

# 一种时间序列的弱非线性检验方法<sup>\*</sup>

姜可宇<sup>†</sup> 蔡志明 陆振波

(海军工程大学电子工程学院, 武汉 430033)

(2007 年 4 月 16 日收到, 2007 年 6 月 18 日收到修改稿)

时间序列的非线性是判定该时间序列具有混沌特性的必要条件. 提出一种基于线性和非线性 AR 模型归一化多步预测误差比值的非线性检验量  $\delta_{\text{NAR}}$ , 采用替代数据法来检测时间序列中的弱非线性. 以 Lorenz 时间序列为例, 分析了估计非线性检验量  $\delta_{\text{NAR}}$  时各相关参数对弱非线性检测性能的影响. 通过混沌时间序列非线性检测试验, 对 4 种混沌时间序列中的 3 种非线性检验量  $\delta_{\text{NAR}}$  都表现出比基于 AIC 模型选择准则的非线性检验量  $\delta_{\text{AIC}}$  和非线性零阶预测误差  $\delta_{\text{ZP}}$  更强的弱非线性检测能力, 表明该非线性检验量具有较强的数据适应性, 而且对于不同的数据, 具有最佳非线性检测效果的参数比较稳定, 这为其他时间序列弱非线性检测的参数选择提供了参考.

关键词: 替代数据, 混沌时间序列, 非线性 AR 模型, 非线性检验

PACC: 0545, 0540

## 1. 引 言

混沌作为一种低维的非线性确定系统, 由于其对初始条件的敏感性, 会产生貌似随机的复杂信号. 由于确定性的混沌时间序列处理方法具有统计信号处理所不具备的优点, 如带内噪声的抑制等<sup>[1-3]</sup>. 20 世纪 90 年代, 众多学者纷纷尝试应用混沌理论来研究和解释各学科中的一些复杂现象. 但在利用混沌理论对时间序列进行处理之前, 应先判断该时间序列是否源于混沌系统. 关联维、最大 Lyapunov 指数等是判断时间序列是否源于混沌系统的重要指标, 但对于含噪声、有限长的实验观测时间序列, 估计得到的关联维数和最大 Lyapunov 指数存在较大误差, 容易出现误判. 由于非线性是时间序列具有混沌特性的必要条件, 在对时间序列进行混沌特性判定前, 时间序列的非线性检验是非常必要的.

时间序列的非线性检验方法主要分模型检验法和替代数据检验法两大类. 模型检验法就是直接比较时间序列的线性模型和非线性模型的建模误差或基于建模误差的统计量. 线性模型一般选择线性 AR 模型或 ARMA 模型<sup>[4]</sup>. 非线性模型可以选择 Volterra 级数模型<sup>[5]</sup>、指数 AR 模型<sup>[6]</sup>或随机非线性 AR 模

型<sup>[7]</sup>等. 如果时间序列基于非线性模型的建模误差明显小于线性模型, 则表明该时间序列含有非线性成分. 但模型本身和模型阶数等不易确定, 时间序列的长度及模型参数的估计方法等也直接影响模型对于时间序列的建模误差. 替代数据检验法<sup>[8]</sup>的主要思想是比较原数据与替代数据的各种非线性检验量, 如三阶自相关统计量<sup>[9]</sup>、时间逆不对称统计量<sup>[9]</sup>、相空间轨道转移误差<sup>[10]</sup>、循环图量化分析统计量<sup>[11]</sup>、非线性零阶预测误差<sup>[10]</sup>等. 如果原数据与替代数据非线性检验量的统计特性存在较大差异, 则表明原时间序列含有非线性成分.

在替代数据非线性检验方法中, 上述非线性检验量一般受噪声尤其是带内噪声的影响比较大, 当待检验时间序列中的噪声大到一定比例时, 非线性检验量对弱非线性的检测能力一般都迅速下降. 本文提出了一种基于线性 AR 模型和非线性 AR 模型归一化多步预测误差比值的非线性检验量, 采用替代数据检验法对时间序列进行非线性检验, 用叠加带内噪声的 Lorenz 混沌时间序列对该非线性检验量的弱非线性检测性能进行分析和研究, 并与两种常用的非线性检验量进行了对比, 用另外 3 种叠加带内噪声的混沌时间序列检验了该方法的数据适用性.

<sup>\*</sup> 国家重点基础研究发展规划项目(批准号 5132102ZZT32)资助的课题.

<sup>†</sup> 通讯联系人. E-mail: jiangky828@163.com

## 2. 替代数据检验法描述

替代数据检验法作为检验时间序列中非线性成分的方法是由 Theiler 等人<sup>[8]</sup>提出的. 其主要思想是首先指定某些特定的线性时间序列作为零假设, 再由待检验时间序列出发产生若干既满足零假设条件又保留待检验序列均值、方差、自相关等统计特性的线性替代数据, 然后分别计算待检验序列和替代数据的关联维数、非线性零阶预测误差、时间逆不对称统计量等非线性检验量, 最后根据原序列和替代数据非线性检验量的显著性差异程度在一定置信度内接受或拒绝最初指定的零假设条件. 如果零假设条件被拒绝, 则表明待检验序列包含非线性成分.

用替代数据检验法进行时间序列非线性检验的关键之一是产生与待检验时间序列统计特性尽量一致的替代数据. 本文所需替代数据和叠加到混沌时间序列的带内噪声均由随机相位法<sup>[8]</sup>产生. 零假设为待检验时间序列由线性相关的高斯噪声所产生. 如果拒绝该零假设, 则待检验时间序列包含非线性成分.

假设原序列为  $\{x_n, n = 1, 2, \dots, N\}$ , 随机相位法产生替代数据步骤如下:

- 1) 对原序列  $x_n$  进行 Fourier 变换求得其频谱  $f_k$ ;
- 2) 对频谱  $f_k$  每个离散频率点乘以一个因子  $e^{i\varphi_k}$ , 其中随机相位  $\varphi_k$  在  $[-\pi, \pi]$  范围内均匀分布;
- 3) 最后对相位随机化后的频谱  $f'_k$  作 Fourier 反变换, 得到满足零假设的替代数据.

随机相位法产生替代数据过程中需要注意的问题可以参考文献 [12]. 按以上步骤可以产生任意多个相同零假设下的替代数据.

## 3. 非线性检验量

### 3.1. 基于最优线性 AR 模型/指数 AR 模型的非线性检验量

文献 [4] 设计了一种基于 AIC 模型选择准则的非线性检验量  $\delta_{AIC}(H)$ . 对于  $k$  阶线性 AR 模型来说,

$$AIC(AR(k)) = N \log \epsilon_k + 2(k + 1), \quad (1)$$

式中  $\epsilon_k$  表示  $k$  阶线性 AR 模型拟合时间序列  $x_n$  的

误差.

$H$  步预测指数 AR 模型表示为

$$\hat{x}_{n-1+H} = \sum_{j=1}^p \left( a_j + b_j \exp\left(-\frac{x_{n-1}^2}{h}\right) \right) x_{n-j}, \quad (2)$$

其中

$$h = -\frac{\max(x_{n-1}^2)}{c}, \quad (3)$$

式中  $c = -30.0$ .

对于  $p$  阶指数 AR 模型来说,

$$AIC(\exp AR(p)) = N \log \epsilon_p + 2(2p + 1). \quad (4)$$

为了使两类模型的参数个数保持相等, 要求  $k_{\max} = 2p_{\max}$ . 非线性检验量  $\delta_{AIC}$  定义为

$$\delta_{AIC} = \min_p AIC(\exp AR(p)) - \min_k AIC(AR(k)). \quad (5)$$

### 3.2. 非线性零阶预测误差

文献 [7][8] 研究表明, 与轨道转移误差、时间逆不对称统计量等非线性检验量相比, 非线性零阶预测误差具有较稳定的微弱非线性检测能力.

根据 Takens 时延重构定理<sup>[11]</sup>, 时间序列  $\{x_n, n = 1, 2, \dots, N\}$  的相空间重构轨道点可以表示为

$$x_n = [x_n, x_{n+\tau}, \dots, x_{n+(k-1)\tau}]^T, \quad n = 1, 2, \dots, N - (k-1)\tau, \quad (6)$$

式中  $k$  为嵌入维,  $\tau$  为嵌入时延. 假设  $x_n$  为相空间中的任一轨道点, 搜索该轨道点的  $L$  个近邻点, 假设  $L$  个近邻点的序号分别为  $n_1, n_2, \dots, n_L$ .  $y_1^H, y_2^H, \dots, y_L^H$  表示这些近邻点演化  $H$  步后的像,  $y_j^H = x_{n_j+(k-1)\tau+H}$ . 轨道点  $x_n$  演化  $H$  步后的像可以用  $L$  个近邻点的  $H$  步像平均值作为预测值, 即

$$\hat{y}_n^H = \frac{1}{L} \sum_{j=1}^L y_j^H. \quad (7)$$

时间序列  $H$  步归一化非线性零阶预测误差表示为

$$\delta_{2p} = \left[ \frac{\sum_{n=1}^{N-(k-1)\tau-H} (\hat{y}_n^H - y_n^H)^2}{\sum_{n=1}^{N-(k-1)\tau-H} (x_n - \bar{x})^2} \right]^{1/2}. \quad (8)$$

### 3.3. 基于线性/非线性 AR 模型归一化误差比值的非线性检验量

非线性检验量的设计应满足两个条件: 1) 对于线性序列来说, 替代数据与原序列的检验量没有显著差异; 2) 对于非线性序列来说, 两者的检验量存在

明显差异。

对于时间序列  $\{x_n, n = 1, 2, \dots, N\}$ , 可以运用如下非线性 AR 模型来对其进行建模:

$$\begin{aligned} \hat{x}_{n-1+H} &= a_0 + a_1 x_{n-1} + a_2 x_{n-2} + \dots + a_k x_{n-k} \\ &\quad + a_{k+1} x_{n-1}^2 + a_{k+2} x_{n-1} x_{n-2} + \dots + a_{M-1} x_{n-k}^d \\ &= \sum_{m=0}^{M-1} a_m z_m(n), \end{aligned} \quad (9)$$

式中,  $k$  是模型阶数;  $d$  表示系统的非线性程度; 函数基  $z_m(n)$  是由  $n$  时刻前的  $k$  个历史值  $(x_{n-1}, x_{n-2}, \dots, x_{n-k})$  的所有  $0-d$  次项构成, 包括常数项在内, 共有  $M = (k+d)! / (d! k!)$  项。系数  $a_m, m = 0, 1, 2, \dots, M-1$  可以由最小二乘法拟合得到。

通过 (9) 式可以计算出  $H$  步预测值  $\hat{x}_{n-1+H}, n = k+1, k+2, \dots, N+1-H$ 。非线性 AR 模型对于序列  $x_n$  的归一化误差为

$$\epsilon(H) = \left[ \frac{\sum_{n=k+1}^{N+1-H} (\hat{x}_{n-1+H} - x_{n-1+H})^2}{\sum_{n=k+1}^{N+1-H} (x_n - \bar{x})^2} \right]^{1/2}, \quad (10)$$

其中  $\bar{x}$  为时间序列  $x_n$  的均值。

对于时间序列  $x_n$ , 在合理选取模型阶数  $k$  后, 令  $d=1$ , 则非线性 AR 模型变为常见的线性 AR 模型, 其归一化误差表示为  $\epsilon_L(H)$ 。令  $d = d_{opt} (> 1)$ , 系统模型在线性 AR 模型的基础上增加了很多非线性项。这些非线性项并非都对时间序列的建模有用, 在未知哪些非线性项对建模有重要意义的情况下, 可以按升幂顺序逐渐增加非线性项, 从而得到非线性项逐渐增加的  $M-k-1$  个非线性模型, 这些非线性模型拟合时间序列  $x_n$  的归一化误差记为  $\epsilon_{NL}^j(H), j = 1, 2, \dots, M-k-1$ , 将其中的最小值记为  $\epsilon_{NL}(H)$ 。

检验时间序列  $x_n$  非线性的检验量定义为

$$\delta_{\text{NAR}}(H) = \log \epsilon_{\text{NL}}(H) - \log \epsilon_L(H), \quad (11)$$

式中  $H$  表示拟合模型的预测步长。如果时间序列  $x_n$  包含非线性成分, 选取合适的  $k$  和  $d$  后, 最优非线性模型的归一化误差  $\epsilon_{\text{NL}}(H)$  应比线性模型的归一化误差  $\epsilon_L(H)$  小很多, 相应的检验量  $\delta_{\text{NAR}}(H) < 0$ 。相反, 如果时间序列  $x_n$  不包含非线性成分, 则  $\epsilon_{\text{NL}}(H) \approx \epsilon_L(H)$ , 相应的检验量  $\delta_{\text{NAR}}(H)$  接近 0。结合替代数据检验法, 即可在一定置信度内据此判断时间序列  $x_n$  是否包含非线性成分。需要指出的是,  $\epsilon_L(H)$  的估计并没有在较大的阶数范围内搜索最佳的

线性模型, 因为对于采样率较高的连续类混沌时间序列, 足够阶数的线性 AR 模型和非线性 AR 模型在时间序列的拟合误差上相差并不大, 这容易将非线性时间序列误判为线性。因此在分别估计线性 AR 模型和非线性 AR 模型的拟合误差时, 两种模型的阶数  $k$  应相等。

## 4. 参数选取分析

基于线性/非线性 AR 模型归一化误差比值的非线性检验量本身受模型阶数  $k$ 、最高幂次  $d$ 、检验数据长度  $N$ 、预测步长  $H$  四个参数选取的影响。下面以 Lorenz 时间序列为待检验数据, 讨论这四个参数的优化选取。

### 4.1. Lorenz 时间序列及待检验数据集

Lorenz 方程如下:

$$\begin{aligned} \frac{dx(t)}{dt} &= \sigma(y(t) - x(t)), \\ \frac{dy(t)}{dt} &= rx(t) - y(t) - x(t)z(t), \\ \frac{dz(t)}{dt} &= x(t)y(t) - bz(t), \end{aligned} \quad (12)$$

式中  $\sigma = 10.0, r = 28.0, b = 8/3$ 。积分步长设为 0.08。采用龙格-库塔算法迭代产生足够长的 Lorenz 时间序列, 从  $x$  分量截取长 400000 的时间序列, 再分成 100 段, 每段长 4000, 构成数据集  $A$ 。采用第 2 节所述方法产生替代数据。

为准确检验本文设计非线性检验量对时间序列中弱非线性成分的检测能力, 仿真产生线性和弱非线性两种待检验数据集, 分别记为数据集  $B$  和数据集  $C$ , 每个数据集由 100 段长为 4000 的时间序列构成。线性数据集  $B$  由数据集  $A$  的 100 段 Lorenz 时间序列的相应替代数据构成, 显然这些替代数据是线性的。非线性数据集  $C$  由数据集  $A$  的 100 段 Lorenz 时间序列叠加放大 1.2 倍的相应线性替代数据构成, 由于每段时间序列中叠加了一定比例的线性成分, 因此数据集  $C$  中的每段时间序列仅包含弱的非线性成分, 便于检验设计非线性检验量对于弱非线性的检测能力。

根据文献 [6] 所述方法, 若置信度为  $\alpha$ , 则对于单边检验来说, 需产生  $1/(1-\alpha)-1$  个替代数据。按 95% 的置信度来检验时间序列的非线性, 则需产生 19 个相应的替代数据, 替代数据产生方法同上。如

果19个替代数据的非线性检验量都比待检验序列的非线性检验量大,则判待检验序列包含有非线性成分,否则判待检验序列是线性的.

#### 4.2. 参数优化分析

为全面考虑模型阶数  $k$ 、最高幂次  $d$ 、检验数据长度  $N$ 、预测步长  $H$  四个参数对设计非线性检验量正确判别时间序列线性/非线性特性的影响,每个参数都在一定范围内选取了几个值,然后采用时间序列替代数据检验法来判别数据集  $B, C$  中 100 段时

间序列的线性/非线性特性,并统计判为非线性的比例  $P_d$ .

表 1 列出了设计非线性检验量在各参数选取下数据集  $B, C$  判为非线性的比例.从线性数据集  $B$  的判别结果来看,在模型阶数  $k$ 、最高幂次  $d$ 、检验数据长度  $N$  这 3 个参数的各种取值组合下,判非线性的比例都随预测步长  $H$  的增大而递增.但根据替代数据检验法置信度预设 95% 的条件,对线性数据集  $B$  错判为非线性的比例应在 5% 左右,从表 1 来看,预测步长取 1 是比较合适的.

表 1 线性/非线性数据集(Lorenz)的判别

模型阶数	最高幂次	检验长度	数据类别	判非线性的比例 ( $H = 1-10$ )
2	3	3000	非线性	0.79 0.81 0.82 0.82 0.82 0.83 0.85 0.86 0.86 0.86
			线性	0.05 0.08 0.08 0.08 0.09 0.11 0.13 0.16 0.17 0.18
3	3	3000	非线性	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
			线性	0.02 0.06 0.06 0.08 0.11 0.11 0.13 0.13 0.15 0.15
4	3	3000	非线性	0.99 0.99 0.99 0.99 0.99 0.99 0.99 0.99 0.99 0.99
			线性	0.01 0.06 0.08 0.1 0.13 0.13 0.14 0.15 0.20 0.20
3	2	3000	非线性	0.05 0.05 0.06 0.07 0.09 0.12 0.15 0.16 0.16 0.16
			线性	0.05 0.06 0.08 0.12 0.13 0.14 0.15 0.16 0.19 0.19
4	2	3000	非线性	0.03 0.03 0.06 0.07 0.08 0.1 0.13 0.14 0.14 0.14
			线性	0.05 0.06 0.11 0.12 0.13 0.14 0.15 0.15 0.17 0.17
3	3	2000	非线性	0.95 0.95 0.95 0.95 0.95 0.97 0.97 0.99 0.99 0.99
			线性	0.07 0.09 0.1 0.11 0.11 0.14 0.15 0.15 0.18 0.21
3	3	4000	非线性	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
			线性	0.06 0.07 0.07 0.09 0.11 0.12 0.13 0.14 0.16 0.17

从非线性数据集  $C$  的判别结果来看,在模型阶数  $k$ 、最高幂次  $d$ 、检验数据长度  $N$  这 3 个参数的各种取值组合下,同数据集  $B$  类似,判非线性的比例也都随预测步长  $H$  的增大而递增或保持不变.结合数据集  $B$  和数据集  $C$  的非线性判别结果来看,为保证时间序列线性/非线性判别的准确性,最高幂次  $d$  应不小于 3,模型阶数  $k$  取为 3 或 4,检验数据长度  $N$  不小于 3000.对于 Lorenz 数据的非线性检验来说,在  $d = 3, L = 3000$  的情况下,  $k = 3$  是非线性检验的最佳模型阶数取值,叠加 1.2 倍线性替代数据的 100 个待检验非线性序列全部判别正确,这与假近邻法<sup>[1]</sup>估计的干净 Lorenz 时间序列的嵌入维是一致的.

#### 5. 与其他非线性检验量的比较

为了比较本文设计非线性检验量与基于最优线

性 AR 模型/指数 AR 模型的非线性检验量和非线性零阶预测误差在弱非线性检测方面的能力,仿真产生叠加线性替代数据比例  $R$  从 0.5 到 2.0 的 Lorenz 时间序列数据集  $D$ ,叠加比例每隔 0.1 设计一组,每组由 100 段长 4000 的时间序列构成.替代数据检验的置信度设为 90%,每判定一个时间序列的非线性,需产生 9 个替代数据.

计算非线性检验量  $\delta_{NAR}$  时,参数  $d = 3, k = 3, N = 3000, H = 3$ ,计算非线性检验量  $\delta_{AIC}$  时,参数  $k_{max} = 20, N = 3000, N = 3$ ,计算非线性零阶预测误差  $\delta_{ZP}$  时,  $k = 5, \tau = 3, N = 3000, H = 3, L = 0.005N$ .对于这三个非线性检验量来说,在保持检验长度不变的情况下,参数的其他取值没有表现出更好的非线性检测效果.3 种非线性检验量对于数据集  $D$  的非线性判别比例  $P_d$  如图 1 所示.图中带圆点的实线是非线性检验量  $\delta_{NAR}$  的检测结果,实线是非线性检验量

$\delta_{AIC}$  的检测结果,虚线是非线性零阶预测误差  $\delta_{ZP}$  的检测结果,以下同.从图 1 可以看出,在相同检验数据长度情况下,本文设计非线性检验量  $\delta_{NAR}$  对于弱非线性的检测能力比另外两个非线性检验量强很多,在叠加带内噪声比例达到 1.3 的情况下,仍能准确判断所有 100 段弱非线性时间序列,基于 AIC 准则的非线性检验量  $\delta_{AIC}$  和非线性零阶预测误差  $\delta_{ZP}$  的弱非线性检测能力比较接近.

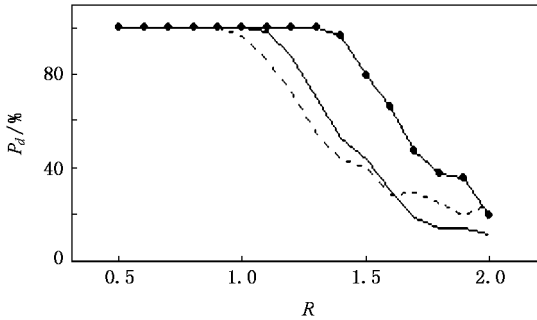


图 1 三种非线性检验量对不同叠加噪声比例 Lorenz 数据的非线性检测结果

## 6. 对其他 3 种混沌时间序列的非线性检验

为了进一步考察本文设计非线性检验量  $\delta_{NAR}$  对于不同弱非线性时间序列检验的适应性,仿真产生 Duffing 时间序列(积分步长为 0.08)、Henon 时间序列和 Logistic 时间序列,按类似方法产生替代数据集,对这 3 种非线性检验量的弱非线性检测能力进行比较.对 3 种混沌时间序列的弱非线性检测结果如图 2、3、4 所示.计算各种时间序列的 3 个非线性检验量时,相应参数都分别进行了优化选取,其他参数组合没有显示出更好的检测结果.图 1—4 所显示的 3 条曲线表明,对于 Lorenz 混沌时间序列和 2 种离散类混沌时间序列的非线性检验,本文设计非线性检验量  $\delta_{NAR}$  比另外 2 种非线性检验量要好很多,对于 Duffing 混沌时间序列的非线性检验,非线性检验量  $\delta_{NAR}$  的弱非线性检测能力在非线形零阶预测误差  $\delta_{ZP}$  和非线性检验量  $\delta_{AIC}$  之间.因此非线性检验量  $\delta_{NAR}$  在检测时间序列的弱非线性能力上具有较好的数据适应性.

本文所采用 3 种非线性检验量的参数优化都是在已知其非线性特性的条件下进行的,而对于待检

验时间序列,参数的选取就存在一定的盲目性,因此非线性检验量参数对于待检验时间序列的稳定性非常重要.对于这 4 种不同的时间序列弱非线性检验,估计非线性检验量  $\delta_{NAR}$  时的参数选取具有较好的稳定性,即  $d=3, k=3$  或  $k=4, N \geq 3000, \tau=1$ , 可以此参数组合来检验其他时间序列的非线性.

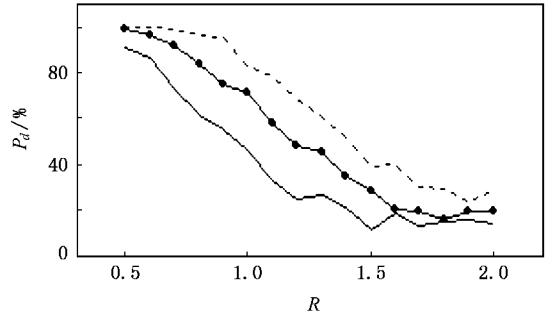


图 2 三种非线性检验量对不同叠加噪声比例 Duffing 数据的非线性检测结果

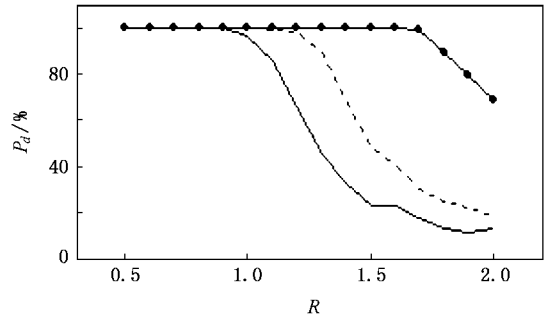


图 3 三种非线性检验量对不同叠加噪声比例 Henon 数据的非线性检测结果

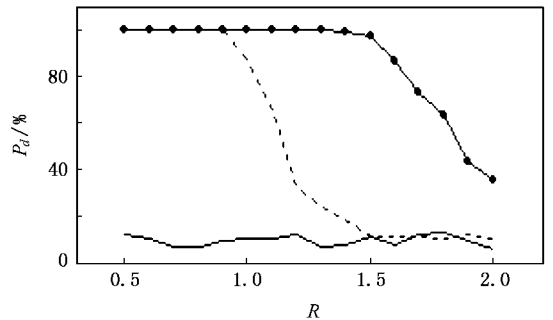


图 4 三种非线性检验量对不同叠加噪声比例 Logistic 数据的非线性检测结果

## 7. 结 论

本文提出了一种对时间序列弱非线性非常敏感

的非线性检验量  $\delta_{\text{NAR}}$ , 采用替代数据检验法检测时间序列的线性/非线性特性. 以 Lorenz 时间序列为例, 分析了非线性检验量  $\delta_{\text{NAR}}$  中模型阶数  $k$ 、最高幂次  $d$ 、检验数据长度  $N$  这 3 个参数的取值对非线性检测性能的影响. 所设计的非线性检验量  $\delta_{\text{NAR}}$  与基于最优线性 AR 模型/指数 AR 模型的非线性检验量  $\delta_{\text{AIC}}$  和非线性零阶预测误差  $\delta_{\text{ZP}}$  相比, 对于 Lorenz 混沌时间序列和 2 种离散类混沌时间序列的非线性检

验具有更高的弱非线性检测能力, 而对于 Duffing 混沌时间序列的弱非线性检测, 非线性检验量  $\delta_{\text{NAR}}$  的检测能力仅比非线性零阶预测误差  $\delta_{\text{ZP}}$  略差. 对这 4 种混沌时间序列弱非线性的检测结果表明非线性检验量  $\delta_{\text{NAR}}$  具有较强的数据适应性, 而且对于不同数据, 具有最佳非线性检测效果的参数比较稳定, 这为其他时间序列弱非线性检测的参数选择提供了参考.

- [ 1 ] Abarbanel H D I 1993 *Review of Modern Physics* **65** 1331
- [ 2 ] Wang F P, Guo J B, Wang Z J, Xiao D C 2001 *Acta Phys. Sin.* **50** 1019 ( in Chinese ) 汪芳平、郭静波、王贇基、萧达川 2001 物理学报 **50** 1019
- [ 3 ] Li X X, Feng J C 2007 *Acta Phys. Sin.* **56** 701 ( in Chinese ) 李雪霞、冯久超 2007 物理学报 **56** 701
- [ 4 ] Bian C H, Ning X B 2004 *Science Letters* **49** 130 ( in Chinese ) 卞春华、宁新宝 2004 科学通报 **49** 130
- [ 5 ] Mauricio B, Poon C S 1996 *Nature*( London ) **381** 215
- [ 6 ] Galka A, Ozaki T 2001 *Physica D* **158** 32
- [ 7 ] Antonis A, Kihwan J 2002 *Annals of Biomedical Engineering* **30** 192
- [ 8 ] Thomas S, Andreas S 2000 *Physica D* **142** 346
- [ 9 ] Thomas S, Andreas S 1997 *Physical Review E* **55** 5443
- [ 10 ] Taeun C, Tim S, Steven J S 1995 *Chaos* **5** 118
- [ 11 ] Cesare M, Ceruso M A 1999 *Physical Review E* 59 992
- [ 12 ] Lei M, Meng G, Feng Z J 2005 *Acta Phys. Sin.* **54** 1059 ( in Chinese ) 雷敏、孟光、冯正进 2005 物理学报 **54** 1059

## A test method for weak nonlinearity in time series<sup>\*</sup>

Jiang Ke-Yu<sup>†</sup> Cai Zhi-Ming Lu Zhen-Bo

( *Electrical Engineering College, Navy Engineering University, Wuhan 430033, China* )

( Received 16 April 2007 ; revised manuscript received 18 June 2007 )

### Abstract

Nonlinearity is necessary for time series to be treated as chaotic time series. A new test statistic for nonlinearity, which is based on the ratio of the multistep normalized prediction error with respect to linear AR models and nonlinear AR models, is used to detect the weak nonlinear components contained in time series by the surrogate data method. Taking example for Lorenz time series, the effect of related parameters for test statistic estimation is analyzed. By the nonlinearity tests for chaotic time series, the proposed test statistic  $\delta_{\text{NAR}}$  has better discrimination power for weak nonlinearity than the test statistic  $\delta_{\text{AIC}}$  based on AIC rules and the zeroth order nonlinear prediction error  $\delta_{\text{ZP}}$ , which shows that the proposed test statistic has strong adaptive abilities for time series. And, for different time series, the parameters with best nonlinearity discrimination performance are kept constant. The stabilization of parameters facilitates the nonlinearity test for other time series.

**Keywords** : surrogate data, chaotic time series, nonlinear AR model, nonlinearity test

**PACC** : 0545, 0540

<sup>\*</sup> Project supported by the State Key Development Program for Basic Research ( Grant No. 5132102ZZT32 ).

<sup>†</sup> Corresponding author. E-mail : jiangky828@163.com