第 19 卷 第 4 期 控 制 与 决 策 2004 年 4 月
Vol 19 No. 4 Control and Decision Apr. 2004

文章编号: 1001-0920(2004)04-0462-03

前馈神经网络的混沌 BP 混合学习算法

李祥飞, 邹 恩, 邹莉华
(株洲工学院 电气工程系, 湖南 株洲 412008)

摘 要: 简要分析由Logistic 映射产生的混沌数以及不同混沌序列之间的概率统计特性, 为混沌全局性搜索提供了依据. 将一种快速BP 算法与混沌优化相结合, 提出了混沌BP 混合算法. 由于混沌Logistic 映射的遍历性、随机性, 使得混合算法收敛速度快, 且具有全局性. 采用混合算法对 XOR 问题和非线性函数进行仿真, 结果表明该算法明显优于标准BP 算法和快速BP 算法.

关键词: 前馈神经网络; 混沌优化; BP 算法; 遍历性
中图分类号: TP11 **文献标识码**: A

Chaos BP hybrid learning algorithm for feedforward neural network

LI Xiang-fei, ZOU En, ZOU Li-hua
(Department of Electrical Engineering, Zhuzhou Institute of Technology, Zhuzhou 412000, China. Correspondent: LI Xiang-fei, E-mail: lixiangfei2002@sina.com)

Abstract: Probabilistic properties are analyzed for chaotic data and different chaotic sequences generated by Logistic map, which provides theoretical basis for chaos global searching. A chaos-BP hybrid algorithm is proposed by means of combination of a new fast BP algorithm and chaos optimization searching. Due to ergodicity and random of chaotic Logistic map, chaos-BP algorithm converges fast and globally, and has no local minimum. The algorithm is applied to XOR problem and nonlinear function approximation. Simulation results show that the chaos-BP algorithm needs shorter learning time than that of the standard BP and fast BP.

Key words: feedforward neural network; chaos optimization; BP algorithm; ergodicity

1 引 言

多层前馈神经网络是研究得最广泛、最成功的一种人工神经网络模型, 这主要是因为BP 算法为前馈神经网络提供了切实可行的学习算法. 但是BP 算法的学习速度慢、收敛时间长、易陷入局部极小, 这些缺陷是BP 算法本身所固有的. 为了解决神经网络的全局性学习, 一些学者将遗传算法和模拟退火算法引入BP 算法[1~3]. 但是这些随机优化算法中的一些参数难以确定, 往往依赖实际经验, 这便带来了很大的盲目性. 混沌是存在于非线性系统中的一种较为普遍的现象, 混沌运动具有遍历性、随机性等特点, 它能在一定范围内按其自身规律不重复地遍历所有状态. 混沌动力学在智能控制领域中的应用已引起人们的重视[4,5], 文献[4]将混沌机制引入BP 算法, 利用混沌的全局遍历性逃逸出权参数的局部极小点, 以获得更优意义的全局网络, 取得了较好的尝试. 本文简要讨论了由Logistic 映射产生的混沌数之间以及混沌序列之间的相关概率统计, 为混沌

收稿日期: 2003-02-23; **修回日期**: 2003-05-15.
基金项目: 湖南省自然科学基金资助项目(01JJY3029).
作者简介: 李祥飞(1969—), 男, 湖南汨罗人, 博士, 从事神经网络、混沌优化和最优控制等研究; 邹恩(1956—), 女, 湖南株洲人, 教授, 从事混沌控制、神经网络等研究.

全局优化提供了理论支持. 在文献[6]提出的快速BP 算法(FBP)的基础上, 结合混沌全局优化设计了一种混沌BP 混合算法(CBP), 以前者梯度下降搜索算法快速收敛到局部最优, 同时依赖后者混沌全局优化摆脱局部极小, 最终收敛到全局最优.

2　混沌 BP 混合算法

Logistic 映射 $x_{n+1} = f(x_n) = \mu x_n(1 - x_n)$; $n = 1, 2, \cdots, N$; $0 < x_0 < 1$; $\mu = 4$. 利用文献[7]的方法讨论由Logistic 映射产生的混沌序列的协方差函数和互协方差函数.

Logistic 映射所产生的混沌数的概率分布密度函数为[7]

$$\rho(x) = \begin{cases} \dfrac{1}{\pi\sqrt{x(1-x)}}, & x \in (0,1); \\ 0, & x \notin (0,1). \end{cases} \tag{1}$$

设

$$f(x) = 4x(1-x), f^{(n)}(x) = \underbrace{f(f(\cdots f(x)))}_{n}$$

混沌序列中的任意混沌数 x_i 与 x_{i+n} 之间的协方差函数为

$$C_x = \lim_{N\to\infty}\frac{1}{N}\sum_{i=0}^{N-1}x_i x_{i+n} - (Ex)^2 = \int x f^{(n)}(x)\rho(x)\mathrm{d}x - (Ex)^2 = 0. \tag{2}$$

任意不同混沌序列 $x_{1,i}$ 与 $x_{2,(i+n)}$ 之间的互协方差函数为

$$C_{x_1,x_2} = \lim_{N\to\infty}\frac{1}{N}\sum_{i=0}^{N-1}x_{1,i}x_{2,(i+n)} - (Ex)^2 = \iint x_1 f^{(n)}(x_2)\rho(x_1)\rho(x_2)\mathrm{d}x_1\mathrm{d}x_2 - (Ex)^2 = 0. \tag{3}$$

混沌序列中不同混沌数的协方差函数值为零表明它们之间的线性联系不密切, 这样的混沌序列曲线的变化既起伏大又很不规则, 使得混沌序列尽可能遍历一维空间的各个区间, 有利于混沌的全局寻优; 同时不同混沌序列的互协方差函数值为零表明它们之间的线性联系也不密切, 这将使得多个混沌序列尽可能遍历多维空间的各个区域, 有利于实现多变量的多维空间全局寻优. 从而使得混沌优化搜索具有全局性.

本文采用文献[6]提出的一种快速BP 算法(FBP), 误差信息在均方误差基础上吸收非线性误差信息来改善收敛速度. 因为3 层前馈神经网络能以任意精度逼近有界非线性函数, 所以采用网络结构为 M-N-L 的3 层前馈网络, 网络的激活函数取 $f(x) = 1/(1+e^{-x})$. 为简单起见, 这里不考虑神经元阈值. 下面给出3 层前馈网络的FBP 算法

$$E = \frac{1}{2}\sum_{t=1}^{L}(e_{1t})^2 + \frac{1}{2}\sum_{t=1}^{L}\lambda(e_{2t})^2. \tag{4}$$

其中: $e_{1t} = d_t - y_t$, $e_{2t} = f^{-1}(d_t) - u_t$, d_t 和 y_t 分别为神经网络的理想输出和实际输出, u_t 为输出层神经元的实际输入.

设 c_j 为中间神经元的输出, 由中间层至输出层权值 V_{jt} 的调整量为

$$\Delta V_{jt} = \mu f'\left(\sum_{j=1}^{N}V_{jt}c_j\right)e_{1t}c_j + \mu\lambda c_j e_{2t}. \tag{5}$$

设 a_i 为神经网络的输入, 则由输入层至中间层权值 W_{ij} 的调整量为

$$\Delta W_{ij} = \mu f'\left(\sum_{i=1}^{M}W_{ij}a_i\right)\hat{e}_{1j}a_i + \mu\lambda a_i\hat{e}_{2j}. \tag{6}$$

其中

$$\hat{e}_{1j} = \sum_{t=1}^{L}e_{1t}f'\left(\sum_{j=1}^{N}V_{jt}c_j\right)V_{jt},$$

$$\hat{e}_{2j} = f'\left(\sum_{i=1}^{M}W_{ij}a_i\right)\sum_{t=1}^{L}e_{2t}V_{jt}.$$

尽管改进的FBP 算法提高了收敛速度, 但由于采用梯度信息, 局部极小问题仍不可避免. 而混沌具有遍历性和随机性, 结合混沌优化搜索来摆脱局部极小可实现前馈网络的全局学习. 因此, 本文提出混沌BP 混合算法. 由于混沌动力系统具有对初始值极其敏感的特点, 取不同的初始值代入Logistic 映射, 便可得到不同轨迹的混沌变量.

混沌BP 混合算法具体步骤如下:

Step1: 在(-1, 1) 区间产生随机数, 赋前馈网络的初始权值.

Step2: 采用FBP 算法对神经网络权参数进行学习, 计算网络误差 $\epsilon = \sum_{k=1}^{n}|d_k - y_k|$, 记为 E_{FBP}, 其中 n 为样本数. 若学习后满足精度要求, 此时前馈网络权值向量记为 W, 算法结束; 否则转 Step3.

Step3: 进入如下混沌优化搜索:

1) 在(0, 1) 区间随机初始化混沌向量 X_0, 且与权值向量 W 的维数一致, 并置 $k = 0$;

2) $X_{k+1} = 4X_k(1 - X_k)$, $k = k + 1$;

3) 对混沌变量 X_k 进行平移、放大, 产生新权值向量 $\hat{W} = W + |W|\rho(2X_k - 1)$, ρ 为常数, 使新权值向量 $\hat{W}$ 在以 W 为中心, $\rho|W|$ 为半径的 $(W - \rho|W|, W + \rho|W|)$ 区间遍历, 计算神经网络误差 ϵ, 并记为 E_{CBP};

4) 若网络误差 E_{CBP} 满足精度要求, 算法结束;

否则, 转 5);

5) 比较 E_{CBP} 和 E_{FBP} 的大小, 若 $E_{CBP} < E_{FBP}$, 则令 $W = \hat{W}$, 转 Step2; 否则, 转 2).

注 1 在网络误差不满足精度要求的情况下, 而混沌优化算法又无法使BP 算法跳出局部极小, 可将 ρ 适当地选大一些, 使权值向量 $\hat{W}$ 在较大区域遍历, 搜索出满足 $E_{CBP} < E_{FBP}$ 的权值向量 $\hat{W}$.

注 2 快速BP 算法的每步迭代都确保误差函数 $E(W)$ 单调递减, 而在混沌优化中每次都保留了使误差函数不增的权向量, 寻找满足 $E_{CBP} < E_{FBP}$ 的权参数向量 W, 帮助 E_{FBP} 跳出局部极小, 保证了混沌BP 混合算法的收敛性.

3 数值算例

为了测试本文提出的CBP 算法的有效性, 作者采用Matlab 5.3 软件编程, 并以 XOR 问题和非线性函数逼近作为仿真算例, 将本文提出的CBP 算法与 SBP 算法、FBP 算法进行比较, 结果如表 1 和表 2 所示. 在本算例中, 参数 $\mu = 0.5$, $\lambda = 0.55$, $\rho = 2.5$. CBP 算法中混沌优化时的CPU 所用时间通过采用Matlab 中的cputime 函数折算成SBP 算法的训练次数来表示.

例 1 XOR 问题

XOR 问题是具有多个极值点的典型非线性问题, 本文选取一个网络结构为 2-2-1 的 3 层前馈网络, 分别采用 3 种算法进行训练, 结果见表 1. 仿真结果表明, 由于前馈网络的权初始值选择不合适, 使得标准BP 算法(SBP 算法)、FBP 算法无法使样本误差控制在精度内, 即陷入局部极小, 而 CBP 算法却能得到较高的误差精度, 避免了陷入局部极小.

表 1 XOR 问题的仿真结果

精度 ϵ	SBP 算法	FBP 算法	CBP 算法
0.840	1 614	737	256
0.150	∞	∞	641
0.041	∞	∞	1 783

例 2 非线性函数逼近 $y = (\sin x_1 + e x_2)/15$. 当 $x_1 \in (0, 20)$ 和 $x_2 \in (0, 2)$ 具有多个极值点时, 在此区间选取训练样本数 15, 选取前馈网络结构为 2-5-1 的 3 层网络, 采用 3 种算法进行训练, 结果如表 2 所示. 从表 2 可以看出, 只有CBP 算法能有效地避免局部极小, 快速达到较高的精度.

表 2 非线性函数逼近问题的仿真结果

精度 ϵ	SBP 算法	FBP 算法	CBP 算法
0.500	4 857	2 892	786
0.170	∞	∞	2 161

4 结 语

本文简要分析了由Logistic 映射产生的混沌数之间以及不同混沌序列之间的相关概率统计特性, 为实现混沌全局优化提供了理论支持, 同时提出一种基于 FBP 算法和混沌优化相结合的CBP 算法. 由于 FBP 算法吸收了误差函数的非线性信息, 明显地提高了网络学习的收敛速度. 而混沌优化方法具有全局性, 在一定程度上避免了陷入局部极小.

参考文献(References):

[1] Carlos Calderon-Macias, Sen Mrinal K, Stoffa Paul L S. Artificial neural networks for parameter estimation in geophysics[J]. *Geophysical Prospecting*, 2000, 48(1): 21-47.

[2] Greeman Roxama M, Stepniewski Slawomir W, Jorgensen Charles C, et al. Designing compact feedforward neural network models with small training data sets[J]. *J of Aircraft*, 2002, 39(3): 452-459.

[3] Treadgold Nicholas K, Gedeon Tamas D. Simulated annealing and weight decay in adaptive learning: The sarprop algorithm[J]. *IEEE Trans on Neural Networks*, 1998, 9(4): 662-668.

[4] 李翔, 陈增强, 袁著祉. 混沌机制在 T-S 模型模糊神经网络的系统辨识研究[J]. 控制与决策, 2001, 16(4): 504-506.
(Li Xiang, Chen Zeng-qiang, Yuan Zhu-zhi. Study on chaotic mechanism in system identification using T-S model fuzzy neural networks[J]. *Control and Decision*, 2001, 16(4): 504-506.)

[5] 潘永湘, 徐前锋, 高红梅. 基于混沌思维的模糊控制算法优化研究[J]. 控制理论与应用, 2000, 17(5): 703-706.
(Pan Yong-xiang, Xu Qian-feng, Gao Hong-mei. The research of the fuzzy control algorithm optimization based on chaos[J]. *Control theory and applications*, 2000, 17(5): 703-706.)

[6] Abid S, Fnaiech F, Najim M. A fast feedforward training algorithm using a modified form of the standard backpropagation algorithm[J]. *IEEE Trans on Neural Networks*, 2001, 12(3): 424-430.

[7] 王亥, 胡健栋. 改进型Logistic-Map 混沌扩频序列[J]. 通信学报, 1997, 18(8): 71-77.
(Wang Hai, Hu Jian-dong. The improved logistic map chaotic spread spectrum sequences[J]. *J of China Institutes of Communications*, 1997, 18(8): 71-77.)