

基于遗传神经网络的超混沌时间序列预测

王永生¹, 肖支才², 叶文¹, 范洪达¹

(1. 海军航空工程学院兵器科学与技术系, 烟台 264001; 2. 海军航空工程学院控制工程系, 烟台 264001)

摘要: 研究了超混沌系统的预测问题。通过分析混沌时间序列, 建立具有多个隐节点的3层前馈网络, 基于泛化性考虑采用剪枝算法训练, 在保证预测精度的基础上消去部分隐节点以降低网络复杂性, 利用遗传算法具有的全局寻优能力重新训练网络, 利用具有局部寻优能力BP算法再次训练该网络。对Mackey-Glass时滞混沌系统预测实验结果表明, 改进算法的泛化性能优于经典BP网络, 归一化预测精度提高10倍多, 能够较好地解决超混沌系统的预测问题。

关键词: 超混沌; 神经网络; 遗传算法; 权消去法; 预测

Hyperchaotic Time Series Forecast Based on Genetic Neural Networks

WANG Yong-sheng¹, XIAO Zhi-cai², YE Wen¹, FAN Hong-da¹

(1. Department of Science and Technology of Weapons, Naval Aeronautical Engineering Institute, Yantai 264001;

2. Department of Control Engineering, Naval Aeronautical Engineering Institute, Yantai 264001)

【Abstract】 This paper studies the forecast of the hyperchaotic system. After analyzing the chaotic time series, a three-layer forward artificial neural network is built up with many nodes in the hidden layer. Considering generalization ability of the net, Weight-Elimination(WE) algorithm is adopted to delete some hidden nodes for reducing complexity of the net and assure the net forecast precision. Genetic Algorithm(GA) is introduced to train the net over again for its global search ability. And the acquired net is trained again by classical BP arithmetic with its localized search. Experiments on the Mackey-Glass time lag chaos system illustrate that the improved method is better than the classical BP arithmetic, and the normalized forecast precision is enhanced by more than 10 times, so it can resolve the prediction of the super chaotic system.

【Key words】 hyperchaos; neural networks; Genetic Algorithm(GA); Weight-Elimination(WE); forecast

1 概述

混沌信号作为一种宽频带类噪声信号, 具有分维特征、对初始条件比较敏感、难以长期预测等特点, 越来越多的研究人员投身于混沌的动态特征及其应用研究, 对混沌系统的建模和预测已成为混沌信号处理的研究热点^[1]。混沌时间序列是指通过对一个混沌系统进行观测采样而得到的一个单变量时间序列。混沌时间序列的预测可看作动力学系统研究的“反问题”, 即给定相空间中的一串迭代序列(轨道的演化过程)或一组观测序列, 要构造一个非线性映射来表达原系统, 这个映射可作为预测模型, 其理论基础是相空间重构理论。然而相空间重构理论仅仅通过单变量信息重构吸引子, 通常用于估计系统的复杂性、确定用于预测的神经网络的结构^[2]。

非线性动力系统可以分为高维系统和低维系统, 偏微分方程和时滞方程描述无穷维系统, 常微分方程描述有限维系统。低维混沌系统可以通过相空间重构理论基本解决网络的结构选择, 但并不完全适用于高维超混沌系统, 而实际中往往遇到混沌程度较高的数据, 如海杂波数据、脑电数据。与双螺旋分类问题类似, 虽然输入数据相对简单, 但是通常难以确定满足问题要求的隐节点的个数, 而且混沌时间序列预测要求在未经训练的数据集上也具有很好的预测性能, 即网络需要具有很好的泛化性能, 简单的增加隐节点数目往往导致网络泛化性能的降低^[3]。

相对于低维非线性系统, 针对无穷维的非线性混沌系统预测的研究还很少。为解决高维非线性超混沌时间序列的预

测问题, 本文首先对获得的混沌时间序列数据进行简单的重构分析, 确定神经网络的输入层节点数目; 构造具有多个隐节点的3层前向网络, 采用权消去剪枝算法处理, 得到一个满足一定预测精度同时保证网络泛化性能的网络结构; 重新对网络训练, 为避免经典BP学习算法易于陷入误差函数的局部极值点, 利用遗传算法的全局寻优特点训练网络的权值; 考虑到遗传算法搜索到最优解附近时, 无法精确地确定最优解的位置, 即它在局部搜索空间不具备微调能力, 再次利用BP算法局部寻优特点训练网络, 以确定最优的网络权值。

2 神经网络剪枝算法

目前大多数神经网络的研究不能较好地解决确定输入与隐层单元数的定量规则问题, 仅有一些通用指导: (1)使网络成为一个完全通用的映射, 必须至少有一个隐层。1989年Robert-Nielson证明一个隐层的BP网可逼近闭区间内任意一个连续函数。(2)网络结构要尽可能紧致, 最好是满足要求的最小网络。为了避免过拟合, 应该设计最小结构的神经网络, 最简单的网络平均意义上泛化能力最好。剪枝方法采用自顶向下的设计方法, 先构造一个足够大的网络, 然后通过训练时删除或合并某些节点或权值, 达到精简网络结构、改进泛化的目的。本文利用权消去法(Weight-Elimination, WE)^[2-4]

作者简介: 王永生(1978-), 男, 讲师、博士研究生, 主研方向: 通信与信息系统, 信号检测; 肖支才, 讲师、硕士; 叶文, 讲师、博士研究生; 范洪达, 教授、博士生导师

收稿日期: 2007-11-20 **E-mail:** hjhywys@sohu.com

剪除网络中的冗余权值，即消除冗余的隐节点。

在神经网络的误差反向传播学习过程中，网络各权值的调整公式为

$$w_j(k) = w_j(k-1) + \Delta w_j^1 + \lambda \Delta w_j^2 \quad (1)$$

$$\Delta w_j^1 = -\eta \frac{\partial Err}{\partial w_j} \quad (2)$$

$$\Delta w_j^2 = -\frac{2w_j/w_0^2}{1+w_j^2/w_0^2} \quad (3)$$

其中， $\frac{\partial Err}{\partial w_j}$ 可按照标准 BP 算法获得； η 为学习速率； λ 为

正则化系数，对 λ 的修改采用文献[3]中的动态修改策略。训练一定次数后，网络中的某些权值将衰减到 0 附近，如果一个隐节点的输出权值都接近 0，就可以删除该节点；如果一个隐节点的输入权值全为 0，则把该隐节点并入输出层的偏移。通过这些操作，即可得到较为精简的神经网络。

3 遗传优化神经网络

利用剪枝法得到的神经网络虽然能够从结构上保证网络具有一定的泛化性能，但剪枝得到的最终权值并不是最佳的，还需要重新训练。而经典的 BP 学习算法易陷入误差函数的局部极值点，而且对于较大的搜索空间、多峰值和不可微函数，也不能搜索到全局极小点。遗传算法(Genetic Algorithm, GA)将达尔文的自然界生物进化思想灵活运用到优化运算领域，能够同时处理空间的若干点，从而有助于搜索全局最优点，免于陷入局部最小，在误差函数不可微或完全没有梯度信息的情况下尤为有效。利用遗传算法的杂交、变异、选择算子可以在全变量空间以较大概率搜索全局解的特点，本文运用遗传算法对神经网络权值进行训练，其中的关键问题是：权值编码方案、编码长度、遗传参数的选择和适应度的确定。并使用了美国 North Carolina State University 开发的遗传算法优化工具箱(Genetic Algorithm Optimization Toolbox, GAOT)[5]对网络权值进行训练。

对于剪枝法得到的 3 层前向网络，假设 I_i 为输入层中第 i 个节点的输出； H_i 为隐含层中第 i 个节点的输出； O_i 为输出层中第 i 个节点的输出； WIH_{ij} 为输入层中第 i 个节点与隐含层第 j 个节点的连接权值； WHO_{ji} 为隐含层中第 j 个节点与输出层第 i 个节点的连接权值。利用遗传算法训练网络权值的步骤如下：

(1) 初始化种群 Pop (包括交叉规模、交叉概率 p_c 、突变概率 p_m) 以及任一 WIH_{ij} 和 WHO_{ji} ；采用实数进行编码，初始种群取 100。

(2) 计算每一个个体评价函数，并将其排序，按下式得到的概率值选择网络个体：

$$p_i = f_i / \sum_{i=1}^N f_i \quad (4)$$

其中， f_i 为个体 i 的适配值，可用误差平方和 E 来衡量，即

$$f(i) = 1/E(i), E(i) = \sum_p \sum_k (V_k - T_k)^2 \quad (5)$$

其中， $i=1,2,\dots,N$ 为种群个体数； k 为输出层节点数； p 为训练样本数。

(3) 以概率 p_c 对个体 G_i 和 G_{i+1} 交叉操作产生新个体 G_i 和 G_{i+1} ，没有进行交叉操作的个体直接复制。

(4) 利用概率 p_m 突变产生 G_i 的新个体 G'_i 。

(5) 将新个体插入到种群 Pop 中，并计算新个体的评价

函数。

(6) 如果找到了满意的个体，结束，否则转(3)。

达到所要求的性能指标或最大指定世代数后，将最终群体中的最优个体解码即可得到优化后的网络连接权系数。

最后考虑到当遗传搜索迅速找到最优解附近时，遗传算法在局部搜索空间不具备微调能力，无法精确确定最优解的位置[5]。为充分发挥 GA 和 BP 的优势，对 GA 训练后的网络使用经典 BP 算法再进行局部寻优，以精确地收敛到最优解。

4 仿真实验与结果分析

实验中用到的混沌时间序列由 Mackey-Galss 时滞微分系统产生。自 1977 年 Mackey 和 Glass 发现时滞系统中的混沌现象以来，时滞混沌系统常常作为检验非线性系统模型性能的标准，对 Mackey-Glass 方程产生的混沌序列进行预测是很多学者进行研究和报告的基准问题。Mackey-Glass 时滞微分方程定义如下[6]：

$$\frac{dx(t)}{dt} = \frac{0.2x(t-\tau_0)}{1+x(t-\tau_0)^{10}} - 0.1x(t) \quad (6)$$

其中，时滞参数 $\tau_0 = 17$ 时呈现混沌特性， τ_0 值越大，混沌程度越高。选择 $\tau_0 = 17$ ，初始值 $x(t)|_{t=0} = 0.1$ ，采用一阶 Euler 算法以步长 1.0 求解 Mackey-Glass 时滞微分方程，从开始的 2 000 点数据后取 3 000 个点作为运算数据，其中，前 1 000 个数据作为训练数据；后 2 000 个数据作为测试数据。仿真中对训练数据和测试数据分别采用均方误差(MSE)和相对误差(Perr)作为评测标准。

预测之前先对 Mackey-Glass 序列进行归一化预处理，利用归一化后的数据重构 Mackey-Glass 系统的二维相平面图，如图 1 所示。这是由 1 步和 17 步延迟数据构造的简单二维相平面图。由于时滞混沌系统理论上无穷维的，无法使用相空间重构理论精细地重塑吸引子形状，因此无法确定用于预测的神经网络输入节点数目及网络结构。本文构造如下用于预测的数据结构： $[x(k+6); x(k), x(k-6), x(k-12), x(k-18)]$ ，其中， $[x(k), x(k-6), x(k-12), x(k-18)]$ 是 BP 网络输入； $x(k+6)$ 是网络预测输出数据，即实施多步预测。

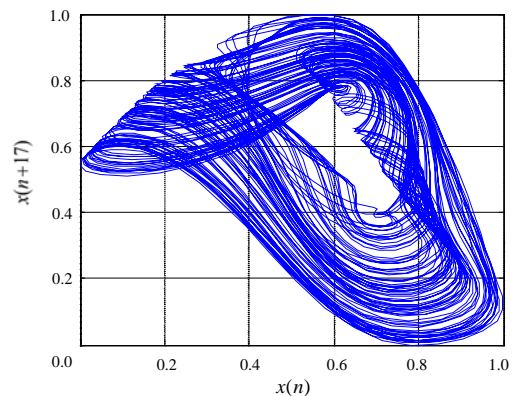


图 1 归一化后重构 Mackey-Glass 二维相平面图

首先利用剪枝法(WE 算法)对混沌时间序列进行预测，设置初始 BP 网络的输入层有 4 个节点，输出层只有 1 个节点，假定开始时隐含层有 30 个隐节点，隐层使用双曲正切 S 型函数，输出层采用线性函数，学习速率为 0.000 02，权消除阈值为 0.003，正则化系数微变量为 $1e-7$ ，进行 40 000 次批处理训练。

图 2 是某次训练后的剪枝网络对测试数据进行预测所得的预测误差曲线，这时网络只有 12 个隐节点。

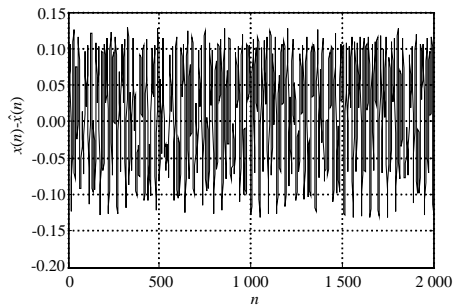


图2 WE算法对测试数据预测的误差序列

对该剪枝网络重新利用 GA 算法训练网络权值, 初始种群为 100, 最大世代数为 500; 采用归一化优先选择(0.09), 算术交叉(2)和非均匀变异(2 gen 3), 利用训练结束后的网络对混沌时间序列进行预测, 结果如图 3 所示。

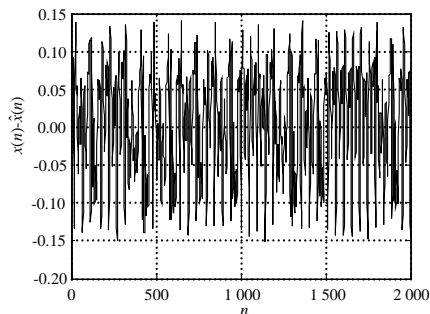


图3 GA算法对测试数据预测的误差序列

对 GA 训练后的网络继续利用经典的梯度下降 BP 算法训练, 学习速率为 0.000 02, 仅进行 5 000 次批处理训练, 利用训练结束后的网络对混沌时间序列进行预测, 结果如图 4 所示。

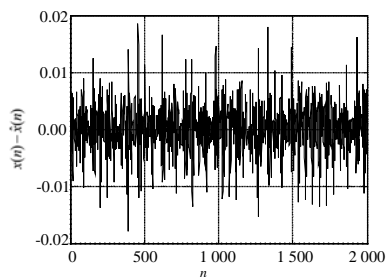


图4 GA_BP算法对测试数据预测的误差序列

最后选用相同的网络结构和传输函数, 赋予随机初始权值, 使用经典的梯度下降 BP 算法进行训练, 性能函数仍为误差平方和(SSE), 学习速率为 0.000 02, 训练 40 000 次, 考察训练后的 BP 网络对训练数据和测试数据的预测性能, 结果如图 5 所示。

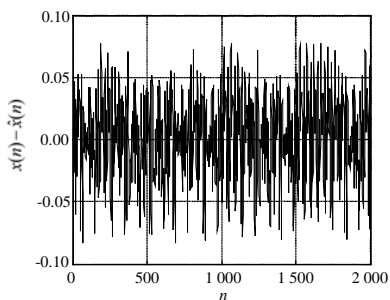


图5 BP算法对测试数据预测的误差序列

上述实验的各项预测统计结果如表 1 所示。

表1 各种方法的预测性能

算法	归一化预测性能			
	误差平方和	均方误差	相对误差	
WE	(训练)	9.096 9	0.009 1	0.023 4
	(测试)	16.992 3	0.008 5	0.022 3
GA	(训练)	3.730 6	0.003 7	0.009 6
	(测试)	6.825 0	0.003 4	0.008 9
GA_BP	(训练)	0.013 4	1.344 1e-005	3.451 6e-005
	(测试)	0.028 9	1.445 3e-005	3.790 4e-005
BP	(训练)	0.870 2	8.701 6e-004	0.002 2
	(测试)	1.657 9	8.289 3e-004	0.002 2

以上实验结果说明, WE 算法能够解决网络隐节点的选择问题, 保证较好的训练效果和泛化性能; WE 算法得到的剪枝网络与相同结构的网络经 BP 训练后的预测结果相比, 性能稍差, 原因是剪枝操作为保证泛化性使网络损失了一定的精度, 因此, WE 算法还需要结合其他方法(如 GA)对得到的网络进行进一步训练。可以看到, GA 算法能够得到一个较好的权值矩阵, 而利用 BP 训练能够进一步改善网络预测性能。另外, 对于相同的训练测试数据, WE 算法每次仿真得到的剪枝网络结构变化较大, 预测统计结果也有变化, 这主要与权消去阈值和初始随机权值的设置有关; 相同结构的 BP 网络的训练结果也有一定变化, 同样与初始随机权值的设置有关。相对于完全使用 BP 算法, 利用遗传算法训练后再使用 BP 算法只需要较少的训练次数就能够达到很高的预测精度。

5 结束语

利用神经网络对混沌时间序列进行预测, 是一种简单有效的方法, 而网络结构及网络权值的确定对网络的预测性能至关重要。针对高维超混沌时间序列预测问题, 本文提出了权消去法与 GA 算法和 BP 网络相结合的预测方法, 在保证一定网络预测泛化性能的基础上降低了网络复杂性, 并且便于利用 GA 算法训练网络权值, 由此提高网络预测的精度, 对典型 Mackey-Glass 混沌序列的预测结果表明了本文方法的有效性。

参考文献

- [1] 张桂英. 基于神经网络的混沌时间序列短期预测[J]. 计算机工程, 2002, 28(11): 197-198.
- [2] Weigend A S, Rumelhart D E, Huberman B A. Back-propagation, Weight-elimination and Time Series Prediction[R]//Proceedings of Connectionist Models Summer School. San Mateo, USA: Morgan Kaufmann, 1990.
- [3] 魏海坤. 神经网络结构设计的理论与方法[M]. 北京: 国防工业出版社, 2005.
- [4] Krogh A, Hertz J A. A Simple Weight Decay Can Improve Generalization[C]//Proc. of Advances in Neural Information Processing Systems. San Matteo, USA: Morgan Kaufmann, 1992.
- [5] Houck C R, Joines J A, Kay M G. A Genetic Algorithm for Function Optimization: A Matlab Implementation. North Carolina State University[EB/OL]. [2007-06-20]. <http://www.ie.ncsu.edu/mirage/GAToolBox/gaot/gaotv5.zip>.
- [6] 贺涛, 周正欧. 基于分形自仿射的混沌时间序列预测[J]. 物理学报, 2007, 56(2): 693-699.