

基于主元分析的 Lyocell 工艺过程故障诊断

安康¹ 张建明²

1 中国纺织科学研究院有限公司, 北京 100025

2 中纺院绿色纤维股份公司, 河南 新乡 453600

[摘要] 文章应用主元分析法对 Lyocell 溶胀溶解工艺进行故障诊断研究。选取溶胀溶解工艺过程的 20 个工艺指标进行分析, 通过数据预处理、建立增广矩阵生成 PDCA 模型并确定主元个数; 应用 Q 统计法和 T2 统计法分析故障的产生; 使用主元贡献图法确定故障源。

[关键词] 主元分析法; 动态主元分析法; 故障诊断; Lyocell

DOI: 10.33142/ec.v4i5.3669

中图分类号: TF046.2

文献标识码: A

Fault Diagnosis of Lyocell Process Based on Principal Component Analysis

AN Kang¹, ZHANG Jianming²

1 China Textile Academy, Beijing, 100025, China

2 Greccell Co., Ltd., Xinxiang, Henan, 453600, China

Abstract: In this paper, principal component analysis (PCA) was used to diagnose the failure of Lyocell swelling dissolution process. 20 process indexes of swelling dissolution process were selected for analysis, and PDCA model was generated and the number of principal components was determined by data preprocessing and augmented matrix. Q statistics and T2 statistics are used to analyze the fault. The principal component contribution diagram method is used to determine the fault source.

Keywords: principal component analysis; dynamic principal component analysis; fault diagnosis; Lyocell

Lyocell 纤维生产工艺主要以植物纤维素浆粕为原料, 以 N-甲基吗啉-N-氧化物 (NMMO) 为溶剂, 浆粕溶解于溶剂制得纤维素纺丝原液, 再通过干喷湿纺制成再生纤维素纤维。Lyocell 纤维不仅原料来源丰富, 而且生产工艺环保、清洁, 又具有优异的特性, 代表了未来纺织新材料的发展方向。随着过程控制技术的不断发展, 对于纤维生产过程的稳定性要求越来越高, 如何及时地检测和诊断故障, 已经成为 Lyocell 纤维生产行业普遍存在且急需解决的问题。

Lyocell 纤维生产工艺复杂, 原液溶胀溶解过程相关变量耦合度高, 工艺过程滞后严重, 工艺模型无法准确建立。而利用主元分析法进行故障诊断, 只需要利用典型生产历史数据建立主元分析模型, 并利用该模型进行故障诊断工作。主元分析法对典型生产历史数据进行处理并进行数据空间投影变换, 在最大限度保留原系统空间特征的基础上, 将原有系统投影为数目减少且相互独立的低维数据空间, 并从新的数据空间中提取符合要求的故障诊断模型。

1 主元分析

主元分析法 (PCA) 把数据系统分解成两个子空间, 通过线性空间降维投影技术, 将源数据空间内的过程趋势数据和随机噪声数据分解为一些列线性正交的矩阵向量, 并按照正交向量对系统的贡献值进行排序。试验表明, 线性的系统故障主要影响过程趋势数据空间, 因此可以使用主元分析法来针对过程趋势数据子空间, 从总体上提高过程故障诊断的效率^[3]。

PCA 法选取降维空间内贡献值较大的主元素, 代表源数据空间的数据变化特征。选取源数据空间样本 $X \in R^{n \times m}$ 。其中 n 表示源数据空间样本个数, m 表示工艺过程变量个数。将源数据空间样本 X 表示为 m 个工艺过程变量的外积之和:

$$X = TP^T = t^1 p^1 + t^2 p^2 + \dots + t^m p^m \quad (1)$$

当源数据空间样本存在线性系统相关性时, 工艺过程变量的数据变化趋势将主要体现在贡献值较大几个负荷向量上, 其它负荷向量上的贡献值将会很小, 因此 PCA 模型如下式所示:

$$X = TP^T + E \quad (2)$$

式中 $T = [t^1 \ t^2 \ \dots \ t^m]$ 表示得分矩阵, $P = [p^1 \ p^2 \ \dots \ p^m]$ 表示负荷矩阵, E 表示噪声子空间。

PCA 基于了过程数据的序列无关性^[1], 认为 t 时刻的过程数据值与 t 时刻之后的过程数据值是一系列相互独立的变量。因此主元分析法仍然是基于静态多元统计理论的方法, 对于过程数据的时域自相关性不能进行有效的分析。对于 Lyocell 纤维生产过程来说, 需要定义更短的采样间隔, 以满足生产过程控制的需要, 因此也加大了数据空间的时域自相关性。动态主元分析法 (DPCA) 通过对 PCA 矩阵增加时间自相关序列来解决时域自相关性问题^[2]。DPCA 通过增加工艺过程变量的前 h 个过程数据值建立增广矩阵:

$$X(h) = \begin{bmatrix} x_t^T & x_{t-1}^T & \cdots & x_{t-h}^T \\ x_{t-1}^T & x_{t-2}^T & \cdots & x_{t-h-1}^T \\ \vdots & \vdots & \ddots & \vdots \\ x_{t+h-n}^T & x_{t+h-n-1}^T & \cdots & x_{t-n}^T \end{bmatrix} \quad (3)$$

在增广矩阵中, X^T 是 t 时刻过程数据值的 m 维向量。实验数据表明, 当用 DPCA 进行生产过程建模时, 选取 $h=1$ 或 2 通常是比较合理的。因此基于主元分析的故障诊断方法可以直接推广到动态主元分析法中, 过程变量可以通过与前 h 个过程数据值进行矩阵计算。

2 故障诊断过程

利用主元分析进行故障诊断的过程如图 1 所示:

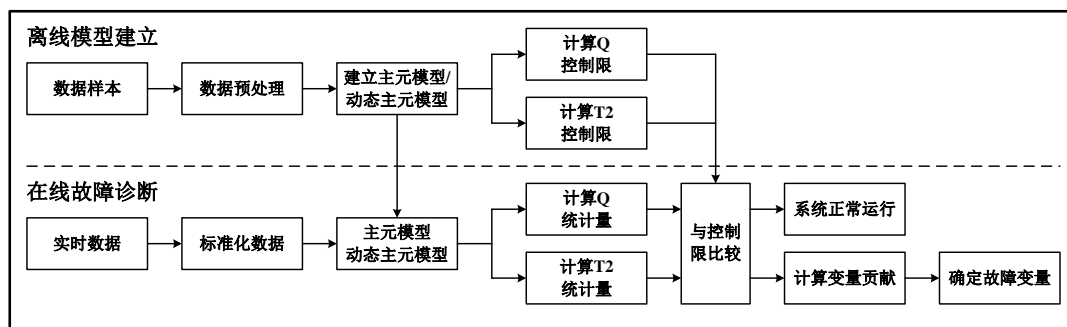


图 1 基于主元分析的故障诊断过程

2.1 数据预处理

主元模型会受到数据维度的影响, 为了从数据采样集中有效地提取信息, 通常需要对数据进行预处理。预处理过程包含: 去除变量、自标定和剔除野点。其中自标定过程使数据集变量标准化, 使每个变量都被赋予相等的权重。将自标定过程应用到源数据空间样本时, 均值和标准差都取源数据空间样本。

$$X^s = [X - (1 \ 1 \ \cdots \ 1)^T M] \text{diag}\left(\frac{1}{s_1}, \frac{1}{s_2}, \cdots, \frac{1}{s_m}\right) \quad (4)$$

其中 $M = [m^1 \ m^2 \ \cdots \ m^m]$ 表示源数据空间样本的均值, $s = [s^1 \ s^2 \ \cdots \ s^m]$ 表示源数据空间样本的标准差。

2.2 主元模型计算方法

计算主元模型需要分解主元向量和负荷向量。在实际应用中通常借助 MATLAB 来完成计算过程。其中 princomp 函数实现对数据采样集的主元分析, 其返回函数即为原始变量的线性组合。

2.3 模型主元个数

在建立主元模型的过程中, 应适当增加主元素的个数, 以减少信息损失和模型误差。通常使用累计方差百分比法确定主元素的个数。

$$CPV = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^m \lambda_i} \quad (5)$$

式中 CPV 为前 k 个主元素的累积方差百分比, k 表示主元素的个数, m 表示过程变量的个数。实验数据表明, 选取 $CPV > 85\%$ 时的主元素个数 k 是比较合理的。

2.4 Q 统计量

Q 统计量表示过程变量值对 PCA 模型的偏离程度。当故障发生时, 该故障投影就会显著增加。对于第 i 个采样点来说 Q 统计量被定义为:

$$Q^i = X^i (I - P^k P^{kT}) X^{iT} \quad (6)$$

其中 $P^k = [p^1 \ p^2 \ \cdots \ p^n]$, I 表示 $n \times n$ 的单位矩阵。Q 统计量的偏离程度如满足下式, 则认为系统发生故障:

$$Q^{new} < Q^\alpha \quad (7)$$

其中 Q^α 为误差空间特征值及检验水相关平的控制限。

2.5 T^2 统计量

T^2 统计量表示测量值在变化趋势和幅值上对 PCA 模型的偏离程度。该统计量是多个故障投影共同作用的结果, 通过主元素的幅值变化的趋势判断故障变化的程度。对于第 i 个采样点来说 T^2 统计量定义为:

$$T^2 = X^i P^k \lambda^{-1} P^{kT} X^{iT} \quad (8)$$

其中 λ 是由前 k 个主元所对应的特征值所组成的对角矩阵。 T^2 统计量的偏离程度如满足下式, 则认为系统发生故障:

$$T^{new} < T^{k,m,\alpha} \quad (9)$$

其中 $T^{k,m,\alpha}$ 为误差空间特征值及检验水相关平的控制限。

2.6 主元贡献图

对 Q 统计量和 T^2 统计量的结果进行分析, 可以判断系统是否出现故障, 进而通过主元贡献图识别故障源。主元贡献就是各变量对残差子空间的贡献值, 其绝对值较大的变量多为引起故障的原因。第 j 个过程变量在第 i 时刻的 Q 统计量的贡献为:

$$Q^{ij} = (X^{ij} - \hat{X}^{ij})^2 \quad (10)$$

其中 X^{ij} 表示第 i 时刻第 j 个过程变量值, \hat{X}^{ij} 表示第 i 时刻第 j 个过程变量的主元模型值。

3 典型工艺过程的故障诊断

选取 Lyocell 溶胀溶解工艺过程的 20 个工艺指标进行分析, 源数据空间样本包含 24h 的生产故障数据信息, 采样周期设定为 1min。建立滞后长度 $h=2$ 的动态主元增广矩阵并进行标准化处理, 使用累计方差百分比法确定源数据空间样本的前 3 个主元素作为主元模型。通过协方差矩阵分别计算 Q 统计量和 T^2 统计量的 90% 控制限。针对生产数据利用主元模型计算统计量, 如果计算结果超出控制限, 则判定生产过程出现故障, 然后利用主元贡献图并结合工艺过程经验确定故障源^[3]。

以 Lyocell 溶胀溶解工艺过程的典型故障数据样本为基础, 建立 3 组源数据空间故障样本。使用 Matlab 进行仿真, 基于主元分析的故障诊断法在源数据空间故障样本 1、2 中都有着不错的表现, 可以准确地对相关工艺故障进行诊断。为了能够比较 PCA 和 DPCA 的优缺点, 以故障数据样本 3 为例, 对故障诊断结果进行比较。故障数据 3 描述了溶胀溶解过程中溶解机真空压力的变化, 压力值在工艺控制点附近上下波动, 并没有大幅度偏离控制范围。主元分析和动态主元分析对于样本 3 的统计结果如图 2 和图 3 所示。虽然两者都能够对故障进行检测, 但由于动态主元分析考虑了时域自相关性问题, 对于检测结果的准确性要优于主元分析。

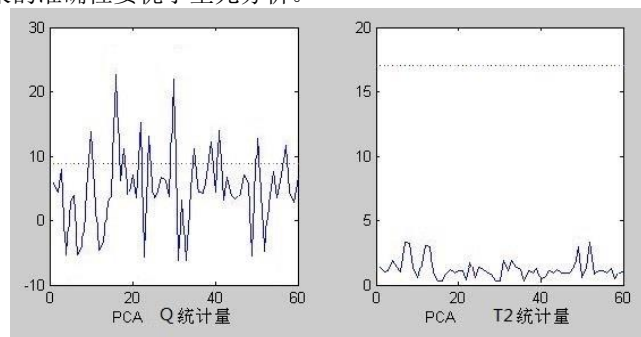


图 2 PCA 对样本 3 的统计量

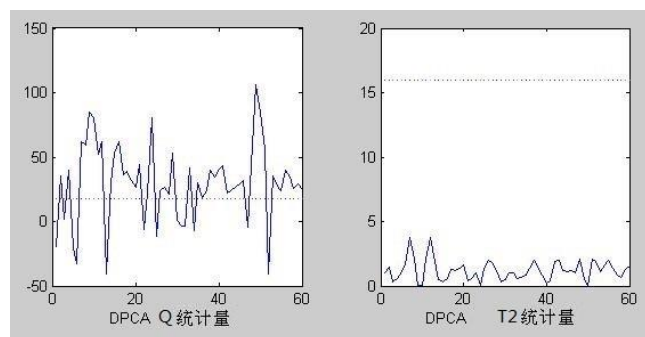


图3 DPCA对样本3的统计量

计算各过程变量的主元贡献值,其绝对值较大的变量即为引起故障的原因。主元分析和动态主元分析对于样本3的主元贡献值如图4和图5所示出。PCA贡献图无法很好的区分故障变量,而DPCA贡献图则具有明显的故障源辨识度。

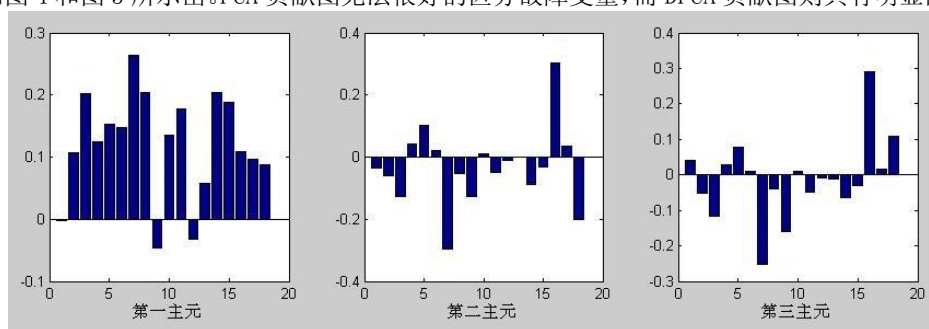


图4 PCA对样本3的主元贡献图

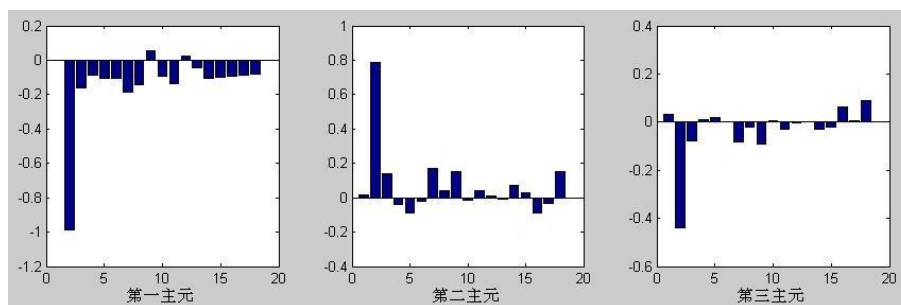


图5 DPCA对样本3的主元贡献图

4 结论

Lyocell 溶胀溶解工艺过程数据不仅包含互相关性,还含有很强的自相关性。使用主元分析法对该工艺过程进行故障诊断,能够很好的实现对于动态过程的故障检测和分离。其中Q统计量对故障的判断更为灵敏,DPCA的故障诊断能力要优于PCA。就多变量故障诊断理论应用来说,主元分析对于非线性过程的故障诊断能力还不够,可以将主元分析法与人工智能相结合^[4],将主元分析法与动态模型故障诊断技术结合起来,也是今后进一步研究的方向。

[参考文献]

- [1]王海清,宋执环,李平.改进PCA及其在过程监测与故障诊断中的应用[J].化工学报,2001(6):471-475.
- [2]费正顺,黄炳强.基于动态主元分析的系统性能评价方法[J].浙江科技学院学报,2015,27(3):197-202.
- [3]李传金,王文标,花义锋,高世泽.基于主元分析的锅炉系统故障检测与诊断[J].机电信息,2016(24):92-93.
- [4]司马锡生.多变量统计过程控制诊断技术的进展[J].南京化工大学学报(自然科学版),2000(6):86-89.

作者简介:安康(1980.3-),男,天津理工大学,控制理论与控制工程专业,硕士,中国纺织科学研究院有限公司,绿纤中心仪电室主任,中级工程师;张建明(1984.12-),男,河南理工大学,自动化专业,学士,中国纺织科学研究院有限公司,设备部技术员,中级工程师。