文章编号：1000－8152(2001)05－0808－03

一种新的非线性回归模型参数估计算法*

陈金山 韦 岗
(华南理工大学电子与信息学院·广州，510640)

摘要：提出一种新的基于混合基因算法(HGA)的非线性回归模型参数估计算法.新算法通过对问题的解空间交替进行全局和局部搜索、达到快速收敛至全局最优解、较好地解决了传统算法通用性差、易陷入局部极小的问题.实验验证了算法的通用性和有效性.

关键词：基因算法；非线性参数估计；最小二乘估计
文献标识码：A

A New Algorithm for Estimation of Parameter of Nonlinear Regression Modeling

CHEN Jinshan and WEI Gang
(College of Electronics and Information, South China University of Technology · Guangzhou, 510640, P.R. China)

Abstract: A new algorithm for parameter estimation of nonlinear regression modeling based on hybrid genetic algorithm (HGA) is proposed. The new algorithm can rapidly converge to the global optima by performing global search and local search alternatively, and achieve better performance than those of traditional algorithms. The experiments verify generalization and effectiveness of the algorithm.

Key words: genetic algorithm; nonlinear parameter estimation; least-squares estimation

1 引言(Introduction)

非线性回归分析在语音信号处理、生物医学和地震预测等方面得到了广泛的应用.然而，由于非线性模型一般都比较复杂，且不容易获得其参数估计，从而极大地限制了它的应用和发展.非线性回归模型参数估计的常见算法[1~4]有：直接搜索法、Hooke-Jeeves 法、Nelder-Mead 法、Gauss-Newton 法、变尺度法和同伦算法等.这些算法通常只对某一类特定问题有效，且要求模型具有连续、可导、单峰等特性.文[5]提出了基于数论的序贯法，此方法不需要导数等辅助信息、但其原则上仍是一种"爬山"搜索法，而且算法的效率很大程度地受限于初始邻域的选取及每一次迭代时邻域加细的精度.基因算法(GA)是一种全局优化算法[6]，但它存在收敛慢的问题.因此，将 GA 与一些传统的寻优方法(如爬山法、模拟退火法、牛顿法等)结合起来，可在保持算法通用性的基础上提高算法的效率.Render[7]等人提出了一种将拟牛顿法与传统 GA 相结合的算法用于非线性函数优化，提高了算法的收敛速度.但由于拟牛顿法需求函数的一阶导数、因而该方法的通用性受到一定的限制.本文提出一种新的基于混合基因算法(HGA)的非线性回归模型参数估计算法，新算法只利用函数值进行搜索，因而适用范围更广.

2 非线性回归模型及其最小二乘估计(Nonlinear regression modeling and least-squares estimation)

非线性回归的一般模型可表示为 $Y = f(X,\theta) + e$，其中 e 服从正态分布 $N(0,\sigma^2)$、$X \in \mathbb{R}^m$，$Y \in \mathbb{R}$，$\theta \in \mathbb{R}^P$，P 为参数个数.

设已知观测值 $\{(x_i, y_i); i = 1,2,\cdots,n\}$，非线性回归模型参数估计的问题就是如何求出 θ 的最小二乘估计，即求 $\hat{\theta} \in \mathbb{R}^P$ 使得对任何 $\theta \in \mathbb{R}^P$ 都有 $S(\hat{\theta}) \leqslant S(\theta)$.

$$S(\theta) = \sum_{i=1}^{n} [y_i - f(x_i,\theta)]^2. \qquad (1)$$

有关最小二乘估计量的存在性，Jennrich[8]给出

* 基金项目：国家自然科学基金重大项目(69896246)；霍英东青年教师基金及国家自然科学基金(69772027)资助项目.
收稿日期：2000－01－03；收修改稿日期：2000－07－04.

了如下定理：

定理 1　假设 Θ 为 $\mathbb{R}^p$ 上的紧子集，$f(X,\theta)$ 关于 θ 在 Θ 上连续，则必存在 $\mathbb{R}^p$ 上的可测函数 $\theta = \hat{\theta}(Y)$ 使得

$$\| Y - f(X,\hat{\theta}(Y)) \|^2 = \inf_{\theta\in\Theta} \| Y - f(X,\theta) \|^2.$$

本文以渐近回归模型(2)为例，讨论非线性回归模型在最小二乘意义下的参数估计.实际上，我们的方法同样适用于其它非线性回归模型.

渐近回归模型

$$f(x) = \alpha - \beta\gamma^x,\ 0 < \gamma < 1. \qquad (2)$$

3　算法描述(Description of algorithm)

基于混合基因算法的非线性回归模型参数估计算法的关键问题及其实现方法：

3.1　编码(Coding)

编码的实质是在问题的解空间与算法的搜索空间之间建立一个映射.采用实数编码时，每个个体用一个 n 维的实向量表示，即 $X = (x_1, x_2, \cdots, x_n)^T$；$a_i \leqslant x_i \leqslant b_i (i = 1,2,\cdots,n)$.定义种群中的任一个体 x_{ji} = IND$[j]$.chrom$[i] \times (b_i - a_i) + a_i; j = 1,2,\cdots,N$.那么，满足此关系的向量 $X_j = (x_{j1}, x_{j2}, \cdots, x_{jn})^T$ 都能满足优化问题的边界约束条件.

3.2　个体评价(Individual evaluation)

根据模型(2)的非线性参数估计的优化要求，建立个体适应值函数：$F(X,t) = [S(\theta,t) + \varepsilon]^{-1}$，$\varepsilon$ 为足够小的正数.对种群中的个体，计算并记录每个个体的适应值，选取最优个体 X_H 和最差个体 X_L.

3.3　选种(Selection)

为了避免算法出现过早收敛和停滞现象，我们选择了线性排名选择策略：首先假设种群成员按适应值大小从好到坏排列为 $X_1, X_2, \cdots, X_N$，然后根据一个线性函数分配选择概率 P_k.

$$P(X_k) =$$

$$P_k = [c - d * k/(N+1)]/N, k = 1,2,\cdots,N. \qquad (3)$$

其中 c, d 为常数，c 和 d 的取值应满足：①$\sum_{k=1}^{N} P_k = 1$；②对任意 $k = 1,2,\cdots,N$ 有 $P_k \geqslant 0$，且 $P_1 \geqslant P_2 \geqslant \cdots \geqslant P_N$.易知，取 $1 \leqslant c \leqslant 2$ 和 $d = 2(c-1)$ 可满足上述要求.

3.4　改造(Transform)

对最差个体进行如下改造：

1) 计算交配池种群的形心 X_C.

2) 对 X_L 进行反射：$X_R = X_C + r(X_C - X_L)$，r 为反射系数.

3) 若 $F(X_R,t) > F(X_L,t)$，则执行扩大操作：$X_E = X_C + q(X_R - X_C)$，q 为扩大系数.

4) 若对多边形中除 X_L 外的任何一点 X_W，均有 $F(X_R,t) < F(X_W,t)$，则执行收缩操作：$X_S = X_C + s(X_H - X_C)$，s 为收缩系数.

5) 若 $F(X_R,t) < F(X_L,t)$，则使所有点向 X_H 点靠近：$X_W = X_H + 0.5(X_W - X_H)$.

6) 令 X_R, X_E 和 X_S 中最好的点代替 X_L.

3.5　杂交(Crossover)

为了维持种群的多样性和避免算法过早收敛，我们采用近邻配对原则[9].这种配对方法不仅可避免较优模式过快地扩散，而且符合基因算法细粒度并行模型的要求，易于获得较大的并行度.

假设 $P(X) > P_c$(P_c 为杂交概率)按近邻配对原则选出的两个父代个体为 Parent[1]和 Parent[2]，其子代个体为 Child[1]和 Child[2]，采用整体算术杂交：先生成 n 个(0,1)区间的随机数 $a_1, a_2, \cdots, a_n$，则

Child[1].chrom$[i]$ =

a_i * Parent[1].chrom$[i]$+(1−a_i) * Parent[2].chrom$[i]$,

Child[2].chrom$[i]$ =

a_i * Parent[2].chrom$[i]$+(1−a_i) * Parent[1].chrom$[i]$,

其中 $i = 1,2,\cdots,n$.

3.6　变异(Mutation)

根据式(3)计算种群中每个个体的入选概率 $P(X_k)$，对于 $P(X_k) < P_m$(P_m 为变异概率)的个体 Parent$[j]$，随机选择其某一位基因 i 进行突变：

Child$[j]$.chrom$[i]$ = Parent$[j]$.chrom$[i]$ + δ.

其中 $\delta = \lambda * N(0,1)$，$N(0,1)$ 为标准正态随机变量，$\lambda = C_M/t$，C_M 为一常数，t 为迭代的代数.

3.7　算法终止准则(Terminal criterion of algorithm)

当进化代数达到最大进化代数 G 或结果在 50 代内没有明显的改进时算法终止，即 $|S(\theta,t) - S(\theta,t+50)| \leqslant \varepsilon$($\varepsilon$ 为一小的正数)或 $t = G$ 时，算法终止.

4　计算实例及结果比较(Examples and results)

计算实例是估计渐近回归模型(2)的参数 $\theta = (\alpha,\beta,\gamma)$，其中 $0 < \gamma < 1$.混合基因算法的参数取值如下：种群规模 $N = 50$；杂交概率 $P_c = 0.6$；变异概率 $P_m = 0.03$；最大进化代数 $G = 2000$.实验算法用 Borland C++3.1 语言编写.

实验数据集如表 1 所示.表 2 给出用 HGA、文[5]和文[10]的方法进行非线性回归参数估计的结

果.这里列出的是进行 10 次随机计算所得到的平均结果.

表 1 实验数据集

Table 1 Datasets for experiment

	数据集 1						数据集 2				
x	2	3	4	5	6	7	0	1	3	5	7
y	18.6	22.6	25.1	27.2	29.1	30.1	20.518	21.138	21.734	22.218	22.286

表 2 估计结果

Table 2 Results of estimation

		HGA 算法	文[5]中序贯法	文[10]中 Gauss-Newton 法
数据集 1	α	33.7909611	33.8022682	33.8023
	β	26.6905654	26.6979828	26.698
	γ	0.7527665	0.7529937	0.752995
	$\hat{\sigma}^2$	0.0349082	0.03491553	0.03491537
数据集 2	α	22.4855205	22.487058	22.487
	β	1.9574835	1.958616	1.9586
	γ	0.704972	0.705544	0.705539
	$\hat{\sigma}^2$	0.00587683	0.005876825	0.005876535

表 2 中 $\hat{\sigma}^2$ 为残差方差估计,定义为

$$\hat{\sigma}^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n-3}.$$

从表 2 可见,对于给定的两组数据,HGA 估计的结果与文献[5]、[10]的估计结果非常接近.这说明本文算法的有效性.

考虑到三种算法每次迭代所需要的计算量差异很大,我们选用算法收敛时间来衡量三种算法的效率.当方差估计 $S(\hat{\theta})$ 小于给定的门限时,算法终止.对三种算法设定同一门限,其收敛时间列于表 3 中.

表 3 估计时间

Table 3 Time of estimation (秒)

	HGA 算法	文[5]中序贯法	文[10]中 Gauss-Newton 法
数据集 1	4.8	7.5	6.3
数据集 2	4.8	5.7	6.3

从表 3 可见,HGA 算法估计的效率最高.序贯法的估计效率与初始邻域的选取及每一次迭代时邻域加细的精度密切相关.Gauss-Newton 法每一次迭代时都要计算偏导数矩阵,因而其收敛时间较长.

5 结论(Conclusion)

由于具有描述简单、通用、较少受初始条件的限制以及全局最优等优点,混合基因算法应用于非线性回归模型的参数估计,克服了传统算法对模型的限制条件较强、通用性较差的缺点,表现出了良好的应用前景.

参考文献(References)

[1] Nash J C and Walker Smith M. Nonlinear Parameter Estimation [M]. New York: Marcel Dekker Inc., 1987

[2] Wei Gang, Man K F and Kwong S. Homotopy theory based nonlinear modeling for time series [J]. Chinese Journal of Electronics, 1997, 6(3): 41 – 45

[3] Ma Ni and Wei Gang. Research on nonlinear modeling of time series [J]. Journal of Electronics, 1999, 16(3): 200 – 207

[4] Wei Gang, Zhang Luqing, Li Xiangwu and Ouyang Jingzheng. Homotopy nonlinear modeling for speech signals [J]. Acta Automatica Sinica, 1997, 23(2): 201 – 206(in Chinese)

[5] Fang Katai and Zhang Jinting. A new algorithm for calculation of estimation of parameters of nonlinear regression modeling [J]. Acta Mathematicae Applicatae Sinica, 1993, 16(3): 366 – 377(in Chinese)

[6] He Qianhua, Wei Gang and Lu Yiqin. Review of genetic algorithms [J]. Acta Electronica Sinica, 1998, 26(10): 118 – 122(in Chinese)

[7] Render J M and Flasse S P. Hybrid methods using genetic algorithms for global optimization [J]. IEEE Transactions on System, Man, and Cybernetics (Part B), 1996, 26(2): 243 – 258

[8] Jennrich R I. Asymptotic properties of nonlinear least squares estimators [J]. Ann. Math. Statist, 1969, 40: 633 – 643

[9] Zhang Qingfu, Peng Wei and Wu Shaoyan, et al. Genetic algorithm + orthogonal design: a new global optimization algorithm [A]. Proceedings of 4th Artificial Intelligence of China [C], Beijing: Tsinghua University Press, 1996, 127 – 133(in Chinese)

[10] Ratkowky D A. Nonlinear Regression Modeling – A unified practical approach [M]. New York: Marcel Dekker Inc., 1983

本文作者简介

陈金山 1963 年生.博士.主要研究方向有神经网络,信号处理理论与 ATM 交换技术.

韦 岗 见本刊 2001 年第 4 期第 486 页.