

doi: 10.3969/j.issn.1006-1576.2011.10.013

基于改进 PSO-LSSVM 的军用工程机械研制费用预测模型

徐波

(南昌陆军学院战术教研室, 南昌 330103)

摘要: 针对传统参数法对装备研制费用进行预测存在的局限性, 采用改进粒子群算法 (particle swarm optimization, PSO) 对 LSSVM 模型进行改进, 构建军用工程机械研制费用预测模型。运用 2 种优化策略改进粒子群算法, 对种群初始化过程进行控制、克服粒子群算法易于早熟的缺点。用改进后的粒子群算法优化最小二乘支持向量机的模型参数和核参数, 以获得更好的预测效果。预测结果表明: 该费用预测模型运用于军用工程机械研制费用预测, 明显优于传统预测模型, 具有很好的预测精度和效率。

关键词: 军用工程机械; 研制费用; 预测; 粒子群算法; 最小二乘支持向量机

中图分类号: TJ02 **文献标志码:** A

Research Costs Forecasting Model of Military Engineering Machinery Based on Improved PSO-LSSVM

Xu Bo

(Staff Room of Tactics, Nanchang Military Academy, Nanchang 330103, China)

Abstract: In order to solve the limitation problem of using the traditional parameter method to predict the research costs, it adopts the improved particle swarm optimization (PSO) to improve the LSSVM model, which constructing the development cost's forecasting model. It uses two kinds of optimization strategy to improve the PSO, which can control the population initialization process, and overcome the shortcomings that the particle swarm algorithm is easy to early maturity. It uses the improved particle swarm algorithm to optimize the model parameters and nuclear parameters of the least square support vector machine (LSSVM) in order to get better prediction effect. The prediction results show that the prediction model used in the cost military engineering machinery, is obviously superior to the traditional forecasting model. The improved prediction model has the very good prediction accuracy and efficiency.

Keywords: military engineering machinery; research costs; forecasting; particle swarm optimization; least square support vector machines

0 引言

军用工程机械研制费用是指在军用工程机械研制过程中, 项目初始论证规划、市场分析、可行性研究、相关技术研究、型号设计、样机和原型机制造、工程设计、设计资料归档、软件、对工程模型的试验评价和相关管理等费用的总和^[1]。其特点主要表现为有效样本数据少, 影响装备研制费用的因素呈非线性。因此, 采用传统的参数法对其进行预测具有局限性。许多文献都有关于 BP 预测模型和 LSSVM 预测模型的论述, 这些模型虽然都具有较强的非线性处理能力, 且适用于小样本数据^[2], 符合军用工程机械研制费用预测的特点, 但是传统 LSSVM 模型本身的缺陷使得训练效率及预测精度并不高。因此, 笔者采用改进粒子群算法 (particle swarm optimization, PSO) 对 LSSVM 模型进行改进, 以构建军用工程机械研制费用预测模型。

1 预测模型构建

1.1 PSO 优化策略

1.1.1 分块种群初始化机制

PSO 是一种全局最优搜索算法^[3]。该算法在种群初始化阶段, 粒子在整个搜索区域内随机产生, 不能保证充分分散于整个搜索空间, 影响优化效果。分块种群初始化机制的主要思想是使每个粒子几乎均匀分布, 假设种群粒子数为 n , 则将整个搜索空间分为 n 个小区域, 每个小区域中随机产生一个粒子。搜索空间可按公式 (1) 进行划分:

$$X_{i,k} \in [a_k + f_k(b_k - a_k), a_k + f_{k+1}(b_k - a_k)]$$
$$k = 1, 2, \dots, D$$
$$\begin{cases} f_k = (i-1) \bmod C^{D-k} \\ f_D = (i-1) - \sum_{j=1}^{D-1} f_j C^{D-j} \\ C = \sqrt[n]{n} \end{cases} \quad (1)$$

收稿日期: 2011-06-16; 修回日期: 2011-07-18

作者简介: 徐波(1983—), 男, 江西人, 硕士, 助教, 从事指挥自动化与战场数字化研究。

式中： a_k 和 b_k 表示粒子位置向量上第 k 维上值的取值范围； mod 表示取模，然后按照式 (2) 产生第 i 个粒子的初始位置：

$$X_{i,k} = a_k + f_k(b_k - a_k) + \frac{1}{\sqrt[n]{n}}(b_k - a_k) \cdot r \quad (2)$$

式中 r 为 $[0,1]$ 上取值的随机数。

1.1.2 无进化代数变异策略

在高维函数的优化问题中，PSO 的求解结果很容易陷入局部最优解。而判断是否进入局部最优解的重要指标是无进化次数。因此将粒子的无进化次数和能引起种群多样性的变异算子一起引入 PSO 中，作为变异时机的选择标准。具体的改进方法为：

1) 计算适应度变化率 $FSlope$ ：适应度变化率是本次迭代与本次迭代前 M 次迭代之间 (M 取 1)，历史最佳粒子位置 p_g 的适应值变化率为：

$$FSlope = \frac{f(p_g(t)) - f(p_g(t-M))}{f(p_g(t))} \quad (3)$$

式中 f 为适应度函数。

2) 统计无进化次数 $stoptime$ ：进化开始时设置无进化次数 $stoptime$ 为 0，适应度变化阈值 $Slopevalue$ (取 $10e^{-10}$)、粒子无进化阈值 $MaxStep$ 和变异概率 p_m (取 0.4)。在迭代过程中，粒子无进化次数根据适应度变化率来确定：

$$\begin{cases} FSlope < Slopevalue & stoptime = stoptime + 1 \\ FSlope > Slopevalue & stoptime = 0 \end{cases} \quad (4)$$

如果无进化次数超过设置的无进化阈值 $MaxStep$ ，则认为算法可能处于停滞状态，此时按照变异概率对种群进行变异操作：

$$p_{gi}(t+1) = p_{gi}(t) + 0.5 \times r \times p_{gi}(t) \quad (5)$$

式中 r 为 $[0,1]$ 内的随机数。

该项改进使 PSO 在训练时避免因陷入局部极小值而出现训练停止的现象，当训练停滞时，算法通过变异使种群粒子跳出局部极小点，继续向全局最优优点逼近。

1.2 改进 PSO 优化 LSSVM 参数

LSSVM 模型中包含模型参数 C, ϵ 和核参数 σ^2 。采用改进 PSO 优化上述参数时，种群在迭代更新过程中不断向本代最优位置和全局最优位置学习。

1) 编码与解码方法：将上述模型参数和核参数构成实数数列，每个数列为 PSO 的一个初始粒子，多个初始粒子构成初始种群。在种群初始化时，要充分考虑到各参数的取值范围，以免种群在更新过程

中出现不符合要求的情况。解码时，依次读取相应位参数值即可；2) 性能评价方法：采用误差平方和作为 PSO 的性能评价函数。在优化过程中，每次迭代学习都需要对粒子进行解码，将获得的参数输入 LSSVM 中计算训练样本的预测值和误差平方和，然后将误差最小粒子更新为当代最优个体或全局最优个体；3) 学习终止方法：粒子群算法优化参数满足以下任何一个条件时停止，一是小于设定的最小误差值；二是达到最大迭代次数。

2 实例分析

将建立的改进 PSO-LSSVM 预测模型运用于军用装载机研制费用的预测中，检验该预测模型对军用工程机械研制费用的效果。

2.1 样本预处理

2.1.1 费用驱动因子选取

军用工程机械对研制费用有影响的特征参数很多。在预测前，必须对相关特征参数进行筛选，确定费用驱动因子。采用关联度分析法对军用装载机的费用驱动因子进行筛选。该方法的优点是可以充分考虑以下因素：一是各特征参数与对应费用的关联度；二是各特征参数间的独立性。选择方法如下：

1) 将各种型号的军用工程机械费用作为参考数列，各特性参数作为对比数列，计算各特征参数与费用的相关性，删除相关性小的特征参数；2) 依次将各特征参数作为参考数列，其他特征参数作为对比数列，计算特征参数之间的相关性，如相关性较大，则删除与费用关联度相对较小且重复的特征参数。

2.1.2 费用时间价值转换

时间价值主要体现为货币的通货膨胀率，而商品零售价格指数能很好地反映该指标。在对军用工程机械各费用进行预测前，要对原费用数据进行时间价值转换，按式 (6) 将所有装备费用转换到基准年。笔者以 2008 年为基准年对样本数据进行转换。

$$\text{基准年费用} = \frac{\text{基准年指数}}{\text{原始年指数}} \times \text{原始费用} \quad (6)$$

2.2 结果分析

军用装载机的特征参数主要有：铲斗容量 X_1 、装载重量 X_2 、装载高度 X_3 、额定转速 X_4 、额定功率 X_5 、液力变距系数 X_6 、油泵压力 X_7 、前进速度 X_8 、后退速度 X_9 、最大牵引力 X_{10} 、爬坡能力 X_{11} 、最小回转半径 X_{12} 、最小离地间隙 X_{13} 、轮距 X_{14} 、轴距 X_{15} 、机器重量 X_{16} 。将所获得的样本数据按照上述方法进行预处理，结果如表 2。

表 1 样本规范化结果数据表

类型	X_1	X_8	X_9	X_{10}	X_{12}	X_{14}	X_{15}	研制费用/万元
1	0.000 0	0.131 6	0.309 5	0.162 8	0.456 5	0.275 9	0.196 4	5.6
2	0.100 0	0.131 6	0.309 5	0.240 3	0.529 9	0.448 3	0.357 1	8.2
3	0.200 0	0.263 2	0.000 0	0.415 3	0.550 5	0.491 4	0.410 7	11.2
4	0.300 0	0.334 2	0.500 0	0.604 7	0.703 0	0.534 5	0.714 3	10.5
5	0.400 0	0.289 5	0.011 9	0.633 5	0.682 9	0.620 7	0.714 3	12.8
6	0.600 0	0.394 7	0.071 4	0.728 7	0.869 6	0.750 0	0.767 9	13.6
7	1.000 0	0.368 4	0.023 8	1.000 0	1.000 0	0.965 5	0.857 1	16.0
8	1.000 0	0.473 7	0.107 1	0.976 7	0.977 7	1.000 0	0.857 1	14.8
9	0.000 0	0.000 0	0.190 5	0.000 0	0.000 0	0.000 0	0.000 0	3.3
10	0.480 0	1.000 0	1.000 0	0.728 7	0.638 6	0.853 4	0.717 9	18.4
11	0.800 0	0.368 4	0.523 8	0.658 9	0.788 0	0.758 6	1.000 0	15.5
12	0.100 0	0.105 3	0.285 7	0.178 3	0.543 5	0.310 3	0.223 2	6.6
13	0.200 0	0.210 5	0.023 8	0.317 8	0.440 2	0.534 5	0.446 4	11.2
14	0.360 0	0.263 2	0.071 4	0.348 8	0.733 7	0.534 5	0.464 3	11.7

2.3 模型预测过程及精度检验

将表 1 数据分别输入 BP 预测模型、LSSVM 预测模型和改进的 PSO-LSSVM 预测模型中进行训练，对比预测结果。其中，BP 预测模型采用带动量因子的 BP 预测模型，选用模型的隐层节点数为 18，训练目标为 0.000 01，LSSVM 预测模型参数 $C=300$ ，核参数为 $\sigma^2=5$ ，LSSVM 预测模型和改进 PSO-LSSVM 预测模型的核函数都采用 RBF 核函数。模型训练过程中，对 BP 预测模型的误差曲线和改进 PSO-LSSVM 的误差曲线进行模拟，结果如图 1、2。

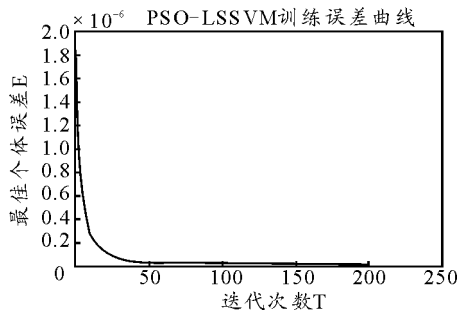


图 1 改进 PSO 训练 LSSVM 误差曲线图

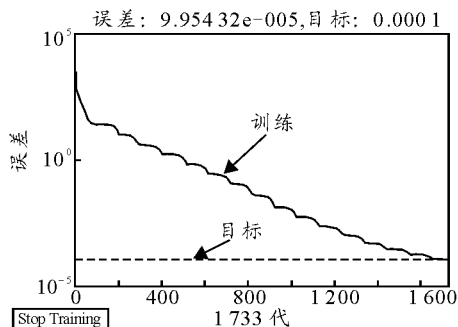


图 2 BP 网络误差曲线图

从图 1、2 可知：改进的 PSO-LSSVM 预测模型训练误差下降很快，经过约 200 代训练趋于停止；而 BP 预测模型经过 1733 代停止，改进 PSO-LSSVM 预测模型的训练效率明显高于 BP 预测模型。

由于 BP 预测模型受初始参数影响较大，为了降低预测的随机性，对 BP 网络连续进行 10 次预测求平均值，并计算预测结果的平均百分比误差，预测结果及误差对比如表 2、3。

表 2 3 种预测模型预测结果对比表 万元

预测样本	真实值	BP 预测模型	传统 LSSVM 预测模型	改进 PSO-LSSVM 预测模型
1	6.6	6.720 4	6.766 4	6.580 9
2	11.2	11.066 6	10.874 8	11.092 4
3	11.7	11.479 4	11.612 0	11.596 0

表 3 3 种预测模型预测误差对比表 %

预测样本	BP 预测模型	传统 LSSVM 预测模型	改进 PSO-LSSVM 预测模型
1	1.82	2.52	0.29
2	1.19	2.90	0.96
3	2.21	0.75	0.89
MPAPE	1.74	2.06	0.71

从表 3 可知，改进 PSO-LSSVM 预测模型预测精度明显优于 LSSVM 预测模型和 BP 预测模型；而且改进 PSO-LSSVM 预测模型在训练样本较少的情况下，预测精度明显高于 BP 预测模型。由此可见，改进的 PSO-LSSVM 预测模型，无论训练效率还是预测精度都比 LSSVM 和 BP 预测模型高，更适用于样本数据较少的军用工程机械研制费用预测。

3 结束语

预测结果表明：该费用预测模型运用于军用工程机械研制费用预测，明显优于传统预测模型，具有很好的预测精度和效率。

参考文献：

[1] 王健, 沈云峰, 申岳国, 等. 工程装备全寿命管理过程评价体系及其应用[J]. 起重运输机械, 2007, 22(9): 25-28.
 [2] 克里斯特安尼, 等. 支持向量机导论[M]. 王猛, 曾华军, 李国正, 译. 北京: 电子工业出版社, 2004.
 [3] 段晓东, 刘向东, 王存睿. 粒子群算法及其在应用[M]. 沈阳: 辽宁大学出版社, 2007.