

惩罚复杂诊断系统混合噪音的模糊小波分类机

吴奇^{1,2} 刘静³ 熊福力¹ 刘晓军²

摘要 针对复杂故障诊断系统特征数据中具有高斯、幅值较大、奇异点等混合噪音,设计一种分段式损失函数,构造基于小波基函数的小波核函数,由此得到一种输出为模糊数的模糊小波 ν -支持向量分类机,即模糊鲁棒小波 ν -支持向量分类机 (FRW ν -SVC). 它可以有效地压制故障特征时序的多种噪音和奇异点,具有很强的鲁棒性,而且它比标准模糊小波 ν -支持向量分类机 (FW ν -SVC) 具有更简洁的对偶优化问题. 最后进行了汽车装配线故障诊断的实例分析,结果表明基于 FRW ν -SVC 的故障诊断模型是有效可行的.

关键词 支持向量分类机,小波核函数,分段式损失函数,故障诊断
中图分类号 TP391

The Fuzzy Wavelet Classifier Machine with Penalizing Hybrid Noises from Complex Diagnosis System

WU Qi^{1,2} LIU Jing³ XIONG Fu-Li¹ LIU Xiao-Jun²

Abstract Aiming at Gaussian noise, greater breadth noise, and oddity point noise from feature series of the complex fault diagnosis system, a segment loss function is designed and wavelet kernel function is constructed on the basis of wavelet base function. Then, a fuzzy wavelet ν -support vector classifier machine whose outputs are fuzzy numbers is proposed, named as fuzzy robust wavelet ν -support vector classifier machine (FRW ν -SVC). FRW ν -SVC, which has stronger robustness and simpler dual optimization problem than standard fuzzy wavelet support vector classifier machine (FW ν -SVC), can inhibit some types of noise and oddity point noise of fault feature series effectively. Finally, FRW ν -SVC is applied to the fault diagnosis of a car assembly line, the results showing the proposed model based on FRW ν -SVC is effective and feasible.

Key words Support vector classifier machine, wavelet kernel function, segment loss function, fault diagnosis

复杂工业系统故障诊断过程是一个动态过程,其诊断行为受很多因素的影响,这些因素大多具有随机性、非线性等特点,这些因素之间复杂的非线性关系,在低维空间很难用一种确定的方法来表达^[1-2]. 而且实际获取的故障信息常受到一些随机误差的影响而退化,通常称这个退化为噪音 (Noise). 在数据的收集、传输或处理过程中都可能出现各种噪音,噪音可能依赖于模式特征时序,也可能与其无关. 理想的噪音,称为白噪音 (White noise),它是退化的最坏估计. 而白噪音的一个特例是高斯噪音 (Gaussian noise). 传统基于模型的故障诊断方法过分依赖系统的数学模型,对于建模误差、参数扰动、

噪声和干扰十分敏感,使得这种方法对复杂非线性系统的故障诊断处理效果不好. 而神经网络主要用于具有大量样本数据的分析方法,但其泛化能力较差,涉及的输入变量较多,使得神经网络模型面临“维数灾难”问题^[3]. 理论上,传统模型的建立都没有考虑对依赖于模式特征时序中的噪音如何进行惩罚 (降噪).

对较少的样本数据,样本数据间的不确定性关系无法在样本数据维空间用一种确定的方式来表达. 但采用结构风险最小原则的支持向量机,将样本空间的数据关系映射到一个高维空间,则在高维空间里,数据间的关系可以用一种确定的表达方式表示. 支持向量机^[4] 是由 Vapnik 提出的一种全新的机器学习方法,由于其具有很好的泛化能力,得到的解是全局最优解,而且能较好地解决小样本、非线性、高维数等实际问题^[5],在国内外学术界已受到广泛重视,并已在模式识别、非线性动态系统辨识、故障诊断、预测建模等方面取得越来越多的应用^[6-12].

由于标准支持向量分类机的 ϵ 不敏感损失函数的缺陷性,不能够有效处理一些高斯、幅值较大和奇异点的噪音^[11],本文设计了一种分段式损失函数,将松弛变量的作用区间分为三个:靠近最优分界的

收稿日期 2008-12-16 收修改稿日期 2009-03-12
Received December 16, 2008; in revised form March 12, 2009
上海市教育发展基金会晨光计划 (2008CG55) 资助
Supported by Shanghai Education Development Foundation Chenguang Project (2008CG55)

1. 复杂工程系统测量与控制教育部重点实验室 (东南大学自动化学院) 南京 210096 2. 东南大学机械工程学院 南京 210096 3. 上海海事大学信息工程学院 上海 200135

1. Key Laboratory of Measurement and Control of Complex Systems of Engineering (School of Automation, Southeast University), Ministry of Education, Nanjing 210096 2. School of Mechanical Engineering, Southeast University, Nanjing 210096 3. College of Information Engineering, Shanghai Maritime University, Shanghai 200135
DOI: 10.3724/SP.J.1004.2009.00773

区间为管道区,中间区间为高斯区,最外层的区间为线性和奇异点区,从而使得落在各个区间内的噪音会被各自区间的损失函数有效压制。

经典分类(两类)问题可用确定的正类(1)或负类(-1)表示。显然训练集中训练点的输出为正类(1)或负类(-1)是由领域专家根据历史资料和自己的经验、知识得出的判决。然而有时由于历史资料不完整或专家的经验、知识不全面,因此专家对某些训练点只能得出模糊判决,如样本点80%属于正类,而20%属于负类。因此,用一个确定数是不能刻化这种模糊决策,只能用一个模糊数来表示。

基于这一思想启发,本文设计一种分段式损失函数,采用小波核函数,提出一种输出为模糊数的模糊小波支持向量分类机,即模糊鲁棒小波 ν -支持向量分类机(FRW ν -SVC)。并将这种模型与制造系统装配线故障诊断的特点相结合,从而得到一种适用于制造系统的故障诊断方法。

1 鲁棒小波 ν -支持向量分类机

在支持向量分类机中,大多数学者仍然使用 ϵ 不敏感损失函数(一次松弛变量)来处理模式特征时序,使用方法不具有针对性^[12-18]。Rossi和Villab^[18]用支持向量机对数据进行分类,使用的是 ϵ 不敏感损失函数,没有针对数据特征设计对应的损失函数。Shim和Hwang等^[19]使用一个平方损失函数的支持向量机作为鲁棒支持向量机,但没有从理论上说明这种损失函数的优点,而且无法处理幅度较大和奇异点的噪音。

一些学者对标准向量机的提法作出修正^[9-10],以便支持向量机所形成的对偶问题更容易求解。这种做法是很值得研究的,因为支持向量机的核心思想就是结构风险最小化,标准支持向量机是实现结构风险的一种途径,但不是唯一途径。

基于以上考虑,本文建立一种新的最优化问题,设计一种分段式损失函数,采用小波核函数,提出一种新的模糊小波支持向量分类机,即模糊鲁棒小波-支持向量分类机(FRW ν -SVC)。

1.1 分段损失函数

样本集的最优损失函数与样本集数据的内在特性之间存在一定的关系。对于高斯噪音,选择高斯函数作为损失函数可以得到最好的降噪效果;对于奇异点和幅值较大的噪音,Laplace损失函数有好的降噪效果^[11]。本文综合各种损失函数的特点,以及样本数据的未知性、随机性、不平稳性等特征,在设计损失函数时将体现上述特征的函数考虑到新的损失函数内,提出分段式损失函数,其数学提法如下:

$$L(\xi) = \begin{cases} 0, & |\xi| \leq \epsilon \\ \frac{1}{2}(|\xi| - \epsilon)^2, & \epsilon < |\xi| \leq \epsilon_\mu \\ \mu(|\xi| - \epsilon) - \frac{1}{2}\mu^2, & |\xi| > \epsilon_\mu \end{cases} \quad (1)$$

其中, $\epsilon + \mu = \epsilon_\mu$, $\epsilon \geq 0$, $\mu \geq 0$ 。鲁棒损失函数分为三部分:

1) $|\xi| \leq \epsilon$ 部分,即 ϵ 不灵敏区,不惩罚小于 ϵ 的偏差,使学习机的解具有稀疏性;

2) $\epsilon < |\xi| \leq \epsilon_\mu$ 部分,选用高斯函数损失函数 $\frac{1}{2}(|\xi| - \epsilon)^2$,可以抑制符合高斯特征的噪音;

3) $|\xi| > \epsilon_\mu$ 部分,选用Laplace损失函数 $\mu(|\xi| - \epsilon)$,可以较好地抑制较大的噪音和一些奇异点。

1.2 小波核函数

根据广义线性判别函数的思路,要解决一个非线性问题,可通过非线性变换将输入空间变换到一个高维特征空间,在此特征空间中寻求最优或广义最优分类面,且只需进行内积运算,而没有必要知道采用的是非线性变换的形式。只要在变换后的特征空间的内积可以用原空间中的变量直接计算得到,则即使特征空间中的维数增加很多,在其中求解最优分类面的问题也并没有增加很多计算复杂度。

统计学习理论指出,根据Hilbert-Schmidt原理,只要一种运算满足Mercer条件,它就可作为内积使用。

引理1 (Mercer条件)^[10,20]。

$L^2(\mathbf{R}^n)$ 下的对称函数 $K(\mathbf{x}, \mathbf{x}')$ 为特征空间中的内积的充分必要条件是,对于任意的 $\varphi(\mathbf{x}) \neq 0$ 且 $\int \varphi^2(\mathbf{x})d\mathbf{x} < \infty$,满足下列条件

$$\iint K(\mathbf{x}, \mathbf{x}')\varphi(\mathbf{x})\varphi(\mathbf{x}')d\mathbf{x}d\mathbf{x}' \geq 0, \quad \mathbf{x}, \mathbf{x}' \in \mathbf{R}^n \quad (2)$$

则满足式(2)的点积核 $K(\mathbf{x}, \mathbf{x}')$ 是一个允许支持向量核。

引理2 (平移不变核)^[10,20]。

$L^2(\mathbf{R}^n)$ 下的对称函数 $K(\mathbf{x}, \mathbf{x}')$, $\mathbf{x}, \mathbf{x}' \in \mathbf{R}^n$,平移函数满足 $K(\mathbf{x}, \mathbf{x}') = K(\mathbf{x} - \mathbf{x}')$,若满足Mercer条件,那么它就是一个允许支持向量核。

引理1和引理2提供了一个判断和构建核函数的方法,但是实际上很难将平移不变核分解为两个函数的内积,引理3给出了平移不变核函数是一个支持向量核的充分必要条件。

引理3^[10,20]。平移不变核函数 $K(\mathbf{x}, \mathbf{x}') = K(\mathbf{x} - \mathbf{x}')$ 是一个允许支持向量核,当且仅当 $K(\mathbf{x})$ 的傅里叶变换成立。

$$F[\mathbf{x}](\boldsymbol{\omega}) = (2\pi)^{-\frac{n}{2}} \int_{\mathbf{R}^n} \exp(-j(\boldsymbol{\omega} \cdot \mathbf{x})) K(\mathbf{x}) d\mathbf{x} \geq 0$$

$$\boldsymbol{\omega} \in \mathbf{R}^n, \mathbf{x} \in \mathbf{R}^n \quad (3)$$

将母小波函数替换成某些具体形式, 即可得到相应的小波核函数, 下面是构造小波核函数的定理.

引理 4^[10, 20]. 设母小波函数为 $\psi(\mathbf{x}_i)$, $\mathbf{x}_i, \mathbf{x}'_i \in \mathbf{R}^d$, \mathbf{x}' 和 \mathbf{x}'_i 都是 d 维列向量, \mathbf{a} 为 l 维尺度因子列向量, \mathbf{b} 和 \mathbf{b}' 是 d 维位移因子列向量, 其中 $\mathbf{a} \in \mathbf{R}^l$, $\mathbf{b}, \mathbf{b}' \in \mathbf{R}^d$, 则满足 Mercer 条件的点积形式的小波核函数

$$K(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^l \psi\left(\frac{\mathbf{x}_i - \mathbf{b}}{a_i}\right) \psi\left(\frac{\mathbf{x}'_i - \mathbf{b}'}{a_i}\right) \quad (4)$$

和满足平移不变核定理的小波核函数

$$K(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^l \psi\left(\frac{\mathbf{x}_i - \mathbf{x}'_i}{a_i}\right) \quad (5)$$

均是可容许的支持向量核.

式 (5) 中没有出现位移因子 \mathbf{b} , 但有一个变量 \mathbf{x}'_i , 该变量的作用等价于位移因子. 因此, 可以用小波函数来构造具有平移不变性的小波核函数.

定理 1. Morlet 母小波函数为

$$\psi(\mathbf{x}) = \cos(\omega_0 \mathbf{x}_i) \exp\left(-\frac{\|\mathbf{x}_i\|^2}{2}\right), \omega_0 \in \mathbf{R}$$

则小波核函数

$$K(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^l \cos\left(\omega_0 \times \frac{\mathbf{x}_i - \mathbf{x}'_i}{a_i}\right) \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}'_i\|^2}{2a_i^2}\right) \quad (6)$$

是一种可容许的支持向量核, 其中 \mathbf{x}' 和 \mathbf{x}'_i 都是 d 维列向量.

证明. 根据引理 3, 只要证明:

$$F[\mathbf{x}](\boldsymbol{\omega}) = (2\pi)^{-\frac{d}{2}} \int_{\mathbf{R}^d} \exp(-j(\boldsymbol{\omega} \cdot \mathbf{x})) K(\mathbf{x}) d\mathbf{x} \geq 0$$

$$\boldsymbol{\omega} \in \mathbf{R}^d, \mathbf{x} \in \mathbf{R}^d \quad (7)$$

其中

$$K(\mathbf{x}) = \prod_{i=1}^l \psi\left(\frac{\mathbf{x}_i}{a_i}\right) =$$

$$\prod_{i=1}^l \cos\left(\frac{\omega_0 \cdot \mathbf{x}_i}{a_i}\right) \exp\left(-\frac{\|\mathbf{x}_i\|^2}{2a_i^2}\right)$$

j 表示虚数单位. 因为

$$\int_{\mathbf{R}^d} \exp(-j(\boldsymbol{\omega} \cdot \mathbf{x})) K(\mathbf{x}) d\mathbf{x} = \int_{\mathbf{R}^d} \exp(-j(\boldsymbol{\omega} \cdot \mathbf{x})) \times$$

$$\prod_{i=1}^l \cos\left(\frac{\omega_0 \cdot \mathbf{x}_i}{a_i}\right) \exp\left(-\frac{\|\mathbf{x}_i\|^2}{2a_i^2}\right) d\mathbf{x} =$$

$$\prod_{i=1}^l \int_{-\infty}^{+\infty} \exp(-j(\omega_i) x_i) \exp\left(-\frac{\|\mathbf{x}_i\|^2}{2a_i^2}\right) \times$$

$$\left(\frac{\exp(j\omega_0 x_i/a_i) + \exp(-j\omega_0 x_i/a_i)}{2}\right) d\mathbf{x}_i =$$

$$\prod_{i=1}^l \frac{1}{2} \int_{-\infty}^{+\infty} \left(\exp\left(-\frac{\|\mathbf{x}_i\|^2}{2a_i^2} + \left(\frac{j\omega_0}{a_i} - j\omega_i a_i\right) \mathbf{x}_i\right) +$$

$$\exp\left(-\frac{\|\mathbf{x}_i\|^2}{2a_i^2} - \left(\frac{j\omega_0}{a_i} + j\omega_i a_i\right) \mathbf{x}_i\right)\right) d\mathbf{x}_i =$$

$$\prod_{i=1}^l \frac{\sqrt{2\pi} |a_i|}{2} \left(\exp\left(-\frac{(\omega_0 - \omega_i a_i)^2}{2}\right) +$$

$$\exp\left(-\frac{(\omega_0 + \omega_i a_i)^2}{2}\right)\right) \quad (8)$$

将式 (8) 代入式 (7), 得

$$F[\mathbf{x}](\boldsymbol{\omega}) = \prod_{i=1}^l \frac{|a_i|}{2} \left(\exp\left(-\frac{(\omega_0 - \omega_i a_i)^2}{2}\right) +$$

$$\exp\left(-\frac{(\omega_0 + \omega_i a_i)^2}{2}\right)\right) \quad (9)$$

当 $a_i \neq 0$ 时, 有

$$F[\mathbf{x}](\boldsymbol{\omega}) \geq 0 \quad (10)$$

□

同样, 可以证明以下 Mexico 小波核函数也是可容许的支持向量核. Mexico 小波为

$$\psi(\mathbf{x}) = (1 - \|\mathbf{x}_i\|^2) \exp\left(-\frac{\|\mathbf{x}_i\|^2}{2}\right) \quad (11)$$

其小波核函数为

$$K(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^l \left(1 - \frac{\|\mathbf{x}_i - \mathbf{x}'_i\|^2}{a_i^2}\right) \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}'_i\|^2}{2a_i^2}\right) \quad (12)$$

其中, \mathbf{x}' 和 \mathbf{x}'_i 都是 d 维列向量.

1.3 模糊鲁棒小波 ν -支持向量分类机

在对标准小波支持向量分类机的求解过程中, 参数 b 的精确确定是困难的, 从而导致在模型参数寻优过程中消耗大量迭代时间. 为此本文从结构风险最小化原则出发, 在建立模型初期, 参数 b 就被考虑进最优化问题的原问题中, 获取一种新的小波支持向量分类机原问题. 它减少了一个对偶问题的约束条件, 参数 b 将不出现在最终决策函数中. 整个求解过程无需对参数 b 进行辩识, 从而提高了计算效率. 具体推导过程如下.

设鲁棒小波 ν -支持向量分类机的最优分界面具有 $\bar{f}(\mathbf{x}) = \bar{\mathbf{w}} \cdot \bar{\mathbf{x}}$ 形式, 标准模糊小波 ν -支持向量分类机的最优分界面有标准形式 $f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b$. 对于样本集 $T = \{\mathbf{x}_i, y_i\}_{i=1}^l$, 其中 $\mathbf{x}_i \in \mathbf{R}^d$, $y_i \in \mathbf{R}$, \mathbf{x}_i 为 d 维列向量. 考虑由 $\bar{\mathbf{x}} = (\mathbf{x}^T, \eta)^T$ 构成的 Hilbert 空间, 并定义 $\bar{\mathbf{x}}_1 = (\mathbf{x}_1^T, \eta)^T$ 和 $\bar{\mathbf{x}}_2 = (\mathbf{x}_2^T, \eta)^T$ 的内积 $\bar{\mathbf{x}}_1 \cdot \bar{\mathbf{x}}_2 = \mathbf{x}_1 \cdot \mathbf{x}_2 + \eta^2$, 令 $\bar{\mathbf{w}} = (\mathbf{w}^T, \frac{b}{\eta})^T$, 其中 \mathbf{w} 为 d 维列向量, $\bar{\mathbf{w}}$ 为 $d+1$ 维列向量, η 为不等于零的实数, 于是有 $\bar{f}(\mathbf{x}) = \bar{\mathbf{w}} \cdot \bar{\mathbf{x}} = \mathbf{w} \cdot \mathbf{x} + \frac{b}{\eta} \eta = \mathbf{w} \cdot \mathbf{x} + b = f(\mathbf{x})$ 成立. 由此可见, 本文所建立的最终决策函数中不出现参数 b 的模糊鲁棒小波 ν -支持向量机的最优化问题 $\bar{f}(\mathbf{x})$, 等价于最终决策函数中出现参数 b 的标准模糊小波 ν -支持向量分类机的最优化问题 $f(\mathbf{x})$. 由于训练点的输出为模糊数, 标准小波支持向量分类机无法对此模糊数进行处理, 为此, 定义 1 建立了模糊数与确定数之间的一种映射关系. 通过这个映射, 一个三角模糊数被转化为一个确定数, 从而方便计算.

定义 1. 对于训练点的输出为模糊数的训练集 $T = \{(\mathbf{x}_i, y_i)\}_{i=1}^l$, 其中 $\mathbf{x}_i \in \mathbf{R}^d$, y_i 为反映其模糊类别模糊数, 样本点属于正类 (用 1 表示) 或负类 (用 -1 表示) 的隶属度分别为 δ^+ 或 δ^- ($\delta^+, \delta^- \in [0.5, 1]$). 为表示方便, 引进 $\delta \in [-1, -0.5] \cup [0.5, 1]$, 使得 $\delta^+ = \delta$, $\delta^- = -\delta$. 于是, y_i 的三角模糊特征可以用下面特殊的三角模糊数表示

$$y_i = (r_1, r_2, r_3) = \begin{cases} \left(\frac{2\delta^2 + \delta - 2}{\delta}, 2\delta - 1, \frac{2\delta^2 - 3\delta + 2}{\delta} \right), & 0.5 \leq \delta \leq 1 \\ \left(\frac{2\delta^2 + 3\delta + 2}{\delta}, 2\delta + 1, \frac{2\delta^2 - \delta - 2}{\delta} \right), & -1 \leq \delta \leq -0.5 \end{cases} \quad (13)$$

模糊正类: 如样本点属于正类的隶属度为 0.9, 而属于负类的隶属度为 0.1, 此模糊特征可用三角模糊数 $\bar{a} = (0.58, 0.8, 1.02)$ 表示, 其中, 0.8 为 \bar{a} 的中心.

模糊负类: 如样本点属于负类的隶属度为 0.8, 而属于正类的隶属度为 0.2, 此模糊特征可用三角模糊数 $\bar{b} = (-1.1, -0.6, -0.1)$ 表示, 其中, -0.6 为 \bar{b} 的中心.

居中: 如样本点属于正类的隶属度为 0.5, 而属于负类的隶属度为 0.5, 此模糊特征可用三角模糊数 $\bar{c} = (-2, 0, 2)$ 表示, 其中 0 为 \bar{c} 的中心.

这样, 任何一个三角模糊数都可以用隶属度 δ^+ 或 δ^- 来表示. 于是, 对于模糊样本集 $T = \{(\mathbf{x}_i, y_i)\}_{i=1}^l$, 其中 $\mathbf{x}_i \in \mathbf{R}^d$, y_i 为形如式 (13) 的三角模糊数, 采用分段式损失函数 (1) 的模糊鲁棒小波 ν -支持向量分类机 (FRW ν -SVC) 的数学提法如下:

$$\begin{aligned} \min_{\mathbf{w}, \xi, b, \rho} & \frac{1}{2} (\|\mathbf{w}\|^2 + b^2) - \nu\rho + \\ & \frac{1}{l} \left(\mu \sum_{i \in I_2} \xi_i + \frac{1}{2} \sum_{i \in I_1} \xi_i^2 \right) \\ \text{s.t.} & y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq \rho - \xi_i \\ & \xi_i \geq 0, \nu \in (0, 1], \rho \geq 0 \end{aligned} \quad (14)$$

其中, \mathbf{w} 为 d 维列向量; $\nu \in (0, 1]$ 用来控制模型复杂性与训练误差之间的平衡, 为控制支持向量的参数; ρ 是给定的参数; ξ_i ($i = 1, \dots, l$) 为松弛变量, 保证了约束的满足; I_1 是表示松弛变量落在 $0 < |\xi_i| \leq \varepsilon_\mu$ 区间内的样本集; I_2 是表示松弛变量落在 $\varepsilon_\mu < |\xi_i|$ 区间内的样本集.

为求上述 FRW ν -SVM 的对偶问题, 定义 Lagrange 函数如下

$$\begin{aligned} L(\mathbf{w}, b, \boldsymbol{\alpha}, \beta, \boldsymbol{\xi}, \rho, \delta) &= \frac{1}{2} (\|\mathbf{w}\|^2 + b^2) - \\ & \nu\rho + \frac{1}{l} \left(\sum_{i \in I_2} \mu \xi_i + \sum_{i \in I_1} \frac{1}{2} \xi_i^2 \right) - \\ & \sum_{i=1}^l (\alpha_i (y_i (\mathbf{w} \cdot \mathbf{x}_i + b) - \rho + \xi_i) + \beta_i \xi_i) - \delta\rho \end{aligned} \quad (15)$$

式中, 乘子 $\alpha_i, \delta, \beta_i \geq 0$. 分别对 $\mathbf{w}, b, \rho, \boldsymbol{\xi}$ 计算偏导, 即 $\nabla_{\mathbf{w}} L = 0, \nabla_b L = 0, \nabla_{\boldsymbol{\xi}} L = 0, \nabla_{\rho} L = 0$. 其中 $\nabla_{\mathbf{w}} L = 0$ 和 $\nabla_b L = 0$ 时, 可得

$$\begin{cases} \mathbf{w} = \sum_{i=1}^l \alpha_i y_i \mathbf{x}_i \\ \sum_{i=1}^l \alpha_i y_i = -b \\ \alpha_i + \beta_i = \frac{1}{l} (\mu + \xi_i) \\ \sum_{i=1}^l \alpha_i - \delta = \nu \end{cases} \quad (16)$$

将上述极值条件代入式 (15), 并对 $\boldsymbol{\alpha}$ 求极大, 以求得优化问题 (14) 的对偶问题如下

$$\begin{aligned} \max_{\boldsymbol{\alpha}} & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j (K(\mathbf{x}_i, \mathbf{x}_j) + 1) - \frac{l}{2} \sum_{i=1}^l \alpha_i^2 \\ \text{s.t.} & \nu \leq \sum_{i=1}^l \alpha_i \\ & 0 \leq \alpha_i \leq \min \left(\frac{1}{l}, \frac{\mu}{l} \right) \end{aligned} \quad (17)$$

FRW ν -SVM 的输出

$$f(\mathbf{x}) = \text{sgn} \left(\sum_{i=1}^l y_i \alpha_i \prod_{j=1}^l \psi \left(\frac{x_j - x_{ij}}{\alpha_i} + 1 \right) \right) \quad (18)$$

其中, x_j 表示输入向量 \mathbf{x} 的第 j 个分量, x_{ij} 表示第 i 个训练样本的第 j 个分量.

从式 (18) 可以发现, FRW ν -SVC 具有更简洁的对偶形式, 简化了计算; 从 FRW ν -SVC 的输出表达式可以发现, 其中不含有参数 b , 因此避免了求参数 b .

用获得的最佳分类模型, 就可以对样本进行分类分析和识别.

2 基于 FRW ν -SVC 的诊断算法

FRW ν -SVM 的训练涉及到 ν , μ 和核函数常量系数 a 等参数, 本文只考虑 $\mu = 1$ 的情况. 利用粒子群算法对 FRW ν -SVM 模型参数进行辨识, 获取最优参数组合 (ν, a) . 由此, 基于 FRW ν -SVM 模型的分类诊断步骤如下:

步骤 1. 对多维样本数据进行归一化处理, 形成样本集.

步骤 2. 对样本时序 y_i ($i = 1, 2, \dots, l$) 的隶属度进行多尺度小波变换, 选择与原始时序相吻合的小波函数, 粗估尺度变化范围.

步骤 3. 按照式 (5) 构造小波函数 $K(\mathbf{x}, \mathbf{x}')$, 建立形如式 (14) 的最优化目标函数.

步骤 4. 利用粒子群算法对参数 (ν, a) 进行辨识, 利用凸二次规划求解优化目标函数问题并判断 KKT 条件, 得到参数 α .

步骤 5. 对新的输入 \mathbf{x} , 根据式 (18) 计算决策值.

3 应用实例

过去, 汽车自动生产线的大多数的故障数据难以通过测量得到. 在实际工业中, 设计师常依靠经验来猜测故障发生的原因, 然后通过大量的实验来验证这些猜测. 这种方法既耗时又不经济, 同时也不可能对系统的每一个组件进行测试. 因此研究出一种系统化、简单且有效的方法来确定故障发生的原因, 并提前排除这些危险因素以达到提高整个系统的可靠性是十分必要的.

以某汽车厂装配生产线故障诊断为例, 影响装配的因素有很多, 有些是数值型变量, 有些是语言型变量. 对于数值型变量, 可以取产生装配线故障的机加工设备: B (齿轮传动 1)、C (齿轮传动 2)、E (传动轴)、G (驱动电机 1) 和 H (驱动电机 2) 的振动信号作为输入征兆. 对于语言型变量 A (辅助工人数)、D (产品需求量) 和 F (上游加工延时), 利用模糊综合评价法获取语言型因素的影响权重, 作为影响因素输入集. 具体方法如下: 建立因素集和评语集, 根据因素集中的评价因素对汽车装配影响程度的不同, 分别赋予不同的权重, 产生权重集, 权重集是因素集上的一个模糊子集, 各因素权重的分配采用层次分析法和调查分析法确定. 这样获取的权重集中某因素对汽车装配影响程度的数据就是数值型了. 该数据可以直接作为支持向量机训练集中的训练数据. 表 1 给出了经过归一化处理的该汽车装配线的模式样本集.

利用 Matlab 7.1 编制相应的模型训练和诊断与预测程序, 运行环境为 Core (TM) 2 CPU 1.80 GHz,

表 1 故障诊断样本集

Table 1 Sample set of fault diagnosis

| 序号 | A | B | C | D | E | F | G | H | 隶属度 |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1 | 0.66 | 0.0154 | 0.0021 | 0.29 | 0.0603 | 0.12 | 0.0289 | 0.0501 | 0.87(+) |
| 2 | 0.98 | 0.0305 | 0.0508 | 0.49 | 0.0485 | 0.35 | 0.0663 | 0.0311 | 0.65(+) |
| 3 | 0.12 | 0.0744 | 0.0159 | 0.69 | 0.0618 | 0.15 | 0.0939 | 0.0068 | 0.74(-) |
| 4 | 0.49 | 0.0015 | 0.0928 | 0.65 | 0.0973 | 0.92 | 0.0208 | 0.0086 | 0.79(-) |
| 5 | 0.82 | 0.0008 | 0.0731 | 0.98 | 0.0493 | 0.25 | 0.0181 | 0.0013 | 0.58(-) |
| 6 | 0.65 | 0.0708 | 0.0548 | 0.55 | 0.0711 | 0.76 | 0.0692 | 0.0621 | 0.55(-) |
| 7 | 0.34 | 0.0901 | 0.0656 | 0.44 | 0.0009 | 0.66 | 0.0608 | 0.0065 | 0.85(-) |
| 8 | 0.97 | 0.0089 | 0.0324 | 0.19 | 0.0623 | 0.33 | 0.0453 | 0.0185 | 0.65(+) |
| 9 | 0.11 | 0.0387 | 0.0049 | 0.63 | 0.0505 | 0.38 | 0.0877 | 0.0214 | 0.91(+) |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| 56 | 0.68 | 0.0756 | 0.0074 | 0.83 | 0.0614 | 0.86 | 0.0171 | 0.0079 | 0.71(-) |
| 57 | 0.94 | 0.0992 | 0.0007 | 0.67 | 0.0295 | 0.73 | 0.0699 | 0.0084 | 0.76(+) |
| 58 | 0.59 | 0.0275 | 0.0022 | 0.94 | 0.0561 | 0.28 | 0.0751 | 0.0016 | 0.52(-) |
| 59 | 0.29 | 0.0784 | 0.0016 | 0.91 | 0.0955 | 0.51 | 0.0234 | 0.0537 | 0.67(+) |
| 60 | 0.53 | 0.0548 | 0.0056 | 0.58 | 0.0287 | 0.45 | 0.0466 | 0.0936 | 0.75(+) |

表 2 最近 10 个模式点的诊断结果

Table 2 The diagnosis results from the latest ten pattern points

| 序号 | 实际隶属度 | 实际模式 | F ν -SVC | 预测模式 | FW ν -SVC | 预测模式 | FRW ν -SVC | 预测模式 |
|----|-------|------|--------------|------|---------------|------|----------------|------|
| 1 | 0.51 | -1 | 0.4936 | 1 | 0.5075 | -1 | 0.5136 | -1 |
| 2 | 0.77 | 1 | 0.7832 | 1 | 0.7736 | 1 | 0.7753 | 1 |
| 3 | 0.82 | -1 | 0.8328 | -1 | 0.8235 | -1 | 0.8217 | -1 |
| 4 | 0.63 | -1 | 0.6419 | -1 | 0.6401 | -1 | 0.6323 | -1 |
| 5 | 0.57 | 1 | 0.4900 | -1 | 0.4987 | -1 | 0.5800 | 1 |
| 6 | 0.71 | -1 | 0.8044 | -1 | 0.7363 | -1 | 0.7223 | -1 |
| 7 | 0.76 | 1 | 0.7577 | 1 | 0.7561 | 1 | 0.7607 | 1 |
| 8 | 0.52 | -1 | 0.4837 | 1 | 0.4984 | 1 | 0.5357 | -1 |
| 9 | 0.67 | 1 | 0.6812 | 1 | 0.6788 | 1 | 0.6735 | 1 |
| 10 | 0.75 | 1 | 0.7340 | 1 | 0.7476 | 1 | 0.7468 | 1 |

内存 1.00 G 的微机. 利用小波理论获取在不同尺度范围内故障模式的小波变换. 本文选择的小波函数有 Morlet, Mexh 和 Gaussian. 粗估计在尺度 $a \in (0, 2]$ 范围内, Morlet 母小波要优于 Mexican 母小波, 更能与决策模式相吻合. 由此, FRW ν -SVC 的核函数选用 Morlet 小波核, 二个参数的变化范围设置如下: $\nu \in [0.001, 1]$, $a \in (0, 2]$. 粒子群参数设定为: $w = 0.9$, $c_1 = c_2 = 2$. 利用粒子群算法获取最优参数组合为: $\nu = 0.9$ 和 $a = 0.4$.

从表 2 中可以发现, 对于靠近模糊隶属度均值的模式, FRW ν -SVC 比 FW ν -SVC 和 F ν -SVC 给出的诊断精度高; 对于远离均值的模式, FRW ν -SVC 给出的诊断精度也比 FW ν -SVC 和 F ν -SVC 高. 这体现出分段损失函数的优越性, 同时也体现出 Laplace 损失函数和高斯损失函数的抑制作用. 将松弛变量落在 $(-\infty, -\varepsilon_\mu) \cup [-\varepsilon_\mu, -\varepsilon) \cup [\varepsilon, \varepsilon_\mu) \cup [\varepsilon_\mu, +\infty)$ 范围内的噪声计算进模型, 故在整体上提高了模糊小波支持向量机分类机的抗噪能力.

最近 10 个模式点的诊断结果如表 2 所示. 其中 F ν -SVC 模型给出的最大诊断误差为 13.3 % (第 6 组), 最小诊断误差为 0.3 % (第 7 组), FW ν -SVC 模型给出的最大诊断误差为 12.5 % (第 5 组), 最小诊断误差为 0.32 % (第 10 组), FRW ν -SVC 模型给出的最大诊断误差为 0.6 % (第 2 组), 最小诊断误差为 0.09 % (第 7 组). 考虑到工厂装配线的现实状况, 数据的收集和估计不可避免地存在着偏差, 因此诊断所得结果还是非常满意的. 由此可见, 小波支持向量分类机对于具有小样本、非线性、不平稳等多维特征时序具有强大的高维映射能力.

为了便于对三种模型的诊断精度进行分析, 表 3 给出了基于一类错误率 (将负类错判为正类的百分比), 二类错误率 (将正类错判为负类的百分比) 和

错误率 (测试误差为 0.03) 的分类效果分析. 可以明显看出, 对于本算例的故障模式, 本文给出的具有鲁棒性的 FRW ν -SVC 的分类指标值均好于标准 F ν -SVC 和 FW ν -SVC 模型分类指标值. 这说明了基于 FRW ν -SVC 模型的诊断方法极具理论意义, 是可行而且有效的, 易于在工程实际中推广应用.

表 3 故障诊断精度分析 (%)

Table 3 Precision analysis of fault diagnosis (%)

| Model | 一类错误率 | 二类错误率 | 错误率 |
|----------------|-------|-------|-----|
| F ν -SVC | 20 | 10 | 20 |
| FW ν -SVC | 10 | 10 | 10 |
| FRW ν -SVC | 0 | 0 | 0 |

4 结论

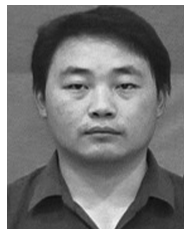
复杂制造系统故障模式受很多因素影响, 有些是数值型因素, 有些是语言型因素, 而且存在小样本问题, 利用传统方法很难建立确切的非线性故障诊断模型. 本文将小波理论、模糊理论和 SVC 三者有机结合, 构造基于小波基函数的小波核函数, 设计一个多段式损失函数, 由此得到一种适合于处理小样本、模糊不确定信息且能够抑制混合噪音的小波核支持向量分类机, 即模糊鲁棒小波 ν -支持向量分类机 (FRW ν -SVC). 实验结果表明: FRW ν -SVC 模型可以有效抑制各种混合噪音, 提高分类精度, 且具有简洁的对偶形式, 不要求参数 b , 决策函数的期望值部分可直接由样本数据和小波核函数给出.

References

- 1 Chai Tian-You, Ding Jin-Liang, Wang Hong, Su Chun-Yi. Hybrid intelligent optimal control method for operation of complex industrial processes. *Acta Automatica Sinica*, 2008,

- 34(5): 505–515
(柴天佑, 丁进良, 王宏, 苏春翌. 复杂工业过程运行的混合智能优化控制方法. 自动化学报, 2008, 34(5): 505–515)
- 2 Chen Mao-Yin, Zhou Dong-Hua. Fault detection and isolation in a class of nonlinear systems. *Acta Automatica Sinica*, 2004, 30(6): 1011–1016
(陈茂银, 周东华. 一类非线性系统的故障诊断. 自动化学报, 2004, 30(6): 1011–1016)
 - 3 Lv Chen, Wang Gui-Zeng, Zhang Ze-Yu. PWM VLSI neural network for fault diagnosis. *Acta Automatica Sinica*, 2005, 31(2): 195–201
(吕琛, 王桂增, 张泽宇. PWM 型 VLSI 神经网络在故障诊断中的应用. 自动化学报, 2005, 31(2): 195–201)
 - 4 Vapnik V N. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 1999, 10(5): 988–999
 - 5 Cao Li-Juan, Wang Xiao-Ming. *Support Vector Machine Based Methods for Financial and Engineering Problems*. Shanghai: Shanghai University of Finance and Economics Press, 2007
(曹丽娟, 王小明. 金融工程的支持向量机方法. 上海: 上海财经大学出版社, 2007)
 - 6 Zhang Li, Xi Yu-Geng. Identification of separable variable nonlinear dynamical system based on SVMs. *Acta Automatica Sinica*, 2005, 31(6): 965–969
(张莉, 席裕庚. 基于支持向量机的可分离非线性动态系统辨识. 自动化学报, 2005, 31(6): 965–969)
 - 7 Cheng Jun-Sheng, Yu De-Jie, Yang Yu. Fault diagnosis approach based on intrinsic mode singular value decomposition and support vector machines. *Acta Automatica Sinica*, 2006, 32(3): 475–480
(程军圣, 于德介, 杨宇. 基于内禀模态奇异值分解和支持向量机的故障诊断方法. 自动化学报, 2006, 32(3): 475–480)
 - 8 Widodo A, Yang B S. Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, 2007, 21(6): 2560–2574
 - 9 Francis E H T, Cao L J. Application of support vector machines in financial time series forecasting. *Omega*, 2001, 29(4): 309–317
 - 10 Wu Q. The forecasting model based on wavelet ν -support vector machine. *Expert Systems with Applications*, 2009, 36(4): 7604–7610
 - 11 Wu Qi, Yan Hong-Sen, Wang Bin. The product sales forecasting model based on robust wavelet ν -support vector machine. *Acta Automatica Sinica*, 2009, 35(7): to be published
(吴奇, 严洪森, 王斌. 基于鲁棒小波 ν -支持向量机的产品销售预测模型. 自动化学报, 2009, 35(7): 1027–1032, 待发表)
 - 12 Yan H S, Xu D. An approach to estimating product design time based on fuzzy ν -support vector machine. *IEEE Transactions on Neural Networks*, 2007, 18(3): 721–731
 - 13 Vapnik V N. *The Nature of Statistical Learning Theory*. New York: Springer, 2000
 - 14 Widodo A, Yang B S. Wavelet support vector machine for induction machine fault diagnosis based on transient current signal. *Expert Systems with Applications: An International Journal*, 2008, 35(1-2): 307–316
 - 15 Francis E H T, Cao L J. Modified support vector machines in financial time series forecasting. *Neurocomputing*, 2002, 48(1-4): 847–861
 - 16 Avci D, Varol A. An expert diagnosis system for classification of human parasite eggs based on multi-class SVM. *Expert Systems with Applications: An International Journal*, 2009, 36(1): 43–48

- 17 Yuan S F, Chu F L. Fault diagnosis based on support vector machines with parameter optimization by artificial immunisation algorithm. *Mechanical Systems and Signal Processing*, 2007, 21(3): 1318–1330
- 18 Rossi F, Villab N. Support vector machine for functional data classification. *Neurocomputing*, 2006, 69(7-9): 730–742
- 19 Shim J, Hwang C, Nau S. Robust LS-SVM regression using fuzzy C-means clustering. In: *Proceedings of the 2nd International Conference*. Xi'an, China: Springer, 2006. 157–166
- 20 Zhang L, Zhou W D, Jiao L C. Wavelet support vector machine. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2004, 34(1): 34–39



吴奇 东南大学博士后. 主要研究方向为计算机视觉, 生物计算, 模式识别和计算机集成制造系统. 本文通信作者.

E-mail: wuqi7812@163.com

(WU Qi Postdoctor at Southeast University. His research interest covers computer vision, biologic computing, pattern recognition, and computer integrated manufacturing system. Corresponding author of this paper.)



刘静 上海海事大学信息工程学院讲师. 主要研究方向为量子信息进化算法和故障诊断.

E-mail: jingliu@cie.shmtu.edu.cn

(LIU Jing Lecturer at the College of Information Engineering, Shanghai Maritime University. Her research interest covers quantum information, evolutionary algorithm, and fault diagnosis.)



熊福力 东南大学自动化学院博士研究生. 主要研究方向为并行工程, 智能控制, 生产计划与调度和计算机仿真.

E-mail: wuqivr@hotmail.com

(XIONG Fu-Li Ph.D. candidate at the School of Automation, Southeast University. His research interest covers concurrent engineering, intelligent control, production planning and scheduling, and computer emulation.)



刘晓军 东南大学机械工程学院博士研究生. 主要研究方向为先进制造技术, CAD/CAPP/CAM, 特征识别和刀具路径优化.

E-mail: liuxiaojunad@gmail.com

(LIU Xiao-Jun Ph.D. candidate at the School of Mechanical Engineering, Southeast University. His research interest covers advanced manufacturing technology, CAD/CAPP/CAM, feature recognition, and tool path optimization.)