Ge Zhiqiang, Song Zhihuan. Distributed PCA model for plant-wide process monitoring [J]. *Industrial & Engineering Chemistry Research*, 2013, **52**: 1947-1957
- [4] Ge Zhiqiang, Song Zhihuan. Mixture Bayesian regularization method of PPCA for multimode process monitoring [J]. *AIChE Journal*, 2010, **56**: 2838-2849
- [5] Jia Mingxing, Xu Hengyuan, Liu Xiaofei, Wang Ning. The optimization of the kind and parameters of kernel function in KPCA for process monitoring [J]. *Computers and Chemical Engineering*, 2012, **46**: 94-104
- [6] Ge Zhiqiang (葛志强), Song Zhihuan (宋执环). New online monitoring method for multiple operating modes process [J]. *Journal of Chemical Industry and Engineering (China)* (化工学报), 2008, **59** (1): 135-141
- [7] Lee D D, Seung H S. Learning the parts of objects by nonnegative matrix factorization [J]. *Nature*, 1999, **401**: 788-791
- [8] Shang Fanhua, Jiao L C, Wang Fei. Graph dual regularization non-negative matrix factorization for co-clustering [J]. *Pattern Recognition*, 2012, **45**: 2237-2250
- [9] Zeng Kun, Yu Jun, Li Cuihua, You Jane, Jin Taisong. Image clustering by hyper-graph regularized non-negative matrix factorization [J]. *Neurocomputing*, 2014, **138**: 209-217
- [10] Lee D D, Seung H S. Algorithms for non-negative matrix factorization // *Advance of Neural Information Process System* [C]. MIT Press, Cambridge, MA, USA, 2001
- [11] Li Xiangbao, Yang Yupu, Zhang Weidong. Fault detection method for non-Gaussian processes based on non-negative matrix factorization [J]. *Asia-Pac. J. Chem. Eng.*, 2013, **8**: 362-370
- [12] Liu Weixiang, Zheng Nanning, Lu Xiaofeng. Non-negative matrix factorization for visual coding // *Proceedings of 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing* [C]. 2003, **3**: 293-296
- [13] Gonzalez E, Zhang Y. Accelerating the Lee-Seung algorithm for nonnegative matrix factorization [R]. Technical Report TR-05-02, Rice University, 2005
- [14] Hoyer P. Non-negative matrix factorization with sparseness constraints [J]. *Journal of Machine Learning Research*, 2004, **5**: 1457-1469
- [15] Olshausen B A, Field D J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images [J]. *Nature*, 1996, **381**: 607-609
- [16] Albright R, Cox J, Duling D, Langville A N, Meyer C D. Algorithms, initializations, and convergence for the nonnegative matrix factorization [R]. NCSU Technical Report Math 81706, North Caroline State University, 2006
- [17] Chen Q, Wynne R J, Goulding P, Sandoz D. The application of principal component analysis and kernel density estimation to enhance process monitoring [J]. *Control Engineering Practice*, 2000, **8**: 531-543
- [18] Botev Z I, Kroese D P. The generalized cross entropy method, with applications to probability density estimation [J]. *Methodology and Computing in Applied Probability*, 2011, **13**: 1-27
- [19] Downs J J, Vogel E. A plant-wide industrial process control problem [J]. *Computers and Chemical Engineering*, 1993, **17**: 245-255
- [20] Yu Jianbo. Local and global principal component analysis for process monitoring [J]. *Journal of Process Control*, 2012, **22**: 1358-1373