

区间自适应遗传算法优化 BP 神经网络 权值阈值研究

朱会霞

(辽宁工业大学 管理学院, 辽宁 锦州 121001)

摘要: 针对 BP 神经网络易陷入局部最优、接近最优解时产生振荡等缺点, 利用区间自适应遗传算法自适应移动搜索区间的功能, 彻底取代 BP 神经网络权值阈值反向调整过程。结果表明该方法可使训练误差达到较高的学习精度, 有效解决了 BP 神经网络的困境。将该算法用于大豆油甲酯合成工艺试验中, 当训练误差达到 2.58×10^{-8} 时, 神经网络的实际输出值与试验值完全拟合, 说明了该方法是一种可靠有效的优化方法。

关键词: 区间自适应遗传算法; 神经网络; 权值阈值

中图分类号: TP301.6 **文献标识码:** A **文章编号:** 1674-3261(2020)05-0328-04

Research on Optimization of BP Weight and Threshold by Interval Adaptive Genetic Algorithm

ZHU Hui-xia

(School of Management, Liaoning University of Technology, Jinzhou 121001, China)

Abstract: Since the weights and the thresholds of BP neural networks obtained via existing optimization methods were usually local optimal, these algorithms oscillate upon approaching the optimal solutions. The objective functions of the interval adaptive genetic algorithm were employed to replace the backward adjustment for the weights and thresholds of the BP algorithm. The results revealed that it was able to render training error to higher learning accuracy. Hence, the above difficulties of the BP algorithm could be coped effectively. In addition, the improved algorithm was also applied to optimizing synthesis of soybean oil methyl ester. When the training error was 2.58×10^{-8} , the actual output of the neural network was completely fitted to the test value, which showed that the method was a reliable and effective optimization method.

Key words: interval adaptive genetic algorithm, neural networks, weights and the thresholds

遗传算法和神经网络是智能算法中两个重要方法。BP (Back-Propagation algorithm) 神经网络本质上是基于梯度下降的一种迭代学习算法, 其误差曲面是非常复杂的多元曲面, 存在许多局部极小解, 当梯度为 0 时, BP 算法则无法辨别极小点性质, 使训练难以收敛到指定误差; 若在误差曲面平坦区, 则调整时间长, 影响收敛速度^[1]。为了克服

这些缺陷, 学者提出了许多改进措施, 如附加动量项、自适应学习率, 与现代优化方法结合等。而建立在遗传学基础上的遗传算法 (genetic algorithms, GA) 主要特点是对目标函数无限制, 广泛用于人工神经网络训练与优化问题^[2-3]。

如何使遗传算法和 BP 算法有机结合, 许多学者做了大量研究。一般用遗传算法优化 BP 神经网络

收稿日期: 2020-01-25

基金项目: 辽宁省教育厅高校科研基金 (JQW201715407)

作者简介: 朱会霞(1978-), 女, 黑龙江佳木斯人, 讲师, 博士。

络结构和初始权值阈值^[4-5]。但遗传算法求解优化问题时, 只能求得给定区间内的优化值^[6]。而神经网络权值和阈值所在区间一般无法预知, 故本文采用能够动态调整权值和阈值所在区间的自适应遗传算法 (IAGA), 彻底取代 BP 算法反向调整权值阈值的过程, 达到自适应调整权值阈值的目的, 从而将 BP 算法和 IAGA 有机结合起来, 充分发挥各自优势, 有效克服 BP 算法学习过程中易陷入局部极小值、振荡等缺陷。

1 BP 算法及 IAGA 算法

1.1 BP 算法原理

三层 BP 算法其结构如图 1 所示。

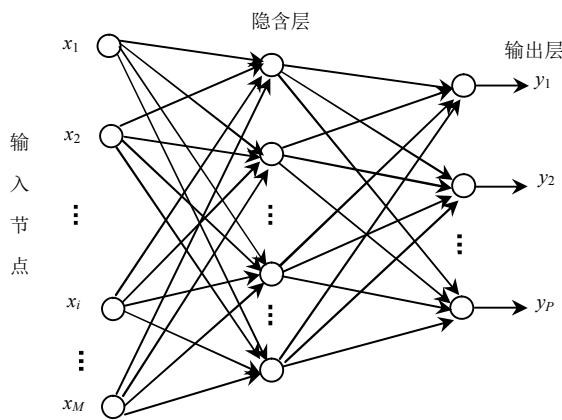


图 1 BP 神经网络的结构

$X=[X_1, X_2, \dots, X_k, \dots, X_N]$ 为 N 个训练样本集, $X_k=[X_{k1}, X_{k2}, \dots, X_{kM}]^T$ 为任一训练样本, ($k=1, 2, \dots, N$), $Y_k=[y_{k1}, y_{k2}, \dots, y_{kp}]^T$ 实际输出为, $d_k=[d_{k1}, d_{k2}, \dots, d_{kp}]^T$ 为期望输出。神经元激励函数均使用 Sigmoid 函数:

$$f(x) = \frac{1}{1 + e^{-x}} \quad -\infty < x < +\infty \quad (1)$$

经神经网络正向传播后工作信号的实际输出为:

$$u_i^l = \sum_{m=1}^M w_{mi} x_{km} + \theta_i, \quad v_i^l = f(u_i^l) \quad (2)$$

$$u_p^p = \sum_{i=1}^I w_{ip} v_i^l + \theta_p, \quad v_p^p = f(u_p^p) \quad (3)$$

$$y_{kp} = v_p^p = f(u_p^p) = f\left(\sum_{i=1}^I w_{ip} v_i^l + \theta_p\right) \quad (4)$$

式中: $i=1, 2, \dots, I$, $p=1, 2, \dots, P$

设 n 为迭代次数, 权值、阈值和网络实际输出是 n 的函数。则输出层第 p 个神经元误差信号为:

$$e_{kp}(n) = d_{kp}(n) - y_{kp}(n) \quad (5)$$

神经元 p 的误差能量定义为 $\frac{1}{2}e_{kp}^2(n)$, 则输出

层所有神经元的误差能量总和为 $E(n)$:

$$E(n) = \frac{1}{2} \sum_{p=1}^P e_{kp}^2(n) \quad (6)$$

在误差信号的反向传播过程中, BP 算法是通过调整权空间中的梯度使 $E(n)$ 值下降的, 详细推导过程见文献[7]。

1.2 IAGA 算法描述

为了能更好的说明 IAGA 自适应移动搜索区间的过程, 以某一权值在初始区间 $[U_1, U_2]$ 中的调整过程为例, 且 $U_1 > U_2$, U_2 点的目标函数值优于 U_1 点。随机产生 pop 个初始种群 $\omega_1^0, \omega_2^0, \dots, \omega_{pop}^0$, 设所有神经元的误差能量总和达到指定较小精度要求时所对应的权值为 ω^* , 且 ω^* 不在 $[U_1, U_2]$ 内, $\omega^* > U_2$ 。经过几次选择、交叉和变异遗传操作后种群中目标函数值最小的个体 ω_1^0 逐渐向 U_1 点趋近, 这时可取

$$\begin{aligned} \Delta_1 &= \omega_1^0 - U_1 \\ U_1' &= U_1 + \Delta \\ U_2' &= U_2 + \Delta \end{aligned} \quad (7)$$

这样, 当前最优解移动至 $[U_1', U_2']$ 区间, 将原区间 $[U_1, U_2]$ 中的个体, 按照既定的解码和编码规则, 映射到新的搜索区间 $[U_1', U_2']$ 中, 完成区间移动, 在新的搜索区间再进行选择、交叉、变异等遗传操作, 最终移动到最优解所在区间。当 $\omega^* < U_1$ 时, 与上述情况类似。

2 IAGA 优化 BP 算法

2.1 IAGA 优化 BP 权值阈值的思想

神经网络最优解所对应权值和阈值的区间范围一般很难预知, 而遗传算法求解优化问题时, 若最优解不在给定的取值范围内, 只能求得一个指定区间内的较优解, 有可能求不到真正使误差能量总和达到指定精度要求的较优解, 使神经网络的拟合效果下降, 为此本文将区间自适应遗传算法用于神经网络权值阈值的调整, 即用 IAGA 的优化功能, 取代误差信号反向传播逐层修改权值和阈值的过程。IAGA 优化 BP 算法的数学模型为:

$$\min(E(n)) = \frac{1}{2} \sum_{k=1}^N \sum_{p=1}^P e_{kp}^2(n) \quad (8)$$

2.2 IAGA 优化 BP 权值和阈值算法描述

步骤 1: 对训练样本进行预处理。

为避免出现奇异样本数据所引起训练时间增加和无法收敛等问题, 将训练样本数据进行归一化处理, 使量纲数据变为无量纲纯量。

步骤 2: 变量的初始化。

输入: 种群规模 pop , 最大迭代次数 T , 交叉

概率 p_c , 变异概率 p_m , 权值与阈值的初始区间。

步骤 3: 计算码长, 随机产生初始种群。

步骤 4: 设置迭代次数 $t=0$ 。

步骤 5: 形成父代种群。

步骤 6: 解码。

步骤 7: 保留 m 个适应度高精英个体。

步骤 8: 选择。

步骤 9: 交叉。采用单点交叉方式。

步骤 10: 变异。

步骤 11: 保留优秀个体^[8]。

步骤 12: 判断是否满足更新区间条件, 不满足, 则执行步骤 13, 满足则移动搜索区间。

步骤 13: 将交叉产生的个体, 按自适应变异方式进行变异。

步骤 14: 精英替代变异后适应度值低的个体。

步骤 15: 形成新一代种群, 遗传进化一代完成, 即 $t=t+1$;

步骤 16: 判断训练误差结束条件, 满足则终止遗传运算, 不满足要求, 则转到步骤 5。

步骤 17: 将计算得到的权值和阈值代入工作信号正向传播的输出公式(4), 并计算神经网络的实际输出。

步骤 18: 反归一化处理。为了与原始数据拟合, 需将神经网络的输出数据反归一化处理。

步骤 19: 与试验数据拟合。

3 仿真结果与分析

3.1 试验数据与预处理

用曹秀丽等^[9]所做的大豆油甲酯合成工艺试验数据为训练样本, 验证区间自适应遗传算法优化神经网络权值阈值, 试验训练数据如表 1 所示。因 Sigmoid 函数在接近 0 或 1 时, 基本达到饱和状态, 增加变慢, 故将试验数据归一化处理到区间[0.1,0.8]之间, 加快网络的收敛速度。

表 1 试验训练样本

序号	醇油摩尔比 A	催化剂用量 $B/\%$	反应温度 $C/^\circ\text{C}$	反应时间 D/min	产率 $y/\%$
1	7:1	1.1	75	60	87.79
2	6:1	1.1	75	50	85.64
3	6:1	1.3	65	50	80.24
4	6:1	1.1	65	60	93.81
5	6:1	1.3	75	60	83.67
6	5:1	1.1	75	60	90.87
7	5:1	1.1	55	60	92.58
8	6:1	0.9	65	70	79.89
9	7:1	1.1	65	70	82.29
10	6:1	1.1	75	70	88.06
11	7:1	1.3	65	60	82.75

12	7:1	1.1	65	50	85.57
13	6:1	0.9	75	60	85.46
14	5:1	1.1	65	50	87.11
15	6:1	1.3	65	70	80.79
16	6:1	1.1	55	50	88.61
17	6:1	1.3	55	60	85.60
18	7:1	1.1	55	60	87.36
19	6:1	0.9	55	60	86.34
20	5:1	1.1	65	70	88.49
21	6:1	1.1	55	70	87.63
22	6:1	0.9	65	50	83.34
23	5:1	1.3	65	60	85.13
24	5:1	0.9	65	60	89.23
25	7:1	0.9	65	60	81.72

3.2 BP 神经网络的训练

网络结构为 4-7-1。权值和阈值初始区间取值为 [-10, 10], 种群规模 80, 编码精度为 1.00e-4。当遗传操作代数 $count=1\ 000$ 时, 进行区间移动。交叉概率 $p_c=0.95$, 变异概率 $p_m=0.05$, 精英个体数 4。用 MATLAB R2016b 编程完成上述思想, 当输出误差达到 2.58×10^{-8} 时, 算法结束, 此时得到的权值阈值结果为:

$$W_M = \begin{bmatrix} 4.76 & -7.28 & 1.08 & 4.52 & 0.96 & 1.94 & -2.03 \\ -1.02 & 7.73 & 2.68 & -4.84 & 8.96 & 0.02 & -3.60 \\ 1.25 & -1.04 & 5.23 & 2.01 & 1.52 & -10.01 & 1.49 \\ -7.28 & -19.43 & 4.19 & 6.90 & 2.35 & -0.58 & 2.41 \end{bmatrix}$$
$$\theta_i = [1.74 \quad 1.46 \quad 10.47 \quad 8.94 \quad 8.97 \quad 2.76 \quad -0.96]$$
$$W_{IP} = [-4.66 \quad -6.55 \quad 7.84 \quad -9.49 \quad -6.48 \quad -28.01 \quad -4.63]^T$$
$$\theta_p = [-4.58]$$

将训练好的权值和阈值代入BP神经网络工作信号正向传播过程的计算公式, 即式(4), 可得网络的实际输出, 并实际输出反归一化处理后得到如表 2所示数据。

表2 网络输出值与实验值拟合精度情况 %

序号	试验值	网络输出值	相对误差
1	87.79	87.79	0.00
2	85.64	85.64	0.00
3	80.24	80.24	0.00
4	93.81	93.81	0.00
5	83.67	83.67	0.00
6	90.87	90.87	0.00
7	92.58	92.58	0.00
8	79.89	79.89	0.00
9	82.29	82.29	0.00
10	88.06	88.06	0.00
11	82.75	82.75	0.00
12	85.57	85.57	0.00
13	85.46	85.46	0.00
14	87.11	87.11	0.00
15	80.79	80.79	0.00
16	88.61	88.61	0.00
17	85.60	85.60	0.00

18	87.36	87.36	0.00
19	86.34	86.34	0.00
20	88.49	88.49	0.00
21	87.63	87.63	0.00
22	83.34	83.34	0.00
23	85.13	85.13	0.00
24	89.23	89.23	0.00
25	81.72	81.72	0.00

从表 2 可知, 用 IAGA 优化方法替代梯度下降法来调整权值和阈值, 可使 BP 神经网络的输出数据与实验数据完全拟合, 说明是一种有效的权值阈值调整方法。

4 结论

用 IAGA 彻底取 BP 神经网络反向调整权值和阈值的方法, 可以两种智能优化方法各自的优点有机地结合起来。IAGA 具有优化 BP 神经网络的能力, 解决了 BP 算法的困境。计算结果表明, 该方法可使训练误差的目标值达到很高的学习精度。在大豆油甲酯合成工艺问题中, 误差可以达到 2.58×10^{-8} , 平均相对误差为 0%, 证明了该方法是一种可靠有效的优化方法。

参考文献:

- [1] Snehamoy Chatterjee, Sukumar Bandopadhyay. Reliability estimation using a genetic algorithm-based artificial neural network: An application to a load-haul-dump machine[J].

Expert Systems with Applications, 2012, 39(12): 10943-10951.

- [2] 张品一, 梁锬. 基于 ADGA-BP 神经网络模型的金融产业发展趋势仿真与预测[J]. 管理评论, 2019, 31(12): 49-60.
- [3] 李帆, 闫献国, 陈峙, 等. 基于遗传算法优化 BP 神经网络的 YG8 硬质合金耐磨性预测模型[J]. 金属热处理, 2019, 44(12): 244-248.
- [4] 苏崇宇, 汪毓铎. 基于改进的自适应遗传算法优化 BP 神经网络[J]. 工业控制计算机, 2019, 32(1): 67-69.
- [5] 李彬, 张云, 王立平, 等. 基于遗传算法优化小波神经网络数控机床热误差建模[J]. 机械工程学报, 2019, 55(21): 215-220.
- [6] 朱会霞, 李微微, 李彤煜, 等. 区间自适应遗传算法优化无约束非线性规划问题[J]. 数学的实践与认识, 2019, 49(4): 110-116.
- [7] 王晓梅. 神经网络导论[M]. 北京: 科学出版社, 2019.
- [8] 王福林, 朱会霞, 王吉权, 等. 遗传算法的一种改进进化策略[J]. 生物数学学报, 2015, 30(1): 69-74.
- [9] 曹秀丽, 王赛丹, 李肖, 等. 响应面试验优化大豆油甲酯合成工艺[J]. 大豆科学, 2017, 36(1): 124-130.

责任编辑: 刘亚兵

(上接第 304 页)

- [5] Amatoul F Z, Lamchich M T, Outzourhit A. Design Control of DC/AC Converter for a grid Connected PV Systems using Matlab/Simulink[J]. Proceedings of the IEEE Power Electronics and Applications (EPE), 2011, 36(2): 1-7.
- [6] 章丽红. 基于重复和 PI 控制的光伏离网逆变器的研究[J]. 电力电子技术, 2012, 46(3): 33-35.
- [7] Tao Huanqi, Liu Suyi. Design and Implementation of Digital Control of Photovoltaic Power Inverter[J]. Procedia Environmental Sciences, 2011, 11: 155-162.
- [8] Samuel Vasconcelos Araújo, Peter Zacharias, Regine Mallwitz. Highly efficient single-phase transformer less inverters for grid-connected photovoltaic systems[J]. IEEE Trans on Industrial Electronics, 2010, 57(9): 3118-3128.
- [9] 黄庆丰, 红培, 翟登辉. 基于 Simulink 的两级式单相光伏逆变器的系统仿真[J]. 电力系统保护与控制, 2014, 42(19): 58-61.

- [10] 董晓鹏, 王兆安. 基于改进周期平均模型的 PWM 整流器控制[J]. 电力电子技术, 1999, 5(2): 11-14.
- [11] 王丰尧. 滑模变结构控制[M]. 北京: 机械工业出版社, 1995: 29-31.
- [12] Chem L G, Moonji Dong, Yuseong G U. Sliding mode controller for the single-phase grid-connected photovoltaic system[J]. IEEE, 2006, 12(5): 1101-1115.
- [13] Meza C, Biel D, Jelesema D, et al. Lyapunov-Based Control Scheme for Single-Phase Grid-connected Inverters[J]. IEEE, 2012, 56(7): 521-525.
- [14] Qun hai Huo, Li Kong, Tong zhen Wei, et al. A New Method for the Photovoltaic Grid-connected Inverter Control[J]. IEEE, 2008, 43(8): 2626-2629.
- [15] 嵇保健, 王建华, 赵剑锋. 不隔离单相光伏并网逆变器系统输入电流低频纹波抑制[J]. 电工技术学报, 2013, 28(6): 140-145.

责任编辑: 孙 林