

基于改进的粒子群算法优化 LSSVM 股价预测研究

刘家旗

西北大学经济管理学院, 陕西 西安 710127

摘要: 为了提高股票价格的预测精度, 针对股票价格数据的非平稳非线性的特性, 本文运用改进的 PSO 实现 LSSVM 的核参数和惩罚系数自适应选择, 提出一种 SAPSO 优化 LSSVM 股价预测模型, 并以此进行实证分析。通过基于 SAPSO-LSSVM 算法的 1 步、3 步、5 步和 7 步预测结果和不同模型的预测时间和预测均方误差的对比结果可知, SAPSO-LSSVM 股价预测模型具有预测精度高, 预测时间短的优点, 同时能够实现预测参数的自适应选择。

关键词: 粒子群算法; 股票预测; LSSVM

中图分类号: F830.91

文献标识码: A

文章编号: 1000-2324(2015)04-0628-04

Study on the Prediction for Stock Price Based on the Optimized LSSVM of the Improved Particle Swarm Algorithm

LIU Jia-qi

School of Economics & Management/Northwest University, Xi'an 710127, China

Abstract: In order to improve the prediction accuracy of the stock price, stock price data for the nonlinear and non-stationary characteristics, this paper used the improved PSO to implement the self-adaptive selection of the LSSVM kernel parameter and penalty coefficient, and proposed a prediction model for stock price on SAPSO optimized LSSVM to analyze a case. The results showed that it had the high prediction accuracy, the advantages of short time and also could realize the self-adaptive selection for forecasting parameters based on prediction results in 1 step, 3 step, 5 step and the 7 on the SAPSO-LSSVM algorithm and the comparison between prediction time and the mean square error of different models.

Keywords: Particle Swarm Algorithm; prediction for stock price; Least Squares Support Vector Machine(LSSVM)

1 引言

随着国内外经济发展环境和投资者预期复杂化等相关影响因素的影响, 如何准确把握股票价格的发展趋势和股票价格的准确预测是当前研究的重点和难点。雷霆等人^[1]结合神经网络理论和小波变换技术, 提出一种基于小波神经网络的股价预测模型, 运用小波变换技术提取股票价格的特征参数。实验结果证明该方法预测精度较高, 但其稳定性较差, 有待提高。Wen Guo 等人^[2]运用 PSO 算法对 BP 神经网络进行全局寻优, 并将其应用于股价预测。仿真结果表明, 寻优模型的预测精度和收敛速度均好于标准 BP 神经网络, 但 PSO 算法存在局部最优的问题。刘渊等人^[3]针对小波神经网络存在的缺陷, 将混沌理论引入小波神经网络, 实证结果表明混沌小波神经网络的预测误差远远小于 RBF 神经网络的股价预测结果, 但对小波基函数的选择和确定难度较大。杨光等人^[4]利用小波核函数的多分辨率的优点, 提出一种基于小波核 LS-SVM 的股价预测, 实验表明该方法具有一定的优越性, 不过参数需要手动确定。

在前人研究的基础上, 本文针对股票价格的非线性特点, 提出一种 SAPSO 算法优化 LSSVM 的核参数和惩罚系数的自适应股价预测模型。通过预测模型的构建和算法流程的优化, 仿真本文算法可以实现股票价格的自适应预测和预测参数的最优化选择。

2 改进的粒子群算法

2.1 粒子群算法

粒子群算法是一种模拟鸟群觅食的群体智能算法, 每个粒子表示一个可能解向量, 粒子的好坏根据适应度函数值进行判断, 并通过向全局和个体最优解学习实现粒子位置和速度的不断更新, 最终实现全局寻优的目的。 $V_i = (v_{i1}, v_{i2}, \dots, v_{id})$ 和 $X_i = (x_{i1}, x_{i2}, \dots, x_{id})$ 分别表示粒子的位置和粒子的速度, 粒子的位置和速度的更新策略如下公式所示^[5]:

收稿日期: 2014-06-05

修回日期: 2014-07-11

作者简介: 刘家旗(1994-), 女, 陕西省咸阳市人, 本科. E-mail:sanyan@263.net

$$V_{id}^{(t+1)} = w \times V_{id}^{(t)} + c_1 r_1 (p_{id}^{(t)} - X_{id}^{(t)}) + c_2 r_2 (p_{gd}^{(t)} - X_{id}^{(t)}) \tag{1}$$

$$X_{id}^{(t+1)} = X_{id}^{(t)} + V_{id}^{(t+1)} \tag{2}$$

$$w = w_{max} - \frac{w_{max} - w_{min}}{t_{max}} \cdot t \tag{3}$$

公式 (1) (2) (3) 中, $p_{id}^{(t)}$, $p_{gd}^{(t)}$ 分别表示迭代次数 t 时刻粒子的个体最优解和全局最优解; $r_1, r_2 \in \hat{I}(0, 1)$ 的随机数; c_1, c_2 表示学习因子; t_{max} 和 t 分别表示最大迭代次数和当前迭代次数; w 为惯性权重, 其中 w_{max}, w_{min} 分别表示惯性权重的最大值和最小值。

2.2 改进的粒子群算法

为了避免粒子群算法的局部最优问题和加快粒子群算法的全局搜索能力, 本文运用非线性的动态惯性权重系数改进标准 PSO 算法, 提出一种 SAPSO 算法, 其公式如下所示:

$$\omega = \begin{cases} \omega_{min} - \frac{(\omega_{max} - \omega_{min}) * (f - f_{min})}{f_{avg} - f_{min}}, & f \leq f_{avg} \\ \omega_{max}, & f > f_{avg} \end{cases} \tag{4}$$

公式 (4) 中, f 表示粒子的当前适应度值; f_{avg} 表示粒子当前所有粒子的平均适应度值, f_{min} 表示粒子当前所有粒子的最小适应度值。由于惯性权重随粒子的适应度函数值的变化而进行自动调整, 因此将改进的粒子群算法称为自适应权重 PSO 算法。

3 SAPSO 优化 LSSVM 预测模型

3.1 LSSVM 支持向量机

Suykens 提出的 LSSVM 可转化为^[6]:

$$Min J(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + C \sum_{k=1}^N \xi_k^2 \quad s.t. \quad y_k = \phi(x_k) \omega^T + b + \xi_k \tag{5}$$

公式 (5) 中, $\xi_k \geq 0, k = 1, 2, \dots, N, C$ 为惩罚因子。

$$L(\omega, b, \xi, \alpha) = \frac{1}{2} \|\omega\|^2 + C \sum_{k=1}^N \xi_k^2 - \sum_{k=1}^N \alpha_k [(\omega^T \phi(x_k) + b + \xi_k) - y_k] \tag{6}$$

公式 (6) 中, $\alpha_k (k = 1, 2, \dots, N)$ 表示拉格朗日乘子。

对 ω, b, ξ, α 求偏导数, 并令其为零, 有:

$$\omega = \sum_{k=1}^N \alpha_k \phi(x_k) \quad \sum_{k=1}^N \alpha_k = 0, \quad \alpha_k = C \xi_k \tag{7}$$

$$\omega^T \phi(x_k) + b + \xi_k - y_k = 0$$

依据 Mercer 条件, 核函数 $k(x_i, x_j)$ 可以用公式 (8) 表达:

$$k(x_i, x_j) = \phi(x_i) \phi(x_j) \tag{8}$$

文中运用 RBF 核函数进行股价预测, 如公式 (9) 所示:

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{9}$$

所以 LSSVM 股价预测模型可用公式 (10) 表达:

$$f(x) = \sum_{i=1}^m \alpha_i \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) + b \tag{10}$$

由公式 (10) 可知, LSSVM 的预测效果主要受 γ, σ 控制, 为了实现 γ, σ 的自适应选择, 运用 SAPSO 算法对 γ, σ 进行自适应选择。

3.2 SAPSO-LSSVM 股价预测模型

由于 LSSVM 模型所要优化的参数是 γ, σ ，所以定义优化模型如公式 (11) 所示：

$$F = \{\gamma, \sigma\} \tag{11}$$

在确保预测精度最优的情况下，通过优化实现 γ, σ 参数的自适应选择。假设 t 时刻的实际股价为 $y(t)$ ，预测股价 $\hat{y}(t)$ ，那么实际股价 $y(t)$ 和预测股价 $\hat{y}(t)$ 的差值 $e(t)$ 可以运用公式 (12) 表示：

$$e(t) = \hat{y}(t) - y(t) \tag{12}$$

针对股价预测问题，实际股价数据为 n ，运用 SAPSO 优化 LSSVM 的核参数和惩罚系数，使得 LSSVM 的实际股价输出和预测股价之间的差值的平方和最小，其适应度函数如公式 (13) 所示：

$$\min Fitness(t) = \frac{1}{2n} \sum_{t=1}^T [e(t)]^2 \tag{13}$$

3.3 算法步骤

基于 SAPSO-LSSVM 的股价预测算法步骤如下，流程图如图 1 所示：

Step1: 归一化股价样本数据，建出训练样本和测试样本；

Step2: 设定 SAPSO 算法的种群大小 $popsiz$ e，学习因子 c_1, c_2 ，最大迭代次数 $\max gen$ ；

Step3: 将构建出的训练样本输入 LS-SVM，根据适应度函数公式 (13) 计算粒子的适应度函数值，寻找粒子个体和全局最优粒子的位置和最优值；

Step4: 粒子速度和位置的更新；

Step5: 计算适应度同时更新位置和速度；

Step6: 若 $gen > \max gen$ ，保存最优解；反之 $gen = gen + 1$ ，转到 Step4；

Step7: 根据最优位置所对应的最优参数 γ, σ 进行股价的预测。

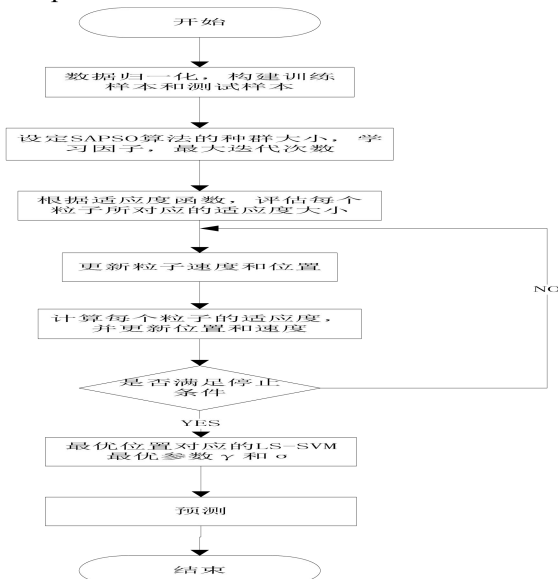


图 1 基于 SAPSO-LSSVM 预测流程图

Fig.1 Flow chart of prediction on SAPSO-LSSVM

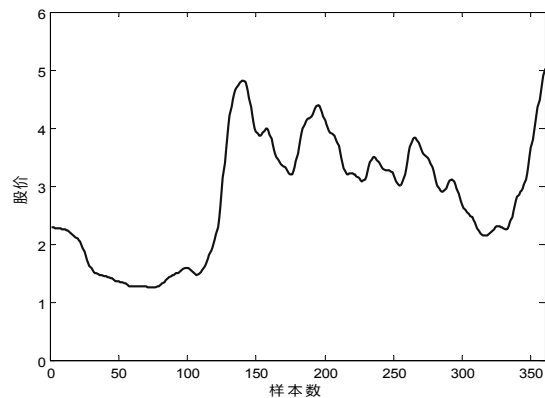


图 2 原始股价序列

Fig.2 Sequences of raw stock price

4 仿真数据

4.1 数据来源

为了本文算法的有效性，运用实际数据进行验证。本文数据来源于上证交易所，以 2013 年 11 月 7 日~2014 年 12 月 31 日青岛啤酒股价数据为研究对象，其股价序列图如图 2 所示。

4.2 评价指标

为了验证本文算法进行股价预测的有效性，采用均方误差用来评价股价预测效果的评价指标，评价指标如公式 (14) 所示：

$$\text{均方误差: } MSE = \sqrt{\frac{1}{K} \sum_{i=1}^K (x_i - \hat{x}_i)^2} \tag{14}$$

公式 (14) 中, x_i, \hat{x}_i 分别表示实际股价和预测股价。

4.3 实验结果

将 360 组股价数据分成训练样本和测试样本, 前 336 组作为训练样本数据, 后 24 组作为测试样本数据, 用于验证预测模型的好坏。设 SAPSO 算法的最大迭代次数为 100, 种群大小为 20, $popmin=-5.12$, $popmax=5.12$, $vmax=1$, $vmin=-1$, 其预测结果如图 3~6 所示。

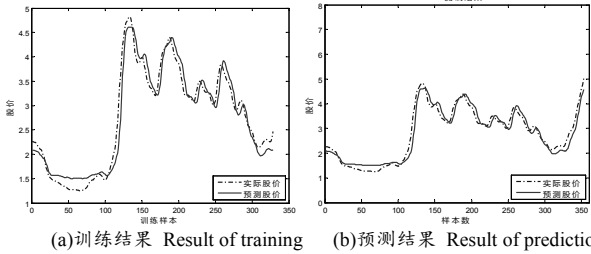


图 3 基于 SAPSO-LSSVM 算法的单步股价预测

Fig.3 One step prediction for stock price based on SAPSO-LSSVM

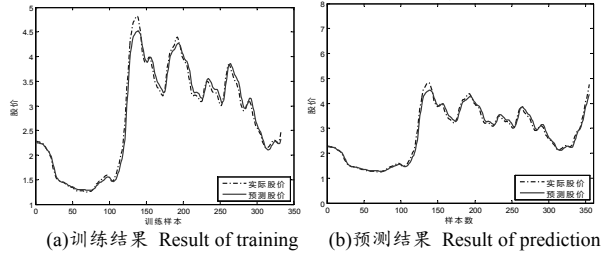


图 4 基于 SAPSO-LSSVM 算法的 3 步股价预测

Fig.4 Three steps prediction for stock price based on SAPSO-LSSVM

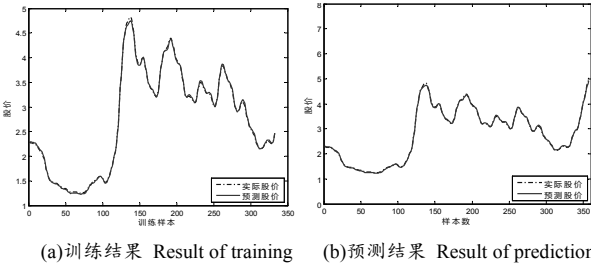


图 5 基于 SAPSO-LSSVM 算法的 5 步网络流量预测

Fig.5 Five steps prediction for stock price based on SAPSO-LSSVM

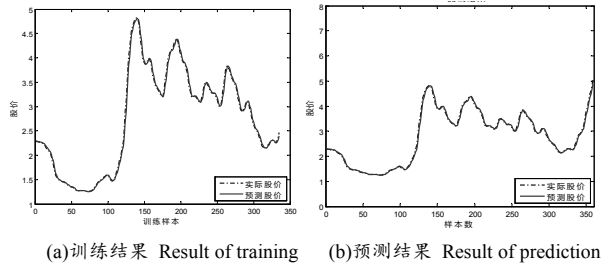


图 6 基于 SAPSO-LSSVM 算法的 7 步股价预测

Fig.6 Seven steps prediction for stock price based on SAPSO-LSSVM

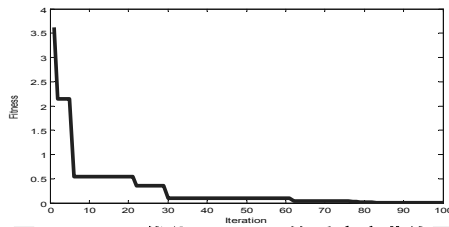


图 7 SAPSO 优化 LSSVM 的适应度曲线图

Fig.7 Curve chart for fitness function based on SAPSO-LSSVM

由 SAPSO-LSSVM 算法的 1 步股价预测、3 步股价预测、5 步股价预测和 7 步股价预测结果图可知, 随着预测步长的增加, SAPSO-LSSVM 算法的股价预测精度也不断提高, 效果很好。图 7 是 SAPSO 算法优化 LSSVM 的适应度曲线。

5 结论

针对 LSSVM 核参数和惩罚系数选择的随机性, 本文运用 SAPSO 算法实现 LSSVM 核参数和惩罚系数自适应选择, 构建出基于 SAPSO-LSSVM 的股价预测模型, 并进行 SAPSO-LSSVM 的 1 步、3 步、5 步和 7 步股价预测。通过对不同股价预测模型预测均方误差和预测时间发现, SAPSO-LSSVM 算法的预测精度和预测效率均优于其他模型, 同时可以实现股价预测参数的自适应选择。

参考文献

[1] 雷霆,余镇危.一种股价预测的小波神经网络模型[J].计算机应用,2012,26(3):526-528

[2] Guo Wen, QiaoYi zheng, Hou Haiyan. BP neural network optimized with PSO algorithm and its application in forecasting[C]. Weihai: Proceedings of the IEEE international Conference on Information Acquisition, 2006:617-621

[3] 刘 渊,戴 悦,曹建华.基于小波神经网络的股票价格时间序列预测[J].计算机工程,2012,34(16):105-110

[4] 杨 光,张国梅,刘星宇.基于小波核 LSSVM 的股价预测[J].微机发展,2011,15(12):125-128

[5] 李欣然.权重自适应调整的混沌量子粒子群优化算法[J].计算机系统应用,2012,21(8):127-130

[6] 叶 青.基于 GARCH 和半参数法的 VaR 模型及其在中国股市风险分析中的应用[J].统计研究,2013(12):18-21