

岩相识别的神经网络计算

王硕儒 范德江 汪丙柱

(青岛海洋大学, 青岛 266003)

提 要 模式识别在定量研究岩相模式中是一种有效的方法,神经网络是模式识别的一种新方法,和其它模式识别方法不同之处在于它能模拟人脑并行处理信息的模式和神经系统的可塑性。

经传统岩相分析,鄂中拗陷二叠系碳酸盐岩确认为碳酸盐岩台坪相,它包含五个亚相,即局限台洼、台坡 B 台洼、台坡 A 和台滩相。它们在岩石的颜色、矿物组分、古生物种属、结构等特征上都有不同的差异。

应用 BP神经网络,特别是与模糊模式识别的结合,对拗陷区的岩相识别是成功的,令人鼓舞,两种方法各自的正确判对率约为 75%,而综合两种方法的判对结果,则岩相全部判对。

关键词 岩相 模式识别 神经网络 BP算法 贴近度

第一作者简介 王硕儒 男 62岁 教授 数学地质和地球物理

前 言

相是沉积环境和岩性特征的综合表现,不同的岩相总是与不同的地质学特征相对应。虽然这些特征随着地理环境的改变而有所差异,但就一个地区而言,总能用一组综合的特征来体现岩相,表示为岩相的标准模式。因此,从根本上讲岩相识别是一个模式识别问题。模式可以是数值特征的,也可以是非数值特征的,后者又总能转换成数值式的。

传统的模式识别采用 Bayes 方法,它是以概率统计为基础,通过先验概率和条件概率估计后验概率来达到确切识别。然而先验概率很少可能知道,而且就某一模式而言,可能还会以不同程度属于多个标准类别,建立在模糊集合论基础上的模糊模式识别,从理论上和实践上已证明了用于岩相识别的可行性和可信性。不过,其标准模式的建立与所选方法有关,具有一定的人为因素⁽¹⁾。神经网络用于模式识别似乎能更确切地抓住和代表人类在处理模式识别问题中的方式。数值式标准模式在一定范围内是可以变动的。1986年由 Rumalhart 等提出的半线性前馈网络使神经网络的研究进入了一个新的时期。这种网络由于能构成所需要的任意复杂的判别函数(非线性函数),这对于诸如岩相这类识别,其输入与输出决非简单关系,可又不知其具体函数关系的问题而言,是特别适用的。

神经网络应用于地质勘探还是近一两年的事,McCormack (1990) 在地球物理学中用它来鉴别地震道大概是世界最早,也是颇具代表性的一种研究⁽²⁾。国外一些地质学家已把神经网络的引入看成是地质勘探的一个前沿学科。

基于上述理由,我们将神经网络与模糊数学结合,利用各自的优点想必会改进岩相识别的能力

1 BP神经网络

神经网络的原理设计思想来源于生物神经系统,是该系统组织和工作原理高度简化的计算机模型,模拟脑神经系统的计算机模型有多种网络,其中分层神经网络颇具代表性,而反向传播网络(BP)又是层状网络模型中最具代表性的,它的机理结构如下所述

神经网络由若干个神经元组成,发生在每个神经元上处理信息的过程如图 1所示。

神经元接收各输入信号 O_i 后,经加权求和 S_j ,并通过一个非线性激励函数而确定此神经元的输出 O_j ; 通常,取 $O_j = (1 + e^{-S_j})^{-1}$ 。

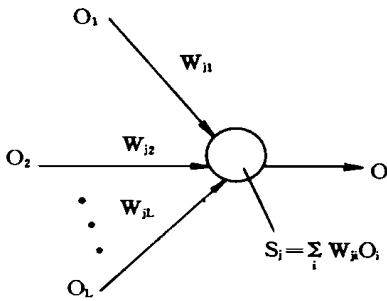
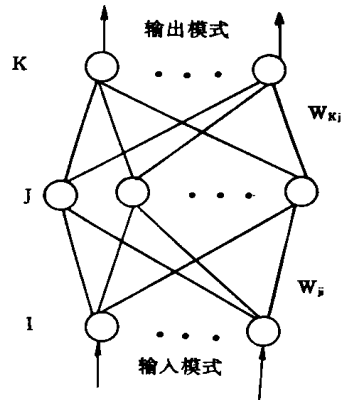


图 1 神经元处理信息过程图

Fig. 1 Circuit map of neuron processing message



K. 输出层 J. 隐层 I. 输入层

K. 输出层 J. 隐层 I. 输入层

图 2 BP神经网络结构示意图

Fig. 2 Texture of the neural network BP

图 2所示为一 m层神经网络, BP神经网络学习算法不同于一般感知机算法,不仅考虑最后一层信号的权重改变外,还考虑网络中其它各层权重的改变。该算法的学习过程是先正向传播,由输入经隐层逐层处理传向输出层。如果在输出层不能得到期望的输出,则转入反向传播,将误差沿原来的连接通路返回,通过修改各层神经元的权值,如此往返,使误差达最小

按推广的 δ 规则,每次连接权重的调整增量为^[3]:

$$\Delta W_{kj} = \eta \delta_k O_j \tag{1}$$

对于非输出层的连接权值

$$\Delta W_{ji} = \eta \delta_j O_i \tag{2}$$

其中

$$\hat{\alpha}_k = (t_k - O_k) O_k (1 - O_k) \quad (3)$$

和

$$\hat{\delta}_j = O_j (1 - O_j) \sum_k \hat{\alpha}_k \cdot W_{kj} \quad (4)$$

在实现这种网络时, 还需考虑诸如选择 η 值的一些因素的作用, 故实用算式为

$$\Delta W_{kj} (n+1) = \eta \hat{\alpha}_k O_j + \alpha \Delta W_{kj} (n) \quad (5)$$

$$\Delta W_{ji} (n+1) = \eta \hat{\delta}_j O_i + \alpha \Delta W_{ji} (n) \quad (6)$$

上式中 t_k 为系统的期望输出值; $n+1$ 表示迭代次数; α 为一比例因子; η 为学习步长; 而 $\alpha \Delta W_{kj}$ (或 $\alpha \Delta W_{ji}$) 是为避免由于采用较大的 η 导致学习速率加快而可能引起的振荡效应增加的一项阻尼项 它可使收敛过程快速 稳定。

神经网络解决问题的方法与一般固定算法不同, 它是采用一种学习技术产生一种解 用重复提供要解决问题的输入和要求的输出的例子和模式来训练神经网络。一直进行到所有训练例子的实际输出与要求的输出在某个特定的误差之内为止 这时, 神经网络的各连接权重已固定, 网络已训练好, 可以接受新的输入, 新输入的输出值就是神经网络计算的解。

2 用神经网络训练和识别碳酸盐岩岩相

取我们在鄂中坳陷松 县采集的二叠系剖面的各层样为例

鄂中坳陷位于扬子地台中部, 北与秦岭褶皱带相连, 南邻华南褶皱带, 东西分别与上、下扬子台地毗邻, 区内海相碳酸盐岩厚度大, 分布广, 岩相类别较多, 具有较广阔的油气远景

经传统的岩相分析后确认鄂中坳陷二叠系为碳酸盐岩台坪沉积, 包括五个相带, 即台洼相、台坡 A 相、台滩相、台坡 B 相及局限台洼相 (图 3)。所采的 36 个样品分属这五个相带, 体现在相标志 (颜色, 成分, 古生物特点) 上有差异。按离古海岸的远近, 顺序组成训练集 = {台洼相 (4 个样), 台坡 A 相 (5 个样), 台滩相 (2 个样), 台坡 B 相 (5 个样), 局限台洼相 (4 个样)}。从 21 个特征标志中筛选出 12 个主要标志, 组成标志集 = {颜色, 方解石, 白云石, 石英, 粘土矿物, 古生物 1-5 类, 硅质结核, 瘤状物} 其中古生物 1 为苔藓, 海棉及粪球粒, 古生物 2 为钙球, 钙藻, 生物屑; 古生物 3 为有孔虫, 海百合; 古生物 4 为腕足类, 棘皮动物 (海百合除外); 古生物 5 为腹足类, 瓣鳃类和介形虫。

颜色的赋值是先将不同颜色 (深浅) 分为不同等级, 再将等级按指数变换后取值 硅质结核及瘤状体按岩石中是否发育这类构造来赋值, 发育取值“1”, 不发育取值“0”; 瘤状体值并作了线性放大。古生物和岩石组份均以样品中百分含量予以赋值。

根据标志集和训练集选输入层神经元 12 个, 输出层 5 个神经元 输出层某一神经元的输出为零值, 表示输入的样不属该岩相; 输出为 1, 则属该岩相。这样, 规定输入样如属台坡 B 相, 要求输出为 (1, 0, 0, 0, 0); 属台洼相, 输出应为 (0, 1, 0, 0, 0); 属台坡 A 相, 输出应为 (0, 0, 1, 0, 0); 局限台洼相为 (0, 0, 0, 1, 0); 台滩相则应为 (0, 0, 0,

0, 1)。据 Lippman 的研究结果认为：“在复杂区域内，只有三层（输入层，隐层，输出层）的网络就能构成所需的任意复杂的判别函数”^[4]。据此，设计隐层为 1 层，8 个神经元，形成 12-8-5 网络结构。

网络训练时单个模式最大允许的均方误差定为 10^{-3} （这个值对于所给有效位为小数点后两位的原始计算数据而言是足够了）， $\eta = 0.9$ ， $\alpha = 0.7$ 经 562 次迭代就使连接权重调整到实际输出与期望输出不超过 5%。从图 4 可以看出，网络稳定，单调收敛。在 MV-6000 机上，完成训练约需 3 分钟。

按训练好的神经网络保存的连接权重对另 16 个样作神经网络识别计算。结果如表 1 所示。一个待识别样其输出也是一个五维向量。如 10[#] 样 (0.98, 0.00, 0.00, 0.01, 0.00)，对它的识别是显而易见的。但如对 18[#] 样 (0.36, 0.05, 0.00, 0.17, 0.00)，输出计算值都不大，可能该样具较明显的随机误差，它的确切判别可以遵循类似模糊教学中最大隶属原则，即该向量中最大的一个元素所对应的岩相即是该样的所属岩相。不过，这类样在整个待识别样中所占比例是很小的，绝大多数元素值相差很大。在这所有的待识别样中，总的判对样有 12 个，判对率 75%。

表 1 神经网络计算结果 (附贴进度计算结果)

Tbale 1 Calculation results of the neural network (with results of similarity)

样品原 编号	神经网络识别		贴进度识别	
	计算输出	确切识别	计算结果	确切识别
8	(0.99, 0.00, 0.00, 0.02, 0.00)	(1, 0, 0, 0, 0)*	(0.19, 0.14, 0.17, 0.19, 0.15)	局限台洼相
9	(0.99, 0.00, 0.00, 0.01, 0.00)	(1, 0, 0, 0, 0)*	(0.18, 0.13, 0.17, 0.17, 0.17)	台坡 B 相*
10	(0.98, 0.00, 0.00, 0.01, 0.00)	(1, 0, 0, 0, 0)*	(0.18, 0.12, 0.17, 0.17, 0.17)	台坡 B 相*
13	(0.01, 0.99, 0.00, 0.02, 0.00)	(0, 1, 0, 0, 0)*	(0.17, 0.29, 0.12, 0.16, 0.11)	台洼相*
18	(0.36, 0.05, 0.00, 0.17, 0.00)	(1, 0, 0, 0, 0)	(0.17, 0.23, 0.14, 0.17, 0.11)	台洼相*
20	(0.00, 0.00, 0.01, 0.00, 0.99)	(0, 0, 0, 0, 1)*	(0.18, 0.11, 0.19, 0.18, 0.18)	台坡 A 相
23	(0.01, 0.01, 0.01, 0.99, 0.00)	(0, 0, 0, 1, 0)*	(0.19, 0.12, 0.17, 0.21, 0.16)	局限台洼相*
26	(0.00, 0.01, 0.00, 0.99, 0.00)	(0, 0, 0, 1, 0)*	(0.17, 0.11, 0.19, 0.20, 0.17)	局限台洼相*
27	(0.00, 0.01, 0.00, 0.99, 0.00)	(0, 0, 0, 1, 0)*	(0.17, 0.08, 0.19, 0.18, 0.21)	台坡 A 相
30	(0.00, 0.00, 0.15, 0.19, 0.09)	(0, 0, 0, 1, 0)*	(0.16, 0.10, 0.18, 0.17, 0.20)	台滩相
33	(0.00, 0.01, 0.00, 0.99, 0.00)	(0, 0, 0, 1, 0)*	(0.18, 0.07, 0.21, 0.18, 0.21)	台滩相
35	(0.00, 0.01, 0.28, 0.82, 0.01)	(0, 0, 0, 1, 0)*	(0.18, 0.08, 0.19, 0.18, 0.21)	台滩相*
39	(0.00, 0.00, 0.00, 0.06, 0.96)	(0, 0, 0, 0, 1)*	(0.16, 0.08, 0.22, 0.16, 0.20)	台坡 A 相
43	(0.23, 0.12, 0.06, 0.00, 0.02)	(1, 0, 0, 0, 0)	(0.19, 0.14, 0.16, 0.22, 0.14)	局限台洼相*
45	(0.00, 0.03, 0.00, 0.99, 0.00)	(0, 0, 0, 0, 1)*	(0.19, 0.10, 0.20, 0.21, 0.16)	局限台洼相*
46	(0.00, 0.00, 0.00, 0.00, 0.99)	(0, 0, 0, 0, 1)	(0.19, 0.12, 0.19, 0.21, 0.17)	局限台洼相*

注：* 为判对样

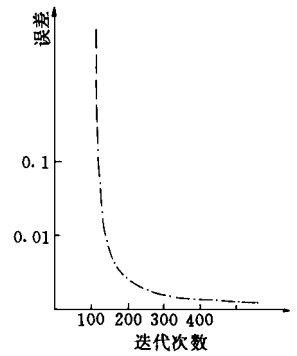


图 4 网络收敛过程图

Fig. 4 Convergence procedure of the neural network

4 与模糊模式识别法的比较

模糊模式识别可以进行无监督学习, 如用 F4 sodata 聚类, 也可实现有监督的学习识别。应用聚类作地质结论须经分析, 还可能需作适当调整, 然而, 使用有监督学习的贴近度计算最简便。

利用已知岩相样的平均值向量建立标准模式, 按如下定义的一种“加权逆距离贴近度”^[5]

$$S(X, Y) = \sum_{k=1}^S W_k (1 - d_k)^P \quad (7)$$

识别, 式中 $d_k = |x^k - y^k|$, x^k 为待识别样 X 第 k 个标志的归一化值, y^k 为标准岩相 Y 第 k 个标志的归一化的平均值, P 为调整未知样从属各岩相的对比度给定值 ($P > 0$), 权重 W_k 可按单标志判别时的判对率确定。

用 (7) 式计算的 16 个样的结果也列于表 1 以便与神经网络计算结果对比, 根据贴近原则, 确切判对 10—11 个样

从表 1 还可以看出, 神经网络与模糊识别判对率大致相当, 前者稍高于后者, 但是, 有意思的是两种判别结果的并集所含样、相对比, 全部判对, 无一漏判, 这决不是一种巧合, 充分说明了岩相识别采用综合方法的必要性

5 结 论

我们的成果表明, 三层神经网络可以解决象碳酸盐岩岩相识别这类比较复杂的问题。隐层的神经元个数不宜过多, 也不宜太少。太少, 网络的信息存储能力低, 难于学到模式; 太多, 由于总会有一些神经元输出所带的权重值很小, 表明这些神经元在决定输出值时所作的贡献很少, 它们的存在是多余的, 反而会传递输入信号所携带的噪声。只要输入或输出层的神经元较多时, 隐层的神经元数可取输入、输出层神经元数的平均值

本文所用的方法当然也适用于比碳酸盐岩岩相较为简单的碎屑岩岩相的识别。我们在对单井或多井测井资料的砂泥岩剖面的岩性识别试验中, 也取得了很好的效果

神经网络用于学习识别的一个很大优点是, 整个过程不需过问, 只需在初始赋值计算时将原始数据归一化处理, 这较之专家系统的识别更具客观性

但是, 神经网络用于识别并非万能, 在 BP 算法上就存在局部极少性问题。虽然, 改进 BP 算法可以采用模拟退火的方法, 但局部极值问题与学习步长的确定是不可分割的。目前尚难将两者直接联系起来找到一种解决方法。因此, 综合其它模式识别方法是很必要的。利用各自的特点, 必将会综合提高模式识别的能力, 本文的实际应用正好说明了这点

参 考 文 献

- [1] 王硕儒, 刘仲衡, 范德江. 海相碳酸盐岩岩相的模糊模式识别. 沉积学报, 1992, 10 (4).
- [2] McCormack M C. Seismic trace editing and First break Picking using neural networks. Extended Abstracts, 60th SEG Annual Meeting, 1990, 321-324.
- [3] 包约翰 (美). 自适应模式识别与神经网络. 科学出版社, 1992.
- [4] Lippman R P. IEEE Assp Magazine. 1987, 4: 4-22.
- [5] 王硕儒等. 模糊模式识别在成矿预测中的应用. 地质与勘探, 1990, 26 (4).

Facies Recognition Using the Neural Networks

Wang Shouru Fan Dejiang and Wang Bingzhu

(Ocean University of Qingdao 26603)

Abstract

Pattern recognition methods are a powerful means in studying facies quantitatively. The neural network is a new method and has many improvements such as parallel processing and plasticity imitating the human brain compared with other pattern recognition methods.

Permian carbonate rocks in the central Hubei basin have been identified as carbonate platform facies after traditional facies analysis. It includes five subfacies, i. e. sub-facies of districted depression, slope B, depression, slope A and shallow out. They are different from each other in rock color, mineral components, paleobiology components, rock structures and so on.

The application of pattern recognition on the basis of the BP neural network for Permian carbonate rock facies studies, particularly comparison with fuzzy pattern recognition is very successful and inspiring. The correct identification ratios of fuzzy sets and the neural network are both about 75%. And the correct identification ratio of the combination of fuzzy sets and neural networks is 100%.

Key words facies pattern recognition neural network BP algorithm similarity