

# 基于小波核偏最小二乘回归方法的混沌系统建模研究<sup>\*</sup>

李 军<sup>†</sup> 董海鹰

(兰州交通大学自动化与电气工程学院, 兰州 730070)

(2008 年 1 月 13 日收到, 2008 年 4 月 22 日收到修改稿)

基于核学习的强大非线性映射能力, 结合用于回归建模的线性偏最小二乘(PLS)算法, 提出一种小波核偏最小二乘(WKPLS)回归方法. 该方法基于支持向量机使用的经典核函数技巧, 将输入映射到高维非线性的特征空间, 在特征空间中, 构造线性的 PLS 回归模型. PLS 方法利用输入与输出变量之间的协方差信息提取潜在特征, 而可允许的小波核函数具有近似正交以及适用于信号局部分析的特性. 因此, 结合它们优点的 WKPLS 方法显示了更好的非线性建模性能. 将 WKPLS 方法应用在非线性混沌动力系统建模上, 并与基于高斯核的核偏最小二乘方法进行了比较. 仿真结果表明, 所提出的 WKPLS 方法能精确地逼近未知非线性混沌动力系统, 且在同等条件下的逼近精度较高.

关键词: 小波核, 偏最小二乘回归, 混沌系统, 建模

PACC: 0545

## 1. 引 言

确定性的动力系统所产生的混沌特性是一种固有的非线性现象. 随着非线性混沌动力学的发展, 混沌系统建模、辨识与控制的研究已经成为一个非常重要的研究方向, 并已在通信与信号处理、非线性控制等领域中得到广泛应用<sup>[1-3]</sup>. 控制非线性混沌动力系统的重要前提之一就是建立混沌系统好的逼近模型. 然而, 对于许多非线性系统, 尤其是混沌系统而言, 建立精确、完美的数学模型通常很困难, 许多学者都对混沌系统的复杂结构和隐含特征参数等进行了分析和研究<sup>[1, 2]</sup>. 因此, 在控制混沌等复杂非线性动力系统之前所进行系统辨识是重要甚至是必要的步骤之一.

选择一个好的辨识模型, 逼近未知系统的非线性动态特性是非线性动力系统辨识中的一个具有挑战性的问题. 前馈神经网络、径向基函数(RBF)神经网络、模糊系统等模型由于对未知非线性动态系统有着良好的逼近特性, 常常被用于混沌系统的时间序列预测、混沌系统建模与控制中, 并取得了成功应用<sup>[3-10]</sup>. 然而, 随着网络规模的增大, 也存在着由于模型参数过多以及使用非线性优化算法所引起的

计算负担过重以及模型“过拟合”等问题.

基于统计学习理论的支持向量机(support vector machine, 简记为 SVM)具有良好的预测能力, 目前它已被证实是强有力的机器学习方法之一, 广泛应用于非线性动力系统辨识和混沌时间序列预测中<sup>[11-14]</sup>, 而其非线性特性完全归因于经典的核变换技术<sup>[15]</sup>. 偏最小二乘(partial least squares, 简记为 PLS)方法作为第二代回归建模方法, 最初在化学计量领域使用<sup>[16, 17]</sup>, 近年来也被应用在机器学习领域, 用于回归和分类研究<sup>[18]</sup>. 它利用输入和输出变量之间的协方差信息提取数据的潜在特征, 能克服观测变量多于观测样本数的情形以及变量之间存在的多重共线性, 因此不会出现神经网络等建模方法所产生的模型“过拟合”情况. PLS 还具有运行快、易完成和便于调整的特点<sup>[19]</sup>. 结合 PLS 和 SVM 各自的优点, 基于核变换技术的核偏最小二乘(kernel partial least squares, 简记为 KPLS)回归方法在文献 [20, 21] 中提出, 将 PLS 线性回归延伸到高维非线性特征空间上.

多维张量积小波核函数具有小波信号拥有的局部分析和特征提取的优点, 能够逼近任意非线性函数, 与常规的高斯核函数相比是正交或是近似正交的<sup>[22]</sup>. 通过构造平移不变小波核函数, 本文提出一

<sup>\*</sup> 甘肃省自然科学基金(批准号 3ZS042-B25-026)和兰州交通大学“青蓝”人才计划资助的课题.

<sup>†</sup> E-mail: lijun691201@mail.lzjtu.cn

种新颖的小波核偏最小二乘(WKPLS)回归方法,在高维特征空间中,应用线性 PLS 方法建模,它与基于高斯核的 KPLS 方法相比,进一步提高了学习性能.在算法实现上,WKPLS 回归方法仅需必备的线性代数知识,避免了神经网络由于参数初始化不当而引起的收敛困难等缺点,也勿需求解复杂非线性优化问题,因此,它比 SVM 和神经网络方法更具有竞争力.将 WKPLS 回归方法应用到复杂的、具有混沌行为的非线性动力系统建模上,对蔡氏混沌电路、Lorenz 混沌吸引子、糖酵解混沌振荡器等进行逼近与辨识,并与基于高斯核的 KPLS 方法进行比较,验证所提出的 WKPLS 方法的有效性.

## 2. WKPLS 回归方法

WKPLS 是建立在 PLS 方法使用的提取潜在特征变量技术基础上完成的<sup>[20,21]</sup>. 因此,首先介绍原始空间上的 PLS 方法,然后将输入映射到高维特征空间(对偶空间),使用非线性核变换技术给出 KPLS 方法.最后基于小波分析技术,在非线性的核函数的构造上给出满足核函数条件的可允许多维张量积小波核函数.

### 2.1. PLS 方法

作为化学计量领域通用的第二代回归建模方法——PLS 方法,其计算简洁,可提取输入、输出数据潜在的综合特征,从而实现高维数据可视化和回归建模.给定训练数据集  $\{(x_1, y_1), \dots, (x_l, y_l)\}$ ,  $x_i \in \mathbb{R}^N, y_i \in \mathbb{R}^M$ , 构成输入数据矩阵  $X \in \mathbb{R}^{l \times N}$  和输出数据矩阵  $Y \in \mathbb{R}^{l \times M}$ . PLS 方法提供了克服样本数低于其维数时建模困难的有效途径.与主成分分析只利用  $X$  的方差信息相比,它同时利用输入  $X$  和输出  $Y$  的协方差选择潜在特征变量(成分).令  $w_j$  和  $\sigma_j$  是  $X_j^T Y$  对应的第一个左奇异矢量和奇异值,其中  $X_j^T$  表示对  $X$  矩阵进行了第  $j$  ( $j = 1, 2, \dots, L$ ) 次压缩映射后的转置,上标符号  $T$  表示矩阵转置. PLS 提取新的特征方向  $w_j$  与特征  $X_j w_j$  是通过计算  $X_j^T Y$  的第一个左奇异矢量及对矩阵  $X_j$  的压缩映射完成的.对  $X_j$  的压缩映射通过将其列矢量投影在与  $X_j w_j$  正交的空间上获得,归一化矢量  $q$  的投影矩阵为  $(I - qq^T)$ , 因此可得

$$X_{j+1} = \left( I - \frac{X_j w_j w_j^T X_j^T}{w_j^T X_j^T X_j w_j} \right) X_j \quad (j = 1, 2, \dots, L). \quad (1)$$

(1) 式表明,所选择的特征  $X_j w_j$  是彼此正交的,因为它们都是  $X_j$  的列矢量的线性组合,而  $X_j$  又被重复投影在以前的  $X_i w_i$  ( $i < j$ ) 的正交补空间中.特征方向  $w_j$  是  $X_j^T Y$  的第一个左奇异矢量,可由迭代法计算,即

$$w_j = X_j^T Y Y^T X_j w_j. \quad (2)$$

由于特征方向  $w_j$  是相对于压缩矩阵  $X_j$  定义的,为使用初始数据矢量直接计算 PLS 所提取的潜在特征,可考虑在高维特征(对偶)空间  $F$  中的测试数据  $\phi(x)$ .

若将  $\phi(x)^T$  看作  $X$  的行矢量,令  $\phi_1(x) = \phi(x)$ , 应用(1)式得

$$\phi_{j+1}(x)^T = \phi_j(x)^T (I - w_j p_j^T). \quad (3)$$

其中

$$p_j = \frac{X_j^T X_j w_j}{w_j^T X_j^T X_j w_j}$$

为载荷矢量.(3)式可改写为

$$\phi(x)^T = \phi_{L+1}(x)^T + \sum_{j=1}^L \phi_j(x)^T w_j p_j^T. \quad (4)$$

对于测试数据所提取的所有特征为  $\hat{\phi}(x)$  其  $L$  个成分应为  $\hat{\phi}(x) = (\phi_j(x)^T w_j)_{j=1}^L$ , 勿需迭代计算  $\phi_j(x)^T$ . 定义  $W$  与  $P$  分别为  $L$  个特征方向矢量  $w_j$  与载荷矢量  $p_j$  作为列矢量所构成的矩阵,由(4)式得

$$\begin{aligned} \phi(x)^T W &= \phi_{L+1}(x)^T W + \sum_{j=1}^L \phi_j(x)^T u_j p_j^T W \\ &= \phi_{L+1}(x)^T W + \hat{\phi}(x)^T P^T W. \end{aligned} \quad (5)$$

PLS 的特征矢量之间正交且共轭,因此

$$\begin{aligned} (I - w_s p_s^T) w_j &= w_j \quad (s > j), \\ (I - w_j p_j^T) w_j &= 0, \end{aligned}$$

则有

$$\begin{aligned} \phi_{L+1}(x)^T W &= \phi_j(x)^T \prod_{i=j}^L (I - w_i p_i^T) w_j \\ &= 0 \quad (j = 1, 2, \dots, L). \end{aligned} \quad (6)$$

由(5),(6)式即可求出关于测试数据矢量所提取的新特征

$$\hat{\phi}(x)^T = \phi(x)^T W (P^T W)^{-1}. \quad (7)$$

多输入、多输出变量的 PLS 回归模型能写成矩阵形式

$$\begin{aligned} Y &= XB + E, \\ \hat{Y} &= XB, \end{aligned} \quad (8)$$

其中  $B$  是  $(N \times M)$  的回归系数矩阵,  $E$  是  $(N \times M)$  的残差矩阵,  $\hat{Y}$  是模型的预测输出.

结合 (7) 式, PLS 回归模型方程的回归系数矩阵可表示为<sup>[15, 18, 19]</sup>

$$B = W(P^T W)^{-1} C^T, \quad (9)$$

其中  $C$  矩阵的维数为  $(M \times L)$ , 由  $c_j$  作为列矢量所构成,

$$c_j = \frac{Y^T X_j w_j}{w_j^T X_j^T X_j w_j}.$$

### 2.2. KPLS 回归方法

基于再生核希尔伯特空间理论, KPLS 方法将输入、输出数据映射到高维特征空间  $F$  中, 将线性 PLS 方法延伸到非线性领域. 它保持了线性 PLS 方法的优点, 其思想是将高维特征空间看作原始空间的对偶. 基于核变换技术, 在对偶空间中建立线性 PLS 回归模型.

定义  $u_j$  为对偶空间的特征方向, 与原始空间的特征方向  $w_j$  相对应, 应保证  $w_j$  位于  $X$  的行空间上, 即

$$a_j w_j = X_j^T u_j,$$

其中  $a_j$  为保证归一化的尺度因子, 表示为如下矩阵形式:

$$W = X^T U \text{diag}(a)^{-1}, \quad (10)$$

其中  $\text{diag}(a)$  表示对角矩阵, 其元素  $\text{diag}(a)_{ii} = a_i$ ,  $U$  由列矢量  $u_j$  构成.

数据核矩阵定义为

$$K = X X^T.$$

相应 (2) 式, 则可得到计算  $u_j$  的递归表达, 即

$$\begin{aligned} u_j &= Y_j Y_j^T X_j X_j^T u_j \\ &= Y_j Y_j^T K_j u_j. \end{aligned}$$

因此, 在高维特征(对偶)空间中计算  $u_j$  需要对输出矩阵  $Y$  进行压缩映射, 而在原始空间中, 提取 PLS 的特征并不需要压缩输出矩阵.

令  $t_j = a_j X_j w_j = K_j u_j$ , 它为原始空间中  $c_j$  的对偶表示. 在对偶空间中, 为计算  $u_j$  和  $t_j$ , 要求出在第  $j$  个阶段的特征提取时的数据核矩阵  $K_j$ . 由 (1) 式知, 对  $X_j$  的压缩映射可表示为

$$X_{j+1} = \left( I - \frac{t_j t_j^T}{t_j^T t_j} \right) X_j,$$

所以  $K_j$  的压缩映射可表示为

$$K_{j+1} = X_{j+1} X_{j+1}^T$$

$$\begin{aligned} &= \left( I - \frac{t_j t_j^T}{t_j^T t_j} \right) X_j X_j^T \left( I - \frac{t_j t_j^T}{t_j^T t_j} \right) \\ &= \left( I - \frac{t_j t_j^T}{\|t_j\|^2} \right) K_j \left( I - \frac{t_j t_j^T}{\|t_j\|^2} \right). \quad (11) \end{aligned}$$

由于特征的正交性, 即

$$(t_j^T) t_i = a_j a_i w_j^T X_j^T X_i w_i = 0 \quad (j > i),$$

$$\left( I - \frac{t_j t_j^T}{t_j^T t_j} \right) t_j = t_j \quad (j > i),$$

得到

$$X_j^T t_j = X^T t_j. \quad (12)$$

因此原始空间上的  $p_j$  可表示为

$$p_j = \frac{X_j^T X_j w_j}{w_j^T X_j^T X_j w_j} = a_j \frac{X_j^T t_j}{t_j^T t_j},$$

写成矩阵形式

$$P = \text{diag}(a) \text{diag}(t_i^T t_i)^{-1} X^T T, \quad (13)$$

其中  $T$  的维数为  $(l \times L)$ , 它是  $t_j$  作为列矢量所构成的矩阵.

考虑  $j > i$  时  $X_j w_j$  与  $t_i$  的正交性, 有

$$\begin{aligned} c_j &= \frac{Y_j^T X_j w_j}{w_j^T X_j^T X_j w_j} \\ &= \frac{Y_j^T X_j w_j}{w_j^T X_j^T X_j w_j} \\ &= a_j \frac{Y_j^T t_j}{t_j^T t_j}. \quad (14) \end{aligned}$$

所以  $Y_j$  的压缩映射可表示为

$$\begin{aligned} Y_{j+1} &= \left( I - \frac{t_j t_j^T}{t_j^T t_j} \right) Y_j \\ &= Y_j - t_j c_j^T. \quad (15) \end{aligned}$$

列矢量  $c_j$  构成  $C$  矩阵, 列矢量  $t_j$  构成  $T$  矩阵,

(14) 式的矩阵形式为

$$C = Y^T T \text{diag}(t_i^T t_i)^{-1} \text{diag}(a), \quad (16)$$

其中  $\text{diag}(t_i^T t_i)$  表示对角矩阵, 其元素  $\text{diag}(t_i^T t_i)_{ii} = t_i^T t_i$ .

将 (10) (13) (16) 式代入 (9) 式, 可得到原始空间中的 PLS 模型回归系数矩阵的对偶表示. 首先得到

$$\begin{aligned} W(P^T W)^{-1} &= (X^T U \text{diag}(a)^{-1}) \{ \text{diag}(a) \text{diag}(t_i^T t_i)^{-1} \\ &\quad \times T^T X X^T U \text{diag}(a)^{-1} \}^{-1}, \\ B &= W(P^T W)^{-1} C^T \\ &= X^T U (T^T X X^T U)^{-1} T^T Y \\ &= X^T U (T^T K U)^{-1} T^T Y. \quad (17) \end{aligned}$$

然后计算对偶空间中的回归系数矩阵  $\alpha$ . 与 (10) 式

类似,应要求  $B = X^T \alpha$ . 对于特征空间上的数据  $\phi(x)$ ,可得

$$\phi(x)^T B = \phi(x)^T X^T \alpha = \kappa^T \alpha,$$

其中  $\kappa$  为核矩阵  $K$  的列矢量,即

$$\kappa_i = \phi(x_i), \phi(x).$$

因此,维数为  $(l \times M)$  的高维特征空间中的回归系数矩阵  $\alpha$  应为

$$\alpha = U(T^T K U)^{-1} T^T Y. \tag{18}$$

为了消除回归模型中偏置项的影响,还应对映射到特征空间中的数据进行中心化处理,这可以通过对训练数据及测试数据构成的核矩阵进行如下变换来实现<sup>[15,18,20]</sup>:

$$K = \left( I - \frac{1}{l} \xi_i \xi_i^T \right) \tilde{K} \left( I - \frac{1}{l} \xi_i \xi_i^T \right), \tag{19}$$

其中  $\tilde{K}$  为未进行中心化前的原始核矩阵,  $\xi_i$  为元素为 1 的列矢量,  $I$  为单位阵.

最终,多维输入、多维输出 KPLS 回归模型的预测输出可写成如下矩阵形式:

$$\begin{aligned} \hat{Y} &= K\alpha \\ &= K U (T^T K U)^{-1} T^T Y. \end{aligned} \tag{20}$$

上述 KPLS 回归方法可概括如下:第一步,对输入、输出数据矩阵及核矩阵进行中心化预处理,并令  $K_1 = K, Y_1 = Y$ . 第二步,顺序提取  $i = 1, 2, \dots, L$  个潜在特征,首先取  $u_i = Y_i$  的第一列,作规范化处理,即  $u_i = u_i / \|u_i\|$ ,重复迭代  $u_i = Y_i Y_i^T K_i u_i$ ,直至收敛.然后计算  $t_i = K_i u_i, c_i = Y_i^T t_i / \|t_i\|^2$ ,由(15)式计算  $Y_{i+1} = Y_i - t_i c_i^T$ ,由(11)式计算  $K_{i+1} = (I - t_i t_i^T / \|t_i\|^2) K_i (I - t_i t_i^T / \|t_i\|^2)$ . 转回至第二步进行下一循环操作.第三步,循环结束后,获得  $U = [u_1, \dots, u_L], T = [t_1, \dots, t_L]$ ,计算  $\alpha = U(T^T K U)^{-1} T^T Y$  最终模型的输出为  $\hat{Y} = K\alpha = K U (T^T K U)^{-1} T^T Y$ .

### 2.3. WKPLS 回归方法

核是定义在原始空间的一个双变量函数,但它却实现了某一高维空间的内积,提供了把线性学习机扩展到非线性学习机的手段. Mercer 定理<sup>[15]</sup>给出了一种判定核函数正定的条件,而现有的 KPLS 方法大都使用高斯 RBF 核函数,即  $k(x, z) = \exp(-\|x - z\|^2 / \sigma^2)$ ,实现具有非线性变换的 PLS 回归算法. 本文基于小波分析技术<sup>[23]</sup>,进一步将应用于 KPLS 方法中的核函数延伸为可允许的多维张量积小波核函数,以提高其学习性能. 在 Mercer 定

理基础上,首先给出判断和构建核函数的定理 1.

定理 1<sup>[24,25]</sup> 平移不变核函数  $k(x, z) = k(x - z)$  是一个允许的支持向量核函数,当且仅当傅里叶变换

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

成立.

依据定理 1,可证明定理 2.

定理 2<sup>[22]</sup> 若  $h(\cdot)$  为母小波函数,伸缩因子为  $\sigma$ ,平移因子分别记为  $m$  与  $d$ ,其中  $m, d, x, z \in \mathbb{R}^N, x^i, z^i, \sigma, m^i, d^i \in \mathbb{R}$ ,则下式给出的满足平移不变性定理的张量积小波核是可允许的多维支持向量核函数:

$$k(x, z) = k(x - z) = \prod_{i=1}^N h\left(\frac{x^i - z^i}{\sigma}\right). \tag{21}$$

为构造一种平移不变小波核函数,不失一般性,可选择墨西哥帽(Mexican hat)母小波,即

$$h(x) = (N - \|x\|^2) \exp(-\|x\|^2 / 2), \tag{22}$$

其中  $N$  为输入  $x$  的维数. 在文献<sup>[22]</sup>的基础上,又可证明定理 3.

定理 3 若给定墨西哥帽母小波函数如(22)式所示,伸缩因子为  $\sigma$ ,其中  $x, z \in \mathbb{R}^N, \sigma, x^i, z^i \in \mathbb{R}$ ,则下式所表示的小波核函数就是一种可允许的多维张量积的支持向量核函数:

$$\begin{aligned} k(x, z) &= \prod_{i=1}^N \left( \left( N - \frac{(x^i - z^i)^2}{\sigma^2} \right) \right. \\ &\quad \left. \times \exp\left( -\frac{(x^i - z^i)^2}{2\sigma^2} \right) \right). \end{aligned} \tag{23}$$

证明 根据定理 1 和定理 2,只需证下列不等式成立即可:

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

其中

$$k(x) = \prod_{i=1}^N h\left(\frac{x^i}{\sigma}\right).$$

由已知条件,有

$$k(x) = \prod_{i=1}^N \left( \left( N - \frac{(x^i)^2}{\sigma^2} \right) \exp\left( -\frac{(x^i)^2}{2\sigma^2} \right) \right).$$

利用

$$\int_{-\infty}^{+\infty} \exp(-\beta x^2) dx = \sqrt{\pi} / \sqrt{\beta},$$

可得

$$\int_{\mathbb{R}^N} \exp(-j\omega \cdot x) k(x) dx$$

$$\begin{aligned}
&= \int_{\mathbb{R}^N} \exp(-\mathbf{j} \boldsymbol{\omega} \cdot \mathbf{x}) \\
&\quad \times \prod_{i=1}^N \left( \left( N - \frac{(x^i)^2}{\sigma^2} \right) \exp\left(-\frac{(x^i)^2}{2\sigma^2}\right) \right) d\mathbf{x} \\
&= \prod_{i=1}^N \int_{-\infty}^{+\infty} \left( N - \frac{(x^i)^2}{\sigma^2} \right) \exp\left(-\frac{(x^i)^2}{2\sigma^2}\right) \\
&\quad \times \exp(-j\omega^i x^i) dx^i \\
&= \prod_{i=1}^N \left( \int_{-\infty}^{+\infty} N \exp\left(-\frac{(x^i)^2}{2\sigma^2}\right) \exp(-j\omega^i x^i) dx^i \right. \\
&\quad \left. - \int_{-\infty}^{+\infty} \frac{(x^i)^2}{\sigma^2} \exp\left(-\frac{(x^i)^2}{2\sigma^2}\right) \exp(-j\omega^i x^i) dx^i \right) \\
&= \prod_{i=1}^N \sqrt{2\pi} |\sigma| \left( (N-1) + (\omega^i)^2 \sigma^2 \right) \\
&\quad \times \exp\left(-\frac{(\omega^i)^2 \sigma^2}{2}\right).
\end{aligned}$$

所以,小波核函数的傅里叶变换为

$$\begin{aligned}
F[\mathbf{K} \mathbf{I} \boldsymbol{\omega}] &= \prod_{i=1}^N |\sigma| \left( (N-1) + \sqrt{2\pi} (\omega^i)^2 \sigma^2 \right) \\
&\quad \times \exp\left(-\frac{(\omega^i)^2 \sigma^2}{2}\right). \quad (24)
\end{aligned}$$

而伸缩因子(尺度参数) $\sigma \neq 0$ ,且 $N \geq 1$ ,因此 $F[\mathbf{K} \mathbf{I} \boldsymbol{\omega}] \geq 0$ .证毕.

由(20),(23)式可得到多维( $N$ 维)输入、多维( $M$ 维)输出的WKPLS的输出表达式,其任一维的输出为

$$\begin{aligned}
f_s(\mathbf{x}) &= \sum_{i=1}^l \alpha_i^s h(\mathbf{x}_i, \mathbf{x}) \\
&= \sum_{i=1}^l \alpha_i^s \prod_{j=1}^N \left( N - \frac{(x_i^j - x^j)^2}{\sigma^2} \right) \\
&\quad \times \exp\left(-\frac{(x_i^j - x^j)^2}{2\sigma^2}\right) \quad (s = 1, 2, \dots, M).
\end{aligned} \quad (25)$$

这里 $x_i^j$ 表示第 $i$ 个训练数据样本的第 $j$ 个分量; $\alpha_i^s$ 由(18)式给出,是对偶空间上的回归系数矩阵中第 $s$ 个列向量中的第 $i$ 个元素.

小波核函数与传统的核函数相比具有如下优点:首先,小波核函数具有对突变信号逐步精细的描述特性,采用小波核的PLS回归模型因而能够以很高精度逼近任意函数,传统核函数则不具备.其次,小波核函数是正交或近似正交的,而传统高斯RBF核函数是相关甚至是冗余的.最后,小波核函数具有小波信号的局部分析、信噪分离和对突变信号检测的多分辨分析能力.因此小波核函数的非线性建

模能力要优于高斯核函数,这也提高了基于小波核函数的WKPLS回归模型的泛化能力.

小波核学习具有强大非线性变换能力,选择多维张量墨西哥帽小波核函数的WKPLS方法与SVM相比,需要调整的参数少、只需确定潜在特征数目和核参数,无大规模凸优化技术求解全局二次优化问题,其求解仅需线性代数知识,适用于高维、样本数据较少的回归建模问题.

### 3. WKPLS 方法在混沌系统建模中的应用

为了对非线性混沌动力系统建模,需要寻找一个合适的辨识模型,能够容易并精确地捕获系统的非线性动态.线性PLS方法能利用协方差信息,产生潜在特征,完成对输入和输出数据很好的低秩逼近,建立最终的回归模型.因而,在此基础上所提出的WKPLS方法是一种有效的、具有竞争力的辨识模型,能用于逼近任意连续时间多变量函数.与神经网络、SVM相比,KPLS的优点如下:算法实现简单、训练开销小、参数调整容易、只需确定潜在特征数目 $L$ 与小波核函数的超参数、模型的推广性好,而WKPLS方法在此基础上进一步提高了对非线性动力系统的逼近性.

首先,考虑如下 $N$ 维连续时间变量的非线性动力系统通用形式:

$$\begin{aligned}
\frac{d\mathbf{x}(t)}{dt} &= \dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), \mathbf{u}(t)), \\
\mathbf{y}(t) &= \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t)),
\end{aligned} \quad (26)$$

其中 $\mathbf{F}(\cdot)$ 实现 $\mathbb{R}^N \times \mathbb{R} \rightarrow \mathbb{R}^N$ 的映射, $\mathbf{h}(\cdot)$ 实现 $\mathbb{R}^N \rightarrow \mathbb{R}$ 的映射.输入 $\mathbf{u} \in \mathbb{R}$ 和状态 $\mathbf{x}(t) \in \mathbb{R}^N$ 是有界的和可量测的.然后,建立一个辨识模型逼近具有(26)式所表示的 $N$ 维非线性动力系统,其中 $\mathbf{F}(\cdot)$ 和 $\mathbf{h}(\cdot)$ 分别用WKPLS模型逼近.

针对具有能产生混沌特性行为的非线性动力系统,下面将以蔡氏电路、Lorenz吸引子模型以及糖酵解混沌振荡器等实例的建模验证本文所提出方法的有效性,其建模效果还与基于高斯核函数的KPLS辨识模型进行比较,以验证选取小波核的优点.应用核变换技术之前,数据应进行归一化预处理,具有零均值和单位方差.采用均方根误差 $E_{\text{RMS}}$ 评价模型辨识效果,定义

$$E_{\text{RMS}} = \sqrt{\frac{1}{l} \sum_{k=1}^l (\mathbf{x}(t_k) - \hat{\mathbf{x}}(t_k))^2}. \quad (27)$$

### 3.1. 双涡卷吸引子蔡氏混沌系统的建模

蔡氏电路是一个非常复杂的混沌动力系统,由如下的微分方程描述<sup>[5,26]</sup>:

$$\begin{aligned} \dot{x}_1 &= a[x_2 - g(x_1)], \\ \dot{x}_2 &= x_1 - x_2 + x_3, \\ \dot{x}_3 &= -bx_2, \end{aligned} \tag{28}$$

其中

$$g(x_1) = m_1 x_1 + \frac{1}{2}(m_0 - m_1) \times (|x_1 + 1| - |x_1 - 1|)$$

具有分段线性特性. 状态变量  $x_1, x_2, x_3$  初始值设为  $[0.1, 0, -0.1]$  时,取  $a = 9, b = 14.286, m_0 = -1/7, m_1 = -2/7$ ,用四阶五级龙格-库塔法求解(28)式,由初始点的数值解得到 800 个数据点. 它们的流向则可构成双涡卷吸引子的混沌状态,其三维相平面轨迹如图 1 所示,吸引子各状态分量  $x_1, x_2, x_3$  由系统的初始点出发随时间  $t$  在状态空间演变的过程. 为校验 WKPLS 的逼近也能展示混沌特性,建立  $\hat{Y}(t) = [\hat{x}_1(t), \hat{x}_2(t), \hat{x}_3(t)] = F(x_1(t), x_2(t), x_3(t))$  的逼近模型. 用图 1 数据点训练的 WKPLS 辨识模型逼近(28)式所表示的非线性动力系统,输入、输出维数均为 3,潜在特征数目  $L$  选为 10,选择核函数中的超参数  $\sigma = 0.1$ . 用训练好的 WKPLS 模型逼近上述混沌系统,其吸引子的逼近解在三维相平面的表示如图 2 所示. 图 3 示出了系统状态轨迹  $x_1(t)$  的实际输出与所建 WKPLS 辨识模型的输出  $\hat{x}_1(t)$  的比较曲线. 各点的预测误差分布见图 4,整个数据集上对状态轨迹  $x_1, x_2, x_3$  逼近的误差  $E_{RMS}$  分别为 0.0098, 0.0014, 0.0128. 从图 1—图 4 中不难看出, WKPLS 辨识模

型完全能够重构(再生)出该混沌系统的吸引子,对其提供很好的逼近. 整个数据集上,用基于高斯核的 WKPLS 辨识模型对状态轨迹  $x_1, x_2, x_3$  逼近的误差  $E_{RMS}$  分别为 0.0135, 0.0023, 0.0181. 由此可见, WKPLS 方法的逼近效果要明显优于高斯核的 KPLS 方法.

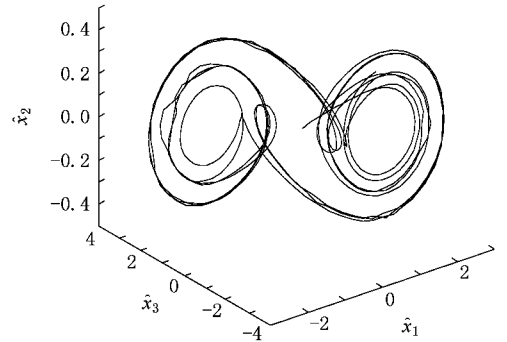


图 2 WKPLS 逼近模型的重构吸引子的相平面图

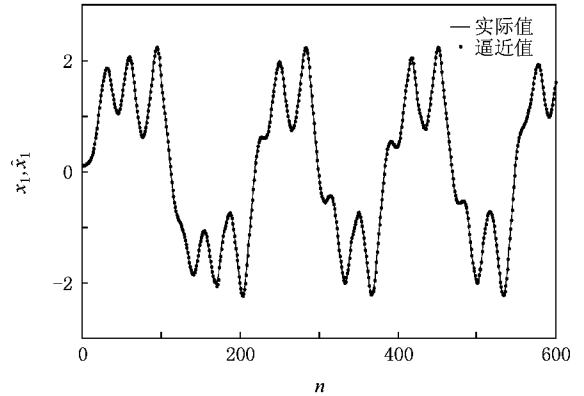


图 3 实际模型状态轨线的输出与 WKPLS 辨识模型的输出

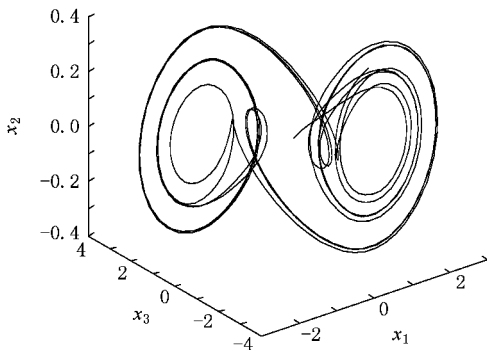


图 1 蔡氏电路双涡卷吸引子的相平面图

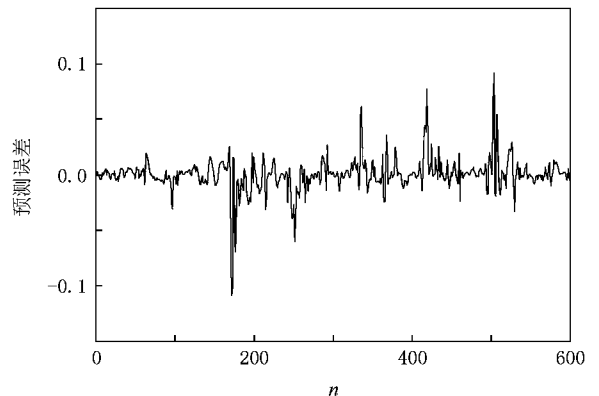


图 4 实际模型状态轨线与 WKPLS 辨识模型的预测误差

### 3.2. Lorenz 混沌系统的建模

Lorenz 吸引子是最著名的三维自治非线性动力学系统之一,能用于流体中湍流的建模.描述它的动力学方程模型为<sup>[27]</sup>

$$\frac{dx_1(t)}{dt} = \sigma(x_2 - x_1),$$

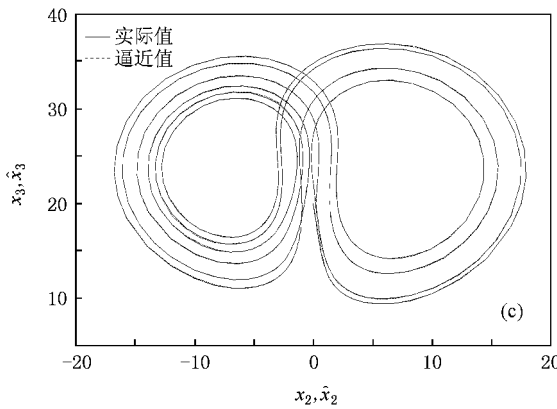
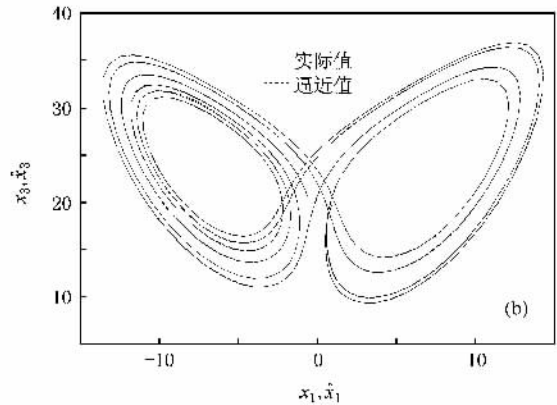
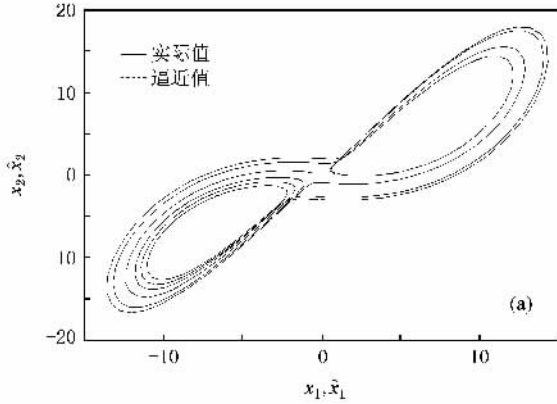


图5 混沌 Lorenz 吸引子与 WKPLS 模型的重构吸引子的二维相平面轨迹状态比较 (a) 投影在  $x_1$  对  $x_2$  的相平面上的比较, (b) 投影在  $x_1$  对  $x_3$  的相平面上的比较 (c) 投影在  $x_2$  对  $x_3$  的相平面上的比较

$$\begin{aligned} \frac{dx_2(t)}{dt} &= (1 + \lambda - x_3)(x_1 - x_2), \\ \frac{dx_3(t)}{dt} &= x_1 x_2 - \gamma x_3. \end{aligned} \quad (29)$$

当 (29) 式中的参数条件满足 (30) 式时,系统的解将展示混沌行为<sup>[26,27]</sup>,即

$$\begin{aligned} \sigma &> \gamma + 1 \\ \lambda &> \frac{(\sigma + 1)(\sigma + \gamma + 1)}{\sigma - \gamma - 1}. \end{aligned} \quad (30)$$

为验证 WKPLS 对 Lorenz 吸引子模型的逼近也能展示混沌行为的情形,取  $(\sigma, \lambda, \gamma) = (10, 24, 2)$ , 系统状态变量  $x_1, x_2, x_3$  初始值设为  $[1, 0, 20]$ . 用四阶五级龙格-库塔法求解 (29) 式,可以由初始点的数值解得到 500 个数据点,其 Lorenz 吸引子的轨迹状态在二维相平面上的流向如图 5 实线所示. 为了校验 WKPLS 逼近模型也能展示混沌特性,建立

$\hat{Y}(t) = [\hat{x}_1(t), \hat{x}_2(t), \hat{x}_3(t)] = F(x_1(t), x_2(t), x_3(t))$  的辨识模型,即用图 5 数据点训练一个辨识器 WKPLS 逼近 (29) 式所表示的非线性动力系统. 模型输入、输出维数均为 3,潜在特征数目  $L$  选为 10,核函数中的超参数  $\sigma = 0.1$ . 用训练好的 WKPLS 模型逼近上述混沌系统,其吸引子轨迹状态的逼近解在相平面的二维表示如图 5 点线所示. WKPLS 辨识模型在整个数据集上对 Lorenz 吸引子状态轨迹  $x_1, x_2, x_3$  逼近的误差  $E_{RMS}$  分别为 0.0087, 0.0093, 0.0182. 图 6 所示为系统的状态轨迹之一  $x_2(t)$  的输出与所建 WKPLS 辨识模型的逼近输出  $\hat{x}_2(t)$  的结果比较,其各点的预测误差分布如图 7 所示. 从图 5、图 6 和图 7 的对照中不难看出, WKPLS 辨识模型完全能够重构出该混沌系统的吸引子,对其提供了极好的逼近. 在整个数据集上,用基于高斯核的

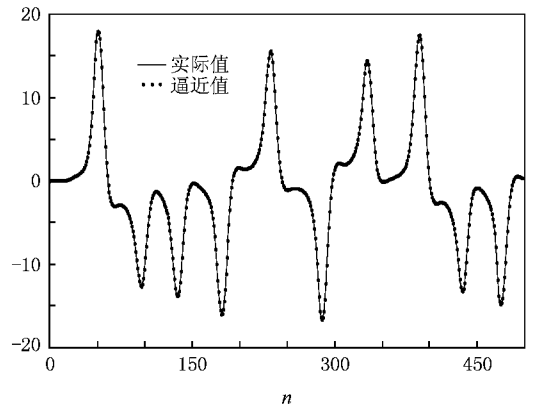


图6 实际模型状态轨迹  $x_2(t)$  与 WKPLS 辨识模型  $\hat{x}_2(t)$  的输出

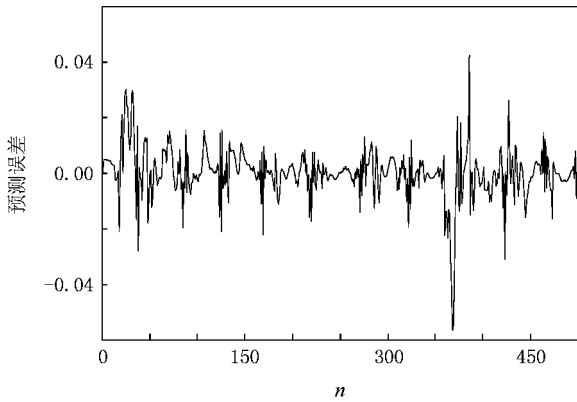


图7 实际模型状态轨线  $x_2(t)$  与 WKPLS 辨识模型  $\hat{x}_2(t)$  的预测误差分布

WKPLS 辨识模型对状态轨迹  $x_1, x_2, x_3$  逼近的误差  $E_{RMS}$  分别为 0.0350, 0.0354, 0.0357. 可见 WKPLS 的逼近效果要优于高斯核的 KPLS 方法, 且高出一个数量级. 基于 WKPLS 的辨识模型极好地复现了系统的混沌行为.

### 3.3. 糖酵解混沌系统的辨识

糖酵解混沌振荡器是具有如下形式的非线性动力系统<sup>[9]</sup>:

$$\begin{aligned} \frac{dx_1(t)}{dt} &= -x_1(t)x_2^2(t) + 0.999 \\ &\quad + 0.42\cos(1.75t), \\ \frac{dx_2(t)}{dt} &= x_1(t)x_2^2(t) - x_2(t), \\ y(t) &= \sin(x_1(t) + x_2(t)), \end{aligned} \quad (31)$$

其中系统状态变量  $x_1, x_2$  初始值设为  $[1.0, 1.0]$ . 用四阶五级龙格-库塔法求解 (31) 式, 由初始点的数值解得到从  $t=0-60$  时间间隔内的 400 个数据点.

为验证 WKPLS 对该具有混沌行为的非线性系统进行逼近也能展示混沌特性, 建立  $\hat{Y}_1(t) = [\hat{x}_1(t), \hat{x}_2(t)] = F(x_1(t), x_2(t))$  及  $\hat{y}(t) = h(x_1(t), x_2(t))$  两个 WKPLS 辨识模型. 用 400 个数据点训练的 WKPLS 辨识模型逼近 (31) 式所表示的非线性动力系统. 辨识模型输入均为二维, 输出分别为二维和一维. 其潜在特征数目  $L$  选为 10, 选择核函数中的超参数  $\sigma$  为 0.1. 用训练好的两个 WKPLS 模型逼近上述非线性动力系统, 其吸引子轨迹状态的逼近解在相平面的二维表示如图 8 中点线所示, 而原系统实际吸引子的轨迹状态的流向则如

图 8 中实线所示. 而实际系统的输出  $y(t)$  与第二个 WKPLS 辨识模型的输出  $\hat{y}(t)$  如图 9 所示, 其各点的预测误差分布见图 10. 在整个数据集上, WKPLS 辨识器对状态轨迹  $x_1(t), x_2(t)$  逼近的误差  $E_{RMS}$  分别为 0.0195, 0.0161. 对系统输出  $y(t)$  逼近的误差  $E_{RMS}$  为 0.0108. 从图 8—10 不难看出, WKPLS 辨识模型完全能够重构出该混沌系统的吸引子以及系统的输出, 对其提供了很好的逼近. 用基于高斯核的 WKPLS 辨识模型, 在整个数据集上对状态轨迹  $x_1, x_2$  逼近的误差  $E_{RMS}$  分别为 0.0237, 0.0193, 对系统输出  $y(t)$  逼近的误差  $E_{RMS}$  为 0.0141. 基于 WKPLS 与 KPLS 的辨识模型均很好复现了系统的动态特性和混沌行为.

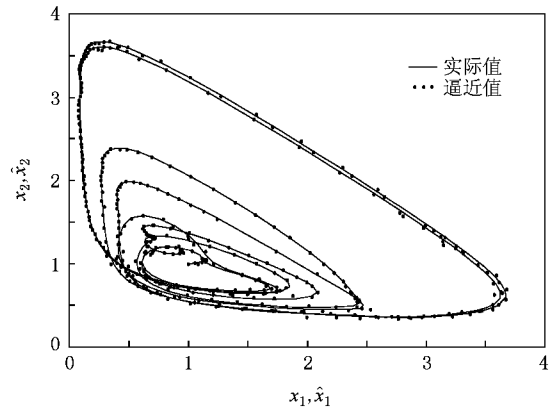


图8 糖酵解混沌系统的吸引子与 WKPLS 辨识模型的重构吸引子的相平面轨迹

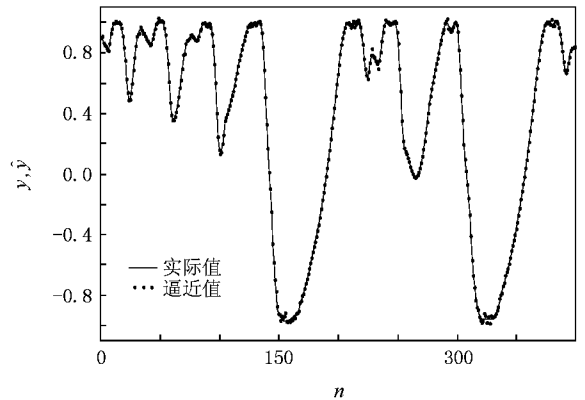


图9 糖酵解混沌系统的输出  $y(t)$  与 WKPLS 辨识模型的输出  $\hat{y}(t)$

上述混沌系统建模实例中, WKPLS 辨识器的潜在特征数目及核参数均可通过交叉验证方法确定. 而且随着潜在特征数目的增加, 其逼近精度也将逐步提高. 与神经网络和 SVM 建模方法中需要确定

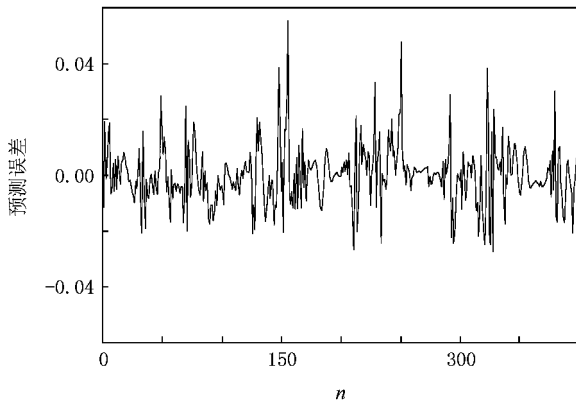


图 10 糖酵解混沌系统的实际模型状态轨线  $x_2(t)$  与 WKPLS 辨识模型  $\hat{x}_2(t)$  的预测误差分布

和求解的大量模型参数相比, WKPLS 的训练速度显然大为提高, 同时不会出现神经网络学习中由于参数随机初始化而导致算法不收敛的严重情形。

## 4. 结 论

基于核学习的强大非线性映射能力, 本文提出

了一种 WKPLS 回归建模方法, 它保持线性 PLS 优点并能够更好地展现输入与输出数据之间的非线性映射关系, 进一步提高了 KPLS 回归方法的泛化能力。在展现非线性特性的核函数选择上, WKPLS 通过构造多维张量积小波核函数, 保持了小波函数近似正交及多分辨分析的优点, 其建模能力优于基于高斯核函数的 KPLS 方法。

神经网络与 SVM 在训练中需要确定很多模型参数, 还需求解一些复杂的非线性优化或全局二次规划的问题, 而 WKPLS 方法只需选择潜在特征数目及核函数的超参数, 并利用矩阵代数求解对偶特征空间中的回归系统矩阵即可。因此, 它是用于数据挖掘领域的一种强有力的回归、分类工具, 可完成高维小样本数据的建模。

针对相当复杂的非线性动力系统, 本文验证了 WKPLS 建模方法的有效性。通过逼近一些未知的混沌动力学系统, 重构其混沌吸引子的实例展示了 WKPLS 强有力的非线性逼近能力。WKPLS 是神经网络与 SVM 建模方法的延伸, 为复杂非线性动力系统的建模、辨识与控制提供了一个有效的新途径。

- [ 1 ] Chen G, Dong X 1998 *From Chaos to Order: Methodologies, Perspectives, and Applications* (Singapore: World Scientific)
- [ 2 ] Cook P A 1986 *Nonlinear Dynamic Systems* (Englewood Cliffs: Prentice-Hall)
- [ 3 ] Haykin S 2001 *Neural Networks: A Comprehensive Foundation* (New Jersey: Prentice-Hall)
- [ 4 ] Sanchez E N, Perez J P, Chen G 2001 *Int. J. Bifur. Chaos* **11** 857
- [ 5 ] Felix R A, Sanchez E N 2004 *IEEE Trans. Neural Networks* **15** 1450
- [ 6 ] Schilling R J, Carroll J J 2001 *IEEE Trans. Neural Networks* **12** 1
- [ 7 ] Li J, Liu J H 2005 *Acta Phys. Sin.* **54** 4569 (in Chinese) [ 李军、刘君华 2005 物理学报 **54** 4569 ]
- [ 8 ] Han M, Shi Z W, Guo W 2007 *Acta Phys. Sin.* **56** 43 (in Chinese) [ 韩敏、史志伟、郭伟 2007 物理学报 **56** 43 ]
- [ 9 ] Wang L X 1995 *IEEE Trans. Aut. Contr.* **40** 11
- [ 10 ] Guan X P, Chen C L, Fan Z P 2002 *Acta Phys. Sin.* **51** 753 (in Chinese) [ 关新平、陈彩莲、范正平 2002 物理学报 **51** 753 ]
- [ 11 ] Thissen U, van Brakel R, de Weijer A P, Melssen W J, Buydens L M C 2003 *Chemom. Intell. Lab. Syst.* **69** 35
- [ 12 ] Cui W Z, Zhu C C, Bao W X, Liu J H 2004 *Acta Phys. Sin.* **53** 3303 (in Chinese) [ 崔万照、朱长纯、保文星、刘君华 2004 物理学报 **53** 3303 ]
- [ 13 ] Ye M Y, Wang X D, Zhang H 2005 *Acta Phys. Sin.* **54** 2568 (in Chinese) [ 叶美盈、汪晓东、张浩然 2005 物理学报 **54** 2568 ]
- [ 14 ] Ren R, Xu J, Zhu S H 2006 *Acta Phys. Sin.* **55** 555 (in Chinese) [ 任韧、徐进、朱世华 2006 物理学报 **55** 555 ]
- [ 15 ] Cristianini N, Shawe-Taylor J 2004 *Kernel Methods for Pattern Analysis* (Cambridge: Cambridge University Press)
- [ 16 ] Helland I S 1999 *Chemom. Intell. Lab. Syst.* **58** 97
- [ 17 ] Wold S 1984 *SIAM J. Sci. Stat. Comp.* **5** 735
- [ 18 ] Saunders C, Grobelnik M, Gunn S, Shawe-Taylor J 2006 *Subspace, Latent Structure and Feature Selection Techniques* (New York: Springer-Verlag) p34
- [ 19 ] Höskuldsson A 1988 *J. Chemometrics* **2** 211
- [ 20 ] Rosipal R, Trejo L J 2001 *J. Machine Lear. Res.* **2** 97
- [ 21 ] Rosipal R 2003 *Neural Networks Worlds* **3** 291
- [ 22 ] Zhang L, Zhou W D, Jiao L C 2004 *IEEE Trans. Syst. Man Cybern. B* **34** 1
- [ 23 ] Daubechies I 1992 *Ten Lectures on Wavelets* (Philadelphia: SIAM Press)
- [ 24 ] Scholkopf B, Smola A J 2002 *Learning with Kernels* (Cambridge: MIT Press)
- [ 25 ] Smola A, Scholkopf B, Muller K R G 1998 *Neural Networks* **11** 637
- [ 26 ] Suykens J A K, Lukas L, Vandewalle J 2000 *IEEE Trans. Circuits Syst. I* **47** 1109
- [ 27 ] Stewart I 2000 *Nature* **406** 948

# Modelling of chaotic systems using wavelet kernel partial least squares regression method<sup>\*</sup>

Li Jun<sup>†</sup> Dong Hai-Ying

( School of Automation and Electrical Engineering , Lanzhou Jiaotong University , Lanzhou 730070 , China )

( Received 13 January 2008 ; revised manuscript received 22 April 2008 )

## Abstract

Based on the powerful nonlinear mapping ability of kernel learning , and in combination with the partial least square ( PLS ) algorithm for linear regression , a wavelet kernel partial least square ( WKPLS ) regression method is proposed . By the method , the input-output data are firstly mapped to a nonlinear higher dimensional feature space , a linear PLS regression model is then constructed by the classic kernel transformation trick used in support vector machines . The PLS approach utilizes the covariance between input and output variables to extract latent features , and the wavelet kernel which is an admissible support vector kernel function is characterized by its local analysis and approximate orthogonality . Hence , the proposed WKPLS method combining PLS approach with wavelet kernel function shows excellent learning performance for modeling nonlinear dynamic systems . The WKPLS is then applied to modelling of several chaotic dynamical systems and compared with the kernel partial least squares ( KPLS ) method using Gaussian kernel function . Simulation results confirm that the WKPLS identifier is fast and can accurately approximate unknown chaotic dynamical system , and its approximation accuracy is higher than the KPLS under the same conditions .

**Keywords :** wavelet kernel , partial least squares regression , chaotic systems , modelling

**PACC :** 0545

---

<sup>\*</sup> Project supported by the Natural Science Foundation of Gansu Province , China ( Grant No. 3ZS042-B25-026 ) and the “ Qing Lan ” Talents Program of Lanzhou Jiaotong University , China .

<sup>†</sup> E-mail : lijun691201@mail.lzjtu.cn