

基于 SVR 算法的零部件尺寸对应力的回归预测以及振动信号的时序预测

韦炯光 S2302W0363

摘要：为了零件轻量化设计和对设备展开振动故障监测，本文将对实验室中的斜轴式柱塞泵进行研究，基于支持向量回归 SVR 算法对关键零部件的两个相关尺寸对应力的影响进行回归预测，同时将检测到的斜轴式柱塞泵的振动位移信号进行时序预测。首先详细地介绍了支持向量机算法 SVM(SVR)的基本原理，然后探讨了 SVR 模型算法的构建过程和模型评价指标，使用 MATLAB 构建 SVR 模型、处理训练与测试数据、选择参数，并通过对针相关尺寸对应力的影响和振动信号的实际问题进行模型应用，最后对模型的性能进行了全面评估。实验结果表明，SVR 模型能够有效预测零件尺寸对应力的影响以及振动信号，具有优良的预测性能。

关键词：液压泵；支持向量回归；应力；振动；预测

Regression Prediction of Component Size Stress and Temporal Prediction of Vibration Signals Based on SVR Algorithm

Abstract: In order to reduce the weight of equipment parts and carry out vibration fault monitoring, this paper will study the inclined shaft piston pump in the laboratory. Based on the support vector regression SVR algorithm, the influence of the two related dimensions of the key parts on the stress is predicted by regression. At the same time, the vibration displacement of the detected inclined shaft piston pump is predicted with time signal. Firstly, the basic principle of support vector machine algorithm SVM (SVR) is introduced in detail. Then, the construction process and model evaluation index of SVR model algorithm are discussed. The SVR model is constructed by MATLAB, parameter selection, training and test data processing. The influence of two related dimensions on stress and vibration signal is applied, and the performance of the model is evaluated comprehensively. The experimental results show that the SVR model can effectively predict the influence of part size on stress and vibration signal, and has excellent prediction performance.

Keywords: hydraulic pump ; support vector regression ; stress ; vibration ; prediction

1 研究背景及意义

本次的课程作业，以课题组所开展的液压泵项目进行不同尺寸对零件应力的回归预测和时间位移振动信号的时序预测。

研究高转速、高压强工况下的斜轴式柱塞泵是当前液压传动技术研究的热点和难点之一，这对于推动我国液压传动技术的发展和提高制造业的核心竞争力，具有十分重要的现实意义和科学价值^[1]。缸体是斜轴式柱塞泵中的关键零部件，作为关键摩擦副的重要组成部分，其尺寸大小直接影响着整机的强度、效率、轻量化等方面。零件尺寸是影响结构受力状态下应力分布的关键因素之一^[2]，不同的零件尺寸会导致相应不同应力以及应力分布不均匀等现象，进而影响结构的强度、刚度和稳定性^[3]。因此探究不同零件尺寸对相应结构受力状态下应力分布影响的回归预测具有重要的意义。

机械振动信号是研究机械系统运行状态的重要指标之一。通过对机械振动信号的预测与分析，可以实现对机械系统状态的准确判断和及时预警，从而提高机械设备的运行效率和可靠性^[4]。

随着人工智能技术的发展，各种工程优化方法和机器学习方法，如处理共轭梯度法 PCG、BP 算法、CNN 卷积神经网络与支持向量机算法 SVM(SVR)等逐渐应用于机械设备的尺寸优化和振动故障信号预测中。相较于其他机器学习算法，支持向量机算法 SVM(SVR)具有以下一些优势^{[5][6]}：①泛化能力强，在合适的参数设置下，能够较好地处理过拟合问题，对新数据具有较好的预测能力。②小样本有效性，对于小样本数据也能构建出较为可靠的模型；③全局最优解，往往能够找到一个全局最优的分类或回归超平面；④对非线性问题的处理，通过核函数可以巧妙地将非线性问题转化为线性可分问题来处理。⑤鲁棒性较好，对噪声和异常值有一定的容忍度。

因此本文将选用支持向量机算法 SVM(SVR)对液压泵进行不同尺寸对零件应力的回归预测和时间位移振动信号的时序预测，并为后续尺寸的优化和开展的故障诊断提供基础，以提高旋转机械的运行效率、降低生产成本。

2 算法介绍

2.1 支持向量回归 SVR 的原理

支持向量回归 (Support Vector Regression, SVR) 是一种基于支持向量机 (Support Vector Machine, SVM) 的回归算法，用于解决回归问题。与传统的回归算法不同，SVR 的目标是通过构建一个预测函数，使得预测值与实际值的差异最小化。SVR 在解决回归问题时的主要思想是通过寻找一个决策边界，使得预测值与实际值之间的差异尽可能小，并且在一定的容忍度内^[7]。

SVR 的核心思想是通过优化问题来找到一个最佳的超平面，使得模型的预测误差最小。

2.2 SVR 模型算法过程

SVR 通过相应的核函数将训练数据集 (包含输入值和输出值) 映射到高维特征空间^{[8][9]}, 在这个特征空间中寻找结构风险最小化的最佳线性回归超平面, 以反映输入数据和输出数据映射前的非线性关系对于给定的训练集 $T = \{ (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n) \}$, 其中 x_i 、 y_i 分别为第 i 个输入向量及其对应的输出值, n 为样本数量, SVR 函数可以写成如下形式:

$$f(x) = \omega^T \varphi(x) + b \quad (1)$$

式中, ω 为权重系数向量; $\varphi(x)$ 为输入空间 x 映射得到的高维特征空间; b 为偏置。

在 SVR 函数两侧生成容忍区间, 两侧间距均为 ε (不敏感参数), 所有落入容忍区间内的样本 (非支持向量) 不计损失, 仅区间外的样本 (支持向量) 才会对函数模型产生影响^[10]。此时寻找最优超平面可以转化为如下的凸优化问题:

$$\min \frac{1}{2} \|\omega\|^2 \quad (2)$$

$$|y_i - (\omega^T \varphi(x_i) + b)| \leq \varepsilon \quad (3)$$

为允许样本出现误差, 引入松弛变量 ξ 和 ξ^* , 使部分样本可以不在容忍区间内。在容忍区间上方时 $\xi_i > 0$, $\xi_i^* = 0$; 在容忍区间下方时 $\xi_i = 0$, $\xi_i^* > 0$ 。则上式可改写为:

$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ y_i - \omega^T \varphi(x_i) - b \leq \varepsilon + \xi_i \\ \omega^T \varphi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0, \xi_i^* \geq 0 \quad i = 1, 2, \dots, n \end{cases} \quad (4)$$

式中, C 为惩罚因子, 用于控制样本误差的惩罚力度。

为了解决带约束优化问题, 利用拉格朗日乘子法将目函数转化为对偶形式, 再利用 KKT 条件求解, 可将 SVR 函数改写成如下形式:

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \alpha_i) K(x_i, x) + b \quad (5)$$

其中, α_i^* 和 α_i 为拉格朗日乘数; $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$ 为核函数, 即为特征空间 $\varphi(x_i)$ 中 2 个向量 x_i 和 x_j 的内积。核函数可使非线性回归问题经过转换后变成近似线性问题, 从而大大减小了 SVR 的计算量^[11]。常见的核函数包括线性核函数、多项式核函数、拉普拉斯核函数、RBF 核函数等, 其中, RBF 核函数性能优越, 能够较好地解决非线性问题, 因此, 本文选取 RBF 核函数来构建 SVR 模型。

2.3 SVR 模型的评价指标

为了从多个角度对模型的预测结果进行评估和比较^{[12][13]}，引入决定系数 R^2 、平均误差MAE、均方误差MSE、均方根误差RMSE对模型预测性能进行量化。

(1) 决定系数 (R^2) 是衡量回归模型拟合优度的统计量，它表示回归模型对观测值的拟合程度，代表了模型中因变量可由自变量解释的百分比。 R^2 的取值范围在 0~1 之间， R^2 越大，说明回归模型对观测值的拟合程度越好。

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2} \quad (6)$$

式中， y_i 为实验值， \hat{y} 为预测值， \bar{y} 为实验值的平均值。

(2) 平均绝对误差值 (MAE) 是真实值与拟合值的平均值差值，用于评价真实值与拟合值的偏离程度，MAE 值越接近于 0，说明模型拟合越好，模型预测准确率越高。

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}| \quad (7)$$

式中， N 为样本的数量。

(3) 均方误差值 (MSE) 反映了模型的误差平方和均值，反映了模型的拟合值与真实值的差异程度，MSE 值越接近于 0，说明模型拟合越好。

$$MSE = \frac{1}{N} \sum_{i=1}^N (f(x_i) - y)^2 \quad (8)$$

(4) 均方根误差 (RMSE)，均方误差 MSE 的算术平方根，回归模型中最常用的评价模型指标，反映预测误差的波动，RMSE 值越接近 0，说明模型拟合越好。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (f(x_i) - y)^2} \quad (9)$$

3 设计实现过程与步骤

3.1 MATLAB 代码实现

根据支持向量回归算法的设计过程编写代码。在 MATLAB 中实现相关支持向量机预测主要是调用 MATLAB 自带的 SVR 工具箱。整体的代码设计实现过程如图 1 所示。

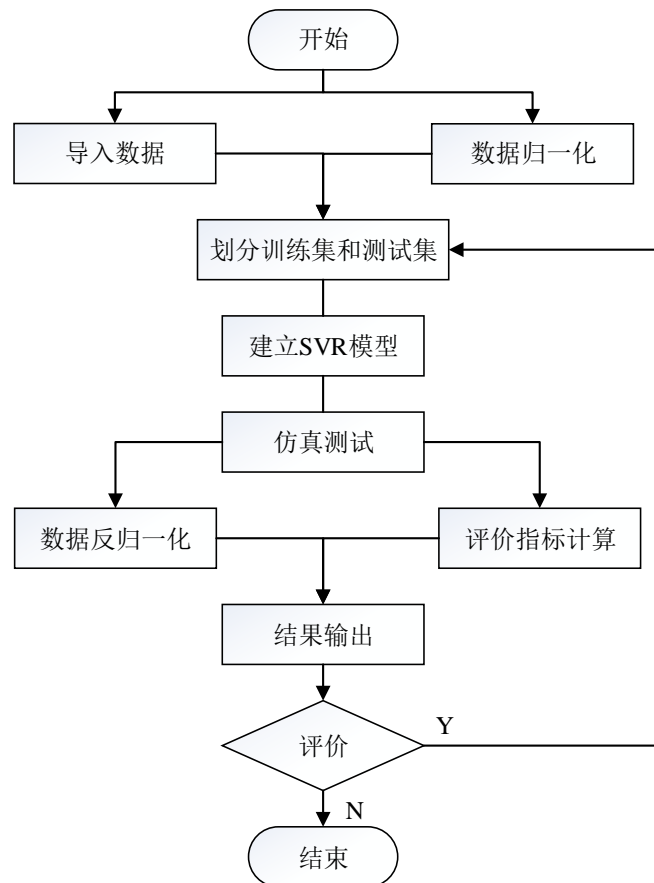


图 1 MATLAB 的 SVR 模型设计流程

3.2 MATLAB 代码设计具体步骤

使用 MATLAB 编写代码，以下是一些主要的设计步骤：

(1) 导入数据，并对数据进行归一化处理，将数据归一到同一量纲上，有利于提高精度。

`imp = readmatrix('svr.xlsx');` %使用 matlab 库中的 `readmatrix()`函数读取，同一路径下的对应名称的表格数据

`X = imp(:,1:end-1);` %将表格中的前两列分别作为两个输入 `x1`、`x2`

`Y = imp(:,end);` %将表格中的最后一列分别作为输出值 `Y`

`x = mapminmax(X', 0, 1);` %`mapminmax` 函数，将数据进行简单归一化处理，将其映射到 0 到 1 之间

`[y, psout] = mapminmax(Y', 0, 1);`

(2) 划分训练集和测试集

`num = size(imp,1);`

`sample = randperm(num);` %将数据样本随机排布，生成一个随机排列的索引向量 `sample`

`ratio = 0.7;` %设置 70%的训练集的占比

```

tr_num = floor(num*ratio); %计算训练集的数量
x_tr = x(:,sample(1: tr_num));%根据索引选取训练集和测试集的输入特征并转置
y_tr = y(sample(1: tr_num));
x_te = x(:,sample(tr_num+1: end));
y_te = y(sample(tr_num+1: end));

```

（3）使用 matlab 中的 fitrsvm()函数进行 SVR 模型的参数设定，以训练支持向量机回归模型。训练后，模型将学习如何根据输入特征来预测目标变量。

```

model=fitrsvm(x_tr,y_tr,"KernelFunction","rbf",..."Solver","ISDA",... "KernelScale","auto");

```

%使用给定的训练数据，采用 RBF 核函数用于计算 Gram 矩阵的核函数，采用 ISDA 优化，支持训练过程中的动态结构调整，并可以降低计算资源需求，减少过拟合风险，设定核尺度参数（KernelScale）

（4）进行仿真测试，使用函数 svmpredict（）进行预测

```

res1=predict(model,x_tr); %进行仿真测试，使用函数 svmpredict（）对训练集和测试集进行预测

```

```

res2 = predict(model, x_te );

```

%实际值

```

Y_tr = Y(sample(1: tr_num)); %获取训练集的实际输出值

```

```

Y_te = Y(sample(tr_num+1:end));

```

%预测值

```

pre1 = mapminmax('reverse', res1, psout); %将训练集和测试集上的预测值反归一化

```

```

pre2 = mapminmax('reverse', res2, psout);

```

（5）计算相关指标

%决定系数 R2

```

R1 = 1 - norm(Y_tr - pre1)^2 / norm(Y_tr - mean(Y_tr))^2;

```

```

R2 = 1 - norm(Y_te - pre2)^2 / norm(Y_te - mean(Y_te))^2;

```

%平方误差 MAE

```

MAE1 = mean(abs(Y_tr - pre1 ));

```

```

MAE2 = mean(abs(pre2 - Y_te ));

```

%均方根误差 RMSE

```

RMSE1 = sqrt(mean((pre1 - Y_tr).^2));

```

```

RMSE2 = sqrt(mean((pre2 - Y_te).^2));

```

4. 模型应用

4.1 零件多输入尺寸对输出应力的回归预测

使用 ANSYS 仿真模块对缸体在相应工况下进行力学分析，缸体的应力情况如图 2 所示。

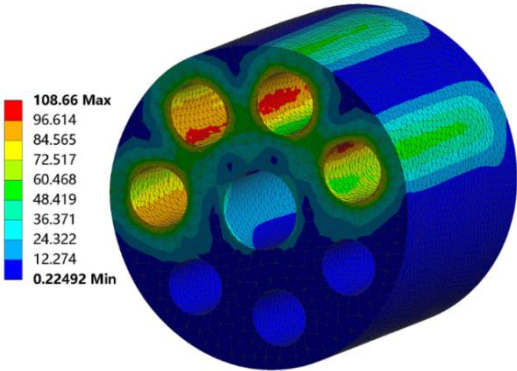


图 2 缸体的应力云图

对缸体的应力应变云图进行分析：缸体的薄弱区域为柱塞孔与缸体外径壁之间，因此缸体的关键尺寸在于缸体的外径与中心孔的直径。

可以设计系列的不同尺寸的试件，在相同外力作用下进行测试，比较不同尺寸试件的应力分布情况。得到不同尺寸下的应力情况，如图 3 所示。

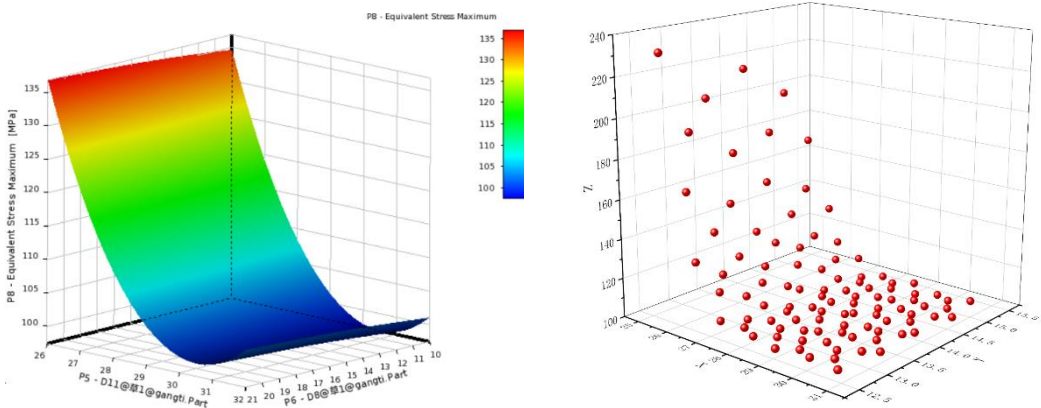


图 3 零件不同尺寸对应的应力影响

对得到的两个不同尺寸对应下的应力数据进行分析 and 处理，采用 100 组系列的数据点，其中输入 x_1 为缸体零件的外径，输入 x_2 为零件的内孔直径，输出 Y 为对应尺寸下所产生的应力大小， $[(x_1, x_2, Y)]$ 。可得到的数据集如表 1 所示（详细数据见文件夹内的表格）。

表 1 不同尺寸下的应力

(缸体外径)输入 x_1	(中心孔径)输入 x_2	(应力)输出 y
25.228	12.614	231.2276193
25.284	14.014	215.544597
25.34	13.314	205.2174236

.....
30.66	14.079625	101.7806011
30.716	13.379625	101.5411083
30.772	14.779625	101.5272827

将数据集文件导入 MATLAB 中，运行程序，得到的结果和评价指标如表 2 所示，预测效果如图 4 和图 5 所示。

表 2 训练集与测试集的评价指标

评价指标	训练集评价结果	测试集评价结果
决定系数 R^2	0.98763	0.97035
平均误差 MAE	1.6981	2.3076
均方根误差 RMSE	3.0016	4.9376

训练集的决定系数 R^2 为 0.98763，和测试集的决定系数 R^2 为 0.97035 较高，说明预测拟合的精度还是较高的。

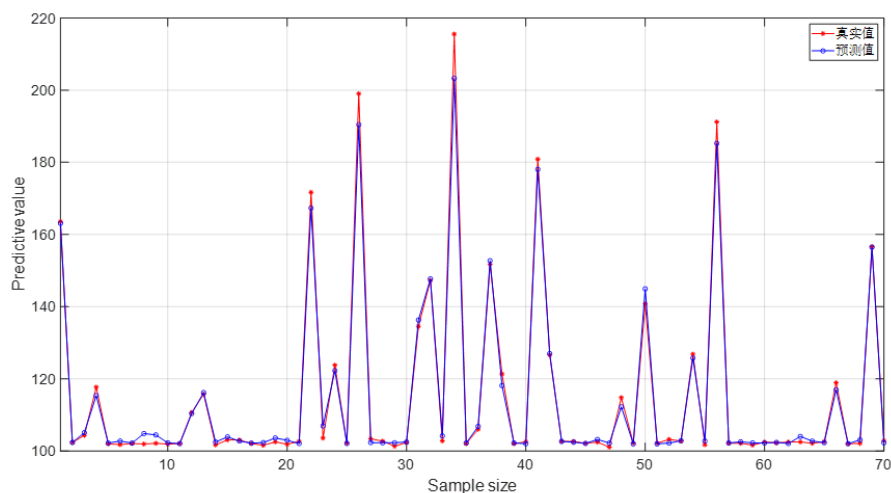


图 4 训练集预测结果对比

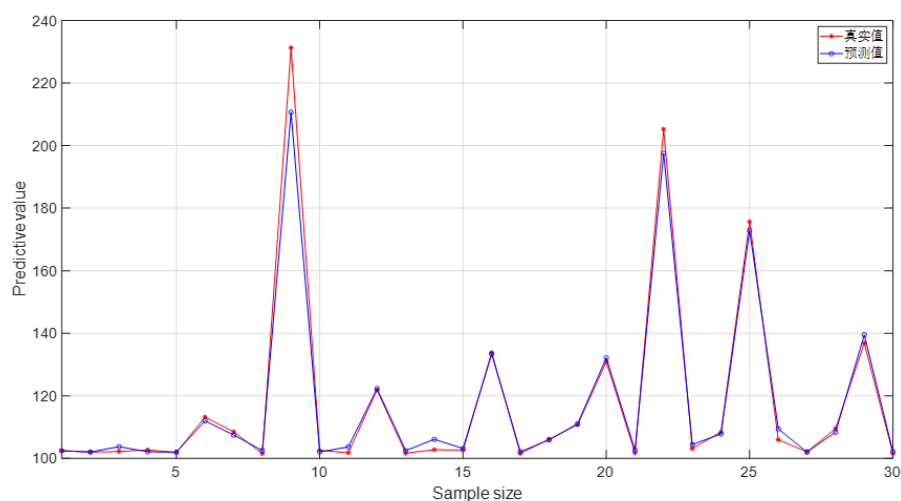


图 5 测试集预测结果对比

由预测结果图可知，存在一些样本处真实值与预测值偏离波动的情况，分析是由于样本数量较少，但预测结果总体趋势与实际影响趋势较为接近，表明 SVR 模型对不同尺寸对应力影响的预测拟合程度较高，结合训练集与测试集的评价指标，这验证了经过训练后的模型对不同尺寸下的应力影响的预测结果较为准确。

根据算法模型，进行优化，处理选择后可选出三个设计点(26.628, 14.276, 114.814)、(25.956, 14.539, 136.860)、(29.596, 13.861, 101.665)，其中(26.628, 14.276, 114.814)相比于原始的设计(28, 14, 136.23)，在降低应力提升结构强度的同时较大程度减少了重量，如表 3 所示。

表 3 预测优化设计点

	设计点 P1	设计点 P2	设计点 P3
缸体的外径/mm	26.628	25.956	29.596
缸体的中心孔径/mm	14.276	14.539	13.861
等效应力/MPa	114.814	136.860	101.665
减重率	12.15%	10.01%	8.82%

4.2 振动信号的时序预测

利用 NI 多功能采集卡可以采集光电传感器的转速信号以及电涡流位移传感器的位移信号。通过位移传感器测量得到随时间变化下的斜轴式柱塞泵缸体零件的振动信号，进行时序预测 $[(X, Y)]$ ，输入 X 为时间信号，输出为对应时间下的振动位移信号 Y。

对测量得到的振动信号进行预处理。原始振动信号的采集数据间隔很小（0.00005s），整体的样本数量过大（809824 个样本数据点），同时过程中会出现其他非正常工况下的干扰噪声信号，需要进行数据平滑和增大采样间隔对振动信号数据进行降噪处理^{[14][15]}，并采用线性插值及二次插值补充缺失数据，最后使用协方差矩阵提升数据稳定性，加强特征性。最终处理完成的振动位移信号随时间的关系如下表所示（详细数据见文件夹内的表格）。

表 4 时间位移振动信号

序号	（输入 X）时间 time	（输出 Y）振动信号 S
1	0	0.518255446
2	0.005	0.518785095
3	0.01	0.519350054
.....
1119	5.585	0.532379418
1120	5.59	0.533120926
1121	5.595	0.533615265

将数据集文件导入 MATLAB 中，运行程序，得到的结果和评价指标如表 5

所示，预测效果如图 6 和图 7 所示。

表 5 时序预测训练集与测试集的评价指标

评价指标	训练集评价结果	测试集评价结果
决定系数 R^2	0.96641	0.96209
平均误差 MAE	0.0015531	0.0016175
均方根误差 RMSE	0.0019943	0.0020911

时序预测的训练集的决定系数 R^2 为 0.96641，测试集的决定系数 R^2 为 0.96209 较高，同时均方根误差 RMSE 也接近于 0，说明预测波动较小，时序预测拟合的精度也是较高的。

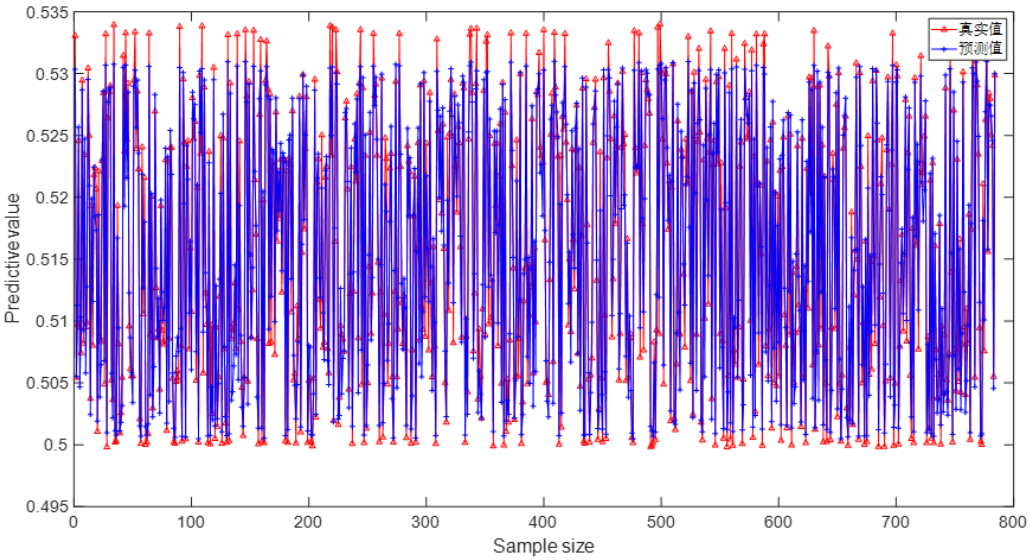


图 6 训练集预测结果对比

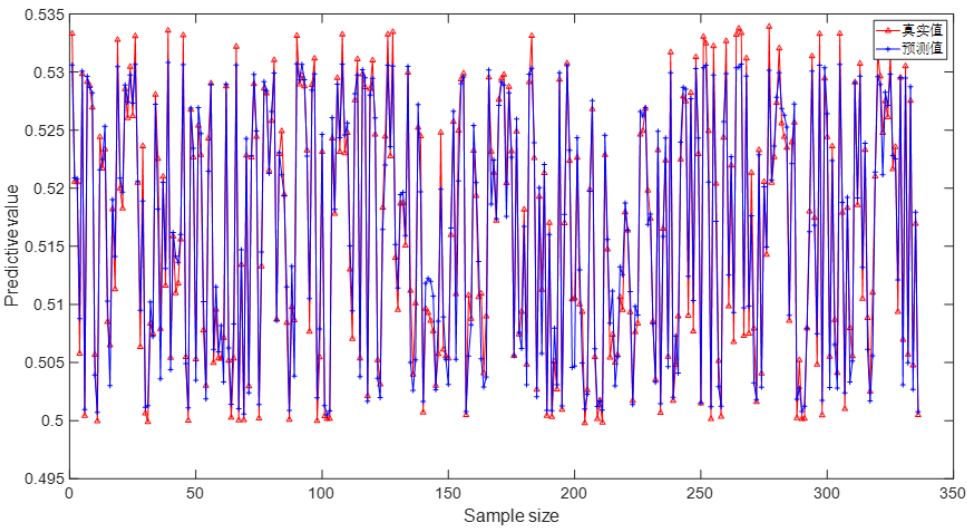


图 7 测试集预测结果对比

由预测结果图可知，样本数量较充足，总体预测波动范围较小，SVR 模型对振动位移的时序预测结果总体趋势与实际影响趋势较吻合一致，表明 SVR 模型对振动位移的时序预测拟合程度较高，结合振动位移的训练集与测试集的评价指标，这验证了经过训练后的模型对振动位移的时序预测结果较为准确，预测效果良好。

5 结论与展望

5.1 结论

本文以斜轴式液压泵为研究对象，采用支持向量回归 SVR 算法，对 SVR 模型进行构建，并使用 MATLAB 实现对关键零部件的两个相关尺寸对应力的影响进行回归预测以及振动位移随时间信号的时序预测模型应用。实验结果表明，SVR 模型能够有效预测拟合不同尺寸对零件应力的影响和时间位移振动信号，具有优良的多输入回归预测和时序预测性能，提高设备的可靠性和性能，为进一步液压泵技术的设计与应用提供了新的思路。

5.2 展望

支持向量回归 SVR 在回归问题中的应用具有很大的潜力和价值。它不仅可以处理线性回归问题，还可以通过核函数处理非线性回归问题。SVR 回归具有鲁棒性强、非线性拟合能力强和泛化能力强等优点，在零件优化设计、设备故障监测和人体运动分析等领域都有广泛的应用前景。

本文也存在一些不足之处，如在模型应用的数据样本中，总体设置的样本数较少，对样本数据的预处理还不够，使得模型训练量有所欠缺，同时预测也存在一定的波动。在后续的学习和研究中，可以结合其他算法如遗传算法 (GA)、粒子群算法(PSO)等深度学习算法对 SVR 回归算法模型进行优化和改进，发挥各自优势，提高预测的能力和对数据处理的精确性，更好地适用在特种装备的优化和故障监测领域。

致谢

由衷地感谢王琥老师通过深厚的专业知识和丰富的教学经验，为我们详细讲解了工程优化中的包含预处理共轭梯度法 PCG、BP 算法、CNN 卷积神经网络与支持向量机算法 SVM(SVR)等各种理论与方法，让我在工程优化这一领域中学到了很多的知识与方法，其中向量机算法 SVM(SVR)使我受益颇多，让我接触并使用到了优化和预测方法，将所学知识运用到实际中，并使用 SVR 回归方法对实验设备进行了研究，得到了零件在不同尺寸输入作用下对应力分布影响的回归预测以及振动信号的时序预测，不仅预测拟合效果良好，为之后所要开展的优化研

究开拓并提供了新的思路。我接下来也会积极学习工程优化中的新知识和方法，为未来的研究努力。

参考文献

- [1] 冯伟, 李红岩, 刘会祥, 等. 旋转配流盘式恒压变量斜轴柱塞泵建模与仿真[J]. 机床与液压, 2022, 50(11): 161-166.
- [2] Xiang L, Pan J, Chen S. Analysis on the stress corrosion crack inception based on pit shape and size of the FV520B tensile specimen[J]. Results in Physics, 2018, 9: 463-470.
- [3] Chu T, Nguyen C T, Yoo H, et al. ESP vibration prediction based on pump operating conditions in laboratory using machine learning[J]. Geoenergy Science and Engineering, 2024, 240: 13039-.
- [4] 邓聪颖, 冯义, 魏博, 等. 基于 SVR-GA 算法的广义加工空间机床切削稳定性预测与优化研究[J]. 仪器仪表学报, 2019, 40(10): 227-236. DOI: 10.19650/j.cnki.cjsi.J1905415.
- [5] Mana A, Majid K, Muhammad F, et al. Predictive modeling for compressive strength of 3D printed fiber-reinforced concrete using machine learning algorithms[J]. Case Studies in Construction Materials, 2024, 20: e02728-.
- [6] 何山, 郝雄博, 赵宇明, 等. 基于遗传算法优化支持向量回归的电池 SOH 预测[J]. 汽车技术, 2024, (05): 31-36. DOI: 10.19620/j.cnki.1000-3703.20230606.
- [7] 杨奕, 刘金源, 陈天民, 等. 基于 PSO-SVR 模型的短期天然气负荷预测[J]. 科学技术与工程, 2023, 23(35): 15210-15216.
- [8] 封周权, 邓佳逸, 华旭刚, 等. 基于贝叶斯优化支持向量回归的流线型箱梁颤振气动外形优化方法[J]. 东南大学学报(自然科学版), 2024, 54(02): 275-284.
- [9] Huo D, Chen J, Wang T. Chaos-based support vector regression for load power forecasting of excavators[J]. Expert Systems With Applications, 2024, 246: 123169-.
- [10] Jia X, Zhou L, Huang H, et al. Improving Electric Vehicle Structural-Borne Noise Based on Convolutional Neural Network-Support Vector Regression[J]. Electronics, 2023, 13(1):
- [11] Zhiqiang L, Sheng D, Timo K. Convex support vector regression[J]. European Journal of Operational Research, 2024, 313(3): 858-870.
- [12] 赖旭东. 基于 MLS-SVR 和 NSGA-II 的多响应参数优化研究[D]. 郑州大学, 2022.
- [13] 付文龙, 周建中, 张勇传, 等. 基于 OVMD 与 SVR 的水电机组振动趋势预测[J]. 振动与冲击, 2016, 35(08): 36-40.
- [14] 施云祺, 王子阳, 杨宏兵. 基于 GA-SVR 混合方法的非标设备设计时间预测[J]. 现代制造工程, 2023, (03): 1-7.
- [15] 崔庆安, 崔楠. 试验设计及参数优化的 LS-SVR 显著性因子筛选[J]. 管理科学学报, 2023, 26(12): 42-61.