

**课程设计（论文）**

题 目 基于SVM的机械

故障模式识别

专 业 机器人工程

学　　 号 171320330

学 生 姓 名 张文栋

课 程 名 称 模式识别

课 程 教 师 赵明航

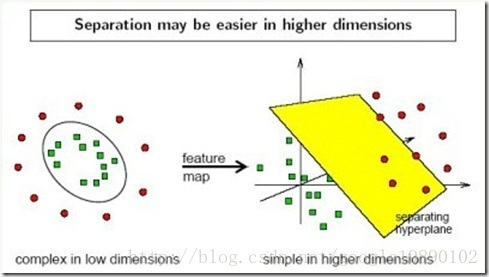
1. 支持向量机的介绍

支持向量机（Support Vector Machine），即SVM，最早是由Vladimir N. Vapnik 和 Alexey Ya. Chervonenkis 在1963年提出的一种机器学习的算法。简单来讲，支持向量机的作用就是对于映射在n维空间上的特征点进行分类，而实现的方法是确定一个超平面，依照特征点的性质将空间划分为几块，每一块代表一个类别。SVM是一种有监督学习的方法。

新版本的MATLAB中增加了classification的工具箱，也内嵌了关于SVM的一些库函数，特别是对于求解二分类问题非常方便。

1. 非线性支撑向量机

在处理非线性问题时，可以通过将分线性问题转化成线性问题，并通过已经构建的线性支持向量机来处理。如下图所示：



通过一种映射可以将输入空间转换到对应的特征空间，体现在特征空间中的是对应的线性问题。核技巧就可以完成这样的映射工作。

设是X输入空间(欧式空间的子集或离散集合)，又H设为特征空间(希尔伯特空间)，如果存在一个从X到H的映射

使得对于所有的x,y,函数K(x,y)满足条件：

则称K(x,y)为核函数，为映射函数。

常用的核函数有多项式核函数（Polynomial Kernel Function）、高斯核函数（Gaussian Kernel Function）等。

1. LibSVM

LIBSVM是一个由台湾大学林智仁(Lin Chih-Jen)教授等开发的SVM模式识别与回归的软件包。LibSVM支持MATLAB、C++等多种语言。在本次课程论文中，我们主要采用了LibSVM中的训练和预测功能。

模型训练函数：

model= libsvmtrain(training\_label\_vector,training\_instance\_matrix[, 'libsvm\_options']);

预测函数：

[predicted\_label, accuracy, decision\_values/prob\_estimates]

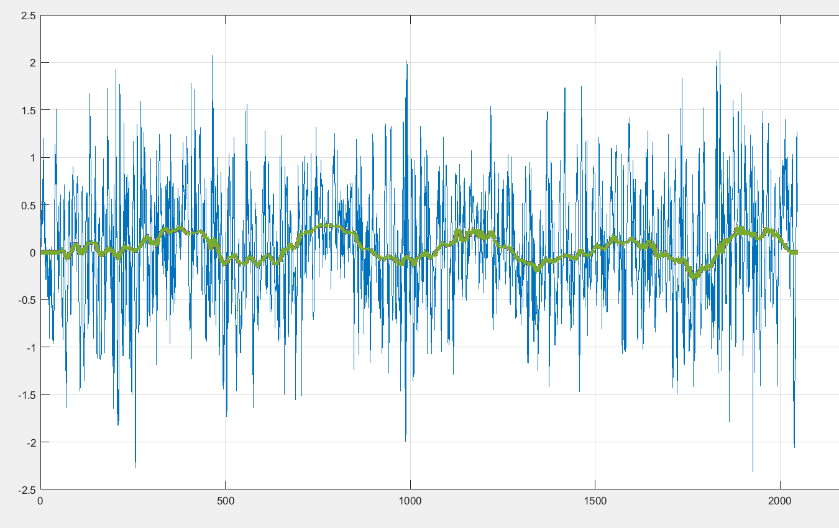
　　　　= libsvmpredict(testing\_label\_vector, testing\_instance\_matrix, model [, 'libsvm\_options'])

1. 基于SVM的故障识别实现过程
   1. 特征提取

针对给定的训练数据，我们对其进行特征的提取和处理。由于数据的形式是一段采样的振动信号，故准备采用以下几个方面进行数据的特征提取，将每一段训练数据样本转化成几个特征值组成的多维向量。首先使用importdata函数进行数据的导入，将每一组数据储存成矩阵的形式，其中矩阵的每一列为一个样本。

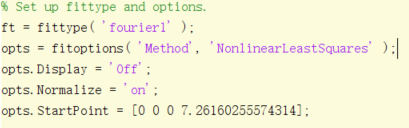
* + 1. 傅里叶拟合曲线

先对每一个样本依次进行中值滤波处理，取阶数为120，采用medfilt1函数进行处理，得到原始数据与滤波后的数据作图如下：

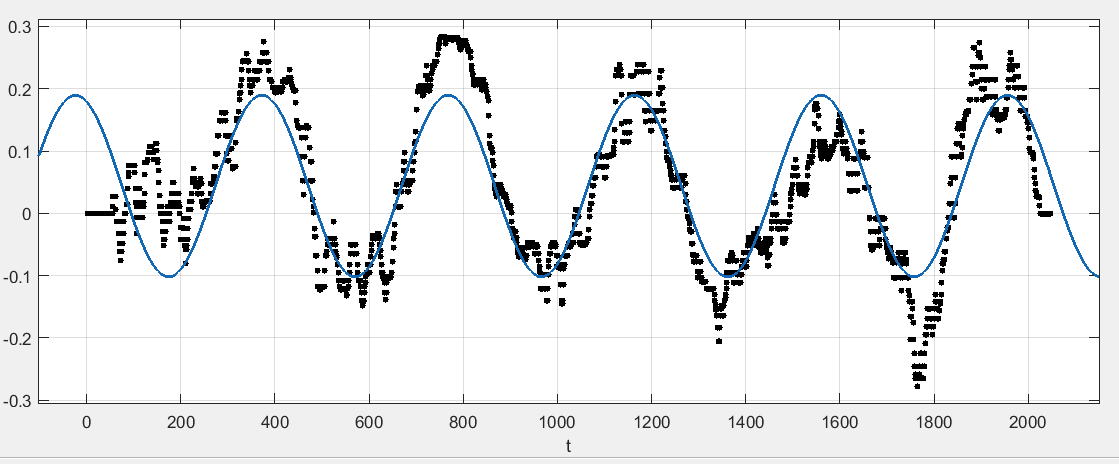


绿色为中值滤波后波形，蓝色为原始振荡。

之后采用fit函数进行拟合，拟合类型为Fourier1，拟合模型设置如下：



以第一类第一个样本为例，拟合结果如图所示：

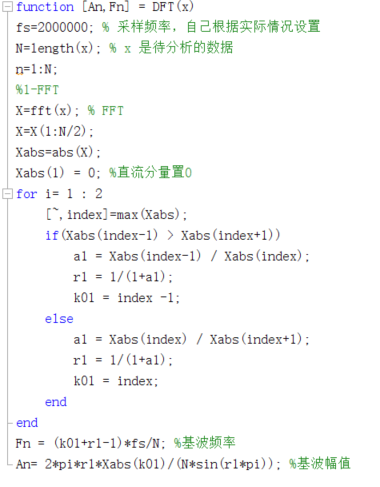


拟合曲线设为f(x) = a0 + a1\*cos(x\*w) + b1\*sin(x\*w)，可得参数：a0,a1,w,b1

取幅度A=,角频率w，初值a0为一组特征值。

* + 1. DFT

对每一组数据进行离散傅里叶变换。构造DFT函数，使输入为离散随机信号时输出为基波频率和基波幅值。



取基波频率和基波幅值为一组特征值。

* + 1. 统计量数字特征

采用MATLAB内嵌的var和mean函数，对每一个样本的每一个采样点求取均值和方差。并以均值为基准采用for循环统计过均值的次数，亦即粗略估计振荡次数。选取均值，方差和振荡次数为一组特征。

* 1. 创建训练样本、测试样本和制定标签

将上述特征整合为一个八维列向量的形式，作为一个样本的总特征来进行模型训练。采用线性代数中矩阵运算的性质对18×500个数据进行标签设置。选取每一组数据中的前450个为训练样本，后50为测试样本。对每一个样本的特征进行归一化处理。

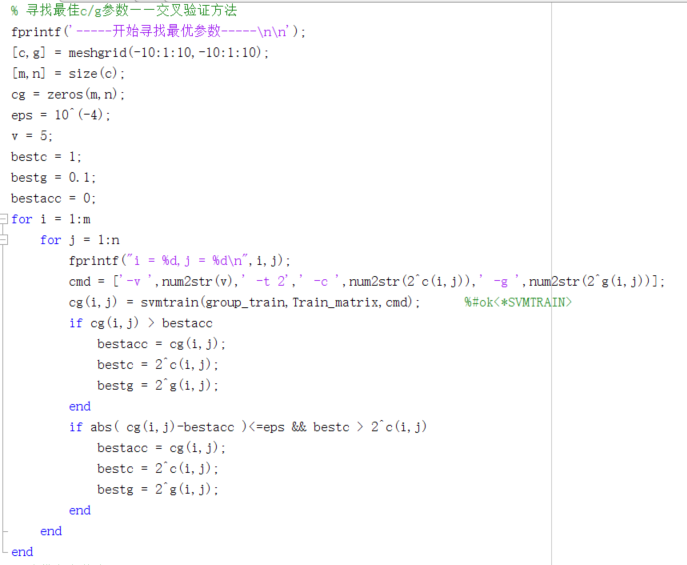
* 1. 使用RBF核函数创建SVM模型并进行训练

采用svmtrain函数对上述的样本进行模型训练，采用RBF核函数，设置svmtrain函数的第三个参数为

cmd = ['-v ',num2str(v),' -t 2',' -c ',num2str(2^c),' -g ',num2str(2^g)];

选取c/g的取值范围为meshgrid(-10:1:10,-10:1:10)，采用交叉验证的方法寻找近似的最优c/g参数。

具体实现如下：



采用最优参数c/g训练出最终模型。

* 1. 模型预测及结果预测

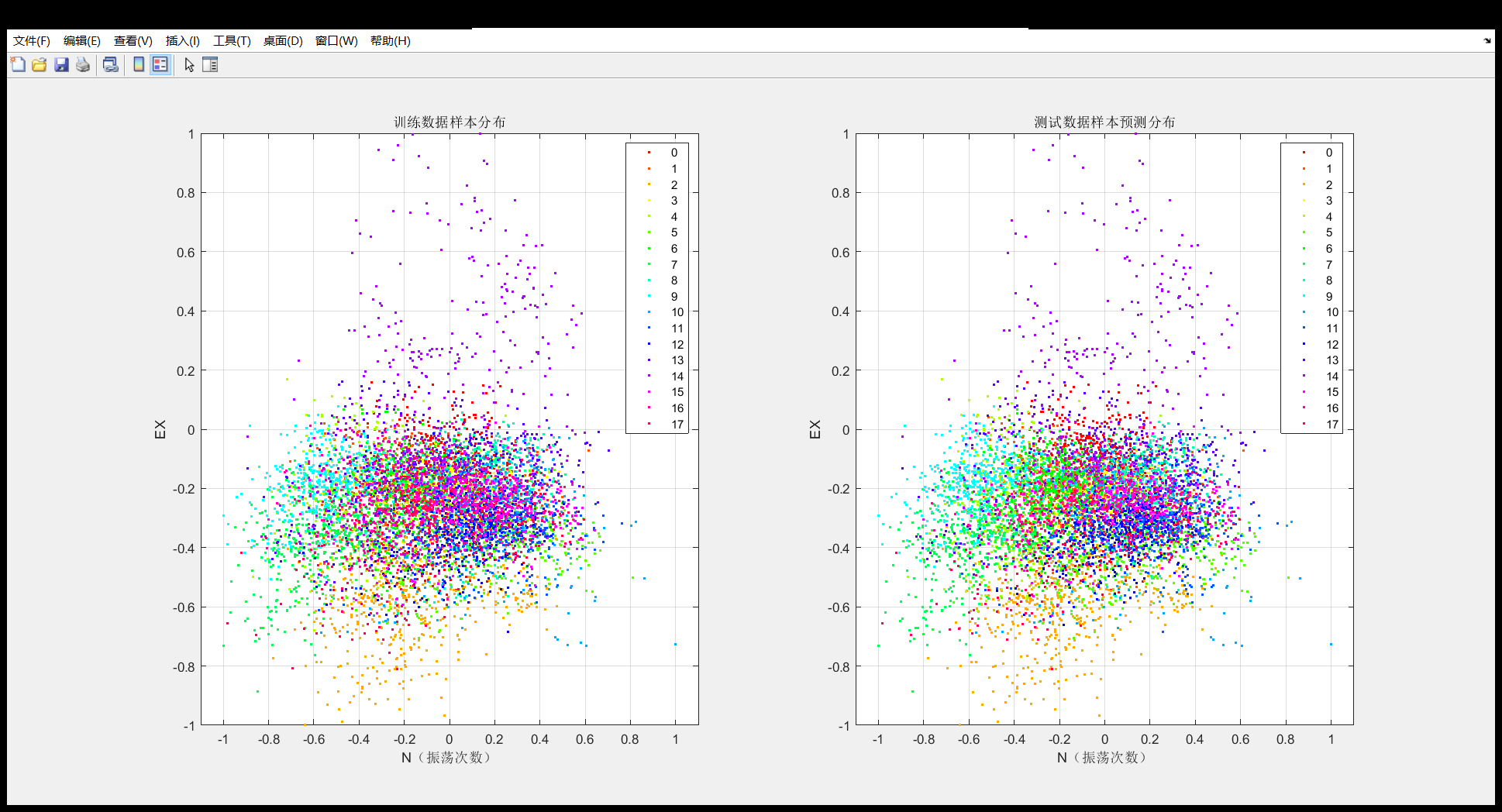
采用svmpredict函数进行模型预测，经检验，获得两组较好的c/g取值：在选取参数c=181.0193和g=8时，测得训练集准确率为97.1975%，测试机准确率约为41.8889%；

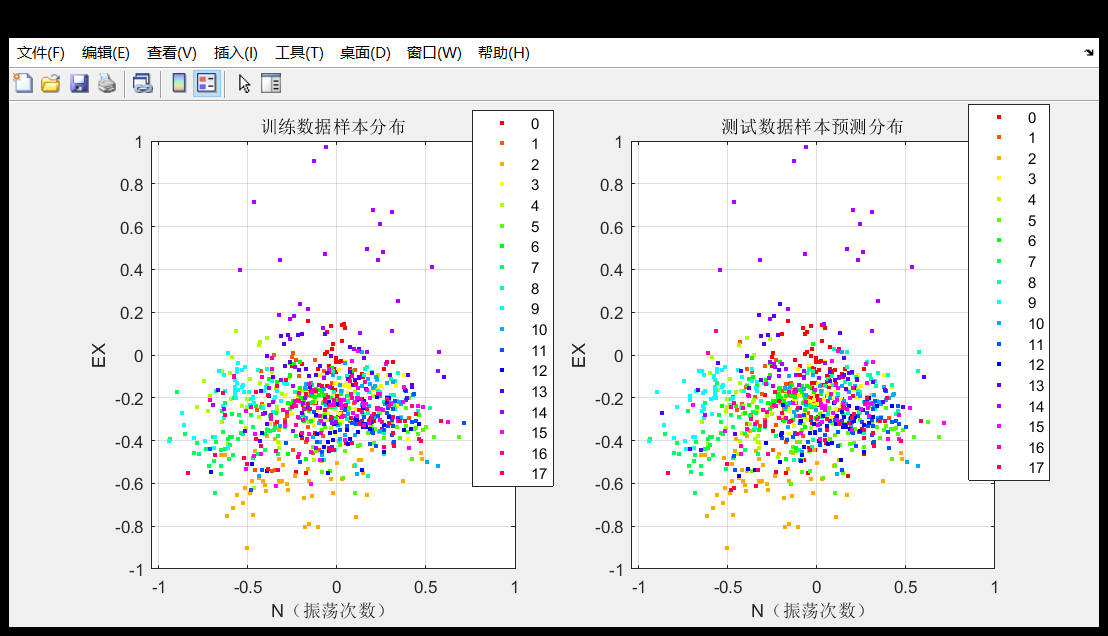


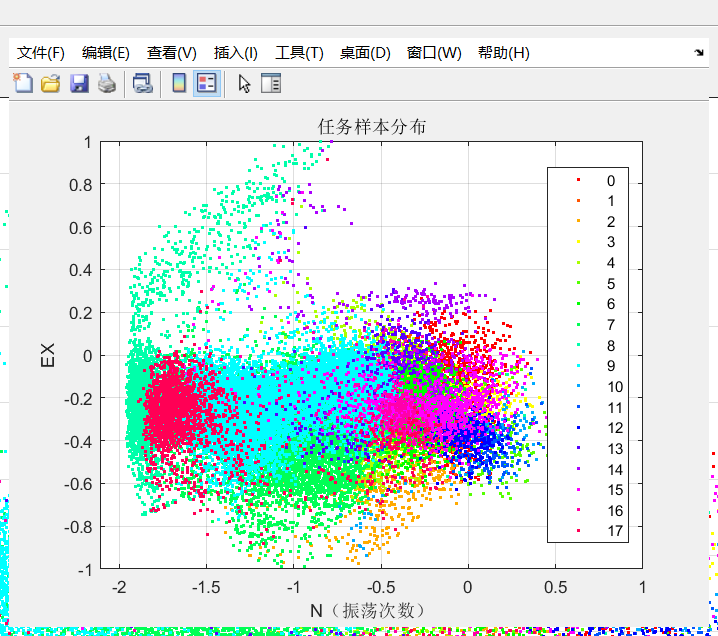
当c=64,g=2时，测得训练集准确率为64.5802%，测试机准确率约为44.3333%；



采用第二组c/g对测试数据样本进行预测，以振荡次数和均值作为XY坐标做出散点图，结果如下：







注：

详细代码实现及分类结果见附件。