

面向样本不平衡的 故障特征提取方法

The Fault Feature Extraction
Method for Sample Imbalance

王 健 韩志艳 著



東北大學出版社
Northeastern University Press

面向样本不平衡的 故障特征提取方法

王 健 韩志艳 著

东北大学出版社

·沈 阳·

© 王健 韩志艳 2016

图书在版编目 (CIP) 数据

面向样本不平衡的故障特征提取方法 / 王健, 韩志艳著. —沈阳: 东北大学出版社, 2016.2

ISBN 978-7-5517-1215-6

I. ①面… II. ①王… ②韩… III. ①语声信号处理—研究 IV. ①TN912.3

中国版本图书馆 CIP 数据核字 (2016) 第 026579 号

内容简介

本书系统地研究了面向样本不平衡的故障特征提取方法。全书共分为 7 章。第 1 章介绍了国内外研究现状、意义和需要解决的难题; 第 2~6 章介绍了半监督主元分析的故障特征提取及诊断算法、局部保持主元分析的故障特征提取及诊断算法、流形学习的联合 FDA 故障特征提取及诊断算法、集成学习的敏感 MFA 故障特征提取及诊断算法、不平衡 SVDD 的故障特征选择及诊断算法等的研究成果; 第 7 章归纳全书内容并对今后工作提出展望。

本书的主要特点是在面向样本不平衡的故障诊断方面提出了开创性的设计和分析方法, 书中的内容来源于作者近年来的创新性研究成果, 新颖实用, 研究方法先进, 尤其注重故障诊断算法的鲁棒性和实用性, 而这些都是当前所属研究领域的前沿问题, 因而具有重要的理论价值与应用价值。

出 版 者: 东北大学出版社

地址: 沈阳市和平区文化路 3 号巷 11 号

邮编: 110819

电话: 024-83680267 (社务部) 83687331 (市场部)

传真: 024-83687332 (总编室) 83680178 (出版部)

网址: <http://www.neupress.com>

E-mail: neuph@neupress.com

印 刷 者: 沈阳航空发动机研究所印刷厂

发 行 者: 东北大学出版社

幅面尺寸: 170mm×240mm

印 张: 8.25

字 数: 193 千字

出版时间: 2016 年 2 月第 1 版

印刷时间: 2016 年 2 月第 1 次印刷

组稿编辑: 孙 锋

责任编辑: 刘乃义

责任校对: 叶 子

封面设计: 刘江旻

责任出版: 唐敏志

ISBN 978-7-5517-1215-6

定 价: 35.00 元

前言

Preface

随着工业系统日益复杂化,基于数据驱动的故障诊断方法正成为当前研究的一个热点方向。区别于传统方法对精确数学模型的依赖,数据驱动式故障诊断通过分析和挖掘过程数据,采用一个反映系统输入和系统状态关系的“黑盒子”对目标系统进行描述,更适用于结构复杂的现代工业系统。

在基于数据驱动的故障诊断方法从理论走向实践应用的过程中,不断遇到新的挑战,诊断样本的不平衡特性就是其中之一。第一,诊断样本是数据驱动故障诊断方法的基础,但是,这些工况数据中绝大部分都是无标签的数据;第二,这些工况数据中相对冗余的传感信息又将造成一定程度的“维数灾难”;第三,这些数据中各种工况的样本数量是不平衡的,它们中大部分是正常工况数据,少部分是故障工况数据。以上特点使得用于诊断的工况数据呈现严重的不平衡的特性(样本标记不平衡,样本价值不平衡,样本类别不平衡)。

针对上述三种样本的不平衡特性,本书探讨了相应的故障特征提取技术,主要研究内容和创新点有以下几方面。

首先,在训练样本标记不平衡的情况下,探讨了如何引入半监督学习方法对传统的主元分析进行改造,提出了一种半监督主元分析故障诊断算法。该算法在计算大量无标注样本方差的基础上,通过加强正常工况样本的影响而减小非正常工况样本的影响,从而提高了主元分析算法的鲁棒性和适应性。

其次,分析了样本价值不平衡,提出了两种故障特征提取算法:局部保持主元分析和联合费舍尔判别分析。

局部保持主元分析从全面提取正常工况样本特征的角度出发,通过引入流形学习算法,将局部保持映射的思想融入到主元分析的目标函数中,使得在原始变量空间投影到低维空间的过程中,不仅实现了整体方差最大化,而且保持了局部近邻结构不变,有利于全局和局部特征的全面提取。



联合费舍尔判别分析从分类的角度出发,设计新的样本类内散度和类间散度的构建方法,克服非高斯性和非线性数据结构对数据几何中心点计算带来的干扰,同时为了克服离群点对几何中心点的影响,算法从能量密度的角度,提出了离群点的处理方法。为了进一步缓解生产过程中的非线性因素对数据全局和局部结构的影响,本书对联合费舍尔判别分析进行了核扩展,最大程度地减少全局和局部数据结构的扭曲,克服非线性数据结构对数据几何中心点计算带来的影响。

最后,针对样本类别不平衡,提出了基于集成学习的敏感边际费舍尔分析和基于不平衡支持向量数据描述的回归特征消去方法。

基于集成学习的敏感边际费舍尔分析结合全局和局部重采样技术,由三个关键部分组成:(1)在全局层面,提出了基于装袋算法的集成模型用来克服数据漂移产生的过拟合;(2)在局部层面,提出了一种基于流形学习的过采样技术——加权合成少数类过采样算法,以解决少数类的小样本问题;(3)提出了敏感边际费舍尔分析,用来解决类重叠问题造成的干扰。

基于不平衡支持向量数据描述的回归特征消去方法是一种特征选择算法,它结合了监督和无监督特征选择方法的优点,并针对监控数据的类别不平衡特性进行了优化,能够在少量故障样本的指导下,更好地描述正常样本的边界分布情况。

本书的内容为作者近年来的研究成果,内容新颖,属于当前所属研究领域的前沿问题,具有重要的理论价值与应用价值。

在本书写作过程中,在导师冯健教授的指导下,我掌握了创造性地解决科研问题的方法,为我今后的学术发展打下了坚实的基础。另外,渤海大学工学院尹作友教授、邵治新博士等提出了许多有价值的建议,博士研究生王申全、李宁、李芳明、韩克镇、卢森骧、韩晓光、黄蕊在书稿整理、仿真实验等方面做了大量的工作,在此向他们表示由衷的感谢。

本书的出版得到了国家自然科学基金(项目编号:61403042,61503038)和辽宁省教育厅一般项目(项目编号:L2013423)的资助,在此一并表示感谢。

由于作者水平有限,不妥之处在所难免,热诚欢迎读者与同行不吝赐教。

王 健

2016年1月于渤海大学

目 录

第 1 章	绪 论	1
1.1	故障诊断的研究背景和意义	1
1.2	故障诊断的研究内容和分类	2
1.3	故障诊断中的样本不平衡特性分析	7
1.4	主要研究内容	18
	本章参考文献	22
第 2 章	基于半监督主元分析的故障特征提取及诊断算法	27
2.1	半监督学习理论	27
2.2	面向样本标记不平衡的半监督主元分析	31
2.3	基于半监督主元分析的故障诊断方法	34
2.4	实验与分析	35
2.5	本章小结	43
	本章参考文献	44
第 3 章	基于局部保持主元分析的故障特征提取及诊断算法	45
3.1	面向样本价值不平衡的局部保持主元分析算法	45
3.2	LPPCA 算法在故障检测中的应用	49
3.3	实验与分析	50
3.4	本章小结	56
	本章参考文献	57

**第4章 基于流形学习的联合FDA故障特征提取及诊断算法** ...58

- 4.1 FDA算法的限制和JFDA算法的提出背景58
- 4.2 面向样本价值不平衡的JFDA算法和它的核扩展62
- 4.3 基于(K) JFDA的故障诊断算法70
- 4.4 实验与分析.....71
- 4.5 本章小结.....80
- 本章参考文献81

第5章 基于集成学习的敏感MFA故障特征提取及诊断算法 ...83

- 5.1 数据类别不平衡对半导体故障特征提取造成的困扰.....83
- 5.2 基于集成学习的敏感边际费舍尔分析.....88
- 5.3 ESMFA算法在故障检测中的应用95
- 5.4 实验和分析.....96
- 5.5 本章小结102
- 本章参考文献103

第6章 基于不平衡SVDD的故障特征选择及诊断算法 107

- 6.1 特征选择研究概述107
- 6.2 无监督和监督特征选择方法在故障诊断中的局限性分析110
- 6.3 面向样本类别不平衡的ISVDD-radius-RFE算法113
- 6.4 实验和分析119
- 6.5 本章小结123
- 本章参考文献124

第7章 总结与展望125



第1章 绪 论

1.1 故障诊断的研究背景和意义

近年来,随着制造技术、电子技术、计算机技术和网络技术的迅速发展,以及市场竞争的日益激烈,各个企业为了求生存、谋发展,不断进行技术引进和技术改革,促使工业生产不断向大型化、连续化、高速化、重载化和智能化等方向发展。这些发展主要表现出三个特点:

① 设备结构的复杂化,各组成部件之间关系复杂,导致影响设备正常运行的因素增加;

② 设备规模的大型化,高度自动化的集成生产系统越来越多,导致系统造价不断升高;

③ 设备速度的极限化,导致安全隐患越来越多。

例如一台600MW的火力发电机组,其I/O点超过15000个,数据库标签超过30000个,运行过程的监控点在6000~7000个之间,控制系统组态图在200~3000张左右,各种控制回路大约为500~600个,其整套设备的投资一般都在数十亿元人民币以上。这一方面提高了系统的自动化水平,满足了工业生产的客观要求,同时降低了生产成本,提高了生产效率;另一方面,影响系统运行的因素骤增,使其产生故障或失效的潜在可能性越来越大,对整个系统的设计、制造、安装、运行、维护等提出了更高的要求。一个部件的故障常常会引起链式反应,若某些微小的故障不能及时排除,那么往往会因为“马蹄钉效应”,导致整个系统甚至整个生产过程不能正常运行乃至瘫痪,轻则降低性能、影响生产,重则造成停机、停产、毁坏设备,甚至机毁人亡的灾难性后果^[1]。

多年来,国内外曾经发生的各种空难、海难、爆炸、断裂、倒塌、泄漏等恶性事故,不仅造成了巨大的人员伤亡和巨大的经济损失,而且还造成了严重的不良社会影响。

国外发生的一些重大事故:1979年,美国三里岛核电站发生泄漏,损失几十亿美元;1984年12月,位于印度博帕尔市的美国碳化物公司农药厂发生农业毒剂泄漏事件,造成2700余人死亡,20多万人中毒,成为世界工业史上



最大的恶性事故；1986年1月28日，美国挑战者号航天飞机由于固体燃料助推火箭密封圈失效，造成泄漏而引起燃料箱爆炸，7名宇航员全部遇难，1986年4月27日，苏联切尔诺贝利核电站因四号发动机剧烈振动，致使反应堆厂房的结构遭到破坏，引起反应堆放射性物质泄漏，32人死亡，13.5万人受害，损失达数十亿美元，并且还污染了周边的一些国家；2000年，法国“协和”客机起飞中机翼起火爆炸，机上109人全部遇难，地面死亡5人；2011年9月25日，智利电力系统发生故障，包括首都圣地亚哥在内的大多数地区漆黑一片，全国1600万人口中有近千万人受到影响，甚至引起骚乱事件；等等。

国内发生的重大事故：1988年，秦岭电厂200MV的汽轮机组转子飞车，损失高达1亿元；1999年，西南航空公司图154M飞机发生撞地事故，11名机组人员和50名乘客全部遇难；2000年，贵州木冲沟煤矿瓦斯爆炸，161名井下矿工遇难；2005年11月13日，吉林石化发生一起特大爆炸事故，爆炸导致8人死亡，60余人受伤，直接经济损失6908万元，同时松花江水体受到污染，引发哈尔滨市停水4天，俄罗斯部分地区用水受到影响，间接损失不可估量；2013年6月2日，中国石油天然气股份有限公司大连石化分公司第一联合车间发生爆炸、泄漏物料着火，造成4人死亡，直接经济损失达697万元；等等。

国内外这些灾难性的事件无时无刻不在提醒人们，为了杜绝灾难性事故，避免严重性事故发生，减少一般性事故，防止和消除故障的发生，现代工程技术系统运行的安全性和可靠性已成为人类必须解决的刻不容缓的问题。过程监控和故障诊断技术可以有效预防安全事故和提高企业的经济效益。英国政府对2000家工厂进行调查的结果表明，设备管理采用故障诊断技术后，每年可节约维修费用3亿英镑，而用于诊断技术的投资费用为0.5亿英镑，直接收益为2.5亿英镑，诊断技术投资与获利之比为1:5，某些工厂投资和获利之比高达1:17^[2]。

1.2 故障诊断的研究内容和分类

1.2.1 故障诊断的研究内容

近20年来，随着科学技术的不断进步与发展，尤其是计算机技术的迅猛发展和普及，故障诊断技术已逐步发展成为一门较为完善的新兴边缘综合性工程学科。故障诊断的本质是状态的识别，即在一定工作环境下，根据系统运行过程中产生的各种信息判断系统是正常运行还是发生了异常现象，对已经发生



或者可能发生的故障进行诊断和分析预报,以确定故障发生的原因和部位,从而提出维修对策,使设备恢复到正常状态的技术^[3]。

从信息提取的角度看,故障诊断的过程就是在不同层次上做不同程度的特征提取以实现从观测空间到特征空间、决策空间和分类空间的变换。这个过程称为故障诊断系统的空间变换,如图1.1所示^[4]。



图1.1 故障诊断系统的空间变换

Fig. 1.1 Spatial transform of fault diagnosis system

从这个意义上说,特征提取是故障诊断研究的“瓶颈”问题,它直接关系到故障诊断的准确性和早期预报的可靠性,受到了工程和学术领域众多专家的广泛关注,具有重要的理论意义和很高的工程实用价值。

1.2.2 故障诊断的分类

故障诊断方法的研究是故障诊断技术的核心,最初于20世纪70年代在美国发展起来。最早的研究成果有:麻省理工学院的Beard在其博士论文中首先提出了用解析冗余代替硬件冗余检测故障信息^[5],Mehra和Peschon在*Automatica*上发表了关于故障检测与诊断的研究论文^[6]。至Patton和Frank等于1989年出版了第一部关于解析冗余的故障诊断技术的专著后^[7],控制系统的故障检测与诊断理论与技术取得了丰富的成果^[8]。一般可以将故障诊断方法划分为三类:基于解析模型的方法、基于知识的方法以及基于数据驱动的方法。

1.2.2.1 基于解析模型的方法

基于解析模型的故障诊断方法一直是人们关注的重点,也是最早发展起来的故障诊断方法。如图1.2所示,文献^[9]总结了基于模型的故障检测技术的发展历程。

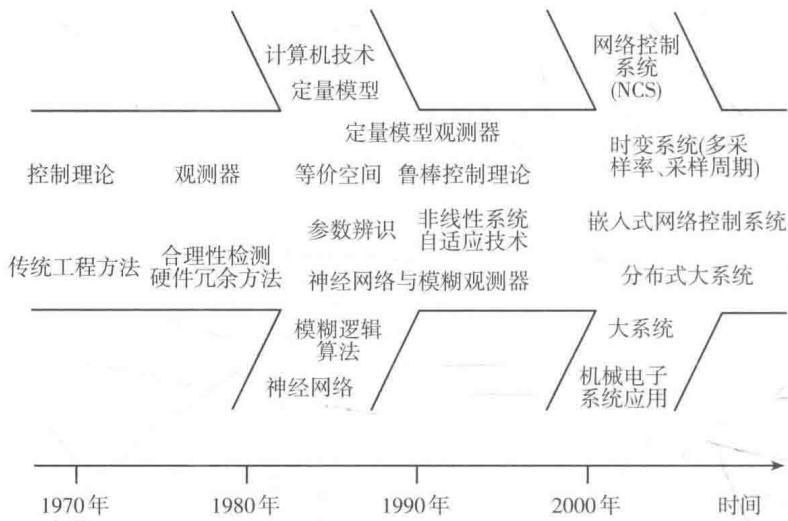


图 1.2 基于模型的故障诊断理论发展

Fig. 1.2 Sketch of the historic development of model-based fault diagnosis theory

事实上，故障检测技术的发展与现代控制理论的发展有着很大的关联，比如 20 世纪 70 年代现代控制领域的观测器理论为基于观测器的故障检测方法建立了基础，再加上近 20 年计算机技术的突飞猛进，这两个因素使得基于模型的故障检测方法得到了快速发展和广泛应用。

基于解析模型的故障诊断方法需要建立系统的精确的数学模型，通过将被诊断对象的可测信息和由模型表达的系统先验信息进行比较，从而产生残差，并对残差进行分析、处理，从而实现故障诊断^[10]，其基本原理如图 1.3 所示。其中的残差，就是与被诊断系统的正常运行状态无关的以及由其输入、输出信息构成的线性或非线性函数。当系统没有故障时，残差为零或近似为零；当系统中出现故障时，残差显著偏离零点。

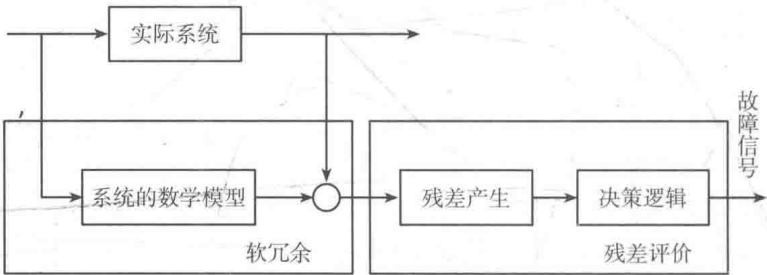


图 1.3 基于模型的故障诊断的原理

Fig. 1.3 Basic model-based fault diagnosis scheme

根据残差产生的形式不同，基于解析模型的方法可以进一步分为基于观测器的故障检测方法、等价空间方法^[11]和参数估计方法^[12]。尽管这三种方法是



各自独立发展起来的,但它们之间却存在着一定的相互联系。现已证明了等价空间方法和基于观测器的故障检测方法是等价的,虽然它们的设计过程有很大区别。而参数估计方法的故障检测则是利用在线的参数估计来判断故障是否发生,通过对系统模型参数的辨识来达到故障检测的目的,即由参数的显著变化来描述故障。很多学者将参数估计方法与基于观测器的故障检测方法进行比较和讨论,这两种方法各有优点和不足,适用于不同的系统环境。相比之下,参数估计方法比基于观测器的故障检测方法更适合非线性系统,因为非线性系统的状态观测器的设计有很大的困难。非线性系统的状态观测器很难设计,目前只对某些特殊的非线性系统有研究,而通常的等价空间方法仅适用于线性系统。基于解析模型的故障诊断方法的主要成果仍集中在线性系统。

1.2.2.2 基于知识的方法

基于知识的方法不需要精确的数学模型,它通过系统的因果模型、专家知识、系统的详细描述或者故障症状举例来获得监控模型,特别适合于现代复杂工业过程中详细数学模型难以获取的系统,因而具有很大的应用前景。基于知识的方法常见的有专家系统法、人工神经网络方法、模糊推理法、故障树等。

专家系统法^[13]被认为是当基于解析模型故障诊断方法无法实现时的替代方法,其中,Stanford大学的Dendral专家系统、EGG公司的Reactor专家系统等都得到了成功的商业应用。国内一些高校和企业也在这方面取得了一些成绩。但是建立这种基于知识的故障诊断系统却不是一件简单的事情,专家系统中的知识获取过程非常困难,智能化水平也不高,种种缺点限制了专家系统这类基于知识的故障检测技术的发展。

人工神经网络理论的应用已经渗透到各个领域,并在智能控制、计算机视觉、传感技术与机器人、生物医学工程等方面取得了令人鼓舞的进展。从仿生学观点来看,由于人工神经网络具有许多优异的性能,它的可塑性、自适应性和自组织性使其具有很强的学习能力;它的并行处理机制使其求解问题时间很短,能满足实时性要求;它的分布式存储使其可靠性和容错性相当好,所以它易于与实时故障诊断结合,为故障信号分析与处理、故障模式识别、故障领域专家知识的组织和推理等方面的研究提供了一种新的途径,并推动了过程监控的智能化。

模糊推理法是从20世纪60年代中期开始发展起来的,用来表示不确定性和不精确性知识。模糊推理利用分级状态,而不指出严格正确或错误,提供了近似但有效地描述复杂不确定系统的方法。模糊故障诊断有两种基本方法:一种是先建立征兆与故障类型之间的因果关系矩阵,再通过相应的模糊合成算子



建立故障与征兆的模糊关系方程,这是一种基于模糊关系及合成算子的诊断方法;另一种是先建立故障与征兆的模糊规则库,再进行模糊逻辑推理的诊断过程,这是一种基于模糊知识技术的诊断方法。

故障树源于20世纪60年代,以其严整的逻辑结构、形象直观的树状图形、揭示故障根源和定量计算失效概率的强大功能,深受可靠性工程界的欢迎,成为系统可靠性和安全性分析的重要工具。基于故障树的故障诊断始于Caceres等^[14]所提出的由系统结构框图建立故障树的方法。故障树的实质就是一种特殊的描述系统因果关系的图形化模型,通过这个图形化模型,可实现由果到因的分析过程,进而从已经发生的故障倒推到故障发生的原因及部位,这是一种完全定性的故障诊断方法。具体来说,基于故障树的故障诊断技术就是从已经检测到的系统故障状态出发,由上至下地逐级推理分析,最终找出故障发生的根本原因、故障发生的概率和发生部位以及故障对于整个系统所造成的破坏的影响程度^[15]。

1.2.2.3 基于数据驱动的方法

随着集散控制系统及智能仪表在工业过程中的广泛应用,大量过程数据被记录下来。另外,计算机、数据库技术的发展为分析大规模数据提供了可能性。但是这些包含过程运行状态信息的数据并没有被有效地利用,以至出现了所谓“数据丰富,信息缺乏”的现象。工业界已越来越意识到将现有数据变为有用的信息,使之服务于系统监控,以增加系统的可靠性并降低维护成本的重要性,因此出现了基于数据驱动的故障诊断方法。

基于数据驱动的方法使用多变量统计、机器学习、模式识别等数据处理方法,利用在正常过程和故障状态下记录的具有代表性的过程历史数据建立正常和故障模式的参考模型。这类方法只依赖于测量数据,特别适用于难以获得准确模型和完整知识描述的大规模复杂过程的监测,在工业应用中最为流行。本书的研究即属于基于数据驱动的过程监测方法。

目前,多元统计过程监控(Multivariate Statistical Process Monitoring, MSPM)方法在数据驱动故障诊断领域得到了广泛的关注。该方法将实测变量所张成的高维空间数据投影到低维特征空间中,实现以较少维数的数据描述监控对象的主要特征。其实质就是将一个高维空间进行降维处理,通过多元投影方法构造一个较小的隐变量空间,以隐变量空间代替原始变量空间,然后在隐变量空间来实现算法的各种功能。这种方法可以较好地解决工业过程数据中存在的高维、高度耦合、共线性、数据缺损以及噪声污染等问题,因此,多元统计过程监控方法近年来得到了长足的发展。



1.3 故障诊断中的样本不平衡特性分析

在基于数据驱动的故障诊断方法从理论走向实践应用的过程中,不断遇到一个又一个新的挑战,诊断样本的不平衡特性就是其中之一。诊断样本是数据驱动故障诊断方法的基础,通常由工业现场使用一定数量的传感器获取。第一,它们主要由恶劣环境下工艺运行过程的部分的工况数据组成,但是,这些工况数据中绝大部分都是无标签的数据,即人们并不知道这些数据所对应的具体工况,人们只有从经过专家标记的数据中才能获取蕴含在其中的诊断信息。第二,这些工况数据中相对冗余的传感信息又将造成一定程度的“维数灾难”。第三,这些数据中各种工况的样本数量是不平衡的,它们中大部分是正常工况数据,少部分是故障工况数据。以上特点使得用于诊断的工况数据呈现严重的不平衡的特性(样本标记不平衡,样本价值不平衡,样本类别不平衡),而这种不平衡的特性在实际应用中是普遍存在的。另外,现代工业系统多由互相关联、紧密耦合的多个工艺环节组成,上述不平衡特性还存在着很大程度的不确定性、非线性、时变性,这对故障诊断的关键环节之一——故障特征的提取——造成了很大的困难。从一定意义上说,特征提取是故障诊断研究的“瓶颈”问题,它直接关系到故障诊断的准确性和早期预报的可靠性。因此,面向不平衡样本的故障特征提取方法研究是一个具有重要理论意义和很高工程实用价值的科学问题。

1.3.1 样本标记不平衡

工业过程中产生的数据大多数是没有类别标注的,只有少数样本有类别标注,而且标注样本的类别需要花费大量的人力、物力、财力,代价非常昂贵也异常困难。传统的故障诊断方法往往只使用少量的有标注样本进行特征提取进而训练分类器,然而大量的未标注样本不仅含有对分类有用的信息,而且可以自动廉价地获得。显然,如果只使用少量的有标注样本而不利用大量“廉价”的无标注样本进行训练,不仅在故障特征提取过程中可能因样本分布不全面而导致训练出来的分类器泛化能力不足,而且也是对资源的一种极大浪费。如果能把大量未标注样本加入到学习算法中去辅助少量的标注样本,不仅能够更全面地刻画样本空间的数据分布情况,而且还可以使分类器获得更好的泛化性能^[16]。

因此,如何合理地利用两者,减少人力标记数据样本的成本,成为近年来基于数据驱动的故障诊断方法研究的热点,半监督学习的方法也应运而生。半



监督学习利用少量的标记数据样本训练系统的分类器,同时可以利用大量的非标记样本来提升分类器的性能和精确性。

最近一些基于半监督学习的故障诊断方法已经被研究人员提出,如文献[17]提出一种基于半监督学习的独立因子分析算法,该算法应用于铁路装置的故障诊断,取得了较好的诊断效果;文献[18]提出一种半监督LDA的方法并应用于建筑能量的故障检测;文献[19]提出了一种结合独立主元分析和高斯混合模型的半监督的故障诊断方法并应用在化工领域。

1.3.2 样本价值不平衡

在20世纪,包括故障诊断在内的许多领域使用的特征不超过40个。近年来,这种情况发生了极大的改变,特别是随着集散控制系统及智能仪表在工业过程中的广泛应用,大量过程数据被记录下来,这类数据集往往样本点数量巨大,用于刻画系统特征的指标变量众多,并且带有动态特性,从而形成了规模宏大、复杂难辨的数据海洋。然而这些刻画系统特征的众多变量对故障诊断的价值却是不同的,它们基本上可以被分为三类:相关特征变量、冗余特征变量和不相关特征变量。不相关特征变量包含的噪声会降低故障诊断精度。冗余特征变量一般不提供额外信息,却增加了模型复杂度和诊断时间。相关特征变量是对故障敏感且对故障诊断有真正贡献的特征。降维技术旨在移除特征集中的不相关特征和冗余特征,只保留相关特征,是故障诊断中数据预处理的重要环节,并且越来越被证明是基于数据的故障诊断方法中必不可少的一部分。降维技术带来了许多好处^[20],例如能更好地解释故障机理,提高故障诊断的精度;降低系统复杂度,减少系统所需的计算资源。因此,降维技术吸引了很多研究者的目光,他们将其应用到了不同的故障诊断系统中,降维技术也是本书的主要研究内容。

1.3.2.1 线性数据降维方法

数据降维算法可以分成线性和非线性数据降维方法两类。线性降维方法包括主元分析(Principal Component Analysis, PCA)^[21]、部分最小二乘(Partial Least Squares, PLS)、独立元分析(Independent Component Analysis, ICA)^[22]、线性判别分析(Linear Discriminant Analysis, LDA)^[23]等。

下面介绍PCA方法和LDA方法,这是两种最基本的MSPM方法,大部分故障诊断方法都遵循与其类似的框架,且对更为复杂的故障诊断方法有重要的参考意义。

(1) PCA故障检测方法

给定一个由具有零均值的数据点构成的矩阵 $X = [x_1, x_2, \dots, x_n] \in \mathbf{R}^{d \times n}$, PCA



首先求一个正交线性变换 $\hat{W} = [\hat{w}_1, \hat{w}_2, \dots, \hat{w}_d] \in \mathbf{R}^{d \times d}$, 使得经过投影后的数据 $\hat{Y} = \hat{W}^T X = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n] \in \mathbf{R}^{d \times n}$ 在第一维度上的投影具有最大的方差, 在第二维度上的投影具有第二大的方差, 以此类推。变换矩阵 $\hat{W}$ 可由样本协方差矩阵 C_X 的特征值分解求得:

$$C_X = \frac{1}{n} X X^T = \hat{W} \hat{\Lambda} \hat{W}^T \quad (1.1)$$

式中, 对角矩阵 $\hat{\Lambda}$ 中的对角元素即特征值从大到小排列。由于投影后数据的协方差矩阵

$$C_Y = \frac{1}{n} \hat{Y} \hat{Y}^T = \frac{1}{n} \hat{W}^T X X^T \hat{W} = \hat{W}^T \hat{W} \hat{\Lambda} \hat{W}^T \hat{W} = \hat{\Lambda} \quad (1.2)$$

所以投影后各个维度上的方差等于 $\hat{\Lambda}$ 中对应的对角元素。通过保留对应的方差最大的那些维度可以实现降维并保持数据的方差结构。假定保留前 r 个维度, 则 PCA 投影 $W = [w_1, w_2, \dots, w_r] \in \mathbf{R}^{d \times r} (r < d)$ 。记 $\tilde{W} = [W, \tilde{W}]$, 则主元空间 $S_p = \text{span}(P)$, 互补的残差空间 $S_{\bar{p}} = \text{span}(\tilde{P})$ 。任意数据点 x 投影到主元空间中的坐标 $y = W^T x$, 位于残差空间内的投影残差为 $x - P^T x$ 。

故障检测方法通常同时监测主元空间和残差空间来捕获故障。 T^2 统计量是典型的主元空间波动程度的度量, 它的物理意义是主元空间中的点与原点间的马氏距离的平方。数据点 x 在主元空间中的 T^2 统计量为

$$T^2 = x^T W \Lambda^{-1} W^T x \quad (1.3)$$

式中, Λ 为训练数据在主元空间中的协方差矩阵, 等于 $\hat{\Lambda}$ 前 r 个对角元素对应的子对角矩阵。假如数据满足多元正态分布, 则 T^2 满足

$$\frac{n(n-d)}{d(n^2-1)} T^2 \sim F_{d, n-d} \quad (1.4)$$

式中, $F_{d, n-d}$ 是自由度为 $(d, n-d)$ 的 F 分布。给定显著性水平 α , T^2 统计量的控制上限为

$$T_{\alpha}^2 = \frac{d(n^2-1)}{n(n-d)} F_{d, n-d; \alpha} \quad (1.5)$$

平方预测误差 (Squared Prediction Error, SPE) 统计量是典型的残差空间波动程度的度量, 它的物理意义是降维残差长度的平方。一个输入空间中的数据点 x 在残差空间中的 SPE 统计量为

$$\text{SPE} = \|x - PP^T x\|^2 \quad (1.6)$$

假如数据满足多元正态分布, 给定显著性水平 α , SPE 统计量的控制上限为

$$\delta_{\alpha}^2 = \eta \chi_{h; \alpha}^2 \quad (1.7)$$



式中, $\eta = \frac{v}{2m}$, $h = \frac{2m^2}{v}$, 其中, m 和 v 分别是根据训练数据估计的 SPE 统计量的平均值和方差; χ_h^2 是自由度为 h 的 χ 分布。

假如数据不满足正态分布, 则 T^2 统计量和 SPE 统计量的控制上限可在给定显著性水平后用核密度估计 (Kernel Density Estimation, KDE) 来计算。

PCA 故障检测方法的离线建模阶段的主要步骤如下:

① 收集过程在正常状态下的训练数据, 经过去中心和方差归一化处理得到数据 $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{d \times n}$;

② 对 X 实施 PCA 得到投影 W , 保留的维数 r 可由交叉验证或主元方差累积百分比等方法确定;

③ 对 X 实施 PCA 投影, 计算训练数据的 T^2 和 SPE 统计量以及 T^2 和 SPE 统计量的控制上限。

在线监测的主要步骤如下:

① 根据建模阶段训练数据的均值和方差对某时刻在线新测量的数据 $x \in \mathbb{R}^d$ 进行平移和伸缩;

② 计算 x 的 PCA 投影 $y = W^T x$ 及 x 的 T^2 和 SPE 统计量, 使用控制图监测它们是否超过控制上限。

(2) LDA 故障识别方法

给定数据矩阵 $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{d \times n}$ 和每个数据点对应的类标识 $l(x_1), \dots, l(x_n) \in \{1, \dots, c\}$ 。LDA 如下定义类内离散度矩阵 S_w 和类间离散度矩阵 S_b :

$$S_w = \sum_{k=1}^c n_k S_k = \sum_{k=1}^c \sum_{x_i \in C_k} (x_i - \bar{x}_k)(x_i - \bar{x}_k)^T \quad (1.8)$$

$$S_b = \sum_{k=1}^c n_k (\bar{x}_k - \bar{x})(\bar{x}_k - \bar{x})^T \quad (1.9)$$

式中, S_k 是类 k 的样本协方差矩阵, n_k 为属于类别 k 的数据点数, 且

$$\bar{x}_k = \frac{1}{n_k} \sum_{x_i \in C_k} x_i \quad (1.10)$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1.11)$$

给定降维维数 r , LDA 投影矩阵 W 定义为

$$J_{\text{LDA}} = \arg \max_{\substack{W \in \mathbb{R}^d \\ W^T W = I}} \frac{W^T S_b W}{W^T S_w W} \quad (1.12)$$

对任意数据点 x , LDA 投影为 $y = W^T x$ 。

记 $W = [w_1, w_2, \dots, w_r]$ 为以下广义特征值问题的与广义特征值 $\lambda_1 > \lambda_2 > \dots > \lambda_r$ 相对应的广义特征向量: