

# 有噪声的多维混沌时序的非线性检验

汤 龙 坤

( 华侨大学数学系, 福建 泉州 362021)

**摘要** 采用多变量时间序列替代数据生成原理, 生成实测多变量时间序列的多组替代时间序列. 综合线性冗余和广义冗余两种检验统计量, 对一种定量检验多维信号非线性方法的抗噪声能力进行分析. 从 Lorenz 和 Chen's 系统获得多维时间序列, 分别对加不同噪声水平的时间序列检验方法的有效性, 并比较不同噪声水平对方法的影响. 仿真实验结果表明, 方法具有很强的抗噪声能力.

**关键词** 多维信号, 广义冗余, 非线性, 噪声

**中图分类号** O 231.2

**文献标识码** A

混沌系统是非线性系统, 从时间序列中检验其非线性特性有必要的. 文〔1~ 3〕是对单个信号进行分析的, 而实际问题中常能同时获得多维信号. 文〔4〕引入广义冗余, 提出了一种定量检验多维信号非线性的方法, 它可用于检验多维时序的非线性, 特别是混沌时序的非线性特性, 但文〔4〕的方法仅对无噪声的信号进行讨论. 本文在此基础上, 对有噪声的多维信号进一步地验证该方法的有效性, 并探讨了其对噪声的容忍度.

## 1 线性和广义冗余的定义和计算

设  $\{x_{m,n}\}_{n=1}^N (m=1, 2, \dots, M)$  为通过实验或观测获得的多个变量的时间序列, 为了便于计算, 对每一变量分别标准化, 即

$$x_{m,n} = \frac{x_{m,n}^* - \frac{1}{N} \sum_{n=1}^N x_{m,n}^*}{\sqrt{\frac{1}{N} \sum_{n=1}^N (x_{m,n}^* - \frac{1}{N} \sum_{n=1}^N x_{m,n}^*)^2}}, \quad m=1, 2, \dots, M; \quad n=1, 2, \dots, N, \quad (1)$$

则对每一个  $m=1, 2, \dots, M$ ,  $\{x_{m,n}\}_{n=1}^N$  的均值为零, 方差为 1. 设  $C$  为时间序列  $\{x_{1,n}\}_{n=1}^N, \{x_{2,n}\}_{n=1}^N, \dots, \{x_{M,n}\}_{n=1}^N$  的关联矩阵,  $\lambda_m (m=1, 2, \dots, M)$  为  $C$  的特征值, 则  $\{x_{m,n}\}_{n=1}^N, m=1, 2, \dots, M$  的线性冗余为

$$L(x_{1,n}; \dots; x_{M,n}) = -\frac{1}{2} \sum_{m=1}^M \log_2 \lambda_m. \quad (2)$$

记  $M$  维向量  $\mathbf{x}_i = (x_{1,n}; \dots; x_{M,n})$  与  $\mathbf{x}_j = (x_{1,n}; \dots; x_{M,n})$  之间的距离为

$$\|\mathbf{x}_i - \mathbf{x}_j\| = \max_{1 \leq m \leq M} |x_{i,m} - x_{j,m}|,$$

定义  $M$  变量时间序列  $\{\mathbf{x}_n\}$  的一阶广义关联积分为

$$C_1^M(r) = \frac{1}{N} \sum_{i=1}^N \log_2 \left[ \frac{1}{N-1} \sum_{j=1, j \neq i}^N H(r - \|\mathbf{x}_i - \mathbf{x}_j\|) \right], \quad (3)$$

$$C_{1,m}^M(r) = \frac{1}{N} \sum_{i=1}^N \log_2 \left[ \frac{1}{N-1} \sum_{j=1, j \neq i}^N H(r - |x_{i,m} - x_{j,m}|) \right], \quad m=1, 2, \dots, M. \quad (4)$$

$\{x_{m,n}\}_{n=1}^N (m=1, 2, \dots, M)$  的一阶广义冗余为

**收稿日期** 2005-12-25

**作者简介** 汤龙坤(1977-), 男, 助教, 硕士, 主要从事混沌时间序列分析的研究. E-mail: emailtom02@sina.com

**基金项目** 华侨大学科研基金资助项目(04HZR07)

$$R(x_{1,n}; \cdots; x_{M,n}) = C^M(r) - \sum_{m=1}^M C_{1,m}^M(r), \tag{5}$$

其中,  $H(x)$  是 Heaviside 函数. 线性冗余只考虑了变量的线性依赖性, 而广义冗余检验了时序的所有依赖性<sup>[4, 5]</sup>.

## 2 多维时序替代数据生成方法

设  $\{x_{m,n}\}_{n=1}^N (m=1, 2, \cdots, M)$  为标准化后的实测时间序列, 对每个变量分别作离散 Fourier 变换, 有

$$y_{m,k} = \sum_{n=1}^N x_{m,n} \cdot \exp\left[\frac{2\pi(n-1)(k-1)}{N} \cdot j\right], \quad k=1, 2, \cdots, N; \quad m=1, 2, \cdots, M. \tag{6}$$

在式(6)中,  $j = \sqrt{-1}$ . 在区间  $[-\pi, \pi]$  内取  $N$  随机数, 使其满足

$$\varphi_1 = 0, \quad \varphi = -\varphi_{N-k-2}, \quad \varphi_{N/2+1} = 0, \quad k=2, 3, \cdots, N/2. \tag{7}$$

然后, 作离散 Fourier 逆变换, 有

$$z_{m,n} = \frac{1}{N} \sum_{k=1}^N (y_{m,k} e^{j\varphi_k}) \cdot \exp\left[-\frac{2\pi(n-1)(k-1)}{N} \cdot j\right], \quad n=1, 2, \cdots, N; \quad m=1, 2, \cdots, M, \tag{8}$$

则得  $\{x_{m,n}\}_{n=1}^N$  的替代时间序列  $\{z_{m,n}\}_{n=1}^N, m=1, 2, \cdots, M$ . 这样生成的替代时间序列是线性高斯多维随机时间序列<sup>[6]</sup>, 但对每个  $m=1, 2, \cdots, M, \{x_{m,n}\}_{n=1}^N$  与  $\{z_{m,n}\}_{n=1}^N$  有相同的幅度分布, 从而有相同的自相关函数. 这样替代时间序列保持了原时间序列的线性性质, 但改变了原时间序列的非线性特性.

## 3 非线性的定量分析

在确定了合适的广义冗余后, 有必要研究延迟时间序列之间的依赖性. 记广义冗余  $R(\tau_1, \tau_2, \cdots, \tau_{M-1}) = R(x_{1,n}; x_{2,n+\tau_1}; \cdots; x_{M,n+\tau_{M-1}})$ , 线性冗余  $L(\tau_1, \tau_2, \cdots, \tau_{M-1}) = L(x_{1,n}; x_{2,n+\tau_1}; \cdots; x_{M,n+\tau_{M-1}})$ . 假设  $R^0(\tau_1, \tau_2, \cdots, \tau_{M-1})$  和  $L^0(\tau_1, \tau_2, \cdots, \tau_{M-1})$  为原时间的广义冗余和线性冗余,  $R^s(\tau_1, \tau_2, \cdots, \tau_{M-1})$  和  $L^s(\tau_1, \tau_2, \cdots, \tau_{M-1})$  为某组替代时间序列的广义冗余和线性冗余, 由替代时间序列的生成原理可知它们是线性的, 因此  $R^s(\tau_1, \tau_2, \cdots, \tau_{M-1}) = L^s(\tau_1, \tau_2, \cdots, \tau_{M-1})$ . 由于替代时间序列和原时间序列有相同的自关联函数, 则

$$\begin{aligned} L^s(\tau_1, \tau_2, \cdots, \tau_{M-1}) &= L^0(\tau_1, \tau_2, \cdots, \tau_{M-1}), \\ R^s(\tau_1, \tau_2, \cdots, \tau_{M-1}) &= L^0(\tau_1, \tau_2, \cdots, \tau_{M-1}) \end{aligned}$$

等价于  $R^0(\tau_1, \tau_2, \cdots, \tau_{M-1}) = R^s(\tau_1, \tau_2, \cdots, \tau_{M-1})$ . 由于比较的量具有相同的数值性质, 只要在估计冗余  $R^0(\tau_1, \tau_2, \cdots, \tau_{M-1})$  和  $R^s(\tau_1, \tau_2, \cdots, \tau_{M-1})$  时用相同的数值参数, 它们的值就可以定量地比较. 定义广义冗余和线性冗余的显著性检验统计量( $D$ )为

$$D_R = \frac{|R^0(\tau_1, \tau_2, \cdots, \tau_{M-1}) - R^s(\tau_1, \tau_2, \cdots, \tau_{M-1})|}{\sigma_R(\tau_1, \tau_2, \cdots, \tau_{M-1})}, \tag{9}$$

$$D_L = \frac{|L^0(\tau_1, \tau_2, \cdots, \tau_{M-1}) - L^s(\tau_1, \tau_2, \cdots, \tau_{M-1})|}{\sigma_L(\tau_1, \tau_2, \cdots, \tau_{M-1})}. \tag{10}$$

在式(9), (10)中,  $\bar{R}(\tau_1, \tau_2, \cdots, \tau_{M-1})$  和  $\bar{L}(\tau_1, \tau_2, \cdots, \tau_{M-1})$  分别是多组替代时间序列广义冗余和线性冗余的均值,  $\sigma_R(\tau_1, \tau_2, \cdots, \tau_{M-1})$  和  $\sigma_L(\tau_1, \tau_2, \cdots, \tau_{M-1})$  分别是它们的标准差. 用  $t$  检验<sup>[3]</sup> 比较实测时间序列与替代时间序列的广义冗余的分布. 如果由多变量时间序列替代数据法生成 39 组替代时间序列, 且以原实测时间序列是线性高斯多变量随机序列为零假设, 则当显著性检验统计量的值大于 1.96 时, 实测时间序列就以大于 95% 置信水平拒绝零假设. 在实际应用中, 利用 Fourier 变换生成的替代时间序列并不总是保持原时间序列的线性性质. 因此, 只有原时间序列和替代时间序列的广义冗余有明显差异, 而它们的线性冗余没有明显差异时, 认为原时间序列具非线性特征才比较合理.

## 4 仿真计算

(1) Lorenz 系统.  $\frac{dx_1}{dt} = 10(x_2 - x_1), \frac{dx_2}{dt} = 28x_1 - x_2 - x_1x_3, \frac{dx_3}{dt} = x_1x_2 - \frac{8}{3}x_3$ , 取初值  $x_{1,0} =$

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

15.34,  $x_{2,0}=13.68, x_{3,0}=37.91$ , 积分步长  $h=0.04$ , 利用四阶 Runge-Kutta 法积分, 迭代生成长度的时间序列  $\{x_{m,n}^*\}_{n=1}^{\bar{N}} (m=1,2,3)$ . 从第 3 001 个点取  $N=1\,024$  的三变量时间序列  $\{x_{m,n}^*\}_{n=1}^N (m=1,2,3)$ , 与加噪声后的时间序列分别用多变量时间序列替代数据法生成 39 组替代时间序列. 线性及广义的显著性检验统计量对应于时间延迟间隔  $\tau$  的函数, 如图 1 所示. 图中实线是无噪声时序的, 其余 9 条虚线是有噪声时序的显著性检验统计量, 它们的噪声水平(方差的百分比  $\sigma\%$ ) 分别为 5, 10, 15, 20, 25, 30, 35, 40 和 45. 在图 1(a) 中, 除带有“ $\times$ ”号(35%, 45%)的有噪声时序有少数个别点大于 1.96 外, 其他都很小且分布在无噪声时序的附近. 可见, 噪声水平在 30% 以内, 原时序与替代时序的线性冗余没有明显差

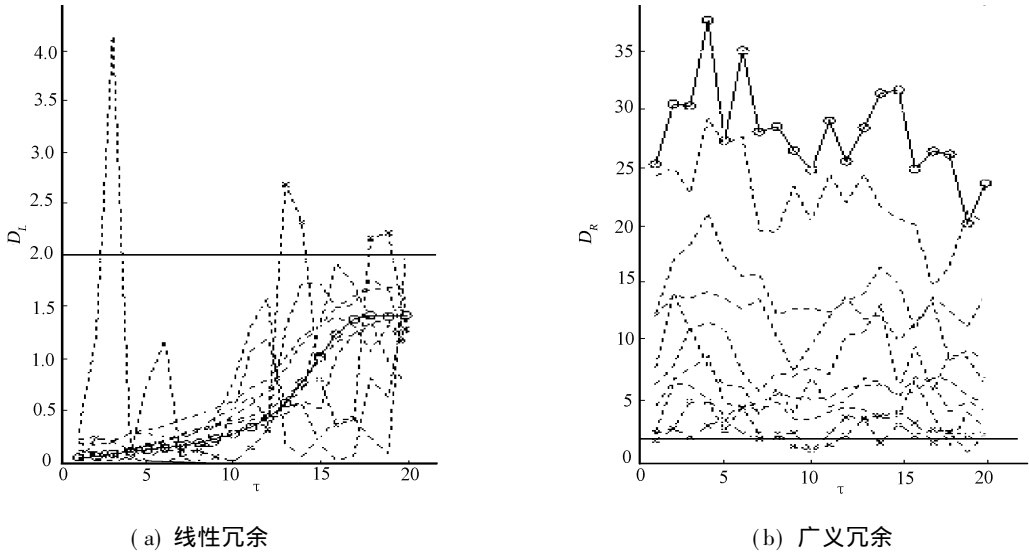


图 1 Lorenz 系统显著性检验统计量

异. 在图 1(b) 中, 自上而下, 随噪声水平的增大, 广义冗余的显著性检验统计量逐渐减小, 当噪声水平达到 40%, 45% (带有“ $\times$ ”号) 时, 统计量的值小于 1.96. 可见噪声水平在 35% 以内, 原时序与替代时序的广义冗余有明显差异. 综合上述两种冗余的显著性统计量可知, 当噪声水平在 30% 以内, 可认为时间序列具有混沌的非线性特征, 与实际情况相吻合.

(2) Chen's 系统.  $\frac{dx_1}{dt}=35(x_2-x_1), \frac{dx_2}{dt}=-7x_1+28x_2-x_1x_3, \frac{dx_3}{dt}=x_1x_2-x_3$ . 初值  $x_{1,0}=0, x_{2,0}=1.00, x_{3,0}=0$ , 积分步长  $h=0.01$ , 利用四阶 Runge-Kutta 法积分, 迭代生成长度  $\bar{N}=20\,000$  的时间序列  $\{x_{m,n}^*\}_{n=1}^{\bar{N}} (m=1,2,3)$ . 从第 3 001 个点取  $N=1\,024$  的三变量时间序列  $\{x_{m,n}^*\}_{n=1}^N (m=1,2,3)$  分别加入不同级别的噪声, 生成 39 组替代时间序列, 计算原时序和替代时序的线性冗余及广义冗余, 结

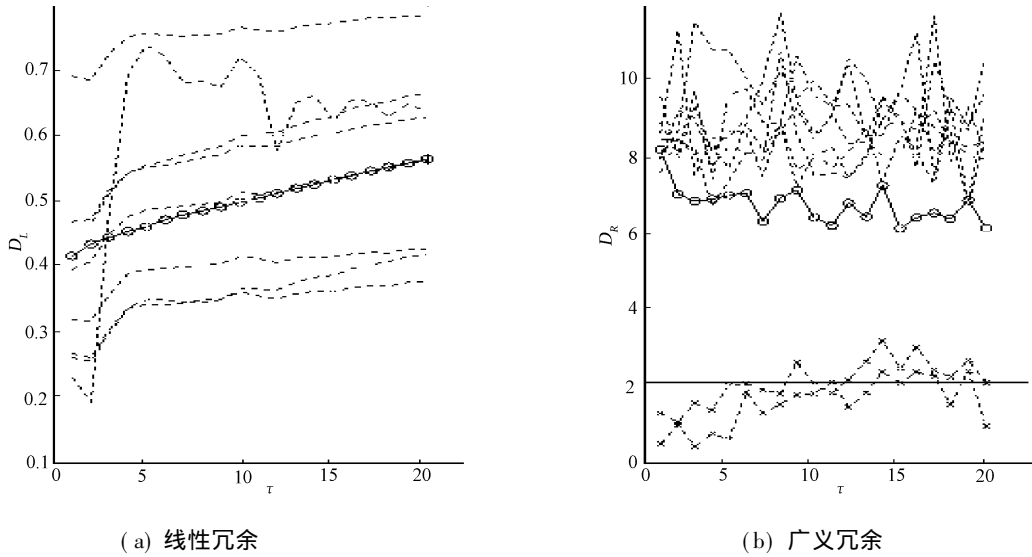


图 2 Chen's 系统显著性检验统计量

果如图 2 所示. 在线性冗余的函数图中, 无噪声时序(实线)和有噪声时序(虚线)的检验统计量数值都很小, 可见原时序与替代时序的线性冗余没有明显差异. 广义冗余的函数图中, 除了噪声水平为 45% 和 50% (带有“×”号)的时序有一部分检验统计量值小于 1.96, 其他的均大于 1.96. 可见噪声水平在 40% 以内, 原时序与替代时序的广义冗余有明显差异. 当噪声水平在 40% 以内, 可认为该时间序列具有混沌的非线性特征, 与实际情况相吻合.

## 5 结束语

仿真结果表明, 建立在广义关联积分基础上的定量检验多维时序非线性的方法, 适用于小数量数据, 具有计算简便、速度快等优点, 而且还有较强的抗噪能力, 能很好地检测到噪声水平高达 30% 的多维混沌时序的非线性特征.

## 参 考 文 献

- 1 Stam C J, Pijn J P M, Pritchard W S. Reliable detection of nonlinearity in experimental time series with strong period component[ J]. Physica, 1998, 112: 361~ 380
- 2 Palus M. Testing of nonlinearity using redundancies: Quantitative and qualitative aspects[ J]. Physica, 1995, 80: 186~ 205
- 3 Theiler J, Eubank S, Longtin A, et al. Testing for nonlinearity in time series: the method of surrogate data[ J]. Physica, 1992, 58: 77~ 94
- 4 王海燕, 汤龙坤. 一种定量检验多维信号非线性的方法[ J]. 信号处理, 2003, 19( 5): 407~ 410
- 5 Prichard D, Theiler J. Generalized redundancies for time series analysis[ J]. Physica, 1995, 84: 476~ 493
- 6 Prichard D, Theiler J. Generating surrogate data for time series with several simultaneously measured variables[ J]. Physical Review Letters, 1994, 73( 7): 951~ 954
- 7 吕金虎, 陆君安, 陈士华. 混沌时间序列分析及其应用[ M]. 武汉: 武汉大学出版社, 2002. 18~ 19

# Testing Nonlinearity in Multidimensional Noisy

## Chaotic Time Series

Tang Longkun

(Department of Mathematics, Huaqiao University, 362021, Quanzhou, China)

**Abstract** Applying the principle of generating surrogate data to generate surrogate time series of measured multivariate time series and using linear and generalized redundancy as test statistics, we discuss the scheme of resisting noise in a proposed quantitative method to detect nonlinearity in multidimensional signals. We obtain differently multidimensional time series from Lorenz and Chen's systems, and test the above method by the time series with different noise level. Computer simulation shows it has a good capability of resisting noise.

**Keywords** multidimensional signals, generalized redundancy, nonlinearity, noise