

引用格式:刘云,易松. 基于双参数最小二乘支持向量机(TPA-LSSVM)的风电时间序列预测模型的优化研究[J]. 北京化工大学学报(自然科学版),2019,46(2):97-102.

LIU Yun, YI Song. Optimization of a wind power time series prediction model based on a two-parameter least squares support vector machine[J]. Journal of Beijing University of Chemical Technology (Natural Science), 2019,46(2):97-102.

基于双参数最小二乘支持向量机(TPA-LSSVM)的风电时间序列预测模型的优化研究

刘 云 易 松*

(昆明理工大学 信息工程与自动化学院, 昆明 650500)

摘 要: 风电时间序列预测模型的优劣直接影响风功率的应用价值,最小二乘支持向量机(least squares support vector machine, LSSVM)在处理风电预测问题上具有明显优势。提出了一种双参数算法(two-parameter algorithm, TPA),从理论上证明了任意初始值均可线性收敛到全局最优值。调用 TPA 算法对 LSSVM 模型的惩罚因子和径向基宽度进行寻优赋值,并将训练好的 TPA-LSSVM 模型应用于风电预测中。仿真结果表明,与 LSSVM 模型、粒子群最小二乘支持向量机(PSO-LSSVM)模型、径向基函数神经网络(RBFNN)模型相比,TPA 算法可以更好地实现 LSSVM 的参数寻优,TPA-LSSVM 模型能有效提高预测精度。

关键词: 最小二乘支持向量机; 时间序列预测; 双参数算法; 参数寻优

中图分类号: TN929.5 **DOI:** 10.13543/j.bhxbzr.2019.02.015

引 言

由于高间歇性和不可调度性,大规模风电一体化对电力系统的稳定运行带来了巨大的挑战,而短期风电预测概念的提出明显缓解了这一问题^[1]。现有的预测主要基于物理模型^[2]、统计模型^[3]和混合模型^[4]3 种模型。物理模型基于地形和风电场的详细物理描述,从数值天气预报中获得模拟结果,而统计模型基于历史气象时间序列。由于数值天气预报的高计算成本,基于数值天气预报的物理模型不适用于短期风电预测(提前预测小于等于 6 h)。混合模型包含物理模型和统计模型,由于同样需要数值天气预报的结果,因此其应用也有较大限制。统计模型与物理模型和混合模型相比,成本更低,复杂度更小,更适合于短期风电预测^[5]。

在统计模型中,通过训练历史风电数据来预测未来风电数据。自回归整合移动平均(ARIMA)模

型是风电预测中使用最为广泛的传统统计模型之一,但该模型通常只能表示线性关系。风电数据具有非平稳性、非线性特点,ARIMA 模型的预测性能受限。Mahnaz 等^[6]提出使用径向基函数神经网络(RBFNN)预测短期风电,RBFNN 能够以任意精度逼近任意的非线性函数,且拓扑结构紧凑,收敛速度快。但 RBFNN 对训练样本依赖程度很高,很难反映系统真正的输入输出关系,且在优选过程中容易出现数据病态现象。Wu 等^[7]提出了支持向量机(support vector machine, SVM),该模型广泛应用于时间序列数据的预测。最小二乘支持向量机(least squares support vector machine, LSSVM)模型^[8]是 SVM 模型的拓展,采用最小二乘线性系统作为损失函数,将 SVM 的二次规划问题转变为求解线性方程,有效提高了运算速度并降低了复杂度。在 LSSVM 模型中,惩罚因子 μ 和径向基宽度 δ 直接影响预测结果。针对 LSSVM 参数的优化问题,Corus 等^[9]使用基于遗传算法(genetic algorithm, GA)优化的 LSSVM 模型预测风电,但 GA 算法存在早熟、欺骗等现象,以及搜索效率低、时间复杂度高问题。Anamika 等^[10]使用基于粒子群优化(particle swarm optimization, PSO)的 LSSVM 模型预测风电,PSO 算

收稿日期:2018-09-13

基金项目:国家自然科学基金(61262040)

第一作者:男,1973 年生,副教授

* 通信联系人

E-mail: 1090037776@qq.com

法使用编码和适应度函数优化控制参数,但容易出现早熟或停滞搜索现象,易陷入局部极值。由于 GA 算法与 PSO 算法在对 LSSVM 的参数进行寻优时均有不同程度的限制,故针对 LSSVM 的参数寻优算法还有待突破。

为了使 LSSVM 预测模型具有更好的预测效果,本文提出一种最优化算法—双参数算法(two-parameter algorithm, TPA),从理论上证明了任意初始值均可线性收敛到全局最优值。TPA 算法可防止早熟和欺骗现象的发生,并且收敛速度快,寻优精度高,不易陷入局部极值。通过调用 TPA 算法对 LSSVM 模型的惩罚因子 μ 和径向基宽度 δ 进行寻优赋值,并将训练好的 TPA-LSSVM 模型应用于风电预测中。仿真结果表明,与 LSSVM 模型、粒子群最小二乘支持向量机(PSO-LSSVM)模型和径向基函数神经网络(RBFNN)模型相比,TPA-LSSVM 模型有效提高了预测精度。

1 模型构建

支持向量机是根据结构风险最小化原则和统计学习理论提出的,最小二乘支持向量机是 SVM 的最小二乘法,即将 SVM 中的非等式约束转化为等式约束,使 SVM 中的二次规划问题变为求解线性方程问题。

首先假设训练样本集为 $(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots, (\mathbf{x}_m, \mathbf{y}_m)$, 其中 $\mathbf{x}_i \in \mathbf{R}^m$ 表示样本输入向量, $\mathbf{y}_i \in \mathbf{R}$ 表示样本输出向量, m 表示样本总数。输入空间 \mathbf{R}^m 被非线性映射函数 $\varphi(\mathbf{x})$ 映射到特征空间 \mathbf{Z} 。根据结构风险最小化原则构建 LSSVM 模型的最优决策函数,可表示为

$$f(\mathbf{x}) = \|\mathbf{w}\|^2 \varphi(\mathbf{x}) + \mathbf{b} \quad (1)$$

式中 $\mathbf{w} \in \mathbf{Z}$ 且 $\mathbf{b} \in \mathbf{R}$, \mathbf{w} 与 \mathbf{b} 的值可以通过风险函数确定。LSSVM 的优化问题可表示为

$$\min_{\mathbf{w}, \mathbf{b}, \mathbf{e}} J(\mathbf{w}, \mathbf{e}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{\mu}{2} \sum_{i=1}^n \mathbf{e}_i^2 \quad (2)$$

$$\text{s. t. } y_i = \mathbf{w}^T \varphi(\mathbf{x}_i) + \mathbf{b} + \mathbf{e}_i \quad (3)$$

式中 \mathbf{w} 表示权重向量, \mathbf{b} 表示偏差向量, \mathbf{e}_i 表示松弛因子向量, μ 表示惩罚因子向量, $\varphi(\cdot)$ 表示从输入空间到高维特征空间的非直接映射。

为了快速找到式(2)中函数的最小值,运用拉格朗日法对该问题进行求解

$$L(\mathbf{w}, \mathbf{b}, \mathbf{e}, \boldsymbol{\alpha}) = J(\mathbf{w}, \mathbf{e}) - \sum_{i=1}^n \alpha_i [\mathbf{w}^T \varphi(\mathbf{x}_i) + \mathbf{b} +$$

$$\mathbf{e}_i - y_i] \quad (4)$$

式中, $\boldsymbol{\alpha}$ 为对应 \mathbf{x}_i 拉格朗日乘子, 可以表示为 $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m)$ 。根据 Kuhn-Tucker 最优化(KKT)条件, 对式(4)进行微分可得出

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum_{i=1}^n \alpha_i \varphi(\mathbf{x}_i) \quad (5)$$

$$\frac{\partial L}{\partial \mathbf{b}} = 0 \Rightarrow \sum_{i=1}^n \alpha_i = 0 \quad (6)$$

$$\frac{\partial L}{\partial \mathbf{e}_i} = 0 \Rightarrow \alpha_i = \mu \mathbf{e}_i \quad (7)$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \Rightarrow \mathbf{w}^T \varphi(\mathbf{x}_i) + \mathbf{b} + \mathbf{e}_i - y_i = 0 \quad (8)$$

根据表达式(5)~(8)可以解出 $\boldsymbol{\alpha}$ 与 \mathbf{b} 的值。

对于一个新样本 \mathbf{x} , LSSVM 模型的输出为

$$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i K(\mathbf{x}, \mathbf{x}_i) + \mathbf{b} \quad (9)$$

式中 $K(\mathbf{x}, \mathbf{x}_i) = \varphi(\mathbf{x}^T) \varphi(\mathbf{x}_i)$, 为核函数。本文选取径向基(RBF)核函数, 可以表示为

$$K(\mathbf{x}, \mathbf{x}_i) = \exp \left[-\frac{(\mathbf{x} - \mathbf{x}_i)^2}{\delta^2} \right] \quad (10)$$

式中 δ 表示基核宽度。

为解决优化问题(2), 应选择合适的惩罚因子 μ 和径向基宽度 δ 来训练 LSSVM, 合理选择 μ 和 δ 可有效提高模型的预测精度。为避免由于经验选择或随机选择参数而造成的误差, 使用 TPA 算法对参数进行优化。

2 基于 TPA 算法优化的最小二乘支持向量机

2.1 TPA 算法描述

首先考虑一个优化问题

$$\underset{\mathbf{x} \in \mathbf{R}^n}{\text{minimize}} f(\mathbf{x}) \quad (11)$$

式中 $f: \mathbf{R}^n \rightarrow \mathbf{R}$, f 为连续可微的强 m -凸函数, Lipschitz 常数为 L 。 f 具有唯一的全局极小值点 \mathbf{x}_* , 且 $\mathbf{x}_* \in \mathbf{R}^n$ 。其中强 m -凸函数定义如下。

设 $(X, \|\cdot\|)$ 是实赋范线性空间, I 是 X 的一个凸子集, J 是 \mathbf{R} 内的一个区间, 且 $(0, 1) \subset J$, 函数 $m: J \rightarrow [0, +\infty)$ 。若函数 $f: I \rightarrow [0, +\infty)$ 满足: 存在常数 $c > 0$, $\forall \mathbf{x}, \mathbf{y} \in I, t \in (0, 1)$, 则称 f 是 I 上的强 m -凸函数

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \geq m(t)f(\mathbf{x}) + m(1-t)f(\mathbf{y}) - ct(1-t) \|\mathbf{x} - \mathbf{y}\|^2 \quad (12)$$

Lipschitz 定义如下。

L 对于在实数集的子集的函数 $f: D \subseteq \mathbf{R}$, 若存在常数, 使得 $|f(a) - f(b)| \leq L|a - b|$, 其中 $\forall a, b \in D$, 则称 f 符合 Lipschitz 条件, 对于 f 最小的常数 L 称为 f 的 Lipschitz 常数。

定义1(函数类) 令函数 f 的集合为 $S_{m,l}$, 其中 f 为连续可微的强 m -凸函数, Lipschitz 常数为 L , 当 $f \in S_{m,l}$ 时, $g = L/m$, g 为条件数。

定义2(TPA 算法) 为解决式(11)的无约束最优化问题, 近年来部分学者提出采用递归法来编写算法^[11-12], 如式(13)所示。

$$\begin{aligned}\xi_{k+1} &= (1 + \beta)\xi_k - \beta\xi_{k-1} - \alpha \nabla f(y_k) \\ y_k &= (1 + \gamma)\xi_k - \gamma\xi_{k-1} \\ x_k &= (1 + \delta)\xi_k - \delta\xi_{k-1}\end{aligned}\quad (13)$$

式中 $\alpha, \beta, \gamma, \delta$ 为变量, $\xi \in \mathcal{L}_e^n$ 表示系统的内部状态, \mathcal{L}_e^n 为所有单边序列 x 的集合, $y \in \mathcal{L}_e^n$ 用于求解梯度, 输出为 $x \in \mathcal{L}_e^n$, 且 $\xi_0, \xi_{-1} \in \mathbf{R}^n$ 表示初始条件。

已知 Lipschitz 常数为 L , $g = L/m$ 。

令 $\rho = 1 - 1/\sqrt{g}$, 则

$$(\alpha, \beta, \gamma, \delta) = \left(\frac{1+\rho}{L}, \frac{\rho^2}{2-\rho}, \frac{\rho^2}{(2-\rho)(1+\rho)}, \frac{\rho^2}{1-\rho^2} \right) \quad (14)$$

式(14)即为 TPA 算法。

2.2 TPA 算法收敛性证明

定理1 令 $x \in \mathcal{L}_e^n$, $x_* \in \mathbf{R}^n$ 且 $k_0 \geq 0$, 如果存在一个序列 $q \in \mathcal{L}_e$, 则

$$\|x_{k+1} - x_*\|^2 \leq \rho^2 \|x_k - x_*\|^2 - q_k, \quad \forall k \geq k_0 \quad (15)$$

$$0 \leq \sum_{j=0}^k \rho^{-2j} q_j, \quad \forall k \geq 0 \quad (16)$$

x_k 以速率 ρ 经过 k_0 次迭代后, 线性收敛到 x_* , 当 $k \geq k_0$ 时, $\|x_k - x_*\| \leq \rho^k$, 其中

$$c = \left(\rho^{-2k_0} \|x_{k_0} - x_*\|^2 + \sum_{j=0}^{k_0-1} \rho^{-2(j+1)} q_j \right)^{\frac{1}{2}} \quad (17)$$

证明: 假设存在一个满足式(15)和式(16)的序列 $q \in \mathcal{L}_e$ 。首先定义一个等式 η_k

$$\eta_k = \|x_{k+1} - x_*\|^2 - \rho^2 \|x_{k+1} - x_*\|^2 + q_k \quad (18)$$

将式(18)代入式(15)可得, 在 $k \geq k_0$ 时, $\eta_k \leq 0$ 恒成立。对其进行求和

$$0 \geq \sum_{j=k_0}^{k-1} \rho^{2(k-j-1)} \eta_j = \|x_k - x_*\|^2 - \rho^{2(k-k_0)}$$

$$\|x_{k_0} - x_*\| + \rho^{2(k-1)} \sum_{j=k_0}^{k-1} \rho^{-2j} q_j \quad (19)$$

根据式(16), 可得

$$- \sum_{j=0}^{k_0-1} \rho^{-2j} q_j \leq - \sum_{j=k_0}^{k-1} \rho^{-2j} q_j \quad (20)$$

结合式(19)和式(20)得到误差的约束条件为

$$\|x_k - x_*\|^2 \leq \rho^{2(k-k_0)} \|x_{k_0} - x_*\|^2 + \rho^{2(k-1)} \sum_{j=k_0}^{k-1} \rho^{-2j} q_j \quad (21)$$

在 $k \geq k_0$ 时, 式(21)成立。从式(21)右边提取 ρ^{2k} , 得

$$\|x_k - x_*\|^2 \leq \rho^{2k} \left(\rho^{-2k_0} \|x_{k_0} - x_*\|^2 + \sum_{j=0}^{k_0-1} \rho^{-2(j+1)} q_j \right) \quad (22)$$

两边同时开方

$$\|x_k - x_*\| \leq \rho \left(\rho^{-2k_0} \|x_{k_0} - x_*\|^2 + \sum_{j=0}^{k_0-1} \rho^{-2(j+1)} q_j \right)^{\frac{1}{2}} \quad (23)$$

令

$$c = \left(\rho^{-2k_0} \|x_{k_0} - x_*\|^2 + \sum_{j=0}^{k_0-1} \rho^{-2(j+1)} q_j \right)^{\frac{1}{2}}$$

可得

$$\|x_k - x_*\| \leq c \rho^k \quad (24)$$

因此, x_k 以速率 ρ 经过 k_0 次迭代后, 线性收敛到 x_* 。

定理2 当 $f \in S_{m,L}$ 且 $0 < m \leq L$ 时, $x_* \in \mathbf{R}^n$ 是函数 f 的唯一最小值。对于任何初始值 $\xi_0, \xi_{-1} \in \mathbf{R}^n$, 双变量算法的迭代都满足

$$\|x_k - x_*\| \leq c \rho^k \quad (25)$$

$$f(x_k) - f(x_*) \leq c^2 \frac{L}{2} \rho^{2k} \quad (26)$$

式(25)和式(26)在 $k \geq 1$ 时成立, 其中 $\rho = 1 - 1/\sqrt{k}$ 且

$$c = \rho^{-1} \left(\|x_1 - x_*\|^2 - \frac{1}{mL} p_m(y_0)^T p_L(y_0) \right)^{1/2} \quad (27)$$

$$p_r(y) = \nabla f(y) - r(y - x_*) \quad (28)$$

为了证明定理2中 x_k 以速率 ρ 线性收敛到 x_* , 需要一个满足式(16)的序列 q_k 且使式(15)成立。本文采用文献[13]中的引理生成序列 q_k 。

引理1 令 $f \in S_{m,L}$ 且 $\nabla f(x_*) = 0$, 定义 $p_r(y) = \nabla f(y) - r(y - x_*)$, 可得到

$$\mathbf{q}_k = \begin{cases} -\mathbf{p}_m(\mathbf{y}_k)^T [\mathbf{p}_L(\mathbf{y}_k) - \rho^2 \mathbf{p}_L(\mathbf{y}_{k-1})], & k \geq 1 \\ -\mathbf{p}_m(\mathbf{y}_0)^T \mathbf{p}_L(\mathbf{y}_0), & k \geq 0 \end{cases} \quad (29)$$

且

$$0 \leq \sum_{j=0}^k \rho^{-2j} q_j, \quad \forall k \geq 0 \quad (30)$$

证明:双变量算法产生序列 \mathbf{x}_k 和 \mathbf{y}_k 且初始条件为 ξ_0 和 ξ_1 。取 $k_0 = 1$, 可得

$$\mathbf{q}_k = \frac{1}{mL} \begin{cases} \mathbf{p}_m(\mathbf{y}_k)^T [\mathbf{p}_L(\mathbf{y}_k) - \rho^2 \mathbf{p}_L(\mathbf{y}_{k-1})], & k \geq 1 \\ -\mathbf{p}_m(\mathbf{y}_0)^T \mathbf{p}_L(\mathbf{y}_0), & k \geq 0 \end{cases} \quad (31)$$

式中 $\mathbf{p}_L(\mathbf{y}_k) = \nabla f(\mathbf{y}_k) - r(\mathbf{y}_k - \mathbf{x}_*)$, 当 $k \geq 1$ 时, 式(16)成立。将式(15)和式(16)代入式(18)中, 可得序列 η_k 的表达式

$$\eta_k = \|\mathbf{x}_{k+1} - \mathbf{x}_*\|^2 - \rho^2 \|\mathbf{x}_k - \mathbf{x}_*\|^2 - \frac{1}{mL} \rho_m (\mathbf{y}_k)^T [\mathbf{p}_L(\mathbf{y}_k) - \rho^2 \mathbf{p}_L(\mathbf{y}_{k-1})] \quad (32)$$

当 $k \geq 1$ 时, 式(32)成立。根据三重动量法的定义, 可以将式(14)进行等效

$$\begin{cases} \mathbf{x}_{k+1} = (1 + \delta)\xi_{k+1} - \delta\xi_k \\ \mathbf{x}_k = (1 + \delta)\xi_k - \delta\xi_{k-1} \\ \mathbf{y}_k = (1 + \gamma)\xi_k - \gamma\xi_{k-1} \\ \mathbf{y}_{k-1} = (1 + \gamma)\xi_{k-1} - \gamma\xi_{k-2} \\ \nabla f(\mathbf{y}_k) = [-\xi_{k+1} + (1 + \beta)\xi_k - \beta\xi_{k-1}]/\alpha \\ \nabla f(\mathbf{y}_{k-1}) = [-\xi_{k-1} + (1 + \beta)\xi_{k-1} - \beta\xi_{k-2}]/\alpha \end{cases} \quad (33)$$

式中,

$$\begin{cases} (\alpha, \beta, \gamma, \delta) = \left(\frac{1+\rho}{L}, \frac{\rho^2}{2-\rho}, \frac{\rho^2}{(2-\rho)(1+\rho)}, \frac{\rho^2}{1-\rho^2} \right) \\ \rho = 1 - 1/\sqrt{L/m} \end{cases} \quad (34)$$

η_k 可由 $\xi_{k-2}, \xi_{k-1}, \xi_k, \xi_{k+1}, \mathbf{x}_*$ 表示, 将式(33)和式(34)代入式(32)中, 可得出等式 $\eta_k \equiv 0$ 的唯一解 $\xi_{k-2}, \xi_{k-1}, \xi_k, \xi_{k+1}, \mathbf{x}_*$ 。

2.3 预测模型参数优化

通过对 TPA 算法及 LSSVM 算法的分析, 提出将 TPA 算法应用于 LSSVM 参数的寻优, 通过对 LSSVM 中的惩罚因子 μ 和径向基宽度 δ 进行寻优赋值, 使 TPA-LSSVM 预测模型具有更好的预测效果。

TPA-LSSVM 预测模型具体应用步骤如下。

(1) 初始化 TPA-LSSVM 预测模型中相关参数。

(2) 使用 TPA 算法搜索惩罚因子 μ 和径向基宽度 δ 的值。

(3) 取惩罚因子 μ 和径向基宽度 δ 的全局最优值, 并对 LSSVM 模型进行训练。

(4) 对训练得到的模型进行测试, 得出测试结果。

3 实验结果与分析

3.1 数据集

采用美国国家海洋和大气管理局 (NOAA) 的“weather data”数据集作为模型的训练数据。NOAA 标准数据集是美国国家海洋和大气管理局通过气象卫星以及其他途径采集形成的实时数据集^[14], 该数据集每 15 min 采样一次, 测试数据和训练数据有不同的概率分布, 使风电预测更具有真实性。本文实验采用 2011~2016 年数据作为训练数据, 2017 年第一季度的风力发电数据作为测试数据。

3.2 TPA 算法参数选取

由式(14)可得, TPA 算法的 4 个参数与函数 f 的 m 与 L 两个参数取值有关。 m 与 L 的选取采用多次试验法, 其中 L 使用文献[5]中的取值, 即 $L = 10^4$ 。为了使训练的数据完整有效, 令 m 的取值范围为 1~10, 对于不同大小组合的 m 值与 L 值, 采用 TPA 算法对训练样本进行预测。表 1 为不同的 m 值对应的预测误差。

表 1 不同 m 值的预测误差
Table 1 Prediction errors for different m values

m	误差
1	0.167
2	0.164
5	0.156
7	0.160
10	0.163

从表 1 中选取预测误差最小的一组值作为函数 f 的维数 m 和 Lipschitz 梯度大小 L 的最终取值。由表 1 可知, m 取 5。由 m 和 L 的取值可以计算出 g 值和 ρ 值, 将 ρ 和 L 代入式(13)中, 可计算出 TPA 算法的具体表达式。

3.3 性能评估

为了评估 TPA-LSSVM 模型的性能, 将其与常

见的 LSSVM 模型、PSO-LSSVM 模型和 RBFNN 模型进行对比。

实验采用归一化平均绝对误差(R_{NMAE})和归一化均方误差(R_{NRMSE})作为评估指标。其中 R_{NMAE} 评估预测值的准确度, R_{NRMSE} 评估偏差的程度与变化。评价指标的定义为^[15]

$$R_{NMAE} = \frac{1}{N} \sum_{i=1}^N \frac{|x_i - \hat{x}_i|}{P} \times 100\% \quad (35)$$

$$R_{NRMSE} = \frac{1}{P} \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2} \times 100\% \quad (36)$$

式中, N 为测试样本大小; P 表示数据集的风力发电容量, 在本文实验中取 967 MW; x_i 为真实值, \hat{x}_i 为预测值。

3.4 预测分析

由图 1 可以看出, TPA-LSSVM 模型的性能明显优于其他预测模型。随着提前预测时间的增加, 模型的 R_{NRMSE} 出现不同程度上升, 但 TPA-LSSVM 模型的增幅最小。仿真结果表明, TPA-LSSVM 模型的偏差最小, 数据更加集中。

由图 2 可知, TPA-LSSVM 模型与其他模型相比, 预测效果更优, 预测精度更高。随着提前预测时间的增加, 模型的误差逐渐变大。仿真结果表明, 在不同的时刻, TPA-LSSVM 模型的 R_{NMAE} 均为最小, 预测精度最高。

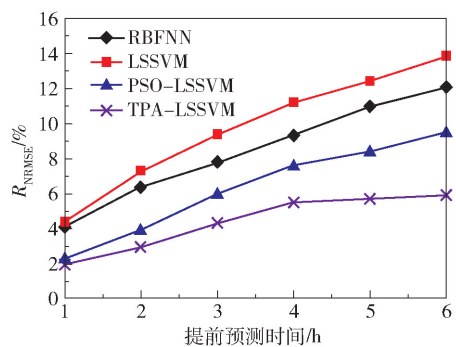


图 1 不同提前预测时间的归一化均方误差
Fig. 1 Normalized mean square errors for different early prediction times

4 结论

惩罚因子和径向基宽度直接影响 LSSVM 模型的预测精度, 本文提出了一种双参数(TPA)算法, 从理论上证明任意初始值均可线性收敛到全局最优值。TPA 算法可防止早熟和欺骗现象的发生, 并且收敛速度快, 寻优精度高, 不易陷入局部极值。通过

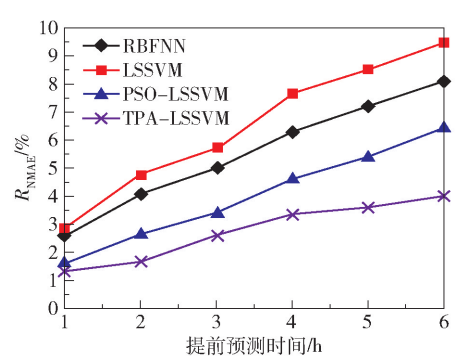


图 2 不同提前预测时间的归一化平均绝对误差
Fig. 2 Normalized mean absolute errors for different early prediction times

调用 TPA 算法对 LSSVM 模型的惩罚因子和径向基宽度进行寻优赋值, 并将训练好的 TPA-LSSVM 模型应用于风电预测中。仿真结果表明, 与 LSSVM 模型、PSO-LSSVM 模型和 RBFNN 模型相比, TPA 算法可以更好地实现 LSSVM 的参数寻优, TPA-LSSVM 模型能有效提高预测精度。在今后的研究工作中可将 TPA 算法与其他寻优算法进行组合, 以提升算法的性能, 从而为商业应用提供更多的选择。

参考文献:

- [1] KUMPF A, TOST B, BAUMGART M, et al. Visualizing confidence in cluster-based ensemble weather forecast analyses [J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 24(1): 109-119.
- [2] WARRINGTON J, DREW D, LYGEROS J. Low-dimensional space- and time-coupled power system control policies driven by high-dimensional ensemble weather forecasts [J]. IEEE Control Systems Letters, 2018, 2(1): 1-6.
- [3] YAN J, ZHANG H, LIU Y Q, et al. Forecasting the high penetration of wind power on multiple scales using multi-to-multi mapping[J]. IEEE Transactions on Power Systems, 2018, 33(3): 3276-3284.
- [4] YAN J, LI F R, LIU Y Q, et al. Novel cost model for balancing wind power forecasting uncertainty [J]. IEEE Transactions on Energy Conversion, 2017, 32(1): 318-329.
- [5] JIAO R H, HUANG X J, MA X H, et al. A model combining stacked auto encoder and back propagation algorithm for short-term wind power forecasting [J]. IEEE Access, 2018, 6: 17851-17858.
- [6] MAHNAZ M, AKHIL K, NALESH S, et al. A hardware architecture for radial basis function neural network classifier[J]. IEEE Transactions on Parallel and Distributed

- Systems, 2018, 29(3): 481–495.
- [7] WU X H, ZUO W M, LIN L, et al. F-SVM: combination of feature transformation and SVM learning via convex relaxation [J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(11): 5185–5199.
- [8] CHENG R J, SONG Y D, CHEN D W, et al. Intelligent localization of a high-speed train using LSSVM and the online sparse optimization approach [J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(8): 2071–2084.
- [9] CORUS D, OLIVETO P S. Standard steady state genetic algorithms can hillclimb faster than mutation-only evolutionary algorithms [J]. IEEE Transactions on Evolutionary Computation, 2018, 22(5): 720–732.
- [10] ANAMIKA, PEESAPATI R, KUMAR N. Electricity price forecasting and classification through wavelet-dynamic weighted PSO-FFNN approach [J]. IEEE Systems Journal, 2018, 12(4): 3075–3084.
- [11] 候旭珂, 杨宏伟, 马方, 等. 一种新的广义鲁棒主成分分析(GRPCA)算法研究及应用 [J]. 北京化工大学学报(自然科学版), 2018, 45(4): 82–85.
- HOU X K, YANG H W, MA F, et al. A new generalized robust principal component analysis (GRPCA) algorithm [J]. Journal of Beijing University of Chemical Technology(Natural Science), 2018, 45(4): 82–85. (in Chinese)
- [12] 易军凯, 王超, 李辉. 面向文本分类的深度置信网络特征提取方法研究 [J]. 北京化工大学学报(自然科学版), 2018, 45(3): 90–94.
- YI J K, WANG C, LI H. A feature extraction method for text categorization based on a deep belief network (DBN) [J]. Journal of Beijing University of Chemical Technology(Natural Science), 2018, 45(3): 90–94. (in Chinese)
- [13] LESSARD L, RECHT B, PACKARD A. Analysis and design of optimization algorithms via integral quadratic constraints [J]. SIAM Journal on Optimization, 2016, 26(1): 57–95.
- [14] KATHRYN D S. National oceanic and atmospheric administration repository of meteorological databases [EB/OL]. [2018-07-28]. <http://www.cdc.noaa.gov/cdc/data.ncep.reanalysis.derived.html>.
- [15] ZHANG Y G, CHEN B, ZHAO Y, et al. Wind speed prediction of IPSO-BP neural network based on Lorenz disturbance [J]. IEEE Access, 2018, 6: 53168–53179.

Optimization of a wind power time series prediction model based on a two-parameter least squares support vector machine

LIU Yun YI Song*

(Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: The advantages and disadvantages of the wind power time series prediction model directly affect the application value of wind power. The least squares support vector machine (LSSVM) has obvious advantages in dealing with wind power prediction. This paper proposes a two-parameter algorithm (TPA), which shows that any initial value can converge linearly to the global optimal value. The two-parameter (TPA) algorithm is employed to optimize the penalty factor and radial base width of the LSSVM model, and the trained TPA-LSSVM model is then applied to wind power prediction. The simulation results show that compared with the LSSVM model, the particle swarm least squares support vector machine (PSO-LSSVM) model and the radial basis function neural network (RBFNN) model, the TPA algorithm can better realize the parameter optimization of LSSVM, and thus the TPA-LSSVM model can effectively improve the prediction accuracy.

Key words: least squares support vector machine; time series prediction; two-parameter algorithm; parameter optimization

(责任编辑:吴万玲)