

RBF-AR 模型在非线性时间序列预测中的应用

甘 敏¹, 彭 辉¹, 陈晓红²

(1. 中南大学 信息科学与工程学院, 长沙 410083; 2. 中南大学 商学院, 长沙 410083)

摘 要 研究了 RBF-AR 模型在非线性时间序列中的建模和预测问题, 并把它与其它一些新近提出的模型或方法加以比较. 一种结构化非线性参数优化方法用来辨识此模型. 数值实验和比较研究表明采用结构化非线性参数优化方法的 RBF-AR 模型在预测精度上要大大优于其它一些模型或方法.

关键词 RBF 网络; RBF-AR 模型; 时间序列预测

RBF-AR model-based nonlinear time series prediction

GAN Min¹, PENG Hui¹, CHEN Xiao-hong²

(1. School of Information Science and Engineering, Central South University, Changsha 410083, China;
2. School of Business, Central South University, Changsha 410083, China)

Abstract This paper investigated nonlinear time series modeling and forecasting problem based on RBF-AR model, and the comparisons between RBF-AR model with other newly developed models. A structured nonlinear parameter optimization method (SNPOM) was applied to estimate the model. It is shown by the simulation tests and comparisons that the performance of RBF-AR estimated by SNPOM is superior to other models.

Keywords RBF networks; RBF-AR model; time series forecasting

1 引言

时间序列分析在工程、经济和自然科学等领域是一个重要研究课题, 它是一种用当前和过去值对未来值进行预测的重要工具. 由于实际系统的复杂性, 在过去三十年中时间序列分析的理论 and 实践已经进入非线性时代. 许多非线性时间序列模型被提出, 比如双线性模型、门限自回归模型和指数自回归模型. 近些年来, 作为一种普遍性的方法, 神经网络已成为一种流行的函数逼近和时间序列预测的工具. 其中, 径向基函数 (Radial basis function, RBF) 网络由于其简单的结构、非线性逼近的精度和快速学习的能力得到了广泛的青睐^[1-3]. 一些新的变体或模型也被大量提出, 如 Rojas 等提出了带回归权重的归一化伪高斯径向基网络 (Normalized PG-RBF network with regression weights) 模型^[4], 与传统的 RBF 网络相比得到了更好的预测精度. Chen 等提出了灵活神经树 (Flexible neural tree, FNT) 模型^[5], 这种模型可以自动的选择合适的输入变量. 随后, 他们又提出了局部线性小波神经网络^[6] (Local linear wavelet neural network, LLWNN). 最近, Valenzuela 等把人工神经网络, 模糊系统和进化算法几种智能技术相结合来提高混合的 ARIMA-ANN 模型的预测精度^[7].

另一类可以普遍描述非线性时间序列的是状态相依模型^[8] (State-dependent models, SDMs). Vesin^[9] 和 Shi 等^[10] 用一组 RBF 网络来逼近状态相依自回归 (State-dependent autoregressive, SD-AR) 模型中的函数系数就得到了 RBF-AR 模型. 直观地, RBF-AR 模型既有 RBF 网络函数逼近的优势, 也有 SD-AR 模型描

收稿日期: 2009-02-21

资助项目: 国家创新研究群体科学基金 (70921001); 湖南省科技计划国际合作重点项目 (2009WK2009); 国家自然科学基金 (60574058)

作者简介: 甘敏 (1982-), 男, 博士研究生, 研究领域为非线性时间序列分析及其应用, 神经网络, 进化计算; 彭辉 (1963-), 男, 教授, 博士生导师, 研究领域为复杂系统建模、控制与优化, 先进控制与智能化系统, 金融时序分析 with 最优投资组合.

述非线性优势. 然而, 应用 RBF-AR 模型的一个难点是其模型的辨识, 包括模型阶次的选择, RBF 网络中心、比例缩放因子以及线性权重的估计. 由于 RBF-AR 模型的参数可分为线性参数和非线性参数, 且线性参数数目大量多于非线性参数, 针对此类 RBF 型模型, 采用结构化的非线性参数优化方法 (Structured nonlinear parameter optimization method, SNPOM)^[11], 可提高收敛速度并得到较优的精度. 数值实验和比较研究表明采用 SNPOM 优化算法的 RBF-AR 模型在预测精度上要大大优于新近提出的一些模型.

2 问题的描述

不失一般性, 考虑非线性时间序列的一步向前预测问题. 对于一个给定的时间序列 $\{y(t) \in \mathbb{R}^1, t = 1, 2, \dots, N\}$, 非线性时间序列建模的主要问题就是构建这样一个函数, $f: \mathbb{R}^p \rightarrow \mathbb{R}^1$, 即

$$y(t) = f(y(t-1), y(t-2), \dots, y(t-p)) + e(t) \quad (1)$$

其中, $f(\cdot)$ 为非线性映射, p 是正整数, $e(t)$ 是与 $\{y(t-i), i > 0\}$ 独立的白噪声.

3 RBF-AR 模型

模型 (1) 的形式非常一般, 不能直接应用. 在参数模型和非参数模型 (1) 之中有许多有用的模型, 本文所感兴趣的是一类叫做状态相依自回归^[8] 或者函数系数自回归^[12] 的模型, 其形式为:

$$y(t) = \phi_0(\mathbf{X}(t-1)) + \sum_{i=1}^p \phi_i(\mathbf{X}(t-1))y(t-i) + e(t) \quad (2)$$

这里, $\mathbf{X}(t-1)$ 叫做状态向量, 一般为由 $y(t)$ 的延迟值所构成的向量; 如有必要, $\mathbf{X}(t-1)$ 也可包含其它解释变量. $\{\phi_i(\mathbf{X}(t-1)), i = 0, 1, \dots, p\}$ 叫做状态相依函数系数.

状态相依自回归模型可视为门限自回归模型和指数自回归模型的推广^[13], 它使得函数系数随状态变量 $\mathbf{X}(t-1)$ 的变化而变化. 虽然状态相依模型 (2) 为非线性时间序列建模提供了一个很好的框架, 但是应用它的主要难点是如何明确函数系数 $\phi_i(\mathbf{X}(t-1))$ 的形式. 实际上, 可以把此问题看成从多维空间 \mathbf{X} 到一维空间 $\phi_i(\mathbf{X})$ 的一个函数逼近问题. RBF 神经网络是解决此问题的一个有效的方法, 它已被从理论上证明能以任意精度逼近任何函数, 因此可用 RBF 神经网络来逼近模型 (2) 中的状态相依函数系数. 高斯 RBF 网络可表示为:

$$\phi(\mathbf{X}) = w_0 + \sum_{k=1}^m w_k \exp(-\lambda_k \|\mathbf{X} - \mathbf{Z}_k\|^2) \quad (3)$$

其中, w_0 为偏置, $w_k (k = 1, 2, \dots, m)$ 为连接隐含层结点和输出结点的线性权重, m 为隐含层结点数目, λ 称为缩放比例因子 (或宽度), \mathbf{X} 和 \mathbf{Z} 分别表示 RBF 网络的输入向量和它的中心向量, $\|\cdot\|$ 表示向量的欧式距离.

用以上形式的 RBF 网络 (3) 来逼近模型 (2) 中的状态相依函数系数就得到了 RBF-AR 模型:

$$\begin{cases} y(t) = \phi_0(\mathbf{X}(t-1)) + \sum_{i=1}^p \phi_i(\mathbf{X}(t-1))y(t-i) + e(t) \\ \phi_i(\mathbf{X}(t-1)) = w_{i,0} + \sum_{k=1}^m w_{i,k} \exp(-\lambda_k \|\mathbf{X}(t-1) - \mathbf{Z}_k\|^2) \\ \mathbf{Z}_k = (z_{k,1}, z_{k,2}, \dots, z_{k,d})^T \end{cases} \quad (4)$$

其中, p, m 和 $d = \dim\{\mathbf{X}(t-1)\}$ 是模型的阶次; T 表示向量或矩阵的转置. 可以看出, RBF 网络 (3) 是 RBF-AR 模型 (4) 的一个组成部分, 因此 RBF-AR 模型可以看作比 RBF 网络更一般的非线性模型. 在一些实际的应用中, RBF 网络可能需要大量的中心以得到满意的建模精度. 然而, RBF-AR 模型把问题的复杂度分散到了各个自回归部分, 因此在达到相似的精度的情况下, RBF-AR 模型中心的个数可能会比 RBF 网络少得多. 这一点将在数值实验中得到验证.

4 模型估计方法

RBF-AR 模型的辨识包括阶次的选择和估计模型中所有的参数. 阶次 (p, m, d) 的选择可以比较其 AIC (Akaike information criterion) 值^[14]. 因此首要的问题是要有一个好的参数优化方法, 然后用不同的阶次重复实验最终选择最好的模型. 在本文中, 我们采用由彭辉等提出的结构化非线性参数优化方法^[11]. 这种方法

把被搜索参数分为线性参数空间和非线性参数空间, 用最小二乘法 (Least-squares method, LSM) 来优化模型中的线性参数, 用 Levenberg-marquardt method(LMM) 方法来优化非线性参数. 此方法极大地提高收敛速度并得到了更高的精度. 这种优化策略按如下步骤进行.

4.1 参数分类

RBF-AR 模型 (4) 的参数可分为线性参数

$$\Theta_L \doteq \{w_{i,k} | i = 0, 1, \dots, p; k = 0, 1, \dots, m\} \quad (5)$$

和非线性参数

$$\Theta_N \doteq \{\lambda_k; \mathbf{Z}_k | k = 1, 2, \dots, m\} \quad (6)$$

为便于下面解释说明, 把模型 (4) 改写成如下两种形式:

$$y(t) = f(\Theta_L, \Theta_N, \mathbf{X}(t-1)) + e(t) \quad (7)$$

$$y(t) = \Psi(\Theta_N, \mathbf{X}(t-1))^T \Theta_L + e(t) \quad (8)$$

4.2 初始化

首先, 在向量空间 $\mathbf{X}(t-1)$ 中随机选择一个子空间 $\mathbf{Z}_k^0 (k = 1, 2, \dots, m)$ 作为 RBF 网络的中心的初始值, 然后用方程 (9) 来计算比例缩放因子的初始值:

$$\lambda_k^0 = -\log \varepsilon / \max_{t-1} \{\|\mathbf{X}(t-1) - \mathbf{Z}_k^0\|^2\} \quad (9)$$

其中, ε 是一个较小的正数. 这样做是为保证: 当状态向量 $\mathbf{X}(t-1)$ 离中心 \mathbf{Z}_k 很远时, $\exp(-\lambda_k \max_{t-1} \{\|\mathbf{X}(t-1) - \mathbf{Z}_k\|^2\})$ 项趋于 0, 使线性权重稳定有界^[10-11]. 选定了初始非线性参数 Θ_N^0 后, 用 LSM 来计算初始线性参数 Θ_L^0 :

$$\begin{cases} \Theta_L^0 = \left[\sum_{t=\tau+1}^M \bar{\Psi}_{t,0} \bar{\Psi}_{t,0}^T \right]^{-1} \sum_{t=\tau+1}^M \bar{\Psi}_{t,0} \bar{y}(t) \\ \bar{\Psi}_{t,0} = \Psi(\Theta_N^0, \bar{\mathbf{X}}(t-1)) \end{cases} \quad (10)$$

这里, $\{\bar{y}(i), \bar{\mathbf{X}}(i-1) | i = \tau+1, \dots, M\}$ 是观测数据集, τ 是模型的最大时间延迟, M 是观测数据的个数.

4.3 优化

定义目标函数为:

$$\begin{cases} V(\Theta_L, \Theta_N) \doteq \frac{1}{2} \|\mathbf{F}(\Theta_L, \Theta_N)\| \\ \mathbf{F}(\Theta_L, \Theta_N) \doteq \begin{bmatrix} f(\Theta_L, \Theta_N, \bar{\mathbf{X}}(\tau)) - \bar{y}(\tau+1) \\ f(\Theta_L, \Theta_N, \bar{\mathbf{X}}(\tau+1)) - \bar{y}(\tau+2) \\ \vdots \\ f(\Theta_L, \Theta_N, \bar{\mathbf{X}}(M-1)) - \bar{y}(M) \end{bmatrix} \end{cases} \quad (11)$$

因而, 参数优化问题即为计算:

$$(\hat{\Theta}_L, \hat{\Theta}_N) = \arg \min_{\Theta_L, \Theta_N} V(\Theta_L, \Theta_N) \quad (12)$$

两个迭代过程用来优化要估计的参数. 记 κ 为迭代的步数, 对于非线性参数向量 Θ_N , 更新的公式为:

$$\Theta_N^{\kappa+1} = \Theta_N^{\kappa} + \beta_{\kappa} \mathbf{d}_{\kappa} \quad (13)$$

其中, \mathbf{d}_{κ} 是搜索方向, β_{κ} 为步长系数. 对于线性参数向量 Θ_L , 与式 (10) 相似:

$$\begin{cases} \Theta_L^{\kappa+1} = \left[\sum_{t=\tau+1}^M \bar{\Psi}_{t,\kappa+1} \bar{\Psi}_{t,\kappa+1}^T \right]^{-1} \sum_{t=\tau+1}^M \bar{\Psi}_{t,\kappa+1} \bar{y}(t) \\ \bar{\Psi}_{t,\kappa+1} = \Psi(\Theta_N^{\kappa+1}, \bar{\mathbf{X}}(t-1)) \end{cases} \quad (14)$$

式 (13) 中每一次迭代的 \mathbf{d}_{κ} 由如下方程决定:

$$\begin{cases} [\mathbf{J}(\Theta_N^{\kappa})^T \mathbf{J}(\Theta_N^{\kappa}) + \gamma_{\kappa} \mathbf{I}] \mathbf{d}_{\kappa} = -\mathbf{J}(\Theta_N^{\kappa})^T \mathbf{F}(\Theta_L^{\kappa}, \Theta_N^{\kappa}) \\ \mathbf{J}(\Theta_N^{\kappa}) = (\partial \mathbf{F}(\Theta_L^{\kappa}, \Theta_N^{\kappa}) / \partial \Theta_N^{\kappa})^T \end{cases} \quad (15)$$

其中, γ_{κ} 控制着 \mathbf{d}_{κ} 的幅度和方向. 当 γ_{κ} 趋近于 0 时, \mathbf{d}_{κ} 将沿高斯-牛顿方向; 当 γ_{κ} 趋近无穷大时, \mathbf{d}_{κ} 将沿最速下降方向. 当 \mathbf{d}_{κ} 求得后, 式 (13) 中的 β_{κ} 由类似与混合二次、三次多项式内插、外插法的线搜索过

程得到. 应当注意在以上的优化过程中线性参数 (或非线性参数) 的更新紧随非线性参数 (或线性参数) 的更新^[11].

5 数值实验

5.1 预测 Mackey-Glass 时间序列

在此节中, RBF-AR 模型将用来预测著名的 Mackey-Glass 时间序列, 来与现存的一些其它模型或方法作比较. Mackey-Glass 微分方程可表示为:

$$\frac{dy}{dt} = \frac{ay(t-\tau)}{1+y^c(t-\tau)} - by(t) \quad (16)$$

这里, 与以前大多数的文献保持一致, 我们选取 $a = 0.2, b = 0.1, c = 10, \tau = 17$. 用四阶的 Runge-Kutta 方法来生成数据, 初始值为 $y(0) = 1.2$. 与文献 [1-7] 相同, 选取 $t = 124$ 到 $t = 1123$ 的 1000 个数据, 前 500 个数据用来训练模型, 后 500 个数据用来检验模型. 为与以前文献 [1-7] 中的工作公平的比较, 我们用固定的向量 $[y(t-18) \ y(t-12) \ y(t-6) \ y(t)]$ 来预测 $y(t+6)$. 用以下 RBF-AR 模型来建模此时间序列:

$$\begin{cases} y(t+6) = \phi_0(\mathbf{X}(t-1)) + \sum_{i=1}^4 \phi_i(\mathbf{X}(t-1))y(t-6(i-1)) + e(t) \\ \phi_i(\mathbf{X}(t-1)) = w_{i,0} + \sum_{k=1}^m w_{i,k} \exp(-\lambda_k \|\mathbf{X}(t-1) - \mathbf{Z}_k\|^2) \\ \mathbf{X}(t-1) = [y(t), y(t-6), y(t-12), y(t-18)]^T \end{cases} \quad (17)$$

表 1 不同结点数目的 RBF-AR 模型对 Mackey-Glass 时间序列的预测结果

模型	结点数	运行时间 (s)	训练数据均方误差	AIC	检验数据均方误差	AIC
RBF-AR(4,10,4)	10	36	3.6489×10^{-6}	-6070.5	3.3547×10^{-6}	-6112.6
RBF-AR(4,12,4)	12	33	3.0088×10^{-6}	-6131.0	2.6286×10^{-6}	-6198.5
RBF-AR(4,14,4)	14	52	1.8856×10^{-6}	-6328.6	1.6735×10^{-6}	-6388.3
RBF-AR(4,16,4)	16	63	1.2865×10^{-6}	-6483.8	1.0805×10^{-6}	-6571.0
RBF-AR(4,18,4)	18	116	6.6598×10^{-7}	-6777.0	6.0959×10^{-7}	-6821.2
RBF-AR(4,20,4)	20	118	5.0704×10^{-7}	-6877.3	5.3484×10^{-7}	-6850.6

用 RBF-AR(p, m, d) 来表示此模型, 显然 $p = 4$. 由于 $\mathbf{X}(t-1)$ 的维数为 4, 故 $d = 4$. 那么可以调节的参数就是 RBF 网络隐含层的结点数目 m 了. 对 $m = 10, 12, 14, 16, 18, 20$ 这些情况, 用第 4 节所介绍的 SNPOM 优化方法做了实验. 实验在 Matlab, CPU 为 3.0GHz, RAM 为 512M 的 windows 系统中完成. 实验结果在表 1 中给出. 从表中可以看到所有模型预测的均方误差 (Mean squared error, MSE) 都非常小, 而且注意到部分模型在检验数据 (Testing data) 上得到的均方误差要比训练数据 (Training data) 要小, 这说明模型的确从训练数据中获取了很好的预测能力. 对于模型优化的计算强度, 由于结点数目比较小, 训练过程在较短的时间内就完成了.

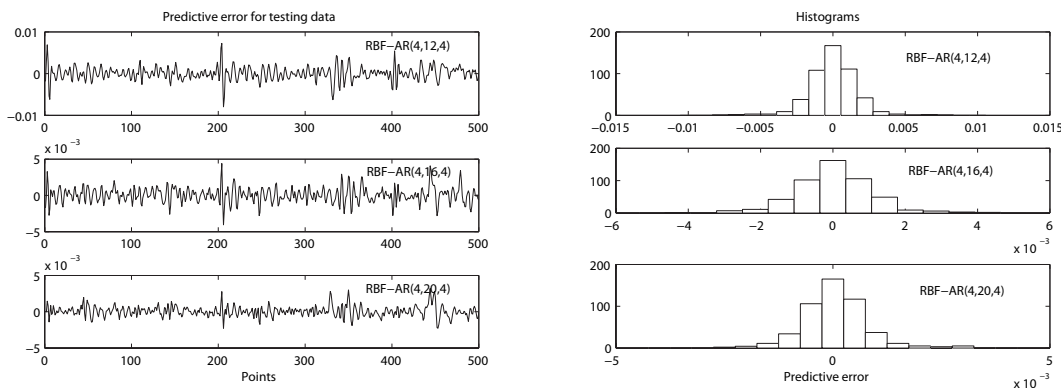


图 1 RBF-AR 模型对 Mackey-Glass 时间序列检验数据的预测误差及其直方图

为了检验模型对 Mackey-Glass 时间序列预测的有效性, 我们在图 1 中分别给出了模型 RBF-AR(4,12,4), RBF-AR(4,16,4) 和 RBF-AR(4,20,4) 对测试数据的预测误差及其直方图. 从图 1 可以看到模型的预测误差序列看起来非常接近白噪声, 且其直方图比较对称的分布在 0 两边, 成高斯状. 图 2 给出了 Mackey-Glass 原始数据和 RBF-AR(4,20,4) 模型的预测数据. 可以看到模型很好地拟合了原始数据.

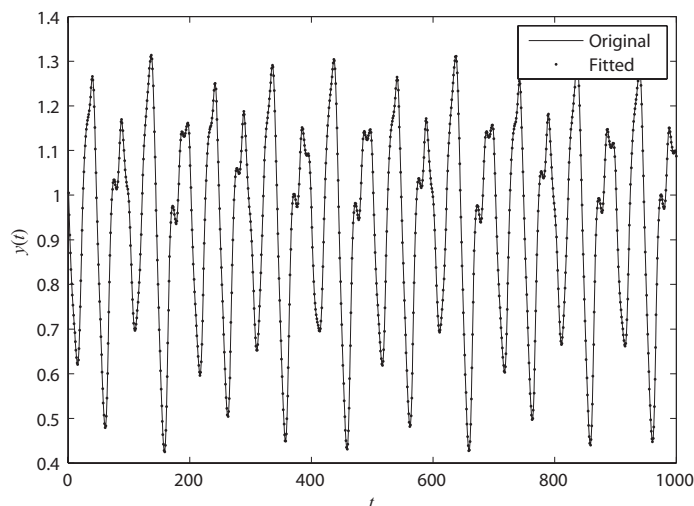


图 2 Mackey-Glass 时间序列原始数据与 RBF-AR(4,20,4) 模型拟合值

下面用以上得到的结果与现存的其它模型或方法作一下比较, 表 2 列出了其它一些方法的实验结果, 其中表的头 7 行结果来自文献 [4]. 从表 2 中可以看到 RBF-AR(20 个结点) 模型得到的结果要好于其它模型一个或几个数量级. 有必要考察一下相同结点数情况下的效果. 文献 [4] 中 PG-RBF 模型用 12 个结点得到了非常好的效果 (8.2369×10^{-6}), 而 RBF-AR(4,12,4) 得到了好得多的结果 (3.0088×10^{-6} , 见表 1). 文献 [6] 的 LLWNN 模型用 10 个结点的预测结果为 1.2960×10^{-5} , RBF-AR(4,10,4) 为 3.3547×10^{-6} (见表 1).

文献 [2] 和 [3] 是用 RBF 网络来做时间序列预测的最新研究成果, 它们都得到了非常好的效果. 但是它们的结点数分别为 238 和 98, 那么就需要大量的时间来优化. 特别是文献 [3] 中用遗传算法来优化 RBF 网络中心 (98 个结点意味着有 392 个 RBF 网络中心), 而遗传算法的缺点是收敛速度慢、计算强度大, 那么必定要花大量的时间. 而 RBF-AR 模型用的中心数要少得多, 优化的方法又是基于梯度的, 因而计算强度大大降低. 如果要达到文献 [3] 中 GA-RBF 的预测精度, RBF-AR 模型只需要用 14 个结点 (见表 1). 因此可以看出, 在达到相似的精度的情况下, RBF-AR 模型中心的个数可能会比 RBF 网络少得多.

表 2 不同模型或方法对 Mackey-Glass 时间序列的预测结果

模型	训练数据均方误差	检验数据均方误差
Autoregressive model	0.0361	Not provided
Cascade correlation NN	0.0036	Not provided
Back-propagation NN	4×10^{-4}	Not provided
Six-order polynomial	0.0016	Not provided
Linear prediction method	0.3025	Not provided
ANFIS and fuzzy system	4.9×10^{-5}	Not provided
Product T-norm	0.0082	Not provided
AFS-RBF (13 neurons) ^[1]	1.4161×10^{-4}	1.7161×10^{-4}
Cubic-RBF (238 neurons) ^[2]	Not provided	1.729×10^{-6}
GA-RBF (98 neurons) ^[3]	1.96×10^{-6}	2.25×10^{-6}
PG-RBF (12 neurons) ^[4]	8.2369×10^{-6}	Not provided
Flexible neural tree ^[5]	4.7624×10^{-5}	5.0737×10^{-5}
LLWNN (10 neurons) ^[6]	1.0890×10^{-5}	1.2960×10^{-5}
AMIMA-ANN ^[7]	7.29×10^{-6}	Not provided
RBF-AR (20 neurons)	5.0704×10^{-7}	5.3484×10^{-7}

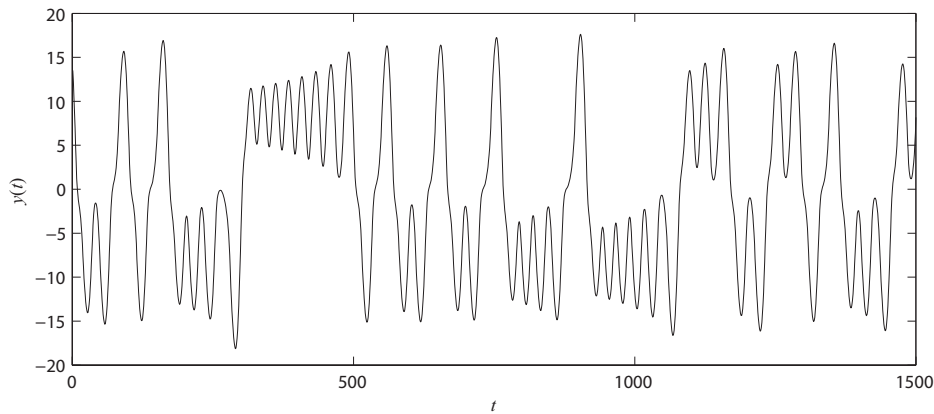


图 3 Lorenz 吸引子混沌时间序列

5.2 预测 Lorenz 吸引子时间序列

Lorenz 吸引子系统可由以下微分方程表示:

$$\begin{cases} \dot{y} = \sigma(z - y) \\ \dot{z} = \rho y - z - yw \\ \dot{w} = yz - \beta w \end{cases} \quad (18)$$

其中, $\sigma = 10$, $\rho = 28$, $\beta = 3/8$. 用 Runge-Kutta 方法解此微分方程来生成数据, 步长为 0.05. 选取了变量 y 的 1500 个数据 (图 3), 前 1100 个数据用来训练模型, 后 400 个数据来检验模型. 用以下 RBF-AR 模型来建模此时间序列:

$$\begin{cases} y(t) = \phi_0(\mathbf{X}(t-1)) + \sum_{i=1}^p \phi_i(\mathbf{X}(t-1))y(t-i) + e(t) \\ \phi_i(\mathbf{X}(t-1)) = w_{i,0} + \sum_{k=1}^m w_{i,k} \exp(-\lambda_k \|\mathbf{X}(t-1) - \mathbf{Z}_k\|^2) \\ \mathbf{X}(t-1) = [y(t-1), y(t-2), \dots, y(t-d)]^T \end{cases} \quad (19)$$

表 3 中给出了 RBF-AR(5,3,3) 模型以及其它一些文献对此时间序列的预测结果. 从表中可以看出 RBF-AR 模型的预测结果要好于其它方法.

表 3 Lorenz 吸引子时间序列预测结果比较

模型	训练数据均方误差	检验数据均方误差
PG-RBF [4]	0.0088	Not provided
AMIMA-ANN [7]	0.0077	Not provided
RBF-AR	0.0068	0.0072

6 结束语

最近几十年大量的非线性时间序列模型被提出. RBF-AR 模型属于一类有突出性能的变系数模型, 然而最近提出的模型并没有和这种模型加以比较. 本文研究了采用结构化非线性参数优化方法的 RBF-AR 模型在非线性的时间序列中的建模与预测问题, 并把它与一些新近模型或方法作了比较. 实验结果表明 RBF-AR 模型在预测精度上要优于其它一些模型或方法. 在达到相似的预测精度的情况下, RBF-AR 模型需要的中心个数比 RBF 网络少得多. 今后的工作是如何选择状态向量的最优组合, 以前的一些研究^[3,5]表明 RBF 网络输入变量的选择对预测结果有重大的影响.

参考文献

- [1] Cho K B, Wang B H. Radial basis function based adaptive fuzzy systems and their applications to system identification and prediction[J]. Fuzzy Sets and Systems, 1996, 83(3): 325-339.

-
- [2] Harpham C, Dawson C W. The effect of different basis functions on a radial basis function network for time series prediction: A comparative study[J]. *Neurocomputing*, 2006, 69(16): 2161–2170.
 - [3] Du H, Zhang N. Time series prediction using evolving radial basis function networks with new encoding scheme[J]. *Neurocomputing*, 2008, 71: 1388–1400.
 - [4] Rojas I, Pomares H, Bernier J L, et al. Time series analysis using normalized PG-RBF network with regression weights[J]. *Neurocomputing*, 2002, 42: 267–285.
 - [5] Chen Y, Yang B, Dong J, et al. Time series forecasting using flexible neural tree model[J]. *Information Sciences*, 2005, 174: 219–235.
 - [6] Chen Y, Yang B, Dong J. Time series prediction using a local linear wavelet neural network[J]. *Neurocomputing*, 2006, 69: 449–465.
 - [7] Valenzuela O, Rojas I, Rojas F, et al. Hybridization of intelligent techniques and ARIMA models for time series prediction [J]. *Fuzzy Sets and Systems*, 2008, 159: 821–845.
 - [8] Priestley M B. State dependent models: A general approach to nonlinear time series analysis[J]. *Journal of Time Series Analysis*, 1980, 1: 57–71.
 - [9] Vesin J. An amplitude-dependent autoregressive signal model based on a radial basis function expansion[C]// *Proceedings of International Conference of ASSP, Minnesota*, 1993, 3: 129–132.
 - [10] Shi Z, Tamura Y, Ozaki T. Nonlinear time series modeling with the radial basis function-based state-dependent autoregressive model[J]. *International Journal of System Science*, 1999, 30(7): 717–727.
 - [11] Peng H, Ozaki T, Haggan-Ozaki V, et al. A parameter optimization method for the radial basis function type models[J]. *IEEE Transactions on Neural Networks*, 2003, 14(2): 432–438.
 - [12] Chen R, Tsay R S. Functional-coefficient autoregressive models[J]. *Journal of the American Statistical Association*, 1993, 88(421): 298–308.
 - [13] Haggan V, Ozaki T. Modeling nonlinear random vibrations using an amplitude-dependent autoregressive time series model[J]. *Biometrika*, 1981, 68: 189–196.
 - [14] Akaike H. A new look at the statistical model identification[J]. *IEEE Transactions on Automatic Control*, 1974, 19: 716–723.