

# 一种基于误差变化率的自适应反向传播算法<sup>1</sup>

谢克明 张建伟

(太原理工大学自动化系 太原 030024)

**摘要** 本文针对 BP 算法存在收敛速度慢的缺点, 提出一种基于网络动态训练误差变化率自动校正学习步长和冲量因子的自适应反向传播算法。异或问题的仿真结果表明, 该方法具有较快的收敛速度。

**关键词** 神经网络, 自适应反向传播算法, XOR 问题

**中图分类号** TN-052

## 1 引言

BP 学习算法收敛速度慢的缺点已经成为神经网络能否广泛应用的关键问题之一。理论分析<sup>[1]</sup>和仿真实验表明: 在网络学习的初始阶段, 较大的学习步长和冲量因子有利于加快收敛速度; 在极小点附近, 过大的学习步长和冲量因子又可能使权重调整幅度过大而发生振荡。由此可知, 在网络训练过程中, 单一的、固定不变的学习步长和冲量因子很难适应不同训练阶段的收敛情况。近年来, 为此所提出的诸多改进算法在兼顾收敛速度和算法复杂性上还存在一定的困难<sup>[2]</sup>。

本文通过引入动态误差变化率和参数细调切换界限, 提出一种基于误差变化率实时校正学习步长和冲量因子的自适应反向传播算法。异或问题的仿真表明, 该方法具有算法简单、收敛速度快等优点。

## 2 基于误差变化率的自适应反向传播算法

基于误差反传的标准 BP 算法的权重和阈值修正公式为

$$\left. \begin{aligned} \Delta w(n) &= -\xi(\partial E/\partial w(n)) + \eta\Delta w(n-1), \\ w(n) &= w(n-1) + \Delta w(n); \end{aligned} \right\} \quad (1)$$

其中  $n$  为迭代次数;  $E$  为误差函数;  $\xi$  为学习步长;  $\eta$  为冲量因子。

为描述网络训练误差的变化情况, 我们定义动态误差变化率:

$$\frac{dE(n)}{dn} = \frac{E(n) - E(n-1)}{E(n-1)}. \quad (2)$$

其大小在一定程度上反映了学习参数对网络收敛速度的动态适应情况。据此, 本文给出如下的学习步长和冲量因子的调整方案:

若本次学习误差小于上次学习误差, 则本次迭代有效, 同时说明当前学习步长适应误差变化的趋势。因此, 在下次迭代中可适当增大学习步长:

$$\xi(n+1) = \xi(n) \cdot [1 + \alpha \cdot e^{-|dE(n)/dn|}]. \quad (3)$$

<sup>1</sup> 1997-01-30 收到, 1997-10-13 定稿  
山西省回国留学人员基金资助项目

反之, 若本次学习误差大于上次的学习误差, 则本次迭代无效。此时应在原调整方向上适当减小学习步长, 重新进行有限次的调整。此时, 学习步长的调整公式为

$$\xi(i) = \xi'(i-1) / [1 + \alpha \cdot e^{-|dE'(i-1)/dn|}], \quad (4)$$

$$dE'(i-1)/dn = (E'(i-1) - E(n-1)) / E(n-1), \quad \xi(0) = \xi(n), \quad E'(0) = E(n), \quad i = 1, 2, \dots, n_0.$$

这里,  $\alpha$  为大于零的步长修正因子, 用来决定学习步长的调整幅度;  $i$  表示误差调整的次数;  $\xi'(\cdot)$ 、 $E'(\cdot)$  和  $dE'(\cdot)/dn$  分别表示误差调整过程中的学习步长、学习误差和学习误差的变化率;  $n_0$  表示误差调整的最大次数。

若经过  $i (i \leq n_0)$  次误差调整后, 学习误差下降, 则进行下一次学习。此时的学习步长和学习误差分别为

$$\left. \begin{aligned} \xi(n+1) &= \xi'(i), \\ E(n+1) &= E'(i). \end{aligned} \right\} \quad (5)$$

反之, 若经过  $n_0$  次误差调整后, 学习误差始终没有下降, 则应进行假饱和点和局部极小点的处理, 即文献 [3] 的方法。

冲量因子的调整方法与学习步长相同, 限于本文的篇幅, 不在赘述。

此外, 由于在学习误差的不同范围内, 学习步长和冲量因子对收敛速度的影响情况不同。因此, 我们进一步引入步长和冲量细调的切换界限  $E_s$ , 由此确定不同误差范围内步长修正因子和冲量修正因子的大小:

$$\left. \begin{aligned} \alpha &= \alpha_0, \quad \beta = \beta_0, \quad E(n) > E_s; \\ \alpha &= \alpha_1, \quad \beta = \beta_1, \quad E(n) \leq E_s; \end{aligned} \right\} \quad (6)$$

这里  $\beta$  为冲量修正因子;  $E_s$  一般取为最大样本误差的 100 ~ 1000 倍;  $\alpha_0$ 、 $\alpha_1$ 、 $\beta_0$  和  $\beta_1$  的取值通过仿真试验确定。一般有  $\alpha_0 > \alpha_1$ 、 $\beta_0 > \beta_1$ 。

### 3 仿真实例

以二维异或 (XOR) 问题为例, 采用 3 层 BP 网络实现异或分类。输入层、隐层和输出层的节点数分别为 2、2、1。仿真过程中使用两组 [0,1] 之间的随机数作为连接权重和阈值的初始值, 取算法的收敛条件为

$$\left. \begin{aligned} n_{\max} &\leq 15000, \\ \max_p (t^p - y^p)^2 &\leq 10^{-5}, \end{aligned} \right\} \quad (7)$$

其中  $n_{\max}$  为最大训练次数;  $t^p$  和  $y^p$  分别为样本  $p$  的教师信号和实际输出信号。

按照表 1 所示的训练参数分别对标准 BP 算法和自适应 BP 算法进行仿真。在自适应算法的仿真中, 取  $E_s=0.01$ 、 $n_0=30$ 。由所得的仿真结果可知, 本文提出的方法明显优于标准 BP 算法。上述仿真过程进一步表明本文算法在极小点附近能较快地收敛。

大量仿真试验表明, 自适应 BP 算法的收敛速度一般比标准 BP 算法的收敛速度高 10~100 倍。

表 1 标准 BP 算法和自适应 BP 算法的仿真

仿真序号	初值	训练算法	训练参数	训练次数 (次)	训练时间 (s)
1	I	标准 BP	$\xi = 0.6, \eta = 0.9$	4660	6034.8
2	I	自适应 BP	$\xi = 0.6, \eta = 0.9, \alpha_0 = \beta_0 = 12, \alpha_1 = \beta_1 = 0.6$	207	271.83
3	II	标准 BP	$\xi = 0.6, \eta = 0.9$	4072	6071.6
4	II	自适应 BP	$\xi = 0.6, \eta = 0.9, \alpha_0 = \beta_0 = 12, \alpha_1 = \beta_1 = 0.6$	151	193.88

#### 4 结 论

标准的 BP 算法由于采用单一的、固定不变的学习步长和冲量因子, 致使算法的收敛速度较慢。本文通过定义动态误差变化率提出一种训练参数调整方案。同时, 在算法实现过程中引入参数细调的切换界限, 有效地改善了网络在极小点附近的收敛速度。异或问题的仿真实例表明, 我们所提出的算法优于标准的 BP 算法。

#### 参 考 文 献

- [1] 管霖, 程时杰, 陈德树. 适用于控制型神经网络的快速学习算法. 华中理工大学学报, 1995, 23(4): 29-33.
- [2] 张建伟. 模糊神经网络 TS 模糊模型辨识及其在预测控制中的应用: [硕士学位论文]. 太原: 太原工业大学自动化系, 1997, 4.
- [3] 孙德保, 高超. 一种实用的克服局部极小的 BP 算法. 信息与控制, 1995, 24(5): 284-287.

## AN ADAPTIVE BACKPROPAGATION ALGORITHM BASED ON ERROR RATE OF CHANGE

Xie Keming    Zhang Jianwei

(*Department of Automation, Taiyuan University of Technology, Taiyuan 030024*)

**Abstract** In order to overcome slow convergence rate of the standard BP algorithm, this paper presents an adaptive backpropagation algorithm which can update learning rate and birr factor automatically based on dynamic training error rate of change. Simulation result of the XOR problem shows much faster convergence rate can be obtained by this algorithm.

**Key words** Neural networks, Adaptive backpropagation algorithm, XOR problem

谢克明: 男, 1944 年生, 教授, 硕士生导师, 目前主要研究方向为复杂工业过程控制、智能控制理论及应用、分散型控制系统。

张建伟: 男, 1973 年生, 硕士, 主要从事神经网络、模糊控制和预测控制的研究。