

# 基于支持向量机的船舶柴油机故障诊断的研究

詹玉龙, 翟海龙, 曾广芳  
(上海海事大学, 上海 200135)

**摘 要:** 通过建立 VC 维统计学习理论, 利用数学建模, 建立并划分最优超平面以取得支持向量来训练, 选取并考虑相关的影响因子以此构建样本数据集, 以期对柴油机故障实现精确的诊断; 而对于非线性空间情况, 可采用核函数的思想来转化为线性空间, 以此降低算法的复杂度; 通过与人工神经网络方法的比较, 表明该方法具有运算速度快, 泛化能力强等优点; 支持向量机(SVM)可以较好解决柴油机故障诊断中的机器过学习、小样本、高维数、非线性等问题。

**关键词:** 船舶、舰船工程; 柴油机; 故障诊断; 机器学习; 支持向量机; 核函数

**中图分类号:** U664.121 **文献标识码:** A

## Feasible Research on Fault Diagnosis of Marine Diesel Engine Based on Support Vector Machine

ZHAN Yu-long, ZHAI Hai-long, ZENG Guang-fang  
(Shanghai Maritime University, Shanghai 200135, China)

**Abstract:** A new machine learning method called SVM which will be applied in the fault diagnosis for marine diesel and its feasibility are introduced in the paper. By establishing VC dimension statistical learning theory and mathematical modeling, an optimized hyperplane can be set up so as to obtain support vector to train sample data, related influential factors can be chosen and considered to construct sample data table for realizing precise fault diagnosis. As for non-linear space, it is effective to be converted to a linear space by adopting kernel function for minimizing complexity of the algorithm. By comparison to the ANN method, SVM has the advantages of strong generalization and fast operation. The SVM can be the good solution for over-learning, small sample, high dimension and non-linear problem etc.

**Key words:** Ship; Naval engineering; Diesel engine; Fault diagnosis; Machine learning; Support vector machine; Kernel function

### 1 最优划分超平面、支持向量和线性支持向量机

支持向量机是从线性可分情况下的最优分类面发展而来的, 也是统计学习理论中最实用的部分。SVM 方法的基本思想是: 定义最优线性超平面, 并把寻找最优线性超平面的算法归结为求解一个凸规划问题, 进而基于 Mercer 核展开定理, 通过非线性映射  $\varphi$ , 把样本空间映射到一个高维乃至无穷维的特征空间(Hibert 空间), 实在特征空间中可以应用线性学习机的方法解决样本空间中的高度非线性分类和回归等问题。设

$$(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l) \quad (1)$$

式(1)中:  $x_i \in R^N$ , 为  $N$  维向量,  $y_i \in \{-1, 1\}$  或者  $y_i \in \{1, 2, \dots, k\}$  或者  $y_i \in R$ ; 通过训练学习寻求模式  $M(x)$ , 使其不但满足训练样本集  $y_i = M(x_i)$ , 而且对于未来的输入数据集:

$$x_{l+1}, x_{l+2}, \dots, x_m \quad (2)$$

同样也能得到满意的对应输出值  $y_i$ 。

当  $y_i \in \{-1, 1\}$  时为最简单的二类划分; 当  $y_i \in \{1, 2, \dots, k\}$  时为  $k$  类划分; 当  $y_i \in R$  时为函数估计, 即回归分析。

当分类机  $M(x)$  为线性函数(直线或线性超平面)时对应线性划分; 否则为非线性分类。

对于式(1)中的训练样本集的线性二类划分问

收稿日期: 2007-03-16

基金项目: 上海市教育委员会基金项目(06EZ032)

作者简介: 詹玉龙(1948-), 男, 江西婺源人, 教授, 轮机长, 主要从事现代轮机工程研究。E-mail: ylzhan@shmtu.edu.cn

题,其实质就是寻求函数

$$y = f(x) = \text{sgn}((w \circ x) + b) \quad (3)$$

使得对于  $i=1, 2, \dots, l$  满足条件:

$$y_i = f(x_i) = \text{sgn}((w \circ x_i) + b) \quad (4)$$

式(4)中,  $w, x, x_i \in R^N, b \in R, w, b$  为有待确定的参数,  $\text{sgn}$  为符号函数,  $(w \circ x)$  为两个向量的点积。显然, 把  $(w \circ x) + b = 0$  作为划分超平面,  $w$  为其法方向向量。

对于线性可分离问题, 满足上述条件如式(3)的线性决策函数不是唯一的。判断优劣要有原则。通常采用误差最小化原则, 即寻求决策函数使对训练样本的分类误差“总和”(有多种汇总方法)最小。

V. N. Vapnik 提出一个边际最大化 (maximal-margin) 原则。所谓边际, 又称为间隔, 是指训练样本集到划分超平面的距离, 它是所有训练样本点到划分超平面的(垂直)距离中的最小者:

$$\min(\|x - x_i\| : x \in R^N,$$

$$(w \circ x) + b = 0, i = 1, \dots, l)$$

所谓边际最大化原则(间隔最大化原则)是指寻求使间隔达到最大的优化为最优, 即是对  $w, b$  寻优, 求得最大间隔。

$$\max_{w, b}(\min(\|x - x_i\| : x \in R^N,$$

$$(w \circ x) + b = 0, i = 1, \dots, l))$$

对应最大间隔的划分超平面称为最优划分超平面, 简称最优超平面。可以证明最大间隔是唯一的, 但达到最大间隔的最优超平面可能不唯一。

最大间隔和最优超平面只由落在边界上的样本点完全确定, 我们称这样的样本点为支持向量。只由少数训练样本点(支持向量)就把最大间隔和最优超平面完全确定, 其余非支持向量的样本点均不起作用, 这具有十分重要的意义。它说明间隔最大化原则下的最优划分不是依赖于所有点, 而只是由支持向量所决定。求最优超平面和最大间隔等同于确定各个样本点是否为支持向量, 有支持向量确定的线性分类机成为线性支持向量机。

对于式(1)给定的训练样本集, 如果样本是线性可分离的, 可建立如下的线性支持向量机: 设图1中的划分超平面  $L$  的方程为  $(w \circ x) + b = 0$ , 两条边界  $l_1, l_2$  的方程(经过恒等变形之后)变为:  $(w \circ x) + b = \pm 1$ , 设在直线  $l_1, l_2$  上分别有任意两点  $x_1, x_2$ , 则有:

$$(w \circ x_1) + b = -1, (w \circ x_2) + b = +1$$

两式相减有:

$$(w \circ (x_2 - x_1)) + b = 2$$

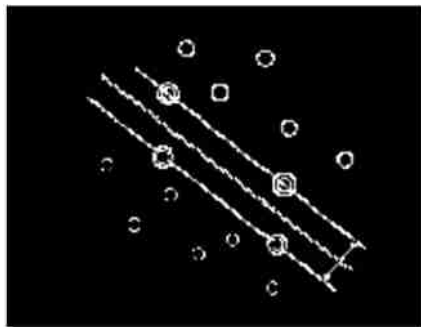


图1 最优划分超平面示意图

Fig. 1 The optimized divided hyperplane chart

进而

$$\left( \frac{w}{\|w\|} \circ (x_2 - x_1) \right) = \frac{2}{\|w\|} \quad (5)$$

式(5)左边恰好就是连接  $x_1, x_2$  的向量在划分超平面法方向上的投影, 它是最大间隔的两倍。这样, 求最大间隔就等价于求  $\|w\|$  或  $\|w\|^2$  或  $\frac{1}{2} \|w\|^2$  的最小值。

考虑到要使所有训练样本以最大间隔分类正确, 应成立:

$$y_i((w \circ x_i) + b) \geq 1$$

因此, 建立线性支持向量机的问题转化为求解一个二次凸规划问题:

$$\begin{cases} \min \frac{1}{2} \|w\|^2 \\ y_i((w \circ x_i) + b) \geq 1 \end{cases} \quad (6)$$

由于目标函数和约束条件都是凸的, 根据最优化理论, 这一问题存在唯一全局最小解。为了求解式(6), 引入拉格朗日函数:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i (y_i((w \circ x_i) + b) - 1)$$

这里的  $\alpha_i \geq 0$  是拉格朗日乘子。应用拉格朗日乘子法求极值。对相应的  $w, b$  求偏导, 可以找到相应的对偶形式:

$$\begin{cases} \frac{\partial L(w, b, \alpha)}{\partial w} = w - \sum_{i=1}^l y_i \alpha_i x_i = 0 \\ \frac{\partial L(w, b, \alpha)}{\partial b} = \sum_{i=1}^l y_i \alpha_i = 0 \end{cases}$$

将得到的关系式:

$$w = \sum_{i=1}^l y_i \alpha_i x_i$$

$\sum_{i=1}^l y_i \alpha_i = 0$  代入到原始的拉格朗日函数

$$\begin{aligned}
L(w, b, \alpha) &= \frac{1}{2} \|w\|^2 - \\
&\sum_{i=1}^l \alpha_i (y_i ((w \circ x_i) + b) - 1) = \\
&\frac{1}{2} \sum_{i,j=1}^l y_i y_j \alpha_i \alpha_j (x_i \circ x_j) - \\
&\sum_{i,j=1}^l y_i y_j \alpha_i \alpha_j (x_i \circ x_j) + \sum_{i,j=1}^l \alpha_i = \\
&\sum_{i,j=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l y_i y_j \alpha_i \alpha_j (x_i \circ x_j) \quad (7)
\end{aligned}$$

在这里考虑 Karush-Kuhn-Tucker 关系

$$\alpha_i (y_i (x \circ x_i) + b - 1) = 0 \quad (8)$$

从而求得最优划分超平面的决策函数

$$\begin{aligned}
M(x) &= \text{sgn}((w^* \circ x) + b^*) = \\
&\text{sgn}(\alpha_i^* y_i (x \circ x_i) + b^*) \quad (9)
\end{aligned}$$

式(9)中,  $\alpha_i^*$ ,  $b^*$  为确定的最优划分超平面的参数, 并且根据式(8)知道: 非支持向量对应的  $\alpha_i$  都为零, 求和只对少数支持向量进行, 对于线性不可分的情况, 通过引入松弛变量  $\xi_i \geq 0$ , 建立新的凸规划问题为:

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + C \sum_i \xi_i \\ y_i ((w \circ x_i) + b) \geq 1 - \xi_i \end{cases} \quad (10)$$

应用完全类似的方法可以求解。若上式中的  $\xi_i$  都为零, 则上述问题就转化成了线性可分问题。式(10)中的大于零的  $\xi_i$  对应错分的样本, 参数  $C$  为惩罚系数。

对于  $k$  类线性分类问题可以建立类似的凸规划问题。也可以把它转化为  $k$  个二类划分(属于  $i$  类或者属于其它类)问题来解决。

## 2 SVM 在船用柴油机故障诊断中的应用研究

### 2.1 柴油机故障诊断的应用研究

对于机器某一种故障, 选定其特征向量作为训练数据。这样若干组训练数据在  $n$  维空间构成了一组区域。不同的故障类型对应不同的区域。所谓故障诊断就变成在  $n$  维空间寻找这些区域的分界面。而分界面的确定和表达必须能够通过训练数据的训练来完成。故障诊断的准确性本质上就变成区域分界的准确性。船舶柴油机的故障诊断也可以依据 SVM 方法, 进行一级故障、二级故障或三级故障的参数分类训练。

#### 2.1.1 特征向量的选取

所选的特征向量应最显著地表征对应故障, 而

且包含与其它故障特征差异最明显的信息。否则, 在特征向量空间中出现故障区域交叉或重迭。在此情况之下, 不论采取什么分类方法。都很难准确地确定分界超平面, 甚至找不到分界超平面。因此, 诊断的准确性都将不会太高。

#### 2.1.2 分类方法的选择

分类算法通过样本的训练能够确定分界面的参数和表达方式, 且训练过程具有单调逼近的特点。但是现有的分类算法均存在不足; 例如, 针对旋转机械故障诊断。神经网络的最优结构和结构参数不能通过样本训练来确定, 而完全依赖于个人经验。

#### 2.1.3 抗噪能力

故障诊断时, 输入的数据同样有噪声, 而噪声的影响可能会使分类出错。因此, 抗噪能力是每个分类算法都要考虑的问题。神经网络是靠大样本训练来抗噪的。而支持向量机分界面主要依靠不同区域最靠边界的训练数据(即所谓的支持向量)来确定, 从而具有良好的抗噪能力。

### 2.2 影响因子的确定

一般来说, 在柴油机的故障诊断过程中, 主要需考虑的因子有: 柴油机本身及各部件结构的抗疲劳强度; 燃油系统、滑油系统、起动系统、操纵系统、调速系统、增压系统和冷却系统等的工作状况和故障参数; 柴油机的部分典型故障如排气阀断裂, 缸套磨损, 曲轴红套滑移, 敲缸、拉缸和咬缸, 连杆螺栓断裂, 调速器失灵, 轴系振动等机械和磨损故障。除此以外, 人为因素也是一个重要的影响因子。如轮机管理人员的操作技术水平、维护管理质量以及敬业精神工作态度等。

#### 2.3 样本数据集的构建

根据影响因子的确定, 综合考虑各影响因素的滞后作用, 则引起柴油机故障的问题就转化为 4 个输入和 1 个输出的支持向量机的函数回归问题。

为了消除各个因子由于量纲和单位不同的影响, 对样本的输入、输出参数分别用式(1)进行规格化处理。

$$y_i \frac{2(z_i - z_{\min})}{z_{\max} - z_{\min}} - 1 \quad (11)$$

式(11)中:  $Z_i$  和  $y_i$  分别为规格化前后的变量;  $z_{\min}$  和  $z_{\max}$  分别为  $Z$  的最小和最大值。

用从船厂得来的相应数据作为检验样本集, 用 SVM 进行学习训练和验证。

#### 2.4 SVM 的学习训练

首先进行核函数的选取, 通过对多项式核函数、径向基函数(RBF)核函数、Sigmoid 核函数的比较分

析,发现  $d=1$  的多项式核函数比较适合某一种故障的预测情况。然后通过不同的  $\epsilon$ 、 $C$  参数的试验,发现  $\epsilon=0.4$ ,  $C=\infty$ , 经过 3.6 s 学习训练,得到 6 个支持向量,各个支持向量的  $\alpha_i-\alpha_i^*$  及其对应的样本序号见表 1,相应的  $w_0$  和  $b$  分别为 2 303.226 275 和 47.990 354。

表 1 对训练样本学习后得到的支持向量及  $\alpha_i-\alpha_i^*$   
Table 1 the value of support vector and the  $\alpha_i-\alpha_i^*$  after the sample learning

支持向量	样本序号	$\alpha_i-\alpha_i^*$	支持向量	样本序号	$\alpha_i-\alpha_i^*$
1	2	3.097 3	4	19	-17.086 1
2	7	28.708 6	5	31	-2.003 5
3	8	17.581 7	6	34	-2.003 5

2.5 结果分析

用得到的  $w_0$  和  $b$  分别对训练和检验资料进行拟合与预测。为了与传统的回归分析和人工神经网络方法进行对比,把 3 种预测的结果一并列入表 2。

表 2 SVM 与人工神经网络模拟与预测结果比较分析  
Table 2 the compare and analysis of the SVM VS. ANN

误差		SVM 方法		人工神经网络方法		回归分析	
		拟合数据	预测数据	拟合数据	预测数据	拟合数据	预测数据
绝对误差	max	0.840	0.233	0.754	0.504	0.911	0.838
	min	0.001	0.003	0.016	0.012	0.004	0.014
	ave	0.272	0.109	0.177	0.220	0.191	0.275
相对误差	max	1.78	0.49	1.6	0.8	1.93	1.75
	min	0.000	0.01	0	0.1	0.01	0.03
	ave	0.57	0.23	0.41	0.4	0.40	0.57

将 SVM 与人工神经网络模拟与预测结果分析比较。人工神经网络在训练过程中容易出现过学习问题,即把各个学习样本的细节,有时甚至错误利用训练的权值保存了下来,难以区分噪声和真实信号。SVM 方法则通过常数  $C$  的调整,一方面使得误差尽可能小,另一方面使回归函数尽可能平滑,从而提高了它的泛化(预测)能力。同时人工神经网络结构是通过大量试验后才得出来的“最优结构”,在隐含层为 8 的最优结构下,仍进行了 24 600 次训练,用时 48 min 23 s 得出这样的预测结果,而 SVM 方法仅需要小样本即可进行计算,运算时间为 3.6s 便取得了可接受的预测值。因此无论从运算速度,还是泛化能力,SVM 方法都表现出了良好的性能。

3 结 语

本文把支持向量机方法引入到柴油机故障诊断研究中来,支持向量机方法能充分利用训练样本的分布特性,根据部分训练样本构建判别函数,不需要过多的先验信息和使用技巧。人工神经网络方法往往需要较大的样本计算、过多的使用技巧才能得出比较满意的结果。SVM 算法最终转化为二次寻优问题,从理论上说,得到的将是全局最优解,有效避免了神经网络易陷入的局部极值问题。同时通过非线性变换和核函数巧妙解决了高维数问题,使得其算法复杂度与样本维数无关,加速了训练学习速度。

参 考 文 献

[ 1 ] 陈永义. 支持向量机方法与模糊系统[ J ]. 模糊系统与数学, 2005, 19( 1 ): 55-59.  
Chen Yongyi. SVM method and blur system[ J ]. Blur system and math, 2005, 19( 1 ): 55-59.

[ 2 ] Qu Liansheng, Liu Xiong, Chen Yaodong. Discovering the Holospectrum Noise and Vibration Control World Wide[ J ]. Safety Science, 1989, 20( 2 ): 58-63.

[ 3 ] Vapnik V, Chapelle O. Bounds on error expectation for support vector machine[ J ]. Neural Computation, 2000, 12 ( 9 ): 78-80.

[ 4 ] 邓乃扬, 田英杰. 数据挖掘中的新方法——支持向量机[ M ]. 北京: 科学出版社, 2004.  
Deng Naiyang, Tian Yingjie. New method for data dig-SVM [ J ]. Beijing: Science Publishing House, 2004, 28( 10 ): 56-61.

[ 5 ] Szladow A, W Ziarko. Rough sets: working with imperfect data[ J ]. AI Expert, 2002, 7( 3 ): 36-41.

[ 6 ] 冯汉中, 陈永义. 处理非线性分类和回归问题的一种新方法( II )——支持向量机在天气预报中的应用[ J ]. 应用气象学报, 2004, 45( 17 ): 13-17.  
Feng Hanzhong, Chen Yongyi. An new method to deal with non-linear classification and regressive problem-SVM in the application of weather forecast[ J ]. Application Weather Transaction, 2004, 45( 17 ): 13-17.

[ 7 ] Vapnik V. Statistical Learning Theory [ J ]. Neural Computation, 2000, 12( 9 ): 19-30.

[ 8 ] Nello Cristianini & John Shawe-Taylor. An introduction to support vector machines and other kernel-based learning methods[ M ]. Canada: Zilog, Inc. 2004.

[ 9 ] 曹龙汉. 柴油机智能化故障诊断技术[ M ]. 北京: 国防工业出版社, 2005.  
Cao Longhan. Diesel intellectualized fault diagnosis technology[ M ]. Beijing: National Defence Industry Publishing House, 2005.