

基于新的改进粒子群算法的 BP 神经网络在拟合非线性函数中的应用

林宇锋 邓洪敏 史兴宇

(四川大学电子信息学院 成都 610065)

摘要 介绍了一种基于新的改进粒子群算法(NIPSO)的 BP 神经网络来解决拟合非线性函数所出现的误差较大的问题。此算法在粒子群优化算法基础上,分别让权重和学习因子非线性和线性变化,建立基于新的粒子群优化算法的新模型,再与 BP 神经网络结合之后来拟合非线性函数。结果表明,新的改进粒子群优化算法更加合理且高效地提高了 BP 神经网络的拟合能力,减小了拟合误差,提高了拟合精度。

关键词 BP 神经网络,粒子群优化算法,函数拟合

中图分类号 TP183 **文献标识码** A

Application of BP Neural Network Based on Newly Improved Particle Swarm Optimization Algorithm in Fitting Nonlinear Function

LIN Yu-feng DENG Hong-min SHI Xing-yu

(College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China)

Abstract A newly improved particle swarm optimization algorithm for the optimization of BP neural network was introduced to solve the problem of large error in fitting the nonlinear function. In this algorithm, a new model based on the newly improved particle swarm optimization algorithm is established through respectively changing the weight non-linearly and learning factor linearly, then it is applied to non-linear function fitting by combining with BP neural network. The results show that the newly improved particle swarm optimization algorithm can more rationally and effectively boost the fitting ability of BP neural network, and improve the accuracy of the fitting.

Keywords BP neural network, Particle swarm optimization algorithm, Function fitting

1 引言

在工程项目中会面临很多复杂难解的非线性函数问题,它们很难用数学公式列写;但 BP 神经网络能通过学习 and 存储大量的输入输出映射关系,来拟合非线性函数模型,使工程项目问题变得相对简单。其拟合的方法是:把需要拟合的系统看成一个黑盒子,用大量存储的训练样本来训练神经网络,通过误差修正建立起神经网络新模型,再去验证和预测系统的输出。在 MATLAB 神经网络工具箱中,可以自主完成网络结构设计、初始权值阈值设定、网络训练等一系列的工作。但实验表明,仅用 BP 神经网络去拟合非线性函数,存在收敛速度很慢、容易陷入局部极小值、精度低等问题。因此,智能优化算法被用于优化 BP 神经网络成为当前项目开发、研究的焦点之一。

粒子群(Particle Swarm Optimization, PSO)算法是智能优化算法中应用较广泛的,其在 1995 年被提出并成功应用于函数拟合、函数优化等领域。PSO 算法是在模拟鸟类捕食行为的基础上提出的,算法中参数少,运算简单,收敛速度快,易实现,但如果把未改进的标准 PSO 算法应用到 BP 神经网络

中,则很容易陷入局部极小值,且精度较低,因此基于粒子群算法的改进受到广泛关注。文献[1]提出一种动态调整粒子群算法,调整方法为: $w = n * w_0$, w_0 随迭代次数的增加而线性减小, n 满足正态分布,以此动态调整权重;文献[2]提出一种改进粒子群算法,主要是与遗传算法的变异相结合,再引入自适应变异算子,对公式中某些变量以一定概率重新初始化,缩小搜索空间,有效提高搜索效率,从而提高搜索到全局最优解的可能性,然后再与 BP 神经网络结合;文献[3]采用改进的粒子群算法,使惯性权重随迭代次数的增加而非线性变化,再与 BP 神经网络结合。实验表明,基于这些改进粒子群算法优化的 BP 神经网络的精度较低,运行时间长,不能完全跳出局部极小值。因此,本文在粒子群优化算法上提出了一种新的改进粒子群算法,不仅让惯性权重非线性变化,而且使学习因子也随迭代次数线性变化,迭代初期权重和学习因子较大,使粒子群有足够的搜索空间,防止陷入局部极小值;迭代后期权重和学习因子较小,使收敛速度加快,精度变高。用得到的全局最优解给 BP 神经网络的权值和阈值赋值,用大量训练样本训练 BP 神经网络,建立起精度相对更高、拟合能力

本文受国家自然科学基金(61174025)资助。

林宇锋(1994—),男,硕士生,主要研究方向为神经网络、人工智能, E-mail: 383714800@qq.com; 邓洪敏(1969—),女,博士,副教授,主要研究方向为非线性动力学、模糊控制、神经网络; 史兴宇(1993—),男,硕士生,主要研究方向为神经网络、人工智能。

更强的BP神经网络,再对非线性函数拟合、预测。实验分析显示:相对改进粒子群算法和标准粒子群算法而言,新的改进粒子群算法BP神经网络对非线性函数的拟合精度更高,能更合理、高效地加快收敛速度。

2 BP神经网络原理

BP(Back Propagation)神经网络由以Rumelhart和McClelland为首的科学家于1986年提出来的,它是一种具有多层结构的前向神经网络^[4]。本文主要介绍3层的BP神经网络,其拓扑结构如图1所示,包括输入层、隐含层和输出层。每一层的神经元与下一层的所有神经元之间采用全互连方式相连,同一层的神经元之间无连接^[5]。

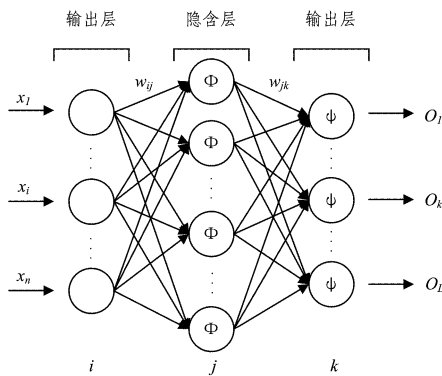


图1 三层网络结构

BP神经网络的学习过程包括如下两个阶段。

(1)第一阶段为正向传播阶段,是一个前向计算的过程。此过程是由输入信号向前传播,经过输入层到隐含层、输出层逐层计算,在输出端得到相应的输出信号。以下为文献^[6]给出的前向计算过程算法:

隐含层第 j 个节点的输入 net_j 为:

$$net_j = \sum_{i=1}^M w_{ij} x_i + \theta_j \quad (1)$$

隐含层第 j 个节点的输出 y_j 为:

$$y_j = \phi(net_j) = \phi\left(\sum_{i=1}^M w_{ij} x_i + \theta_j\right) \quad (2)$$

输出层第 k 个节点的输入 $innet_k$ 为:

$$\begin{aligned} innet_k &= \sum_{j=1}^q w_{jk} y_j + \alpha_k \\ &= \sum_{j=1}^q w_{jk} \phi\left(\sum_{i=1}^M w_{ij} x_i + \theta_j\right) + \alpha_k \end{aligned} \quad (3)$$

输出层第 k 个节点的输出 o_k 为:

$$\begin{aligned} o_k &= \psi(innet_k) = \psi\left(\sum_{j=1}^q w_{jk} y_j + \alpha_k\right) \\ &= \psi\left(\sum_{j=1}^q w_{jk} \phi\left(\sum_{i=1}^M w_{ij} x_i + \theta_j\right) + \alpha_k\right) \end{aligned} \quad (4)$$

(2)第二阶段为反向阶段,是一个误差反向传播过程,即如果实际输出与期望输出的差值(误差)大于所设定的误差值,则误差反向传播,并逐次调整连接的权值和阈值,直到满足误差要求。网络对 p 个训练样本的总误差准则为:

$$E_p = \frac{1}{2} \sum_{p=1}^n \sum_{k=1}^l (T_k^p - O_k^p)^2 \quad (5)$$

其中, O_k^p 是输出节点 k 在样本 p 作用下的实际输出值, T_k^p 是输出节点在样本 p 作用下的期望值。BP神经网络的学习主要是通过更新权值和阈值使得系统的精度变高。BP算法简

单、计算量小,是目前最成熟的算法之一^[7];但BP神经网络在训练之前各个层之间连接的权值和阈值都是随机化的,使得BP神经网络收敛速度慢,且易陷入局部极小值,可能使得到的值非全局最优值。采用新的改进粒子群算法对初始的权值阈值进行优化,从而更加合理、高效地提高BP神经网络的精度,加快收敛速度,得到全局最优解。

3 粒子群优化算法原理

粒子群算法是由Kennedy和Eberhart两位科学家于1995年最早提出的一种群体智能算法^[8-9],其本质思想来源于对鸟类捕食的研究。在鸟类捕食过程中,为了找到食物最多的位置,就要考虑鸟类种群中的最优位置以及自身所经历的最优位置,以此方法可高效地接近种群最优的位置。对于粒子群算法,在初始状态下每个粒子的位置都是潜在的最优解,即把每一个粒子的初始位置设定为个体最优解。当比较每个粒子的优良程度时,采用了适应度函数来表征,其中每个粒子的特征状态用速度、位置和适应度值来体现。一般而言,在将BP神经网络应用到拟合非线性函数中时,将适应度函数设为实际值和期望值之差,以便通过反向过程来调节权值和阈值。假设在一个 D 维的空间中有 m 个粒子,则第 i 个粒子的位置为: $X_i = (X_{i1}, X_{i2}, \dots, X_{iD})$,第 i 个粒子的速度为 $V_i = (V_{i1}, V_{i2}, \dots, V_{iD})$ ^[10]。在粒子群迭代进化过程中,粒子根据自身的位置和同伴的最优位置来调整自身移动的方向,再结合实际位置和速度,共同来决定移动的下一位置。在每个粒子移动的过程中,都有一个最优解;在整个粒子群移动过程中,也有一个最优解。前一个称为个体极值(P_{best}),其个体最优解为 $P_i = (P_{i1}, P_{i2}, \dots, P_{iD})$;后一个称为全局极值(g_{best}),其全局最优解为 $g_i = (g_{i1}, g_{i2}, \dots, g_{iD})$ 。在粒子群迭代过程中,要考虑到个体极值、全局极值以及上一状态的速度和位置,从而得到下一状态的位置,其迭代的速度和位置公式分别为^[11]:

$$V_i = \omega * V_i + C_1 * rand * (P_{best} - X_i) + C_2 * rand * (g_{best} - X_i) \quad (6)$$

$$X_i = X_i + 0.5 * V_i \quad (7)$$

其中, C_1 和 C_2 为常数,称为学习因子,一般取为2; $rand$ 为 $[0, 1]$ 中的随机数; ω 为惯性权重(inertia weight)。需要注意的是:在粒子群迭代过程中,根据速度大小和方向的变化,不断地更新粒子位置,而其速度受到最大速度 V_{max} ^[12] 和最小速度 V_{min} 的限制;即当速度高于 V_{max} 时,粒子速度就限定为 V_{max} ;当速度低于 V_{min} 时,粒子速度就限定为 V_{min} 。式(6)由3部分组成,第一部分为记忆部分,反映了上一状态的速度大小和方向,其目的为权衡全局搜索和局部搜索的作用;第二部分为认知部分(Cognition Model),体现了粒子自身的能力,对自身状态的思考,使得粒子具有较强的全局搜索能力,防止陷入局部极小值;第三部分为社会部分,让粒子间进行信息的交流,主要是考虑全局极值来调整自身位置,体现了粒子之间的团结协作。3个部分共同来得到粒子的下一个位置向量,共同决定粒子的空间搜索能力^[13]。

4 新的改进粒子群优化算法优化神经网络

BP神经网络运算简单,计算量小,但是学习效率低,收敛的速度慢,且容易陷入局部极小值,运算结果的精度不高;粒

子群优化算法的结构比较简单,用粒子群优化算法优化 BP 神经网络,其目的为设置初始状态的权值和阈值,减少随机性,相对 BP 算法运行的速度有所提高,运算结果的精度变高,但在初始权值和阈值寻优的过程中也易陷入局部极小值,从而导致失去全局寻优能力,函数拟合的精度降低。为了降低陷入局部极小值的概率,且更加合理、高效地提高拟合精度,提出了新的改进粒子群优化算法(Newly Improved Particle Swarm Optimization, NIPSO),其主要对学习因子 C_1, C_2 和惯性权重 w 进行改进。

4.1 惯性权重的改进

在标准的粒子群算法中,惯性权重选定为一固定值,不能随粒子进化过程中的状态而产生变化,这将导致在进化初期,如果惯性权重过小,则容易陷入局部极小值,搜索空间减小;在后期,如果惯性权重过大,则不利于算法收敛。因此,惯性权重的变化对优化的性能影响很大。在许多研究中,研究者提出过线性递减权重(LDIW)的方法,使算法初期惯性权重较大,易于跳出局部极小值,提高粒子的全局搜索能力;在后期惯性权重较小,有助于提高收敛精度,即让惯性权重呈线性变化,其算法公式如式(8)所示^[13]。但由于粒子位置和速度的变化是非线性的,因此本文基于惯性权重的改进采用非线性递减权重策略,以更有效地加快粒子的变化速度,从而提高算法的收敛精度,其算法公式如式(9)所示。

$$w = (w_{start} - w_{end}) \frac{Maxiter - iter}{Maxiter} + w_{end} \quad (8)$$

$$w = (w_{start} - w_{end}) \frac{(Maxiter - iter)^2}{Maxiter^2} + w_{end} \quad (9)$$

其中, w_{start} 和 w_{end} 分别为惯性权重的初始值和最终值,其取值范围是: $0.4 \leq w_{end} < w_{start} \leq 0.9$, $Maxiter$ 为迭代次数的最大值; $iter$ 为当前的迭代次数。

4.2 学习因子的改进

在粒子群优化算法中,学习因子分别控制着粒子的认知部分和社会部分。理想状态下,我们希望粒子群能够在算法的初期具备全局的搜索能力,在整个寻优空间进行搜索,不至于陷入局部极小值;在算法后期能够尽快地收敛到全局最小值,加快收敛的速度,提高收敛精度。因此,本文就通过动态调整 C_1 和 C_2 学习因子的大小来达到此目的,其基本的操作方法为:在算法的初始阶段,有大的认知部分和小的社会部分,有助于提高全局搜索能力;在算法的后期,有小的认知部分和大的社会部分,更有助于算法收敛于全局最优解,提高算法的精度和收敛速度。 C_1 ^[13] 和 C_2 的变化规律为:

$$C_1 = (C_{start} - C_{end}) \frac{Maxiter - iter}{Maxiter} + C_{end} \quad (10)$$

$$C_2 = 3 - C_1 \quad (11)$$

其中, $Maxiter$ 和 $iter$ 分别为算法的最大迭代次数和当前迭代次数; C_{start} 为学习因子的初始值, C_{end} 为学习因子的最终值,两者的取值范围是: $1 \leq C_{end} < C_{start} \leq 3$ 。

4.3 优化 BP 神经网络

用本文提出的新的粒子群优化算法(NIPSO)优化神经网络的方法为:粒子群中每个粒子表示的是 BP 神经网络的权重和阈值,每个粒子的维度由网络中权值的数量和阈值个数决定,在设定的训练样本下, BP 神经网络输出的实际值和期望值的误差作为 BP 神经网络训练问题的适应度函数,适应

度值表示神经网络的误差,误差值小表示粒子在空间中具有更好的性能。粒子位置和速度更新的目的,就是去改变初始的权值和阈值,减小适应度值,减小输出的误差,直至其小于所设定的误差值,从而提高 BP 神经网络拟合非线性函数的精度。NIPSO 优化 BP 神经网络的步骤如下:

步骤 1 设定初始化群体属性(网络的拓扑关系、种群大小、迭代最大次数、训练误差等)。

步骤 2 每个粒子 i 在 BP 神经网络中进行第一阶段的正向传播,将输出层的输出值与期望值比较,得到误差 E_i , 如果 $E_i < pbi$, 则 $pbi = E_i, P_i = X_i$; 如果 $E_i < g_{best}$, 则 $g_{best} = E_i, P_g = X_i$; (pbi 表示个体极值, X_i 表示当前解, P_i 表示个体最优解, P_g 表示全局最优解, g_{best} 表示全局极值)。

步骤 3 对每一个粒子,根据式(6)和式(7)更新其位置和速度。

步骤 4 迭代次数加 1, 如果没有达到最大迭代次数, 转向步骤 2; 如果达到最大迭代次数, 转向步骤 5。

步骤 5 得到初始的权值和阈值为 P_g 。

5 仿真实验

为了测试新的基于改进粒子群算法的 BP 神经网络在拟合非线性函数上的应用精度, 本文选取非线性拟合函数为: $y = x_1^2 + x_2^2$, 此拟合函数有两个输入变量和一个输出; BP 神经网络的拓扑结构是 2-10-1(输入层有 2 个神经元, 隐含层有 10 个神经元, 输出层有 1 个神经元)。随机产生 1500 组输入输出值并将其归一化处理, 用 1400 组作为训练样本, 100 组作为测试样本。分别对 3 种算法的参数进行设置, 如表 1 和表 2 所列。

表 1 BP 神经网络测试参数的设置

算法	学习速率	训练次数	训练目标
PSO-BP	0.1	100	10^{-5}
IPSO-BP	0.1	100	10^{-5}
NIPSO-BP	0.1	100	10^{-5}

表 2 粒子群算法测试参数的设置

算法	种群大小	迭代次数	C_1, C_2	惯性权重
PSO-BP	30	10	2, 2	0.5
IPSO-BP	30	10	2, 2	线性
NIPSO-BP	30	10	线性	非线性

实验分别用标准 PSO 算法优化 BP 神经网络、NIPSO 算法优化 BP 神经网络和 IPSO 算法优化 BP 神经网络之后用于拟合非线性函数。将标准 PSO 算法优化 BP 神经网络与 NIPSO 算法优化 BP 神经网络之后用于拟合非线性函数所得到的误差来进行比较, 如图 2 所示; 将用 IPSO 算法优化 BP 神经网络与 NIPSO 算法优化 BP 神经网络之后用于拟合非线性函数所得到的误差进行比较, 如图 3 所示。

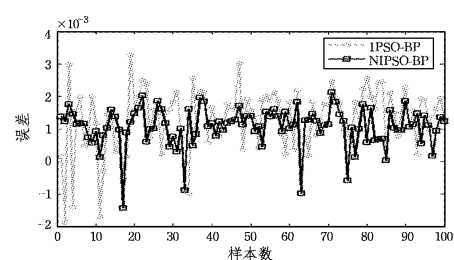


图 2 NIPSO-BP 与 IPSO-BP 的预测误差对比

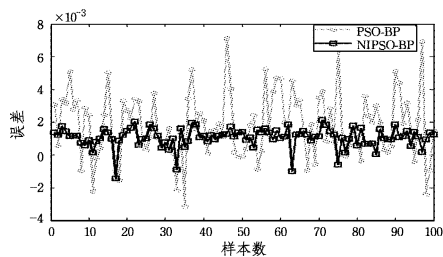


图3 PSO-BP与NIPSO-BP的预测误差对比

由图2和图3可知,新的改进粒子群算法(NIPSO)优化BP神经网络对非线性函数的拟合和预测取得了良好的效果,其精度高于标准PSO算法和改进的粒子群(IPSO)算法,误差也更小,收敛速度有很大改善,达到了我们改进的要求。

结束语 本文对粒子群算法中的权重进行非线性变换,对学习因子进行线性变换,目的在于使得在粒子群位置和速度变化的前期具有较大的权重和学习因子,有较强的全局搜索能力,减小陷入局部极小值的概率;在后期具有较小的权重和学习因子,使得粒子能较快地收敛到最小值。将上述的算法优化BP神经网络之后用于拟合非线性函数,实验表明,新的改进粒子群算法可以更合理有效地提高BP神经网络的预测精度,比标准的PSO算法、IPSO算法的BP神经网络在收敛速度、计算精度、算法稳定性等方面有明显优势,具有更小的训练误差和检测误差。

参考文献

[1] 皮倩瑛,叶洪涛.一种动态调节惯性权重的粒子群算法[J].广西科技大学学报,2016,27(3):27-31.
[2] 李松,刘力军,等.改进粒子群算法优化BP神经网络的短时交

通流预测[J].系统工程理论与实践,2012,32(9):2045-2049.

[3] 王慧,刘希玉.基于最具影响粒子群优化的BP神经网络训练[J].计算机工程与应用,2007,43(18):69-71.
[4] BASHIR Z A, EL-HAWARY M E. Applying Wavelets to Short-Term Load Forecasting Using PSO-Based Neural Networks[J]. IEEE Transaction on power system February, 2009, 24(1): 20-26.
[5] 王德明,王莉,张广明.基于遗传BP神经网络的短期风速预测模型[J].浙江大学报(工学版),2012,46(5):837-841.
[6] 刘冰,郭海霞. MATLAB神经网络超级学习手册[M].北京:人民邮电出版社,2014:159-163.
[7] 许以山,曾碧,尹秀文,等.基于改进粒子群算法的BP神经网络及其应用[J].计算机工程与应用,2009,45(35):233-235.
[8] 李炳宇,萧蕴诗.新的进化计算算法-粒子群优化算法[J].计算机科学,2003,30(6):19-21.
[9] GUDISE V G, VENAYAGAMOORTHY G K. Comparison of Particle Swarm Optimization and backpropagation as training algorithms for neural networks[C]//IEEE Department of Electrical and Computer Engineering University of Missouri-Rolla. USA,2003:110-116.
[10] 赵宏伟,李圣普.基于粒子群算法和RBF神经网络的云计算资源调度方法研究[J].计算机科学,2016,43(3):113-117.
[11] 吕振肃,侯志荣.自适应变异的粒子群优化算法[J].电子学报,2004,32(3):416-420.
[12] CLERC M, KENNEDY J. The Particle Swarm-Explosion, stability, and convergence in a Multidimensional Complex Space[J]. IEEE Transaction on Evolutionary Computation, 2002, 6(1): 58-71.
[13] 沈学利,张红岩,张纪锁.改进粒子群算法对BP神经网络的优化[J].计算机系统应用,2010,19(2):57-61.

(上接第38页)

[27] MING Z, SHANBHAG N R. Soft-Error-Rate-Analysis (SERA) Methodology [J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2006, 25(10): 2140-2155.
[28] FISHMAN G. Monte carlo: concepts, algorithms and applications [M]//Operations Research. New York: Springer-Verlag, 1995.
[29] CHOUDHURY M R, MOHANRAM K. Accurate and scalable reliability analysis of logic circuits [C]//Proceedings of IEEE/ACM Conference on Design, Automation & Test in Europe Conference & Exhibition, 2007 (DATE'07). 2007: 1-6.
[30] FLAUQUER J T, DAVEAU J M, NAVINER L, et al. Fast reliability analysis of combinatorial logic circuits using conditional probabilities [J]. Microelectronics Reliability, 2010, 50(9-11): 1215-1218.
[31] KRISHNASWAMY S, VIAMONTES G F, MARKOV I L, et al. Accurate reliability evaluation and enhancement via probabilistic transfer matrices [C]//Proceedings of IEEE/ACM Conference on Design, Automation and Test in Europe. Orlando, USA: IEEE Computer Society, 2005: 282-287.
[32] 王真,江建慧,员春欣.高性能处理器的差错校正技术[J].计算机研究与发展,2008,45(2):358-366.
[33] 肖杰,江建慧,等.一种基于迭代PTM模型的电路可靠性评估方法[J].计算机学报,2014(7):1508-1520.

[34] 肖杰,江建慧,等.面向晶体管级广义门电路的PTM可靠性计算[J].中国科学:信息科学,2014(10):1226-1238.
[35] 欧阳城添,江建慧,等.触发器可靠性计算的F-PTM方法[J].电子学报,2016(9):2219-2226.
[36] 黄海林,唐志敏,许彤.龙芯1号处理器的故障注入方法与软差错敏感性分析[J].计算机研究与发展,2007,43(10):1820-1827.
[37] MUKHERJEE S S, WEAVER C, EMER J, et al. A systematic methodology to compute the architectural vulnerability factors for a high-performance microprocessor [C]// Proceedings of 36th Annual IEEE/ACM International Symposium on Microarchitecture. San Diego, CA, USA: IEEE Computer Society, 2003: 29-40.
[38] KRISHNASWAMY S, PLAZA S M, MARKOV I L, et al. Signature-Based SER Analysis and Design of Logic Circuits [J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2009, 28(1): 74-86.
[39] 胡谋.计算机容错技术[M].北京:中国铁道出版社,1995.
[40] 徐拾义.可信计算系统设计和分析[M].北京:清华大学出版社,2006.
[41] GeNIe & SMILE[OL]. http://genie.sis.pitt.edu. 2010.
[42] 王真,江建慧.基于概率转移矩阵的串行电路可靠性计算方法[J].电子学报,2009,37(2):241-247.