

Principal Component Analysis: Mechanical Fault Prediction Based on Data-Driven Technique

Luhui LIN¹, Jie MA²

School of Automation, Beijing Information Science & Technology University, Beijing, China

E-mail: 82072847@163.com; mjbeijing@163.com

Abstract: This paper describes the technique based on data-driven. This method does not require a priori knowledge of the process, solely on the basis of process data collected, and can explore the hidden information in the data. And principal component analysis (PCA) is one of the analysis methods of the data-driven. This paper aims to propose a fault prediction method based on PCA which is a multivariate statistical analysis method. It does not need to understand the structure and principles of the system in-depth and can be based on the sensor measurement data of the system running process entirely; and the PCA algorithm, which is simple and easy to implement, can also reduce the noise interference and simplify the data processing.

Keywords: principal component analysis (PCA); data-driven; fault prediction

主元分析：基于数据驱动的机电设备故障预测方法

林禄辉¹, 马洁²

北京信息科技大学自动化学院, 北京, 中国, 100192

E-mail: 82072847@163.com; mjbeijing@163.com

摘要: 本论文介绍的是基于数据驱动的方法, 其不需要过程的先验知识, 仅仅以采集的过程数据为基础, 而且还能够发掘数据中的隐藏信息; 主元分析方法(Principal Component Analysis, PCA)就是其中的一种分析方法。旨在提出基于 PCA 这种多元统计分析的故障预测的方法, 这种方法可以不需要对系统的结构和原理有深入的了解, 完全基于系统运行过程中传感器的测量数据, 而且可以减少噪声的干扰, 可简化数据的处理。

关键词: 主元分析; PCA; 数据驱动; 故障预测

1 引言

随着设备的现代化程度日益提高, 结构日益复杂, 在对设备状态进行分析中需要考虑更复杂的运行状态、工艺参数及工作环境, 较简便的预测技术已不足以指导设备的状态预测以及预知维护; 对大型变工况机电系统非平稳、时变的实际预测问题尚有待深入系统地研究; 混沌分形可预测问题、隐马尔可夫模型预测、基于支持向量机预测方法、基于经验模态分解预测方法、基于主元分析(Principal Component Analysis, PCA)的故障预测技术等以及相关的智能方法^[1]值得在大型复杂机电系统状态预测领域研究探讨。

故障预测是保障机械设备长期安全、满负荷运行

的关键技术, 是机械故障诊断学研究的重点之一。深入开展大型旋转机械设备故障预测关键技术及设备状态发展敏感特征提取技术, 预报设备故障和维护时间, 保障设备可靠运行; 为发展清洁能源和实现能源与经济、环境协调发展提供安全技术保障。本论文主要目的是提出基于 PCA 的这种故障预测的方法, 因为 PCA 方法能有效地进行特征提取, 同时还有降噪特性, 能简化数据的处理^[2]。

PCA 最早由 Pearson 在 1901 年提出^[3]。随后, 大量作者对 PCA 作了比较深入的研究, 使其理论逐渐完善, 其中主要包括 Fisher 和 MacKenzie^[4]、Hotelling^[5]等人的工作。

由于 PCA 方法具有不依赖于过程数学模型的特点, PCA 方法经过近几年的研究和发展, 已在各领域得到广泛的应用^[6-8]。

基金项目: 国家自然科学基金项目(50375017)“基于数据的机电设备多变换域非线性故障预测理论方法研究”项目

2 基于数据驱动方法

2.1 传统的预测方法

传统的预测方法应用于大型机电设备时,是选用以振动烈度值、频域幅值等振动量来构成特征参量,通过实时或定时的信号采集和数据处理,对历史以及现时数据进行拟合或联想推测,根据外推的发展趋势以及设定的判断阈值来预测设备或设备部件故障发生和维修的时间;但是这种方法一般适用于机电设备稳定工况或短历程非稳定工况。而且,大型机电设备发生的大部分故障都是可预知的趋势性故障,因此,为了对设备施行智能预知维护,需要对设备工作状态特别是对早期故障状态进行发展趋势预示,其中的前提是确定智能预知维护的敏感特征信息并能有效地进行特征信息的选择和提取;在确定和提取敏感特征信息中,可以将高维特征空间变换(或映射)为低维空间,进而提取设备状态发展的敏感特征信息。

因此,当设备负荷工况变化明显时,现有的特征量就不能有效地反映故障变化趋势,在面对变负荷的设备的状态特征提取时,可以探讨一种高维数据投影到低维子空间上,且能反映高维数据结构或特征的投影技术的方法。

2.2 基于数据驱动的方法

基于数据驱动的方法不需要过程的先验知识,仅仅以采集的过程数据为基础,通过各种数据分析处理方法发掘数据中的隐含信息,获取正常操作和与故障情况的不同数据特征模式,从而判断过程的运行状况^[9]。由于基于数据驱动的方法只依赖过程数据,因而通用型较强,比较适合机电设备对象的故障检测和诊断。同时,通过各种数据处理与分析方法(如多元统计的方法、聚类分析、频谱分析、小波分析等)挖掘出数据中隐含的信息,从而指导操作工业进行生产。根据对象是否存在非线性,基于数据驱动的建模方法可分为线性回归方法、人工神经网络方法、支持向量机方法、模糊建模方法等。

因此,在采用基于数据驱动的故障预报方法时,就是对所采集的过程运行数据进行分析处理,从而可以在不需知道系统精确解析模型的情况下完成系统的故障预报,完全从系统的历史数据出发,因此在实际系统中更容易直接应用。数据驱动技术的优势在于它们能将高维的数据变换成低维的数据,并从中获取重要的信息。

2.3 主元分析 (PCA) 的应用

数据驱动本身是计算机技术名词,即依靠大量数据的分析处理来支持决策,在系统实现中,强调以时间序列访问和操纵系统的内部或外部数据,对大规模历史数据建立数据仓库而实现各种功能。数据驱动最主要最广泛的表达方式是统计方法,几乎所有统计方法都可以用于“数据驱动”的描述。

统计理论和算法诞生得很早,在科学、经济、社会等方面也一直发挥着重要作用,但是直到信息时代,由于计算机和网络技术的快速进步,获取大量数据已非难事,统计方法才在工厂过程的应用中得到真正意义上地实现和发展,从上世纪八十年代后期开始到现在,不到二十年间,其方法、内容、范围都得到了极大的创新和拓展,正因为如此,数据驱动技术应用统一框架的概念也才能随之而生。最早将统计方法应用于工厂过程的是 Wise BM 和他的同伴^[10],这也是第一次将主元分析(Principal Component Analysis, PCA)和过程监控(Process monitoring)联系起来,从此主元分析成为过程监控的基本方法之一,多变量统计方法也开始在过程监控的应用上翻开了篇章,而在此之前多变量统计方法只用于分析数据及分类与模式识别领域。在 Wise 等人的工作之后,多变量统计过程故障监控和诊断研究蔓生出许多分枝并相互交错^[11],在利用主元分析的特点及突破局限性上,研究者在后续的研究中不断提出了新的或改进的方法。

数据是多变量统计技术的基础,但通常系统分析不能就每一个数据进行分析,否则计算量和复杂度会呈爆炸式增长,因此多变量统计方法在工厂过程应用时的基本思路和形式是依据原始数据抽取表征过程的特征信息,使过程得到降维或其他简约处理,最普遍的方法就是构造较小的隐变量空间,以隐变量空间作为原始变量空间的代表,然后利用隐变量空间来实现各种功能,构造隐变量空间的目的不同,算法也可以不同。虽然统计技术在应用时的条件性限制较强,但其应用空间涉及了很多方面。1995年, D.R. Lewin 就将 PCA 成功地应用于离心泵的故障预测中^[12],这说明 PCA 在故障预测中的应用,已经很早便开始了,只是发展的速度比较慢,还未普及起来。

应用基于 PCA 等多元统计分析的故障预测的方法,可以不需要对系统的结构和原理有深入的了解,完全基于系统运行过程中传感器的测量数据,而且 PCA 算法简单,易于实现。但是,并且由于实际系统的复杂性,这类方法中还有许多问题有待进一步的研究。

究, 比如过程变量之间非线性, 以及过程的动态性和时变性等。

另外, PCA 可用于含有噪声的和高度相关的测量数据进行分析。它试图从原系统变量中提取具有最佳解释能力的新综合变量(成分), 从而克服变量间多重相关性造成的信息(数据集中的变异, 方差)重叠, 并有效区分系统信息与噪声, 降低异常点和错误样本对建模的影响。它是一种将高维信息投影到低维子空间, 并保留主要过程信息的方法。在对原始数据进行主元分析后, 可以将原始的多维数据矩阵转化为几个能提供直观信息的主元图。

2 主元分析

现代工业过程中的大量实时数据可以被采集及存储。但是, 由于一些变量之间通常存在着复杂的相关性, 往往使得操作人员难以通过这些数据判断真正的过程变化。这种情况一般称为“数据过载, 信息缺失”^[13]。PCA 可用于含有噪声的和高度相关的测量数据进行分析。它试图从原系统变量中提取具有最佳解释能力的新综合变量(成分), 从而克服变量间多重相关性造成的信息(数据集中的变异, 方差)重叠, 并有效区分系统信息与噪声, 降低异常点和错误样本对建模的影响。

2.1 主元分析的理论基础

主成分分析(PCA)就是把多个特征映射到少数几个综合特征的一种统计分析方法。在多特征研究中, 往往由于特征过多, 且彼此之间存在着一定的相关性, 因而使获取的信息在一定程度上存在信息重叠, 导致在高维空间中研究模式识别问题更加复杂。主成分分析采取一种降维的方法, 找出一个综合因子来替代原来众多的特征, 使这些综合因子尽可能多地反映原来多个特征的信息, 而且彼此之间互不相关, 从而达到简化的目的。其算法的原理是: 借助于一个正交变换 T , 将分量相关的原随机向量 $x = (x_1, x_2, \dots, x_p)^T$ 转化成分量不相关的新随机向量 $z = (z_1, z_2, \dots, z_p)^T$, 这在代数上表现为将 x 的协方差矩阵变换成对角矩阵, 在几何上表现为将原坐标系变换成新的正交坐标系, 使之指向样本点散布最开的 p 个正交方向, 然后对多维变量系统进行降维处理, 使之能以一个较高的精度转换成低维变量系统。

从数学的角度而言, 主元分析通过对测量数据矩阵的协方差矩阵进行奇异值分解(Singular value

decomposition, SVD)而得到。其原理如下, 采集处于正常操作条件下的过程数据, 对其进行标准化处理, 得到均值为 0, 方差为 1 的数据矩阵 $X_{n \times m}$ (n 为采样点数, m 为测量变量数)。然后进行主元分析, 设主元模型为:

$$X = TP^T + \tilde{T}\tilde{P}^T = TP^T + E \quad (1)$$

式中, $T \in R^{n \times k}$ 和 $P \in R^{m \times k}$ 分别为主元得分与载荷矩阵, $k \leq m$ 为选取的主元数; $\tilde{T} \in R^{n \times (m-k)}$ 和 $\tilde{P} \in R^{m \times (m-k)}$ 分别为残差的得分和载荷矩阵, $E \in R^{n \times m}$ 为预测残差矩阵。主元和载荷矩阵可以通过对测量数据矩阵 X 的协方差矩阵 $R = X^T X / n - 1$ 进行奇异值分解得到, 即:

$$R = [P, \tilde{P}] \Lambda [P, \tilde{P}]^T \quad (2)$$

$$T = XP \quad (3)$$

$$\tilde{T} = X\tilde{P} \quad (4)$$

式中, $\Lambda = \text{diag}(\lambda_i)$, 其中, λ_i ($i=1 \dots m$) 为主元得分 T 和残差子空间 \tilde{S} 的和。即 $S = \hat{S} \oplus \tilde{S}$ 式中 \oplus 表示两空间的直和, 子空间 \hat{S} 和 \tilde{S} 互为交补空间, 且 $\dim(\hat{S}) = k$, $\dim(\tilde{S}) = m - k$ 。在对过程性能进行监控时, 一般是利用 Hotelling T^2 和平方预测误差 SPE(Square Predicted Error)两个统计量。其中, Hotelling T^2 统计量在主元子空间 \hat{S} 定义, 即:

$$T^2 = \sum_{i=1}^k \frac{t_i^2}{\lambda_i} \quad (5)$$

其置信限可由 F 分布确定。而 SPE 统计量位于残差子空间, 其定义为:

$$SPE = \|\tilde{C}x\|^2 = \|\tilde{P}\tilde{P}^T x\|^2 \leq Q_\alpha \quad (6)$$

式中, Q_α 为 SPE 统计量的控制限。PCA 通过检测 T^2 和 SPE 是否超限, 确定过程是否处于正常工况。Dunia 和 Qin(1998)对 PCA 在故障诊断方面的应用作了详细系统的理论分析^[14]。他们采用子空间的方法分析了残差子空间的故障可检测性、重构性、区分性, 并给出了相应的充要条件。国内学者王海清(2002)等在此基础上, 按照类似的方法, 从主元子空间和残差子空间两个角度全面分析了 PCA 在检测过程中出现的各种情况^[15,16], 即: (1) T^2 和 SPE 统计量超过控制界限; (2) T^2 统计量超过控制界限, SPE 统计量没有超过; (3) T^2 统计量没有超过控制界限, SPE 统计量超过; (4) T^2 和 SPE 统计量均未超过控制界限, 给出了故障

可检测性的充要条件。

在实际工作中, 由于 n 个特征变量的相关性, 往往造成了分析数据的困难。主元分析的目的在于:

- a) 选择少数无关的新变量来概括原来的 n 个特征。
- b) 通过对观测数据和几个主分量的数据的整理和分析, 提取出有用的信息。
- c) 利用这些信息进行决策。

2.2 主元个数的选取

PCA 中保留主元个数的选取对得到的主元模型的质量有重要的影响。根据特定问题的不同侧重点, 可以采用不同的指标: 如果关心对数据矩阵的解释程度, 可以采用主元的累积贡献率或平均值特征值来衡量; 如果关心模型的预测性能, 可以采用交叉有效性校验(Cross-Validation, CV); 根据特定的统计假设, 可以采用 Akaike 信息准则(Akaike Information Criterion, AIC)来确定需要保留的主元数。特别地, 可以通过最小化传感器重构误差(Variance of Reconstruction Error, VRE)来选取。

本论文主要是关心模型的预测性能, 因此通常采用交叉检验法或累积方差贡献率(CPV) 来确定主元个数。交叉检验法是将数据分为两部分, 一部分用于建立主元模型, 一部分用来测试各主元模型, 然后从中选取检验数据测试误差最小的那个主元模型; 累积方差贡献率是选择百分比大于 85% 的主元个数。在基于主元分析的状态监测和故障诊断中, 一般用来描述正常运行过程的主元个数不超过 3 个^[17]。

2.3 主元分析的基本过程

PCA方法首先将原始数据变换到一个新的正交空间。在这个新数据空间里, 第一坐标轴方向上的方差最大, 即变换后得到的第一个新变量变异最大; 第二主元与第一主元正交且方差第二大, 第三主元与第一、二主元都正交且方差第三大, 其余变量依此类推。虽然新空间的变量数与原始数据的变量数一样, 但实践发现在大多数情况下只需要前面少数几个新变量就占了方差总和的绝大部分; 所以可以将其余的变量忽略, 而将保留下来的变量作为进一步分析对象。

简单的说就是PCA把 p 个相关的观测变量 y_1, y_2, \dots, y_p 变成 p 个不相关的主元变量 v_1, v_2, \dots, v_p , 其中, 主元 v_i 是将观测数据 y 在载荷向量 b_i 方向上的投影, 即 $v_i = b_i^T y$, b_i 是观测数据协方差矩阵的特征向量。

然后, 构造主元统计量 T^2 和 Q , 通过检测 T^2 和 Q 的样本取值是否超过其相应的控制限, 确定系统是否处于正常工况。其中, T^2 表示了系统为稳态时数据点到模型中心的距离 (建模时经过归一化后中心在原点), 因此 T^2 的大小反映了新数据偏离正常状态的程度, 是主元模型内部变化的一种测度。建模时根据正常状况下的数据求出 T^2 统计量的置信区间, 然后通过判断新数据的 T^2 统计量是否在该区间内就可以粗略的判断系统是否运行在正常状态。 Q 统计量与 T^2 统计量的作用在一定程度上恰好互补, Q 统计量主要监测输入变量的数据结构是否变化, PCA模型是否仍适用, 是衡量模型外部数据变化的测度。当 Q 统计量发生大的变化时, 说明PCA统计模型代表的正常工况下的变量之间的关系被破坏, 有过程故障 (或传感器故障) 发生。

3 结 语

主元分析(PCA)等统计方法是基于数据建立起系统的统计模型, 根据假设检验等多变量统计分析技术进行异常检测与故障诊断。在复杂过程的故障诊断与预测维护中有着潜在的应用前景。基于 PCA 的状态监测和故障诊断方法是利用过程变量间的相关关系, 在低维空间建立正常工况下的主元模型, 通过检验新的数据样本相对于主元模型的背离程度, 从而发现异常和故障。

主元分析法不依赖于精确数学模型, 通过对高维相关变量空间进行降维映射处理, 将其转化为相互独立的低维变量空间, 实现对复杂过程数据的特征抽取, 并建立相应过程的主元模型。主元模型舍弃了部分残差而保留体现数据变异的主要方向, 从而达到抽取系统信息、清除系统干扰的目的。因此, 比较适合像机电设备这种复杂的系统。

近几年发展起来的时间序列分析方法, 对预报和诊断某些故障效果较好, 但建模需要时间长, 做到实时处理很困难。而 PCA 方法具有减少噪声干扰, 简化数据处理, 以及处理速度快等特性, 而且还能降低异常点和错误点的影响。而且, 是基于数据建立起系统的统计模型, 不需要精确的数学模型, 算法简单, 易于实现。

目前, 主元分析(PCA)广泛的应用与模式识别、信号和图像处理、故障诊断、控制领域和其他领域中, 并且都取得了较好的成果。在 2009 年国际自动控制联合会(IFAC)的 SAFEPROCESS 会议上就提及到了基于

PCA 的故障预测技术, 这说明国际上已经开始很重视将 PCA 用于故障预测方面的研究, 这又为基于 PCA 的故障预测技术研究与应用掀开了另一篇章。期望在该方面扩展研究并缩短理论与实际应用间的距离。

致谢

感谢国家自然科学基金给予的资助, 本国家自然科学基金项目编号为: 50375017。

References (参考文献)

- [1] Zhao zhenghua, Study on Fault Prediction[J], *Journal of Reform and Explore*, 2008(10), P22-24(Ch).
赵正华, 故障预测技术研究[J], 改革探索, 2008(10), P22-24.
- [2] Xiong Li, Data-driven Based Process Monitoring and Optimization[D], *Zhejiang: Zhejiang University*, 2008(Ch).
熊丽, 基于数据驱动技术的过程监控与优化方法研究[D], 浙江: 浙江大学, 2008.
- [3] Pearson K. On lines and Planes of closer fit to systems of points in space[J]. *Philos. Mag.* 1901, 2: 559-572.
- [4] Fisher R, MacKenzie W. Studies in crop variations. II: The manorial response of different potato varieties [J]. *J. Agr.Sci.*, 1923, 13: 311- 329.
- [5] Hotelling H. Analysis of a complex of statistical variables into principal components [J]. *J.Edu. Psych.*, 1933, 24: 417-441.
- [6] Xia C M, John Howell. Isolating multiple sources of plant-wide oscillations via independent component analysis[C]// *IFAC SafeProcess*. 2003: 873-878.
- [7] Wang Lifang, Zhu qunxiong, Research and application of PCA in process fault diagnosis based on wavelet theory [J], *Journal of Control and Instruments in Chemical Industry*, 2004, 31 (6):25-27(Ch).
王丽舫, 朱群雄, 基于小波理论的主元分析在故障诊断中的研究与应用[J], 化工自动化及仪表, 2004, 31(6), 25-27.
- [8] Zhang Jie, Yang Xianhui, Multivariate statistical process control [M], *Beijing: Chemical Industry Press*, 2000 (Ch).
张杰, 阳宪惠, 多变量统计过程控制[M], 北京: 化学工业出版社, 2000.
- [9] Guo Ming, Researches on Performance Monitoring and Fault Diagnosis for Process Industry Based on Data-driven Technique[D], *Zhejiang: Zhejiang University*, 2004(Ch).
郭明, 基于数据驱动的流程工业性能监控与故障诊断研究 [D], 浙江: 浙江大学, 2004.
- [10] Wise, B.M., D.J. Veltkamp, et al.. Principal component analysis for monitoring the West Valley liquid fed ceramic melter [J]. *Waste Management* '88 Proc:811-818,1988.
- [11] Qin, S.J. Statistical process monitoring: basics and beyond [J]. *Journal of Chemometrics*, 17:480-502, 2003.
- [12] D.R. Lewin. Predictive Maintenance Using PCA [J]. *Journal of Control Eng. Practice*, 3(3):415-421, 1995.
- [13] B.M. Wise and N.B. Gallagher. The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*, 6(6):329- 348, 1996.
- [14] Dunia, R and Qin, S.J. Subspace approach to multidimensional fault identification and reconstruction [J]. *AIChE.J.*, 44(8), 1813-1831,1998.
- [15] Wang Haiqing, Song Zhihuan, Li Ping, Study on the Fault Detectability of Principal Component Analysis[J], *Journal of Instrument*, 2002, 23(3): 232- 235(Ch).
王海清, 宋执环, 李平, 主元分析方法的故障可检测性研究 [J], 仪器仪表学报, 2002, 23(3), 232-235.
- [16] Wang Haiqing, Song Zhihuan, Wang Hui, Fault Detection Behavior Analysis of PCA-Based Process monitoring Approach [J], *Journal of Chemical Industry and Engineering*, 2002,53 (3):297-301(Ch).
王海清, 宋执环, 王慧, PCA 过程监测方法的故障检测行为分析[J], 化工学报, 2002, 53(3), 297-301.
- [17] Zhu Songqing, Shi Jinfei, PCA Approach to Condition Monitoring and Fault Diagnosis [J], *Journal of Machine Tool & Hydraulics*, 2007, 35(1):241-243 (Ch).
朱松青, 史金飞, 状态监测与故障诊断中的主元分析法[J], 机床与液压, 2007, 35(1), 241-243.