

# 基于小波匹配的新小波支持向量机预测模型

彭献永<sup>1</sup>, 吴奇<sup>2</sup>

(1. 艾默生过程控制有限公司, 上海 201206)

(2. 上海交通大学电子信息与电气工程学院, 上海 200240)

**[摘要]** 针对产品销售时序具有多维、小样本、非线性、随机性等特征, 已有的支持向量核不可能精确逼近任意的销售时序曲线。将小波理论应用于支持向量核函数, 并对标准支持向量机进行修正, 形成一种新的小波支持向量机(WN- $\nu$ -SVM)。设计了自适应正态变异粒子群算法(ANPSO)对小波支持向量机模型参数进行辨识, 并进行了汽车销量预测的实例分析。结果表明, 基于 WN- $\nu$ -SVM 模型的短期预测方法是有效可行的, 具有理论意义和实用价值。

**[关键词]** 自适应变异, 正态变异, 小波核, 支持向量机, 短期预测

**[中图分类号]** TP391 **[文献标志码]** A **[文章编号]** 1672-1292(2019)02-0050-09

## The Forecasting Model Based on Matched Wavelets and Wavelet Kernel Support Vector Machine

Peng Xianyong<sup>1</sup>, Wu Qi<sup>2</sup>

(1. Emerson Process Management Co., Ltd., Shanghai 201206, China)

(2. School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

**Abstract:** Aiming at the characters of multi-dimension, small sample, nonlinearity, randomness of the time series of product sales, the existing support vector kernel does not accurately approximate any time series curve of the sales. A new wavelet support vector machine(WN- $\nu$ -SVM) is proposed on the basis of the combination between wavelet theory and the modified support vector machine. An adaptive and normal mutation particle swarm optimization(ANPSO) algorithm is designed to select the best parameter of WN- $\nu$ -SVM model. The application results of vehicle sales prediction case show that the short-term forecasting approach based on the WN- $\nu$ -SVM model is more effective and feasible.

**Key words:** adaptive mutation, normal mutation, wavelet kernel, support vector machine, short-term forecasting

短期产品销售预测通常是指以日、周、月为时间单位对销售指标进行预测, 是企业生产规划、库存控制的前提与基础, 对于企业的“市场需要多少, 生产多少”的“零库存”策略意义重大。

产品销售预测方法大致分为两大类, 对于线性时序, 以自回归方法为主, 其代表模型主要有 ARMA、ARIMA、ARCH、GARCH 等<sup>[1-3]</sup>; 对于非线性时序, 以神经网络为主的人工智能方法, 其代表模型主要有 BP、RBF、FNN 等<sup>[4-5]</sup>。在这些方法中, 对于含有季节性、非线性、随机性等特征的时序, 神经网络由于其可以逼近非线性目标函数, 预测效果比自回归模型要好<sup>[6-7]</sup>, 因而受到广泛关注, 也是研究最经常使用的方法之一。然而, 神经网络由于其存在学习速度慢、泛化能力差、往往会陷入局部极小等缺点, 正逐渐被支持向量机 SVM<sup>[8]</sup>所取代。

作为一种较新的机器学习算法, 支持向量机能够较好地解决小样本、非线性、高维数和局部极小等实际问题<sup>[9-11]</sup>。值得指出的是, Yan 和 Xu 建立了一种新的模糊支持向量机, 成功运用于产品开发时间估计的研究。但在实际应用中, 也体现出其不足之处, 目前已有的支持向量核都是  $L^1$  空间的, 因此, 理论上现有的支持向量机是不可能逼近平方可积空间  $L^2$  上的任意函数。这种基的不完备性导致了支持向量机也不能

收稿日期: 2019-02-21.

基金项目: 国家自然科学基金(61671293).

通讯联系人: 彭献永, 博士, 工程师, 研究方向: 复杂系统的性能分析、建模与优化控制. E-mail: pengxy2003@163.com

逼近任意的目标曲线<sup>[12]</sup>.

本文将小波理论应用于支持向量机的核函数,形成小波核函数,可通过平移生成  $L^2$  空间上的一组完备基,并对标准支持向量机表达方法进行修正,获取对偶问题更简洁、求解效率更高的一种新的小波支持向量机(Wavelet New  $\nu$ -Support Vector Machine, WN- $\nu$ -SVM). 为了对 WN- $\nu$ -SVM 模型参数进行优选,本文设计了一种改进粒子群算法,即自适应正态变异粒子群算法(ANPSO). 该算法的主要思想为:将自适应变异和正态变异引入标准粒子群算法,形成一种基于迭代进程和适应度值综合的自适应变异粒子群算法,算法按照自适应方式从迭代进程和目标函数适应度值上调整速度惯性因子,同时结合正态变异算子调整粒子的搜索方向,最终获取全局最优解. 本文将 WN- $\nu$ -SVM 模型与产品销售预测的特点相结合,得到一种新的产品销售的预测模型.

## 1 WN- $\nu$ -SVM 模型

对于小波支持向量机的研究,分为两个研究领域:一方面 Stanislaw 和 Konrad<sup>[13]</sup>等利用小波将时序分解成低频系数和低频系数,然后利用支持向量回归机分别对低频系数和高频系数进行预测,最后对低频和高频的预测值进行重构. 该方法充分考虑到数据的内在特性,尽可能将数据分解成高频部分和低频部分,给出的预测结果较好,但该方法存在分解与重构的误差,且没有充分连续小波函数在构建高维空间中的作用;另一方面 Liu<sup>[14]</sup>等用母小波函数作为支持向量机的核函数,这样就不存在分解与重构的误差,其研究主要集中在标准最小二乘支持向量机的基础上,且选择小波核函数具有主观性,没有从时序的内在特征出发,选择与当前时序相匹配的母小波函数. 目前对这两个领域的研究主要集中在以小波分解和重构为基础的支持向量机预测方面.

本文将小波与原始时序特征进行匹配,遴选与当前时序最吻合的母小波函数,在此基础上将小波理论与  $\nu$ -支持向量机方法相融合,发挥各自特长,形成一种新的支持向量机(WN- $\nu$ -SVM).

### 1.1 支持向量核函数条件

根据广义线性判别函数的思路,要解决一个非线性问题,可通过非线性变换将输入空间变换到一个高维特征空间,在此特征空间中寻求最优或广义最优分类面,且只需进行内积运算,而没有必要知道采用的是非线性变换的形式. 只要在变换后的特征空间的内积可以用原空间中的变量直接计算得到,则即使特征空间中的维数增加很多,在其中求解最优分类面问题并未增加多少计算复杂度.

根据 Hilbert-Schmidt 原理,只要一种运算满足 Mercer 条件,其就可作内积使用.

**引理 1**<sup>[15-17]</sup> 点积核  $L^2(\mathbf{R}^n)$  下的对称函数  $K(\mathbf{x}, \mathbf{x}')$  为特征空间中的内积的充分必要条件是,对于任意的  $\varphi(\mathbf{x}) \neq 0$  且  $\int \varphi^2(\mathbf{x}) d\mathbf{x} < \infty$ , 条件成立.

$$\iint K(\mathbf{x}, \mathbf{x}') \varphi(\mathbf{x}) \varphi(\mathbf{x}') d\mathbf{x} d\mathbf{x}' \geq 0, \quad \mathbf{x} \text{ 为 } n \text{ 维列向量.} \quad (1)$$

**引理 2**<sup>[15-17]</sup> 平移不变核  $L^2(\mathbf{R}^n)$  下的对称函数  $K(\mathbf{x}, \mathbf{x}')$ , 平移函数满足  $K(\mathbf{x}, \mathbf{x}') = K(\mathbf{x} - \mathbf{x}')$ , 若满足 Mercer 条件,则其为一个允许支持向量核.

引理 1 提供了判断和构建核函数的方法. 但实际上很难将平移不变核分解成两个函数的内积,引理 2 给出了平移不变核函数是一个支持向量核的充分必要条件.

**定理 1** 平移不变核函数  $K(\mathbf{x}, \mathbf{x}') = K(\mathbf{x} - \mathbf{x}')$  是一个允许支持向量核,当且仅当  $K(\mathbf{x})$  的傅立叶变换成立.

$$F[\mathbf{x}](\omega) = (2\pi)^{-n/2} \int_{\mathbf{R}^n} \exp(-j(\omega \cdot \mathbf{x})) K(\mathbf{x}) d\mathbf{x} \geq 0, \quad \omega \in \mathbf{R}, \mathbf{x} \text{ 为 } n \text{ 维列向量.} \quad (2)$$

### 1.2 小波核函数

将母小波函数替换成某些具体形式,即可得到相应的小波核函数.

**定理 2** 设母小波函数为  $\psi(\mathbf{x})$ ,  $a$  为尺度因子,  $b$  为位移因子,其中  $\mathbf{x} \in \mathbf{R}^n$ ,  $a, b \in \mathbf{R}$ . 则满足 Mercer 定理的点积形式的小波核函数:

$$K(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^n \psi\left(\frac{\mathbf{x}_i - b_i}{a}\right) \psi\left(\frac{\mathbf{x}'_i - b'_i}{a}\right) \quad (3)$$

和满足平移不变核定理的小波核函数:

$$K(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^n \psi\left(\frac{\mathbf{x}_i - \mathbf{x}'_i}{a}\right) \tag{4}$$

均是可容许的支持向量核.

式(4)中没有出现位移因子  $b$ , 但有一个变量  $\mathbf{x}'$ , 该变量的作用等价于位移因子. 因此, 可用小波函数来构造具有平移不变性的小波核函数.

**定理 3** Morlet 母小波函数为  $\psi(x) = \cos(\omega_0 x) \exp(-x^2/2)$ ,  $\omega_0, x \in \mathbf{R}$ , 则小波核函数:

$$K(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^n \cos\left(\omega_0 \frac{\mathbf{x}_i - \mathbf{x}'_i}{a}\right) \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}'_i\|^2}{2a^2}\right), \mathbf{x} \text{ 为 } n \text{ 维列向量} \tag{5}$$

是一种可容许的支持向量核.

**证明** 根据定理 1, 只需证明

$$F[\mathbf{x}](\omega) = (2\pi)^{-n/2} \int_{R^n} \exp(-j(\omega \cdot \mathbf{x})) K(\mathbf{x}) d\mathbf{x} \geq 0, \tag{6}$$

式中,  $K(\mathbf{x}) = \prod_{i=1}^n \psi\left(\frac{\mathbf{x}_i}{a}\right) = \prod_{i=1}^n \cos\left(\frac{\omega_0 \mathbf{x}_i}{a}\right) e^{-\left(\frac{\|\mathbf{x}_i\|^2}{2a^2}\right)}$ .

因为,

$$\begin{aligned} \int_{R^n} \exp(-j\omega \mathbf{x}) K(\mathbf{x}) d\mathbf{x} &= \int_{R^n} \exp(-j\omega \mathbf{x}) \left( \prod_{i=1}^n \cos\left(\omega_0 \frac{\mathbf{x}_i}{a}\right) \exp\left(-\frac{\|\mathbf{x}_i\|^2}{2a^2}\right) \right) d\mathbf{x} = \\ &= \prod_{i=1}^n \int_{-\infty}^{\infty} \exp(-j\omega_i \mathbf{x}_i) \left( \frac{\exp\left(j\omega_0 \frac{\mathbf{x}_i}{a}\right) + \exp\left(-j\omega_0 \frac{\mathbf{x}_i}{a}\right)}{2} \right) \exp\left(-\frac{\|\mathbf{x}_i\|^2}{2a^2}\right) d\mathbf{x}_i = \\ &= \prod_{i=1}^n \frac{1}{2} \int_{-\infty}^{\infty} \left( \exp\left(-\frac{\|\mathbf{x}_i\|^2}{2} + \left(\frac{\omega_0 j}{a} - j\omega_i a\right) \mathbf{x}_i\right) + \exp\left(-\frac{\|\mathbf{x}_i\|^2}{2} - \left(\frac{\omega_0 j}{a} + j\omega_i a\right) \mathbf{x}_i\right) \right) = \\ &= \prod_{i=1}^n \frac{|a| \sqrt{2\pi}}{2} \left( \exp\left(-\frac{(\omega_0 - \omega_i a)^2}{2}\right) + \exp\left(-\frac{(\omega_0 + \omega_i a)^2}{2}\right) \right), \end{aligned} \tag{7}$$

将式(7)代入式(6)可得:

$$F[\mathbf{x}](\omega) = \prod_{i=1}^n \left( \frac{|a|}{2} \right) \left( \exp\left(-\frac{(\omega_0 - \omega_i a)^2}{2}\right) + \exp\left(-\frac{(\omega_0 + \omega_i a)^2}{2}\right) \right), \tag{8}$$

当  $a \neq 0$  时, 有

$$F[\mathbf{x}](\omega) \geq 0, \tag{9}$$

证毕.

同理, 可以证明以下 Mexico 小波核函数也是可容许的支持向量核. Mexico 小波为:

$$\psi(x) = (1-x^2) \exp\left(-\frac{x^2}{2}\right), x \in \mathbf{R}. \tag{10}$$

其小波核函数为:

$$K(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^n \left( 1 - \frac{\|\mathbf{x}_i - \mathbf{x}'_i\|^2}{a^2} \right) \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}'_i\|^2}{2a^2}\right), \mathbf{x} \text{ 为 } n \text{ 维列向量}. \tag{11}$$

与普通核函数相比较, 小波核函数的特点使用低维的计算量得到了高维的结果, 在未增加计算复杂度的同时得到了性质更好的高维投影.

### 1.3 WN- $\nu$ -SVM

对标准  $\nu$ -SVM 最优化问题加以改造, 将  $b$  的影响放入置信范围, 获取新的支持向量机, 即 W- $\nu$ -SVM. 对样本集  $T = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)\}$ , 其中  $\mathbf{x}_i \in \mathbf{R}^d, y_i \in \mathbf{R}, \mathbf{x}_i$  为  $d$  维列向量,  $i = 1, 2, \dots, l$ , WN- $\nu$ -SVM 的数学表示为:

$$\min_{\mathbf{w}, \xi^{(*)}, b, \varepsilon} \tau(\mathbf{w}, \xi^{(*)}, \varepsilon) = \frac{1}{2} (\|\mathbf{w}\|^2 + b^2) + C \cdot \left( \nu \cdot \varepsilon + \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*) \right), \tag{12}$$

$$\text{s.t. } ((\mathbf{w}^T \cdot \mathbf{x}_i) + b) - y_i \leq \varepsilon + \xi_i, \quad (13)$$

$$y_i - ((\mathbf{w}^T \cdot \mathbf{x}_i) + b) \leq \varepsilon + \xi_i^*, \quad (14)$$

$$\xi^{(*)} \geq 0, v \in (0, 1], \varepsilon \geq 0, i = 1, 2, \dots, l. \quad (15)$$

式中,  $\mathbf{w}$  为  $d$  维列向量;  $C > 0$  为惩罚系数, 用以控制模型复杂性与训练误差之间的平衡;  $v \in (0, 1]$  为控制支持向量个数的参数;  $\varepsilon$  为控制管道大小的参数;  $\xi^{(*)} = \{\xi_1, \xi_1^{(*)}, \dots, \xi_i, \xi_i^{(*)}, \dots, \xi_l, \xi_l^{(*)}\}$ , ( $i = 1, \dots, l$ ) 为松弛变量, 保证了约束的满足。

采用对偶原理、KKT 条件和小波核函数技巧, 可以求得上述优化问题的对偶问题:

$$\max_{\alpha^{(*)}} \mathbf{w}(\alpha^{(*)}) = -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) (K(\mathbf{x}_i \cdot \mathbf{x}_j) + 1) + \sum_{i=1}^l (\alpha_i^* - \alpha_i) y_i, \quad (16)$$

$$\text{s.t. } 0 \leq \alpha_i^{(*)} \leq \frac{C}{l}, \quad (17)$$

$$\sum_{i=1}^l (\alpha_i + \alpha_i^*) \leq C \cdot v, \quad i = 1, 2, \dots, l. \quad (18)$$

选择适当的  $C$  和  $v$ , 选择在某个尺度范围内能够与目标时序变化特征相吻合的小波函数作为 WN- $v$ -SVM 的核函数, 则 WN- $v$ -SVM 输出函数:

$$f(\mathbf{x}) = \sum_{i=1}^m (\alpha_i - \alpha_i^*) \left( \prod_{i=1}^n \psi \left( \frac{\mathbf{x}^j - \mathbf{x}_i^j}{a} \right) + 1 \right), \quad (19)$$

式中,  $\mathbf{x}^j$  表示输入向量第  $j$  个分量;  $\mathbf{x}_i^j$  表示第  $i$  个训练样本的第  $j$  个分量。

从式(12)~(19)可以发现 WN- $v$ -SVM 与标准  $v$ -SVM 的区别:

- (1) WN- $v$ -SVM 最优化问题置信范围比标准  $v$ -SVM 的最优化问题置信范围多一个  $\frac{1}{2}b^2$  项;
- (2) WN- $v$ -SVM 对偶问题比标准  $v$ -SVM 的对偶问题少了一个等式约束;
- (3) WN- $v$ -SVM 和标准  $v$ -SVM 的结构相同, 其区别在于其所用到的核函数不同. 标准支持向量机常用的核函数有多项式核函数、径向基核函数和感知器核函数 3 种; 而小波支持向量机所用的核函数为满足定理 2 的所有小波函数, 如 Morlet 母小波函数和 Mexico 母小波函数;
- (4) WN- $v$ -SVM 输出的决策函数比标准  $v$ -SVM 输出的决策函数少了一个参数  $b$ , 节约了模型参数寻优过程中的迭代时间, 提高了求解效率。

## 2 ANPSO 算法

### 2.1 问题的粒子群表示

WN- $v$ -SVM 模型的未知参数有小波尺度范围  $a$ 、惩罚参数  $C$ 、控制支持向量  $v$ , 这些参数的变化区间都是连续的, 目标函数的自变量具有多维、小样本、非线性、多峰等特点, 因而普通的进化算法无法处理好, 由此本文提出一种智能优化算法, 即自适应正态变异粒子群算法 (ANPSO)。

粒子群算法 (particle swarm optimization, PSO) 具有模型简单、可调参数少、实现方便等特点, 目前已被广泛应用于函数优化、神经网络训练、先进控制、模式识别、信息融合、组合优化等领域<sup>[18-19]</sup>, 其中连续空间优化问题一直是应用热点。

与其他进化算法类似, PSO 算法在解决连续空间优化问题尤其是优化具有高维复杂地形函数时, 往往存在着容易陷入局部最优、过早收敛而导致进化停滞的问题. 其根本原因是由于进化初期, 种群中的超级粒子吸引其他粒子越来越接近, 种群迅速向其周围聚集, 使得种群多样化大幅降低. 同时由于粒子越来越接近种群局部最优点, 并不断更新自身的最优位置, 粒子的记忆部分和社会认知部分的作用逐渐减弱, 粒子的飞行速度逐渐趋近于 0, 于是收敛速度大大降低甚至出现进化停滞, 因此很难发现分布在局部最优领域内的全局最优点. 而许多优化函数往往具有高维、多峰、地形复杂等特点, 表现为全局最优点附近分布着许多局部最优点 (实际工程应用中的局部最优点可能是数值噪声等多种原因造成), 面对这类函数优化问题, 基本粒子群算法很容易早熟收敛到局部解, 因此优化精度难以提高. 针对产品销售时序具有多维、小样本、非线性、多峰等特点, 将 WN- $v$ -SVM 模型参数组成的向量作为粒子位置, 利用自适应正态变异

粒子群算法对销售时序进行训练,获取最优 WN- $\nu$ -SVM 模型参数组合  $C, \nu, a$ .

粒子群优化算法首先初始化一群随机粒子,通过迭代找到最优解.在每一次迭代中粒子通过两个极值来更新自己:一个是粒子到目前为止发现自己的最好位置( $p_{best}$ ),另一个是整个种群到目前为止发现的最好位置( $g_{best}$ ).

粒子找到上述 2 个极值以后,根据以下公式来更新自己的速度和位置:

$$\mathbf{v}_{ij}^{k+1} = \mathbf{w}_{ij}^k \cdot \mathbf{v}_{ij}^k + c_1 \cdot \text{rand}() \cdot (\mathbf{p}_{ij} - \mathbf{x}_{ij}^k) + c_2 \cdot \text{rand}() \cdot (\mathbf{p}_{g_j} - \mathbf{x}_{ij}^k), \quad (20)$$

$$\mathbf{x}_{ij}^{k+1} = \mathbf{x}_{ij}^k + \mathbf{v}_{ij}^{k+1}, \quad (21)$$

式中,  $i=1, 2, \dots, N$ ,  $N$  为该群体中粒子的总数;  $\mathbf{v}_{ij}^k$  为第  $k$  次迭代粒子  $i$  飞行速度矢量的第  $j$  维分量;  $\mathbf{x}_{ij}^k$  为第  $k$  次迭代粒子  $i$  位置矢量的第  $j$  维分量;  $\mathbf{p}_{ij}$  为粒子  $i$  个体最好位置  $p_{best}$  的第  $j$  维分量;  $\mathbf{p}_{g_j}$  为群体最好位置  $g_{best}$  的第  $j$  维分量;  $c_1, c_2$  为权重因子;  $\text{rand}()$  为随机函数,产生  $[0, 1]$  的随机数;  $\mathbf{w}_{ij}^k$  为第  $k$  次迭代第  $i$  个粒子速度的惯性权重矢量第  $j$  维分量.

### 2.2 自适应正态变异粒子群算法(ANPSO)

针对标准粒子群算法的不足,引进自适应变异算子<sup>[19]</sup>,从目标函数适应度值上调整速度惯性因子,同时结合正态分布变异算子<sup>[20]</sup>调整搜索方向.自适应变异是实数编码下的一种高效的变异算子,其将解的质量与变异算子相关联.关联的方式有 2 种,其一是按照适应度值进行关联,适应度大的个体在较小的范围内变异,而适应度较小的个体在较大的范围内变异.其二是按照迭代进程进行关联,这种适应度是根据迭代次数来改变某些参数,即在迭代次数小时,参数在较大的范围内变异;随着迭代次数增加,被变异参数逐渐变小,在较小的范围内变异,粒子在小范围内搜索.基于这种考虑,具有自适应正态变异功能的粒子群算法中的速度、位置迭代公式为:

$$\mathbf{v}_{ij}^{k+1} = (1-\lambda) \cdot \mathbf{w}_{ij}^k \cdot \mathbf{v}_{ij}^k + \lambda \cdot N(0, \sigma_i^k) + c_1 \cdot \text{rand}() \cdot (\mathbf{p}_{ij} - \mathbf{x}_{ij}^k) + c_2 \cdot \text{rand}() \cdot (\mathbf{p}_{g_j} - \mathbf{x}_{ij}^k), \quad (22)$$

$$\mathbf{x}_{ij}^{k+1} = \mathbf{x}_{ij}^k + \mathbf{v}_{ij}^{k+1}, \quad (23)$$

$$\mathbf{w}_{ij}^k = \beta \cdot \left( 1 - \frac{f(\mathbf{x}_i^k)}{f(\mathbf{x}_m^k)} \right) + (1-\beta) \cdot \mathbf{w}_{ij}^0 \cdot \exp(-\alpha k^2), \quad (24)$$

$$\sigma_i^{k+1} = \sigma_i^k \cdot \exp(N_i(0, \Delta\sigma)), \quad (25)$$

式中,  $f(\mathbf{x}_i^k)$  为第  $k$  代第  $i$  个粒子的适应度值;  $f(\mathbf{x}_m^k)$  为粒子群在第  $k$  代最佳适应度;  $\mathbf{w}_{ij}^0$  为第  $i$  个粒子速度的惯性权重矢量第  $j$  维分量的初值;  $\lambda$  为增量系数  $\Delta\sigma$  为二次步长控制参数;  $N_i(0, \Delta\sigma)$  为均值为 0、方差为  $\Delta\sigma$  的符合正态高斯分布的随机数;  $\sigma_i^k$  为第  $k$  代时摄动向量的第  $i$  个分量;  $\alpha$  为控制粒子衰减速度的控制因子;  $\beta$  为控制迭代进程变异与每一代的基于适应度变异的自适应因子.

式(24)右端第一个式子是表达基于适应度迭代的变异,适应度较大的个体在较小的范围内变异,适应度较小的个体在较大的范围内变异.第二个式子是表示基于迭代次数的变异,迭代次数小的时候,参数在较大的范围内变异,粒子在大范围内搜索;随着迭代次数增大,被变异的参数逐渐减小,在较小的范围内变异,粒子在小范围内搜索,逐渐到达全局最优值.

式(25)是修正速度变化的正态变异算子<sup>[20]</sup>,在正态变异策略中,群体的个体是由一个解向量  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  和一个摄动向量  $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_n)$  描述,该摄动向量是速度向量的控制向量,且其自身也要进行不断的变异.

采用实数自适应正态变异算子的粒子群算法可以对群体多样性损失进行恢复,提高了算法的全局搜索能力.

## 3 相关算法

根据上述自适应正态变异算子优化粒子群算法的思想,本文提出的 APSO 算法步骤描述如下:

**算法 1** 自适应正态变异粒子群算法(PSO Algorithm)

**Step 1** 初始化算法的基本参数:种群规模 pop\_size,种群的维度 pop\_nvar,最大迭代次数 MaxIt,精度要求 ErrGoal,初始惯性权重  $\mathbf{w}_i^0$ ,最初权重因子  $c_1$  和  $c_2$ ,初始正态分布的方差  $\Delta\sigma$ ,初始摄动量  $\sigma_i^0$ ,控制粒子衰减速度的控制因子  $\alpha$ ,自适应因子  $\beta$ ,增量系数  $\lambda$ .

**Step 2** 初始化粒子的位置  $\mathbf{x}_i^0$ , 计算各个粒子的适应度值, 把当前粒子作为各粒子的个体极值点, 把粒子群适应度值最小的粒子作为全局极值点.

**Step 3** 采用自适应变异(式(24))和具有二级步长的正态变异(式(25))对粒子群进行变异操作, 用式(22)和(23)更新粒子的位置.

**Step 4** 根据更新后的位置, 计算粒子群中每一个粒子的适应度值, 把本次更新后的最优个体与上次迭代的最优个体进行比较, 更新全局最优初始适应度值.

**Step 5** 当满足搜索精度或者达到最大搜索次数则转入 Step 6; 否则转入 Step 3.

**Step 6** 退出循环, 输出全局极值和最优粒子.

**算法 2** 利用 ANPSO 算法对 WN- $\nu$ -SVM 模型参数进行辨识, 获取最优参数组合  $(C, \nu, a)$ . 基于 ANPSO-WN- $\nu$ -SVM 模型的产品销售预测的步骤如下:

**Step 1** 对由销售时序和影响因素组成的多维样本数据进行归一化处理, 形成样本集.

**Step 2** 对销售时序进行多尺度小波变换, 选择与原始时序相吻合的小波函数和尺度变化范围.

**Step 3** 按照式(4)构造小波函数  $K(\mathbf{x}, \mathbf{x}')$ , 建立形如式(12)的最优化目标函数.

**Step 4** 调用算法 1, 求解优化目标函数问题并判断 KKT 条件, 得到参数  $\alpha^{(*)}$ .

**Step 5** 对新的输入  $\mathbf{x}$ , 根据式(19)计算预测值.

## 4 算例研究

以汽车销售时序预测为例, 考虑 6 个影响因素: 品牌知名度 BF、性能参数 PP、外形美观度 FB、销售人员经验数 SC、居民可用于购买存款数  $n_d$  和油价  $n_p$ , 前 4 个为语言型数据, 后 2 个为数值型数据. 选取 60 个汽车销售时序实例进行研究, 其中 48 个作为训练样本, 其余 12 个作为测试数据.

利用 Matlab7.1 编制相应的模型训练和预测程序, 运行环境为 Core(TM)2 CPU 1.80 GHz、内存 1.00 G 的微机. 为比较不同预测方法的效果, 自适应粒子群的适应度函数采用均方误差 MSE 表示. 最大迭代次数为 100, 归一化后样本的预测精度为 0.005, 初始惯性权重为  $w^0 = 0.9$ , 初始正态分布方差为  $\Delta\sigma = 0.5$ , 自适应因子  $\beta = 0.8$ , 控制粒子衰减速度的控制因子  $\alpha = 2$ , 权重因子  $c_1 = c_2 = 2$ .

利用算法 2 的步骤 1 和 2 获取在不同尺度范围内汽车销售时序的小波变换. 本文选择的小波函数有 Morlet, Haar, Mexh 和 Gaus. 为精减篇幅, 本文只给出有代表性的 Morlet 和 Mexican 母小波函数在不同尺度下的小波变换, 如图 1~3 所示. 研究发现在尺度  $a \in [0.3, 2]$  范围内, Mexican 母小波要优于 Morlet 母小波, 更能与原始销售时序相吻合.

因此, 可以确定 WN- $\nu$ -SVM 模型的核函数为 Mexican 母小波, 3 个模型参数范围可确定如下:  $\nu \in [0, 1]$ ,  $C \in \left[ \frac{\max(x_i) - \min(x_i)}{l}, a \in [0.3, 2] \right]$ . 利用算法 1 获取 WN- $\nu$ -SVM 最优参数组合为:  $C = 269.48$ ,  $\nu = 0.82$  和  $\alpha = 0.56$ . 利用算法 2 对汽车销售时序进行预测, 其结果如图 4 所示.

为检验 ANPSO-WN- $\nu$ -SVM 模型的预测能力, 本文将其与 ARMA、PSO-W- $\nu$ -SVM、PSO-WN- $\nu$ -SVM 这 3 个模型预测结果进行比较. 为反映误差分布, 采用最近 12 个月的预测结果作为误差统计的基础. 4 种不

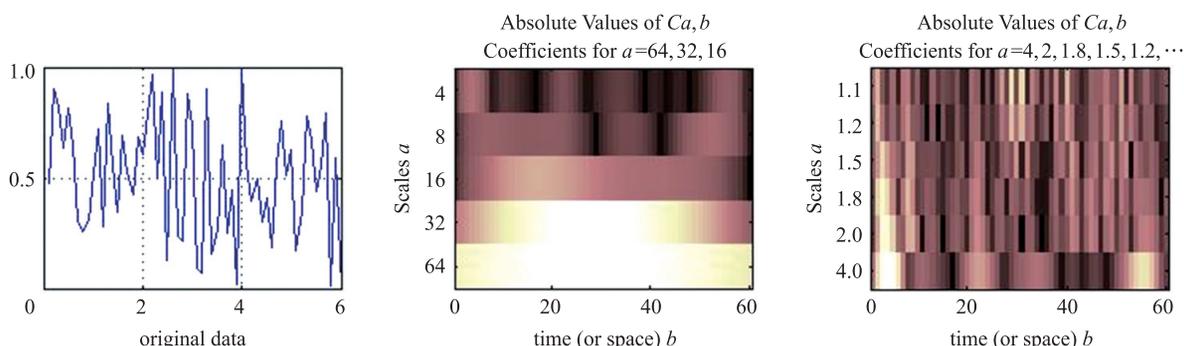


图 1 销售时序在不同尺度下的 Mexican 小波变换

Fig. 1 Mexican wavelet transform of sales time series in the scope of different scale

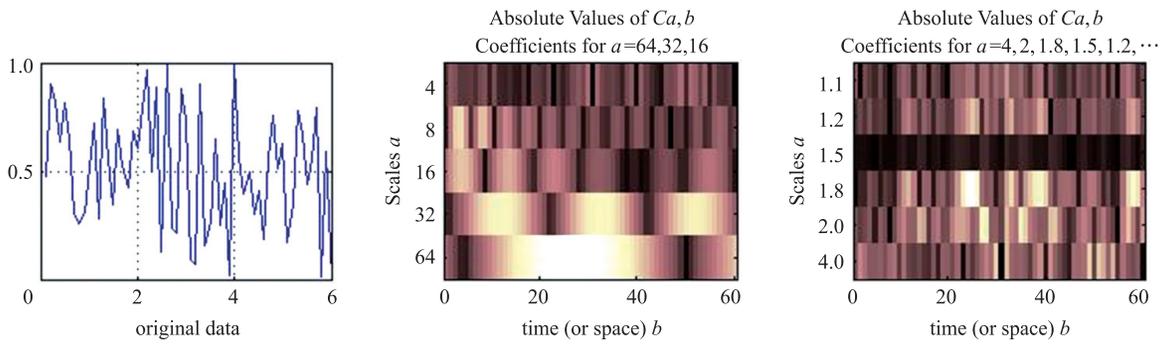


图 2 销售时序在不同尺度下的 Morlet 小波变换

Fig. 2 Morlet wavelet transform of sales time series in the scope of different scale

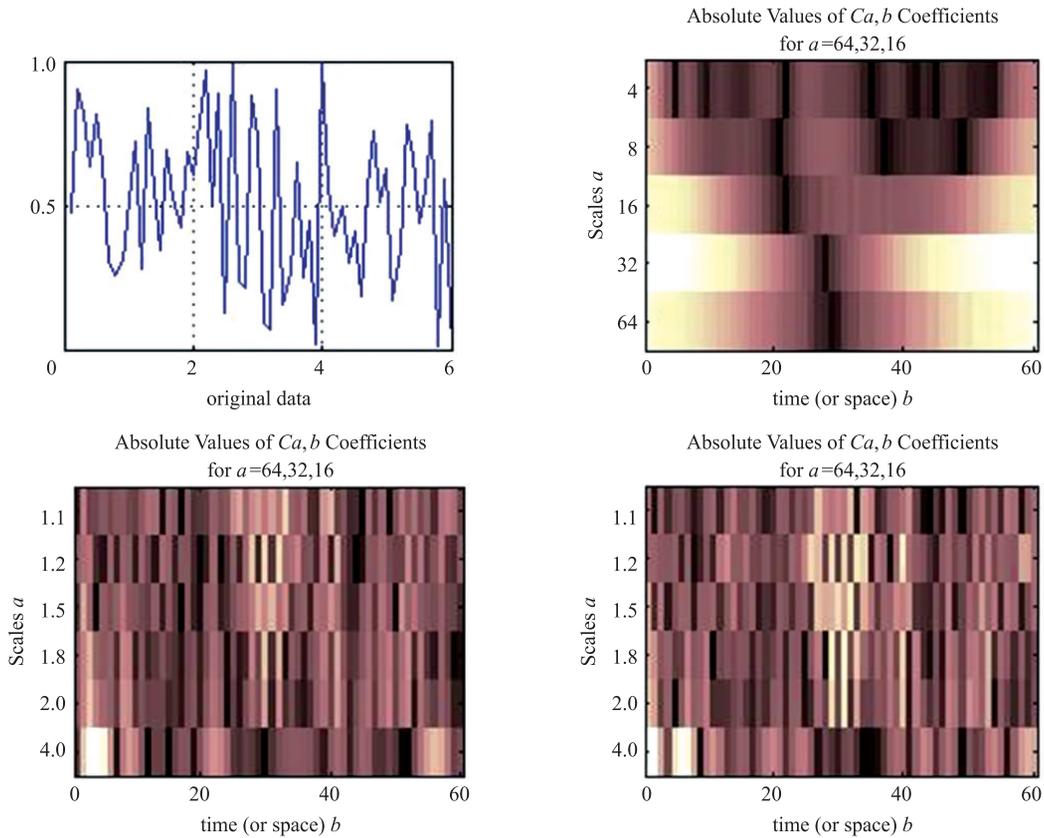


图 3 销售时序在不同尺度下的 Gaussian 小波变换

Fig. 3 Gaussian wavelet transform of sales time series in the scope of different scale

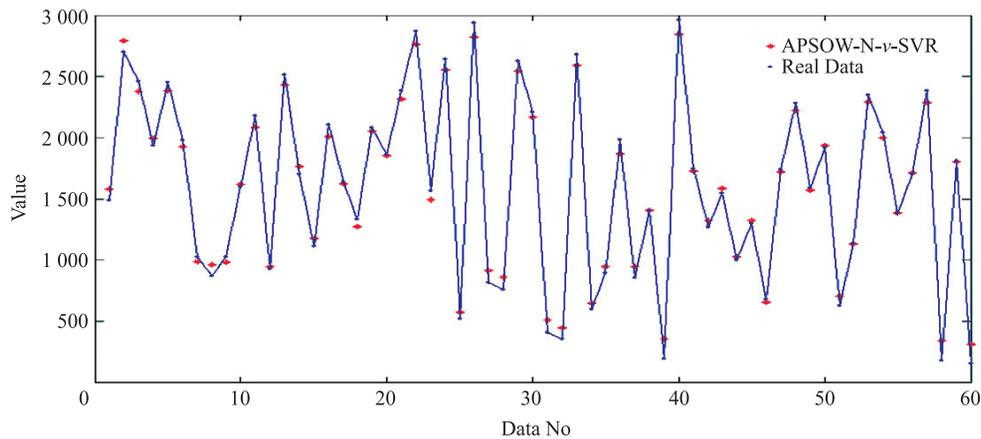


图 4 基于 ANPSO-W-ν-SVM 模型的汽车销售预测效果

Fig. 4 The car sales forecasting result based on ANPSO and WN-ν-SVM model

同模型的预测结果如表 1 所示. 其中 PSO-W- $\nu$ -SVM 模型是指惯性权重线性变化的粒子群算法优选标准小波支持向量机(W- $\nu$ -SVM)模型参数,  $w \in [0.1, 0.9]$ ; PSO-WN- $\nu$ -SVM 是指惯性权重线性变化的粒子群算法优选小波支持向量机(WN- $\nu$ -SVM),  $w \in [0.1, 0.9]$ ; ANPSO-WN- $\nu$ -SVM 是指自适应正态变异粒子群算法优选小波支持向量机(WN- $\nu$ -SVM).

表 1 4 种不同模型预测比较

Table 1 Forecasting result comparison of four different models

最近 12 月	实际值	预测值			
序号	原始时序	ARMA	PSO-W- $\nu$ -SVM	PSO-WN- $\nu$ -SVM	ANPSO-WN- $\nu$ -SVM
1	151	1 588	330	322	308
2	1 808	1 824	1 789	1 819	1 805
3	177	1 554	359	352	337
4	2 382	1 616	2 309	2 301	2 287
5	1 703	1 651	1 733	1 725	1 711
6	1 383	1 581	1 407	1 399	1 385
7	2 042	1 529	2 037	2 124	1 997
8	2 350	1 659	2 315	2 307	2 293
9	1 139	1 748	1 123	1 131	1 134
10	627	1 599	725	717	703
11	1 922	1 634	1 930	1 922	1 936
12	1 589	1 548	1 548	1 556	1 570

对表中 4 种不同模型的预测结果进行误差指标处理, 获取的误差指标分布如表 2 所示. 误差指标有平均绝对误差 MAE、均方误差 MSE、平均绝对百分比误差 MAPE, 其公式分别为:  $MAE = \sum |y_i - \hat{y}_i|/n$ ;  $MSE = \sum (y_i - \hat{y}_i)^2/n$ ;  $MAPE = \sum |y_i - \hat{y}_i| \cdot 100/(n \cdot |\hat{y}_i|)$ , 其中  $y_i$  表示真实值,  $\hat{y}_i$  表示预测值. 从表 2 可发现, PSO-W- $\nu$ -SVM 模型的误差指标 MAE 要好于 PSO-WN- $\nu$ -SVM 模型, 误差指标 MAPE 和 MSE 要稍差于 PSO-WN- $\nu$ -SVM 模型. 而采用自适应正态变异的 ANPSO-WN- $\nu$ -SVM 模型的误差指标 MAE, MAPE 和 MSE 均好于 PSO-W- $\nu$ -SVM, PSO-WN- $\nu$ -SVM 这 2 个模型. 可见, 自适应正态变异算子能够提高粒子的搜索能力和求解精度.

表 2 各种预测模型的误差统计

Table 2 Error statistic of four forecasting models

模型	MAE	MAPE	MSE
ARMA	580.00	1.717	560 910
PSO-W- $\nu$ -SVM	59.200	0.209	7 099
PSO-WN- $\nu$ -SVM	61.000	0.201	7 093
ANPSO-WN- $\nu$ -SVM	53.400	0.182	5 915

## 5 结语

产品销售预测是企业精益生产和库存控制的基础. 随着市场竞争加剧, 面对日益增加的产品复杂性和多样性, 企业迫切需要系统化的产品销售预测方法. 传统的模型无法精确逼近具有小样本、随机性、季节性、非线性等特征的产品销售时序曲线. 本文提出的基于 ANPSO-WN- $\nu$ -SVM 模型的短期预测方法, 可以在反映数据内在特性原貌的基础上, 在  $L^2$  空间逼近任意销售时序. 算例分析表明本文的短期预测方法具有较好的性能, 易于在工程实际中应用.

### [参考文献] (References)

- [1] BOX G E P, JENKINS G M. Time series analysis: forecasting and control[M]. 3rd ed. Englewood Cliffs, NJ: Prentice-Hall Inc, 1994.
  - [2] ENGLE R F. Combining competing forecasts of inflation using a bivariate ARCH model[J]. Journal of economic dynamics and control, 1984, 18(2): 151-165.
  - [3] 林燕平, 窦万峰. 基于 ARIMA 模型的城市公共自行车需求量短期预测方法研究[J]. 南京师范大学学报(工程技术版), 2016, 16(3): 36-40.
- LIN Y P, DOU W F. Research on short-term prediction method of demand number in urban public bicycle based on the ARIMA

- model[J]. Journal of Nanjing normal university(engineering and technology edition),2016,16(3):36-40.(in Chinese)
- [4] TANG Z,FISHWICK P A. Feedforward neural nets as models for time series forecasting[J]. ORSA journal of computing,1993,5(4):374-385.
- [5] HILL T,CONNOR M O,REMUS W. Neural network models for time series forecasts[J]. Management science,1996,42(7):1082-1092.
- [6] ZHANG G P. An investigation of neural networks for linear time-series forecasting[J]. Computers and operations research,2001,28(12):1183-1202.
- [7] TEALAB A. Time series forecasting using artificial neural networks methodologies;a systematic review[J]. Future computing and informatics journal,2018,3(2):334-340.
- [8] VAPNIK V. The nature of statistical learning[M]. New York;Springer,1995.
- [9] JEDLINSKI L,JONAK J. Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform[J]. Applied soft computing,2015,30:636-641.
- [10] MORALES N,TOLEDO J,ACOSTA L. Path planning using a multiclass support vector machine[J]. Applied soft computing,2016,43:498-509.
- [11] GUO J H,LIU Z,HUANG W, et al. Short-term traffic flow prediction using fuzzy information granulation approach under different time intervals[J]. IET intelligent transport systems,2018,12(2):143-150.
- [12] YAN H S,XU D. An approach to estimating product design time based on fuzzy  $\nu$ -support vector machine[J]. IEEE transactions on neural networks,2007,18(3):721-731.
- [13] STANISLAW OSOWSKI,KONRAD G. Forecasting of the daily meteorological pollution using wavelets and support vector machine[J]. Engineering applications of artificial intelligence,2007,20:745-755.
- [14] LIU B S,LI Y J,XING Z W, et al. Research on freight traffic forecast based on wavelet and support vector machine[C]//International Conference on Machine Learning and Cybernetics. Dalian,China:IEEE,2006:2524-2530.
- [15] ZHANG L,ZHOU W D,JIAO L C. Wavelet support vector machine[J]. IEEE transactions on systems,man and cybernetics,2004,34(1):34-39.
- [16] KHANDOKER A H,LAI D T H,BEGG R K, et al. Wavelet-based feature extraction for support vector machines for screening balance impairments in the elderly[J]. IEEE transactions on neural systems and rehabilitation engineering,2007,15(4):587-597.
- [17] WIDODO A,YANG B S. Wavelet support vector machine for induction machine fault diagnosis based on transient current signal[J]. Expert systems with applications,2008,35(1/2):307-316.
- [18] MEZAI L,HACHOUF F. Score-level fusion of face and voice using particle swarm optimization and belief functions[J]. IEEE transactions on human-machine systems,2015,45(6):761-772.
- [19] XU F,CHEN H,GONG X, et al. Fast nonlinear model predictive control on FPGA using particle swarm optimization[J]. IEEE transactions on industrial electronics,2016,63(1):310-321.
- [20] KRUSIENSKI D J. A modified particle swarm optimization algorithm for adaptive filtering[C]//International Symposium on Circuits and Systems. Kos,Greece:IEEE,2006.

[责任编辑:严海琳]