

粒子群算法在主震断层面参数估计中的应用^{*}

王福昌, 万永革, 胡顺田

(防灾科技学院, 河北三河 065201)

摘要: 提出了利用粒子群全局优化算法, 由余震震源参数确定主震断层面的方法, 从而克服了常见的 Gauss-Newton 法对初值的敏感性。同时, 讨论了 L_1 和 L_2 范数准则对参数估计的影响。将该方法用于 2003 年 7 月 21 日云南大姚 6.2 级地震断层面参数估计, 取得了较好的结果。

关键词: 粒子群算法; 参数估计; 非线性优化模型; 余震; 主震断层面

中图分类号: P315.3⁺3 文献标志码: A 文章编号: 1000-0666(2008)02-0149-06

0 引言

人类对地球内部物理性质和矿产资源分布的了解, 大多来自地表地质和地球物理、地球化学资料的反演和解释 (王家映, 2000)。在科学研究和工程技术中, 在由观测数据反推模型或体系时, 都无法回避反演问题。绝大多数地球物理反演问题都属于非线性问题。一部分非线性反演问题可以通过线性化方法迭代求解, 如 Newton 法、拟 Newton 法及其变形等; 另一类方法则是直接搜索的方法, 这类方法又分为非启发式和启发式方法。非启发式的反演方法, 如蒙特卡罗法 (Monte Carlo method), 计算量很大, 效率较低。近年来, 人们不断发展启发式的优化方法, 并用于地球物理反演问题中, 如遗传算法 (万永革等, 1997)、模拟退火算法 (曹小林等, 2000; 于鹏等, 2007)、量子退火算法 (魏超等, 2006)、差分演化算法 (李志伟等, 2006) 等。

粒子群算法 (Particle Swarm Optimization) (Kennedy 等, 1995; 谢晓锋等, 2003) 是近年来发展起来的一种新的、有效的现代启发式算法, 具有计算简便、收敛速度快、收敛准确等特点, 在最优化问题中有一定的应用, 但尚未见到用于地球物理反演问题的文献报道。本文中笔者首先介绍了粒子群算法的基本原理和步骤, 然后建立利用余震数据拟合主震断层面参数的范数准则下

的数学模型, 最后根据 2003 年 7 月 21 日云南大姚 6.2 级地震及其余震的双差定位法定位结果 (华卫等, 2006), 使用粒子群算法求解, 通过研究震源点在断层面附近的分布规律, 说明模型和算法的有效性。

1 粒子群优化算法描述

粒子群优化算法是一种基于群智能的演化计算技术, 它是 Kennedy 等 (1995) 受人工生命研究结果的启发, 于 1995 年首先提出来的。该算法与遗传算法类似, 它也是一种基于群体的优化工具, 但与遗传算法相比, 该算法具有收敛速度快, 容易实现, 而且又具有深刻智能背景的优点。

粒子群算法是根据对环境的适应度将群体中的个体移动到好的区域; 将每个个体看作 D 维搜索空间中的一个没有体积的粒子 (点), 所有的粒子都有一个由被优化的函数决定的适应值, 主要特点为: ① 每一粒子都被赋予了初始随机速度并在解空间中流动; ② 个体具有记忆功能; ③ 个体的进化主要依靠它本身的“飞行经验”以及同伴的“飞行经验”进行动态调整, 最终飞向最优位置, 即通过迭代找到最优解 (Parsopoulos 等, 2002)。

设在 D 维空间中有 m 个粒子, 第 i 个粒子表示为 $x_i = (x_{i1}, x_{i2}, \dots, x_{iD}) \in R^D$, 它所经历的最好位置 (对应最优的适应值) 为 $p_i = (p_{i1}, p_{i2}, \dots, p_{iD})$

* 收稿日期: 2007-07-30.

基金项目: 中国地震局地震科学联合基金 (A07015) 资助.

在群体中所有粒子所经过的最好位置的索引号用符号 S 表示, 即 P_g 粒子的速度 $V_i = (V_{i1}, V_{i2}, \dots, V_{id})$ 。对于每一代粒子, 其第 d ($1 \leq d \leq D$) 维第 i 个粒子的速度和位置根据如下方程进行变化 (Shi 等, 1998):

$$V_{id} = \omega \cdot V_{id} + \zeta \cdot \text{rand}_d() \cdot (P_{id} - X_{id}) + \xi \cdot \text{rand}_d() \cdot (P_{gd} - X_{id}); \quad (1)$$

$$X_{id} = X_{id} + V_{id}. \quad (2)$$

其中, $\text{rand}_d()$ 、 $\text{rand}_d()$ 是 $[0, 1]$ 内的均匀分布的相互独立随机数; ζ 和 ξ 为取值于 $[0, 2]$ 内的加速常数 (Suganthan, 1999), 通常解释为“自身学习率”和“社会学习率”; ω 为惯性权重因子, 当权值较大时, 粒子对未探测空间搜索能力强, 局部细调能力弱, 反之, 权值较小时, 算法的局部搜索能力增强, 而搜索新空间的能力减弱, 利用这样的特性, 在计算的初期选取较大的权值以加大搜索区域, 而在计算的后期, 为了更好地进行局部搜索, 选取较小的权值。Shi (2000) 提出了一个模糊系统, Chatterjee (2006) 提出一个规则来动态调节 ω , 但方法都过于复杂。一般认为, ω 可以在 $[W_{min}, W_{max}]$ 内线性递减取值, 即

$$\omega = \left(\frac{Iter_{max} - Iter}{Iter_{max}} \right) \cdot (W_{max} - W_{min}) + W_{min}. \quad (3)$$

其中 $Iter_{max}$ 表示最大迭代次数, 由所研究的问题确定, $Iter$ 表示当前迭代次数。

此外, 为了防止搜索发散 (爆炸), 当前粒子的加速导致它在某一维的速度超过了最大速率, 则粒子速度 V_{id} 会被一个最大速率 V_{max} 所控制, 即

$$V_{id} = \begin{cases} V_{max}, & V_{id} > V_{max}; \\ -V_{max}, & V_{id} < -V_{max}; \end{cases}$$

粒子群算法的基本步骤如下:

(a) 初始化设置: 种群规模为 N , 随机初始化它们的位置 X_i 和速度 V_i , 最大迭代次数为 $Iter_{max}$;

(b) 设置每个粒子目前最优位置为 $P_i = X_i$ 并记 $P_g = \max_i P_i$ 为全局最优位置;

(c) 计算每个粒子的适应度值, 即对每个粒

子计算其目标函数值;

(d) 对每个粒子, 将其适应度值与其历史最优位置进行比较, 更新 P_i ;

(e) 将每个粒子的当前最优位置 P_i 与群体历史最优位置 P_g 进行比较, 更新 P_g 并更新 P_g 的索引号;

(f) 根据粒子群速度和位置更新方程 (1) 和 (2) 来调整微粒的速度和位置;

(g) 检查终止条件通常为达到最大迭代次数或者足够好的适应值或者最优解停滞不再变化。若上述条件得以满足, 则终止迭代, 否则返回第 2 步。

2 非线性模型

地震震源机制解是描述地震的重要参数, 在地震学研究中占有相当重要的位置, 该参数又几乎是确定地壳深部应力场的唯一途径。目前地震的震源机制解往往通过 P 波初动 (Reasenber 等, 1985)、体波波形 (姚振兴等, 1985)、面波波形 (马淑田等, 1999) 和大地测量数据 (陈运泰等, 1979) 来测定。

大地震发生后, 大量余震在发震断层面及其附近发生。因此余震震源位置的空间分布可以较为精确地勾画出发震断层面的形状和位置。假定地震发震断层可以用一个平面来模拟, 且大多数余震发生在这个断层面的附近, 则可以通过余震震源位置参数来求解发震断层的走向、倾角及位置。早在 1992 年, 王鸣等 (1992) 就采用了这种原则, 可是他们采用的 Gauss-Newton 法对初始解的选取比较敏感。万永革等 (2008) 利用唐山地震序列余震资料双差定位法精确定位的结果, 采用模拟退火算法对唐山地震序列破裂面的几何参数进行求解, 虽然他们的算法不依赖于初始解, 但仍然采用 L_2 范数得到评价解的总残差。由于 L_2 范数准则在可用数据较少且分布较为散乱时, 受野值点的影响较大, 本文讨论分别采用 L_1 范数和 L_2 范数来评价解的总残差的结果。

本文给出根据余震位置确定主震断层面走向、倾角的粒子群算法。如果在地震之后能够快速确定余震的震源位置, 就可以给出地震断层面的一种约束。

下面给出根据余震震源位置参数来求解发震

断层参数的非线性模型。

设在地理坐标系中, (x_i, y_i, z_i) 表示第 i 个余震的震源位置, 地震断层面的走向为 φ , 倾角为 δ 到坐标原点距离为 ρ 则断层面法向量为 $(\sin\varphi \sin\delta, -\cos\varphi \sin\delta, \cos\delta)$ (万永革等, 2000)。于是断层面在地理坐标系中的方程 (王福昌等, 2007) 为

$$x_i \sin\varphi \sin\delta + y_i (-\cos\varphi) \sin\delta + z_i \cos\delta - \rho = 0 \quad (4)$$

所以震源点到平面的垂直残差为

$$d_i = x_i \sin\varphi \sin\delta + y_i (-\cos\varphi) \sin\delta + z_i \cos\delta - \rho \quad (5)$$

假定共有 n 个可以利用的余震数据样本, 这可以建立以所有余震到断层面垂直距离残差向量 $d = (d_1, d_2, \dots, d_n) \in R^n$ 的 L^p 范数为目标函数的优化问题, 即

$$E(\rho, \varphi, \delta) = \|d\|_p = \left[\sum_{i=1}^n |d_i|^p \right]^{1/p}, \quad p > 0 \quad (6)$$

我们的目标是找到一组符合地震学解释要求的值 $(\rho^*, \varphi^*, \delta^*)$, 使得 $E(\rho, \varphi, \delta)$ 在 L^p 范数准则下在此处取最小值。在式 (6) 中, $p=2$ 时, 该式即为王鸣等 (1992) 在研究中使用的最小二乘准则模型; $p=1$ 时, 可以削弱较大残差的影响 (Prueger等, 1988; Guittou等, 2003), 从而在可用数据较少且数据分布较为散乱时, 给出一个比最小二乘准则更稳健的参数估计值。

3 利用地震余震数据确定断层面参数

为了验证算法和准则的有效性, 我们利用 2003 年 7 月 21 日云南大姚 6.2 级地震余震双差定

位数据来确定主震断层面参数。使用双差定位数据, 在式 (6) 中, 求出一组值 $(\rho^*, \varphi^*, \delta^*)$, 使得 $E(\rho, \varphi, \delta)$ 在 L^p 范数准则下在此处取到最小值, 再根据地震学解释, 确定断层面的走向 φ 和倾角 δ 最后确定断层的区域。

由于粒子群算法不要求计算目标函数的导数, 因此无论 $p=1$ 或 $p=2$ 时, 粒子群算法都同样有效。在粒子群算法中, 种群大小设为 50 个, 粒子的速率上界不超过 20, 最大迭代次数 200, 自身学习率为 0.3, 社会学习率为 0.8, 最大惯性权重因子为 0.8, 最小惯性权重因子为 0。表 1 给出了 $p=1$ 或 $p=2$ 时, 不同样本情形下的反演结果。

表 1 不同样本和准则下的断层面参数估计

L^p 范数最小准则	走向 $\varphi/(^\circ)$	倾角 $\delta/(^\circ)$	样本数	L^p 范数	运行时间/s
L^1 范数	106.4	88.9	424	325.0	6.50
L^2 范数	103.7	84.6	424	655.9	7.47
L^1 范数	107.1	71.0	82	66.3	1.40
L^2 范数	103.6	63.4	82	91.7	1.69

根据哈佛大学给出的该地震的震源机制解结果, 断层面走向为 291° , 倾角为 84° 。我们取 424 个样本值, 在范数 1 的情况下得到的断层面走向为 106.4° , 倾角为 88.9° , 两者的走向相差 4.6° , 倾角相差 6.1° (根据 Ak 等 (1980) 关于走向和倾角的定义), 结果非常接近, 说明我们得到的结果是可靠的。

根据粒子群算法求出的参数, 可以确定断层面的几何形态。从图 1 到图 4 分别给出了 4 种情形下余震分别在水平面 (图 1a, 2a, 3a, 4a)、断层面 (图 1b, 2b, 3b, 4b) 和垂直于断层面的横断面 (图 1c, 2c, 3c, 4c) 上的投影, 及余震距断层面距离的分布 (图 1d, 2d, 3d, 4d)。其中, 空

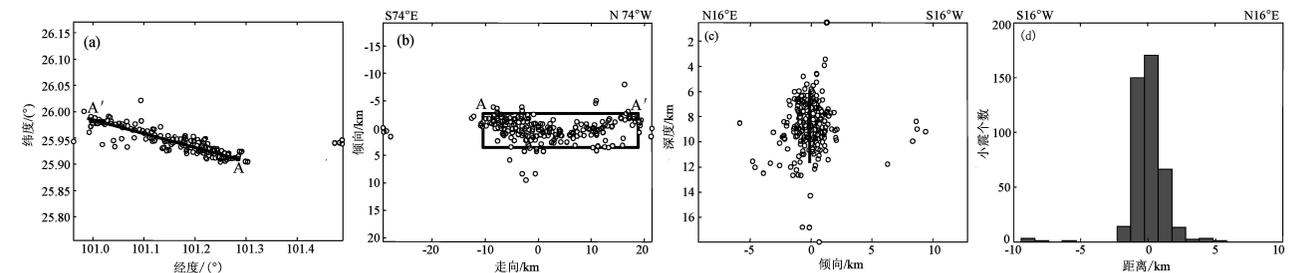


图 1 取 424 个样本时 L^1 范数最小准则结果

(a) 余震在水平面上的投影; (b) 余震在断面上的投影;

(c) 余震在垂直于断层面的横断面上的投影; (d) 余震距断层面距离的分布

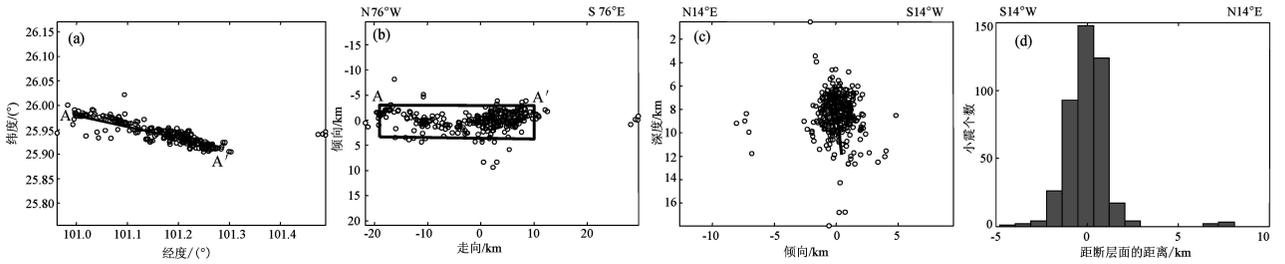


图 2 取 424 个样本时 $\frac{1}{2}$ 范数最小准则结果

(a) 余震在水平面上的投影; (b) 余震在断层面上的投影;
(c) 余震在垂直于断层面的横断面上的投影; (d) 余震距断层面距离的分布

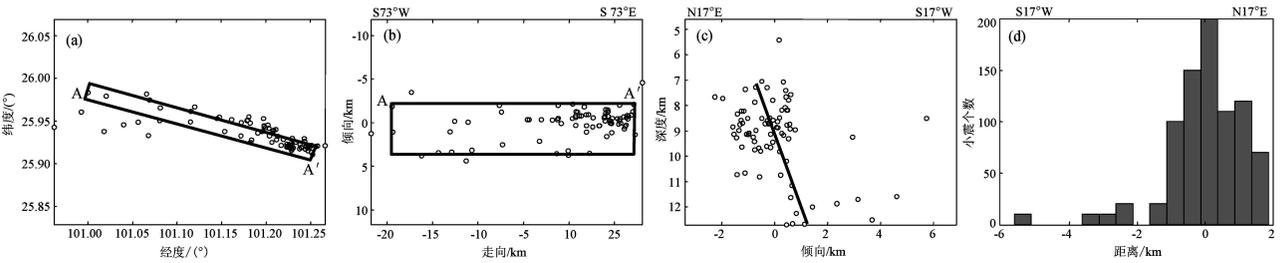


图 3 取 82 个样本时 $\frac{1}{2}$ 范数最小准则结果

(a) 余震在水平面上的投影; (b) 余震在断层面上的投影;
(c) 余震在垂直于断层面的横断面上的投影; (d) 余震距断层面距离的分布

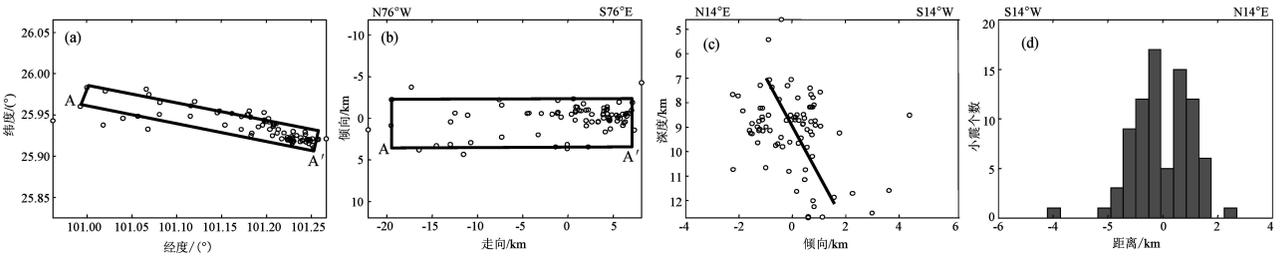


图 4 取 82 个样本时 $\frac{1}{2}$ 范数最小准则结果

(a) 余震在水平面上的投影; (b) 余震在断层面上的投影;
(c) 余震在垂直于断层面的横断面上的投影; (d) 余震距断层面距离的分布

心圆表示余震震源点的位置, 黑色线框所包围的区域表示由算法所确定的断层面区域, 直线段表示垂直于横断面的断层投影。

将图 1 和图 2、图 3、图 4 进行对比, 可以看出, $\frac{1}{2}$ 范数最小准则下求得的断层面受野值点数据的影响较小 (图 1c、3c), 余震距断层面距离的分布呈现宽尾巴的对称分布 (图 1d) 或者偏态分布 (图 3d), 而 $\frac{1}{2}$ 范数最小准则下求得的断层面受野值点数据的影响较大 (图 2c、4c), 余震距断层面距离的分布呈现近似正态分布 (图 2d、4d)。由于在取 $\frac{1}{2}$ 范数时, 野值点对应的残差求平方之后“夸大”了它对断层面的作用, 因此当数据样

本中野值点较多时, 使用 $\frac{1}{2}$ 范数更为合理。 $\frac{1}{2}$ 范数最小准则是一种比较稳健的方法, 以前由于它对应的目标函数不可导, 在数学上处理起来不方便, 而且由于当时计算机没有普及, 导致了它很少被使用。现在随着优化算法的发展和计算能力的提高, $\frac{1}{2}$ 范数最小准则的应用比较容易了, 而且人们还提出了其他的稳健估计准则, 如 Guittou 等 (2003) 使用的可以看成 $\frac{1}{2}$ 和 $\frac{1}{2}$ 范数混合的 Huber 范数最小准则等。

笔者的计算机配置为 Pentium N3.0G 的芯片, 512MB 内存, 计算软件为 MATLAB7.0 环境, 文中所有结果均在此环境下计算得出。

4 结论

粒子群算法的优势在于算法的简洁性, 它易于实现, 参数调整少, 无需目标函数的导数信息, 是一种有效的非线性全局优化算法, 具有很强的搜索能力和很高的准确性与稳定性。由于很多地球物理问题都是非线性的, 因此粒子群算法具有一定的适用性。尽管粒子群算法仍然是一种随机搜索算法, 无法估计参数的精度, 但是这种直接搜索的算法也不存在如 Gauss-Newton 迭代法过程中的初始点选择、病态矩阵处理和误差积累等问题。因此对于许多非线性、多极值的非线性反演问题, 采用这种简洁而又易于实现的启发式随机优化算法求解是一种不错的选择。

参考文献:

曹小林, 洪学海, 曹俊兴. 2000. 面波波形反演中的模拟退火法 [J]. 成都理工学院学报, 47 (3): 296—301.

陈运泰, 黄立人, 林邦慧, 等. 1979. 用大地测量资料反演的 1976 年唐山地震的位错模式 [J]. 地球物理学报, 22 (3): 201—217.

华卫, 刘杰, 郑斯华, 等. 2006. 2003 年云南大姚 6.2、6.1 级地震序列特征分析及地震触发研究 [J]. 中国地震, 22 (1): 10—23.

李志伟, 胥颐, 郝天珧, 等. 2006. 利用 DE 算法反演地壳速度模型和地震定位 [J]. 地球物理学进展, 21 (2): 370—378.

马淑田, 姚振兴, 纪晨. 1999. 用长周期面波波形拟合及 P 波初动方向估计中等地震的震源机制 [J]. 地球物理学报, 42 (6): 785—799.

万永革, 刘瑞丰, 李鸿吉. 1997. 用遗传算法反演京津唐张地区的三维地壳结构和震源位置 [J]. 地震学报, 19 (6): 623—633.

万永革, 吴忠良, 周公威, 等. 2000. 根据震源的两个节面的走向角和倾角求滑动角 [J]. 地震地磁观测技术, 21 (5): 26—30.

万永革, 沈正康, 刁桂苓, 等. 2008. 利用小震分布和区域应力场确定大震断层参数方法研究及其在唐山地震序列中的应用 [J]. 地球物理学报, 51 (3) (待刊).

王福昌, 万永革. 2007. 利用余震震源分布确定主震断层面的方法研究 [J]. 内陆地震, 21 (1): 39—43.

魏超, 朱培民, 王家映. 2006. 量子退火的反演和实现 [J]. 地球物理学报, 49 (2): 577—583.

王家映. 2000. 地球物理反演理论 [M]. 北京: 高等教育出版社.

王鸣, 王培德. 1992. 1989 年 10 月 18 日大同—阳高地震的震源机制和发震构造 [J]. 地震学报, 14 (4): 407—415.

谢晓锋, 张文俊, 杨之廉. 2003. 微粒群算法综述 [J]. 控制与决策, 18 (2): 129—134.

姚振兴, Hembeger D V. 1985. 测定断层面解的地震波形反演方法 [J]. 地震, 5 (3): 46—53.

于鹏, 王家林, 吴健生, 等. 2007. 重力与地震资料的模拟退火约束联合反演 [J]. 地球物理学报, 50 (2): 529—538.

Aki K, Richards P G. 1980. Quantitative seismology Volume I [M]. San Francisco: W. H. Freeman and company press.

Chatterjee A, Sianty 2006. Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization [J]. Computers & Operations Research, 33 (3): 859—871.

Guignon A, Symes W. 2003. Robust inversion of seismic data using the Huber norm [J]. Geophysics, 68 (4): 1310—1319.

Kennedy J, Eberhart R C. 1995. Particle Swarm Optimization [C] // Hoboken: Proceedings of the IEEE International Conference on Neural Networks, Piscataway: IEEE Press, 1942—1948.

Parsopoulos K E, Vrahatis M N. 2002. Recent approaches to global optimization problems through particle swarm optimization [J]. Natural Computing, (1): 235—306.

Prügger A, Gendzwil D J. 1988. Microearthquake location: A nonlinear approach that makes use of simplex stepping procedure [J]. BSA, 78 (2): 799—815.

Reasenbergs P, Oppenheimer D. 1985. FPFIT, FPLOT, FPPAGE, Fortran computer programs for calculating and displaying earthquake fault plane solutions [R]. Menlo Park: United States Geological Survey, 85—739.

Shi Y, Eberhart R. 1998. A modified particle swarm optimizer [C] // Anchorage: International Conference on Evolutionary Computation, Piscataway: IEEE Press, 69—73.

Shi Y, Eberhart R. 2000. Fuzzy adaptive particle swarm optimization [C] // Hoboken: Proceedings IEEE International Conference on Evolutionary Computation, Piscataway: IEEE Press, 84—88.

Stegun P N. 1999. Particle swarm optimizer with neighborhood operator [C] // Tejsseyre: Proceedings of the Congress Evolutionary Computation, Piscataway: IEEE Press, 1951—1957.

Application of Particle Swarm Optimization to the Estimation of Main Shock Fault Plane Parameters

WANG Fu.chang, WAN Yong.ge, HU Shun.tian

(Institute of Disaster Prevention Science and Technology, Shanhe 065201, Hebei, China)

Abstract

A new method to determine the fault plane of main shock by using Particle Swarm Global Optimization Algorithm was proposed, which overcomes the sensitivity of Gauss-Newton Method to initial value. The effect of L_1 - and L_2 -criteria on parameter estimation was also discussed. This method was used to evaluate the fault plane parameters of the Dayao, Yunnan earthquake on July 21, 2003, and the results show that the method is valid.

Key words: Particle Swarm Optimization, parameter estimation, Nonlinear Optimization Model, after shock, fault plane of main shock



王福昌 中国地震局防灾科技学院副教授。1997年毕业于曲阜师范大学数学系,获学士学位;2000年毕业于大连理工大学应用数学系应用数学专业,获硕士学位。主要从事多元统计分析和现代最优化方法的应用研究。



郭晓 中国地震局兰州地震研究所助理研究员。2005年毕业于中国地震局兰州地震研究所固体地球物理专业,获硕士学位。主要从事卫星遥感应用及地震预报研究工作。



赖加成 福建省地震局泉州地磁台助理工程师。1994年毕业于防灾技术高等专科学校(现为防灾技术学院)。主要从事地震地磁监测工作。



曾宪伟 中国地震局兰州地震研究所在读硕士研究生。2005年毕业于青岛大学物理系,获学士学位。主要从事数字化观测数据的处理和分析工作。



莘海亮 中国地震局兰州地震研究所在读硕士研究生。2005年毕业于河北大学物理科学与技术学院,获学士学位。研究方向为地震精定位与地壳三维速度结构。

注:陈立德、刘翔、叶建庆、常祖峰的个人简介已分别刊登在本刊的 Vol.30 No.1, Vol.28 No.4, Vol.30 No.3, Vol.29 No.2