

新型 SVM 对时间序列预测研究^{*}

朱家元 段宝君 张恒喜

(空军工程大学工程学院飞机与发动机工程系 西安710038)

Prediction of Time Series Based on Least Squares Support Vector Machines

ZHU Jia-Yuan DUAN Bao-Jun ZHANG Heng-Xi

(Department of Aircraft and Engine Engineering, Air Force Engineering University, Xi'an 710038)

Abstract In this paper, we present a new support vector machines—least squares support vector machines (LS-SVMs). While standard SVMs solutions involve solving quadratic or linear programming problems, the least squares version of SVMs corresponds to solving a set of linear equations, due to equality instead of inequality constraints in the problem formulation. In LS-SVMs, Mercer condition is still applicable. Hence several type of kernels such as polynomial, RBF's and MLP's can be used. Here we use LS-SVMs to time series prediction compared to radial basis function neural networks. We consider a noisy (Gaussian and uniform noise) Mackey—Glass time series. The results show that least squares support vector machines is excellent for time series prediction even with high noise.

Keywords Machine learning, Support vector machines, Statistical learning theory, Neural networks, Time series prediction

1 引言

时间序列预测在工程、经济、工业制造、金融和管理等许多领域有着广泛而重要的应用价值。国内外许多学者采用了各种方法进行时间序列预测研究。近年来,神经网络在模式识别、决策和控制等领域得到了广泛的应用。一些学者采用了神经网络方法对时间序列预测进行了分析。文[1~4]分别探讨了BP网络、Hebbian网络、源分配网络和模糊神经网络在时间序列预测中的应用问题。但是神经网络受网络结构复杂性和样本复杂性的影响较大,有时会出现过学习或低泛化能力现象,影响预测精度和收敛速度。

Vapnik^[5]在1995提出一种新型统计学习方法—支持向量机(SVMs, Support Vector Machines),支持向量机具有完备的统计学习理论基础和出色的学习性能,已成为机器学习界的研究新热点,并在很多领域都得到了成功的应用^[6],如人脸检测、手写体数字识别、文本自动分类等。Vapnik提出支持向量机解决非线性函数估计问题,采用不敏感损失函数和Huber损失函数^[7]。文[6]提出一种新型支持向量机方法—最小二乘支持向量机(LS-SVMs, Least Squares Support Vector Machines)用于解决分类和函数估计问题。这种方法采用最小二乘线性系统作为损失函数,代替传统的支持向量机采用二次规划方法,并应用到模式识别和非线性函数估计,取得了较好的效果,运算简单、收敛速度快、精度高。因此,本文尝试采用最小二乘支持向量机对时间序列预测进行研究,并通过与采用RBF神经网络进行预测的结果进行对比,说明方法的有效性。

2 最小二乘支持向量机

采用最小二乘支持向量机进行函数估计的算法如下所

述。

设训练样本集 $D = \{(x_k, y_k) | k = 1, 2, \dots, N\}$, $x_k \in R^n$, $y_k \in R$, x_k 是输入数据, y_k 是输出数据。在权 w 空间(原始空间)中的优化问题可以描述为:

$$\min_{w, b, e} J(w, e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{k=1}^N e_k^2 \quad (1)$$

约束条件: $y_k = w^T \varphi(x_k) + b + e_k, k = 1, \dots, N$ 。

其中, $\varphi(\cdot): R^n \rightarrow R^h$ 是核空间映射函数, 权向量 $w \in R^h$ (原始空间), 误差变量 $e_k \in R$, b 是偏差量。损失函数 J 是SSE误差和规则化量之和, γ 是可调常数。核空间映射函数的目的是从原始空间中抽取特征, 将原始空间中的样本映射为高维特征空间中的一个向量, 以解决原始空间中线性不可分的问题。

根据优化函数(1)式, 定义拉格朗日函数:

$$L(w, b, e, a) = J(w, e) - \sum_{k=1}^N a_k \{w^T \varphi(x_k) + b + e_k - y_k\} \quad (2)$$

其中, 拉格朗日乘子 $a_k \in R$ 。对上式进行优化:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{k=1}^N a_k \varphi(x_k) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{k=1}^N a_k = 0 \\ \frac{\partial L}{\partial e_k} = 0 \rightarrow a_k = \gamma e_k \\ \frac{\partial L}{\partial a_k} = 0 \rightarrow w^T \varphi(x_k) + b + e_k - y_k = 0 \end{cases} \quad (3)$$

其中, $k = 1, \dots, N$ 。消除变量 w, e , 可得以下矩阵方程:

$$\begin{bmatrix} 0 & 1^T \\ 1, & \Omega + \frac{1}{\gamma} I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (4)$$

其中, $y = [y_1; \dots; y_N]$, $1 = [1; \dots; 1]$, $a = [a_1; \dots; a_N]$, $\Omega_{kl} = \varphi(x_k)^T \varphi(x_l)$

^{*} 此论文受国防预研资助基金资助(项目编号:98J19.3.2.JB3201)和空军重点型号工程课题资助。朱家元 博士研究生,研究方向:装备发展规划与管理,数字化综合保障工程,支持向量机与智能信息处理,便携式维修计算机研究。段宝君 博士,研究方向:装备发展规划与管理,数字仿真技术等。张恒喜 教授,博士生导师,研究方向:飞机系统工程,数字化综合保障工程。

$(x_k)^T \varphi(x_l), k, l=1, \dots, N$. 根据 Mercer 条件^[8], 存在映射函数 φ 和核函数 $K(\cdot, \cdot)$ 使得:

$$K(x_k, x_l) = \varphi(x_k)^T \varphi(x_l) \quad (5)$$

LS-SVM 最小二乘支持向量机的函数估计为:

$$y(x) = \sum_{k=1}^N a_k K(x, x_k) + b \quad (6)$$

其中, a, b 由式(4)求解出。核函数可以有不同形式, 如: 多项式核、多层感知高斯器(MLP)核、样条生成核、RBF 核等。在本文的时间序列预测研究中, 取 RBF 核函数, 即 $K(x_k, x_l) = \exp(-\frac{\|x_k - x_l\|^2}{2\sigma^2})$ 。最小二乘支持向量机的函数估计精度和收敛速度受 (γ, σ) 选择的影响。

3 时间序列预测分析

混沌是自然界与人类社会普遍存在的运动形态, 对于混沌动力系统的研究现已成为动力系统研究的中心内容之一。在这里, 我们采用最小二乘支持向量机对典型混沌时间序列, 即 Mackey-Glass 时间序列进行预测研究。Mackey-Glass

时间序列由差分延迟方程产生, 该方程定义为:

$$\frac{dx}{dt} = \beta x(t) + \frac{\alpha x(t-\delta)}{1+x(t-\delta)^{10}} \quad (7)$$

其中: $\beta = -0.1, \alpha = 0.2, \delta = 17$ 。本文采用四阶龙格-库塔法获得时间序列。按照如(8)式格式, 抽取1000个输入输出对数据, 构成训练向量集(实际是194个模式向量), 抽取1000个输入输出对数据, 构成测试向量集, 其中, $m=6, \tau=6$ 。然后分别对下一步 $x_{t+\tau}$ 和第100步 $x_{t+100\tau}$ 进行预测, 预测误差设定为 RMS。在最小二乘支持向量机中, 参数 γ 取10, 核函数取 RBF 核函数, 即 $K(x_k, x_l) = \exp(-\frac{\|x_k - x_l\|^2}{2\sigma^2})$, $\sigma^2 = 0.75$ 。

$$X(t) = [x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau}]^T \quad (8)$$

为了研究最小二乘支持向量机在混沌时间序列预测中的稳定性和抗干扰能力, 我们引入不同噪声率的高斯和均匀分布两种噪声到训练集中(见表1), 而测试向量集不含噪声。同时, 我们在相同条件下, 采用 RBF 神经网络进行了时间序列的预测, 以用于对比分析。对混沌 Mackey-Glass 时间序列预测结果如表1。

表1 LS-SVM 与 RBF 神经网络对混沌 Mackey-Glass 时间序列预测结果

噪声	高斯分布				均匀分布					
	22.15%		44.3%		6.2%		12.4%		18.6%	
RMS 误差	1步	100步	1步	100步	1步	100步	1步	100步	1步	100步
LS-SVM's	0.016	0.165	0.032	0.302	0.005	0.026	0.010	0.064	0.018	0.136
RBF	0.017	0.124	0.048	0.294	0.010	0.058	0.016	0.098	0.030	0.312

从 LS-SVM 与 RBF 神经网络对混沌时间序列预测结果可以得出, 对于高斯分布噪声情况, LS-SVM 与 RBF 神经网络的预测效果非常相似, 对于均匀分布噪声情况, LS-SVM 对时间序列的预测精度明显高于 RBF 神经网络, 特别是在高噪声的情况下。

结论 本文采用新型支持向量机—最小二乘支持向量机, 对时间序列预测进行了研究。通过对典型混沌 Mackey-Glass 时间序列进行的预测结果表明, 最小二乘支持向量机收敛速度快, 具有良好的预测能力和抗噪能力。因此, 采用最小二乘支持向量机进行时间序列是可行、有效的。可以适当引入到工程、经济、工业制造、金融和管理等许多领域的时间序列研究当中。

参考文献

- 1 Gencay R, Tung Liu. Nonlinear Modeling and Prediction with Feedforward and Recurrent Networks. *Physica D*, 1997, 108: 119~134

- 2 Deco G, Obradovic D. Decorrelated Hebbian Learning for Clustering and Function Approximation. *Neural Computation*, 1995(7): 338~348
- 3 Platt J. Learning by Combining Memorization and Gradient Descent. In *Advances in Neural Information Processing Systems*, 1996
- 4 Maguire L P, Roche B, McGinnity T M, Mcdaid L J. Predicting a Chaotic Time Series Using a Fuzzy Neural Network. *Information Sciences*, 1998, 112: 125~136
- 5 Vapnik V N. *Statistical Learning Theory*. New York, Springer, 1998
- 6 Suykens J A K, Vandewalle J, Moor D B. Optimal Control by Least Squares Support Vector Machines. *Neural Networks*, 2001, 14(1): 23~35
- 7 朱家元, 虞健飞, 张恒喜. 基于嵌入局部模型的 SOM 网络对混沌时间序列预测研究[J]. *控制与决策*, 2003, 1
- 8 朱家元, 张恒喜. 基于支持向量机的 R&D 项目中止决策研究. *计算机科学*, 2002, 29(9)