

文章编号: 1000-0887(2001)05-0441-08

低维混沌时序非线性动力系统的 预测方法及其应用研究*

马军海¹, 陈予恕²,

(1. 天津大学 管理学院, 天津 300072; 2. 天津大学 力学系, 天津 300072)

(本刊编委陈予恕来稿)

摘要: 主要研究由低维混沌时序所确定的非线性动力系统的预测方法及其应用。在国外学者研究工作的基础上, 应用一种非线性混沌模型在相空间内对时序进行重构工作, 先通过改进的最小二乘方法来估计模型的参数, 满足一定精度后, 再采用最优化方法来估计模型的参数, 并用所得的混沌时序模型在其相空间内对时序的未来值进行预测。给出了非常有代表性的实例对文中模型和算法进行验证。结果发现采用该算法能较准确地求得模型的参数, 在相空间中对混沌时序进行预测, 将传统方法中的外推变成了相空间中的内插, 及选取最佳的模型阶数等工作都能增加预测的准确程度, 且混沌时序不可能进行长期的预测。

关键词: 非线性; 混沌模型; 参数识别; 时序预测

中图分类号: O175.14; O241.81 **文献标识码:** A

引 言

现实问题中大量的时序问题都归属于非线性混沌时序问题, 实践证明无论是用低阶的还是用高阶的线性模型都无法去描述非线性混沌时序。所以广泛地开展用非线性混沌模型去研究低维混沌时序就变得越来越重要, 最近十几年来国内外学者用神经网络理论^[1], 小波理论^[2]等对非线性时序开展了研究, 并取得了一定的成果。Martin Casdagli^[3], 和 M E Davies^[4] Alex et Potapov^[5] 等人^[6~15] 分别对混沌时序的建模及预测开展了初步的研究。由于混沌吸引子的内在行为具有相当的不规则性及混沌吸引子具有十分复杂的几何结构, 且不同的混沌吸引子具有的复杂结构也各不相同。所以一般来说, 不同的混沌实测数据应该建立不同的混沌模型。但至今对各种描述不同混沌时序的混沌模型的优劣还未完全弄清楚, 这一问题还有待于进一步的研究。本文主要研究由低维混沌时序所确定的非线性动力系统的预测方法及其应用。在国外学者研究工作的基础上, 应用一种非线性混沌模型在相空间内对时序进行重构工作, 先通过改进的最小二乘方法来估计模型的参数, 满足一定精度后, 再采用最优化方法来估计模型的

* 收稿日期: 2000_02_22; 修订日期: 2001_01_08

基金项目: 国家自然科学基金资助项目(19990510); 国家重点基础研究专项经费资助项目(G1998020316)

作者简介: 马军海(1965—), 男, 教授, 二站博士后;

陈予恕(1931—), 山东肥城人, 教授, 博导, 俄国科学院外籍院士。

参数,并用所求得的混沌时序模型在其相空间内对时序的未来值进行预测。我们给出了非常有代表性的实例对本文模型^[16]和算法进行验证。结果发现采用本文的算法能较准确地求得模型的参数,在相空间中对混沌时序进行预测将传统方法中的外推变成了相空间中的内插,及选取最佳的模型阶数等工作都能增加预测的准确程度,且混沌时序不可能进行长期的预测。

1 模型的提出

令我们所得到的混沌时序 $x_1, x_2, x_3, \dots, x_N$ 其中 x_i 的采样时间为 $\tau_i, i = 1, 2, 3, \dots, N \cdot \Delta \tau = \tau_{i+1} - \tau_i$ 满足最佳时间采样间隔,且根据文献[8]的结论假设其最佳嵌入维数为 m , 则可以在其相空中重建 m 维向量 $v(n)$, 其满足(1)式:

$$v(n) = (x_n, x_{n+1}, x_{n+2}, \dots, x_{n+m-1}) \quad (n = 1, 2, \dots, N - m + 1), \quad (1)$$

基坐标框架由文献[7]中(47)式确定。 $v(n)$ 确定以后,即:

$$v(1), v(2), v(3), \dots, v(N - m + 1) \quad (2)$$

确立以后,怎样用(2)式来预测 $v(N - m + 2), v(N - m + 3), \dots$ (其对应混沌时序中的 x_{N+1}, x_{N+2}, \dots), 是必须要解决的核心问题。假设 \mathbf{R}^m 上的函数 F 和参数

$$a = (a_1, a_2, a_3, \dots, a_p) \quad (3)$$

满足(4)式:

$$v(n+1) = F(v(n), a) \quad (4)$$

如果已经得到了,映射 $F(v, a)$ 和已知 \mathbf{R}^m 上的向量 v , 我们就可以得到:

$$\left. \begin{aligned} v &\rightarrow v(1) = F(v, a), \\ v(1) &\rightarrow v(2) = F(v(1), a) = F(F(v, a), a) = F^2(v, a), \\ v(2) &\rightarrow v(3) = F(v(2), a) = F(F(v(1), a), a) = F^2(v(1), a) = F^3(v, a), \\ v(3) &\rightarrow v(4) = F(v(3), a) = F(F^2(v(1), a), a) = F^3(v(1), a) = F^4(v, a), \\ &\dots \quad \dots \quad \dots \\ v(n) &= F(v(n-1), a) = F(F(v(n-2), a), a) = \dots = F^{n-1}(v(1), a) = F^n(v, a) \end{aligned} \right\} \quad (5)$$

选择映射函数 $F(v, a)$ 如(6)式所示:

$$F(v, a) = \sum_{n=1}^{N-1} v(n+1) f_{\sigma}(v, v(n), a), \quad (6)$$

这里 $v(n)$ 是由式(1)所确定的 m 维时序向量。 $f_{\sigma}(v, v(n), a)$ 是由式(7)所确定的 \mathbf{R}^m 维空间上的迭代映射。

$$f_{\sigma}(v, v(n), a) =$$

$$\frac{\exp[-|v - v(n)|^2/\sigma] \left[a_1 + a_2 v(n)(v - v(n)) + \sum_{k=3}^p a_k (|v - v(n)|^2/\sigma)^{m_k} \right]}{\sum_{n=1}^{N-1} \exp[-|v - v(n)|^2/\sigma] \left[a_1 + \sum_{k=3}^p a_k (|v - v(n)|^2/\sigma)^{m_k} \right]}, \quad (7)$$

a 是 p 维时序向量, m_k 为正整数, σ 为时序 $x(n)$ ($n = 1, 2, \dots, N$) 的均方差。

建立如下的预测映射:

$$v(m+1) = \sum_{k=1}^L X_k F(v(m), a) = \sum_{k=1}^L X_k F^k(v(m-k+1), a), \quad (8)$$

这里 L 的值可取成时序的最佳嵌入维数 m^* (8) 式中的每一迭代项为 $X_k v_n(m+1)$, 这就要求 X_k 应满足(9) 式

$$\sum_{k=1}^L X_k = 1 \cdot \quad (9)$$

为了保证预测的准确性这里要求: $X_j \geq X_{j+1}$ (10)

通过(11) 式来求参数 $\mathbf{a} = (a_1, a_2, a_3, \dots, a_p)$ 和参数 $\mathbf{X} = (X_1, X_2, \dots, X_L)$

$$C(\mathbf{X}, \mathbf{a}) = \frac{\sum_{n=L}^{N-1} \left[| \mathbf{v}(n+1) - \sum_{k=1}^L X_k F^k(\mathbf{v}(n-k+1), \mathbf{a}) |^2 \right]}{\sum_{n=1}^N | \mathbf{v}(n) \cdot \mathbf{v}(n) |^2} \cdot \quad (11)$$

由(7) 式可以看出如果已知 \mathbf{R}^m 上的向量 \mathbf{v} 和已知的相空间中的第 n 个向量 $\mathbf{v}(n)$ 相等则 $f_{\sigma}(\mathbf{v}, \mathbf{v}(n), \mathbf{a}) = 1$; 反之, 若 $| \mathbf{v} - \mathbf{v}(n) |$ 非零, 则 $f_{\sigma}(\mathbf{v}, \mathbf{v}(n), \mathbf{a})$ 下降的很快。

2 模型参数辨识方法

为了求得(11) 式中的各参数本文采用改进的非线性系统最小二乘算法^[7] 进行参数辨识, 其具体过程如下: 令带有白噪声及多维控制输入的非线性系统模型写为:

$$\mathbf{v}(m+1) = f[\mathbf{v}(m), \mathbf{U}(m), \mathbf{D}(m), m] + \varepsilon_m, \quad (12)$$

这里 $\mathbf{D} = [\mathbf{X}, \mathbf{a}] = [X_1, X_2, \dots, X_L, a_1, a_2, a_3, \dots, a_p]^T$, (13)

$$\mathbf{D}(m) = [\mathbf{X}(m), \mathbf{a}(m)] = [X_{1m}, X_{2m}, \dots, X_{Lm}, a_{1m}, a_{2m}, a_{3m}, \dots, a_{pm}]^T, \quad (14)$$

其中 $\mathbf{v}(m)$, $\mathbf{U}(m)$ 分别表示到时刻 m 为止的系统输出集合和系统输入向量集合, \mathbf{D} 表示系统的未知参数向量, $f[\dots]$ 表示已知形式的非线性函数, ε_m 是白噪声, 上标 T 表示转置, N 为样本值。一般的最小二乘准则为:

$$J_1(\mathbf{D}) = \sum_{i=1}^{N-1} \left\{ f_i[\mathbf{v}(m-1), \mathbf{U}(m-1), \mathbf{D}(m-1), m-1] - \mathbf{v}(i) \right\}^2, \quad (15)$$

但在对(15) 式目标函数 $J_1(\mathbf{D})$ (非线性情况下) 对 \mathbf{D} 求极小值的过程中会出现病态情况, 会使算法无法进行或收敛速度过慢等。为此文献^[7] 建议将 $J_1(\mathbf{D})$ 改写成(16) 式形式。

$$J_2(\mathbf{D}) = \sum_{i=1}^{N-1} \left\{ f_i[\mathbf{v}(m-1), \mathbf{U}(m-1), \mathbf{D}(m-1), m-1] - \mathbf{v}(i) \right\}^2 + \mu \| \mathbf{D} - \mathbf{D}(m) \|^2, \quad (16)$$

由(16) 式可以求得关于 \mathbf{D} 的极小值。

$$\mathbf{D}(m+1) = \mathbf{D}(m) + [\mathbf{W} + (\Psi^T(\mathbf{D}(m)))^2 \mathbf{J}^{-1} \Psi(\mathbf{D}(m))]^T [V(N+1) - F_{N+1}(\mathbf{v}(N), \mathbf{U}(N), \mathbf{D}(m), N)], \quad (17)$$

其中 $\Psi(\mathbf{D}(m))$ 是 F_{N+1} 在 $\mathbf{D}(m)$ 处的 Jacobi 矩阵。不难看出当 $\mu \rightarrow \infty$ 时算法变成最速下降法。

$J_1(\mathbf{D})$ 达到一定精度后再将所得参数值作为新的识别问题的初值, 计算表明, 这样一来能增加参数识别的精度。则参数识别问题变为问题。

$$\min_{\mathbf{D}} f(\mathbf{D}) = C(\mathbf{X}, \mathbf{a}) \quad (\text{这里 } \mathbf{D} \text{ 见(13) 式}) \quad (18)$$

耗散函数 $f(\mathbf{D})$ 是向量 \mathbf{D} 的连续可微函数, 其导数为:

$$\therefore \dot{f}(\mathbf{D}) = \left[\frac{\partial f(\mathbf{D})}{\partial X_1}, \frac{\partial f(\mathbf{D})}{\partial X_2}, \dots, \frac{\partial f(\mathbf{D})}{\partial X_L}, \frac{\partial f(\mathbf{D})}{\partial a_1}, \frac{\partial f(\mathbf{D})}{\partial a_2}, \dots, \frac{\partial f(\mathbf{D})}{\partial a_p} \right]^T, \quad (19)$$

$$\text{记 } \dot{f}(\mathbf{D}_k) = \dot{f}(\mathbf{D})|_{\mathbf{D}=\mathbf{D}_k} \quad (20)$$

参数的识别过程如下:

步骤 1: 置初值 \mathbf{D}_0 , $k = 0$, 拟对称正定的 Hessian 矩阵的初值矩阵 \mathbf{H}_0 , 控制误差 ε (取 $\mathbf{H}_0 = \mathbf{E}$), 转步骤 2;

步骤 2: 计算 $\dot{f}(\mathbf{D}_k)$, 并求搜索方向 $P_k = -\mathbf{H}_k \dot{f}(\mathbf{D}_k)^T$ 转步骤 3;

步骤 3: 以 \mathbf{D} 为新的起始点, 求一维搜索问题 $\min_{\lambda} f(\mathbf{D} + \lambda P_k)$ 得 λ_k 置 $\mathbf{D}_{k+1} = \mathbf{D}_k + \lambda_k P_k$, 转步骤 4;

步骤 4: 若 $\|\mathbf{D}_{k+1} - \mathbf{D}_k\| \leq \varepsilon$, 置 $\mathbf{D} = \mathbf{D}_{k+1}$, 迭代结束, 则求出了最优解. 否则转步骤 5;

步骤 5: 计算 $\dot{f}(\mathbf{D}_{k+1})$, 令: $y_k = \dot{f}(\mathbf{D}_{k+1}) - \dot{f}(\mathbf{D}_k)^T$,

$$s_k = \mathbf{D}_{k+1} - \mathbf{D}_k,$$

实验数据的修正 \mathbf{H}_k 阵为 $\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{\mathbf{H}_k y_k y_k^T \mathbf{H}_k}{y_k^T \mathbf{H}_k y_k} \quad k+1 \Rightarrow k$;

步骤 6: 若 $k = n$, 则 $\mathbf{D}_k \Rightarrow \mathbf{D}_0$, 转到步骤 1, 若 $k < n$, 则转向步骤 3;

步骤 6 中保持共轭方向的优越性, 一般迭代 n 次后就要重新开始, 即令 $\mathbf{H}_0 = \mathbf{E}$, $\mathbf{D}_0 = \mathbf{D}_n$.

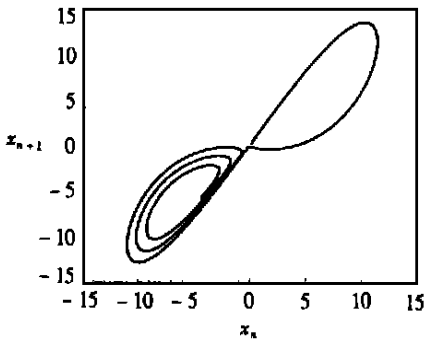
3 计算结果

取如下 2 组数据(限于篇幅)对本文的算法进行验证

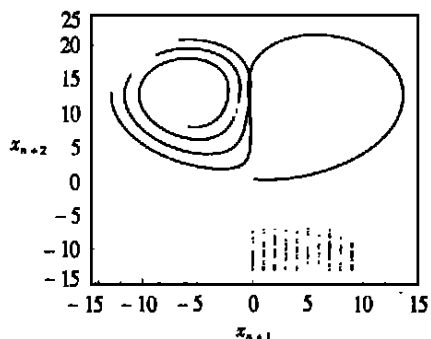
$$\text{第 1 组对 Lorenz 系统} \begin{cases} \dot{x} = \sigma(y - x), \\ \dot{y} = -xz + rx - y, \\ \dot{z} = xy - bz, \end{cases}$$

取 $\sigma = 10$, $r = 13.926$, $b = 8/3$ 同缩分岔的情况, 前 1 000 点作为暂态点去掉, 把后 6 000 点作为第 1 组原始数据点. 其 (x_n, x_{n+1}) 相图如图 1(a), 其 (x_{n+1}, x_{n+2}) 相图如图 1(b), 其 (x_n, x_{n+2}) 相图如图 1(c), 其 (x_n, x_{n+1}, x_{n+2}) 相图如图 1(d). 在此参数状态下, Lorenz 系统为标准同缩分岔的情况, 两同缩轨道为边界的两片叶子组成, 这两片叶子只交于唯一的平衡点 p , 因此任何其中一片叶子的轨道都不可能到达另一片叶子, 因此由此轨道组成的集合 A_p 不是吸引集, 这在图 1(d) 中表现的非常清楚, 其它复杂的情况如:

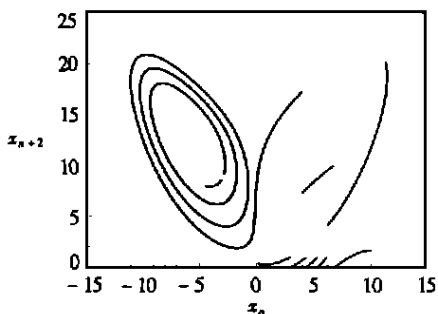
- 集合 A_p 含有两个吸引子的情况,
- 集合 A_p 含有倍周期分岔的情况,



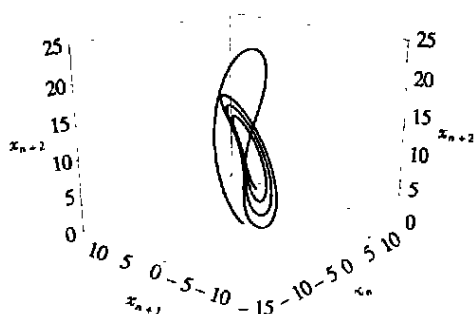
(a) 第 1 组数据的 (x_n, x_{n+1}) 相图



(b) 第 1 组数据的 (x_{n+1}, x_{n+2}) 相图



(c) 第 1 组数据的 (x_n, x_{n+2}) 相图



(d) 第 1 组数据的 (x_n, x_{n+1}, x_{n+2}) 相图

图 1

c) 集合 A_p 含有吸引盆的情况,

d) 集合 A_p 含有预湍流区域等的情况因涉及到更深的现代数学及优化理论知识, 处理起来比较困难, 我们将在以后的工作中陆续报道.

第 2 组 用我们所得到的混沌经济数据 6 000 点作为实验数据, 其时间历程图如图 2, 其 (x_n, x_{n+1}, x_{n+2}) 相图如图 3. 经多次反复计算比较发现: 对本文选取的非常有代表性的 2 组数据, 预测模型(8)中取(9)式模型阶数 L 为最佳嵌入维数 m , 或 $m + 1$, 其预测结果较为准确, 反之, 若 L 再增加或减少, 则误差 $C(X, a)$ 都再直线增加.

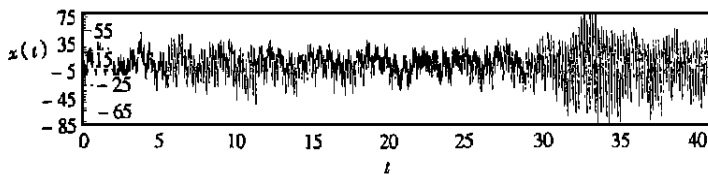


图 2 第 2 组经济数据的时间历程图

1) 对 Lorenz 系统其 $m = 3$, 本文取 $L = 3, p = 4, m_k = 2$, 样本点 $N = 4096$ (点), 计算中得到: $C(X, a) = 2.3724 \times 10^{-5}, X_1 = 0.68568, X_2 = 0.185762, X_3 = 0.128558, a_1 = 43.67978, a_2 = 0.034977, a_3 = 6.927302, a_4 = -0.109752$. 用这组参数进行预测得到相对误差不超过 35% 的预测值长度 $N_Y = 103$, 其预测值 (x_n, x_{n+1}, x_{n+2}) 的相图如图 4(a) 所示, 其样本值、预测值 (x_n, x_{n+1}, x_{n+2}) 的混合相图如图 4(b) 所示.

2) 对第 2 组经济数据混沌的情况其 $m = 4$ (根据文献[9]的结论), 本文取 $L = 5, p = 6, m_k = 3$, 样本点 $N = 4000$ (点), 计算中得到:

$$C(X, a) = 6.193576 \times 10^{-4}, X_1 = 0.414437, X_2 = 0.25796, X_3 = 0.150832, X_4 = 0.106023, X_5 = 0.070748, a_1 = 37.06423, a_2 = 0.040201, a_3 = 6.29325, a_4 = -0.7763, a_5 = 2.72248, a_6 = 0.025936.$$

用这组参数进行预测得到相对误差不超过 35% 的预测值长度 $N_Y = 93$, 其预测值 $(x_n,$

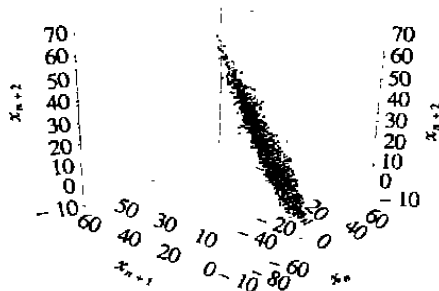
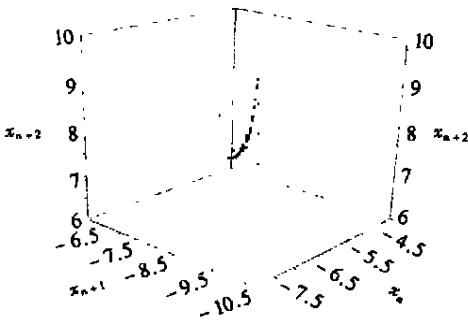
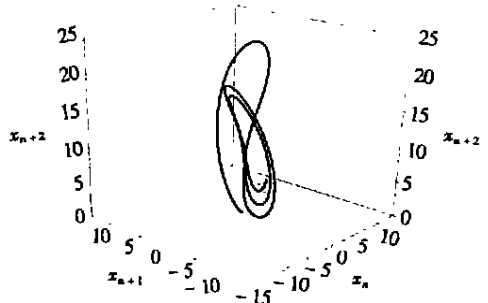


图 3 第 2 组经济数据的 (x_n, x_{n+1}, x_{n+2}) 相图

x_{n+1}, x_{n+2} 的相图如图 5(a) 所示, 其样本值、预测值(x_n, x_{n+1}, x_{n+2}) 的混合相图如图 5(b) 所示。

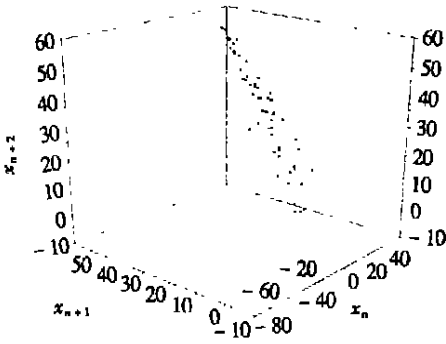


(a) 第 1 组数据预测值
(x_n, x_{n+1}, x_{n+2}) 的相图

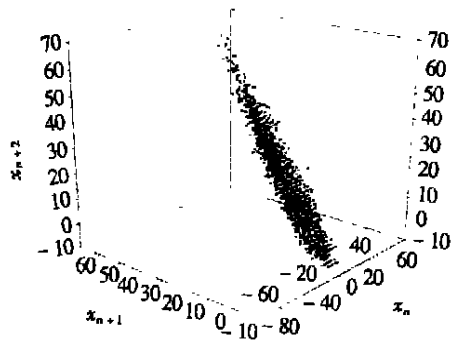


(b) 第 1 组数据样本值, 预测值
的(x_{n+1}, x_{n+2}) 混合相图

图 4



(a) 第 2 组数据预测值
(x_n, x_{n+1}, x_{n+2}) 的相图



(b) 第 2 组数据样本值, 预测值的
(x_n, x_{n+1}, x_{n+2}) 混合相图

图 5

4 结 论

1) 随着样本值 N 的变化, 预测模型中参数 X_L, a_p 也发生一定的变化, 但只要 N 大于低维混沌时序样本值所要求的粗略的估计值(见文献[8]), $C(X, a)$ 的变化就在所允许的良好范围内。

2) 模型阶数 L 取最佳嵌入维数 m , 或 $m+1$, 其预测结果较为准确, 反之, 若 L 再增加或减少则误差 $C(X, a)$ 都再直线增加, 预测值便不再准确。

3) p 值和 m_k 值的选择目前在理论上无法给出定论, 但作者反复研究了多个低维混沌时序发现该模型中 p 一般在 3 到 6 范围内取值; m_k 一般在 2 到 4 范围内取值, 否则会导致预测精度下降、预测长度减小以及再次加大参数辨识的难度等相关问题的发生。

4) 若用预测模型中的同一参数 X_L, a_p 值则不可能进行长期的时序预测, 反之, 当预测误差大于实际问题所允许的范围时, 样本值中加入新产生的时序真值, 则可以重复前一过程继续预测。

5) 对各种描述不同混沌时序的混沌模型的优劣还未完全弄清楚, 这一问题还有待于进一

[参 考 文 献]

- [1] ZHANG Qing_hua. Wavelet networks[J]. IEE Transaction on Neural Networks, 1992, **6**(11): 889—898.
- [2] CAO Liang_yue, HONG Yi_guang, FANG Hai_ping, et al. Predicting chaotic timeseries with wavelet networks[J]. Phys D, 1995, **85**(6): 225—238.
- [3] Casdagli Martin. Nonlinear prediction of chaotic time series[J]. Phys D, 1989, **35**(11): 335—356.
- [4] Davies M E. Reconstructing attractors from filtered timeseries[J]. Phys D, 1997, **101**(5): 195—206.
- [5] Potapov Alexei. Distortions of reconstruction for chaotic attractors[J]. Phys D, 1997, **101**(8): 207—226.
- [6] Abarbanel Henry D I, Brown Reggie, Kadtke James B. Prediction and system identification in chaotic nonlinear systems: time series with broadband spectra[J]. Phys Lett A, 1989, **138**(8): 401—407.
- [7] 马军海, 陈予恕, 刘曾荣. 动力系统实测数据的非线性混沌特性的判定[J]. 应用数学和力学, 1998, **19**(6): 481—488.
- [8] 马军海. 混沌时序动力系统非线性重构[D]. 天津: 天津大学力学系, 1997.
- [9] Diambra L, Plastino A. Modeling time series using information theory[J]. Phys Lett A, 1996, **216**(12): 278—236.
- [10] Judd Kevin, Mees Alistair. Modeling chaotic motions of a string from experimental data[J]. Phys D, 1996, **92**(10): 221—236.
- [11] McGuire Gary, Azar Nabeel B, Shelhamer Mark. Recurrence matrices and the preservation of dynamical properties[J]. Phys Lett A, 1997, **237**(8): 43—47.
- [12] Schroer Christian G, Sauer Tim, Ott Edvard, et al. Predicting chaotic most of the time from embeddings with self_intersections[J]. Phys Rev Lett, 1998, **80**(7): 1410—1412.
- [13] Castillo E, Gutierrez J M. Nonlinear time series modeling and prediction using functional networks extracting information masked by chaos[J]. Phys Lett A, 1998, **244**(8): 71—84.
- [14] Judd Kevin, Mees Alistair. Embedding as a modeling problem[J]. Phys D, 1998, **120**(7): 273—286.
- [15] Kugiuntzis D, Lingjxrde O C, Christophersen N. Regularized local linear prediction of chaotic time series[J]. Phys D, 1998, **112**(4): 344—360.
- [16] Abarbanel Henry D I, Brown Reggie, Kadtke James B. Prediction in chaotic nonlinear system: methods for time series with broadband Fourier spectra[J]. Phys Rev A, 1990, **41**(4): 1782—1807.
- [17] CHEN Yu_shu, MA Jun_hai, LIU Zeng_rong. The state space reconstruction technology of different kinds of chaotic data obtained from dynamical system[J]. Acta Mechanica Sinica, 1999, **15**(1): 82—92.
- [18] Chen C H. Applied Timeseries Analysis [M]. Beijing: World Scientific Publishing Cor, 1989.
- [19] YANG Shu_zi, WU Ya. Applied Timeseries Analysis in Engineering [M]. Beijing: World Scientific Publishing Cor, 1992.

Study on the Prediction Method of Low_Dimension Time Series That Arise From the Intrinsic Nonlinear Dynamics

MA Jun_hai¹, CHEN Yu_shu²

(1. School of Management, Tianjin University, Tianjin 300072, P R China;

2. Department of Mechanics, Tianjin University, Tianjin 300072, P R China)

Abstract: The prediction methods and its applications of the nonlinear dynamic systems determined from chaotic time series of low dimension are discussed mainly. Based on the work of the foreign researchers, the chaotic time series in the phase space adopting one kind of nonlinear chaotic model were reconstructed. At first, the model parameters were estimated by using the improved least square method. Then as the precision was satisfied, the optimization method was used to estimate these parameters. At the end by using the obtained chaotic model, the future data of the chaotic time series in the phase space was predicted. Some representative experimental examples were analyzed to testify the models and the algorithms developed in this paper. The results show that if the algorithms developed here are adopted, the parameters of the corresponding chaotic model will be easily calculated well and true. Predictions of chaotic series in phase space make the traditional methods change from outer iteration to interpolations. And if the optimal model rank is chosen, the prediction precision will increase notably. Long term superior predictability of nonlinear chaotic models is proved to be irrational and unreasonable.

Key words: nonlinear; chaotic model; parameter identification; time series prediction