BP 网络自动搜索法结构寻优及其在基坑监测中的应用

陈艳国,袁宝远,李爱国,朱旭芬,赵燕容,王　江

(河海大学土木工程学院,江苏 南京　210098)

摘要:针对 BP 网络的缺陷,采用模拟退火法与自动搜索过程相结合的方法优化 BP 网络,提出了一种较为实用的自动搜索法来确定 BP 网络隐层节点数,利用模拟退火法确定其初始权值与偏置值,并通过 MATLAB 软件平台实现.经润扬大桥深基坑工程南锚锭基坑支撑轴力与排桩位移多维非线性建模验证,此方法预测结果具有极高的精度.

关键词:基坑监测;BP 神经网络;初始权值;隐层节点数;模拟退火;自动搜索

中图分类号:TP274　　**文献标识码**:A　　**文章编号**:1006-7647(2005)S1-0030-04

深基坑工程具有造价高、施工难度高、不稳定因素多等特点,基坑监测信息是基坑稳定性状况的直观反映,因此监测系统的优劣性对于掌握深基坑稳定性状况具有重要影响.但深基坑是由支护体系与周边土体相互结合的多种介质组合的高度复杂的空间系统,其变形和安全性受地质条件、岩土体性质、场地环境、气候变化、地下水动态等因素影响.常规统计方法无法利用监测信息建模预测,BP 神经网络具有并行处理、联想记忆、分布式知识存储与鲁棒性强及自组织、自适应学习功能,使其在复杂非线性系统的分析和预测中得到了广泛应用.但 BP 网络也具有明显缺陷,易收敛于局部极小值,过拟合,网络结构选择具有主观性,初始权值和偏置值难于确定等问题制约了它的应用.本文采用模拟退火法及所提出的自动搜索法来优化 BP 网络,针对基坑监测信息建立非线性"隐式"模型,来预测基坑稳定性发展趋势.

1　BP 网络

BP 网络本质是基于误差反向传播算法的前馈网络,BP 网络的层数一般指中间层和输出层,中间层即为隐含层(常用 1 层或 2 层),其模型见图 1.对于输入数据,要先向前传播到隐节点,经过激活函数后,再把隐节点的输出信息传播到输出节点,经误差函数分析,然后误差逐层反向传播"分摊"给各隐层神经元并逐层调整权值与偏置值,这样前后循环,直至误差满足要求,最后给出输出结果.隐层节点的激活函数通常选取标准 Sigmoid 型函数,BP 算法是一个很有效的算法,许多问题都可由它来解决.

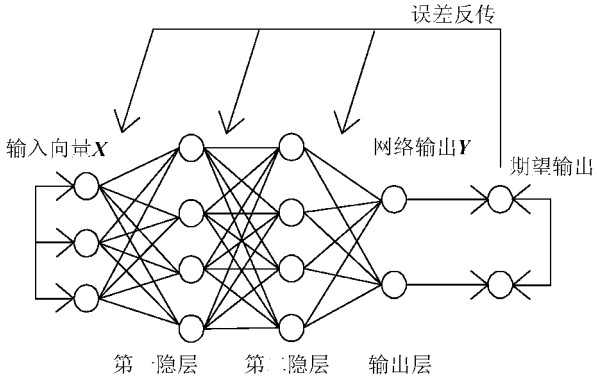

图 1　BP 网络结构模型

现代科学计算软件 MATLAB 具有强大的矩阵运算功能及各种数学算法函数,对于编程具有很大的优越性,本文算法全部基于 MATLAB 软件 M 语言编程实现,文中函数命令,如接点转移函数 logsig,tansig 等,均为 MATLAB 工具箱内部函数,这里不再赘述.

2　BP 网络性能因素

影响 BP 神经网络效果的主体因素有:原始数据的处理,初始权值 W 与初始偏置值 B,BP 网络结构(网络层数、隐层节点数与层间连接函数).这些因素对网络的优越与否影响重大,在深基坑工程中,运用 BP 网络的目的主要是建模预测和反分析,若网络结构不当,面对大量的监测数据重复建模将是一项十分繁琐的工作,本文就 BP 网络的关键影响因素简要予以总结,并提出模拟退火法结合新的自动

作者简介:陈艳国(1980—),男,黑龙江同江人,硕士研究生,从事岩土工程与地质工程研究.

搜索法来优化网络结构.

2.1 原始数据预处理

实测资料往往伴随着许多不确定的影响因素,在进行 BP 网络建模时,对数据进行预处理是必要的,目前常用的方法有归一化、主成分分析法、特征提取法等,还有研究者提出用 $y' = y/(y_{\max} + y_{\min})$ 来处理.

2.2 初始权值 W 与偏置值 B

BP 网络的训练效果强烈依赖于初始权值 W 与偏置值 B,目前对于 W 与 B 学术界并没有明确的理论来说明应如何取值,各研究者针对各自的研究领域提出了不同的取值方法,如利用几何数学提出的复合法取初始权值[1];初始权值 $W = K^{-1/2}$(其中 K 为前一层的节点数目)[2];权退化法;遗传算法;感受野的 Gabor 函数来优化网络初始权值[3].

总结前人的经验成果,目的都是希望得到 ANN 均方误差性能曲面的最小值点,达到全局最优目的,本文提出用基于金属降温理论的模拟退火法(Simulated Annealing)[4,5]来优化 BP 网络的初始权值与偏置值.其实现步骤如下:

a. 选定网络结构,包括网络层数(以 3 层,双隐层为例)、各层节点数(设输入层为 R,输出层为 S_3,隐层 1 为 S_1,隐层 2 为 S_2)及层间连接函数(以 logsig,logsig,purelin 为例).

b. 依据网络结构建立网络传播方程和误差方程,如两隐层 BP 算法,网络传播方程为 $f = \text{purelin}(W_3\text{logsig}(W_2\text{logsig}(W_1P_n + B_1) + B_2) + B_3)$,误差函数为 $E = f - T_n$,现在要做的就是对误差函数求极小值.

c. 对误差方程进行编码,要优化的自变量为初始权值与偏置值,其个数为 $L = S_1R + S_2S_1 + S_3S_2 + S_1 + S_2 + S_3$,在模拟退火法中 L 是误差函数 $E = f - T_n$ 的自变量个数,利用网络传播方程 $f(x)$ 确定哪些 x 的值对应哪些权值与偏置值.

d. 选定足够高的初始温度 T_0,随机产生初始解 x_0,计算当前初始函数能量值 E_{old},选定温度衰减函数,一般用 $T = \lambda T$(依经验 λ 为 0.8~0.98).

e. 对当前最优点作一随机扰动(扰动值 Δx),产生一新的最优点 $x = x_0 + \Delta x$,计算新的目标函数值 E_{new}.

f. 若 $E_{\text{new}} < E_{\text{old}}$,则接受新产生的最优点 x 作为当前最优点;若 $E_{\text{new}} \geqslant E_{\text{old}}$,则以概率 $p = \exp\left(-\dfrac{E_{\text{new}} - E_{\text{old}}}{T}\right)$ 接受该新产生的最优点 x 为当前最优点.

g. 在当前温度 T 下,重复步骤 e 抽样足够多次数,找出当前温度 T 下的能量函数最小值.

h. 温度冷却 $T = \lambda T$,重复步骤 e.

i. 若温度冷却次数达到要求,则输出当前最优点,优化结束.

j. 转至步骤 c,解译出 W 与 B.

2.3 BP 网络结构

BP 网络结构包括网络层数、隐层节点数及层间连接函数.网络层数与层间连接函数目前争议不大,已有证明,3 层网络能精确逼近任意函数.输入层到隐层及隐层之间一般采用 Sigmoid 型函数,也有文章提出采用双极性 S 型函数 $y = f(a) = -\dfrac{1}{2} + \dfrac{1}{1 + e^{-a}}$,还有的提出在 S 型函数中加入可调系数 $y = f(a) = \dfrac{1}{\alpha_0 + \beta_0 e^{-\gamma_0 a}}$[6],隐层到输出层则用线性传输函数 purelin.

目前争议较多的是 BP 网络隐层节点数的确定,过少的节点数不足以建立好的映射网络,过多又会出现过拟合,影响 BP 网络的泛化能力.目前隐层节点数的确定方法有多种,如利用遗传算法来确定[7],利用模拟退火法来确定[8],还有人提出隐层节点数的经验取值范围,即对于含 1 个隐层的网络,单输入单输出网络,取隐层节点数 4~8 个;多输入多输出的网络,取隐层节点数为输入节点的 2~4 倍,还有研究者提出用搭积木的形式来一步步增减隐层节点数对所研究问题进行分析,取优舍劣.

经大量试验得出,当隐层节点数达到收敛要求后,再增加 1~2 个节点并不会十分影响网络的收敛性及泛化能力,基于这种经验和高频率原则(即某种结果出现的频率高),提出确定隐层节点数目的一种新方法——自动搜索法,即自动寻找网络最优结构.

3 BP 网络非线性建模过程

a. 原始输入输出数据资料预处理,$P \to P_n$,$T \to T_n$.

b. W,B 及 ANN 的网络结构采用由模拟退火法及自动搜索法联合起来的混合搜索法来确定,首先建立 K 次循环,然后确定网络结构.输入输出层节点数首先由输入输出数据直接定出,然后确定两隐层节点数目的上限,隐层 1、隐层 2 节点数目的上限分别为 M 和 N,再确定层间连接函数.隐层节点数的确定采用提出的新方法'自动搜索法'来确定,其步骤如下:①首先建立 M 次子循环,再嵌套建立 N 次子循环,这相当于建立了 $M \times N$ 个网络结构;②在每一结构下运用模拟退火法寻找最优的 W 和

B，让 BP 网络以此为起点进行训练，得出每一结构训练误差 $E_{(i,j)}$；③找出 $M \times N$ 个结构的 $E_{(i,j)}$ 的最小值 $E_{\min}$；④找出 $E_{\min}$ 所对应的 S_1 与 S_2 的值，这样通过最初的 K 次循环就能找出 K 组 S_1 与 S_2 的值，把其命名为最优结构组；⑤然后把其中 S_1 与 S_2 都相同的分组，依据‘高频率’原则选取组中 S_1 与 S_2 数目最多的作为最优的网络结构.

c. 经步骤 b 找到最优结构后，便建立了针对所研究问题的最优 BP 网络模型.

4 工程实例

润扬大桥是大跨径悬索桥，且是国内首例采用排桩冻结法作为止水帷幕的深大基坑工程，掌握其基坑稳定性发展趋势是保障基坑安全的必要任务. 南锚锭基坑排桩位移是基坑稳定性状况的直观反映，而实测位移随着基坑开挖施工扰动的影响，呈现出高度非线性变化的现象. 本文基于上面所述模拟退火自动搜索方法改进的 BP 网络对润扬大桥南锚锭第一道支撑轴力与基坑 I3 号测斜孔排桩位移进行非线性建模. 南锚第一道支撑轴力有 4 个测点：Z1-1，Z1-2，Z1-3，Z1-4，即 4 组轴力数据，基坑 I3 孔排桩位移为 1 组数据. 采用 2001 年 12 月 25 日到 2002 年 2 月 18 日的数据进行分析，此时冻涨力正处于明显变化之中，时刻注意排桩位移显得尤为重要，排桩位移的预测对于指导施工进程，掌握冻结状况乃至基坑稳定性有着重要意义.

基于上述监测信息建模过程如下：

a. 初始数据归一化处理，选用 MATLAB 工具箱函数 premnmx.

b. 由于所建模型为四维空间，所以网络结构选用较为复杂的双隐层结构，层间连接函数分别为 logsig，logsig 与 purelin.

c. 进行自动搜索过程，其间包含有利用模拟退火法来取得 BP 网络权值与偏置值的初始值. 建立 7 次主循环，隐层节点数 S_1，S_2 的上限均为 8，这样就建立了两层 8—8 子循环，即自动循环（$8 \times 8 \times 7$）次即可自动寻找出较优的网络结构，省去了诸多麻烦，见表 1. 综合选择结果为 $S_1 = 4$，$S_2 = 7$，即隐层 1 与隐层 2 的优化结构数为 4—7. 由表 1 知 BP 网络最优结构应为 4—7—1 结构，当然也可以上下浮动 1 ~ 2 个节点数目，但为了防止过拟合，当满足收敛要求后，节点数目以少为妙.

表 1 BP 网络结构优化结果

主循环次数 K	隐层 1 节点数 S_1	隐层 2 节点数 S_2	主循环次数 K	隐层 1 节点数 S_1	隐层 2 节点数 S_2
1	4	7	5	8	4
2	5	3	6	5	3
3	4	7	7	4	8
4	4	6			

d. 由所得网络优化结构，对数据进行建模分析，结果见图 2 ~ 5. 图 2 显示此优化后的网络经 107 步便可达到收敛误差 1×10^{-20}，说明优化后的网络结构利用模拟退火法求解初始权值与偏置值后收敛速度快. 图 3 显示基于模拟退火法与自动搜索法南锚锭第一道支撑 I3 孔排桩位移与 4 个轴力测点的 BP 网络建模仿真误差数量级为 10^{-11}，模型精确. 图 4 选取训练数据后的 22 个实测值作为预测数据，将预测值与实测值作比较，可以看出，误差很小. 图 5 显示实测值与预测值的相近性为 0.98，表明所建模型对此问题预测较为精确.

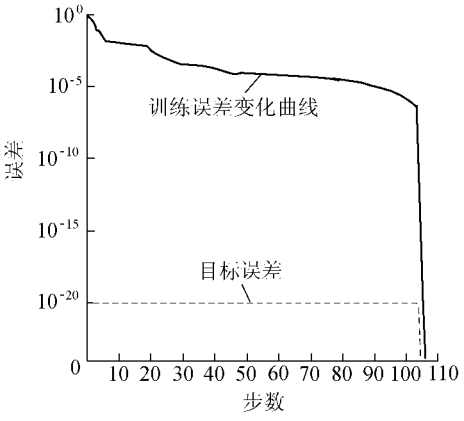

图 2 网络训练过程

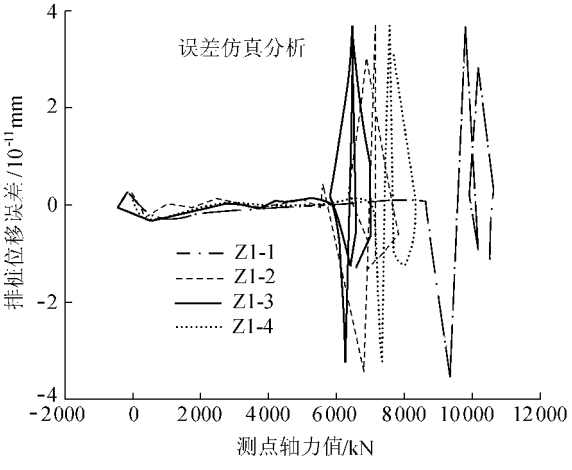

图 3 网络仿真误差分析

结果表明，基于模拟退火法与自动搜索法优化的 BP 网络用于深基坑的监测信息非线性建模，具有较高的精度.

5 结 论

本文针对基坑工程的高度复杂性，利用 BP 网络对监测信息进行非线性建模来掌握基坑稳定性的发展趋势. 对于当前 BP 网络的问题进行了总结，利用模拟退火法确定网络初始权值，并对其隐层节点

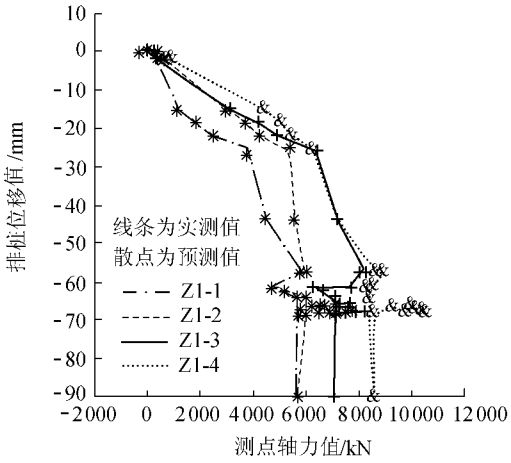

图 4　预测值与实测值比较

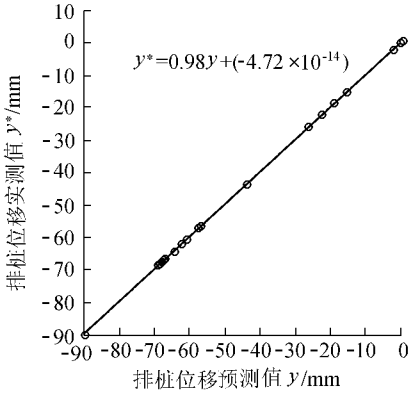

图 5　预测值与实测值相关分析

数的确定提出了一种简单有效的自动搜索方法. 基于 MATLAB 软件平台进行程序设计,应用于润扬大桥南锚锭基坑轴力-位移预测检验,结果表明以此方法改进的 BP 网络具有较好的非线性建模能力.

参考文献:

[1] 江艳君,李柠,黄道. 修正初始权值的 BP 网络在 CSTR 故障诊断中的应用[J]. 华东理工大学学报,2004,30(2):207-210.

[2] 李宇峰,裴旭东,黄聪明. BP 神经网络实际应用中的若干问题[J]. 兵工自动化,1998(1):1-4.

[3] 景晓军,余农. 基于形态学变权神经网络的数据精炼[J]. 电子学报,2005,33(3):397-401.

[4] 邢文训,谢金星. 现代优化计算方法[M]. 北京:清华大学出版社,1999.

[5] 康立山,谢云,罗祖华. 非数值并行算法—模拟退火法[M]. 北京:科学出版社,1994.

[6] 陈详光,裴旭东. 人工神经网络技术与应用[M]. 北京:中国电力出版社,2003.

[7] 郭晓婷,朱岩. 基于遗传算法的进化神经网络[J]. 清华大学学报:自然科学版,2000,40(10):116-119.

[8] 邹恩,李祥飞,刘耦耕,等. 最优模糊神经网络参数的设计-混沌模拟退火学习法[J]. 中南大学学报,2004,35(3):443-447.

(收稿日期:2005-07-18　编辑:骆　超)

(上接第 29 页)

比较两种方法得出的不均匀系数和曲率系数,由式(11)求 C_u 的误差率 δ_{C_u} 和 C_c 的误差率 δ_{C_c}(见表 2):

$$\delta_{C_u}=\frac{C'_u-C_u}{C_u}\qquad \delta_{C_c}=\frac{C'_c-C_c}{C_c}\qquad(11)$$

进一步分析可知 C_u 的平均误差率 $\delta_{C_u}=11.35\%$,C_c 的平均误差率 $\delta_{C_c}=8.77\%$,两者均小于 15%,表明分维 D 可以表示软弱层带的颗粒级配. 而且,D 值越大,表明颗粒大小越不均匀;反之,D 值越小,则颗粒大小越均匀.

当粗粒土的级配同时满足 $C_u\geqslant5$,$C_c=1\sim3$ 时称为良好级配;当 $C_u\geqslant5$ 时,由式(9)得 $D\geqslant1.887$;当 $C_c=1\sim3$ 时,由式(10)得 $D\leqslant2.631$. 所以,当粗粒土的坡度分布的分维数满足 $D=1.887\sim2.631$ 时,其级配称为良好级配,故只需计算出分维值,就可判定级配良好与否.

由此判据,试件 1 和 2 的不均匀系数 C_u 小于 5,分维 D 小于 1.887,属于非良好级配;试件 4 ~ 20 的不均匀系数 C_u 大于 5,曲率系数 C_c 介于 1 ~ 3 之间,分维 D 介于 1.887 ~ 2.631 之间,属于良好级配;只有试件 3 不符合上述判据.

3　结　论

a. 坝区软弱层带粒度成分的分形结构是客观存在的,可以把分维作为描述其粒度成分特征的定量参数. 分维值在 1.881 ~ 2.355 之间,平均值为 2.121.

b. 坝区软弱层带的级配特征可以用分维单独描述,当 $D=1.887\sim2.631$ 时,其级配称为良好级配.

参考文献:

[1] FALCONER K J. 分形几何:数学基础与应用[M]. 谢和平,张永平,宋晓秋,等译. 重庆:重庆大学出版社,1991.

[2] 张水平,徐汉诧. 编译. 分形集几何学[M]. 徐州:中国矿业大学出版社. 1991.

[3] WEISBERG S. 应用线形回归(Applied Linear Regression)[M]. 王静龙,梁小筠,李宝慧,译. 北京:中国统计出版社. 1998.

[4] 陈伟庆,孔书祥. 土力学[M]. 北京:中国铁道出版社,2002.

[5] 田堪良,张会礼. 论天然沉积砂卵石粒度分布的分形结构[J]. 西北水资源与水工程,1996,7(4):26-31.

(收稿日期:2005-09-14　编辑:高建群)