

电测深曲线的遗传算法反演

王兴泰 李晓芹¹⁾ 孙仁国

(长春地质学院地球物理系, 长春 130026)

摘 要

电测深曲线作为地下介质电阻率和深度的非线性函数, 其解具有高度的非唯一性。常规的基于局部线性化的最优化反演方法易使解估计陷入局部极大值中, 而且严重地依赖初始模型的选择。遗传算法作为一种全局最优化方法, 对初始模型的依赖性大为减弱, 且不易陷入局部极大值之中, 从而能有效地解决这类非线性最优化问题。本文首次将遗传算法用于电测深解释并对实测曲线进行反演, 效果很好, 显示了遗传算法独特的优越性。

关键词 遗传算法, 电测深, 反演。

1 引 言

在水文工程及环境地质调查中, 经常需要开展电测深工作, 以了解电性随深度的纵向变化。早期, 对电测深曲线的定量解释主要依赖于理论量板, 随着计算机在地学中的广泛应用, 出现了数字解释法。目前, 最常用的反演方法有阻尼最小二乘法、梯度法及变尺度法, 它们都属于局部最优化方法, 需要将非线性问题线性化, 通过解方程而求得局部极值。这三种方法都严重地依赖于初始模型的选择, 解释结果与工作人员的实际经验关系密切, 人为因素较多, 而且不能保证解估计是全局最优的。

近年来发展起来的遗传算法作为一种全局最优化方法, 运用随机而非确定性的规则对模型空间中的一族而非一个点进行全局性而非局部搜索, 它仅利用目标函数而不要求导数或其它附加限制, 更重要的是遗传算法大大放宽了对初始模型的选择要求, 由于这些独特优点, 近几年来日益受到国内外广大地球物理学者的关注, 最先将其引入地球物理学领域的是用来反演一维地震波模型^[1], 本文首次尝试将其引入电法勘探中, 用于反演电测深曲线。

2 遗 传 算 法

2.1 基本思想

遗传算法 (Genetic algorithm, 简称 GA 算法): 由 Holland 于 1975 年模拟自

1) 现在国家地震局地球物理研究所, 北京 100080。

本文 1994 年 11 月 8 日收到, 1995 年 9 月收到修改稿。

然选择和遗传学理论,依据“适者生存”原理而创立的。它作为一种随机全局搜索算法,对模型参数用二进制编码形成“串”,利用转移概率规则帮助指导搜索。首先生成一个模型群体(集),然后由一个称之为“选择(Selection)”的过程根据它们的拟合函数值进行繁殖(Reproduction),即根据群体中各成员的拟合好坏进行排队,并根据拟合程度分布情况选择一个新的群体,保留拟合好的成员而去掉拟合差的成员,拟合函数好的模型被繁殖的概率大,然后是“交换(Crossover)”和“变异(Mutation)”,通过随机部分交换和改变串中的某位来更新模型,这样就生成了新一代模型,它将被评价、选择和繁殖(保证新一代群体的成员数与老一代群体的成员数相同)。重复上述过程直至模型群体变得均匀一致,即模型群体的拟合函数值变得很小,拟合函数的平均值趋于群体的最大拟合函数值。

GA 算法并不能绝对保证收敛到拟合函数的全局最优解,但由于这种算法是平行地评价模型空间各部分的拟合函数并对它们进行比较,所以只要模型群体的成员数大小、交换和变异的概率选择合适,这种算法一般不会使解陷入拟合函数的局部最优解中。

2.2 具体实现中的有关问题

2.2.1 参数编码 GA 算法的特点之一就是二进制编码实现模型的参数化,即将参数空间中一个点映射到一条染色体上,二进制染色体上每一位即一个基因。染色体长度取决于参数取值范围和模型分辨率,即取

$$2^K = [\max(m) - \min(m)]/\Delta m, \quad (1)$$

式中, K 为染色体长度, m 表示模型, Δm 为模型参数的采样间隔,即分辨率。在参数编码时,可将 k 个参数分别单独编码,也可将全部参数编入一条染色体中,各参数只占染色体的一段。本文采用后一种方法。

2.2.2 初始模型集生成 初始模型集是随机产生、均匀分布于模型空间的。为保证算法收敛,有必要在初始模型集中包含足够的遗传物质,使其可以得到参数空间的任意点,而且在运算过程中不丢失这种能力。为此,本文采用 Berg^[2] 提出的方案,即在初始模型群体中增加各位都为“0”和各位都为“1”的成员。

2.2.3 模型选择 选择模型的方法有多种,本文用模型集内各模型的目标函数对所有模型目标函数值之和的比值定义选择概率

$$P_i(m) = \frac{E(m)}{\sum_A E(m)}, \quad (2)$$

式中 $A \subset M$, M 表示模型空间, A 为上一代模型群体组成的子空间, $E(m)$ 为对应模型得到的目标函数。根据这些概率作选择,直至原模型子集都配成对。

2.2.4 交换 对配对后的模型各编码参数设定一个位交换概率,分别随机地交换两个编码参数中选定位置右边的各位,得到两个新参数。

2.2.5 变异 按一定概率改变随机选定的参数编码中的某位,以产生新一代群体,若新一代模型的拟合程度好,则保留,否则仍保留上一代模型。

2.2.6 目标函数的尺度变换 为避免在迭代初期一个或若干个偶发的优越模型(但它们未必接近全局最佳模型),利用初始条件控制了整个迭代搜索过程,也要避免后期在模型大体收敛到最佳值附近时,因各模型目标函数相差不大,造成搜索速度下降,需要对目

标函数作尺度变换。Stoffa 和 Sen^[1] 采用模拟退火法中的方法来拉伸目标函数, 即引入退火温度 T_d , 使生存概率为

$$P_d(m) = \frac{\exp[E(m)/T_d]}{\sum_d \exp[E(m)/T_d]}, \quad (3)$$

这样随着温度的降低, 选择的是对目标函数影响最大的模型参数值, 从而使原始模型群体中的好模型控制着算法的过程。

3 正演计算及目标函数的选取

3.1 正演计算

对于电测深曲线的正演计算, 本文采用众所周知的递推公式:

$$T_i = \frac{T_{i+1} + \rho_i \tanh(\lambda h_i)}{1 + T_{i+1} \tanh(\lambda h_i) / \rho_i}, \quad (i = n-1, n-2, \dots, 1) \quad (4)$$

其中, T_i 为第 i 层的电阻率转换函数, ρ_i 为第 i 层电阻率值, h_i 为第 i 层厚度, λ 为常数。

然后利用数字滤波方法, 将最后得到的电阻率转换函数 T_1 变为所需视电阻率 ρ_e ,

$$\rho_{ej} = \sum_{k=k_1}^{k_2} T_1(j-k)C(k), \quad (5)$$

这里 j 为极距编号, $C(k)$ 为滤波系数, k_1 、 k_2 分别为滤波系数的起止编号。

3.2 目标函数的确定

目标函数的选择对收敛速度有重要影响, 恰当地选择目标函数是 GA 算法中的一个重要问题。针对电测深曲线, 为提高模型参数质量, 加速收敛, 本文采用如下目标函数作为拟合尺度

$$\sigma = \left(\frac{1}{N} \sum_{j=1}^N (\ln \rho_{ej} - \ln \rho_{ij})^2 \right)^{1/2}, \quad (6)$$

ρ_{ij} 为第 j 个极距的实测视电阻率值, N 为极距总个数。

4 数值模拟及结果分析

对于 n 层地电结构, 其层参数可用 \mathbf{d} 表示

$$\mathbf{d} = [\rho_1, \rho_2, \rho_3, \dots, \rho_n, h_1, h_2, h_3, \dots, h_{n-1}]^T, \quad (7)$$

\mathbf{d} 共有 $(2n-1)$ 个模型参数, 其中最后一层为均匀半空间。

在用 GA 算法进行反演时, 分两步进行: (1) 根据实测曲线确定层数, 这一般是通过实测曲线进行定性分析得到的; (2) 给出每一层电阻率及厚度的搜索范围, 即对模型参数 d_i , 确定其搜索范围

$$d_i^{\min} \leq d_i \leq d_i^{\max}, \quad (8)$$

同时还要根据精度及计算时间的合理性确定采样间隔 Δd_i 、模型群体大小 W , 一般要反

演的模型参数越多, W 值越大。

以下所有计算均在加有协处理器的 386 微机上进行, 电极装置形式为对称四极。

4.1 二层 G 型理论曲线

表 1 给出了模型参数, GA 算法的搜索范围。模型群体 W 取 32 个, 经过 40 次迭代, 用时 20s, 得到的反演结果(表 1)非常理想, 这里的最佳模型拟合方差 $\sigma = 0.003$ 。从图 1a 也可看出, 理论模型的正演曲线和由反演所得模型的正演曲线完全重合, 两者的拟合程度极高。

表 1 两层地电模型搜索范围及反演结果

模型参数	ρ_1 Ω_m	ρ_2 Ω_m	h_1 m
理论值	50	500	3
搜索范围	40—60	400—600	1—6
反演结果	49.88	498.53	2.97

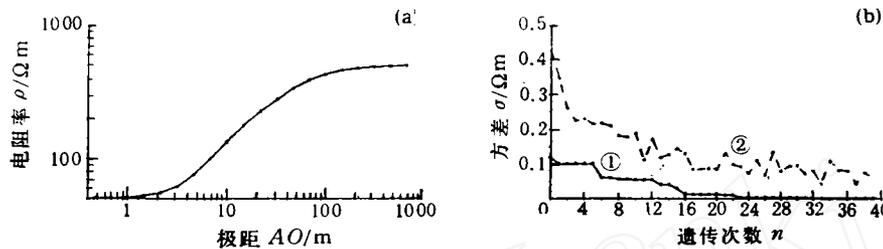


图 1 两层 G 型电测深曲线

(a) 理论(实线)与反演结果(虚线); (b) 拟合函数随遗传次数的收敛情况;
① 最佳模型; ② 平均模型。

原则上讲, 随着遗传次数的增加, 平均模型的拟合差与最佳模型的拟合差将收敛于同一值, 即模型群体变得均匀一致, 此时反演结果最真实可靠。但在计算中, 因受计算时间的限制, 一般在满足精度要求的情况下, 只要最佳模型参数及其拟合差随遗传次数的增加而稳定时, 即可停止迭代。如图 1b 中, 实际上迭代超过 30 次后, 最佳模型的参数及其拟合差基本上不再随遗传次数增加而变化, 我们就认为该最佳模型即是反演的最终结果。

4.2 三层 H 型理论曲线

这里的模型参数引自文献 1), 以便将 GA 算法、变尺度法和阻尼最小二乘法加以对比。从表 2 中可看出, 变尺度法和阻尼最小二乘法对中间层电阻率和厚度的初始选择较苛刻, 而该层参数往往又是最难确定的, 相对而言, GA 算法不仅大大放宽了对初始模型选择的要求, 而且在较短的时间内能够得到与变尺度法、阻尼最小二乘法同样理想的结果(如图 2a)。这里 $W = 64$, 60 次迭代, 用时 60s, 拟合方差 $\sigma = 0.018$ 。由于中间层电阻率较低, 厚度较小, 迭代次数增多(图 2b)。

1) 中南矿业学院, 直流电法推断解释中的若干问题(内部资料), 1978。

表 2 三层地电模型及反演结果

模型参数		ρ_1 Ωm	ρ_2 Ωm	ρ_3 Ωm	h_1 m	h_2 m
理论值		10	1	15	3	15
变尺度法	初值	10	3	40	1	5
	结果	10.02	1.10	14.46	2.90	16
阻尼最小二乘法	初值	10	0.5	25	1.5	30
	结果	10	1.1	14.5	2.98	14.8
遗传算法	范围	5—20	0.5—3	5—50	1—5	5—30
	结果	9.98	0.99	15.03	3.02	14.97

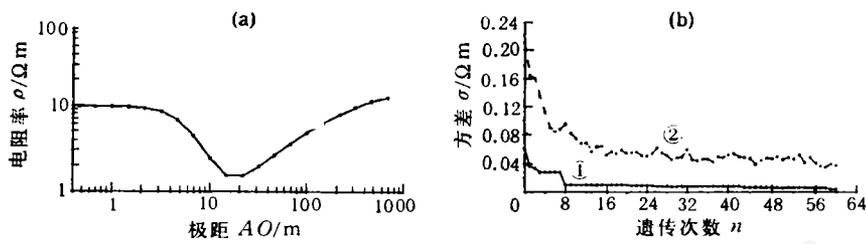


图 2 三层H型电测深曲线
(a) 理论(实线)与反演结果(虚线); (b) 同图 1(b).

表 3 四层实测钻孔资料及 GA 算法反演结果

模型参数	ρ_1 Ωm	ρ_2 Ωm	ρ_3 Ωm	ρ_4 Ωm	h_1 m	h_2 m	h_3 m
搜索范围	65—75	50—300	10—60	2000—5000	3—20	5—40	20—100
结果	69.67	152.84	26.58	4392.15	8.18	22.05	81.26
钻孔资料					8.2	22.3	80.1

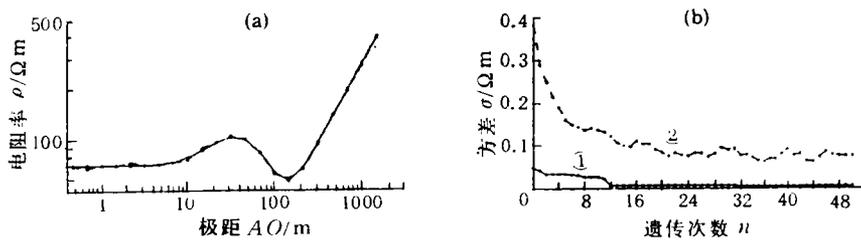


图 3 四层电测深曲线
(a) 实测(实线)与反演结果(虚线); (b) 同图 1(b).

4.3 四层 KH 型实测曲线的反演

从表 3 可以看出, GA 算法的反演结果与实际钻孔资料¹⁾非常吻合, 图 3a 中除个别点外, 曲线的拟合也很好。这里 $W = 128$, 经过 50 次迭代, 用时 250s, 拟合方差 $\sigma = 0.005$ 。

5 关于等值现象的讨论

电测深曲线的反演属于典型的非唯一性问题, 即在一定的误差范围内, 电测深曲线常出现等值现象, 具体表现在反演时, 目标函数出现很多几乎相等的次极大。

在用 GA 算法反演电测深曲线时, 因其采样是在整个模型空间均匀进行, 而且随着迭代次数的增加, 拟合方差小至 10^{-2} 到 10^{-4} , 甚至更小, 使得出现 S 等值的条件 ($\nu_2 = h_2/h_1 \ll 1, \mu_3 = \rho_3/\rho_2 \gg 1$) 或 T 等值的条件 ($\nu_2 \ll 1, \mu_3 \ll 1$), 相对地就很难得以满足, 所以在层数给定的情况下, 即使在很大的搜索范围内, GA 算法都能保证解估计收敛到真实模型, 当层数较多时, 还可结合对最后迭代模型参数的直方图分析, 而得到最优解。作者对大量理论模型的反演, 充分证明了这一点, 可以说 GA 算法在很大程度上克服了电测深反演中固有的等值现象, 这也是它不同于阻尼最小二乘法和变尺度法等常规反演方法的独特之处。用 GA 算法反演电测深曲线, 是在已知地电结构层数的情况下进行, 所以对于混层等值现象, GA 算法不能有效地克服。

6 结 论

GA 算法作为一种随机搜索方法, 具有全局搜索, 不用求导, 收敛较快, 更重要的是它不需要良好的初始模型, 从而减少了电测深反演解释中的人为因素, 比已有的反演方法具有更大的自动性, 而且它在一定程度上有效地解决了电测深反演中固有的等值现象, 大大提高了解释结果的精度, 是一种很有潜力的方法, 完全可以取代已有的局部最优化反演方法。

参 考 文 献

- [1] Stoffa, P. L., Sen, M. K., Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms, *Geophysics*, **56**, 1794—1810, 1991.
- [2] Berg, E., Simple convergent genetic algorithm for multiparameter data, 60th SEG meeting, American, Expanded Abstracts, 312—316, 1990.
- [3] 赵改善编译, 求解非线性最优化问题的遗传算法, *地球物理学进展*, **7**(1), 90—96, 1992.

1) 林君琴, 吴至善, 电测深曲线联合反演解释在水工物探中的应用(内部资料), 1989.

THE INVERSION OF RESISTIVITY SOUNDING CURVE USING GENETIC ALGORITHMS

WANG XING-TAI LI XIAO-QIN SUN REN-GUO
(Changchun Geology College, Changchun 130026, China)

Abstract

The sounding curve is the nonlinear function of depth and resistivity, which has high nonunique solution. The normal optimum inversion methods based on local linearization are usually lost in local maximum values, and they seriously depend on the selection of initial model. The genetic algorithm is a global search one, which is hardly lost in local maximum and has less dependence on initial values, so it can effectively solve the problem. In this paper, a lot of sounding curves about geoelectric models or field examples are inverted using GA. The results are ideal, which prove it is available for GA to inverse the resistivity sounding data.

Key words Genetic algorithms (GA), Resistivity sounding, Inversion.

作者简介 王兴泰,男,1937年生,1961年毕业于长春地质学院物探系,现为长春地质学院地球物理系工程物探研究室主任,教授。主要从事电法勘探及工程物探的研究工作,曾发表《高密度电阻率法及其应用技术研究》等40多篇学术论文。