

改进的 BP 神经网络模型预测充填体强度

魏 微^{1,2}, 高 谦¹

(1. 北京科技大学 土木与环境工程学院, 100083 北京; 2. 承德石油高等专科学校, 067000 河北 承德)

摘 要: 为了准确评价和预测充填体强度,采用改进的 BP 神经网络算法,利用实验室做的 18 组充填体单轴抗压强度试验结果,建立了充填体强度与影响因素之间的 5-7-1 网络模型结构(输入层为 5 个神经元,隐含层为 7 个神经元,输出层为 1 个神经元,输入为胶砂比及各胶凝材料掺量,输出为充填体 28 d 单轴抗压强度)。结果表明,改进的 BP 神经网络对于充填体的强度具有良好的预测能力,建立的网络模型不仅收敛速度快而且训练精度高,对充填体强度的预测结果与训练数据和测试数据的最大相对误差仅为 4.23%。

关键词: 充填体;强度;预测;改进的 BP 神经网络

中图分类号: TD853.34

文献标志码: A

文章编号: 0367-6234(2013)06-0090-06

Strength prediction of backfilling body based on modified BP neural network

WEI Wei^{1,2}, GAO Qian¹

(1. School of Civil and Environment Engineering, University of Science and Technology Beijing, 100083 Beijing, China;

2. Chengde Petroleum College, 067000 Chengde, Hebei, China)

Abstract: To evaluate and predict the strength of backfilling body, a new method is provided to establish a model of the relationship between backfilling body strength and influence factors. The modified BP neural network algorithm is used to establish the model based on 18 groups results of the compressive strength tests of the backfilling in laboratory. The structure of the model is 5-7-1 type, that is to say 5, 7 and 1 neurons are the input, hidden and output layers respectively, where the input is including the cement-sand ratio and quantity of the cemented material and the output is the 28 days compressive strength of the backfilling body. The results show that BP neural network model exhibits excellent prediction ability for the prediction of the strength of backfilling body, the prediction model speeds up the convergence rate of BP network and improves the training accuracy. The maximum relative error between the predicted results and the test data is 4.23%.

Key words: backfilling body; strength; prediction; modified BP neural network

随着我国采矿事业的发展,无废开采已成为采矿技术发展的必然趋势^[1-2]。充填采矿法是将矿山尾矿废物充填于井下采空区,不仅可节省尾矿废物的堆存场地、减少环境污染,而且可提高回采率、降低贫化率、有效控制采区地压、维护采空区稳定、保障安全作业,因此充填采矿技术越来越

受到人们的重视^[3-5]。影响充填体强度的因素有很多^[6],既有确定的、可量化的因素,也有一些模糊的或不确定性的因素,这些影响因素之间具有复杂的非线性关系。迄今为止,国内外学者提出了许多预测充填体强度的方法,如相似试验、经验公式估算、数值模拟、弹性力学分析等^[7]。但是由于充填体是一种多相复合力学介质,用以上方法往往得不到精确的预测结果。人工神经网络不仅具有较强的非线性动态处理能力,可实现输入样本和输出样本的非线性映射,而且具有较强的自适应性,在整个学习过程中无需外界参与,能从已有数据中自动归纳出规则,取得知识,可以利用经过训练的神经网络直接进行推理^[8-9]。对于充填体

收稿日期: 2012-09-04.

基金项目: 国家重点基础研究发展计划项目(2010CB731501);长江学者和创新团队发展计划项目(IRT0950);河北省钢铁产业技术升级专项资金项目(SJGS-KJ-12-03).

作者简介: 魏 微(1976—),女,博士;

高 谦(1956—),男,教授,博士生导师.

通信作者: 高 谦, gaoqian@ces.ustb.edu.cn.

的强度预测,可以从已有的试验数据中获取影响因子的规律,避免传统强度预测方法的不足,提高了预测的准确性和快速性.

BP 网络是应用最多的多层网络,是人工神经网络最精华的部分. 与 BP 网络相比,改进的 BP 神经网络具有更强的非线性映射能力,训练速率更快,不易陷入局部极小. 本文采用改进的 BP 神经网络模型预测充填体强度.

1 改进学习算法的 BP 网络

BP 网络是按误差逆传播算法 (Error back-propagation training) 学习的一种多层网络,一般由输入层、隐含层和输出层组成. 输入层神经元个数由数据样本维数决定;隐含层神经元是整个网络的计算核心,计算输入样本;输出层输出其计算结果.

改进 BP 网络算法的基本原理与标准 BP 网络算法的基本原理相同:学习过程由信号的正向传播与误差的反向传播两个过程组成. 正向传播时,信息从输入层经隐含层逐层处理后传向输出层. 如果在输出层得不到期望的输出,则转向误差的反向传播,误差反传是将误差信号通过隐含层逐层向输入层反传,并分摊给各层的所有单元,从而获得各层单元的参考误差 (或称误差信号),以作为修改各层神经元连接权值的依据. 信号的正向传播与误差的反向传播是周而复始进行的,权值不断修改的过程也就是网络学习 (或训练) 的过程,此过程一直进行到网络输出误差达到允许的范围或达到设定的训练次数为止.

标准 BP 网络算法采用梯度下降法,在实际应用中存在网络收敛速度慢、容易震荡和陷入局部极小等问题^[10-12]. 本文通过引入附加冲量项和自适应学习率来改进标准 BP 学习算法^[13-14],从而提高网络收敛速度,避免震荡和陷入局部极小等问题.

1.1 引入附加冲量项

所谓的附加冲量项是指在网络每次的权值和阈值改变量中加入正比例与前一次的改变量. 附加冲量项的引入加快了网络的收敛,有效地避免了局部最小问题的出现,同时保证了训练过程的稳定性. 具体算法为

$$\Delta w(t+1) = -(1-\beta)\eta g_w(t) + \beta\Delta w(t), \tag{1}$$

$$\Delta\theta(t+1) = -(1-\beta)\eta g_b(t) + \beta\Delta\theta(t). \tag{2}$$

式中: $\Delta w(t)$ 、 $\Delta w(t+1)$ 分别为当前和修正后的

权值差; $\Delta\theta(t)$ 、 $\Delta\theta(t+1)$ 分别为当前和修正后的阈值差; η 为学习率; $g_w(t)$ 、 $g_b(t)$ 分别为当前性能函数对权值和阈值的梯度; β 为动量因子, $0 < \beta < 1$, 等式右边第二项为附加冲量项.

1.2 引入自适应学习率

在引入附加冲量项的 BP 算法中,学习率 η 是一个固定的常数,它的取值将直接影响到网络的训练性能. 在误差平坦区域, η 太小,会增加训练次数、延长训练时间;在误差变化剧烈的区域, η 太大会跨过较窄的“坑凹”处,使训练出现震荡,降低网络的稳定性. 针对固定 η 存在的弊端,引入自适应学习率 η , 使其该大时大,该小时小,本文采用的修正公式为

$$\eta(t+1) = \eta(t) \times 2^{-\text{sign}[e(t)-e(t-1)]}. \tag{3}$$

式中: $\eta(t)$ 、 $\eta(t+1)$ 分别为当前和修正后的学习率值; $e(t)$ 、 $e(t-1)$ 分别为当前和上时刻的学习误差.

2 充填体强度试验

对于神经网络建模来说,训练数据的获取是重要前提,训练数据的完备性、有效性决定了神经网络模型的精度和泛化能力^[15]. 为了给预测模型提供真实、有效的训练数据,本文开展了充填体的单轴抗压强度实验室测试工作.

本试验所用原料主要有唐龙新型建材有限公司的水淬高炉矿渣,唐山周边地区普通石灰窑烧制的石灰,唐山丰润新区热电厂的脱硫石膏,石人沟铁矿的全尾砂和钻牌 42.5R 水泥.

2.1 试件制备

胶砂制备. 使用 JJ-5 型行星式水泥胶砂搅拌机,将胶结材料 (以石灰、脱硫石膏和水泥为复合激发剂,主要材料为工业废渣水淬高炉矿渣)、水、尾砂按设计的质量比搅拌 180 s 制成胶砂.

试块成型. 将搅拌好的胶砂倒入 7.07 cm × 7.07 cm × 7.07 cm 三联试模,为了保证试块浇注过程中料浆不发生沉淀,采用边搅拌、边注模的浇注方式.

标准养护. 将三联试模放入温度 (20 ± 1) °C、湿度不低于 90% 的 YH-40B 型标准恒温恒湿养护箱中进行养护.

试块脱模. 三联试模在养护箱养护 48 h, 然后进行脱模,脱模后试块继续放入 YH-40B 型养护箱养护,直到 28 d 龄期.

2.2 单轴压缩试验及结果

利用 SANS 数显固定位移压机以 0.5 mm/s 连续匀速加荷进行单轴抗压强度的测定. 每龄期

强度测试 3 块,取其平均值,作为该龄期充填体的单轴抗压强度值. 充填体 28 d 单轴抗压强度试验结果如表 1 所示.

表 1 试验结果和预测结果

样本	编号	胶砂比	胶结材料的质量分数/%				试验值/ MPa	预测值/MPa			相对误差/%		
			水泥	石灰	石膏	矿渣		改进 BP 网络	标准 BP 网络	二次多项式回归	改进 BP 网络	标准 BP 网络	二次多项式回归
网络训练样本	1	1:5	2.5	16.0	1.0	80.5	5.42	5.46	5.37	5.42	0.74	0.98	0.00
	2	1:5	2.5	10.0	7.5	80.0	4.66	4.68	4.37	4.66	0.43	6.17	0.00
	3	1:8	0.5	11.0	1.5	87.0	2.80	2.79	3.08	2.80	0.36	9.88	0.00
	4	1:8	0	12.0	9.5	78.5	2.64	2.63	2.84	2.64	0.38	7.41	0.00
	5	1:8	4.0	13.0	2.0	81.0	2.47	2.47	2.52	2.47	0.00	1.92	0.00
	6	1:8	6.0	10.5	2.0	81.5	2.46	2.47	1.80	2.46	0.41	26.83	0.00
	7	1:8	4.0	9.0	8.0	79.0	2.44	2.43	2.64	2.44	0.41	8.13	0.00
	8	1:8	3.0	10.0	3.5	83.5	2.44	2.47	1.98	2.44	1.23	19.02	0.00
	9	1:7	2.0	12.0	2.0	84.0	4.19	4.20	4.05	4.19	0.24	3.40	0.00
	10	1:7	2.5	15.0	4.0	78.5	3.30	3.30	3.04	3.30	0.00	7.81	0.00
	11	1:10	0	15.0	4.0	81.0	2.72	2.71	2.60	2.72	0.37	4.58	0.00
	12	1:10	2.0	14.0	3.0	81.0	2.54	2.51	2.96	2.54	1.18	16.72	0.00
	13	1:12	2.5	16.0	1.0	80.5	1.80	1.85	1.83	1.80	2.78	1.87	0.00
	14	1:12	3.0	10.0	3.5	83.5	1.81	1.84	1.95	1.81	1.66	7.70	0.00
网络测试样本	15	1:5	2.0	12.0	2.0	84.0	5.74	5.62	4.04	6.86	2.09	29.68	19.51
	16	1:7	5.0	10.0	4.0	81.0	3.31	3.45	2.17	1.91	4.23	34.57	42.30
	17	1:8	2.5	10.0	7.5	80.0	2.78	2.71	4.15	2.27	2.52	49.14	18.35
	18	1:10	3.5	17.0	4.5	75.0	2.50	2.55	1.80	-0.39	2.00	28.00	115.60

注: 1) 胶砂比为胶结材料与尾砂的质量比; 2) 相对误差指预测值对试验值的相对误差.

3 改进的 BP 神经网络预测模型

影响充填体强度的主要因素有料浆含量、胶砂比、胶结材料配方及用量、骨料、温度、粒级等. 本研究重点研究胶砂比、胶凝材料配方对充填体强度的影响, 试验时固定料浆质量分数为 68%, 充填骨料为石人沟全尾砂. 为了建立充填体强度预测模型, 实验室实测了 18 组胶砂比不同、胶凝材料配方不同条件下的充填体 28 d 单轴抗压强度值.

3.1 网络模型的选取

根据 Kolmogorov 定理^[16], 3 层前向网络可以任意精度逼近任意连续函数, 所以确定网络总层数为 3 层. 第 1 层为输入层, 因为本研究主要分析 5 个因素对充填体强度的影响, 所以输入层为 5 个神经元, 分别代表胶砂比和水泥、石灰、石膏、矿渣的质量分数. 第 2 层为隐含层, 隐含层神经元个数的选取对网络非常重要, 若神经元数设置太多, 则耗机时间过长, 使网络失去概括判断能力, 且对提高精度无突出贡献, 若神经元数设置太少, 则网

络求解问题精度提高慢, 易产生误差率波动. 目前, 隐含层神经元个数的选取都是经验性的规则, 并运用试错法进行确定. 目前较常用的规则如下: Homik 提出的隐含层神经元数量为 $\sqrt{(2m + 1)}$ 至 $(2m + n)$, Hecht Nielsen 提出隐含层神经元数量为 $(2m + 1)$, David Shepard Associate 提出的经验公式为 $\sqrt{mn}$. 在以上这些经验公式中, m 为输入变量数量, n 为输出变量数量. 分别对它们进行计算, 得到网络模型的隐含层神经元数在 3 ~ 11 之间. 用表 1 中前 14 组试验数据作为训练样本, 后 4 组试验数据作为测试样本建立隐含层神经元数可变的网络模型. 第 3 层为输出层, 本研究是预测充填体 28 d 的单轴抗压强度, 所以输出层选取 1 个神经元, 代表充填体 28 d 的单轴抗压强度. 充填体强度的影响因素多且比较复杂, 不是简单的线性关系, 所以隐含层神经元传递函数采取非线性变换的双极性 Tansig 函数, 输出层神经元传递函数为 Logsig 函数. 具体表达式为

$$f_1(x) = \frac{1 - e^{-x}}{1 + e^{-x}},$$

$$f_2(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

为了观察网络性能,选择合适的隐含层单元数,表 2 给出隐含层不同单元数对应网络对测试样本的预测误差,图 1 列举出一些隐含层不同神经元数对应网络的训练误差曲线。

表 2 网络对测试样本的预测误差

样本 编号	隐含层不同单元数的网络对测试样本的 预测相对误差/%									
	3	4	5	6	7	8	9	10	11	
15	3.4	18.1	8.1	3.9	2.1	2.5	3.0	2.7	3.6	
16	40.0	25.0	14.0	7.1	4.2	4.9	8.3	16.7	20.4	
17	11.2	21.6	16.2	4.8	2.5	4.4	5.2	10.4	16.9	
18	27.4	17.8	7.5	4.3	2.0	3.6	7.1	11.4	13.5	

从图 1 可以看出隐含层神经元数为 7 时,网络的收敛速度已经很快,再增加隐含层神经元数,网络的收敛速度提高不大;另外,表 2 给出的预测误差表明:隐含层神经元数超过 7 时,对测试样本的预测误差反而增大,说明隐含层神经元数过多,会降低了网络的泛化能力.综合考虑,隐含层神经元数选取 7 时,网络的性能最好,所以预测模型的隐含层神经元数选为 7.

本文建立的预测充填体的神经网络结构如图 2 所示.

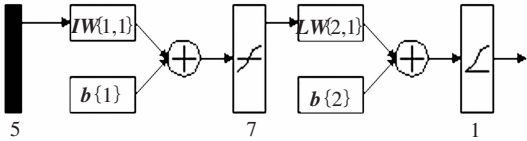


图 2 预测充填体强度的神经网络结构图

3.2 数据的归一化

网络的各个输入数据通常具有不同的物理意义和不同的量纲;而且网络采用的传递函数会因输入的绝对值过大使权值调整进入误差曲面的平坦区.为了使网络训练一开始就给各输入分量以同等重要的地位,加快网络学习和计算收敛效率,在对网络进行训练之前,所有样本数据,包括输入和输出数据归一化至[0,1],使不同物理量的数据具有平等地位,消除奇异样本数据.同理,预测结果需进行反归一化处理,处理后的数据才是真实结果.本文对数据进行归一化处理和反归一化处理时采用的公式分别为

$$x_n = (x - x_{\min}) / (x_{\max} - x_{\min}) \tag{5}$$

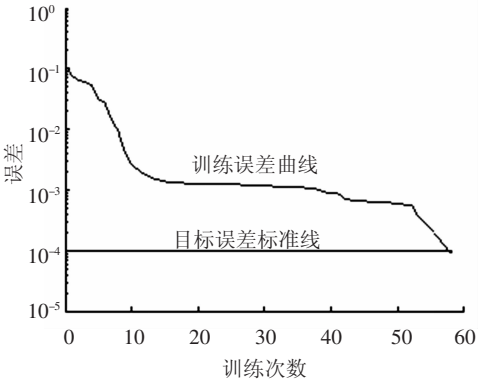
$$y = y_n (y_{\max} - y_{\min}) + y_{\min} \tag{6}$$

式中: x 为输入的试验数据, $x_{\min}$ 、 $x_{\max}$ 分别为输入数据中的最小值和最大值, x_n 为归一化后的数据; y 为反归一化处理值; y_n 为网络预测结果输出值; $y_{\min}$ 、 $y_{\max}$ 分别为待预测数据中的最小值和最大值.

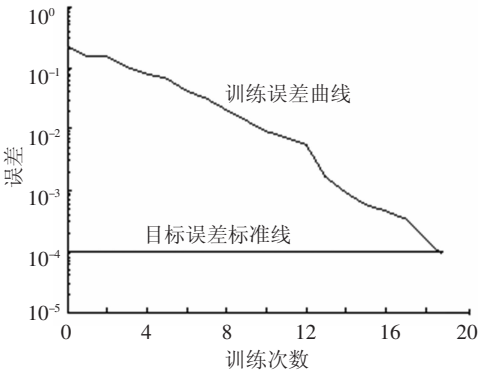
3.3 网络训练与预测

选取表 1 的前 14 组试验数据作为训练样本,其余 4 组作为测试样本.网络训练目标误差选取 0.000 1,网络允许训练次数为 1 000.为了比较改进 BP 网络和 BP 网络的训练效果,分别采用相同网络结构 5×7×1 的改进 BP 网络模型和标准 BP 网络模型对试验数据进行训练.利用 Matlab 语言编程,在 Matlab6.5 软件平台运行,网络的训练效果分别如图 1(b)和图 3 所示.

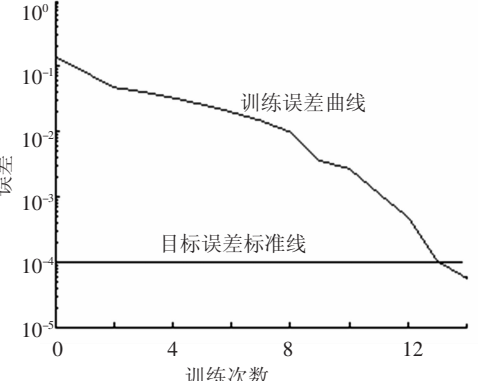
由图 1(b)可以看出,改进的 BP 网络的训练速率非常快,随着训练次数的增加,网络的误差迅速下降,训练 19 次时误差就已达到了目标误差值,图中的直线为目标误差标准线 0.000 1.网络



(a) 隐含层 5 个神经元



(b) 隐含层 7 个神经元



(c) 隐含层 9 个神经元

图 1 改进的 BP 神经网络模型训练误差曲线

模型隐含于输入层和隐含层间、隐含层和输出层间的权值、隐含层和输出层各神经元的阈值中。

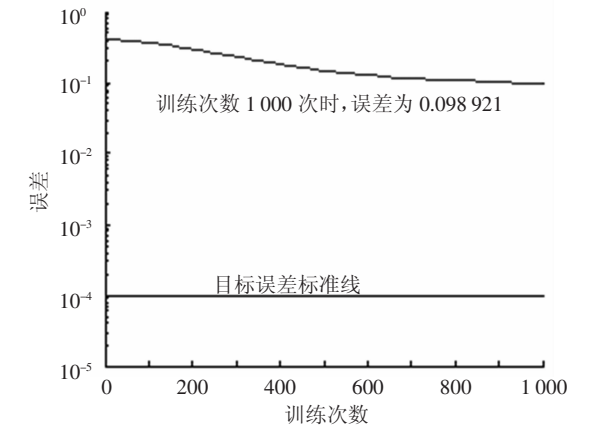


图3 标准BP神经网络模型训练误差曲线

输入层与隐含层间的连接权值矩阵 $IW\{1,1\}$ 为

$$IW\{1,1\} = \begin{bmatrix} 2.697 & -0.280 & -0.596 & 2.223 & -1.304 \\ 2.117 & 3.748 & -1.117 & -1.842 & -2.997 \\ -1.582 & 1.841 & -3.131 & -3.861 & 0.456 \\ -4.402 & 0.726 & -1.183 & -0.888 & -2.518 \\ 3.662 & 1.708 & 0.584 & 1.398 & -1.033 \\ 7.883 & -0.398 & -0.067 & 1.138 & -1.647 \\ -1.145 & -3.276 & 2.631 & 2.330 & 3.142 \end{bmatrix}$$

隐含层和输出层间的权值矩阵 $LW\{2,1\}$ 为

$$LW\{2,1\} = [2.158, 2.413, -1.147, -6.262, -1.833, 5.600, -2.676].$$

隐含层神经元的阈值矩阵 $b\{1\}$ 为

$$b\{1\} = [-4.238, -0.637, -3.504, 2.590, 3.073, 0.721, 3.907].$$

输出层神经元的阈值矩阵 $b\{2\}$ 为

$$b\{2\} = [-1.309].$$

图3显示出BP网络的训练速率非常慢,随着训练次数的增加,网络的误差缓慢下降,达到限定的训练次数1000次时,误差为0.098921,远远未达到目标误差值0.0001.这说明改进的BP网络明显优于标准BP网络,原因是改进网络中引入的附加冲量项和可变学习率提高了网络的训练速度.所以本实验采用改进的BP网络模型.

为检验网络性能,分以下两种情况进行仿真:

1)样本经学习训练后,用训练样本进行测试,检验网络的数据拟合能力;2)利用训练好的网络对测试样本进行预测,检验网络的泛化能力.具体仿真结果见表1,可以看出,标准BP网络建立的网络模型对训练样本预测的最大相对误差为26.83%,对测试样本预测的最大相对误差为49.14%;改进BP网络建立的网络模型对训练样本预测的最大相对误差为

2.78%,对测试样本预测的最大相对误差为4.23%.利用改进的BP网络建立的充填体预测模型远远优于标准BP网络建立的模型,改进的BP网络建立的模型不仅具备优良的拟合能力,同时也具备优良的泛化能力.

4 神经网络预测精度与训练结果分析

为了证明改进的BP网络在建立充填体强度模型方面的优越性,利用表1中的训练样本数据,采用二次多项式逐步回归方法得到充填体强度与影响因子之间的模型为

$$y = -36.982 + 104.458x_1 + 673.510x_2 - 329.753x_3 + 91.281x_4 + 10.370x_1^2 - 213.838x_2^2 - 319.711x_3^2 - 237.918x_4^2 - 3060.762x_1x_2 + 6.886x_1x_4 - 1932.409x_2x_3 + 679.094x_3x_5. \quad (6)$$

式中: y 为28d单轴抗压强度, x_1 为胶砂比, x_2 、 x_3 、 x_4 、 x_5 分别为水泥、石灰、石膏、矿渣的质量分数.

式(6)的相关系数 $R = 1.00000$, 剩余标准差 $S = 0.00326$, 调整后相关系数 $R_a = 0.99999$, 可以看出利用二次多项式逐步回归方法得到的回归模型已达到了完全相关的程度,说明充填体抗压强度与影响因子之间存在密切的联系.此回归模型对训练样本的拟合相对误差均为0,但是对测试样本的预测相对误差却在18.35%~115.6%之间,说明回归方法对训练样本以外的样本不具备预测能力.改进的BP网络对训练样本预测的最大相对误差为2.78%,对测试样本预测的最大相对误差为4.23%,说明改进的BP网络不仅对训练样本具有高精度的仿真能力,更重要的是对训练样本以外的预测样本同样具有高精度的预测能力.

为了进一步检验改进BP网络的网络性能,绘制了测试样本的网络训练结果分析图,如图4所示.

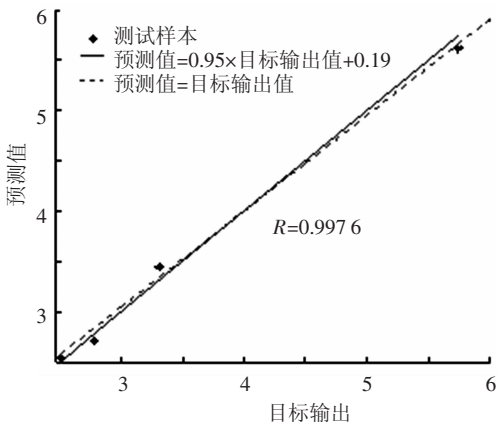


图4 神经网络训练结果分析(测试样本)

图4中,横坐标为目标输出即试验值,纵坐标为网络预测值(反归一化后的结果),实线为仿真拟合线,虚线为网络预测值等于试验值。 R 为网络输出与目标输出(即试验值)的相关系数,它越接近1,表明网络输出与目标输出越接近,网络性能越好。本预测得到的相关系数 $R = 0.9976$,说明该神经网络模型对充填体强度有较好的解释,可以用于充填体的强度预测,并且预测精度高。

5 结 论

1)改进的BP神经网络模型适用于从实例样本中提取特征、获取知识,可实现充填体强度和影响因子之间的非线性映射。

2)采用改进的BP神经网络模型,在石人沟铁矿建立了充填体强度的预测模型。该模型具有自动学习和充填体强度高精度预测的功能。

3)采用改进的BP神经网络模型的预测方法,可快速、准确、有效地确定充填体的强度,为充填体的强度预测提供了新的方法。

参考文献

- [1] 周爱民. 矿山废胶结充填[M]. 2版. 北京:冶金工业出版社,2010.
- [2] 李夕兵,刘志祥,古德生. 矿业固体尾废与采空区互为资源战略的思考[J]. 矿冶工程,2005,25(6):1-15.
- [3] 李一帆,张建国,邓飞,等. 深部采空区尾砂胶结充填体强度特性试验研究[J]. 岩土力学,2005,26(6):865-868.
- [4] 邓代强,高永涛,吴顺川,等. 水泥尾砂充填体劈裂拉伸破坏的能量耗散特征[J]. 北京科技大学学报,

2009,31(2):144-148.

- [5] 任奋华,蔡美峰,来兴平,等. 采空区覆岩破坏高度监测分析[J]. 北京科技大学学报,2004,26(2):115-117.
- [6] 赵才智,周华强,柏建彪,等. 膏体充填材料强度影响因素分析[J]. 辽宁工程技术大学学报,2006,25(6):904-906.
- [7] CHANG Q L, ZHOU H Q, HOU C J. Using particle swarm optimization algorithm in an artificial neural network to forecast the strength of paste filling material [J]. Journal of China University of Mining & Technology, 2008, 18(4):551-555.
- [8] 周开利,康耀红. 神经网络模型及其MATLAB仿真程序设计[M]. 北京:清华大学出版社,2005.
- [9] 韩力群. 人工神经网络教程[M]. 北京:北京邮电大学出版社,2006.
- [10] 丛爽. 面向MATLAB工具箱的神经网络理论与应用[M]. 合肥:中国科学技术大学出版社,1998.
- [11] 王旭,王宏,王文辉. 人工神经网络原理与应用[M]. 沈阳:东北大学出版社,2000.
- [12] HAGAN M T, DEMUTH H B, BEALE M H. 神经网络设计[M]. 戴葵,译. 北京:机械工业出版社,2002.
- [13] 钟珞,饶文碧,邹承明. 人工神经网络及其融合应用技术[M]. 北京:科学出版社,2007.
- [14] 崔阳,徐龙,刘艳,等. 基于改进BP神经网络的煤催化气化预测模型研究[J]. 燃料化学学报,2011,39(2):90-93.
- [15] YAO X. Evolutionary artificial neural networks [J]. International Journal of Neural Systems, 1993, 9(3):203-222.
- [16] 何玉斌,李新忠. 神经网络控制技术及其应用[M]. 北京:科学出版社,2000.

(编辑 魏希柱)