

# 加速 BP 网络在单桩沉降预估中的动态适应<sup>\*</sup>

韦 鹏 生

(华侨大学土木工程系, 泉州 362011)

**摘要** 提出一种学习率自适应加速算法, 即采用动态的网络学习率代替在整个学习过程中根据经验选取的传统的学习率, 使 BP 神经网络的收敛速度得到显著的提高. 应用这种神经网络, 对单桩沉降预估过程以及根据信息扩散原理所取得的样本-教师模式匹配, 可进行动态的学习. 数值模拟表明, 加速的学习算法使这种神经网络对问题的解具有动态适应性. 采用它作单桩沉降预估, 能缩短网络的学习时间. 同时, 在经过相同的学习轮数或学习时间之后, 其预估精度要优于标准的 BP 网络.

**关键词** BP 网络, 学习率, 沉降预估, 信息扩散法, 动态适应

**分类号** TU 470: TP 18

信息扩散法<sup>[1]</sup>是一种能在一组具有特定物理意义的、结构较为复杂的数据和另一组结构较为简单的数据之间, 建立一种可逆的映射. 从而, 在此基础上求解更复杂问题的一种数值分析方法. 文献[2]用信息扩散方法建立了用长径比预估工作荷载下单桩的沉降模型, 然后用 BP 网络来对上述模型中的样本-教师模式匹配进行学习. 神经网络法预估单桩沉降的准确度要优于一般常用的弹性理论法、三维线弹性有限元法和半理论半经验法. 本文的工作是将一种加速的学习算法<sup>[3]</sup>应用于上述 BP 网络, 使其收敛速度得到了显著的改善, 从而缩短了网络的学习时间. 在经过相同的学习时间后, 它对单桩沉降的预估精度要高于文献[2]所构造的网络.

## 1 BP 网络模型的一般形式

标准 BP 网络模型如图 1 所示. 其中  $x^{(1)}$  为输入层的输入向量,  $x^{(n)}$  为输出层的输出向量(设第  $n$  层为输出层).  $x^{(2)}, x^{(3)}, \dots, x^{(n-1)}$  为隐层的输出向量;  $y^{(1)}, y^{(2)}, \dots, y^{(n)}$  为各层神经元的阈值;  $M^{(1)}, M^{(2)}, M^{(3)}, \dots, M^{(n-1)}$  为网络各层

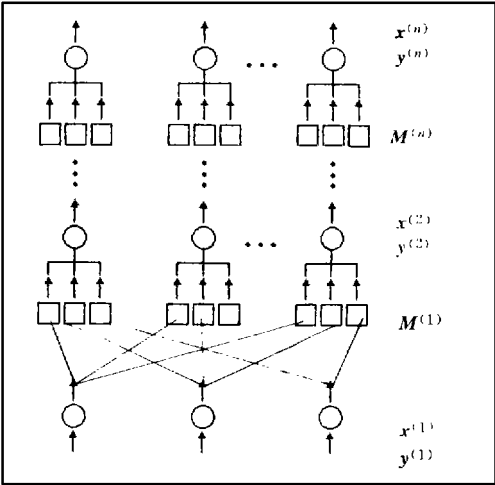


图 1 标准 BP 神经网络模型

间的连接权值. BP 网络采用非线性规划理论中的最快下降法对网络的权值和阈值进行调整, 以达到样本和教师的模式匹配<sup>[8]</sup>.

## 2 加速的学习算法<sup>[8,5]</sup>

神经网络的收敛速度与网络学习率  $\alpha$  和动量项  $\eta$  密切相关. 不同的网络具有不同的最佳学习率, 甚至同一网络对于不同的样本和(或)教师也存在不同的最佳学习率. 但到目前为止的实际应用中,  $\alpha$  和  $\eta$  的取值几乎全部是通过经验来确定的<sup>[2,6~8]</sup>, 故使用者经验丰富与否就决定了神经网络的收敛速度, 显然是不合理的. 因此, 应提出一个动态的网络学习率来代替不变的经验值, 使网络对问题具有解的自适应能力.

文献 [8] 构造的 SCNN 网络中, 神经元转换函数采用了全导通函数, 即

$$v_i^{k+1} = f(\sum_j \omega_j^{k,k+1} v_j^k) = \sum_j \omega_j^{k,k+1} v_j^k, \quad (1)$$

其中  $v_i^{k+1}$  为第  $k+1$  层第  $i$  神经元的输出,  $\omega_j^{k,k+1}$  为第  $k$  层第  $j$  神经元到第  $k+1$  层第  $i$  神经元的连接权值,  $v_j^k$  为第  $k$  层第  $j$  神经元的输出. 因而, 对于文献 [8] 所构造的 3 层网络(仅包含输入层、1 个隐层和输出层), 应当采用式 (1) 作神经元转换函数, 并令输入层节点数等于输出层节点数. 此时因

$$\begin{aligned} x_i^{(3)}(t+1) &= \sum_{j=1}^m [M_{ij}^{(2)}(t) + \alpha(x_i^{(1)} - x_i^{(3)})x_j^{(2)}(t)] \quad x_j^{(2)}(t+1) = \\ &= \sum_{j=1}^m M_{ij}^{(2)} x_j^{(2)} + \alpha \sum_{j=1}^m (x_i^{(1)} - x_i^{(3)})(x_j^{(2)}(t))^2 = \\ &= x_i^{(3)}(t) + \alpha(x_i^{(1)} - x_i^{(3)}) \sum_{j=1}^m (x_j^{(2)}(t))^2, \end{aligned} \quad (2)$$

若取

$$\alpha = 1 / \sum_{j=1}^m (x_j^{(2)}(t))^2, \quad (3)$$

可得

$$x_i^{(3)}(t+1) = x_i^{(3)}(t) + (x_i^{(1)} - x_i^{(3)}). \quad (4)$$

式(2)~(4)中,  $m$  为隐层节点数目,  $t+1$  代表  $t$  时刻的下次迭代, 其余各参量意义如图 1 所示.  $\alpha$  是关于  $x_i^{(2)}$  ( $i=1, 2, \dots, m$ ) 的函数, 因而也是关于  $x_i^{(1)}$  ( $i=1, 2, \dots, n$ ) (设  $n$  为输入层的节点数) 的函数. 可见,  $\alpha$  不仅能随具体问题的不同而自动变化, 而且在同一问题的演变过程中也能自动进行调整. 式(4)表明, 当网络进行样本与教师的模式匹配时, 实际输出与教师的误差始终以最大的权值和阈值修正率从输出层倒传递到输入层. 然而, 式(3)与输入层和输出层的节点数目无关, 故可将其推广到输入层节点数不等于输出层节点数的情形, 同时令  $\eta = \alpha$ . 如神经元转换函数取 Sigmoid 函数, 则式(3)改写为

$$\begin{aligned} \alpha &= \text{sig}[1 / \sum_{j=1}^m (x_j^{(2)}(t))^2] = \\ &= 1 / \{1 + \exp[-1 / \sum_{j=1}^m (x_j^{(2)}(t))^2]\}. \end{aligned} \quad (5)$$

数值模拟表明, 这种动态的学习算法能使网络以更快速度收敛, 故称之为加速的网络学习算法. 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

### 3 加速 BP 网络的动态学习效果

本文用 Borland C<sup>++</sup> 实现上述的加速 BP 网络<sup>[1]</sup>. 其中, 输入层节点为 8 个, 隐层节点为 25 个(只有 1 个隐层), 输出层节点为 6 个. 网络学习率可由式(5)确定. 用信息扩散法得到的样本和教师如表 1 所示<sup>[1]</sup>, 网络收敛过程如表 2 所示.

表 1 网络样本和教师

| 样    本   | 教    师           |
|--|------------------|
| 0.000, 0.000, 0.000, 0.000, 0.019, 0.267, 0.504, 0.209 | 1, 0, 0, 0, 0, 0 |
| 0.000, 0.000, 0.100, 0.101, 0.269, 0.241, 0.249, 0.041 | 0, 1, 0, 0, 0, 0 |
| 0.000, 0.000, 0.325, 0.337, 0.337, 0.000, 0.000, 0.000 | 0, 0, 1, 0, 0, 0 |
| 0.441, 0.000, 0.559, 0.000, 0.000, 0.000, 0.000, 0.000 | 0, 0, 0, 1, 0, 0 |
| 0.222, 0.222, 0.333, 0.222, 0.000, 0.000, 0.000, 0.000 | 0, 0, 0, 0, 1, 0 |
| 0.333, 0.333, 0.333, 0.000, 0.000, 0.000, 0.000, 0.000 | 0, 0, 0, 0, 0, 1 |

表 2 网络的收敛过程

| 迭代次数  | 网络实际输出                                   | $\alpha$ | 与教师的总误差     |
|-------|--|----------|-------------|
| 1 297 | 0.658, 0.711, 0.225, 0.379, 0.464, 0.179 | 0.736    | 0.083 602 7 |
| 1 298 | 0.591, 0.751, 0.187, 0.306, 0.372, 0.156 | 0.738    | 0.059 731 8 |
| 1 299 | 0.496, 0.665, 0.459, 0.257, 0.297, 0.139 | 0.744    | 0.101 783 1 |
| 1 300 | 0.410, 0.605, 0.244, 0.509, 0.232, 0.123 | 0.757    | 0.084 041 1 |
| 1 301 | 0.334, 0.522, 0.196, 0.297, 0.404, 0.107 | 0.754    | 0.074 898 4 |
| 1 302 | 0.273, 0.437, 0.164, 0.242, 0.285, 0.496 | 0.753    | 0.078 231 1 |
| 5 185 | 0.972, 0.020, 0.000, 0.000, 0.000, 0.012 | 0.677    | 0.000 107 7 |
| 5 186 | 0.022, 0.969, 0.015, 0.000, 0.002, 0.000 | 0.709    | 0.000 131 6 |
| 5 187 | 0.000, 0.024, 0.972, 0.016, 0.022, 0.000 | 0.682    | 0.000 148 6 |
| 5 188 | 0.000, 0.000, 0.011, 0.973, 0.004, 0.011 | 0.716    | 0.000 060 6 |
| 5 189 | 0.000, 0.003, 0.022, 0.006, 0.973, 0.020 | 0.753    | 0.000 161 6 |
| 5 190 | 0.013, 0.000, 0.000, 0.015, 0.030, 0.975 | 0.719    | 0.000 151 4 |
| 7 777 | 0.990, 0.010, 0.000, 0.000, 0.000, 0.002 | 0.676    | 0.000 053 6 |
| 7 778 | 0.008, 0.994, 0.002, 0.000, 0.001, 0.000 | 0.707    | 0.000 064 0 |
| 7 779 | 0.000, 0.008, 0.990, 0.004, 0.010, 0.000 | 0.679    | 0.000 072 1 |
| 7 780 | 0.000, 0.000, 0.008, 0.991, 0.002, 0.001 | 0.712    | 0.000 031 1 |
| 7 781 | 0.000, 0.002, 0.005, 0.003, 0.991, 0.007 | 0.748    | 0.000 076 1 |
| 7 782 | 0.002, 0.000, 0.000, 0.004, 0.012, 0.998 | 0.715    | 0.000 074 5 |

神经网络学习完毕后, 可将单桩长径比经正向信息扩散, 得到神经网络的输入向量. 尔后, 将此向量输入训练好的神经网络, 得到相应的输出向量. 最后, 再将输出向量经反向信息扩散, 即得到该单桩的沉降预估值<sup>[1,2]</sup>.

### 3 结束语

本文用动态的网络学习算法代替传统的经验取值法, 使神经网络的收敛速度得到了显著的提高. 因此, 网络在经过相同的学习时间后, 其样本-教师模式匹配更为精确, 从而提高了它

的参数预估精度. 表2数据表明, 基于动态学习算法的加速BP网络对样本-教师模式匹配, 具有较高的自适应能力. 另外, 由于神经网络具有良好的联想和容错功能, 故用它作参数预报, 可解决其它方法所不能或很难解决的难题. 诸如可解决个别样本包含矛盾信息, 甚至包含错误信息等的情形. 场地土具有高度的离散性质, 因而应用具有离散信息处理特性的神经网络, 以处理岩土工程问题是十分恰当的.

## 参 考 文 献

- 1 黄崇福. 模糊人工神经网络在地震工程中的应用. 地震工程与工程振动, 1993, 13(2): 63~67
- 2 王建华. 神经网络法预估水泥搅拌单桩沉降. 土木工程学报, 1996, 29(1): 55~61
- 3 韦鹏生. 用结构可控神经网络作结构分析. 华侨大学学报(自然科学版), 1998, 19(3): 275~279
- 4 庄镇泉, 王煦法, 王东生. 神经网络与神经计算机. 北京: 科学出版社, 1992. 18~30
- 5 焦李成. 神经网络计算. 西安: 西安电子科技大学出版社, 1995. 296~432
- 6 文 毅, 武广号. 遗传算法与神经网络协作求解结构优化问题. 土木工程学报, 1996, 29(5): 24~29
- 7 姜绍飞, 刘之洋. 人工神经网络在外包钢混凝土框架柱设计参数预报中的适应性. 土木工程学报, 1997, 30(1): 73~77
- 8 Yeh Y C, Kuo Y H, Hsu D S. Building a KBES for diagnosing PC pile with artificial neural network. J. Computing in Civil Engineering, ASCE, 1993, 7(1): 53~59
- 9 严桂兰, 刘甲耀. C语言图形化生成技术(I)CPAS-I系统的研制. 华侨大学学报(自然科学版), 1994, 15(1): 97~101

# Dynamic Adaptation of Accelerated BP Network in Preestimating the Settlement of Single Pile

Wei Pengsheng

(Dept. of Civil Eng., Huaqiao Univ., 362011, Quanzhou)

**Abstract** For enhancing convergence speed of BP neural network, an adaptive accelerated algorithm of learning rate is advanced. That is to say, the conventional learning rate which is chosen according to experience during the entire process of learning will be replaced by the dynamic learning of the network. By applying this neural network, the settlement of single pile is preestimated; and dynamic learning of sample-teacher pattern matching based on information diffusion principle can be carried out. As indicated by numerical simulation, the accelerated learning algorithm makes the problem settlement of this neural network to have dynamic adaptability. When it is adopted to preestimate the settlement of single pile, the learning of network can be shortened, moreover, its preestimate is better than standard BP network in accuracy after similar round or time of learning.

**Keywords** BP network, learning rate, preestimate of settlement, information diffusion method, dynamic adaptation