

# 基于随机参数调整的改进反向传播学习算法

贾立山<sup>1,2</sup>, 谈至明<sup>1</sup>, 王 知<sup>1</sup>

(1. 同济大学 交通运输工程学院, 上海 201804; 2. 中国民航大学 航空自动化学院, 天津 300300)

**摘要:** 针对 BP(反向传播)神经网络学习易陷入局部极小的缺陷, 提出了一种改进 BP 神经网络学习算法——RMBP 算法. RMBP 算法在学习参数调整中增加了随机性, 使其方便地跳出局部极小点, 并沿梯度下降方向到达全局极小点. 异或问题的学习试验结果表明, RMBP 算法较 BP 学习算法和其他常见的改进 BP 算法具有学习速度快、学习精度高、资源占用少的优势. 最后, 结合民航飞机实时飞行仿真系统研究, 对一组飞机空气动力参数样本进行了学习, 以说明 RMBP 算法的有效性.

**关键词:** 改进 BP 算法; 随机参数调整; 神经网络; 全局极小; 飞行仿真

**中图分类号:** TP 183; U 8; TP 391.9 **文献标识码:** A

## Modified BP Algorithm Based On Random Adjustment of Parameters

JIA Lishan<sup>1,2</sup>, TAN Zhiming<sup>1</sup>, WANG Zhi<sup>1</sup>

(1. College of Transportation Engineering, Tongji University, Shanghai 201804, China; 2. College of Aeronautical Automation, Civil Aviation University of China, Tianjin 300300, China)

**Abstract:** A modified BP algorithm of neural network, random adjustment of parameters (RMBP) algorithm, is proposed to overcome the defect of easy going into local minimum of BP neural network. By adding randomization to the progress of parameters adjustment, RMBP helps BP neural network jump out of local minimum point easily, and get to global minimum point along gradient descent direction. Result of study experiment of XOR problem shows that RMBP has advantages of high study speed, high study precision and low resource occupation in comparison with other modified or unmodified BP algorithm. Finally, in combination with study of real-time flight simulation system of civil aviation airplane, a study is made of a group of aircraft aerodynamic force character parameters to verify the effectiveness of RMBP algorithm.

**Key words:** modified BP algorithm; random adjustment of parameters; neural network; global minimum; flight simulation

BP(back propagation, 反向传播)神经网络产生于 20 世纪 80 年代, 是一种应用反向传播算法进行学习的包含隐层的多层神经网络. BP 算法应用隐层产生输入模式自身的内部表示, 克服了感知器神经网络只能解决 1 阶谓词逻辑, 对高阶谓词逻辑无能为力缺陷; 解决了多层神经网络的学习问题; 拓宽了神经网络的应用范围. BP 神经网络可以进行线性及非线性分类、线性函数逼近、联想记忆等工作, 在模式识别、故障诊断、智能控制、通信等领域具有非常广泛的应用.

BP 学习算法采用梯度下降算法调整系统参数. 但是梯度下降算法具有收敛速度慢、易陷入局部极小的缺陷. 因此出现了许多 BP 算法的改进算法.

动量算法能够滤除神经网络学习时的高频振荡, 从而提高学习速度, 改善学习效果. 但对存在局部极小的神经网络, 使用该算法仍然要求选取能够进入全局极小点的初始参数, 否则也会陷入局部极小. 改变学习速度算法使网络学习在不同区域采用不同学习速度, 可提高学习精度. 但增加了许多参数, 提高了初始值选择的工作量和难度, 同时该算法增加了学习工作复杂度, 对一些普通 BP 网络能够达到全局极小点的问题却不能收敛. 采用共轭梯度法和 LM(Levenberg - Marquardt)算法虽然能够减小学习的迭代次数, 但算法复杂, 计算工作量较普通 BP 算法大了许多, 所以在学习时间上并不占优<sup>[1-2]</sup>.

带有网络复杂性的改良算法, 其网络复杂性的

收稿日期: 2010-01-27

第一作者: 贾立山(1976—), 男, 助理研究员, 博士生, 主要研究方向为载运工具运用工程以及计算机控制与仿真技术.  
E-mail: jlshfd163@yeah.net

通讯作者: 谈至明(1960—), 男, 教授, 博士生导师, 工学博士, 主要研究方向为交通运输工程. E-mail: tanjk@online.sh.cn

计算也要消耗额外的计算时间. 模拟退火算法虽然能够促使神经网络收敛到全局极小, 但收敛速度仍然缓慢<sup>[3-4]</sup>.

文献[5-6]均采用自动选择学习步长的变步长BP学习算法, 通过试探法选择BP神经网络的学习步长, 虽然能够减小网络学习迭代次数, 提高学习精度, 但是试探法本身要用去许多迭代次数, 整个网络的学习时间有可能会较长.

文献[7]的基于局部及全局误差的BP改进算法综合考虑了局部和全局误差进行学习速率调节, 虽然能够在误差精度要求高的时候减少迭代次数, 节省学习时间, 但在误差精度较低时迭代次数与普通BP算法相差不大, 并且大大高于基于全局误差的变步长BP算法. 所以, 其优势并不明显.

文献[8]的粒子群优化-反向传播(particle swarm optimization-back propagation, PSO-BP)算法和文献[9-10]的粒子群优化-反向传播算法采用粒子群优化(particle swarm optimization, PSO)算法代替传统梯度下降算法进行网络参数的调整, 避免了传统梯度下降算法收敛速度慢, 易陷入局部极小的缺陷. 但是, PSO算法本身也存在初始参数选择难度大, 易陷入局部极小的问题.

文献[11]的基于改进差异进化(modified differential evolution, MDE)算法的神经网络学习算法应用改进的差异进化(differential evolution, DE)算法代替梯度下降算法进行学习, 能够在一定程度上减小迭代次数, 提高学习精度.

BP网络学习陷入局部极小, 有时与权值与阈值的初值选取不当有关. 为此, 有些改进BP算法着眼于优化神经网络的参数初值. 文献[12]的带粒子群优化的改进BP神经网络(modified BP neural network with particle swarm optimization, MBPPSO)算法和文献[13]的粒子群优化改进BP神经网络学习算法应用粒子群优化算法寻找合适的权值与阈值的初值, 然后采用BP算法进行学习, 以保证BP算法在学习之初就能够找到梯度下降方向, 到达全局极小点. 文献[14]的多种群自适应遗传反向传播(multi-population adaptive genetic back propagation, MAGBP)算法和文献[15]的遗传优化反向传播(genetic algorithm-back propagation, GA-BP)算法均采用自适应遗传算法寻找权值与阈值的初值. 这些方法使神经网络在适当的参数初值基础上进行BP算法学习, 能够保证神经网络参数收敛于全局极小点, 并且减小了BP网络学习的迭代次数, 提

高了BP算法的学习效率. 但是, 神经网络参数初值的优化同样要占用计算机的工作时间, 其对系统资源的需求甚至大于BP算法对系统资源的需求. 因此, 这些改进算法不一定能够减小神经网络的学习时间.

根据以上现有各改进BP学习算法的对比分析, 在分析普通BP学习算法的缺陷的成因基础上, 提出了一种基于参数随机调整的改进BP学习算法, 称为RMBP(random modified back propagation)算法.

为了不失一般性, 应用异或问题对普通BP算法、RMBP算法、LM算法和改进差异进化法进行了对比. 证明RMBP能够有效克服普通BP算法易陷入局部极小的缺陷, 相比其他改进BP算法具有学习速度快、学习精度高和资源占用少的优势. 最后, 结合民航飞机实时飞行仿真系统的研究, 利用RMBP算法对一组参数数据进行了学习记忆, 证明了该算法的有效性.

## 1 BP神经网络

BP神经网络是多层网络, 由网络节点和节点间的连接组成, 包括1个输入层、1个输出层和若干中间层, 每层包括若干个节点. 除输入层的网络节点外, 其他各层网络节点都具有阈值. 网络节点间的连接具有权值. 阈值、权值和输入通过转换函数计算输出, 利用误差函数计算输出与样本的误差, 并据此进行反向传播, 纠正各阈值和权值参数, 直到误差达到可以接受的范围之内. 图1为BP神经网络的结构示意图.

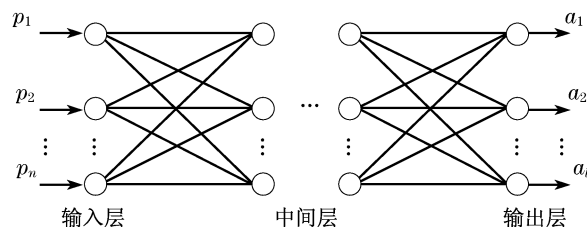


图1 BP神经网络的结构

Fig.1 Structure of BP neural network

图1中,  $p_d$  ( $d = 1, 2, \dots, n$ ) 表示输入层第  $d$  个节点的输入,  $a_g$  ( $g = 1, 2, \dots, l$ ) 表示输出层第  $g$  个节点的输出.

BP神经网络中间层和输出层节点的输入依赖于前一层节点的输出, 其公式如下:

$$x_{m+1} = \mathbf{W}_{m+1} \mathbf{a}_{m+1} + \mathbf{b}_{m+1}, \quad m = 0, 1, \dots, M-1 \quad (1)$$

式中:  $m$  为层数;  $\mathbf{x}_{m+1} = (x_{m+1,1}, x_{m+1,2}, \dots, x_{m+1,j})^T$  为第  $m+1$  层节点的输入值向量;

$$\mathbf{W}_{m+1} = \begin{bmatrix} W_{m+1,11} & \cdots & W_{m+1,i1} \\ \vdots & & \vdots \\ W_{m+1,1j} & \cdots & W_{m+1,ij} \end{bmatrix}$$

为第  $m+1$  层节点与第  $m$  层节点的连结权值矩阵, 矩阵的每一行为第  $m+1$  层某一个节点与第  $m$  层所有节点的连结权值向量.  $\mathbf{a}_m = (a_{m,1}, \dots, a_{m,i})^T$  为第  $m$  层节点的输出值向量.  $\mathbf{b}_{m+1} = (b_{m+1,1}, \dots, b_{m+1,j})^T$  为第  $m+1$  层节点的阈值向量.

第  $m+1$  层的输出为

$$\mathbf{a}_{m+1} = \mathbf{F}(\mathbf{x}_{m+1}) \quad (2)$$

其中  $\mathbf{F}(\mathbf{x}_{m+1}) = (f(x_{m+1,1}), \dots, f(x_{m+1,j}))^T$  为第  $m+1$  层的转换函数向量, 转换函数可以根据需要指定, 如 sigmoid 转换函数为

$$f(x) = 1/(1 + e^{-x}) \quad (3)$$

线性转换函数为

$$f(x) = x \quad (4)$$

网络的输入向量为  $\mathbf{a}_0 = \mathbf{p} = (p_1, \dots, p_n)^T$ , 输出向量为  $\mathbf{a} = \mathbf{a}_M = (a_1, \dots, a_l)^T$ .

BP 神经网络的误差函数为平方误差函数

$$\mathbf{E} = \frac{1}{2}(\mathbf{t} - \mathbf{a})^T(\mathbf{t} - \mathbf{a}) \quad (5)$$

其中  $\mathbf{t} = (t_1, \dots, t_l)^T$  为一个样本值向量.

BP 神经网络进行学习时首先由输入进行前向传播计算出网络输出, 然后计算出总体误差. 再进行反向传播, 根据误差值逐层修正权值和阈值, 直到误差满足要求为止. 网络学习反向传播公式为

$$\mathbf{F}_m(\mathbf{x}_m) = \begin{bmatrix} f_m(x_{m,1}) & 0 & \cdots & 0 \\ 0 & f_m(x_{m,2}) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & f_m(x_{m,s_m}) \end{bmatrix}, \quad m = M, \dots, 2, 1 \quad (6)$$

$$\mathbf{s}_M = -\mathbf{F}_M(\mathbf{x}_M)(\mathbf{t} - \mathbf{a}) \quad (7)$$

$$\mathbf{s}_m = \mathbf{F}_m(\mathbf{x}_m)(\mathbf{W}_{m+1})^T \mathbf{s}_{m+1}, \quad m = M-1, \dots, 1 \quad (8)$$

$$\Delta \mathbf{W}_m(k) = -\alpha \mathbf{s}_m (\mathbf{a}_{m-1})^T, \quad m = M, M-1, \dots, 1 \quad (9)$$

$$\Delta \mathbf{b}_m(k) = -\alpha \mathbf{s}_m, \quad m = M, M-1, \dots, 1 \quad (10)$$

$$\begin{cases} \mathbf{W}_m(k+1) = \mathbf{W}_m(k) + \Delta \mathbf{W}_m(k) \\ \mathbf{b}_m(k+1) = \mathbf{b}_m(k) + \Delta \mathbf{b}_m(k) \end{cases}, \quad m = M, M-1, \dots, 1 \quad (11)$$

## 2 RMBP 算法

BP 网络学习时, 如果网络初始参数选择不当并且系统存在局部极小点, 系统容易陷入局部极小. 在局部极小点, 系统误差仍然较高, 使系统误差减小的网络参数调整会被使系统误差增大的网络参数调整改过来; 然后参数调整又使误差减小, 接下来的参数调整又会使误差增大. 网络参数调整会因为这种反复而陷入死区, 网络也就无法跳出局部极小. 同时, 如果某一节点的输出进入了变换函数的饱和区, 该节点的参数调整将不再对节点的输出敏感, 这意味着该节点参数调整的速度大大降低, 进而导致系统学习速度的降低. 在反向传播参数调整算法中加入小的随机扰动, 打乱系统参数调整规律, 同时采用相应方法使随机扰动的影响随系统误差的减小而减小, 这样就可以使系统无法在局部极小点稳定, 引导系统进入全局极小区域.

在学习算法加入随机扰动, 会增加网络参数的振荡. 这样一方面会增加网络的学习次数, 另一方面当扰动过大时也会使网络跳出极小区域而发散. RMBP 中解决这一问题的方法主要有: ①随机扰动值选择较小的值, 这样能降低网络参数的振荡; ②在学习算法中加入动量项, 以滤除网络参数调整的高频振荡, 减小参数调整次数, 提高学习速度; ③在判断网络参数已经进入极小区域后去除随机扰动, 保证网络能够在全局极小区域稳定收敛到全局极小点.

RMBP 的随机扰动值加在网络参数的修正值上. 结合动量算法, 式(9)、(10)修改如下:

$$\Delta \mathbf{W}_m(k) = \gamma \Delta \mathbf{W}_m(k-1) - (1 - \gamma) \alpha \mathbf{s}_m ((\mathbf{a}_{m-1})^T + h(\xi_m(k)))^T \quad (12)$$

$$\Delta \mathbf{b}_m(k) = \gamma \Delta \mathbf{b}_m(k-1) - (1 - \gamma) \alpha (h \eta_m(k) + \mathbf{I}) \mathbf{s}_m \quad (13)$$

式中:  $m = M, M-1, \dots, 1$ ;  $\gamma \in [0, 1]$ ;  $\xi_m(k) = (\xi_{m,1}(k), \dots, \xi_{m,i}(k))^T$  为随机扰动值向量, 并有  $\xi_{m,j}(k) = \text{rand}[0, 0.8]$ ,  $j = 1, \dots, i$ ;

$$\eta_m(k) = \begin{bmatrix} \eta_{m,11}(k) - 1 & 0 & \cdots & 0 \\ 0 & \eta_{m,22}(k) - 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \eta_{m,ii}(k) - 1 \end{bmatrix}$$

为随机扰动值矩阵, 并有  $\eta_{m,jj} = \text{rand}[0, 0.8]$ ,  $j = 1, \dots, i$

$$h = \begin{cases} 0, & \mathbf{E} < \theta \\ 1, & \mathbf{E} \geq \theta \end{cases} \quad (14)$$

其中  $\theta$  为网络误差门限值;

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

对于节点连接权值的调整,随机扰动值具有两个功能.首先,随机扰动值能够放大权值的调整值,从而加快权值的调整速度并加快网络学习速率.即使当一些节点参数处于饱和区的时候,只要在输出数据和样本数据间存在较大误差,随机扰动值就能使节点的权值参数得到有效的调整.其次,随机扰动值能够增大权值调整的随机性,这样可以帮助网络跳出局部极小点.

对于节点阈值的调整,使阈值的调整值乘上随机扰动值也能够增加网络参数调整的随机性,有助于防止网络陷入局部极小区域.

令随机扰动值小于 1,可以降低网络的振荡,加快学习速度.经过多种样本的学习实验,发现如果随机扰动值的上限大于 0.8,会引起网络参数调整的剧烈震荡而无法收敛;如果上限小于 0.8,又会降低网络学习速度.所以随机扰动值取  $[0, 0.8]$  的范围,读者应用 RMBP 算法时针对具体情况可作适当修改.同时, RMBP 的参数修正值结合了动量算法,目的是消除高频振荡,进一步提高网络学习速度,式(12)、(13)中的  $\gamma$  即为动量因子.

但是,由于随机扰动值会增大网络参数的振荡,可能会使网络在全局最优区域也无法稳定,所以在改进 BP 学习算法中采用阶段式学习方法.开始时,在网络学习算法中加入随机扰动值,这将防止网络停留在局部极小区域并向全局最优区域靠近.当网络误差小于某一个设定的门限值  $\theta$  时,网络被认为进入了全局最优区域,则去除随机扰动值,减小网络参数振荡.如果网络参数的调整已经找到了梯度下降方向,则误差会逐渐减小并最终到达全局最小点.但如果还未找到梯度下降方向,网络可能会跳出全局最优区域,这时网络误差会高于门限值  $\theta$ ,则再加入随机扰动,使网络能够进入全局最优区域.如此反复调整,直至达到全局最小点为止.

综上所述,改进 BP 学习算法的学习过程如下:

- (1) 初始化神经网络参数;
- (2) 设定学习速率  $\alpha$  值、门限值  $\theta$  和动量因子  $\gamma$ , 令  $h = 1$ ;
- (3) 应用 RMBP 算法学习样本数据;
- (4) 如果误差达到学习要求,则停止学习并保存

网络参数;否则,如果误差小于门限值,则令  $h = 0$ ,即在学习算法中去除随机扰动值,并跳到步骤(3).否则,如果误差大于门限值,则令  $h = 1$ ,即在学习算法中加入随机扰动值,并跳到步骤(3).

### 3 异或问题学习试验

RMBP 算法通过随机参数调整,使网络无法在局部极小点稳定,并最终到达全局极小点.对于存在局部极小点的数据,如果网络参数初值选取不当,普通 BP 算法会陷入局部极小而无法收敛, RMBP 算法却能使网络跳出局部极小点,向全局极小点收敛. RMBP 采用分阶段学习方法,在全局极小区域,去除随机参数调整,减小网络振荡,保证网络能够收敛到很高的精度.同时,随机参数值的加入加大了网络参数调整力度,可以加快网络的学习速度.当然,不当的初值虽然不会影响 RMBP 算法向全局极小点收敛,却会增加网络的调整次数,延长学习时间.

随机参数调整加强了网络参数调整的振荡,所以当学习的数据区域比较平坦,容易达到全局极小点时,只使用随机参数调整方法并不能使 RMBP 算法在学习速度上比普通 BP 算法有较大优势,有时甚至慢于普通 BP 算法.因此, RMBP 算法在随机参数调整的基础上加入了动量算法.动量算法相当于低通滤波器的作用,滤除了网络参数的高频振荡.动量算法的加入使 RMBP 算法可以设置很高的学习速率而不会出现剧烈振荡,大大提高了网络的学习速度.而且,当网络参数调整进入某一一致方向时,动量因子还会起到加快网络学习速度的作用.正确地设置动量因子、学习速率可以使 RMBP 算法达到很高的学习速度.

下面以异或问题为例,对 RMBP 算法的学习性能加以试验验证.

异或(exclusive-or, XOR)问题,即 XOR 问题,是比较经典的神经网络学习问题.其用单层感知器无法实现,只能采用多层神经网络解决.转换函数采用硬极限函数的多层神经网络可以解决该问题.但是,该网络的学习方法实际上为多层感知器学习算法,与典型的 BP 算法有一定区别,这里不做讨论.

应用普通 BP 算法, XOR 问题是一个典型的带有局部极小点的问题. XOR 问题的训练样本有 2 个输入值 1 个输出值,如表 1 所示.表 1 中  $X_1$  和  $X_2$  分别表示输入值 1 和输入值 2.

表 1 XOR 问题训练样本  
Tab.1 Training samples of XOR problem

输入		输出
$X_1$	$X_2$	
0	0	0
0	1	1
1	0	1
1	1	0

为了进行对比,选择普通 BP 算法、RMBP 算法、LM 算法和 MDE 算法分别进行 XOR 问题的学习.

普通 BP 算法、RMBP 算法、LM 算法均选择一个 2-2-1 神经网络,即输入层 2 个节点,一个有 2 个节点的中间层,输出层有 1 个节点.中间层的转换函数为 sigmoid 函数,输出层的转换函数为线性转换函数.网络初始权值为 0.8,初始阈值为 0.5. RMBP 算法学习速率  $\alpha = 3$ ,门限值  $\theta = 10^{-6}$ ,动量因子  $\gamma = 0.9$ .所用计算机配置为 AMD Athlon 4200 + 双核处理器,主频  $2.2 \text{ GHz} \times 2$ , 2 G 内存,操作系统为 Windows XP. RMBP 算法软件采用 Visual C++ .net 2003 实现,普通 BP 算法和 LM 算法采用 Matlab 6.5.1 的神经网络工具箱实现.

MDE 算法采用文献[9]中的学习结果对比.文献[9]中的神经网络采用了 2-3-1 网络.中间层的转换函数为 tan-sigmoid 函数,输出层的转换函数为线性函数.种群数量为 25,比例因数  $F$  为 0.5,交叉因数  $R_c$  为 0.9.计算机采用 Intel Pentium D 820 处理器,操作系统为 Windows XP.学习算法采用 Matlab 7.0 实现.

在多种目标误差下的学习结果对比如表 2 所示.

表 2 XOR 问题学习结果对比  
Tab.2 Comparison of study result of XOR problem

目标 误差	BP 算法		RMBP 算法		LM 算法		MDE 算法
	迭代 次数	时间/s	迭代 次数	时间/s	迭代 次数	时间/s	迭代 次数
$10^{-1}$	1 461	0.32	31	0.01	2	0.01	26
$10^{-2}$	2 757	0.61	42	0.01	2	0.01	70
$10^{-3}$	3 593	0.79	56	0.01	3	0.01	102
$10^{-4}$	4 405	0.97	68	0.02	3	0.01	138
$10^{-5}$	5 376	1.18	83	0.02	4	0.01	158
$10^{-6}$	6 769	1.49	87	0.02	4	0.01	190
$10^{-10}$	13 425	2.95	177	0.04	5	0.02	320
$10^{-15}$	21 783	4.79	265	0.06	6	0.02	497

因为 MDE 算法所用计算机机型和软件环境与其他几种算法不同,在计算时间上不具可比性,故表 2 中没有列出其计算时间.但是从其网络结构、种群

数量和计算算法上可知,MDE 算法单次迭代计算量不会小于 RMBP 算法的单次迭代计算量.从表 2 可以看到,普通 BP 算法的迭代次数是其他 3 种算法迭代次数的几百到几千倍,MDE 算法次之,RMBP 算法居中,LM 算法最少.当目标误差为  $10^{-15}$  时,MDE 算法迭代次数将近 2 倍于 RMBP 算法的迭代次数,因此其学习时间长于 RMBP 算法;而 RMBP 和 LM 算法所用时间均不到 1 s,且二者仅相差 0.04 s. LM 算法每次迭代都要求一个  $9 \times 9$  矩阵的逆矩阵,单步计算量较 RMBP 算法大,所以虽然迭代次数少,但所用时间与 RMBP 算法相当. LM 算法学习时占用系统资源比普通 BP 算法多很多,不适合大规模神经网络的学习. RMBP 算法学习时占用系统资源与普通 BP 算法相同,能够用于大规模神经网络的学习.因此,RMBP 算法的适用范围比 LM 算法大.

由学习结果可见,RMBP 算法能够有效克服普通 BP 算法易陷入局部极小和收敛速度慢的缺陷,而且学习速度快,精度高.与其他改进算法相比,它具有学习速度快、学习精度高、占用资源少的优势.

#### 4 应用举例

飞机是交通运输的重要工具.飞机飞行仿真是系统仿真的重要分支,在民航,飞行仿真是进行飞行员及其他相关飞机操作使用人员培训的主要工具.为了满足训练要求,需要飞行仿真系统具有很高的实时性和逼真度.为了满足逼真度要求,飞行仿真系统每个运算周期都要进行大量的查表和插值运算.据研究,在固定翼飞机飞行仿真系统中,气动系数的计算时间是解算六自由度运动方程所用时间的 21.4 倍,而气动系数计算时间主要花费在查表和插值运算上<sup>[16]</sup>.

BP 神经网络具有记忆和联想功能.如果用 BP 神经网络学习记忆表格数据,使用时通过前向传播计算所需数据,在合理选择网络结构的情况下,其实时性和计算精度都将优于查表插值法.但因为气动参数数据的复杂性,普通 BP 学习算法不一定能够找到学习数据的全局极小点.

气动系数包括阻力系数、侧力系数、升力系数、滚转力矩系数、俯仰力矩系数和偏航力矩系数.下式为固定翼喷气式飞机阻力系数的计算公式:

$$C_d = C_{Dc} + \Delta C_{D\beta} + \Delta C_{D\epsilon} + \Delta C_{D\zeta} + \Delta C_{Dbo} + \Delta C_{Dgear} \quad (15)$$

式中: $C_{Dc}$ 为基本阻力系数; $\Delta C_{D\beta}$ 为侧滑引起的阻力

系数增量; $\Delta C_{D\epsilon}$ 为副翼偏转引起的阻力系数增量; $\Delta C_{D\zeta}$ 为方向舵偏转引起的阻力系数增量; $\Delta C_{D\delta_0}$ 为地面效应引起的阻力系数增量; $\Delta C_{Dgear}$ 为起落架收放引起的阻力系数增量.

以侧滑引起的阻力系数增量为例,表 3 为 Boeing 737-300 飞机  $\Delta C_{D\beta}$  随侧滑角  $\beta$  变化的数值表.

表 3  $\Delta C_{D\beta}$  数值表  
Tab.3 Data table of  $\Delta C_{D\beta}$

$\beta/\text{rad}$	-1.57	-0.26	0	0.26	1.57
$\Delta C_{D\beta}$	1.23	0.05	0	0.05	1.23

选择 1-2-1 结构的神经网络,隐层转换函数为 sigmoid 函数,输出层转换函数为线性转换函数.以侧滑角  $\beta$  为输入, $\Delta C_{D\beta}$  为输出,分别用普通 BP 算法和 RMBP 算法对该表格数据进行学习. RMBP 网络初始权值为 0.008,初始阈值为 0.005,学习速率  $\alpha=3$ ,门限值  $\theta=10^{-6}$ ,动量因子  $\gamma=0.9$ .在不同目标误差下的学习结果如表 4 所示.

表 4 表格 3 数据学习结果  
Tab.4 Study result of data of table 3

目标误差	迭代次数	
	BP 算法	RMBP 算法
$10^{-1}$	28	28
$10^{-2}$	失败	47
$10^{-15}$	失败	34 165

由表 4 可知,表格 3 的数据样本空间存在局部极小点,普通 BP 算法在目标误差小于等于  $10^{-2}$  时陷入局部极小而无法收敛.而 RMBP 算法有效地跳出了局部极小,并以很快的速度收敛到全局极小点.需要指出,表 4 中普通 BP 算法的网络初值与 RMBP 算法选取的不同,这是因为普通 BP 算法使用 RMBP 算法的初值即使目标误差为  $10^{-1}$  也无法收敛.表 4 中普通 BP 算法的结果是进行了多次随机初值选取后的最优结果.而且,RMBP 算法的迭代次数也会受到网络初值的影响,如果取初始权值为 0.8,初始阈值为 0.5,则误差为  $10^{-1}, 10^{-2}, 10^{-15}$  时 RMBP 算法的迭代次数分别为 37,50,13 842.对比表 4 结果,误差为  $10^{-1}, 10^{-2}$  时迭代次数略有增加,误差为  $10^{-15}$  时迭代次数却显著减少.但是采用不同的初值,RMBP 算法均能收敛到全局极小,表现出随机参数调整能够避免网络陷入局部极小的优良特性.

应用 RMBP 算法训练出的神经网络可以根据当

前飞机侧滑角的大小直接计算出侧滑引起的阻力系数增量,然后与其他阻力系数增量结合计算出当前的阻力系数.而且该神经网络节点少,计算速度快,精度高,明显优于查表插值算法.

5 结论

RMBP 算法通过在 BP 算法的参数调整中增加随机扰动值,可以避免神经网络学习陷入局部极小,促使网络找到梯度下降方向并到达全局极小点.同时,随机扰动值还具有提高网络学习速度的功能.

RMBP 算法还结合了动量算法和采用了分阶段学习的方法,有效抑制了随机扰动值造成的网络参数调整震荡,确保网络学习收敛于全局极小点,加快了网络的学习速度.

通过 XOR 问题的学习试验,证明 RMBP 算法相对普通 BP 算法和其他常用 BP 改进算法,具有学习速度快、学习精度高、占用资源少和适用范围广的优势.最后,通过飞机空气动力参数样本的学习,证明了 RMBP 算法的有效性.

参考文献:

[1] Hagan M T, Demuth H B, Beale M H. 神经网络设计[M]. 戴葵,宋辉,谭明峰,等译.北京:机械工业出版社,2002.  
Hagan M T, Demuth H B, Beale M H. Neural network design [M]. Translated by DAI Kui, SONG Hui, TAN Mingfeng, et al. Beijing: China Machine Press, 2002.

[2] 蒲春,孙政顺,赵世敏. Matlab 神经网络工具箱 BP 算法比较[J]. 计算机仿真, 2006, 23(5): 142.  
PU Chun, SUN Zhengshun, ZHAO Shimin. Comparison of BP algorithms in Matlab NN toolbox [J]. Computer Simulation, 2006, 23(5): 142.

[3] 罗四维. 人工神经网络建造[M]. 北京:中国铁道出版社,1998.  
LUO Siwei. Building of artificial neural network[M]. Beijing: China Railway Publishing House, 1998.

[4] 王静伟. BP 神经网络改进算法的研究[J]. 中国水运, 2008, 8(1): 157.  
WANG Jingwei. Study of modified algorithm of BP neural network[J]. China Water Transport, 2008, 8(1): 157.

[5] 潘昊,钟珞,陈杰. BP 神经网络训练的函数变步长搜索调整法[J]. 湖北工学院学报, 1997, 12(2): 1.  
PAN Hao, ZHONG Luo, CHEN Jie. Probing modification of BP neural network learning-rate[J]. Journal of Hubei Polytechnic University, 1997, 12(2): 1.

[6] 许廷发,张敏,顾海军,等.改进的 BP 算法在多目标识别中的应用[J]. 光学精密工程, 2003, 11(5): 513.  
XU Tingfa, ZHANG Min, GU Haijun, et al. Multi-target recognition with improved BP algorithm [J]. Optics and Precision Engineering, 2003, 11(5): 513.

- [7] 马成前,王庆喜. 基于局部及全局误差的 BP 神经网络研究[J]. 武汉理工大学学报,2009,31(20):99.  
MA Chengqian, WANG Qingxi. BP neural network research based on local and global error[J]. Journal of Wuhan University of Technology,2009,31(20):99.
- [8] 侯清兰. 神经网络等智能方法的优化及应用[D]. 武汉:武汉理工大学计算机科学与技术学院,2006.  
HOU Qinglan. The optimization and application of the neural network together with intelligence methods[D]. Wuhan:Wuhan University of Technology. College of Computer Science & Technology,2006.
- [9] LIU Suyi, ZHANG Leduo, WANG Qian, et al. BP neural network in classification of fabric defect based on particle swarm optimization[C]//Proceedings of the 2008 International Conference on Wavelet Analysis and Pattern Recognition, ICWAPR. Hong Kong, IEEE,2008:216-220.
- [10] LIU Suyi, LIU Jingjing, ZHANG Leduo. Classification of fabric defect based on PSO-BP neural network[C]//Proceedings-2nd International Conference on Genetic and Evolutionary Computing, WGEC 2008. Jingzhou. IEEE,2008:137-140.
- [11] GAO Yuelin, LIU Junmin. A modified differential evolution algorithm and its application in the training of BP neural network [C]//Proceedings of the 2008 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. Xi'an: IEEE,2008:1373-1377.
- [12] SHEN Xianjun, CHEN Caixia, HE Tingting, et al. Postgraduate entrant and employment forecasting using modified BP neural network with PSO[C]//Proceedings of 2009 4th International Conference on Computer Science and Education, ICCSE 2009. Nanning: IEEE,2009:1699-1703.
- [13] 张银雪,贾振红,刘子建. 基于改进 BP 神经网络和粒子群优化算法的图像滤波方法的研究[J]. 光电子·激光,2009,20(3):406.  
ZHANG Yinxue, JIA Zhenhong, LIU Zijian. Image filtering based on modified BP neural network and PSO[J]. Journal of Optoelectronics Laser,2009,20(3):406.
- [14] XIONG Zhibin. A modified adaptive genetic BP neural network with application to financial distress analysis[C]//Proceedings-2nd International Conference on Genetic and Evolutionary Computing, WGEC 2008. Jingzhou: IEEE,2008:149.
- [15] CHENG Qiming, DU Xufeng, ZHENG Yong, et al. The control system of GA-BP fuzzy neural network controller and elman network identifier[C]//Proceedings of the IEEE International Conference on Automation and Logistics, ICAL 2008. Qingdao: IEEE,2008:1653-1657.
- [16] 卢惠民. 飞行仿真数学建模与实践[M]. 北京:航空工业出版社,2007.  
LU Huimin. Mathematical modeling and practice of flight simulation[M]. Beijing:Aviation Industry Press,2007.

• 下期文章摘要预报 •

## 融合粗糙集和证据理论的车地无线通信设备故障诊断

涂继亮,潘洪亮,董德存,罗艳芬

针对车地无线通信设备故障诊断信息不一致的情况进行故障推理和诊断,提出了一种融合粗糙集与证据理论综合集成的车地无线通信设备故障诊断方法.该方法首先利用粗糙集剔除故障特征数据的冗余成分,提取互不相关的必须特征,将车地无线通信设备故障诊断问题用一个不同简化层次的故障决策网络表示,由网络节点根据定义出的规则置信度和覆盖度可推导出对应有效的故障诊断规则集合.在建立规则库的基础上,引入基于证据理论的信息融合技术,解决多传感器故障监测数据与诊断规则准确匹配的问题,建立故障推理机制,便可方便得出车地无线通信单元故障诊断结果.实例分析表明,该方法能有效提高模型故障诊断能力,具有较强实用性.