

一种提高 BP 神经网络泛化能力的改进算法

夏 玫, 陈立潮, 王新波

(太原科技大学 计算机科学与技术学院, 山西 太原 030024)

摘 要:针对 BP 神经网络中样本数据复杂和容易陷入到局部极值的不足, 通过利用主成分分析法对样本数据进行简化, 并采用遗传算法优化神经网络的初始权值, 提出了一种改进 BP 神经网络泛化能力的算法。该算法兼顾了主成分分析法剔除冗余的特性和遗传算法能够全局优化的特点, 既具备了较快的收敛速度又不易陷入到局部极值中, 具有较强的泛化能力。实验结果表明, 改进后的算法在数据降维、冗余信息的剔除、预测能力等方面都得到了一定程度上的改善。

关键词:BP 神经网络; 泛化能力; 主成分分析; 遗传算法

中图分类号: TP389.1

文献标识码: A

文章编号: 1673-629X(2009)09-0062-03

A Modified Algorithm to Improve Generalization Ability of BP Neural Network

XIA Mei, CHEN Li-chao, WANG Xin-bo

(Institute of Computer Science and Technology, Taiyuan University of Science and
Technology, Taiyuan 030024, China)

Abstract: To solve the complexity of the sample data of BP neural network and easily getting into a local value, an improved algorithm is proposed in this paper. This algorithm uses the principle components analysis to predigest the sample and the genetic algorithms to find the best point, and it combines the character of PCA to reduce redundancy with that of GA to optimize the whole, so it can not only pick up convergence and can avoid to get into a local value. With the result of the experiment of the improved algorithm of BP neural network, it is valid to reduce the dimensions of data, delete redundancies, and increase forecast ability.

Key words: BP neural network; generalization; principle components analysis; genetic algorithms

0 引 言

自 20 世纪 80 年代以来, 人们对神经网络的研究一直热情不减, BP 神经网络因其自身结构简单、易于学习而得到了广泛的应用。在利用神经网络解决问题时, 通常会给出 BP 神经网络的训练样本数据、输出数据和一定的学习算法, 再对其参数进行调整以达到学习样本的目的。利用检验样本集合对这个调整好的网络进行检验, 最终检验 BP 网络的性能。而神经网络的泛化能力即网络识别训练集合以外样本的能力, 被认为是衡量神经网络性能最重要的指标, 没有泛化能力的网络是没有实用价值的^[1,2]。近些年人们对 BP 网络的泛化能力做了一些研究, 包括提高样本的质量、改进学习算法、利用网络的集成等等^[3]。文中利用主

成分分析的方法简化样本, 再利用遗传算法(GA)优化 BP 网络的初始权值, 加快网络的收敛速度, 以达到最终提高 BP 神经网络泛化能力的目的。

1 主成分分析和遗传算法

主成分分析(principle components analysis, 简称 PCA)是统计学中的一个概念, 主要是把分散在一组变量上的信息集中到某几个综合指标(也就是说主成分)上, 每个主成分都是原始变量的线性组合, 主成分之间互为正交关系, 继而降低数据的维数、删减冗余信息, 对数据进行合理的压缩。当样本数据维数较多, 结构复杂的时候, 采用主成分分析的方法可以简化输入样本, 减少训练时间, 提高训练的效率, 达到提高神经网络泛化能力的目的^[4,5]。

遗传算法(Genetic Algorithms)是在生物进化论的背景下产生和发展的一种高效的随机搜索与优化的方法, 作为一种新的全局优化搜索算法, 以其简单通用、鲁棒性强、适于并行处理等显著特点奠定了它作为 21

收稿日期: 2008-12-15; 修回日期: 2009-03-02

基金项目: 山西省科技攻关项目资助(20080321012-01)

作者简介: 夏 玫(1980-), 女, 内蒙古呼伦贝尔人, 硕士研究生, CCF 会员, 研究领域为智能软件技术; 陈立潮, 博士, 教授, 研究领域为模式识别、数据挖掘、数据库与智能软件技术。

世纪关键智能算法之一的地位^[6]。它的搜索不依赖梯度信息,而是利用群体搜索策略和群体中个体之间的信息交换。遗传算法具有较好的全局最优求解能力,能够有效地跳出局部极值;因为对目标函数的性态没有明显的要求,因此,遗传算法也有良好的适应性;此外,遗传算法的搜索效率比较高,具有很高的并行性,并且有一定的扩充能力。基于以上的优点,遗传算法通常会在神经网络中连接权的进化、网络结构的进化、学习规则的进化等方面得到广泛的应用。

2 BP神经网络算法的改进

由于BP网络有着强大的非线性映射功能,并具有一定的推广、概括能力,而被广泛应用于函数逼近、模式识别/分类、数据压缩等领域,并取得了很好的效果。但在BP神经网络被广泛应用的过程中,人们也发现了BP神经网络自身有两个不可回避的不足:一是收敛速度较慢;二是容易陷入到局部极值中。由于BP神经网络对各种参数比较敏感,稍小的变动就会引起拟合和泛化能力的振荡,所以当网络收敛速度慢时,容易引起过度训练,从而影响到网络的泛化能力。而BP神经网络采用了梯度下降法进行学习,不同的初始权值会导致不同的结果,加之网络在训练过程中存在着多个局部极值点,使得BP神经网络容易陷入到局部极值点,这一点也严重影响到BP神经网络的泛化能力^[7]。

2.1 算法改进的思想

针对网络收敛速度慢和容易陷入到极值中的不足,分别从样本集合和初始权值两方面进行改进。对于样本集合通常会利用归一化方法对它进行压缩,使它的输入样本限定在一定的区间之内,但是这种方法无法对样本数据进行维数的简化和冗余的减少,当遇到样本维数较多的情况,样本的输入变得繁琐,而利用主成分分析方法可以有效避免这种情况。而在网络训练初始权值的选取上,传统的权值获取方法都是随机给定的,很容易陷入局部极值无法逃逸,利用遗传算法在随机点上集中遗传迭代出一个优化值可以有效避免局部极值的陷入^[8,9]。因此改进算法的基本思想为:利用主成分分析的方法对神经网络的样本集合进行压缩、降维和去噪,用特征向量和原始样本乘积的*i*个主成分代替原始样本作为输入样本,然后用遗传算法对随机的初始权值进行优化,最后将降维的样本集合和优化了的初始权值带入,进行BP神经网络训练,并用检验样本集合对这个算法进行检验。

2.2 算法的流程

改进算法的具体流程如下:

Step1:对初始样本集合预处理,包括相关矩阵的计算和相关矩阵的特征值、特征向量。

相关矩阵 *R*:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix} \quad (1)$$

$$\text{其中: } r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}}$$

利用特征方程 $| \lambda I - R | = 0$ 求得特征值 λ_i 和正交单位化特征向量 e_i 。

Step2:计算贡献率和累计贡献率。

第 *j* 个主成分贡献率为:

$$\lambda_j / \sum_{i=1}^p \lambda_i \quad (2)$$

前 *m* 个主成分的累计贡献率为:

$$\sum_{i=1}^m \lambda_i / \sum_{i=1}^p \lambda_i \quad (3)$$

Step3:取主成分累计贡献率占到 80% ~ 90% 的前 *m* 个主成分替代原始样本数据。

Step4:初始化种群,并根据个体评价函数选择网络个体。

采用实数编码方式,设有 *N* 个个体的实数编码种群,它的染色体长度 $L = L1 * R + R * L2 + R + L2$ 。其中 *L1*、*L2* 为BP神经网络输入、输出数据向量的维数,*R* 为常数。随机初始化之后,选取其中性能好的作为初始种群。设定适应函数为 $f(i) = 1/E(i)$,其中 $E(i) = \sum_p \sum_k (V_k - T_k)^2$, *N* 为染色体数, *k* 为输出层节点数, *p* 为学习样本数, *T_k* 为教师信号。

Step5:利用概率重新产生新个体,将新个体放入到种群中,并计算新个体的评价函数。

个体的选择概率为:

$$p_s = f_i / \sum_{i=1}^N f_i \quad i = 1, 2, 3, \dots, n \quad (4)$$

当群体中个体适应值的差异非常大时,最优个体与最差个体被选择的几率就会成指数增长,最优个体在下一代生存的几率将大大加强,而最差的个体就没有了生存的机会。

Step6:考察网络误差平方和 *E_i*,如果未达到预期值则返回 Step5,达到预期值则继续 Step7。

Step7:以提取出来的主成分分析值作为输入样本,以GA遗传出的优化初值作为初始权值,训练三层BP神经网络。

Step8:选取多组检验样本对网络进行检验。

3 实验分析

为了验证改进算法的有效性,采用土壤中各种元素的含量来判断土壤适合生长的植被类型和所处位置。

按照植被的类型和所处的位置,将其进行分类,所属类别分别由 1~5 来表示,其中:

植被:弃耕地-1、蔬菜地-2、农田-3、林地-4、草地-5;

位置:工业区-1、住宅区-2、郊区耕地-3、草坪-4、绿化区-5,样本数据如表 1 所示。

表 1 样本数据

序号	Pb	Cu	Ni	As	Cd	植被	位置
1	50.8	35.3	25.2	12.9	0.289	1	1
2	24.1	26.6	23.2	9.1	0.173	2	1
3	20.9	24.1	21	8.2	0.16	1	1
4	27.1	27.3	22.9	8.1	0.176	3	2
5	29.8	34	26.1	10.3	0.194	4	3
6	28	27.5	23.8	8.1	0.211	1	1
7	28.1	27.9	22.9	8.6	0.208	5	4
8	28	29.5	23.7	8.8	0.239	5	4
9	30.4	29.5	23.9	9.5	0.221	2	2
.....

先利用特征方程求得特征值及方差贡献率,再根据主成分分析法取占总体 80%~90% 的前七组主成分分析值,如图 1 前七组的累计贡献率接近 90%,最后将得到的 PCA1,PCA2,...,PCA7 作为输入样本集合。与归一法相比,主成分分析方法在降维方面有一定的优势,为简化输入、加快网络的收敛速度做了充分的准备。

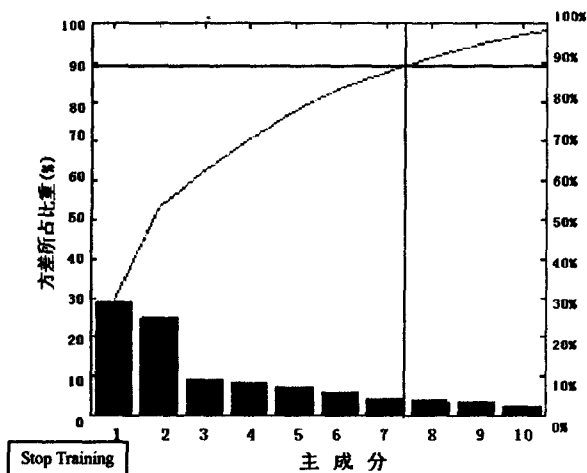
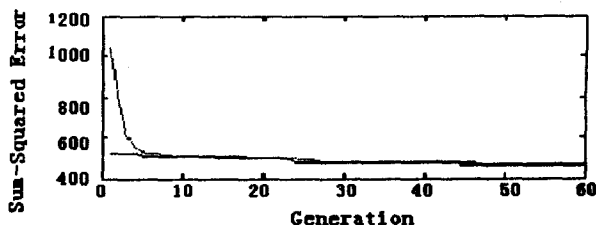


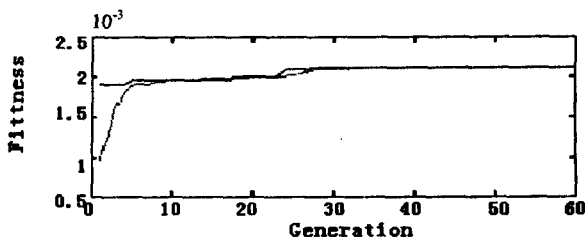
图 1 主成分比例图

确定了输入样本之后,利用遗传算法优化 BP 神经网络的初始权值,采用精度较高的实数编码形式,初始种群数为 30,遗传代数数为 60,按照给定的适应值 $f(i)$ 进行迭代,误差平方和曲线和适应值曲线如图 2

所示,从中可以看到在 50 次迭代之后,染色体的平均适应度基本趋于稳定,60 代的迭代之后得到一个最优的初始权值。



(a) 遗传算法误差平方和曲线



(a) 遗传算法适应值曲线

图 2 遗传算法误差平方和曲线、适应值曲线

将得到的主成分分析值作为训练样本数据,GA 所得到的最优值作为初始权值,建构一个三层 BP 神经网络。隐含层的节点个数取 15,设定误差目标为 $goal=0.02$, $lr=0.01$,进行 BP 神经网络训练,从图 3 看出,经过 13 步收敛后,BP 神经网络达到训练目标,可见收敛速度得到了提高。而与标准 BP 神经网络算法比较所得的结果如表 2,从中可以看出,选取 3 组不同的检验样本,改进 BP 神经网络算法比标准 BP 神经网络算法都有更高的预测准确度,这表明比较标准 BP 神经网络,改进的算法具有更好的泛化能力。

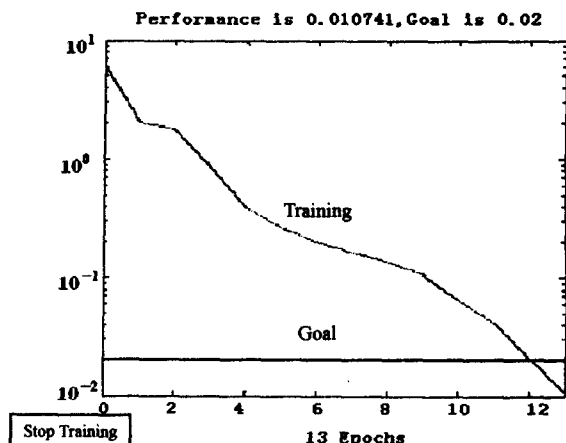


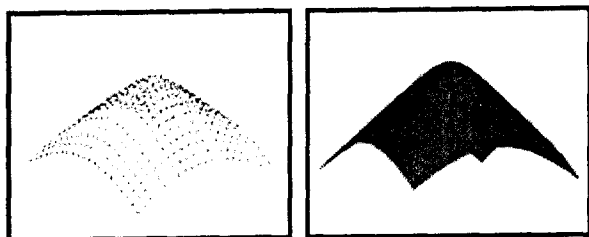
图 3 训练结果

表 2 标准 BP 神经网络、改进 BP 神经网络结果比较

神经网络的类型	输入层节点数	隐层节点数	预测准确率 (3 组检验样本)		
标准 BP 神经网络	12	25	83.33%	60.00%	75.00%
改进 BP 神经网络	7	15	92.86%	80.00%	87.50%

4 仿真结果

图 2 表示用文中方法做仿真实验得到的结果。(a)表示空间中矩形排列的原始数据,共有 441 个三维数据点,是从解析方程 $z = \sqrt{3 \times (1 + x^2/4 + y^2/4)}$ 采样得到的。(b)是用文中方法得到的双三次 NURBS 曲面的光照图。



(a) 原始数据点

(b) 拟合后的 NURBS 曲面

图 2 实验仿真

5 结束语

曲面重建技术是逆向工程中的热点研究课题,广泛应用于航空、汽车、船舶以及其它工业造型领域。NURBS 技术作为国际标准,具有统一、通用、有效的标准算法和相应的配套技术,可以表示自由曲线曲面,因此在重构技术中被广泛采用。而神经网络具有强大的非线性学习能力和一定的抗噪性。文中正是考虑到这些特点提出了一种基于神经网络的 NURBS 曲面重建方法,把参数化后的数据交由神经网络去学习进而拟合出相应的曲面。

仿真实验表明该方法是可行的,对原始数据的拟

合度较高。然而该方法处理的是规则的矩形排列的数据点集,因此在下一步的研究中将考虑大规模散乱数据点集的压缩和神经网络曲面重建。同时神经网络的训练算法也会进一步改进。

参考文献:

- [1] Piegl L, Tiller W. The NURBS Book[M]. Berlin: Springer Verlag, 1997: 353-363.
- [2] 祁伟丽, 秦新强. 基于 B 样条的平面轮廓重构闭合曲面算法[J]. 计算机技术与发展, 2008, 18(7): 112-115.
- [3] Ma Weiyin, Kruth J P. Parameterization of randomly measured points for least squares fitting of B-spline curves and surfaces[J]. Computer Aided Design, 1995(9): 663-675.
- [4] 武剑洁, 王启付, 黄运保, 等. 逆向工程中曲面重建的研究进展[J]. 工程图学报, 2004(2): 133-142.
- [5] Barhak J, Fischer A. Adaptive reconstruction of freeform objects with 3D SOM neural network grids[J]. Computers & Graphics, 2002(26): 745-751.
- [6] 周金字, 谢里阳. 基于神经网络预拟合的 B 样条曲面反求[J]. 东北大学学报: 自然科学版, 2003(6): 556-559.
- [7] 杨广全, 李思益, 曹西京. CAGD 中的 NURBS 曲面——理论与应用[J]. 西北轻工业学院学报, 2001(3): 62-65.
- [8] Cheng Ming-Yang, Wu Hung-Wen, Su A Wen-Yu. On Non-Uniform Rational B-spline Surface Neural Networks[J]. Neural Process Lett, 2008(28): 1-15.
- [9] 韩力群. 人工神经网络理论、设计及应用[M]. 北京: 化学工业出版社, 2005: 43-49.
- [10] 杜华英, 赵跃龙. 人工神经网络典型模型的比较研究[J]. 计算机技术与发展, 2006, 16(5): 97-99.

(上接第 64 页)

4 结束语

文中将主成分分析、遗传算法融合在 BP 神经网络中, 先利用主成分分析进行样本降维, 再利用遗传算法优化初始权值, 最后进行 BP 神经网络训练, 形成一种提高 BP 神经网络泛化能力的改进算法, 为实际生活中样本复杂的神经网络提出了新的思路。通过对输入样本集进行主成分分析处理, 达到了数据降维、剔除冗余信息的目的, 借助遗传算法能够全局寻优的特点对初始权值进行优化, 弥补了标准 BP 算法容易陷入到局部极值中的不足, 增加了网络对训练集合以外样本的识别能力, 提高了 BP 神经网络的泛化能力。

参考文献:

- [1] Lehtokangas M. Modified Cascade - Correlation Learning for Classification[J]. IEEE Trans. on Neural Networks, 2000, 11(3): 795-798.
- [2] 武 妍, 张立明. 神经网络的泛化能力与结构优化算法研究[J]. 计算机应用研究, 2002(6): 21-25.
- [3] 周开利, 康耀红. 神经网络模型及其 MATLAB 仿真程序设计[M]. 北京: 清华大学出版社, 2006: 69-100.
- [4] 杨 静, 毛宗源. 基于 PCA 和神经网络的识别方法研究[J]. 计算机工程与应用, 2007, 43(25): 246-248.
- [5] Ruan Qing, Wang Yi-qiang. PCA Approach to BP Learning[J]. Journal of Fudan University: Natural Science, 2005, 44(2): 318-322.
- [6] 王智平, 刘在德, 高成秀, 等. 遗传算法在 BP 网络权值学习中的应用[J]. 甘肃工业大学学报, 2001, 27(2): 20-22.
- [7] Zorriassatine F, Tannock J D T. A review of neural networks for statistical process control[J]. Journal of Intelligent Manufacturing, 1998, 19: 209-224.
- [8] 贾丽会, 张修如. BP 算法分析与改进[J]. 计算机技术与发展, 2006, 16(10): 200-203.
- [9] 熊 凌, 赵明旺. 基于遗传算法的 BP 网络全局收敛的混合智能学习算法[J]. 武汉科技大学学报: 自然科学版, 2000, 23(2): 183-186.