

## 混沌系统中参数估计的演化建模方法\*

王柳<sup>1)2)</sup> 何文平<sup>2)†</sup> 万仕全<sup>3)</sup> 廖乐健<sup>1)</sup> 何涛<sup>4)</sup>

1)(北京理工大学计算机学院, 北京 100081)

2)(国家气候中心, 北京 100081)

3)(扬州市气象局, 扬州 225009)

4)(常州市环境监测中心站, 常州 213000)

(2013年5月15日收到; 2013年9月18日收到修改稿)

借助于演化算法的自组织、自适应和自学习特征, 本文提出了基于演化算法的参数辨识方案, 并利用经典的 Lorenz 方程进行了数值仿真试验, 研究了参数辨识方案对于单参数、双参数以及 Lorenz 系统三个参数完全未知时的性能. 数值试验结果表明, 新方法能够很好的对未知参数进行较为快速、准确的辨识, 但存在对多个参数同时搜索时速度较慢的缺陷. 鉴于此, 将演化算法变异操作中的参数变异范围附加一种约束机理, 试验结果表明, 这一约束机理有效地提高了多参数估计中算法的收敛速度.

**关键词:** 混沌系统, 参数估计, 演化算法

**PACS:** 92.60.Wc

**DOI:** 10.7498/aps.63.019203

## 1 引言

参数估计是动力系统研究中一个典型的反问题<sup>[1]</sup>. 在大气科学中的许多动力学模式中, 由于模式中离散格点尺度较粗糙, 许多时空尺度较小的物理过程通常无法被直接准确的模拟, 如云的微物理过程(包括云的形成), 因此, 许多与云相关的物理现象在模式中常常采取参数化的形式<sup>[2]</sup>, 与此类类似的还有边界层参数化、陆面过程中的参数化等, 这些参数值的选取对于数值模式性能有着直接的影响, 而通过理论计算或者是观测资料得到的大部分参数都存在着很大的不确定性. 又如海气振子系统中的许多重要物理参数都是随时间变化和可调的, 如何根据某个时期的历史观测数据合理确定厄尔尼诺和南方涛动模型中的未知参数是一个重要问题, 这对于认识其动力学机理具有重要的意义<sup>[3]</sup>. 在化学反应、流体力学和电路等实际系统中,

通常已知系统的动力学方程, 但某些参数不可测或是难以测量的, 而这些参数对深入理解系统的动力学特性和对系统实施适当的控制具有重要作用. 混沌控制和混沌同步已广泛应用于保密通信、信息科学、生命科学等领域, 估计这类非线性动力学系统的参数是一个具有重要现实意义的科学问题. 但由于混沌系统的复杂性或出于保密的需要, 系统的某些参数无法给出<sup>[4,5]</sup>. 因此, 对非线性动力学方程中参数进行有效估计是一个非常重要的研究课题.

近年来, 针对非线性系统参数估计问题, 已有学者先后提出了多种方法. 李丽香等<sup>[4,5]</sup>提出了混沌蚂蚁群算法, 但其算法的收敛速度与蚁群系统的组织因子有很大关系; 关新平等<sup>[6]</sup>提出了未知参数辨识观测器的概念; 戴栋等<sup>[7]</sup>基于遗传算法提出了一种混沌系统参数估计的方法, 但其收敛速度较慢; 曹小群等将变分原理用于混沌系统的参数估计

\* 全球变化研究国家重大科学研究计划(批准号: 2012CB955902和2013CB329605)、国家自然科学基金(批准号: 41275074, 41005041和41175067)和公益性行业(气象)科研专项(批准号: GYHY201106015和GYHY201106016)资助的课题.

† 通讯作者. E-mail: wenping\_he@163.com

得到了较好的结果<sup>[8]</sup>. 值得指出的是, 基于变分方法的误差辨识技术对于简单非线性模型比较有效, 但对于更为复杂的非线性系统, 该算法会出现临界变分的现象. Li等<sup>[9]</sup>和Li等<sup>[10]</sup>也先后提出了混沌系统的参数估计方法. 诚然, 任何方法都不是万能的, 需要针对特定的参数辨识问题, 发展不同的参数辨识方法.

演化算法是模拟自然界的生物演化过程, 借鉴生物界的自然选择和自然遗传机理而发展起来的一类求解问题的策略和方法<sup>[11]</sup>. 由于它本身具有的自组织、自适应和自学习等智能特征, 目前已被成功应用于机器学习、模式识别、经济预测以及气候预测模式的误差订正等领域.

借助于演化算法的自组织、自适应和自学习特征, 本文首先提出了基于演化算法的参数辨识方案, 随后利用一个典型的非线性系统—经典的Lorenz方程<sup>[12,13]</sup>进行了数值仿真试验, 研究了参数辨识方案对于单参数、双参数以及Lorenz系统三个参数完全未知时的性能. 数值试验结果表明, 无论是Lorenz系统部分参数未知, 还是其参数全部未知, 基于演化算法的参数辨识技术均能够对其进行快速、准确的识别. 但在数值试验的过程中发现, 当同时对多个参数进行估计时, 参数估计算法存在收敛时间长、搜索速度较慢的缺陷, 针对此问题, 本文提出了相应的改进措施: 在演化程序的变异操作中, 对参数的变异范围采取了一种约束机理, 即演化时间越长, 参数的变异范围越小. 试验结果表明, 改进方案极大地提高了演化算法搜索参数的速度. 由于演化算法采用群体的方式组织搜索, 使得它可以同时在解空间内不同区域进行搜索, 特别适用于大规模并行计算. 它不受搜索空间的限制条件, 如是否连续、可微分、峰值等, 也不需要其他辅助信息, 这使得该算法具有搜索效率高、简单且易于操作、程序通用性较强等特点.

## 2 参数辨识方法

### 2.1 演化建模算法

演化算法的主要思想是采用遗传程序设计优化模型结构, 以遗传算法优化模型的参数, 通过与模型的预处理(如通过滤波消除原始数据的部分噪声)、模型的简洁化和规范化、模型的预测等辅助步骤相结合, 根据观测数据建立常微分方程. 演化建模算法的一般流程是, 首先对观测数据进行预处理,

如滤波去噪、数据缺测插值等; 其次是初始化群体, 然后是对这些群体进行演化操作, 包括群体间的杂交以及个体的遗传和变异等, 在此基础上再对各个群体根据一定的评价函数来进行客观、定量的评价, 从而能够选择和生成下一代种群, 如此, 则一次完整的种群演化结束. 如果演化建模过程中某一代群体中输出的最优个体或者某几个个体满足了预设的建模停机条件, 则演化程序过程终止.

在本文的研究过程中, 由于已知方程的动力学形式, 只需要对方程的参数进行辨识和估算. 因此, 在演化建模过程中, 不需要进行种群间的杂交, 只需要对参数进行遗传和变异操作. 显然, 利用演化建模方法来估算参数要较直接根据观测数据来建立其微分方程要更加容易.

### 2.2 基于演化建模的参数估计算法

基于演化算法的参数估计流程图已由图1给出, 该算法的实现方案具体如下:

1) 设置演化建模过程中演化算法模块的参数, 包括遗传概率(probability of replicate, Pr), 变异概率(probability of mutation, Pm),  $\xi$ 为随机数, 作为参数的变异范围; 设置演化建模的停机条件.

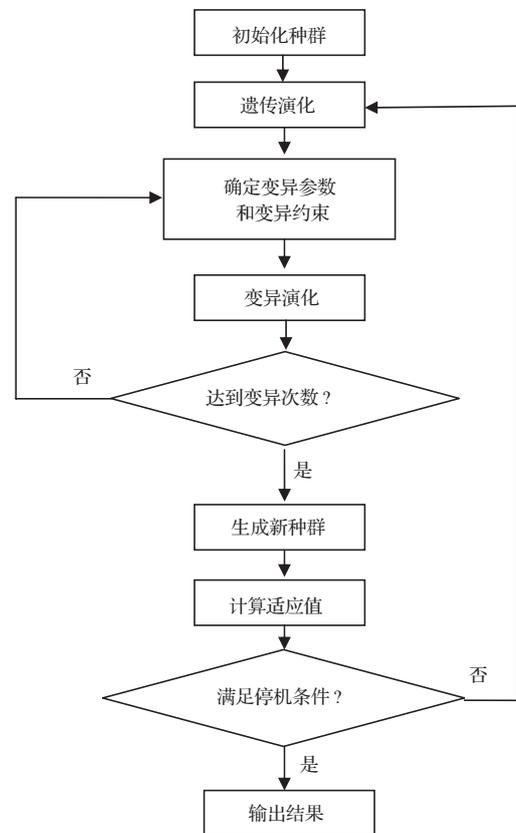


图1 基于演化算法的参数估计流程图

2) 初始化微分方程中待估计参数的第一代种群, 种群规模为  $M$  ( $M$  为正整数).

3) 对初始化种群进行演化操作, 即遗传和变异.

4) 计算种群的适应值, 并根据一定的规则 (详见下文) 从中挑选出下一代种群.

5) 对新种群再度进行演化操作.

6) 不断重复 4)—5) 步让种群逐代演化, 直到所得到的种群其适应值满足预先设定的停机条件为止.

7) 最后一代种群中适应值最优 (如方差最小) 的参数即为最佳估计参数.

演化算法中的遗传操作: 在演化建模的过程中, 当完成一次演化算法后, 随机从  $M$  个个体中选择 60% 的个体 (即  $0.6M$  个个体) 组成一个新的子群, 将其中最好的个体作为下一代种群的个体. 而后再次随机从  $M$  个个体中选择 60% 的个体组成一个新的子群, 再次选择其中最好个体作为下一代种群的个体. 重复这一过程, 直至选满  $M/2$  (或  $(M-1)/2$ ) 个个体.

变异的具体操作如下: 随机从经过遗传操作得到的新群体中选择一个个体, 令待估计的参数的随机初值  $A = A + \xi$ ,  $\xi$  为一定参数搜索范围内的随机数, 为此则生成一个新的个体, 将其作为演化建模的下一代个体. 重复变异操作, 直到通过变异操作得到的个体数目与遗传操作得到的个体数目之和达到预设的种群规模  $M$  为止.

为了便于读者能够应用新参数估计算法, 本文给出了程序实现该算法的伪代码, 具体如下 (“//” 后面的文字为对伪代码的简要解释):

```
Pr=0.5; Pm=0.5; //设置遗传和变异操作的
概率
INI_Group(M); ///初始化种群;
while (!conditions) //如果未满足停机条件;
{
  A. 遗传演化程序模块:
  while (!conditions) //如果未满足遗传个数,
  则遗传演化操作程序循环执行.
  {
    for (int j = 0, size = 0.6* (M.size); j < size;
    j++);
```

```
//从种群 M 中选择 60% 的个体为一组, 再从
中选择最优的个体遗传;
```

```
{
  Ditem=M.randomSeed; //从种群 M 中随机选
  择一个个体;
  TList.push_back(Ditem); //将随机选取的个
  体加入到新的子群;
}
TList.sort(); //对新子群进行排序;
Gooditem=TList.pop_front(); //选取新子群
  中的最优个体;
M1.push_back(Gooditem); //将最优个体加
  入到下一代群体 M1 中.
```

B. 变异演化程序模块

(注: 该程序模块的作用在于: 在保留遗传操作后的个体的前提下, 再对已经执行过遗传操作的所有个体进行变异操作, 增加优秀个体的种群多样性.)

```
while (!conditions)//如果未满足变异个数, 则
  程序循环执行.
{
  DList=M1; //将已经执行遗传操作的个体复
  制放入种群 Copy_M1 中;
  Ditem=DList.randomSeed(); //随 机 从
  Copy_M1 选择一个个体;
  parameters=Ditem.selectNumNode(); //选择
  个体中需要变异的参数;
  delta=rand(2000); //参数变异区间 [-2000.0,
  2000.0];
  T=runTime(); //记录演化程序的运行时间;
  DAMPING=delta*(1/T); //计算变异约束;
  parameters=parameters+DAMPING; //约束
  条件下的参数变异演化;
  M1.push_back(Ditem); //将变异后的个体加
  入新种群 M1 中.
}
M1.calculateFitness(); //计算适应值;
M=M1+Copy_M1; // 经过遗传和变异操作
  后的新产生的个体总数为 M.
}
伪代码结束.
```

### 3 参数估计实例

Lorenz 系统是一个经典的混沌模型(Lorenz, 1963)<sup>[12]</sup>, 已经得到了广泛的研究<sup>[13]</sup>. 众多参数识别方法常采用其作为理想非线性动力学模型进行数值仿真试验, 该系统可由下面的常微分方程组来描述:

$$\begin{aligned} \frac{dx}{dt} &= -\sigma x + \sigma y, \\ \frac{dy}{dt} &= rx - y - xz, \\ \frac{dz}{dt} &= xy - bz, \end{aligned} \quad (1)$$

其中参数  $\sigma, b$  分别取为 10.0, 8/3,  $r$  为混沌控制参数, 取为 28.0. 在数值试验的过程中, 从 Lorenz 系统的吸引子上任意选取一点  $(-7.60979, 4.79228, 37.33112)$  作为积分初值, 数值求解采用四阶龙格-库塔法, 积分步长为 0.01.

本文采用方差来定量衡量某次参数估计值的好坏. 设待辨识参数的系统变量的准确数值解为  $(X, Y, Z)$ , 其中  $X = (X_1, X_2, \dots, X_n), Y = (Y_1, Y_2, \dots, Y_n), Z = (Z_1, Z_2, \dots, Z_n)$ ; 辨识参数带入方程组(1)后的数值解为  $(x, y, z)$ , 其中  $x, y, z$  为含辨识参数的 Lorenz 系统的状态向量  $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n), z = (z_1, z_2, \dots, z_n)$ . 辨识参数的准确性可以方差 (Variance) 的大小来定量判断, 方差定义如下:

$$\begin{aligned} \text{Variance} &= \sqrt{(X - x)^2 + (Y - y)^2 + (Z - z)^2}. \end{aligned}$$

在参数辨识过程中, 演化程序执行一代, 则对种群中的每个个体计算一次方差.

在参数估计实验中, 演化算法种群大小  $M = 200$ , 复制概率和变异概率均为 0.5. 本文首先考虑参数  $r$  未知时的情况. 假定参数  $r$  的初值为  $r_0 = 40$ , 参数的搜索半径为 200, 即  $\xi \in [-200, 200]$ , 基于演化算法的参数估计结果已由图 2 给出. 由于本文数值仿真试验的目的是检验演化算法搜索到的最优参数与真值的接近程度, 因此图中展示的参数演变结果及相应的误差演变曲线仅仅包含了每一代种群中最优的 20 个个体(即方差最小的前 20 个个体, 下文同).

从图 2(a)可以看出, 每一代最优的 20 个个体其方差的离散度, 随着演化程序的进行, 迅速减小. 在大约演化 100 代后, 20 个个体的方差均小于 1.0. 从其 20 个最优种群的参数逐代演化来看(图 2(b)), 经过约 20 代的演化后, 参数搜索的范围迅速的从  $[2, 55]$  骤然缩小到  $[26, 30]$ . 从每一代种群中最优参数值的逐代演化图中可以发现(图 2(c)), 仅仅经过一代演化, 估计到的参数值便已快速搜索到 28.4209, 且演化算法约执行 7 代以后估计到的参数为 28.0958, 演化至 20 代后最优估计参数为 28.0131. 当继续演化至 350 代时, 参数的精度进一步提高, 最优估计参数为 28.0021, 这已经非常接近真实值 28.0. 这表明对于单个参数的辨识, 演化算法不仅收敛速度快, 而且还具备很强的辨识能力.

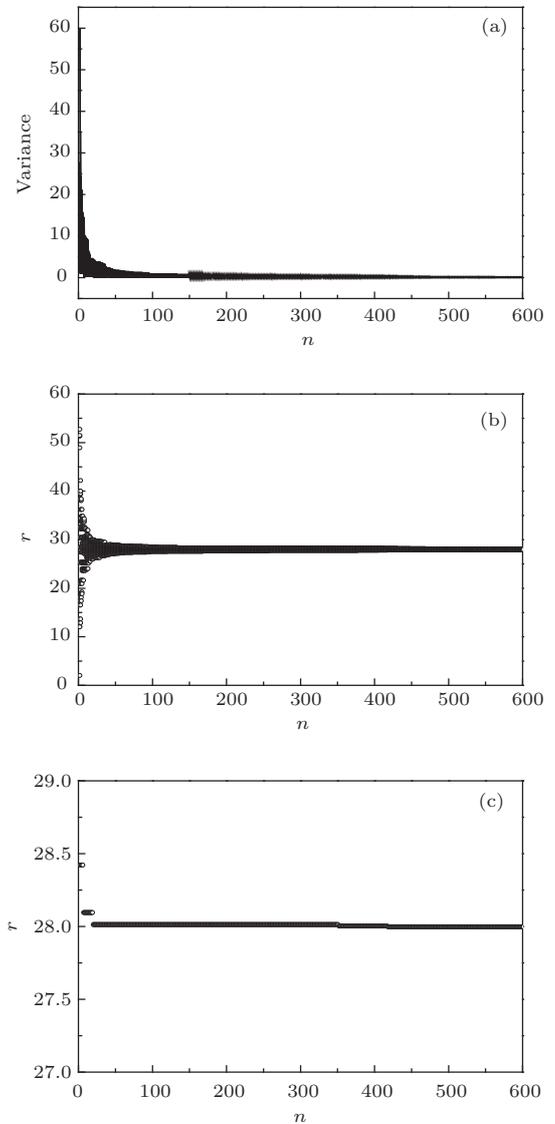


图 2 基于演化算法的单参数 ( $r$ ) 估计, 参数搜索范围为  $[-200, 200]$  (a) 方差演变曲线; (b) 20 个最优参数估计的逐代演化进程; (c) 最优参数估计的逐代演化进程

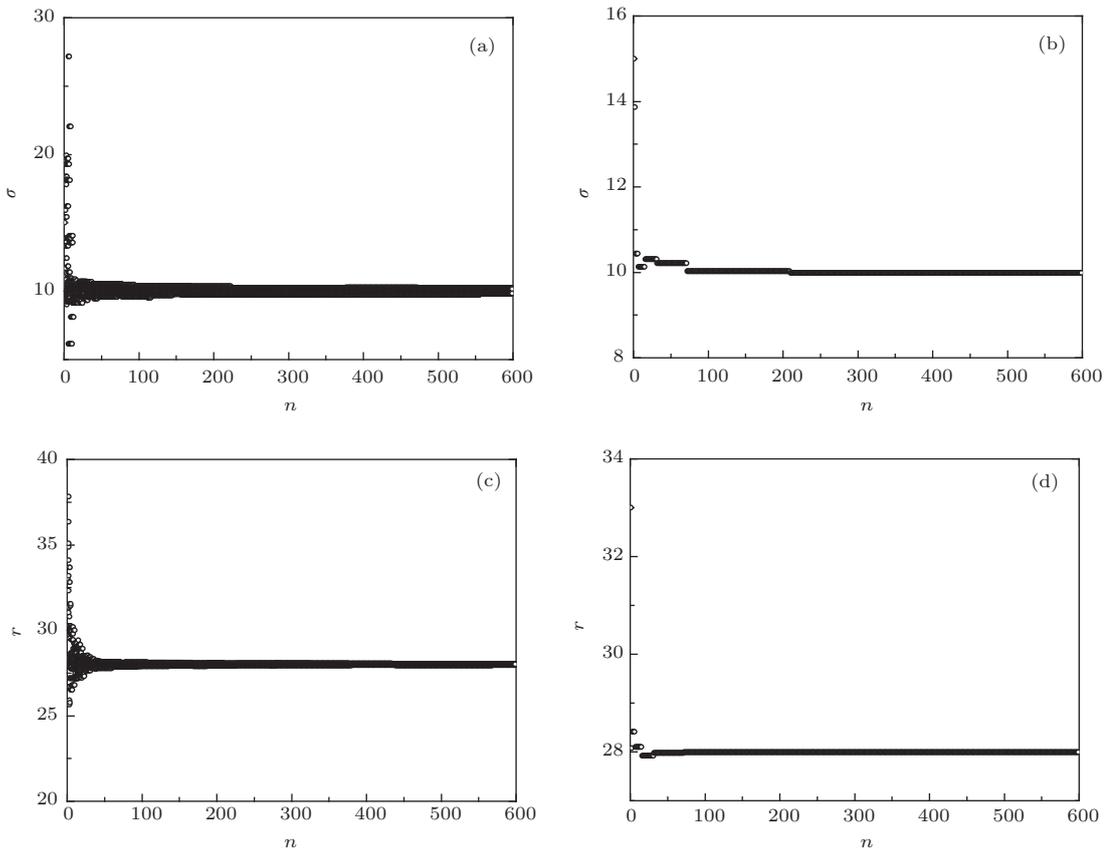


图3 基于演化算法的双参数估计, 参数搜索范围为  $[-20, 20]$  (a) 20个最优估计(参数  $\sigma$ ) 的逐代演化进程; (b) 参数  $\sigma$  的最优估计的逐代演化进程; (c) 20个最优参数估计(参数  $r$ ) 的逐代演化进程; (d) 参数  $r$  的最优估计的逐代演化进程

为了进一步考察演化算法在多参数估计中的性能, 图3给出了当Lorenz系统两个参数  $\sigma, r$  未知时的参数辨识结果. 在双参数辨识过程中, 仅参数搜索范围变为  $\xi \in [-20, 20]$ , 其他设置与单参数试验时完全相同. 试验中令待辨识参数  $\sigma, r$  的初值分别取为  $\sigma_0 = 15, r_0 = 33$ . 对于参数  $\sigma$ , 当演化算法运行至14代时, 参数搜索范围已经从  $[0, 30]$  缩小至  $[9, 11]$ (图3(a)); 从其最优参数估计的演化来看(图3(b)), 当演化算法运行至72代时, 搜索到的参数为10.0211, 已经非常接近真实参数值10.0. 对于参数  $r$ , 当演化算法运行至14代时, 参数搜索范围已经从  $[25, 38]$  缩小至  $[27, 29]$  (图3(c)); 从其最优参数估计的演化来看(图3(d)), 当演化算法运行至7代时, 搜索到的参数为28.0956; 进一步运行至72代时, 搜索到的最优参数为27.9901, 这已经非常接近真实参数值28.0(与真值的差异仅为  $-0.0099$ ). 这一结果表明演化算法对于双参数估计也具有很强的辨识能力.

在实际中可能会存在方程形式已知, 但方程的参数完全未知的情形(如保密通信). 鉴于此, 假设Lorenz系统的三个参数完全未知, 设这三个未

知参数  $\sigma, r, b$  的初始估计值随机取为  $\sigma_0 = -15.0, r_0 = 33.0, b_0 = 5$ . 在试验中, 参数搜索范围为  $\xi \in [-10, 10]$ , 其他设置与单参数试验时完全相同. 从图4中不难发现, 三个参数的收敛速度较单参数和双参数估计试验明显变慢, 尤其是参数  $\sigma$  和  $r$ . 这可能是由于需要对三个参数同时进行搜索, 而同时更为准确的找到近似参数的难度显然会随着搜索参数个数的增多而明显增大, 进而造成了演化算法的收敛速度明显变慢. 从参数  $\sigma$  的估计中可以发现, 当演化算法运行至696代时, 得到的最优估计参数值为10.0451, 较双参数试验中搜索到同等精度的参数值(只需运行72代, 10.0211)明显变慢(图4(b)). 当演化算法运行至696代时, 得到的最优估计参数( $r$ ) 值为28.0193, 搜索速度亦显著的慢于双参数和单参数试验. 从参数  $b$  的估计试验中可以发现, 当演化算法运行至23代时, 便能够搜索到非常接近真实值的参数了(2.695, 与真值  $8/3$  相差约0.028333). 总体来看, 演化算法能够较好的对三个参数同时进行有效估计, 但是参数的搜索速度较单参数和双参数估计明显减小, 耗费的计算机时明显增大.

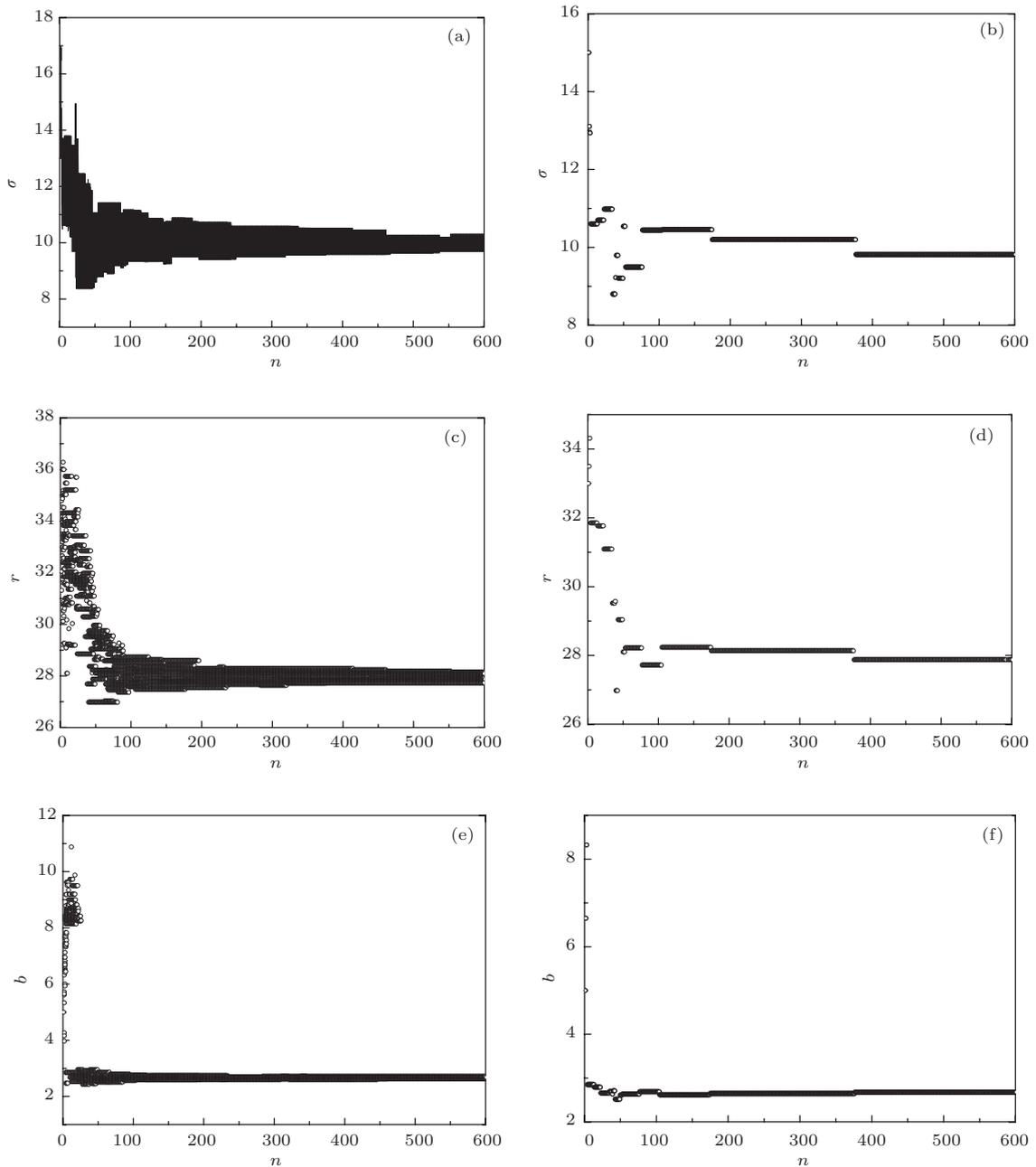


图4 Lorenz 系统三个参数均未知时基于演化算法的参数估计, 参数搜索范围为  $[-10, 10]$  (a) 20 个最优估计 (参数  $\sigma$ ) 的逐代演化进程; (b) 参数  $\sigma$  的最优估计的逐代演化进程; (c) 20 个最优估计 (参数  $r$ ) 的逐代演化进程; (d) 参数  $r$  的最优估计的逐代演化进程; (e) 20 个最优估计 (参数  $b$ ) 的逐代演化进程; (f) 参数  $b$  的最优估计的逐代演化进程

针对上述方案对于多参数同时估计中算法收敛过慢的问题, 本文对演化算法的变异操作作了如下修正: 即令待估计参数的随机初值  $A = A + T \times \xi$ ,  $T = 1/t$  ( $t$  为演化算法执行的时间), 且在每一代种群中参数变异过程中都采用  $T$  对参数的搜索范围进行约束. 这样处理的好处在于, 随着演化程序的运行, 参数种群逐步在优化, 总体而言, 被估计的参数值也越来越靠近真实值; 但在变异的过程中如果对参数的变异范围不加以约束, 可能会导致变异后的参数离真实值差别更大. 而加上约束后, 随着演

化程序的运行, 新估计参数的变异范围越来越小, 这样相当于使得参数的搜索方向更加明确, 减小了参数在变异过程中的不确定性. 设三个未知参数的初始估计值为  $\sigma_0 = 50$ ,  $r_0 = 100$ ,  $b_0 = 50$ , 参数搜索范围为  $\xi \in [-2000, 2000]$ , 其他设置与单参数试验时完全相同. 显然, 改进后的多参数估计实验中, 无论是参数初始猜测值与原始值的差异, 还是参数的搜索范围都明显变得更大. 因此, 本参数估计试验的难度显然明显提高. 从数值试验的结果来看 (图 5), 三个参数的 20 个最优估计的离散度随着演

化程序的运行迅速的减小, 均在约10代后收敛于真值附近. 从三个未知参数的逐代最优估计值的演变可以发现, 演化算法仅运行9代即可找到参数 $\sigma$ 的最优值10.00327, 运行至15代可以找到雷诺数 $r$ 的最优估计值27.99836, 继续运行至20代可以搜索到参数 $b$ 的最优估计值2.665519, 这三个最优估计

值与真实值之间的差异均缩小到0.001的精度. 对比参数估计方案改进前后的结果可知(表1), 改进后的参数估计方案无论是估计值的精度, 还是搜索速率均明显的得到大幅度的提高. 显然, 在演化算法搜索参数的过程中, 对参数的变异采用一定的约束措施能够极大的提高算法的搜索效率.

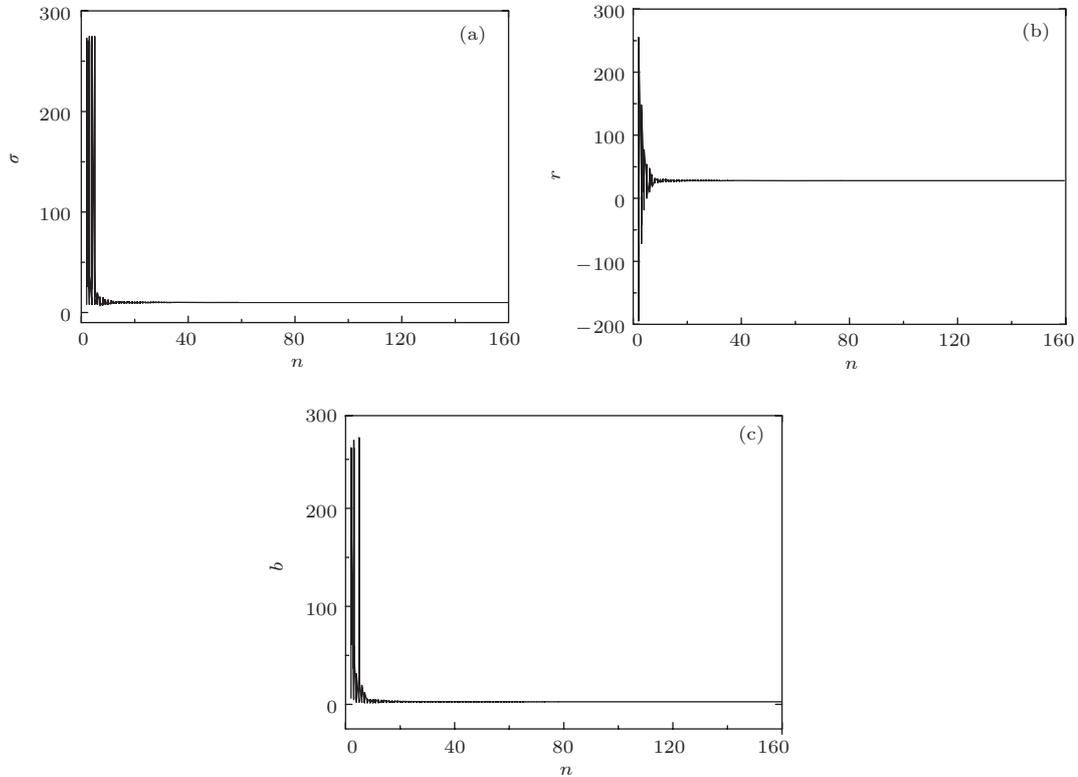


图5 Lorenz 系统三个参数均未知时基于演化算法的参数估计, 参数搜索范围为  $[-2000, 2000]$  (a) 20 个最优估计 (参数  $\sigma$ ) 的逐代演化进程; (b) 20 个最优估计 (参数  $r$ ) 的逐代演化进程; (c) 20 个最优估计 (参数  $b$ ) 的逐代演化进程

表1 参数估计方案改进前后的性能比较

变异操作	最优值		估计误差		最优值的演化代数	
	无约束	有约束	无约束	有约束	无约束	有约束
$\sigma$	10.0451	10.00327	0.0451	0.00327	696	9
$r$	28.0193	27.99836	0.0193	0.001637	696	15
$b$	2.695	2.665519	0.028333	0.0011477	23	20

李丽香等<sup>[5]</sup>提出了采用混沌蚂蚁群算法对混沌系统进行参数估计, 并以典型的Lorenz混沌系统为例进行了计算机模拟, 数值结果表明, 在单参数和多参数情况下, 这种方法都能得到较好的参数估计结果. 鉴于此, 将本文提出的参数估计新方法 with 混沌蚂蚁群算法的多参数估计结果进行了比较.

图6展示了利用混沌蚂蚁群算法对Lorenz混沌系统的三个参数同时进行参数辨识的结果, 从图中不难发现, 混沌蚂蚁群算法能够非常有效的收敛

到被辨识参数的值. 但将其与演化算法的参数估计结果(图5)对比可知, 在混沌蚂蚁群算法的参数估计试验中, 参数 $\sigma$ 的搜索范围是 $[0, 15]$ , 参数 $r$ 的搜索范围是 $[0, 50]$ , 参数 $b$ 的搜索范围是 $[0, 10]$ , 而基于演化算法的三参数估计试验中其搜索范围均为 $[-2000, 2000]$ , 本文所提出新方法的搜索范围明显较混沌蚂蚁群算法更大. 尽管如此, 新方法的收敛速度明显快于混沌蚂蚁群算法: 演化算法基本在20代以内收敛, 而混沌蚂蚁群算法基本上在100代

以后才趋于收敛. 因此, 本文基于演化建模方法的参数估计方案较混沌蚂蚁群算法有着明显的优势.

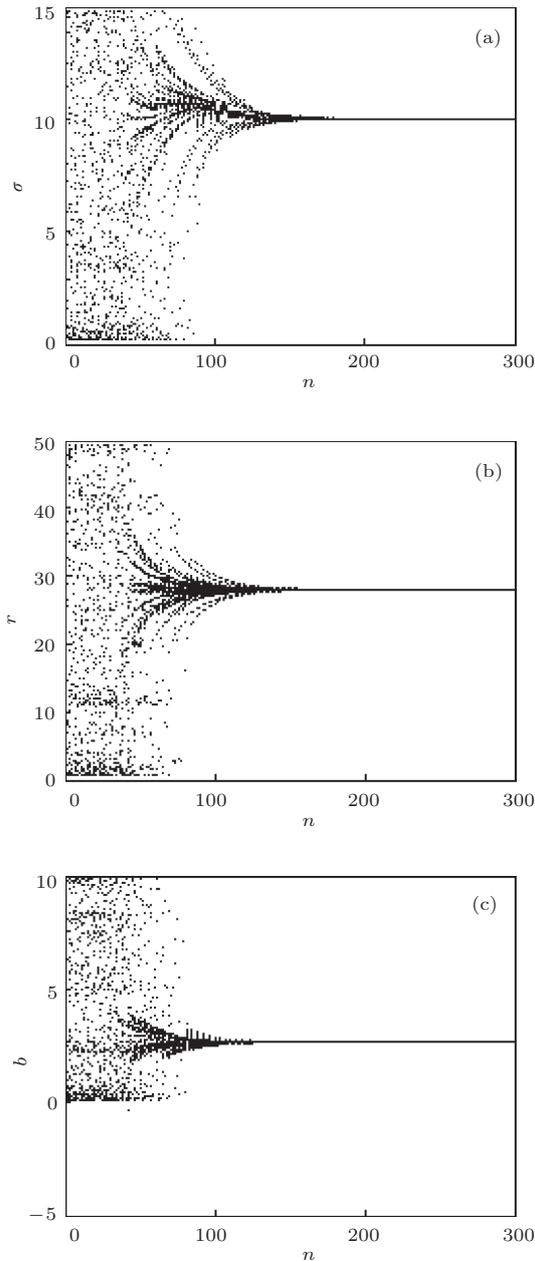


图6 Lorenz 系统三个参数均未知时基于混沌蚂蚁群算法的参数估计结果 (三个参数的真实值与图5 完全相同), 参数 $\sigma$  的搜索范围是  $[0, 15]$ , 参数 $r$  的搜索范围是  $[0, 50]$ , 参数 $b$  的搜索范围是  $[0, 10]$ , 蚂蚁群的个体总数为 20 (a) 参数 $\sigma$  的辨识结果曲线; (b) 参数 $r$  的辨识结果曲线; (c) 参数 $b$  的辨识结果曲线<sup>[5]</sup>

## 4 结论与讨论

本文基于演化算法提出了一种新的非线性动力学方程参数估计方法, 利用典型的混沌模型 Lorenz 方程进行了数值仿真试验. 数值结果表明, 基于演化算法的参数辨识方案不仅能够对 Lorenz

方程部分参数未知时的情形进行准确、快速的辨识, 对于方程中所有参数完全未知时的情况也能够进行非常有效的估计. 由于演化算法采用群体的方式组织搜索, 使得它可以同时在解空间内不同区域进行搜索, 特别适用于大规模并行计算. 与混沌蚂蚁群算法的多参数估计试验结果比较发现, 新方法无论是在参数搜索范围, 还是在算法的收敛速度, 均有着很大的优势, 这进一步证明了新方法的有效性.

必须指出的是, 对于多个参数估计的问题, 由于本文是同时对多个参数进行估计, 通过一个总的方差来对算法搜索到的参数进行识别, 耗费的机时较多, 参数的搜索速度明显减慢. 由于随着演化程序的逐代推进, 参数种群整体上应该是逐步得到优化, 即被估计的参数值越来越靠近真实值, 因此, 可以对演化算法中参数变异的范围进行适当的约束, 以避免好的参数个体可能会由于变异范围过大而更加远离真实值, 进而使得参数的搜索方向更加明确, 提高参数的搜索速度. 鉴于此, 本文对参数估计的演化算法提出了相应的改进措施: 在演化程序的变异操作中, 对参数的变异范围采取了一种约束机理, 即演化时间越长, 参数的变异范围越小. 数值仿真试验结果表明, 改进方案极大地提高了演化算法的搜索效率和精度, 节约了大量的机时, 证实了改进方案的有效性.

已有的大量研究表明, 种群规模过小, 随着演化的进行, 会导致大量相似的个体出现, 使得种群的多样性下降, 不利于最优种群的获得; 如果种群规模过于庞大, 可以同时处理更多的个体, 因而容易找到全局最优解, 其缺点是增加了每次演化的时间, 尤其是对一些复杂的、精度要求比较高的问题, 种群的大量增加会导致计算量呈指数增长. 本文参数试验中的种群规模大小的选择主要参考了以往大量相关研究结果, 同时我们对种群规模大小分别为 100, 300, 400, 1000 时进行了性能测试, 发现适当的改变种群大小, 对于本文算法的搜索速度以及精度的影响并不大. 种群规模的大小选择与参数估计中算法的搜索速度以及求解问题的精度之间的定量关系非常重要, 确实值得进一步的深入研究, 这将有利于新算法向大型的数值模式推广应用. 当然这个问题已经超出了本文现有的研究范畴, 在将来的后续研究中可以对这一问题进行进一步的探讨.

在已有演化算法的相关研究中, 个体遗传、变

异、杂交等操作的概率主要是基于统计经验和研究者的对相关问题的理解程度,而且在不同的研究对象中三种演化操作的概率也不尽相同,这是一个非常难以定量确定的问题.由于参数估计问题的特殊性,三种演化操作减少为遗传和变异两种.在自然界生物的演化过程中,相对而言变异的概率是较小的,但是作者考虑到,由于本文试验中选择的种群规模毕竟并不庞大,如果让遗传的概率过大,会导致遗传在固定的种群中进行,遗传的结果过度的依赖于初始选择的种群,不利于种群的进化.而大幅度增加变异的概率后,可以极大提高种群中的个体活力,即使得种群中新增了新型基因,相当于在不增加计算资源的情况下增加了种群的规模.在本文参数估计试验中,个体遗传和变异的概率均等,都取为0.5,主要是依据经验所选择,适当的改变遗传和变异的概率(如  $P_r=0.4$ ,  $P_m=0.6$ ),不会对本文结果产生实质性的影响.在遗传操作的具体执行中,为了尽量将每一次演化过程中参数值较优的遗传给下一代,需要将适应值越小的保留下来,以确保较优的参数值在后面的演化中有更多遗传和变异的概率.

本文中确定挑选出子群体的范围为种群的60%,主要考虑到,如果将50%的个体认定为不好的估计值,而剩余的个体作为较好的估计值,至少每次有10%较好的个体能够有遗传的概率.这样处理的好处在于,一方面保证了种群中较好的个体能够遗传,较差的个体没有遗传的概率,另一方面,也避免了在整个种群中进行遗传操作,单个个体遗传的概率太大,破坏了种群的多样性.参数估计试验的结果表明,适当增加挑选子群体的概率对于结果的影响不大,就本文试验结果而言,只要挑选子种群的样本量超过种群规模的一半,就能达到较好的遗传效果.对于最优的子群体范围的选择,需要

大量的研究才能确定,而且对于不同的问题,亦可能还存在不确定性,即有可能不存在普适的最优阈值.从本文的数值仿真试验结果来看,尽管在演化建模的各个参数的量化研究中,还有许多需要继续深入探讨和研究的问题,但其结果依然是令人鼓舞的.在将来的相关研究工作中我们会继续对此问题进行探索与完善.

## 参考文献

- [1] Huang J P, Yi Y H 1991 *Sci. Chin. B* **3** 331 (in Chinese) [黄建平, 衣育红 1991 中国科学B辑 **3** 331]
- [2] Shi N 2009 *Meteorology statistical forecasting* (Beijing: Metrological press) pp30–65 (in Chinese) [施能 2009 气象统计预报 (北京: 气象出版社) 第30–65页]
- [3] Song J Q, Cao X Q, Zhang W M, Zhu X Q 2012 *Acta Phys. Sin.* **61** 110401 (in Chinese) [宋君强, 曹小群, 张卫民, 朱小谦 2012 物理学报 **61** 110401]
- [4] Li L X, Peng H P, Wang X D, Yang Y X 2006 *Int. J. Bifurcat. Chaos* **16** 2351
- [5] Li L X, Peng H P, Yang Y X, Wang X D 2007 *Acta Phys. Sin.* **56** 55 (in Chinese) [李丽香, 彭海朋, 杨义先, 王向东 2007 物理学报 **56** 55]
- [6] Guan X P, Peng H P, Li L X, Wang X Q 2001 *Acta Phys. Sin.* **50** 26 (in Chinese) [关新平, 彭海朋, 李丽香, 王益群 2001 物理学报 **50** 26]
- [7] Dai D, Ma X K, Li F C, You Y 2002 *Acta Phys. Sin.* **51** 2459 (in Chinese) [戴栋, 马西奎, 李富才, 尤勇 2002 物理学报 **51** 2459]
- [8] Cao X Q, Song J Q, Zhang W M, Zhao J, Zhang L L 2011 *Acta Phys. Sin.* **60** 070511 (in Chinese) [曹小群, 宋君强, 张卫民, 赵军, 张理论 2011 物理学报 **60** 070511]
- [9] Li N Q, Pan W, Yan L S, Luo B, Xu M F, Jiang N 2011 *Chin. Phys. B* **20** 060502
- [10] Li X T, Yin M H 2012 *Chin. Phys. B* **21** 050507
- [11] Cao H, Kang L, Chen Y, Yu J 2000 *Genet. Program. Evol. M.* **1** 309
- [12] Lorenz E N 1963 *J. Atmos. Sci.* **20** 130
- [13] Feng G L, He W P 2007 *Chin. Phys. B* **16** 2825

# Evolutionary modeling for parameter estimation for chaotic system\*

Wang Liu<sup>1)2)</sup> He Wen-Ping<sup>2)†</sup> Wan Shi-Quan<sup>3)</sup>  
Liao Le-Jian<sup>2)</sup> He Tao<sup>4)</sup>

1) (*National Climate Center, China Meteorological Administration, Beijing 100081, China*)

2) (*School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China*)

3) (*Yangzhou Meteorological office, Yangzhou 225009, China*)

4) (*Changzhou Environment Monitoring Center, Changzhou 213000, China*)

( Received 15 May 2013; revised manuscript received 18 September 2013 )

## Abstract

On the basis of evolutionary algorithm, a novel method for parameter estimation of nonlinear dynamic equations is given in the present paper. Numerical tests indicate that the unknown parameters all can be estimated quickly and accurately whether the partial parameters are unknown or all parameters are unknown in the classic Lorenz equation. However, it is found that the convergence rate of the new algorithm is relatively slow when multiple unknown parameters are estimated simultaneously. To solve this problem, a corresponding improvement of measure is proposed, namely, a constraint mechanism is taken during the variation operation of evolutionary algorithm. The improvement is mainly based on the characteristic that the longer the running time of the evolutionary algorithm, the smaller the range of variation of the estimated parameters. Results indicate that the searching speed of the algorithm is greatly improved by using the improved estimation parameter project.

**Keywords:** chaotic system, parameter estimation, evolutionary algorithm

**PACS:** 92.60.Wc

**DOI:** 10.7498/aps.63.019203

---

\* Project supported by the National Basic Research Program of China (Grant Nos. 2012CB955902, 2013CB329605), the National Natural Science Foundation of China (Grant Nos. 41275074, 41005041, 41175084), and the Special Scientific Research Fund of Meteorological Public Welfare Profession of China (Grant Nos. GYHY201106015, GYHY201106016).

† Corresponding author. E-mail: [wenping\\_he@163.com](mailto:wenping_he@163.com)