

基于人工神经网络的地震活动性研究

冯利华

(浙江师范大学地理系, 浙江 金华 321004)

摘要: 人工神经网络通过神经元之间的相互作用来完成整个网络的信息处理, 具有自学习和自适应等一系列优点, 因而用它来进行地震活动性研究是可行的. 针对地震活动性问题, 初步建立了基于人工神经网络的计算分析系统, 给出了应用实例.

主题词: 地震活动性; 人工神经网络; BP 算法

中图分类号: P315.5 文献标识码: A 文章编号: 1000-0844(2000)04-0402-05

0 前言

在地震活动性研究中, 已经提出了许多重要的模型^[1,2], 古登堡的频度-震级关系($\lg N = a - bM$)就是其中一例. 但是, 频度-震级关系有一定的适用范围. 统计表明, 在单对数坐标系中, 小震频度系统地低于该直线关系. 曾有人把这一现象归结为小震的漏计, 但实际情况并非完全如此. 事实上, 这是分布本身的特点^[3], 因此只能说古登堡提出的这一直线关系还不能完全描述客观存在的频度与震级之间的函数关系. 由于近年来发展起来的人工神经网络(Artificial Neural Networks, 简称为 ANN)具有自学习、自组织、自适应和容错性等一系列优点, 在模式识别中已有成功的应用^[4,5], 为此, 拟根据 ANN 的原理和方法^[6,7]来进行地震活动性研究, 以供商榷.

1 ANN 的原理和方法

ANN 是由大量简单神经元广泛连接而成的复杂网络. 它是在现代生物学研究人脑组织的基础上提出来的, 可用来模拟人类大脑神经的思维活动^[8]. 它具有并行分布的信息处理结构, 通过对非线性函数的复合来逼近输入和输出之间的映射. 它不需要设计任何数学模型, 只靠过去的经验来学习, 通过神经元的模拟、记忆和联想, 处理各种模糊的、非线性的、含有噪声的数据, 采用自适应的模式识别方法来进行计算分析.

神经网络的算法包括 Hebb、Delta、Kohonen、BP 算法等. 其中 BP 算法即误差反传训练算法(Error Back Propagation), 是由 Rumelhart 等人于 1985 年提出来的, 它实现了 Minsky 多层网络的设想. 作为前馈式神经网络典型代表的多层感知机, 通常由输入层、输出层和若干隐含层组成, 其拓扑结构如图 1 所示.

采用误差反传训练算法的神经网络称为 BP 网络, 其学习过程由正向传播和反向传播组成. 在正向传播过程中, 样本信号经过 Sigmoid 函数[$f(x) = 1/(1 + e^{-x})$]的作用, 逐层向前

传播, 每一层神经元的状态只影响到下一层神经元的状态。如果在输出层不能得到期望的输出信号, 那么修改各层神经元的权值, 同时使输出信号的误差沿原路返回。经过反复传播, 最后使信号误差达到所要求的范围。

设网络具有 m 层, 令 y_j^m 表示第 m 层中第 j 个结点的输出, 而 y_j^0 就等于 x_j , 即第 j 个输入。令 W_{ij}^m 表示从 y_j^{m-1} 到 y_i^m 间的连接加权, θ_j^m 表示第 m 层第 j 结点的阈值。BP 网络训练步骤如下:

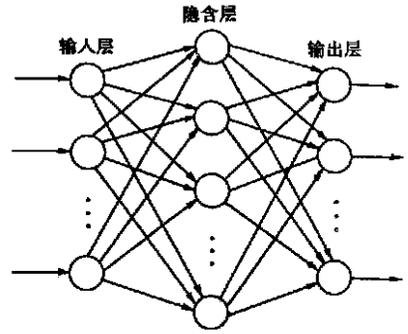


图 1 神经网络拓扑结构示意图
Fig. 1 The topological structure of artificial neural networks.

① 将各权值和阈值赋予 $(-1, 1)$ 间的随机数。

② 从训练数据组中选一数据对 (x^k, T_k) , 将输入变量加到输入层 ($m = 0$), 使得

$$y_i^0 = x_i^k \quad (\text{对所有 } i \text{ 点}) \quad (1)$$

式中 k 表示训练图形号。

③ 信号通过网络向前传播, 即利用关系式

$$y_j^m = F(s_j^m) = F\left(\sum_i W_{ij}^m y_i^{m-1} + \theta_j^m\right) \quad (2)$$

计算从第 1 层开始的各层内每个结点 j 的输出 y_j^m , 直至全部完毕。其中 $F(s)$ 取 Sigmoid 函数。

④ 计算输出层每个结点的误差值

$$\delta_j^m = y_j^m (1 - y_j^m) (T_j^k - y_j^m) \quad (3)$$

这个误差是由实际输出值和目标要求值之差获得的。

⑤ 计算前面各层每个结点的误差值

$$\delta_i^{m-1} = F'(s_i^{m-1}) \sum_j W_{ij}^m \delta_j^m \quad (4)$$

这要靠逐层反传误差算得 ($m = m, m - 1, \dots, 1$)。

⑥ 反向逐层修正权值和阈值

$$W_{ij}^m(t+1) = W_{ij}^m(t) + \eta \delta_j^m y_i^{m-1} + \alpha [W_{ij}^m(t) - W_{ij}^m(t-1)] \quad (5)$$

$$\theta_j^m(t+1) = \theta_j^m(t) + \eta \delta_j^m + \alpha [\theta_j^m(t) - \theta_j^m(t-1)] \quad (6)$$

式中: t 为迭代次数, η 为学习速率 [$\eta \in (0, 1)$], α 为动量因子 [$\alpha \in (0, 1)$]。

⑦ 返回到 ②, 转入下一个图形, 重复 ② ~ ⑦, 直至网络全局误差

$$E = \sum_k \sum_j (T_j^k - y_j^m)^2 / 2 \quad (7)$$

达到预设的精度为止。

神经网络训练好后, 其权值和阈值就已确定, 此时即可进行计算分析。

2 实例与测试

现以江苏及其邻区 ($31^\circ \sim 35^\circ \text{N}$, $118^\circ \sim 122^\circ \text{E}$) 的地震预报为例来说明 ANN 在地震活动性研究中的应用问题, 原始数据来源于参考文献 [9]。为了对地震的活动性指标择优劣汰, 这里利用灰色关联分析^[10]来进行。经过计算, 取关联度 $r > 0.984$ 的指标作为输入变量: x_1 为上一年发生次数最多的地震震级; x_2 为上一年度的 b 值 [$b = 0.4343 / (M - 0.5)$]; x_3 为上一年地震的平均震级; x_4 为上一年地震的平均纬度偏差; x_5 为上一年地震的平均经度偏差; x_6 为上一年地震的平均纬度; x_7 为上一年地震的平均经度。输出变量为江苏及其邻区的年最大震级 M_{\max}

(表 1). 因而输入层的结点数为 7, 输出层的结点数为 1. 隐含层的结点数根据 Kolmogorov 定律取为 15. 由此得到用于地震预报的神经网络的拓扑结构为 (7, 15, 1).

表 1 江苏及其邻区的地震活动性指标和年最大震级

	年份	x_1	x_2	x_3	$x_4/(\text{°})$	$x_5/(\text{°})$	$x_6/(\text{°})$	$x_7/(\text{°})$	M_{\max}	拟合震级
训练 样本	1976	2.6	0.391	2.611	1.010	1.105	32.900	121.800	4.3	4.305
	1977	1.8	0.641	2.436	1.225	1.095	34.011	121.101	5.0	5.004
	1978	1.8	0.608	2.214	1.910	0.744	33.060	119.300	3.9	3.900
	1979	2.3	0.452	2.461	1.014	0.934	32.270	120.300	6.3	6.148
	1980	2.1	0.498	2.372	1.118	0.608	31.161	119.101	3.5	3.556
	1981	1.8	0.657	2.161	1.181	0.957	33.022	121.172	3.9	3.890
	1982	1.8	0.600	2.224	1.099	0.937	31.150	119.101	5.0	5.001
	1983	2.1	0.498	2.373	0.891	0.804	32.303	120.243	3.5	3.636
	1984	1.8	0.615	2.206	1.052	0.957	32.303	120.041	6.3	6.208
	1985	2.4	0.357	2.718	0.459	0.527	32.303	121.221	4.1	4.088
	1986	2.6	0.435	2.499	0.587	0.727	32.273	120.254	5.1	5.103
	1987	2.3	0.449	2.468	1.056	1.063	32.172	121.194	5.4	5.413
	1988	2.3	0.432	2.504	0.491	1.055	33.210	120.172	3.6	3.616
	测试	1989	2.3	0.440	2.486	0.986	0.884	32.183	121.251	4.0

为了加快收敛速度, 需要对原始数据进行规一化, 即

$$x_i = (x_i - x_{\min}) / (x_{\max} - x_{\min}) \quad (8)$$

式中: x_{\max} 和 x_{\min} 分别为系列中的最大值和最小值. 可知 x_i' 位于 $[0, 1]$ 之内.

经过规一化处理的原始数据加载到 BP 网络后, 即可选取一定的训练参数对其进行训练和学习. 这里学习速率 η 取 0.85, 动量因子 α 取 0.60. 为了对训练和学习过的网络进行测试, 将 1976 ~ 1988 年的资料作为训练样本, 1989 年的资料作为测试样本. 训练样本经过 10 000 次的训练和学习, 网络全局误差 $E = 0.003$, 小于预设精度, 故网络收敛. 由网络输出的拟合值列于表 1. 从表中可知, 震级拟合值的平均误差只有 0.038 级, 最大误差也只有 -0.152 级, 故拟合情况良好.

由于训练过的网络已模拟并记忆了输入变量和输出变量之间的函数关系, 因而可以用它来进行地震预报. 测试结果也列于表 1, 1989 年预报的年最大震级为 3.923 级, 而实际发生的年最大震级为 4.0 级, 可见计算结果是好的.

3 讨论与比较

前已述及, 在地震活动过程中, 许多因变量 y 与自变量 x 之间存在着一定的关系, 即 $y = f(x)$, 但这种关系无法用一个具体的函数来描述, 即使用最简单的频度 - 震级关系 $[N = f(M)]$ 也难以表示出来. 而神经网络的最大特点在于不需要设计任何数学模型, 只要通过对历史数据的训练和学习, 网络就能够模拟并记忆输入变量和输出变量之间的任何复杂的函数关系, 并通过联想来实现计算分析.

就频度 - 震级关系 $[N = f(M)]$ 而言, 表 2 列出了中国 1901 ~ 1976 年震级 $M \geq 6$ 的 502 次强震^[11]、宁夏及其邻区 1966 ~ 1980 年 $M \geq 1$ 的 2 480 次地震^[12]. 同时列出了由神经网络计算的拟合值 [拓扑结构为 (1, 3, 1), 学习速率 $\eta = 0.85$, 动量因子 $\alpha = 0.60$, 学习和训练次数为 400 万次]、用古登堡的直线关系计算的拟合值 (最小二乘法). 从表中可以看到, 除了大震频

度一侧的 2 个点以外, 神经网络的计算结果均优于直线关系的计算结果.

表 2 中国强震与宁夏及其邻区地震

中国强震				宁夏及其邻区地震			
震级 M	累积频度 N /次	神经网络拟合值/次	直线关系拟合值/次	震级 M	累积频度 N /次	神经网络拟合值/次	直线关系拟合值/次
6.0	502	501.84	654.47	1	2480	2478.34	7089.61
6.5	222	222.00	218.29	2	1999	1999.00	1291.52
7.0	95	94.97	72.81	3	581	581.06	235.28
7.5	29	29.25	24.28	4	89	88.46	42.86
8.0	9	7.95	8.10	5	4	7.61	7.81
8.5	2	3.54	2.70	6	1	1.56	1.42

在地震活动性分析中, 古登堡的直线关系被用来外推地震平均次数

$$N_0 = N/m = 10^{a-bM}/m \tag{9}$$

式中: m 为统计年数, N_0 即为某一地区未来 1a 内震级 $\geq M$ 的地震平均次数. 就宁夏及其邻区 1966 ~ 1980 年发生的地震(表 2)而言, 根据最小二乘法, 求得

$$N_0 = N/m = 10^{4.590 - 0.7395M}/15 \tag{10}$$

由式(10)外推的宁夏及其邻区 1981 年震级 $\geq M$ 的地震次数列于表 3. 由神经网络外推的结果和宁夏及其邻区 1981 年实际发生的震级 $\geq M$ 的地震次数也列于表 3. 从表中可以看到, 由神经网络外推的地震次数和实际情况比较接近, 而直线关系由于不能完全描述频度与震级之间的函数关系, 因而外推的地震次数和实际情况相差甚远^[13].

4 结语

综上所述, 在地震活动过程中, 许多因变量与自变量之间的关系无法用具体的函数来描述, 而神经网络不需要设计任何数学模型, 它通过神经元之间的相互作用来完成整个网络的信息处理, 根据对非线性函数的复合来逼近输入和输出之间的映射, 具有自学习、自组织、自适应和容错性等一系列优点, 因而用它来进行地震活动性研究是可行的, 计算结果也表明了这一点.

本文针对地震活动性问题, 初步建立了基于 ANN 的计算分析系统, 开发了实用的功能软件. 这是地震活动性研究的一种新的尝试, 当然, 还有其它一些有关问题需要深入探讨, 但可以相信, 神经网络技术在地震活动性研究中具有广阔的应用前景.

表 3 两种方法外推结果的比较

震级 M	直线关系外推的地震次数	神经网络外推的地震次数	实际发生的地震次数
1	472.62	165.22	183
2	86.10	133.27	106
3	15.69	38.74	34
4	2.86	5.90	5
5	0.52	0.51	0
6	0.09	0.10	0

[参考文献]

[1] 陆远忠, 陈章立, 王碧泉, 等. 地震预报的地震学方法[M]. 北京: 地震出版社, 1985. 236—267.
 [2] 冯利华. 物元分析在地震预报中的应用试验[J]. 地震学报 1998, 20(6): 635—639.
 [3] 李全林, 于泳, 郝柏林, 等. 地震频度-震级关系的时空扫描[M]. 北京: 地震出版社, 1979. 73.
 [4] 夏元友, 朱瑞庚, 李新平. 基于神经网络的岩质边坡稳定性评估系统研究[J]. 自然灾害学报, 1996, 5(1): 98—104.
 [5] 李祚泳, 邓新民. 神经网络在台风预报中的应用初探[J]. 自然灾害学报, 1995, 4(2): 86—90.
 [6] 周继成, 周青山, 韩飘扬. 神经网络——第六代计算机的实现[M]. 北京: 科学普及出版社, 1993. 47—51.
 [7] 陈天平. 神经网络及其在系统识别中的逼近问题[J]. 中国科学(A), 1994, 24(1): 1—7.

- [8] Lippmann R P. An introduction to computing with neural nets[J]. IEEE ASSP Magazine, 1987, 4(2): 4—22.
- [9] 冯德益, 林命周, 顾瑾平, 等. 模糊地震学[M]. 北京: 地震出版社, 1992. 267.
- [10] 邓聚龙. 灰色系统(社会, 经济)[M]. 北京: 国防工业出版社, 1985. 49—52.
- [11] 国家地震局. 中国强震简目[M]. 北京: 地震出版社, 1977. 6—19.
- [12] 宁夏回族自治区地震局. 宁夏地震目录[M]. 银川: 宁夏人民出版社, 1982. 172—274.
- [13] 冯利华, 吴樟华. 地震活动趋势的逐步回归周期分析[J]. 地震学刊, 1997, (1): 42—45.

ON THE SEISMIC ACTIVITY BASED ON THE NEURAL NETWORK

FENG Li-hua

(*Department of Geography, Zhejiang Normal University, Jinhua 321004, China*)

Abstract: The artificial neural network can complete the news treatment of the whole network by means of a mutual function between the neural units, and has a lot of merits of self-learning, self-adjusting and suchlike. Therefore it is feasible that the seismic activity is researched by the neural network. The analysis system based on the neural network is initially set up and an applied example is put forward in the light of the problem of seismic activity.

Key words: Seismic activity; Artificial neural network; Back-propagation algorithm

(上接 375 页)

ON CHARACTERISTICS OF SEISMIC WINDOW ANOMALY DURING THE JIASHI EARTHQUAKES SWARM

LI Ying-zhen, QU Yan-jun

(*Xinjiang Seismological Bureau, Urumqi 830011, China*)

Abstract: A large number of seismologic windows in a great area, especially in the Tian shan earthquake belt, showed clear anomaly during the Jiashi strong earthquake swarm from 1997 to 1998. Strong earthquakes of the swarm can be divided into five groups according to the occurrence time. The first strong earthquake of each group can be checked out by the monthly and ten-days microseism frequencies of Kashi and Atushi seismic windows. The succedent strong earthquakes of each group can be checked out by the daily microseism frequency of the Kashi and Atushi seismic windows. It can be found from the anomalous features that strong earthquakes occurred usually after a high value of monthly frequency or after a low value of ten-days frequency and after a low value of daily frequency. It can be also found from the temporal process of the window anomalies that the duration and interval of the anomalies of Kashi and Atushi seismic windows during Jiashi earthquake swarm are longer than those of a single event. The duration of the anomaly during the first strong shock of each group is longer than that of the succedent strong earthquake. The daily frequencies of the two windows notably increased before the succedent strong earthquakes occurred from the exponential attenuating background in 1997, but in 1998 the daily frequencies increased before the strong earthquake occurred in the normal background. The results of the analyses indicate that the earthquake window is a valuable method in earthquake prediction.

Key words: Seismic window; Seismic frequency; Anomalous character; Strong earthquake sequence; Jiashi earthquake