文章编号: 1000−8152(2007)05−0719−06

主成分分析与神经网络的结合在多变量序列预测中的应用

席剑辉[1, 2], 韩 敏[2]

(1. 沈阳航空工业学院 自动控制系, 辽宁 沈阳 110034; 2. 大连理工大学 电子与信息工程学院, 辽宁 大连 116023)

摘要: 目前预测方法的研究主要集中在单变量时间序列上, 本文建立起一种针对多元变量非线性时间序列建模和预测的方法框架. 首先, 同时考虑序列状态间的线性相关性和非线性相关性, 建立初始延迟窗以包含充分的预测信息; 然后, 利用主成分分析(PCA)方法寻找不同变量在数据空间中的最大方差方向, 扩展PCA应用于提取多个变量的综合信息, 重构多元变量输入状态相空间; 最后, 利用神经网络逼近不同变量之间以及当前状态和将来状态之间的函数映射关系, 实现多元变量预测. 对Rössler混沌方程和大连降雨、气温序列的预测仿真说明了本文方法的有效性, 为多元变量时间序列分析提供了一条新的途径.

关键词: 多元变量时间序列; 神经网络; 预测; 主成分分析

中图分类号: TP13 **文献标识码**: A

Prediction of multivariate time series based on principal component analysis and neural networks

XI Jian-hui[1, 2], HAN Min[2]

(1. Department of Automation, Shenyang Institute of Aeronautical Engineering, Shenyang Liaoning 110034, China;
2. School of Electronic and Information Engineering, Dalian University of Technology, Dalian Liaoning 116023, China)

Abstract: Most of previously published prediction methods are concentrated on the modeling of univariate time series. The main purpose of this paper is to study a new methodology to model and predict multivariate nonlinear time series. Firstly, both the linear correlations and the nonlinear correlations are detected to initialize an embedding delay window, which could contain enough information for prediction. Then, the principal components analysis (PCA) method is expanded to extract the joint information of multiple variables in a complex system since PCA could find the uncorrelated directions of maximum variance in the data space of different variables. The multivariate phase space is reconstructed. Furthermore, neural network makes predictions on the basis of approximating both the functional relation between different variables and the map between current state and future state. Finally, two simulation examples, one is from the typical Rössler equation and the other is from the practically observed values of rainfall and temperature of Dalian, are used to explain the validity of the proposed method. It provides a new way to analyze the multivariate time series.

Key words: multivariate time series; neural network; prediction; principal component analysis

1 引言(Introduction)

复杂系统是由大量相互作用或相互分离的子系统结合在一起, 具有非线性、混沌或不确定动态行为的系统, 在实际问题中广泛存在, 例如经济管理系统、生理系统、水文系统等. 因此如何从获取的时间序列去刻画原复杂系统的动态特征成为复杂系统研究中的一项重要工作. 但是, 目前针对多元变量时间序列模型描述和预测的研究还十分缺乏, 已有研究大部分集中在单变量时间序列上. 其基本思路是首先基于观测数据, 重构系统状态空间, 即确定预测模型输入向量; 然后采用适当的建模方法构造系统模型[1,2]. 实际上, 该思路同样适用于多变量时间序列的研究. 一旦在第1步中构造出多元变量输入向量, 第2步的预测过程, 多变量和单变量的情况极为类似, 都是利用预测模型对系统输入输出关系的逼近.

第1步有效构造多元变量输入向量, 使其能够包含多个变量的综合信息, 将有利于提高预测精度. Cao等人[3]首先计算每个变量的嵌入延迟参数, 然后选取不同的嵌入维数组合建立输入向量, 根据预测精度选择最佳的输入状态组合. 王海燕等人[4]则扩展了上述方法, 引入广义关联积分和广义关联维数

收稿日期: 2005−05−08; 收修改稿日期: 2006−10−27.
基金项目: 国家自然科学基金资助项目(60374064).

的计算公式, 在不同的广义关联维数下计算虚假邻近点的个数, 最少时对应最佳输入状态组合. 这些方法都需要对每个变量求取嵌入延迟和嵌入维数, 在大量的组合方式中通过比较做出选择, 计算量较大. 而且嵌入延迟、嵌入维数等参数的选取较具经验性, 目前还没有一个统一的方法, 尤其针对多元变量系统, 计算较为困难. 主成分分析(PCA)[5,6]是利用统计原理建立系统低维模型的方法, 可以获得具有较大方差的新输入向量, 并且减少输入维数. 扩展PCA方法应用于多元变量输入状态相空间重构, 可以在保留输入向量主要信息的同时, 避免嵌入维数等参数的计算问题, 简化输入状态选取过程.

第2步构造系统多个变量之间以及变量状态之间的函数关系模型. 由于系统的复杂性和非线性, 限制了常规方法的应用. 近年来, 神经网络模型辨识方法[2,7]因为其具有良好的函数逼近能力而愈来愈受到重视. 而神经网络所具有的多输入多输出处理能力使其可以良好地用于多元变量时间序列模型描述.

本文沿用单变量时间序列建模的思想, 结合主成分分析和神经网络方法进行多元变量建模和预测. 为使输入变量中包含充分的预测信息, 在计算状态之间线性相关性的同时, 引入一种非线性函数检测状态间的非线性相关性, 合理确定初始的嵌入延迟窗. 然后利用PCA构造有效输入向量, 利用神经网络建立系统输入输出模型.

2 多元变量输入状态空间重构(Input state space reconstruction for multivariate time series)

令$\boldsymbol{x}(t)=[x_1(t),\cdots,x_n(t)]^{\mathrm{T}}, t=1,2,\cdots,N$表示要研究的复杂系统多元变量离散时间序列, $x_i(t)$对应系统第i个变量在t时刻的值. 针对每个变量, 选择适当的嵌入延时τ_i, 嵌入维数m_i, $i=1,2,\cdots,n$,就可得到多元变量嵌入延迟向量:

$$\begin{aligned}\overrightarrow{X}(t)=[&x_1(t),x_1(t-\tau_1),\cdots,x_1(t-(m_1-1)\tau_1)\\ &\cdots\\ &x_i(t),x_i(t-\tau_i),\cdots,x_i(t-(m_i-1)\tau_i)\\ &\cdots\\ &x_n(t),x_n(t-\tau_n),\cdots,x_n(t-(m_n-1)\tau_n)]^{\mathrm{T}},\end{aligned} \tag{1}$$

其中τ_i和m_i的选择没有统一的评价标准存在, 计算较为困难. Rosenstein[8]等人认为建立输入状态空间时, 确定嵌入延迟窗:

$$T_{im}=\tau_i(m_i-1), \tag{2}$$

比单独直接求取m_i或τ_i更为适合. 所以确定式(1)的基本步骤为:

Step i 对每个变量序列, 根据观测点元素之间的相关性适当选择$T_{im}, i=1,\cdots,n$;

Step ii 采用平均位移、互信息等方法计算τ_i, 因为预测时较近的历史状态对预测的影响较大, 通常选取较小的τ_i;

Step iii 利用式(2)计算m_i;

Step iv 将m_i和τ_i代入式(1), 构造多元变量嵌入延迟向量.

其中, 如果嵌入延迟向量元素之间的关系被噪声覆盖, 说明T_{im}过小; 如果嵌入延迟向量第1个元素和最后一个元素之间几乎没有相关性, 说明T_{im}过大. 所以, T_{im}需要适当选取, 以包含足够的预测信息. 常用的自相关函数只能提取序列间的线性相关性, 为了能同时提取状态之间的非线性相关性[9], 本文采用以下方法确定T_{im}初值. 计算线性相关函数:

$$\phi_{xx}(T_{im})=\mathrm{E}\{[x_i(t)-\overline{x}_i(t)][x_i(t-T_{im})-\overline{x}_i(t)]\}, \tag{3}$$

检测状态之间的线性相关性. 同时计算非线性相关函数:

$$\phi_{x^2x^2}(T_{im})=\mathrm{E}\{[x_i^2(t)-\overline{x}_i^2(t)][x_i^2(t-T_{im})-\overline{x}_i^2(t)]\}, \tag{4}$$

检测状态之间的非线性相关性. 设T_{im}^{x}和$T_{im}^{x^2}$分别对应$\phi_{xx}(T_{im})$和$\phi_{x^2x^2}(T_{im})$ 的第1个极大值, 则

$$T_{im}=\max(T_{im}^{x},T_{im}^{x^2})-1. \tag{5}$$

这样计算出的延迟窗T_{im}, 其中的点能够充分反映序列的线性和非线性关系. 因为预测时, 通常τ_i选值较小, 则导致m_i较大. 式(1)表现出高维性质, 提高了非线性建模的难度. 为了解决此问题, 引入PCA方法提取不同元素的有效信息.

设第i个变量的输入样本矩阵为$\boldsymbol{X}_{l\times h}$, 其中$h=m_i, l=N-\max\limits_i[(m_i-1)\tau_i], i=1,2,\cdots,n$,标准化样本矩阵为$\tilde{\boldsymbol{X}}_{l\times h}$, 对$\tilde{\boldsymbol{X}}_{l\times h}$进行奇异值分解

$$\tilde{\boldsymbol{X}}=\boldsymbol{U}\Sigma\boldsymbol{V}^{\mathrm{T}}, \tag{6}$$

其中$\boldsymbol{U}\in F^{l\times l}, \boldsymbol{V}\in F^{h\times h}$都为正交阵,

$$\Sigma=\mathrm{diag}[s_1,s_2,\cdots,s_r,0,\cdots,0], \tag{7}$$

$\Sigma\in F^{l\times h}, s_1\geqslant s_2\geqslant\cdots\geqslant s_r$为对应的$r$个奇异值, 按降序排列. 因为奇异值的平方对应为$\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^{\mathrm{T}}$的特征值, 反映的是不同变量的方差大小. 方差越大, 所代表的系统信息就越有效. 显然后$h-r$维对应方差为零的变量可以舍弃, 从而将原h维输入向量转化为r维向量, 同时又不会损失原系统的物理信息.

但是, 实际系统中由于噪声和其他干扰量的存在, 可能不会存在严格为零的奇异值. 一般做法是舍弃较小的奇异值, 对应输入样本矩阵中的噪声成分; 保留较大的奇异值, 对应输入样本矩阵中的主元成分. Valle等人[10]比较了多种方法判断需舍弃的奇异值成分, 这里采用一种简单直接的方法. 计算每个主元的方差贡献率为

$$\eta_j = \frac{s_j^2}{\sum\limits_{k=1}^{r} s_k^2}, j = 1, 2, \cdots, r, \tag{8}$$

保留前p个主元, 当满足

$$\sum_{j=1}^{p} \eta_j > \eta_0, \tag{9}$$

η_0为常数, 一般取$0.8 < \eta_0 < 1$. 如果正交阵

$$V_{h\times h} = [v_1\, v_2 \cdots v_h], \tag{10}$$

可取

$$\hat{V}_{h\times p} = [v_1\, v_2 \cdots v_p], \tag{11}$$

则得到主元矩阵为

$$Z \approx \tilde{X}\hat{V}, \tag{12}$$

$Z \in F^{l\times p}$, 变量的输入降为p维, 而且输入元素间的协方差近似为零, 可以认为近似独立. 吴春国等人[6]证明奇异值分解满足正交分解最优性, 与PCA具有等价性, 为上述过程提供了理论基础.

3　多元变量时间序列预测模型(Multivariate time series prediction model)

本文沿用单变量时间序列预测的思想, 首先采用PCA方法构造多元变量输入向量; 然后选择非线性建模方法对系统输入输出进行匹配, 即建立模型$F_\eta(\cdot)$实现

$$\overrightarrow{Y}(t+\eta) = F_\eta(\overrightarrow{X}(t)), \tag{13}$$

其中$\overrightarrow{Y}(t+\eta) = [y_1(t+\eta) \cdots y_n(t+\eta)]^{\mathrm{T}}$, $y_i(t+\eta)$为模型对未来$t+\eta$时刻第i个变量状态$x_i(t+\eta)$的预测值, η为预测步数, $i = 1, \cdots, n$. 结合PCA和神经网络(NN)建立预测模型的基本结构如图1所示. 其基本步骤为:

Step 1　建立式(1)所示的多元变量嵌入延迟向量, 如节2中的Step i至Step iv;

Step 2　提取式(1)的主元信息, 对每一个变量利用PCA方法降维, 如式(6)到式(12)所示, 最终建立新的多元变量嵌入延迟向量;

Step 3　采用神经网络方法根据当前输入预测未来时刻的状态.

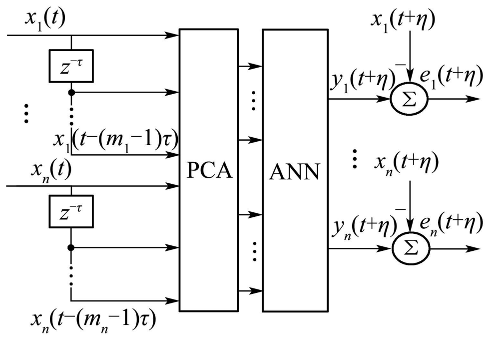

图 1　多元变量预测模型基本结构

Fig. 1　The structure of basic multivariate prediction model

神经网络应用于时间序列预测, 主要包含两种类型: 1) 采用静态神经网络, 引入延迟输入向量表现系统的动态特性, 其作用类似一个非线性自回归模型; 2) 采用递归神经网络, 网络带有反馈分支, 能够自动反映系统动态特性. 本文以前馈网络为例说明多元变量时间序列预测的过程. 网络输入层和中间层节点函数为正切Sigmoid函数, 输出层采用线性输出函数. 网络训练算法为带动量因子的自适应学习率反向传播(BP)算法[11]. 其权值调整过程为

$$w_{ij}^k(t) = w_{ij}^k(t-1) + \Delta w_{ij}^k(t), \tag{14}$$

$$\Delta w_{ij}^k(t) = -\gamma\frac{\partial E}{\partial w_{ij}^k(t-1)} + \lambda\Delta w_{ij}^k(t-1). \tag{15}$$

其中: $w_{ij}^k(t)$是k层的第i个节点指向$k+1$层第j个节点的权值, $\Delta w_{ij}^k(t)$为权值调整量, γ为自适应调整的学习率, λ是动量因子, $0 < \lambda < 1$. E为网络评价函数, 训练网络的目的就是通过调整网络权值使E达到最小. 令l表示网络训练样本的个数, n为预测变量个数, 则

$$E = \frac{1}{2}\sum_{t=1}^{l}\sum_{j=1}^{n}(y_j(t) - \hat{y}_j(t))^2, \tag{16}$$

其中: $y_j(t)$为第j个变量的理想输出, $\hat{y}_j(t)$ 为网络输出. E足够小并趋于稳定时可结束训练.

4　实例仿真(Simulations cases)

为定量说明网络预测性能的好坏, 引入评价指标均方根误差E_{RMSE}和预测精度E_{PA}[12], 从不同侧面反映预测的效果. 设$y(t)$对应实际观测值, $\hat{y}(t)$为网络预测值, y_m和σ_y分别对应$y(t)$的均值和标准方差, $\hat{y}_m$和$\sigma_{\hat{y}}$分别为$\hat{y}(t)$的均值和标准方差, 有

$$E_{\mathrm{RMSE}} = (\frac{1}{l-1}\sum_{t=1}^{l}[\hat{y}(t) - y(t)]^2)^{1/2}, \tag{17}$$

$$E_{\mathrm{PA}} = \frac{\sum\limits_{t=1}^{l}[(\hat{y}(t) - \hat{y}_m)(y(t) - y_m)]}{(l-1)\sigma_{\hat{y}}\sigma_y}, \tag{18}$$

E_{RMSE}反映了预测值对观测值的平均偏离程度. 取值大于或等于零, 预测无误差时等于零. E_{PA}反映了预测值和观测值在其均值附近的偏离之间的相关性, 取值在+1和−1之间. 预测无误差时为1.

将本文方法应用于两个二变量时间序列仿真实例, 分别来自Rössler混沌方程产生的理想混沌序列和大连降雨、气温实际观测序列. 预测模型结构如图1所示. 神经网络为一4层前馈神经网络, 包括输入层、输出层和两层隐层. 分述如下.

4.1 Rössler混沌方程$x(t)$和$y(t)$时间序列预测(Predicting $x(t)$ and $y(t)$ time series of Rössler equation)

Rössler混沌方程如式(19)所示, $a = 0.398, b = 2.0, c = 4.0, x(0) = y(0) = 1.0, z(0) = 31.0$, 此时方程表现出混沌特性.

$$\begin{cases} \dot{x} = -y - z, \\ \dot{y} = -x + ay, \\ \dot{z} = b + z(x - c). \end{cases} \tag{19}$$

取其中的$x(t)$和$y(t)$时间序列作为仿真对象.

采用本文方法对$x(t)$和$y(t)$二变量时间序列进行建模和预测, 计算序列的线性相关性和非线性相关性如式(3)和式(4)所示, 得到$m_1 = m_2 = 25$, 选择$\tau_1 = \tau_2 = 1$, 即初始输入样本向量为50维:

$$\overrightarrow{X}(t) = [x(t), \cdots, x(t-24), y(t), \cdots, y(t-24)]^{\mathrm{T}},$$

η步预测时, 网络输出为

$$\overrightarrow{Y}(t) = [x(t+\eta), y(t+\eta)]^{\mathrm{T}},$$

利用4阶Runge-Kutta法产生1000个输入样本.

首先, 分别对输入样本矩阵X_1^{T}和X_2^{T}进行PCA, 计算各个奇异值的方差贡献率, 并从大到小降序排列, 列出前5个如表1所示. 如果令$\eta_0 = 0.96$, 可以看出$x(t)$序列前4个主元的方差贡献和已经大于$\eta_0, p_1 = 4$; 同理$y(t)$序列$p_2 = 3$. 所以新的输入向量维数可以降低到$p = p_1 + p_2 = 7$维.

表 1 Rössler方程$x(t)$和$y(t)$序列奇异值及方差贡献

Table 1 Eigenvalues and their energy contribution of $x(t)$ and $y(t)$ of Rössler equation

序号j		1	2	3	4	5
$X(t)$	s_j	50.127	45.101	10.877	4.586	1.984
$X(t)$	η_j	0.4388	0.3948	0.0952	0.0401	0.0085
$Y(t)$	s_j	54.938	45.395	10.266	2.516	0.934
$Y(t)$	η_j	0.4796	0.3963	0.0896	0.0220	0.0082

根据文献[13], 选择神经网络参数如表2所示, 利用降维后的样本训练网络, 取前500个样本为训练样本, 后500个样本为校验样本, 初始权值随机选取. 一步预测的网络学习误差和校验误差曲线如图2所示, 具有较快的收敛性. 学习样本和校验样本的E值都可达到一个较小的值, 保证了网络的学习精度和泛化精度. 图3给出了1000个样本的预测值和实际值的比较, 上图对应$x(t)$序列, 下图对应$y(t)$序列. E_p为网络输出值和实际值之间的差值. 可以看出, 所建模型的预测精度较高, 预测值和观测值相近, 符合同一发展规律. 定量的性能指标计算见表3, 表征预测值对观测值偏离程度的E_{RMSE}值接近于0, 表征预测值与观测值相关程度的E_{PA}值接近于1, 都说明了预测模型的高精度.

为进行结果比较, 分别采用文献[3]和单变量建模方法对上述序列仿真. 其中, 文献[3]中采用互信息法分别确定序列的延迟时间和嵌入维数, 建立起一个10维的输入向量; 单变量预测模型仍然采用PCA和前馈网络结合, 只是输入只由相应的变量状态及其延时组成, 所涉及的计算参数选取同前. 分别取$\eta = 1$和$\eta = 20$, 计算结果见表3. 可以看出对于本例中不含噪声的理想数据, 两种多变量预测模型取得较好的结果, 但PCA确定输入相空间时计算更为方便快捷, 所建模型更为简单; 而单变量预测模型的性能随预测时间下降较快, 一方面是混沌特性的影响, 另一方面也是单变量序列提供给预测的信息缺乏导致.

表 2 4层神经网络仿真条件

Table 2 Simulation conditions of four-layer neural network

网络初始权值	随机确定的小数值
输入层节点个数	p
第1隐层节点个数	$2p$
第2隐层节点个数	p
学习速率γ初值	0.1
动量因子λ	0.7

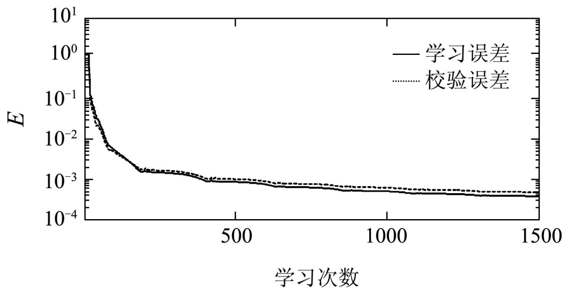

图 2 Rössler方程$x(t)$和$y(t)$建模学习和校验误差曲线

Fig. 2 Learning and testing curves of modeling $x(t)$ and $y(t)$ of Rössler equation

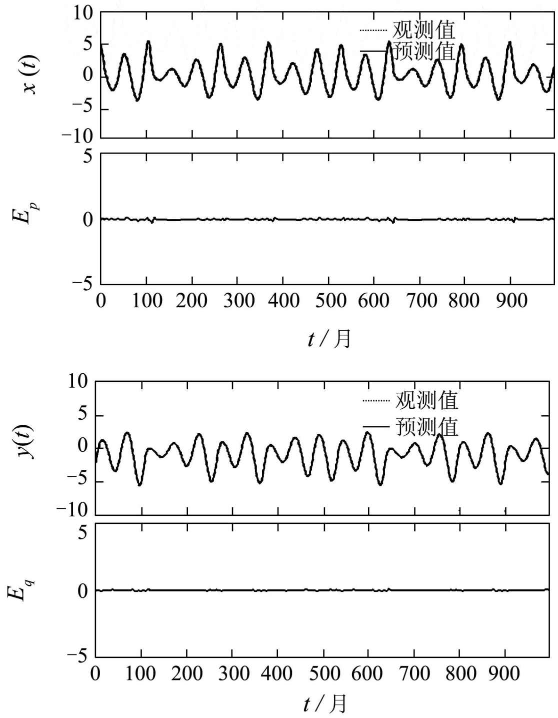

图 3　Rössler方程$x(t)$和$y(t)$序列一步预测的仿真结果

Fig. 3　One-step prediction simulation results of $x(t)$ and $y(t)$ of Rössler equation

表 3 Rössler方程$x(t)$和$y(t)$序列预测模型性能比较

Table 3　Prediction model performance comparison of $x(t)$ and $y(t)$ of Rössler equation

序列	$x(t)$		$y(t)$	
性能指标	E_{RMSE}	E_{PA}	E_{RMSE}	E_{PA}
$\eta = 1$ PCA+NN	0.0581	0.9995	0.0432	0.9997
$\eta = 1$ 文献[3]	0.0494	0.9998	0.0309	0.9998
$\eta = 1$ 单变量	0.0615	0.9996	0.0506	0.9997
$\eta = 20$ PCA+NN	0.0869	0.9992	0.1319	0.9986
$\eta = 20$ 文献[3]	0.1158	0.9985	0.1763	0.9993
$\eta = 20$ 单变量	0.3364	0.9875	0.2013	0.9869

4.2 大连降雨和气温二变量时间序列预测(Predicting rainfall and temperature time series of Dalian city)

气温和降雨是水文序列中关系极为密切的两个变量, 本节数据资料采用大连地区1951 年至2001年共612个数据点的月降雨和气温序列作为研究对象. 设$x(t)$代表降雨序列, $y(t)$代表气温序列, 采用前文方法计算可得$m_1 = m_2 = 12$, 选择$\tau_1 = \tau_2 = 1$, 则初始输入样本向量为24维

$$\overrightarrow{X}(t) = [x(t), \cdots, x(t-11), y(t), \cdots, y(t-11)]^{\mathrm{T}},$$

η步预测时, 网络输出为

$$\overrightarrow{Y}(t) = [x(t+\eta), y(t+\eta)]^{\mathrm{T}},$$

对降雨和气温序列分别进行主成分分析. 令$\eta_0 = 0.96$, 表4列出每个变量的奇异值和方差贡献率, 可以看出$x(t)$序列可降维至$p_1 = 9$; $y(t)$序列$p_2 = 5$. 新的输入向量维数为$p = p_1 + p_2 = 14$维.

一步预测时, 经整理总共可产生600组输入输出样本对. 神经网络仿真条件同表2, 取前300组样本为训练样本, 后300个样本为校验样本. 大连气温和降雨序列一步预测的网络学习误差和校验误差曲线如图4所示. 网络预测曲线和实际观测曲线的比较如图5所示.

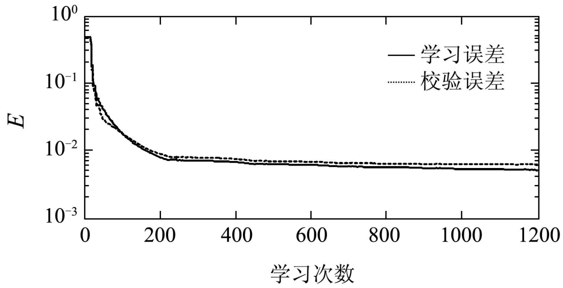

图 4　大连气温和降雨序列建模的学习和校验误差曲线

Fig. 4　Learning and testing curves of modeling rainfall and temperature series of Dalian

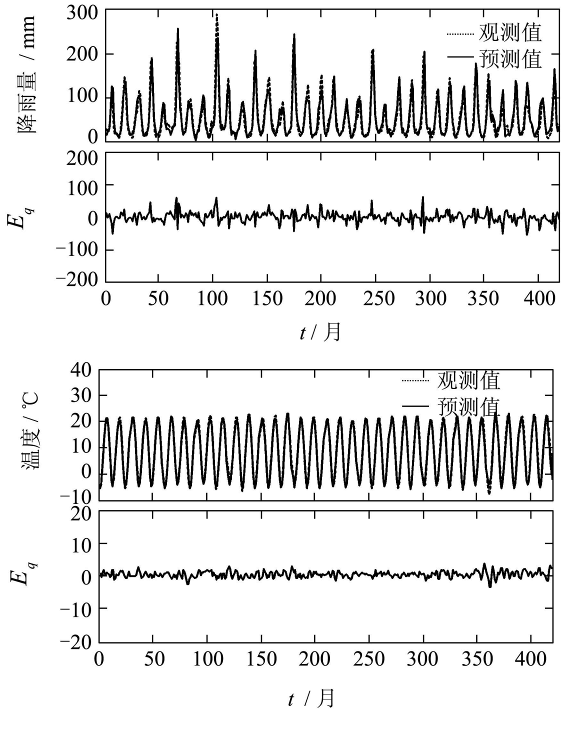

图 5　大连降雨和气温序列一步预测的仿真结果

Fig. 5　One-step prediction simulation results of rainfall and temperature series of Dalian

同例1, 单变量预测模型和文献[3]方法的结果见表5. 采用文献[3]方法, 计算可得$\tau_1 = \tau_2 = 5, m_1 = m_2 = 8$, 即输入向量维数为16. 可以看出, 对实际观测序列, 由于环境干扰等因素, 高维输入的不同状态之间相关性较强, 存在冗余信息. 本文方法利

用PCA增大输入变量间的方差, 降低冗余信息的影响, 具有更好的模型辨识能力. 对气温序列, 其周期规律性较强, 3种模型的预测精度相近, 而且受预测期的影响较小. 对降雨序列, PCA结合神经网络建立的多变量预测模型表现出较强的优越性, 随着预测期的增长, 这种优越性愈加明显.

表 4 大连降雨和气温时间序列的奇异值及方差贡献

Table 4 Singular values and their variance contribution of rainfall and temperature series

序号j		1	2	3	4
$x(t)$	s_j	8.2452	8.0538	4.1887	3.7475
$x(t)$	η_j	0.2452	0.2395	0.1246	0.1114
$y(t)$	s_j	30.8307	17.6706	17.6629	1.3168
$y(t)$	η_j	0.4386	02514	0.2512	0.0187
序号j		5	6	7	8
$x(t)$	s_j	3.3335	2.0734	1.4497	1.0119
$x(t)$	η_j	0.0991	0.0617	0.0431	0.0301
$y(t)$	s_j	1.2191	0.5734	0.3836	0.2544
$x(t)$	η_j	0.0173	0.0082	0.0055	0.0036
序号j		9	10	11	12
$x(t)$	s_j	0.6153	0.4352	0.2834	0.1893
$x(t)$	η_j	0.0183	0.0129	0.0084	0.0056
$y(t)$	s_j	0.1650	0.1133	0.0693	0.0413
$x(t)$	η_j	0.0023	0.0016	0.0010	0.0006

表 5 大连降雨和气温序列预测模型性能比较

Table 5 Prediction model performance comparison of rainfall and temperature series of Dalian

序列		$x(t)$		$y(t)$	
性能指标		$E_{\rm RMSE}$	$E_{\rm PA}$	$E_{\rm RMSE}$	$E_{\rm PA}$
$\eta=1$	PCA+NN	15.5745	0.9457	1.0069	0.9946
$\eta=1$	文献[3]	19.0334	0.9288	1.4797	0.9901
$\eta=1$	单变量	17.4778	0.9297	0.8948	0.9954
$\eta=5$	PCA+NN	26.8938	0.8258	1.1283	0.9929
$\eta=5$	文献[3]	28.4662	0.8227	1.6162	0.9849
$\eta=5$	单变量	31.2086	0.7533	0.9543	0.9948

5 结论(Conclusions)

构建能体现多组序列内部关系的模型结构, 以便深入掌握复杂系统的行为, 进行有效预测是本文的主要研究目的. 结合PCA和神经网络方法, 对多元变量序列进行建模和预测, 结果说明:

1) 主成分分析方法重构多元变量输入状态相空间, 可以避免嵌入维数等参数的计算, 建立有效的网络输入输出样本对. 建立初始嵌入延迟窗时, 同时考虑序列状态间的线性和非线性关系, 使输入向量能够包含充分的预测信息;

2) 仿真结果说明多个变量联合信息的输入可能改善网络的预测性能, 尤其是中、长期预测的精度可以得到提高, 为多元变量时间序列研究提供了一条新的思路.

参考文献(References):

[1] LEUNG H, LO T, WANG S. Prediction of noisy chaotic time series using an optimal radial basis function neural network[J]. *IEEE Trans on Neural Networks*, 2001, 12(5): 1163 – 1172.

[2] HAN M, XI J, XU S, et al. Prediction of chaotic time series based on the recurrent predictor neural Network[J]. *IEEE Trans on Signal Processing*, 2004, 52(12): 3409 – 3416.

[3] CAO L, MEES A, JUDD K. Dynamics from multivariate time series[J]. *Physica D*, 1998, 121: 75 – 88.

[4] 王海燕, 盛昭瀚, 张进. 多变量时间序列复杂系统的相空间重构[J]. 东南大学学报(自然科学版), 2003, 33(1): 115 – 118.
(WANG Haiyan, SHENG Zhaohan, ZHANG Jin. Phase space reconstruction of complex systems based on multivariate time series[J]. *Journal of Southeast University (Natural Science Edition)*, 2003, 33(1): 115 – 118.)

[5] BRAUNER N, SHACHAM M. Considering precision of data in reduction of dimensionality and PCA[J]. *Computers and Chemical Engineering*, 2000, 24(12): 2603 – 2611.

[6] 吴春国, 梁艳春, 孙延风, 等. 关于SVD与PCA等价性的研究[J]. 计算机学报, 2004, 27(2): 286 – 288.
(WU Chunguo, LIANG Yanchun, SUN Yanfeng, et al. On the Equivalence of SVD and PCA[J]. *Chinese Journal of Computers*, 2004, 27(2): 286 – 288.)

[7] YAZDIZADEH A, KHORASANI K. Adaptive time delay neural network structures for nonlinear system identification[J]. *Neurocomputing*, 2002, 47(1/4): 207 – 240.

[8] ROSENSTEIN M T, COLLINS J J, CARLO J de LUCA. Reconstruction expansion as a geometry-based framework for choosing proper delay time[J]. *Physica D*, 1994, 73: 82 – 98.

[9] LUIS A A. A nonlinear correlation function for selecting the delay time in dynamical reconstructions[J]. *Physics Letters A*, 1995, 203: 88 – 94.

[10] VALLE S, LI W, JOE Q S. Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods[J]. *Industrial & Engineering Chemistry Research*, 1999, 38(11): 4389 – 4401.

[11] 王永骥, 涂健. 神经元网络控制[M]. 北京: 机械工业出版社, 1998: 32 – 54.
(WANG Yongji, TU Jian. *Neural Cell Networks Control*[M]. Beijing: Machine Industry Press, 1998: 32 – 54.)

[12] CHEN J L, ISLAM S, BISWAS P. Nonlinear dynamics of hourly ozone concentrations: nonparametric short term prediction[J]. *Atmospheric environment*, 1998, 32(11): 1839 – 1848.

[13] KENYA A de OLIVEIRA,ÁLVARO V, ELTON C da SILVA. Using artificial neural networks to forecast chaotic time series[J]. *Physica A*, 2000, 284(1/4): 393 – 404.

作者简介:

席剑辉 (1975—), 女, 沈阳航空工业学院副教授, 博士, 主要从事混沌序列预测研究, E-mail: xjhui_01@163.com;

韩 敏 (1959—), 女, 大连理工大学电子与信息工程学院教授, 博士, 主要从事混沌序列分析、神经网络研究, E-mail: minhan@dlut.edu.cn.