

多元混沌时间序列的多核极端学习机建模预测

王新迎 韩敏

Multivariate chaotic time series prediction using multiple kernel extreme learning machine

Wang Xin-Ying Han Min

引用信息 Citation: *Acta Physica Sinica*, 64, 070504 (2015) DOI: 10.7498/aps.64.070504

在线阅读 View online: <http://dx.doi.org/10.7498/aps.64.070504>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2015/V64/I7>

您可能感兴趣的其他文章

Articles you may be interested in

短期风速时间序列混沌特性分析及预测

Chaotic characteristics analysis and prediction for short-term wind speed time series

物理学报.2015, 64(3): 030506 <http://dx.doi.org/10.7498/aps.64.30506>

交通流突变点的无标度特征分析

Analysis of scale-free characteristic on sharp variation point of traffic flow

物理学报.2014, 63(24): 240509 <http://dx.doi.org/10.7498/aps.63.240509>

基于压缩感知的振动数据修复方法

Vibration data recovery based on compressed sensing

物理学报.2014, 63(20): 200506 <http://dx.doi.org/10.7498/aps.63.200506>

用于混沌时间序列预测的组合核函数最小二乘支持向量机

Combination kernel function least squares support vector machine for chaotic time series prediction

物理学报.2014, 63(16): 160508 <http://dx.doi.org/10.7498/aps.63.160508>

基于能力区域的交通状态预测方法

Traffic state prediction based on competence region

物理学报.2014, 63(14): 140504 <http://dx.doi.org/10.7498/aps.63.140504>

多元混沌时间序列的多核极端学习机建模预测*

王新迎 韩敏†

(大连理工大学, 电子信息与电气工程学部, 控制科学与工程学院, 大连 116023)

(2014年7月29日收到; 2014年9月5日收到修改稿)

多元混沌时间序列广泛存在于自然、经济、社会、工业等领域. 对多元混沌时间序列进行建模预测有助于人类更好地管理、控制与决策. 针对多元混沌时间序列的建模预测问题, 本文提出一种基于多核极端学习机的预测方法. 首先对多元混沌时间序列进行相空间重构, 将多元混沌时间序列的时间相关性转化为空间相关性. 提出一种结合多核学习算法与核极端学习机模型的多核极端学习机建立相空间中输入输出数据的非线性映射. 多核极端学习机模型结合了多核学习算法的数据融合能力以及核极端学习机的训练简便优势. 基于 Lorenz 混沌时间序列预测和 San Francisco 河流月径流量预测的仿真实验表明, 与其他常见混沌时间序列预测方法相比, 本文提出的基于多核极端学习机的多元混沌时间序列预测方法具有更小的预测误差.

关键词: 混沌时间序列, 神经网络, 核方法, 预测**PACS:** 05.45.Tp, 07.05.Mh, 87.85.dq**DOI:** 10.7498/aps.64.070504

1 引言

非线性混沌时间序列预测是当今科学研究领域和工程技术领域中的一个研究热点, 并已在信号处理、自动控制等方面取得了广泛的应用. 虽根据嵌入延时理论仅需单变量时间序列重构相空间即可描述系统动态, 但实际复杂系统往往由多个变量描述, 采用多变量时间序列建模可包含更多的系统动态信息, 并可得到更高的预测精度^[1,2]. 因此, 多变量混沌时间序列预测成为目前混沌时间序列预测的一个重要研究方向^[3,4]. 在混沌时间序列预测领域, 具有强大的非线性映射能力的神经网络已成为混沌时间序列预测的一种有效方法^[5,6]. 但传统神经网络基于梯度下降的学习算法收敛速度慢, 易陷入局部最优等缺陷制约了其预测精度及更广泛应用.

另一方面, 以支持向量机 (support vector machine, SVM) 为代表的核方法在统计学习理论的基础上发展起来, 表现出较好的预测性能^[7,8]. 核极

端学习机 (extreme learning machine with kernels, KELM)^[9] 是最近几年发展起来的一种有效的机器学习方法, 其将传统极端学习机^[10] 的隐含层节点映射扩展为核方法中的核函数映射, 有效结合了极端学习机训练简便和核方法的泛化性能好的优势, 且表现出较传统支持向量机更好的学习能力. 然而, 由于不同的核函数具有的特性并不相同, 从而使得在不同的应用场合, 核函数的性能表现差别很大, 且核函数的构造或选择至今没有完善的理论依据. 此外, 当样本含有异构信息, 样本规模很大, 多维数据的不规则或数据在高维特征空间分布的不平坦, 采用单个简单核进行映射的方式对所有的样本进行处理并不合理. 针对该问题, 近年来, 出现了大量关于核组合方法的研究, 即多核学习方法^[11,12]. 胡寿松等^[13] 将多核学习方法由分类领域推广到回归领域并用以解决单元混沌时间序列预测问题, 取得了较支持向量机更好的预测效果.

基于以上分析, 本文提出一种将核极端学习机^[9] 与多核学习算法^[14] 相结合的多核极端学习机模型, 在保留多核学习的非同源数据融合能力和核

* 国家自然科学基金 (批准号: 61374154) 和国家重点基础研究发展计划 (973 计划) (批准号: 2013CB430403) 资助的课题.

† 通信作者. E-mail: minhan@dlut.edu.cn

函数选择能力的同时有效利用核极端学习机训练简便快速的优势,对多元混沌时间序列进行建模预测.

2 多元混沌时间序列相空间重构和核极端学习机

2.1 多元混沌时间序列相空间重构

在进行混沌时间序列建模和预测前,通常需要对混沌时间序列进行相空间重构.根据 Takens 嵌入延时定理^[15],单元混沌时间序列的延迟变量可以作为重构动态系统吸引子的代理变量.单元混沌时间序列相空间重构已经成为分析在自然界和人类世界广泛存在的复杂非线性系统的有利工具.然而,研究表明:通过在重构中增加多元耦合时间序列可以有效提高这些单元混沌时间序列的重构精度^[1].

给定 M 维多元混沌时间序列 $x(1), x(2), \dots, x(N)$, 其中 $x(i) = (x_1(i), x_2(i), \dots, x_M(i))$, $i = 1, 2, \dots, N$. 对该多元混沌时间序列进行相空间重构有

$$\begin{aligned} D(n) &= \{x_1(n), x_1(n - \tau_1), \dots, x_1(n - (d_1 - 1)\tau_1), \\ & \quad x_2(n), x_2(n - \tau_2), \dots, x_2(n - (d_2 - 1)\tau_2), \\ & \quad \dots \\ & \quad x_M(n), x_M(n - \tau_M), \dots, \\ & \quad x_M(n - (d_M - 1)\tau_M)\}^T, \end{aligned} \quad (1)$$

其中, $\tau_i, d_i, i = 1, \dots, M$ 是 i 维混沌时间序列的延迟时间和嵌入维数.由嵌入延时定理^[15]可知:若 d 或 d_i 充分大,则存在一个映射

$$F: \mathbb{R}^d \rightarrow \mathbb{R}^d \left(d = \sum_{i=1}^M d_i \right)$$

使得

$$D(n+1) = F(D(n)). \quad (2)$$

上式也可写为

$$\begin{aligned} x_1(n+1) &= F_1(D(n)), \\ x_2(n+1) &= F_2(D(n)), \\ &\vdots \\ x_M(n+1) &= F_M(D(n)). \end{aligned} \quad (3)$$

由(3)式,延迟时间 $\tau_i, i = 1, \dots, M$ 和嵌入维数 d_i 确定好以后,重构相空间内的多元混沌时间序列即可用于建模预测.可采用信息熵方法等方法确定延迟时间 τ 和嵌入维数 d .

2.2 核极端学习机

目前用于时间序列预测的神经网络模型,多是采用基于梯度下降的学习算法,存在收敛速度慢,易陷入局部最优等缺陷.为克服这些缺陷,提出了一种新的神经网络——极端学习机^[10],其输入权值和隐含层偏置值随机生成且在训练中保持不变,仅需采用线性回归的方法确定输出权值.进一步的,核极端学习机^[9]模型将极端学习机与核方法相结合,进一步解决了极端学习机需要确定隐含层个数的问题,且具有更好的泛化性能.

极端学习机^[10]是一种新的前馈神经网络模型,其可看成满足等式约束条件的优化问题:

$$\begin{aligned} \min L_{\text{KELM}} &= \frac{1}{2} \|\mathbf{w}\|^2 + C \frac{1}{2} \sum_{i=1}^N \xi_i^2, \\ \text{s.t. } \mathbf{h}(\mathbf{x}_i) \cdot \mathbf{w} &= t_i - \xi_i, \quad i = 1, \dots, N, \end{aligned} \quad (4)$$

其中, \mathbf{w} 为连接隐含层与输出层的输出权值, ξ_i 为样本 x_i 对应的误差值, t_i 为样本 x_i 对应的期望输出值, C 为正则化参数, N 为样本个数. $\mathbf{h}(\mathbf{x}_i) = g(\mathbf{W}_{\text{in}} \cdot \mathbf{x}_i + \mathbf{b})$ 为样本 x_i 激发的隐含层输出行向量,其中, \mathbf{W}_{in} 为连接输入层和隐含层的输入权值, \mathbf{b} 为隐含层偏置值, g 为激活函数.在极端学习机中,输入权值 \mathbf{W}_{in} 和偏置值 \mathbf{b} 随机初始化,并在训练过程中保持不变,因而隐含层输出向量 $\mathbf{h}(x)$ 仅与输入 \mathbf{x} 相关,称为隐含层节点映射.

极端学习机优化问题(4)式相应的对偶优化问题为

$$\begin{aligned} L_{\text{DELIM}} &= \frac{1}{2} \|\mathbf{w}\|^2 + \frac{C}{2} \sum_{i=1}^N \|\xi_i\|^2 \\ & \quad - \sum_{i=1}^N \alpha_i (\mathbf{h}(\mathbf{x}_i) \cdot \mathbf{w} - t_i + \xi_i), \end{aligned} \quad (5)$$

其中, α_i 为样本 x_i 对应的拉格朗日乘子.(5)式的KKT条件为

$$\begin{aligned} \frac{\partial L_{\text{DELIM}}}{\partial \mathbf{w}} = 0 &\rightarrow \mathbf{w} = \sum_{i=1}^N \alpha_i \mathbf{h}(\mathbf{x}_i)^T \\ &\rightarrow \mathbf{w} = \mathbf{H}^T \boldsymbol{\alpha}, \end{aligned} \quad (6)$$

$$\frac{\partial L_{\text{DELIM}}}{\partial \xi_i} = 0 \rightarrow \alpha_i = C \xi_i, \quad i = 1, \dots, N, \quad (7)$$

$$\frac{\partial L_{DELIM}}{\partial \alpha_i} = 0 \rightarrow \mathbf{h}(\mathbf{x}_i) \cdot \mathbf{w} - t_i + \xi_i = 0, \quad i = 1, \dots, N, \quad (8)$$

其中, $\boldsymbol{\alpha} = [\alpha_1 \dots \alpha_N]^T$, $\mathbf{H} = \begin{bmatrix} \mathbf{h}(\mathbf{x}_1) \\ \vdots \\ \mathbf{h}(\mathbf{x}_N) \end{bmatrix}$ 为极端

学习机隐含层输出矩阵. 将(6)式和(7)式代入到(8)式, 可得

$$\left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T\right) \boldsymbol{\alpha} = \mathbf{T}, \quad (9)$$

其中, \mathbf{I} 为单位矩阵, $\mathbf{T} = [t_1 \dots t_N]^T$.

将(9)式代入到(6)式, 可得输出权值 w 的计算公式如下:

$$\mathbf{w} = \mathbf{H}^T \left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T\right)^{-1} \mathbf{T}. \quad (10)$$

相应的回归函数可写为

$$f(\mathbf{x}) = \mathbf{h}(\mathbf{x}) \mathbf{H}^T \left(\mathbf{H}\mathbf{H}^T + \frac{\mathbf{I}}{C}\right)^{-1} \mathbf{T}. \quad (11)$$

考虑到支持向量机中的核函数映射 $\phi(\mathbf{x})$ 与极端学习机中的隐含层节点映射 $\mathbf{h}(\mathbf{x})$ 的相似性, 将 $\mathbf{h}(\mathbf{x})$ 扩展为未知的特征映射, 构建核极端学习机 (extreme learning machine with kernels, KELM) [9] 模型. 定义如下的核矩阵:

$$\boldsymbol{\Omega}_{KELM} = \mathbf{h}(\mathbf{x}) \cdot \mathbf{h}(\mathbf{x}') = K(\mathbf{x}, \mathbf{x}'). \quad (12)$$

相应的输出函数可以重写为

$$f(\mathbf{x}) = \begin{bmatrix} K(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ K(\mathbf{x}, \mathbf{x}_l) \end{bmatrix}^T \left(\boldsymbol{\Omega}_{KELM} + \frac{\mathbf{I}}{C}\right)^{-1} \mathbf{T}. \quad (13)$$

在该极端学习机的特殊核实现中, 隐层特征映射 $\mathbf{h}(\mathbf{x})$ 可以是未知的, 但通常给出相应的核 $K(\mathbf{u}, \mathbf{v})$ (例如 $K(\mathbf{u}, \mathbf{v}) = \exp(-\gamma \|\mathbf{u} - \mathbf{v}\|^2)$). 因此, 不再需要确定隐层节点数 (隐层特征空间的维数). 与传统支持向量机方法相比较, 核极端学习机在求解时约束更少, 因而具有更好的性能 [9].

3 多核极端学习机预测模型

核极端学习机方法结合了极端学习机训练简便快速与支持向量机方法泛化性能强的优势, 但本质上仍为单核学习方法, 其预测性能受制于核

函数的类型和核参数的选择. 同时, 考虑到多核学习获得的混合核可以使数据信息在特征空间得到更为充分的表示, 一方面提高回归预测模型的预测性, 另一方面减少了支持向量的个数. 支持向量数个数的减少不仅可以提高泛化性能, 还有利于缩短预测时间. 本文将多核学习方法与核极端学习相结合, 提出一种多核极端学习机 (multi kernel extreme learning machine, MKELM) 的模型, 并推导出相应的求解算法.

将(12)式所示的极端学习机核扩展为

$$\boldsymbol{\Omega}_{MKELM} = \sum_{k=1}^M \mu_k K_k(\mathbf{x}, \mathbf{x}'), \quad (14)$$

其中, M 为核的个数, μ_k 为核加权系数, 满足 $\sum_{k=1}^M \mu_k = 1, \forall \mu_k > 0$. (14) 式中的每个核 $K_k(\mathbf{x}, \mathbf{x}')$ 可以是不同类型的核, 如多项式核和高斯核; 也可以是同一类型的具有不同核参数的核.

采用(14)式的多核替换核极端学习机中(12)式所示的核矩阵, 即得到多核极端学习机 (multi kernel extreme learning machine, MKELM) 模型. 多核极端学习机模型需同时求解核加权系数和权值.

多核极端学习机的优化问题可写为

$$\begin{aligned} \min L_{MKELM} &= \frac{1}{2} \sum_k \frac{1}{\mu_k} \|\mathbf{w}_k\|^2 + C \frac{1}{2} \sum_{i=1}^N \xi_i^2, \\ \text{s.t.} \quad \sum_k \phi_k(\mathbf{x}_i) \mathbf{w}_k &= t_i - \xi_i, \quad i = 1, \dots, N, \\ \sum_k \frac{1}{\mu_k} &= 1, \end{aligned} \quad (15)$$

其中, \mathbf{w}_k 为定义解的核 k 对应的特征权值, ξ_i 为松弛因子, C 为平衡模型复杂度和预测性能的正则化参数.

(15) 式的对偶优化问题为

$$\begin{aligned} J(\mu) &= -\frac{1}{2} \sum_{i,j} \alpha_i \alpha_j \sum_{k=1}^M \mu_k K_k(\mathbf{x}_i, \mathbf{x}_j) \\ &\quad - \frac{1}{2C} \sum_i \alpha_i^2. \end{aligned} \quad (16)$$

可采用两步迭代优化算得求解上式: 第一步, 保持参数 μ 固定, 将 $K(\mathbf{x}_i, \mathbf{x}_j) = \sum_k \mu_k K_k(\mathbf{x}_i, \mathbf{x}_j)$ 作为一个单独核求解相应的核极端学习机优化问题. 第二步, 对参数 μ 进行更新. 对给定的 μ , $J(\mu)$ 是核极端学习机的目标值, 其中核 K 是由 μ 加权的单个核组成的. 因此, $J(\mu)$ 是凸

的,因而保证了全局收敛和对 μ 可微^[14]. $J(\mu)$ 对 μ_k 的微分可以写为

$$\frac{\partial J}{\partial \mu_k} = -\frac{1}{2} \sum_{i,j} \alpha_i \alpha_j K_k(\mathbf{x}, \mathbf{x}'). \quad (17)$$

与多核学习算法类似,对多核极端学习机算法可以总结如下:

初始化: 令 $\mu_k = 1/M, k = 1, \dots, M$

While 未达到终止条件 do

求解 $K = \sum_k \mu_k K_k$ 的核极端学习机计算 $J(\mu)$

利用 (17) 计算 $\frac{\partial J}{\partial \mu_k}, k = 1, \dots, M$ 以及相应的下降方向 D .

令 $d = \arg \max_k \mu_k, J^\dagger = 0, \mu^\dagger = \mu, D^\dagger = D$

While $J^\dagger < J(\mu)$ do

$\mu = \mu^\dagger, D = D^\dagger$

$v = \arg \min_{\{k|D_k < 0\}} -\mu_k/D_k, \gamma_{\max} = -\mu_v/D_v$

$\mu^\dagger = \mu + \gamma_{\max} D, D_d^\dagger = D_d - D_v, D_v^\dagger = 0$

求解 $K = \sum_k \mu_k^\dagger K_k$ 的核极端学习机计算 J^\dagger

End while

对 $\gamma \in [0, \gamma_{\max}]$ 在 D 上进行线搜索

$$\mu \leftarrow \mu + \gamma D$$

End while

其中, 停止条件可以是预先设置的最大迭代次数, 或者 μ 的变化小于某个阈值.

4 仿真实验

为验证所提算法的有效性, 将其应用于 Lorenz 混沌时间序列预测、San Francisco 河流月径流量时间序列预测仿真实验中, 并引入均方根误差 (root mean square error, RMSE) 来定量说明预测性能

$$\text{RMSE} = \left(\frac{1}{N-1} \sum_{t=1}^N [p_i - d_i]^2 \right)^{1/2},$$

其中, N 为测试样本个数, d_i 是某个实际观测值, p_i 为对应预测输出.

为说明本文所提多核极端学习机 (MKELM) 方法的有效性, 与极端学习机 (ELM)^[10], 在线极端学习机 (OSELM)^[16], 支持向量回归 (SVR), 多核学习算法 (MKL^[14]) 和核极端学习机 (KELM)^[9] 等方法得到的结果进行比较. 每组数据仿真 50 次, 取仿真结果的平均值.

4.1 Lorenz 混沌时间序列预测

Lorenz 混沌方程是一组三元常微分方程, 如 (18) 式所示. 当取 $a = 10, b = 8/3, c = 28, x(0) = y(0) = z(0) = 1.0$ 时, 系统产生混沌 (见图 1).

$$\begin{aligned} \dot{x} &= a(y - x), \\ \dot{y} &= (c - z)x - y, \\ \dot{z} &= xy - bz. \end{aligned} \quad (18)$$

利用四阶 Runge-Kutta 法迭代产生混沌时间序列, 使用 $x(t), y(t)$ 和 $z(t)$ 序列共同预测 $x(t+\eta)$. 取延迟时间和嵌入维数分别为 $\tau_1 = 8, \tau_2 = 7, \tau_3 = 8, m_1 = 6, m_2 = 6, m_3 = 6$, 得到相应的 18 维延迟坐标向量为

$$d(t) = \{x(t), \dots, x(t - 5\tau_1), y(t), \dots, y(t - 5\tau_2), z(t), \dots, z(t - 5\tau_3)\}.$$

期望输出为 $x(t+1)$. 相空间重构产生 2000 组数据, 其中前 1500 组作为训练样本, 后 500 组作为测试样本.

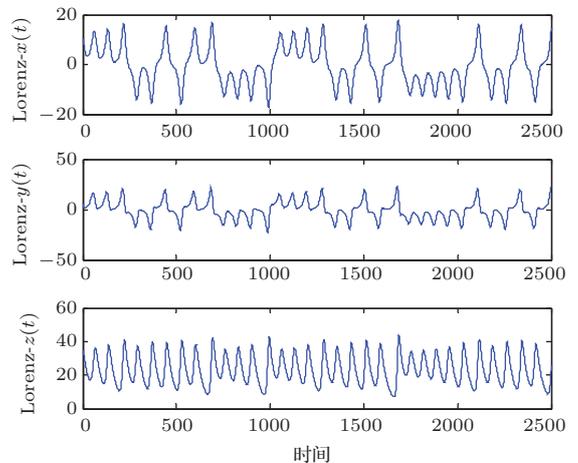


图1 Lorenz 混沌时间序列 $x(t), y(t), z(t)$

采用 10 折交叉检验的方式确定预测模型的超参数. 极端学习机和在线极端学习机的隐含层节点个数设置为 5 到 200 之间间隔为 5 进行变化. 选择 50 个样本点初始化在线极端学习机, 每 5 个样本点进行权值更新. 支持向量回归和核极端学习机的核函数均选为高斯核函数, 正则化参数 $C = [2^{-5}, 2^{-4}, \dots, 2^4, 2^5]$ 和核宽参数 $\gamma = [2^{-5}, 2^{-4}, \dots, 2^4, 2^5]$ 采用网格法确定最优参数对.

多核学习算法的核函数选择线性核, 多项式核 ($K(\mathbf{u}, \mathbf{v}) = (\mathbf{u} \cdot \mathbf{v} + 1)^d, d = 1, 2, 3$) 和高斯核 (核宽 $\gamma = [0.01:0.05:0.2 \ 0.5 \ 1 \ 2 \ 5 \ 7 \ 10 \ 12 \ 15 \ 17 \ 20]$), 正则化参数 $C = [2^{-5}, 2^{-4}, \dots, 2^4, 2^5]$ 进行寻优. 多核极端学习机的核函数同样包括高斯核函数, 线性核函数以及多项式核函数, 正则化参数 $C = [2^{-5}, 2^{-4}, \dots, 2^4, 2^5]$ 进行寻优. 表 1 列出了几种方法的仿真预测结果比较, 表 2 给出了几种方法在多元数据上的训练时间的比较.

表 1 Lorenz 混沌时间序列预测仿真结果 (RMSE)

| 选择的变量 | ELM | OSELM | SVR | MKL | KELM | MKELM |
|--------|--------|--------|--------|--------|--------|--------|
| [1—6] | 0.0582 | 0.4278 | 0.9950 | 0.1726 | 0.1019 | 0.0973 |
| [1—18] | 0.0549 | 0.1551 | 0.8848 | 0.1056 | 0.0652 | 0.0312 |

表 2 Lorenz 混沌时间序列预测仿真时间

| 运算时间/s | ELM | OSELM | SVR | MKL | KELM | MKELM |
|--------|---------|---------|---------|-------|----------|----------|
| 训练 | 36.8826 | 44.3053 | 33.8020 | 17554 | 263.6217 | 947.2922 |

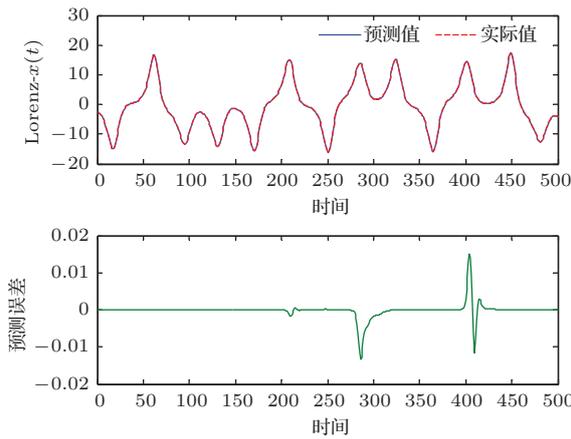


图 2 (网刊彩色) Lorenz 混沌时间序列 $x(t)$ 预测结果

从表 1 中可以看出, 表中所列几种方法使用 x, y, z 三元序列进行相空间重构数据的预测结果要好于仅使用单元 x 序列进行相空间重构数据的预测结果. 采用核方法的支持向量回归, 核极端学习机, 多核学习算法和本文所提多核极端学习机的预测效果要好于极端学习机和在线极端学习机的预测效果. 本文所提的多核极端学习机算法的预测结果要好于核极端学习机和多核学习算法的预测结果, 证明所提方法的有效性. 此外, 表 2 中可以看出, 本文所提算法的训练时间较传统的多核学习算法有很大优势; 与核极端学习机算法相比, 本文所提算法的训练时间有所增加, 但考虑到核极端学习机算法仅对单个高斯核的核宽和正则化系数进

行寻优, 而本文所提算法则可采用线性核、多项式核、高斯核等的组合, 具有更好的非线性表达能力.

图 2 给出了基于多核极端学习机的预测曲线和相应的预测误差曲线, 从图中可以看出, 预测曲线可以较好的拟合实际曲线, 误差曲线也在 0 附近的较小范围内波动.

4.2 San Francisco 河流月径流量预测

为进一步验证方法的有效性, 将所提预测方法应用于一组实际的多元混沌时间序列仿真实验中. 根据 Hipel 和 McLeod 等 [17] 的研究成果可知, 河流月径流量与月降雨量和月平均温度具有相关性. 本文采用 1932 年 1 月至 1966 年 12 月的月平均气温、月降雨量、San Francisco 河流月径流量三组时间序列 (见图 3) 进行相空间重构, 共同预测下一月的 San Francisco 河流月径流量. 三组序列的重构参数均设为: 嵌入维数 12, 延迟时间 1. 重构生成 408 组数据, 前 300 组数据作为训练样本, 后 108 组数据作为测试样本. 几种方法的参数设置与 4.1 节 Lorenz 混沌时间序列预测的参数设置相同. 仿真结果如表 3 所示, 在多元数据上的训练时间比较如表 4 所示.

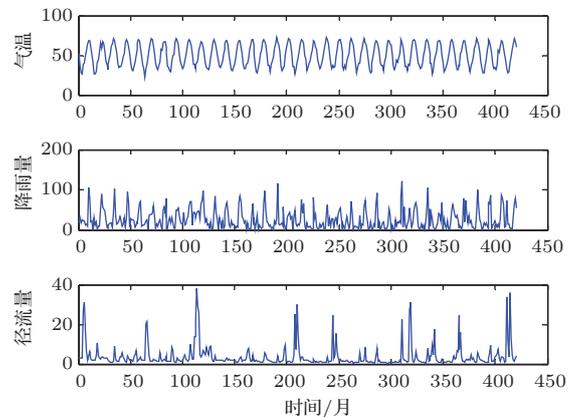


图 3 月平均气温、降雨量和河流径流量时间序列

表 3 San Francisco 河流月径流量预测仿真结果 (RMSE)

| 选择的变量 | ELM | OSELM | SVR | MKL | KELM | MKELM |
|--------|--------|--------|--------|--------|--------|--------|
| [1—12] | 8.8296 | 8.7790 | 8.3847 | 8.5274 | 8.4868 | 8.4305 |
| [1—36] | 8.2137 | 7.8840 | 7.3423 | 7.9665 | 7.3353 | 7.3657 |

表 4 San Francisco 河流月径流量预测仿真时间

| 运行时间/s | ELM | OSELM | SVR | MKL | KELM | MKELM |
|--------|---------|---------|--------|----------|---------|---------|
| 训练 | 11.1942 | 29.8158 | 4.0715 | 863.4688 | 24.7276 | 48.4627 |

表3中列出了仅使用San Francisco河流月径流量进行相空间重构与使用月平均气温、月降雨量、San Francisco河流月径流量三元时间序列进行相空间重构的仿真结果比较. 可以看出, 在比较的6种方法中, 使用三元时间序列进行相空间重构的结果均要较仅使用月径流量单元时间序列的预测结果要好, 说明月平均气温和月降雨量与河流月径流量具有相关性, 采用这三元时间序列进行预测可明显提高预测精度. 从表3中还可以看出, 基于核的预测方法如支持向量回归, 核极端学习机, 多核学习算法, 多核极端学习机在实际数据中的预测结果也要好于极端学习机和在线极端学习机预测方法. 本文所提多核极端学习机方法, 预测精度与核极端学习机和多核学习算法相当. 计算仿真50次, 多核极端学习机和核极端学习机的标准差, 可以看到多核极端学习机的标准差: 0.0398(单元)和0.0862(三元)比核极端学习机的标准差: 0.0732(单元)和0.1466(三元)要更小. 以上分析说明本文所提多核极端学习机方法的有效性和稳定性. 在表4所示的训练时间方面, 本文所提算法较多核学习算法有了较大提升; 与其他几种方法相比, 训练时间未表现出优势, 但本文所提算法可采用不同的核函数实现更丰富的非线性映射, 且具有更好的稳定性.

图4给出了基于本文所提多核极端学习机多元时间序列预测算法对San Fransico河流径流量进行预测的结果. 从图中可以看出, 预测曲线能较好的跟踪拟合实际曲线, 且误差曲线在0附近的小范围波动, 说明本文所提方法能对San Fransico河流径流量进行有效建模预测.

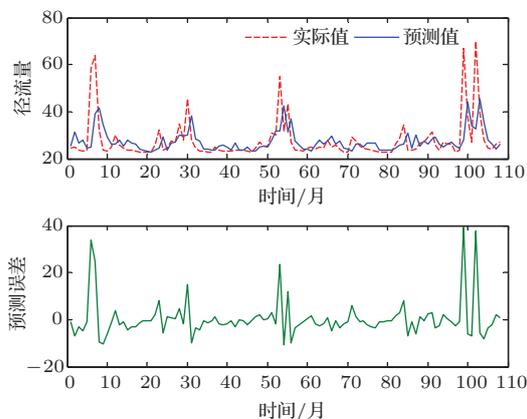


图4 (网刊彩色) San Fransico 河流径流量时间序列预测结果

5 结 论

多元混沌时间序列广泛存在于自然、经济、社会与工业等领域, 其包含较单元时间序列更多的动态信息. 针对多元混沌时间序列预测问题, 本文提出一种基于多核极端学习机的预测方法. 采用多元相空间重构技术将多元混沌时间序列的时间关系转化为空间关系. 提出一种多核极端学习机算法, 将多核学习算法与核极端学习机相结合, 有效结合了多核学习算法的数据融合能力和核极端学习机的高精度建模能力. 基于人工和实际多元混沌时间序列数据的仿真实验证明所提方法的有效性.

参考文献

- [1] Deyle E, Sugihara G 2011 *PLoS ONE* **6** e18295
- [2] De Gooijer J G, Hyndman R J 2006 *Int. J. Forecasting* **22** 443
- [3] Frenzel S, Pompe B 2007 *Phys. Rev. Lett.* **99** 204101
- [4] Wang X Y, Han M 2012 *Acta Phys. Sin.* **61** 80507 (in Chinese) [王新迎, 韩敏 2012 物理学报 **61** 80507]
- [5] Grigorievskiy A, Miche Y, Ventela A M, Severin E, Lendasse A 2014 *Neural Netw.* **51** 50
- [6] Butcher J B, Verstraeten D, Schrauwen B, Day C R, Haycock P W 2013 *Neural Netw.* **38** 76
- [7] Tang Z J, Ren F, Peng T, Wang W B 2014 *Acta Phys. Sin.* **63** 50505 (in Chinese) [唐舟进, 任峰, 彭涛, 王文博 2014 物理学报 **63** 50505]
- [8] Sapankevych N, Sankar R 2009 *IEEE Comput. Intell. M.* **4** 24
- [9] Huang G B, Zhou H M, Ding X J, Zhang R 2012 *IEEE Trans. Syst. Man Cybern. B Cybern.* **42** 513
- [10] Huang G B, Zhu Q Y, Siew C K 2006 *Neurocomputing* **70** 489
- [11] Wang H Q, Sun F C, Cai Y N, Chen N, Ding L G 2010 *Acta Autom. Sin.* **36** 1037 (in Chinese) [汪洪桥, 孙富春, 蔡艳宁, 陈宁, 丁林阁 2010 自动化学报 **36** 1037]
- [12] Gönen M, Alpaydin E 2011 *J. Mach. Learn. Res.* **12** 2211
- [13] Zhang J F, Hu S S 2008 *Acta Phys. Sin.* **57** 2708 (in Chinese) [张军峰, 胡寿松 2008 物理学报 **57** 2708]
- [14] Rakotomamonjy A, Bach F, Canu S, Grandvalet Y 2008 *J. Mach. Learn. Res.* **9** 2491
- [15] Takens F 1981 *Dynam. Syst. Turbul.* **898** 366
- [16] Liang N Y, Huang G B, Saratchandran P, Sundararajan N 2006 *IEEE Trans. Neural Netw.* **17** 1411
- [17] Hipel K W, McLeod A I 1994 *Time series modelling of water resources and environmental systems.* (Amsterdam: Elsevier) p553

Multivariate chaotic time series prediction using multiple kernel extreme learning machine*

Wang Xin-Ying Han Min[†]

(Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116023, China)

(Received 29 July 2014; revised manuscript received 5 September 2014)

Abstract

Multivariate chaotic time series is widely present in nature, such as in economy, society, industry and other fields. Modeling and predicting multivariate time series will help human to better manage, control, and make decision. A prediction method based on multiple kernel extreme learning machine is proposed in this paper to model the complex dynamics of multivariate chaotic time series. First, the multivariate chaotic time series is reconstructed in phase space, transforming the temporal correlation into spatial correlation. Then, a prediction model-multiple kernel extreme learning machine, which combines the multiple kernel learning and extreme learning machine with kernels, is proposed to approximate the nonlinear function of the input – output data in phase space. The proposed multiple kernel extreme learning machine could effectively combine the simple training of extreme learning machine with kernels and the data fusion capabilities of multiple kernel learning. Simulation results based on Lorenz chaotic time series prediction and San Francisco monthly runoff prediction demonstrate that, compared with other state-of-art chaotic time series prediction methods, the proposed multiple kernel extreme learning machine could get a better prediction accuracy.

Keywords: chaotic time series, neural networks, kernel methods, prediction

PACS: 05.45.Tp, 07.05.Mh, 87.85.dq

DOI: [10.7498/aps.64.070504](https://doi.org/10.7498/aps.64.070504)

* Project supported by the National Natural Science Foundation of China (Grant No. 61374154), and the State Key Development Program for Basic Research of China (973 Program) (Grant No. 2013CB430403).

[†] Corresponding author. E-mail: minhan@dlut.edu.cn