

# 余震区长度与震级关系的模糊神经网络模型

王家鼎

(西北大学城市与资源学系, 西安 710069)

黄崇福

(北京航空航天大学, 北京 100083)

**摘要** 本文试图将模糊人工神经网络及模糊信息分析理论引入到余震区长度与震级关系的识别之中。采用信息扩散与BP型人工神经网络相结合的方法建立模型,该模型有较强的自适应能力及处理矛盾样本的功能。最后将该模型的识别结果与统计结果进行了比较,结果表明该模型具有一定的优越性。

**关键词:** 模糊识别 信息处理 余震 震级 模型

## 1 引言

有关地震断层与震级的关系问题,早已引起国内外学者的关注,郭增建等于1965年统计了各级地震地表破裂的最大长度与震级之间的关系<sup>(1)</sup>,提出了地震断层长度与震级之间的关系式。以后又有许多学者从不同的角度得到一些统计公式<sup>(2)</sup>。

地震的地表破裂的最大长度是反映地震断层规模或长度的重要依据。但是,在实际中有的地表破裂带可能是由于局部场地条件造成的,地表破裂带的资料不仅数量少,而且其可靠性也并不理想,因此,很难取得断层最大破裂长度的资料。

强震的余震区是地震断层发生应力调整的表现形式,其范围的大小即可反映震源体的大小,也可反映地震断层的规模。因此,人们用余震区长轴的长度来代表地震断层的长度。由于余震区长度与震级的关系是一种较复杂的非线性关系,因此,本文尝试用模糊人工神经网络及模糊信息分析理论研究它们之间的关系。

## 2 模糊人工神经网络

### 2.1 人工神经元

人工神经元结构模型如图1所示。此人工神经元一般称为神经元,有时也称为节点。其中 $u_i$ 为神经元的内部状态,有时也用来代表神经元本身; $\theta_i$ 为阈值; $X_i$ 为输出信号; $\omega_j$ 表示从 $u_j$ 到 $u_i$ 的连接权值; $S_i$ 表示外部输入信号。上述模型可描述为

$$\delta_i = \sum_j \omega_{ij} X_j + S_i - \theta_i \quad (1)$$

$$u_i = f(\delta_i) \tag{2}$$

$$y_i = g(u_i) = h(\delta_i), \quad h = g \circ f \tag{3}$$

当神经元没有内部状态时,可令 $y_i = u_i, h = f$ 。一般将 $n_{e,t_i} = \sum_{j=1}^n \omega_{ij} X_j$ 记为节点的输入。当 $X_j$ 来自于另一神经元输出时也可记为 $O_j$ ,在没有内部状态时,常常把阈值也可看作一权值,

则 $\lambda$ 值记为 $n_{e,t_i} = \sum_{j=0}^n \omega_{ij} \theta_j$ ,其中 $\omega_{i0} = -\theta_i, O_0 = 1$ 。一般称 $f$ 为激励函数。

2.2 BP 网络

BP (Back-Propagation learning Algorithm)网络如图 2 所示。它由输入层、隐层和输出层构成。其基本的算法过程是:将样本中的自变量参数值赋予输入节点,依权值和激励函数的作用在输出节点得到输出值。

设归一化后的知识样本为:

$$W = (x_1, x_2, \dots, x_s, \dots, x_s, \quad y_1, y_2, \dots, y_k, \dots, y_m) \tag{5}$$

一般可称输入层为 S 层,输出层为 K 层。在 S 层内,节点的输入值为:

$$n_{e,t_s} = X_s \tag{6}$$

其中左边的 S 依节点不同而变,节点的输出值为:

$$O_s = n_{e,t_s} \tag{7}$$

在第 i 层( $i \neq s$ ),节点的输入值记为

$$n_{e,t_i} = \sum \omega_{ij} O_j \tag{8}$$

输出值记为

$$O_i = f(n_{e,t_i}) = 1 / (1 + \exp(-n_{e,t_i})) \tag{9}$$

式(9)一般称为激励函数。

均方误差表征的能量函数为

$$E = \frac{1}{2} \sum_{k=1}^m (y_k - O_k)^2 \tag{10}$$

网络的学习就是调整 $\omega_{ij}$ ,使 E 趋于最小。

设输出层为 k 层,记

$$\delta_k = (Y_k - O_k) O_k (1 - O_k) \tag{11}$$

对其它层,有递推公式:

$$\delta_j = f_j(n_{e,t_j}) \sum_k \delta_k \omega_{kj} \tag{12}$$

得权值修正公式:

$$\omega_{ij}^{(t+1)} = \omega_{ij}^{(t)} + \eta_j \delta_j O_i + \alpha \Delta \omega_{ij}(t-1) \tag{13}$$

其中 $\Delta \omega_{ij}(t-1)$ 为前一步的修正量, $\omega_{ij}$ 的初值随机地在 $(-1, 1)$ 中选取; $\eta$ 称为学习速率; $\alpha$ 称为惯性系数。取 $\eta, \alpha \in (0, 1)$ ,依具体情况而定。

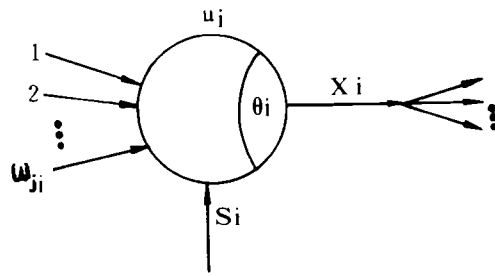


图 1 神经元结构模型  
Fig. 1 Model of neural structure.

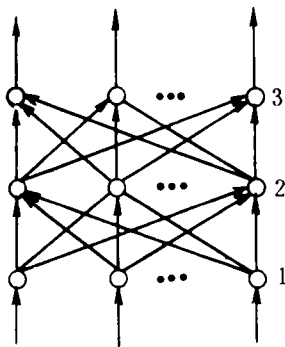


图 2 BP 网络

1. 输入节点; 2. 隐节点; 3. 输出节点

Fig. 2 BP neural network.

### 2.3 实可测 BP 网络中的信息扩散

如知识样本所在的因素空间为实可测空间,并且没有矛盾样本,也不涉及知识的非完备性问题,则可直接用传统 BP 网络学习样本知识,否则,需要构造状态网络,并用信息扩散转化样本点。

作者在信息分配的基础上曾提出过数种信息扩散方法,如有限双向累次分配信息扩散法<sup>[3]</sup>,Laplace 方程信息扩散法<sup>[4]</sup>,二次型信息扩散法<sup>[5]</sup>等。本文在此推导出一个二维正态型信息扩散公式。

设输入输出节点集  $U_r$  和  $O_s$  分别为

$$\begin{cases} U_r = \{u_1, u_2, \dots, u_n\} \\ O_s = \{o_1, o_2, \dots, o_m\} \end{cases} \quad (14)$$

$(x, y)$  是样点  $W$  的观测状态。二维正态信息扩散公式为:

$$\mu(u_i, o_j, x, y) = \frac{1}{2\pi h^2} \exp\left[-\frac{(u_i - x)^2 + (o_j - y)^2}{2h^2}\right] \quad (15)$$

用式(15)求得  $A = U_r \times O_s$  上的一个模糊关系矩阵,其中

$$h = \frac{1}{2}(h_r + h_s) \quad (16)$$

而  $h_r$  和  $h_s$  可由下式决定

$$h = \begin{cases} 1.6987 d/N & 1 \leq N \leq 5 \\ 1.4456 d/N & 6 \leq N \leq 7 \\ 1.4230 d/N & 8 \leq N \leq 9 \\ 1.4208 d/N & 10 \leq N \end{cases} \quad (17)$$

其中  $d = \|a - b\|$ , 即观测样本最大和最小状态  $(a, b)$  的距离。N 为样本容量。

令 
$$\mu'(u_i, o_j, x, y) = \mu(u_i, o_j, x, y) / \sum_x \mu(u_i, o_j, x, y) \quad (18)$$

记 
$$\mu'(u_i, o_j) = \sum_y \mu'(u_i, o_j, x, y) \quad (19)$$

则

$$\mu(u_i, o_j) = \mu'(u_i, o_j) / \sum_x \mu'(u_i, o_j) \quad (20)$$

(20)式为  $W$  在  $(u_i, o_j)$  点的联合落影。而

$$\mu_{o_j}(u_i) = \mu(u_i, o_j) \quad (21)$$

(21)式为  $W$  在  $o_j$  点的条件落影。以

$$\omega_j = (\mu_{o_j}(u_1), \dots, \mu_{o_j}(u_n), o_j) \quad j=1, 2, \dots, n \quad (22)$$

为样点,共有  $n$  个,即可用来训练 BP 网络。

## 3 余震区长度与震级关系的模糊神经网络模型

对于余震区长度与震级的关系,前人采用了以下统计公式经最小二乘法估计地震震级

$$M = a + b \ln L \quad (23)$$

式中  $M$  为地震震级,  $L$  为余震区长度,  $a, b$  为统计系数。

实际资料证明,  $M$  和  $L$  是一种较复杂的非线性关系。所以采用模糊信息处理与人工神经网络相结合的方法来处理这类较复杂的非线性关系是可行的。本文使用李玉龙先生收

集的我国 22 次地震的余震区长度资料作为样本(表 1)\*,并用另外 10 次地震的资料(表 2)作为检验样本。

表 1 强震震级及余震区长度(训练网络用)

序号	时间	地点	余震区长度(km)	震级(M <sub>s</sub> )
1	1920-12-16	宁夏海原	220	8.5
2	1927-05-23	甘肃古浪	160	8.0
3	1961-10-01	甘肃岷县	30	5.7
4	1962-05-21	青海崑布逊湖	90	6.8
5	1966-03-08	河北邢台	75	7.2
6	1970-01-05	云南通海	110	7.7
7	1971-03-23	新疆乌什	27	6.3
8	1972-01-16	新疆柯坪	40	6.2
9	1972-04-09	新疆轮台	26	5.6
10	1972-12-19	新疆西克尔	26	6.1
11	1973-06-03	新疆精河	52	6.0
12	1973-08-11	四川南坪	30	6.3
13	1975-02-04	辽宁海城	70	7.3
14	1976-04-06	内蒙和林格尔	15	6.2
15	1976-07-28	河北唐山	90	7.8
16	1976-08-16	四川平武	72	7.2
17	1976-08-17	四川盐源	40	6.7
18	1976-09-23	内蒙巴音木仁	35	6.2
19	1978-04-22	新疆库尔勒	36	5.8
20	1979-03-26	新疆库车	26	6.0
21	1979-08-25	内蒙五原	26	6.0
22	1981-01-24	四川道孚	45	6.9

表 2 强震震级及余震区长度(检验样本)

序号	时间	地点	余震区长度(km)	震级
1	1936	甘肃康乐	38.5	6 $\frac{3}{4}$
2	1963-04-25	云南云龙	19	6.0
3	1970	宁夏西吉	35	5.5
4	1973-02-06	四川炉霍	158	7.9
5	1974-05-11	云南昭通	60	7.1
6	1974-05-11	云南大关	65	7.1
7	1976-05-29	云南龙陵	100	7.5
8	1975-01-12	云南楚雄	12	5.6
9	1982	宁夏海原	51.3	5.7
10	1986	青海门源	47	6.7

\* 李玉龙,新编陕、甘、宁、青地震区划图原则与方法的研究,国家地震局兰州地震研究所研究报告,1986年

### 3.1 网络的构成和对网络的训练

由于余震区长度  $L$  与震级  $M$  皆可为实可测因素,因此可采用实可测 BP 网络进行处理。由式(22)知构造学习样本知识的 BP 网络输出节点只有一个,节点的输出值就是状态的预测值,由 BP 网络的激励函数特性可知,节点输出为 0 和 1 是极限情况。因此,将观测值应压缩到比  $[0,1]$  更小的区间,以便不包括 0,1 这两个点。

取  $C_0=0.1, C_1=0.8$ 。设  $L$  的最大最小值分别为  $a_1, b_1$ ;  $M$  的最大、最小值分别为  $a_2, b_2$ , 用下式对表 1 样本点压缩:

$$l = \frac{C_0 + C_1(L - b_1)}{a_1 - b_1}; m = \frac{C_0 + C_1(M - b_2)}{a_2 - b_2} \quad (24)$$

从表 1 可知,  $L$  归一化后的最大、最小值  $u_n, u_1$  及震级  $M$  的最大、最小值  $o_m$  和  $o_1$  分别为

$$u_1 = 0.0005, u_n = 0.8005; o_1 = 0.0345, o_m = 0.8345$$

按等步长取归一化控制状态

$$\left. \begin{aligned} U_l &= \{u_1, u_2, \dots, u_{10}\} = \{0.0005, 0.0894, \dots, 0.8005\} \\ O_m &= \{o_1, o_2, \dots, o_{10}\} = \{0.0345, 0.1234, \dots, 0.8345\} \end{aligned} \right\} \quad (25)$$

对归一化样本点在  $U_l \times O_m$  上用式(15)进行扩散,其  $h$  由下式求得:

$$h = \frac{1}{2} \left[ \frac{1.4208(u_{10} - u_1)}{22} + \frac{1.4208(o_{10} - o_1)}{22} \right] \quad (26)$$

按式(18)~(21)可求得表 1 所提供资料在  $U_l \times O_m$  上的条件落影,见表 3。

表 3  $U_l \times O_m$  上的条件落影

$u_i \backslash o_j$	0.0345	0.1234	0.2123	0.3011	0.3900	0.4789	0.5678	0.6567	0.7456	0.8345
0.0005	0.2634	0.2527	0.2826	0.1609	0.0260	0.0034	0.0001	0.0000	0.0000	0.0000
0.0894	0.5987	0.5430	0.5458	0.4725	0.2669	0.0589	0.0095	0.0003	0.0000	0.0000
0.1783	0.1327	0.1889	0.1603	0.1968	0.2927	0.3489	0.1795	0.0204	0.0056	0.0001
0.2671	0.0054	0.0152	0.0102	0.0953	0.2757	0.5054	0.3684	0.2211	0.0707	0.0007
0.3560	0.0000	0.0001	0.0011	0.0707	0.1322	0.0777	0.2980	0.3365	0.0815	0.0007
0.4449	0.0000	0.0000	0.0001	0.0039	0.0067	0.0057	0.1140	0.1188	0.0409	0.0014
0.5338	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0184	0.1610	0.2875	0.0160
0.6227	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0118	0.1298	0.2376	0.0167
0.7116	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0007	0.0077	0.0675	0.1995
0.8005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0018	0.2093	0.7655

由式(22)知,表 3 可提供给输入为 10 个节点,输出为 1 个节点的 BP 网络。取一个隐层,5 个隐节点,得  $10 \times 5 \times 1$  的 BP 网络。其中每一输入节点对应于样本输入的一个控制状态,输出节点对应  $M$  归一化后的因素名称,该节点的输出值直接为归一因素的状态值。此 BP 网络的输入节点为因素状态神经元,输出节点为因素神经元。用表 3 对该网络进行训练,取  $\eta=0.9, \alpha=0.7$ ,用几十分钟进行学习后便可达到精度。训练后网络的权值和阈值如表 4、表 5 所示。

### 3.2 用训练好的网络进行震级估计

设余震区长度  $L$  已知,用式(24)将其转化为  $[a, b]$  区间内的数,用线性信息分配公式<sup>[6]</sup>

$$u_i = 1 - \frac{|l - l_i|}{A} = 1 - \frac{|l - l_i|}{l_{i+1} - l_i} \quad (27)$$

表4 输入层到隐含层的权值、隐含层的阈值

输入层节点数	隐含层节点				
	1	2	3	4	5
1	-1.188	-1.092	0.578	0.837	1.567
2	-0.909	-0.824	0.926	0.557	0.891
3	-0.596	-0.457	0.058	-0.399	-0.126
4	-0.041	-0.072	-0.353	-0.271	0.036
5	0.281	0.751	-0.204	-0.143	-0.884
6	-0.082	-0.211	0.256	0.287	0.004
7	0.284	0.509	-0.574	-0.510	-0.648
8	0.254	0.632	-0.798	0.123	-0.361
9	0.698	0.569	-0.462	-0.234	0.114
10	0.772	1.051	-1.187	-1.021	-0.864
隐含层阈值	0.044	0.164	-0.636	-0.518	-0.867

将归一化L分配给式(25) $U_1$ 上的有关控制状态,其值就是节点的输入值,通过网络,得一个单变量输出。设输出为 $O(M)$ 。由式(24)的逆公式

$$M = [O(M)(a_2 - b_2) - c_0] / c_2 + b_2 \quad (28)$$

可得最终的震级估计值。由网络法对表1的22组资料进行内验,其相对平均误差为 $4.149862E-02$ 。李玉龙先生曾对这22组资料进行统计得出如下公式

$$M = 2.3 + 1.13 \ln L \quad (29)$$

我们利用式(29)对22组资料内验,其相对平均误差为 $4.399709E-02$ 。

表6 对10组样本检验结果

序号	L(km)	$M_s$	统计法( $M_s$ )	网络法( $M_s$ )
1	38.5	$6\frac{3}{4}$	6.4	6.2
2	19.0	6.0	6.6	5.9
3	35.0	5.5	6.3	6.0
4	158.0	7.9	8.0	8.1
5	60.0	7.1	6.9	7.0
6	65.0	7.1	7.0	7.2
7	100.0	7.5	7.5	7.7
8	12.0	5.6	5.1	5.7
9	51.3	5.7	6.7	6.8
10	47.0	6.7	6.6	6.7

表5 隐含层到输出层的权值、输出层的阈值

隐含层节点	输出层节点
1	1.813
2	2.056
3	-1.819
4	-1.379
5	-2.194
输出层阈值	-0.198

另外,我们同时利用两种方法对表 2 的 10 组资料进行检验(非母体检验),结果见表 6。结果表明模糊网络法的相对平均误差为 0.048,统计法的相对平均误差为 0.060。

在此需要说明的是,本问题的原始数据离散性较小,从而统计法结果和实际震级的相对平均误差与网络法结果和实际震级的相对平均误差的差别不是很大。我们曾对很多离散性较大的数据作过分析,两者的差别很大,显然网络法的精度高于统计法。因此,实可测 BP 网络法与模糊信息分析法相结合是解决复杂非线性问题的良好途径之一。

(本文 1993 年 12 月 10 收到)

### 参考方献

- 1 郭增建,等. 震源孕育模式的讨论,地球物理学报,1973,16:43—48
- 2 刘庆民. 震级与震源断裂长度关系的讨论. 西北地震学报,1979,1(1):60—65
- 3 Huang Chongfu, Wang Jiading. Principle of Information Diffusing Relevant to Fuzzy Speech Sound Recognition. BUSEFAL(法),1992,49:73—82
- 4 王家鼎,黄崇福. 模糊信息处理中的信息扩散方法及其应用. 西北大学学报(自然科学版),1992,22(4):383—392
- 5 黄崇福. 信息扩散原理与计算思维及其在地震工程中的应用. 北京师范大学博士论文. 北京:北京师范大学出版社,1992. 28—56
- 6 黄崇福,王家鼎. 模糊信息分析与应用. 北京:北京师范大学出版社,1992. 133—211

## A MODEL OF FUZZY NEURON ON THE RELATIONSHIP BETWEEN LENGTH OF AFTERSHOCK AREA AND MAGNITUDE

Wang Jiading

(Northwest University, Xi'an 710069)

Huang Chongfu

(Beijing University of Aeronautics and Astronautics, Beijing 100083)

### Abstract

By using information diffusion method combined with BP neural network, the authors suggest a kind of fuzzy neuron in present paper. The model of relationship between length of aftershock area and magnitude is discussed by the fuzzy neuron. This model has strong adaptability and can treat contradictory samples. Comparison of the results of this model with those of statistical method shows the advantage of the model.

**Subject words:** Fuzzy identification, Information processing, Aftershock, Magnitude, Model