

非线性映射参数辨识的离散变分方法*

曹小群[†] 张卫民 宋君强 朱小谦 赵军

(国防科学技术大学计算机学院, 长沙 410073)

(2011年3月18日收到; 2011年5月3日收到修改稿)

提出一种辨识非线性映射系统中未知参数的离散变分方法, 对以 $bx_{k+1} = \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})$ 为状态控制方程的所有映射混沌系统具有通用性. 对典型的 Logistic 映射和 Henón 映射中的未知参数进行了估计, 仿真结果表明了该方法的有效性.

关键词: 映射, 参数辨识, 离散变分方法, 伴随方程

PACS: 05.45.Pq, 04.20.Fy

1 引言

自 Lorenz 于 1963 年发现第一个混沌吸引子后, 混沌控制和同步已成为非线性科学中重要的研究方向之一, 已经提出很多有效的控制与同步方法^[1-5]. 由于混沌系统的复杂性, 某些参数经常无法测量或确定; 或者出于某些特殊原因, 系统的某些参数不可知, 例如保密通信的要求. 因此为了实现混沌系统的控制或同步, 近几年研究人员不断提出辨识混沌系统参数的新方法: 文献 [6] 通过参数自适应方法对目标系统的参数进行估计, 达到了广义同步的目的; 文献 [7] 提出了未知参数辨识观测器的概念, 并对 Lorenz 系统的参数进行了辨识; 文献 [8] 基于遗传算法对混沌系统进行参数估计, 并以典型的 Lorenz 系统为例进行了研究; 文献 [9] 基于具有群集智能的改进粒子群优化算法成功估计了混沌系统的参数; 文献 [10] 和 [11] 分别基于混沌蚂蚁群算法和混合差分进化算法的全局优化搜索的能力, 对混沌系统进行了参数估计. 在混沌控制和同步领域, 非线性映射是一类重要的混沌系统, 估计它们的未知参数是也是一个必须解决的关键问题.

非线性映射的参数辨识问题属于离散动力系统研究中典型的反问题, 和正问题不同, 它是通过

试验或运行中的观测数据来反求模型中的未知参数. 由于测量信息含有噪声和系统的强非线性, 导致反问题的求解经常是不适定的. 即解不一定存在、即使解存在也不唯一、在解存在唯一条件下也不稳定 (即解不连续依赖于观测数据)^[12]. 因此非线性映射系统参数辨识问题的求解必须采用特殊方法^[6-11]. 本文基于离散变分原理来估计非线性映射中的未知参数, 该方法对以 $\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})$ 为状态控制方程的所有映射具有通用性. 首先引入拉格朗日乘子向量序列, 将以非线性离散系统为约束的最优化问题转换为无约束最优化问题; 接着利用离散变分原理导出伴随方程和目标泛函关于待辨识参数的梯度; 然后给出了离散变分方法辨识非线性映射中未知参数的算法流程. 最后对 Logistic 映射和 Henón 映射中的未知参数进行了估计, 仿真结果表明该方法具有很高的估计精度, 同时具有较好的抗噪声性能. 需要说明的是, 本文引入 Lagrange 乘子来解决问题, 但正如 He^[14-18]所指出的, 对复杂系统 Lagrange 乘子可能会失效和产生临界变分^[16]. 为了克服临界变分现象, He 提出了一种新理论——半反推方法 (semi-inverse method)^[14-18]. 因此, 对某些非线性映射, 采用 He 提出的半反推方法可取得更好的效果, 这是以后需要进一步开展的工作.

* 国家自然科学基金 (批准号: 41105063, 61070041) 资助的课题.

[†] E-mail: caoxiaoqun@nudt.edu.cn

2 问题描述

考虑如下 n 维映射混沌系统的参数辨识问题:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta}), \\ \mathbf{x}|_{k=0} &= \mathbf{x}_0, \\ (k &= 0, 1, 2, \dots, N-1), \end{aligned} \quad (1)$$

其中 $\mathbf{x}_0$ 和 $\mathbf{x}_k \in R^n$ 分别表示非线性映射的 n 维状态向量序列和初值, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)^T \in R^m$ 表示非线性映射中需要辨识的物理参数向量. 目标是利用离散混沌系统状态的测量数据估计不可测的未知参数向量 $\boldsymbol{\theta}$. 假设非线性映射系统的结构已知, 即矢量函数 $\mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta}) = (f_1(\mathbf{x}_k, \boldsymbol{\theta}), f_2(\mathbf{x}_k, \boldsymbol{\theta}), \dots, f_n(\mathbf{x}_k, \boldsymbol{\theta}))^T$ 的表达式是已知的, 要对非线性映射系统的待辨识参数进行估计, 则可以通过对下面的目标泛函进行最小化而获得

$$J(\boldsymbol{\theta}) = \frac{1}{2} \sum_{k=1}^N \|\mathbf{x}_k - \mathbf{x}_k^{\text{obs}}\|^2, \quad (2)$$

其中 $\mathbf{x}_k^{\text{obs}}$ 和 $\mathbf{x}_k (k = 1, 2, 3, \dots, N)$ 分别表示非线性映射系统的观测序列和状态向量序列. 显然, 非线性映射系统的参数辨识问题变成在离散状态方程 (1) 式约束下的目标泛函的最优化问题, 决策变量为 $\boldsymbol{\theta}$, 优化目标为 $\min_{\boldsymbol{\theta} \in R^m} J(\boldsymbol{\theta})$. 但是, 由于非线性映射系统的强非线性和不稳定性, 系统参数很难估计; 另外, 目标泛函 $J(\boldsymbol{\theta})$ 中可能存在许多局部极值点, 以上原因使准确计算目标泛函关于决策变量 $\boldsymbol{\theta}$ 的梯度 $\nabla_{\boldsymbol{\theta}} J = (\nabla_{\theta_1} J, \nabla_{\theta_2} J, \dots, \nabla_{\theta_m} J)^T$ 变得非常重要. 下面利用离散变分方法准确获得 $\nabla_{\boldsymbol{\theta}} J$ 的表达式及计算梯度需要的伴随方程.

3 非线性映射参数辨识的离散变分方法

首先引入拉格朗日乘子向量序列 $\boldsymbol{\lambda}_k \in R^n (k = 0, 1, 2, \dots, N)$, 将以非线性映射方程为约束条件的最优化问题转化为如下形式的无约束最优化问题:

$$\begin{aligned} J(\boldsymbol{\theta}) &= \frac{1}{2} \sum_{k=1}^N \|\mathbf{x}_k - \mathbf{x}_k^{\text{obs}}\|^2 \\ &+ \sum_{k=1}^{N-1} \boldsymbol{\lambda}_{k+1}^T \cdot [\mathbf{x}_{k+1} - \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})], \end{aligned} \quad (3)$$

(3) 式的一阶变分为

$$\begin{aligned} \delta J(\boldsymbol{\theta}) &= \sum_{k=1}^N (\mathbf{x}_k - \mathbf{x}_k^{\text{obs}})^T \cdot \delta \mathbf{x}_k \\ &+ \sum_{k=0}^{N-1} \left\{ \delta \boldsymbol{\lambda}_{k+1}^T \cdot [\mathbf{x}_{k+1} - \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})] \right. \\ &\left. + \boldsymbol{\lambda}_{k+1}^T \cdot [\delta \mathbf{x}_{k+1} - \delta \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})] \right\}, \end{aligned} \quad (4)$$

式中 $\delta \mathbf{x}_k$ 和 $\delta \boldsymbol{\lambda}_k$ 分别表示非线性映射系统状态向量序列和拉格朗日乘子向量序列的一阶变分. 而目标泛函 $J(\boldsymbol{\theta})$ 和矢量函数 $\mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})$ 的一阶变分按定义可以进一步写为

$$\begin{aligned} \delta J(\boldsymbol{\theta}) &= (\nabla_{\boldsymbol{\theta}} J)^T \cdot \delta \boldsymbol{\theta} \\ \delta \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta}) &= \left(\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \mathbf{x}_k} \right)^T \cdot \delta \mathbf{x}_k \\ &+ \left(\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^T \cdot \delta \boldsymbol{\theta}. \end{aligned} \quad (5)$$

将 (5) 式中第二式代入 (4) 式右端第三项, 有

$$\begin{aligned} \sum_{k=0}^{N-1} \boldsymbol{\lambda}_{k+1}^T \cdot [\delta \mathbf{x}_{k+1} - \delta \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})] &= \sum_{k=1}^N \boldsymbol{\lambda}_k^T \cdot \delta \mathbf{x}_k \\ &- \sum_{k=0}^{N-1} \boldsymbol{\lambda}_{k+1}^T \cdot \left[\left(\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \mathbf{x}_k} \right)^T \cdot \delta \mathbf{x}_k \right. \\ &\left. + \left(\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^T \cdot \delta \boldsymbol{\theta} \right] \\ &= (\boldsymbol{\lambda}_k^T \cdot \delta \mathbf{x}_k)|_{k=0}^{k=N} + \sum_{k=0}^{N-1} \left\{ \left[\boldsymbol{\lambda}_k \right. \right. \\ &- \left. \left(\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \mathbf{x}_k} \right) \cdot \boldsymbol{\lambda}_{k+1} \right]^T \cdot \delta \mathbf{x}_k \\ &- \left. \boldsymbol{\lambda}_{k+1}^T \cdot \left(\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^T \cdot \delta \boldsymbol{\theta} \right\}. \end{aligned} \quad (6)$$

将 (5) 式中第一式和 (6) 式分别代入 (4) 式的左右两端, 通过合并同类项, 而且由于一阶变分 $\delta \mathbf{x}_k, \delta \boldsymbol{\lambda}_k$ 和 $\delta \boldsymbol{\theta}$ 可以取任意值, 同时 $\delta \mathbf{x}_0 = 0$ 和 $\delta \mathbf{x}_N \neq 0$, 因此可导出下面的等式方程组:

$$\begin{aligned} \mathbf{x}_{k+1} - \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta}) &= 0, \\ -\boldsymbol{\lambda}_k + \left[\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \mathbf{x}_k} \right] \cdot \boldsymbol{\lambda}_{k+1} &= (\mathbf{x}_k - \mathbf{x}_k^{\text{obs}}), \\ \nabla_{\boldsymbol{\theta}} J &= - \sum_{k=0}^{N-1} \frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \cdot \boldsymbol{\lambda}_{k+1}, \\ \boldsymbol{\lambda}_N &= 0, \\ (k &= 0, 1, 2, \dots, N-1), \end{aligned} \quad (7)$$

式中的第一个表达式就是原非线性映射的状态控制方程组. 第二个表达式是非线性映射所对应的伴随方程, 最初引入的拉格朗日乘子向量序列成为了

伴随状态向量序列; 从表达式中不难看出, 伴随方程的求解需要预先计算原映射系统的演化轨迹序列 $\mathbf{x}_k$, 而且 λ_k 前面的负号表示需要对伴随方程按照 $k = N - 1, N - 2, \dots, 2, 1, 0$ 的顺序进行逆向迭代计算. 而 (7) 式中第四式表明伴随方程的初值条件是 $\lambda_N = 0$. (7) 式中的第三个表达式是目标泛函关于未知参数向量梯度的表达式, 易知梯度的计算需要预先求解非线性映射的状态控制方程和伴随方程, 获得映射系统的状态量和伴随量的演化序列 $\mathbf{x}_k$ 和 λ_k . (7) 式中的第二个和第三个表达式对所有以 $\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})$ 为状态控制方程的非线性映射系统具有普适性, 是获得伴随方程和目标泛函梯度的通用公式. 在未知参数的泛函梯度求出之后, 可以选择合适的下降算法 (如拟牛顿算法和共轭梯度算法), 按照下式对每个未知参数进行迭代:

$$\theta_j^{i+1} = \theta_j^i - \nabla_{\theta_j} J|_{\theta^i} \cdot \rho_j^{i+1}, \quad j = 1, 2, \dots, m. \quad (8)$$

最终可以辨识出非线性映射的未知参数值. 当 $i = 0$ 时 θ^i 代表未知参数向量的初始猜测值; ρ_j^{i+1} , ($j = 1, 2, \dots, m$) 表示第 i 次迭代步长, 其具体大小由下降算法决定. 离散变分方法估计非线性映射中未知参数的具体流程如图 1 所示, 算法详细说明如下.

第一步 设定非线性映射系统中各个未知物理参数的初始猜测值;

第二步 利用新的猜测值, 对非线性映射方程 (1) 式进行正向迭代计算, 获得系统状态向量的演化序列 $\mathbf{x}_k$ ($k = 1, 2, 3, \dots, N$) 并加以存贮;

第三步 利用 $\mathbf{x}_k$ 和 $\mathbf{x}_k^{\text{obs}}$, 根据 (2) 式计算目标泛函 $J(\boldsymbol{\theta})$ 的值;

第四步 利用 $\mathbf{x}_k$, $\mathbf{x}_k^{\text{obs}}$ 和伴随量的初值条件, 对伴随模式从 $k = N$ 到 $k = 0$ 进行反向迭代计算, 求得伴随向量的演化序列 λ_k 并加以存贮;

第五步 利用 $\mathbf{x}_k$ 和 λ_k 按照 (7) 式中第三个表达式求各未知参数的泛函梯度值;

第六步 利用标准下降算法, 求出步长 ρ_j^{i+1} , 根据 (8) 式对各未知参数进行迭代, 求出新估计值. 若不满足程序终止条件 (如梯度模变化小于设定的阈值或达到预先指定的最大迭代次数), 则利用新的物理参数向量 $\boldsymbol{\theta}$ 值从第二步开始进行新一轮的迭代循环; 若满足, 则终止程序, 同时获得非线性映射的

参数估计值.

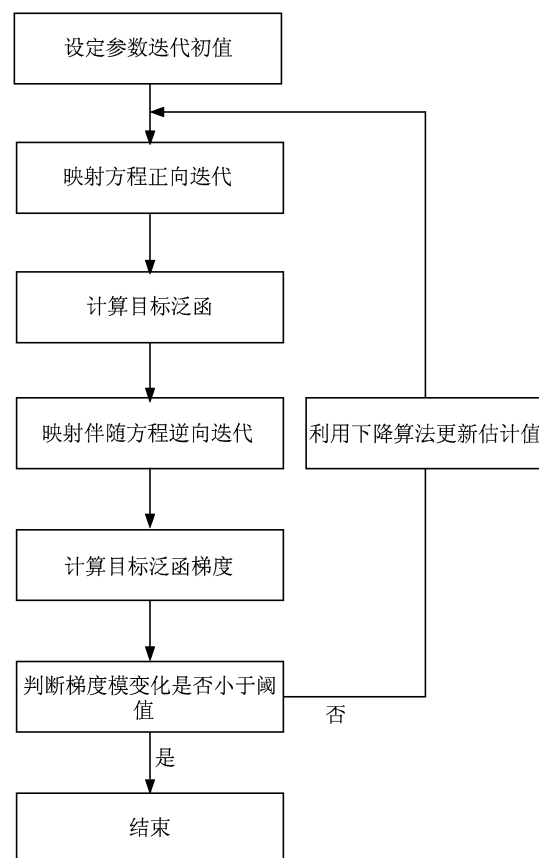


图 1 离散变分方法辨识非线性映射参数的流程图

4 数值仿真结果及分析

4.1 Logistic 混沌映射系统参数的辨识

首先以典型的 Logistic 映射为例, 说明利用离散变分方法辨识非线性映射系统未知参数的有效性. Logistic 映射可用离散方程表示如下:

$$x_{k+1} = \mu x_k (1 - x_k), \quad (9)$$

当 $3.57 < \mu \leq 4$ 时, Logistic 映射处于混沌状态. 目标是在获得状态向量的观测序列 x_k^{obs} ($k = 1, 2, \dots, N$) 的情况下, 估计未知参数向量 $\boldsymbol{\theta} = \mu$. 由 (9) 式与 (1) 式对比可知, $\mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta}) = \mu x_k (1 - x_k)$, 其一阶变分表达式为

$$\begin{aligned} \delta \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta}) &= \mu(1 - 2x_k) \cdot \delta x_k \\ &\quad + x_k(1 - x_k) \cdot \delta \mu, \end{aligned} \quad (10)$$

其中 δx_k 和 $\delta \mu$ 分别是 x_k 和 μ 的一阶变分. 引入伴随量序列 $[\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_N]$, 并将 $x_k, x_k^{\text{obs}}, \lambda_k, \boldsymbol{\theta} = \mu$, $\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial x_k} = \mu(1 - 2x_k)$ 和 $\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} =$

$x_k(1-x_k)$ 代入 (7) 中的第二式和第三式, 分别得到 Logistic 映射的伴随方程和目标泛函关于未知参数 μ 的梯度表达式如下:

$$\begin{aligned} -\lambda_k + \mu(1-2x_k)\lambda_{k+1} &= (x_k - x_k^{\text{obs}}), \\ \lambda_k|_{k=N} &= 0, \\ (k &= 0, 1, 2, \dots, N-1), \end{aligned} \quad (11)$$

$$\nabla_{\mu} J = - \sum_{k=0}^{N-1} x_k(1-x_k)\lambda_{k+1}. \quad (12)$$

在仿真试验中, 观测数据的采集过程如下: 设定参数的真值为 $\mu = 3.6$, 让 Logistic 映射自由演化, 在经历暂态过程之后任取混沌状态的某一点作为初值, 并以此时刻作为初始时刻; 然后任其迭代演化至第 N 步, 可得到 Logistic 映射在标准离散状态序列 $x(k)$, $k = 0, 1, 2, \dots, N$, 而且这些状态量都处于系统的混沌状态. 在试验中, 系统初始状态量设置为 $x_0 = 0.01892$, 其是 Logistic 映射混沌状态中的一点. 试验中只抽取前面 $N = 4$ 个状态量作为观测量, 可以表示为 x_k^{obs} , $k = 1, 2, \dots, N$; 然后可在所选状态量上叠加高斯随机数作为观测噪声, 观测噪声的均值取为 0、标准偏差取为 σ_o . 离散形式的目标泛函表示如下:

$$J(\mu) = \frac{1}{2} \sum_{i=1}^N (x_k - x_k^{\text{obs}})^2. \quad (13)$$

在获得 Logistic 映射系统的伴随方程、目标泛函及其梯度表达式后, 利用标准下降算法可以对 μ 进行最优估计, 未知参数的迭代初值设定为 $\mu = 5.2$. 在对 Logistic 映射系统的未知参数进行变分迭代辨识过程中, 未知参数 μ 的每次更新, 都要对 Logistic 映射系统的控制方程进行一次正向 ($k = 0 \rightarrow N$) 迭代计算, 以便计算目标泛函的大小, 同时对相应的伴随方程进行一次逆向 ($k = N \rightarrow 0$) 迭代计算, 从而计算目标泛函关于未知向量的梯度的值.

图 2 中给出了不存在观测噪声时未知参数 μ 的迭代估计序列值; 图 3 表示的是目标泛函梯度模 $\|\nabla_{\theta_j} J\|$ 的下降图. 理论上当目标泛函关于未知参数的梯度模 $\|\nabla_{\theta_j} J\| = 0$ 时, 未知参数向量取最优估计值, 但本文中为了减小计算量, 采用的是通常的迭代收敛标准: 前后两次迭代的目标泛函梯度模的变化小于预先指定的阈值: $\|\nabla_{\theta_j} J\| - \|\nabla_{\theta_{j+1}} J\| < 10^{-8}$, 此时的计算误差小于机器误差, 可以忽略. 从图中可以看到, 未知参数 μ

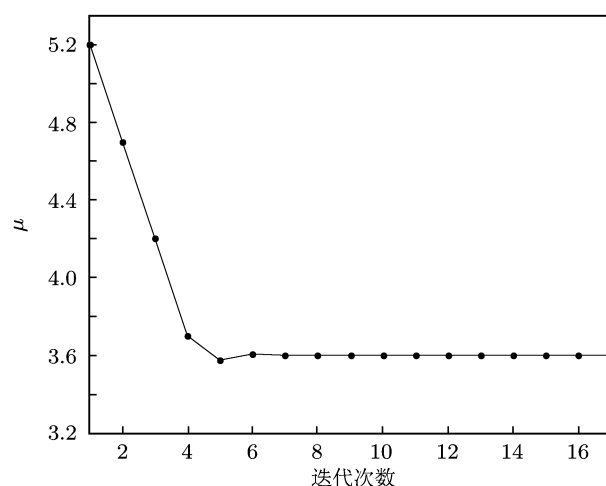


图 2 参数 μ 的辨识结果曲线图

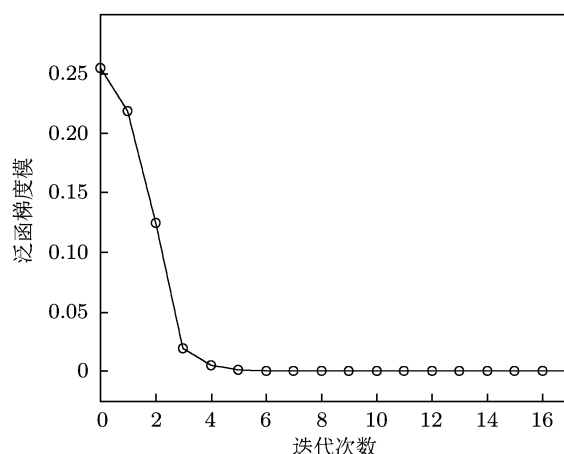


图 3 泛函梯度模的变化图

表 1 不同观测噪声水平情况下 Logistic 映射的参数估计结果

σ_o	0.0	0.1	0.2
真值	3.600000	3.600000	3.600000
估计值	3.599999	3.599989	3.599782
迭代次数	14	23	42
目标泛函	1.8936×10^{-15}	0.006565	0.023118

只经历 17 次迭代就收敛到真实值, 同时目标泛函和梯度模都非常接近 0. 分析图中所示的试验结果可得如下结论: 由于不存在离散误差, 采用离散变分方法辨识 Logistic 映射混沌系统的未知参数具有较高的准确性. 另外离散变分方法的优势是通过伴随方程能够准确计算出目标泛函梯度, 因此使目标泛函快速下降, 同时达到收敛. 表 1 中给出了在不同观测噪声水平情况下离散变分方法对 Logistic 映射系统中未知参数进行估计的结果. 从表中可以看出, 在不存在观测误差时, 未知参数的估计精度最

高, 估计结果可以精确到小数点后第六位, 基本上可以视为真值; 随着观测误差增加, 参数的估计精度下降, 但迭代估计值仍然收敛到真值附近, 进一步验证了离散变分方法辨识非线性映射中未知参数的有效性, 同时说明离散变分方法具有较好的抗噪声性能. 另外, 从表 1 中还可归纳出一个特点: 当不存在观测误差时, 迭代收敛时的目标泛函值基本上为 0; 随着观测误差的增加, 达到收敛时的迭代次数、计算开销和目标泛函值都有所增大.

4.2 Henón 映射系统参数的辨识

本部分中以 Henón 映射系统为例, 进一步说明利用离散变分方法对非线性映射的未知参数进行辨识的有效性. Henón 映射可由如下的非线性离散方程表示:

$$\begin{aligned} x_{k+1} &= y_k - ax_k^2 + 1, \\ y_{k+1} &= bx_k, \\ (k &= 0, 1, 2, \dots, N-1). \end{aligned} \quad (14)$$

当 $a = 1.4$, $b = 0.3$ 时, Henón 映射系统呈现混沌状态. 设定状态向量 $\mathbf{x} = (x, y)^T$ 和物理参数向量 $\boldsymbol{\theta} = (a, b)^T$, 且 $\boldsymbol{\theta}$ 是未知的. 目标是在获得状态向量 $\mathbf{x}$ 的观测量 $\mathbf{x}_k^{\text{obs}} = (x_k^{\text{obs}}, y_k^{\text{obs}})^T$ 的情况下, 辨识未知参数向量 $\boldsymbol{\theta}$. 通过比较 (14) 式与 (1) 式, 易知 $\mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta}) = [y_k - ax_k^2 + 1, bx_k]^T$, 其一阶变分表达式为

$$\begin{aligned} \delta \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta}) &= \begin{pmatrix} -2ax_k & b \\ 1 & 0 \end{pmatrix}^T \cdot \delta \mathbf{x}_k \\ &+ \begin{pmatrix} -x_k^2 & 0 \\ 0 & x_k \end{pmatrix}^T \cdot \delta \boldsymbol{\theta}, \end{aligned} \quad (15)$$

其中 $\delta \mathbf{x}_k = (\delta x_k, \delta y_k)^T$, $\delta \boldsymbol{\theta} = (\delta a, \delta b)^T$. 引入与 $\mathbf{x}_k$ 对应的伴随向量序列 $\boldsymbol{\lambda}_k = (\lambda_k, p_k)^T$, 并将 $\mathbf{x}_k$, $\mathbf{x}_k^{\text{obs}}$, $\boldsymbol{\lambda}_k$, $\boldsymbol{\theta}$, $\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \mathbf{x}_k} = \begin{pmatrix} -2ax_k & b \\ 1 & 0 \end{pmatrix}$ 和 $\frac{\partial \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} -x_k^2 & 0 \\ 0 & x_k \end{pmatrix}$ 等代入 (7) 式中的第二式和第三式, 同时利用伴随初始条件 (7) 式中第四式, 分别得到 Henón 映射混沌系统的伴随方程和目标泛函关于未知参数的梯度公式如下:

$$-\lambda_k - 2ax_k\lambda_{k+1} + bp_{k+1} = (x_k - x_k^{\text{obs}}),$$

$$-p_k + \lambda_{k+1} = (y_k - y_k^{\text{obs}}), \quad (16)$$

$$(\lambda_k, p_k)|_{k=N} = 0;$$

$$\begin{aligned} \nabla_a J &= \sum_{k=0}^{N-1} x_k^2 \lambda_{k+1}, \\ \nabla_b J &= - \sum_{k=0}^{N-1} x_k p_{k+1}; \end{aligned} \quad (17)$$

$$(k = 0, 1, 2, \dots, N-1).$$

在仿真试验中, 观测数据的采集过程采用与 (a) 中辨识 Logistic 映射系统未知参数相同的方法. 模式初始状态量 (x_0, y_0) 取为 $(0.5423, 0.1871)$, 其是 Henón 映射混沌系统状态中一点. 离散形式的目标泛函表示如下:

$$\begin{aligned} J(\boldsymbol{\theta}) &= \frac{1}{2} \sum_{k=1}^N [(x_k - x_k^{\text{obs}})^2 \\ &+ (y_k - y_k^{\text{obs}})^2]. \end{aligned} \quad (18)$$

在获得 Henón 映射系统的伴随方程、目标泛函及其梯度表达式后, 利用标准下降算法可以对 $\boldsymbol{\theta}$ 进行最优估计, 未知参数的迭代初值设定为 $\boldsymbol{\theta}^0 = (2.0, -0.8)$. 在对 Henón 映射系统未知参数进行变分迭代估计过程中, 参数向量 $\boldsymbol{\theta}$ 的每次更新, 都要对 Henón 映射的状态控制方程 (14) 式进行一次正向 ($k = 0 \rightarrow N$) 迭代, 以便计算目标泛函 $J(\boldsymbol{\theta})$ 的大小, 同时对 Henón 映射的伴随方程 (16) 式进行一次逆向 ($k = N \rightarrow 0$) 迭代, 以便计算目标泛函梯度 $\nabla_{\boldsymbol{\theta}} J$ 的值.

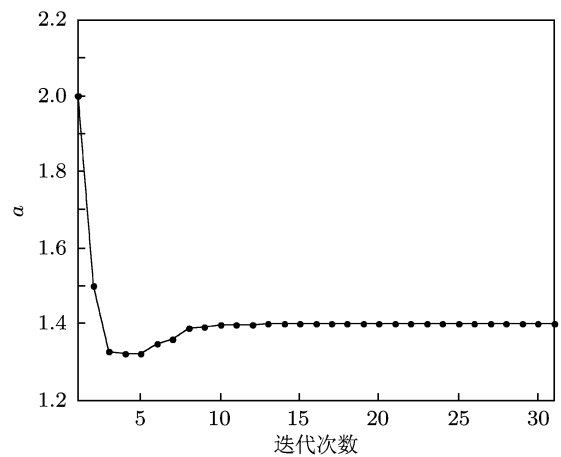


图 4 参数 a 的辨识结果曲线图

利用图 1 中的算法流程可以得到 $\boldsymbol{\theta} = (a, b)^T$ 的迭代估计序列值, 图 4 至图 7 显示了在不存在观测噪声 (即 $\sigma_o = 0$) 时的仿真结果. 图 4 和图 5 显示的是两个未知参数在迭代估计过程中值的变化;

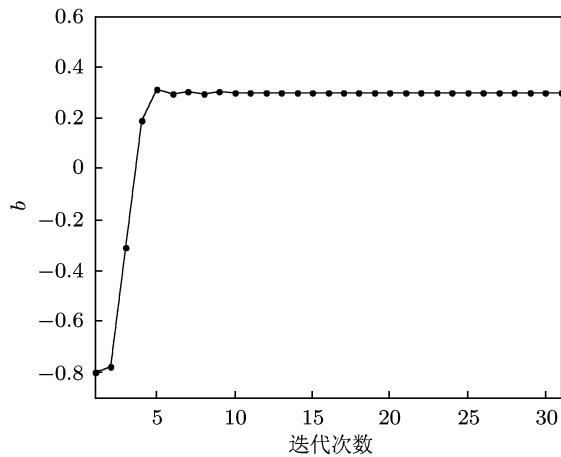
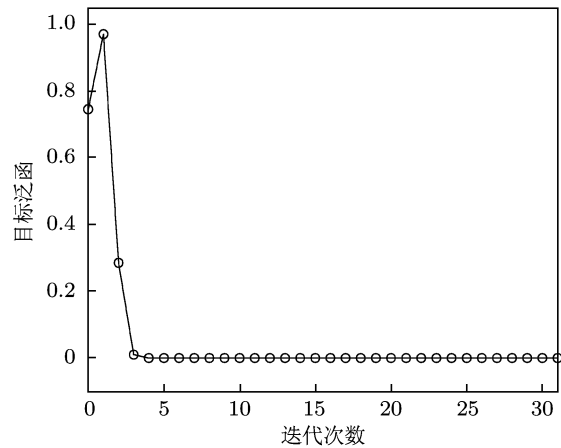
图5 参数 b 的辨识结果曲线图

图6 目标泛函变化曲线图

图6表示的是目标泛函值的下降图. 从图中可以看到, 未知参数向量 θ 在经历 31 次迭代后已经非常接近真实值, 同时目标泛函基本上不再下降, 达到收敛. 分析图中所示的试验结果可得如下结论: 采用离散变分方法估计 Henón 映射混沌系统的未知参数具有较高的准确性, 所有参数的估计精度达到了 $O(10^{-6})$. 另外离散变分方法的优势是通过伴随方程能够准确的计算出目标泛函梯度, 因此使目标泛函能快速下降, 参数辨识值只经过几步就非常接近真实值. 表2给出了在不同观测噪声水平情况下离散变分方法对 Henón 非线性映射系统中未知参数辨识的结果, 从表中可以看出, 在不存在观测误差时, 未知参数的估计精度最高. 由于模式不存在离散误差, 各个参数的估计结果可以精确到小数点后第六位, 表明了离散变分方法估计非线性映射系

统未知参数的有效性. 随着观测误差增加, 参数的估计精度下降, 但辨识值仍然收敛到真值附近, 说明离散变分方法具有较好的抗噪声性能.

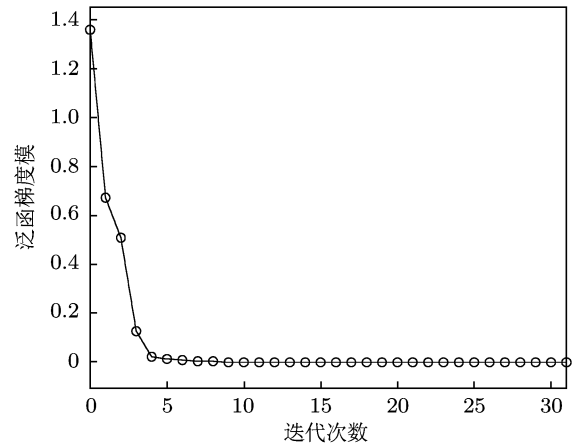


图7 泛函梯度模收敛曲线图

表2 不同观测噪声水平情况下 Henón 映射系统的参数估计结果

σ_o	0.0	0.1	0.2
参数 a	1.399999	1.3995125	1.3976859
参数 b	0.299999	0.2993769	0.2988312
迭代次数	28	39	57
目标泛函	3.0173×10^{-16}	0.0046178	0.0325870

5 结论

本文利用离散变分方法来解决非线性映射系统中未知参数辨识问题. 首先引入拉格朗日乘子向量序列将非线性映射系统的状态控制方程引入到目标泛函中, 将原来以非线性离散系统为约束的最优化问题转换为无约束最优化问题; 接着利用离散变分原理导出了非线性映射系统的伴随方程和目标泛函关于未知参数梯度的通用公式; 然后给出了离散变分方法估计非线性映射中未知参数的算法流程. 最后对典型的 Logistic 映射和 Henón 映射的未知物理参数进行了辨识, 数值仿真结果验证了方法的有效性: 在不存在观测噪声时方法具有很高的估计精度; 当存在观测误差时, 迭代估计值序列仍能收敛到真值附近, 说明离散变分方法具有较好的抗噪声性能.

- [1] Pecora L M, Carroll T L 1990 *Phys. Rev. Lett.* **64** 821
- [2] Li G H 2005 *Chaos. Soliton and Fractals.* **32** 1454
- [3] Tu L L 2011 *Chin. Phys. B* **20** 030504
- [4] Jia L X, Dai H, Hui M 2010 *Chin. Phys. B* **19** 100501
- [5] Zhang R X, Yang S P 2009 *Chin. Phys. B* **18** 3295
- [6] He M F, Mu Y M, Zhao L Z 2000 *Acta Phys. Sin.* **49** 830 (in Chinese) [贺明峰, 穆云明, 赵立中 2000 物理学报 **49** 830]
- [7] 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]
- [8] 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]
- [9] He Q, Wang L, Liu B 2007 *Chaos. Soliton and Fractals.* **34** 654
- [10] Li L X, Peng H P, Yang Y X, Wang X D 2007 *Acta Phys. Sin.* **56** 51 (in Chinese) [李丽香, 彭海朋, 杨义先, 王向东 2007 物理学报 **56** 51]
- [11] Li N Q, Pan W, Yan L S, Luo B 2011 *Chin. Phys. B* **20** 060502
- [12] Huang S X, Wu R S 2001 *Mathematical Physics Problems in Atmosphere Science* (Beijing: Meteorology Press) (in Chinese) [黄思训, 伍荣生 2001 大气科学中的数学物理问题 (北京: 气象出版社)]
- [13] Cai G L, Huang J J 2006 *Acta Phys. Sin.* **55** 3997 (in Chinese) [蔡国梁, 黄娟娟 2006 物理学报 **55** 3997]
- [14] He J H 2008 *Int. J. Modern. Phys. B* **22** 3487
- [15] He J H 1997 *Int. J. Turbo. Jet-Eng.* **14** 23
- [16] He J H 2000 *Appl. Math. Mech.* **21** 797
- [17] He J H 2001 *Int. J. Nonlin. Sci. Numer.* **2** 309
- [18] He J H, Lee E W M 2009 *Phys. Lett. A* **373** 1644

Parameter identification of map chaotic system with discrete variational method*

Cao Xiao-Qun[†] Zhang Wei-Min Song Jun-Qiang Zhu Xiao-Qian Zhao Jun

(School of Computer Science, National University of Defense Technology, Changsha 410073, China)

(Received 18 March 2011; revised manuscript received 3 May 2011)

Abstract

In this paper a method is proposed to estimate the unknown parameters of nonlinear map based on the discrete variational principle, which can be applied to all map chaotic systems governed by the following equation: $\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k, \boldsymbol{\theta})$. Numerical simulations on the well-known Logistic map and Henón map are conducted and all unknown parameters of the two discrete chaotic systems are identified respectively. Simulation results show that the discrete variational method is effective for parameter identification of the map chaotic system.

Keywords: chaotic system, parameter identification, discrete variational method, adjoint equation

PACS: 05.45.Pq, 04.20.Fy

* Project supported by the National Natural Science Foundation of China (Grant Nos. 41105063, 61070041).

[†] E-mail: caoxiaoqun@nudt.edu.cn