

应用人工神经网络实现潜在震源区的定量划分*

姜效典 张维岗

(中国青岛 266003 青岛海洋大学)

摘要

本文提出了应用人工神经网络原理, 确定反映潜在震源区特征的地震构造和地震活动因素、定量划分潜在震源区的方法。对华北地区的实际资料计算表明, 该方法充分反映了潜在震源区与其控制特征之间的规律。原理简明, 便于应用, 计算结果合理, 分区细致, 资料具实用性。

关键词 神经网络; 潜在震源区; 定量划分

引言

潜在震源区是指未来具有发生破坏性地震潜在可能的地区。它以地震构造和地震活动为主要特征, 描述了未来地震活动强度在空间上的分布特点。它的划分是地震区划和地震危险性分析的基础。随着地震工程的需要和地震研究的进展, 潜在震源区的划分显得越来越重要。

近年来, 人们对定量划分潜在震源区, 进行了方法研究和探索, 取得了一些成果(杜兴信、丁韫玉, 1990; 叶洪等, 1988)。这些方法的基本思想是根据专家经验总结特征和规律, 建立一些判别规则, 用计算机来进行计算、判别。这是一种传统人工智能的设计思想, 方法的效果直接依赖于设计者所能考虑到的规律, 而且当规则增多时计算速度会迅速下降。笔者提出的人工神经网络方法划分潜在震源区的基本思想是用样本进行训练, 网络自行学习规律, 这便大大增加了方法的灵活性, 使其性能不依赖于设计者对潜在震源区划分的具体经验。对于新增加的样本, 可在原有基础上对网络进一步训练, 而识别的计算量并不增加。

人工神经网络是一种自动模式识别形式, 是目前认识水平上对人类(或动物)神经系统的某种模拟。网络一般由许多简单计算单元(即神经元)的大量相互连接构成, 而这些连接的强度(权值)是通过训练自动调整的, 即所谓学习。网络无需人们事先给出特征和规律, 通过样本学习, 便可以自己总结规律, 并将学到的规律分布于连接权值中。可

* 1993 年 7 月 17 日收到初稿, 1994 年 6 月 7 日决定采用。

见, 神经网络对于解决那类已知输入和输出, 但不清楚其相互间规律的问题有着比其它任何算法无可比拟的优势. 目前, 人工神经网络已广泛应用于图象处理和识别领域, 在地球物理学中已用于合成测井曲线估计岩性(许建华, 1993)、地震资料的初至拾取和道编辑(张学工、李衍达, 1992; McCormack *et al.*, 1993)、测定物探数据中地下目的层位置(Poulton *et al.*, 1992)等.

我们将选择性能相对稳定的反向传播神经网络, 经过适当的训练和学习, 完成潜在震源区的定量划分. 方法分两部分实现: 一是用模式识别的 ID3 方法, 确定出最能反映震源区特征的地震地质因素, 包括地震构造和地震活动标志; 二是用误差反向传播的神经网络模型, 通过对连接网络的权系数进行训练, 完成潜在震源区的定量划分.

1 原 理

1.1 模式识别 ID3 方法(包约翰, 1992)

ID3 方法是一种利用模式集给出顺序测试特征因素, 从而形成由特征因素顺序组合决定模式归类的判决树方法. 它将模式集作为一个原始系统, 计算其熵值, 然后对每个特征因素进行测试, 将模式集划分为不同的分支子模式集, 计算其系统熵降低, 产生最大熵降低的特征因素为最有效的分类因素, 以此从众多的特征因素中选出最能反映震源区特征的地震地质因素.

设有一模式集 X 含有 N 个模式, 分别属于 $c_i (i=1, 2, \dots, C)$ 种类别, 类别 c_i 中模式的数目为 N_i , 每个模式有 k 个特征因素, 每个特征因素有 J_k 个值. ID3 方法如下:

(1) 计算初始熵值:

模式集 X 的系统初始熵为

$$\text{熵}(X) = \sum_{i=1}^c - \left(\frac{N_i}{N} \right) \cdot \log_2 \left(\frac{N_i}{N} \right) \quad (1)$$

(2) 选择判决树的根节点:

1) 对每一个特征因素 $A_k, k=1, 2, \dots, K$, 依次按其取值 $k_j, j=1, 2, \dots, J$, 把原始模式集分成 J 个子模式集, 构成第一级分支子模式集群, 对应于 a_{kj} 的子模式集中的模式数为 n_{kj} .

2) 每个分支子模式集的 n_{kj} 个模式, 属于 c_i 的模式数是 $n_{kj}(i)$, 计算各分支子模式集的熵

$$\text{熵}(X, A_{kj}) = \sum_{i=1}^c - \left[\frac{n_{kj}(i)}{n_{kj}} \right] \cdot \log_2 \left[\frac{n_{kj}(i)}{n_{kj}} \right] \quad (2)$$

对应于特征因素 A_k 的分支子模式集群系统的熵

$$\text{熵}(X, A_k) = \sum_{i=1}^J \left(\frac{n_{kj}}{N} \right) \cdot \text{熵}(X, A_{k,i}) \quad (3)$$

由特征因素 A_k 引起的熵降低为

$$\Delta \text{熵}(k) = \text{熵}(X) - \text{熵}(X, A_k) \quad (4)$$

3) 选择导致最大熵降低的特征因素 A_{k_0} :

$$\text{若 } \Delta \text{熵}(k_0) > \Delta \text{熵}(k) \quad k=1, 2, \dots, K \quad k \neq k_0 \quad (5)$$

则 A_{k_0} 即为判决树的根.

(3) 将这一级子模式集群作为原始模式集, 其熵作为初始熵, 再由 A_k , $k \neq k_0$ 中选择一个特征因素 A'_k 作为判决树下一级的节点, 使新得到的下一级分支子模式集群的系统熵降低最大.

(4) 重复(2)和(3), 直到系统的熵等于零. 这时每一分支子模式集均为同一类别的模式所构成. 判决树各级节点处的特征因素及其取值的顺序组合, 即为不同模式类别属性的判据.

1.2 误差反向传播的人工神经网络(Lippmann, 1987)

神经元类型(输入输出关系)、网络构成形式(各神经元之间的连接方式), 以及连接权重的学习规则是一个神经元网络的三个要素. 本文选择的网络结构如图1所示. 图中,

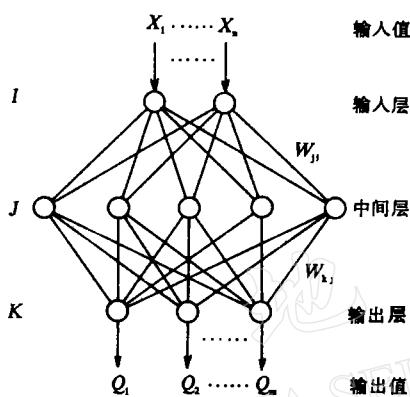


图 1 网络结构图

X_i 是输入向量的各分量, 以固定强度 1 连接到各神经元, Q_i 是相应的神经元输出各分量, 网络的神经节点用空心圆表示, 互相连接的权 W_{ji} 用线表示, W_{ji} 为第 j 个神经元到第 i 个神经元的连接强度(权值), 每一层的节点输出送到下一层节点, 这些输出值由于连接权值不同而被放大、衰减或抑制. 除输入层外, 每一节点的输入为前一层所有节点输出值的加权和, 每一节点的输出由节点输入、激励函数和偏置量(阈值) θ 决定.

对第 j 层网络, 节点输入为

$$net_j = \sum W_{ji} O_i + \theta_j \quad (6)$$

输出值为

$$O_j = f(net_j) \quad (7)$$

激励函数

$$f(x) = \frac{1}{1 + e^{-x}} \quad (8)$$

网络的实际输出 O_k 与期望输出 Q_k 之间存有差异 δ .

$$\text{输出层} \quad \delta_k = O_k(1 - O_k) \cdot (Q_k - O_k) \quad (9)$$

$$\text{中间层} \quad \delta_j = O_j(1 - O_j) \cdot \sum_k (\delta_k W_{kj}) \quad (10)$$

可见, 误差是由最上层(输出层)开始计算, 逐渐向下“传播”, 故称之为“误差反向传播”神经网络模型.

修改权值和阈值

$$W_{ji}(t+1) = W_{ji}(t) + \eta \cdot \delta_j \cdot O_j + \alpha \cdot [W_{ji}(t) - W_{ji}(t-1)] \quad (11)$$

$$\theta_j(t+1) = \theta_j(t) + \eta \cdot \delta_j + \alpha \cdot [\theta_j(t) - \theta_j(t-1)] \quad (12)$$

这里, η 为学习速度, 即步长. η 越大, W , θ 改变越剧烈. $\alpha > 0$ 为惯性系数, 影响着网络学习过程的收敛情况.

通过调整权值和阈值, 使 δ 逐渐减小. 当 δ 小于要求精度时, 便认为网络已训练完毕, 可以进行识别计算.

2 人工神经网络划分潜在震源区的方法

以华北地区为例 (32° — 42° N, 108° — 123° E). 首先从众多的表征华北地区潜在震源区特征的地震地质参数中, 用 ID3 方法确定出最佳控制因素.

选取已知地震震级的、有代表性的震源点作为研究样本; 选取能表征震源点地震特征并可测得的地震地质因素为样本的特征因素. 据此, 参考国家地震局分析预报中心、地质研究所 1986 年编制的华北地区 1987—2000 年地震区划判定图册和中国地震目录, 在研究区内选择 24 个样本模式 ($N=24$), 构成模式集 X , 样本地理点见图 2. 每个样本可测得的特征因素有 10 个 ($K=10$), 它们是: {NNNE-NE 向活动断裂, NW 向活动断裂, 盆地边缘主控断裂, 新生代沉积厚度, 地形高差, 活动断裂交汇点数, 重力区域异常梯度, 航磁 ΔT_a 上延 10 km 异常梯度, 地热流量, 地壳厚度}. 模式集 X 中有类别 5 种 ($c=5$, 以地震震级划分), c_1, c_2, \dots, c_5 各表示震级 $M \geq 8, 7 \leq M < 8, 6 \leq M < 7, 5 \leq M < 6, M < 5$ 的 5 类潜在震源区, 如图 2 所示. 属类别 c_1 的模式 $N_1=5$, 属类别 c_2, c_3, \dots, c_5 的模式分别为 $N_2=5, N_3=4, N_4=4, N_5=6$. 样本的 10 个特征因素分别以变

量 A_1, A_2, \dots, A_{10} 记, 表示构造特征的 A_1, A_2, A_3 因素赋二元数值量化. 如 A_{i2} 为“0”表示第 i 个样本无 NW 向断裂经过, 为“1”表示有 NW 向断裂经过, A_7, A_8 的异常梯度以 $10 \text{g.u. (20 nT)}/3 \text{ km}$ 计算, 其余因素赋实测值. 对这 10 个因素用文中所述 ID3 原理计算, 得到模式集 X 的判决树, 可知因素 A_5, A_{10} 对研究区的控震作用不明显. A_3 和 A_9, A_7 和 A_8 表示的是同一属性数据, 可以合并. 因此认为, 控制研究区内潜在震源区特征因素的组合以 {NNNE-NE 向活动断裂, NW 向活动断裂, 盆地边缘主控断裂, 新生代沉积厚度, 活动断裂交汇点数, 航磁 ΔT_a 上延 10 km 异常梯度} 最显著.

控制潜在震源区的地震地质因素确定后, 便可用误差反向传播网络对研究区进行定量划分.

为了完成区划, 要将研究区域划分成若干个单元. 本文以构造线、地层界线、岩性界线作为单元的边界, 将华北地区划分成 44 个计算单元 (见图 3). 同一单元内的每个地震地质因素具一致性, 其均值作为该单元的取值, 所有单元都用文中据 ID3 方法选出的 6 个显著参数作为特征因素. 现在我们要识别这些单元各属于震级 $M \geq 8, 7 \leq M < 8, 6 \leq M < 7, 5 \leq M < 6, M < 5$ 等 5 种类别的哪一类.

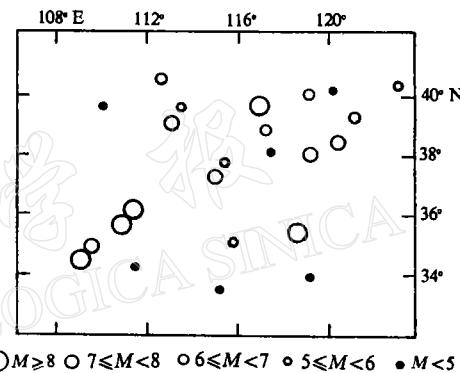


图 2 样本模式分布

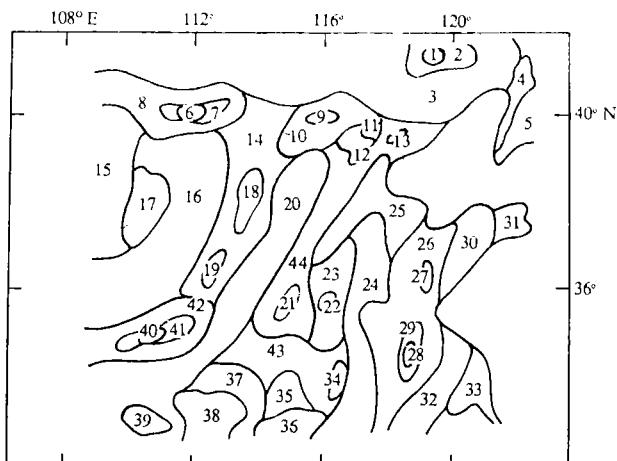


图3 单元划分图

所述原理,置各权值和阈值的初始值 W_{ij} 和 θ_j 为小的随机数,但不应完全相等,否则它们会始终相等。据式(7)计算网络中间单元状态和实际输出,用式(9)、(10)得到训练误差。依此误差,按式(11)、(12)修改权值和阈值,如此反复计算、迭代,使权值收敛于一组稳定值,网络学习成功。

表1 网络学习样本取值

单元号	x_1	x_2	x_3	x_4	x_5	x_6	单元属性	期望输出
11	0	1	0	9	8	0.80	$M \geq 8$	(1, 0, 0, 0, 0)
40	1	0	1	20	4	0.67	$M \geq 8$	(1, 0, 0, 0, 0)
9	1	0	1	10	4	0.60	$7 \leq M < 8$	(0, 1, 0, 0, 0)
19	1	1	0	10	1	0.40	$7 \leq M < 8$	(0, 1, 0, 0, 0)
27	1	0	0	4	6	0.60	$7 \leq M < 8$	(0, 1, 0, 0, 0)
41	0	0	1	15	8	0.80	$7 \leq M < 8$	(0, 1, 0, 0, 0)
10	0	1	0	5	1	0.96	$6 \leq M < 7$	(0, 0, 1, 0, 0)
23	1	0	1	0	5	0.60	$6 \leq M < 7$	(0, 0, 1, 0, 0)
26	1	0	1	0	6	0.57	$6 \leq M < 7$	(0, 0, 1, 0, 0)
33	0	1	1	0	4	0.80	$6 \leq M < 7$	(0, 0, 1, 0, 0)
5	1	0	0	10	0	0.50	$5 \leq M < 6$	(0, 0, 0, 1, 0)
20	1	0	1	12	0	0.60	$5 \leq M < 6$	(0, 0, 0, 1, 0)
32	1	0	1	8	1	0.56	$5 \leq M < 6$	(0, 0, 0, 1, 0)
36	1	0	0	5	4	0.80	$M < 5$	(0, 0, 0, 0, 1)
37	1	0	0	4	2	0.92	$M < 5$	(0, 0, 0, 0, 1)

从公式(8)可以看出,由 f 表示的激励函数,只有在权值为无穷时,节点的输出才会是 0 或 1。因此,对于本文使用的二值化目标输出,在网络训练时,输出只要达到 0.9 或 0.1,就认为学习成功。步长 η 取 0.2, 惯性系数 α 取 0.5, 在 MV-6 000 超小型机上(程序用 C 语言编制)迭代 2 086 次,运行近 40 min,网络训练完毕。

构造如图 1 所示的网络结构,使网络有 6 个输入节点,分别输入每个单元的 6 个特征因素,以 x_1, x_2, \dots, x_6 记;网络有 5 个输出节点,对应于 5 种可能震级类别的一种。某一输出节点的输出信号若是 0,则表示该类别震级不存在;若是 1,则表示该类别震级存在。网络中间层由 8 个神经元组成。

首先对网络进行训练,从图 3 中选择 15 个研究程度较高、能代表 5 种地震震级的单元作为学习样本,取值列于表 1。按文中

为进一步验证所设计网络的有效性, 我们选取了样本集以外的两组数据参加网络的训练: 一组是 6 个特征因素加重力区域异常梯度 (A_7)、地热流量 (A_9), 以符号 wy_1 记; 另一组是 6 个特征因素加地形高差 (A_5)、地壳厚度 (A_{10}), 以符号 wy_2 记。用 wy_1 组数据对网络训练, 迭代时间增加 1.5 倍; 用 wy_2 组数据对网络训练, 网络不收敛, 见图 4。因此, 可以认为网络设计合理, 对特征因素的选择反映灵敏。

在训练好的网络中, 将所有单元数据

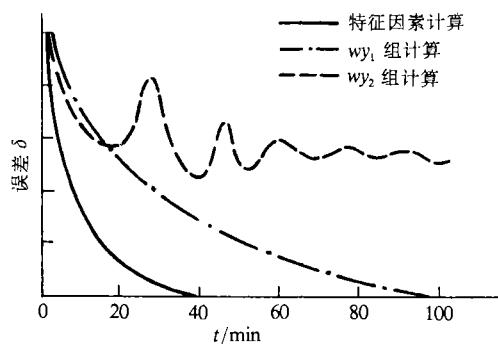


图 4 不同数据训练网络的收敛性

一一送入, 进行识别、分类。这时, 由于连接权值已确定, 无误差的反向时间消耗, 因此网络运行非常快, 仅 9 min, 整个区域识别划分完成, 结果绘于图 5。计算结果与已知点上的地震资料对比, 符合率在 90% 以上。

如原理和实际资料计算所示, 本文提出的求解方法快速、简捷、稳定, 计算效果好。理论上认为, 它可以将震源区在空间上的分布特点刻划得足够细致。实际计算时, 只需增加一些神经元即可, 这样处理的资料具很大的实用参考价值。但应该指出, 笔者在计算中也发现了这种方法

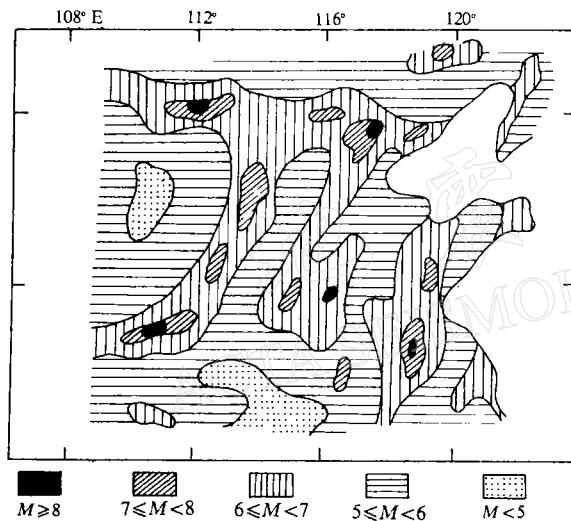


图 5 华北地区潜在震源区区划图

存在的问题。首先是单元划分, 划得越小, 对未来地震活动强度的分布特点刻划得越详细, 但单元内所能有的地震统计数据也越少。因此, 合理适度的单元选取要依研究区域的具体地震地质情况和计算者的经验, 不易统一划线。而选取单元不同, 自然是区划结果不同; 其次, 当因素多、数据量大, 参加网络训练的样本又不典型时, 网络学习时间较长, 迭代次数增加, 收敛速度减慢。因此, 网络训练的方式还有待改进。尽管如此, 该方法的确明显地优于专家系统、其它人工智能技术及常规的编程方法。特别要指出的是, 通过笔者的尝试, 可以认为, 神经网络新技术的及时引进, 将可能解决一些地震预报领域中不易解决的问题。

参 考 文 献

包约翰, 1992. 自适应模式识别与神经网络, 83—91. 科学出版社, 北京。

- 杜兴信、丁韫玉, 1990. 应用模糊推理及幕态思维进行潜在震源区定量划分的探索. 地震学报, 12, 257—264.
- 许建华, 1993. 反向传播神经网络模型及其在测井资料岩性自动识别中的应用. 石油物探, 32, 3, 53—59.
- 叶洪、于之水、张渤海、刘大健, 1988. 海南岛及广东、广西大陆沿海地区 6 级以上潜在震源区的图象识别研究, 海南岛北部地震研究文集, 98—105. 地震出版社, 北京.
- 张学工、李衍达, 1992. 用人工神经网络实现地震记录中的废道自动切除. 地球物理学报, 35, 5, 637—643.
- Lippmann, R. P., 1987. An introduction to computing with neural nets. *IEEE ASSP Magazine*, 37, 4, 4—22.
- McCormack, M. D., Zaucha, D. E. and Dushek, D. W., 1993. First-break refraction event picking and seismic data trace editing using neural networks. *Geophysics*, 58, 67—78.
- Poulton, M. M., Sternberg, B. K. and Glass, C. E., 1992. Location of subsurface targets in geophysical data using neural networks. *Geophysics*, 57, 1534—1544.

地 震 学 报
ACTA SEISMOLOGICA SINICA