

文章编号: 1001-0920(2008)03-0357-04

## 基于小波变换和 AR-LSSVM 的非平稳时间序列预测

王晓兰<sup>a,b</sup>, 张万宏<sup>a</sup>, 王慧中<sup>a</sup>

(兰州理工大学 a. 电气工程与信息工程学院, b. 甘肃省有色金属新材料国家重点实验室, 兰州 730050)

**摘要:** 提出一种基于二进正交小波变换和 AR-LSSVM 方法的非平稳时间序列预测方案. 首先利用 Mallat 算法对非平稳时间序列进行分解和重构, 分离出非平稳时间序列中的低频信息和高频信息; 然后对高频信息构建自回归模型, 对低频信息则用最小二乘支持向量机进行拟合; 最后将各模型的预测结果进行叠加, 从而得到原始序列的预测值. 研究表明, 该方法不仅能充分拟合低频信息, 而且可避免对高频信息的过拟合.

**关键词:** 小波变换; 非平稳时间序列; 最小二乘支持向量机; 自回归; 预测

中图分类号: TP18

文献标识码: A

## Non-stationary time series prediction based on wavelet analysis and AR-LSSVM

WANG Xiaolan<sup>a,b</sup>, ZHANG Wan-hong<sup>a</sup>, WANG Hui-zhong<sup>a</sup>

(a. School of Electrical and Information Engineering, b. State Key Laboratory of Gansu Advanced Non-ferrous Metal Materials, Lanzhou University of Technology, Lanzhou 730050, China. Correspondent: ZHANG Wan-hong, E-mail: zhangwanhong@mail2.lut.cn)

**Abstract:** A non-stationary time series prediction method using the wavelet analysis and AR-LSSVM is proposed. By wavelet decomposition and reconstruction, the non-stationary time series are decomposed into a low frequency signal and several high frequency signals. The high frequency signals are predicted with auto-regression models, and the low frequency is predicted with least square support vector machines. The prediction result of the original time series is the superimposition of the respective prediction. This new method avoids the over-fitted for high frequency signals, and adequately fits the low signal of the non-stationary time series, so better predicting performance can be obtained. Experiments show the effectiveness of the predicting method.

**Key words:** Wavelet transform; Non-stationary time series; Least squares support vector machines; Auto-regression; Prediction

### 1 引言

时间序列是动力系统演化过程的外在表现,也是动力系统分析和控制的基础.在实际问题中,时间序列几乎都是非平稳和非线性的,尽管无法知道那些系统变量如何相互作用来支配系统的演化,但可以肯定的是系统变量具有不同的时频特性.对于动力学系统而言,建立良好的系统模型非常重要,它不仅具有重要的理论意义,而且具有广阔的应用前景.

目前,国内外学者从不同途径对动力学系统的辨识与预测进行研究,统计分析是比较常用的辨识与预测方法.在传统的统计学中,各种方法和结论只

有样本趋向无穷大时,其性能才有理论上的保证,而在众多工程实际中,样本数目常是有限的,甚至样本数很少,这导致很多传统的统计分析方法难以取得理想的效果.神经网络以其较好的自学习和函数逼近能力,较多地应用于时间序列预测<sup>[1,2]</sup>,但由于训练过程遵循经验风险最小化准则,存在过拟合,估计参数多,训练过程受局部极小点的困扰,网络结构的选择过分依赖于经验等缺陷,影响和限制了网络的泛化能力<sup>[3]</sup>.基于统计学习理论的支持向量机(SVM)<sup>[4,5]</sup>,根据结构风险最小化原则,克服了上述方法的不足.研究发现,支持向量机既能较好地解决小样本、过学习、高维数、局部最小等实际难题,又有

收稿日期: 2006-11-30; 修回日期: 2007-03-08.

基金项目: 甘肃省自然科学基金项目(ZS022-A25-039).

作者简介: 王晓兰(1963—),女,兰州人,教授,从事计算机控制、时间序列预测等研究;张万宏(1983—),男,甘肃会宁人,硕士生,从事时间序列预测控制等研究.

很强的泛化能力。

对于均值具有趋向性的非平稳时间序列预测,除了模型的选择外,关键取决于如何提取时间序列中的低频和高频成分并进行建模,以及如何避免对高频信息的过拟合.为了解决这类非平稳时间序列的预测问题,考虑到二进正交小波分解对非平稳性时间序列的适应性<sup>[6]</sup>,对低频的分离作用及支持向量机较好的泛化能力,本文提出一种基于小波分析和 AR-LSSVM 的非平稳时间序列预测方法,采用二进小波变换提取原始序列中的高频信息和低频信息,对高频信息和低频信息分别建模,在充分拟合低频信息的同时,避免对高频信息的过拟合,从而提高了非平稳时间序列的预测精度.

### 2 预测原理及方法

#### 2.1 最小二乘支持向量机

最小二乘支持向量机(LSSVM)是由标准支持向量机演变而来的一种算法<sup>[7,8]</sup>.标准支持向量机需要求解凸二次规划问题,虽然所得的解是唯一的最优解,但算法的复杂度依赖于样本数据的个数,样本数据量越大,计算速度越慢,占用的内存也越大.LSSVM 通过解一组线性方程组取代标准支持向量机中二次规划优化,提高了收敛速度,比标准支持向量机的计算复杂性低.回归型 LSSVM 的提法如下:

对于一个给定的训练样本集  $(x_i, y_i), x_i \in R^n, y_i \in R, i = 1, 2, \dots, l$ ,利用非线性映射  $\phi$ 把数据集从输入空间映射到高维特征空间,以使输入空间中的非线性拟合问题变成高维特征空间中的线性拟合问题.高维特征空间的线性回归函数为

$$f(x, w) = w^T \phi(x) + b. \tag{1}$$

其中:  $w$  和  $\phi(x)$  为  $n$  维向量,  $b$  为阈值.

根据结构风险最小化原理,综合考虑函数复杂度和拟合误差,回归问题可表示为约束优化问题

$$\min_{w,b,e} J(w, e) = \frac{1}{2} w^T w + \frac{1}{2} \sum_{i=1}^l e_i^2; \tag{2}$$

s. t.  $y_i = w^T \phi(x_i) + b + e_i, i = 1, 2, \dots, l.$

其中:  $e_i$  表示误差, 为正则化参数.为求解上述优化问题,把约束优化问题变成无约束优化问题,建立 Lagrange 函数

$$L(w, b, e, \lambda) = \frac{1}{2} w^T w + \frac{1}{2} \sum_{i=1}^l e_i^2 - \sum_{i=1}^l \lambda_i [w^T \phi(x_i) + b + e_i - y_i]. \tag{3}$$

于是,最优解的条件为

$$\begin{cases} w = \sum_{i=1}^l \lambda_i \phi(x_i), & \lambda_i = 0, & \lambda_i = re_i, \\ y_i = w^T \phi(x_i) + b + e_i. \end{cases} \tag{4}$$

其中  $\lambda_i$  是 Lagrange 乘子.根据 KKT 最优条件,对于

$i = 1, 2, \dots, l$ ,消去  $e_i$  和  $w$  后,得到线性方程组

$$\begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} 0 & e_i^T \\ e_i & + \frac{1}{r} I \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ y \end{bmatrix}. \tag{5}$$

其中:  $Y = [y_1, y_2, \dots, y_l]^T, \lambda = [\lambda_1, \lambda_2, \dots, \lambda_l]^T, e_i = [1, 1, \dots, 1]^T, \Phi = (\phi(x_i) \phi(x_j))_{l \times l}$  为核函数矩阵,  $I$  为  $l \times l$  单位矩阵.

根据 Hilbert-Schmidt 原理,通过引入满足 Mercer 条件的核函数  $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ ,将变换空间中的内积转化为原空间中某个函数的计算,从而间接求解输入空间向高维特征空间的映射  $\phi$ .有多种核函数  $K(x_i, x_j)$  可供选择,典型的核函数有多项式核函数、径向基核函数(RBF)、Sigmoid 核函数等,本文采用最常用的 RBF 核函数.解式(5)可得最小二乘支持向量机回归模型

$$y(x) = \sum_{i=1}^l \lambda_i K(x, x_i) + b. \tag{6}$$

LSSVM 只需确定核函数的形状参数和惩罚系数,而不需选取不敏感损失函数的值,这既简化了计算,又便于实际使用.

#### 2.2 二进正交小波变换的 Mallat 算法

Mallat 从空间概念方面形象地说明了多分辨率分析特性,并给出了信号分解不同频率通道的算法及重构算法<sup>[9]</sup>.

##### 2.2.1 分解算法<sup>[9]</sup>

若将  $\omega$  理解为待分解的离散信号,则根据分解算法有

$$c_{j+1} = Hc_j, d_{j+1} = Gc_j, j = 0, 1, \dots, J. \tag{7}$$

其中:  $H$  和  $G$  分别为低通滤波器和高通滤波器;  $c_{j+1}$  和  $d_{j+1}$  分别为原始信号在分辨率  $2^{-(j+1)}$  下的低频信号和高频信号,是原始信号在相邻不同频率段上的成分;  $J$  表示最大的分解层数.最终将待分解信号  $\omega$  分解为  $d_1, d_2, \dots, d_j$  和  $c_j$ .

该分解算法利用二抽取,使每层分解比分解前的信号数据长度减半,而总输出数据长度与输入数据  $\omega$  长度保持一致.信号个数的减少对预测是不利的,但经 Mallat 算法分解后的信号可采用重构算法进行二插值重构.

##### 2.2.2 重构算法<sup>[9]</sup>

$$c_j = H^* c_{j+1} + G^* d_{j+1}. \tag{8}$$

其中:  $H^*$  和  $G^*$  是对偶算子;  $j = J - 1, J - 2, \dots, 0$ .采用式(8)对小波分解后的信号进行重构可以增加信号个数.对  $d_1, d_2, \dots, d_j$  和  $c_j$  分别进行重构,得到  $D_1, D_2, \dots, D_j$  和  $C_j$ ,有

$$X = D_1 + D_2 + \dots + D_j + C_j. \tag{9}$$

其中:  $D_1 = \{d_{1,1}, d_{1,2}, \dots\}, \dots, D_j = \{d_{j,1}, d_{j,2}, \dots\}$  为

第一层至第  $J$  层的高频信号重构;  $C_i : \{c_{i,1}, c_{i,2}, \dots\}$  为第  $J$  层低频信号重构. 在 Mallat 重构算法中利用二插值,即在输入数据序列每相邻数据之间补一个零,以使数据长度增加一倍,从而恢复二抽取前的数据长度.

### 2.3 基于 Mallat 算法的 AR-LSSVM 预测方法

该方法首先利用二进正交分解与重构的 Mallat 算法,对非平稳时间序列进行分解与重构,然后对低频和高频信息分别建模,而原始序列的预测结果为低频和高频预测结果的叠加. 预测框架如图 1 所示.

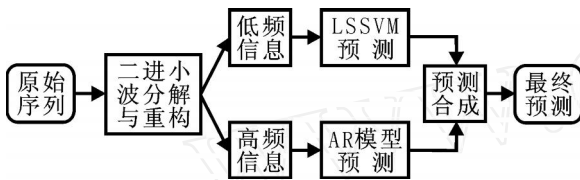


图 1 基于 Mallat 算法的 AR-LSSVM 预测框架

令  $X_i = \{x_1, x_2, \dots\}$  为具有趋向性的非平稳时间序列,对其进行二进小波分解,并对分解后的各层时间序列分别进行重构,得到

$$X = G_1 + G_2 + \dots + G_N + X_N. \quad (10)$$

其中:  $G_1 = \{g_{1,1}, g_{1,2}, \dots\}, \dots, G_N = \{g_{1,N}, g_{2,N}, \dots\}$  为第一层至第  $N$  层高频信号的重构;  $X_N = \{x_{N,1}, x_{N,2}, \dots\}$  为第  $N$  层低频信号的重构. 有

$$X_i = g_{1,i} + g_{2,i} + \dots + g_{N,i} + x_{N,i}. \quad (11)$$

若已知  $\{t_i / i = M\}$  时刻的  $X_i$  值,要预测  $k$  步以后的状态值,则需求解  $\bar{X}_{M+k}$ ,即

$$\bar{X}_{M+k} = \bar{g}_{1,M+k} + \bar{g}_{2,M+k} + \dots + \bar{g}_{N,M+k} + \bar{x}_{N,M+k}. \quad (12)$$

由于  $G_1, G_2, \dots, G_N$  可近似看作平稳时间序列,可分别建立  $AR(p)$  模型,并对  $\bar{g}_{1,M+k}, \bar{g}_{2,M+k}, \dots, \bar{g}_{N,M+k}$  进行预测. 其步骤如下:

- 1) 对  $G_j$  建立  $AR(p)$  模型,并利用已知的  $g_{j,i}$  对模型进行参数估计,其中  $1 \leq j \leq N, i \leq M$ ;
- 2) 对上述所建模型进行适用性检验;
- 3) 依据检验后适合的  $AR(p)$  对  $g_{j,M+k}$  进行预测,得到  $g_{j,M+k}$  的预测值  $\bar{g}_{j,M+k}$ .

由于低频信号  $X_N$  表示长期趋势,且具有非线性,可采用 LSSVM 方法对  $X_N$  进行学习、建模和预测. 根据上述 LSSVM 原理,其预测模型的具体学习算法如下:

- 1) 算法初始化:选择适合的序列嵌入维数  $m$ , 并利用 Cross-validation 法选择核函数的超参数及调节常数,这里的序列嵌入维数  $m$  定义为:预测序列某一点值时预测模型需用此点以前的数据点的个数;
- 2) 数据预处理:对低频信号  $X_N$  进行归一化预

处理,有

$$\hat{x}_i = 2 \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} - 1, \quad (13)$$

归一化后  $\hat{x}_i \in [-1, 1]$ ;

3) 训练集和测试集的选取:按照选择的嵌入维数  $m$ ,选取 LSSVM 预测模型所需的输入向量  $X_{i \times m}$  和输出向量  $Y_{i \times 1}$ ;

4) 模型求解:利用式 (5) 求解模型的参数  $(a, b)^T$ .

最后将各个预测结果进行叠加,便得到原始时间序列的预测结果.

### 3 预测实例

上证指数是具有趋向性的非平稳时间序列,也是用于比较和判别预测方法的典型数据. 采用上证指数来验证该预测算法,并将其与单纯的 AR 模型、BP 神经网络预测结果进行比较. 预测验证前,首先采用 DB4 小波对原始序列进行 3 层分解与重构,然后进行归一化处理,并使用 AR-LSSVM 进行预测. 为了评价模型的预测性能,采用预测均方根误差为性能指标,即

$$MSE = \sqrt{\frac{1}{M} \sum_{k=1}^M [y(k) - \hat{y}(k)]^2}. \quad (14)$$

其中:  $M$  为预测样本数,  $y(k)$  为序列的真实值,  $\hat{y}(k)$  为预测结果.

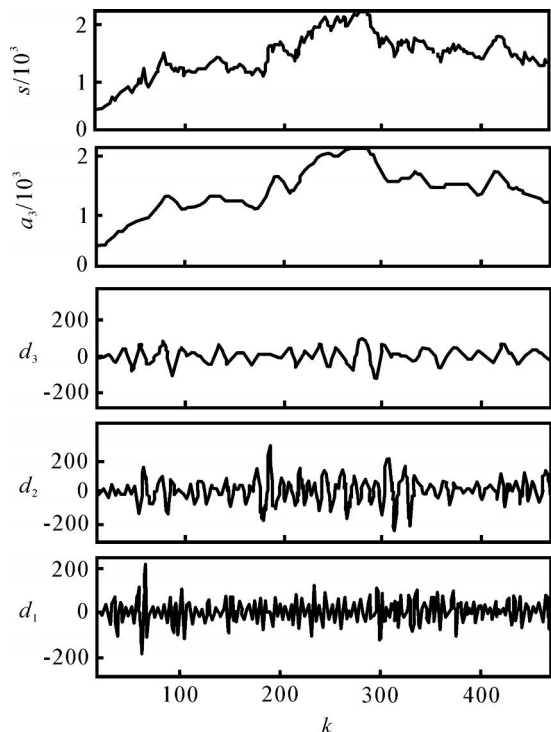


图 2 上证指数原始序列及 3 层小波分解

上证指数原始序列及 3 层分解的结果如图 2 所示. 其中训练样本 400 点,测试样本 60 点,分别采用基于小波变换的 AR-LSSVM, AR 模型和 BP 神经网络

络进行一步前向预测,预测结果如表1所示.其中BP神经网络的输入层节点为6,隐层节点为10,输出节点为1.AR模型的阶数通过反复实验,从中选取预测性能最好的模型来确定.

表1 AR-LSSVM,AR模型和BP神经网络预测结果

预测方法	AR-LSSVM	AR(6)	BP神经网络
MSE	0.012	0.061	0.095

从表1可以得出:AR-LSSVM的预测精度优于AR(6)和BP神经网络的预测精度,从而说明该方法泛化能力较强,预测效果较好,不仅能充分拟合低频信息,而且可避免对高频的过拟合.

#### 4 结 语

本文将基于统计学习理论的支持向量机和小波变换应用于非平稳时间序列的预测.首先根据趋向性非平稳时间序列固有的确定性、非线性和波动性,采用小波变换分离出非平稳时间序列中的非线性低频趋势成分和波动成分;然后利用对平稳时间序列具有较好预测能力的统计推断建模,并结合最小二乘支持向量机,实现对均值具有非平稳特性的时间序列的精确预测.研究结果表明,该方法是这类非平稳时间序列的有效预测方法,对进一步解决实际工程问题具有良好的应用前景.

#### 参考文献(References)

- [1] Cristea P, Tuduce R, Cristea A, et al. Time series prediction with wavelet neural networks[J]. Seminar Neural Network Application in Electrical Engineering, 2000, 5(9): 25-27.
- [2] Ashraf A M Khalaf, Kenji Nakayama. Time series prediction using a hybrid model of neural network and FIR filter[J]. Neural Networks Proc, 1998, 3(5): 1975-1980.
- [3] Vapnik V N. The nature of statistical learning theory [M]. 2nd ed. New York: Springer-Verlag, 1999.
- [4] Cortes C, Vapnik V N. Support vector network[J]. Machine Learning, 1995, 20(3): 273-295.
- [5] 许建华, 张学工, 李衍达. 支持向量机的新发展[J]. 控制与决策, 2004, 19(5): 481-484.  
(Xu Jian-hua, Zhang Xue-gong, Li Yan-da. Advances in support vector machines [J]. Control and Decision, 2004, 19(5): 481-484.)
- [6] Zhang Bai-ling, Richard Coggins. Multiresolution forecasting for future trading using wavelet decomposition[J]. Neural Networks, 2001, 12(4): 765-775.
- [7] 叶美盈, 汪晓东, 张浩然. 基于在线最小二乘支持向量机回归的混沌时间序列预测[J]. 物理学报, 2005, 54(6): 2568-2573.  
(Ye Mei-ying, Wang Xiao-dong, Zhang Hao-ran. Chaotic time series forecasting using online least squares support vector machine regression [J]. Acta Physica Sinica, 2005, 54(6): 2568-2573.)
- [8] 阎威武, 邵惠鹤. 支持向量机和最小二乘支持向量机的比较及应用研究[J]. 控制与决策, 2003, 18(3): 358-360.  
(Yan Wei-wu, Shao Hui-he. Application of support vector machines and least squares support vector machines to heart disease diagnoses [J]. Control and Decision, 2003, 18(3): 358-360.)
- [9] 胡昌华, 李国华, 刘涛, 等. 基于MATLAB 6.x的系统分析与设计——小波分析[M]. 西安: 西安电子科技大学出版社, 2004.  
(Hu Chang-hua, Li Guo-hua, Liu Tao, et al. Design and analysis of series based on MATLAB 6.x: Wavelet analysis[M]. Xi'an: Xidian University Press, 2004.)
- [13] Makowski L, Ostroy J M. Perfect competition and the creativity of the market [J]. J of Economic Literature, 2001, 39(2): 479-535.
- [14] 杜义飞, 李仕明, 林光平. 讨价还价与供应链的利润最大化均衡[J]. 中国管理科学, 2006, 14(1): 37-42.  
(Du Yi-fei, Li Shi-ming, Lin Guang-ping. Bargaining processes and equilibrium of profit-maximizing for supply chain [J]. Chinese J of Management Science, 2006, 14(1): 37-42.)
- [15] 杜义飞, 李仕明. 基于Rubinstein-Stahl模型的供应链中间产品定价研究[J]. 系统工程学报, 2006, 21(1): 33-37.  
(Du Yi-fei, Li Shi-ming. Research on pricing intermediate product of supply chain based on Rubinstein-Stahl model [J]. J of System Engineering, 2006, 21(1): 33-37.)

(上接第356页)