

管理科学:面对复杂性 III — 经济时序动力系统最佳时序采样间隔分析研究^①

盛昭瀚, 马军海

(东南大学管理学院管理系, 南京 210096)

摘要:给出一种根据经济时序动力系统实测数据采样时间间隔 τ 的最佳范围选出方法, 并给出了相应的计算程序, 其对低维时序动力系统和高维时序动力系统同样有效, 算例表明本文建议的方法是有效的。

关键词:分形; 混沌经济时序; 最佳嵌入维数; 最佳采样间隔 τ 动力系统,

分类号: O175.14, O241.81 **文献标识码:** A **文章编号:** 1007-9807(2000)01-0100-06

0 引言

在现实生活中, 在工程界、自然科学界及社会科学界广泛地存在着各种各样的时序问题, 这些时序的形态往往呈现出复杂性和多样性。而现实中时序往往是通过实验或现场用等时采样的方法来得到:

$$x(t_1), x(t_2), x(t_3), \dots, x(t_i), \dots, x(t_N),$$

$$x(t_i) \in R \quad (\text{实数域})$$

其中 $t_i = t_0 + i\tau$ 。动力系统实测时间序列 $x(t_i)$ 是一维的, 为了根据它来构造非线性混沌时序经济动力系统模型和预测系统未来的动力学行为, 则需根据所测的一维时间序列来计算系统的结构参数, 为此必须解决的问题是: 研究采样间隔 τ 对分维数计算等的影响问题, 本文给出了一种根据经济时序动力系统实测数据的最佳嵌入维数选出的新方法, 在此基础上提出了采样时间间隔 τ 的最佳选出范围, 并给出了相应的计算程序, 其对低维时序动力系统和高维时序动力系统同样有效, 算

例表明, 本文建议的方法基本上可以克服以上方法的缺陷, 是有效的。

1 最佳采样间隔 τ 的选取

由于 Hausdorff 维数 D_H , 计盒维数 D_b , 信息维数 D_I , 关联维数 D_2 , 广义维数 D_q 计算出的分维数值基本上是一致的, 为了其计算时的便于实现和快速性, 目前, 广泛采用 G-P 算法计算关联维数, 而采样间隔 τ 对分维数的计算值有着非常重要的影响, 选择不好会使计算所得分维值的误差偏大, 甚至得到错误的分维计算值。

假设通过实验或现场所得到的等时采样的时序为

$$x_1, x_2, x_3, \dots, x_N \quad (1)$$

由于动力系统实测数据间不可避免地存在着数据之间的自相关性, 为了减少或基本根除由于数据间的自相关性所造成的影响, $C(r, N, w)$ 常采用式(2):

$$C(r, N, w, \tau) = \frac{2}{(N-w+1)(N-w)} \sum_{N=w}^N \sum_{i=1}^{N-n} H\{r - |V_{i+n} - V_i|\} \quad (2)$$

收稿日期: 1998-12-25.

基金项目: 国家自然科学基金资助项目(69874004).

作者简介: 盛昭瀚(1944-), 男(汉族), 江苏镇江人, 东南大学管理学院教授, 博士生导师。

H 为 Heaviside 函数,这里要求 W 应满足条件:

$$W \geq \tau \left(\frac{2}{N}\right)^{\frac{1}{m}} \quad (3)$$

τ 为时序之间的时间间隔, m 为嵌入维数,采用时序的关联维数为

$$D_2(m) = \frac{\sum_{i=1}^l \log_2(r_i) \log_2 C_m(r_i) - \sum_{i=1}^l \log_2(r_i) \sum_{i=1}^l \log_2 C_m(r_i)}{l \sum_{i=1}^l (\log_2(r_i))^2 - \left(\sum_{i=1}^l \log_2(r_i)\right)^2} \quad (5)$$

由式(2)~(5)计算分维数的公式可知时序之间的时间间隔 τ 对分维数的值有着非常重要的影响, τ 过大会使时序所描述的动力系统信号失真(会使数据之间失去原动力系统的本质固有联系),由此所算出的分维数根本是不可信的; τ 过小则由于动力系统实测数据间不可避免地存在着数据间的自相关性,这样算出来的分维数也会带有偏差。

为了正确地获取连续信号中各种频率成份的信息,在最高频率谱波的一个周期里至少采样两次.根据 Shannon 采样定理,可以得到数据之间采样的时间间隔 τ 的粗略的取值范围: $\tau \leq 1/(2f_{\max})$. 这里 f_{\max} 为连续信号中 $x(t)$ 出现的最高频率成份. Shannon 给出了 τ 取值的上界,但未完全解决最佳 τ 的取值问题.

$$d = \lim_{m \rightarrow \infty} \lim_{N \rightarrow \infty} \frac{d[\log_2 C(r, N, W)]/dr}{d[\log_2(r)]/dr} \quad (4)$$

根据式(2)及最小二乘估计可得到分维的近似估计公式:

基于 Theiler 的自相关函数法 Liebert and Schuster^[1], Kember and Fowler^[2] 又发展了此方法. Buzug and Pfister^[3] 建议可以通过分析动力系统的奇怪吸引子的全局或局部性态来确定最佳的时序采样间隔 τ .

Ying Cheng Lai, David Lerner^[4] 认为可同时使嵌入维数 M 和时序采样间隔 τ 变动,但总体上保证量 $M\tau$ 保持不变来确定最佳的时序采样间隔 τ .

可以先通过文献[5]求得最佳的嵌入维数,然后求得最佳的时序采样间隔 τ .

为了求得最佳的时序采样间隔 τ , Ying Cheng Lai, David Lerner^[4] 给出了 $\tau = 1, M = 2$ 的 $C(r, N, w, M, \tau)$ 展式

$$C(r, N, w, M, \tau) = C(r, N, w, 2, 1) = C\left(\frac{1}{2}r, 1, 1\right) + R(r) = r - \frac{1}{4}r^2 + \begin{cases} \frac{1}{2}r^2 & 0 < r < \frac{2}{3} \\ 3r - \frac{7}{4}r^2 & \frac{2}{3} \leq r < 1 \end{cases} \quad (6)$$

及 $\tau = 2, M = 2$ 的 $C(r, N, w, M, \tau)$ 展式:

$$C(r, N, w, M, \tau) = C(r, N, w, 2, 2) = \begin{cases} \frac{1}{2}r + \frac{17}{16}r^2 & 0 \leq r < \frac{2}{5} \\ -\frac{3}{4} + \frac{7}{2}r - \frac{7}{4}r^2 & \frac{2}{5} \leq r < \frac{2}{3} \\ \frac{3}{2}r - \frac{7}{16}r^2 & \frac{2}{3} \leq r < \frac{4}{5} \\ -1 + 4r - 2r^2 & \frac{4}{5} \leq r \leq 1 \end{cases} \quad (7)$$

当 $D_2 = 1$ 时,

$$\log_2 C(r, \tau, M) = \log_2 r - M\tau + \tau + 1 \quad (8)$$

对于更普遍的情况有

$$\log_2 C(r, \tau, M) \approx D_2 \log_2 r - M \tau K_2 \log_2 e \quad (9)$$

K_2 为 2 柄, 计算时 K_2 可取最大的 Lyapunov Exponent, 这里 T 通过式(10) 确定.

$$\begin{aligned} x_n &= x(t_0 + nT) = g(z_n) \\ &= g(z(t_0 + nT)) = g(\Phi_{nT}(z_0)) \end{aligned} \quad (10)$$

这时落入半径为 r 的 D_2 维球体内的点集个数 J 为

$$J \approx \frac{1}{2} N^2 r^{D_2} e^{-M \tau K_2} \quad (11)$$

由于 J 必须大于等于 1, 则最少的搜索半径 r 应满足式(12)

$$\log_2 r_{\min}(\tau, M) > (1 - 2 \log_2 N + \tau M T K_2 \log_2 e) / D_2 \quad (12)$$

由于 $J \leq N$, 则最大的搜索半径 r 应满足式(13)

$$\log_2 r_{\max}(\tau, M) \leq (1 - \log_2 N + \tau M T K_2 \log_2 e) / D_2 \quad (13)$$

结合 Ying Cheng Lai, David Lerner^[4] 所证得式(14)

$$\log_2 r_{\max}(\tau, M) \leq \tau M T K_2 \log_2 e / (D_2 - M) \quad (14)$$

可以进一步得到

$$\begin{aligned} \Delta &= \log_2 r_{\max}(\tau, M) - \log_2 r_{\min}(\tau, M) \\ &< \tau M T K_2 \log_2 e / (D_2 - M) \end{aligned}$$

$$- (1 - 2 \log_2 N + \tau M T K_2 \log_2 e) / D_2 \quad (15)$$

根据文献[4] 的研究结论 Δ 应满足

$$\Delta \geq 4 \quad (16)$$

于是可以进一步得到

$$\tau_{\max} < \frac{(2 \log_2 N - 1 - 4 D_2)(M - D_2)}{M^2 K_2 T \log_2 e} \quad (17)$$

对于 Logistic map $\lambda = 4$ 标准混沌的情况, $D_2 = 1.129, K_2 = \lambda_1 = 0.4461, T = 1$, 计算时取 $N = 3000$, 则可算得

$$\tau_{\max}(M) < 27.32422 \frac{(M - D_2)}{M^2} \quad (18)$$

对于 Henon map $\alpha = 1.4, \beta = 0.3$ 标准混沌的情况, $D_2 = 1.26, K_2 = \lambda_1 = 0.408, T = 1$, 计算时取 $N = 3000$, 则可算得

$$\tau_{\max}(M) < 26.51003431 \frac{(M - D_2)}{M^2} \quad (19)$$

2 计算结果

取如下三组数据分别对其理论进行研究:

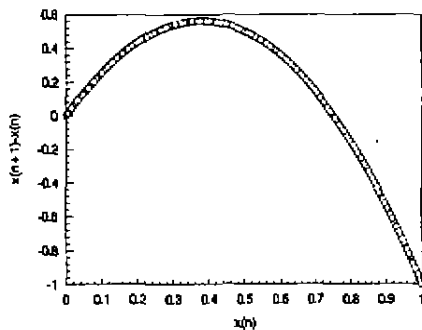


图1 第1组混沌数据的相图

(1) 对 Logistic map $x_{n+1} = \lambda x_n(1 - x_n)$ 取 $\lambda = 4$ 标准混沌的情况, 前 1000 点作为暂态点去掉, 把后 8000 点作为(1) 组原始数据点, 其相图如图 1, 其维数图(时序间隔 τ 对维数的影响图)如图 5. (计算维数时取 $N = 3000$)

$$(2) \text{ 取 Henon map } \begin{cases} x_{n+1} = 1 - \alpha x_n^2 + y_n \\ y_{n+1} = \beta x_n \end{cases}$$

$\alpha = 1.4, \beta = 0.3$ 标准混沌的情况 10000, 前 1000 点作为暂态点去掉, 把后 8000 点作为(2) 组原始数据点, 其相图如图 2, 其维数图(时序间隔 τ 对维数的影响图)如图 6. (计算维数时取 $N = 3000$)

(3) 用所得到的混沌经济数据 8000 点作为

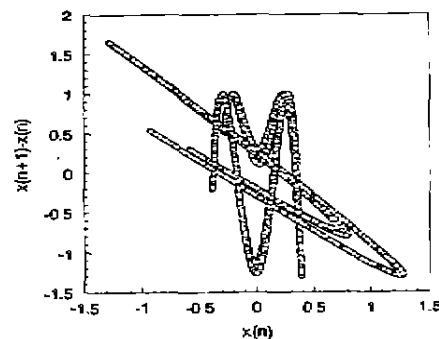


图2 第2组混沌数据的相图

实验数据, 其时间历程图如图 3, 其相图如图 4, 其维数图(时序间隔 τ 对维数的影响图)如图 7. (计算维数时取 $N = 3000$)

按照文中理论编制的计算维数的程序对(1) 组数据 Logistic 映射、(2) 组数据 Henon 映射、(3) 组实测混沌经济数据均 3000 点进行计算, 从理论上分析由式(7) 可以算得(1)、(2) 组数据的最佳嵌入维数均是 3; (3) 组数据的最佳嵌入维数均是 4.

按照式(18) 的计算结果取最佳嵌入维数等于 3 对 Logistic map 其最佳时序间隔 τ 应满足:

$$\tau_{\max}(M) < 5.68041736 \quad (20)$$

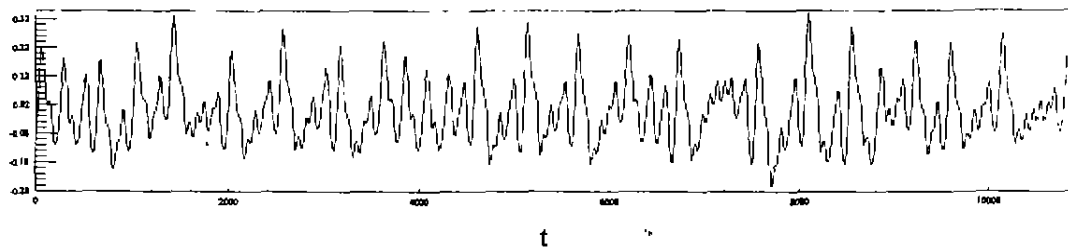


图3 第3组实测经济混沌数据的时间历程图

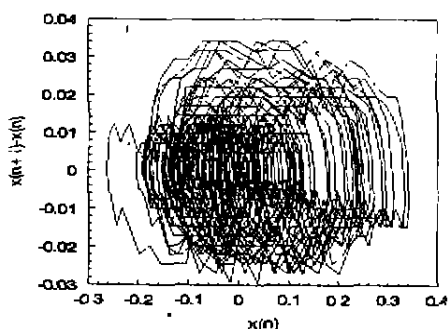


图4 第3组实测经济混沌数据的相图

图5对时间间隔 τ 取不同的值(嵌入维数 $M=3$)进行了研究,其中:

从上数第4条线为 $\tau=1, N=3000, W=12, M=3, R_{\min}=0.16, R_{\max}=1.6$,最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点30个,算得其分维数为 $D=1.08629$.

从上数第1条线为 $\tau=3, N=3000, W=12, M=3, R_{\min}=0.16, R_{\max}=1.6$,最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点30个,算得其分维数为 $D=1.13$.

从上数第2条线为 $\tau=5, N=3000, W=12, M=3, R_{\min}=0.15, R_{\max}=1.55$,最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点30个,算得其分维数为 $D=1.19$.

从上数第3条线为: $\tau=6, N=3000, W=12, M=3, R_{\min}=0.16, R_{\max}=1.6$,最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点30个,算得其分维数为 $D=2.137$.

从计算结果来看, τ 取1,3,5时(满足式(20) τ 的取值范围),算得的分维数 D 和真值相差较小,是比较准确的;而当 $\tau=6$ 时(不满足式(20) τ 的取值范围),算得的分维数 D 和真值相差较大是不

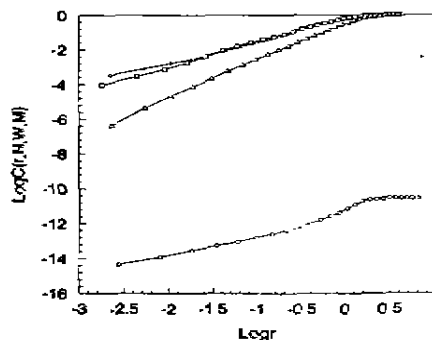


图5 第1组实测经济混沌数据的维数图

准确的,这说明本例的计算结果和理论分析是相吻合的.

按照式(19)的计算结果对Henon map其最佳时间间隔 τ 应满足

$$\tau_{\max}(M) < 5.1252733 \quad (21)$$

图6对时间间隔 τ 取不同的值(嵌入维数 $M=3$)进行了研究,其中:

从上数第1条线为 $\tau=1, N=3000, W=12, M=3, R_{\min}=0.35, R_{\max}=2.0$,最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点30个,算得其分维数为 $D=1.284$.

从上数第2条线为 $\tau=3, N=3000, W=12, M=3, R_{\min}=0.65, R_{\max}=2.1$,最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点30个,算得其分维数为 $D=1.2751$.

从上数第3条线为 $\tau=5, N=3000, W=12, M=3, R_{\min}=0.65, R_{\max}=2.1$,最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点30个,算得其分维数为 $D=1.391$.

从上数第4条线为 $\tau=6, N=3000, W=12, M=3, R_{\min}=0.65, R_{\max}=2.2$,最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点30个,算得其

分维数为 $D = 2.01$.

从计算结果来看, τ 取 1, 3, 5 时(满足式(21) τ 的取值范围), 算得的分维数 D 和真值相差较小, 是比较准确的; 而当 $\tau = 6$ 时(不满足式(21) τ 的取

值范围), 算得的分维数 D 和真值相差较大, 是不准确的. 这说明本例的计算结果和理论分析是相吻合的.

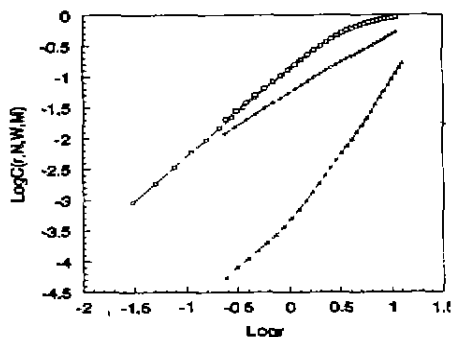


图6 第2组实测经济混沌数据的维数图

对第3组混沌经济数据其最佳时序间隔 τ 应满足:

$$\tau_{\max}(M) < 0.024 \text{ (单位: 秒)} \quad (22)$$

图7对时序间隔 τ 取不同的值(嵌入维数 $M = 4$) 进行了研究, 其中:

从上数(下端点开始)第4条线为 $\tau = 0.01$, $N = 3000$, $W = 12$, $M = 4$, $R_{\min} = 35$, $R_{\max} = 75$, 最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点 30 个, 算得其分维数为 $D = 2.26$.

从上数(下端点开始)第3条线为 $\tau = 0.02$, $N = 3000$, $W = 12$, $M = 4$, $R_{\min} = 25$, $R_{\max} = 80$, 最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点 30 个, 算得其分维数为 $D = 2.38$.

从上数(下端点开始)第1条线为 $\tau = 0.03$, $N = 3000$, $W = 12$, $M = 3$, $R_{\min} = 30$, $R_{\max} = 83$, 最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点 30 个, 算得其分维数为 $D = 3.56$.

从上数(下端点开始)第2条线为 $\tau =$

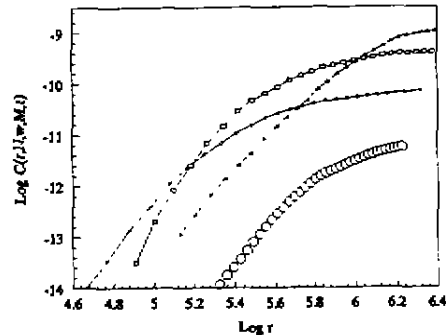


图7 第3组实测经济混沌数据的维数图

0.05, $N = 3000$, $W = 12$, $M = 3$, $R_{\min} = 25$, $R_{\max} = 80$, 最小搜索半径 R_{\min} 到最大搜索半径 R_{\max} 内取点 30 个, 算得其分维数为 $D = 3.74$. 从计算结果来看 τ 取 0.01, 0.02 时(满足式(22) τ 的取值范围), 算得的分维数 D 和真值相差较小, 是比较准确的; 而当 $\tau = 0.03, 0.05$ 时(不满足式(22) τ 的取值范围), 算得的分维数 D 和真值相差较大是不准确的. 这说明本例的计算结果和理论分析是相吻合的.

3 结论

- 1) 采样间隔 τ 对计算分维数的值有着重要的影响, 因此, 实际计算中只有选出的最佳范围内的采样间隔 τ , 才能够得到正确的分维数的计算值.
- 2) 从计算结果来看, 本文给出的最佳时序采样间隔 τ 在算法是好的、切实可行的、易于实现的.

参考文献:

- [1] Liebert W, Schuster H G. An improved estimator of dimension and some comments on providing confidence intervals [J]. Phys. Lett. A, 1989, 142: 107~112
- [2] Kemper G, Fowler A C. Some comments on the correlation dimension of noise [J]. Phys. Lett. 1993, 179: 72~78
- [3] Buzug T, Pfister G. Problems in estimation dynamics from data [J]. Phys. Rev. 1992, 45: 7073~7081
- [4] Ying Cheng Lai, Lerner D. Effective scaling regime for computing the correlation dimension from chaotic time series [J]. Phys. D, 1998, 115(5): 1~18
- [5] 马军海. 混沌时序动力系统非线性重构[D]. 天津大学博士学位论文, 1997

- [6] 盛昭瀚,马军海.管理科学:面对复杂性I—混沌时序经济动力系统重构技术[J].管理科学学报,1998,1(1):31~42
- [7] Takens F, Manc.etal. In dynamical systems of turbulence[J], V Lecture Notes in Mathematics,edited by D. A. rand and L. S. Berlin; Young Springer. 1981.898;366
- [8] 马军海,陈予恕,刘曾荣.动力系统实测数据的非线性混沌特性的判定[J].应用数学和力学,1998,19(6):481~488
- [9] Broomhead D S,King G P. Extracting qualitative dynamics from experimental data[J]. Phys D. 1986;20(4):217~236
- [10] Mees A I,Rapp P E. Jennings L S. Singular-value decomposition and embedding dimension. Phys. RevA,1987,36(1):340~346
- [11] Theiler J. Spurious dimension from correlation algorithms applied to limited time series data[J]. Phys. Rev A34 1986,34:2427~2432
- [12] Luis Cabrera J,Javier F. Numerical analysis of transient behavior in the discrete random logistic equation with delay [J]. Phys. Lett. 1995,197(1):19~24
- [13] Ma Junhai,Chen Yushu,Liu Zengrong. Threshold values for diagnosis of chaotic nature of the data obtained in non-linear dynamic analysis[J]. Applied Mathematics and Mechanics,1998,19(6):513~520
- [14] 马军海,陈予恕,刘曾荣.高斯分布的相位随机化对实测数据临界值影响的分析研究[J].应用数学和力学,1998,19(11):955~963

Management science to face with complicated problems Ⅲ — study on some problems about fractal and chaotic of the economic dynamical time series

SHENG Zhao-han, MA Jun-hai

Institute of Systems Engineering, Southeast University, Nanjing 210096

Abstract: The paper presents the quality is constrained by the delay time τ . We derive an upper bound for this maximum delay time. Numerical calculating results on several model chaotic time series support the theoretical argument. According to the theory the concrete paradigm, as well as the selected method of calculating length of the economic time series were proposed. The proposed method in this paper was proved by calculating examples to be effective.

Keywords: nonlinear; delay time τ ; chaotic economic time series; fractal