

小波分解求取含噪混沌时间序列的关联维数

陈 铿 韩伯棠

(北京理工大学管理与经济学院 北京 100081)

摘 要 关联维数是一个混沌时间序列分析中很重要的参数,在分析了噪声对关联维数求取的影响后,提出了一种应用小波分解求取含噪混沌时间序列关联维数的方法,数值试验证实了我们这种方法的正确性与可行性。

关键词 关联维数,噪声,混沌时间序列,小波分解

Determining the Correlation Dimension of Noisy Chaotic Time Series with Wavelet Decomposition

CHEN Keng HAN Bo-Tang

(School of Management & Economics, Beijing Institute of Technology, Beijing 100081)

Abstract Correlation dimension is an important parameters in the analysis of chaotic time series. With analyzing the influence of the noise in chaotic time series on determining the correlation dimension, a method of determining the correlation dimension of noisy chaotic time series is proposed, the numeric result has shown that the method is valid.

Keywords Correlation dimension, Noise, Chaotic time series, Wavelet decomposition

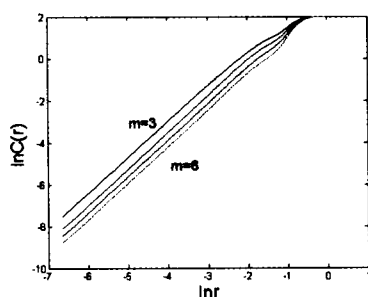
1 前言

关联维数是混沌时间序列分析中一个很重要的参数, Takens 定理^[1]指出:对于一无限长、无噪声的 d 维混沌吸引子的标量时间序列 $\{x(t)\}$, 我们总可以在拓扑不变的意义上找到一个 m 维的嵌入相空间, 只要 $m \geq 2d+1$ 。在没有任何对混沌时间序列的先验信息的情况下, $[d]+1$ 也是我们确定混沌吸引子最小嵌入维的下界, 其中 $[]$ 表示取整。

Grassberger 和 Procaccia 提出的 G-P 算法^[2]是确定关联维数的经典算法, 在混沌时间序列的分析中被广为采用。但噪声对关联维数求取的影响及消除却少有人考虑, 这是因为混沌时间序列与白噪声一样, 其频谱都是广谱, 它们在很大的范围内与白噪声的频谱几乎是重合的, 传统的 Fourier 滤波方法很难将两者分开^[3], 文[4, 5]采用了投影法消除混沌时间序列中的噪声, 但仅限在 5% 的噪声水平内, 不能消除高噪声水平。小波分析的出现, 为我们消除噪声对混沌时间关联维数的影响提供了一个有力的手段, 文[6, 7]研究了用小波对含噪混沌时间序列进行去噪, 但未讨论噪声对关联维数的影响。

2 G-P 算法

经典的 G-P 算法可描述如下, 对于混沌时间序列 $\{x(t)\}$



(a) 无噪声的关联维数曲线

(1) 给定不同的小正数 r

(2) 计算对应的关联积分 $C(r) = \frac{1}{N^2} \sum_{i,j=1, i \neq j}^N H(r - \|X_i - X_j\|)$

$X_j\|)$

其中, $\|X_i - X_j\|$ 为相点 X_i 与 X_j 间的距离, H 为 Heaviside 函数, N 为观测点数。

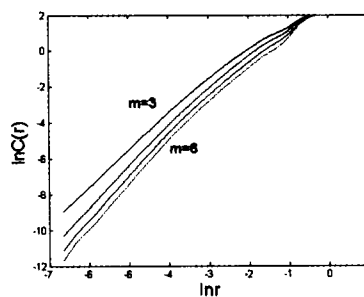
(3) 对 $\ln C(r)$ 与 $\ln r$ 的关系进行线性回归, 其斜率即为关联维数。

3 噪声对关联维数的影响

我们以 Lorenz 系统^[8]为例, 研究了噪声对混沌时间序列的关联维数求取的影响, 具体作法是对按(1)式用四阶 Runge-Kutta 法进行数值积分求出的 Lorenz 吸引子的 x 轴分量加上不同比例的 Gauss 白噪声, 再用 G-P 算法计算其关联维数曲线。

$$\begin{cases} \dot{x} = -\sigma(x-y) \\ \dot{y} = -xz + rx - y \\ \dot{z} = xy - bz \end{cases} \quad (1)$$

其中, $\sigma=10$, $r=28$, $b=8/3$, 数值积分时, 初值为 $(0, 1, 0)$, 时间步长为 $\Delta t=0.01$ 秒, 去掉了前面 10000 个暂态点。



(b) 去噪 (10%) 后的关联维数曲线

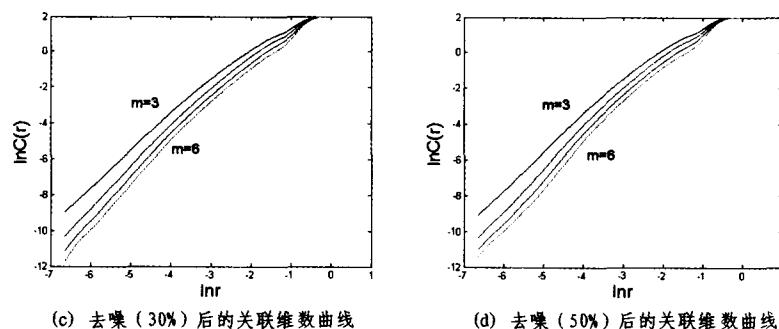


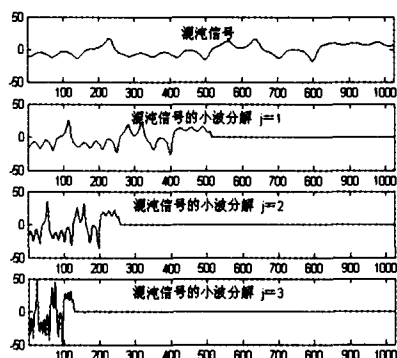
图1 噪声对 Lorenz 系统关联维求取的影响

图1显示了噪声水平对关联维数求取的影响。(a)为对无噪声的 Lorenz 混沌时间序列求取的关联维曲线图,(b)、(c)、(d)分别为对含噪 10%、30%和 50%的含噪 Lorenz 混沌时间序列所得关联维曲线图,计算中, $r_0=0.01$, $dr=0.0025$,延迟时间 $\tau=25\Delta t$,计算点数 $N=5000$,嵌入维数分别为 $m=3,4,5,6$ 。

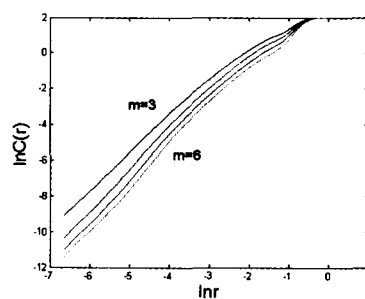
从图1中可以看出,无噪声的时候,不同嵌入维的关联维数曲线在无标度区近似平行,含噪情况下,不同嵌入维的关联维数曲线相互不平行,尾部呈散开状,且含噪水平越高,尾部散开得越厉害;同时,在无标度区,直线的斜率(即关联维的数值)也发生了较大的变化。显然,在这种情况下,无论对哪一条嵌入维的曲线求取关联维,所得到的结果都是错误的。

4 小波分解去噪

小波分析中的多分辨率分析(Multi-Resolution Analysis, 简称 MRA)^[9]将 $L^2(R)$ 按分辨率 2^{-j} 先分解成一嵌套的闭子空间 $\{V_j\}_{j \in Z}$,然后通过正交补的塔式分解,将 $L^2(R)$ 分解成



(a) 混沌时间信号的多尺度的小波分解



(b) 高斯白噪声的多尺度的小波分解

图2 混沌信号和高斯白噪声的多尺度小波分解图

从图2中可以看出,随着小波分解尺度的增加,混沌信号越来越集中于低频系数部分,而高斯白噪声却是一直分布在整个小波系数中,因此随着小波分解的进行,在某个尺度上,小波系数的高频部分将几乎全是噪声,将其清零,对信号进行重构,即可得到滤去噪声的混沌信号。

5 小波分解去噪求取关联维数

我们对不同含噪水平的 Lorenz 系统采用小波分解的方法进行了去噪处理,分解尺度均为 4,分解过程中使用的是 Daubechies 小波 db4^[11],小波分解去噪后用 G-P 算法求取的关联维数曲线见图3,从图中可以看出,经过小波分解去噪

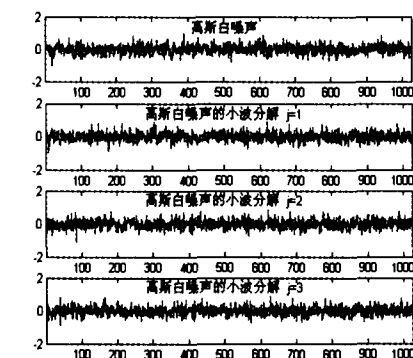
一连串的正交小波子空间序列 $\{W_j\}_{j \in Z}$ 的直和,小波分析即是函数 $f(t) \in L^2(R)$ 分别投影到不同分辨率的正交小波空间 $\{W_j\}_{j \in Z}$ 上进行分析与研究。

信号 $f(t) \in L^2(R)$ 的正交小波分解可以表示为:

$$f(t) = \sum_{k \in Z} c_{j,k} \varphi_{j,k}(t) + \sum_{i=1}^j \sum_{k \in Z} d_{i,k} \phi_{i,k}(t) \quad (2)$$

其中, $\varphi_{j,k}(t) = 2^{-j/2} \varphi(2^{-j}t - k)$ 与 $\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j}t - k)$ 分别为尺度空间 V_j 和小波空间 W_j 的一组正交基,称之为分辨率 2^{-j} 下的尺度函数和小波函数, j 称为小波分解的尺度, k 为时间平移, $c_{j,k} = \langle f(t), \varphi_{j,k}(t) \rangle$ 与 $d_{j,k} = \langle f(t), \phi_{j,k}(t) \rangle$ (符号 $\langle \cdot \rangle$ 表示内积, $j \in Z$) 分别为 $f(t)$ 在分辨率 2^{-j} 下的正交小波变换的系数,也称离散逼近与离散细节。

混沌时间信号的小波变换在每个尺度是时间的函数,其概率分布随时间的平移而变化,而噪声的小波变换在每个尺度上其概率分布不随时间的平移而变化^[10],图3给出了混沌信号和高斯白噪声的多尺度小波分解图。



后,不同含噪水平的关联维数曲线基本上均消除了尾部散开的情形,同时,不同嵌入维的曲线在无标度区也都已近似平行,选取各关联曲线中的无标度区(见图4),可计算出去噪后的 Lorenz 系统关联维数,计算结果见表1。

从表1中可以看出,去噪后所计算出的关联维数,已全部都在 2.0~2.7 间。从计算结果来看,正确的嵌入维($m=3$)的效果要好于其它嵌入维数;嵌入维数越大,效果相对越差;同时;噪声水平越高,去噪后的计算效果越差,这是显然的,因为去噪过程中总是要损失有效成分。值得注意的是,从计算结果来看,混沌时间序列中的噪声会加大计算出的关联维数值,这点在实际混沌时间序列关联维数的计算中必须予以考虑到。

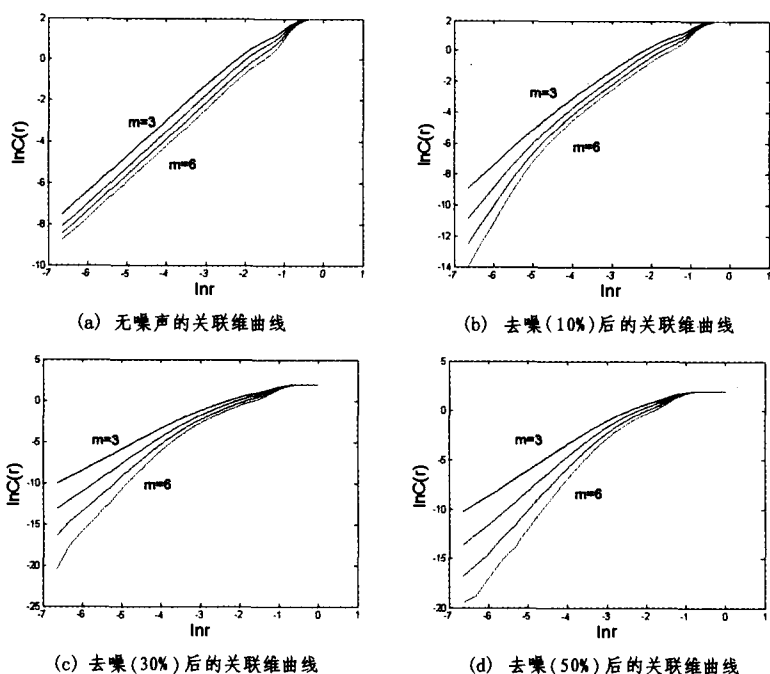


图3 小波去噪后的 Lorenz 系统求取的关联维

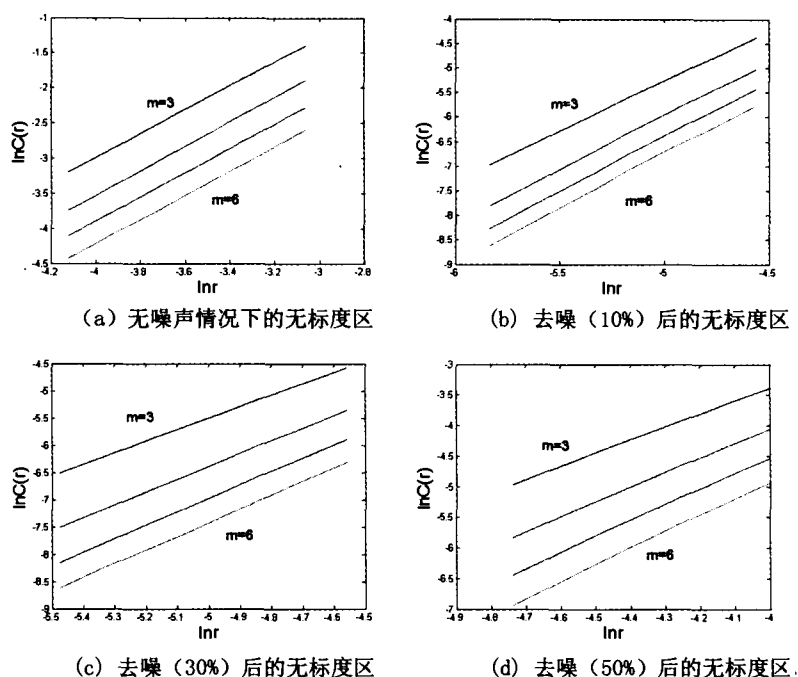


图4 小波分解去噪后选取的无标度区

表1 小波分解去噪后的计算所得关联维数

嵌入维	精确值	无噪声	去噪1 (10%)	去噪2 (30%)	去噪3 (50%)
3	2.06	2.0154	2.0256	2.1130	2.1367
4		2.0439	2.1820	2.3563	2.4086
5		2.0307	2.2212	2.4707	2.5932
6		2.0287	2.2310	2.5169	2.7112

结论 我们采用小波分解的方法对含噪 Lorenz 混沌时间序列进行了去噪处理,通过去噪处理后的关联维数的计算,说明我们所提出的这种方法是可行的,对实际混沌时间序列的去噪处理结果,我们将另文发表。

参考文献

- 1 Takens F. Detecting strange attractors in Turbulence. Lecture Notes in Mathematics, Berlin; Springer, 1981, 898: 366~381
- 2 Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. Physica 9D, 1983, 9: 189~208
- 3 Abarbanel H D I, Brown R, Sidorowich J J, et al. The analysis of observed chaotic data in physical systems. Reviews of Modern Physics, 1993, 65(4): 1331~1392
- 4 Hammel SM. A noise reduction method for chaotic system. Phys Lett, A, 1990, 148(8-9): 421~428

(下转第 241 页)

$\mathfrak{S}_B$ 表示基本对象: $\mathfrak{S}_B = \{\Theta_j | \Theta_j \in \mathfrak{S}_C\}$, $\mathfrak{S}_C$ 表示复合对象: $\mathfrak{S}_C = \{\Theta^j, \Theta^2, \dots, \Theta^m\}$ 。逻辑结构是人类对档案内容的理解, 可由公式(2)中的 $(\Phi, \delta, \alpha, \beta)^T$ 来描述。

另外, 可以定义档案的结构强度^[4]。假设一页文档可分为 n 个对象, 对应着 n 个变量, 用 H_i 表示第 i 个变量的偏熵 (partial entropy), H 表示整个文档的熵, 则档案的结构强度定义为:

$$S_S = \sum_{i=1}^n H_i - H \quad (4)$$

比如, 某档案包含 4 个复合对象, 对应的变量为 x_1, x_2, x_3, x_4 。则该档案的结构强度为

$$\begin{aligned} S_S &= \sum_{i=1}^n H_i - H \\ &= \sum_{i=1}^4 [-\sum_{j=1}^n p_j(x_i) \log p_j(x_i)] - \{-[\sum_{j=1}^n p_j(x_1, x_2, \\ &\quad x_3, x_4) \log p_j(x_1, x_2, x_3, x_4)]\} \\ &= -\sum_{i=1}^4 \sum_{j=1}^n p_j(x_i) \log p_j(x_i) + \sum_{j=1}^n p_j(x_1, x_2, x_3, x_4) \log \\ &\quad p_j(x_1, x_2, x_3, x_4) \end{aligned}$$

3 档案图像的虚拟化解释

3.1 档案定义的解释

档案是知识的载体, 它表达了人类的思想, 这种思想广泛涉及政治、经济、历史、文化、教育、艺术、科学、工程等^[3]。其中“思想”指出了档案的概念性含义, 对应于公式(1)定义的概念属性, 而“载体”暗示了档案是一具体的二维图像, 对应于公式(2)的具体属性。这种定义与文[5]的定义不谋而和, 即此处的“思想”就是指档案的历史联系, 此处的“载体”就是指文件实体集合, 并且“思想”与“载体”相辅相成, 二者不能独立构成档案, 它们的有机结合是现代档案学的逻辑起点。当人们将自己的思想以符号形式书写时, 概念结构就经过人脑编码成了具体结构, 即 $\omega \rightarrow \Omega$; 当人们对档案图像进行处理时, 具体结构就解码成了概念结构, 即 $\Omega \rightarrow \omega$ 。

按照公式(1)定义的概念档案, 具有语言不变性的重要性质, 即 $\frac{\partial \omega}{\partial \Sigma} = 0$ 。它意味着当档案从一种语言转换为另外一种语言时, 其概念结构保持不变。比如, 当档案从英文翻译为中文时, 档案的思想是相同的。

3.2 数码化过程的解释

档案数码化过程是档案具体结构 Ω 形成的过程, 它可视为档案的一种数字化虚拟。数字化虚拟来源于虚拟的本义——“假定”、“设想”。“数字化虚拟”不是假设了档案, 而是假设了数码符号并把它法定化。由这种法定化符号表达出来的档案信息就与假设无关了。

在档案的虚拟化过程中, 由于信息载体发生了变化, 引起信息与载体的关系也发生变化, 即信息对载体的依附性减弱,

信息的流动性和相对独立性加强。人们随之把注意的重心从物理载体逐步移向载体上的信息, 从物理管理为主转化为逻辑管理为主, 强调信息在不同空间位置 and 不同物理载体之间的永恒的流动性和信息的相对独立性。因此, 档案虚拟的本质意义是档案信息的独立与自由。

档案虚拟包括广义的虚拟和狭义的虚拟。广义的虚拟是指规则文明或符号文明, 是人类对各种规则的合成、选择及其演化。狭义的虚拟是指当代的数字化的表达方式和构成方式、实践方式和创造方式。

3.3 档案图像处理的解释

档案处理分为两个阶段: 档案分析和档案理解。从原始档案中提取几何结构的过程对应档案分析, 将几何结构映射为逻辑结构的过程称为档案理解。因此, 档案处理就是构建等式(2)和(3)表示的五元组的过程。档案分析就是提取等式(3)中的元素 $\mathfrak{S}$ 、 Θ 和 Θ_j 的过程, 即 Ω 的几何结构。档案理解就是发现等式(3)中 Φ 、 δ 、 α 和 β 的过程, 即 Ω 的逻辑结构。逻辑结构是人类对档案内容的理解, 一旦得到档案的逻辑结构, 就意味着档案可以采用人工智能技术或其它技术进行解码。

对于等式(4)定义的档案结构强度 S_S , 它代表各对象熵之和与整幅档案图像熵的差, 即各对象所包含信息量之和与整幅档案图像信息量之差, 因此, S_S 的值越大, 表明档案中各对象之间包含的信息量重叠较多, 即对象之间的“思想”相关度较高, 则档案图像的结构强度就越高; 反之, S_S 的值越小, 表明档案中各对象之间的独立性越强, 即对象之间的“思想”相关度较低, 则档案图像的结构强度就越低。

结论 档案图像是实态档案的虚拟化, 它提供的信息具有极大的空间流动性, 信息能以极高的速度、极大的容量、跨越广阔的地域自由流动, 为全社会共享, 这是传统档案绝对做不到的。从档案信息与载体的“二位一体”到信息的独立与自由, 是档案虚拟化的最具有本质意义的变化。从信息利用的角度来看, 档案图像还提供了利用的共享性、利用的复用性、利用的交互性和利用的灵活性。

参 考 文 献

- 1 丁海斌. 档案虚拟论[J]. 档案学通讯, 2004(2): 25~28
- 2 章毓晋. 中国图像工程. 中国图像图形学报[J], 1995, 1(1): 78~83
- 3 Tang Y Y, Cheriet C D Y, et al. Automatic Analysis and Understanding of Documents. Handbook of Pattern Recognition and Computer Vision. World Scientific, Singapor, 1993. 625~654
- 4 Watanabe S. Pattern Recognition, Human and Mechanical[M]. Wiley Interscience, New York, 1985
- 5 冯湘君, 刘新安. 现代档案学理论的逻辑起点[J]. 浙江档案, 2005(7): 7~9
- 9 Mallat S. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Trans on PAMI, 1989, 11(7): 674~693
- 10 Mallat S, Hwang WenLian. Singularity detection and processing with wavelets. IEEE Transaction on Information Theory, 1992, 38(2): 617~643
- 11 Daubechies I. Orthonormal Bases of Compactly Supported Wavelets. Commun. on Pure and Appl. Math, 1988, XLI: 909~996

(上接第 192 页)

- 5 余建祖, 苏楠. 混沌时序的噪声降低技术研究. 航空学报, 1999, 20(6): 498~502
- 6 黄显高, 徐健学, 何岱海, 等. 利用小波多尺度分解算法实现混沌系统的噪声缩减. 物理学报, 1999, 48(10): 1810~1816
- 7 马丽萍, 石炎福, 余华瑞. 含噪声混沌信号的小波去噪方法研究. 信号处理, 2002, 18(1): 83~87
- 8 Lorenz E N. Deterministic Nonperiodic Flow. J. Atmos. Sci, 1963, 20: 130~141