

## 【兵器与装备】

## 基于支持向量机的柴油机燃油系统故障诊断\*

贾志杰, 李国璋, 罗 亮

(军械工程学院, 石家庄 050003)

**摘要:**介绍了支持向量机(SVM)的机理,应用SVM对柴油机燃油系统进行故障诊断,通过试验确定了SVM参数的选择方法.实践诊断结果表明,SVM具有较好的诊断效果,对故障样本诊断的准确度较RBF神经网络高.

**关键词:**柴油机;故障诊断;支持向量机

**中图分类号:** U664

**文献标识码:** A

**文章编号:** 1006-0707(2008)04-0012-03

柴油发动机燃油系统是柴油机的重要系统之一,其故障规律很难用简单的数学模型描述.尽管神经网络在智能故障诊断系统中获得了广泛的应用,但神经网络普遍存在着在线学习速度慢、网络结构难以确定、局部最优、泛化能力差等缺点.最近几年发展起来的支持向量机(Support Vector Machine, SVM)建立在统计学习理论的VC维理论和结构风险最小原理上,其核心思想是使样本之间的分类间隔最大化.它能较好地解决小样本、非线性、高维数和局部极小点等实际问题,比传统神经网络具有更强的分类能力和泛化能力,学习算法效率较高,并且巧妙地解决了维数灾难问题<sup>[1]</sup>.因此SVM目前已成为机器学习界研究的热点,并成功地应用于分类、函数逼近和故障诊断等方面<sup>[2-4]</sup>.本研究尝试在小样本条件下采用SVM对柴油机燃油系统进行故障诊断.

## 1 基于支持向量机的故障诊断机理

设有  $n$  个训练样本,将支持向量机用于故障诊断,就是在样本特征空间内寻找中心为  $a \in \mathbf{R}^1$ ,半径为  $R \in \mathbf{R}^d$  的超球,使正常样本点位于超球内,异常点位于超球之外,从而实现分类目的.超球体积要尽量小,以尽量减少接受异常样本点的机会.因此,故障诊断问题可用下述优化问题来描述,并求解最小的超球半径  $R$ :

$$\min(R^2 + C \sum_i \xi_i) \quad (1)$$

$$\text{s.t. } x_i - a^2 \leq R^2 + \xi_i \quad (2)$$

$$\xi_i \geq 0, i = 1, 2, \dots, n$$

式中:  $x_i$  为第  $i$  个样本点;  $\xi_i$  为松弛变量,是考虑到有一些样本不能被超平面正确分类专门设置的;  $C$  为权值系数,它的作用是控制对错分样本(个别样本点出现在超球以

外)惩罚的程度,  $C$  越大,表示要求实现较小的分类间隔,  $C = \infty$  时,表示要求无错分样本.

可将上述优化问题转化为一个二次规划问题,它有唯一的极小点,可用Lagrange乘子法解决.定义Lagrange函数  $L(R, a, \xi, \alpha, \gamma)$  为:

$$L(R, a, \xi, \alpha, \gamma) = R^2 + C \sum_i \xi_i - \sum_i \alpha_i (R^2 + \xi_i - \|x_i - a\|^2) - \sum_i \gamma_i \xi_i \quad (3)$$

式中:  $\alpha_i \geq 0, \gamma_i \geq 0$  为Lagrange乘子.Lagrange函数对  $R, a, \xi_i$  分别求偏导并令求导结果等于0,得:

$$\sum_i \alpha_i = 1 \quad (4)$$

$$a = \sum_i \alpha_i x_i \quad (5)$$

$$\gamma_i = C - \alpha_i \quad (6)$$

将以上3个公式代入Lagrange函数(3),得到对偶优化问题的目标函数为:

$$W(\alpha) = \sum_i \alpha_i K(x_i, x_i) - \sum_{i,j} \alpha_i \alpha_j (x_i, x_j) \quad (7)$$

式中:  $(x_i, x_j)$  表示欧式空间中的内积;  $(x_i, x_j)$  为核函数.这样,原问题的对偶问题转化为:

$$\max W(\alpha) \quad (8)$$

$$\text{s.t. } \sum_i \alpha_i = 1, 0 \leq \alpha_i \leq C \quad (9)$$

其KKT条件为:

$$\alpha_i (R^2 + \xi_i - \|x_i - a\|^2) = 0 \quad (10)$$

$$\gamma_i \xi_i = 0$$

采用顺序最小优化方法对上式求解,可得到  $\alpha_i$ .由式(5)知,球心  $a$  为所有样本点  $x_i$  的加权平均,加权系数就是  $\alpha_i$ .注意到只有少部分  $\alpha_i$  的数值大于零,其他均为零,称对应  $\alpha_i > 0$  的样本点为支持向量.

支持向量可分为2类<sup>[3]</sup>.若  $C = \alpha_i$ ,可得:  $\xi_i > 0, R^2 +$

\* 收稿日期:2008-04-21

作者简介:贾志杰(1978—),男,天津武清人,硕士研究生,主要从事车辆试验分析与故障诊断.

$\xi_i = \|x_i - a\|^2$ , 从而  $\|x_i - a\|^2 \geq R^2$ , 此时对应的样本数据位于超球面上或超球外部; 若  $C > \alpha_i$ , 则  $\|x_i - a\|^2 = R^2$ , 说明对应的样本数据位于超球面上. 通过调节  $C$  的大小可调整位于超球面外部的样本点数,  $C$  越大, 位于超球面外部的样本数据就越少.

在通过优化计算得到超球半径  $R$  后, 计算样本数据到球心的距离  $\|x_i - a\|$ , 若该距离不超过  $R$ , 即样本位于超球内, 则说明该样本为正常值, 否则为故障点.

## 2 柴油机燃油系统故障诊断

燃油系统是柴油机的一个重要组成部分, 该系统的工作状态直接影响柴油机的燃烧过程, 决定柴油机的性能, 对柴油机燃油系统及时进行检测和故障诊断对保持柴油机良好的工作性能有重要意义. 目前, 检测燃油系统的主要方法是用串接在高压油管上的压力传感器测量油压波形. 这种方法要求改装高压油路, 很不方便. 笔者采用夹在高压油管上的夹持式传感器, 通过测量高压油管的膨胀变形量来反映油压波形, 对油压波形进行分析处理, 提取有效的特征向量, 应用支持向量机技术, 从而诊断出燃油系统的故障.

### 2.1 基本原理

当高压油泵柱塞开始关闭进油孔时, 高压泵内的燃油被压缩, 当压力升高超过高压油管中的剩余压力时, 燃油进入高压油管. 设高压油管与喷油器构成的燃油容积为  $V_c$ , 高压油泵将体积为  $\Delta V_c$  的燃油压入高压油管, 则燃油被压缩时能量为

$$\epsilon = E_p \cdot \Delta V_c = P_E \cdot V_c \quad (11)$$

式中:  $\epsilon$  为燃油喷射能量;  $E_p$  为燃油弹性系数;  $P_E$  为燃油的压力. 式(11)两边同时除以  $V_c$  得:

$$P_E = \epsilon / V_c = E_p \cdot \Delta V_c / V_c \quad (12)$$

式(12)表明, 压力受燃油系统容积  $V_c$  和燃油增加量  $\Delta V_c$  的影响, 且是油管内单位体积燃油喷射能量的一种度量, 它反映燃油喷射能量的水平. 当燃油系统出现漏油、喷油嘴积炭和喷油开启压力降低等故障时, 都能在油压波形中反映出来, 因此检测油压波形是一种有效的燃油系统故障诊断方法. 油压波形可以直接测量, 也可以通过检测高压油管的膨胀变形量间接得到. 喷油过程中, 在高压油管内燃油压力的作用下, 油管外径膨胀量为:

$$\mu = P_E R_0^2 [(1-\nu)R + (1+\nu)R_0^2/R] / [E(R_0^2 - R^2)] \quad (13)$$

式中:  $R_i$ ,  $R_o$  和  $R$  分别为油管的内径、外径和中径;  $\nu$  为油管材料的泊松系数;  $E$  为油管的弹性模量. 由式(13)可以看出, 油管外径的膨胀量与油管内部的燃油压力成比例, 通过检测油管的膨胀量来诊断燃油系统的故障在理论上是可行的.

笔者在 12150L 型柴油机上 (转速为  $n = 500$  r/min, 采样频率为 50 kHz), 进行了 4 种典型故障的模拟故障诊断试验, 测量的油压波形如图 1 所示, 记为  $F = \{F_0, F_1, F_2, F_3, F_4\}$ , 其中:  $F_0$  为无故障;  $F_1$  为柱塞磨损;  $F_2$  为高压油管渗

漏;  $F_3$  为针阀弹簧弹力下降;  $F_4$  为开启压力降低.

故障样本集的确定是 SVM 进行准确诊断的关键. 根据柴油机运行中的实际状况, 针对上述 5 种模式, 每个模式采集了 40 个样本, 并将每个模式的前 20 个数据样本用于训练, 后 20 个数据样本作为检验样本.

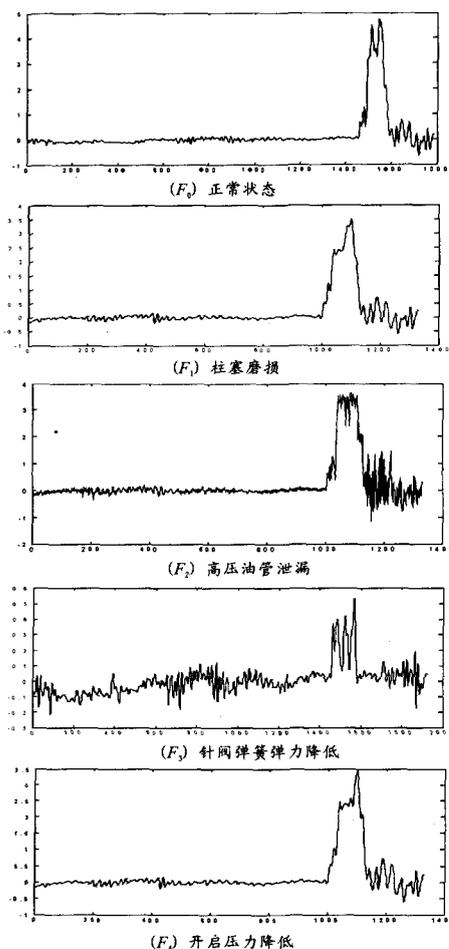


图 1 燃油系统故障油压波形 (转速  $n = 500$  r/min)

### 2.2 基于支持向量机的故障诊断方法

关于如何设置 SVM 的各个参数, 理论上还没有统一的论述, 往往要针对待解决的实际问题来选择. 在本研究的故障诊断中, 采用 RBF 核函数, 也称为高斯函数:

$$k(x_i, x_j) = \exp\left\{-\frac{\|x_i - x_j\|^2}{\sigma^2}\right\}, i = 1, 2, \dots, n \quad (14)$$

其中  $\sigma$  是感知变量, 它决定了高斯函数的宽度.  $\sigma$  的数值可以控制支持向量机的个数: 当支持向量机个数太多时, 可以适当减小  $\sigma$  的数值, 反之亦然. 目前还没有统一的方法来确定  $\sigma$  的数值, 一般根据特征数据分析的结果取经验值. 为简单起见, 本研究采用以下经验公式确定  $\sigma$  的数值:  $\sigma^2 = E(\|x_i - x_j\|^2)$ , 其中  $E$  表示数学期望<sup>[5]</sup>.

$C$  为惩罚参数, 一般也是通过试验来找到对特定训练集最好的  $C$  的数值, 本研究取  $C = 150$ .

应用 SVM 进行故障诊断的步骤: ① 在获得了样本数

据后,首先对数据进行规一化处理,以消除量纲的影响;②选择非线性映射函数  $\varphi$ ,满足  $K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j)$ ,将训练数据通过  $\varphi$  映射到高维特征空间;③选择上述的核函数  $K(x_i, x_j)$ 和惩罚参数  $C$ ,利用训练样本以及式(10)所示的 KKT 条件,求解式(8)、(9)所示的二次优化问题,以获得对应的支持向量  $\alpha_i$ 、球心  $a$  及超球半径  $R$ ;④根据步骤③得到的参数值,计算样本数据到球心的距离,若该距离不超过  $R$ ,即样本点位于超球内,则说明该样本数据为正常值,否则为故障点;⑤对所有的训练样本,重复步骤②~④。

### 2.3 诊断结果

基于各种故障对油压波形的影响,选择峰值( $X_p$ )、方根幅值( $X_{ms}$ )、方差( $\sigma_x^2$ )、斜度( $\alpha$ )、峰值指标( $C_f$ )、重心频率( $FC$ )、均方频率( $MSF$ )、均方根频率( $RMSF$ )为特征参量,进行故障诊断,诊断结果见表1。

表1 特征参量与诊断结果

故障	$F_0$	$F_1$	$F_2$	$F_3$	$F_4$	
部分归一化后的故障特征	$X_p$	0.501	0.152	0.201	0.781	0.915
	$X_{ms}$	0.090	0.010	0.551	0.121	0.961
	$\sigma_x^2$	0.000	0.010	0.010	0.015	0.101
	$\alpha$	0.081	0.000	0.000	0.815	0.001
	$C_f$	0.011	0.511	0.851	0.925	0.501
	$FC$	0.019	0.505	0.890	0.701	0.955
	$MSF$	0.309	0.998	0.751	0.201	0.201
	$RMSF$	0.551	0.995	0.809	0.435	0.500
	诊断结果	$F_0$	$F_1$	$F_2$	$F_3$	$F_4$

表1中的结果表明,SVM能够正确地分类故障样本,具有较强的数据处理能力。为了进一步验证SVM抵抗噪声的能力,在测试样本中加入了噪声。设 $D_1$ 为加入噪声前的故障样本数据矩阵, $D_2$ 为加入噪声后的故障样本数据矩阵,实验需要的含有噪声的样本数据矩阵元素通过下式获得。

$$D_2(i, j) = D_1(i, j) \cdot (1 + \text{rands}(1)) \quad (15)$$

式中: $\text{rands}(1)$ 为随机函数,可生成一个-1到1之间的随机函数。利用上式对每种故障情况分别产生50组的测量参

数,共250组样本,用150组作为训练集,用100组作为测试集,分别采用SVM和RBF网络,学习同一批训练样本和测试样本,在未经任何预处理的情况下直接供SVM和RBF网络进行诊断,得到表2所示的诊断结果。

表2 误诊样本数

故障类型		$F_1$	$F_2$	$F_3$	$F_4 F_1$	$F_2 F_4$
误诊样本数	RBF	0	0	2	3	4
	SVM	0	0	0	1	2

诊断结果表明,SVM较RBF网络具有较强的抗噪声能力,对复合故障样本有较高的诊断准确度,具有很强的分类能力。

### 3 结束语

本研究探讨了SVM的机理,应用SVM对柴油机燃油系统进行了故障诊断,讨论了SVM参数的选择方法。故障诊断结果表明,SVM具有较好的诊断效果,与RBF神经网络相比,SVM具有较强的抗噪声能力和分类能力,对复合故障样本诊断的准确度较高。

### 参考文献:

- [1] Vapik V N. 统计学习理论的本[M]. 张学工,译. 北京:清华大学出版社,2000.
- [2] Tay Francis E H, Cao Li-juan.  $\epsilon$ -descending support vector machines for financial time series forecasting[J]. Neural Processing Letters, 2002, 15(2): 179 - 195.
- [3] 翟永杰,韩埃,王东风,等. 基于损失函数的SVM算法及其在轻微故障诊断中的应用[J]. 中国电机工程学报, 2003, 23(9): 198 - 203.
- [4] 胡寿松,王源. 基于支持向量机的非线性系统故障诊断[J]. 控制与决策, 2001, 15(5): 617 - 620.
- [5] 朱志宇,刘维亭. 基于支持向量机的船舶柴油机故障诊断[J]. 船舶工程, 2006, 28(5): 31 - 33.