

# 基于二阶离散变分方法的非线性映射参数估计\*

曹小群†

(国防科学技术大学计算机学院, 长沙 410073)

(2012年11月11日收到; 2012年12月27日收到修改稿)

提出一种估计非线性映射未知参数的二阶离散变分方法. 首先针对非线性离散混沌系统, 利用变分方法导出了伴随方程和目标泛函梯度, 以此为基础利用二阶离散变分方法给出了二阶伴随方程和精确计算 Hessian 矩阵 - 向量乘积的显式表达式; 其次设计了估计非线性映射未知参数的新算法, 并以此对 Hyperhenón 映射和二维抛物映射中的未知参数进行了精确的估计. 数值仿真结果表明了该方法的有效性和优点.

**关键词:** 非线性映射, 参数估计, 二阶离散变分方法, 伴随方程

**PACS:** 05.45.Pq, 04.20.Fy, 45.10.Db

**DOI:** 10.7498/aps.62.080506

## 1 引言

近年来混沌控制和同步已经成为非线性科学研究中的重要方向之一, 在已知精确参数的前提下相应的理论和方法已经得到广泛发展<sup>[1-3]</sup>. 在实际中, 由于混沌系统的复杂性, 某些参数经常无法测量或确定; 或者出于保密通信等特殊原因, 系统的某些参数不可知. 此时, 要实现混沌系统的控制或同步, 上述方法有相当的局限性. 因此参数估计是混沌控制与同步中首先必须解决的课题, 具有重要的现实意义. 文献 [4] 引入基于反馈的同步和自适应控制方法, 对若干混沌系统进行了参数估计; 文献 [5] 通过最小化平均同步误差, 对一个给定的混沌动态系统进行了参数估计; 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] 提出了一种基于变分原理的混沌系统参数估计方法, 并对 Lorenz 系统和超混沌 Chen 系统的参数进行了估计. 近年来一系列智能优化算法已经成功应用于混沌系统参数估计中, 如改进粒子群优化算法<sup>[9]</sup>、混沌蚂蚁群算法<sup>[10]</sup>、混合差分进化算法<sup>[11]</sup>和混合交叉进化算法<sup>[12]</sup>等.

非线性映射是混沌控制和同步领域中一类重要的混沌系统, 其未知参数估计也是一个必须解决的关键问题. 非线性映射参数估计是离散动力系统研究中典型的反问题, 和正问题不同, 它一般通过观测数据来反求系统的参数或初/边值. 因为观测中不可避免地包含有噪声以及系统固有的非线性, 所以反问题求解过程中经常出现不适定现象: 解不一定存在、即使解存在也不惟一、在解存在惟一条件下也不稳定 (即解不连续依赖于观测数据)<sup>[13,14]</sup>. 因此非线性映射系统参数估计问题的求解通常采用特殊方法<sup>[8-13]</sup>. 本文基于二阶离散变分原理<sup>[14-18]</sup>来估计非线性映射中的未知参数. 首先引入拉格朗日乘子向量, 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, 并利用离散变分方法导出了伴随方程和目标泛函梯度表达式; 其次采用二阶离散变分方法导出了二阶伴随方程和 Hessian 矩阵 - 向量乘积的通用公式; 然后设计了二阶离散变分方法估计非线性映射中未知参数的算法流程, 并以此估计了 Hyperhenón 映射<sup>[19]</sup>和二维抛物映射<sup>[20]</sup>中的未知参数. 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.

\* 国家自然科学基金 (批准号: 41105063) 资助的课题.

† 通讯作者. E-mail: caoxiaoqun@nudt.edu.cn

## 2 问题描述

考虑如下  $n$  维非线性映射系统的未知参数估计问题:

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

其中  $\mathbf{x}_0$  和  $\mathbf{x}_k \in \mathbb{R}^n$  分别表示  $n$  维初值和状态向量,  $\mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta}) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  表示非线性映射的模式函数向量,  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)^T \in \mathbb{R}^m$  表示需要估计的未知物理参数向量. 目标是利用非线性映射的观测数据准确辨识出不可测的未知参数向量  $\boldsymbol{\theta}$ , 在观测噪声满足高斯分布的情况下, 该问题可通过对下面具有最小二乘意义的目标泛函求极小值而解决:

$$\begin{aligned} J(\boldsymbol{\theta}) &= \sum_{k=1}^N J_k(\boldsymbol{\theta}) \\ &= \frac{1}{2} \sum_{k=1}^N \langle [\mathbf{y}_k^o - \mathbf{h}(\mathbf{x}_k)], \mathbf{R}_k^{-1} [\mathbf{y}_k^o - \mathbf{h}(\mathbf{x}_k)] \rangle, \end{aligned} \quad (2)$$

其中  $\mathbf{y}_k^o$  和  $\mathbf{R}_k (k=1, 2, 3, \dots, N)$  分别表示映射系统的观测向量和观测误差协方差矩阵,  $\langle \cdot, \cdot \rangle$  表示欧氏空间的内积. 显然, 非线性映射的未知参数估计问题变成在离散状态方程 (1) 式约束下的目标泛函最优化问题. 但是, 由于非线性映射物理系统的强非线性和不稳定性将导致目标泛函  $J(\boldsymbol{\theta})$  中可能存在多个局部极值点, 从而使得仅仅依靠一阶导数 (即目标泛函梯度  $\nabla_{\boldsymbol{\theta}} J = (\nabla_{\theta_1} J, \nabla_{\theta_2} J, \dots, \nabla_{\theta_m} J)^T$ ) 信息很难准确估计未知参数值, 进而使得如何精确计算目标泛函的高阶信息 (例如, 与 Hessian 矩阵有关的信息) 就成为一个重要问题. 同时考虑到未知参数向量  $\boldsymbol{\theta}$  是通过非线性映射系统与目标泛函间接相关的, 因此要准确地得到  $J$  关于  $\boldsymbol{\theta}$  的二阶导数信息是非常困难的. 本文利用二阶离散变分方法给出了切线性方程、二阶伴随方程和准确计算 Hessian 矩阵 ( $\nabla_{\boldsymbol{\theta}}^2 J$ )- 向量乘积的公式, 既可以避免 Hessian 矩阵直接求解的巨大计算代价, 又可以加快最优化算法的收敛速度和提高未知参数的估计精度.

## 3 非线性映射参数估计的二阶变分方法

### 3.1 一阶变分方法

首先引入  $n$  维拉格朗日乘子向量  $\boldsymbol{\lambda}_k \in \mathbb{R}^n$  ( $k=0, 1, 2, \dots, N$ ), 将以非线性映射系统为约束条

件的最优化问题 (2) 式转化为下面 (3) 式表示的无约束最优化问题:

$$J(\boldsymbol{\theta}) = \sum_{k=1}^N J_k(\boldsymbol{\theta}) + \sum_{k=0}^{N-1} \langle \boldsymbol{\lambda}_{k+1}, \mathbf{x}_{k+1} - \mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta}) \rangle. \quad (3)$$

$J_k(\boldsymbol{\theta})$  的一阶变分可由下面的表达式给出:

$$\begin{aligned} \delta J_k(\boldsymbol{\theta}) &= \frac{1}{2} \delta [(\mathbf{y}_k^o - \mathbf{h}(\mathbf{x}_k))^T \mathbf{R}_k^{-1} (\mathbf{y}_k^o - \mathbf{h}(\mathbf{x}_k))] \\ &= \langle -\mathbf{f}_k, \delta \mathbf{x}_k \rangle, \end{aligned} \quad (4)$$

其中

$$\mathbf{f}_k = (\mathbf{D}_{\mathbf{h}}^{\mathbf{x}_k})^T \mathbf{R}_k^{-1} [\mathbf{y}_k^o - \mathbf{h}(\mathbf{x}_k)], \quad (5)$$

而  $\mathbf{D}_{\mathbf{h}}^{\mathbf{x}_k} \in \mathbb{R}^{m \times n}$  表示观测算子  $\mathbf{h}(\mathbf{x}_k)$  关于  $\mathbf{x}_k$  的 Jacobian 矩阵. 结合 (2) 式和 (4) 式可得

$$\sum_{k=1}^N \delta J_k(\boldsymbol{\theta}) = \sum_{k=1}^N \langle -\mathbf{f}_k, \delta \mathbf{x}_k \rangle. \quad (6)$$

另外 (3) 式的一阶变分为

$$\begin{aligned} \delta J(\boldsymbol{\theta}) &= \sum_{k=1}^N \delta J_k(\boldsymbol{\theta}) \\ &+ \sum_{k=0}^{N-1} \left[ \langle \delta \boldsymbol{\lambda}_{k+1}, \mathbf{x}_{k+1} - \mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta}) \rangle \right. \\ &\left. + \langle \boldsymbol{\lambda}_{k+1}, \delta \mathbf{x}_{k+1} - \delta \mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta}) \rangle \right], \end{aligned} \quad (7)$$

上式中  $\delta \mathbf{x}_k$  和  $\delta \boldsymbol{\lambda}_k$  分别表示非线性映射状态向量和拉格朗日乘子向量的一阶变分. 由 (1) 式显然有  $\langle \delta \boldsymbol{\lambda}_{k+1}, \mathbf{x}_{k+1} - \mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta}) \rangle = 0$ . 而目标泛函  $J(\boldsymbol{\theta})$  和向量函数  $\mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta})$  的一阶变分按定义可以进一步写为

$$\delta J(\boldsymbol{\theta}) = \langle \nabla_{\boldsymbol{\theta}} J, \delta \boldsymbol{\theta} \rangle,$$

$$\delta \mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta}) = (\mathbf{D}_{\mathbf{M}}^{\mathbf{x}_k}) \delta \mathbf{x}_k + (\mathbf{D}_{\mathbf{M}}^{\boldsymbol{\theta}}) \delta \boldsymbol{\theta}, \quad (8)$$

其中  $\mathbf{D}_{\mathbf{M}}^{\mathbf{x}_k} \in \mathbb{R}^{n \times n}$  和  $\mathbf{D}_{\mathbf{M}}^{\boldsymbol{\theta}} \in \mathbb{R}^{m \times n}$  分别表示模式向量函数  $\mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta})$  关于  $\mathbf{x}_k$  和  $\boldsymbol{\theta}$  的 Jacobian 矩阵. 将 (8) 式中第二式代入 (7) 式右边第三项, 同时利用恒等式  $\sum_{k=0}^{N-1} \langle \boldsymbol{\lambda}_{k+1}, \delta \mathbf{x}_{k+1} \rangle = \sum_{k=1}^N \langle \boldsymbol{\lambda}_k, \delta \mathbf{x}_k \rangle$  和伴随性质  $\langle \boldsymbol{\lambda}, \mathbf{A} \mathbf{x} \rangle = \langle \mathbf{A}^T \boldsymbol{\lambda}, \mathbf{x} \rangle$  可得

$$\begin{aligned} &\sum_{k=0}^{N-1} \langle \boldsymbol{\lambda}_{k+1}, \delta \mathbf{x}_{k+1} - \delta \mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta}) \rangle \\ &= \sum_{k=0}^{N-1} \langle \boldsymbol{\lambda}_k - (\mathbf{D}_{\mathbf{M}}^{\mathbf{x}_k})^T \boldsymbol{\lambda}_{k+1}, \delta \mathbf{x}_k \rangle \\ &- \left\langle \sum_{k=0}^{N-1} (\mathbf{D}_{\mathbf{M}}^{\boldsymbol{\theta}})^T \boldsymbol{\lambda}_{k+1}, \delta \boldsymbol{\theta} \right\rangle \\ &+ \langle \boldsymbol{\lambda}_{k+1}, (\mathbf{D}_{\mathbf{M}}^{\mathbf{x}_k}) \delta \mathbf{x}_k \rangle \Big|_{k=0}^{k=N}. \end{aligned} \quad (9)$$

将 (8) 式中第一式与 (6) 式和 (9) 式代入 (7) 式并合并同类项. 考虑到一阶变分  $\delta \mathbf{x}_k$ ,  $\delta \boldsymbol{\lambda}_k$  和  $\delta \boldsymbol{\theta}$  可以取任意值, 且  $\delta \mathbf{x}_0 = 0$ ,  $\delta \mathbf{x}_N \neq 0$ , 进而可导出下面的方程组:

$$\begin{aligned} \boldsymbol{\lambda}_k &= [D_M^{\mathbf{x}_k}(\mathbf{x}_k, \boldsymbol{\theta})]^T \boldsymbol{\lambda}_{k+1} + \mathbf{f}_k, \\ \boldsymbol{\lambda}_{N+1} &= 0, \\ \nabla_{\boldsymbol{\theta}} J &= - \sum_{k=0}^{N-1} [D_M^{\boldsymbol{\theta}}(\mathbf{x}_k, \boldsymbol{\theta})]^T \boldsymbol{\lambda}_{k+1}, \\ &(k = N, \dots, 3, 2, 1), \end{aligned} \quad (10)$$

(10) 式中的第一式称为伴随方程,  $\boldsymbol{\lambda}_k$  表示  $k$  时次的伴随状态向量; 从表达式易知, 伴随方程积分之前必须求得非线性映射的状态向量序列  $\mathbf{x}_k$ , 而且需按照顺序  $k = N, N-1, \dots, 1$  进行逆向积分. (10) 式中第二式给出了伴随方程的初值条件  $\boldsymbol{\lambda}_{N+1} = 0$ , 因为是逆向积分, 所以初值条件在末端时刻  $k = N+1$  给出. (10) 式中的第三式是目标泛函关于未知参数向量梯度的表达式, 计算梯度之前需要确定状态向量  $\mathbf{x}_k$  和伴随向量  $\boldsymbol{\lambda}_k$ . 在本小节中利用一阶变分方法获得了伴随方程和目标泛函梯度的公式, 通过计算目标泛函梯度向量的值, 进而利用基于一阶导数信息的最优化算法 (共轭梯度方法和拟牛顿方法) 可以估计未知参数 [8,13]. 但是, 在混沌系统参数估计中, 完全依赖一阶导数信息的最优化算法通常收敛速度较慢, 对于具有较高维数的强非线性混沌系统尤其如此. 因此为了加快收敛速度, 有必要研究能同时利用导数和精确 Hessian 矩阵信息的非线性映射未知参数估计方法.

### 3.2 二阶变分方法

为了导出 Hessian 矩阵 - 向量乘积及相关的二阶伴随方程的通用表达式, 首先对 (1) 式求一阶和二阶变分, 则分别得到相应的切线性系统 (TLS):

$$\begin{aligned} \delta \mathbf{x}_{k+1} &= D_M^{\mathbf{x}_k}(\mathbf{x}_k, \boldsymbol{\theta}) \delta \mathbf{x}_k + D_M^{\boldsymbol{\theta}}(\mathbf{x}_k, \boldsymbol{\theta}) \delta \boldsymbol{\theta}, \\ \delta \mathbf{x}_0 &= 0, \end{aligned} \quad (11)$$

和二阶切线性系统:

$$\begin{aligned} \delta^2 \mathbf{x}_{k+1} &= \delta [D_M^{\mathbf{x}_k}] \delta \mathbf{x}_k + D_M^{\mathbf{x}_k} \delta^2 \mathbf{x}_k \\ &\quad + \delta [D_M^{\boldsymbol{\theta}}] \delta \boldsymbol{\theta} + D_M^{\boldsymbol{\theta}} \delta^2 \boldsymbol{\theta}, \\ \delta^2 \mathbf{x}_0 &= 0, \end{aligned} \quad (12)$$

(11) 式和 (12) 式中的  $\delta \mathbf{x}_k$  和  $\delta \boldsymbol{\theta}$  表示一阶变分或切线性向量, 而  $\delta^2 \mathbf{x}_k$  和  $\delta^2 \boldsymbol{\theta}$  表示二阶变分或二阶切线性向量.

其次对 (10) 式中的伴随方程直接求一阶变分, 并将伴随向量  $\boldsymbol{\lambda}_k$  的一阶变分  $\delta \boldsymbol{\lambda}_k$  用  $\boldsymbol{\zeta}_k$  表示 (即  $\boldsymbol{\zeta}_k = \delta \boldsymbol{\lambda}_k$ ), 则可得

$$\begin{aligned} \boldsymbol{\zeta}_k - \delta [(D_M^{\mathbf{x}_k})^T] \boldsymbol{\lambda}_{k+1} - (D_M^{\mathbf{x}_k})^T \boldsymbol{\zeta}_{k+1} - \delta \mathbf{f}_k &= 0 \\ &(k = 1, 2, \dots, N-1, N). \end{aligned} \quad (13)$$

$\boldsymbol{\zeta}_k$  称为二阶伴随向量, 考虑到  $\boldsymbol{\lambda}_{N+1} = 0$ , 进而可确定二阶伴随量的末端值为  $\boldsymbol{\zeta}_{N+1} = 0$ . 下面推导 Jacobian 矩阵  $D_h^{\mathbf{x}_k}$ ,  $D_M^{\mathbf{x}_k}$  和  $D_M^{\boldsymbol{\theta}}$  的一维变分表达式.

因为

$$\begin{aligned} \delta (D_h^{\mathbf{x}_k}) &= \delta \mathbf{x}_k \begin{bmatrix} (\nabla_{\mathbf{x}_k} h_1)^T \\ (\nabla_{\mathbf{x}_k} h_2)^T \\ \vdots \\ (\nabla_{\mathbf{x}_k} h_N)^T \end{bmatrix} = \begin{bmatrix} (\delta \mathbf{x}_k)^T \nabla_{\mathbf{x}_k}^2 h_1 \\ (\delta \mathbf{x}_k)^T \nabla_{\mathbf{x}_k}^2 h_2 \\ \vdots \\ (\delta \mathbf{x}_k)^T \nabla_{\mathbf{x}_k}^2 h_N \end{bmatrix} \\ &= [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{h}] \in \mathbb{R}^{N \times N}, \end{aligned}$$

所以有

$$\delta [(D_h^{\mathbf{x}_k})^T] = [\delta (D_h^{\mathbf{x}_k})]^T = [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{h}]^T. \quad (14)$$

同理可得非线性模式函数向量 Jacobian 矩阵的一维变分表达式:

$$\delta (D_M^{\mathbf{x}_k}) = [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 M] + [\delta \boldsymbol{\theta}, \nabla_{\mathbf{x}_k \boldsymbol{\theta}}^2 M], \quad (15)$$

$$\delta (D_M^{\boldsymbol{\theta}}) = [\delta \mathbf{x}_k, \nabla_{\boldsymbol{\theta} \mathbf{x}_k}^2 M] + [\delta \boldsymbol{\theta}, \nabla_{\boldsymbol{\theta}}^2 M]. \quad (16)$$

将 (15) 和 (16) 式代入 (12) 和 (13) 式, 可得:

$$\begin{aligned} \delta^2 \mathbf{x}_{k+1} &= ([\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 M] + [\delta \boldsymbol{\theta}, \nabla_{\mathbf{x}_k \boldsymbol{\theta}}^2 M]) \delta \mathbf{x}_k \\ &\quad + D_M^{\mathbf{x}_k} \delta^2 \mathbf{x}_k + ([\delta \mathbf{x}_k, \nabla_{\boldsymbol{\theta} \mathbf{x}_k}^2 M] \\ &\quad + [\delta \boldsymbol{\theta}, \nabla_{\boldsymbol{\theta}}^2 M]) \delta \boldsymbol{\theta} + D_M^{\boldsymbol{\theta}} \delta^2 \boldsymbol{\theta}, \end{aligned} \quad (17)$$

$$\begin{aligned} \boldsymbol{\zeta}_k &= [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 M]^T \boldsymbol{\lambda}_{k+1} + [\delta \boldsymbol{\theta}, \nabla_{\mathbf{x}_k \boldsymbol{\theta}}^2 M]^T \boldsymbol{\lambda}_{k+1} \\ &\quad + (D_M^{\mathbf{x}_k})^T \boldsymbol{\zeta}_{k+1} + \delta \mathbf{f}_k. \end{aligned} \quad (18)$$

(17) 式被称为非线性映射的二阶切线性模式, 刻画的是二阶切线性向量从时刻  $k=0$  到  $k=N$  的演化; 相应地, (18) 式称为二阶伴随模式, 刻画的是二阶伴随向量从时刻  $k=N$  到  $k=0$  的演化. 对 (17) 和 (18) 式左右两端分别用  $-\boldsymbol{\lambda}_{k+1}$  和  $\delta \mathbf{x}_k$  求内积, 并将两式所得到的结果相加, 则可得:

$$\begin{aligned} &\langle -\boldsymbol{\lambda}_{k+1}, \delta^2 \mathbf{x}_{k+1} \rangle + \langle \delta \mathbf{x}_k, \boldsymbol{\zeta}_k \rangle \\ &= \langle -\boldsymbol{\lambda}_{k+1}, [D_M^{\mathbf{x}_k} \delta^2 \mathbf{x}_k] \rangle \\ &\quad - \langle \boldsymbol{\lambda}_{k+1}, [\delta \boldsymbol{\theta}, \nabla_{\mathbf{x}_k \boldsymbol{\theta}}^2 M] \delta \mathbf{x}_k \rangle \\ &\quad - \langle \boldsymbol{\lambda}_{k+1}, D_M^{\mathbf{x}_k} \delta^2 \mathbf{x}_k \rangle - \langle \boldsymbol{\lambda}_{k+1}, [\delta \mathbf{x}_k, \nabla_{\boldsymbol{\theta} \mathbf{x}_k}^2 M] \delta \boldsymbol{\theta} \rangle \\ &\quad - \langle \boldsymbol{\lambda}_{k+1}, [\delta \boldsymbol{\theta}, \nabla_{\boldsymbol{\theta}}^2 M] \delta \boldsymbol{\theta} \rangle - \langle \boldsymbol{\lambda}_{k+1}, D_M^{\boldsymbol{\theta}} \delta^2 \boldsymbol{\theta} \rangle \end{aligned}$$

$$\begin{aligned}
 & + \langle \delta \mathbf{x}_k, [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{M}]^T \boldsymbol{\lambda}_{k+1} \rangle \\
 & + \langle \delta \mathbf{x}_k, [\delta \boldsymbol{\theta}, \nabla_{\mathbf{x}_k}^2 \mathbf{M}]^T \boldsymbol{\lambda}_{k+1} \rangle \\
 & + \langle \delta \mathbf{x}_k, (\mathbf{D}_M^{\mathbf{x}_k})^T \boldsymbol{\zeta}_{k+1} \rangle + \langle \delta \mathbf{x}_k, \delta \mathbf{f}_k \rangle. \quad (19)
 \end{aligned}$$

为了简化 (19) 式, 利用伴随性质后有以下等式成立:

$$\begin{aligned}
 & \langle -\boldsymbol{\lambda}_{k+1}, [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{M}] \delta \mathbf{x}_k \rangle \\
 & + \langle \delta \mathbf{x}_k, [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{M}]^T \boldsymbol{\lambda}_{k+1} \rangle = 0, \\
 & \langle -\boldsymbol{\lambda}_{k+1}, [\delta \boldsymbol{\theta}, \nabla_{\mathbf{x}_k}^2 \mathbf{M}] \delta \mathbf{x}_k \rangle \\
 & + \langle \delta \mathbf{x}_k, [\delta \boldsymbol{\theta}, \nabla_{\mathbf{x}_k}^2 \mathbf{M}]^T \boldsymbol{\lambda}_{k+1} \rangle = 0. \quad (20)
 \end{aligned}$$

另一方面, 由一阶切线性模式和一阶伴随模式可导出以下等式:

$$\begin{aligned}
 \mathbf{D}_M^{\mathbf{x}_k} \delta \mathbf{x}_k & = \delta \mathbf{x}_{k+1} - \mathbf{D}_M^\theta \delta \boldsymbol{\theta}, \\
 [\mathbf{D}_M^{\mathbf{x}_k}]^T \boldsymbol{\lambda}_{k+1} & = \boldsymbol{\lambda}_k - \mathbf{f}_k. \quad (21)
 \end{aligned}$$

将 (20) 式和 (21) 式都代入 (19) 式, 进行化简后可得:

$$\begin{aligned}
 & \langle \delta \mathbf{x}_k, \delta \mathbf{f}_k \rangle + \langle \mathbf{f}_k, \delta^2 \mathbf{x}_k \rangle \\
 & - \langle \boldsymbol{\lambda}_{k+1}, [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{M}] \delta \boldsymbol{\theta} \rangle \\
 & - \langle \boldsymbol{\lambda}_{k+1}, [\delta \boldsymbol{\theta}, \nabla_{\mathbf{x}_k}^2 \mathbf{M}] \delta \mathbf{x}_k \rangle \\
 & - \langle (\mathbf{D}_M^\theta)^T \boldsymbol{\lambda}_{k+1}, \delta^2 \boldsymbol{\theta} \rangle \\
 & - \langle \delta \boldsymbol{\theta}, (\mathbf{D}_M^\theta)^T \boldsymbol{\zeta}_{k+1} \rangle = 0. \quad (22)
 \end{aligned}$$

将链式法则应用到 (4) 式上, 可以获得  $J_k(\boldsymbol{\theta})$  的二阶变分表达式:

$$\delta^2 J_k(\boldsymbol{\theta}) = \langle -\delta \mathbf{f}_k, \delta \mathbf{x}_k \rangle + \langle -\mathbf{f}_k, \delta^2 \mathbf{x}_k \rangle. \quad (23)$$

另外对  $\delta J(\boldsymbol{\theta}) = \langle \nabla_{\boldsymbol{\theta}} \mathbf{J}, \delta \boldsymbol{\theta} \rangle$  两边进行一阶变分运算, 可以导出下式:

$$\delta^2 J(\boldsymbol{\theta}) = \langle \nabla_{\boldsymbol{\theta}}^2 J \delta \boldsymbol{\theta}, \delta \boldsymbol{\theta} \rangle + \langle \nabla_{\boldsymbol{\theta}} J, \delta^2 \boldsymbol{\theta} \rangle. \quad (24)$$

将 (10) 式中目标泛函梯度表达式代入 (24) 式, 同时对 (23) 式左右两边从  $k=1$  到  $N$  求和, 再考虑到  $\delta^2 J(\boldsymbol{\theta}) = \sum_{k=1}^N \delta^2 J_k(\boldsymbol{\theta})$ , 进而得到:

$$\begin{aligned}
 & \sum_{k=1}^N (\langle -\delta \mathbf{f}_k, \delta \mathbf{x}_k \rangle + \langle -\mathbf{f}_k, \delta^2 \mathbf{x}_k \rangle) \\
 & = \langle \nabla_{\boldsymbol{\theta}}^2 J \delta \boldsymbol{\theta}, \delta \boldsymbol{\theta} \rangle - \sum_{k=0}^{N-1} \langle (\mathbf{D}_M^\theta)^T \boldsymbol{\lambda}_{k+1}, \delta^2 \boldsymbol{\theta} \rangle. \quad (25)
 \end{aligned}$$

对 (22) 式左右两边从  $k=0$  到  $N-1$  求和, 并与 (25) 式相加, 化简后得到

$$\left\langle -\sum_{k=0}^{N-1} \left\{ [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{M}]^T \boldsymbol{\lambda}_{k+1} \right. \right.$$

$$\left. \left. + [\delta \boldsymbol{\theta}, \nabla_{\boldsymbol{\theta}}^2 \mathbf{M}]^T \boldsymbol{\lambda}_{k+1} + (\mathbf{D}_M^\theta)^T \boldsymbol{\zeta}_{k+1} \right\}, \delta \boldsymbol{\theta} \right\rangle = \langle \nabla_{\boldsymbol{\theta}}^2 J \delta \boldsymbol{\theta}, \delta \boldsymbol{\theta} \rangle, \quad (26)$$

考虑到一阶变分  $\delta \boldsymbol{\theta}$  的任意性, 从而有

$$\begin{aligned}
 & \nabla_{\boldsymbol{\theta}}^2 J \delta \boldsymbol{\theta} \\
 & = -\sum_{k=0}^{N-1} \left\{ [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{M}]^T \boldsymbol{\lambda}_{k+1} \right. \\
 & \quad \left. + [\delta \boldsymbol{\theta}, \nabla_{\boldsymbol{\theta}}^2 \mathbf{M}]^T \boldsymbol{\lambda}_{k+1} + (\mathbf{D}_M^\theta)^T \boldsymbol{\zeta}_{k+1} \right\}. \quad (27)
 \end{aligned}$$

(27) 式是目标泛函关于未知物理参数的 Hessian 矩阵 - 向量乘积的表达式, 其显式给出了可以改善最优化算法性能的二阶导数信息, 与 Hessian 矩阵相乘的向量  $\delta \boldsymbol{\theta}$  表示在每次最优化迭代过程中未知参数向量逼近物理参数真值时的增量. 在本小节中利用二阶变分方法给出了切线性方程、二阶伴随方程以及计算目标泛函关于未知物理参数的 Hessian 矩阵 - 向量乘积的公式. 从 (27) 式可知, 在计算  $\nabla_{\boldsymbol{\theta}}^2 J \delta \boldsymbol{\theta}$  的值之前需要确定状态向量  $\mathbf{x}_k$ , 一阶切线性向量  $\delta \mathbf{x}_k$ , 伴随向量  $\boldsymbol{\lambda}_k$  和二阶伴随向量  $\boldsymbol{\zeta}_k$  的值, 因此需要依次求解非线性映射正模式 (1), 切线性模式 (11), 一阶伴随模式 (10) 和二阶伴随模式 (18). 其中, 非线性映射系统和切线性模式需要进行正向积分, 一/二阶伴随模式需要进行逆向积分.

### 3.3 算法流程设计

在导出目标泛函梯度  $\nabla_{\boldsymbol{\theta}} J$  和 Hessian 矩阵 - 向量乘积  $\nabla_{\boldsymbol{\theta}}^2 J \delta \boldsymbol{\theta}$  的公式之后, 可以选择合适的牛顿型最优化算法对未知物理参数向量进行迭代估计. 本文中选用的是牛顿 - 共轭梯度方法 (Newton-conjugate gradient method, NCG)<sup>[21,22]</sup>. 在 NCG 方法中使用共轭梯度方法近似求解牛顿方程  $\nabla^2 J(\boldsymbol{\theta}_k) \mathbf{s}_k = -\nabla J(\boldsymbol{\theta}_k)$ , 当牛顿方程的残余量足够小 (即满足 (28) 式) 或者遇到负曲率方向时, 对共轭梯度方法能进行足够精确地截断.

$$\|\nabla^2 J(\boldsymbol{\theta}_k) \mathbf{s}_k + \nabla J(\boldsymbol{\theta}_k)\|_2 \leq \eta_k \|\nabla J(\boldsymbol{\theta}_k)\|_2, \quad (28)$$

在 (28) 式中  $\eta_k \in (0, 1)$ , 确定方向矢量  $\mathbf{s}_k$  后, 使用简单 Armijo 线搜索方法<sup>[21]</sup> 能计算出步长大小  $\alpha_k$ . 因为完整 Hessian 矩阵的直接计算和存储代价太高, 所以不适合快速算法. 而 NCG 方法的优点是只需要使用 Hessian 矩阵 - 向量乘积的值, 因此本文 3.2 部分中利用二阶离散变分方法给出的 (27) 式正好能满足此要求. 二阶离散变分方法估计非线性映

射系统的未知参数的具体计算流程如图 1 所示, 具体步骤说明如下.

**第 1 步** 设定未知物理参数向量的初始猜测值;

**第 2 步** 利用参数向量的新估计值, 对非线性映射模式 (1) 进行正向积分, 获得系统状态量的演化轨迹  $\mathbf{x}_k$  并加以存贮;

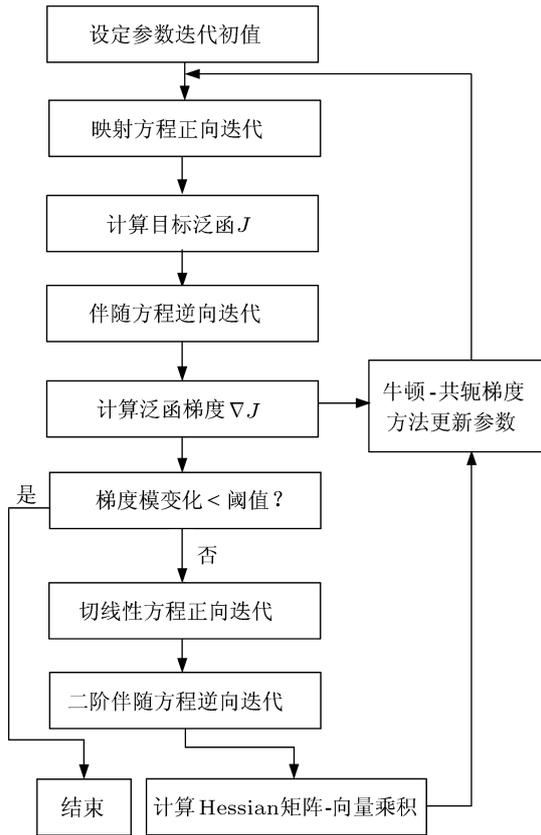


图 1 二阶变分方法估计非线性映射参数的流程图

**第 3 步** 利用  $\mathbf{x}_k$  和  $\mathbf{y}_k^o$  根据 (2) 式计算目标泛函  $J(\theta)$  的值;

**第 4 步** 利用  $\mathbf{x}_k$  和  $\mathbf{y}_k^o$  和伴随初值条件, 对伴随模式从  $k = N$  到  $k = 1$  进行逆向积分, 求得伴随向量的演化轨迹  $\lambda_k$  并加以存贮;

**第 5 步** 利用  $\mathbf{x}_k$  和  $\lambda_k$  根据 (10) 式中第三式求目标泛函关于各个未知物理参数的梯度向量值;

**第 6 步** 如果满足程序终止条件 (如达到所要求的收敛精度或预先指定的最大迭代次数), 则终止程序, 同时获得非线性系统的参数估计值; 若不满足, 则继续进行下面的步骤;

**第 7 步** 利用演化轨迹  $\mathbf{x}_k$  和切线性量初值条件, 对切线性模式 (11) 进行正向积分, 获得切线性量或扰动量的演化轨迹  $\delta\mathbf{x}_k$  并加以存贮;

**第 8 步** 利用  $\mathbf{x}_k, \delta\mathbf{x}_k, \lambda_k$  和二阶伴随初值条件  $\zeta_N = 0$ , 对二阶伴随模式 (18) 从  $k = N$  到  $k = 1$  进行反向积分, 求得二阶伴随向量的演化轨迹  $\zeta_k$  并加以存贮;

**第 9 步** 利用  $\mathbf{x}_k, \delta\mathbf{x}_k, \lambda_k$  和  $\zeta_k$  按照 (27) 式计算 Hessian 矩阵与未知参数向量增量之间的乘积;

**第 10 步** 将目标泛函梯度值和 Hessian 矩阵 - 向量乘积值输入到基于 NCG 方法的最优化程序中, 计算出方向向量和最优步长, 进而利用  $\theta_{k+1} = \theta_k + \alpha_k \mathbf{s}_k$  对未知物理参数向量进行迭代更新. 利用新估计值从第二步开始进行新一轮的迭代循环.

## 4 数值仿真结果及分析

### 4.1 Hyperhenón 映射系统的参数估计

首先以典型的 Hyperhenón 映射为例, 研究二阶离散变分方法估计非线性映射系统中未知物理参数的有效性. Hyperhenón 映射的控制方程可表示如下:

$$\begin{aligned} x_{k+1} &= a - x_k^2 - by_k, \\ y_{k+1} &= x_k, \quad z_{k+1} = y_k, \\ k &= 0, 1, 2, \dots \end{aligned} \quad (29)$$

当  $a = 1.76$  和  $b = 0.1$  时, 该系统处于超混沌状态 [19]. 状态向量表示为  $\mathbf{x} = (x, y, z)^T$ ,  $\theta = (a, b)^T$  表示未知物理参数向量. 目标是在获得  $\mathbf{x}$  的观测量  $\mathbf{y}_k^o = (x_k^o, y_k^o, z_k^o)^T$  的情况下, 辨识未知参数向量  $\theta$ . 通过比较 (29) 式与 (1) 式, 易知非线性模式函数向量可表示为  $\mathbf{M}(\mathbf{x}_k, \theta) = [a - x_k^2 - by_k, x_k, y_k]^T$ , 它的一阶变分为

$$\begin{aligned} \delta\mathbf{M}(\mathbf{x}_k, \theta) &= (\mathbf{D}_{\mathbf{M}}^{\mathbf{x}_k})\delta\mathbf{x}_k + (\mathbf{D}_{\mathbf{M}}^{\theta})\delta\theta \\ &= \begin{pmatrix} -2x_k & -b & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \delta x_k \\ \delta y_k \\ \delta z_k \end{pmatrix} \\ &\quad + \begin{pmatrix} 1 & -y_k \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \delta a \\ \delta b \end{pmatrix}. \end{aligned} \quad (30)$$

(30) 式中  $\delta\mathbf{x}_k = (\delta x_k, \delta y_k, \delta z_k)^T$  和  $\delta\theta = (\delta a, \delta b)^T$  分别表示  $\mathbf{x}_k$  和  $\theta$  的一阶变分. 根据 (11) 式可得切线

性方程组:

$$\begin{aligned} \delta x_{k+1} &= -2x_k \delta x_k - b \delta y_k + \delta a - y_k \delta b, \\ \delta y_{k+1} &= \delta x_k, \quad \delta z_{k+1} = \delta y_k, \\ \delta x_0 &= 0, \quad \delta y_0 = 0, \quad \delta z_0 = 0. \end{aligned} \quad (31)$$

首先引入与  $\mathbf{x}_k$  对应的伴随向量序列  $\boldsymbol{\lambda}_k = (\lambda_k, p_k, q_k)^T$ , 并将  $\mathbf{x}_k, \mathbf{y}_k^o, \boldsymbol{\lambda}_k, \boldsymbol{\theta}, (\mathbf{D}_M^{\mathbf{x}_k})^T$  和  $(\mathbf{D}_M^{\boldsymbol{\theta}})^T$  等的表达式代入 (10) 式中的第一和第三式, 同时考虑到伴随初始条件, 进而得到 Hyperhenón 超混沌映射系统的伴随方程和目标泛函关于未知参数的梯度公式, 分别如 (32) 和 (33) 式所示:

$$\begin{aligned} \lambda_k &= -2x_k \lambda_{k+1} + p_{k+1} + (x_k^o - x_k), \\ p_k &= -b \lambda_{k+1} + q_{k+1} + (y_k^o - y_k), \\ q_k &= (z_k^o - z_k), \\ (\lambda_{N+1}, p_{N+1}, q_{N+1}) &= 0, \\ (k &= 1, 2, \dots, N). \end{aligned} \quad (32)$$

$$\nabla_a J = - \sum_{k=0}^{N-1} \lambda_{k+1}, \quad \nabla_b J = \sum_{k=0}^{N-1} y_k \lambda_{k+1}. \quad (33)$$

其次将  $\boldsymbol{\lambda}_k$ ,

$$\begin{aligned} \boldsymbol{\zeta}_k &= (\zeta_k, \xi_k, \varsigma_k)^T, \\ \delta \mathbf{f}_k &= -(\delta x_k, \delta y_k, \delta z_k)^T, \\ [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{M}] &= \begin{pmatrix} -2\delta x_k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ [\delta \boldsymbol{\theta}, \nabla_{\boldsymbol{\theta}}^2 \mathbf{M}] &= \begin{pmatrix} 0 & -\delta b & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

和  $\mathbf{D}_M^{\mathbf{x}_k}$  代入 (18) 式中, 则得到 Hyperhenón 映射系统的二阶伴随方程:

$$\begin{aligned} \zeta_k &= -2x_k \zeta_{k+1} + \xi_{k+1} - 2\lambda_{k+1} \delta x_k - \delta x_k, \\ \xi_k &= -b \zeta_{k+1} + \varsigma_{k+1} - \delta y_k, \\ \varsigma_k &= -\delta w_k, \\ \zeta_{N+1} &= 0, \xi_{N+1} = 0, \varsigma_{N+1} = 0. \end{aligned} \quad (34)$$

最后将  $\boldsymbol{\lambda}_k, \boldsymbol{\zeta}_k, [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{M}] = \begin{pmatrix} 0 & -\delta y_k \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ ,  $[\delta \boldsymbol{\theta}, \nabla_{\boldsymbol{\theta}}^2 \mathbf{M}] = 0$  和  $\mathbf{D}_M^{\boldsymbol{\theta}}$  代入 (27) 式中, 得到 Hyper-

henón 映射的 Hessian 矩阵 - 向量乘积的计算公式:

$$\nabla_{\boldsymbol{\theta}}^2 J \delta \boldsymbol{\theta} = \left[ - \sum_{k=0}^{N-1} \boldsymbol{\zeta}_{k+1}, \sum_{k=0}^{N-1} (\lambda_{k+1} \delta y_k + y_k \boldsymbol{\zeta}_{k+1}) \right]^T. \quad (35)$$

在仿真试验中, 观测数据的采集过程如下: 首先设定物理参数的真实值, 让 Hyperhenón 映射自由演化, 在经历暂态过程之后任取混沌状态的一点作为初值, 并以此时刻作为初始时刻; 其次任其迭代演化至第  $N$  步, 可获得 Hyperhenón 映射的离散混沌状态序列  $\mathbf{x}_k (k = 0, 1, 2, \dots, N)$ . 在本试验中, 初始条件设定为混沌状态中一点:  $x_0 = 0.8350073$ ,  $y_0 = -0.8732880$  和  $z_0 = 1.6236088$ , 抽取前面 4 个状态量作为观测数据  $\mathbf{y}_k^o (k = 1, 2, \dots, N)$ ; 然后可在所选状态量上叠加高斯随机数作为观测噪声, 噪声的均值取为 0, 标准偏差取为  $\sigma_o$ . 因为在这种情况下观测误差协方差矩阵  $\mathbf{R}_k$  为  $3 \times 3$  大小的单位矩阵, 所以离散形式的目标泛函表示如下:

$$J(\boldsymbol{\theta}) = \frac{1}{2} \sum_{i=1}^N \|\mathbf{y}_k^o - \mathbf{x}_k\|^2. \quad (36)$$

根据伴随方程、目标泛函及其梯度表达式, 利用共轭梯度算法能对非线性映射的未知物理参数进行最优估计, 这里称之为“一阶变分方法”; 而根据一/二阶伴随方程、切线性方程、目标泛函、梯度及 Hessian 矩阵 - 向量乘积的表达式, 利用 NCG 算法也能对非线性映射的未知物理参数进行最优估计, 称之为“二阶变分方法”. 在两个方法中, 未知参数的迭代初值都设定为:  $\boldsymbol{\theta}^0 = (0, 0, 0)^T$ .

图 2 和图 3 中分别给出了在采用一阶和二阶变分方法时目标泛函对数值和梯度模对数值随迭代次数的下降图. 理论上当目标泛函关于未知参数的梯度模  $\|\nabla_{\boldsymbol{\theta}_j} J\| = 0$  时, 未知参数向量取最优估计值. 但本文中为了减小计算代价, 采用通常的迭代收敛标准, 即前后两次迭代的目标泛函梯度模变化小于预先指定的阈值:  $\|\nabla_{\boldsymbol{\theta}_j} J\| - \|\nabla_{\boldsymbol{\theta}_{j+1}} J\| < 10^{-8}$ . 此时的计算误差小于机器误差, 可以忽略. 从图 2 和图 3 中可以看出, 就目标泛函及其梯度值的变化而言, 二阶变分方法的下降速度明显快于一阶变分方法. 由表 1 进一步可知, 前者达到收敛只需要经历 8 次迭代, 而后者需要 72 次迭代; 而且二阶变分方法达到收敛后的目标泛函值和梯度值比一阶变分方法都要小数个量级. 图 4 和图 5 表示的是物理参数  $a$  和  $b$  的估计值在最小化过程中随迭代次数的变化以及逼近真实值的程度, 从图中易知, 采用

二阶变分方法时估计值收敛到真值的速度明显高于一阶变分方法. 分析和总结图 2—5 中所示的试验结果, 可得如下结论: 由于二阶离散变分方法能够同时利用目标泛函梯度信息和准确的 Hessian 矩阵信息, 因此目标泛函能够快速下降, 进而达到收敛. 表 1 中给出了在无观测噪声情况下一/二阶离散变分方法对 Hyperhenón 映射系统中未知物理参数进行估计的比较结果, 容易得出结论: 二阶离散变分方法用非常少的迭代次数获得了更高的估计精度. 随着观测误差增加, 虽然未知参数的估计精度下降, 但在相同误差水平情况下二阶变分方法无论是在目标泛函迭代收敛速度还是在参数估计精度

上都优于一阶变分方法, 因篇幅限制具体情况不再赘述.

### 4.2 二维抛物映射系统的参数估计

在本部分中以二维抛物映射系统为例, 进一步说明利用二阶离散变分方法估计非线性映射中未知参数的有效性. 二维抛物映射可用离散方程表示如下:

$$\begin{aligned} x_{k+1} &= ay_k^2, \\ y_{k+1} &= by_k(1-x_k), \\ k &= 0, 1, 2, \dots \end{aligned} \quad (37)$$

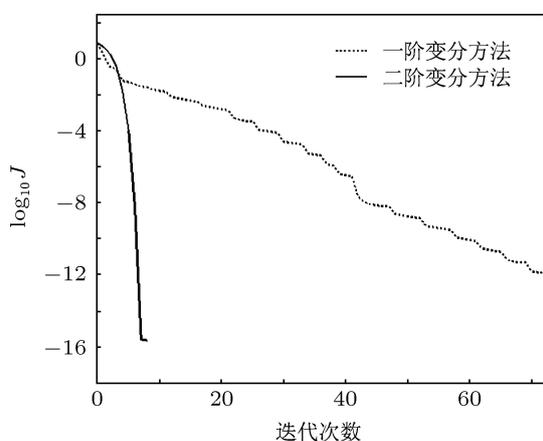


图 2 Hyperhenón 映射目标泛函值变化图

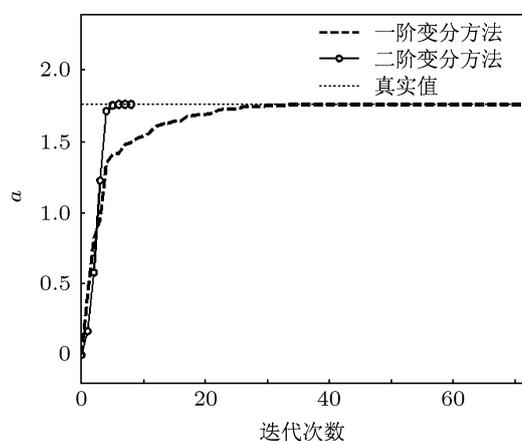


图 4 Hyperhenón 映射参数 a 辨识曲线图

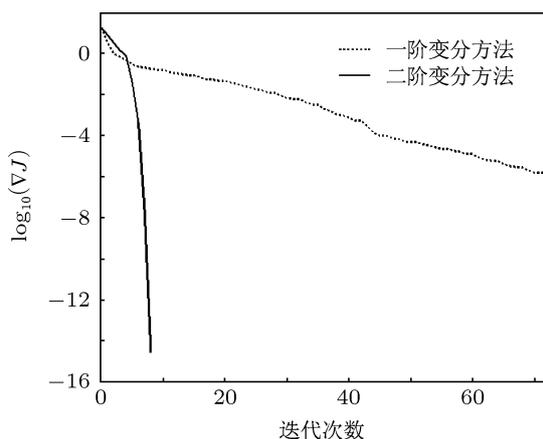


图 3 Hyperhenón 映射泛函梯度值变化图

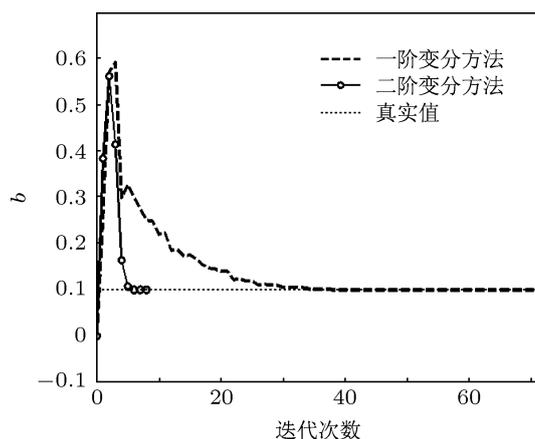


图 5 Hyperhenón 映射参数 b 辨识曲线图

表 1 无观测噪声时 Hyperhenón 映射的参数估计结果

方法	$a$	$b$	迭代次数	$J$	$\nabla J$
一阶变分	1.759998026	0.1000010779	72	$1.1910 \times 10^{-12}$	$1.4694 \times 10^{-6}$
二阶变分	1.760000051	0.1000003441	8	$2.4007 \times 10^{-16}$	$2.4992 \times 10^{-15}$

(37) 式表示的是一个二维双参数离散迭代映射, 两个控制参数  $a$  和  $b$  必须为正实数.  $a$  具有全局动态幅度调节功能, 称之为幅度调节参数; 而  $b$  具有动力学特性调节功能, 称之为特性调节参数 [20]. 当  $a = 1.00$  和  $b = 1.72$  时, 二维抛物映射系统呈现混沌状态. 设定状态向量  $\mathbf{x} = (x, y)^T$  和未知物理参数向量  $\boldsymbol{\theta} = (a, b)^T$ , 目标是在获得  $x$  的观测序列  $\mathbf{y}_k^o = (x_k^o, y_k^o)^T$  的情况下, 估计未知参数向量  $\boldsymbol{\theta}$ .

首先通过比较 (37) 式与 (1) 式, 易知  $\mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta}) = [ay_k^2, by_k(1-x_k)]^T$ , 其一阶变分为

$$\delta \mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta}) = \begin{pmatrix} 0 & 2ay_k \\ -by_k & b(1-x_k) \end{pmatrix} \begin{pmatrix} \delta x_k \\ \delta y_k \end{pmatrix} + \begin{pmatrix} y_k^2 & 0 \\ 0 & y_k(1-x_k) \end{pmatrix} \begin{pmatrix} \delta a \\ \delta b \end{pmatrix}. \quad (38)$$

根据 (11) 式可得切线性方程组:

$$\begin{aligned} \delta x_{k+1} &= 2ay_k \delta y_k + y_k^2 \delta a, \\ \delta y_{k+1} &= b(1-x_k) \delta y_k - by_k \delta x_k \\ &\quad + y_k(1-x_k) \delta b, \\ \delta x_0 &= 0, \quad \delta y_0 = 0. \end{aligned} \quad (39)$$

其次引入与  $\mathbf{x}_k$  对应的伴随向量序列  $\boldsymbol{\lambda}_k = (\lambda_k, p_k)^T$ , 并将  $\mathbf{x}_k, \mathbf{y}_k^o, \boldsymbol{\lambda}_k, \boldsymbol{\theta}, (\mathbf{D}_M^{\mathbf{x}_k})^T$  和  $(\mathbf{D}_M^{\boldsymbol{\theta}})^T$  等的表达式代入 (10) 式中的第一和第三式, 同时考虑到伴随初始条件, 进而得到二维抛物映射系统的伴随方程和目标泛函梯度公式, 分别如 (40) 和 (41) 式所示:

$$\begin{aligned} \lambda_k &= -by_k p_{k+1} + (x_k^o - x_k), \\ p_k &= 2ay_k \lambda_{k+1} + b(1-x_k) p_{k+1} + (y_k^o - y_k) \\ (\lambda_{N+1}, p_{N+1}) &= 0, \\ (k &= 1, 2, \dots, N). \end{aligned} \quad (40)$$

$$\begin{aligned} \nabla_a J &= -\sum_{k=0}^{N-1} y_k^2 \lambda_{k+1}, \\ \nabla_b J &= -\sum_{k=0}^{N-1} y_k(1-x_k) p_{k+1}. \end{aligned} \quad (41)$$

然后将  $\boldsymbol{\lambda}_k, \boldsymbol{\zeta}_k = (\zeta_k, \xi_k)^T, \delta \mathbf{f}_k = -(\delta x_k, \delta y_k)^T$ ,

$$\mathbf{D}_M^{\mathbf{x}_k}, [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{M}] = \begin{pmatrix} 0 & 2a\delta y_k \\ -b\delta y_k & -b\delta x_k \end{pmatrix} \text{ 和}$$

$$[\delta \boldsymbol{\theta}, \nabla_{\boldsymbol{\theta}}^2 \mathbf{M}] = \begin{pmatrix} 0 & 2y_k \delta a \\ -y_k \delta b & (1-x_k) \delta b \end{pmatrix} \text{ 代入 (18) 式}$$

中, 则得到二维抛物映射的二阶伴随方程:

$$\begin{aligned} \zeta_k &= -by_k \xi_{k+1} - p_{k+1}(b\delta y_k + y_k \delta b) - \delta x_k, \\ \xi_k &= 2ay_k \zeta_{k+1} + b(1-x_k) \xi_{k+1} \end{aligned}$$

$$\begin{aligned} &+ 2(a\delta y_k + y_k \delta a) \lambda_{k+1} \\ &+ (1-x_k) p_{k+1} \delta b - b p_{k+1} \delta x_k - \delta y_k, \\ \zeta_{N+1} &= 0, \quad \xi_{N+1} = 0. \end{aligned} \quad (42)$$

$$\text{最后将 } \boldsymbol{\lambda}_k, \boldsymbol{\zeta}_k, [\delta \mathbf{x}_k, \nabla_{\mathbf{x}_k}^2 \mathbf{M}] = \begin{pmatrix} 0 & 2y_k \delta a \\ -y_k \delta b & (1-x_k) \delta b \end{pmatrix},$$

$[\delta \boldsymbol{\theta}, \nabla_{\boldsymbol{\theta}}^2 \mathbf{M}] = 0$  和  $\mathbf{D}_M^{\boldsymbol{\theta}}$  代入 (27) 式中, 则得到二维抛物映射的 Hessian 矩阵 - 向量乘积的计算公式:

$$\begin{aligned} \nabla_{\boldsymbol{\theta}}^2 J \delta \boldsymbol{\theta} &= - \begin{pmatrix} \sum_{k=0}^{N-1} (2y_k \lambda_{k+1} \delta y_k + y_k^2 \zeta_{k+1}) \\ \mathcal{X} \end{pmatrix}, \\ \mathcal{X} &= \sum_{k=0}^{N-1} [(1-x_k) p_{k+1} \delta y_k - y_k p_{k+1} \delta x_k \\ &\quad + y_k(1-x_k) \xi_{k+1}]. \end{aligned} \quad (43)$$

在仿真试验中, 观测数据的生成采用与 4.1 部分中相同的方法. 模式初始状态量  $(x_0, y_0)$  取为  $(0.0147653, -0.1697946)$ , 其是二维抛物映射混沌系统状态中一点. 考虑到观测误差协方差矩阵是  $2 \times 2$  大小的单位矩阵, 目标泛函表示如下:

$$J(\boldsymbol{\theta}) = \frac{1}{2} \sum_{k=1}^N [(x_k - x_k^{\text{obs}})^2 + (y_k - y_k^{\text{obs}})^2]. \quad (44)$$

根据伴随方程、目标泛函及其梯度表达式, 利用共轭梯度算法能对二维抛物映射的未知物理参数进行最优估计, 称为一阶变分方法; 而根据一/二阶伴随方程、切线性方程、目标泛函、梯度及 Hessian 矩阵 - 向量乘积的表达式, 利用 NCG 算法能够对二维抛物映射的未知物理参数进行最优估计, 称为二阶变分方法. 两种方法中未知参数向量的迭代初值都设定为  $\boldsymbol{\theta}^0 = (0, 0)^T$ . 最小化迭代收敛采用与 4.1 部分中相同的准则.

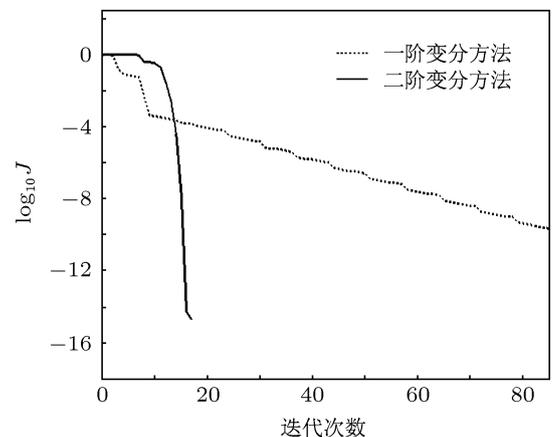


图 6 二维抛物映射目标泛函值变化图

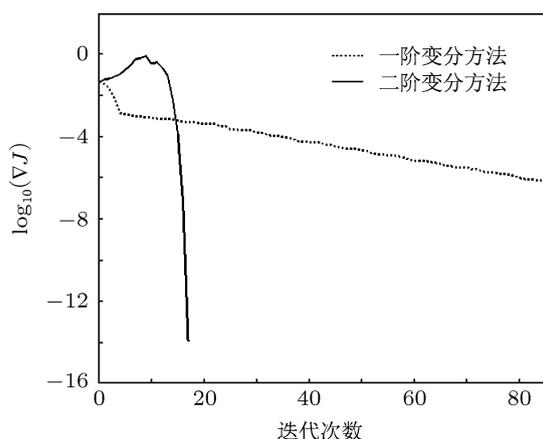


图7 二维抛物映射泛函梯度值变化图

在二维抛物映射系统的未知物理参数估计中, 图6和7中分别给出了在采用一阶和二阶变分方法时目标泛函对数值和梯度模对数值随迭代次数的下降图. 从图6和7中易知: 就目标泛函及其梯度模对数值的变化而言, 二阶变分方法的下降速度明显快于一阶变分方法. 由表2进一步可得: 前者达到收敛只经历17次迭代, 而后者需要85次迭代. 同时, 二阶变分方法达到收敛后的目标泛函值和梯度模值比一阶变分方法都要小数个量级. 图8和图9展示的是二维抛物映射系统中参数 $a$ 和 $b$ 的估计值随迭代次数的变化及其逼近真实值的程度, 从图中易知: 在利用二阶变分方法估计两个未知物理参数时, 收敛速度都快于一阶变分方法, 而对参数 $a$ 估计的收敛速度差异更明显. 表2中分别给出了在无观测噪声情况下一/二阶离散变分方法对二维抛物映射系统中未知参数进行估计的比较结果: 一阶变分方法对参数 $a$ 和 $b$ 的估计精度分别是小数点后3位和4位, 而二阶变分方法对两个参数的估

计精度都能达到小数点后7位. 根据上面的分析归纳出如下结论: 因为二阶离散变分方法能够同时利用准确的目标泛函梯度和Hessian矩阵信息, 所以能够用更少的迭代次数获得更高的估计精度, 明显优于一阶变分方法. 上述结论在不同观测误差水平情况下仍然成立, 由于篇幅所限不再赘述.

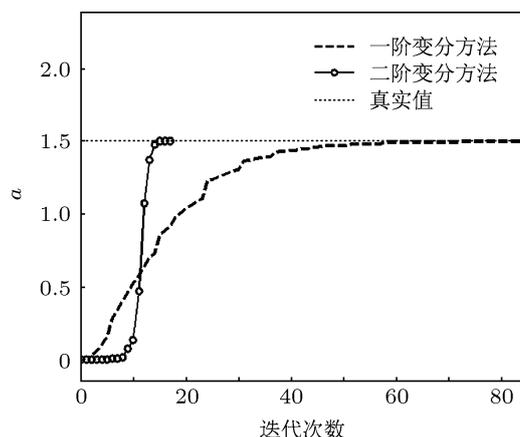


图8 二维抛物映射参数a的辨识曲线图

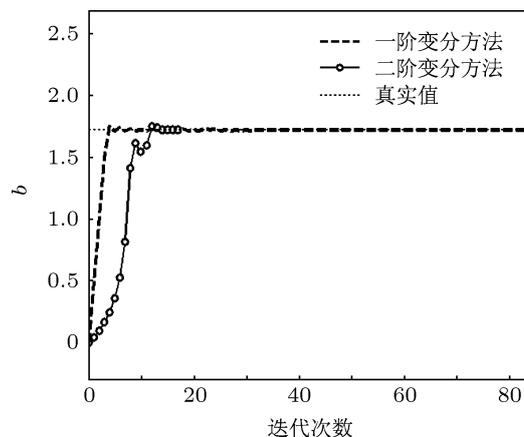


图9 二维抛物映射参数b的辨识曲线图

表2 无观测噪声时二维抛物映射的参数估计结果

方法	$a$	$b$	迭代次数	$J$	$\nabla J$
一阶变分	1.499243388	1.7199841621	85	$2.4142 \times 10^{-10}$	$7.6938 \times 10^{-7}$
二阶变分	1.500000052	1.7200000157	17	$2.0320 \times 10^{-15}$	$1.0945 \times 10^{-14}$

## 5 结论

本文提出一种估计非线性映射未知物理参数向量的二阶离散变分方法. 首先针对  $\mathbf{x}_{k+1} = \mathbf{M}(\mathbf{x}_k, \boldsymbol{\theta})$  形式的非线性离散混沌系统, 通过拉格朗日乘子将系统的状态控制方程引入到目标泛函中,

将映射方程约束最优化问题转换为无约束最优化问题; 其次利用变分方法给出了一/二阶伴随方程、目标泛函梯度和Hessian矩阵-向量乘的显式表达式; 然后设计了估计非线性映射未知参数的新算法, 并以此对Hyperhenón映射和二维抛物映射中的未知参数进行了估计. 数值仿真结果表明了二阶离散

变分方法的有效性和优点: 因为能够同时利用准确的目标泛函梯度信息和与 Hessian 矩阵相关的二阶信息, 所以二阶变分方法在无论有无观测噪声的情况下都能够用较少的迭代次数获得更高的估计精

度, 明显优于一阶变分方法. 该方法可推广应用于其他类型非线性映射和混沌系统的未知物理参数估计中.

- [1] Liu F C, Liang X M 2005 *Acta Phys. Sin.* **54** 4584 (in Chinese) [刘福才, 梁晓明 2005 物理学报 **54** 4584]
- [2] Wang X Y, Wu X J 2006 *Acta Phys. Sin.* **55** 605 (in Chinese) [王兴元, 武相军 2006 物理学报 **55** 605]
- [3] Wu Z Q, Tan F X, Wang S X 2006 *Acta Phys. Sin.* **55** 1651 (in Chinese) [吴忠强, 谭拂晓, 王绍仙 2006 物理学报 **55** 1651]
- [4] Maybhat A, Amritkar R E 1999 *Phys. Rev. E* **59** 284
- [5] Parlitz U 1996 *Phys. Rev. Lett.* **76** 1232
- [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 Y 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] 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] Gao F, Tong H Q 2006 *Acta Phys. Sin.* **55** 577 (in Chinese) [高飞, 童恒庆 2006 物理学报 **55** 577]
- [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] Wang J Y, Huang D X 2008 *Acta Phys. Sin.* **57** 2755 (in Chinese) [王钧炎, 黄德先 2008 物理学报 **57** 2755]
- [12] Long W, Jiao J J 2012 *Acta Phys. Sin.* **61** 110507 (in Chinese) [龙文, 焦建军 2012 物理学报 **61** 110507]
- [13] Cao X Q, Zhang W M, Song J Q, Zhu X Q, Zhao J 2012 *Acta Phys. Sin.* **61** 020507 (in Chinese) [曹小群, 张卫民, 宋君强, 朱小谦, 赵军 2012 物理学报 **61** 020507]
- [14] 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)]
- [15] He J H 2008 *Int. J. Modern. Phys. B.* **22** 3487
- [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
- [19] Chen E D 2005 *Ph. D. Dissertation* (Dalian: Dalian University of Science and Technology) (in Chinese) [陈尔东 2005 博士学位论文 (大连: 大连理工大学)]
- [20] Meng J D, Bao B C, Xu Q 2011 *Acta Phys. Sin.* **60** 010504 (in Chinese) [孟继德, 包伯成, 徐强 2011 物理学报 **60** 010504]
- [21] Kelley C T 1999 *Iterative Methods for Optimization* (Philadelphia: SIAM) p95
- [22] Nocedal J, Wright S J 2006 *Numerical Optimization* (2nd Ed.) (Verlag, Berlin, Heidelberg, New York: Springer) p198

# Parameter estimation of nonlinear map based on second-order discrete variational method\*

Cao Xiao-Qun<sup>†</sup>

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

(Received 11 November 2012; revised manuscript received 27 December 2012)

## Abstract

In this paper a method is proposed to estimate the unknown parameters of nonlinear map based on the second-order discrete variational method. On the basis of adjoint equations and gradient expressions of cost functional derived from variational method, the second-order adjoint equations and the exact formulas for the Hessian-vector product are deduced using the second-order discrete variational method. A new algorithm is given for estimating the unknown parameters of nonlinear maps. The numerical simulations show that all unknown physical parameters of Hyperhenón map and two-dimensional parabolic map are estimated successfully and precisely. Simulation results also demonstrate the validity and advantages of second-order discrete variational method of estimating the parameters of the discrete chaotic systems.

**Keywords:** nonlinear map, parameter estimation, second-order discrete variational method, adjoint equation

**PACS:** 05.45.Pq, 04.20.Fy, 45.10.Db

**DOI:** 10.7498/aps.62.080506

---

\* Project supported by the National Natural Science Foundation of China (Grant No. 41105063).

<sup>†</sup> Corresponding author. E-mail: caoxiaoqun@nudt.edu.cn