

基于支持向量机的故障模式识别研究

齐保林^{1,2}, 李凌均¹, 李志农¹

(1. 郑州大学 振动工程研究所, 河南 郑州 450001; 2. 郑州牧业工程高等专科学校 食品工程系, 河南 郑州 450011)

摘要: 支持向量机为因缺乏大量故障样本受到制约的智能诊断提供了一个全新的途径. 从振动信号中提取特征向量作为支持向量机的输入, 对滚动轴承故障模式进行识别. 实验表明, 在含噪声条件下支持向量机对滚动轴承故障模式仍然具有优秀的分类性能.

关键词: 支持向量机(SVM); 模式识别; 故障诊断

中图分类号: TH 17; TP 18

文献标识码: A

0 引言

智能故障诊断是从被测数据中寻求规律, 对机械运行状态进行判别和预测. 传统统计理论是基于数据的学习机器的理论基础, 研究样本数趋于无穷大时的统计规律. 而故障诊断中, 获得大量的故障样本作为机器学习的先验知识常常受到条件限制. Vladimir N. Vapnik 等研究的统计学习理论(Statistical Learning Theory, SLT), 专门研究有限样本下的统计规律^[1,2]. 支持向量机(Support Vector Machine, SVM)是在 SLT 结构风险最小化原则上建立起来的机器学习方法. 与传统的统计理论不同, SVM 不追求样本数趋于无穷大时的最优解, 而追求在有限样本下的最优解, 是一种兼顾经验风险和置信范围的折中思想^[3]. SVM 克服了困扰机器学习的诸多问题, 如神经网络过学习、维数灾难、局部极小值等, 并具有更好的学习、泛化性能.

由于 SVM 优秀的学习性能, 成为近十年来机器学习的研究热点, 在手写字体识别、语音识别等领域得到了成功的应用. 笔者将该方法用于机械故障诊断领域, 讨论了 SVM 方法在故障模式分类中的应用, 并以滚动轴承振动信号为对象, 实验验证了该方法在有限故障样本和含噪声条件下对故障模式分类具有良好的分类性能. SVM 方法的应用为设备监测、智能诊断提供了一种新途径.

1 支持向量机算法

1.1 最优分类面

设样本集 $\{(x_i, y_i)\}$ 线性可分, $i = 1, 2, \dots, n$, $x_i \in R^d$, $y_i \in \{-1, +1\}$ 是样本的类别符号. 为使两类样本无错分开, 构造方程 $w \cdot x + b = 0$ 的最优分类超平面 H , 如图 1 所示. 图中 $H_1: w \cdot x_i + b = 1$, $H_2: w \cdot x_i + b = -1$ 为过两类样本点、距离 H 最近且平行于 H 的超平面. H_1 和 H_2 之间的距离为分类间隔 Margin. 最优分类面是将两类样本无错分开, 且使分类间隔最大的超平面. 将 H_1 和 H_2 的方程归一化:

$$y_i [w \cdot x_i + b] - 1 \geq 0 \quad (i = 1, 2, \dots, n) \quad (1)$$

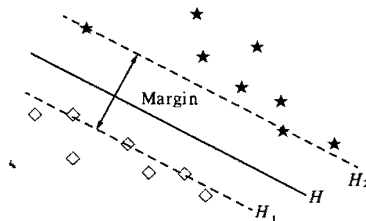


图 1 最优分类面

Fig. 1 Optimal separating hyperplane

根据 Vapnik 统计学习理论, 将两类样本无错分开保证了经验风险值最小, 而分类间隔最大则使得推广性界的置信范围最小, 从而使真实风险最小^[2].

收稿日期: 2006-07-21; 修订日期: 2006-09-22

基金项目: 河南省自然科学基金资助项目(0611022400)

作者简介: 齐保林(1963-), 男, 河南新野人, 郑州牧业工程高等专科学校副教授, 郑州大学硕士研究生, 主要从事设备故障智能诊断、支持向量机等方面的研究工作.

由图 1 可知,两类样本集到分类面 H 的最近距离分别为 $1/\|\mathbf{w}\|$ 。欲使分类间隔 $2/\|\mathbf{w}\|$ 最大,需使下式最小。

$$\Phi(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 = \frac{1}{2} \mathbf{w}^T \mathbf{w} = \frac{1}{2} (\mathbf{w} \cdot \mathbf{w}) \quad (2)$$

即求解最优分类面,是在式(1)约束下求函数式(2)的极小值。过超平面 H_1, H_2 的样本点就是方程式(2)达到极值的样本点,它们共同支持最优分类超平面 H ,称为支持向量(Support vectors, SVs)。为求解方程式(2)的极小值,定义下列 Lagrange 函数^[4]

$$L_p(\mathbf{w}, b, \alpha) = \frac{1}{2} (\mathbf{w} \cdot \mathbf{w}) - \sum_{i=1}^n \alpha_i \{y_i [\mathbf{w} \cdot \mathbf{x}_i + b] - 1\} \quad (3)$$

式中: $\alpha_i (\alpha_i \geq 0)$ 是 Lagrange 系数。这里的目标是对 \mathbf{w} 和 b 求函数式(3)的最小值。根据约束优化理论的 Kuhn - Tucker(KT) 条件^[4],式(3)的极值问题可转化为其较为简单的对偶问题,即在条件

$$\sum_{i=1}^n y_i \alpha_i = 0 \quad (\alpha_i \geq 0) \quad (i = 1, 2, \dots, n) \quad (4)$$

约束下求

$$L_D(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \quad (5)$$

的最大值。这是一个在不等式约束条件下求解二次函数的极值问题,有唯一解,且满足

$$\alpha_i [y_i (\mathbf{w} \cdot \mathbf{x}_i) + b] - 1 = 0 \quad (i = 1, 2, \dots, n) \quad (6)$$

式中: α_i 对应的样本 \mathbf{x}_i 是支持向量。若 α^* 为优化问题的最优解,由式(5)得到最优分类面的权系数向量为

$$\mathbf{w}_0 = \sum_{i=1}^n \alpha_i^* y_i \mathbf{x}_i = \sum_{i \in \{i | \alpha_i^* > 0\}} \alpha_i^* y_i \mathbf{x}_i \quad (7)$$

最优分类面的分类,用判别函数表示

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i \in \{i | \alpha_i^* > 0\}} y_i \alpha_i^* (\mathbf{x} \cdot \mathbf{x}_i) - b_0 \right) \quad (8)$$

式中: b_0 常数可以根据方程式(6)由两类样本中任意一个支持向量求得。

1.2 广义最优分类面

对于线性不可分样本,引入松弛因子 $\zeta_i (\zeta_i \geq 0)$ 。此时,不等式(1)的约束条件为

$$y_i [\mathbf{w} \cdot \mathbf{x}_i + b] - 1 + \zeta_i \geq 0 \quad (i = 1, 2, \dots, n) \quad (9)$$

则求解广义最优分类面,就是在式(9)和 $\zeta_i \geq 0$ 约束下求函数

$$L_p(\mathbf{w}, \zeta) = \frac{1}{2} (\mathbf{w} \cdot \mathbf{w}) + C \sum_{i=1}^n \zeta_i \quad (10)$$

的极小值。

式中: C 为惩罚系数 ($C > 0$),是对越界进入超平面 H_1, H_2 样本点的惩罚,实现算法复杂度和样本错分率之间的折中。当 $C \rightarrow \infty$,是两类样本无错分开的理想情况。方程式(10)转变成和方程(5)形式相同的对偶问题,约束条件为

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

1.3 支持向量机^[3]

判别函数式(8)只包含待分类样本和支持量的内积。对原输入空间中线性不可分的样本,通过非线性映射 T 将原空间样本 \mathbf{x} 映射到某高维线性特征空间 F 中,即 $R^d \rightarrow F, \mathbf{x} \rightarrow T(\mathbf{x})$,然后用广义最优分类面分类。高维线性特征空间的内积定义为 $K(\mathbf{x}_i, \mathbf{x}_j) = T(\mathbf{x}_i) \cdot T(\mathbf{x}_j)$,称为核函数,其选择需要满足 Mercer 条件^[1]。SVM 优势在于不必知道映射 T 的具体函数,在最优分类面中采用适当的核函数 $K(\mathbf{x}_i, \mathbf{x}_j)$ 可实现从低维空间向高维空间的映射,实现非线性变换后的线性分类,而计算复杂度并不增加。核函数不同, SVM 算法不同。常用核函数^[4]有线性核函数 $K(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;多项式核函数 $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + 1)^d, d = 1, 2, \dots$;径向基核函数 $K(\mathbf{x}, \mathbf{y}) = \exp \left[-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2} \right]$;多层感知器核函数 $K(\mathbf{x}, \mathbf{y}) = \tanh[b(\mathbf{x} \cdot \mathbf{y}) + c]$ 。

2 SVM 的应用

2.1 实验数据

在振动实验台上,采集正常滚动轴承和滚动体点蚀轴承的振动加速度信号^[5]。轴承载荷 2.2 kW,工作转速 1 730 r/min,采样频率 12 kHz,每个样本采样长度为 4 096 点。

2.2 试验与分析

2.2.1 选择学习样本和测试样本

选取 18 个样本作为学习样本集,12 个样本作为测试样本集。学习样本集中,10 个为正常轴承样本,8 个为故障样本。测试样本集中,6 个为正常轴承样本,6 个为滚动体点蚀故障样本。从原始样本数据中提取峰值、绝对均值、方差、峭度、斜度、高阶累积量等统计量,组成特征向量^[6,7],输入 SVM 分类器中进行训练、测试。使用 2 阶径向基核函数 SVM。定义无故障样本符号 $y = +1$,有故障符号 $y = -1$ 。选取惩罚系数 $C = 100$ 。

对学习样本集,计算得 SVM 分类器模型参数为 $\|\mathbf{w}\|^2 = 391.601\ 786, \alpha_i = [0, 0, 0, 18.291\ 6,$

0,0,0,0,177.562 5,0,0,0,0,0,0,195.747 7,0]^T. α_i 中的 3 个不为 0 的点为支持向量(占训练样本 16.7%),训练误差为 0,分类间隔 0.101 067, $b=0$.分类结果见表 1,12 个测试样本被正确分类,分类正确率 100%.

表 1 12 个测试样本分类结果
Tab. 1 Classification results

编号	已知	无噪声		含噪声	
		分类距离	分类	分类距离	分类
1	正常	1.289 104	1	1.483 556	1
2	正常	1.660 493	1	1.719 390	1
3	正常	1.075 224	1	2.506 155	1
4	正常	1.264 488	1	2.317 857	1
5	正常	0.895 937	1	2.542 922	1
6	正常	0.373 476	1	1.426 066	1
7	故障	-1.240 820	-1	-1.937 457	-1
8	故障	-1.360 918	-1	-0.981 129	-1
9	故障	-1.364 671	-1	-1.020 855	-1
10	故障	-1.654 474	-1	-1.381 972	-1
11	故障	-1.559 839	-1	-2.336 147	-1
12	故障	-0.992 593	-1	-10.517 655	-1

在上述学习样本和测试样本中加入白噪声信号,重新计算.得分类器模型参数: $\|w\|^2=1\,121.517\,447$, $\alpha_i=[0,0,0,0,0,0,0,0,560.758\,7,0,0,0,0,0,560.758\,7,0,0,0]^T$,支持向量数为 2(11.1%),训练误差为 0, $b=0$.由于噪声干扰,分类距离(负值取绝对值)有所增加(见表 1),分类间隔减小到 0.059 721.12 个含噪声的测试样本仍然被正确分类,分类正确率 100%.这表明,SVM 在有噪声环境下仍然具有良好的分类性能.

2.2.2 减少学习样本数,再次实验

重新选取 9 个样本作为学习样本集,其中 5 个为正常轴承样本,4 个为故障样本.选取 7 个样本为测试样本集,其中 4 个正常样本,3 个故障样本.选用线性核函数 SVM 进行实验,仍然取惩罚系数 $C=100$.输入训练样本,得到 SVM 模型参数: $\|w\|^2=47.489\,884$, $\alpha_i=[0,0,0,12.145\,0,11.599\,9,0,0,23.744\,9,0]^T$,支持向量个数为 3(33.3%),训练误差为 0, $b=4.042\,9$.分类间隔为 0.290 221.分类结果见表 2,7 个测试样本被正确分类,故障分类正确率为 100%.

在上述学习样本和测试样本中加入白噪声,测试 SVM 分类性能.运行结果: $\|w\|^2=77.809\,461$, $\alpha_i=[38.904\,7,0,0,0,0,0,0,0,38.904\,7]^T$,即支持向量数 2(22.2%),训练误差为 0, $b=20.257\,8$.在噪声干扰下,分类间隔减小

到 0.226 733.从表 2 分类结果可知,7 个含有噪声的测试样本被正确分类,故障分类正确率 100%.这表明,在小样本、含噪声条件下,SVM 具有优秀的分类性能.

表 2 7 个测试样本分类结果
Tab. 2 Classification results

编号	已知	无噪声		含噪声	
		分类距离	分类	分类距离	分类
1	正常	0.795 477	1	1.472 408	1
2	正常	1.228 493	1	1.449 044	1
3	正常	1.064 442	1	1.148 149	1
4	正常	1.028 672	1	1.250 096	1
5	故障	-0.864 476	-1	-0.791 373	-1
6	故障	-0.775 076	-1	-0.722 408	-1
7	故障	-0.900 385	-1	-0.844 911	-1

3 结论

SVM 是在 SLT 基础上发展起来的新的机器学习算法.实验表明:对小样本、含噪声的轴承振动信号所包含的技术状态信息,具有出色的分辨性能.SVM 优秀的模式分类能力,能够满足机械故障诊断的工程实际要求,在机械故障监测、智能诊断领域具有广阔的应用前景.

参考文献:

[1] VAPNIK V N. The nature of statistical learning theory [M]. 2nd Edition, New York: Spring - Verlag, 1999.

[2] VAPNIK V N. 统计学习理论[M]. 许建华,张学工,译.北京:电子工业出版社,2004.

[3] CRISTANINI N, SHAWE - TAYLOR J. An introduction to support vector machines and other kernel - based learning methods [M]. New York: Cambridge University Press, 2000.

[4] STEVE R, GUNN. Support vector machines for classification and regression [R]. Southampton: University of Southampton, 1998.

[5] 李凌均,张周锁,何正嘉.支持向量机在机械故障诊断中的应用研究[J].计算机工程与应用,2002,38(19):19-21.

[6] 韩捷,张瑞林.旋转机械故障机理及诊断技术[M].北京:机械工业出版社,1997.

[7] 毕果,韩捷,梁川.基于矢量振动信号的 AR 功率谱分析及应用[J].郑州大学学报(工学版),2003,24(3):80-83.

- 器人航迹推算方法研究[J]. 河南理工大学学报, 2005, 24(3): 210-216.
- [6] BORENSTEIN J, FENG L. A method for measuring, comparing, and correcting dead-reckoning errors in mobile robots[R]. Copies of this report are available from the University of Michigan as: Technical Report UM-MEAM-94-22. 20-22.
- [7] 柳方舟, 李满华, 马培荪. 管道形轮腿式月球探测机器人的运动学建模[J]. 东南大学学报(自然科学版), 2003, 33(6): 741-745.
- [8] 丁宝新. 关于定位与测磁中的坐标转换问题[J]. 水雷战与舰艇防护, 1994, (2): 5-14.

Study on Dead Reckoning for a Deep-sea Cobalt-rich Mining Vehicle

BU Ying-yong, LIU Guang-hua, SHI Chun-xue, WANG Ji-chan

(Mechanical and Electrical Engineering College, Central South University, Changsha 410083, China)

Abstract: Taking into account uneven terrain, complex working environment, rolling slip, we design a driven wheel machine with the ability to apperceive changing of terrain and measure the velocity. We present the kinematics model for a deep-sea mining vehicle operating in uneven terrain combining different kinematics models of the vehicle on different slopes at different times. Sensors such as tilt, compass, coder and angular displacement sensors measure the dynamic parameters information, then the dead reckoning of mining vehicle is realized on complex deep-sea environment according to its kinematics model. It can work under complex working conditions. The result of experiment shows that these technologies are viable.

Key words: deep-sea mining vehicle; cobalt-rich; uneven terrain; dead reckoning; driven wheel

(上接第11页)

Research on Fault Pattern Recognition Based on Support Vector Machines

QI Bao-lin^{1,2}, LI Ling-jun¹, LI Zhi-nong¹

(1. Institute of Vibration Engineering, Zhengzhou University, Zhengzhou 450001, China; 2. Food Engineering Department, Zhengzhou College of Animal Husbandry Engineering, Zhengzhou 450011, China)

Abstract: SVM provides people with a completely new means of artificial intelligent diagnosis which was restricted by shortage of large quantity of fault datasets. In this paper, feature vectors extracted from vibration signals of rolling bearing were selected as inputs of support vector machines to recognize fault patterns of rolling bearing. The results show that SVMs have excellent classification performance to fault patterns of rolling bearing, even under noises and interference.

Key words: support vector machines (SVMs); pattern recognition; fault diagnosis