

火箭发动机故障检测的快速增量单分类支持向量机算法

张万旋,张箭,卢哲,薛薇,张楠*
(北京航天动力研究所,北京 100076)

摘要:为解决液体火箭发动机故障诊断正负样本不平均问题,以及实现发动机稳态工作段自适应故障检测,建立了基于快速增量单分类支持向量机的异常检测模型。采取特征工程方法,对传感器获得的多变量时间序列进行特征提取。通过增量学习方法,对单分类支持向量机模型进行改进,并应用于液体火箭发动机异常检测,使单分类支持向量机检测模型具备对不同台次、不同工况的自适应性,提高了模型的计算速度。对多台次热试车数据的分析结果表明,该模型十分有效,训练速度快,具备实用价值。

关键词:单分类支持向量机;特征提取;自适应检测;增量学习;异常检测

中图分类号:V434 **文献标志码:**A **文章编号:**1001-2486(2024)02-115-08

Fast incremental one-class support vector machine algorithm for rocket engine fault detection

ZHANG Wanxuan, ZHANG Jian, LU Zhe, XUE Wei, ZHANG Nan*
(Beijing Aerospace Propulsion Institute, Beijing 100076, China)

Abstract: In order to solve the problem of imbalance between positive and negative samples in liquid rocket engine fault diagnosis, and to enable adaptive fault detection during engine steady working state, a anomaly detection model based on fast incremental one-class support vector machine was established. Feature engineering method was adopted to extract features from sensor-obtained multivariate time series. Through incremental learning, the one-class support vector machine model was improved and applied to liquid rocket engine anomaly detection. The one-class support vector machine detection model was endowed with adaptability for various engine individuals and multiple working conditions, while increasing computing speed. The analysis results of multiple hot test data show that the model is effective, fast-training and practically valuable.

Keywords: one-class support vector machine; feature extraction; adaptive detection; incremental learning; anomaly detection

液体火箭发动机是航天运载器故障敏感多发部位^[1-2],发动机故障诊断面临故障样本稀缺、正负样本不平均、故障模式不完备的问题。故障数据获取的代价是高昂的,因此实际热试车数据以正常数据为主。另一方面,由于发动机工况变化、不同台次稳态参数偏差,要求检测算法具有一定自适应性。

针对发动机自适应检测需求,研究者提出了多种自适应检测算法。朱恒伟等^[3]提出了自适应阈值算法(adaptive threshold algorithm, ATA),该算法利用滑窗内样本的均值和标准差构造检测阈值。杨雪等^[4]将自回归(auto regressive, AR)模型对均值进行估计,提出了基于自回归模型的自适应阈值算法。聂饶^[5]针对过程神经网络预测模型逼近发动机的动态过程,根据预测残差进

行故障检测,并基于增量学习方法提出了过程神经网络检测算法的自适应更新方法。

异常是指“少而不同”的样本^[6],异常检测^[7]是指无监督或者弱监督条件下的分类任务。异常检测技术十分适合液体火箭发动机故障检测以正样本为主的情况。

单分类异常检测模型可以在仅有正常样本数据的情况下发现异常或未知状态样本,在此基础上发展的自适应算法,可以为不完备信息故障检测、自适应在线检测提供有效途径。自Schölkopf等^[8]和Tax等^[9]分别提出 ν -支持向量机(ν -support vector machine, ν -SVM)和支持向量机数据描述(support vector data description, SVDD)两种典型单分类支持向量机以来,研究者们对其进行了诸多改进^[10-11]。胡雷^[12]针对液体火箭发动

机涡轮泵故障检测问题,将序贯最小优化算法^[13]引入单类支持向量机,提高了单分类问题求解效率。彭小辉^[14]针对航天器推进系统故障检测类别不平衡以及误分代价不等等问题提出了一种边界可调整的代价敏感支持向量数据描述方法,并通过一种增量学习算法节省了训练时间。Zhu 等^[15]针对某燃气发生器循环发动机,提出以卷积神经网络构建自编码器对 22 路性能参数进行特征提取,再将提取后的特征输入 ν -SVM 模型进行故障检测。虽然上述方法有效降低了单分类问题的时间和空间复杂度,但是仍然依赖梯度下降法求解二次规划问题,计算速度仍有一定提升空间。

本文在传统 SVDD 模型基础上,引入增量学算法,开发了快速增量单分类支持向量机检测算法,并将其应用于某型发动机的自适应异常检测。这里,异常检测是指发动机在缺少故障样本情况下的故障检测手段。因此,本文用异常检测概念替代该条件下的故障检测概念,不再赘述。

1 快速增量单分类支持向量机检测算法

1.1 原始 SVDD 模型

可以证明^[16],两种典型单分类支持向量机模型 ν -SVM 和 SVDD 在采用高斯核的情况下是等价的,本文采取 SVDD 模型作为单分类器。SVDD 通过寻找高维空间正常样本点集的最小包络球来确定检测边界,若待测样本落在包络球内部或边界上,则认为待测样本正常,反之判定为异常。SVDD 模型为:

$$\begin{aligned} \min_{a, R, \xi_i \geq 0} & R^2 + C \sum_{i=1}^N \xi_i \\ \text{s. t. } & \|x_i - a\|^2 \leq R^2 + \xi_i \\ & \forall i = 1, 2, \dots, N \end{aligned} \quad (1)$$

式中: R 表示超球半径; ξ_i 表示松弛量; N 表示样本点总数; $C = \frac{1}{\nu N}$ 表示惩罚系数,由人为设定, $0 < C \leq 1$,由大数定律可保证,当 N 足够大,至少有 $1 - \nu$ 比例的正常数据将落在优化解对应的超球内部或边界上^[11]; x_i 表示原始特征空间样本点; a 表示包络球中心。通常引入映射函数 $\phi(x)$ 将样本点从特征空间映射到高维空间,使得原始空间中不可分的点集在高维空间中可分,并引入核函数 $K(x)$ 使得高维空间的内积可在原空间中进行计算,即 $\phi(x) \cdot \phi(y) = K(x \cdot y)$ 。

利用拉格朗日乘子法^[17]将式(1)转化为:

$$\begin{aligned} \min_{a, R, \xi_i, \alpha_i, \beta_i, C \geq 0} & L = R^2 + C \sum_{i=1}^N \xi_i - \\ & \alpha_i (R^2 + \xi_i - \|x_i - a\|^2) - \beta_i \sum_{i=1}^N \xi_i \end{aligned} \quad (2)$$

式中, α_i, β_i 表示拉格朗日乘子, L 表示目标函数。

对式(2)求偏导可得:

$$\begin{cases} \frac{\partial L}{\partial R} = 0 \Rightarrow \sum_{i=1}^N \alpha_i = 1 \\ \frac{\partial L}{\partial a} = 0 \Rightarrow a = \sum_{i=1}^N \alpha_i x_i = \mathbf{1} \\ \frac{\partial L}{\partial \xi_i} = 0 \Rightarrow \beta_i = C - \alpha_i \end{cases} \quad (3)$$

可将式(2) 转化为如下二次规划问题对偶形式:

$$\begin{aligned} \max_{0 \leq \alpha_i \leq C} & L = \sum_{i=1}^N \alpha_i K(x_i, x_i) - \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(x_i, x_j) \\ \text{s. t. } & \sum_{i=1}^N \alpha_i = 1 \end{aligned} \quad (4)$$

由于式(1) 存在不等式约束,因此式(4) 需满足 KKT(Karush-Kuhn-Tucker) 条件^[17]:

$$\begin{cases} \alpha_i (R^2 + \xi_i - \|x_i - a\|^2) = 0 \\ (C - \alpha_i) \xi_i = 0 \end{cases} \quad \forall i \quad (5)$$

由式(4) 和约束条件 $0 \leq \alpha_i \leq C$ 可得:

$$\|x_i - a\|^2 < R^2 \Rightarrow \alpha_i = 0 \quad (6)$$

$$\|x_i - a\|^2 = R^2 \Rightarrow 0 < \alpha_i < C \quad (7)$$

$$\|x_i - a\|^2 > R^2 \Rightarrow \alpha_i = C \quad (8)$$

式(6) 对应的样本点落在超球内部,称为非支持向量。式(7) 对应的样本点落在超球边界上,称为边界支持向量。式(8) 对应的样本点落在超球边界之外,称为非边界支持向量。图 1 给出了 SVDD 算法超球和样本点示意图。

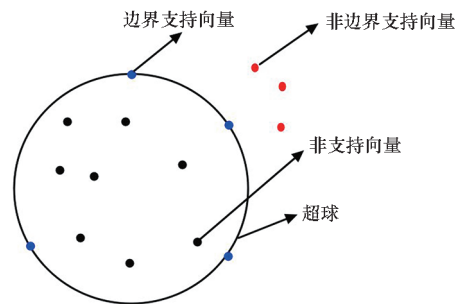


图 1 SVDD 模型示意图

Fig. 1 Sketch of SVDD model

通常采用梯度下降类方法^[18]求解式(4),如拟牛顿法、共轭梯度法、信赖域法,该类方法的时间复杂度可以表示为 $O(N^3)$,其中, N 为样本数量。当样本数量较大时,采用该类方法将导致计算异常缓慢,这对于发动机离线和在线故障诊断

都是不可接受的。

1.2 快速增量算法数学基础

在模型训练阶段,假设输入模型的样本点均在超球内部或边界上,由惩罚系数 C 的物理意义可以设定 $C = 1$ 。当采用高斯函数作为核函数 $K(\mathbf{x})$ 时,即

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{(\mathbf{x}_i - \mathbf{x}_j)^2}{2\sigma^2}\right) \quad (9)$$

式中, σ 表示高斯核带宽。

式(2)可以简化为:

$$\begin{aligned} \min_{0 \leq \alpha_i \leq C} L &= \sum_{i=1}^N \sum_{j=1}^N \alpha_i k(\mathbf{x}_i, \mathbf{x}_j) \\ \text{s. t.} \quad &\sum_{i=1}^N \alpha_i = 1 \end{aligned} \quad (10)$$

对于待测样本点 \mathbf{z} , 由式(3)和 KKT 条件式(5), 其到超球中心 \mathbf{c} 的距离 $d(\mathbf{z})$ 为:

$$\begin{aligned} d^2(\mathbf{z}) &= \|\mathbf{z} - \mathbf{c}\|^2 \\ &= 1 - 2 \sum_{i=1}^M \alpha_i k(\mathbf{z}, \mathbf{x}'_i) + \sum_{i=1}^M \sum_{j=1}^M \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}'_j) \end{aligned} \quad (11)$$

式中, \mathbf{x}'_i 表示边界支持向量, M 表示边界支持向量数目。

对于任一边界支持向量 \mathbf{x}'_k , 以下关系式成立:

$$\begin{aligned} R^2 &= \|\mathbf{x}'_k - \mathbf{c}\|^2 \\ &= 1 - 2 \sum_{i=1}^M \alpha_i K(\mathbf{x}'_k, \mathbf{x}_i) + \sum_{i=1}^M \sum_{j=1}^M \alpha_i \alpha_j K(\mathbf{x}'_i, \mathbf{x}'_j) \end{aligned} \quad (12)$$

于是, 样本 \mathbf{z} 的异常分数为:

$$Q(\mathbf{z}) = (d^2 - R^2)/2 = \sum_{i=1}^M \alpha_i K(\mathbf{x}'_k, \mathbf{x}_i) - \sum_{i=1}^M \alpha_i K(\mathbf{z}, \mathbf{x}'_i) \quad (13)$$

式中, $Q(\mathbf{z})$ 表示异常分数, $Q(\mathbf{z}) \leq 0$ 则认为样本正常, 反之认为异常。

下面, 为给本文提出的算法做铺垫, 不加证明地引入必要定理和推论, 其证明过程见文献[19]。

定理1 设 $\mathbf{A}_k = [K(\mathbf{x}'_i, \mathbf{x}_j)]_{k \times k}$ 为核函数矩阵, 则 \mathbf{A}_k 为对称正定矩阵, 且满足

$$\mathbf{A}_{k+1} = \begin{bmatrix} \mathbf{A}_k & \mathbf{v} \\ \mathbf{v}^T & 1 \end{bmatrix} \quad (14)$$

若 \mathbf{A}_k 已知, 则 \mathbf{A}_{k+1}^{-1} 可按式(14)计算:

$$\mathbf{A}_{k+1}^{-1} = \begin{bmatrix} \mathbf{A}_k^{-1} + \mathbf{p}\mathbf{p}^T/\beta & -\mathbf{p}/\beta \\ -\mathbf{p}^T/\beta & 1/\beta \end{bmatrix} \quad (15)$$

式中, $\mathbf{p} = \mathbf{A}_k^{-1} \mathbf{v}$, $\beta = 1 - \mathbf{v}^T \mathbf{p}$ 。

定理2 假设 \mathbf{A}_{k+1} 为对称正定矩阵, 且其逆满足

$$\mathbf{A}_{k+1}^{-1} = \begin{bmatrix} \mathbf{P}_{k \times k} & \mathbf{u} \\ \mathbf{u}^T & \lambda \end{bmatrix} \quad (16)$$

则 \mathbf{A}_k^{-1} 可按式(16)计算:

$$\mathbf{A}_k^{-1} = \mathbf{P} - \mathbf{u}\mathbf{u}^T/\lambda \quad (17)$$

定理3 样本点 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 全部为边界支持向量, 当且仅当式(18)有正解。

$$\mathbf{A}_k^{-1} \boldsymbol{\alpha} = \mathbf{e} \quad (18)$$

式中, \mathbf{e} 为 $k \times 1$ 元素全为 1 的矩阵。

定理4 假设样本点 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 构成一个边界支持向量集合, 则新样本点 \mathbf{x}_{k+1} 为非支持向量, 当且仅当方程 $\mathbf{A}_{k+1}^{-1} \boldsymbol{\alpha} = \mathbf{e} \Rightarrow \alpha_{k+1} \leq 0$

推论1 设 $\boldsymbol{\alpha}_0$ 是方程(18)的根, 式(10)中的目标函数 L 可按式(19)求解:

$$L = \frac{1}{\|\boldsymbol{\alpha}_0\|_1} \quad (19)$$

并且 L 满足

$$L = \sum_{i=1}^M \alpha_i k(\mathbf{z}, \mathbf{x}'_i) \quad (20)$$

上述定理和推论给出了快速单分类算法的数学基础。定理1和定理2是实现效率提升的关键, 即可通过核函数矩阵逆的增量式运算增广或缩减核函数矩阵, 因此该算法将时间复杂度由 $O(N^3)$ 减小到 $O(k_s^2)$, k_s 为支持向量数目。并且增量运算通过不断引入新数据更新模型, 非常适合在线检测等新数据不断加入的场景。定理3和定理4给出了边界支持向量的判别方法。推论1表明可利用方程(16)的解计算目标函数 L , 从而计算式(13)所示的异常分数。

1.3 双层决策球

由于发动机系统误差与测量误差带来固有噪声, 传感器测得的数据不可避免地存在正常状态离群点, 如果采用单层 SVDD 超球面进行检测, 这些离群点将被判定为异常点并且不会对单分类器进行更新, 导致较大的虚警率。因此, 引入缩放因子 $\delta \in (0, 1)$ 对异常分数进行松弛, 其意义为对超球面进行放大得到外层超球面, 样本点到外层超球面中心的距离与半径之差的 $1/2$, 表示为:

$$Q_c(\mathbf{z}) = \delta \sum_{i=1}^M \alpha_i K(\mathbf{x}'_k, \mathbf{x}'_i) - \sum_{i=1}^M \alpha_i K(\mathbf{z}, \mathbf{x}'_i) \quad (21)$$

将新增样本点 \mathbf{z} 输入单分类器, 如果 \mathbf{z} 落在内层超球面内部, 则不对单分类器进行更新; 如果 \mathbf{z} 落在内层超球面与外层超球面之间或边界上, 则对单分类器进行更新; 如果 \mathbf{z} 落在外层超球面之外, 则判定为异常点, 如图2所示。

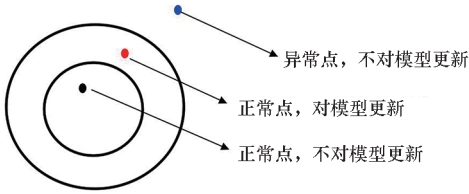


图 2 双层决策球示意图

Fig. 2 Sketch of double decision sphere

1.4 快速增量单分类支持向量机算法流程

快速增量单分类支持向量机算法包括支持向量扩展、缩减、再验证过程,如图 3 所示。

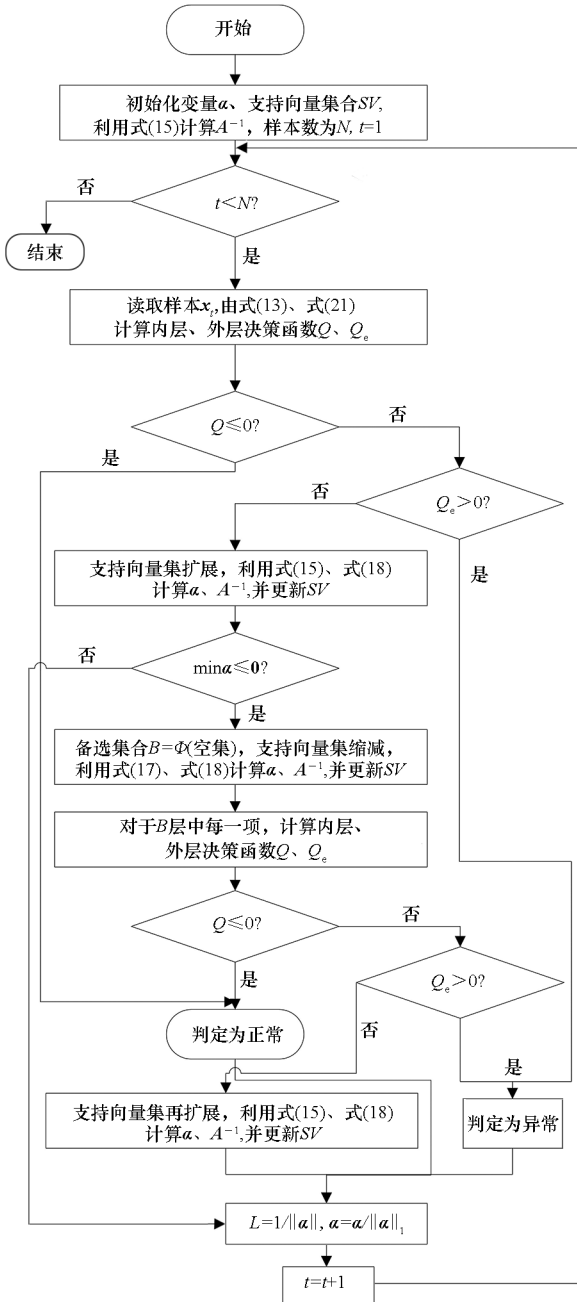


图 3 算法流程图

Fig. 3 Algorithm flow chart

支持向量集扩展: 假设已经有作为支持向量的样本点集合 $\{x_1, x_2, \dots, x_k\}$, 当新的样本点 x_{k+1} 输入模型, 则可由定理 1 计算核函数矩阵的 A_{k+1}^{-1} , 并由定理 3 计算拉格朗日乘子向量 α 。如果 $\alpha_{k+1} \leq 0$, 则由定理 4, x_{k+1} 不能成为支持向量, 支持向量集合保持不变。如果 $\alpha_{k+1} > 0$, 但是至少有一个 $i (1 \leq i \leq k)$, 使得 $\alpha_i \leq 0$, 则 x_1, x_2, \dots, x_k 至少有一个是非支持向量, 进入缩减步骤。

支持向量集缩减: 由式(13)、式(14)可得:

$$\alpha_{k+1} = \frac{d^2 - R^2}{2(1 - v^T A_k^{-1} v)} \quad (22)$$

式中, α_i 被放在 α_{k+1} 的位置。对于不同的非负 α_i , 式(22) 等号右侧分母数值相差不大, 于是绝对值越大的非负 α_i 对应样本点距离超球中心越近, 删除它之后的集合更可能满足支持向量条件。因此, 按照绝对值由大到小的顺序依次删除 $\alpha_1, \alpha_2, \dots, \alpha_k$ 中的非正项, 直到剩下的拉格朗日乘子为正数。

支持向量集再验证: 由于样本点不断加入, 在第 i 时刻被缩减的 α_i 对应的样本点 x_i 可能在以后时刻重新成为支持向量。因此, 引入备选集合存储这些在之前时刻被缩减的样本点, 并在每次缩减步骤之后, 对备选集合中的样本点进行再验证, 再验证过程与扩展过程相同, 从而筛选出可能的支持向量。

初始化采用如下思路: 若待测发动机具备各工况离线数据, 则采集低、中、高工况各 100 个样本点进行离线初始化; 若待测发动机无离线试车数据, 则在发动机进入各工况稳态工作时, 采集前 100 个样本点 (1 s) 的数据进行在线初始化。

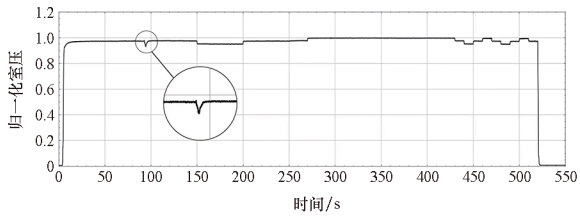
2 试车数据分析

氢氧火箭发动机在试车过程中积累了大量数据, 选取发生典型故障的 4 次试车建立故障数据集进行验证, 如表 1 所示, 由于调节元件更换、外界扰动等因素, 这 4 次试车稳态工作水平略有不同。根据试后分析, 试验 1~3 发生的故障为突变型故障, 试验 4 发生的故障事先有征兆, 存在 10 s 以上的故障发展过程。图 4 给出了 4 次试车归一化推力室压力曲线。

表1 试车数据集

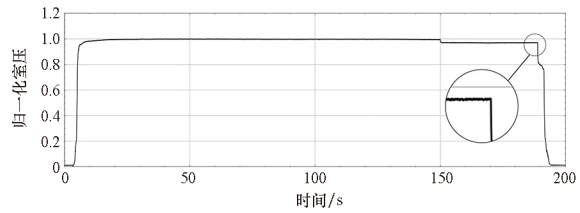
Tab.1 Test dataset

试验序号	现象	故障定位
1	93 s 氢泵和氧泵的转速、流量突然大幅下降,98 s 后基本恢复之前工况水平	涡轮泵密封环故障
2	188 s 参数大幅下降	涡轮泵烧蚀故障
3	302 s 参数大幅下降	轮盘振动破坏
4	273 s 参数缓慢下降,290 s 参数大幅下降	涡轮泵叶片故障



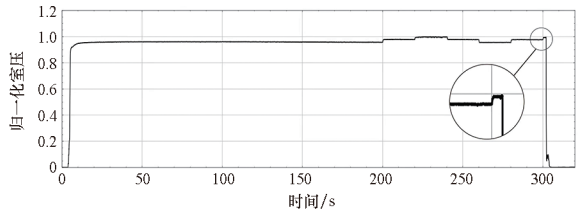
(a) 试验 1

(a) Test 1



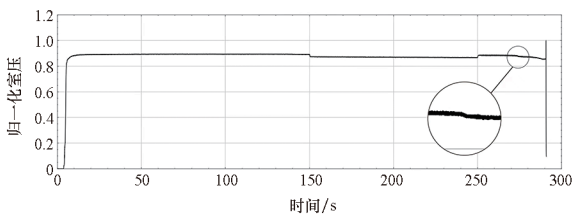
(b) 试验 2

(b) Test 2



(c) 试验 3

(c) Test 3



(d) 试验 4

(d) Test 4

图4 试车室压曲线

Fig.4 Test pressure of combustion chamber

2.1 特征构造

发动机故障检测不仅要依靠传感器直接测得

的瞬时参数,还要挖掘隐藏在参数序列中的故障信息。因此,根据物理意义和统计学意义,在原始静态参数的基础上,构造新特征量。

由于发动机压力传感器和转速传感器测量精度较好,选取 21 路传感器性能参数作为原始诊断数据,主要包括关键测点压力参数与氢泵、氧泵转速。

根据文献[20],发动机流路分段的压差、压比特征能够表征泄漏等缓变故障,因此,根据发动机系统结构,计算发动机 17 个组件分段的进出口压比、压差共 34 个物理特征。由于该型发动机采用副路燃气阀调节方式阶跃式调节混合比,其存在低、中、高三个工况,因此低、中、高工况分别用 -1、0、1 表示,并作为工况特征。

采用滑动窗口将上述特征量构成的时间序列截取为片段,并计算上述特征除工况特征以外的均值、方差等 8 个时域统计特征,见表 2,时域统计特征的表达式见文献[12]。滑窗长度应为特征量计算提供足够的信息,又不会使特征丢失实时性,综合考虑选取滑窗长度为 50(0.5 s)。将上述所有特征量拼接构成长度为 496 的特征向量。

表2 特征提取

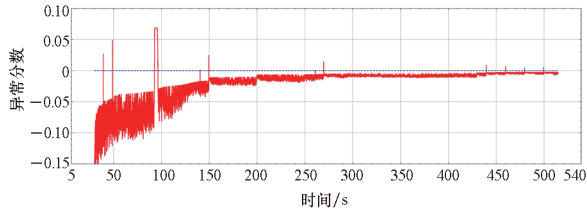
Tab.2 Feature extraction

特征类别	特征量	数量
原始参数	压力、泵转速	21
物理特征	各分段压差、压比	34
工况特征	表征低、中、高工况的数值	1
统计特征	均值、方差、偏度、峰度、峰度因子、波形因子、脉冲因子、裕度因子	440

由于支持向量机是一种距离度量的方法,量纲对分类结果影响较大,因此,采取标准化方法,对上述方法提取的特征向量进行无量纲化处理。

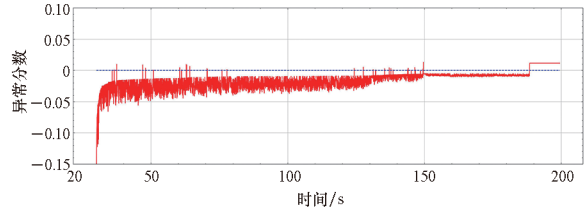
2.2 试车数据异常检测

采用快速增量单分类支持向量机算法对 4 次试车数据进行检测。 σ 控制了决策边界的松紧程度^[9],缩放因子控制了每个时间步用于更新决策边界的样本点的数量,综合考虑虚警与漏检风险,经多次尝试调整,选取式(9)中 $\sigma = 0.3$,式(21)中 $\delta = 0.75$ 。根据图 2 所示流程对试车时序进行检测,得到异常分数随时间变化曲线,如图 5 所示。由式(21)可知,若该曲线超过纵轴即可判定为异常。



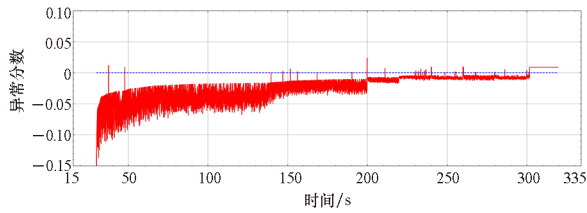
(a) 试验 1

(a) Test 1



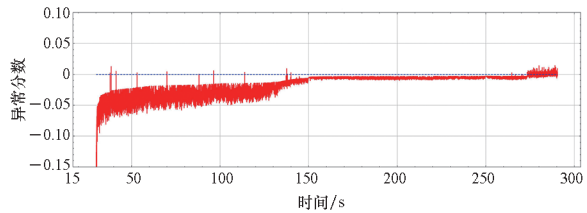
(b) 试验 2

(b) Test 2



(c) 试验 3

(c) Test 3



(d) 试验 4

(d) Test 4

图 5 试车异常分数

Fig.5 Anomaly score of test

2.3 检测结果分析

根据图 5 可知快速增量单分类支持向量机对 4 次试车典型故障可以有效进行检测;对于不同台次调节元件造成的不同稳态工作水平,以及同一台次不同工况,该模型表现出良好的自适应性。异常分数曲线波动幅值随时间减小,这是因为随着自适应学习过程的深入,支持向量集合趋于固定,模型趋于稳定。

为了进一步对检测结果进行量化验证,将试车数据打上标签,对于故障发展阶段发动机状态不明确时刻数据,不纳入评估范围,如图 6 所示。

结合分类结果和样本标签,有 4 种可能分类

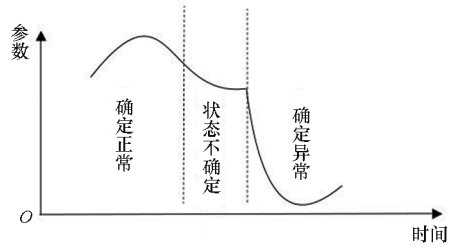


图 6 试车数据打标签

Fig.6 Test data labeling

结果:①真阳性——实际为阳,预测为阳;②假阴性——实际为阳,但预测为阴;③真阴性——实际为阴,预测为阴;④假阳性——实际为阴,但预测为阳。从上述 4 种分类结果出发,可以定义如下评价指标^[21]:

$$\text{准确率} = \frac{TP + TN}{TP + TN + FP + FN} \quad (23)$$

$$\text{精确率} = \frac{TP}{TP + FP} \quad (24)$$

$$\text{召回率} = \frac{TP}{TP + FN} \quad (25)$$

其中: TP 表示真阳性样本数, TN 表示真阴性样本数, FP 表示假阳性样本数, FN 表示假阴性样本数。准确率表示所有预测正确的样本占总体样本的比例,该值越大越好。精确率表示预测阳性的样本中,实际为阳性的比例,该值越大,则虚警率越低,检测效果越好。召回率表示实际为阳性的样本中,预测为阳性的比例,该值越大,则漏警率越低,检测效果越好。以试验 1 为例,其真阳 375 次,假阳 22 次,真阴 48 054 次,假阴 0 次,则试验 1 准确率为 0.999,精确率为 0.945,召回率为 1。表 3 给出了对试验 1~4 检测结果的评估。

表 3 检测结果评估

Tab.3 Detection result evaluation

试验序号	准确率	精确率	召回率	单样本检测用时/ms
1	0.999	0.945	1.000	0.216
2	0.998	0.965	1.000	0.223
3	0.997	0.952	1.000	0.192
4	0.970	0.963	0.615	0.264
综合	0.991	0.958	0.911	0.222

由表 3 可知,本文提出的算法具有较快计算速度,同时保持了较高的准确率。较高的精确率

表明虚警较少,结合图5可知虚警主要发生在变工况阶段,多为孤立离群点,可通过设置多点连续检测策略对虚警进行过滤。对于试验1发生故障又恢复的情况,模型准确识别了故障,恢复之后并未发生虚警。

试验1~3中模型保持了零漏警率,与试验1~3发生突变故障不同,试验4在发生故障破坏前,参数存在缓慢恶化过程,其异常分数曲线逐渐逼近阈值0,并在恶化过程多次突破阈值,导致总体召回率减小。因此,异常分数曲线不仅能对试车数据进行正常/异常二元检测,还可以反映发动机相对正常工作区间的偏离情况,从而反映发动机健康状态。

3 结论

1)快速增量学习支持向量机异常检测算法可以仅利用正常数据对发动机进行异常检测,热试车数据表明,该算法具有99%准确率,虚警、漏警率小,对于突变、缓变故障以及故障后恢复情况检测效果良好。

2)基于增量学习、双层决策球思想实现了模型实时更新,该算法具有对发动机不同台次存在参数偏差、不同工况异常检测问题的良好自适应性。

3)通过增量学习算法提高单分类支持向量机计算速度,将时间复杂度由 $O(N^3)$ 降低到 $O(k_s^2)$,单样本检测平均用时约为0.2 ms,为基于该算法的自适应在线诊断提供了有力支持。

参考文献 (References)

- [1] 吴建军,朱晓彬,程玉强,等.液体火箭发动机智能健康监控技术研究进展[J].推进技术,2022,43(1):7-19. WU J J, ZHU X B, CHENG Y Q, et al. Research progress of intelligent health monitoring technology for liquid-propellant rocket engines[J]. Journal of Propulsion Technology, 2022, 43(1): 7-19. (in Chinese)
- [2] 张振臻,陈晖,高玉闪,等.液体火箭发动机故障诊断技术综述[J].推进技术,2022,43(6):20-38. ZHANG Z Z, CHEN H, GAO Y S, et al. Review on fault diagnosis technology of liquid rocket engine[J]. Journal of Propulsion Technology, 2022, 43(6): 20-38. (in Chinese)
- [3] 朱恒伟,王克昌,陈启智.液体火箭发动机地面试车故障检测的自适应阈值算法[J].推进技术,2000,21(1):1-4. ZHU H W, WANG K C, CHEN Q Z. Adaptive thresholds algorithm for fault detection of liquid rocket engine in ground test[J]. Journal of Propulsion Technology, 2000, 21(1): 1-4. (in Chinese)
- [4] 杨雪,张振鹏,向世勇,等.基于自回归模型的自适应阈值算法[J].航空动力学报,2005,20(6):1088-1092. YANG X, ZHANG Z P, XIANG S Y, et al. Adaptive threshold algorithm based on autoregressive model for fault detection of LRE in ground test[J]. Journal of Aerospace Power, 2005, 20(6): 1088-1092. (in Chinese)
- [5] 聂尧.基于过程神经网络的液体火箭发动机故障预测方法研究[D].长沙:国防科学技术大学,2017. NIE Y. Investigation on fault prediction methods based on process neural network for liquid-propellant rocket engines[D]. Changsha: National University of Defense Technology, 2017. (in Chinese)
- [6] 王思齐.智能异常检测及其应用[D].长沙:国防科技大学,2019. WANG S Q. Intelligent anomaly detection and its applications[D]. Changsha: National University of Defense Technology, 2019. (in Chinese)
- [7] 刘鑫.无监督异常检测方法研究及其应用[D].成都:电子科技大学,2018. LIU X. Research on unsupervised anomaly detection algorithm and application[D]. Chengdu: University of Electronic Science and Technology of China, 2018. (in Chinese)
- [8] SCHÖLKOPF B, WILLIAMSON R, SMOLA A, et al. Support vector method for novelty detection[C]//Proceedings of the 12th International Conference on Neural Information Processing Systems, 1999: 582-588.
- [9] TAX D M J, DUIN R P W. Support vector data description[J]. Machine Learning, 2004, 54: 45-66.
- [10] IRANMEHR A, MASNADI-SHIRAZI H, VASCONCELOS N. Cost-sensitive support vector machines [J]. Neurocomputing, 2019, 343: 50-64.
- [11] GU B, SHENG V S, TAY K Y, et al. Incremental support vector learning for ordinal regression[J]. IEEE Transactions on Neural Networks and Learning Systems, 2015, 26(7): 1403-1416.
- [12] 胡雷.面向飞行器健康管理的新异类检测方法研究[D].长沙:国防科学技术大学,2010. HU L. Research on novelty detection methods oriented to flight vehicle health management [D]. Changsha: National University of Defense Technology, 2010. (in Chinese)
- [13] PLATT J C. Sequential minimal optimization: a fast algorithm for training support vector machines[R/OL]. (1998-04-21)[2021-11-01]. <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/tr-98-14.pdf>.
- [14] 彭小辉.航天器推进系统自主健康管理关键技术研究[D].长沙:国防科技大学,2017. PENG X H. Study on some key techniques of autonomous health management for spacecraft propulsion system [D]. Changsha: National University of Defense Technology, 2017. (in Chinese)
- [15] ZHU X B, CHENG Y Q, WU J J, et al. Steady-state process fault detection for liquid rocket engines based on convolutional auto-encoder and one-class support vector machine[J]. IEEE Access, 2020, 8: 3144-3158.
- [16] GÖRNITZ, LIMA L A, MÜLLER K R, et al. Support vector data descriptions and k -means clustering: one

- class? [J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(9): 3994 – 4006.
- [17] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 403 – 408.
ZHOU Z H. Machine learning [M]. Beijing: Tsinghua University Press, 2016: 403 – 408. (in Chinese)
- [18] 陈宝林. 最优化理论与算法[M]. 2版. 北京: 清华大学出版社, 2005: 281 – 322.
CHEN B L. Optimization theory and algorithm[M]. 2nd ed. Beijing: Tsinghua University Press, 2005: 281 – 322. (in Chinese)
- [19] JIANG H S, WANG H Y, HU W H, et al. Fast incremental SVDD learning algorithm with the Gaussian kernel [C]// Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, 2019: 3991 – 3998.
- [20] 李艳军. 新一代大推力液体火箭发动机故障检测与诊断关键技术研究[D]. 长沙: 国防科学技术大学, 2014.
LI Y J. Study on key techniques of fault detection and diagnosis for new generation large-scale liquid-propellant rocket engines[D]. Changsha: National University of Defense Technology, 2014. (in Chinese)
- [21] DE HOND A A H, STEYERBERG E W, VAN CALSTER B. Interpreting area under the receiver operating characteristic curve[J]. Lancet Digital Health, 2022, 4(12): e853 – e855.

(编辑: 梁慧, 罗茹馨)