

介绍了支持向量机回归算法,对某型飞机主油泵故障数据进行了拟合训练和外推预测分析,为飞机部件故障数预测、可靠性评估和备件需求预测提供了借鉴和依据。

基于支持向量机回归算法的飞机主油泵故障数预测

Failure Number Prediction of Aircraft Master Oily Pump Based on SVM Regression Algorithms

◎祝华远 赵经成 徐伟勤 / 海军航空工程学院青岛分院

对于飞机部件的故障数预测、可靠性评估、备件需求预测等通常采用样本数目趋于无穷大时的渐近理论,即传统统计学及神经网络理论等进行分析,但在航空维修保障领域,维修数据通常为小样本数据,这些方法存在着固有的算法缺陷。

统计学习理论是一种专门研究小样本情况下机器学习规律的理论,它不仅考虑了对渐进性能的要求,而且追求在现有有限信息的条件下得到最优结果。支持向量机(SVM)是基于统计学习理论的机器学习算法,它采用结构风险最小化原则,在最小化样本点误差的同时,缩小模型预测误差的上界,从而提高了模型的泛化能力和抗噪声扰动能力。主要应用领域为模式识别、回归预测、概率密度函数估计等。目前,支持向量机算法及其应用尚在发展阶段。

支持向量机回归算法

1. 线性回归算法

给定一个线性可分样本集

$$S = ((X_1, Y_1), \Lambda, (X_n, Y_n))$$

式中: $X_i \in R^d$ 为输入值, $y_i \in R$ 为对应的目标值。

线性回归问题即求回归函数

$$f(X) = (W \cdot X) + b$$

式中: $W \in R^n, b \in R$, 且满足结构风险最小化原理。

根据统计学习理论,回归函数的估计可转换成对优化目标函数求最小值

$$Q(W) = \frac{1}{2} \|W\|^2 + CR_{emp}(f)$$

式中: C 为惩罚因子,实现经验风险与置信范围的折中, $R_{emp}(f)$ 为损失函数。

常用的损失函数有拉普拉斯(Laplace)函数、虎泊(Huber)函数和线性 ε 不敏感损失函数,其中线性 ε 不敏感损失函数因具有较好的性质而得到广泛应用。当引入 ε 线性不敏感损失函数时,回归函数的估计可进一步转换成如下的最优化问题:

$$\min_{w, b, \zeta} \frac{1}{2} \|W\|^2 + C \sum_{i=1}^n (\zeta_i + \zeta_i^*)$$

式中: $\zeta_i \geq 0, \zeta_i^* \geq 0$ 为约束松弛因子。

约束条件为

$$y_i + W \cdot X_i - b \leq \varepsilon + \zeta_i$$

$$W \cdot X_i - y_i + b \leq \varepsilon + \zeta_i^*$$

上述最优化问题中,通过约束条件及对不满足约束条件的惩罚得到的估计函数在训练样本上尽可能满足误差要求,同时最小化与置信范围值相关的 $\|W\|^2/2$ 。

对于该最优化问题,一般采用拉格朗日乘子法转换成对偶最优化问题,然后进行最优化计算,得到回归函数

$$f(X, \alpha_i, \alpha_i^*) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) (X_i \cdot X) + b$$

式中: b 可由约束条件按等号求出, α_i, α_i^* 为拉格朗日乘子。

2. 非线性回归算法

对于非线性回归问题,根据 V. Vapnik 提出的核函数理论,可用核函数 $K(X_i, X_j)$ 代替内积运算,实现由低维空间到高维空间的映射,从而使低维输入空间的非线性回归问题转化为高维特征空间的线性回归问题。引入核

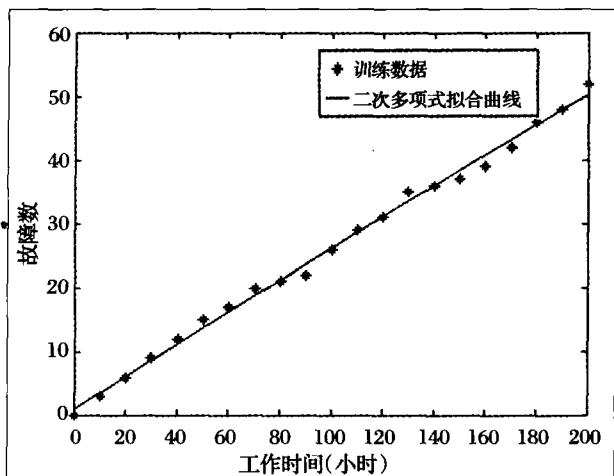


图1 最小二乘二次多项式拟合曲线

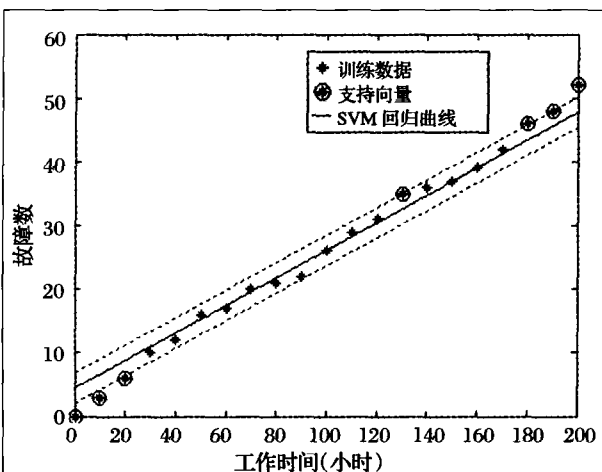


图2 支持向量机回归拟合曲线

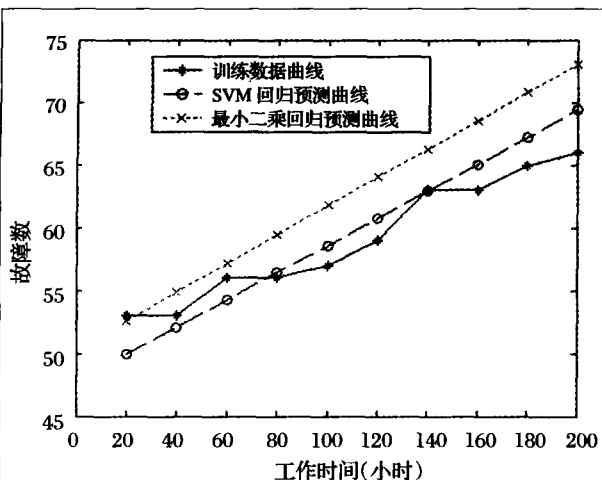


图3 预测结果与实际值比较

函数后,线性回归函数式中的内积($X_i \cdot X$)可用核函数 $K(X_i \cdot X)$ 来代替。

核函数的种类较多,常用的核函数有线性核函数、多项式核函数、径向基核函数、Sigmoid 核函数等。

故障数预测分析

据统计,当某型飞机主油泵工作时间为 300 小时时,共有 66 台发生故障。

设 $t_i=10i(i=1, \Lambda, 30)$, 将各时间区间 $(0, t_i)$ 内主油泵的故障数分成两组,取前 20 个数据作为训练数据,后 10 个数据作为测试数据,分别采用最小二乘、支持向量机回归模型对训练数据进行拟合。

1. 最小二乘拟合

采用二次多项式对训练数据进行最小二乘拟合,拟合后的二次多项式为 $y=1.0265+0.2573x-5.5277 \times 10^{-6}x^2$, 拟合曲线如图 1 所示。

2. 支持向量回归拟合

采用支持向量机解决非线性回归问题,首先必须确定 3 个自由参数: ε 不敏感值、惩罚因子 C 以及核参数。这三个参数在很大程度上决定了支持向量机的学习能力。为此,在支持向量机中通过输出最小学习误差试验,优化参数设置为: ε 不敏感值取 2.4、惩罚因子 C 取 10000,核参数采用径向基核函数 $K(X_i \cdot X) = \exp(-\|X - X_i\|^2 / 2\sigma^2)$, 宽度 σ 取 499,拟合曲线如图 2 所示。

3. 故障数预测

采用由训练数据得到的最小二乘拟合二次多项式和支持向量机回归预测模型,分别对各时间区间 $(0, t_i)$, $t_i=10i(i=21, \Lambda, 30)$ 内的主油泵故障数进行预测,将测试数据作为校验样本,最小二乘与支持向量机回归预测的预测误差平均值分别为 3.80、1.41,最大预测误差分别为 7.1、3.2,预测结果如图 3 所示。

从图 3 可以看出,与最小二乘回归预测方法相比,采用支持向量机回归算法对主油泵故障数预测,可有效提高预测精度,具有较好的预测效果。

结论

将支持向量机回归算法应用到航空维修保障领域,对某型飞机主油泵故障数进行了拟合训练和外推预测分析,通过与最小二乘回归预测方法比较,结果表明:支持向量机回归算法对于飞机部件故障数预测及备件需求预测等具有较好的借鉴和应用价值。 □