

混沌时间序列的时滞判定^{*}

田玉楚

(浙江大学工业控制技术研究所工业控制技术国家重点实验室, 杭州 310027)

(1996 年 4 月 4 日收到; 1996 年 5 月 21 日收到修改稿)

运用重构混沌吸引子的嵌入技术, 开发了提取混沌时间序列信息的宏观信息熵方法, 用以判定时滞的存在性和时滞的大小. 仿真实验说明了本文方法的有效性.

PACC: 0545; 0250

1 时滞混沌系统

许多实际系统需要用时滞方程来描述. 时滞方程通常具有丰富的动态特性. 比如, 在无时滞的连续时间自治系统中, 混沌行为必须用至少三阶的微分方程来描述. 但若存在时滞, 一阶微分方程也可能产生混沌. 考虑如下时滞系统^[1]:

$$\epsilon \dot{x} = -x + f(x_{t-\tau}). \quad (1)$$

它含一指数恢复项 $-x$ 和一时滞非线性外力驱动项 $f(x_{t-\tau})$, ϵ 为一很小的无量纲参数, 它表征恢复时间与滞后时间 τ 的比值. 该系统的动态行为取决于非线性函数 f 的形式和时滞 τ 的大小. 因 ϵ 很小, 故该系统与如下的离散映射密切相关:

$$x_t = f(x_{t-\tau}). \quad (2)$$

这种映射形式在混沌理论中已得到广泛研究. 可见, 在许多情况下, (1) 式将表现为混沌态. 现考虑 f 为 Mackey-Glass 方程^[2]

$$f(x) = \lambda x / (1 + x^c), \quad (3)$$

取 $\lambda = 2$, $c = 10$ 和 $\tau = 1$, 可得 (2) 和 (3) 式的映射图 (如图 1). 容易证明, 若 $x_0 > 0$, 则对 $\forall t > 0$, 有 $x_t > 0$; 相反, 若 $x_0 < 0$, 则对 $\forall t > 0$, 有 $x_t < 0$; 若 $x_0 = 0$, 则对 $\forall t > 0$, 恒有 $x_t \equiv 0$. 故 x_t 的运动区域取决于初值 x_0 的符号. 根据问题的实际意义, 取 $x_0 > 0$, 则系统的运动限于 $[0, \infty)$, 系统映射图为图 1 右上区域的单峰映射. 借助于几何方法^[3], 很容易分析这类系统的动态行为.

由 $x^* = f(x^*)$ 求得 (2) 和 (3) 式的不动点 $x_1^* = 0$, $x_2^* = 1$. 因 $|f'(x_1^*)| = 2 > 1$, $|f'(x_2^*)| = |-8| > 1$, 说明不动点 x_1^* 和 x_2^* 均不稳定. 另外, 对 $\forall x > x_2^*$, 均有 $f(x) < x_2^*$, 说明系统并不发散, 因而必然呈现出比较复杂的动态特性. 由图 2 的一段时间曲线已可看到混沌运动的表现特征.

在 (1) 式中, 利用 (3) 式的 Mackey-Glass 非线性函数, 并取 $\tau = 1$, 结果导致混沌吸引

^{*} 中国博士后科学基金资助的课题.

子. 对于所选参数 $\lambda=2, c=10$ 和 $\epsilon<1$, Farmer 发现^[4], 这时的混沌吸引子具有信息维 $D \approx 1/\epsilon$. 图 3 和图 4 分别示出 $\epsilon=0.05$ 时该混沌吸引子的相图和时间曲线.

现设(1)和(3)式中的非线性函数未知, 希望根据图 4 的采样时间序列判定系统中的时滞. 这对系统建模、分析、估计、预测和控制等都有很重要的意义. Fowler 和 Kember 曾对此作过讨论^[1]. 本文运用 Takens 嵌入技术^[5], 从信息理论观点出发, 用信息熵方法提取混沌时间序列的时滞信息, 从而达到判定时滞的目的.

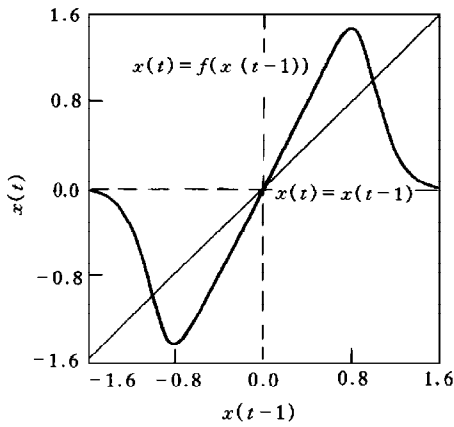


图 1 迭代映射(3)式的映射图

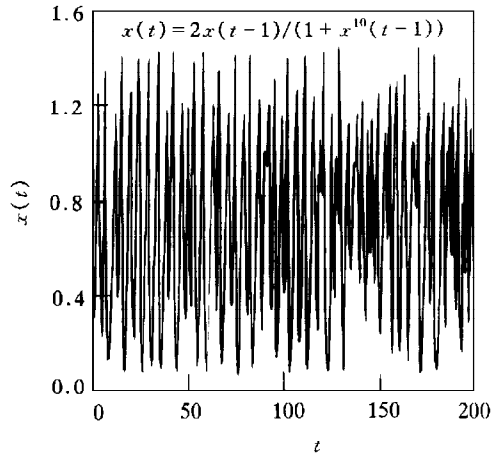


图 2 迭代映射(3)式的时间曲线

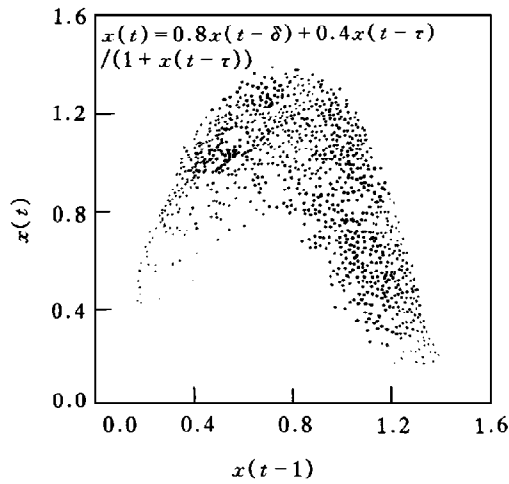


图 3 混沌系统(1)和(3)式的相图
 $\epsilon=0.05, \delta=\epsilon/4, \tau=1$

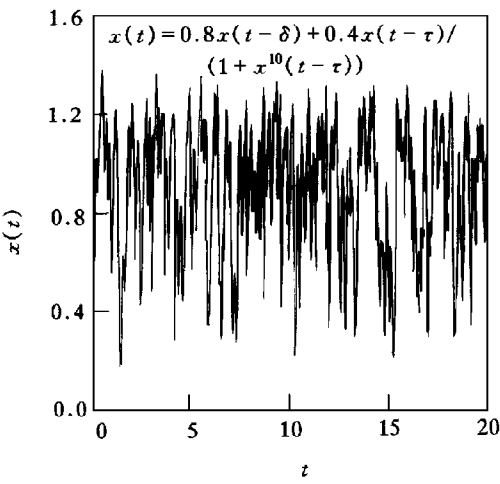


图 4 混沌系统(1)和(3)式的时间曲线
参数同图 3

2 Shannon 信息熵

Shannon 信息熵与概率分布联系在一起. 考虑具有离散取值空间的随机变量 $X \in \{x_1, x_2, \dots, x_n\}$, $Y \in \{y_1, y_2, \dots, y_m\}$ 和 $Z \in \{z_1, z_2, \dots, z_d\}$, 其概率分布分别为 $P(X) = [P(x_i)]_{1 \times n}$, $P(Y) = [P(y_j)]_{1 \times m}$ 和 $P(Z) = [P(z_k)]_{1 \times d}$, 则 Shannon 熵定义为 (如 X)

$$H(X) = - \sum_{i=1}^n P(x_i) \log P(x_i). \quad (4)$$

它描述了 X 的总体不定性, 亦即为消除 X 的不定性所需获取的信息. 类似地, 由联合分布 $P(Y, Z) = [P(y_j, z_k)]_{m \times d}$ 和 $P(X, Y, Z) = [P(x_i, y_j, z_k)]_{n \times m \times d}$ 也可定义具有类似意义的联合熵

$$H(Y, Z) = - \sum_{j=1}^m \sum_{k=1}^d P(y_j, z_k) \log P(y_j, z_k), \quad (5)$$

$$H(X, Y, Z) = - \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^d P(x_i, y_j, z_k) \log P(x_i, y_j, z_k). \quad (6)$$

另外, 由条件分布 $P(X|Y, Z) = [P(x_i|y_j, z_k)]$, 还可定义已知 Y 和 Z 后 X 的 Shannon 条件熵

$$H(X|Y, Z) = - \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^d P(x_i, y_j, z_k) \log P(x_i|y_j, z_k). \quad (7)$$

它表明已知 Y 和 Z 之后随机变量 X 的总体不定性. 由(5)–(7)式可直接导出

$$H(X|Y, Z) = H(X, Y, Z) - H(Y, Z). \quad (8)$$

从本质上说, 混沌系统是确定性的而不是随机的, 但又与随机系统有类似的宏观行为. 由此可在不变测度意义下, 将混沌系统视为一类特殊的随机系统加以处理. 作者曾研究过这种等价变换及等价概率分布的求取方法^[6]. 其基本方法是先把连续的状态空间量化为 Q 级, 再利用遍历特性以时间平均近似集平均而得到有关的概率分布, 然后就可利用(4)–(8)式定量计算特定空间划分下的 Shannon 信息熵, 并进一步分析系统的信息特性.

3 嵌入技术与信息熵判据

为了从图 4 的混沌时间序列重构系统动力学, 需运用 Takens 的嵌入技术^[5]. 为此, 先利用后向差分 $\dot{x} \approx (x_t - x_{t-\delta})/\delta$, 把时滞微分方程(1)化为差分方程

$$x_t = \frac{\epsilon}{\epsilon + \delta} x_{t-\delta} + \frac{\delta}{\epsilon + \delta} f(x_{t-\tau}), \quad (9)$$

其中 δ 为离散时间间隔, 要求 $\delta < \epsilon$. 由此可见, $x_{t-\delta}$ 和 $x_{t-\tau}$ 可近似确定 x_t . 故若把图 4 的时间序列嵌入到 R^3 空间 $(x_t, x_{t-\delta}, x_{t-\tau})$, 则运动轨迹必靠近一个曲面. 另一方面, 对 $\forall \Delta \neq \tau$, 且 $\Delta > \delta$, 由 $x_{t-\delta}$ 和 $x_{t-\Delta}$ 不能确定 x_t 的运动状态. 因此, 若把图 4 的时间序列嵌

入到 R^3 空间 $(x_t, x_{t-\delta}, x_{t-\Delta})$, 只要 $\Delta \neq \tau$, 则运动轨迹必然充满该 R^3 空间中的一个立方体. 据此可定性判别系统中的时滞.

现在从信息理论的观点出发, 将上述思想定量化. 考虑已知 $x_{t-\delta}$ 和 $x_{t-\Delta}$ 后状态 x_t 的不定性, 这可由 Shannon 条件熵 $H(x_t | x_{t-\delta}, x_{t-\Delta})$ 来描述. 根据(8)式, 它是三维总体的联合熵 $H(x_t, x_{t-\delta}, x_{t-\Delta})$ 与二维总体的联合熵 $H(x_{t-\delta}, x_{t-\Delta})$ 之差, 即

$$H(x_t | x_{t-\delta}, x_{t-\Delta}) = H(x_t, x_{t-\delta}, x_{t-\Delta}) - H(x_{t-\delta}, x_{t-\Delta}), \quad (10)$$

表明通过测量 $x_{t-\delta}$ 和 $x_{t-\Delta}$ 之后, 仍未消除的 $x_t, x_{t-\delta}$ 和 $x_{t-\Delta}$ 的总体不定性.

因为由 $x_{t-\delta}$ 和 $x_{t-\tau}$ 可近似确定 x_t , 亦即通过测量 $x_{t-\delta}$ 和 $x_{t-\tau}$ 可基本消除 x_t 的不定性, 所以条件熵 $H(x_t | x_{t-\delta}, x_{t-\tau})$ 必为一很小的数值. 而对于 $\forall \Delta \neq \tau$, 且 $\Delta > \delta$, 因为根据 $x_{t-\delta}$ 和 $x_{t-\Delta}$ 不能确定 x_t 的状态, 测量 $x_{t-\delta}$ 和 $x_{t-\Delta}$ 之后, x_t 仍具有很大的不定性, 故这时的条件熵 $H(x_t | x_{t-\delta}, x_{t-\Delta})$ 也一定比 $H(x_t | x_{t-\delta}, x_{t-\tau})$ 大得多. 从这个意义而言, 由小到大不断改变 Δ 的数值, 用 Shannon 条件熵 $H(x_t | x_{t-\delta}, x_{t-\Delta})$ 的大幅度减小, 作为判定时滞的准则是非常合适的.

4 仿真实验

考虑(1)和(3)式. 取 $\epsilon = 0.05$, $\tau = 1$, $\lambda = 2$, $c = 10$, 各种状态初值均为 0.6. 数值求解(1)和(3)式, 略去前 1000 步的过渡过程, 得到图 4 所示的时间序列. 设 $\delta = \epsilon/4$, 数据样本为 3000 个, 并将包含状态取值范围的区间 $[0.11, 1.42]$ 量化为 $Q = 16$ 级, 由小到大逐渐改变 Δ 的数值, 计算相应的 Shannon 条件熵 $H(x_t | x_{t-\delta}, x_{t-\Delta})$, 得到图 5 所示的曲线. 可见, 在 $\Delta = 1$ 处, $H(x_t | x_{t-\delta}, x_{t-\Delta})$ 急剧减小, 从而可判定系统存在时滞, 且 $\tau = 1$. 这与实际情况完全相符.

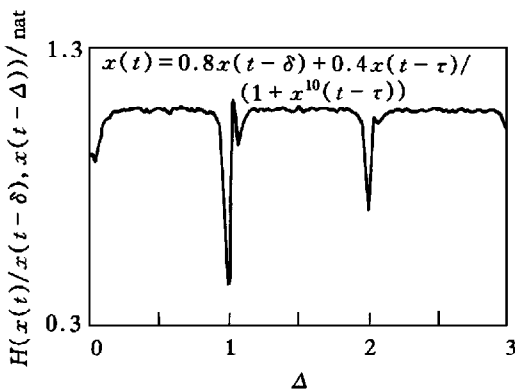


图 5 $H(x_t | x_{t-\delta}, x_{t-\Delta})$ 曲线 ($Q = 16$) 参数同图 3

5 关于空间量化的讨论

为了讨论空间量化对 Shannon 信息熵数值的影响, 下面先证明熵的聚类 (clustering) 性质. 以 $H(P_1, P_2, \dots, P_n)$ 记概率总体 $\{P_1, P_2, \dots, P_n\}$ 的熵, 其中 $P_i \in [0, 1]$ ($i = 1, 2, \dots, n$), $\sum_{i=1}^n P_i \leq 1$. 根据对数函数的性质, 对 $\forall P_i, P_j \in \{P_1, P_2, \dots, P_n\}$,

$$-P_i \log P_i - P_j \log P_j \geq -(P_i + P_j) \log (P_i + P_j), \quad (11)$$

$$\text{所以 } H(P_1, P_2, \dots, P_n) \geq H(P_1, \dots, P_i + P_j, \dots, P_n) \geq \dots \geq H\left(\sum_{i=1}^n P_i\right). \quad (12)$$

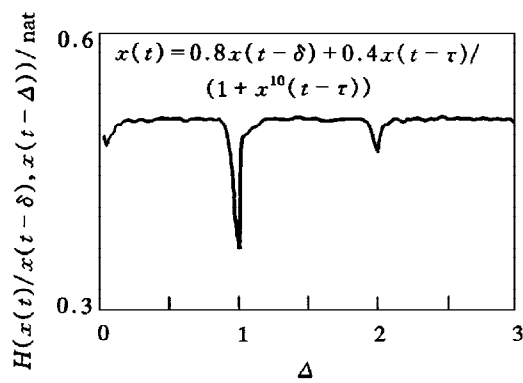


图 6 $H(x_t | x_{t-\delta}, x_{t-\Delta})$ 曲线 ($Q=4$) 参数同图 3

空间量化级数越多(即 Q 越大), 数值计算精度就越高, 由上述熵的聚类性质可知计算的熵值也越大. 但增大 Q 必然加重计算和存贮负担(本文方法的计算量和存储空间均在 Q^3 数量级). 幸好, 本文判别时滞的 Shannon 条件熵准则并不注重熵的绝对值大小, 而是寻找熵的突变点. 因此, 量化级数的多少对问题求解并无实质影响. 所以我们往往选取较少的量化级数, 通常取 $Q=5$ 左右即可. 例如取 $Q=4$ 重做上节的仿真实验, 得到图 6 所示的熵曲线, 同样在 $\Delta=1$ 处观察

到条件熵的大幅度减小, 与上节取 $Q=16$ 时所得结论完全一样.

6 结 语

运用嵌入技术和宏观信息熵方法, 能从低维时间序列中提取高维动力系统的有关信息, 由此改进系统建模、分析、估计、预测和控制等. 本文运用吸引子重构技术, 开发了一种判定混沌时间序列时滞的信息熵方法, 讨论了空间离散级数对判定方法的影响, 给出了具体的仿真实例. 本文的方法可在许多情况下进行扩展, 也适用于有测量噪声的场合, 有关这些更进一步的问题拟另文讨论.

- [1] A. C. Fowler and G. Kember, *Phys. Lett.*, **A175**(1993), 402.
- [2] M. C. Mackey and L. Glass, *Science*, **197**(1977), 287.
- [3] R. M. May, *Nature*, **261**(1976), 459.
- [4] J. D. Farmer, *Physica*, **D4**(1982), 366.
- [5] F. Takens, *Lecture Notes in Math.*, **898**(1981), 366.
- [6] 田玉楚、符雪桐、吕勇哉, *控制与决策*, **8**(1993), 345.

DELAY IDENTIFICATION IN CHAOTIC TIME SERIES

TIAN YU-CHU

(*State Key Laboratory of Industrial Control Technology,
Institute of Industrial Process Control, Zhejiang University, Hangzhou 310027*)

(Received 4 April 1996; revised manuscript received 21 May 1996)

ABSTRACT

By using the embedding techniques for reconstructing chaotic attractors, this paper proposes a macroscopic information entropy approach for identifying the delay in a chaotic time series sampled from a continuous-time system. The effectiveness of the proposed approach is illustrated through numerical experiments.

PACC: 0545; 0250