

基于神经模糊系统的储层参数反演

杨 斌, 肖慈珣, 王 斌, 彭真明

(成都理工学院信息工程与地球物理系, 四川成都 610059)

摘要: 神经模糊系统, 即把神经网络与模糊逻辑结合起来, 用神经网络来构造模糊系统, 使建立的储层参数反演模型既能处理输入信息, 又能嵌入专家的模糊性知识, 提高了模型的抗干扰能力和预测精度, 同时, 也克服了人工神经网络技术的储层参数反演与预测在实际应用中暴露出的一些难点。通过基于神经模糊系统的网络结构、建模步骤及应用的研究, 与常规多元线性回归分析、模糊系统建模相比较, 获得了较高的储层参数反演预测精度。

关键词: 储层参数; 神经网络; 模糊逻辑; 反演预测; 应用

第一作者简介: 杨斌, 男, 32 岁, 副教授(博士), 石油地质

中图分类号: P618.130 文献标识码: A

80 年代末以来, 对于储层参数的定量反演计算, 应用神经网络技术(BPNN)已取得了大量成果。但是, 随着研究的逐步深入, 人们发现, 基于神经网络技术的储层参数定量计算在实际应用中存在如下难点。(1) 已建立的 BPNN 反演模型对输入信息的抗干扰能力差, 对所选学习训练集中的学习样本可靠性要求高, 且要求输入单元值准确。而实际中, 由于操作、经验、井眼环境、数据预处理等诸多复杂因素的影响, 输入到预测模型中的单元值均带有噪声, 使预测结果出现偏差。(2) 仅能处理数值输入, 无法对专家知识等的语言信息进行处理。而在挑选学习样本时专家经验起着重要作用, 它直接决定着 BP 建立的模型的外推预测能力。

由此, 我们设想, 应当使建立的反演模型既能对数值信息进行处理, 又能嵌入专家的模糊性知识, 来提高模型的抗干扰能力和预测精度。由神经网络与模糊逻辑结合形成的神经模糊系统就能适应这样的要求。模糊逻辑接近于人的形象思维方式, 擅长于定性分析和推理, 处理不确定的信息, 可充分利用专家经验知识, 并将知识存在规则集中, 但模糊系统的规则集和隶属度函数等设计参数只能靠经验来选择, 很难自动设计和调整, 这是模糊系统的主要缺点, 神经网络是根据人脑的生理结构和信息处理过程来设计的, 具有很好的自组织、自学习的能力, 但其内部是典型的“黑箱”学习模式,

令人无法理解学习过程中的数据意义, 也无法用人们常用的方式来解释其输入输出关系。因此, 将二者有机地结合, 相互取长补短, 形成的神经模糊系统, 既能定性地表示领域专家的经验知识, 又具有自学习和处理定量数据的能力, 集学习、联想、识别、自适应及模糊信息处理于一体。

神经模糊系统的实现方式有许多种, 这里我们采用基于神经网络的模糊推理系统 NNFR^[1] (Neural Networks driven Fuzzy Reasoning) 来实现, 即用神经网络来构造模糊系统, 利用神经网络的学习方法, 根据输入输出样本来自动设计和调整模糊系统的设计参数, 实现模糊系统的自学习和自适应功能, 这就是 NNFR 基本思路。

1 NNFR 网络结构

NNFR 网络结构如图 1 所示, 它在结构上像多层前馈网络, 但在功能上是模糊系统。在图 1 所示的 NNFR 系统中, 共有 $R+1$ 个神经网络, 其中 $NN_1 \sim NN_R$ 分别用来表示 R 条规则的结论部分中的函数 $f_j(x)$, 而 NN_{nf} 则是给出每条规则对于输入 X 的适应度, 则 NNFR 的输出由式(1)决定, 其中 g_j 为第 j 个网络 NN_j 的输出, μ_j 为第 j 个网络 NN_j 的适应度。

$$y = \sum_{j=1}^R \mu_j g_j \quad (1)$$

NNFR 是利用神经网络来表示多维非线性隶属函数, 其优点是能直接产生多维非线性隶属函数。

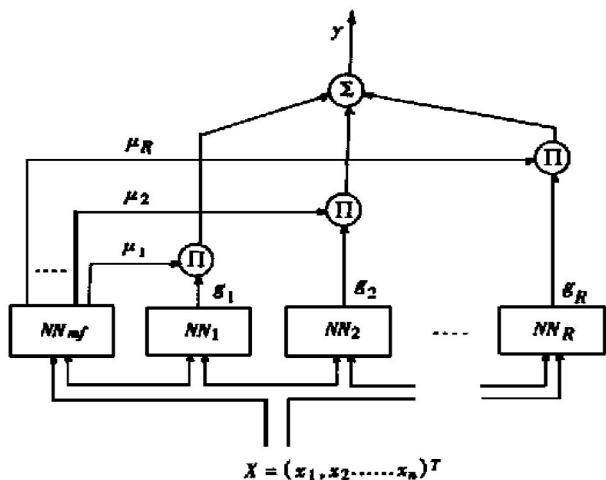


图1 基于神经网络的模糊推理网络结构

Fig.1 Texture of fuzzy reasoning networks based on neural networks

2 NNFR 设计及实现步骤

NNFR 的三个设计步骤是: (1) 决定模糊规则的个数 m ; (2) 获取计算规则适应度的神经网络 NN_{mf} ; (3) 获取生成结论部分非线性函数的 m 个神经网络 $NN_j (j=1, 2, \dots, m)$ 。

在专家知识无法用语言表达时, 采用无导师的规则聚类算法从经验数据中获取知识是十分必要的。这里我们采用 (FCM, Fuzzy C-Means clustering) 模糊 C -均值聚类算法^[2] 来获取模糊规则的个数。神经网络均采用我们新改进的多层前馈网络结构优化和学习算法。

第一步: 选择网络训练样本数据, 每个样本对有如下关系 $(x_1, x_2, \dots, x_m) \rightarrow Y$ 。

第二步: 将输入输出数据对 (x_i, y_i) 分成两部分, 一部分作为网络的训练数据, 共有 n_t 个, 另一部分作为检验已训练好的网络模型的评估数据, 共有 n_c 个, 则 $n = n_t + n_c$ 。

第三步: 用 FCM 聚类方法将训练数据分成 r 类, 用 $R^s (s=1, 2, \dots, r)$ 表示。 R^s 中的训练数据用 (X_i^s, Y_i^s) 表示, $i=1, 2, \dots, (n_t)^s$, $(n_t)^s$ 表示 R^s 中的训练数据的个数, 在这一步, 把 m 维空间分成了 r 部分意味着模糊规则数为 r 。

第四步: 决定 NN_{mf} 网络, 就是决定各规则的“如果”部分, 使每个“如果”部分获得一个适应度 (隶属度)。如果 X_i 是 NN_{mf} 的输入层上的一个矢量, 网络输出层上的期望输出为

$$W_i^s = \begin{cases} 1 & x_i \in R^s \\ 0 & x_i \notin R^s \end{cases} \quad i=1, 2, \dots, (n_t)^s, s=1, 2, \dots, r \quad (2)$$

这样 NN_{mf} 的输入层节点为 m 个, 输出层上节点数为 r 个, 中间为隐含层。 NN_{mf} 经过学习之后, 就能对任意给定的输入 x_i , 输出 x_i 对 R^s 的适应度 $\bar{W}_i^s$ 。这个适应度就是 x_i 属于各个规则的“如果”部分的隶属度

$$\mu_A^s(x_i) = \bar{W}_i^s \quad i=1, 2, \dots, n \quad (3)$$

第五步: 完成每条规则的后半部分, 即“那么”部分的网络构造。

第 s 条规则对应的网络为 NN_s , 将训练数据 $X_{i1}^s, X_{i2}^s, \dots, X_{im}^s$ 作为它的输入值, $y_i^s (i=1, 2, \dots, (n_t)^s)$ 作为它的输出值, 然后训练神经网络 NN_s 。 NN_s 训练完之后, 对任意的输入矢量 x_i , 由 NN_s 就可得到一个控制值, 这相当于是第 s 条规则对输入 x_i 作出的反应。为防止网络的过学习, 引入下面的准则函数

$$I^s = \frac{n_c}{(n_t)^s + n_c} \sum_{i=1}^{(n_t)^s} [y_i - g_s(x_i)]^2 + \frac{(n_t)^s}{(n_t)^s + n_c} \sum_{j=1}^{n_c} [y_j - g_s(x_j)] \mu_A^s(x_j)^2 \quad (4)$$

当 I^s 的值在网络 NN_s 的训练中取得最小值时, 停止训练。当 NN_s 构造完后, 用检验数据 $(x_{i1}, \dots, x_{im}) (i=1, 2, \dots, n_c)$ 作为这些网络的输入, 由式 (5) 计算出该规则网络的误差量, μ_A 是由 NN_s 网络的输出得到的值。

$$\Theta_m = \sum_{i=1}^{n_c} [y_i - g_s(x_i)] \mu_A^s(x_i)^2 \quad (5)$$

第六步: 简化所构成的 NN_s 网络, 消去那些无关的输入变量。

将输入变量任意消除一个 X^p , 还剩 $m-1$ 个输入变量, 这样尽管输入变量少了一维, 但输入输出对应关系没变。用这些数据重复第五步, 得到模糊规则的“那么”部分的神经网络, 并由式 (5) 计算出在消除 X^p 的情况下, 第 s 条规则的控制量的误差平方和 Θ_{m-1}^{sp} 。比较 Θ_m 和 Θ_{m-1}^{sp} , 若

$$\Theta_m > \Theta_{m-1}^{sp} \quad (6)$$

则说明输入变量 X^p 在所有输入变量中对于第 s 条规则所起作用很小, 可以忽略 X^p 分量。

第七步: 回到第五步, 重新构造 $NN_s (s=1, 2, \dots, r)$, 并计算 Θ_{m-1}^{sp} , 并循环执行步骤五和步骤六, 直到式 (6) 对剩下的任意变量都不成立为止。

第八步: 由式 (1) 导出最终的预测值或控制量 y_i 。

3 应用实例

数据取自 Bloch^[3]文中的中国南海崖埕气田的中新统-渐新统砂岩测量值, 共有 43 个砂岩样本, 5 个测量值作为孔隙度和渗透率的控制因素。它们分别是: (1) 刚性颗粒百分含量; (2) 塑性颗粒百分含量; (3) 碎屑骨架百分含量; (4) 粒径; (5) 分选系数。Bloch 用多元线性回归(MLR) 技术得到了在该区预测孔隙度和渗透率的线性回归方程。最近, Fang^[4]等人应用 Bloch 给出的数据建立了预测该区孔隙度和渗透率的模糊模型, 并与 Bloch 结果进行了对比, 其方法原理和预测精度上都优于 MLR。但是, FM 预测法仍未解决模糊系统靠经验来设计推理用的模糊规则问题, 且在确定变量隶属函数时使用分段线性拟合方式, 具有主观随意性, 所建立的 FM 模型不具有学习能力, 难以适应环境条件变化和实际大量应用。另一方面, 用目前流行的 BPNN 法对该数据集进行学习训练, 迭代 50 000 步以上, 但网络仍无法收敛, 学习误差较大, 难以建立满意的模型。

我们用 NNFR 进行了建模和预测。首先从该数据集中随机选取 30 个数据做为训练样本, 剩下的样本做为检验样本, 用 FCM 聚类法对学习样本聚类, 结果对孔隙度和渗透率参数各自都划分为 7 组。根据分组后的样本训练 NN_{mf} , 此网络结构为 5-7-1, 当用样本反复训练 NN_{mf} 38 次后, 误差平方和为 0.018 4。用前述步骤 6, 7 试消除那些无关的输入量, 结果发现对于每条规则对应的任一输入变量均难以去掉, 因此在各子网络中输入层单元数均为 5 个。完成了 NNFR 模糊系统模型的建立后, 对该数据集 43 个样本用前述步骤 8 预测得到 NNFR 模糊推理系统的孔隙度和渗透率数值。经与 FM 法 MLR 法预测结果对比(表 1), 结合孔隙度、渗透率预测值与实测值交会图(图 2, 3)可以看出, NNFR 的孔隙度预测与实测结果间相关系数 R 为 0.965 9, 较 MLR 法(R 为 0.75) 和 FM 法(R 为 0.90) 的相关系数都高, 效果最好。NNFR 的渗透率预测结果相关系数(R 为 0.951 6)较 MLR 法(R 为 0.86) 高, 与 FM 法(R 为 0.96) 预测结果相近, 但 NNFR 法人为因素少, 建模的效率和自动化程度比模糊模型法要高得多, 它还能克服 BPNN 对于复杂条件和数据难以建模的困难, 更有利于软件实现和实际推广应用。

表 1 用三种方法预测的样本孔隙度及渗透率值与实测值统计对比表

Table. 1 Comparison between predicted porosity and permeability obtained by three methods and measured porosity and permeability of samples

43 个样本 误差统计	孔隙度/ %			渗透率/ $10^{-3}\mu\text{m}^2$		
	MLR 预测值	FM 预测值	NNFR 预测值	MLR 预测值	FM 预测值	NNFR 预测值
平均绝对 误差	1.109 3	1.188 4	0.390 7	153.404 7	59.434 9	70.679 1
平均相对 误差/ %	8.25	10.13	2.96	80.313 7	32.20	32.41
相关系数	0.75	0.90	0.965 9	0.86	0.96	0.951 6

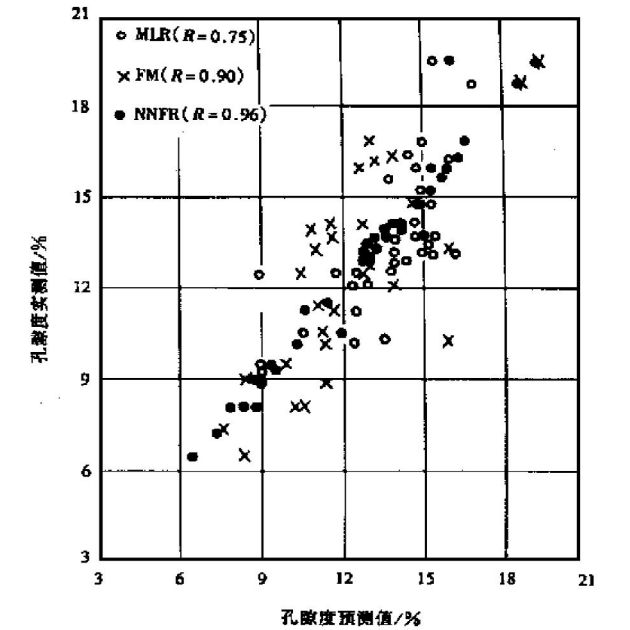


图 2 样本孔隙度预测值与实测值交会图

Fig. 2 Cross plot of predicted and measured values for sample porosity

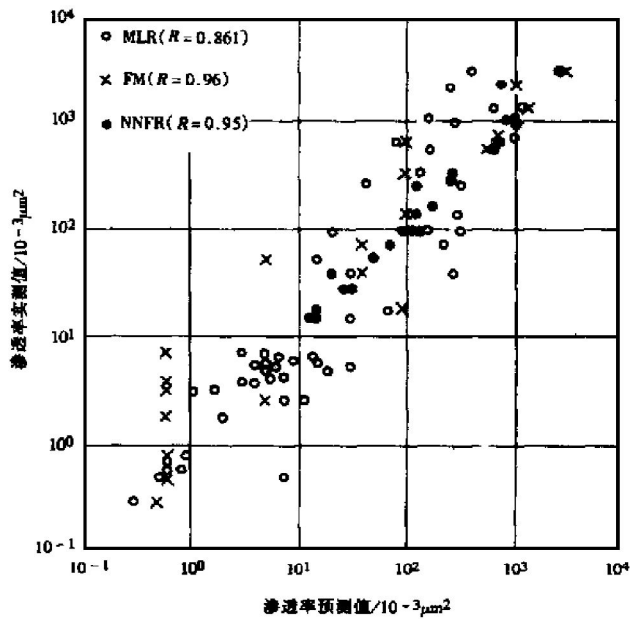


图 3 样本渗透率预测值与实测值交会图

Fig. 3 Cross plot of sample permeability predicted and measured

参 考 文 献

- 1 Takagi H. NN-driven fuzzy reasoning[J]. *Journal of Approximate Reasoning*, 1991, 5(1): 191~ 213
- 2 Bezdek J C. Pattern recognition with fuzzy objective function algorithm [M]. New York: Plenum, 1981. 191~ 203
- 3 Bloch S. Empirical prediction of porosity and permeability in sandstones [J]. *AAPG Bull*, 1991, 75(10): 1 145~ 1 160
- 4 Fang J H, Chen H C. Fuzzy modelling and the prediction of porosity and permeability from the compositional and textural attributes of sandstone [J]. *Journal of Petroleum Geology*, 1997, 20(2): 185~ 204

RESERVOIR PARAMETER INVERSION USING NEURAL FUZZY SYSTEM

Yang Bin Xiao Cixun Wang Bin Peng Zhenming

(Department of Information Engineering and Geophysics, Chengdu College of Technology, Chengdu, Sichuan)

Abstract: Although a lot of developments have been gained on the reservoir parameters inversion with the artificial neural Networks(ANN), but the problem of the finite training data has been exist for a long time, and ANN is difficult to handle fuzzy expert knowledge. In order to make full use of expert knowledge, a nonlinear intelligent inversion approach based on neural networks driven fuzzy reasoning(NNFR) system is presented in this paper to achieve the ability of self learning, self adaption and fuzzy information input. With some practical application of proposed methods, much better results of reservoir parameter inversion from analysis data are obtained.

Key Words: reservoir parameter; neural networks; fuzzy logic; inversion prediction; application

第 15 届国际 HKT 学术会在成都召开

第 15 届国际喜马拉雅-喀喇昆仑-西藏学术讨论会(Himalaya-Karakoram-Tibet Workshop, 简称 HKT 会议)于 2000 年 4 月 22~ 24 日在成都西藏饭店国际会议厅举行。来自澳大利亚、加拿大、中国、法国、希腊、德国、印度、意大利、日本、尼泊尔、新西兰、巴基斯坦、俄罗斯、瑞士、英国、美国等 16 个国家及中国香港、台湾等 2 个地区的 205 名学者出席了这次国际盛会。中国青藏高原研究会理事长孙鸿烈院士主持了会议开幕式,四川省副省长李进和科技部副司长邵立勤作了重要讲话。国土资源部副部长寿嘉华出席了闭幕式,并发表重要讲话。

这次大会的主题是新世纪的青藏高原研究,包括喜马拉雅地质、喀喇昆仑及帕米尔地质、缝合带地质、青藏高原腹地地质、青藏高原北缘地质、青藏高原东缘地质、青藏高原周边盆地、青藏高原与全球气候、青藏高原与资源、青藏高原的环境科学等 10 个专题。刘东生、肖序常、P. Tapponnier、Michael P. Searle、Mark T. Harrison 等 11 名著名学者分别作了“亚洲季风与青藏高原”、“青藏高原西部的构造演化”、“青藏高原隆升观念的演化”、“喜马拉雅的形成方式”等内容丰富的学术报告。80 余名学者宣读了论文,40 余名学者用展版形式交流了自己的最新研究成果。

这次会议将对我国实施西部大开发战略,合理开发青藏高原的资源,保护生态环境,促进经济社会的可持续发展,开展国际科技合作,起到积极推动作用。

经与会各国代表协商决定,第 16 届 HKT 学术会议由澳大利亚主办。

第 15 届 HKT 组委会办公室