

支持向量机用于液体火箭 发动机的故障诊断

何 浩, 胡小平, 姜志杰, 刘伟强

(国防科技大学 航天与材料工程学院, 湖南 长沙 410073)

摘 要: 支持向量机 (Support Vector Machine, 简称 SVM) 是一种基于机器学习的模式分类算法, 其在解决小样本、非线性及高维模式识别等问题中都表现出许多特有的优势。用 SVM 对液体火箭发动机的故障数据进行检测和诊断。通过对发动机仿真模型的 9 种故障数据的学习, 能检测出 18 组故障数据中的 17 组, 但有 4 组出现误报, 对误报故障进行二次学习和再检测, 能对这 4 种故障正确检测。经过对 C75 试车 4 种故障数据的学习, 能正确检测其故障类型, 进一步验证了该方法的正确性和可行性。

关键词: 支持向量机; 液体火箭发动机; 故障诊断; 模式识别

中图分类号: V434

文献标识码: A

文章编号: (2008) 03-0007-06

SVM implemented in fault diagnosis of liquid rocket engine

He Hao, Hu Xiaoping, Jiang Zhijie, Liu Weiqiang

(Inst. of Aerospace and Material Engineering, National Univ. of Defense Technology, Changsha 410073, China)

Abstract: SVM which is based on machine learning algorithm is a method for pattern classification. The advantage of SVM is to solve the small samples, no-linear and pattern recognition with high dimension. In this paper, the method of SVM is used in fault diagnosis for the data of practical LRE trial run and simulated model. The SVM classifier detects the four faults of heat run completely. Among the eighteen groups of simulated model data, seventeen groups of them can be detected, although incorrect warnings are happened during four groups detecting. Through learning and detecting on the four groups again, the faults are divided and detected correctly.

Key words: SVM; liquid rocket engine; fault diagnosis; pattern recognition

收稿日期: 2008-01-08; 修回日期: 2008-03-10。基金项目: 国家自然科学基金(50376073)资助项目。

作者简介: 何浩 (1982—), 男, 硕士, 研究领域为火箭发动机故障诊断, 数据挖掘。

1 引言

液体火箭发动机是一种复杂的动力系统,它是由大量的、相互之间存在广泛联系的组件和部件组成。随着运载火箭的有效载荷不断增加,发动机的工作参数值随之不断提高,以至发动机的工作环境越来越恶劣,于是对其可靠性和安全性提出了越来越高的要求,加之相关科学技术的快速发展所提供的实现可能,使得对液体火箭发动机健康监控的研究受到普遍重视^[1]。在推进技术领域,用于液体火箭发动机的故障诊断可分为三类:基于测量信号、基于模型及基于人工智能。由于试车数据有限,且发动机结构复杂,检测和诊断的有效性还有待提高^[2]。用一种新方法 SVM 实现对液体火箭发动机的故障进行检测和诊断。

SVM 是 Vapnik 等根据统计学习理论提出的一种模式分类方法^[3,5]。相比其他传统算法,如神经网络^[2]、遗传算法等的实现中带有很大的经验成分,而 SVM 有严格的理论和数学基础。其采用核函数思想把非线性问题转化为线性问题,降低了算法的复杂度。由于 SVM 的最优解基于已有的样本信息,而不是传统算法的样本趋于无穷大时的最优解,因此其专门针对小样本问题,而实际中样本数目往往是有限的, SVM 相比其他算法更具有推广能力。

目前, SVM 广泛应用于各领域。如人脸识别、网络入侵检测及视频字幕提取等。Rychetsky 将其用于内燃机的故障诊断,杨琦等用 SVM 对液压设备故障进行诊断^[6]。

2 支持向量机

2.1 支持向量机

2.1.1 线性可分

SVM^[4]最初是针对模式识别中的线性可分问题。

图 1 中,空心圆和圆中心有十字的点分别代表两类样本, H 为分类线, H₁、H₂ 分别为过各类中离分类线最近的样本且平行于分类线的直线,

它们之间的距离叫做分类间隔。所谓最优分类线就是要求分类线不但能将两类正确分开,即错分率为 0,而且使分类间隔最大。

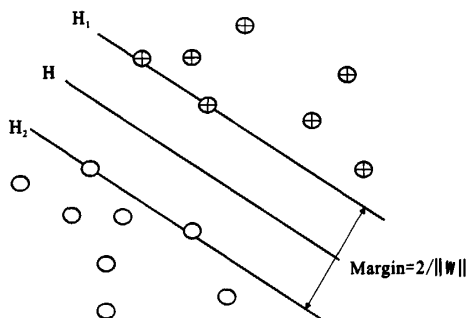


图 1 线性可分情况下的最优分类线

Fig.1 Linear separable optimum classification line

设有 n 个样本 x_i $i=1,2,\dots,n$, 每个样本为 d 维向量。 y_i 为每个样本的所属类别 (即每个样本的标号)。如果存在超平面方程为

$$W_i \cdot x_j + b = 0 \quad (1)$$

使得

$$y_i [W_i \cdot x_j + b] \geq 1, i=1,2,\dots,n \quad (2)$$

此时分类间隔为 $2/\|W\|$, 要使间隔最大等价于使 $\|W\|$ 最小或使 $1/2\|W\|^2$ 最小。这样, 满足公式 (2) 且使 $1/2\|W\|^2$ 最小的分类面就叫做最优分类面, 把过两类样本中离分类面最近的点, 且平行于最优分类面的超平面 H₁、H₂ 上的训练样本点称作支持向量。

定义 Lagrange 函数如下:

$$L(W,b,\alpha) = \frac{1}{2}(W_j \cdot W_j) - \sum_{i=1}^n \alpha_i ((y_i (W_j \cdot x_i) + b) - 1) \quad (3)$$

式中, α_i 为 Lagrange 系数。将上式对 W 和 b 求 Lagrange 函数的极小值。为此分别对 W 和 b 求偏微分并令他们等于 0, 就可以把上述最优分类面问题转化为其对偶问题。再加上原约束条件可以把原问题转化为如下凸问题二次规划的对偶问题:

$$\begin{cases} \max \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_j \alpha_i y_i y_j (x_i^T x_j) \\ \alpha_i \geq 0, i=1,2,\dots,n \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{cases} \quad (4)$$

式中, α_i 为与每个样本对应的 Lagrange 乘子。解上述问题后得到的最优分类函数是

$$f(x_k) = \text{sgn}\{(\mathbf{W}_j \cdot x_i) + b\} = \text{sgn}\left\{\sum_{i=1}^n \alpha_i y_i (x_i \cdot x_k) + b\right\} \quad (5)$$

式中, b^* 是分类阈值, 可以用任一个支持向量(满足公式(2)中的等号)求得, 或通过两类中任意一对支持向量取中值求得。

2.1.2 线性不可分

在线性不可分的情况下, 可在公式(2)中增加一个松弛项 $\xi_i \geq 0$, 成为

$$y_i [(\mathbf{W}_j \cdot x_i) + b] \geq 1 - \xi_i \quad (6)$$

其中

$$i=1, 2, \dots, n$$

将目标改为求下式最小值

$$(\mathbf{W}_j, \xi_i) = \frac{1}{2} \|\mathbf{W}_i\|^2 + C \left[\sum_{i=1}^n \xi_i \right]$$

式中, $C > 0$ 是一个常数, 为惩罚因子。对于线性不可分的最优分类面与线性可分情况下几乎完全相同, 只是公式(5)约束条件有如下改变:

$$0 \leq \alpha_i \leq C, i=1, 2, \dots, n \quad (7)$$

2.1.3 非线性问题

对非线性问题, 可以通过非线性变换转化为某个高维空间中的线性问题, 再变换空间求解最优分类面。根据泛函的有关理论, 只要一种核函数 $K(x_i, x_j)$ 满足 Mercer 条件, 它就对应某一变换空间中的内积。因此, 在最优分类面中采用适当的内积函数 $K(x_i, x_j)$ 可以实现某一非线性变换后的线性分类, 此时目标函数变为

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \quad (8)$$

而相应分类函数也变为

$$f(x_k) = \text{sgn}\{(\mathbf{w}_j \cdot x_i) + b\} = \text{sgn}\left\{\sum_{i=1}^n \alpha_i y_i K(x_i \cdot x_k) + b\right\} \quad (9)$$

2.2 核函数

对于非线性不可分问题, SVM 首先通过一个非线性映射将原空间样本映射到高维特征空间(核空间), 变为线性可分或近似线性可分, 实现对原空间的非线性算法, 其关键是引入了核函数, 因此核函数的选择对 SVM 有重要影响。常用核函数有: 多项式核函数、径向基核函数及

Sigmoid 核函数。

目前还没有很好的办法指导针对具体问题的核函数选择。齐志泉^[7]提出一种基于混合遗传算法求解 LOO 上界极小点的核函数选择方法。针对模式识别问题, 李盼池、许少华等^[8]指出对于规模小、维数低的模式识别对核函数类型不敏感; 但对于维数较高时应采用多项式和径向基核函数; 对维数较低、规模较大的模式识别问题应采用径向基和 Sigmoid 核函数。针对液体火箭发动机试车数据特点及试算对比采用多项式核函数。

3 基于支持向量机的故障诊断和检测

用林智仁博士开发的 Libsvm 设计一个分类器来进行故障检测, 分为数据训练和故障预测两部分。训练模型时, 每组数据对应一个标号(代表其类别)。训练时对应要检测的故障数据标号为+1, 其余为-1。待检测数据用训练好的模型依次预测, 产生一个标号(为-1或+1)。若预测标号+1, 表明检测数据与检测模型有相同的故障类型。当多组同时预测时, 预测为+1的组数占总组数的比例即为预测精度, 即可能发生与检测模型相同故障的概率。核函数取多项式核函数, SVM 算法选择的是 c-svc (一种 SVM 分类算法)。

3.1 数据预处理

由于火箭发动机试车经历了起动、稳态及关机几个过程, 监测参数众多, 产生大量试车数据, 且不同阶段参数有显著差异。如 C7527 试车时, 起动段氢涡轮泵转速在 0~2.6 秒内由 0 增加到 45,000 转/分, 稳态段氢涡轮泵转速稳定在 40,000 转/分左右。因此有必要对数据进行适当的预处理。针对火箭发动机的试车数据的特点, 采取以下几种方法对数据进行预处理。

3.1.1 数据离散及一致化

火箭发动机各参数值随时间不断变化(尤其在起动及关机时), 为充分利用试车数据的有用信息, 将连续时间离散为一个时刻, 根据需要选取不同时刻参数。另外, 由于每次试车数据采集步长不一致, 且监测参数各不同, 因此需对数据采集, 提取相同时刻参数值一致化数据。

3.1.2 数据选取

为了保证检测结果的正确性，检测前需进行数据分析，剔除有明显错误的数据。对剧烈波动的参数要分析其是确实发生故障还是由于传感器误差导致。若为故障数据则保留并做平滑处理来用于检测。另外监测参数众多，选取时可选关键参数数据，选取 C75 的 15 个关键参数，仿真模型的 5 个参数来监测。

3.2 对仿真模型故障数据的检测和诊断

由于试车数据有限，尤其是故障试车数据。通过对仿真模型模拟故障数据的检测，检验方法的正确性和可行性是十分必要的。对参考文献[2]发动机仿真模型数据，模拟 9 种故障类型，监测其 5 个参数，如表 1 和表 2。参数值做如下处理，参数值 = (故障时值-正常值) / 正常值。

表 1 仿真模型故障类型

Tab.1 Fault types of simulation model

序号	故 障 名 称
1	氧化剂主阀故障
2	燃料主阀故障
3	两主阀同时有故障
4	燃气发生器氧化剂管路过滤网阻塞
5	燃气发生器燃料入口处泄漏
6	氧化剂喷注器阻塞
7	推力室喉部烧蚀
8	泵效率损失
9	涡轮入口燃气泄漏

表 2 仿真模型监测参数

Tab.2 Monitoring parameters of simulation model

序号	符号	故 障 名 称
1	q_{mo}	氧化剂流量
2	q_{mf}	燃料剂流量
3	p_{α}	氧化剂喷前压力
4	p_n	燃料主管分支处压力
5	n	涡轮泵转速

每种故障有 5 组数据，用其中三组来训练，剩余两组作为待检测数据。用 9 种故障每种中的三组数据训练 9 个模型。将待检测数据依次用训练好的 9 个 model 来检测，每次生成一个预测精度。这样就形成一个形式上为一个 9×9 的矩阵(定义为精度矩阵)。矩阵的每个元素为一预测精度 (accuracy)，每个元素的下标代表所检测的故障数据组数及所用的 model 数。例如， $accuracy_{43}$ 代表用第 3 个训练的 model 来预测第 4 组故障数据的精度。分析可知，所生成的精度矩阵对角线上的元素应为 100% (若不为 1，则为漏报)，且矩阵中每列中只有对角线上元素不为 0 (否则为误报)。精度矩阵如下：

$$\begin{bmatrix} accuracy_{11} & \cdots & accuracy_{1n} \\ \vdots & \ddots & \vdots \\ accuracy_{n1} & \cdots & accuracy_{nn} \end{bmatrix}$$

生成的精度矩阵如表 3 所示。表格中的值即为预测精度 (空白处预测精度为 0)。由上述分析可得出的结果，如表 4 所示。

表 3 仿真模型故障检测精度表

Tab.3 Detection precision table of simulation model

种类	1	2	3	4	5	6	7	8	9
1	1								
2		1							
3			0.5						
4				1				0.5	
5					1				1
6						1			
7							1		
8					1			1	
9					1				1

表 4 首次检测结果

Tab.4 Result of fault detection for the first time

精度	检测结果	说 明
检测率	0.9474	18 组数据检测出 17 组
误报率	0.6316	有 12 组数据误报
漏报率	0.0526	有 1 组没检测出

该方法需解决的是降低误报率, 进一步提高检测性能。一种解决方案为, 在检测时若发生误报, 再对该组数据作进一步检测, 直至确定该组数据的故障类型。如第 9 种故障被同时检测出三种类型故障 (第 5、8、9 种故障), 为进一步确定其故障类型, 设计一种多故障分类器, 先用误报的三种故障的数据来训练, 得到三个新模型, 训练过程与第一步相同。这样再对第 9 种故障用这三个模型来检测, 可得三个预测精度, 精度最高的即为最后要检测的故障类型。经计算, 生成的三个精度分别为 $accuracy_1=0.5$, $accuracy_2=0$, $accuracy_3=1$, 可见该故障为第 9 种类型故障, 与实际相符。对其他的误报可用同样方法检测出其故障。

3.3 对实际试车数据的检测和诊断

选取 C75 试车的一种故障数据和一种正常数据 (看作第四种故障)。监测其中的 15 个参数, 如表 5、6 所示。

表 5 实际试车监测参数

Tab.5 Monitoring parameters of practical test	
序号	故障名称
1	氢泵转速
2	氧泵转速
3	氢泵流量
4	氧泵流量
5	燃烧室压力
6	氢涡轮入口压力
7	氢涡轮出口压力
8	氧涡轮入口压力
9	氧涡轮出口压力
10	氢涡轮入口温度
11	氢涡轮出口温度
12	氧涡轮入口温度
13	氧涡轮出口温度
14	氧贮箱压力
15	氢贮箱压力

表 6 实际试车故障数据说明

Tab.6 Illustration of practical test fault data	
故障名称	训练数据说明
正常试车	60.5 秒至 80 秒为正常数据
氢泵故障	氢泵欠速在 275.8 秒左右关机
参数过低	二次启动工作参数低
氧副文氏管堵塞	在 42.64 至 43.20 秒时发生堵塞

对氧副文氏管发生堵塞的试车数据进行检测, 结果如表 7 所示。检测过程如图 2 所示。

表 7 对实际试车数据的检测结果

Tab.7 Detection result of the practical test data		
精度号	精度	结果说明
$accuracy_1$	0%	正常可能性为:0
$accuracy_2$	0%	氢泵故障的可能性为:0
$accuracy_3$	0%	参数过低的可能性为:0
$accuracy_4$	100%	氧副文氏管堵塞可能性为:100%

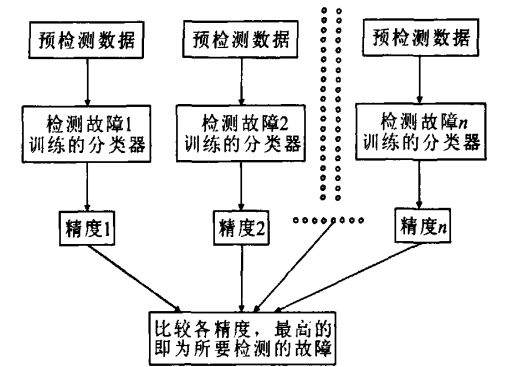


图 2 多故障检测示意图

Fig.2 The flow chat of faults detection

由表 7 可知, 第 4 类故障预测的精度最高, 即所检测的故障为氧副文氏管入口处堵塞。用同样方法也能正确检出另外三种故障。该方法的检测正确率偏高, 有几方面原因: 首先, 检测时选用参数多 (15 个), 就能充分提取故障数据的有用信息; 其次, 由于试车数据有限, 诊断的故障数据仍是训练数据的试车数据, 虽为不同时刻, 但出自同一次试车; SVM 本身的优越性也是一个重要的因素。

3.4 讨论

SVM 对小样本有很好的检测能力, 在对 C75 试车故障检测中若只取 10 个参数数据训练, 采样步长加倍, 仍能正确检测。在仿真模型故障检测中, 只取每种故障五组数据中的两组来训练, 检测率无显著下降。可见其在小样本学习、分类方面有良好的性能。

对于数据噪声 SVM 具有较好的抗干扰能力。在仿真模型故障检测中, 加入一组错误数据训练, SVM 仍保持较高的检测率。但加入两组时, 错分率就会增大 (噪声淹没原始数据)。可见, SVM 对于小噪声有很好抗干扰能力。

4 结束语

SVM 对 9 种仿真模型故障数据经过两次检测得到 94.4% 检测率, 而且误报率也很小。很好地解决了小样本、非线性的分类问题, 表现出较高的检测效率, 对于相同的仿真模型故障数据, 神经网络的检测率仅为 50%^[2]; 对 4 种实际试车故障能够完全正确检测; 且具有良好的小样本学习能力和抗干扰能力。

参考文献:

- [1] 张育林. 液体火箭发动机健康监控技术[M]. 长沙: 国防科技大学, 1998.
- [2] 吴建军. 液体火箭发动机故障检测与诊断[D]. 长沙: 国防科技大学, 1995.
- [3] 张学工. 关于统计学习理论与支持向量机 [J]. 自动化学报, 2000, 26(1): 33-42.
- [4] 祁享年. 支持向量机及其应用研究综述 [J]. 计算机工程, 2004, 30(10): 6-9.
- [5] Vapnik V N. Estimation of Dependences Based on Empirical Data [J]. Berlin: pringer-Verlag, 1982, 13(2): 46-49.
- [6] Davide Mattera, Francesco Palmieri, Simon Haykin. An Explicit Algorithm for training Support Vector Machines [J]. Signal processing letters, 1999, 24(5): 53-57.
- [7] 齐志泉, 田英杰, 等. 支持向量机中的核参数选择问题[J]. 控制工程, 2005, 12(4): 379-381.
- [8] 李盼池, 许少华. 支持向量机在模式识别中的核函数特性分析 [J]. 计算机工程与设计, 2005, 26(2): 302-304.

(编辑: 马 杰)

(上接第 38 页)

- [5] Little R R, Chelner H, Buswell H J. Development, Testing and Application of Embedded Sensors for Solid Rocket Motor Health Monitoring [C]. 37th International Annual Conference of ICT, Karlsruhe, Germany, June, 2006.
- [6] Cunliffe A V. Fraction Measurements-A Tool to Study Cross-Linking and Ageing in Composite Propellants and PBXs [C]. 37th International Annual Conference of ICT, Karlsruhe, Germany, June, 2006.
- [7] Nevieri R, Guyader M. DMA: A Powerful Technique to Assess Ageing of MED [C]. 37th International Annual Conference of ICT. Karlsruhe, Germany, June, 2006.
- [8] Gottlieb L. Migration of Plasticizer between Bonded Propellant Interfac [J]. Propellants, Explosives, Pyrotechnics, 2003, 28(1): 12-17.
- [9] Grythe K F, Hansen F K. Diffusion Rates and the Role of Diffusion in Solid Propellant Rocket Motor Adhesion[J]. Journal of Applied Polymer Science, 2007, 103 (3): 1529-1538.
- [10] Freitas de Moraes A M, Ribeiro Pinto J, Aguiar, J. Optimization of Bondline Properties of Solid Rocket Motors [C]. 37th International Annual Conference of ICT. Karlsruhe, Germany, June, 2006.
- [11] HO S Y. Viscoelastic Response of Solid Rocket Motor Components for Service Life Assessment [J]. Journal of Materials Science, 1997, 32(19): 5155-5161.

(编辑: 王建喜)