

# 最小二乘支持向量域的混沌时间序列预测

任 韧<sup>1)†</sup> 徐 进<sup>2)</sup> 朱世华<sup>1)</sup>

1) 西安交通大学电子与信息工程学院, 710049 西安)

2) 西安交通大学医工所, 710049 西安)

3) 西安交通大学理学院, 710049 西安)

(2005 年 5 月 17 日收到, 2005 年 6 月 13 日收到修改稿)

从支持向量域(SVD, Support Vector Domain)出发,根据 Takens 延时相空间重构思想,利用支持向量机非线性映射,建立了混沌时间序列和混沌非线性相轨迹运动的 SVD 预测模型.采用数据集作为支持对象元素,机器学习缩小模型泛化误差的上界.利用最小二乘支持向量域(SVD),预测了 Henon/Lorenz/Rosler 三种混沌时间序列.预测结果表明,三种预测模型将集合映射到一个更高维特征空间,通过嵌入维数,实现了序列预测,误差随嵌入维数变化趋于恒定.与支持向量机(SVM)相比,SVD 所需支持向量少,收敛速度快,鲁棒性强,核函数选择容易灵活,且存在自适应方法.网格点数提高了 10—20 倍,序列预测在小样本、非线性、未知概率密度条件下,预测和实际值取得了一致.

关键词:支持向量域,混沌,最小二乘,时间序列预测

PACC: 0545

## 1. 引 言

支持向量域(SVD)是 20 世纪末出现的一种机器学习算法.它在支持向量机(SVM)基础上,利用已有的对象集合,采用域描述模型,通过已知小样本集合产生对象的一种区域分类方法,它以 Vapnik 等提出的统计学习理论为基础,是一种推广的机器学习方法<sup>[1-4]</sup>.统计学习理论认为低风险的规则是最优的,在一个大的测试样本中统计,高概率、损失均值小的是搜索的目标,这种技术为函数集提供了元素最小化风险结构,在最小化样本点误差的同时,缩小模型泛化误差的上界,从而提高模型的泛化能力.在针对小样本的学习中效果尤为突出.

支持向量域的支持向量机模型已经在模式识别、统计判决和函数模型回归上取得了应用. SVD 比 SVM 支持向量数少.与神经网络相比, SVD 针对小样本使用分类和识别仍然取得了成功. SVD 的支持向量机网络隐层节点数选择受经验风险影响,但不存在神经网络的过拟合与训练极小点的问题.并且结构简单,不依赖于数学模型,具有自学习自调整特点,在解决小样本、非线性、密度估计和高维模式识

别上具有较好的预测和识别能力.

大多数模式识别问题是分类和辨识问题,泛化能力越强越难给出其精确的分类归属,但是却可以给出一组对象的描述.这组对象数据经过训练可以归结出对象空间特征,并拒绝非本集合数据.这种由支持向量划分的支持域模型在时间序列的预测上具有显著的优点.除具有支持向量机的优点外,支持向量域是一种鲁棒性和分类性好的方法,它不需要概率密度估计.在部分支持对象的基础上,支持对象新预测的数据只需与超球体分类面边界内的数据比较,比支持向量机单纯的超球形边界使用更容易和灵活.其参数的选择可以依据训练序列错误次数和超球形分界面半径确定.支持向量域的参数选择可以调整,并且不严格.本文采用径向基高斯核函数,核函数的宽度可调整,最小超球形分类面可以获得灵活的描述.由于支持向量域关注边界描述而非数据密度,对象数明显减小,对一个数据集,支持向量域不但具有很高的稳健性、鲁棒性,在预测时间上还具有高效率.

混沌时间序列是一种介于确定性和随机性之间的随机敏感充分不规则的非线性动力现象.宏观上表现为无序无律的混乱运动,以及对初值十分敏感

\* 国家自然科学基金(批准号: 60372055, 30400101)资助的课题.

† 通讯联系人. E-mail: renren@mail.xjtu.edu.cn

的蝴蝶效应. 微观上呈现无穷嵌套几何自相似性. 运动具有普适规律, 是有序和无序、确定性和随机性的中间态. 混沌时间序列区别于随机时间序列, 它具有短期的可预测性和长期的不可预测性. 混沌时间序列的进一步研究表明非线性系统参数变化, 运动分岔是出现混沌时间序列的先兆, 将引起时间序列的随机性、永不重复和周期无穷大以及序列性质等发生变化. 1980年由 Packard 等人提出了混沌时间序列的预测重构思想<sup>[5-7]</sup>, 1981年 Takens<sup>[5]</sup>在拓扑学嵌入理论上, 提出了时间序列的延迟相空间重构和嵌入定理. 并且将 Lyapunov 指数大于零, 作为研究判断一个时间序列是否混沌的依据. 1990年 Ott 等<sup>[6]</sup>基于参数扰动法成功实现了混沌系统的预测控制. Shukla<sup>[7]</sup>认为大气参数是确定系统产生的对初始状态敏感的混沌序列, 从而完成了大气温度预测<sup>[8]</sup>. 已往对混沌时间序列的预测主要有 Volterra 自适应预测法, 但级数增多记忆单元增大, 滤波系数和计算幂次增加过快. Takens 提出的嵌入定理和相空间重构思想<sup>[5-8]</sup>, 克服了上述缺点, 采用支持向量和支向量域实现非线性映射, 通过核函数选择, 隐层节点数和自适应调整参数, 实现回归预测上的自学习, 完成预测. 目前混沌序列预测已在基因、蛋白质和外太空大气预报等方面开始应用, 并取得良好效果<sup>[9]</sup>.

本文基于混沌时间序列固有的确定性和非线性性质, 利用混沌动力系统相空间延迟坐标重构理论, 在支持向量与统计学习理论的基础上, 首先建立了混沌序列的支持向量域预测模型和对两类问题处理方法, 其次采用最小二乘法支持向量域表述, 预测了 Henon, Lorenz 混沌时间序列和 Rossel 混沌时间序列, 讨论了嵌入维数和模型预测误差的关系.

## 2. 支持向量域预测模型

### 2.1. 支持向量域模型

支持向量域(SVD)是一种数据描述方法, 它是通过计算包含一组数据最小超球形边界对该组数据进行描述. 在运用分类点进行训练时, 从最小超球形边界得到该类数据描述边界, 进一步由描述边界对其他未知的两类数据进行分类.

支持向量域分类器对一组样本集进行描述时, 通过找到包含该组数据的最小半径超球体分界面进行描述, 为消除奇点的影响, 采用松弛变量  $\xi_i$  把奇

点排除在超球体分界面外, 在约束条件为  $(x_i - a)^T \times (x_i - a) \leq R^2 + \xi_i^2$  时, 最小化半径  $\min F(R, a, \xi_i) = \min |R^2 + C \sum_i \xi_i|^{[1]}$ , 其中  $\xi_i \geq 0$ ,  $x_i$  为分类数据,  $R$  为超球体分界面半径,  $a$  为超球体中心,  $C$  是系数变量通过超球体分界面和被拒绝域来选择确定, 这个优化问题的解由拉格朗日泛函的鞍点给出, 泛函为  $L(R, a, \alpha_i, \xi_i)$ , 其中  $\alpha_i$  为支持向量即拉格朗日系数, 通过  $L(R, a, \alpha_i, \xi_i)$  定义了非线性映射  $\phi(x)$ , 可以将输入空间映射到一个高维的 Hilbert 空间中  $y(x) = w^T \phi(x) + b$ ,  $w$  为权矢量,  $b$  为偏差, 并且使训练样本中距离最优分类超平面最近的点距分类面的距离最大. 混沌序列的延迟相空间理论定义状态空间  $R^S$ , 嵌入空间  $R^m$ , 一个延迟坐标嵌入为  $R^S \rightarrow R^{S+m}$  映射,  $m$  为嵌入维数, 通过相空间嵌入维数可以实现在原序列延迟下, 重构出未来系统演化的预测模型和预测值,  $x(t+T) = \phi(x(t)) + \xi(t)$ , Vapnik 根据泛函理论认为在构造最优超平面时, 满足 Mercer 条件下, 不需要显式表征特征空间. 而只需要计算特征向量与特征空间中向量的内积(泛函)即可, 在 Hilbert 空间非线性映射可通过核函数分类器完成,  $\phi(x)$  用核函数来映射, 实现混沌序列在高维空间的映射表示和分离.

混沌序列非线性模型中长期动态特性呈现了类随机现象的内在行为, 源于系统的确定有界和全局稳定性, 混沌时间序列对短期可预测, 可以通过一些特征实现. 1) 对初值的敏感性, 序列中后出现的值对先前值的依赖性. 2) 遍历性和稳定性. 3) 存在奇异吸引子, 及奇异吸引子关联维数, 并存在吸引域, 整体非线性混沌序列是有界和全局稳定的. SVD 支持向量机对非线性问题具有强大的处理能力, 将核函数作为混沌非线性时间序列映射中的映射  $\phi(x)$ , 它能够以任意精度逼近任意连续函数, 而且没有训练序列极小点和过拟合, 核函数的选择灵活, 避免了进行非线性变换  $\phi(x)$ , 而且收敛速度快, 只需小样本数据, 就可以在高维进行分类预测.

### 2.2. 支持向量域分类器的二类问题的原理

SVD 从部分数据出发构造最优分类超平面, 通过线性变换将输入空间变换到一个高维空间, 对于数据训练序列  $(x_1, y_1), \dots, (x_l, y_l)$ ,  $x_i \in R^n$ ,  $y_i \in \{+1, -1\}$ ,  $i = 1, 2, \dots, l$ , 如果  $x_i$  为第一类, 则  $y_i = +1$ , 否则记为  $y_i = -1$ , 对于  $y_i = +1$  样本集合, 优化问题由 Lagrangian 泛函鞍点表示, 泛函为

$$L(R, a, \alpha_i, \xi_i) = \sum_i \{ \alpha_i [ R^2 + \xi_i - (x_i^2 - 2ax_i + a^2) ] \} + \sum_{ij} \alpha_i \alpha_j K(x_i \cdot x_j) \leq R^2. \quad (3)$$

$$+ \sum_i v_i \xi_i - C \sum_{i=1}^l \xi_i - R^2 \quad (1)$$

$\alpha_i$  是拉格朗日系数即支持向量,  $\alpha_i \geq 0, v_i \geq 0$ , 其中限制条件  $0 \leq a_i \leq C, \sum a_i = 1$ , 建立最优分类超球面(1)式, 对 Lagrangian 泛函的参量  $R, a, \xi$  求导, 根据 Karush-Kuhn-Tucker 最优化条件可以推出  $\alpha, R$  和  $\xi$ ,

$$\frac{\partial L}{\partial R} = 0 \rightarrow \sum_i \alpha_i = 1,$$

$$\frac{\partial L}{\partial a} = 0 \rightarrow \sum_i \alpha_i x_i = a \text{ 且 } a = \frac{\sum_i \alpha_i x_i}{\sum_i \alpha_i},$$

$$\frac{\partial L}{\partial \xi} = 0 \rightarrow C = \alpha_i + v_i \rightarrow 0 < \alpha_i < C,$$

令核函数  $K(x_i \cdot x_j) = (x_i \cdot x_j)$ , 支持向量  $\alpha_i, L = \sum_i \alpha_i (x_i \cdot x_i) - \sum_{ij} \alpha_i \alpha_j (x_i \cdot x_j)$  非线性变换用核函数代替内积, 一种新的数据域描述及约束条件为

$$L(\alpha) = \sum_{ij=1}^l \alpha_i \alpha_j K_i K_j (x_i \cdot x_j) - \sum_{i=1}^l \alpha_i K_i (x_i \cdot x_i), \quad (2)$$

设被预测的支持对象数据为  $Z$  满足  $(z - a) \cdot (z - a) \leq R^2$  被接收, 用支持向量  $\alpha_i$  表示的接受条件区域为

$$(z \cdot z) - 2 \sum_i \alpha_i (z \cdot x_i) + \sum_{ij} \alpha_i \alpha_j (x_i \cdot x_j) \leq R^2$$

即

$$K(Z \cdot Z) - 2 \sum_i \alpha_i K(z \cdot x_i)$$

上式表明最优超球体分类面内训练数据组成的一个线性数据集合, 满足(3)式则被接受,  $z$  为支持对象. 用  $K(z \cdot z)$  代替  $\phi(x)$  将数据映射到高维 Hilbert 空间. 权重因子  $\alpha_i \neq 0$  被称为支持向量. 可以用权重因子小于  $C$  的支持向量描述  $R$ , 对于所有  $\alpha_i = C$  的向量, 被认为处于边界和超球体之外, 被拒绝, 通过  $C$  讨论两类问题.

非线性回归函数分类器的解为

$$f(x) = \text{sgn} \left| k(x \cdot x) + \sum_{ij} \alpha_i \alpha_j K_i K_j k(x_i \cdot x_j) - 2 \sum_i \alpha_i K(x \cdot x_i) - R^2 \right|, \quad (4)$$

其中  $i, j$  为支持向量的个数,  $x_k$  为支持对象分类数据,  $\alpha_i$  是支持向量,  $x_i$  为输入向量.

核函数可采用多项式核函数或高斯核函数

$$K(x_i \cdot x_j) = \exp \left( - \frac{(x_i - x_j)^2}{\sigma^2} \right).$$

我们设计了一个二层神经网络来构造这样一个分类器或非线性的决策函数, 支持向量机由已知的对象集合和核函数来描述.

决策规则取

$$y = \text{sgn} \left| k(x \cdot x) + \sum_{ij} \alpha_i \alpha_j K_i K_j k(x_i \cdot x_j) - 2 \sum_i \alpha_i K(x \cdot x_i) - R^2 \right|,$$

权值  $y_1 \alpha_1, \dots, y_n \alpha_n$ , 输入向量为  $x_i$ , 其支持向量域模型表达为见图 1.

### 3. 函数回归的最小二乘支持向量域 SVD 表述

支持向量机通过映射将输入空间映射到高维空

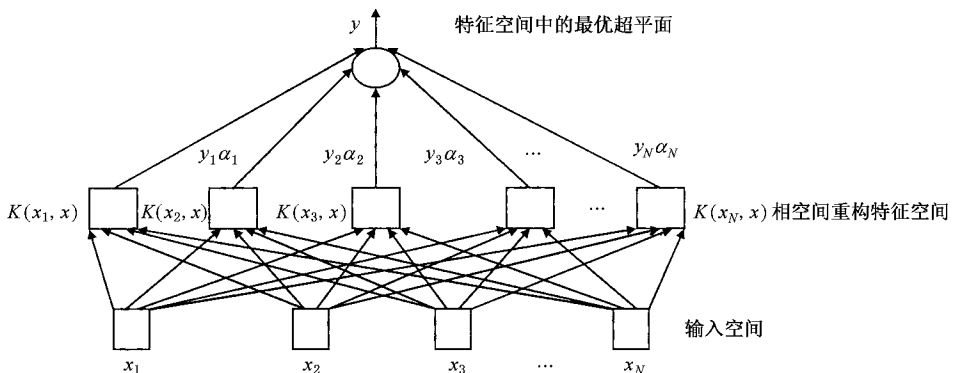


图 1 相空间重构模型

间,而且使分类间隔最大,通过使间隔最大化来控制分类器的复杂度,进而实现较好的推广能力.给定训练样本  $\{x_1, y_1\}, \dots, \{x_N, y_N\}$ ,  $k = 1, 2, \dots, N$ , 函数形式为  $y(x) = w^T \Phi(x) + b$ ,  $w$  为权矢量,采用核函数代换  $K(x_i, x_j) = \phi(x_i)\phi(x_j)$ . 对于训练数据集  $\{x_k, y_k\}^N$ , LM-SVD 定义为如下优化问题.

$$\begin{aligned} \min_{w, b, v} J(w, v) &= \frac{1}{2} w^T w + v \frac{1}{2} \sum_{k=1}^N e_k, v > 0 \\ &= \frac{1}{2} w^T w + v \frac{1}{2} \sum_{k=1}^N (y_k - w^T \phi(x_k) - b)^2 \\ & \quad k = 1, 2, 3, \dots, N \end{aligned}$$

上式优化解为

$$\begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 & I^T \\ I & \Omega + vI \end{bmatrix} \begin{bmatrix} 0 \\ y \end{bmatrix},$$

$I = [1, \dots, 1]^T$ ,  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_N]^T$ ,  $\Omega = K(x_i, x_j)$ , 通过上式建立 SVM 的预测模型,并对未来数据预测.

$$\begin{aligned} \text{分类函数为 } f(x) &= \text{sgn} \left| \sum_{i=1}^i \alpha_i y_i K(x_i, x) + b \right| = \\ \text{sgn} \left| \sum_{i=1}^i \alpha_i \exp \left\{ -\frac{\|x - x_k\|^2}{\sigma^2} \right\} + b \right|. \end{aligned}$$

$\Omega = K(x_i, x) = \Phi(x_i)\Phi(x)$  是满足 Mercer 条件的核函数.

上式由输入特性构成训练数据集  $\{x_k, y_k\}$ , 支持向量的个数为  $N$ , 由函数逼近和相空间重构可以得到其预测值.

## 4. 混沌时间序列的支持向量域预测

采用 SVD 最小二乘法对混沌时间序列可以进行精确的预测.首先混沌时间序列可以看作是一系列确定系统产生的有序时间序列的组合,具有短期的可预测性,其次针对混沌系统非线性动力性质、分岔和分形特点,可以通过相空间重构法预测.根据 Takens 理论,相空间是一个高维空间,可以建立适当的相空间,将被预测数据嵌入到高维的相空间中,研究在高维空间运动序列的表现,分类提取出时间序列中掩盖起来的信息.通过嵌入维数和嵌入空间,建立一个与原映射拓扑等价的非线性模型.同时建立预测的评价体系,采用相对误差和均方误差来表示预测水平<sup>[8-12]</sup>.

$$\text{均方误差 MSE} = \sqrt{\frac{1}{K} \sum_{k=1}^k (x_k - \hat{x}_k)^2}, \text{ 相对误差}$$

$$\alpha(n) = \frac{|x_i - \hat{x}_i|}{|x_i|}.$$

假设重构后的轨道空间由状态矢量  $X$  构成,  $\phi$  为重构空间的非线性预测模型,则非线性预测模型为  $x(t+T) = \phi(x(t)) + \epsilon(t)$ , 嵌入维数为  $m$ ,  $x_i = [x(i), x(i+\tau\Delta t), \dots, x(i+(m-1)\tau\Delta t)]^T$  其中  $m \geq 2d+1$ ,  $d$  为关联维数即吸引子维数,反映吸引子所需的变量个数,可由 Grassberger-Procaccia 求得<sup>[7]</sup>.

对于一个实际的混沌时间序列  $x(t)$ ,  $t_s$  作为采样间隔,采样值记为离散的时间序列  $x(t_0), x(t_0 - t_s), x(t_0 - 2t_s), x(t_0 - 3t_s), \dots, x(t_0 - kt_s)$ , 其中  $k = 1, 2, \dots, m$ , 根据相空间理论分别用 SVD 支持向量域的支持向量,建立对  $n$  个混沌时间序列的预测模型.将产生的混沌时间序列作为神经网络的输入,即  $x(t)$  时间序列作为混沌轨道的支持对象,按如下方法构造输入输出:

$$y = \begin{bmatrix} y_1^T \\ y_2^T \\ \vdots \\ y_n^T \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \dots & x_d \\ x_2 & x_3 & \dots & x_{d+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n+1} & \dots & x_{n+d-1} \end{bmatrix}, \quad (5)$$

其中  $x_j = x_j(t_0 + (j-1)\tau)$ ,  $\tau$  是延时因子,嵌入的维数  $d$  通过比较相邻点在高维相空间的距离判断.在嵌入不同的关联维数下,关联积分随尺度趋于确定值.预测的步骤为:1)初始化,2)时间序列生成,3)相空间延迟重构预测,4)数据处理误差分析.

### 4.1. Henon 混沌时间序列预测

Henon 混沌系统有别于 Hamilton 系统相空间容积不变的系统(相空间维数在运动中保持不变)<sup>[13]</sup>, Henon 在运动时相空间维数低于原来相空间维数,相空间收缩到吸引子(Attractors)上,是耗散系统中混沌的独特现象<sup>[13]</sup>.

Henon 混沌时间序列定义为  $x_{n+1} = 1 + Ax_n^2 + Bx_{n-1}$ , 采用径向基函数  $\phi(x_i, x_j) = \exp\left(-\frac{(x - x_i)^2}{\sigma^2}\right)$  作为核函数,最小二乘法 SVM 进行辨识预测.

Henon 映射是关联维数为 1.25 的混沌,  $A = 1.29$ ,  $B = 0.3$  时,产生混沌,对于  $A$  加扰动,初始值小的扰动差会造成轨道充分大的分离(称为指数分离).采用 Henon 映射产生混沌时间序列,可以看出其具有复杂的非线性混沌效应和自相似性.加之能量不守恒,耗散系统特点的影响,其混沌时间预测模

型建立较困难.采用 SVD 支持向量域模型,可以在对模型未知的情况下,通过 SVD 的支持向量机最小二乘,由较小的训练样本集对未知的 Henon 序列预测.

方法为对 SVD 模型采用 Henon 时间序列 3000 个数据点作为训练集,嵌入维数  $d = 3$ ,在调节量、超参数  $\{\gamma, \sigma\} = \{50, 2\}$  进行训练建模,取初值产出 300 个数据,并利用将来的 500 个数据点验证,图 2 显示了预测值与实际值结果.

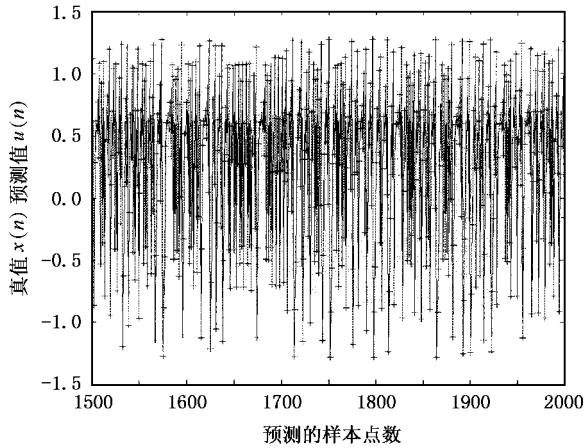


图 2 Henon 时间序列一步预测实际值与预测值的比较 (MSE = 0.003 相对误差  $e_{err} = 0.0071$ )

固定  $\{\gamma, \sigma\} = \{50, 2\}$ ,改变嵌入维数  $m$ ,预测 Henon 时间序列.可以观察到在嵌入维数小于 2 时,均方误差 MSE 随嵌入维数增大而减小,在嵌入维数为 2,3,4,5 时,MSE 变化较小.可以看出相空间的关联维数随嵌入维数的增加而减小,在大于最小嵌入维数后,继续增加嵌入维数,MSE 不会减小,如图 3.对于混沌时间序列,随着嵌入维数的增大,关联维数将收敛于某一确定值.

Henon 时间序列预测样本在训练 50 次后均方误差在 0.0012 以下,如图 4.

#### 4.2. Lorenz 混沌时间序列预测

Lorenz 系统是一个相空间容积收缩的耗散系统<sup>[14,15]</sup>表达式为

$$\begin{aligned} \dot{x} &= a(y - x), \\ \dot{y} &= rx - xz - yx, \\ \dot{z} &= xy - bz. \end{aligned}$$

微分方程中参数  $a = 10, b = 8/3, r = 28$  或  $r = 34$  下,表现为混沌,并存在奇异吸引子,并且  $r < 24.74$  时,出现混沌,  $24.07 < r < 24.74$ ,表现为混沌

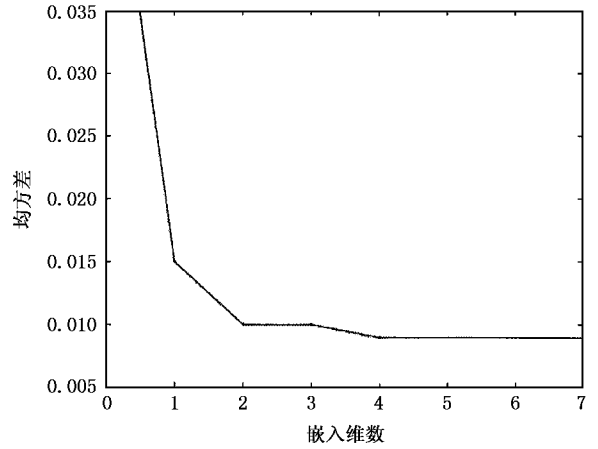


图 3 Henon 时间序列的嵌入维数与 MSE 关系

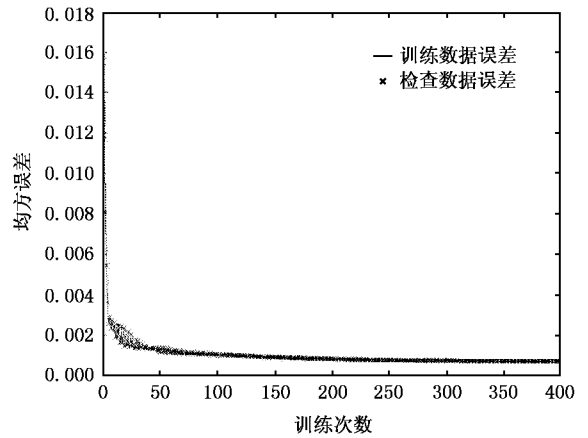


图 4 Henon 时间序列预测的样本随训练次数的误差

或稳定流.采用上述预测方法对其  $x, y, z$  坐标方向时间序列预测,预测值重构的相轨迹如图 5,其最小嵌入维数为 3,关联维数  $D = 2.06$ .图 5 预测的 Lorenz 吸引子是双叶形,空间相轨道折叠.

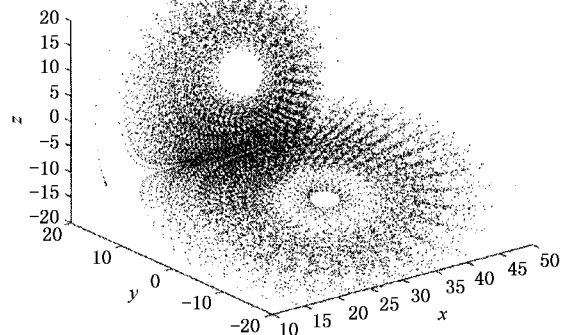
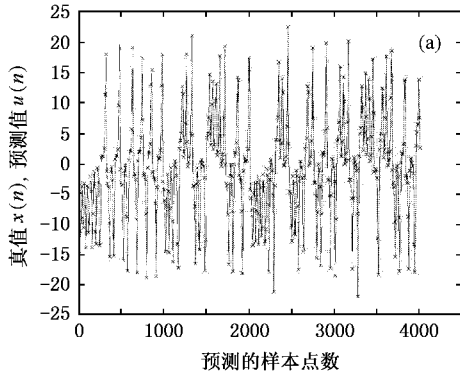


图 5 基于三维序列预测值重构的 Lorenz 混沌轨道相图

利用 Lorenz 系统表达式  $r = 34$  , 产生  $x$  分量的 Lorenz 序列 , 采用 SVD 模型 , 取 5000 个数据 , 前 1000 作为训练集 , 预测后 4000 , 并抽取 150 个数据作为检验数据 , 嵌入维数取为 4 , 在调节量 , 超参数  $\{\gamma, \sigma\} = \{50, 2\}$  进行训练建模 , 比较预测 4000 点与实际值的



误差图 6 , 预测结果 , 图 (a) 是一步预测值 0 至 4000 点预测结果的比较 , 图 (b) 是 8 步预测的结果 .

建模固定  $\{\gamma, \sigma\} = \{50, 2\}$  , 改变嵌入维数  $d$  预测 Henon 时间序列 . 可以观察到在嵌入维数小于 2 时 , MSE 随嵌入维数增大 .

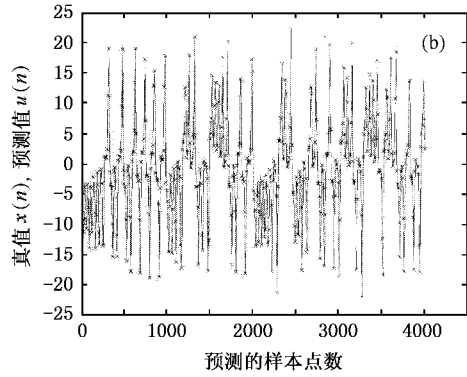


图 6 Lorenz 时间序列预测值 ( × ) 与实际值 ( — ) 的比较 ( a ) 一步预测的比较 ( MSE = 0.0022 , eerr = 0.00127 ) ( b ) 8 步预测的比较 ( MSE = 0.0031 , eerr = 0.00157 )

通过检验 , 结果相对误差 eerr = 0.00157 , 均方误差 MSE = 0.0031 , 误差和均方误差分别如图 7 , 8 .

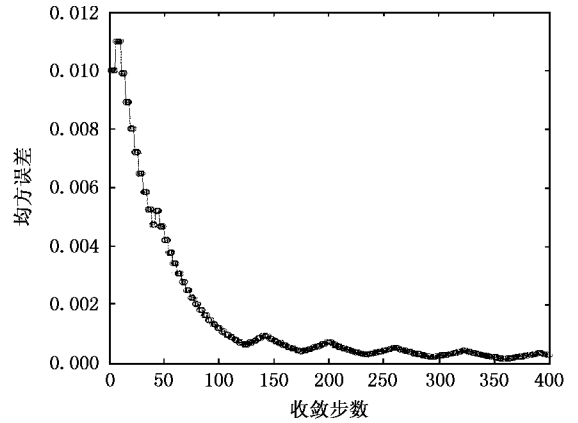
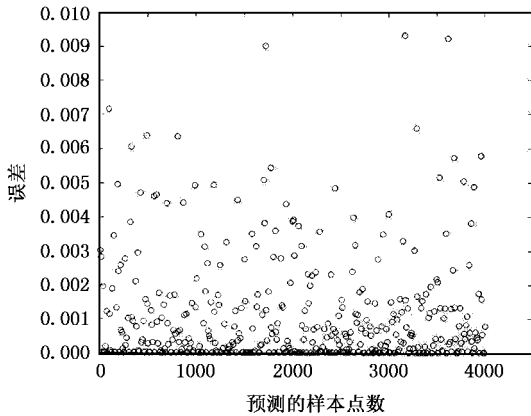


图 7 Lorenz 时间序列 8 步预测值与实际值的误差

图 8 均方误差随步长的收敛

### 4.3. 微分定义的 Rossler 混沌时间序列预测

Rossler 混沌方程是 1976 年由 Rossler 提出<sup>[16]</sup> , 是一种最低维数方程的混沌现象 , 形式是一个常微系统方程  $\dot{x} = f(x, \lambda)$  , 将  $x$  看作是  $t$  的函数 , 据此 Rossler 定义了一个  $x$  分量的时间序列 , 混沌从  $c = 5.7$  时出现 , 其关联维数为 2.1 , 最小嵌入维数 3 , 表达式为

$$\begin{aligned} \dot{x} &= -y - z, \\ \dot{y} &= x + ayx, \\ \dot{z} &= b - z(x - c). \end{aligned}$$

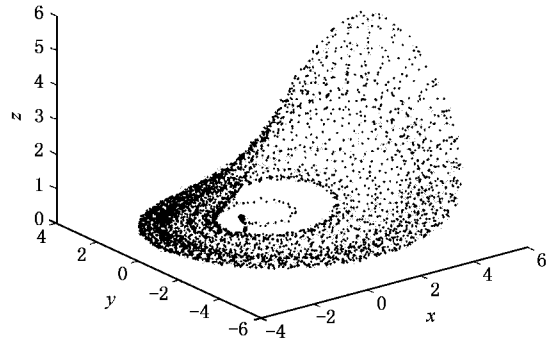


图 9 Rossler 时间序列重构的相轨道

通过系统的方程解预测  $x, y, z$  方向时间序列 , 相空间重构的序列几何相轨道表示如图 9 . 空间

Rosser 混沌运动具有局部不稳定性和整体的确定性,在 3 维空间中,Rosser 的轨道存在折叠,但不正交,并且可以得到近似相同的运动状态行为,可以看出这里 Rosser 系统  $a = 0.398, b = 2, c = 4$  时,利用 Rosser 表达式产生 Rosser 的  $x$  分量时间序列,采用 SVD 模型,取 3000 个数据,前 1000 作为训练集,后

2000 个数据作为检验数据,嵌入维数取为 4,在调节量、超参数  $\{\gamma, \sigma\} = \{50, 2\}$  进行训练建模,1 步预测和 8 步预测如图 10(a, b),8 步预测收敛速度如图 11,8 步预测值误差如图 12,实际值的与预测结果在 1 步预测和 8 步预测情况下误差在  $10^{-2}$  量级。

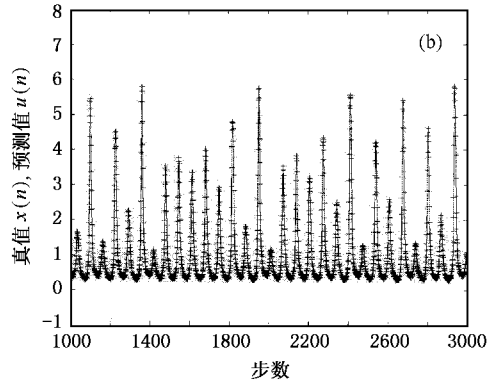
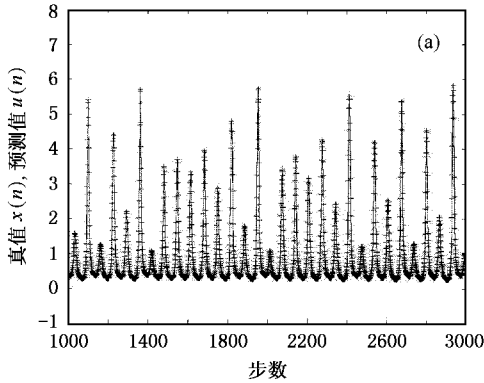


图 10 Rosser 的  $x$  分量时间序列预测值(+)与真实值(-)的不同 (a)1 步预测(MSE = 0.008,  $E_{err} = 0.0181$ )(b)8 步预测(MSE = 0.00223,  $E_{err} = 0.00854$ )

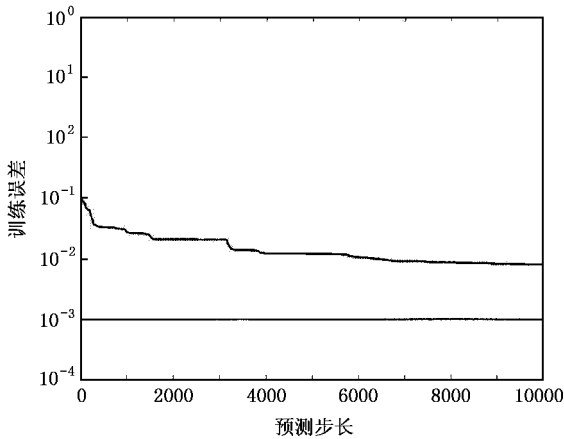


图 11 Rosser 的  $x$  分量时间序列的 8 步预测随步长收敛(MSE = 0.00223,  $E_{err} = 0.00854$ )

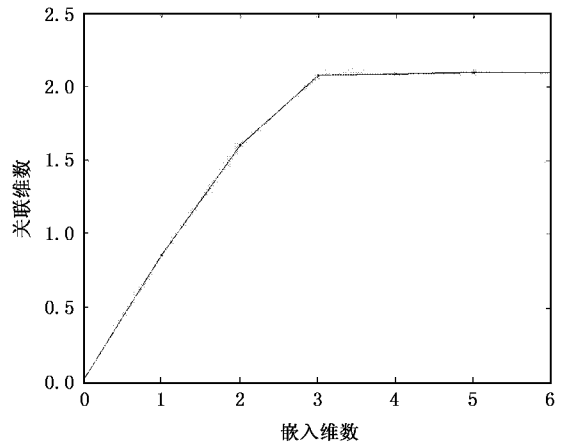


图 13 嵌入维数与关联系数的关系曲线

对于 Rosser 序列关联维数,随着嵌入维数增大,关联维数趋于某一确定值,而影响上述三种混沌序列的噪声随嵌入维数增大,其关联维数增加,如图 13,随嵌入维数,关联维数趋近 2.1<sup>[17-19]</sup>。

在采样时间间隔  $\Delta t = 1$  下,预测采用数据样本 1000,训练样本 1800,三种混沌序列相对误差分析和均方误差作对比分析如表 1。

表 1 三种混沌序列预测误差分析

混沌时间序列	嵌入维数	收敛系数	$\mu$ 有效样本率	均方误差/%
Henon	$\mathfrak{X}(p = 3)$	0.01	81.5	0.0177
Lorenz	$\mathfrak{X}(p = 3)$	0.01	93.5	0.0169
Rosser	$\mathfrak{X}(p = 4)$	0.01	89.5	0.0119

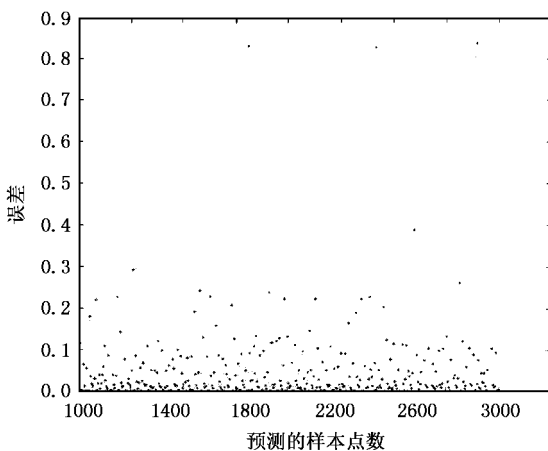


图 12 Rosser 时间序列 8 步预测值与实际值的误差

支持向量域的出错率和核函数的选择有关,分别进行多项式核函数和径向基核函数和两层神经网络核函数试验对比,比较不同类型 SVD 机器学习决策规则的性能,试验结果如表 2. SVD 与多项式和两层神经网络比较出错概率较小,可以获得对函数逼近的更多支持向量数的支持.逼近的精度要求越高,

要求支持向量数量越多. SVD 径向基支持向量方法,可以用较小的支持向量数数目,构建更多的信号预测的网络点.

应用上述径向基 SVM 对 Henon(  $\delta = 0.35, m = 1.5$  ), Lorenz 和 Rossler 的预测比较.

表 2 SVD 与多项式和两层神经网络支持向量数及出错率比较

分类器类型	分类器	支持向量数	出错率/%
多项式核函数	$K(x, x_i) = \left(\frac{x \cdot x_i}{256}\right)^3$	274	4.1
SVD 径向基核函数	$K(x, x_i) = \exp\left(-\frac{\ x - x_i\ }{256 \times 0.3}\right)$	291	4.0
NN 神经网络核函数	$K(x, x_i) = \tanh\left(\frac{2\ x - x_i\ }{256} - 1\right)$	254	4.2

表 3 支持向量域 SVD 支持向量数、收敛步数、预测时间和正确率比较

分类器类型	支持向量数	MSE	收敛步数	预测时间	正确率/%
多项式	274	0.01	720	153.8	96
SVM 径向基函数 RBF	291	0.0037	404	128.7	96.1
NN 两层神经网络	254	0.0039	400	131.5	95.8
支持向量域 SVD	146	0.004	200	108.5	94

表 3 显示: 1) 支持向量域 SVD 需要的支持向量数比支持向量、多项式和神经网络需要少一半,需要的样本数也下降为原来的一半,其正确率仅下降不到 2%. 2) 收敛步数和预测时间上分别比原来有较大的提高,约提高 15.7%. 结果表明 SVD 方法比支持向量机的 SVM 方法的学习效率高,这一方面由于支持向量域搜索的样本及支持向量数减小,搜索范围小,另一方面也由于支持向量 SVD 与 SVM 学习过程比较,避免了 SVM 中大量样本的递归最小二乘中的矩阵运算.

## 5. 结 论

根据相空间重构、Vapnik 的统计学习支持向量机和 Tax 的支持向量域理论,对确定性系统产生的随机混沌时间序列采用最小二乘法支持向量域方

法,通过小样本的训练,1) 将一个已知的训练集合通过非线性方法,映射到一个更高维的特征空间中,利用混沌时间序列特点,采用嵌入维数方法完成了混沌时间序列预测. 2) 获得了混沌时间序列的非线性预测结论,预测结果表明小于最小嵌入维数的嵌入,随维数提高其预测的相对误差和均方误差迅速减小,超出最小嵌入维数的嵌入,嵌入维数的增加对预测性能没有明显的改善. 3) 文中对 SVM 的几种判别决策规则进行了对比, SVD 法降低了一半分类辨识的支持向量机数目,预测的网格点数却增加 10—20 倍,在少一半的样本数下,收敛速度提高 2 倍,优化时间是单纯 SVM 的一半,正确率损失 2%,是实践中可以接受的. 4) 通过小样本建立的多分类器,可以极大地提高运算速度,减小分类错误率. 所以混沌时间序列的 SVD 方法在实践上是有意义的可行的方法.



- [ 1 ] Tax D , Duin R 1999 *Data Domain Description Using Support Vectors* , *Pro European Symposium Artificial Neural( Brussel :D Facto )* 251
- [ 2 ] Louvier H , Rico M. 2002 *Phy. Rev. E* **65** 016203
- [ 3 ] Stark J , Hardy K 2003 *Science* **301** 1192
- [ 4 ] David M J 1999 *Pattern Recognition Letter* **20** 1191
- [ 5 ] Takens F 1981 *Detecting strange attractors in turbulence Lecture notes in Mathematics* ( Berlin :Springer-Verlag ) p366
- [ 6 ] Ott E , Grebogi C , Jorke J 1990 *Phys. Rev. Letter* **64** 1196
- [ 7 ] Shukla J 1998 *Science* **282** 728
- [ 8 ] Grassberger P , Procaccia I 1983 *Phys. Rev. Letter* **50** 346
- [ 9 ] 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 ]
- [ 10 ] Gan J C , Xiao X C 2003 *Acta Phys. Sin.* **52** 1097 ( in Chinese )
- [ 11 ] Gan J C , Xiao X C 2003 *Acta Phys. Sin.* **52** 1102 ( in Chinese )  
[ 甘建超、肖先赐 2003 物理学报 **52** 1102 ]
- [ 12 ] Zhang J S , Xiao X C 2000 *Acta Phys. Sin.* **49** 2333 ( in Chinese )  
[ 张家树、肖先赐 2000 物理学报 **49** 2333 ]
- [ 13 ] Henon M 1976 *Commun. Math. Phys.* **50** 69
- [ 14 ] Lorenz E N 1963 *J. Atmos. Sci.* **20** 130
- [ 15 ] Lorenz E N 1993 *The Essence of Chaos* ( Washington :University of Washington Press ) p13
- [ 16 ] Rossler O E 1976 *Phys Lett. A* **57** 397
- [ 17 ] Zhang J S , Li H C , Xiao X C 2005 *Chin. Phys.* **14** 49
- [ 18 ] Zhang S 2003 *Chin. Phys.* **12** 594
- [ 19 ] Gan J C 2004 *Chin. Phys.* **13** 312

## Prediction of chaotic time sequence using least squares support vector domain \*

Ren Ren<sup>1)†</sup> Xu Jin<sup>2)</sup> Zhu Shi-Hua<sup>1)</sup>

1) *School of Electronics and Information Engineering , Xi 'an Jiaotong University ,Xi 'an 710049 ,China )*

2) *Institute of Biomedical Engineering , Xi 'an Jiaotong University ,Xi 'an 710049 ,China )*

3) *School of Science , Xi 'an Jiaotong University ,Xi 'an 710049 ,China )*

( Received 17 May 2005 ; revised manuscript received 13 June 2005 )

### Abstract

Starting from the support vector domain model , the paper establishes SVD predictive models of chaos time series as well as chaos phase trace of non-linear map , based on Takens phase space delay reconstructing theory . We adopted the method of data set as support object elements . Machine self-learning reduces error upper limit of the generalized model . The three chaos time series , Henon/Lorenz/Rosler are predicted by least square . The prediction result indicates that the predictive model makes the set to be mapped into an eigen space of higher dimensions , and the series is predicted by embed dimensions . The predictive error changes with the increase of embed dimension to a constant . Compared with SVM , the SVD requires smaller support vector , and has faster convergence rate . It has robustness characteristics with adaptive flexible kernel function choice . The predicted net points are ten to twenty times more than SVM . Under the conditions of small sample , non-linear , and unknown probability density , the predicted series is in concordance with the series' true value .

**Keywords** : support vector domain , chaos , least squares , time series prediction

**PACC** : 0545

\* Project supported by the National Natural Science Foundation of China ( Grant Nos. 60372055 , 30400101 ) .

† Corresponding author . E-mail : renren@mail.xjtu.edu.cn