

基于混沌优化的多层前馈神经网络

蒋 凭

(杭州市西湖区人民检察院, 浙江 杭州 310013)

摘要: BP 算法是应用广泛的一种多层前馈神经网络模型, 针对算法求解精度低、搜索速度慢、易于陷入局部极值点等问题, 根据混沌理论的全局优化思想, 提出采用“多次载波”技术将混沌优化和前馈神经网络相结合, 利用已找到的近似最优解来启发搜索全局最优解的方法训练神经网络, 以布尔函数识别、曲线逼近、模式识别 3 个典型应用对算法进行验证。研究结果表明, 算法具有较好的泛化能力和快速全局收敛的性能, 特别是针对中小规模的网络, 混沌优化算法在训练时间、全局收敛率等指标方面优于 BP 算法。

关键词: 混沌优化算法; 神经网络; 多次载波

中图分类号: TP391

文献标识码: A

文章编号: 1001-4551(2010)10-0109-03

Multi-layer feed-forward neural network based on chaos optimization algorithm

JIANG Ping

(Hangzhou West Lake District Prosecution Office, Hangzhou 310013, China)

Abstract: BP algorithm is widely used as a multi-layer feed-forward neural network model. Aiming at low precision, long search time, prone to local optimum problems, and according to chaos theory of global optimization idea, chaos optimization and combination of feed-forward neural network based on “multiple-carrier” technology was proposed. Using the approximate optimal result has been found in heuristic search method of global optimization to train the neural network, boolean function recognition, curve approximation, pattern recognition were validated. The results indicate that the algorithm has better generalization ability and fast global convergence of performance, especially for small and medium-scale networks, chaos optimization algorithm in training time, the global convergence rate are better than BP algorithm.

Key words: chaos optimization algorithm; neural networks; multiple-carrier

0 引 言

多层前馈神经网络具有较强的非线性映射、函数逼近和大规模并行分布处理能力, 已广泛应用于系统建模、模式识别、特征压缩等领域。作为一种具有学习能力的智能系统神经网络, 学习算法是其重要的组成部分, 并且一直是神经网络的研究重点。反向传播算法(BP)算法是被广泛应用的学习算法^[1-3], 该算法学习过程的目的是使误差能量函数值降低到给定的精度, 但由于误差函数的高维复杂性和局部极值的存在性, 使以梯度下降为基础的 BP 学习算法存在难以克服的缺点, 如学习过程易陷入局部极值点、学习算法的收敛速度慢、学习过程易出现振荡现象等^[4]。

混沌是存在于非线性系统中的一种较为普遍的现象, 混沌并不是一片混乱, 而是有着精致内在结构的一类现象。混沌运动具有遍历性、随机性、“规律性”等特点, 混沌运动能在一定范围内按其自身的“规律”不重复地遍历所有状态^[5-8]。因此, 将混沌应用于非线性优化问题全局最优的求解引起了人们的广泛重视, 并取得了较好的成果。

本研究将混沌优化算法^[9]和神经网络结合起来, 可兼有神经网络的广泛映射能力和混沌算法的快速、全局收敛等特点, 在某种程度上避免了神经网络收敛速度慢、易于陷入局部极小点的问题。利用混沌优化算法学习神经网络的权值, 将二者结合起来可解决非线性模型的辨识问题及模式识别、分类等问题。

1 基于混沌优化算法的神经网络训练

大多数神经网络都使用 BP 算法进行训练,这种算法具有收敛时间长、易于陷入局部极值等缺陷。混沌优化算法利用混沌变量的遍历性、随机性、“规律性”来进行全局优化,因此用它来训练神经网络的权值,可避免 BP 算法的缺陷。

算法的基本思想是:设多层前馈神经网络共有 L 层,第 L 层神经元的个数为 N_L ,第 l 层与第 $l+1$ 层的连接权为 $W^{(l)}$,则需优化的变量为权值向量。设用 P 个样本训练网络,并设第 p 个样本的输出为 Y_p ,则优化的目标函数可定义为:

$$\min (f(W)) = \min \left(\frac{1}{2} \sum_{p=1}^P \sum_{k=1}^{N_L} (d_{pk} - y_{pk})^2 \right) \quad (1)$$

为了提高效率,在搜索过程中采用“多次载波”技术,利用已找到的近似最优解来启发搜索全局最优解。某一次载波经过一段搜索后可以很快找出一个近似最优解,它往往处于真正最优解的区域内。然后,再在近似最优解的基础上再次载波,再次载波以速率 α 缩小搜索空间,相当于在近似最优解的领域内进行细搜索,这样就可以很快找到全局最优解,大大提高搜索速度。这一过程反复进行,当达到给定的目标函数精度,或达到给定的迭代次数时优化结束。

混沌优化算法的主要步骤如下:

Step 1:首先选择用于载波的混沌变量。本研究选用式(2)所示的 Logistic 映射,其中 μ 是控制参数,取 $\mu = 4$ 。设 $0 \leq x_i \leq 1, i \leq n$,此时系统(2)完全处于混沌状态。利用混沌对初值敏感的特点,赋给式(2) n 个微小差异的初值后即可得到 n 个混沌变量。

$$x_i = \mu x_i (1 - x_i) \quad (2)$$

Step 2:算法初始化。 $k=1$,将 f^* 设置为一个较大的数值。

Step 3:根据式(3)把 x_i 映射到优化变量的取值区间成为 x_i' ,其中 a_i 和 b_i 分别为变量取值的上、下限。

$$x_i' = a_i + x_i (b_i - a_i) \quad (3)$$

Step 4:用混沌变量进行迭代搜索。令 $w_i = x_i'$,得向量 W_k ,计算式(1)中 $f(W)$ 的值得 $f(W_k)$ 。If $f(W_k) \leq f^*$ then $f^* = f(W_k), W^* = W_k$; Else if $f(W_k) > f^*$ then 放弃 W_k 。

Step 5: $k = k + 1$ 。

Step 6:如果经过 Step 4 和 Step 5 的若干步搜索 f^* 都保持不变,则进行以下步骤。

Step 7:计算新的搜索范围。根据式(4)和式(5)

确定新的搜索范围 $[a_i', b_i']$ 。

$$a_i' = w_i - \alpha (b_i - a_i) \quad (4)$$

$$b_i' = w_i + \alpha (b_i - a_i) \quad (5)$$

式中, $\alpha \in (0, 0.5)$, $W = (w_{01}^{(1)}, w_{02}^{(1)}, \dots, w_{N_{L-1}N_L}^{(1)})$, 向量长度 $n = (N_1 + 1)N_2 + (N_2 + 1)N_3 + \dots + (N_{L-1} + 1)N_L$ 。产生 n 个用于载波的混沌变量,选用 Logistic^[6] 进行映射。

Step 8:对新的搜索范围进行越界检查。If $a_i' > a_i$ then $a_i = a_i'$; Else a_i 保持不变。If $b_i' < b_i$ then $b_i = b_i'$; Else b_i 保持不变。

Step 9:继续执行 Step 3 ~ Step 8 的迭代搜索。直到 f^* 的值达到允许的网络误差精度或达到给定的迭代次数。

2 计算机仿真实验

为了验证利用混沌优化算法训练神经网络的可行性,扩展这一方法的实际工程应用,进行了 XOR 布尔函数识别、正弦函数逼近及随机模式识别实验。

2.1 XOR 布尔函数识别

XOR 布尔函数经常用于测试神经网络的非线性识别能力^[10]。前馈神经网络的拓扑结构取 2-2-1,即输入层、中间层为 3 个节点,输出层为一个节点,隐层和输出层的激发函数为非线性 Sigmoid 函数,权值向量 W 的长度 $n=9$ 。分别用 BP 算法及混沌优化算法进行 20 次随机选取初值的重复学习,最大学习次数为 1 000 次,误差函数的精度为 $E \leq 0.01$,其中混沌优化算法权值的区域为 $[-10, 10]$,空间缩小率 $\alpha = 0.2$ 。对训练结果取平均值,结果如表 1 所示。

表 1 混沌优化算法和 BP 学习算法的计算结果对比

算法	平均学习次数	运行时间/s	收敛率
BP 算法	230	2	85%
混沌优化算法	800	1.5	96%

由表 1 可知,采用混沌优化训练后,对于布尔函数的识别,学习次数相比 BP 算法有所增加,但是总的训练时间缩短 25%,收敛率提高 11%,算法性能提高显著。

2.2 曲线逼近问题

对三层前馈神经网络进行学习来逼近非线性函数 $f(x) = e^{-x} \sin(2\pi x)$ 。输入样本区间为 $[0, 1]$,步长为 0.1,则共有 11 对输入输出样本对。神经网络结构中输入、输出层为 1 个节点,隐层为 4 个节点,隐层和输出层的激发函数为非线性 Sigmoid 函数,权值向量 W 的长度 $n=13$ 。分别用 BP 算法及混沌优化算法进行 10 次随机选取初值的重复学习,最大学习次数为

2 000 次,误差函数的精度为 $E \leq 0.01$,其中混沌优化算法权值的区域为 $[-5, 5]$,空间缩小率 $\alpha = 0.1$ 。通过混沌优化算法训练得到的神经网络对曲线的逼近效果如图 1 所示,实线代表样本输出,虚线代表实际输出,网络的实际输出与样本输出基本吻合。混沌优化算法与 BP 算法之间的性能比较如表 2 所示。

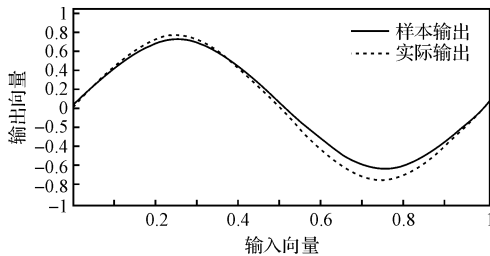


图 1 基于混沌优化算法的神经网络泛化性能

表 2 曲线逼近问题的学习结果

算法	平均学习次数	运行时间/s	收敛率
BP 算法	500	4	78%
混沌优化算法	1 500	3.8	97%

由图 1 可知,曲线逼近问题上,混沌优化算法的泛化性能表现良好;根据表 2 可知,相对 BP 算法而言,运行时间相差不大,但是收敛率提升明显。

2.3 随机模式识别问题

该问题的目的是建立 5 维输入与输出向量之间的映射关系,其中向量的每个分量为区间 $[0.05, 0.95]$ 上的随机数。6 个这样的样本如表 3 所示。神经网络结构中输入层和输出层都为 5 个节点、隐层为 6 个节点,隐层和输出层的激发函数为非线性 Sigmoid 函数,权值向量 W 的长度 $n = 71$,网络的误差函数精度为 $E \leq 0.05$ 。分别用 BP 算法及混沌优化算法进行训练,最大迭代次数为 5 000 次,其中混沌优化算法权值的区域为 $[-20, 20]$,空间缩小率 $\alpha = 0.3$ 。混沌优化算法与 BP 算法对随机映射关系问题的性能比较如表 4 所示。

表 3 输入和输出向量之间为随机映射的关系问题

序号	输入向量	输出向量
1	0.05 0.95 0.05 0.7 0.7	0.4 0.6 0.8 0.95 0.6
2	0.7 0.6 0.3 0.2 0.8	0.05 0.2 0.6 0.9 0.05
3	0.6 0.3 0.1 0.7 0.1	0.94 0.2 0.1 0.3 0.8
4	0.5 0.4 0.3 0.05 0.7	0.95 0.9 0.95 0.8 0.8
5	0.9 0.6 0.6 0.5 0.9	0.5 0.3 0.3 0.1 0.1
6	0.2 0.5 0.8 0.95 0.05	0.3 0.5 0.6 0.8 0.7

表 4 随机模式识别问题的学习结果

算法	平均学习次数	运行时间/s	收敛率
BP 算法	1 800	83	70%
混沌优化算法	3 700	81	92%

根据表 4 可知,在网络规模增大、学习次数增多的情况下,两种算法的运行时间接近一致,但是混沌优化算法收敛率明显高于 BP 算法。

3 结束语

从以上实验可知:①虽然混沌优化算法的迭代次数多于 BP 算法,但由于不需要进行反向计算,总的运行时间小于 BP 算法。②由于混沌优化算法借助了混沌变量“有规则的遍历性”,算法较易跳出局部极值点,网络训练的收敛率远远高于 BP 算法。③随着网络规模的增大,混沌优化算法由于受其遍历特性的限制,运行时间延长。若网络规模进一步扩大,算法效率将会有所下降,但是明显优于 BP 算法。

总而言之,对于中小型规模的神经网络而言,混沌优化算法具有运算效率高、不易陷入局部极值及泛化性能突出等优点,具有广阔的工程应用前景。

参考文献 (References):

- [1] 刘筱玲,宋顺成,史洪刚.用神经网络 BP 算法预测钨合金材料抗拉强度[J].材料科学与工艺,2006,14(1):63-65.
- [2] 李晓峰,刘光中.人工神经网络 BP 算法的改进及其应用[J].四川大学学报:工程科学版,2000,32(2):105-109.
- [3] 房振勇,游文虎,冯汝鹏.改进 BP 算法在模糊神经网络中的应用[J].北京航空航天大学学报,2007,33(11):1321-1324.
- [4] HAGAN M T, DEMUTH H B. Neural Network Design[M]. PWS Publishing Company,1996.
- [5] 卢 侃.混沌动力学[M].上海:上海翻译出版社,1990.
- [6] OTT E, GREBOGI C, YORKE J A. Controlling chaos[J]. Physical Review Letters,1990,64(11):1196-1199.
- [7] SHINBROT T, GREBOGI C. Using small perturbations to control chaos[J]. Nature,1993,363(3):411-417.
- [8] 李 兵,蒋慰孙.混沌优化方法及其应用[J].控制理论与应用,1997,14(4):613-615.
- [9] 梁久祯.智能计算-若干理论问题及其应用[M].北京:国防工业出版社,2007.
- [10] YAN Yi, ZHANG Hang-ping, ZHOU Bin. A New Learning Algorithm for Neural Networks with Integer Weights and Quantised Non-linear Activation Functions[C]//IFIP International Conference on Artificial Intelligence in Theory and Practice(IFIP AI 2008). Milano: [s. n.],2008:[s. n.].