

基于在线小波支持向量回归的混沌时间序列预测^{*}

于振华[†] 蔡远利

(西安交通大学电子与信息工程学院, 西安 710049)

(2005 年 6 月 8 日收到, 2005 年 12 月 6 日收到修改稿)

混沌时间序列预测是非线性动力学研究中一个十分重要的问题, 支持向量回归方法为其提供了一种有效的解决思路. 通过分析新样本加入训练集后支持向量集的变化情况, 建立了一种混沌时间序列预测的支持向量回归算法, 具备了在线学习的特点. 同时, 针对混沌信号提出了一种满足小波框架的小波核函数, 它不但能以较高的精度逼近任意函数, 而且适合于混沌信号的局部分析, 提高了支持向量回归的泛化能力. 最后就 Mackey-Glass 混沌时间序列在线预测问题进行了大量仿真. 结果表明, 本文算法与现有的算法相比具有训练时间短、预测精度高等特点, 有一定的理论及实用价值.

关键词: 混沌时间序列, 支持向量回归, 在线学习, 小波核

PACC: 0545

1. 引 言

随着非线性混沌动力学理论和应用研究的不断深入, 混沌时间序列的建模与预测已成为混沌信息处理领域中的一个研究热点^[1,2], 并已在信号处理、自动控制等领域中得到了广泛的应用. 由于神经网络具有较强的非线性映射能力, 许多研究者利用神经网络对混沌时间序列进行预测与控制^[3-8], 但因神经网络存在过拟合、训练过程中易陷入局部极小、网络结构的选择过分依赖于经验等缺陷^[9], 直接影响了建模的精度及可靠性.

近年来, Vapnik^[9]提出的支持向量机(support vector machine, SVM)是在统计学习理论上建立起来的一种新的通用机器学习方法, 其特点是利用核函数将输入空间通过非线性变换映射到高维空间, 在这个新空间中求取最优线性分类超平面. 与神经网络相比, 其优点是不存在局部极小和维数灾难问题、泛化能力强, 可较好地克服神经网络固有的缺陷.

SVM 最初是用来解决模式识别问题, 现已推广到回归估计中, 称为支持向量回归(support vector regression, SVR), 并已应用在混沌系统的辨识与预测中^[10-13]. 但在现有的混沌时间序列预测的 SVR 训

练算法中, 训练样本以批处理的方式提供, 而对于实时在线的时间序列预测, 训练样本的总数未知, 样本是顺序获得的, 这样每次新样本的引入都必须从第一个样本开始重新学习. 在样本数比较少的时候, 训练还能够较短时间内完成; 当样本数目较多的时候, 学习效率就会急剧下降. 因此如何采用迭代方式进行在线学习是将 SVR 应用于实时混沌时间序列预测的关键.

近年来, 已提出了几种在线训练算法^[14-19], 其中大部分用于 SVM 分类, 对 SVR 的在线训练算法研究较少. 在现有的在线训练算法中, 有的算法只保留支持向量, 仅仅提供了一种近似解; 有的算法在支持向量数目较多时, 求解过程缓慢, 不适宜在线调整; 有的算法需要用户选择很多参数, 而确定这些参数的值也是一个非常复杂的问题. 本文以 SVM 分类算法为基础^[14,20], 提出一种用于混沌时间序列预测的在线 SVR 算法. 此外, 针对混沌信号的特点, 本文在多维小波核函数的基础上, 建立了一种满足小波框架条件的小波核函数, 将其应用于 SVR 中. 最后, 基于小波核函数的在线支持向量回归应用于 Mackey-Glass 混沌时间序列预测, 并与采用高斯核函数的传统 SVR 和文献 [19] 中的增量算法进行了比较, 仿真结果表明, 本文给出的算法具有预测精度高、训练时间短等优点, 适合于在线混沌时间序列预测.

^{*} 国家高技术研究发展计划(批准号 2003AA721070)资助的课题.

[†] E-mail: zhuayu@gmail.com

2. 支持向量回归

给定样本集 $\{(x_1, y_1), \dots, (x_l, y_l)\}$, 其中 $x_i \in R^n, y_i \in R$, 所要求的回归函数形式为

$$f(x) = (w^T \Phi(x)) + b, \quad (1)$$

式中 $w \in R^n, b \in R, \Phi(\cdot)$ 把输入样本从输入空间映射到高维特征空间. w 和 b 可通过解下列最优化问题得到:

$$\min_{w, b} R = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*), \quad (2)$$

(2) 式的约束条件为

$$y_i - (w^T \Phi(x) + b) \leq \varepsilon + \xi_i,$$

$$(w^T \Phi(x) + b) - y_i \leq \varepsilon + \xi_i^*,$$

$$\xi_i, \xi_i^* \geq 0 \quad (i = 1, \dots, l).$$

这里 ξ_i, ξ_i^* 为引入的松弛变量, $C > 0$ 为惩罚系数, ε 为不敏感损失函数.

利用拉格朗日函数和对偶原理, 得到 (2) 式的对偶问题,

$$\min_{\alpha, \alpha^*} W = \frac{1}{2} [\alpha^T (\alpha^*)^T] \begin{bmatrix} Q & -Q \\ -Q & Q \end{bmatrix} \begin{bmatrix} \alpha \\ \alpha^* \end{bmatrix} + [\varepsilon I^T + y^T, \varepsilon I^T - y^T] \begin{bmatrix} \alpha \\ \alpha^* \end{bmatrix}, \quad (3)$$

(3) 式的约束条件为

$$0 \leq \alpha,$$

$$\alpha^* \leq C,$$

$$[I^T, -I^T] \begin{bmatrix} \alpha \\ \alpha^* \end{bmatrix} = 0.$$

这里,

$$Q_{ij} = \Phi(x_i)^T \Phi(x_j) = K(x_i, x_j),$$

其中 $K(x_i, x_j)$ 为满足 Mercer 条件的核函数. 这样,

(1) 式所描述的回归函数可以描述为

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b. \quad (4)$$

(3) 式在约束 $I^T, -I^T \begin{bmatrix} \alpha \\ \alpha^* \end{bmatrix} = 0$ 下的拉格朗日函数为

$$L = \frac{1}{2} \sum_{i=1}^l (\alpha_i - \alpha_i^*) Q_{ij} (\alpha_j - \alpha_j^*) + \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) - \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) + \delta \sum_{i=1}^l (\alpha_i - \alpha_i^*). \quad (5)$$

优化此方程可以导出 Karush-Kuhn-Tucker (KKT 条件),

$$g_i = \frac{\partial L}{\partial \alpha_i} = \sum_{j=1}^l Q_{ij} (\alpha_j - \alpha_j^*) + \varepsilon - y_i + \delta = 0,$$

$$g_i^* = \frac{\partial L}{\partial \alpha_i^*} = - \sum_{j=1}^l Q_{ij} (\alpha_j - \alpha_j^*) + \varepsilon + y_i - \delta = 0,$$

$$\frac{\partial L}{\partial \delta} = \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0, \quad (6)$$

注意 (6) 式中的 δ 等于 (1) 和 (4) 式中的 $b^{[21]}$. 令 $\theta_i = \alpha_i - \alpha_i^*$, 则 θ_i 由 α_i 和 α_i^* 决定. 由 (6) 式可以把 SVR 的训练样本集划分成以下的集合^[20]: 标准支持向量集合 S , 边界支持向量集合 E , 保持向量集合 R (非支持向量集合). 以 g_i 为中心, KKT 条件如下:

$$2\varepsilon < g_i \rightarrow g_i^* < 0$$

$$(\theta_i = -C, i \in E),$$

$$g_i = 2\varepsilon \rightarrow g_i^* = 0$$

$$(-C < \theta_i < 0, i \in S),$$

$$0 < g_i < 2\varepsilon \rightarrow 0 < g_i^* < 2\varepsilon$$

$$(\theta_i = 0, i \in R), \quad (7)$$

$$g_i = 0 \rightarrow g_i^* = 2\varepsilon$$

$$(0 < \theta_i < C, i \in S),$$

$$g_i < 0 \rightarrow g_i^* > 2\varepsilon$$

$$(\theta_i = C, i \in E).$$

3. 基于小波核函数的在线支持向量回归

小波分析适合于信号的局部分分析和突变信号的检测. 结合小波分析建立一种满足框架条件的小波核函数, 能充分提高 SVR 的学习精度. 在线 SVR 适合于处理大规模样本集的情况, 也适合于处理样本集随着时间不断变化的情况, 其学习精度可以随着应用过程中样本集的积累而逐步提高.

3.1. 支持向量核条件

SVR 通过核函数将样本数据映射到高维特征空间. 在特征空间中实现线性回归, 核函数必须满足 Mercer 条件.

定理 1^[22] 令 R^N 为输入样本集空间,

$$\forall g(x) \in L_2(R^N),$$

$$\int g^2(x) dx < \infty \quad (g \neq 0),$$

并且

$$K(x, x') \in L_2(R^N \times R^N),$$

如果

$$\iint_{L_2 \otimes L_2} K(x, x')g(x)g(x')dx dx' \geq 0 \quad (8)$$

成立, 则 $K(x, x')$ 为特征空间中的点积.

定理 1 提供了判断和构造核函数的简单方法.

对于平移不变核函数, 如 $K(x, x') = K(x - x')$ 并满足 Mercer 条件, 则它也是一个容许支持向量核, 定理 2 给出了平移不变核函数的充要条件.

定理 2^[23] 若平移不变核函数 $K(x, x') = K(x - x')$ 是一个容许支持向量核, 当且仅当其傅里叶变换

$$F[k](\omega) = (2\pi)^{-N/2} \int_{R^N} \exp(-j\omega x) K(x) dx \geq 0 \quad (9)$$

成立.

3.2. 小波核函数

若 $\psi(x) \in L_2(R)$ 是一个母小波, $a \in R$ 和 $c \in R$ 分别为尺度和平移因子, $x, x' \in R^N$, 则可以构造一个满足 Mercer 条件的小波核函数为^[24]

$$K(x, x') = \prod_{i=1}^N \psi\left(\frac{x_i - c_i}{a_i}\right) \psi\left(\frac{x'_i - c'_i}{a_i}\right). \quad (10)$$

选择合适的小波函数作为核函数也是一个关键. 除了要满足 Mercer 条件外, 还要考虑小波函数的性质. 对于正交小波基, 虽然基函数具有正交性, 对函数的逼近更有效, 但由于训练数据的稀疏性^[25], 不能充分利用正交基的优点, 而且正交基构造复杂. 对于满足框架条件的小波, 由于基函数选取灵活, 在 $L_2(R^d)$ 中只用一个母小波就能构造出一个单尺度的小波框架^[25], 而如果采用正交基, 需 $2^d - 1$ 个母小波, 因此本文采用满足框架条件的小波来构造核函数. 采用如下的小波函数构造平移不变核函数:

$$\psi(x) = \cos(1.75x) \exp\left(-\frac{x^2}{2}\right). \quad (11)$$

推论 1^[24] 给定母小波函数(11)式, $x_i, x'_i, a_i \in R, a_i \neq 0, x, x' \in R^N$, 则小波核函数

$$\begin{aligned} K(x, x') &= \prod_{i=1}^N \psi\left(\frac{x_i - x'_i}{a_i}\right) \\ &= \prod_{i=1}^N \left(\cos\left(1.75 \times \frac{x_i - x'_i}{a_i}\right) \right. \\ &\quad \left. \times \exp\left(-\frac{\|x_i - x'_i\|^2}{2a_i^2}\right) \right) \quad (12) \end{aligned}$$

是一种容许支持向量核.

证明 由定理 2, 只需证明(9)式对所有 x 成立

即可, 其中

$$\begin{aligned} K(x) &= \prod_{i=1}^N \psi\left(\frac{x_i}{a_i}\right) \\ &= \prod_{i=1}^N \cos\left(1.75 \frac{x_i}{a_i}\right) \exp\left(-\frac{\|x_i\|^2}{2a_i^2}\right). \end{aligned}$$

根据文献[24], 首先计算积分

$$\begin{aligned} &\int_{R^N} \exp(-j\omega x) K(x) dx \\ &= \int_{R^N} \exp(-j\omega x) \left(\prod_{i=1}^N \cos\left(1.75 \frac{x_i}{a_i}\right) \right. \\ &\quad \left. \times \exp\left(-\frac{\|x_i\|^2}{2a_i^2}\right) \right) dx \\ &= \prod_{i=1}^N \int_{-\infty}^{\infty} \exp(-j\omega_i x_i) \exp\left(-\frac{\|x_i\|^2}{2a_i^2}\right) \\ &\quad \times \left(\frac{\exp(1.75x_i j/a_i) + \exp(-1.75x_i j/a_i)}{2} \right) dx_i \\ &= \prod_{i=1}^N \frac{1}{2} \int_{-\infty}^{\infty} \left(\exp\left(-\frac{\|x_i\|^2}{2a_i^2} + \left(\frac{1.75j}{a_i} - j\omega_i a\right)x_i\right) \right. \\ &\quad \left. + \exp\left(-\frac{\|x_i\|^2}{2a_i^2} - \left(\frac{1.75j}{a_i} + j\omega_i a\right)x_i\right) \right) dx_i \\ &= \prod_{i=1}^N \frac{|a| \sqrt{2\pi}}{2} \left(\exp\left(-\frac{(1.75 - \omega_i a)^2}{2}\right) \right. \\ &\quad \left. + \exp\left(-\frac{(1.75 + \omega_i a)^2}{2}\right) \right). \quad (13) \end{aligned}$$

将(13)式代入(9)式得

$$\begin{aligned} F[k](\omega) &= \prod_{i=1}^N \left(\frac{|a|}{2} \right) \left(\exp\left(-\frac{(1.75 - \omega_i a)^2}{2}\right) \right. \\ &\quad \left. + \exp\left(-\frac{(1.75 + \omega_i a)^2}{2}\right) \right). \quad (14) \end{aligned}$$

显然 $F[k](\omega) \geq 0$. 推论 1 得证.

将(12)式代入(4)式中, 则回归函数为

$$\begin{aligned} f(x) &= \sum_{k=1}^N \alpha_k \prod_{i=1}^N \left(\cos\left(1.75 + \frac{x_k^i - x_k'^i}{a_k^i}\right) \right. \\ &\quad \left. \times \exp\left(-\frac{\|x_k^i - x_k'^i\|^2}{2a_k^i{}^2}\right) \right) + b, \quad (15) \end{aligned}$$

式中 x_k^i 表示的 k 个训练样本的 i 个分量. 为了简单起见, 令 $a_k^i = a$, 这样核函数的参数就变成 1, 并且可以利用交叉验证法来选择参数^[24].

3.3. 在线支持向量回归学习算法

当在样本集中加入一个新样本 x_n 时, 其系数 θ_n 的变化将会影响现有样本的 g_i, g_i^* 或 θ_i , 可能会使 S, E 和 R 集合中的现有样本发生迁移. 我们首先考虑集合 S, E 和 R 中的样本在没有发生变化的情

况下,新样本 θ_n 的变化对现有样本造成的影响.由(7)式 KKT 条件知,当样本在集合 S 中保持不变时,其 g_i 值保持不变;当样本在集合 E 和 R 中保持不变时,其 θ_i 值保持恒定.根据文献 [20],下面分别对这两种情况进行讨论.

当引入新样本 x_n 时,在集合 S, E 和 R 中的样本没有发生迁移的情况下,根据(6)和(7)式,样本系数的变化情况可以通过下列公式进行计算:

$$\Delta g_i = Q_{in}\Delta\theta_n + \sum_{j \in S} Q_{ij}\Delta\theta_j + \Delta b, \quad (16)$$

$$\Delta g_i^* = -\Delta g_i, \quad (17)$$

$$\Delta\theta_n + \sum_{j \in S} \Delta\theta_j = 0. \quad (18)$$

当样本 x_i 位于集合 S 中, $\Delta g_i = 0$ 则可得

$$\sum_{j \in S} Q_{ij}\Delta\theta_j + \Delta b = -Q_{in}\Delta\theta_n, \quad (19)$$

$$\sum_{j \in S} \Delta\theta_j = -\Delta\theta_n. \quad (20)$$

定义集合 S 中的样本为 $S = \{S_1, S_2, \dots, S_k\}$ (19)

和(20)式可表示为

$$\begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & Q_{S_1 S_1} & \dots & Q_{S_1 S_k} \\ \dots & \dots & \dots & \dots \\ 1 & Q_{S_k S_1} & \dots & Q_{S_k S_k} \end{bmatrix} \begin{bmatrix} \Delta b \\ \Delta\theta_{S_1} \\ \dots \\ \Delta\theta_{S_k} \end{bmatrix} = - \begin{bmatrix} 1 \\ Q_{S_1 n} \\ \dots \\ Q_{S_k n} \end{bmatrix} \Delta\theta_n, \quad (21)$$

则

$$\begin{bmatrix} \Delta b \\ \Delta\theta_{S_1} \\ \dots \\ \Delta\theta_{S_k} \end{bmatrix} = \beta \Delta\theta_n, \quad (22)$$

式中,

$$\beta = \begin{bmatrix} \beta \\ \beta_{S_1} \\ \dots \\ \beta_{S_k} \end{bmatrix} = -U \begin{bmatrix} 1 \\ Q_{S_1 n} \\ \dots \\ Q_{S_k n} \end{bmatrix}, \quad (23)$$

$$U = \begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & Q_{S_1 S_1} & \dots & Q_{S_1 S_k} \\ \dots & \dots & \dots & \dots \\ 1 & Q_{S_k S_1} & \dots & Q_{S_k S_k} \end{bmatrix}^{-1}.$$

从(22)式可得

$$\Delta b = \beta \Delta\theta_n, \quad (24)$$

$$\Delta\theta_j = \beta_j \Delta\theta_n \quad (\forall j \in S). \quad (25)$$

(24)和(25)式描述了新样本 x_n 的 θ_n 变化对集合 S 中样本的影响.当集合 S 发生变化时(23)式中的 U 必须进行更新.根据文献 [14]的方法可以高效地更新 U .当一个新样本 x_n 加入集合 S 时, U 可以根据下式进行更新:

$$U = \begin{bmatrix} & & & 0 \\ & U & & \\ & & & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix} + \frac{1}{\gamma_n} \begin{bmatrix} \beta \\ 1 \end{bmatrix} \begin{bmatrix} \beta^T & 1 \end{bmatrix}, \quad (26)$$

$$\gamma_n = Q_{in} + \sum_{j \in S} Q_{ij}\beta_j + \beta.$$

当样本 x_i 位于集合 E 和 R 中,由(16)(24)和(25)式可得

$$\begin{aligned} \Delta g_i &= Q_{in}\Delta\theta_n + \sum_{j \in S} Q_{ij}\Delta\theta_j + \Delta b \\ &= Q_{in}\Delta\theta_n + \sum_{j \in S} Q_{ij}\beta_j \Delta\theta_n + \beta \Delta\theta_n \\ &= (Q_{in} + \sum_{j \in S} Q_{ij}\beta_j + \beta) \Delta\theta_n \\ &= \gamma_i \Delta\theta_n. \end{aligned} \quad (27)$$

(27)式描述了新样本 x_n 的 θ_n 变化对集合 E 和 R 中样本的影响.

(24)(25)和(27)式只有在集合 S, E 和 R 中的样本不发生迁移时才成立,从中可以看出,只要给定增量 $\Delta\theta_n$ 就可以在满足 KKT 条件的同时更新训练样本的系数 g_i 和 θ_i .但在有些情况下,当增加了新样本 x_n 时,新样本影响了现有的样本,集合 S, E 和 R 中的样本必须经过迁移才能满足 KKT 条件,这可以通过改变 θ_n 的值来实现.样本的迁移过程如下:对于集合 S 中的样本 x_i ,如果 θ_i 从 $0 < |\theta_i| < C$ 变化到 $|\theta_i| = C$ 则 x_i 从集合 S 迁移到集合 E ;如果 θ_i 变化到 $\theta_i = 0$,则 x_i 从集合 S 迁移到集合 R .对于集合 E 中的样本 x_i ,如果 g_i 或 g_i^* 变化到 0,则 x_i 从集合 E 迁移到集合 S .对于集合 R 中的样本 x_i ,如果 g_i 或 g_i^* ($g_i > 0$ 和 $g_i^* > 0$)变化到 0,则 x_i 从集合 R 迁移到集合 S .根据以上分析结果及文献 [20],由此得出在线 SVR 算法的步骤.

第一步:令新样本 x_n 的系数 $\theta_n = 0$.

第二步:如果 $g_n > 0$ 和 $g_n^* > 0$,则把新样本 x_n 添加到集合 R 中,算法结束.

第三步:如果 $g_n < 0$ (或 $g_n^* < 0$),则增加(或减小) θ_n 的值,同时更新集合 S 中样本 i 的 θ_i 值或集合 E, R 中样本 i 的 g_i 和 g_i^* 值,直到下列条件之一

成立：

(1) $g_n = \alpha$ (或 $g_n^* = 0$), 把新样本 x_n 添加到集合 S 中, 同时更新矩阵 U , 算法结束.

(2) $\theta_n = \alpha$ (或 $\theta_n = -C$), 把新样本 x_n 添加到集合 E 中, 算法结束.

(3) 根据 g_i, g_i^* 或 θ_i 的变化, 集合 S, E, R 中的样本 i 根据上述的迁移过程将会发生迁移, 同时更新矩阵 U .

第四步: 返回第三步, 计算下一个样本, 直至样本计算完毕.

从以上的算法可以看出, 通过改变新样本 x_n 的 θ_n 值保证新样本 x_n 满足 KKT 条件. 同时 x_n 的加入也会影响到原来样本的 g_i 和 θ_i 值, 并使其在 S, E 和 R 中发生迁移以满足 KKT 条件. 这个过程是以迭代方式实现的, 而不是一次性计算的, 因此适合于大规模样本集的在线训练, 且求解过程比较快, 这也是将在线 SVR 应用于混沌时间序列预测的关键.

4. 混沌动力系统的预测

基于小波核函数的在线 SVR 适用于一般 SVR 批量学习算法应用的领域及大规模数据集和实时学习的情况, 而且提供的精度和训练速度要优于一般的批量学习算法. 在本文中, 我们通过对一个标准的混沌时间序列 Mackey-Glass 预测问题来检验基于小波核函数的在线学习算法的性能, 并与具有高斯核函数的序列最小优化(sequential minimal optimization, SMO)算法及文献[19]的增量学习算法性能进行比较, 仿真中采用均方根误差作为评价指标.

Mackey-Glass 混沌时间序列产生方程为

$$\frac{dx}{dt} = \frac{0.2x(t-\tau)}{1+x(t-\tau)^{10}} - 0.1x(t), \quad (28)$$

式中 τ 为时滞参数, 当 $\tau \geq 17$ 时产生的数据是混沌的, τ 值越大混沌程度越高. 图 1 为 $\tau = 17$ 时 Mackey-Glass 混沌时间序列, 从中可以看出时间序列具有复杂的非线性混沌特征. 我们以 Mackey-Glass 混沌时间序列中 1200 个数据作为训练样本. 用本文的在线算法、增量算法和 SMO 算法对样本进行在线训练, 其中 $C = 100, \epsilon = 0.05$, 增量算法和 SMO 算法均采用高斯核函数

$$K(x_i, x_j) = \exp(-\lambda \|x_i - x_j\|^2) \quad (\lambda = 1).$$

将混沌数据样本分成两部分, 取 200 个样本用作训练样本, 其中初始训练集包括 10 个样本, 另外 1000

个样本作为在线预测的样本. 三种算法的预测结果分别如图 2、图 3、图 4 和表 1 所示.

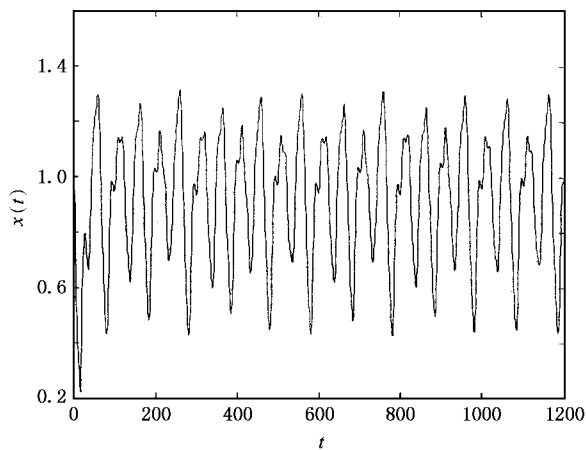


图 1 $\tau = 17$ 时 Mackey-Glass 混沌时间序列

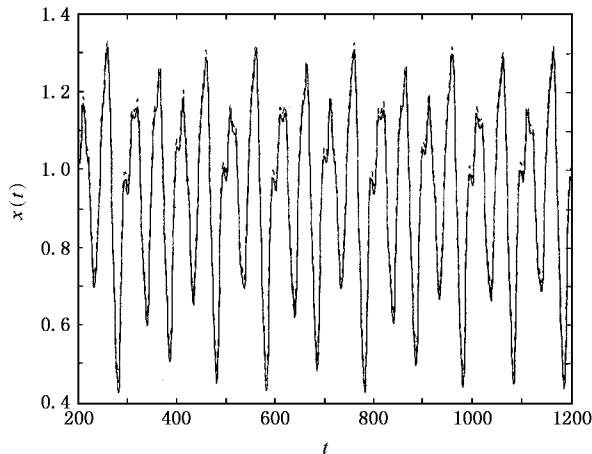


图 2 在线小波 SVR 预测 Mackey-Glass 仿真曲线 实线为实际值, 虚线为预测值

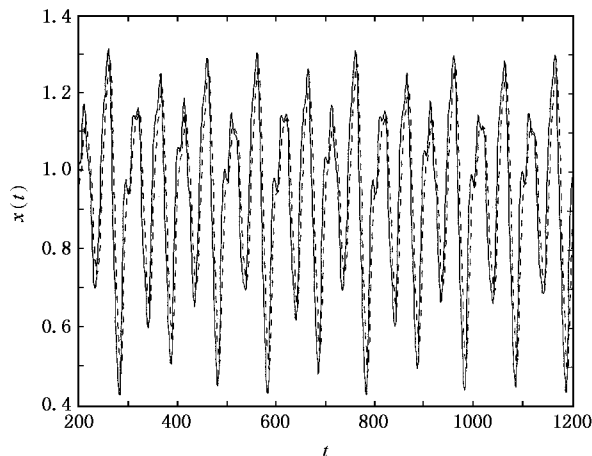


图 3 增量 SVR 预测 Mackey-Glass 仿真曲线 实线为实际值, 虚线为预测值

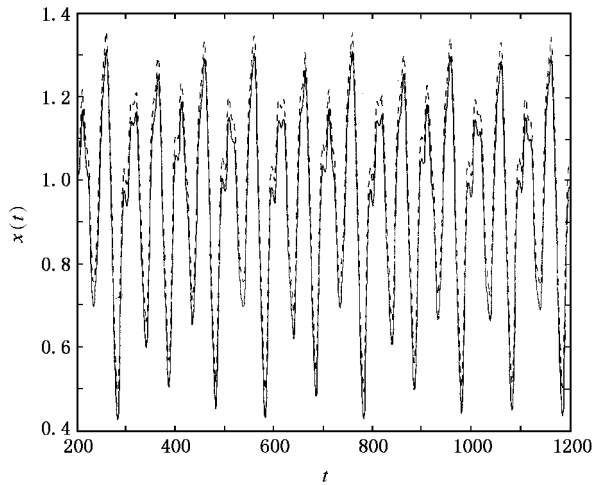


图4 SMO SVR 预测 Mackey-Glass 仿真曲线 实线为实际值,虚线为预测值

从仿真结果可以看出,由于小波核函数具有精细描述的特性,基于小波核函数的在线 SVR 的预测值与系统的实际输出值符合很好,预测精度要高于基于高斯核函数的 SVR。由于基于增量学习的 SVR 只是提供了一种近似解,因此预测精度较低。由表 1 可以看出,在线 SVR 的学习效率较高,这是因为在线学习并没有像传统方法那样抛弃前一次的计算结果,机械地从头开始训练,而是根据新样本的特性采用不同策略来迭代修改回归函数,充分提高了预测效率。

表 1 Mackey-Glass 预测结果

预测方法	训练时间	均方根误差
在线小波 SVR	21.26	0.0012
增量 SVR	45.28	0.0074
SMO SVR	146.45	0.0031

5. 结 论

由于传统的 SVR 算法不适合在线混沌时间序列预测,本文建立了一种在线 SVR 算法。该算法充分考虑了新增样本与 KKT 条件之间的关系,在过去训练结果的基础上建立新的有效回归函数。由于该算法是以迭代方式实现的,避免了新样本引入时必须从第一个样本开始重新学习的问题,因此学习效率较高。同时利用在线 SVR 求解回归估计问题时,选择合适的核函数也是一个关键。本文在小波分析理论的基础上,建立了一种满足小波框架条件的小波核函数。由于小波具有时频局部特性和变焦特性,小波核函数能以较高的精度逼近任意函数。最后,我们将基于小波核函数的在线 SVR 应用于 Mackey-Glass 混沌时间序列预测。仿真结果表明:基于小波核函数的在线 SVR 具有训练时间短、预测精度高、泛化能力强等特点,在混沌时间序列的在线预测中具有较强的实用价值。

- [1] Gershenfeld N, Schoner B, Metois E 1999 *Nature* **397** 329
- [2] Gan J C, Xiao X C 2003 *Acta Phys. Sin.* **52** 2995 (in Chinese) [甘建超、肖先赐 2003 物理学报 **52** 2995]
- [3] Tan W, Wang Y N 2003 *Acta Phys. Sin.* **52** 795 (in Chinese) [谭文、王耀南 2003 物理学报 **52** 795]
- [4] Zhang S, Liu H X, Gao D T 2003 *Chin. Phys.* **12** 594
- [5] Zhang J S 2000 *Chin. Phys.* **9** 408
- [6] Liu D, Kong Z Q, Ren H P 2003 *Acta Phys. Sin.* **52** 531 (in Chinese) [刘丁、孔志强、任海鹏 2003 物理学报 **52** 531]
- [7] Tan W, Wang Y N, Liu Z R 2002 *Acta Phys. Sin.* **51** 2463 (in Chinese) [谭文、王耀南、刘祖润 2002 物理学报 **51** 2463]
- [8] Wang Z Y, Cai Y L, Jia D 1999 *Acta Phys. Sin.* **48** 206 (in Chinese) [王忠勇、蔡远利、贾冬 1999 物理学报 **48** 206]
- [9] Vapnik V N 2000 *The Nature of Statistical Learning Theory* (New York : Springer)
- [10] Ye M Y, Wang X D 2004 *Chin. Phys.* **13** 454
- [11] Kim K 2003 *Neurocomputing* **55** 307
- [12] Cui W Z, Zhu C C, Bao W X *et al* 2004 *Acta Phys. Sin.* **53** 3303 (in Chinese) [崔万照、朱长纯、保文星等 2004 物理学报 **53** 3303]
- [13] Cui W Z, Zhu C C, Bao W X *et al* 2005 *Acta Phys. Sin.* **54** 3009 (in Chinese) [崔万照、朱长纯、保文星等 2005 物理学报 **54** 3009]
- [14] Cauwenberghs G, Poggio T 2001 *Advances in Neural Information Processing Systems* (Cambridge : MIT Press)
- [15] Kivinen J, Smola A J, Williamson R C 2004 *IEEE Trans. Sign. Proc.* **52** 2165
- [16] Tax D M J, Laskov P 2003 *IEEE Workshop on Neural Networks for Signal Processing* (Toulouse : IEEE)
- [17] Gentile C 2001 *J. Mach. Learn. Res.* **2** 213
- [18] Xiao R, Wang J C, Sun Z X *et al* 2002 *J. Software* **12** 1818 (in Chinese) [萧嵘、王继成、孙正兴等 2002 软件学报 **12** 1818]
- [19] Ralaivola L, d'Alche-Buc F 2001 *Lect. Notes Comput. Sci.* **2130** 322
- [20] Martin M 2002 *Lect. Notes Comput. Sci.* **2430** 282
- [21] Chang C C, Lin C J 2002 *Neural Comput.* **14** 1959
- [22] Mercer J 1909 *Philos. Trans. R. Soc. A* **209** 415
- [23] Smola A J, Scholkopf B, Muller K R 1998 *Neural Networks* **11** 637
- [24] Zhang L, Zhou W, Jiao L 2004 *IEEE Trans. Syst. Man Cyb. B* **34** 34
- [25] Zhang Q 1997 *IEEE Trans. Neural Networks* **8** 227

Prediction of chaotic time-series based on online wavelet support vector regression^{*}

Yu Zhen-Hua[†] Cai Yuan-Li

(*School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China*)

(Received 8 June 2005; revised manuscript received 6 December 2005)

Abstract

Support vector regression (SVR) is an effective method for the predication of chaotic time-series, which is a fundamental topic of nonlinear dynamics. Through analyzing the possible variation of support vector sets after new samples are inserted to the training set, a novel SVR algorithm is proposed; thus an online learning algorithm is set up. In connection with the specific characteristics of chaotic signals, a wavelet kernel satisfying wavelet frames is also presented. The wavelet kernel can approximate arbitrary functions, and is especially suitable for local processing; hence the generalization ability of SVR is improved. To illustrate the good performance of the online wavelet SVR, a benchmark problem, i. e. the online prediction of chaotic Mackey-Glass time-series, is considered. The simulation results indicate that the online wavelet SVR algorithm outperforms the existing algorithms in higher efficiency of learning as well as better accuracy of prediction.

Keywords : chaotic time-series, support vector regression, online learning, wavelet kernel

PACC : 0545

^{*} Project supported by the National High Technology Development Program of China (Grant No. 2003AA721070).

[†] E-mail : zhuayu@gmail.com