

基于构造性神经网络的石油测井解释方法*

纪福全,程国建,刘淑英

(西安石油大学 计算机学院,陕西 西安 710065)

摘要: 为了克服传统的 BP 神经网络算法收敛速度慢、隐层数以及隐层节点数难以确定等缺点,研究了一种构造性神经网络学习算法——级联算法 CCA 及其在石油工程中的应用。采用该算法进行了储层参数预测的研究,通过与 BP 神经网络的研究结果进行比较,进一步体现出构造性神经网络的优越性。

关键词: 构造性神经网络 BP 算法 级联算法 石油测井 储层参数

石油测井解释的重要目的就是利用测井获得的信息来确定储层岩性、储层流体性质、储层参数(孔隙度、渗透率和饱和度)^[1-2],以便计算地层的含油储量,以确定合适的开采规模等。因此,测井是石油工业的第一个重要环节。然而,测井数据量非常大,主要包括声、电、核等各种属性的测井信息,且测井数据往往存在模糊性和噪声影响,常用的解释方法明显存在不足。

神经网络作为一种非线性数学方法,已被广泛应用于石油地质研究领域,如利用神经网络的模式识别功能对岩性、岩相、储层敏感性及储层流体性质进行识别^[3-6]和利用其参数预测功能进行测井储层参数计算^[7-10]等。神经网络在测井储层参数计算中比常规的线性回归法精度高得多,但是由于传统的 BP 神经网络算法是一个梯度下降优化过程,所以存在自身的限制与不足,例如,它需要的训练时间较长、存在局部极小值、隐元数目难以确定等。本文介绍的构造性神经网络学习方法——级联算法^[11]CCA(Cascade Correlation Algorithm),它为了解决 BP 算法中遇到的收敛速度及隐元数目难于确定等问题提供了行之有效的办法。该算法从一个很小的网络开始,通过增加和训练神经元,使得网络动态地增长,直到得到满意的结果。

1 石油测井解释原理

测井属性从不同角度反映了储层的物体特征,根据测井理论与实践可知,它与储层岩性、储层流体性质、储层参数之间存在如下关系:

$$y_i = f_i(x_1, x_2, x_3, \dots) \quad (1)$$

式中, f_i 表示第 i ($i=1, 2, 3, \dots$) 个非线性映射关系; y_i 表示声波时差(AC)、自然伽玛(GR)、补偿中子(CNL)、密度(DEN)、自然电位(SP)、井径(CAL)、电阻率(RT)、深侧向电阻率(RD)、浅侧向电阻率(RS)、微电极电阻率

(RMLL)、自然伽玛能谱(NGS)、感应(IL)等测井信息; x_1, x_2, x_3, \dots 分别表示岩性、流体性质、储层参数等。

因为 y_i ($i=1, 2, 3, \dots$) 是已知的测井信息,若期望能从式(1)中求出岩性、流体性质、储层参数等,由多元函数理论有:

$$x_j = g_j(y_1, y_2, y_3, \dots) \quad (2)$$

从形式上看,可以由式(2)求出岩性、流体性质和储层参数等,但式(2)关系复杂,难以用解释式表达。当地下地质情况比较简单且了解比较透彻时,可凭经验选择几种测井属性,再采用数理统计、模糊数学、专家系统及神经网络等方法自动进行岩性识别或区分流体性质等。

虽然较完善的测井系列一般都提供十种以上类型的测井信息,但对于不同的识别或映射目的,有用的测井信息不尽相同。这要看这些测井信息对于所识别或判断的地层特性能否有所反映而定,只有当各测井属性间相互独立时才能获得最优效果。因此,必须考虑到测井属性之间存在相关性,必须进行信息简化。例如,对于岩性分析,一般选用 GR、AC、DEN 和 CNL 四种测井信息即可获得良好的效果;对于孔隙度分析,一般选用 AC、DEN、CNL 等测井信息;对于饱和度分析,一般选用 RD、DEN、CNL、GR 等测井信息;对于渗透率分析,一般选用 RT、GR、CNL、AC 等测井信息。可见,信息简化在测井解释中是必不可少的。

2 CC 神经网络基本原理及算法

CC 神经网络是 20 世纪 80 年代末 90 年代初提出的一种神经网络^[11],开始只是一个具有最小拓扑结构的网络,仅由输入输出层构成。在最小化网络误差的过程中,级联算法一步步地添加新元到隐层当中。级联属于监督类学习结构,经级联算法构建的网络是具有最小拓扑的多层神经网络,它有两个明显的优点:用户不必担心网络拓扑结构的设计和级联算法的学习速度比常规算法要

* 国家自然科学基金资助项目(编号:40572082)

快。以下简要介绍 CC 算法的数学背景及其学习工程。

2.1 CC 算法的数学背景

本文的目标是训练输出权重使误差平方和最小, 即:

$$\min: \varepsilon = \sum_p \frac{1}{2} \sum_0 (y_{p0} - t_{p0})^2 \quad (3)$$

式中, t_{p0} 、 y_{p0} 是对应于模式 P 的在输出元 O 上的期望值与观察值。

通过梯度下降法可使 ε 最小化:

$$\frac{\partial \varepsilon}{\partial \omega_{io}} = \sum_p e_{p0} I_{ip} \quad (4)$$

(4)式中, I_{ip} 是输入元或隐元 i 的值(对第 P 个模式而言), ω_{io} 代表输入或隐元 i 到一个输出元 O 的连接权值。

$$e_{p0} = (y_{p0} - t_{p0}) f_p'(net_0) \quad (5)$$

(5)式中, f_p' 是输出元 O 激励函数的一阶导数, $net_0 = \sum_j \omega_{0j} x_j - \theta_0$

式中, ω_{0j} 代表输入或隐元 O 到输出元 x_j 的连接权值, θ_0 是阈值)。

训练阶段结束后再调整候选元, 以使候选元输出值与一个输出元残留误差 e_{p0} 之间的关联度 C 成为最大。Fahlman^[11]给出的关联度为:

$$c = \sum_0 \left| \frac{\sum_p (y_{p0} - \bar{y}_0)(e_{p0} - \bar{e}_0)}{\sum_p (y_{p0} - \bar{y}_0)^2 + \sum_p (e_{p0} - \bar{e}_0)^2} \right| = \frac{\sum_0 \left| \sum_p y_{p0} (e_{p0} - \bar{e}_0) \right|}{\sqrt{\sum_0 \left(\sum_p y_{p0}^2 + \sum_p (e_{p0} - \bar{e}_0)^2 \right)}} \quad (6)$$

式中, \bar{y}_0 是某个候选元的平均激励, \bar{e}_0 是某个输出元的平均误差(对全体模式的平均 \bar{P})。采用梯度下降法进行关联度的最大化, 即:

$$\frac{\partial C}{\partial W_i} = \sum_p \delta_p I_{pi} \quad (7)$$

式中, $\delta_p = \sum_0 \sigma_0 (e_{p0} - \bar{e}_0) f_p'$ (8)

式中, δ_p 表示候选元输出与输出端残留误差关联度的符号。

2.2 CC 算法的学习过程描述

CC 算法有两层含义: 一是级联式结构(每次仅添加一个新元, 并且在它们全面添加后不再发生变化)。二是学习算法 (即创建并添加新元于网络中)。对每个新隐元, CC 算法试图使新元输出信号与网络残留误差信号之间关联的幅度最大化。

CC 算法可描述如下:

(1)用仅有一个输入层、一个输出层(输入输出层全连接)的最小网络初始化。

(2)用一个常规学习算法(如 BP)训练连接权值, 直到网络误差不再减小。

(3)产生所谓的候选元(Candidate Units)。每个候选元均与全体输入元连接, 并与业已存在的隐层

元连接, 但候选元与输出元之间没有权值连接。

(4)通过训练连入的候选元的权值使候选元激励与网络残留误差之间的关联度最大化。学习算法采用常规方法。关联度不再改善时停止训练。

(5)选择具有最大关联度的候选元, 冻结其输入权值并将其连入网络中。此时, 候选元变成一个隐元, 并且将其与全体输入元相连。因为新隐元的引入权冻结, 一个新的永久性特征检测子即由此获得。

(6)返回(2), 重复算法直到网络的全体误差小于某个给定值时, 则停止训练。

基于 CC 算法的网络拓扑的生长过程如图 1 所示。

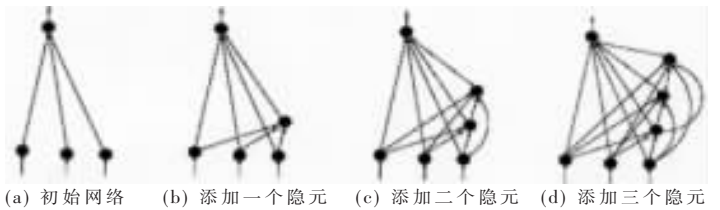


图 1 级联网络拓扑的生长过程

3 储层参数定量计算

选取某油田台参 2 井的测井资料作为研究对象, 主要根据测井信息来定量计算孔隙度(有效孔隙度 EPOR、有效含水孔隙度 EPRW)和饱和度(有效含水饱和度 ESW、含水饱和度 SW)的大小。将深度为 2740~2770 米的 150 个具有代表性的样本(包括九种类型的测井资料)进行预处理, 优选出反映孔隙度和饱和度变化的 AC、GR、CNL、DEN 和 RD 等五种测井属性信息。表 1 列出了本实验中作为输入模式的测井数据和目标输出的最大与最小值。

采用 CC 神经网络, 输入层有五个神经元分别代表五种测井属性; 输出层有四个神经元, 分别代表四种储层参数: EPOR、EPRW、ESW 和 SW。级联网络模型初始结构与训练结构如图 2 所示。

训练初始参数设置为: 学习步长 $\eta_1=0.1, \eta_2=1$; 动量参数 $\mu_1=0.1, \mu_2=0.1$; 单调点消去值 $c=0.1$; 极限误差为 0.001。进行网络学习训练过程是: 当网络的迭代满足精度要求时, 记录此时的权值, 作为知识保存; 将该井中 2785~2805 米处的 A 层、2844~2853 米处的 B 层、2868~2870 米处的 C 层、2905~2921 米处的 D 层的测井数据分别输入训练好的神经网络进行定量计算, 最后输出计算结果。与岩芯分析数据比较, 其均方差结果如表 2 所示。由表 2 可以看出, 网络 A、B、C 层解释精度很高, D 层的解释精度稍微低些, 但均优于 BP 以及其他的传统方法, 其

表 1 用于归一化处理的数据

属性	AC($\mu\text{s}/\text{m}$)	GR(API)	CNL(%)	DEN(g/cm^3)	RD($\Omega \cdot \text{m}$)	EPOR(%)	EPRW(%)	ESW(%)	SW(%)
Min	170.0	50.0	0.00	2.40	2.00	0.00	0.00	0.00	0.00
Max	240.0	100.0	20.0	2.80	2000.0	20.0	20.0	100.0	100.0

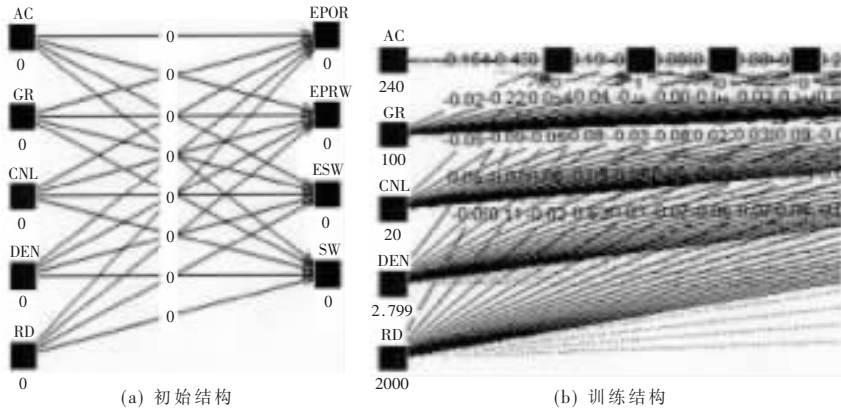


图2 级联网络模型初始结构与训练结构

表2 均方差计算结果对比

层号	深度(m)		EPOR(%)	EPRW(%)	ESW(%)	SW(%)
A	2785~2805	CC	3.34	6.67	6.48	3.43
		BP	5.64	8.72	6.32	4.56
B	2844~2853	CC	9.23	13.21	15.09	9.73
		BP	17.12	10.21	17.11	10.68
C	2863~2870	CC	9.75	5.58	3.32	5.89
		BP	12.48	6.84	2.15	5.73
D	2905~2921	CC	41.05	20.32	12.08	17.81
		BP	52.88	21.44	15.11	16.41

解释效果令人满意。

本文采用CC算法克服了BP算法收敛缓慢、隐层节点数难以确定以及局部极小值等问题,加快了神经网络的收敛速度。采用CC神经网络,对于一个构造区域,只要掌握了关键井的数据,就可以对本地区测井进行储层参数预测,方法既简单且易操作,准确率比较高。这为神经网络在测井资料地质解释中提供了一种新的方法,提高了探寻和鉴别含油气地层的精确性,在油气资源开发

领域具有实用价值。

参考文献

- [1] 《测井学》编写组. 测井学[M]. 北京:石油工业出版社,1998.
- [2] 雍世和,张超谟. 测井数据处理与综合解释[M]. 东营:中国石油大学出版社,1996.
- [3] 肖慈,杨斌,马维炎. 利用测井录井信息识别水淹层[J]. 测井技术,1998,22(4):267-272.
- [4] 孙建孟,谭未一,李召成. 应用测井和BP神经网络算法预测储层敏感性[J]. 石油钻探技术,2001,29(2):37-40.
- [5] 席道瑛,张涛. BP神经网络模型在测井资料岩性自动识别中的应用[J]. 物探化探计算技术,1995,17(1):42-48.
- [6] 阳文生,赵力民,侯守探等. 精细储层描述在荆丘油田调整挖潜中的初步实践[J]. 石油实验地质,2000,22(4):375-381.
- [7] 刘争平,何永富. 神经网络在测井解释中的应用[J]. 地球物理学报,1995,35(增刊1):323-330.
- [8] 陶淑娴,肖慈,杨斌等. 神经网络在测井解释中的应用[J]. 石油物探,1995,34(3):90-102.
- [9] 夏宏泉,张贤辉,范翔宇等. 基于神经网络法的逐点渗透率测井解释研究[J]. 西南石油学院学报,2001,23(1):11-13.
- [10] 金燕,张旭. 测井裂缝参数估算与储层裂缝评价方法研究[J]. 天然气工业,2002,22(增刊):64-67.
- [11] FAHLMAN S. Fast learning variations on back propagation: An eEmpirical study[M]. Proceedings of the Connectionist Models Summer School,1988.

(收稿日期:2006-10-10)