

基于 EMD 方法的混沌时间序列预测^{*}

杨永锋[†] 任兴民 秦卫阳 吴亚锋 支希哲

(西北工业大学, 西安 710072)

(2007 年 12 月 25 日收到 2008 年 1 月 18 日收到修改稿)

将经验模态分解(EMD)方法引入到非线性数据处理中,提出用 EMD 分解后的数据进行混沌预测的方法.通过 Duffing 方程和 Lorenz 系统的非线性响应预测实例表明,EMD 分解后的信号和原始信号相比具有较小的最大 Lyapunov 指数,可提高预测时间和长时预测精度.

关键词: EMD, 混沌, 预测

PACC: 0545

1. 引言

混沌理论表明即使系统初始状态条件存在细微差异,随着系统演化也可能导致显著差异,这便是混沌系统的蝴蝶效应^[1].非线性是产生混沌序列的根本原因,在寻找混沌时间序列的预测方法时,应该注意混沌运动服从一定的规则,但混沌又是类似随机的.因此,混沌预测方法只具有有限的预测能力.虽然混沌系统对初始条件极为敏感,不可能得到完美的预测值,但是在它的可预测性失去前,即在一个相对短的时间长度内,混沌预测的精度还是可以保证的.混沌的上述性质一方面指出了原本认为不可预测的复杂事物具有可预测性,另一方面也指出了原本认为可预测的简单事物的预测具有局限性.

混沌时间序列的预测通常是以重构相空间理论为基础^[2],这类问题可以理解为动力学系统研究的逆问题.它是给定相空间中的一串迭代序列,如何构造一个非线性映射来表示这一动力系统,这样的非线性映射就作为预测模型.到目前为止,已经发展了许多的混沌时间序列建模和预测方法,这些预测方法大致分为全局预测法^[3]、局部预测法^[4]、自适应预测法^[5]和基于 Lyapunov 指数的预测法^[6].其中基于 Lyapunov 指数的时间序列预测是根据数据序列本身所计算出来的客观规律,即 Lyapunov 指数作预测,而不是像传统方法那样通过选取主观模型预测,因而预测的精度和可信度较高^[6].

根据混沌时间序列预测理论,若系统的最大 Lyapunov 指数为 λ_1 ,则最大可预测时间约为 $1/|\lambda_1|^{[7]}$.因此 λ_1 越大,运动可预测时间 t_0 越短,运动的可预测性越差.对于混沌系统,当预测时间小于最大预测时间尺度 t_0 时,系统预测误差随预测步长的增加而增大,但是比较平稳,一旦超过这个界限,误差将会倍增,便失去了预测的意义.因此若存在一种有效的方法,可将系统的响应分解为若干个响应的线性叠加,并且保证各分解后的响应的频率成分仅为原信号的一部分,此时,各分解信号的最大 Lyapunov 指数均小于原信号的最大 Lyapunov 指数(实践获得),这将对于提高最大可预测时间有着明显的意义.EMD 方法可对信号进行自适应分解,它不仅适用于平稳信号也适用于非平稳信号^[8],因此被引入到分解混沌信号中.本文采用 EMD 方法对混沌信号进行分解,然后基于最大 Lyapunov 指数对分解后信号进行混沌预测,并对比了原始信号以及分解后信号的预测误差.

2. 基于最大 Lyapunov 指数的混沌预测

Wolf 等人^[9]根据在混沌时间序列重构相空间中,邻近轨道之间的距离随时间演化呈指数形式分离的研究成果,提出根据最大 Lyapunov 指数进行预测的混沌时间序列预测方法.其思想是在历史时间序列样本中寻找相似点,根据相似点的演化行为和最大 Lyapunov 指数的物理意义,运用一定的数学模

^{*} 陕西省自然科学基金(批准号:2007A08)资助的课题.

[†] E-mail: yyfnpu@hotmail.com

型获取预测值.

不妨设 Y_M 为预测的中心点,相空间中 Y_M 的最邻近点为 Y_k ,其距离为 $d_M(0)$,最大 Lyapunov 指数为 λ_1 ,即

$$d_M(0) = \min_{\hat{j}} \|Y_M - Y_{\hat{j}}\| = \|Y_M - Y_k\|, \quad (1)$$

为了避免参考点 Y_M 和最近邻点 $Y_{\hat{j}}$ 位于同一轨线上,这里采用限制短暂分离,即要求 $|M - \hat{j}| > P$,其中 P 为时间序列的平均周期.根据最大 Lyapunov 指数定义为

$$\|Y_M - Y_{M+1}\| = \|Y_k - Y_{k+1}\| e^{\lambda_1}, \quad (2)$$

其中,点 Y_{M+1} 只有最后一个分量未知,故可预测.

3. 时间序列的重构相空间

一般时间序列主要是在时域或变换域中进行研究,而在混沌时间序列处理中,无论是混沌不变量的计算、混沌模型的建立和预测都是在相空间中进行,因此相空间重构是混沌时间序列处理中非常重要的第一步.互信息法可判断系统的非线性相关性,并能更准确地表明两个坐标之间的独立性^[10],因此本文采用互信息法求延迟时间.对于嵌入维数选取则采用目前公认比较有效的改进虚假最邻近点法,即 Cao 方法^[11].

4. 经验模态分解

经验模态分解(EMD)可对信号进行自适应的分解,不仅适用于平稳信号也适用于非平稳信号.目前发现它在机械、交通、医学、电力等许多领域都具有很高的应用价值^[12-14].

EMD 方法分解信号基于以下三条假定:1)信号具有至少两个极值点:一个极大值点和一个极小值点;2)特征时间尺度定义为相邻极值点之间的时间间隔;3)如果信号没有极值点而仅有拐点,那么在对它分解之前首先将其微分一次或者多次以获得极值点.然后对所得结果进行积分就可以得到相应的分量.特征模式函数(IMF)满足以下两个条件:

1)整个时间历程内,穿越零点的次数与极值点个数相等或至多相差 1.

2)信号上任意一点,由局部极大值定义的上包络线和由局部极小值定义的下包络线的均值为 0,即信号关于时间轴局部对称.

EMD 具体实现是通过一种被称为“筛分”处理的过程实现对信号的分解.具体的处理过程为

1)对任一给定信号 $X(t)$,首先确定出 $X(t)$ 上的所有极值点,用三次样条曲线连接所有极大值点形成上包络线,同样的方法形成下包络线.数据 $X(t)$ 与上下包络线的均值 m_1 的差记为 h_1 ,则

$$h_1 = X(t) - m_1, \quad (3)$$

将 h_1 视作新的 $X(t)$,重复过程 1),直到 h_1 满足 IMF 的两个条件时,则其成为从原始信号筛选出的第一阶 IMF.通常第一阶 IMF 分量 h_1 包含信号的最高频成分.

2)将 h_1 从 $X(t)$ 中分离出来,得到一个去掉高频分量的差值信号 r_1 ,有

$$r_1 = X(t) - h_1, \quad (4)$$

把 r_1 作为“新”信号重复以上筛分步骤,直到第 n 阶的残余信号成为单调函数,不再能筛分出 IMF 分量

$$r_n = r_{n-1} - h_n. \quad (5)$$

3)数学上, $X(t)$ 可表示为 n 个 IMF 分量和一个残余项的和,即

$$X(t) = \sum_{j=1}^n h_j(t) + r_n(t), \quad (6)$$

式中, $r_n(t)$ 为残量,代表了信号中的平均趋势,而各 IMF 分量 $h_j(t)$ 则分别代表了信号从高到低不同频率段的成分,每一频率段所包含的频率成分是不同的,同一 IMF 分量中,不同时刻处的瞬时频率也是不同的,这种不同频率成分的局部时间分布式随信号本身的变化而变化.

Huang 给出的筛分过程结束的标准是一种类似于 Cauchy 收敛准则的理论上的标准^[8],它定义了如下标准偏差:

$$SD = \frac{1}{T} \int_0^T \frac{|h_k(t) - h_{k-1}(t)|^2}{|h_{k-1}(t)|^2} dt, \quad (7)$$

通常将 SD 的取值定在 0.2 到 0.3 之间,即满足 $0.2 < SD < 0.3$ 时筛分过程即可结束.此标准的物理考虑为,既要使得 $h_k(t)$ 足够接近 IMF 的要求,又要控制筛分的次数,从而使所得到 IMF 分量保留原始信号中幅值调制的信息.经多次试算,本文的 SD 取值为 0.28.

5. Duffing 系统的混沌预测

研究如下具有负线性刚度的 Duffing 方程^[15,16]:

$$\ddot{x} + c\dot{x} - x + x^3 = f\cos\omega t, \quad (8)$$

本节针对 $f = 0.32$ 时的响应进行混沌预测. 采用 Newmark- β 法积分, 积分步长为 0.005, 积分初值为 0. 利用小数据量计算系统的最大 Lyapunov 指数为 0.7941, 经过 EMD 分解后其主要分量 Imf1 和 Imf2 的最大 Lyapunov 指数分别为 0.2711 和 0.1787, 均明显小于原信号的最大 Lyapunov 指数(见图 1).

根据表 1 的参数, 采用最大 Lyapunov 指数预测法对 Duffing 系统的响应和经 EMD 分解后的 Imf1, Imf2 进行预测, 原数据为 90000 点, 为消除 Imf 端点效应影响, 截取 10000—79800 点的数据作为已知样本, 79801—80000 点作为校验样本, 测试预测的准确度. 发现对于 Duffing 系统的混沌响应在 160 点以前, 预测结果与校验样本吻合较好, 在 160 点以后, 随着预测点明显超过最大可预测时间 1.259, 预测误差急剧放大, 预测失去意义. 对于 EMD 分解后的数据 Imf1 和 Imf2, 在 160 点以前预测结果误差较原

数据预测结果大, 但在 160 点以后, 由于两个分量的可预测时间比原始数据大 1 倍多, 因此, 预测误差和原始数据的误差相比在逐渐减小, 在 200 点时可发现 EMD 分解后的预测明显优于原始数据的直接预测(见图 2(d)).

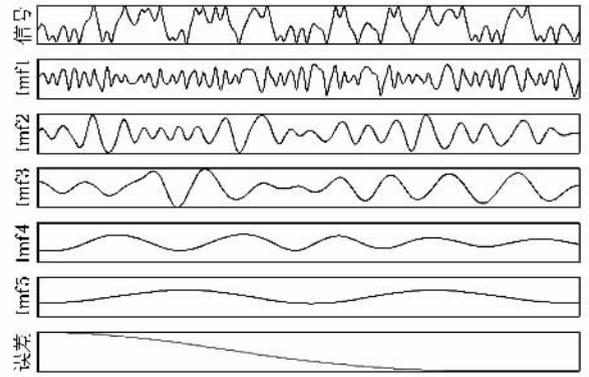


图 1 Duffing 系统混沌响应及其 Imf 分量

表 1 Duffing 系统响应及其主要 Imf 分量的最大 Lyapunov 指数

数据	τ	m	P	λ_1	t_0
Duffing	23	4	264.5	0.7941	1.259
Imf1	26	6	143.1	0.2711	3.689
Imf2	51	6	253.6	0.1787	5.596

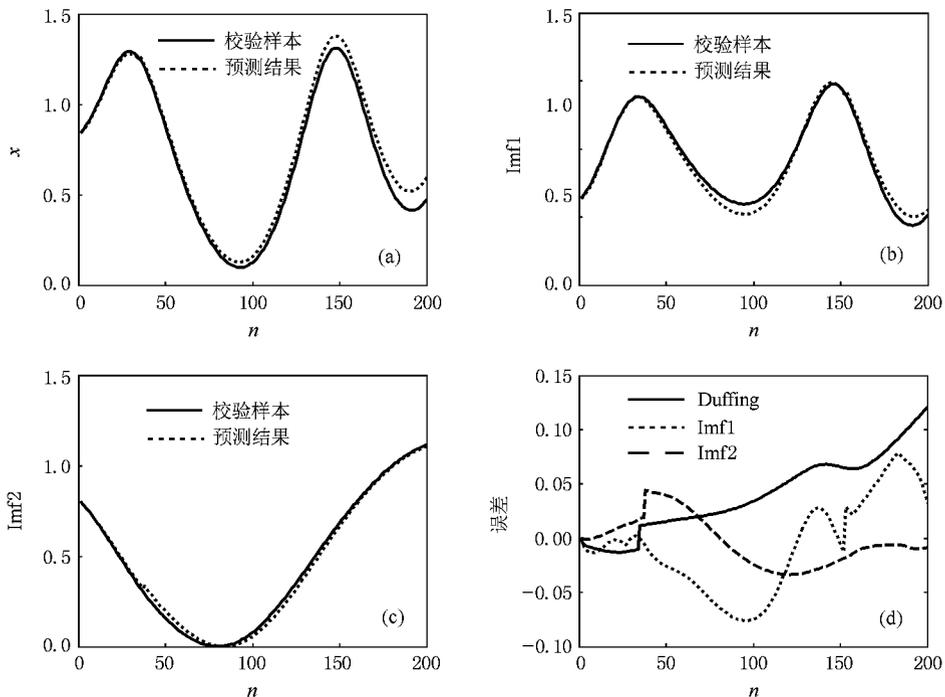


图 2 Duffing 系统响应及其 Imf 分量预测 (a)Duffing 混沌的响应及其预测结果 (b)Imf1 及其预测结果; (c)Imf2 及其预测结果 (d) 预测误差比较

6. Lorenz 系统的混沌预测

无量纲 Lorenz 方程为

$$\begin{aligned} \dot{x} &= a(y - x), \\ \dot{y} &= cx - xz - y, \\ \dot{z} &= xy - bz, \end{aligned} \quad (9)$$

选取参数 $a = 16, b = 4, c = 45.92$, 采用四阶 Runge-Kutta 算法积分, 积分步长为 0.01, 积分初值为 $[-1, 0, 1]$. 利用小数据量计算系统的最大 Lyapunov 指数为 1.5015 (文献 [9] 参考值为 1.5) 经过 EMD 分解后其主要分量 Imf1, Imf2 和 Imf3 的最大 Lyapunov 指数分别为 0.7875, 0.0231 和 0.0088, 均明显小于原信号 (见图 3).

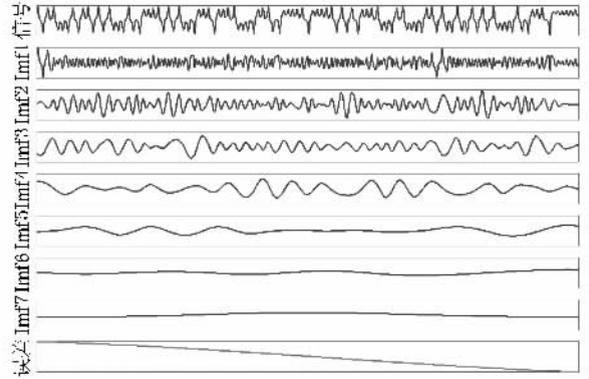


图 3 Lorenz 系统 x 分量响应及其 Imf 分量

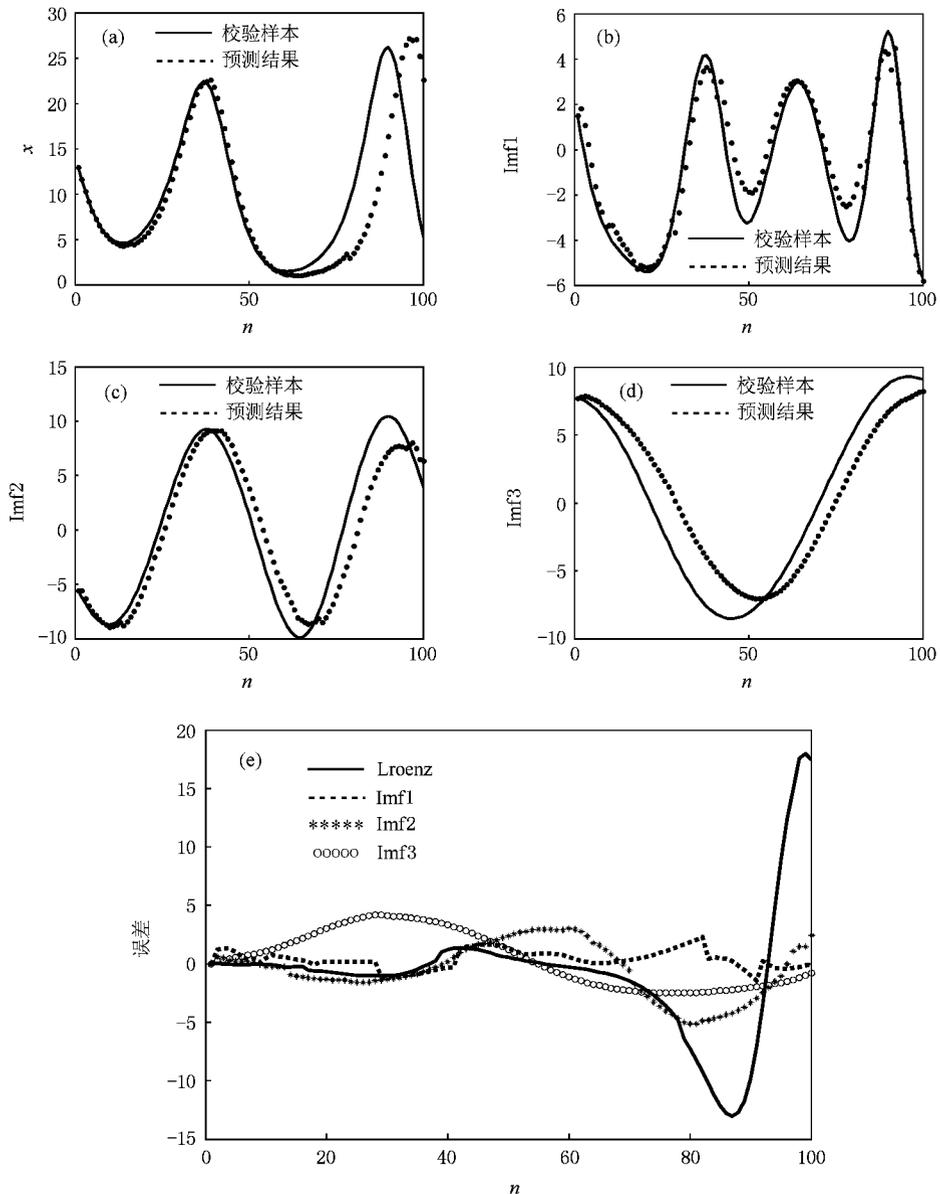


图 4 Lorenz 系统 x 分量及其 Imf 分量预测 (a) Lorenz 系统 x 分量及其预测结果 (b) Imf1 及其预测结果; (c) Imf2 及其预测结果 (d) Imf3 及其预测结果 (e) 预测误差比较

表 2 Lorenz 系统 x 分量响应及其主要 Imf 分量的最大 Lyapunov 指数

数据	τ	m	P	λ_1	t_0
Lorenz	10	3	71.9	1.5015	0.666
Imf1	10	7	39.3	0.7875	1.267
Imf2	18	6	77.3	0.0231	43.29
Imf3	27	6	117.3	0.0088	113.6

根据表 2 的参数,采用最大 Lyapunov 指数预测法对 Lorenz 系统的 x 分量和经 EMD 分解后的 Imf1, Imf2 和 Imf3 进行预测,原数据为 50000 点和上节类似,为消除 Imf 端点效应影响,截取 10000—39900 点的数据作为已知样本,后 100 点作为校验样本,测试预测的准确度.发现对于 Lorenz 系统的 x 分量在 70 点以前,预测结果与校验样本吻合较好,在 70 点后,随着预测点超过最大可预测时间 0.666,预测误差急剧放大,预测失去意义.对于 EMD 分解后的数据 Imf1, Imf2 和 Imf3,在 70 点以前预测结果误差较原始数据预测结果大,但在 70 点以后,由于三个分量的

可预测时间比原始数据大 1 倍左右,因此,预测误差和原始数据的误差相比在逐渐减小,在 100 点时可发现 EMD 分解后的预测明显优于原始数据的直接预测(见图 4(e)).

7. 结 论

本文采用 EMD 方法对 Duffing 方程和 Lorenz 系统的混沌响应进行了预测,结果表明,分解后的信号和原始信号相比具有较小的最大 Lyapunov 指数,这对于提高混沌系统可预测时间有着重要意义. EMD 分解后信号在短时混沌预测中可能没有原信号的预测精度高(主要由于 EMD 分解一定程度上破坏了原始信号的吸引子形状),但在长时混沌预测中表现出了明显的优势.在绘图时,因 Imf 信号后几级一般为类周期信号或单调函数,且幅值很小,为清楚显示主要 Imf 分量的预测效果,并未给出后几级 Imf 信号预测的误差.

- [1] Huang R S 2000 *Chaos and Its Application* (Wuhan: Wuhan University Press) (in Chinese) [黄润生 2000 混沌及其应用(武汉:武汉大学出版社)]
- [2] Takens F 1981 *Dynamical Systems and Turbulence* (Berlin: Springer) p366
- [3] Su S F, Lin C B, Hsu Y T 2002 *IEEE Transactions on Systems* **32** 416
- [4] Linsay P S 1991 *Phys. Rev. Lett.* A **153** 353
- [5] Zhang J S, Xiao X C 2001 *Acta Phys. Sin.* **50** 2092 (in Chinese) [张家树、肖先赐 2001 物理学报 **50** 2092]
- [6] Lü J H 2000 *Proceedings of the Chinese Society of Electrical Engineering* **20** 80 (in Chinese) [吕金虎 2000 中国电机工程学报 **20** 80]
- [7] Lü J H, Lu J A, Chen S H 2002 *Chaos Time Series Analysis and Its Application* (Wuhan: Wuhan University Press) (in Chinese) [吕金虎、陆君安、陈士华 2002 混沌时间序列分析及其应用(武汉:武汉大学出版社)]
- [8] Huang N E, Shen Z, Long S R 1998 *Proceedings of Royal Society (London)* A **454** 903
- [9] Wolf A, Swift J B, Swinney H L 1985 *Phys. D* **16** 285
- [10] Fraser A M, Swinney H L 1986 *Phys. Rev. A* **33** 1134
- [11] Cao L Y 1997 *Phys. D: Nonlinear Phenomena* **110** 43
- [12] Yang Y, Yu J D, Cheng J C 2005 *Journal of Vibration and Shock* **24** 70 (in Chinese) [杨宇、于德介、程军圣 2005 振动与冲击 **24** 70]
- [13] Datig M, Schlurmann T 2004 *Ocean Engineering* **31** 1783
- [14] Gagarin N, Huang N E, Oskard M S, Sixbey D G, Mekemson J R 2004 *Int. J. Vehicle Design* **36** 87
- [15] John G, Philip H 1997 *Non-Linear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (New York: Springer-Verlag)
- [16] Yeng Y F, Ren X M, Qin W Y 2008 *Nonlinear Analysis—Theory Methods & Applications* **68** 582

Prediction of chaotic time series based on EMD method^{*}

Yang Yong-Feng[†] Ren Xing-Min Qin Wei-Yang Wu Ya-Feng Zhi Xi-Zhe

(*Northwestern Polytechnical University, Xi'an 710072, China*)

(Received 15 December 2007 ; revised manuscript received 18 January 2008)

Abstract

In order to improve the nonlinear response prediction precision in a long period , the empirical mode decomposition(EMD) method is introduced in the nonlinear prediction. Here , the EMD method is used to decompose the signal , the rosenstein method is used to calculate the largest Lyapunov exponent(LLE) , and then the prediction results are obtained on the basis of the LLE. The simulation results of Duffing equation , Lorenz system and cracked rotor system show that the EMD 's signals have smaller LLE than the original signal. In this way , the maximum prediction time of a nonlinear signal can be obtained.

Keywords : EMD , chaos , prediction

PACC : 0545

^{*} Project supported by the Natural Science Foundation of Shaanxi Province , China (Grant No.2007A08).

[†] E-mail : yyfnpu@hotmail.com