

人工神经网络用于岩体工程的方法改进

刘勇健¹, 刘义建²

(1. 广东工业大学 建设学院, 广东 广州 510643;

2. 湖南地勘局 4073 队, 湖南 怀化 418000)

摘要: 岩体工程问题往往具有高度的复杂性和非线性, 而人工神经网络是非线性的动力系统, 是解决这类问题最有效方法之一. 但经典 BP 网络模型存在着收敛速度慢、稳定性差、易陷入局部极小的缺陷. 本文针对神经网络拓扑结构、数据的预处理、学习算法三方面的改进作了较系统的介绍. 实例研究表明改进后的 BP 网络性能良好, 该方法在工程中具有广阔的应用前景.

关键词: 人工神经网络; BP 网络; 拓扑结构; 岩体工程

中图分类号: TP301.6; TU441.4

文献标识码: A

文章编号: 1007-7162(2002)01-0021-05

人工神经网络是由大量模拟生物神经元功能的简单处理单元相互连接而成的巨型复杂网络. 目前已被广泛应用于模式识别、信号处理、知识工程、优化组合和智能控制等领域. 近年来, 它开始用于解决岩体工程问题. 由于岩体工程常包含一个巨大天然或人工的岩体, 岩体的工程性质受地质作用、岩土体的结构及所处环境条件等因素的影响, 常常存在很大的差异性, 岩体工程同时还涉及结构与岩土体的相互作用等复杂问题. 因此, 岩体工程问题往往是一类复杂的高度非线性问题. 而人工神经网络作为一门迅速兴起的非线性科学, 尤其在处理信息十分复杂、背景知识不清楚、推理规则不明确的问题具有独特的优越性.

人工神经网络在诸如边坡的稳定性分析、砂土的液化评判、地基沉降预测、地基及单桩承载力计算、基坑变形预测、岩土工程的可靠性分析、支护结构的优化设计等方面取得了非常理想的效果. 在众多神经网络模型中, BP 网络模型应最广泛, 笔者将针对应用 BP 网络模型过程中常遇到的问题和改进方法作进一步探讨.

1 BP 网络模型及存在问题

1.1 BP 网络模型

如图 1 所示, 一般 BP 网络是由输入层(n 个神经元)、隐含层、输出层(m 个神经元)组成, BP 网络的学习过程是一种误差修正型学习算法, 由正向传播和反向传播组成. 在正向传播过程中, 输入信号从输入层通过作用函数后, 逐层向隐含层、输出层传播, 每一层神经元状态只影响下一层神经元状态. 如果在输出层得不到期望的输出, 则将误差反向传播以调整连接权值, 使得输出误差信号最小.

通过样本训练就完成从输入的 n 维欧氏空间到输出的 m 维欧氏空间的映射. 因此, 可用

收稿日期: 2001-05-30

基金项目: 广东工业大学青年基金资助项目(982033); 重点学科资助项目(20216)

作者简介: 刘勇健(1968), 女, 副教授, 主要研究方向为岩土工程.

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.>

于非线性分类、预测等问题, 作为函数计算器, 能以任意精度逼近任何非线性函数^[1].

1.2 BP 网络模型存在的主要问题

经典 BP 网络主要问题表现为: ①由于采用误差梯度下降法, 当学习率选取不当或学习过程中出现“平台”时, 易落入局部极小而不是全局最小; ②当样本太多或学习率太小, 收敛速度慢, 有时甚至出现“瘫痪”现象; ③网络的泛化与推广能力较差, 主要是由于样本太少, 或输入的参数代表性不强造成的.

经典 BP 网络具有收敛速度慢、稳定性差, 易陷入局部极小的缺陷, 大大限制了神经网络在工程中的应用和推广, 深入研究其原因和改进的方法具有一定的理论意义和实用价值.

2 BP 网络模型改进方法探讨

根据笔者对 BP 网络模型的多年研究发现, 影响 BP 网络性能的主要因素可归纳为三个方面: 即神经网络的拓扑结构、数据的处理、BP 的学习算法, 下面针对三方面存在的问题, 提出行之有效的改进措施.

2.1 优化网络结构

BP 算法对网络结构非常敏感, 不同的网络结构解决问题的能力不同, 神经网络的结构越复杂, 它处理复杂非线性问题的能力越强, 但训练时间也长; 如果神经网络结构过于简单, 则网络很难收敛或即使收敛但时间太长. 长期以来, 由于所建立神经网络结构比较庞大, 而使求解问题的速度和效率不高.

神经网络的拓扑结构由网络的层数、各层的节点数及节点之间的连接等组成. 如图 1 所示, 对于经典 BP 网络来说只有相邻层上的节点相连接, 同一层上的节点不连接. 在前向式网络拓扑结构中, 输入节点和输出节点的个数是由求解问题本身决定的, 且与参数的选取和预处理相关. 因此, 网络结构的优化设计, 重点在于确定隐含层的层数隐节点的个数.

隐含层的选取需根据问题的复杂性而定. 据研究表明, 隐含层层数增加, 可形成更加复杂的决策域, 使网络解决复杂及非线性问题的能力增强, 但太多的隐含层会使学习时间延长, 对于一般的非线性问题, 用一个隐含层就能满足要求. 隐节点的选取通常较困难. 隐节点过少, 学习过程可能很难收敛或时间过长, 隐节点太多, 网络性能下降且学习时间长. 为了找到合适的节点数, 一种方法是采用自构学习算法^[2], 自构学习分两个阶段, 在预估阶段, 网络根据问题的复杂程度给定节点数($2N+1$), 在自构阶段, 网络根据学习的情况调动节点数, 最终得到一个最优网络结构. 另一种方法是用遗传算法^[3,4]对神经网络进行优化.

2.2 数据的预处理

网络学习前的数据处理对神经网络具有至关重要的作用, 它影响到网络的学习速度和精度. 资料的整理包括网络输入变量的确定和输入参数的预处理.

当所收集的网络输入变量较多, 而且各变量之间又有一定的相关性时, 应首先对这些变量进行相关分析, 力争变量之间的独立性, 建立一个简捷有效的网络模型. 例如震级和烈度两参数具有相关性, 则可选取其中之一即可.

为了使输入变量的绝对值不影响网络的精度和收敛速度, 在样本学习之前应根据变量的类

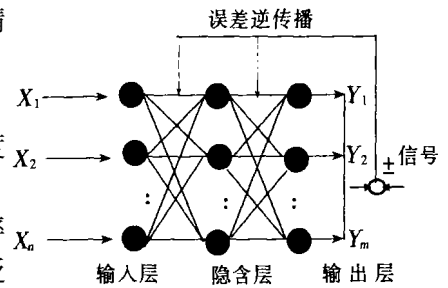


图 1 BP 网络的拓扑结构图

型作相应预处理.

2.2.1 连续型变量

连续变量在岩土工程问题中较常见, 对于此类变量, 进行归一化处理, 即

$$X^{\#} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}. \quad (1)$$

式中: X 为真实值, $X_{\max}$ 、 $X_{\min}$ 分别为真实值中的最小值与最大值, $X^{\#}$ 为归一化后的取值. 如果考虑到 BP 算法的激励函数, S 型函数在自变量接近于 0, 1 时曲线平缓, 收敛速度慢, 文献[3]建议输入变量与输出数据变换在 $[0.1 \sim 0.9]$ 或 $[0.2 \sim 0.8]$ 之间, 这样 S 型函数在该区间内变化梯度增大, 网络收敛速度提高, 网络性能改善. 其变化为

$$\begin{aligned} X^{\#} &= 0.1 + \frac{0.8(X - X_{\min})}{X_{\max} - X_{\min}} & X^{\#} &\in [0.1, 0.9], \\ X^{\#} &= 0.2 + \frac{0.6(X - X_{\min})}{X_{\max} - X_{\min}} & X^{\#} &\in [0.2, 0.8]. \end{aligned} \quad (2)$$

2.2.2 定性变量

如岩体结构类型、地基处理方法、节理粗糙程度等, 对于定性变量首先分类, 再参照专家的经验知识, 按照各自的实际意义转换成整数范围内分布的数字, 再按连续型变量来处理. 如岩体结构类型取值, 整体块状结构取 1、层状结构取 2、碎裂结构取 3、散体结构取 4.

2.2.3 区间变量

在变量变化范围较小的区域内采样时, 其离散值具有相同的数量级, 且前几位有效数字相同, 仅后几位有效数字不同, 这样的密集型数据构成的样本集, 采用文献[4]方法处理. 变化较为复杂, 但此类变量在岩土工程中不多见.

2.3 改进 BP 学习算法

2.3.1 改进激励函数

激励函数的作用是将可能的无限域变换到一指定的有限域范围内输出. 经典 BP 算法常选用 S 型函数作为激励函数, S 型函数的动态范围为 $[0, 1]$, 我们将函数值接近 0 或 1 的定义域两端定义为饱和区, 中间为非饱和区. 在饱和区函数值对自变量的变化不敏感. 所以, 当网络训练处于饱和区, 且网络的实际输出与期望输出相差较大时, 而激励函数的导数 $f'(x) = f(x)(1 - f(x))$ 已很小或接近 0, 此时梯度近于 0, 即处于误差梯度平坦区, 就会落入局部极小点或网络呈现“瘫痪”状态.

为了加快学习速度, 一种改进方法是将输出层的激励函数改为 $f(x) = \frac{1}{1 + e^{-\beta x}}$, 该函数能扩大 S 型函数的非饱和区, 改善网络性能; 有学者提出隐含层激励函数可取反正切函数、双曲正切函数或高斯函数等, 原因在于它们的导数恒大于 S 函数的导数, 故优于 S 型函数. 第二种办法则是在对数据预处理是避开饱和区或缩小饱和区. 如前所述, 实验表明收敛时间平均减少 30% ~ 50%.

2.3.2 改进代价函数

经典的 BP 网络采用误差平方和作为代价函数, 即 $E = \frac{1}{2} \sum_i (y_i^* - y_i)^2$, 随着迭代次数的增多, 实际输出与期望输出的逼近, 误差越来越小, 使得网络的收敛速度越来越慢, 有学者提出了采用绝对逼近精度与相对逼近精度相结合来检验网络的学习精度, 即 $E = \frac{\theta_1}{2} \sum_i (y_i^* - y_i)^2 + \frac{\theta_2}{2} \sum_i (1 - \frac{y_i^*}{y_i})^2$, 其中

θ_1, θ_2 为可变系数. 在网络训初期, 误差较大时, 以绝对误差来指导权值的调整; 在学习后期, 误差较小, 则以相对误差来指导调节.

2.3.3 引入动量项, 自适应调整网络参数

经典 BP 网络采用误差曲面梯度下降法, 该算法在峡谷区会引起振荡, 影响收敛速度. 可引入动量项平滑梯度下降方向的剧烈变化, 加快收敛速度. 此时, 网络的权值调整按 $w_{ij}(k+1) = w_{ij}(k) + \eta \delta y_i + \alpha [w_{ij}(k) - w_{ij}(k-1)]$ 计算. 网络的学习率 η 和动量系数 α 的大小对网络的学习速度和收敛度有较大的影响, η 和 α 较小时, 权值修正量小, 学习缓慢, 而 η 较大, 学习速度提高, 但可能引发振荡; Jacek 指出 η 的最佳范围是在 $[0, 10]$ 之间, 且允许值范围内的较大的取值, 可使学习速度上百倍的提高, 在稳定点附近产生振荡说明隐藏着最佳的 η . α 取值一般在 $[0, 1]$ 之间. 有文献表示分阶段动态调整学习率 η 和动量系数 α 可加快收敛速度.

2.3.4 全局最优

经典 BP 算法存在局部极小问题, 而非所期望的全局极小. 彻底解决这个问题, 除上述改进措施外, 我们可以将全局优化方法引入 BP 算法之中, 即在原来学习算法中加入一个全局优化方法形成的初始点的模块, 以选择好的初始值, 从而自动避免了局部极小问题. 为了求得全局极小问题, 可以采用一些优化理论的方法, 如卡尔曼滤波、同伦算法、遗传算法等. 它不仅可以避免局部极小还可以提高学习速度.

3 实例应用

边坡抗震抗滑稳定性受众多因素的影响^[5], 从各影响因素中选取烈度(I)、地面运动加峰值(α)、坡高(H)、坡角(θ)、岩石重度(γ)、内聚力(C)、内摩擦(φ)、岩体结构类型(T)、结构面(S)、水文地质条件(W)共 10 个具有代表性的指标作为 BP 网络的输入. 而对于定性变量, 参照专家经验知识, 首先按其实际意义转换成整数据区间上的分布数字, 如表 1, 在此基础上再进行归一化处理. 例中选用边坡抗震抗滑稳定性作为网络的输出, 判别结果采用二进制表示(即输出层为 2 个神经元), 见表 2 根据西部 70 个边坡实例作为神经网络学习样本, 采用改进 BP 算法进行学习和检验. 如表 3 列出部分样本输入与输出值, 可见网络的实际输入与期望入基本吻合, 表明改理的 BP 算法性能良好. 与改进前的 BP 网络算法相比, 在迭代次数、收敛速度、预测精度上均有明显的改善.

表 1 定性变量赋值表		
类型	变量	取值
结构类型	整体块状结构	1
	层状结构	3
	碎裂结构	5
	散体结构	7
结构面形态	无	2
	含反倾结构面	4
	含低缓结构面	6
	含高陡结构面	8
水文地质条件	干燥	1
	潮湿	2

4 结语

通过对 BP 型神经网络进行三大方面改进后, 改进 BP 网络具有很强的自组织、自适应、容错性和学习功能, 能很好地掌握岩土体的各种信息间复杂的非线性关系. 利用实测资料, 对高度复杂和高度非线性的土工结构进行直接建模^[5], 更能反映客观规律. 该方法计算简便, 精度高, 因而在工程中具有广泛的应用前景.

表 2 评判结果的二时制表示		
稳定性等级	二进制表示	代号
稳定	(0, 0)	I
较稳定	(0, 1)	II
不稳定	(1, 0)	III
极不稳定	(1, 1)	IV

表 3 部分学习样本的输入与输出

样 本 号	烈 度 <i>I</i>	地面加 速度峰 值 α	坡高 H/m	坡角 $\theta/^\circ$	岩石重 度 $\gamma/(\text{kN}\cdot\text{m}^{-3})$	内聚 力 C/kPa	内摩擦 角 $\varphi/^\circ$	结构 类型 T	结构 面 S	水文 地质 条件 W	期望 输出 代号	实际 输出 代号
1	6.2	0.20	80	45	22.30	10.0	40.0	5	4	2	III	III
2	4.9	0.10	23	30	22.0	0.0	41.0	3	2	1	I	I
3	7.6	0.25	320	42	25.5	24.1	37.6	3	6	2	IV	IV
4	5.3	0.10	22	33	27.3	31.5	29.4	1	2	1	II	II
5	4.4	0.10	50	40	23.8	46.3	35.2	5	4	2	II	II
6	5.7	0.10	298	41	27.2	10.1	39.0	3	6	2	II	II
7	4.6	0.10	34	50	25.1	16.8	38.9	3	2	1	I	I
8	7.6	0.30	87	55	26.0	46.3	35.6	5	6	2	IV	IV

参考文献:

[1] 靳蕃, 范俊波, 谭永东. 神经网络与神经计算机原理及应用[M]. 成都: 西南交通大学出版社, 1991. 96-120.
[2] 杨晓红, 刘乐善. 用遗传算法优化神经网络结构[J]. 计算机应用与软件, 1997(1): 59-65.
[3] 姜绍飞. 人工神经网络用于建筑工程领域的数据处理方法[J]. 哈尔滨建筑大学学报, 1999, 32(5): 24-32.
[4] 梁艳春. 人工神经网络 BP 处法密集型数据的预处理[J]. 吉林大学学报, 1995(3): 19-22.
[5] 刘勇健. 基于人工神经网络的边坡搞滑稳定性分析[J]. 西北地震, 2001(1): 41-45.

The Improving Methods of Arificial Neural Network in
Geo-technical Engineering

LIU Yong-jian¹, LIU Yi-jian²

(1. College of Construction, GDUT, Guangzhou 510643, China;
2. 407 Brigade of Geo-technical Invesitigation Bureau of Hunan Province, Huaihua 418000, China)

Abstract: Artificial neural network, a nonlinear dynamic system, is one of the most effective methods to solve problems of geo- technical engineering which has high complication and nonlinear. However, since traditional BP model has the shortcomings, such as slow convergence rate, poor stability and partial minimum, an improvement in the fields of neural network topological structure, data pre- processing and study calculation method is systematically discussed in this paper. Real example research reveals that BP network has good quality after being improved, and thes method has wide practicing in the engineering.

Key words: artificial neural network; BP network; topological structure; geo- technical engineering