

岩石可爆性神经网络研究*

冯夏庭

(东北大学采矿系 沈阳 110006)

摘要 应用人工神经网络系统理论,采用机器学习的方法,建立了岩石的可爆性指数与岩体的爆炸漏斗体积 V 、大块率 K_1 、平均合格率 K_2 、小块率 K_3 和波阻抗 Z 之间的非线性映射关系,并将其用神经网络、网络连接权值矩阵和节点阈值向量分布式表达出来。对于新的岩石,网络采用并行推理的方法预报出其可爆性。实践表明,神经网络方法科学、具有较强的非线性动态处理的能力。

关键词 岩石 可爆性 神经网络 非线性映射 自学习 并行推理

1 引言

岩石可爆性研究是爆炸力学和爆破工程中的一项重要基础性工作,它可为爆破设计、施工、矿山企业制定定额和科学管理提供数据,为爆破工程的优化提供基本参量。多年来,国内外在岩石可爆性方面做了大量的研究,提出了岩石爆破性分级表和经验公式^[1]。回归分析、模糊综合评判^[2]、灰色关联度分析^[3]等,已应用到岩石可爆性研究上来。大量的实践表明,岩石对爆破作用的抵抗能力主要取决于岩石性质、作用于岩石的方式以及岩石和炸药性能和匹配等。影响岩石的可爆性能的因素众多,关系十分复杂,存在大量的定性描述和不确定性信息,这些因素难以在一个精确的数学方程中都包括进来。同一种岩石,由于其赋存条件不同,地质结构构造的复杂性和变化性不同,其性能参数差异很大,这就导致岩石可爆性的数据有时很难服从经典统计规律,因而应用统计方法就带来一定的困难。在模糊、灰色分析中,虽然能考虑一些定性描述,但权值、隶属度函数、功效函数需要靠统计方法或人为而定,这仍然无法摆脱人为因素的影响。

作者意在另辟新径,即利用人工神经网络系统理论,从积累的工程实例中学习知识,建立岩石的可爆性与其各种影响因素的关系,然后对其进行推广,由网络推理出新的岩石的可爆性能参数。

2 神经网络原理

人工神经网络系统理论是八十年代后期迅速发展起来的,目前已成为国际上的研究热点。人工神经网络是人脑神经网络结构的简化,它具有与人脑神经网络某些相似的特性,如自学习、自组织、非线性动态处理、分布式知识存储、联想记忆等,因而它特别适合于处理岩石爆炸力学等类似的复杂问题。

* 国家八·五重点攻关项目。

1994年1月10日收到原稿,4月13日收到修改稿。

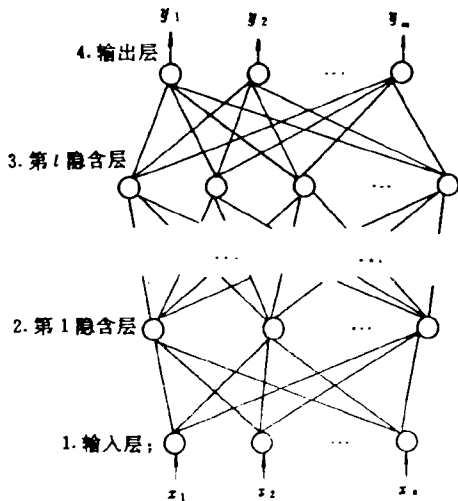


图 1 多层前馈神经网络模型

- 1. 输入层; 2. 第 1 隐含层;
- 3. 第 l 隐含层; 4. 输出层

Fig. 1 The model of a multilayer feedforward neural network

- 1. Input layer; 2. 1st hidden layer;
- 3. lth hidden layer; 4. Output layer

图 1 给出的是一个多层前馈神经网络模型。它由一个输入层、一个输出层和若干个隐含层组成,输入层的输入节点用于接受输入变量信息,输出层用于信息的输出。隐含层的作用是在输入信息之间建立一个复杂的非线性映射关系。这种非线性映射是网络通过对大量的工程实例样本的自学习而掌握的。

网络的学习过程是先从输入层开始正向计算各层神经元(节点)的输入和输出。当输出层节点的实际输出与事先给定的信号不一致或误差不满足要求时,则执行误差反向传播,逐层修改网络的连接权值,直至误差信号满足要求为止。

提供给网络学习用的 P 个实例样本集 $(x_1, d_1), (x_2, d_2), \dots, (x_p, d_p)$, 输入矢量 X 的期望输出矢量 $d = F(X)$, 并有 n 维欧氏空间到 m 维欧氏空间的映射

$$F: R^n \rightarrow R^m$$

网络学习的过程就是寻找一个 G

$$G: R^n \rightarrow R^m$$

使 G 与 F 最佳逼近。

对于当前网络 G , 其结构为 $N_i \rightarrow N_{u_1} \rightarrow N_{u_2} \rightarrow \dots \rightarrow N_{u_l} \rightarrow N_o$, 即输入层 N_i 个输出层 N_o 和 l 个隐含层 $N_{u_1} \dots N_{u_l}$, 网络实际输出 $G(x)$ 与输出信号矢量 $F(y)$ 有关

$$G(X) = F(Y) = [F_1(y_1), F_2(y_2), \dots, F_p(y_p)] \quad (1)$$

其中 $F_i(y_i)$ 为节点激励函数, 可取如下形式

$$F_i(x) = \frac{2}{1 + \exp[-(x + \theta)]} - 1 \quad (2)$$

式中, θ 为节点的阈值

对于当前随机地提供给网络学习的样本 x_k, d_k 第一隐含层节点 j 的输入为

$$h_j^k = \sum_{i=1}^n x_i^k W_{ij} \quad (3)$$

式中: x_i^k 为样本 k 的第 i 个输入变量值; W_{ij} 为输入节点 i 与其上层(隐含层 N_{u_1}) 节点 j 的连接权值。节点 j 的输出

$$o_j^k = F(h_j^k) = \frac{2}{1 + \exp[-(h_j^k + \theta_j)]} - 1 \quad (4)$$

对于第 l 隐含层 ($l=2, \dots, l$) 和输出层, 节点 r 的输入可统一写成:

$$h_r^k = \sum_q F_q(o_q^k) W_{rq} \quad (5)$$

输出写成

$$o_r^k = \frac{2}{1 + \exp[-(h_r^k + \theta_r)]} - 1 \quad (6)$$

式中: g_q^k 为样本 k 在节点 r 所在的层的下一层节点 q 的输出; W_{rq} 为节点 q 与节点 r 之间的连接权值。

定义一个平方误差能量 e_k :

$$e_k = d_k - F(y_k) = (d_1^k - F_1(y_1^k), d_2^k - F_2(y_2^k), \dots, d_m^k - F_m(y_m^k)) \quad (7)$$

式中: d_j^k 为样本 k 在 j 输出节点的期望输出; $F_j(y_j^k)$ 为输出节点 j 的网络计算输出。样本 k 的误差

$$E_k = \frac{1}{2} \sum_{j=1}^m [d_j^k - F_j(y_j^k)]^2 = \frac{1}{2} e_k e_k^T \quad (8)$$

系统的累计误差

$$E = \sum_k E_k \quad (9)$$

若 E (或 E_k) 不小于一允许值, 则进行误差反向传播, 逐层修改网络的连接权值。

对于输出节点 j 与第 l 隐含层的节点 q , 有连接权值修正量:

$$\begin{aligned} \Delta W_{qj}(k) &= - \frac{\partial E_k}{\partial W_{qj}} = - \frac{\partial E_k}{\partial y_j^k} \frac{\partial y_j^k}{\partial W_{qj}} = - \frac{\partial E_k}{\partial y_j^k} F'_q(h_q^k) = - \frac{\partial E_k}{\partial F_j(y_j^k)} \frac{\partial F_j(y_j^k)}{\partial y_j^k} F'_q(h_q^k) \\ &= - [d_j^k - F_j(y_j^k)] F'_q(h_q^k) \end{aligned} \quad (10)$$

对于第 l 隐含层的节点 q 与第 $l-1$ 隐含层的节点 r , 有权值修正量:

$$\begin{aligned} \Delta W_{rq}(k) &= - \frac{\partial E_k}{\partial W_{rq}} = - \frac{\partial E_k}{\partial h_q^k} \frac{\partial h_q^k}{\partial W_{rq}} = - \frac{\partial E_k}{\partial F_q} \frac{\partial F_q(h_q^k)}{\partial h_q^k} F'_r(g_r^k) \\ &= - \left[\sum_{j=1}^m \frac{\partial E_k}{\partial y_j^k} \frac{\partial y_j^k}{\partial F_q} \right] F'_q(h_q^k) F'_r(g_r^k) \\ &= - \left[\sum_{j=1}^m [d_j^k - F_j(y_j^k)] F'_j(y_j^k) W_{rj} \right] F'_q(h_q^k) F'_r(g_r^k) \end{aligned} \quad (11)$$

式中: g_r^k 为第 $l-1$ 隐含层 r 节点的输入; $F_r(g_r^k)$ 为第 $l-1$ 隐含层 r 节点的输出。

对于第 i 隐含层的节点 p 与第 $i-1$ 隐含层的节点 s ($2 \leq i \leq l-1$), 连接权值修正量为:

$$\begin{aligned} \Delta W_{sp}(k) &= - \frac{\partial E_k}{\partial W_{sp}} = - \frac{\partial E_k}{\partial t_s^k} \frac{\partial t_s^k}{\partial W_{sp}} = - \frac{\partial E_k}{\partial F_s} \frac{\partial F_s(t_s^k)}{\partial t_s^k} F'_p(f_p^k) \\ &= - \left[\sum_q \frac{\partial E_k}{\partial h_q^k} \frac{\partial h_q^k}{\partial F_s} \right] F'_s(t_s^k) F'_p(f_p^k) = - \left[\sum_q \frac{\partial E_k}{\partial h_q^k} W_{sq} \right] F'_s(t_s^k) F'_p(f_p^k) \end{aligned} \quad (12)$$

式中: $f_p^k, F_p(f_p^k)$ 为节点 p 的输入与输出; $t_s^k, F_s(t_s^k)$ 为节点 s 的输入与输出; W_{sq} 为第 i 隐含层的节点 p 与第 $i+1$ 隐含层节点 q 和连接权值。

对于第 1 隐含层的节点 s 与输入节点 i 的连接, 有权值修正量:

$$\begin{aligned} \Delta W_{is}(k) &= - \frac{\partial E_k}{\partial W_{is}} = - \frac{\partial E_k}{\partial f_s^k} \frac{\partial f_s^k}{\partial W_{is}} = - \frac{\partial E_k}{\partial F_s} \frac{\partial F_s(f_s^k)}{\partial f_s^k} x_i^k \\ &= - \left[\sum_r \frac{\partial E_k}{\partial t_r^k} \frac{\partial t_r^k}{\partial f_s^k} \right] F'_s(f_s^k) x_i^k = - \left[\sum_r \frac{\partial E_k}{\partial t_r^k} W_{rs} \right] F'_s(f_s^k) x_i^k \end{aligned} \quad (13)$$

在 $k+1$ 时刻, 权值 W_{ij} 被更新为

$$W_{ij}(k+1) = W_{ij}(k) + \eta_k \Delta W_{ij}(k) + \alpha_k \Delta W_{ij}(k-1) \quad (14)$$

式中: η_k 为 k 时刻的学习率, $\eta_k \in (0, 1)$; α_k 为 k 时刻的动量系数, $\alpha_k \in (0, 1)$ 。

3 岩石可爆性非线性映射的建立

大量的试验表明:在一定爆破技术、工艺、炸药性能要求的前提下,岩石的可爆性能与岩石的爆破漏斗体积 V 、大块率 K_1 、平均合格率 K_2 、小块率 K_3 和波阻抗等参量有关^[1]。因此,在建立岩石可爆性神经网络结构时,可将 V 、 K_1 、 K_2 、 K_3 、 Z 等参量用网络的输入节点表达,网络的输出节点输出反映岩石可爆性能的一个综合指数。在输入层与输出层之间构造几个隐含层,用以充分表达这些输入参量对岩石可爆性能参数的影响关系,按文献[7]的方法自适应地确定出该网络有两个隐含层,各有隐节点 5 和 4 个,由此建立了 5-5-4-1 的网络结构。

用 38 个实例样本数据训练上述网络,在经过 1835 次学习循环后,网络误差达到 0.001251,即认为网络对 38 个实例样本有了基本的掌握。表 1 列出了这些样本数据和网络对其的记忆输出。建立的输入参量 V 、 K_1 、 K_2 、 K_3 、 Z 与岩石可爆性能综合指数之间的非线性映射分布地表达在网络的连接权值矩阵(见表 2)和节点阈值向量(见表 3)上。

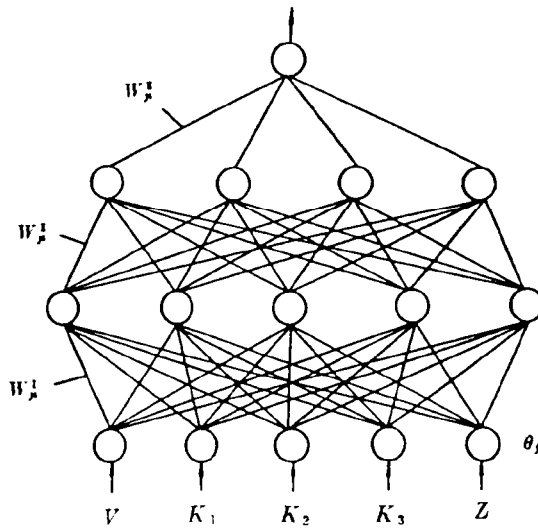


图 2 岩石可爆性神经网络结构

V —— 岩体爆破漏斗体积, K_1 —— 大块率, K_2 —— 平均合格率, K_3 —— 小块率, Z —— 波阻抗

Fig. 2 The structure of network for rock blastability

V —— Explosion carter volume of rock mass; K_1 —— Ratio of large-sized rock;

K_2 —— Average qualified ratio; K_3 —— Ratio of small sized rock; Z —— Wave impedence

用自学习方法学得岩石可爆性能与其各个输入参量之间的关系可以作为预测新的岩石的可爆性能的知识。网络利用这个知识进行并行推理预测出了未经网络学习过的样

表 2 岩石可爆性网络的连接权值
Table 2 The connection weights of rock blastability network learned from case histories

W_{μ}^j		输 入 节 点				
		x_1	x_2	x_3	x_4	x_5
第一 隐 含 层 的 节 点	1	-8.773455	-1.611980	0.567283	0.565747	2.177203
	2	1.816645	-2.593535	2.705406	-0.097287	-4.618602
	3	4.624455	-3.771679	1.075124	2.892252	-1.551670
	4	4.355222	0.706448	-2.491257	1.397529	-3.868519
	5	2.975601	-2.947345	1.449581	1.595527	3.960672
W_{μ}^j		第 一 隐 含 层 节 点				
		1	2	3	4	5
第 二 隐 含 层 的 节 点	1	-3.354825	-3.929960	-2.441555	-3.138778	2.185150
	2	-3.185389	-1.679069	0.727710	-2.657152	3.661368
	3	-1.145701	-1.406629	2.176790	0.089133	-0.454686
	4	2.493385	1.894585	-0.235594	2.683317	2.682216
W_{μ}^k		第 二 隐 含 层 节 点				
		1	2	3	4	
输出节点 y		0.602348	-0.575500	-0.892232	-2.095083	

表 3 岩石可爆性网络的节点阈值
Table 3 The node threshold of rock blastability network

θ_j	节 点 编 号				
	1	2	3	4	5
输入层	0	0	0	0	0
第一隐含层	-1.807019	-0.753823	3.106155	-1.934059	-1.910627
第二隐含层	0.645095	-0.983133	-1.053638	-0.913118	
输 出 层	-1.649433				

表 4 岩石可爆性的网络预报值

Table 4 The prediction of network for rock blastability

样本 编号	爆炸漏斗 体积 V/m^3	大块率 K_1 (%)	平均合格率 K_2 (%)	小块率 K_3 (%)	波阻抗 $Z/$ ($g/cm^2 \cdot s$)	岩石可爆性指数			
						实际值	网络预报值	绝对误差	相对误差
1	0.414	16.7	34.0	49.3	7.9	8.0	7.618554	0.381445	0.047681
2	0.356	11.2	33.8	55.0	8.5	7.0	7.223604	0.223604	0.031943
3	0.508	10.0	22.0	68.0	5.3	5.0	5.208771	0.208771	0.041754
4	0.293	2.7	16.1	81.2	10.0	6.0	5.843318	0.156682	0.026114
5	0.765	12.8	34.5	52.7	8.4	4.0	4.154635	0.154635	0.038659
6	0.057	30.3	47.4	22.3	9.4	14.0	14.132376	0.132376	0.009455
7	0.058	22.4	38.8	38.8	13.5	14.0	13.978539	0.021461	0.001533
8	0.097	39.2	48.4	12.4	7.7	15.0	14.827836	0.172164	0.011478
9	0.819	49.2	40.4	10.4	8.1	9.0	8.989563	0.010437	0.001160
10	0.122	35.0	54.3	10.7	12.6	15.0	14.740869	0.259131	0.017275
11	0.395	9.9	33.4	56.7	7.3	6.0	5.971676	0.028324	0.004721

4 讨 论

多次试验表明,网络学习参数 η 和 α 值选取的是否合理将影响到网络的收敛速度和收敛性能。网络在学习过程中,系统误差可能会停止在局部极小 $\{W\}_{local}$ 或某个稳定点 $\{W\}_{stationary}$ 或在这些点之间振荡,如图 3 所示。

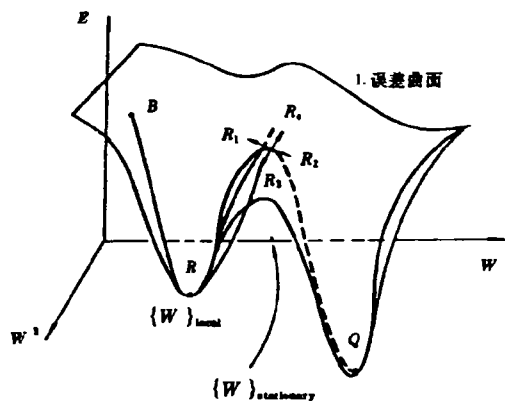


图 3 网络系统误差曲面

Fig. 3 Curvic surface of network system errors

1. Curvic surface of error

造成系统误差在局部极小 $\{W\}_{local}$ 点和某个稳定点 $\{W\}_{stationary}$ 之间振荡的一个主要原因可能是 η 或 α 值取的太大,适当地降低 η 或 α 值即解决这个问题。但过小的 η 将有可能使网络学习速度慢到无法忍受的程度。为了保证系统误差不出现振荡,又不致于使网络学习速度太慢,可以采取 η_t 可调的方法以适应系统误差的变化,网络开始学习 η_t 取大值,随

着学习过程的进行逐渐减小 η_k 值。

系统在学习过程中还有可能出现另一种现象,即系统误差停止在某个局部极小点或某个稳定点,由图3可知,系统误差曲线由B点出发到S点(极大值点或极小值点或其它点)。在S处,本来应该继续沿SQ到Q点(全局极小点),但由于S处各样本的调整值要求恰好反对称,致使其代数和为零或 $\eta\delta_{i,o_j} + \Delta W(n) = 0$,因而网络的权值调整为0,故网络系统误差永远走不出S点而陷于“原地踏步”的境地。避免系统误差在某个点稳定而不降低的一种作法,是将节点激励函数改为“温度”控制的Sigmoid型函数,这相当于给 $f(x)$ “升温”,从而改变现行数值对样本的对称关系。对于本文的连续值问题,取

$$f(x) = \frac{1 - \exp[-(x + \theta)/\theta_i]}{1 + \exp[-(x + \theta)/\theta_i]} \quad (15)$$

5 结 论

(1)应用人工神经网络方法,可以先不考虑岩石可爆性能与其各种影响因素之间到底存在什么样的函数关系,通过机器学习可以找出它们蕴含在实例样本中的内在联系——高度非线性映射。

(2)学得的岩石可爆性分析的知识是用网络的连接权值矩阵和节点阈值向量分布式存储的,因此该方法具有很高的容错性,它不会因为某个变量的输入数据丢失或含有噪声而影响网络的推理。

(3)神经网络方法的可靠性取决于学习样本的代表性,学习参数 η 、 α 对网络的收敛速度和收敛性能有影响,隐含层的数目直接影响网络处理非线性问题的能力。因此,提供良好的学习样本集,采用合适的网络结构和学习参数都是提高网络性能的关键。

(4)信息的传播由输入向输出并行进行,具有并行推理、高速度的特点。

参 考 文 献

- 1 东北工学院岩石爆破性分级科研组. 我国矿山爆破性分级的研究. 爆破, 1984, (1): 20~22
- 2 黄萍萍. 岩体可爆性的模糊综合评判分级法. 长沙矿山研究院季刊, 1989, (4): 63~72
- 3 于亚伦, 王德胜, 璩世杰. 水厂铁矿的岩石爆破性分区. 岩石力学与工程学报, 1990, 9(3): 195~201
- 4 Rumelhart D E. Parallel Distributed Processing: Exploration in the Microstructure of Cognition. MIT Press, 1986, 178~235
- 5 Lippmann R P. An Introduction to Computing with Neural Network. IEEE ASSP, 1987, 4(22): 4~22
- 6 Michalski R S. Machine Learning. Morgan Kaufmann Publishing Inc, 1986, 1~178
- 7 冯夏庭, 等. 神经网络在矿岩设计参数预测中的适应性研究. 东北工学院学报, 1992, 14(s1): 163~166
- 8 冯夏庭. 林韵梅, 岩石力学与工程专家系统. 沈阳: 辽宁科技出版社, 1993, 43~51

A STUDY ON NEURAL NETWORK ON ROCK BLASTABILITY

Feng Xiating

(*Northeastern University, Shenyang 110006*)

ABSTRACT With the application of artificial neural network theory and machine learning method, this paper establishes a nonlinear mapping between the blastability of rock and its affected factors such as volume of explosion crater, mass ratio of big rock blocks, ratio of small rock blocks, qualified ratio of blasting and wave impedance and represents them distributedly on neural network, connection weights and threshold of nodes. The blastability of new type of rock mass is predicted by means of the method of parallel inference. The results show that the proposed method has some more important advantages than traditional ones and it has strong ability for nonlinear dynamic processing.

KEY WORD rock mass, blastability, neural network, nonlinear mapping, self-learning, parallel inference.