

界面位场的遗传反演方法研究*

石琳珂

(国家地震局地球物理勘探中心, 郑州 450003)

摘 要 本文分析了 Parker—Oldenburg 位场反演方法中存在的问题, 提出了用位场频率域快速正演方法与遗传算法相结合的位场遗传反演方法。针对遗传算法中收敛速度慢的问题, 提出了两方面的改进措施, 使这种位场遗传反演方法更为完善。经理论模型验证与实际资料的处理, 表明该方法是有用的。

主题词: 地震勘探; 三维地震法; 地震反演; 迭代法

0 引 言

由地面已知位场资料反演物性界面, 由此了解地球主要物性分界面如莫氏面、居里面的变化起伏情况, 是一个有着重要现实意义的课题, 无论在地球物理研究, 还是在矿产预测方面都有重要应用。

目前, 在利用位场资料反演物性界面的研究、应用工作中, 最常用的方法是 Parker—Oldenburg 反演方法。其中, 正演方法由 Parker 于 1973 年提出^[1], 次年, Oldenburg 把它推广成迭代反演形式^[2]。这种方法于 80 年代引入我国, 从那以后国内的很多研究人员对 Parker 正演公式又做了改进, 建立了物性参数三维分布的位场频率域快速正演理论公式^{[3][4]}, 即物性参数不仅在平面上可以任意变化, 在深度上也可以呈线性或指数函数形式变化, 使这种方法的应用范围更宽。

但是, Parker—Oldenburg 反演方法中也存在一些问题, 主要在于反演方面。该方法的反演公式中含有一项对高频干扰成分有放大作用的向下延拓因子, 它往往使迭代过程发散。为克服向下延拓因子的不良影响, 不同的研究人员在各自的应用研究中采取了不同的措施, 如用补偿圆滑滤波、正则化滤波、逐次向下延拓等, 以便使反演过程稳定收敛。

但是, 作者认为这里仍有一些问题值得考虑。首先, 向下延拓计算本身在位场数据处理中就是一个难度很大的问题, 解决它并非易事。其次, 有时在选择滤波参数时也会遇到困难, 很难做到恰如其分。最后一点, 也是最重要的一点, 即在反演过程中反复使用滤波手段, 必然严重影响数据质量, 使有用信息损失严重。而快速富里叶变换带来的高速度并不能弥补有用信息的损失, 所以用 Parker—Oldenburg 反演方法做位场反演得

* 国家地震局合同制项目 (90410806) 资助

不偿失。

针对这些问题,本文提出一种位场频率域快速正演方法与遗传算法相结合的位场反演方法——位场遗传反演方法。这种反演方法充分利用了正演方面采用 FFT 带来的高速度,同时又不存在 Parker—Oldenburg 反演方法中的种种困难,是一种空间域与频率域相结合、扬长避短的位场反演方法。下面对位场遗传反演方法的正演计算、反演计算以及应用中的具体情况分别予以介绍。

1 三维密度界面的频率域快速正演方法

设密度函数 $\sigma = \sigma(\xi, \eta)$ 在平面上是变化的,则三维密度单一界面的快速正演公式为:

$$F[\Delta g(x, y)] = 2\pi f \exp(-\omega^* Z_0) \sum_{n=1}^{\infty} \frac{(-\omega)^{n-1}}{n!} F[\sigma(\xi, \eta) * \Delta h(\xi, \eta)^*] \quad (1)$$

其中, f 是万有引力常数, ω 是圆频率, Z_0 是界面深度调整参数, n 是级数求和项数, $\Delta h(\xi, \eta)$ 是界面相对于界面深度调整参数的起伏幅度, 定义为:

$$\Delta h(\xi, \eta) = h(\xi, \eta) - Z_0 \quad (2)$$

$\Delta g(x, y)$ 是重力异常值, $F[\]$ 为富里叶变换运算符号, 由 (1) 式得到 Δg 的频谱后, 再做一次富里叶反变换, 即可得到理论重力异常值。

2 位场遗传反演方法

遗传反演方法的基础是遗传算法, 属于随机搜索法。它没有一个解析的反演迭代公式, 其反演过程是通过一系列简单的数字处理来实现的^[5]。下面结合位场反演的具体情况, 给出遗传反演方法的具体步骤。

以单一密度界面为例。首先假设观测点为 mn 个, 观测值 Δg_j ($j=1, mn$) 为已知, 并设模型群体大小为 0。

2.1 确定密度界面的初始值及搜索范围

由经过处理后的布格重力异常数据 $\Delta g(x, y)$, 根据公式 [6]

$$h(\xi, \eta) = \frac{\Delta g(x, y)}{\epsilon \pi f \Delta \sigma} \quad (3)$$

可得到界面的初始值 $h(\xi, \eta)$ 。其中, f 是万有引力常数, $\Delta \sigma$ 是界面上下密度差, $1 < \epsilon < 2$ 。得到 $h(\xi, \eta)$ 后, 即可得到界面深度搜索范围的上下限 $h_{\perp}(\xi, \eta)$ 和 $h_{\top}(\xi, \eta)$ 。

2.2 初始模型群体的产生

在上述指定的参数搜索范围及分段基础上, 由随机数发生器产生一随机数 a ($0 \leq a \leq 1$), 由它来随机地选择每一个模型参数, 这些模型参数组成一个完整的界面模型。重复这个步骤 Q 次, 得到 Q 个模型, 形成遗传算法的初始模型群体。

2.3 正演计算、目标函数

有了初始模型群体后, 就可以用正演公式依次计算各个模型的理论场值 f_{ij} ($i=1, 0$, $j=1, mn$) 和目标函数 $F(i)$ ($i=1, Q$), 其中目标函数可根据具体情况定义。 $F(i)$ 的大小表示第 i 个模型对观测值 (环境) 的不适应程度 (misfitness), 其值越大表示适应程

度越差。

2.4 目标函数的转换

这一步骤是把衡量模型对环境不适应程度的量(目标函数值) $F(i)$, 转换成表示适应程度的量 $Fitness(i)$ 。最简单的方法是取 $F(i)$ 的倒数, 即

$$Fitness(i) = 1/F(i) \quad i=1, Q \quad (4)$$

它表示目标函数越大, 其适应程度越小。反之, 目标函数越小, 其适应程度就越大。

到此, 各个模型对环境的适应程度, 就有相应的适应程度值 $Fitness(i)$ 唯一表示, 它们在以后遗传演化过程中的结果, 也完全由 $Fitness(i)$ 来决定。

2.5 确定各模型的再生(reproduction)

这一步骤是对模型空间进行全局搜索的一种方式。再生模型的选择是根据各模型对环境的适应程度由随机数决定。因为随机数的产生是均匀的, 所以适应程度值大的模型被选择获得再生的概率相应地大些。这种按模型对环境的适应程度决定其再生机会的方法, 体现了优胜劣汰、适者生存的生物进化原理。

2.6 基因交换(crossover)

这一步骤是对模型空间进行局部搜索的一种方式。

在上述获得再生的模型群体中随机选择两个染色体(模型)组成一对双亲(parents), 那么有 Q 个模型的群体可以组成 $Q/2$ 对双亲。对这 $Q/2$ 对双亲分别按概率 P_c 进行基因交换, 可产生一对后代。由此产生下一代模型群体, 而且下一代的模型群体中仍有 Q 个模型。

2.7 基因变异(mutation)

这是对模型空间进行局部搜索的另一种方式。

对模型群体中各个模型的全部二进制位按概率 P_m 做随机的改变, 即 $0 \rightarrow 1, 1 \rightarrow 0$ 。

到此, 第一代遗传演化过程结束, 产生了新一代模型群体。转到步骤 3 可继续进行下一代遗传演化。

2.8 演化中止条件

目标函数的最小值小于给定精度; 或演化到指定次数。

以上就是位场遗传反演方法的基本内容。从以上讨论可以看到, 遗传算法在其演化过程中, 除了随机数的产生和二进制运算之外, 不涉及复杂的数学运算。它只对由正演得到的目标函数值做了一些简单的处理, 如目标函数到适应程度的转换, 适应程度的求和, 基因交换和变异等, 而这些运算不会在过程中引入不稳定的因素, 使过程发散。这样, 就从根本上避免了 Parker—Oldenburg 反演方法中存在的问题。

3 遗传算法加速收敛方法

3.1 问题

在用遗传算法解反演问题的过程中, 经常可以看到这种情况, 即在开始阶段, 目标函数的下降很快, 随着遗传演化的进行, 目标函数的下降越来越慢, 最后成为一条近乎水平的直线, 但这是并未达到所要求的精度。出现这种情况的原因有两个, 分别存在于模型参数方面和适应程度函数方面。

1) 模型参数方面的原因是, 给定的模型参数初始搜索精度太低。对于这种情况有两个解决办法。①给定较高的初始模型搜索精度。但这个办法会大量地增加模型搜索空间的大小, 对于多参数的反演问题可能引起计算机内存方面的问题。②先给定较低的初始搜索精度, 在遗传演化的过程中不断缩小搜索范围。这样, 在保持搜索区间分段数不变的情况下, 相当于逐步提高了搜索精度。这就是所谓逐步缩小搜索范围的方法^[7]。

2) 适应程度函数方面的原因是, 演化进行到后期, 群体中各模型都已接近其真值, 它们的目标函数也比较接近, 因此它们的适应程度值差别不大, 于是造成在选择再生模型时出现困难。

通过对原因的分析可以发现, 解决这个问题关键是要把这些再生概率相差不大的模型区别开来, 而要做到这一步, 就要首先把它们的适应程度数值上的微小差别区别开来。下面用 S 型函数来解决这一问题。

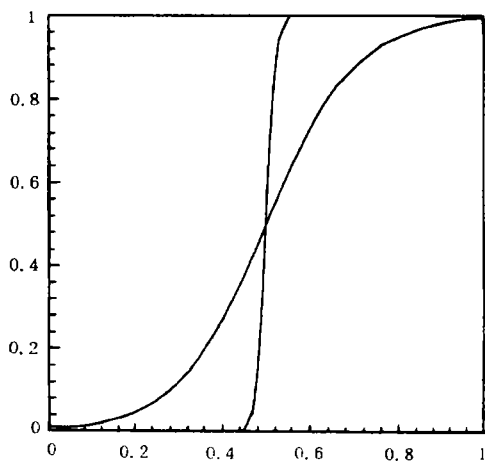


图 1 具有不同 β 值的 S 型函数 ($\alpha=1$)

增加了 S 型函数后, 遗传反演方法中的第 5 步“确定各模型的再生”要稍微做些变动。首先把各模型的 Fitness (i) 代入 S 型函数中, 得到的结果仍用 Fitness (i) 表示, 然后再做下面的计算。

应用中发现, S 型函数加速收敛的效果在前期比逐步缩小搜索范围的方法还要好, 但在后期, 后者的效果更好些。如果把这两种方法结合起来, 可以收到更好的加速收敛效果。因此, 在以后的应用中, 都用这种两者结合的方法做为加速收敛的基本方法。

4 数值模拟

4.1 理论模型

理论界面模型见图 2, 由理论界面模型得到的重力异常见图 3, 这些数据在反演中做为观测数据使用。

4.2 参数

正反演所用的各种参数为:

3.2 S 型函数

S 型函数的作用是放大大于阈值的量而压制小于阈值的量。其数学表达式为

$$y = \frac{1}{1 + \alpha * e^{-\beta(x-\theta)}} \quad (5)$$

其中, x 是自变量, 表示各模型的适应程度值。 y 表示模型的适应程度值经 S 型函数作用后的结果。 α 是比例控制参数, 其值为 2 或 3。 β 是曲线形态控制参数, $\beta=900 \sim 1300$ (据经验)。 θ 是阈值, 控制曲线的位置和对 x 的作用结果, 其值可取为群体平均适应程度值的 $(1+\epsilon)$ 倍, 其中 $\epsilon=0.0 \sim 0.3$ 。其图形见图 1。这些参数确定之后, S 型函数就完全确定了。

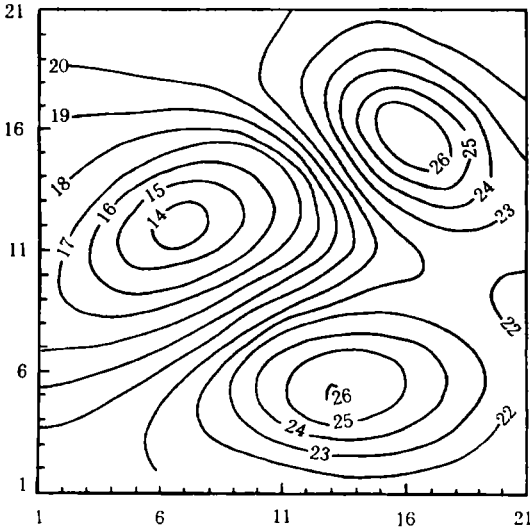


图 2 理论界面等值线图 (单位: km)

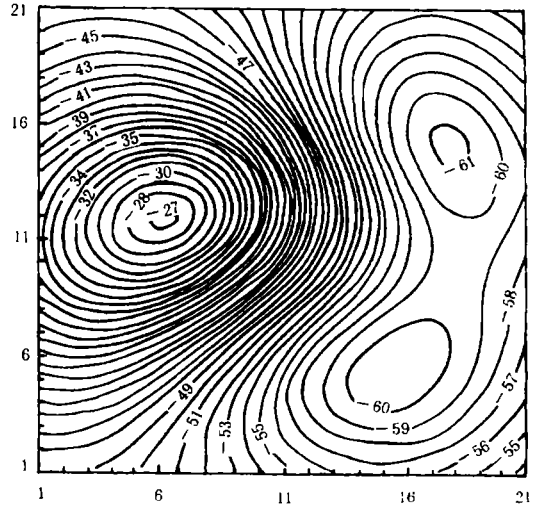


图 3 理论布格重力异常 (单位: 10^{-5}ms^{-2})

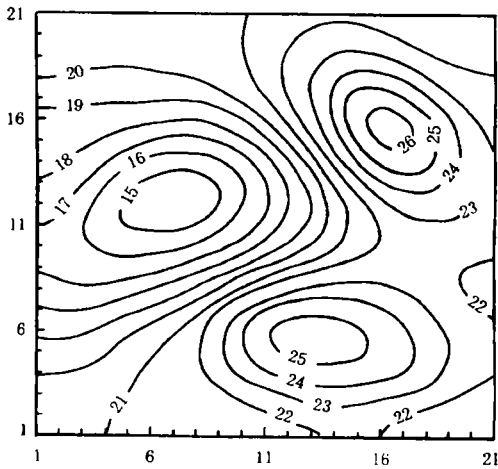


图 4 数值模拟反演结果 (单位: km)

数值模拟结果是令人满意的。

5 反演实例

为检验本文提出的遗传反演方法的效果,利用山西省地质矿产局 1988 年所做的山西高原 1:100 万布格重力异常资料^①,对山西省及其邻近地区的莫氏面做了反演计算。

①测网参数:线数、点数均为 21,线距、点距均为 5km。

②正演参数: $\Delta\sigma = 0.25 * 10^3 \text{kg/m}^3$, $h_0 = 20 \text{km}$ 。

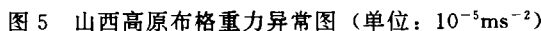
③遗传算法参数: $p_c = 0.90$, $p_m = 0.04$, $Q = 20$ 。

④加速收敛方法:逐步压缩搜索范围与 S 型函数相结合。

数值模拟的结果见图 4。把图 4 与图 2 比较可以看到,在形态上反演结果与其理论值非常相似,在数值上除三个极值点处数值稍有差别外,其它都比较接近。

根据这些情况,可以认为这里所提出的位场遗传反演方法,其数

① 山西省地质矿产局物探队,山西省 1:50 万区域重力调查成果报告,1988。



5.4 反演计算

从图 7 可见, 本区莫氏面的埋深有如下特征: ①在位于河北、河南省的东南边界附近, 莫氏面埋深较浅, 深度为 33~37km。向西北方向莫氏面逐渐加深, 到位于内蒙古自治区的西北边界附近, 莫氏面深度增加到 43~45km。②在与太行山 S 形重力梯级带相对应的区域, 莫氏面深度变化剧烈, 从 35km 增加到 41km。③在太原~临汾一带, 莫氏面

计算区域范围为 $N: 34^\circ$

5.2 布格重力异常

5.3 深部异常信息提取

从图 6 可见, 其主要特征

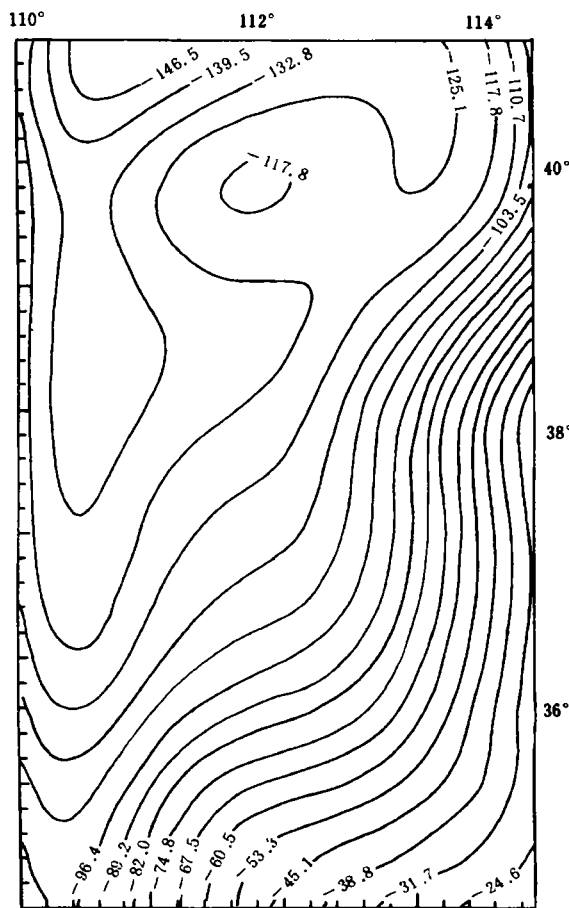


图6 布格重力异常经正则化滤波后的结果
(单位: 10^{-5}ms^{-2})

变化比较平缓, 约为 41~42km。④ 浑源~繁峙和兴县一带为局部凹陷区, 其深度分别为 43km 和 44km。⑤ 在繁峙西部有一局部隆起区, 最浅处小于 41km。

根据我国莫氏面分区, 莫氏面深度在我国东部约为 30~40km 之间, 在中部约为 40~50km, 东部与中部的分界线位于大兴安岭~太行山~武陵山一线。本区的反演结果总体上与上述分区结果一致。

在局部上, 本次反演结果与陕西省地矿局的反演结果在基本趋势和局部隆起及凹陷的部位上也大体相同, 只是在数值上有些差别。造成这些差别的原因是, 后者是由若干条剖面上的二维反演结果绘制等值线得到的, 而前者是直接由三维反演方法得到, 因此, 两者之间存在一定的差异在所难免。

6 总 结

本文分析了目前位场资料反演中应用最为广泛的 Parker—Oldenburg 反演方法中的优点及不足, 弘

扬其优点, 避开其不足, 结合目前在国外发展迅速的解非线性问题的有效工具——遗传算法, 提出了三维位场资料反演的遗传反演方法。主要内容归纳如下。

(1) 本文首先分析了 Parker—Oldenburg 反演方法的优缺点。其优点是正演采用频率域快速计算公式, 速度快, 精度高, 而且使用方便。后经国内外研究人员的不断改进及发展, 已经十分完善。物性参数不仅在平面上可以变化, 在深度上也可以呈线性或指数函数形式变化。该方法的这些优点已得到国内外研究人员的一致认可, 被公认为是位场正演计算中最有效的方法。为充分利用它的这些优点, 本文提出的位场遗传反演方法也用它做为其正演计算手段:

(2) Parker—Oldenburg 方法的反演部分存在一些问题。其中最主要的问题是反演公式中含有向下延拓因子, 该因子对高频干扰成分有放大作用, 结果导致反演过程不稳定。为解决这一下延拓因子带来的麻烦, 人们在应用中采取了很多不同的措施, 效果也不尽相同, 但没有一个公认的有效解决办法。这是因为, 向下延拓计算本身在位场数据处理中就是一大难题。鉴于这种情况, 作者认为, 与其花费很大精力去解决一个很难解决的问

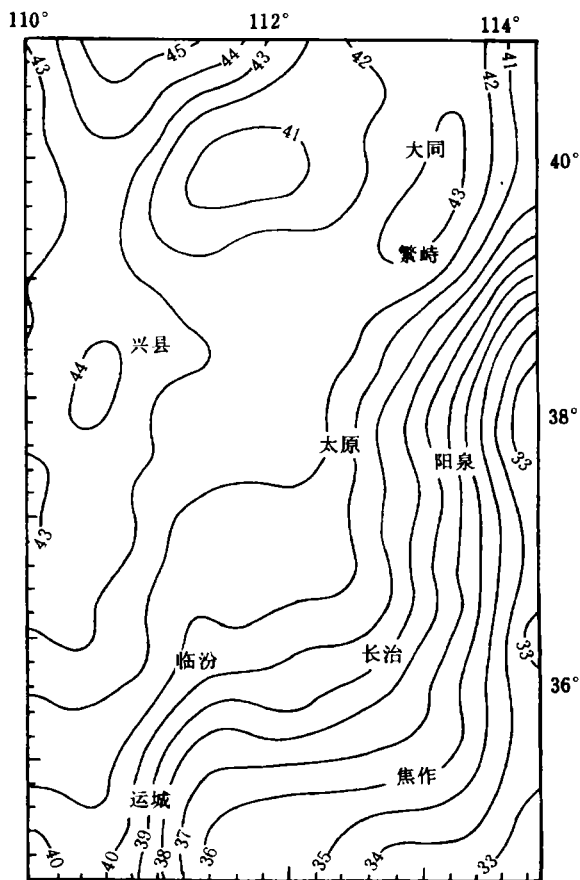


图 7 用遗传反演方法得到的山西高原
莫氏面埋深等值线图 (单位: km)

题, 倒不如彻底避免这个问题, 最有效的办法是去寻找不含有这类问题的新的反演方法。基于这种考虑, 提出了位场资料的遗传反演方法。做为遗传反演方法核心部分的遗传算法, 它的运算非常简单, 这种运算上的简单性使它的反演过程非常稳定、健全。这样, 就从根本上避免了 Parker—Oldenburg 反演方法中的种种问题。

(3) 遗传反演方法的基础是遗传算法。遗传算法近几年在国外发展非常迅速, 应用范围遍及科学技术的各个主要领域, 被认为是解决非线性优化问题的有效方法。只要有待解决问题的正演计算程序, 就很容易与遗传算法结合, 构成反演系统。而做为遗传算法核心部分的遗传三步运算——再生、交换和变异, 其数学计算又非常简单, 其中只涉及到随机数的产生、十进制与二进制的相互转换及二进制的取反运算, 不涉及任何其它复杂的数学方法。这些简单的运算使反演过程

很稳定。另外, 与其它随机搜索方法相比, 遗传算法具有经验积累能力, 其搜索效率比其它方法高。

(4) 在遗传反演方法中增加了加速收敛的措施。遗传算法的一个弱点是当待求参数增多时收敛速度慢。针对这种情况, 在分析了造成这一问题的原因后, 提出了两方面的改进措施。一方面是从参数搜索的角度, 另一方面是从参数的适应程度分辨的角度。这些措施有很强的针对性, 理论计算表明, 效果显著。

(5) 用本文提出的这种遗传反演方法, 分别做了理论上的数值模拟和实际资料的反演计算。经与理论模型及前人的反演结果对比, 效果良好。据现有资料, 用遗传算法解这种多参数的反演问题是前所未有的。可以说, 这是一次有意义的探索和尝试, 可以为今后的进一步研究、应用提供经验和借鉴。

本项研究工作得到了中国地质大学(北京)研究生院应用地球物理系管志宁教授、国家地震局地球物理勘探中心孙武城研究员的指导, 在此表示衷心感谢!

(1995 年 9 月 25 日收到初稿)

参 考 文 献

- 1 Parker, R. L., The rapid calculation of potential anomalies, *Geophys. J. R. Astr. Soc.*, 1973, 31: 447—455
- 2 Oldenburg, D. W., The inversion and interpretation of gravity anomalies, *Geophysics*, 1974, 39 (4): 526—536
- 3 潘作枢. 位场资料处理及解释问题. 北京: 地质出版社, 1992. 16—22
- 4 冯锐. 三维物性分布的位场计算. *地球物理学报*, 1986, 29 (4): 399—406
- 5 D. E. Goldberg, *Genetic Algorithms in Search, Optimization & Machine Learning*, Addison—Wesley publishing Company, Inc., 1989. 10—18
- 6 孙德梅, 等. 三维密度界面反演的一个近似方法. *物探与化探*. 1984, 8 (2): 89—98
- 7 石琳珂. 逐步缩小搜索范围的遗传算法. *地球物理学进展* (1995, 待刊).

THE STUDY OF POTENTIAL GENETIC INVERSION METHOD

Shi Linke

(Geophysical Prospecting Center SSB, Zhengzhou 450003)

Abstract

The problems in Parker—Oldenburg potential data inversion method has been analyzed, the Genetic inversion method which combining both fast forward method in frequency domain and Genetic Algorithms has been put forward in this paper. To counter the problem that convergence is slow in Genetic Algorithms, two improved measures were adopted, which made this Genetic inversion more perfection. After computing on theoretical models and using real data, the effect of this potential Genetic inversion method was verified.

Subject words: Fast forward method in frequency domain, Genetic Algorithms, potential Genetic inversion method, moho interface.