

基于 LS-SVM 的铜铬合金挤压加工挤压力预测

舒服华

(武汉理工大学 机电工程学院, 武汉 430074)

摘要: 提出一种最小二乘支持向量机的 Cu-0.75Cr 铜合金反挤压力预测新模型。以断面缩减率、凸模锥角和挤压温度这 3 个主要工艺参数作为影响因素, 以反挤压过程的挤压力为影响对象, 通过最小二乘支持向量机模型建立影响因素和影响对象之间的复杂非线性关系。以正交实验数据为样本对模型进行训练, 用训练好的模型预测在一定反挤压条件下 Cu-0.75Cr 铜合金的挤压力。结果表明: 该模型不仅预测精度和处理速度大大高于人工神经网络预测模型, 而且建模速度也比标准支持向量机快, 实际预测误差小于 3%。

关键词: Cu-0.75Cr 铜合金; 反挤压; 挤压力; 预测; 最小二乘支持向量机

中图分类号: TG 146.11

文献标识码: A

Prediction model for extrusion force of Cu-0.75Cr alloy based on least square support vector machine

SHU Fu-hua

(School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan 430074, China)

Abstract: A novel prediction model for extrusion force of Cu-0.75Cr alloy reverse extrusion process based on least square support vector machine (LS-SVM) was proposed. With fault plane contraction rate, convex model awl angle and extrusion temperature as influence factors, and with extrusion force as influence object, the complex nonlinear relations among the influence factors and influence object were fitted by LS-SVM model. Orthogonal experiment was taken to obtain data samples, and LS-SVM model was established through the data samples, so that the extrusion forces of Cu-0.75Cr alloy under different backward extrusion process conditions can be predicted by this model. The results show that not only the prediction accuracy and treatment speed by this model are much higher than those of artificial neural networks (ANN), but also the construction speed is higher than that of standard SVM, and the practical prediction errors are less than 3.0%.

Key words: Cu-0.75Cr alloy; backward extrusion process; extrusion force; prediction; least square support vector machine

铜铬合金具有较高的硬度和强度, 耐磨性和减摩性好, 经时效处理后强度、硬度、导电性和导热性均显著提高, 易于电焊和钎焊, 在大气和淡水中耐蚀, 高温抗氧化性好。因此, 在许多工业领域得到广泛的应用^[1-2], 如 Cu-0.75Cr 用于制造电机整流子、电焊机电极、高速列车导线, 以及其他要求高温高强度、高硬度、导电性和导热性的零件等。挤压是 Cu-0.75Cr

铜合金常用的加工形式, 设计合理的挤压工艺规程是挤压加工的关键, 而正确计算挤压过程中的挤压力是前提条件。影响挤压力的因素多种多样, 如挤压温度、凸模锥角、凸模直径、润滑等, 这些因素对挤压力的影响呈一种复杂的非线性关系。常规挤压力的计算方法有图算法、经验公式法和理论计算法等^[3-4], 虽然这些方法在一些特定的场合有一定的预测能力, 但一般

而言都存在一定不足，如主观因素大、建模复杂、计算困难、精度低等。近年来，许多学者将人工神经网络技术运用于预测Cu-0.75Cr铜合金挤压力，取得了一定的突破^[2-3]，解决了传统方法建模难、计算复杂等问题，使预测精度和预测速度大大提高，但也存在先天性不足，如建立模型时存在训练时间长、易陷入局部极值点、外插能力弱、泛化能力低等缺点，尤其是需要大量的数据样本为基础。支持向量机(Support vector machine, SVM)是由统计学习理论发展起来的一种新型学习机器，它以结构风险最小化原理为理论基础，不同于基于经验风险最小化原理的神经网络方法，能根据有限的样本信息在模型的复杂性和学习能力之间寻求最佳折衷，适合小样本情况下机器的学习，具有较强的学习泛化能力^[5-9]。最小二乘支持向量机(Least square SVM, LS-SVM)是标准SVM的一种新扩展，它用等式约束代替标准SVM的不等式约束，将二次规划问题转化为线性方程组求解，降低了计算复杂性，具有更快的求解速度和更好的鲁棒性^[10]。本文作者利用最小二乘支持向量机预测Cu-0.75Cr铜合金反挤压工程的挤压力，以克服神经网络方法存在的泛化能力不强、处理速度慢、所需样本数量多等不足。

1 最小二乘支持向量机原理

支持向量机的基本思想是通过用内积函数定义的非线性变换将输入空间变换到一个高维空间，在这个高维空间中寻找输入变量和输出变量之间的一种非线性关系，即通过非线性映射 $\varphi(\cdot)$ 把样本空间映射到一个高维特征空间(Hilbert空间)，把寻找最优线性回归超平面的算法归结为求解一个凸约束条件下的一个凸规划问题，并可以求得全局最优解，在特征空间中可以应用线性学习机的方法解决样本空间中的高度非线性回归问题。支持向量机非线性回归原理如图 1 所示^[4]。

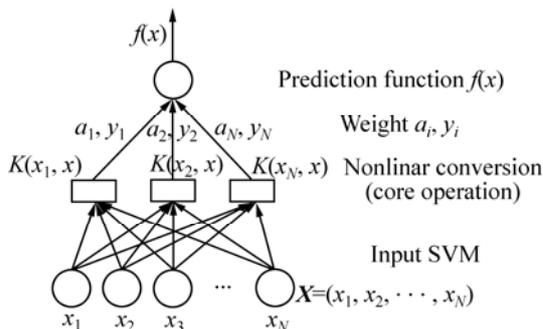


图 1 支持向量机原理示意图
Fig.1 Schematic diagram of principle for SVM

LS-SVM通过最小化经验风险与结构风险之和来求解SVM。设某一待回归系统的输入、输出样本数据集为 $\{x_k, y_k\} (k=1, 2, \dots, N)$ 。其中， x_k 为 n 维系统输入向量， $x_k \in \mathbf{R}^n$ ； y_k 为系统输出， $y_k \in \mathbf{R}$ 。LS-SVM回归模型可以表示为以下形式：

$$f(x) = \omega^T \varphi(x) + b \tag{1}$$

式中 $\varphi(\cdot)$ 为将输入空间映射为高维特征空间的映射函数，即 $\varphi(\cdot) : \mathbf{R} \rightarrow \mathbf{R}^{nh}$ ； ω 为超平面的权值向量， $\omega \in \mathbf{R}^n$ ； b 为偏置量。

LS-SVM 回归算法选择误差 e_k 的 2-范数为损失函数，其优化问题为

$$\begin{cases} \min J(\omega, e) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{k=1}^N e_k^2 \\ k=1, 2, \dots, N \end{cases} \tag{2}$$

式中 $\gamma > 0$ 为惩罚系数，用于调节误差所取的作用，它能够使训练误差和模型复杂度之间取一个折衷，以便使所求的函数具有较好的泛化能力，并且 γ 值越大，模型的回归误差越小。

引入Lagrange函数进行求解^[5]：

$$\begin{aligned} L(\omega, b, e, a) &= J(\omega, e) - \\ &\sum_{k=1}^N a_k (\omega^T \varphi(x_k) + b + e_k - y_k) \end{aligned} \tag{3}$$

式中 $a_k (k=1, 2, \dots, N)$ 为 Lagrange 乘子。

最优的 a 和 b 可以通过KKT条件获得^[6]，即：

$$\begin{cases} \frac{\partial L}{\partial \omega} = 0 \\ \frac{\partial L}{\partial \omega} = 0 \\ \frac{\partial L}{\partial b} = 0 \\ \frac{\partial L}{\partial e_k} = 0 \\ \frac{\partial L}{\partial a_k} = 0 \end{cases} \Rightarrow \begin{cases} \omega = \sum_{k=1}^N a_k \varphi(x_k) \\ \sum_{k=1}^N a_k = 0 \\ a_k = \gamma e_k \\ \omega^T \varphi(x_k) + b + e_k - y_k = 0 \end{cases} \tag{4}$$

消去式(4)中的变量 ω 和 e_k ，优化问题转化为求解如下线性方程组：

$$\begin{pmatrix} 0 & \bar{1}^T \\ \bar{1} & \Omega + \gamma^{-1} I \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix} \tag{5}$$

式中 $y = [y_1, y_2, \dots, y_N]^T$ ； $\bar{1} = [1, 1, \dots, 1]$ ； $a = [a_1, a_2, \dots, a_N]^T$ ； I 为单位矩阵； Ω 为方阵。

根据 Mercer 条件可知，存在映射 $\varphi(\cdot)$ 和核函数

$K(\cdot)$:

$$\Omega_{ij} = \varphi(x_i)^T \varphi(x_j) = K(x_i, x_j), i, j = 1, 2, \dots, N \quad (6)$$

由式(5)求出 a, b 后, 可得到 LS-SVM 的非线性函数, 估计为:

$$y(x) = \sum_{k=1}^N a_k K(x, x_k) + b \quad (7)$$

选择不同形式的核函数 $K(\cdot)$, 可以生成不同的支持向量机。

由此可见, LS-SVM 辨识算法选用与标准 SVM 不同的损失函数, 并用等式约束将求解二次规划问题转化为求解线性方程, 显然可以极大地降低算法的复杂性, 并提高辨识过程的抗干扰能力。LS-SVM 模型的学习训练主要是对线性方程组(5)进行求解, 而用 LS-SVM 模型进行预测只需计算各个训练样本与待测样本之间的核函数 $K(x_k, x)$, 不涉及函数 $\varphi(\cdot)$ 的具体形式。

2 铜铬合金挤压力预测

2.1 挤压力变化规律

以 Cu-0.75Cr 铜合金反挤压为例, 反挤压模具工作部分见图 2。挤压设备: 液压驱动的铜及铜合金挤压机, 挤压能力: 500 t; 锁模能力: 60 t; 模具材料: 5CrNiMo; 铜锭直径: 100 mm; 压缩气体压力: 0.4~0.6 MPa; 冷却器水温: 28~30; 挤压速度: 0.04 m/s, 挤压速度由操作箱上的速度表设定和调整, 可以控制近模速度和挤压速度; 盛锭筒温度由自动调整器控制。主要考虑挤压过程中断面缩减率(ε_F)、凸模锥角(α)和挤压温度(t)等工艺因素对挤压力的影响规律。在挤压过程中, 变形温度对挤压力的影响是通过变形抗力 σ_K 的大小反映出来的, 随着变形温度的升高, 坯料的变形抗力下降, 所需挤压力下降, 即当变形抗力随着温度的升高而线性减小时, 则挤压力随温度的升高而线性下降。实际上, 大多数金属和合金的变形抗力随温度升高而下降是非线性的, 从而挤压力与变形温度一般也为非线性关系^[1, 11]。在反挤压过程中, 总挤压力(F)的大小是单位挤压力(p)与凸模和挤压件接触部分在水平面上投影部分面积(S)的乘积, 因此当凸模直径增加时, 总的挤压力是增加的。反挤压时凸模锥角的大小对挤压力的影响也是比较明显的, 凸模锥角对挤

压力的影响与变形程度有关^[1-2], 根据研究, 在 Cu-0.75Cr 反挤压过程中, 其他条件一定, 凸模锥角从 120°到 180°变化时, 挤压力是逐渐增大的。

由于影响挤压力的因素多, 且为非线性关系, 加之各因素相互作用, 所以, 很难用一种精确的数学模型描述, 最小二乘支持向量机模型是解决一种问题的理想方法, 它能精确拟合这种复杂非线性影响关系。

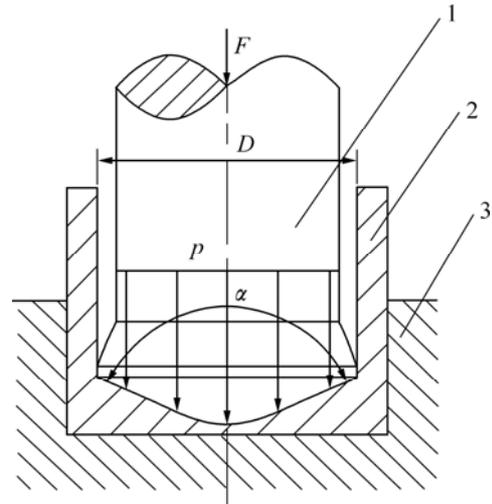


图 2 Cu-0.75Cr 合金反挤压模具工作部分

Fig.2 Backward extrusion dies of Cu-0.75Cr alloy: 1—Convex model; 2—Extrusion part; 3—Concave model

2.2 建立 LS-SVM 模型

进行 Cu-0.75Cr 铜合金反挤压实验获取样本数据^[1]。取断面缩减率(ε_F)、反挤压温度(t)、凸模锥角(α)为影响因子, 凹模直径为固定因素取为 64 mm。采用正交实验设计, 每个实验因子取 3 个水平, 如表 1 所示, 共安排实验 25 组, 其中 16 组实验数据作为训练样本, 其余 9 组为测试样本。

表 1 正交实验的因素水平

Table 1 Factors and levels of orthogonal test

Level	$\varepsilon_F/\%$	$\alpha/^\circ$	$t/$
1	39	120	700
2	52	150	740
3	62	170	820

以 Cu-0.75Cr 铜合金反挤压过程中的断面缩减率(ε_F)、凸模锥角(α)和挤压温度(t)作为输入, 将挤压力 F 作为输出。建立 LS-SVM 系统, 采用径向基核函数。模型性能评价指标采用平均误差计算公式

$$\varepsilon = \frac{1}{m} \sum_{i=1}^m \frac{|Q_i - f_i|}{f_i} \quad (8)$$

式中 Q_i 为实际的测量值； f_i 为预测值； m 为验证次数。

核宽度 σ 、惩罚系数 γ 是建立 LS-SVM 模型的重要问题。 σ 和 γ 的选择通常采用交叉验证方法，但是交叉验证方法由于 σ 和 γ 的参数集参数有限的原因，经常出现不能满足达到误差精度的情况。为此，本文作者采用一种自适应的 σ 和 γ 选择来建立 LS-SVM 模型，具体步骤如图 3^[5, 12-13]。

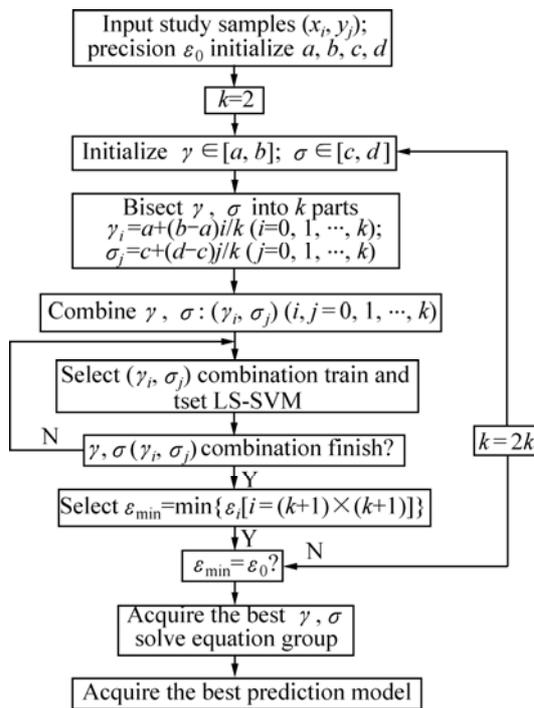


图 3 LS-SVM 预测模型流程图

Fig.3 Flow chart of LS-SVM model

表 2 ANN 模型预测结果

Table 2 Test results of samples by ANN model

Sample No.	Fault plane contraction rate, ε_f (%)	Convex model awl angle, α (°)	Extrusion temperature/	Extrusion process, F/kN				
				Test value	ANN Prediction value	Error/%	LS-SVM Prediction value	Error/%
1	120	39	700	869.4	846.8	2.60	888.53	2.20
2	150	39	740	975.2	935.6	4.06	958.9	1.67
3	170	39	820	1197.2	1255.6	4.89	1231.2	2.84
4	120	52	700	687.9	722.8	5.07	706.1	2.65
5	150	52	740	732.5	772.2	5.42	748.7	2.21
6	170	52	820	781.2	821.1	5.11	797.3	2.06
7	120	62	700	1 008.5	1 071.2	6.21	989.1	1.92

2.3 系统的训练和测试

用训练样本对 LS-SVM 系统进行学习训练，完成训练后，用测试样本对系统进行测试。步骤如下^[14-15]：

- 1) 输入训练样本；
- 2) 选定径向基核函数，初始化核宽度 σ 和惩罚因子 γ ；
- 3) 按自适应选择方法求解核宽度 σ 和惩罚因子 γ ；
- 4) 根据 LS-SVM 算法求解回归参数 a 和 b ；
- 5) 将模型参数引入 LS-SVM 预测机，输入测试样本预测挤压力。

预测结果如表 2 所列。由表 2 可知，最大预测误差小于 3%，平均误差为 1.97%。为了检验本预测方法与神经网络预测方法的性能，设计一个 BP 神经网络，网络结构及参数见文献[1]。ANN 模型预测的结果也列于表 2，最大预测误差大于 6%，平均误差为 4.38%。这表明本文 LS-SVM 预测模型预测挤压力精度显著提高，同时，LS-SVM 预测模型学习训练时间大大缩短，仅为 ANN 模型的千分之一。几种预测方法的结果比较如图 3 所示。

从测试结果还可以得出，断面缩减率、凸模锥角、挤压温度三者对挤压力的影响基本规律。当断面缩减率不变时，挤压力随凸模锥角增大和挤压温度的降低而增大；当凸模锥角不变时，挤压力随断面缩减率的增大和挤压温度的降低而增大；当挤压温度不变时，挤压力随凸模锥角增大和断面缩减率的增大而增大。

8	150	62	740	1 102.7	1 145.1	3.85	1 228.6	2.35
9	170	62	820	1 285.2	1 256.8	2.22	1 261.8	1.82

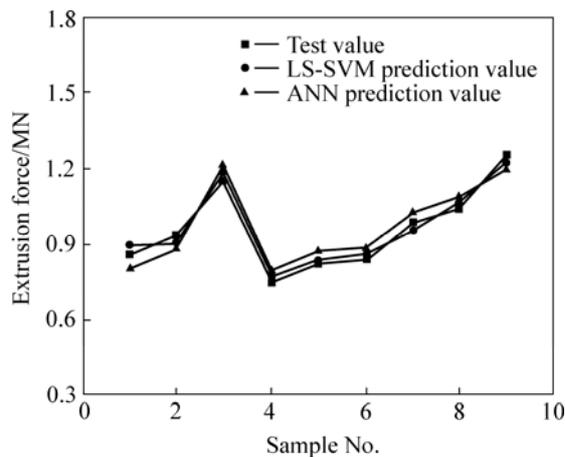


图 4 几种预测方法结果比较

Fig.4 Comparison of prediction results by different methods

3 结论

1) 利用最小二乘支持向量机模型预测 Cu-0.75Cr 铜合金反挤压加工挤压力, 不需建立方程式, 模型的建立和修改容易, 模型具有自学习和自适应能力和较强的概括性, 可处理被干扰的数据。

2) 以 Cu-0.75Cr 铜合金反挤压受挤断面缩减率、凸模锥角、挤压温度为影响因素, 通过 LS-SVM 能自动预测挤压力, 预测结果精确可靠, 预测误差在 3% 以内, 对现实生产具有一定的指导作用。

3) SVM 预测模型, 所需样本数量少, 适合小样本状态下的机器学习; 而 LS-SVM 预测模型, 求解简便快捷、预测精度高, 性能优于 ANN 和标准 SVM 模型, 较好解决了高维数、局部极小、小样本等机器学习问题。

REFERENCES

[1] 苏娟华, 陈 钢, 吴朋越. 基于 BP 神经网络的 Cu-0.75Cr 挤压过程中的挤压力预测[J]. 有色金属, 2005, 57(1): 35-38.
SU Jun-hua, CHEN Gang, WU Peng-yue. Extrusion force prediction of Cu-0.75Cr alloy with BP network[J]. Nonferrous Alloy, 2005, 57(1): 35-38.

[2] Qi L H, Hou J J, Cui P L. Research on prediction of the processing parameters of liquid extrusion by BP network[J]. Journal of Material Processing Technology, 1999, 95: 232-237.

[3] 顾小兰, 叶以富, 田秋红, 程勇锋, 施利旦. 铜铬合金等径角挤压后的组织与性能[J]. 机械工程材料, 2006, 30(3): 51-63.

GU Xiao-lan, YE Yi-fu, TIAN Qiu-hong, CHENG Yong-feng, SHI Li-dan. Microstructure and properties of Cu-Cr alloy processed by ECAP[J]. Materials Mechanical Engineering, 2006, 30(3): 51-63.

[4] 李慧丽, 吴朋越, 谢水生. 利用 BP 神经网络预测铜合金挤压力[J]. 铸造装备与制造技术, 2005, 21(2): 67-69.
LI Hui-li, WU Peng-yue, XIE Shui-sheng. Application of BP network to predict the copper alloy extrusion pressure[J]. Cast Equipment and Manufacture Technique, 2005, 21(2): 67-69.

[5] 曹一波, 谢小鹏. 基于最小二乘支持向量机的磨损预测[J]. 润滑与密封, 2007, 32(2): 138-141.
CAO Yi-bo, XIE Xiao-peng. Wear loss prediction using least square support vector machine[J]. Lubrication Engineering, 2007, 32(2): 138-141.

[6] MOSER E M, FALLER C H, EGG-MANN F. Modeling the functional performance of plasma polymerized thin films[J]. Thin Solid Films, 1999, 355/356: 49-54.

[7] 吴德会. 基于最小二乘支持向量机的表面粗糙度预测模型铣削加工[J]. 中国机械工程, 2007, 18(7): 838-840.
WU De-hi. A prediction model for surface roughness in milling based on least square support vector[J]. Chinese Mechanical Engineering, 2007, 18(7): 838-840.

[8] SUYKENS J T, VAN G L. Least squares support vector machines[M]. Singapore World Scientific, 2002.

[9] CHAPELLE O, VAPNIK V N. choosing multiple parameters for SVM[J]. Mach Learn, 2002, 46: 131-146.

[10] SUYKENS JA K, VANDEWALLE J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3): 293-300.

[11] RAWTANI L, RANA J L. Modeling of material behavior data in a functional form suitable for neural network representation[J]. Computational Materials Science, 1999, 15: 493-503.

[12] 张 猛, 付丽华, 张 维. 一种新的最小二乘支持向量计算方法[J]. 计算机工程与应用, 2007, 43(11): 33-34.
ZHANG Meng, FU Li-hua, ZHANG Wei. Modified least squares support vector machines[J]. Computer Engineering and Applications, 2007, 43(11): 33-34.

[13] HSU C W, LIN C J. A comparison of methods for multiclass support vector machines[J]. IEEE Transactions on Neural Networks, 2002, 13(2): 415-425.

[14] ENGEL Y, MANNOR S, MEIR R. The kernel recursive least-squares algorithm[J]. IEEE Transactions on Signal Processing, 2004, 52(8): 2275-2285.

[15] ZHANG Xue-gong. Introduction to statistical learning theory and support vector machines[J]. Acta Automation Science, 2000, 26(1): 32-41.

(编辑 龙怀中)