

多层感知器的一种快速网络训练法及其应用*

宋宜斌

(烟台大学计算机科学与工程系 264005)

摘要 从多层感知器原理分析出发,提出一种自适应学习速率因子方法,用于对多层感知器中 BP 算法的改进,并将改进算法用于 XOR 问题的学习及某分类器实例样本的学习。仿真结果表明,改进的 BP 算法可显著加速网络训练速度,学习过程具有较好的收敛性和较强的鲁棒性。

关键词 多层感知器, BP 算法, 自适应学习速率因子

分类号 TP 18

Quick Training Method for Multi-layer Perception and Its Application

Song Yibin

(Yantai University)

Abstract Based on the mechanism of multi-layer perception (MLP), a method with adaptive learning rate factors for the improvement of BP algorithm is presented. The improved algorithm is applied to the learning process of the XOR question and some classifier. The simulation results show the presented quick training algorithm can speed up the learning process of MLP, and improve the learning properties on convergence and robust performance.

Key words multi-layer perception, BP algorithm, adaptive learning rate factor

1 引言

多层感知器在结构上是一种多层前向神经网络,反向传播误差(BP)是神经网络完成学习过程的主要方法之一。其基本思想是将样本模式从网络的输入层输入,经隐含层的逐层处理后传入输出层。若输出层未得到期望的输出结果,则进行误差的反向传播,误差信号沿原连接路径返回。网络则根据反向传播的误差信号修改各层的连接权,使误差达到最小。然而,BP 算法在具体实现中常会出现一些问题,如收敛速度缓慢,局部极小等。为此,人们进行了广泛研究^[1-3],提出各种不同的改进算法。

本文从多层感知器原理分析出发,提出一种自适应学习强度因子方法,用于对 BP 算法的改进,并将改进算法用于 XOR 问题的学习及某分类器实例样本的学习。仿真结果表明,改进的 BP 算法可显著

改善和加速学习过程,并具有良好的收敛性和较强的鲁棒性。

2 多层感知器原理与 BP 算法分析

2.1 多层感知器网络结构分析

多层感知器(MLP)在网络结构上通常采用典型的三层网络结构,即由输入层、隐含层和输出层组成。它主要应用于模式识别、函数逼近及逻辑功能运算等方法。MLP 输入层中任一神经元的输出可表示为输入模式分量的加权和,这种加权和的概念同样适合于其余各层。

设某层中任一神经元 j 的输入为 net_j ,其输出为 y_j ,与此层相邻的低层中任一神经元 i 的输出为 y_i 。则有

$$net_j = \sum_i w_{ji} y_i \quad (1)$$

$$y_j = f(net_j) \quad (2)$$

其中, w_{ji} 为神经元 j 与神经元 i 之间的连接权; $f(\bullet)$ 为神经元的输出函数,通常取 S 函数

* 山东省教委科研基金项目(J98F04)

1998-08-31 收稿, 1998-11-23 修回

$$f(\text{net}_j) = \frac{1}{1 + e^{-(\text{net}_j + h_j)/\theta}} \quad (3)$$

其中, h_j 为神经元 j 的阈值, θ 用以调整 S 函数的陡度^[4]。

2.2 MLP 中的 BP 算法

设 MLP 输出层第 k 个神经元的实际输出为 y_k , 输入为 net_k , 相邻隐含层中第 j 个神经元的输出为 y_{jo} 则有

$$y_k = f(\text{net}_k) = f\left(\sum_i w_{ki}y_{ji}\right) \quad (4)$$

反向误差传播学习的目的是修改连接权 w 值, 使学习均方差达到最小值。这就要求连接权 w_{kj}, w_{ji} 沿输出层输出方差的负梯度方向学习。

当输入模式为 x_p 时, w_{kj} 的修正量应为

$$\Delta_p w_{kj} = \eta \delta_{pk} y_{pj} \quad (5)$$

$$\delta_{pk} = (o_{pk} - y_{pk}) y_{pk} (1 - y_{pk}) \quad (6)$$

其中 η 为学习速率调整因子(又称学习强度因子)。

同理可推导出, 对于在输入模式 x_p 下与输出层相邻的隐含层中, 第 j 个神经元与比该层低一层中的第 i 个神经元之间连接权 w_{ji} 的修正值为

$$\Delta_p w_{ji} = \eta \delta_{pj} y_{pi} \quad (7)$$

$$\delta_{pj} = y_{pj} (1 - y_{pj}) \delta_{pk} w_{kj} \quad (8)$$

2.3 改进的 BP 算法

由上述讨论知, 权值的收敛 ($\Delta_p w_{kj}$ 或 $\Delta_p w_{ji}$)

0) 速度主要取决于学习速度调整因子 η 学习规律 δ 以及神经元的实际输出。各层神经元连接权值的学习应沿输出层的输出方差 E_p 的负梯度方向变化。当从负梯度方向推导确定出学习规律 δ 后, 学习速率因子 η 即可对学习过程的收敛速度产生显著影响。可以根据需要来适时调整 η 甚至可采用强学习过程 ($\eta \gg 1$) 来完成权值的学习。

据此, 本文采用学习因子自适应调速法对 BP 算法进行改进, 改进的 BP 算法中采用自适应调整的学习速率因子 $f(\eta)$, 以改善学习质量, 加速学习过程。即

$$w_{ji}(t+1) = w_{ji}(t) + f(\eta) \delta_{yj} + \alpha [w_{ji}(t) - w_{ji}(t-1)] \quad (9)$$

其中, $0 < \alpha < 1$ 为势能因子^[4]。学习速率的分段自适应系数 $f(\eta)$ 的选取原则为

$$f(\eta) = \begin{cases} \eta, & t < k_1 T \\ \eta e^{-pt}, & k_1 T \leq t < k_2 T \\ \eta_b, & t \geq k_2 T \end{cases} \quad (10)$$

其中, η 为强学习因子, η_b 为常规学习因子, η_b 为弱学习因子, p 为学习速率衰减系数, T 为采样周期,

k_1, k_2, k_3 均为正整数。

3 学习过程的仿真分析

3.1 XOR 问题的学习

感知器选取 2-2-1 三层网络结构对 XOR 问题进行学习, 即输入层/隐含层/输出层的神经元个数依次为 2, 2, 1。权重初值可在 $[-4, 4]$ 内任意选取。经典 BP 算法与改进 BP 算法的各自学习结果列于表 1。

表 1 XOR 问题的学习结果

经典 BP 算法 ($\eta = 0.95$, 迭代 200 次)			
样本输入	期望输出	实际输出	均方差
0 0	0	0.156 1	0.009 7
0 1	1	0.968 4	
1 0	1	0.968 4	
1 1	0	0.226 5	
改进 BP 算法 (取分段自适应 η 值, 迭代 10 次)			
样本输入	期望输出	实际输出	均方差
0 0	0	6.644 0E - 4	5.518 2E - 8
0 1	1	1.000 0	
1 0	1	1.000 0	
1 1	0	4.634 6E - 6	

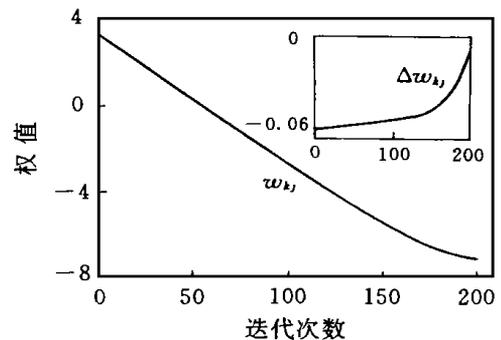


图 1 经典 BP 算法中神经元连接权因子 w_{kj} 的学习过程

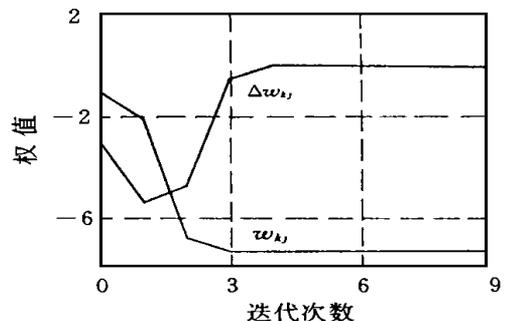


图 2 改进 BP 算法中神经元连接权因子 w_{kj} 的学习过程

表 2 MLP 学习结果(迭代计算 20 次, 均方差为 $E = 4.7234E - 024$)

期望 输出	实 际 输 出 值								
	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9
$O_1 = 1$	1.000 0 E + 000	6.761 2 E - 018							
$O_2 = 1$	2.518 1 E - 022	1.000 0 E + 000	2.518 1 E - 022						
$O_3 = 1$	2.518 1 E - 022	2.518 1 E - 022	1.000 0 E + 000	2.518 1 E + 000					
$O_4 = 1$	2.518 1 E - 022	2.518 1 E - 022	2.518 1 E - 022	1.000 0 E + 000	2.518 1 E - 022				
$O_5 = 1$	1.599 2 E - 020	1.000 0 E + 000	1.599 2 E - 020						
$O_6 = 1$	3.321 7 E - 020	1.000 0 E + 000	3.321 7 E - 020	3.321 7 E - 020	3.321 7 E - 020				
$O_7 = 1$	3.321 7 E - 020	1.000 0 E + 000	3.321 7 E - 020	3.321 7 E - 020					
$O_8 = 1$	2.518 1 E - 022	1.000 0 E + 000	2.518 1 E - 022						
$O_9 = 1$	1.030 9 E - 012	1.000 0 E + 000							

图 1, 图 2 分别给出了 BP 算法改进前后, MLP 输出层神经元与隐含层神经元之间连接权因子的学习过程, 图中 Δw_{kj} 为对应权值的变化增量。可以清楚地看到, 改进后的算法使加权因子学习过程的收敛速度明显提高。

3.2 分类器应用实例

为了进一步验证改进 BP 算法的有效性, 本文以某智能故障诊断系统为例进行讨论与分析。已知故障输入样本及对应的诊断期望输出(略), 输入样本模式数为 $m = 9$, 输入变量数为 $i = 13$, 输出变量数为 $k = 9$ 。仍对输入样本进行经典 BP 算法和改进 BP 算法的学习。为了同时研究改进后算法的鲁棒性, 对输入样本加入随机噪声污染, 仿真结果列于表 2。限于篇幅, 只给出改进 BP 算法的学习结果。

仿真结果表明, 改进后的 BP 算法可有效地提高多层感知器网络的学习速度, 迅速而高精度地完成学习任务, 同时学习过程还兼备良好的鲁棒性与稳

定性。

参 考 文 献

- 1 Hush D, Salas J. Improving the learning rate of back-propagation. In: Proc of IEEE Int Conf on Neural Networks San Diego, 1988. 441- 447
- 2 Ochiai K, Toda N, Uau S. Kick-out learning algorithm to reduce the oscillation of weights. Neural Networks, 1994, 7(5): 797- 807
- 3 徐嗣鑫, 戴友元. 前向神经网络的一种快速学习方法及其应用. 控制与决策, 1993, 8(4): 284- 288
- 4 程相君, 王春宁, 陈生潭. 神经网络原理及其应用. 北京: 国防工业出版社, 1995

作 者 简 介

宋宜斌 男, 1957 年生, 1988 年于合肥工业大学获硕士学位, 现为烟台大学计算机系副教授。目前主要从事计算机控制工程, 神经网络控制等方向的研究。