

遗传算法及其在地震定位中的应用

GENETIC ALGORITHMS METHOD AND ITS APPLICATION IN EARTHQUAKE LOCATION

栾锡武*

(中国科学院海洋研究所 青岛 266071)

* 传统的搜索和最优化方法可以分为:微分为基础的方法、枚举法和蒙特卡罗法等三类随机算法,它们在实际应用所要求的条件、搜索效率、初始模型对搜索全局最优值影响等方面各有优缺点。遗传算法是模拟生物的优胜劣汰遗传变异而新发展起来的一种优化算法,近

一两年来国内外学者开始探索它在地球物理学特别是在地震科学中的应用,取得了较为满意的结果^[1]。

* 中国科学院海洋研究所现代海底热液活动研究青年实验室工作人员。
收稿日期:1996年4月22日

1 遗传算法

遗传算法首先将问题的待求解用二进制数来表示,并将所有这些二进制数串联在一起形成一个由 01 组成的字符串,将一个二进制数看作一个基因,将一个字符串看作一个染色体,一个染色体即为一个个体,问题的解包含在一个字符串中,随机地生成 N 个字符串,即为各包含有不同染色体的个体, N 个个体生成一个种群。遗传算法要求一个衡量解优劣的目标函数,在种群的演化过程中,其中的个体根据目标函数的大小进行淘汰、保留。人为地保持种群的大小不变,保留下来的个体通过基因的遗传交配而形成新的种群,如此反复,最终得到最优的种群。这个种群中字符串所代表的参数即为问题的解。

下面通过一个简单例子,即求 $y = 1\,000 - (x - 2)^2$ 在 $0 < x < 32$ 区间的极大值,来说明遗传算法的交配、繁殖、变异过程。

首先将 x 表示为一个二进制数,由 4 个二进制数构成一个种群,第一个种群随机生成,应该指出,初始模型集是均匀分布于模型空间中的。由初始模型计算 y 值和生存概率,用轮盘赌的方法从期待概率得到父种群中的个体在子种群中出现的数目,由表 1 可见,只有生存概率大的个体才得以生存并繁殖,差的个体如 11101 则被淘汰。

新一代染色体按一定的概率随机地选择两两一组的双亲,分别随机地交换部分染色体,如表 2 中 1, 2 号染色体在第四位交换基因, 3, 4 号在第二位交换基因,产生新一代种群。

本例正好在第 3 步找到最佳解, $x = 2$ 。

表 1 繁殖前后的种群状况

编号	染色体	x	$y = f(x)$	生存概率	期待数目	实际数目	新一代
1	01100	12	900	0.32	1.32	1	01100
2	00110	6	984	0.36	1.45	2	00110
3	10111	23	559	0.21	0.82	1	10111
4	11101	29	271	0.10	0.40	0	00110

表 2 交配前后的染色体状况

编号	交配前的染色体	交配后的染色体	变异后的染色体
1	01 100	01110 ¹⁾	01010 ¹⁾
2	00 110	00100	00100
3	0011 0	00111	00111
4	1011 1	10110	10110

1) 染色体按一定的概率发生变异,不断重复以上步骤。

2 目标函数的选取和选择淘汰的原则

如何选择目标函数最佳地判断解的好坏,是一个非常重要的问题,一般目标函数有很多峰值,而在峰值处解的可能性很大,所以应选择这样的目标函数:使其在真实解附近有较大的峰值,而且在峰值附近比较光滑,又使它的极值相对较小,这样能避免解的不稳定性并使其尽快收敛,同时保证不会陷入局部极大值中^[2],在最优化问题中,观测值和模型正演计算值是比较的基础。设有 N 组观测数据,则一般目标函数取为二者差值的平方和,目标函数越小的个体是较优的个体。

计算时人为地保持种群的大小不变,种群中的个体依照其生存概率的大小决定在下一代中被选择或淘汰,在选择淘汰的原则时也有多种可能。一种是轮盘赌的方法:按照各个个体生存概率的比例分配绘制一个理想的赌盘,旋转它并根据其停止的位置确定选取的模型,直到选出需要的种群为止。一般利用这种方法优秀的个体会被保留下来并繁殖后代,劣等的个体会被淘汰。

3 地震时空参数的反演

除了专门研究震源机制和震源物理等问题外,在一般的分析工作中,都把震源看成是一个几何点,并认为地震是在某个瞬间发生的。这就是“点源模型”。正是由于自觉或不自觉地采用了这个模型,才能去测定地震的空间位置 and 发震时间(总称为地震的时空参数),即进行地震定位^[3]。

设地震的时空参数为 x, y, h, o 。对于均匀介质,直达波 P 的走时方程是一个非线性方程

$$(X - x_k)^2 + (Y - y_k)^2 + h^2 = (P_k - o)^2 V^2 \quad (1)$$
$$k = 1, 2, \dots, m; \text{一般地 } m \geq 5$$

其中 V 为地震诱发的直达波 P 的波速即 V_P ; x_k, y_k 为台站的位置; P_k 为第 k 个台站的 P 波到时。

表 3 第 2 次繁殖前后的种群状况

编号	染色体	x	$y = f(x)$	生存概率	期待数目	实际数目	新一代
1	01010	10	936	0.27	1.07	1	01010
2	00100	4	996	0.28	1.13	1	00100
3	00111	7	975	0.28	1.11	1	00111
4	10110	22	600	0.17	0.69	1	10110

表 4 第 2 次交配前后的染色体状况

编号	交配前的染色体	交配后的染色体	变异后的染色体
1	01 010	01100	01100
2	00 100	00010	00010
3	0 0111	00110	00110
4	1 0110	10111	10111

表 5 第 3 次繁殖前后的种群状况

编号	染色体	x	$y = f(x)$
1	01100	12	900
2	00010	2	1000
3	00110	6	984
4	10111	23	559

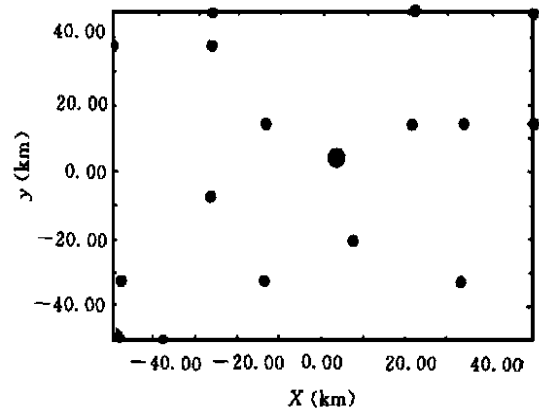


图 1 台站及震中位置

通过修定方法或将非线性方程组线性化的方法可以对方程(1)求解, 现在用遗传算法来探讨这一问题。其中 x_k, y_k, P_k 是实测资料, X, Y, h, O, V 为待反演的参数, 采用二进制码进行编码, 种群的大小定为 32, 交换概率和变异概率分别定为 0.9, 0.02, 取标准误差 δ_p 作目标函数

$$\delta_p = \frac{\sum (P'_k - P_k)^2}{m - 5} \quad (2)$$

式中 P_k 是观测到时, P'_k 是正演到时

$$P'_k = \frac{(X - x_k)^2 + (Y - y_k)^2 + h^2}{V} + O \quad (3)$$

$P'_k - P_k$ 即为到时残差。

为检验方法的可行性, 在一个平面坐标系中位置为 (2.5 km, 3.5 km), 深 4.5 km 处, O 时刻设计了一次地震, 布设 16 个台站进行观测, 台站的位置如图 1, 设 P 波速度为 1, 在此情况下用遗传算法程序进行计算, 结果很快就得到预期的结果。图 2 为解的分布情况。

4 讨论

遗传算法采用一种自适应演化过程的方法来逼近问题的解答, 效率较高, 它最适于解决复杂问题。当然, 地震定位问题远不止如此简单, 本文只是给出了一个简单的例子。但可以看出它在地震定位等地球物理学领域中有着广阔的应用前景。

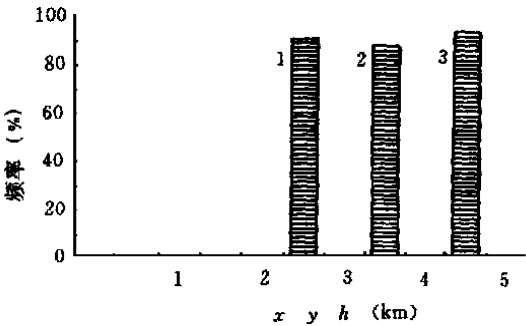


图 2 地震的空间参数初次搜索后的分布直方图 (1)震中的 X 坐标; (2)震中的 Y 坐标; (3)震深(单位 km), 搜索次数为 100 次

参考文献

[1] 石耀霖, 1993. 地球物理学会年刊. 地震出版社, 58.
[2] 陶春辉, 1993. 长春地质学院学报 23(4): 447~ 454.
[3] 张少泉, 1977. 近震分析. 地震出版社, 1~ .