第 4 卷第 1 期
2006 年 3 月

动力学与控制学报
JOURNAL OF DYNAMICS AND CONTROL

Vol.4 No.1
Mar. 2006

基于动态标架分割法的动力学系统复杂性分析

李建康 张春利 解幸幸 李 舒

(江苏大学理学院,镇江 212013)

摘要 根据符号动力系统与真实动力学系统拓扑共轭的特性,本文提出动态标架分割法,把动力学系统的某时间变量序列转化成符号序列.运用 Lemple-Ziv 复杂度算法计算该符号序列的复杂度值,据此对动力学系统的复杂性进行分析,从而可以对动力学系统的性质进行定性地判断.以杜芬振子为例,数值模拟结果表明基于动态标架分割法计算得到的复杂度能够很好地描述系统的复杂性,并可定性地判断系统的性质.

关键词 符号时间序列,动态标架分割法,Lemple-Ziv 复杂度,动力学系统

引言

近年来复杂性、非线性科学的不断发展,人们逐渐对一些存在的复杂现象进行研究,而不是以往的线性化处理.20 世纪 60 年代,柯尔摩戈若夫(Kolmogorov),索莫洛夫(Solomonoff),柴廷(Chaitin)等人分别对复杂性作了相近的定义[3]:一个系统的复杂程度与该系统行为(空间结构或时间序列所表示的变化)的最小描述有关.其基本思想是:某事物的算法复杂度等于产生该事物的图形结构或符号序列的最短程序的长度与该图形结构或符号序列本身大小之比的极限(当后者趋于极限时).简言之,描述系统行为的程序语句数越长,系统越复杂.1921 年莫尔斯[4](M. Morse)首先注意到符号动力学方法对于研究动力系统的重要性.1938 年莫尔斯和黑德隆(G. A. Hedlund)研究动力系统拓扑理论时提出,许多动力系统与符号序列拓扑共轭[1].因此,可以借助符号动力系统理论来研究真实动力系统的内在特性.为了具体地用由符号序列表示的真实系统,以便对系统的复杂性进行计算,兰帕尔(A. Lemple)[2]和齐夫(J. Ziv)[2]提出一种简单算法,由此算法得到的复杂度为 Lemple-Ziv 复杂度,记为 C_{LZ}.

实际工程中,一些复杂动力学系统都具有非线性因素,从而会导致系统出现极为复杂的运动现象.传统分析方法:统计分析法(如求关联系数)、频闪采样法、功率普分析法、自相关函数法等,这些方法对这些复杂问题的解决具有一定的局限性.现代工程师如果能够事先对系统性质进行整体定性判断,然后根据判断结果再进行问题的分析,那将会大大提高工作效率.虽然通常情况下,人们对系统知之甚少,甚至不知道其动力学规律,但是系统变量的时间序列容易得到.因此,本文根据符号动力学理论,把得到的实际动力学系统的时间序列按照动态标架分割法(Dynamic Frame Partition,简记为 DFP)符号化,把时间序列转化成符号时间序列并计算其 Lemple-Ziv 复杂度,将复杂度作为描述动力学系统复杂性的一个特征量,从而对动力学系统的复杂性作定性描述.

1 时间序列符号化

符号动力系统[5]是形式上最简单的一种动力系统,它是实际动力学系统的一种高度概括和抽象.符号动力学[6,7]的状态空间或相空间就是符号序列空间,动力学则由移位操作实现.上述表明真实动力学系统与符号动力系统可以建立对应关系.在建立真实动力学系统与符号动力系统的对应关系时,必须注意到动力学过程的几何或物理特点,采取能够反映这种特点的粗粒化方法.下文介绍的动态标架分割法能够准确地反映真实动力学系统的本质特性.

1.1 常用符号化方法说明

引入划分 $P = \{P_1, P_2, \cdots, P_n\}$,并把系统状态空间划分为 $m = (n+1)d$ 个单元,其中 d 是状

2005-09-13 收到第 1 稿,2005-11-06 收到修改稿.

态空间的维数，q 是划分个数，$q = 1$ 为最简单的二进制划分. 每个单元用一个符号 $s_r = \{s_1, s_2, \cdots, s_m\}$ 进行标记. 从而，在相空间为连续曲线的系统的轨迹，将被转换为符号序列. 这些符号标记了被轨迹所访问的单元. 这一符号模型以最简单地可能方式完全地描述了系统的动力学特征. 它为研究复杂动力系统的复杂性行为提供了一种有力手段.

将原始时间序列转化为符号序列的形式，通常采用两种方法：静态分割区间法和一阶差分法. 如图 1 所示为二进制分割区间法，测量值在分割线之上取符号 1，之下取符号 0. 如图 2 所示为一阶差分法，规定相邻的两点差值为正，则取符号 1，为负则取符号 0.

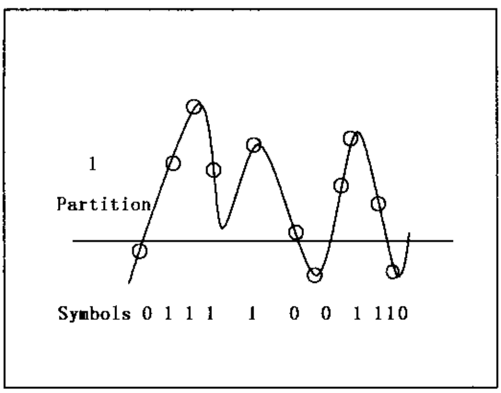

图 1 分割区间法符号化示意图

Fig. 1 Fixed partition transformation of a data series into a symbol series

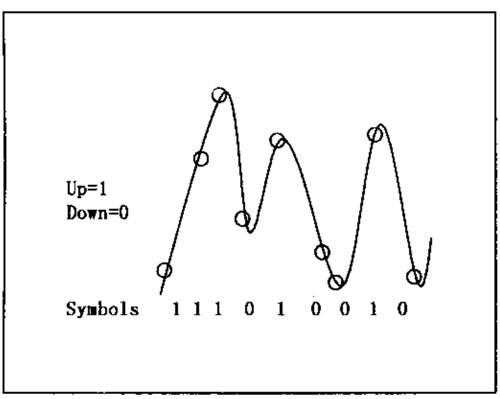

图 2 一阶差分法(前差分) 符号化示意图

Fig. 2 First difference transformation of a data series into a symbol series

这两种方法都有不足之处[8]：一阶差分法对于突发的大噪声不敏感，故通常采用分割区间法. 对于分割区间法，分割区间越多，原始数据转化为符号时间序列就越细化，但是会带来一些负面影响，例如对噪声的抑制作用减弱，因此需要选择合适的分割区间. 此外两种方法不能对各个时刻状态的细部特征进行描述. 作者根据复杂动力学系统的运动特性，在上述两种符号化方法基础上提出动态标架分割法.

1.2 动态标架分割法

通常复杂动力系统的动力学规律在某量的时间变化数据(或曲线) 上表现的很不明显，甚至杂乱无章. 分割法从静态上对系统进行划分，一阶差分法仅是简单意义上的动态. 上述两种方法在复杂动力系统的时间序列分析中，会丢失很多关键的信息. 因此也就失去了符号动力系统的意义. 动力学系统时间序列数据(或曲线) 上任何不同时刻的状态在几何、物理上都有一定的内在联系. 为此引进标架(如图 3 所示)，每一状态对应一个标架，标架横轴是静态的分割标准；纵轴是系统对应的状态线.

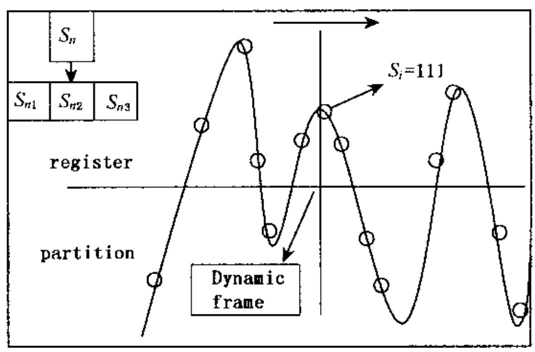

图 3 动态标架分割法示意图

Fig. 3 Dynamic frame partition transformation of a data series into a symbol series

借助计算机存储理论，我们把动力学系统的每一时刻的状态变量(注意是动力系统中的任一个个变量) 采用 3 位寄存器格式表示，如图 3，第 t_n 时刻的状态的符号表示为：$S_n = s_{n1} s_{n2} s_{n3}$. 第一位寄存符号 s_{n1} 和第三位寄存器符号 s_{n3} 代表状态变量数据值 V_n 与标架纵轴前后的数据值 V_{n-1}、V_{n+1} 之间关系的符号，s_{n2} 表示状态变量数据值与标架横轴($\overline{V}$，所有数据的平均值) 之间的关系. 其符号化具体规则如下

$$s_{n1} = \begin{cases} 1 & v_n \geqslant v_{n-1} \\ 0 & v_n < v_{n-1} \end{cases}$$

$$s_{n2} = \begin{cases} 1 & v_n \geqslant \bar{v} \\ 0 & v_n < \bar{v} \end{cases}$$

$$s_{n3} = \begin{cases} 1 & v_n \geqslant v_{n+1} \\ 0 & v_n < v_{n+1} \end{cases} \tag{1}$$

那么寄存器 $S_n = s_{n1}s_{n2}s_{n3}$ 就是动力系统状态在 t_n 时刻的符号状态表示,整个动力系统就可表示为 $S = S_1S_2\cdots S_i\cdots$. 整个符号空间的每个元素共有 8 种情况,其集合 s_k 表示如下

$$s_k = \{000,001,011,111,010,100,110,111\} \tag{2}$$

为了表达的方便可进一步根据二进制关系把编码(其实就是符号化力求其简洁) 如(表 1) 所示. 因此实际的动力系统就可以简洁地被只含 8 个符号的符号时间序列 $S = S_1S_2\cdots S_i\cdots, S_i \in (0,1,2,3,4,5,6,7)$ 来替代(也可以用其他符号代替). 这样就可以运用 Lemple-Ziv 复杂度算法对一个复杂动力系统的内部复杂特性进行度量.

表 1　二进制编码规则

Table 1　The code rule of binary system

状态	000	001	011	010	100	101	110	111
编码	0	1	2	3	4	5	6	7

2　Lemple-Ziv 复杂度算法

根据 Kolmogorov 理论,复杂性可认为是产生某给定"0,1"序列最少的计算机程序的比特数,但是这一表述没有一般的算法能够实现,有时根本就不能计算. 1976 年,A · Lemple 和 J · Ziv 提出一种度量符号序列复杂性的简单算法,称为 Lemple-Ziv 复杂度.

Lemple-Ziv 复杂度算法描述如下:

令 $c(n)$ 为某一给定的(一般为"0,1"序列) 符号序列 $S = s_1s_2\cdots s_n$(s_i, $i \in 1,2,\cdots,n$ 为一个字符) 的复杂性计数(即复杂度). 令 S、Q 分别代表两个字符串, SQ 表示把 S、Q 两个字符串相加成组成的总字符串, SQP 表示把 SQ 中最后一个字符删去所得的字符串(P 表示删去最后一个字符的操作). 令 $V(SQP)$ 表示 SQP 的所有不同的子串的集合. $c(n)$, S、Q 的初始化为 $c(n) = 1$, $S = s_1$, $Q = s_2$, 因此 $SQP = s_1$. 现假定 $S = s_1s_2\cdots s_r$, $Q = Q_{r+1}$ 若 $Q \in V(SQP)$,则表示 s_{r+1} 是 $S = s_1s_2\cdots s_r$ 字符串的一子串,那么 S 不变,只将 Q 更新为 $Q = s_{r+1}s_{r+2}$,再判断 Q 是否属于 $V(SQP)$(此时,因为 S 不变, Q 更新了,所以 SQP 也要更新),如此反复进行,直到发现 $Q \notin V(SQP)$ 时为止. 设此时 $Q = s_{r+1}s_{r+2}\cdots s_{r+i}$,即表明 $s_{r+1}s_{r+2}\cdots s_{r+i}$ 不是 $s_1s_2\cdots s_rs_{r+1}\cdots s_{r+i-1}$ 的子串. 因此 $c(n)$ 值将加 1. 然后将上述 Q 组合到 S 中,使 S 更新为 $S = s_1s_2\cdots s_rs_{r+1}\cdots s_{r+i}$,而取为 Q 为 $Q = s_{r+i+1}$,重复以上步骤,直到 Q 取到最后一位为止. 这样就把 $s_1s_2\cdots s_n$ 分成了 $c(n)$ 个不同的子串. 根据 A · Lemple 和 J · Ziv 的研究,对几乎所有属于[0,1] 区间, x 对应的二进制分解所表示的序列都会趋向一个定值

$$\lim_{n\to\infty} c(n) = b(n) = n/\log_2 n \tag{3}$$

$b(n)$ 是随机序列的渐进行为,可以用它来使 $c(n)$ 归一化,称为"归一化复杂度"

$$C_{LZ}(n) = c(n)/b(n) \tag{4}$$

通常就用这个函数来表达时间序列的复杂度的变化. 从该算法可以看出完全随机序列复杂度趋于 1,而规则序列则趋于 0[2], C_{LZ} 值越大动力系统越复杂,相反越小动力学系统的规律性就越明显.

例如,序列 $S = 10101010$ 的复杂度可由以下步骤得到

(1) 先取 $S = s_1 = 1$, $s_2 = 0$, $SQ = 10$, $SQP = 1$, $Q \notin V(SQP)$, Q 为插入, $SQ = 1 \cdot 0 \cdot$.

(2) $S = s_1s_2 = 10$, $Q = s_3 = 1$, $SQ = 101$, $SQP = 10$, $Q \notin V(SQP)$,故第二步 Q 是复制, $SQ = 1 \cdot 0 \cdot 1$.

(3) $S = s_1s_2 = 10$, $Q = s_3s_4 = 10$, $SQ = 1010$, $SQP = 101$, $Q \in V(SQP)$,故第三步 Q 仍为复制, $SQ = 1 \cdot 0 \cdot 10$.

(4) 以后几步都是重复(2) 和(3),即都是复制. 于是得: $S = 1 \cdot 0 \cdot 101010$,整个序列分为三段,故 $c(8) = 3$.

(5) $b(8) = 8\log_2 8 = 24$,故得归一化复杂度为: $C_{LZ} = c(8)/b(8) = 3/24 = 0.125$ 结果表明此序列 S 的复杂度较小,原因在于此序列是一周期序列.

3　数值分析 杜芬动力系统

动力学方程

$$\begin{cases} \ddot{x} + 0.3\dot{x} - x + x^3 = F\cos 1.2t \\ x(0) = \dot{x}(0) = 0 \end{cases} \tag{5}$$

该动力学系统中只有一个参变量 F,根据非线性理论可知,F 取不同的值,系统的运动将出现不同的运动性质(周期,拟周期,混沌等).分别运用静态分割法和动态标架分割法把所得到的上述杜芬系统的时间序列符号化,并计算其 Lemple-Ziv 归一化复杂度.分别取 F 为 :$F = 0.2, 0.28, 0.32, 0.6, 0.7, 0.8$,计算其数值解,采样间隔为,解区间,绘制运动时间曲线图(左)(x-t) 和相图(右)(x-$\dot{x}$) 如图 4 所示.从杜芬系统的相图(右),可知当 $F = 0.2, 0.8$ 时,杜芬系统作拟周期运动,而当 $F = 0.28, 0.32, 0.6$ 时系统在相图间中的轨迹逐渐出现‘混乱”现象,并且混乱程度逐渐‘增大”即系统出现了复杂的混沌运动.直接从得到的数据曲线图(x-t) 上可以确定,当 $F = 0.2, 0.8$ 时,系统具有良好的周期性质 ;可是当 $F = 0.28, 0.32, 0.6, 0.7$ 时情况就复杂了,根本不能从(x-t) 曲线图判断系统的复杂性.动力学系统在其相空间中的轨迹可以粗略地反映系统的某些运动性质,所以从相图的‘混乱”情况可定性地判断 :当 $F = 0.28, 0.32, 0.6, 0.7$ 时,杜芬系统运动的规律性明显小于 $F = 0.2, 0.8$ 时对应的系统运动的规律性.但是这还不能对它们之间的具体关系给以准确的描述.

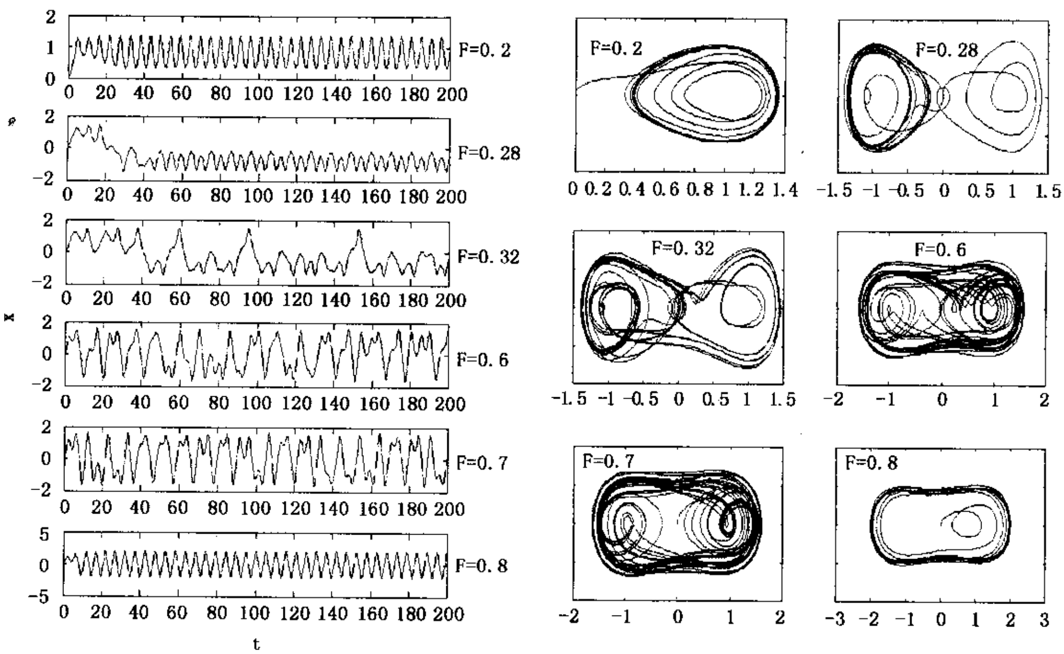

图 4 杜芬系统的 x-t 曲线(左) 和其在(x,$\dot{x}$) 相平面上的轨线(右)

Fig.4 Curve in the x-t coordinates and phase trajectory of the Duffing system

为满足实际问题的需要,取 Lemple-Ziv 归一化复杂度值作为系统的特征量并对杜芬系统的运动性质进行分析.分别取上述杜芬系统在各种情况下的数值解 $x(t)$,$t \in k \cdot \Delta t$ 为杜芬系统的拓扑离散化(取其 1024 点数据所表示的离散化动力系统),分别运用静态分割区间法和动态标架分割法符号化得到各种情况下的拓扑共轭的符号动力系统 :$\boldsymbol{Z}^p = s_1 s_2 \cdots s_k$ 和 $\boldsymbol{Z}^{DFP} = r_1 r_2 \cdots r_k$.根据上文介绍的简单易行的 Lemple-Ziv 复杂度算法分别计算符号动力系统 $\boldsymbol{Z}^p = s_1 s_2 \cdots s_k$ 和 $\boldsymbol{Z}^{DFP} = r_1 r_2 \cdots r_k$ 的归一化复杂度(简记为 :C_{LZC} 和 C_{LZD}).(其对应数值如表 2 所示) 从表中杜芬系统在各种情况下 Lemple-ziv 归一化复杂度值可知 :采用静态分割法符号化时,计算得到的结果为 :$C_{LZC}^{F=0.2} > C_{LZC}^{F=0.28}$,$C_{LZC}^{F=0.2} > C_{LZC}^{F=0.32}$,$C_{LZC}^{F=0.6} > C_{LZC}^{F=0.8}$,由相图可知,$F = 0.2$ 和 $F = 0.32$,$F = 0.32$、0.6 与 $F = 0.8$ 时的复杂度与实际不符,这种符号化方法不能准确地反映系统的复杂性.当采用动态标架分割法时,计算得到得各种情况下的系统复杂度基本清楚地显示系统的真实复杂性.由它们的复杂度值可以判断 :$F = 0.2, 0.8$ 时系统的复杂性较小、周期性较好以及运动较为稳定 ;而 $F = 0.32, 0.6, 0.7$ 时系统的复杂度较大,系统运动较为复杂、具有强非线性 ;系统复杂度值的大小也反映出 :$F = 0.6, 0.7$

时系统运动状态更为复杂,与实际情况基本相符.由以上分析知,动态标架分割法能够较准确地把真实动力系统转化成与其拓扑的符号动力系统.根据 *Lemple-Ziv* 复杂度值可以描述动力学系统的复杂性,据此可定性地判断系统是否存在非线性因素和定性地分析系统运动的动力学规律性.为工程师在实际问题中是否考虑非线性因素时提供了简单可行的方法.

表 2　两种符号化方法在不同情况下对应的 Lemple-Ziv 归一化复杂度

Table 2　The Lempel-Ziv normalized complexity of different conditions with two methods of transformation

F	0.2	0.28	0.32	0.6	0.7	0.8
C_{LZC}	0.1172	0.06836	0.0977	0.1270	0.1465	0.1270
C_{LZD}	0.2246	0.24410	0.2734	0.3516	0.3711	0.2148

4　结论

动态标架分割法是把真实动力系统时间序列拓扑转化成符号动力系统的符号时间序列的一种强有力的数据处理方法,它能够准确反映动力学系统的真实态性、并且操作简单快捷.在复杂动力学系统定性分析分析中有很大的优势.

Lemple-Ziv 复杂度能够对复杂动力学系统的复杂性进行较为准确的描述,根据系统变量的时间序列得到的复杂度值,可以对动力学系统的复杂性进行定性地分析,从而对系统的性质有有一个整体把握.在实际问题中,直接运用测量得到的时间序列基于动态标架分割法计算复杂度值,使实际问题的解决变得较为简单,且结果可靠.本文方法在动力学系统前期分析中,具有对系统作定性地分析的良好特性;但是如何根据复杂度值对系统某个具体性质定量分析还需要进一步研究;进一步的研究结果可以在动力学结构损伤检测中发挥作用以及对非线性系统进行非线性检测和分析诊断.

参　考　文　献

1　Morse M,Hedlund GA. Symbolic dynamics. *Journal of the American Mathematical society*,1938,60(1):815～816

2　Lempel A,Ziv J. On the complexity of finite sequences. *IEEE Trans Inf Theory*,1976,22:75～85

3　刘秉正,彭建华.非线性动力学.北京:高等教育出版社,2004(liubingzheng,pengjianhua. Nolinear dynamics. Beijing:Higher Education Press,2004(in Chinese))

4　Morse M. Recurrent geodesies on a surface of negative curvature. *Journal of the American Mathematical Society*,1921,22:84～100

5　郑伟谋,郝柏林.实用符号动力学.上海:上海科技教育出版社,1994(Zheng Weimou,HAO Bai-lin. Shanghai:Shanghai Scientific and Technological Education,1994(in Chinese))

6　JACLSON EA. Perspectives of Nonlinear Dynamics. Cambridge England:Cambridge University Press,1989

7　周作领.符号动力系统.上海:上海科技教育出版社,1997(Zhou Zuoling. Shanghai:Shanghai Scientific and Technological Education,1997(in Chinese))

8　金宁德,李伟波.非线性时间序列的符号化分析方法.动力学与控制学报,2004,2(3):54～59(Jin Ningde,Li Weibo. Study On The Analysis Method Of Nonlinear Symbolic Time Series. *Journal of Dynamics and Control*,2004,2(3):54～59(in Chinese))

ANALIZING COMPLEXITY OF DYNAMICAL SYSTEMS BASED ON DYNAMICAL FRAME PARTITION

Li Jiankang Zhang Chunli Xie Xingxing Li Shu
(*Faculty of Science, Jiangsu University, Zhenjiang* 212013, *China*)

Abstract There is a good property of topology conjugated between the dynamics system and the symbolic dynamics. In this paper, dynamic frame partition method, which is used to transform the time series of the dynamical system into the symbolic time series, is proposed. We compute the Lempel-Ziv complexity value of the symbolic series. And then the intrinsic complexity of the system can be made qualitative analysis. At last, a numerical simulation result indicates that the proposed method can describe the complexity of the dynamics system and can qualitatively estimate the property of the system.

Key words symbolic time series, dynamical frame partition, Lemple-Ziv complexity, dynamics system

Received 13 September 2005, revised 06 November 2005.