·煤田物探·

文章编号: 1001-1986(2001)06-0048-04

人工神经网络专家系统在瞬变电磁法反演中的应用

李　实[1], 闫　述[2], 李创社[2], 张立新[3], 陈明生[4]
(1. 山西大学电子系, 山西 太原　030006; 2. 西安交通大学电信工程学院, 陕西 西安　710049;
3. 山西平阳机械厂水暖分厂, 山西 侯马　043000; 4. 煤炭科学研究总院西安分院, 陕西 西安　710054)

摘要: 目前在瞬变电磁法反演方法中, 人们使用较多的是电磁场反演拟合的方法。但往往存在以下不足: 首先这些反演方法理论复杂, 不易掌握; 其次它要求工作人员有丰富的工作经验, 缺乏延续性, 很难推广; 再次数据处理量大, 计算复杂。本文提出一个基于人工神经网络的专家系统。它避开了具体复杂的电磁场计算, 只要经过适当的学习训练就能够解决复杂的实际问题; 而且还具有学习记忆功能, 使得瞬变电磁法的反演工作具有延续性和继承性。随着专家系统的不断完善, 该方法将有广阔的发展前景。

关　键　词: 瞬变电磁法; 反演方法; 人工神经网络; 专家系统
中图分类号: P631.3+25　　**文献标识码**: A

1　引言

瞬变电磁法(The transient electromagnetic method, 简称 TEM)是目前地球物理探测领域里一种应用前景非常广阔的新的勘探方法[1,2,3]。与地球物理勘探领域中的其他方法相比, 瞬变电磁勘探法有简单易行、信息丰富、精度较高、投资小、见效快和应用广等优点。所以应用领域不断扩大, 近年已经在固体矿藏、油气田、土地含盐量勘探和地质填图等方面得到了广泛应用, 并且步入工程、地下水以及地热勘查等行列, 取得了可喜的成果[4,5]。

瞬变电磁法属时间域电磁感应方法, 它利用接地电极或不接地回线通以脉冲电流而在地下建立起一次脉冲磁场, 在一次磁场间隙期间, 利用探测线图观测二次涡流场, 该二次涡流场中包含了丰富的地质信息。瞬变电磁法电磁资料处理的目的就是根据这些实测的二次涡流场曲线(或 TEM 曲线)来反演地下地层的电阻率和厚度变化。

目前人们较多使用对电磁场的反演拟合方法来进行这项工作。经过人们多年的努力, 电磁场的反演拟合方法已取得了较大的进展, 但同时也遇到了很多困难, 且目前无法解决。另外这些反演方法还存在以下不足: 首先这些反演方法理论十分复杂, 不易学习掌握, 要进行反演必须对很深的电磁场理论有相当的熟悉和了解; 其次这些反演方法都要求工作人员有较丰富的经验, 工作人员经验的多少直接影响反演的结果, 所以使反演工作缺乏延续性, 很难推广; 再次后期数据处理量大, 计算复杂, 工作量大, 且要专业人员来完成, 很难进行实时反演。因此我们拟设计一个基于人工神经网络的专家系统。该系统避开了具体复杂的电磁场计算, 只要经过适当的学习训练就能够解决那些复杂的实际问题。而且该专家系统还具有学习记忆功能, 它一边工作一边学习, 使得瞬变电磁法的反演工作具有延续性和继承性。我们使用该专家系统对 TEM 曲线的反演工作做了初步尝试, 结果表明该方法是切实可行的, 随着专家系统的不断完善, 该方法将有广阔的发展前景。

2　方法和原理

人工神经网络是模拟人脑机理和功能的一种新型计算机和人工智能技术。人工神经网络在数据处理中避免了数据分析和建模中的困难, 采用拟人化的方法进行处理, 所以它特别适合于不确定性和非结构化信息处理。地质勘探和开发中的大量信息就是这种信息。神经网络适合研究和解决这些问题, 因此在地质勘探和开发中具有重要意义[6]。

用基于人工神经网络的专家系统反演 TEM 曲

收稿日期: 2001-02-05
基金项目: 国家自然科学基金项目部分成果(编号: 49974026)
作者简介: 李实(1965—), 男, 山西太原人, 博士, 从事电磁场与微波技术的研究.

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

线, 我们采用的人工神经网络是改进的BP 人工神经网络, 即自适应调节学习速率的误差反传BP 人工神经网络[7]。误差反传(BP)算法是目前应用最为广泛的并且比较成功的一种算法, 因其误差通过网络逆向传播而得名。BP 算法实际上是把一组样本输入/输出问题变为一个非线性问题, 使用了优化中最普遍的梯度下降法, 用迭代运算求解权。学习过程可用求权向量函数 $f(w)$ 在集合 W 上的极小问题来表示, 即

$$f_M=\begin{cases}\min f(w);\\ s.t \quad w\in W。\end{cases}$$

这里 w、W 和 f 分别为问题的变量, 可行集(可取整个实数空间)和目标函数。$s.t$ 意为受约束于…。传统的BP 算法可以用以下模型算法表示:

a. 给定 $\varepsilon>0$, 选定 W, 学习率 $\eta>0$;

b. 计算网络输出, 若所有模式的实际输出与网络实际目标(教师)输出之差小于 ε, 则结束, 转 **e**, 否则执行 **c**;

c. 求每个部件的梯度 $\nabla f(w_k)$ 及下降方向 $d_k=-\nabla f(w_k)$;

d. 修正每个部件的权值向量, 使得 $w_{k+1}=w_k+\eta d_k$, 转 **b**;

e. 结束。

从上述模型算法可以清楚地看出, η 的取值在一定程度上可以加快 $f(w)$ 的求解速度, 但是不能保证 w_k 的修正量 ηd_k 使得 $f(w)$ 求解速度达到最佳, 若对(4)作如下设定:

$$F(\eta)=f(w_k+\eta d_k)-f(w_k)。$$

显然, 要使 $f(w)$ 的求解达到最快速度, 必须合理选择 η, 使得 $f(\eta)$ 沿方向 d_k 达到最小。

经过以上分析, 这里给出学习率自调整模型算法。具体描述如下:

a. 给定 $\varepsilon>0$, 选定 W, 学习率 η;

b. 计算网络输出, 若所有模式的实际输出与网络实际目标(教师)输出之差小于 ε, 则结束, 转 **e**, 否则执行 **c**;

c. 求每个部件的梯度 $\nabla f(w_k)$ 及下降方向 $d_k=-\nabla f(w_k)$;

d. 修正每个部件沿方向 d_k 求极小, 即对变量 η 求解问题:

$$F_d=\begin{cases}\min f(w_k+\eta d_k)-f(w_k);\\ s.t \quad \eta\geq 0。\end{cases}$$

则取 $w_{k+1}=w_k+\eta d_k$ 转 **b**;

e. 结束。

由模型算法可以看出, 学习率自调整算法的核心是 F_d 这个单变量函数的极小值问题, 也就是一维搜索问题。目前解决此类问题方法很多, 这些方法通常要搜索到问题的最优解, 至少是局部的最优解。在求解 F_d 时没有必要得到精确的解 η, 只需要 w_{k+1} 满足一些较宽的规则, 如 f 要满足下降性等。这些规则虽不能保证 w_{k+1} 为对应 F_d 问题的最优解, 但是比较容易得到, 从而改善学习效率。

这种不精确的一维搜索包括两部分: 规则和求得符合规则点的步骤。以下具体描述规则和步骤。

设 $f:R^n\to R$ 连续可微, 在点 x, 取定搜索方向为 d, 有

$\nabla f(x)^{\mathrm{T}}d\leq 0$ 要找到一点 $x\leftarrow x+\lambda d$, $\lambda>0$。

规则:

a. $f(x+\lambda d)\leq f(x)+\alpha\lambda\nabla f(x)^{\mathrm{T}}d$;

b. $\nabla f(x+\lambda d)^{\mathrm{T}}d\geq\beta\nabla f(x)^{\mathrm{T}}d$;

c. $\lambda=1$ 优先。

其中选择 $\alpha\in(0,1/2)$, $\beta\in(\alpha,1)$。实用经验提供的一种好的选择是 $\alpha=0.1$, $\beta=0.7$, 或在这附近稍加调整。

步骤:

a. 计算 $f(x)$, $\nabla f(x+\lambda d)^{\mathrm{T}}d$, 置 $\lambda=1$;

b. 计算 $f(x+\lambda d)$, 如果规则 **a** 成立则转 **c**, 否则 $\lambda=1/2$ 重复执行 **b**;

c. 计算 $\nabla f(x+\lambda d)^{\mathrm{T}}d$, 如果 **b** 成立, 则得解 λ, $x\leftarrow x+\lambda d$, 否则转 **b**。

在实际问题处理过程中, 常会出现这样的情况: 函数在 x_k 附近波动; $\|\nabla f(x)\|$ 太小。此时在进行一维搜索时会耗费过多的时间, 这样对整体效率的提高不利, 且容易进入局部最小。为此在设计算法时, 可对 λ 的值加以限制, 即当经过多次迭代值过小且仍未能满足规则时, 令 λ 为一定值, 退出一维不精确搜索。(图 1)

3　具体方法

在解实际问题中, 不能直接采用上述改良的BP 算法来编写反演TEM 曲线的人工神经网络的专家系统, 否则就会出现以下几个问题: 首先, 实测TEM 曲线一般是由 20(甚至更多)个采样点组成, 那么对于人工神经网络的专家系统来说就相当于有 20 个输入变量。为了让人工神经网络的专家系统得到充分的学习训练, 如果每个输入变量取 10 个训练节点, 那么从理论上讲对于 20 个输入变量就有 10^{20} 个训练节点, 即需要 10^{20} 个学习样本该人工神经网络的专家系统才能得到充分的训练, 这是个天文数字, 几乎是不可能实现的, 这就是所谓的维数灾难问题;

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

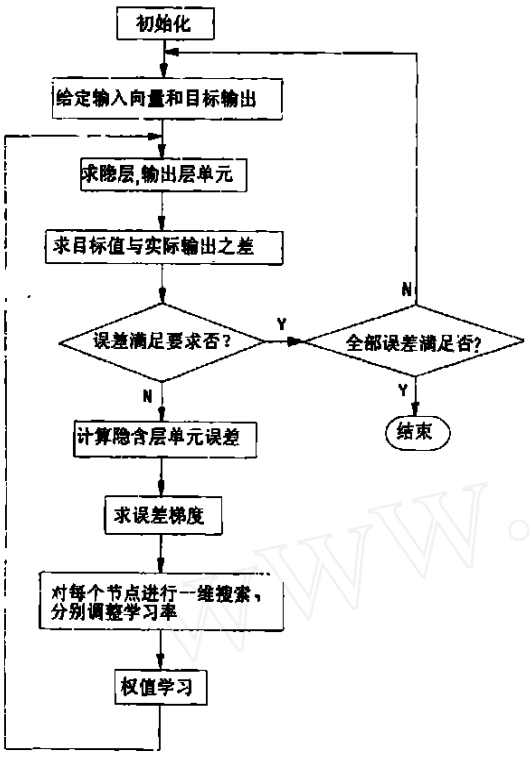


图 1 学习率自调节算法框图

其次, 对于不同的 TEM 系统或不同的采样方法, 这 20 个采样点的采样时刻不同。而即使是对同一条 TEM 曲线, 这 20 个采样点的采样时刻不同, 人工神经网络将视为来自不同 TEM 曲线输入变量。这将造成人工神经网络训练学习的困难, 影响人工神经网络的专家系统的学习训练和通用性; 再次, 对于实测 TEM 曲线, 总是存在一定的噪声干扰, 如将这 20 个输入变量用来学习训练, 则噪声干扰也将进入人工神经网络的专家系统, 这对后期的反演工作是十分有害的。我们解决以上这些问题的方法是提取实测 TEM 曲线的特征不变量。

我们知道在近区瞬变场测深方法中, 晚期测量结果能很好地给出地电断面的分层信息。其物理过程是, 在上部导电层中晚期阶段刚刚出现, 即开始出现涡流的衰减过程, 并以其纵向电导 S_1 来表征该地层的存在时, 在更深的导电层中由于"烟圈"效应, 涡流还在产生和增加的早期阶段。但是, 由于第一层中很强的衰减涡流的屏蔽作用, 在地表测量中很难或很微弱地出现第二层的影响。随着时间的推移, 在地表上可观测到第一层和第二层共同影响的瞬变场结果, 并以 S_1+S_2 来表征其综合影响。在更晚的时间上出现 $S_1+S_2+S_3$ 的综合影响, 依此类推。这样, 随着时间的往后推移, 在晚期测量的结果上逐次附加更深处的涡流影响。于是晚期测量的结果可表示为:

$$V = A_1 e^{-at} + A_2 e^{-a_2(t-\tau_2)} + \cdots + A_n e^{-a_n(t-\tau_n)}, \quad (1)$$

化简得:

$$\begin{aligned} V &= A_1 e^{-at} + A_2 e^{\tau_2} e^{-a_2 t} + \cdots + A_n e^{\tau_n} e^{-a_n t} \\ &= A_1 e^{-at} + A_2 e^{-a_2 t} + \cdots + A_n e^{-a_n t}, \end{aligned} \quad (2)$$

于是对于简单两层地质结构就有:

$$V = A_1 e^{-at} + A_2 e^{-a_2 t}, \quad (3)$$

其中 V、t 已知, 而 A_1、a_1、A_2 和 a_2 是待定系数, 它们可以通过解非线性方程组(4)得到

$$\begin{cases} V_1 = A_1 e^{-a_1 t_1} + A_2 e^{-a_2 t_1}; \\ V_2 = A_1 e^{-a_1 t_2} + A_2 e^{-a_2 t_2}; \\ V_3 = A_1 e^{-a_1 t_3} + A_2 e^{-a_2 t_3}; \\ V_4 = A_1 e^{-a_1 t_4} + A_2 e^{-a_2 t_4}。 \end{cases} \quad (4)$$

可见对于两层地质结构我们可以将 20 个输入变量减少为 4 个输入变量, 同理地质结构每再增加一层就再增加 2 个输入变量; 另外, 一个 TEM 曲线就对应一组待定系数 A_1、a_1、A_2 和 a_2, 与采样时刻、采样速度和采样密度无关。这样就从根本上解决了上述人工神经网络的专家系统所面临的困难。

4 非线性方程组的求解方法

设 n 个未知数的非线性方程组有形式[8]

$$F(X) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix} = \begin{bmatrix} f_1(x_1, x_2, \cdots, x_n) \\ f_2(x_1, x_2, \cdots, x_n) \\ \vdots \\ f_n(x_1, x_2, \cdots, x_n) \end{bmatrix} = 0, \quad (5)$$

设 $f_1(x_1, x_2, \cdots, x_n)$ $(i = 1, 2, \cdots, n)$ 在所讨论的区域内关于自变量各分量的二阶偏导数连续, 利用多元函数的 Taylor 公式, 在 $X^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^T$ 处展开并略去二次以上的项, 可得近似公式:

$$f_i(X) \approx f_i(X^{(k)}) + \sum_{j=1}^{n} \frac{\partial f_i(X^{(k)})}{\partial x_j} (x_j - x_j^{(k)}), \quad i = 1, 2, \cdots, n, \quad (6)$$

以此近似式代入非线性方程组, 其矩阵向量形式是

$$F(X^{(k)}) + F'(X^{(k)})(X - X^{(k)}) = 0。 \quad (7)$$

当矩阵 $F'(X^{(k)})$ 非奇异时, 线性方程组(7)的解存在且唯一, 取其解作为第 $k+1$ 次近似, 于是

$$X^{(k+1)} = X^{(k)} - [F(X^{(k)})]^{-1} F(X^{(k)}), \quad k = 0, 1, 2, \cdots。 \quad (8)$$

这就是解非线性方程组的 Newton 迭代公式, 因公式(8)中需要求逆矩阵, 具体实现时有困难。所以变形为

$$X^{(k+1)} = X^{(k)} - F(X^{(0)})^{-1} F(X^{(k)}), \quad k = 0, 1, 2, \cdots, \quad (9)$$

其各次迭代都要求解具有相同系数矩阵 $F'(X^{(0)})$ 的线性方程组

$$F'(X^{(0)})(X^{(k+1)} - X^{(k)}) = -F(X^{(k)})。 \quad (10)$$

因此, 可将 $F'(X^{(0)})$ 作某种分解并保存起来, 在各次迭代中使用。

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

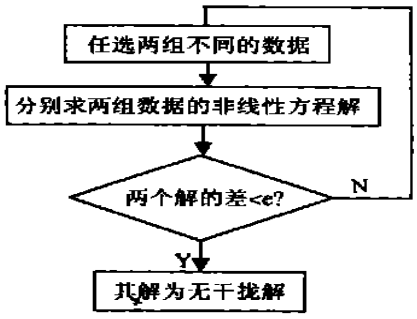


图 2　用解非线性方程组法消除干扰的工作流程图

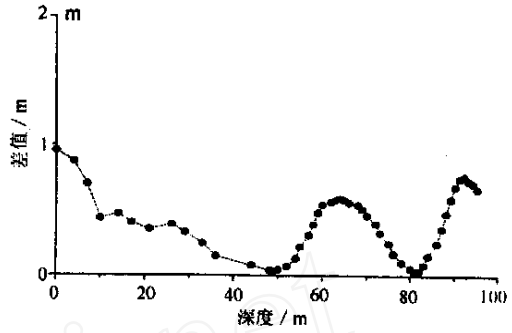


图 3　计算深度与实际深度在各采样点之差

5　用解非线性方程组法消除干扰的原理和方法

在瞬变电磁法中, 干扰一般是没有规律的, 即使有规律也不可能遵循公式(1)所示的规律。因此我们可以认为: 在 20 个输入变量中任选两组不同的数据作为非线性方程组的已知量, 如果其中存在噪声干扰, 即只要任选的两组数据中包含干扰, 则这两个解就不可能相同; 同理, 如果这两个解相同, 则说明该解为无干扰的解。因此根据以上原理, 我们编写了图 2 所示的用解非线性方程组法消除干扰的工作流程图。

6　结果与讨论

图 3 是依据上述方法对一个厚度变化为 0 到 100 m 的测线所做的数值计算结果。它们是经过 30 万次训练后数值计算结果与实际值之差。从中可以看出, 在 100 m 深度的两端误差在 1 m 左右, 而在 100 m 深度的中部误差仅有 0.2 m 左右, 即两端误差大, 中间误差小, 这完全符合神经网络的训练规律。与煤炭科学研究总院西安分院物探所实际工作中所使用的方法相比较, 该精度已经达到实用的要求, 所以结果表明该方法是切实可行的。由于神经网络专家系统的方法具有操作简单, 工作具有延续性和继承性等优点, 所以随着专家系统的不断完善, 该方法将有广阔的发展前景。

参考文献

[1] 牛之琏. 时间域电磁法原理[M]. 中南工业大学出版社, 1992.

[2] 方文藻. 瞬变电磁测深法原理[M]. 西北工业大学出版社, 1993.

[3] 朴化荣. 电磁测探法原理[M]. 地质出版社, 1990.

[4] Spies B P, Frischknecht F C. Electromagnetic sounding, In: Electromagnetic Methods in Applied Geophysics (ed. M N Nabighian) [M]. Society of Exploration Geophysicists, 1991: 285—425.

[5] Raiche A P. The effect of ramp function turn-off on the TEM response of layered earth [J]. Exploration Geophysics, 1984, 15: 37—41.

[6] 冯敬英等. 用人工神经网络技术预测油气藏[J]. 中国地球物理学会年刊, 北京: 中国建材工业出版社, 1996: 173.

[7] 罗四维等. 人工神经网络建造[M]. 中国铁道出版社, 2000: 103—105.

[8] 陈明逵等. 计算方法教程[M]. 西安交通大学出版社, 1992: 185—196.

The application of artificial neural network expert system to the transient electromagnetic method inversion

LI Shi[1], YAN Shu[2], LI Chuang-she[2], ZHANG Li-xin[3], CHEN Ming-shen[4]

(1. Department of Electronics & Information science Shanxi University, Taiyuan 030006, China;
2. Xi'an Jiaotong University, Xi'an 710049, China;
3. Shanxi Pingyang Machine Shop, Houma 043000, China; 4. Xi'an Branch of CCRI, Xi'an 710054, China)

Abstract: Fitting is used mostly in the transient electromagnetic method inversion at present. However there is a few of deficiencies as following. First, its complex theoretical system is no easy to be mastered; secondly, it is difficult to be popularized in the shortage of inheritance. Therefore, scientific researcher is requested to have rich work experience; thirdly, computation complicated for its huge data package. In this paper, an expert system based on artificial neura network is put forward, which avoids detailed and knotted electromagnetic computation. Once afte appropriate training and learning, it can solve practical problem effectively. Furthermore, it memory capacity enhances the continuity and inheritance in the field of TEM inversion. A brigh outlook will be brought along with the consummation of this expert system.

Key words: TEM; inversion; artificial neural network; expert system

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net