

1-2-2019

## Fault Detection Based on GPNMF for Industrial Process

Yuguang Niu

*1.State Key Laboratory for Alternate Electric Power System with Renewable Energy Source, North China Electric Power University, Beijing 102206, China; ;*

Shilin Wang

*2.School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; ;*

Zhongwei Lin

*1.State Key Laboratory for Alternate Electric Power System with Renewable Energy Source, North China Electric Power University, Beijing 102206, China; ;2.School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; ;*

Xiaoming Li

*3.School of Automation Engineering, Northeast Dianli University, Jilin 132012, 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.

---

## Fault Detection Based on GPNMF for Industrial Process

### Abstract

**Abstract:** As a newly dimension reduction technique, non-negative matrix factorization (NMF) has been applied in varying research areas. NMF methods require the original data non-negative. However, the operating data of industrial process maybe not satisfy this restriction. To resolve the problem, *a new method is presented, which can be called as generalized projection non-negative matrix factorization (GPNMF). We use GPNMF to extract the latent variables that drive a process and to combine them with process monitoring techniques for fault detection. The corresponding contribution plots are defined for fault isolation.* The proposed method is applied to a 1 000 MW unit boiler process. The simulation results clearly illustrate the feasibility of the proposed method.

### Keywords

fault detection, fault isolation, generalized projection non-negative matrix factorization, boiler process

### Recommended Citation

Niu Yuguang, Wang Shilin, Lin Zhongwei, Li Xiaoming. Fault Detection Based on GPNMF for Industrial Process[J]. Journal of System Simulation, 2018, 30(2): 521-532.

# 基于 GPNMF 的工业过程故障检测

牛玉广<sup>1</sup>, 王世林<sup>2</sup>, 林忠伟<sup>1,2</sup>, 李晓明<sup>3</sup>

(1.华北电力大学新能源电力系统国家重点实验室, 北京 102206; 2.华北电力大学控制与计算机工程学院, 北京 102206;  
3.东北电力大学自动化工程学院, 吉林 132012)

**摘要:** 非负矩阵分解(NMF)作为一种新的矩阵降维技术已经广泛应用于不同的科学领域。NMF 要求待分解矩阵元素均为非负值, 但是, 实际工业过程所产生的运行数据并不能保证都是非负的。针对这一问题, 提出一种新算法——广义投影非负矩阵分解(GPNMF)。利用 GPNMF 提取测量矩阵中包含过程运行特征的隐变量信息, 使之与过程监控技术相结合来实现工业过程的故障检测, 并构建相应的贡献图法来实现故障分离。将所提算法应用于国内某电厂 1 000 MW 机组锅炉系统, 实验结果验证了新方法对故障检测及分离的有效性。

**关键词:** 故障检测; 故障分离; 广义投影非负矩阵分解; 锅炉过程

中图分类号: TP277

文献标识码: A

文章编号: 1004-731X (2018) 02-0521-12

DOI: 10.16182/j.issn1004731x.joss.201802020

## Fault Detection Based on GPNMF for Industrial Process

Niu Yuguang<sup>1</sup>, Wang Shilin<sup>2</sup>, Lin Zhongwei<sup>1,2</sup>, Li Xiaoming<sup>3</sup>

(1. State Key Laboratory for Alternate Electric Power System with Renewable Energy Source, North China Electric Power University, Beijing 102206, China; 2. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; 3. School of Automation Engineering, Northeast Dianli University, Jilin 132012, China)

**Abstract:** As a newly dimension reduction technique, non-negative matrix factorization (NMF) has been applied in varying research areas. NMF methods require the original data non-negative. However, the operating data of industrial process maybe not satisfy this restriction. To resolve the problem, a new method is presented, which can be called as generalized projection non-negative matrix factorization (GPNMF). We use GPNMF to extract the latent variables that drive a process and to combine them with process monitoring techniques for fault detection. The corresponding contribution plots are defined for fault isolation. The proposed method is applied to a 1 000 MW unit boiler process. The simulation results clearly illustrate the feasibility of the proposed method.

**Keywords:** fault detection; fault isolation; generalized projection non-negative matrix factorization; boiler process

## 引言

随着现代化工业过程日趋复杂, 每天都会产生



收稿日期: 2017-01-10 修回日期: 2017-02-28;  
基金项目: 国家自然科学基金青年基金(51606033);  
作者简介: 牛玉广(1964-), 男, 河南, 博士, 教授, 研究方向为新能源电力系统建模与控制, 大型火电机组优化控制与故障诊断; 王世林(1987-), 男, 河北, 博士生, 研究方向为大型火电机组控制系统故障诊断。

海量的包含过程运行特征的历史数据。如何从海量数据中及时发现工业过程在运行过程中所出现的异常工况、设备失灵、传感器故障等故障信息已成为工业过程故障检测与诊断研究领域中的热点问题。近年来, 基于多元统计过程监控(MSPM)技术的故障检测方法已经成功的在工业过程中得到应用<sup>[1]</sup>。常用的 MSPM 方法有主元分析<sup>[2]</sup>(Principal

Component Analysis, PCA)、独立主元分析<sup>[3]</sup>(Independent Component Analysis, ICA)和非负矩阵分解<sup>[4]</sup>(Non-negative Matrix Factorization, NMF)等。

PCA 通过对过程变量的协方差矩阵进行分解以实现数据降维的目的,分解所得的各主元变量之间是互不相关的,每一个主元变量都可以通过过程变量的线性组合得到。PCA 方法得到的主元子空间反映了过程变量的主要变化,而残差空间则主要反映了过程的噪声和干扰等<sup>[5]</sup>。基于 PCA 的过程监控方法已经取得了众多的研究成果<sup>[6-9]</sup>。基于 PCA 的故障检测方法假设过程变量服从多元正态分布,然而有些情况下过程变量并不完全是正态分布的。针对具有非高斯分布的多个过程变量, Kano 等<sup>[10]</sup>提出了基于 ICA 故障检测方法, ICA 方法的旨在提取过程变量中相互独立且非高斯的少数几个本质变量的信息。基于 ICA 的故障检测方法自提出以来也逐渐成为一个研究热点。

NMF 是一种新兴的矩阵分解技术,与传统的 MSPM 方法假设过程变量服从某一特定分布不同, NMF 除了要求测量矩阵满足非负条件之外没有对过程变量施加任何假设。由于在机理上具有隐变量的正向纯加性的特点, NMF 方法在实现数据压缩时,能够基于数据的局部特征来描述数据的信息<sup>[11]</sup>。因此,相比于传统的 MSPM 方法 NMF 表现出了更好的数据压缩和解释能力。

NMF 对原始数据矩阵和分解因子的非负性约束也导致了分解结果具有天然的稀疏性。稀疏性是数据压缩技术中的一个关键属性。Saito<sup>[12]</sup>认为在矩阵分解过程中应优先考虑稀疏性而不是统计独立性。Ding 等<sup>[13]</sup>和 Yang 等<sup>[14]</sup>证明如果给 NMF 的分解因子施加一个正交性约束就可以显著提高算法的稀疏性。Yuan 和 Oja<sup>[15]</sup>通过引入一个线性变换矩阵,改进了 NMF 的基本框架,提出了投影非负矩阵分解(Projection NMF, PNMF)算法, PNMF 算法可以得到稀疏程度更高的基矩阵。

在处理大规模数据中所展现的优越性使得 NMF 受到了越来越多的关注,目前,已经广泛应

用在数据挖掘、模式识别、聚类及谱分析等研究领域<sup>[16]</sup>。但是,在工业过程故障检测领域, NMF 的应用成果还相对较少。Chikhi 等<sup>[17]</sup>的研究表明 NMF 方法在特征提取方面要优于传统的 MSPM 方法。Li 等<sup>[18]</sup>提出了一种基于传统 NMF 的故障检测方法,并验证了其监控性能要优于基于 PCA 和 ICA 的方法。王帆等<sup>[19]</sup>在基本的 NMF 算法中引入了稀疏编码的概念并对系数矩阵  $H$  进行正交化处理,提出了基于稀疏性非负矩阵分解的故障监测方法,并成功的应用在 TE 过程上。

一般情况下, NMF 算法要求过程变量满足非负条件。然而,在实际工业过程中由于单位选择的不同测量数据有时并不一定都是非负的。同时,为了去除不同量纲和不同数量级大小的数据对检测结果的影响,在进行数据分析之前要对原始数据进行归一化处理,数据预处理后测量矩阵中将包含负数,不再满足 NMF 算法中对数据的非负约束条件。针对以上问题,为了使 NMF 算法的应用范围更广泛,本文在 PNMF 算法的基础上适当放宽对原始数据的非负约束,提出了一种新的分解算法——广义投影非负矩阵分解(Generalized Projection NMF, GPNMF)算法,并以 GPNMF 为基础构造了新的统计监控模型。

## 1 相关 MSPM 算法介绍

### 1.1 主元分析法

PCA 法将多个过程变量构成的测量空间分解为主元子空间(PCS)和残差子空间(RS)。它们均由测量矩阵  $X = [x_1, x_2, \dots, x_m]^T \in R^{m \times n}$  ( $m$  表示样本个数,  $n$  表示过程变量个数)的协方差矩阵的特征向量来定义。基于 PCA 的过程监控模型可以描述为:

$$X = TP^T + E \quad (1)$$

式中:  $T \in R^{m \times k}$  为得分矩阵;  $P \in R^{n \times k}$  为负载矩阵;  $E$  为残差矩阵;  $k$  表示主元的个数。

为了利用测量数据来判断过程运行状态,通常采用平方预测误差(Squared Prediction Error, SPE)和 Hotelling's  $T^2$  来检测过程是否发生异常。SPE 和

$T^2$  计算如下:

$$T^2 = t^T \Lambda^{-1} t = x^T P \Lambda^{-1} P^T x \quad (2)$$

$$SPE = (x - \hat{x})^T (x - \hat{x}) = x^T (I - PP^T) x \quad (3)$$

式中:  $\Lambda$  为由  $X$  协方差矩阵的前  $k$  个特征值所组成的对角矩阵。其中,  $T^2$  统计量的控制限可以由  $F$  分布求得,  $SPE$  统计量的控制限可以根据残差变量的近似分布来求得<sup>[20]</sup>。

### 1.2 非负矩阵分解法

对一个由  $m$  个变量组成的列向量  $x(x \geq 0)$  进行  $n$  次采样, 记每次采样所得样本向量为  $x_i(i=1, 2, \dots, n)$ , 取  $X = [x_1, x_2, \dots, x_n] \in R^{m \times n}$ , NMF 的目标就是找到两个非负矩阵  $W \in R^{m \times k}$  和  $H \in R^{k \times n}$  使得

$$X \approx WH \quad (4)$$

成立,  $W$  称为基矩阵,  $H$  称为系数矩阵, 一般情况下  $k$  的取值应该满足  $(m+n)k < mn$ <sup>[4]</sup>。

NMF 问题的求解可以归结为一个非线性优化问题, 通过定义目标函数来度量分解结果对原始矩阵的逼近程度。Lee 和 Seung<sup>[11]</sup>采用欧几里德距离的平方来描述  $X$  与  $WH$  间的差异, 其数学形式如下:

$$\min_{W, H} F = E(X \| WH) = \frac{1}{2} \|X - WH\|_F^2 \quad (5)$$

$s.t. \quad W \geq 0, H \geq 0$

研究表明上述目标函数单独以  $W$  或  $H$  为变量时是凸函数, 但同时以  $W$  和  $H$  为变量时却不是凸函数。因此求解上述优化问题不能得到全局最优解只能得到局部最优解。Lee 和 Seung<sup>[11]</sup>利用类似于 EM(Expectation maximization)算法中使用的优化策略对目标函数  $F_X$  进行交替求解得到一组迭代规则, 并在理论上证明了目标函数在该迭代规则下是单调不增的, 其数学形式如下:

$$H_{kj} \leftarrow H_{kj} \frac{(W^T X)_{kj}}{(W^T WH)_{kj}} \quad (6)$$

$$W_{ik} \leftarrow W_{ik} \frac{(XH^T)_{ik}}{(WHH^T)_{ik}} \quad (7)$$

上述迭代规则被称为乘法迭代算法, 作为迄今最为经典和使用最为广泛的单调算法, 乘法迭代算法被当做一种基准算法来检验新算法的性能。

### 1.3 嵌入投影的 PNMF 法

多数 NMF 的改进算法都是着眼于基本 NMF 算法的目标函数, 通过增加不同的正则项来对 NMF 算法施加不同的约束条件。然而, Yuan 和 Oja<sup>[15]</sup>从不同的角度出发, 提出了基于线性投影的 PNMF 算法, 该算法的目标函数为:

$$\min_W F = \frac{1}{2} \|X - WW^T X\|_F^2 \quad (8)$$

$s.t. \quad W \geq 0$

将式(8)展开成  $W$  的函数形式并去掉常数项后为:

$$F(W) = \frac{1}{2} \text{Tr}(-2XX^T WW^T + WW^T XX^T WW^T) \quad (9)$$

式(9)中  $\text{Tr}()$  表示矩阵的迹。

由矩阵微分可得:

$$dF(W) = \frac{1}{2} d(\text{Tr}[-2XX^T WW^T + WW^T XX^T WW^T]) = \text{Tr}((-2W^T XX^T + W^T XX^T WW^T + W^T WW^T XX^T) dW) \quad (10)$$

即得式(8)的梯度矩阵为:

$$\frac{\partial F(W)}{\partial W} = -2XX^T W + WW^T XX^T W + XX^T WW^T W \quad (11)$$

采用梯度下降法来求解式(8)中的优化问题, 可以得到如下迭代规则:

$$W_{ij} = W_{ij} - \delta_{ij} \frac{\partial F}{\partial W_{ij}} \quad (12)$$

式中:  $\delta_{ij}$  为步长, 为了保证在计算过程中  $W$  满足非负约束,  $\delta_{ij}$  的取值被定义为:

$$\delta_{ij} = \frac{W_{ij}}{(WW^T XX^T W)_{ij} + (XX^T WW^T W)_{ij}} \quad (13)$$

将式(13)带入式(12)可得 PNMF 算法的迭代规则为:

$$W_{ij} = W_{ij} \frac{2(XX^T W)_{ij}}{(WW^T XX^T W)_{ij} + (XX^T WW^T W)_{ij}} \quad (14)$$

与传统的 NMF 方法相比, PNMF 的迭代规则只有一个未知变量, 简化了求解的难度, 同时降低了解的不确定性以及迭代初值对求解结果的影响。

## 2 广义投影非负矩阵分解

### 2.1 GPNMF 算法

PNMF 算法通过嵌入线性变换矩阵来提高传统 NMF 算法的性能,这一思路非常值得借鉴。在建立故障检测模型的过程中,从原始数据空间到特征数据空间采用线性映射的方式实现将有效的提高算法的泛化能力。然而,PNMF 算法并不能保证目标函数在迭代过程中是单调递减的,有可能会出现问题震荡,这一点会在下一小节中利用具体实例来进行验证。同时,在实际工业过程中,过程变量所产生的数据不一定是非负的。因此,为了使 NMF 算法更好的适应实际工业过程,本文放宽了对原始数据矩阵  $X$  的非负限制,提出了一种新算法——广义投影非负矩阵分解(GPNMF),其分解形式如下:

$$X \approx XH^T H \quad (15)$$

在 GPNMF 算法中同样采用欧氏距离来衡量原始数据矩阵与分解结果之间的差异,其数学形式如下:

$$\min_H F = \frac{1}{2} \|X - XH^T H\|_F^2 \quad (16)$$

s.t.  $H \geq 0$

上述约束优化问题可以看作是 Tikhonov 于 1963 年提出的正则化最小二乘问题<sup>[21]</sup>的一个典型应用。将式(16)中目标函数  $F$  展开成关于变量  $H$  的函数,并去掉其中有关  $X^T X$  的常数项后,可得:

$$F(H) = \frac{1}{2} \text{Tr}(-2H^T HX^T X + H^T HX^T XH^T H) \quad (17)$$

由矩阵微分可得式(17)的梯度矩阵为:

$$\frac{\partial F(H)}{\partial H} = (-2HX^T X + 2HX^T XH^T H) \quad (18)$$

令  $\frac{\partial F(H)}{\partial H} = 0$ , 由此得到式(16)中最小二乘问题的解为:

$$HX^T X = HX^T XH^T H \quad (19)$$

由于在 GPNMF 算法中要放宽对原始数据矩阵  $X$  的非负性要求,现在考虑  $X$  包含正元素和负

元素的情况,定义  $X_+$  为  $X$  中所有正元素的绝对值,  $X_-$  为  $X$  中所有负元素的绝对值,计算方法如下:

$$X_+ = \frac{1}{2}(|X| + X), \quad X_- = \frac{1}{2}(|X| - X) \quad (20)$$

式中  $|X|$  表示对矩阵  $X$  所有元素取绝对值,这样  $X$  就可以表示为  $X_{\pm} = X_+ - X_-$ 。因此式(19)可以改写为如下形式:

$$\frac{(H[X^T X]_+ + H[X^T X]_- H^T H)_{ij}}{(H[X^T X]_- + H[X^T X]_+ H^T H)_{ij}} = 1 \quad (21)$$

从式(17)可以看出目标函数  $F$  是关于系数矩阵  $H$  的 4 次函数,因此,提出了一种新的迭代规则:

$$H_{ij} = H_{ij} \sqrt[4]{\frac{(H[X^T X]_+ + H[X^T X]_- H^T H)_{ij}}{(H[X^T X]_- + H[X^T X]_+ H^T H)_{ij}}} \quad (22)$$

下面将证明式(22)中迭代规则的收敛性,为了证明式(16)中的优化问题在式(22)的迭代规则下是单调不增的,在此我们引入辅助函数的概念。

定义 1: 称  $G(h, h')$  为  $F(h)$  的一个辅助函数,如果  $G(h, h')$  和  $F(h)$  满足如下条件<sup>[11]</sup>。

$$G(h, h') \geq F(h), \quad G(h, h) = F(h) \quad (23)$$

引理 1: 如果  $G(h, h')$  为  $F(h)$  的一个辅助函数,则  $F(h)$  在如下迭代规则下是单调不增的。

$$h^{t+1} = \arg \min_h G(h, h') \quad (24)$$

证明:

$$F(h^t) = G(h^t, h^t) \geq G(h^{t+1}, h^t) \geq F(h^{t+1})。$$

定理 1: 目标函数  $F(H)$  的值在式(22)的迭代规则下是单调不增的。

证明: 根据  $X_{\pm} = X_+ - X_-$ , 将式(17)改写为如下形式(忽略式中的常数系数):

$$F(H) = \text{Tr}(2H^T H[X^T X]_- + H^T H[X^T X]_+ H^T H - 2H^T H[X^T X]_+ - H^T H[X^T X]_- H^T H) \quad (25)$$

由于  $X \in R^{m \times n}$ ,  $H \in R^{k \times n}$  且  $H_{ij} = (H^T)_{ji}$ , 因此将式(25)写成求和的形式:

$$\begin{aligned}
 F(H) = & 2 \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( [X^T X]_{-} \right)_{jl} H_{ij}^2 + \\
 & \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( [X^T X]_{+} H^T H \right)_{jl} H_{ij}^2 - \\
 & 2 \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( [X^T X]_{+} \right)_{jl} H_{ij}^2 - \\
 & \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( [X^T X]_{-} \right)_{jl} H_{ij}^2 H_{il}^2 \quad (26)
 \end{aligned}$$

寻找到一个合适的辅助函数是证明定理 1 成立的关键因素, 为此根据式 (26) 的数学结构定义了一个新的函数  $G(H, H')$ , 其数学形式如下:

$$\begin{aligned}
 G(H, H') = & \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( [X^T X]_{-} \right)_{jl} \frac{H_{ij}^4 + H_{ij}'^4}{H_{ij}^2} + \\
 & \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( [X^T X]_{+} H^T H' \right)_{jl} \frac{H_{ij}^4}{H_{ij}^2} - \\
 & 2 \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( [X^T X]_{+} \right)_{jl} H_{ij}'^2 \left( 1 + \log \frac{H_{ij}^2}{H_{ij}'^2} \right) - \\
 & \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( [X^T X]_{-} \right)_{jl} H_{ij}^2 H_{il}^2 \left( 1 + \log \frac{H_{ij}^2 H_{il}^2}{H_{ij}'^2 H_{il}'^2} \right) \quad (27)
 \end{aligned}$$

由式 (26) 和 (27) 可以看出  $G(H, H) = F(H)$  是成立的, 只需再证明  $G(H, H') \geq F(H)$  就可说明  $G(H, H')$  是  $F(H)$  的辅助函数。

令  $G(H, H')$  和  $F(H)$  做差可得:

$$\begin{aligned}
 G(H, H') - F(H) = & \sum \left( \left( [X^T X]_{-} \right)_{jl} \frac{H_{ij}^4 + H_{ij}'^4}{H_{ij}^2} - 2H_{ij}^2 \right) + \\
 & \sum \left( \left( [X^T X]_{+} H^T H' \right)_{jl} \frac{H_{ij}^4}{H_{ij}^2} - \left( [X^T X]_{+} H^T H \right)_{jl} H_{ij}^2 \right) + \\
 & 2 \sum \left( \left( [X^T X]_{+} \right)_{jl} \left( H_{ij}^2 - H_{ij}'^2 \left( 1 + \log \frac{H_{ij}^2}{H_{ij}'^2} \right) \right) \right) + \\
 & \sum \left( \left( [X^T X]_{-} \right)_{jl} \left( H_{ij}^2 H_{il}^2 - H_{ij}'^2 H_{il}'^2 \left( 1 + \log \frac{H_{ij}^2 H_{il}^2}{H_{ij}'^2 H_{il}'^2} \right) \right) \right) \quad (28)
 \end{aligned}$$

通过简单的数学运算可以得出结论, 当  $a \geq 0$  时  $a \geq 1 + \log a$  是成立的, 同时  $a^2 + b^2 \geq 2ab$  显然也

是成立的。因此, 在式(28)中第一、三、四项均是大于等于 0 的, 只要再证明第二项是  $\geq 0$  的, 就能证明  $G(H, H')$  是  $F(H)$  的辅助函数。令  $H_{ij} = a_{ij} H'_{ij}$ ,  $L(H)$  等于式(28)中的第二项。将  $H_{ij} = a_{ij} H'_{ij}$  带入  $L(H)$  可得:

$$\begin{aligned}
 L(H) = & \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( [X^T X]_{+} H^T H' \right)_{jl} H_{ij}'^2 a_{ij}^4 - \\
 & \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( [X^T X]_{+} H^T H' \right)_{jl} H_{ij}'^2 a_{ij}^2 a_{il}^2 = \\
 & \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( H^T H' [X^T X]_{+} H^T H' \right)_{jl} (a_{ij}^4 + a_{il}^4) - \\
 & \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( H^T H' [X^T X]_{+} H^T H' \right)_{jl} a_{ij}^2 a_{il}^2 = \\
 & \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n \left( H^T H' [X^T X]_{+} H^T H' \right)_{jl} (a_{ij}^2 - a_{il}^2)^2 \geq 0 \quad (29)
 \end{aligned}$$

因此  $G(H, H') \geq F(H)$  是成立的, 即  $G(H, H')$  是  $F(H)$  的辅助函数。其满足引理 1 中的性质, 因此最小化  $G(H, H')$  的值可以求得 H 的全局最优解, 因为  $G(H, H')$  是关于 H 的严格凸函数, 所以  $G(H, H')$  在一阶偏导数为 0 时取得最小值, 对式 (27) 求关于变量 H 的一阶偏导数并令其等于 0, 可得:

$$\begin{aligned}
 \frac{\partial G(H, H')}{\partial H} = & 4 \left( H [X^T X]_{-} + H [X^T X]_{+} H^T H \right) \frac{H_{ij}^2}{H_{ij}^2} - \\
 & 4 \left( H [X^T X]_{+} + H [X^T X]_{-} H^T H \right) \frac{H_{ij}'^2}{H_{ij}'^2} = 0 \quad (30)
 \end{aligned}$$

整理后可得:

$$H_{ij} = H'_{ij} \sqrt[4]{\frac{\left( H [X^T X]_{+} + H [X^T X]_{-} H^T H \right)_{jl}}{\left( H [X^T X]_{-} + H [X^T X]_{+} H^T H \right)_{jl}}} \quad (31)$$

可以看出式 (31) 的结构与式 (22) 是完全一样的, 式 (31) 是在式 (22) 的迭代规则下所能求得的极限值, 也是构造式 (22) 中迭代规则的理论依据。综上所述, 式 (16) 中约束优化问题的目标函数在式 (22) 中的迭代规则下是单调不减的。

## 2.2 GPNMF 与 PNMF 和 NMF 的对比分析

### 2.2.1 GPNMF 与 PNMF 的收敛性对比

在上一节中提到 PNMF 算法并不能保证目标函数在迭代过程中是单调递减的, 有可能会发生震荡。由于 PNMF 算法采用梯度下降法来求解式(8)中的约束优化问题, 为了确保在迭代过程中每一步的结果都满足非负约束, PNMF 为此对迭代步长的结构进行了专门的设计。然而, 对于梯度下降法来说, 步长的选择是至关重要的, 如果步长的选择不合适的话就会引起算法的震荡, 按照 PNMF 算法中对迭代步长的设计, 当算法迭代初值确定时迭代步长也就确定了, 迭代步长的这种选取方式是不能保证目标函数单调下降的。反观 GPNMF 算法, 通过附录中对定理 1 的推导证明了式(16)中约束优化问题的目标函数在式(22)中的迭代规则下是单调不增的。下面将通过一个实例来验证两种算法的区别, 以国内某火力发电厂 1000MW 机组锅炉过程为实例, 在机组负荷稳定运行在 800MW 时对锅炉过程中给水系统、主汽压系统、主汽温系统 3 个系统共计 33 个主要过程变量进行 1000 次采样, 采样时间为 3s。分别采用 GPNMF 和 PNMF 法对上述数据矩阵进行分解, 分解结果如图 1 和图 2 所示。

由图 1 可以看出 PNMF 算法从迭代开始到结束目标函数的值一直处于剧烈震荡过程中, 这与上文的分析结果是一致的。图 2 中 GPNMF 算法目标函数的值是单调、稳定、收敛的, 并且 GPNMF 算法目标函数的收敛速度非常快, 从迭代一开始目标函数的值就迅速减小, 在进行 5 次迭代之后目标函数的值就趋于平稳不再发生变化。

概括来说, GPNMF 算法的模型性质要优于 PNMF 算法的模型性质。PNMF 算法不能保证目标函数在迭代求解过程中是单调、稳定、收敛的, 不便于实际应用, 而 GPNMF 算法依据定理 1 进行求解, 保证了算法的单调性、稳定性和收敛性, 实用效果好。

### 2.2.2 GPNMF 与 NMF 的稀疏性对比

将式(4)改写成向量的形式为  $x_i = Wh_i$  ( $x_i$  和  $h_i$

分别为 X 和 H 的列向量), 因此, 可以将每个数据样本看作是所有基向量(W 的列向量)的加权和, 权重系数就是 H 相应的列向量  $h_i$ 。如果将 W 的每一列看作是一个局部特征, 则提高 H 每列的稀疏度就意味着可以用更少的局部特征来表征数据样本, 从而提高数据的压缩性和可解释性。

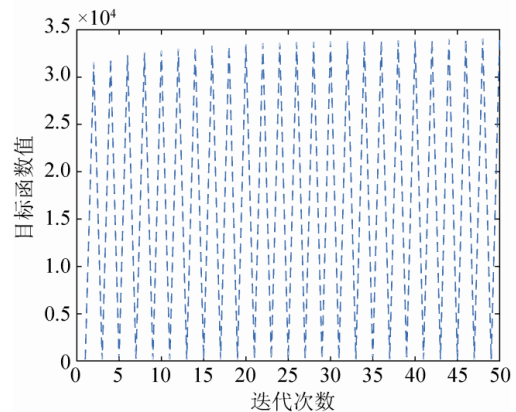


图 1 PNMF 算法迭代曲线

Fig. 1 Iterative curve of PNMF algorithm

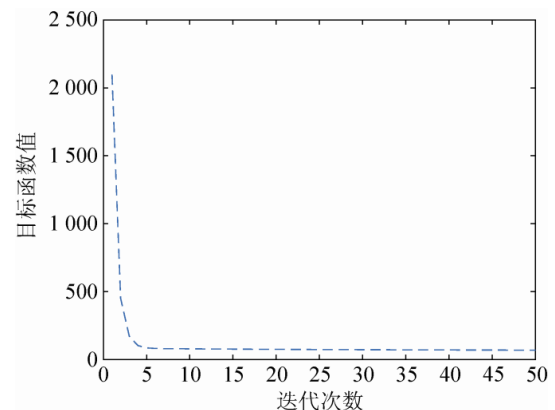


图 2 GPNMF 算法迭代曲线

Fig. 2 Iterative curve of GPNMF algorithm

在 GPNMF 算法中将式(4)改写为式(15)的形式, 要想式(15)成立则要求投影矩阵  $H^T H$  尽可能的接近单位矩阵, 即在求解过程中系数矩阵 H 逐渐趋近于正交。由于 H 是非负的, 因此正交就意味着稀疏, 因为对于一组非负正交向量来说, 每个向量只能有一个元素为非零值。

为了定量的分析 GPNMF 和 NMF 算法分解结果中系数矩阵 H 的稀疏程度, 本文引入 L1 范数和 L2 范数之间的关系来度量一个向量的稀疏程度,



定义如下:

$$sparseness(H_i) = \frac{\sqrt{k} - \|H_i\|_1 / \|H_i\|_2}{\sqrt{k} - 1} \quad (32)$$

式中:  $H_i$  为矩阵 H 的第 i 列; k 为矩阵 H 的维度。由式(32)可以看出当  $H_i$  中只有一个元素为非零其余元素均为零时向量稀疏度取得最大值为 1, 当  $H_i$  中所有元素均不为零时向量稀疏度取得最小值为 0。

采用 GPNMF 和 NMF 算法分别对上一节中锅炉过程原始数据矩阵进行分解, 图 3 给出了分解结果中系数矩阵 H 每一列的稀疏度, 从定量的角度说明了 GPNMF 算法可以得到比 NMF 算法更稀疏的结果。

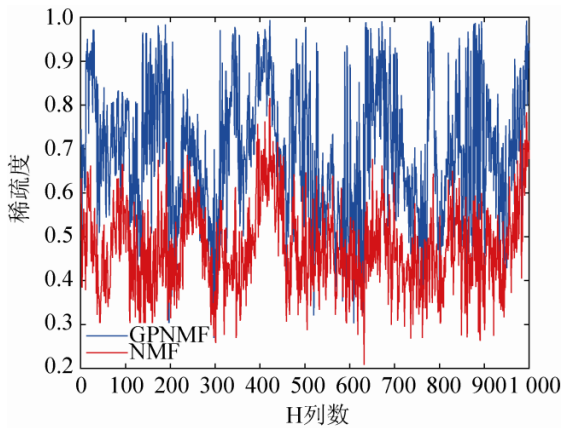


图 3 GPNMF 和 NMF 算法中 H 稀疏度对比  
Fig. 3 Sparseness comparison of H in GPNMF and NMF

综上所述, 本文所提出的 GPNMF 算法, 不仅模型性质比 PNMF 算法好, 而且分解结果比传统的 NMF 算法具有更好的数据描述能力。

### 3 基于 GPNMF 的故障检测方法

#### 3.1 GPNMF 初始化

目前, NMF 及其改进算法都是通过迭代求解的, 众所周知一个好的迭代初值可以提升 NMF 算法的收敛速度和准确性。许多研究者虽然在文章中指出了一个好的初始值对于 NMF 算法的重要性, 但是, 在实际应用过程中多数人仍然采用随机法对 NMF 算法进行初始化。由于 NMF 只能收敛于局部最优解, 不同的初始值将导致不同的结果。在文献[22]中 Langville 等人通过比较几种常用的初

始化方法得知基于 SVD 的初始化方法得到的解要优于其他方法所得到的解。本文在 SVD 算法的基础上提出了一种新的非负双奇异值分解(NDSVD)初始化方法, 该方法包含两个 SVD 过程, 一个用来生成 H 的初始近似矩阵, 另一个利用单位秩矩阵的代数性质使近似矩阵逼近初始矩阵  $H_0$ 。

NDSVD 算法的步骤如下:

GPNMF 初始化方法

1. 输入测量矩阵  $X \in R^{m \times n}$ , 常数  $k < mn / (m+n)$ ;
2. 用 SVD 法分解 X 并取前 k 个最大的奇异值所对应的矩阵为  $[U_{m \times k}, \Sigma_{k \times k}, V_{n \times k}]$ ;
3. 令  $H(:, 1) = \sqrt{S(1, 1)} \times V(:, 1)^T$ ;
4. 从  $j=2$  开始重复执行以下步骤直至  $j=k$ :
  - (1) 令  $U(:, j) = x_+ - x_-$ ,  $V(:, j) = y_+ - y_-$  (其中  $x_+$  和  $y_+$  代表 U 和 V 第 j 列中所有正元素的绝对值,  $x_-$  和  $y_-$  代表 U 和 V 第 j 列中所有负元素的绝对值);
  - (2) 令  $\mu = \|x_+\| \cdot \|y_+\|$ ,  $\xi = \|x_-\| \cdot \|y_-\|$ ;
  - (3) 如果  $\mu > \xi$ , 则  $H(j, :) = \sqrt{S(j, j)} \times \mu \times (y_+ / \|y_+\|)^T$ ;
  - (4) 如果  $\mu \leq \xi$ , 则  $H(j, :) = \sqrt{S(j, j)} \times \xi \times (y_- / \|y_-\|)^T$ 。

#### 3.2 基于 GPNMF 的统计监控模型

基于 GPNMF 的统计过程监控模型可以描述为:

$$X = WH + E \quad (33)$$

式中  $WH$  所有列张成的空间为特征空间, E 表示与特征空间互补的残差空间。  $\hat{H} = (W^T W)^{-1} W^T X$  为系数矩阵 H 的重构值, 它反映了过程变量的变化情况。由于 GPNMF 算法将测量空间分解所得的特征空间和残差空间与 PCA 算法分解所得主元子空间和残差子空间类似, 因此, 本文参照 PCA 方法中监控统计量  $T^2$  和 SPE 的定义, 构造基于 GPNMF 统计监控模型的监控统计量来监控特征空间和残差空间的变化, 新的监控统计量定义如下:

$$T_G^2 = \hat{H}(i)^T \hat{H}(i) \quad (34)$$

$$SPE_G = (x(i) - \hat{x}(i))^T (x(i) - \hat{x}(i)) \quad (35)$$

式中:  $\hat{H}(i)$  和  $x(i)$  分别为重构矩阵  $\hat{H}$  和测量矩阵 X 的第 i 列。  $\hat{x}(i)$  为  $x(i)$  的重构值, 其计算式如下:

$$\hat{x}(i) = W\hat{H}(i) = WW^T x(i) \quad (36)$$

关于监控统计量  $T_G^2$  和  $SPE_G$  控制限的计算, 假设其是服从某一固定分布来确定统计量的控制限显然是不合理的。因此, 本文采用更为一般的方法核密度估计(KDE)来计算监控统计量的实际分布信息, 进而确定其控制限。在计算过程中设定置信区间  $\theta$  为 99%。

### 3.3 基于 GPNMF 的故障检测流程

基于 GPNMF 的故障检测方法主要包括离线建模和在线诊断两部分, 其详细流程如下。

离线建模:

(1) 采集工业过程系统正常运行状态下过程变量的样本矩阵  $X$ , 并对样本矩阵  $X$  做归一化处理。而后用 NDSVD 得到初始系数矩阵  $H_0$ 。

(2) 利用 GPNMF 方法分解  $X$  得到相应的系数矩阵  $H$ 。

(3) 通过式(25)和(26)计算得到相应的监控统计量  $T_G^2$  和  $SPE_G$ 。

(4) 利用 KDE 法分别计算监控统计量  $T_G^2$  和  $SPE_G$  的控制限  $T_{G\theta}^2$  和  $SPE_{G\theta}$ 。

在线诊断:

(1) 将新的测试样本矩阵  $X_{new}$  做归一化处理, 并通过  $\hat{H} = W^T X$  计算得到  $\hat{H}_{new}$ 。

(2) 通过式(25)和(26)计算得到新的监控统计量

$T_{Gnew}^2$  和  $SPE_{Gnew}$ 。

(3) 将  $T_{Gnew}^2$  和  $SPE_{Gnew}$  的值分别与  $T_{G\theta}^2$  和  $SPE_{G\theta}$  作比较, 若超出控制限说明有故障发生, 反之说明系统运行正常。

## 4 仿真实验验证

### 4.1 故障检测

本节以国内某火力发电厂 1 000 MW 机组锅炉过程为实例, 来验证 GPNMF 算法在故障检测中的应用情况。实验过程所需过程数据为 2.2.1 节中所采集的某电厂锅炉过程实际运行数据。

下面我们将所采集的锅炉系统运行数据分为训练集和测试集两部分, 每部分包含 500 个采样点。在测试集中分别模拟传感器的两类故障, 然后采用 GPNMF 算法、传统 NMF 算法以及 PCA 算法分别对其进行检测, 并对检测结果进行对比, 进而说明 GPNMF 算法的有效性。

故障 1: 选取主蒸汽温度 A 作为实验对象, 从测试集第 51 个采样样本开始对该信号叠加实测值的 1% 来模拟传感器恒偏差故障, 三种方法的检测结果如图 4 所示。

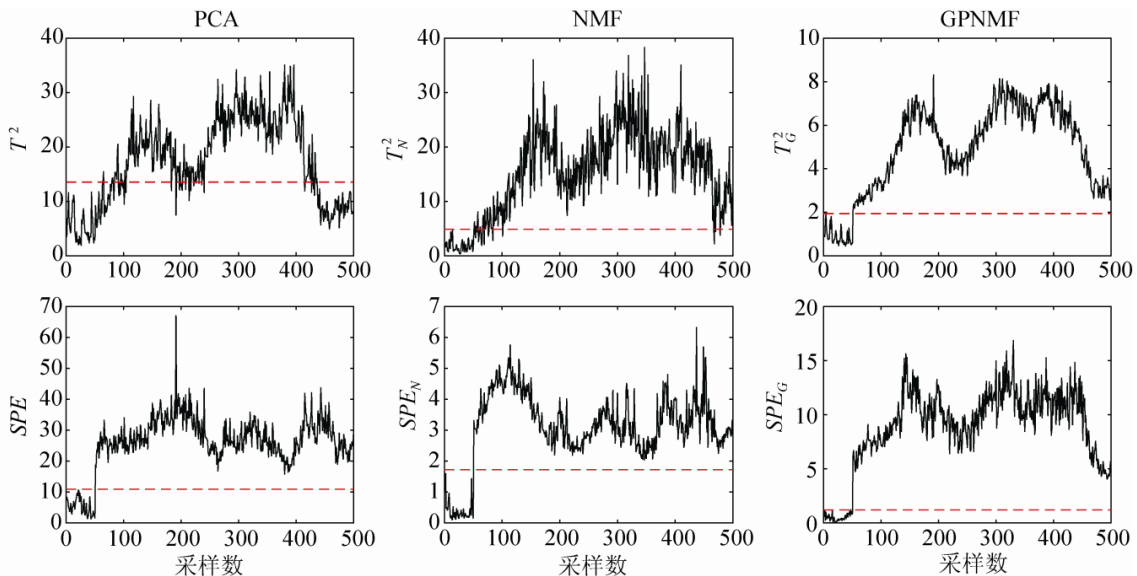


图 4 故障 1 的检测结果

Fig. 4 Fault detection results for fault 1

由于在故障 1 中只对主蒸汽温度 A 一个变量进行了信号叠加, 其他变量均保持稳定不变, 并且叠加的信号幅度很小, 这就使得变量的均值和标准差波动的幅度很小, 从而使得对该故障的检测具有相当大的难度。将所检测出的故障样本占总体故障样本的百分比定义为检测准确率, 由图 4 可以看出, 对于传感器的恒偏差故障, 三种方法的 SPE 统计量均检测出了 100% 的故障样本, 但是 GPNMF 算法统计量越过控制限的幅度比其他两种方法更高, 说明新算法故障检测灵敏度更高。对于另一个监控统计量来说, 由表 1 可以看出, PCA 方法的统计量检测出了 77.75% 的故障样本, 传统 NMF 方法的统计量检测出了 95.56% 的故障样本, 而 GPNMF 方法的统计量则检测出了 100% 的故障样本。进一步的实验表明当故障程度提高到 1.5% 时 NMF 方法的统计量就能完全检测出故障样本, 当故障程度

提高到 3% 时 PCA 方法的统计量才能完全检测出故障样本。这个例子表明, 当只有局部变量发生故障时, GPNMF 可以准确地捕捉到系统局部信息的变化, 同时面对故障幅度较小的情况 GPNMF 也展现出更好的监控性能。

表 1 故障 1 的检测精度  
Tab. 1 Detection accuracy for fault 1

方法	指标	准确率(%)
PCA	$T^2$	77.75
NMF	$T_N^2$	95.56
GPNMF	$T_G^2$	100

故障 2: 选取主蒸汽流量作为实验对象, 从测试集第 21 个采样样本开始对该信号叠加一个斜坡信号(信号上限为主蒸汽流量所有样本平均值的 2%)来模拟传感器的漂移故障, 两种方法的检测结果如图 5 所示。

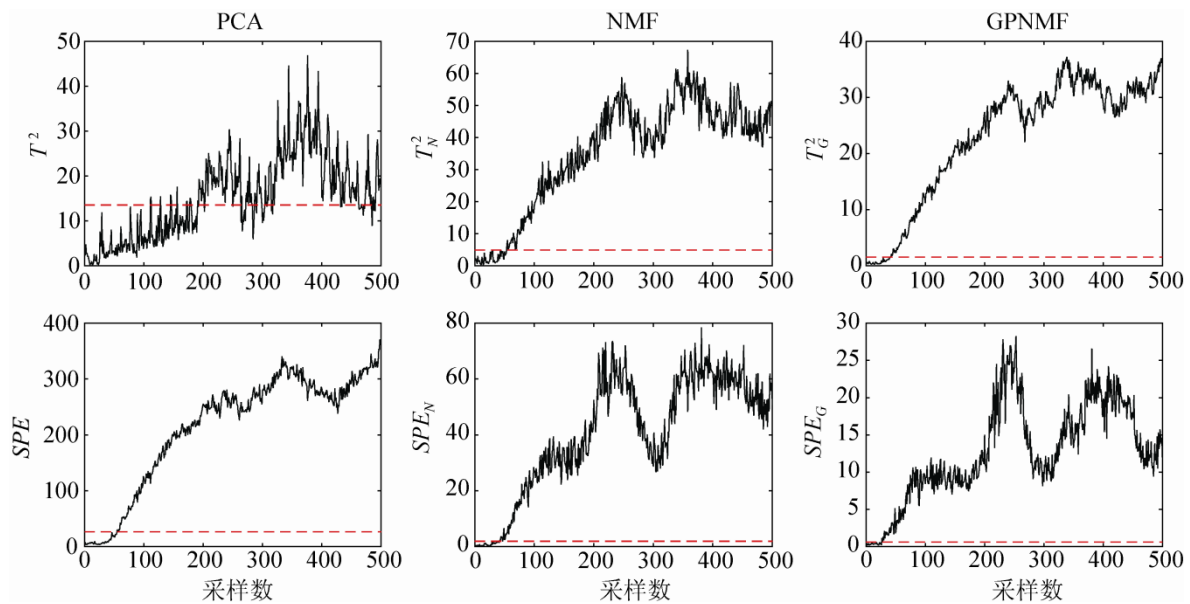


图 5 故障 2 的检测结果  
Fig. 5 Fault detection results for fault 2

传感器的漂移故障属于一种渐变型故障, 在故障刚刚发生时对数据的相关性影响很小, 但随着时间的推移, 故障对数据间相关性的影响将越来越大, 因此三种方法对故障 2 的检测都存在一定程度上的延迟。由图 5 可以看出, 三种方法的监控统计量均只检测出了部分故障样本, 但是, 除 PCA 法的  $T^2$

统计量之外, 其余监控统计量均在第一次检测到故障样本之后就检测出了所有的故障样本, 说明 GPNMF 和 NMF 故障模型的检测精度要优于传统的 PCA 故障模型。从检测速度的角度来说, PCA 方法的  $T^2$  统计量在第 91 个故障样本时第一次检测出故障, SPE 统计量在第 35 个故障样本时第一次

检测到故障。NMF 方法的  $T_N^2$  统计量在第 34 个故障样本时第一次检测出故障,  $SPE_N$  统计量在第 22 个故障样本时第一次检测到故障。而 GPNMF 方法的  $T_G^2$  统计量则是在第 22 个故障样本时第一次检测出故障, 所用时间比 PCA 法缩短了 75.82%, 比 NMF 法缩短了 35.29%,  $SPE_G$  统计量在第 7 个故障样本时第一次检测到故障, 所用时间比 PCA 法缩短了 80%, 比 NMF 法缩短了 68.18%。由此说明, 在面对渐变型故障时, GPNMF 能够比 NMF 和 PCA 更早的检测出故障的存在, 为及时排除系统故障赢得更多宝贵时间。

综上所述, GPNMF 和 NMF 故障模型的检测效果要明显优于传统的 PCA 故障模型。由 GPNMF 算法的分解式可以看出, 要想分解式成立则投影矩阵  $H^T H$  要尽可能的接近单位矩阵, 即系数矩阵  $H$  接近于正交, 正交则意味着稀疏, 因此 GPNMF 算法可以得到更加稀疏的系数矩阵  $H$ , 使得 GPNMF 算法可以用更少的基向量的组合来表征数据样本, 增强了数据的压缩性和可解释性。因此, GPNMF 的故障检测结果较之 NMF 算法有了一定程度的提升。

## 4.2 故障分离

在上一节中通过 GPNMF 故障模型的  $T_G^2$  和  $SPE_G$  监控统计量对系统故障实现了较为精准的检测后, 接下来将对故障进行分离以判断故障发生的位置。本文采用传统的贡献图法对系统故障进行分离。贡献图法是 PCA 算法中常用的一种故障分离方法<sup>[23]</sup>。监控统计量  $T^2$  和  $SPE$  经常被用于贡献图中。其中基于  $T^2$  的贡献图定义为:

$$T^2 = (x^T D x) = \|D^{1/2} x\|^2 = \sum_{i=1}^m Cont_i^{T^2} \quad (37)$$

$$Cont_i^{T^2} = (\xi_i^T D^{1/2} x)^2 = x^T D^{1/2} \xi_i \xi_i^T D^{1/2} x \quad (38)$$

式中:  $m$  为 PCA 中过程变量个数;  $D = P \Lambda^{-1} P^T$ ;  $\xi_i$  表示单位矩阵  $I_m$  的第  $i$  列;  $Cont_i^{T^2}$  表示每个过程变量对  $T^2$  的贡献值。

基于  $SPE$  的贡献图定义为:

$$SPE = \|\tilde{C}x\|^2 = \sum_{i=1}^m Cont_i^{SPE} \quad (39)$$

$$Cont_i^{SPE} = (\xi_i^T \tilde{C}x)^2 \quad (40)$$

式中:  $\tilde{C} = I - PP^T$ ;  $Cont_i^{SPE}$  表示每个过程变量对  $SPE$  的贡献值。

Alcala 和 Qin<sup>[24]</sup>对贡献图法做了细致的分析和总结, 并给出了一个构造贡献图法的基本框架。基于此框架来构造本文所提出的 GPNMF 故障模型中监控统计量  $T_G^2$  和  $SPE_G$  的贡献图, 具体如下:

$$Cont_j^{T_G^2} = (\delta_j^T (WW^T)^{1/2} x)^2 \quad (41)$$

$$Cont_j^{SPE_G} = (\delta_j^T (I - WW^T) x)^2 \quad (42)$$

式中:  $Cont_j^{T_G^2}$  和  $Cont_j^{SPE_G}$  分别表示每个过程变量对监控统计量  $T_G^2$  和  $SPE_G$  的贡献值;  $\delta_j$  表示单位矩阵  $I_m$  ( $m$  为 GPNMF 算法中过程变量的个数) 的第  $i$  列。

下面将采用本节所提出的基于  $T_G^2$  和  $SPE_G$  的贡献图法来确定 4.1 节中引起传感器两类故障的过程变量的位置。由于贡献图法对故障变量的分离准确度与故障检测准确率有着密切的关系, 只有当  $T_G^2$  和  $SPE_G$  能够将系统故障准确的检测出来时, 再对每个变量的贡献值进行计算才是有意义的。因此, 在对故障 1 进行分离时选取第 51 个采样样本进行贡献值的计算, 对故障 2 进行分离时选取第 30 个采样样本进行贡献值的计算。计算结果如图 6~7 所示。

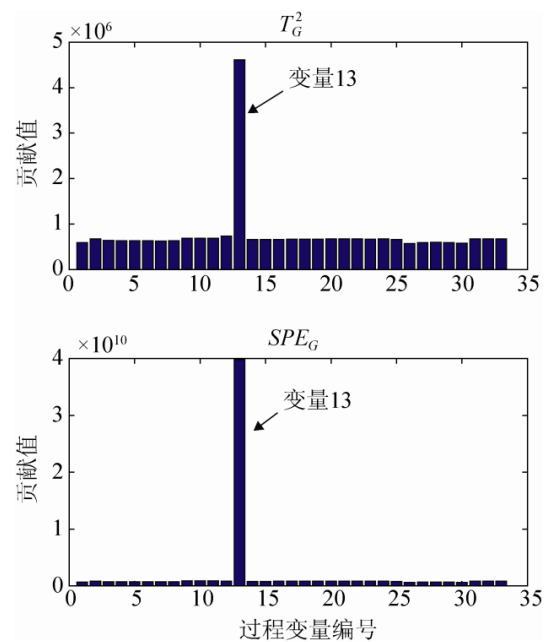


图6 故障1发生时, 所有变量的贡献图  
Fig.6 Contribution plots when fault 1 occurs



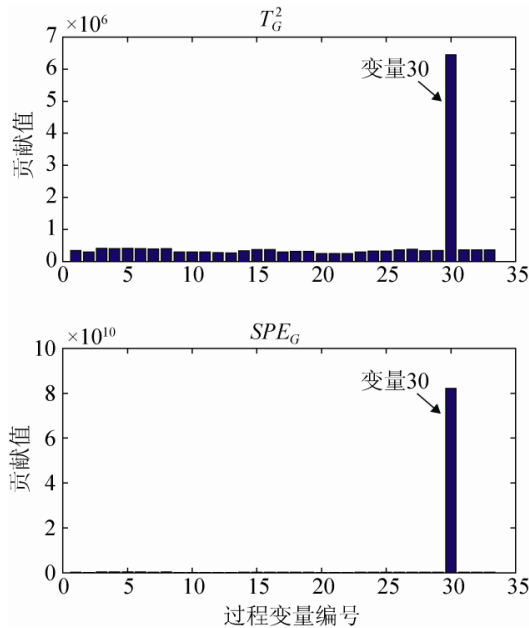


图7 故障2发生时, 所有变量的贡献图  
Fig.7 Contribution plots when fault 2 occurs

在故障1中只对主蒸汽温度A(变量13)一个变量进行了信号叠加, 其他变量均保持稳定不变, 在故障2中只对主蒸汽流量(变量30)一个变量进行了信号叠加, 其他变量均保持稳定不变, 因此可以认为变量13和变量30分别为对故障1和故障2贡献最大的变量, 图6~7中对 $T_G^2$ 和 $SPE_G$ 贡献值的计算也清晰的显示了这个结果。因此可以得出结论: 利用基于 $T_G^2$ 和 $SPE_G$ 的贡献图法对故障变量进行分离是有效的。

## 5 结论

近年来NMF算法已经在统计过程监控领域展现出了巨大的发展潜力。本文提出了一种新的NMF改进算法——广义投影非负矩阵分解算法。该算法不仅解决了PNMF算法在分解过程中震荡不收敛的问题, 而且对系数矩阵H隐含了一个正交约束, 使得GPNMF算法可以得到比NMF算法更加稀疏的H, 增强了数据的压缩性和可解释性。为了将GPNMF算法应用于工业过程故障检测本文提出了一个基于GPNMF算法的故障模型, 并构建了适用于该模型的监控统计量 $T_G^2$ 和 $SPE_G$ , 随后提出了基于 $T_G^2$ 和 $SPE_G$ 的贡献图法用以分离故障

变量。国内某火电厂锅炉过程的故障实例表明本文所提出的GPNMF故障模型以及基于 $T_G^2$ 和 $SPE_G$ 的贡献图法都展现出了良好的检测和故障分离性能。这说明本文所提出的基于GPNMF算法的故障模型是值得作进一步的研究和应用的。

## 参考文献:

- [1] Ge Z, Song Z, Gao F. Review of recent research on data-based process monitoring[J]. Industrial & Engineering Chemistry Research (S0888-5885), 2013, 52(10): 3543-3562.
- [2] Jiang Q, Yan X, Zhao W. Fault detection and diagnosis in chemical processes using sensitive principal component analysis[J]. Industrial & Engineering Chemistry Research (S0888-5885), 2013, 52(4): 1635-1644.
- [3] Mahvash A, Lakis A A. Independent component analysis as applied to vibration source separation and fault diagnosis[J]. Journal of Vibration and Control (S1077-5463), 2016, 22(6): 1682-1692.
- [4] Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization[J]. Nature (S1476-4687), 1999, 401(6755): 788-791.
- [5] 周东华, 胡艳艳. 动态系统的故障诊断技术[J]. 自动化学报, 2009, 35(6): 748-758.  
Zhou Donghua, Hu Yanyan. Fault diagnosis techniques for dynamic systems[J]. Acta Automatica Sinica, 2009, 35(6): 748-758.
- [6] Deng X, Tian X. Entropy principal component analysis and its application to nonlinear chemical process fault diagnosis[J]. Asia-Pacific Journal of Chemical Engineering (S1932-2143), 2014, 9(5): 696-706.
- [7] Shao R, Hu W, Wang Y, et al. The fault feature extraction and classification of gear using principal component analysis and kernel principal component analysis based on the wavelet packet transform[J]. Measurement (S0263-2241), 2014, 54(4): 118-132.
- [8] Zhao Q. The Study on Rotating Machinery Early Fault Diagnosis based on Principal Component Analysis and Fuzzy C-means Algorithm[J]. Journal of Software (S1000-9825), 2013, 8(3): 709-715.
- [9] 高东, 吴重光, 张贝克, 等. 基于PCA和SDG的传感器故障诊断方法研究及应用[J]. 系统仿真学报, 2011, 23(3): 567-573.  
Gao Dong, Wu Chongguang, Zhang Beike, et al. Method combining PCA and SDG for fault diagnosis of sensors and its application[J]. Journal of System Simulation, 2011,

- 23(3): 567-573.
- [10] Kano M, Tanaka S, Hasebe S, et al. Monitoring independent components for fault detection[J]. *AICHE Journal* (S1547-5905), 2003, 49(4): 969-976.
- [11] Lee D D, Seung H S. Algorithms for non-negative matrix factorization[C]// *Advances in neural information processing systems*, 2001: 556-562.
- [12] Saito N. *Modern signal processing*[M]. California: Mathematical Sciences Research Institute, 2004: 317-340.
- [13] Ding C, Li T, Peng W, et al. Orthogonal nonnegative matrix t-factorizations for clustering[C]// *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2006: 126-135.
- [14] Yang Z, Laaksonen J. Multiplicative updates for non-negative projections[J]. *Neurocomputing* (S0925-2312), 2007, 71(1): 363-373.
- [15] Yuan Z, Oja E. Projective nonnegative matrix factorization for image compression and feature extraction[C]// *Scandinavian Conference on Image Analysis*. Springer Berlin Heidelberg, 2005: 333-342.
- [16] Li Z, Wu X, Peng H. Nonnegative matrix factorization on orthogonal subspace[J]. *Pattern Recognition Letters* (S0167-8655), 2010, 31(9): 905-911.
- [17] Chikhi N F, Rothenburger B, Aussenac-Gilles N. A comparison of dimensionality reduction techniques for web structure mining[C]// *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*. IEEE Computer Society, 2007: 116-119.
- [18] Li X, Yang Y, Zhang W. Fault detection method for non-Gaussian processes based on non-negative matrix factorization[J]. *Asia-Pacific Journal of Chemical Engineering* (S1932-2143), 2013, 8(3): 362-370.
- [19] 王帆, 杨雅伟, 谭帅, 等. 基于稀疏性非负矩阵分解的故障监测方法[J]. *化工学报*, 2015, 66(5): 1798-1805.  
Wang Fan, Yang Yawei, Tan Shuai, et al. Fault detection method based on sparse non-negative matrix factorization[J]. *CIESC J*, 2015, 66(5): 1798-1805.
- [20] Jackson J E, Mudholkar G S. Control procedures for residuals associated with principal component analysis[J]. *Technometrics* (S0040-1706), 1979, 21(3): 341-349.
- [21] Tikhonov A. Solution of incorrectly formulated problems and the regularization method[C]// *Soviet Math. Dokl.* 1963, 5: 1035.
- [22] Albright R, Cox J, Duling D, et al. Algorithms, initializations, and convergence for the nonnegative matrix factorization[R]. Tech. rep. 919. NCSU Technical Report Math 81706. North Carolina State University. 2006.
- [23] 周东华, 李钢, 李元. 数据驱动的工业过程故障诊断技术: 基于主元分析与偏最小二乘的方法[M]. 北京: 科学出版社, 2010: 28-30.  
Zhou Donghua, Ligang, Li Yuan. *Data driven fault diagnosis for industrial processes: Based on principal component analysis and partial least squares*[M]. Beijing: Science Press, 2010: 28-30.
- [24] Alcalá C F, Qin S J. Reconstruction-based contribution for process monitoring[J]. *Automatica* (S0005-1098), 2009, 45(7): 1593-1600.

## 《系统仿真学报》荣获“2017 中国国际影响力优秀学术期刊”证书

由中国学术期刊（光盘版）电子杂志社与清华大学图书馆联合成立的中国学术文献国际评价研究中心，发布了 2017 版《中国学术期刊国际引证年报》，《系统仿真学报》荣获“2017 中国国际影响力优秀学术期刊”。

《年报》（2017 版）采用的统计源期刊为 20192 种，涵盖 WoS 收录的 SCI 期刊 8874 种、SSCI 和 A&HCI 期刊 4645 种，ESCI 期刊 5578 种；增补期刊 1762 种。参照中外文学术期刊总被引频次、影响因子、半衰期等各项国际引证指标，计算期刊影响力指数(CI)，对国内 6210 种学术期刊排序，遴选了人文社科、自然科学与工程技术两个类别的 TOP10%为国际影响力品牌学术期刊。TOP5%以内的期刊为“最具国际影响力学术期刊”、TOP5-10%之间的为“国际影响力优秀学术期刊”。

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

• 532 •