

一种改进的人工神经网络学习算法及其在超声检测中的应用

刘镇清

(同济大学声学研究所, 上海 200092)

摘要: 本文用多层感知器(MLP)与误差反向传播算法(error back-propagation algorithm)构造训练人工神经网络,提出了新的误差反向传播改进算法。试验结果表明,改进的BP算法收敛速度较之常规BP算法明显加快,因而在工业现场的超声检测领域有广阔应用前景。

关键词: 人工神经网络; 学习算法; 改进; 超声检测

中图分类号: TB553 文献标识码: A

An improved learning algorithm for artificial neural network and its application in ultrasonic testing

LIU Zhen-qing

(Institute of Acoustics, Tongji University, Shanghai 200092, China)

Abstract: The model of multilayer perceptron (MLP) and back-propagation (BP) training algorithm in artificial neural network are employed in this paper. New ideas are proposed to improve learning algorithm in aspects of learning rate for BP training. Experiment results are also presented to demonstrate the effect of improvement, which has a widely applied future for ultrasonic testing in industry.

Key words: artificial neural network; learning algorithm; improvement; ultrasonic testing

1 引言

人工神经网络在各行各业得到了愈来愈多的应用,误差反向传播算法(以下简称BP算法)由于其算法简洁性、容易理解,常被用于训练多层感知器,使人工神经网络具有预定的功能。BP算法的主要问题是网络训练耗时较长,特别在学习条件复杂、出现难以分辨的模式时,大量的样本常导致漫长的训练时间,这对工业现场的应用不利。例如,在关键设备及构件的超声无损检测工作中,需要不断将现场得到的确定的自然缺陷信息输入到人工神经网络训练,提高网络的分辨率与正确率,由于工业检修现场允许检测的时间常常很短,采用BP算法训练时,加快收敛速度、减少训练时间成为改进BP的主要目标。

人工神经网络BP算法的具体做法在若干工具软件(如:Matlab科学计算软件)都有介绍^[1],且也有人提出过诸如增加动量项(又称:冲量项)、变步长(学习率)、模拟退火方法等^[2]。本文通过理论分析

与推导,提出了一种BP算法的改进途径,它将固定不变的学习率修改成可以根据不同训练状态自适应调节的学习率,因而大大加快了BP算法的收敛速度,使之更实用化。对模拟及超声检测信号的试验表明,该自适应调节的学习率训练算法达到了预期的效果。

2 人工神经网络模型

2.1 神经元模型

本文采用的神经元模型见图1,它分成3部分,1.连接键,其权值用 w_{ji} 表示,下标 j 和 i 表示从神经元 i 到 j ,或者从输入信号 i 到神经元 j ;2.累加器;3.非线性激励函数 $\varphi(\cdot)$,它的输出值一般在 $[0, 1]$,其中 $y_0, y_1 \dots y_i$ 是输入信号; v_j 为神经元的内部激励水平; z_j 是神经元的输出。

2.2 网络结构

网络结构采用多层感知器(multilayer perceptron)简称MLP,它是前馈网络中最重要的一种,包括一个输入层(感觉单元层)、一个或多个隐层及一个输出层,每层有若干个神经元,其模型图见图2。

多层感知器每个神经元的输出端都有一个光滑(处处可导)的非线性激励函数 $\varphi(\cdot)$,本文采用输出值介于 $(-1, 1)$ 之间的双曲正切函数 $\varphi(x) = \tanh(x)$ 作为神经元的非线性激励函数,其形状见图3。

收稿日期: 2000-07-10; 修订日期: 2000-09-06

国家自然科学基金资助课题

庆祝上海市声学学会成立20周年征文

作者简介: 刘镇清(1962),男,博士,教授,博士生导师,主要从事超声波及其应用、现代信号处理与系统的研究。

3 修正学习率

经典的 BP 算法的学习率是固定的,一般在 (0, 1) 之间。但固定的学习率并非最佳选择,学习率在训练中应能自适应地调整,并且不同的网络部分应允许有不同的学习率。

设 n 代表迭代次数,也代表第 n 个输入信号; $w_{ji}(n)$ 代表第 n 次迭代时的从神经元 i 到 j 的连接键权值,令 $\eta_{ji}(n)$ 代表连接键权值 $w_{ji}(n)$ 的学习率,定义学习率修正量为:

$$\Delta \eta_{ji}(n+1) = -\gamma \frac{\partial E(n)}{\partial \eta_{ji}(n)} \quad (1)$$

其中 γ 为一正的常数,是学习率调整过程中的步幅调整参数。设:代价函数 $E(n)$ 为:

$$E(n) = \frac{1}{2} \sum_j e_j^2(n) = \frac{1}{2} \sum_j [d_j(n) - z_j(n)]^2 \quad (2)$$

式中, $e_j(n)$ 代表第 n 次迭代时输出神经元 j 的误差信号; $d_j(n)$ 代表输出神经元 j 对应第 n 个输入样本的期望输出; $z_j(n)$ 代表神经元 j 的输出信号。

$\frac{\partial E(n)}{\partial \eta_{ji}(n)}$ 可表示为:

$$\frac{\partial E(n)}{\partial \eta_{ji}(n)} = \frac{\partial E(n)}{\partial z_j(n)} \frac{\partial z_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial \eta_{ji}(n)} \quad (3)$$

$$\text{其中 } v_j(n) = \sum_{i=0}^p w_{ji}(n) y_i(n) \quad (4)$$

p 是神经元输入的个数,学习过程中的 $w_{ji}(n)$ 为:

$$w_{ji}(n) = w_{ji}(n-1) - \eta_{ji}(n) \frac{\partial \varepsilon(n-1)}{\partial w_{ji}(n-1)} \quad (5)$$

上式中, $\varepsilon(n-1)$ 代表第 $n-1$ 次迭代时总的输出误差,即:

$$\varepsilon(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (6)$$

将式(5)代入式(4)可得:

$$v_j(n) = \sum_{i=0}^p y_i(n) \left[w_{ji}(n-1) - \eta_{ji}(n) \frac{\partial \varepsilon(n-1)}{\partial w_{ji}(n-1)} \right] \quad (7)$$

$$\text{因此, } \frac{\partial v_j(n)}{\partial \eta_{ji}(n)} = -y_i(n) \frac{\partial \varepsilon(n-1)}{\partial w_{ji}(n-1)} \quad (8)$$

$$\text{同时, } \frac{\partial E(n)}{\partial z_j(n)} = -[d_j - z_j(n)] = -e_j(n) \quad (9)$$

根据图 1 有:

$$\frac{\partial z_j(n)}{\partial v_j(n)} = \varphi'_j[v_j(n)] \quad (10)$$

将式(8)、式(9)和式(10)代入式(3)可得:

$$\frac{\partial E(n)}{\partial \eta_{ji}(n)} = -\varphi'_j[v_j(n)] e_j(n) y_i(n) \left[-\frac{\partial \varepsilon(n-1)}{\partial w_{ji}(n-1)} \right] \quad (11)$$

根据式(2)、式(4)、式(6)和图 1 中的 $z_j(n) = \varphi_j[v_j(n)]$, 有

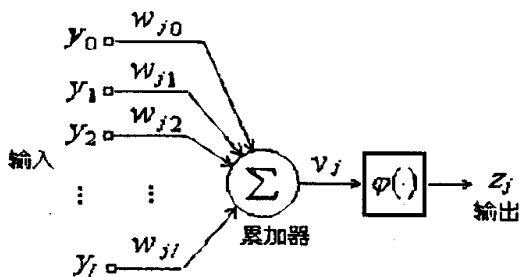


图 1 神经元模型图

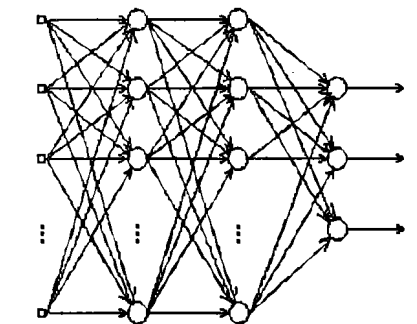


图 2 多层感知器的模型图

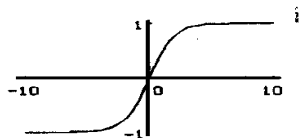


图 3 双曲正切函数

该函数相当于把常规的 Sigmoid 函数沿着纵轴向下平移至反对称的结果,理论与实验均表明,该函数较之常规的 Sigmoid 函数在处理声学信号时具有优势^[3]。

2.3 误差反向传播算法(BP)

BP 算法能对多层感知器进行监督训练,使其得以发挥多方面的能力。

总的来说, BP 算法包括前向和反向两个过程。在前向过程中,一个输入模式(矢量)被加到网络的输入层,并且沿着网络逐层向前传播,最后一个输出矢量作为对输入的响应,在此期间网络的连接键权值保持固定;在反向过程中,网络的实际输出与期望输出相比较,产生一个误差信号;误差信号从输出层开始逐层反向传播以修改权值,修改后的网络趋向于产生与期望输出更接近的信号。如此前向与反向交替进行,不断循环迭代,最终在一定的标准下使网络的输出收敛到期望输出。在经典的 BP 算法中,采用梯度算法修正权值^[2],即权矢量沿着误差信号对权矢量梯度的相反方向改变。

$$\frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} = -e_j(n) \phi_j' [v_j(n)] y_i(n) \quad (12)$$

所以式(11)可简化为:

$$\frac{\partial E(n)}{\partial \eta_i(n)} = -\frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} \frac{\partial \varepsilon(n-1)}{\partial w_{ji}(n-1)} \quad (13)$$

因此,导出的学习率修正公式为:

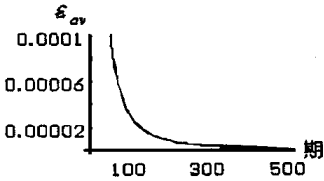
$$\Delta \eta_i(n+1) = \gamma \frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} \frac{\partial \varepsilon(n-1)}{\partial w_{ji}(n-1)} \quad (14)$$

从该学习率修正公式中可以看出,如果误差函数对权值的导数在连续两次迭代中保持同号的话,说明权值修正的步幅太小,应增大学习率,则学习率修正量为正值;如果误差函数对权值的导数在连续两次迭代中为异号的话,说明权值修正的步幅太大,应减小学习率,则学习率修正量为负值。如此起到了学习率在训练中自适应地调整的目的。

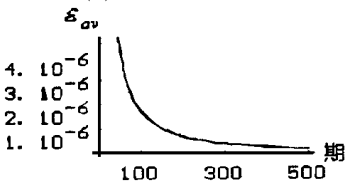
4 试验验证

为了验证新的激励函数的作用,本文按图 2 所示的网络模型,用双曲正切函数作非线性激励函数构建了网络。网络设为三层,第一层 120 个神经元,第二层 20 个,第三层 1 个。用计算机模拟数据进行网络训练的对比实验,模拟数据分为两个模式,第一个模式为 12 组 120 点长的随机数,取值范围为 $[-1, 1]$,期望输出为 1;第二个模式为 12 组叠加了幅值为 0.1 的正弦信号的 120 点长的随机数,期望输出为 -1。BP 算法的收敛曲线见图 4,横轴为训练期数,纵轴为所有训练集误差的平均值 ε_{av} 。训练时,固定学习率取 $\eta = 0.4$,式(14)中的 $\gamma = 0.3$ 。可见采用本文学习率自适应调节的收敛速度较之学习率固定的收敛速度有相当大的提高。

同样根据式(14)修改的 BP 算法,用文献[4]的超声检测数据及目标进行训练,其结果见图 5,从图

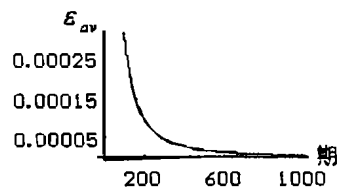


(a) 修正学习率前

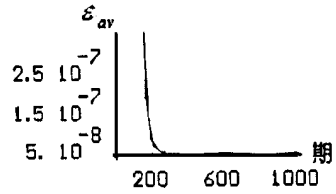


(b) 修正学习率后

4 修正学习率对训练速度的影响



(a) 修正学习率前



(b) 修正学习率后

5 改进 BP 算法对超声检测数据训练速度的影响

中可见收敛速度有了很大的提高。因此在工业现场超声无损检测的应用中我们就采用这种修改的 BP 算法作为正式的训练算法,达到了快速有效的目的。

5 结束语

本文根据常用的多层感知器网络模型 BP 算法训练速度较慢的不足,提出了学习率在训练中自适应地调整的修正方案。实验证明这个修正方案大大加快了 BP 算法训练速度,使得人工神经网络在若干时间紧迫的工业现场能够即刻训练采集到的数据,并立即用于现场分析测试,由此提高了人工神经网络在实际应用中的生命力。采用本文的人工神经网络训练方法,已成功应用于强散射材料超声无损检测的信号处理与识别问题^[4,5],并实际应用于工业生产现场的限时性检修工作^[6],取得了良好的效果。

参考文献

- [1] 张宜华. 精通 MATLAB5[M]. 北京: 清华大学出版社, 1999.
- [2] 张立明. 人工神经网络的模型及其应用[M]. 上海: 复旦大学出版社, 1993.
- [3] 李成林. 复合材料超声检测信号的小波分解与人工神经网络分析[D]. 同济大学声学专业硕士学位论文, 1997.
- [4] Z. Q. Liu, M. D. Lu and M. A. Wei. Structure noise reduction of ultrasonic signals using artificial neural network adaptive filtering[J]. Ultrasonics, 1997, 35(4): 325-329.
- [5] 刘镇清, 李成林, 刘江韦等. 超声探伤信号的时延神经网络处理[J]. 声学学报, 1997, 22(4): 297-302.
- [6] 刘镇清, 梁穗, 李成林等. 奥氏体不锈钢铸件的超声探伤信号处理与识别系统[J]. 无损检测, 1996, 18(10): 271-274.