归一化 KLD 系数及多维序列相关和同步的检测*

谢小平 陈宏平 曹志彤 何国光†

(浙江大学物理系, 杭州 310027)

(2011 年 10 月 17 日收到; 2011 年 11 月 28 日收到修改稿)

提出了 KLD 系数和归一化 KLD 系数来刻画多维序列的相关结构, 以解决 KLD 维密度固有的局限性. 利用完全相关和完全不相关的多维序列, 导出 KLD 维密度的上界和下界函数, 进而导出 KLD 系数的上界和下界, 在此基础上提出归一化 KLD 系数. 解析分析和数值仿真都证明, 多维序列相关结构的变化会引起归一化 KLD 系数线性的变化. 数值仿真还证明, 即使多维序列中仅有其中的两个时间序列的相关结构发生改变, 归一化 KLD 系数仍能灵敏地检测到. 不仅如此, 归一化 KLD 系数还可用于非平稳时间序列的分析. 耦合映象格子的数值仿真结果表明, 归一化 KLD 系数还能够分析非线性系统的相关结构.

关键词: 多维序列分析, 相关结构, 检测, KLD

PACS: 05.45.−a, 05.45.Tp

1 引 言

自然界的现象, 无论简单和复杂, 都是在一定的时间和空间中出现的. 一个时空动力系统, 其各组成部分总是不间断地参与系统各种独立的瞬态过程, 使得系统的动力学展现出非常复杂的时空相关模式. 这种复杂的时空相关模式常常表现为系统连接程度的不断变化, 或者多维数据集内相应的时间序列之间同步程度的连续不断变化 [1]. 目前, 同步现象已成为时空动力学研究的一个热点 [1−5]. 为了检测和量化这种时空相关模式和同步的变化, 人们提出了许多方法, 如傅里叶分析 [6]、主成分分析 (PCA) [7], Karhunen-Loève 分析 (KLD) [8] 和小波分析 [9] 等线性方法, 也有诸如维密度 [10,11]、Lyapunov 指数谱分析 [12−14]、非线性主成分分析 (NLPCA) [15]、局部线性嵌入分析 (LLE)[16] 和独立成分分析 (ICA)[17] 等非线性方法.

近年来, KLD 维度和 KLD 维密度 [8,18−20] 等被用来刻画时空混沌系统, 然而, 由于两者都依赖于某个可以任意取值的分数, 因此严格地说, 两者都只是系统描述相对的而不是绝对的测度 [19].

本文在分析 KLD 维密度局限性的基础上, 提出了 KLD 系数; 利用完全不相关和完全相关的多维序列, 导出了 KLD 系数上界和下界的解析结果, 由此提出归一化的 KLD 系数 Ns_{KLD}. 然后利用 Ns_{KLD} 检测和刻画多维序列的相关结构和同步特性.

2 归一化 KLD 系数

众所周知, 由于 KLD 刻画的是协方差矩阵, 从本质上来说 KLD 是一种纯粹的线性方法. 通过引入 KLD 维度, 它的应用范围拓展到非线性的时空混沌系统 [8,18]. 目前, 这种方法已经广泛地应用于量化高维非线性系统的复杂度 [20].

2.1 KLD 维密度

用 $m \times n$ 维矩阵 $\boldsymbol{A}$ 表示一个由 n 个变量和 m 个时间点所组成的零均值多维时间序列. 不失一般性, 假设 $m \geqslant n$, 则 $N = \min\{m, n\} = n$. 其相关矩阵 $\boldsymbol{C} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的 N 个特征值可以表示为 $\sigma_1^2 \geqslant \sigma_2^2 \geqslant \cdots \geqslant \sigma_N^2 \geqslant 0$.

* 国家自然科学基金 (批准号：61175060) 和浙江省教育厅项目 (批准号: Y200909251) 资助的课题.

† E-mail: gghe@zju.edu.cn

© 2012 中国物理学会 Chinese Physical Society

矩阵 $\boldsymbol{A}$ 的 KLD 维度 D_{KLD} 定义为:

$$D_{\mathrm{KLD}} = \max\left\{p : \sum_{i=1}^{p}\sigma_i^2 \Big/ \sum_{i=1}^{N}\sigma_i^2 \leqslant f\right\}. \quad (1)$$

它反映了用一个给定的精确度 f $(0 < f < 1)$ 去近似描述一个时空数据时所需要的 KLD 特征模式的数目.

时空混沌系统的数值仿真表明, 对任意的分数 f, D_{KLD} 和系统的大小呈现出线性的尺度关系 [8], 这个结果导致了归一化的 KLD 维度 —KLD 维密度 [18] 的引入: $\delta_{\mathrm{KLD}} = D_{\mathrm{KLD}}/N$.

图 1 为高斯分布随机多维序列的 KLD 维密度与分数 f 之间的关系. 从图 1 中可以看出, 虽然 KLD 维密度的数值介于 0 和 1 之间, 但 KLD 维密度是分数 f 的单调递增阶跃函数. 如果 f 增加, 则 KLD 维密度也随之增加, 因此 f 取不同的值, 其对应的 KLD 维密度也会不同. 其次对于任意一个小于 1 的分数 f, KLD 维密度都不能达到其最大值 1.

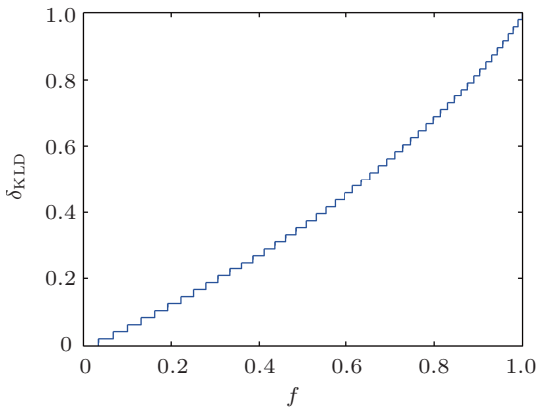

图 1 高斯分布随机多维序列的 KLD 维密度与 f 的关系

尽管随机矩阵 (RTM) 理论认为, 用随机矩阵的特征谱描述的应该是由随机相关或噪声控制的部分, 只有那些明显偏离 RTM 预测的特征值才包含关于系统真实相关结构的信息. 一般较大的特征值会明显偏离解析的 RTM 结果, 于是随机矩阵理论认为除了较大的特征值外, 特征值谱的其余部分并不反映任何相关结构信息 [21]. 然而, 协方差矩阵的特征值和特征矢量随时间演化的研究表明, 一般情况下, 特征值谱较小特征值部分并不完全由噪声和随机相关所控制, 相反它们包含关于系统动力学的本质信息. 有时还存在着这样的情形, 从统计意义上来说, 特征值谱较小特征值部分包含比最大特征值更多的系统相关结构信息 [1]. 对于任意一个小于 1 的分数 f, KLD 维密度反映了某个方差比例的特征模式, 而只有比较大的特征值才能引起 KLD 维密度的变化, 这造成多维序列相关结构较大的变化才能影响 KLD 维密度, 这样必然会影响某些情形下相关信息的获得 [19].

2.2 归一化 KLD 系数

为了解决这些问题, 我们利用 KLD 维密度, 提出 KLD 系数. KLD 系数 s_{KLD} 定义如下:

$$s_{\mathrm{KLD}} = \int_0^1 \delta_{\mathrm{KLD}}(f)\mathrm{d}f. \quad (2)$$

很显然, KLD 系数与分数 f 无关.

众所周知, 矩阵 $\boldsymbol{C}$ 的特征值分布与多维序列中各通道时间序列之间的相关程度直接相关 [1]. 如果多维序列中任意两个时间序列之间都不相关或统计正交, 则矩阵 $\boldsymbol{C}$ 的非对角线元素都为 0, 而所有对角线元素都等于 σ^2(假设多维序列中每一个时间序列都有相同的方差 σ^2) . 此时, 矩阵 $\boldsymbol{C}$ 的所有特征值都相同, 即 $\lambda_i = \sigma^2, \forall\, i$.

在这种情况下, 对任意的一个数 $0 \leqslant f < 1$, 存在一个整数 i $(1 \leqslant i < N)$, 使得

$$\frac{(i-1)}{N} \leqslant f < \frac{i}{N}. \quad (3)$$

于是, 我们可以求得这种情况下的 δ_{KLD}

$$\delta_{\mathrm{KLD}}(f) = \frac{i-1}{N}. \quad (4)$$

当 $f = 1$ 时, $\delta_{\mathrm{KLD}}(1) = 1$.

从而, 在完全不相关的条件下, 其 KLD 系数 s_{KLDup} 可以表示如下:

$$\begin{aligned} s_{\mathrm{KLDup}} &= \int_0^1 \delta_{\mathrm{KLDup}}(f)\mathrm{d}f = \sum_{i=1}^{N}\int_{\frac{i-1}{N}}^{\frac{i}{N}} \frac{i-1}{N}\mathrm{d}f \\ &= \sum_{i=1}^{N}\frac{i-1}{N^2} = \frac{N-1}{2N}. \end{aligned} \quad (5)$$

与前面相反, 如果多维序列中任意两个时间序列之间完全相关, 则矩阵 $\boldsymbol{C}$ 的所有元素都相同. 此时, 矩阵 $\boldsymbol{C}$ 的特征值中仅有一个不为 0, 即 $\lambda_{\max} = N\sigma^2$, 其它特征值均为 0. 在这种完全相关的情况下, 对任意的一个数 $0 \leqslant f < 1$, 我们可以求得 δ_{KLD} 为:

$$\delta_{\mathrm{KLD}}(f) = \begin{cases} 0, & 0 \leqslant f < 1, \\ 1/N, & f = 1, \end{cases} \quad (6)$$

于是, 在完全相关的条件下, 其 KLD 系数 s_{KLDlow} 可以表示如下:

$$s_{\mathrm{KLDlow}} = 0. \tag{7}$$

很显然, 由于实际的多维序列既不是完全相关, 也不是完全不相关的, 故它们的特征值谱既不是只有一个不为 0, 也不是所有特征值都相等. 很容易证明, s_{KLDup} 和 s_{KLDlow} 分别为 KLD 系数 s_{KLD} 的上界和下界.

为了更合理地表示多维序列的相关结构, 我们提出归一化的 KLD 系数 Ns_{KLD}.

$$Ns_{\mathrm{KLD}} = \frac{s_{\mathrm{KLD}} - s_{\mathrm{KLDlow}}}{s_{\mathrm{KLDup}} - s_{\mathrm{KLDlow}}}. \tag{8}$$

事实上, 由于 $s_{\mathrm{KLDlow}} = 0$, 上述定义又可以表示为:

$$Ns_{\mathrm{KLD}} = \frac{s_{\mathrm{KLD}}}{s_{\mathrm{KLDup}}} = \frac{2N}{N-1} s_{\mathrm{KLD}}. \tag{9}$$

为了研究多维序列中各时间序列之间的相关程度对 Ns_{KLD} 的影响, 我们考虑相关矩阵 $\boldsymbol{C}$ 为式 (10) 所示形式的 N 阶矩阵, 即:

$$\boldsymbol{C} = \begin{pmatrix} 1 & b & \cdots & b \\ b & 1 & \cdots & b \\ \cdots & \cdots & \cdots & \cdots \\ b & b & \cdots & 1 \end{pmatrix}, \tag{10}$$

其中 $0 < b < 1$.

显然, 对于式 (10) 所示的对称矩阵, 其 N 个特征值可表示为 [22]:

$$\begin{aligned} &\sigma_1^2 = (1-b) + bN = 1 + (N-1)b, \\ &\sigma_2^2 = \sigma_3^2 = \cdots = \sigma_N^2 = 1 - b. \end{aligned} \tag{11}$$

于是, 对应于式 (10) 所示相关矩阵的多维序列, 其 KLD 维密度为:

$$\delta_{\mathrm{KLD}} = \begin{cases} 0, & 0 \leqslant f < [1 + (N-1)b]/N, \\ \dfrac{i}{N}, & [1 + (N-1)b + (i-1)(1-b)]/N \\ & \leqslant f < [1 + (N-1)b + i(1-b)]/N; \\ & i = 1, 2, \cdots, N-1. \end{cases} \tag{12}$$

因此, 对应的 KLD 系数为:

$$\begin{aligned} s_{\mathrm{KLD}} &= \int_0^1 \delta_{\mathrm{KLD}}(f)\,\mathrm{d}f \\ &= \sum_{i=1}^{N-1} \int_{[1+(N-1)b+(i-1)(1-b)]/N}^{[1+(N-1)b+i(1-b)]/N} \frac{i}{N}\,\mathrm{d}f \\ &= \sum_{i=1}^{N-1} \frac{i}{N} \cdot \frac{1-b}{N} \\ &= \frac{N(N-1)/2}{N} \cdot \frac{1-b}{N} \\ &= \frac{(N-1)}{2N}(1-b). \end{aligned} \tag{13}$$

不难得到, 多维序列的 Ns_{KLD} 为:

$$Ns_{\mathrm{KLD}} = \frac{s_{\mathrm{KLD}}}{s_{\mathrm{KLDup}}} = (1-b). \tag{14}$$

由上式可知, 当多维序列中各时间序列之间的相关结构发生改变时, 其 Ns_{KLD} 也会相应地线性变化, 因此 Ns_{KLD} 变化可以刻画多维序列中相关结构的变化.

3 仿真结果和验证实例的讨论

3.1 相关系数对 Ns_{KLD} 的影响

为了构造如 (10) 式所示相关矩阵的多维序列, 我们首先产生 $n+1$ 维不相关 (任一通道的时间序列方差均归一化) 或统计标准正交的多维序列 (m 个时间点) , 将其中的 n 维序列组成一个不相关或统计正交的多维序列, 然后结合剩余的那个一维时间序列, 按照以下的线性组合方式产生多维序列:

$$\begin{aligned} X_i(t) =& \sqrt{1-b}\,x_i(t) + \sqrt{b}\,x_{n+1}(t), \\ & i = 1, 2, \cdots, n,\ t = 1, 2, \cdots, m. \end{aligned} \tag{15}$$

将构造的多维序列分别求 Ns_{KLD}, 结果如图 2 所示. 图 2(a) 所示的是估计的 Ns_{KLD} 与相关系数 b 之间的关系, 其中, $m = 500$, $n = 48$, $b \in [0, 1]$ (步长为 0.001). 不难看出, 多维序列的 Ns_{KLD} 与相关系数 b 之间呈现非常好的线性关系. 图 2(b) 所示的是估计的 Ns_{KLD} 与式 (14) 所示的解析值之间的差值 $\Delta Ns_{\mathrm{KLD}} = Ns_{\mathrm{KLD}} - (1-b)$. 由图 2(b) 可以看出, Ns_{KLD} 估计值和解析值之间的误差非常小, 小于 2×10^{-4}. 图 2(c) 是时间点数 m 从 100 变化到 1000 时估计的 Ns_{KLD} ($n = 48$, $b = 0.3$), 图 2(d) 为时间序列数目 n 从 32 变化到 256 时估计的 Ns_{KLD} ($m = 500$, $b = 0.3$). 很显然, 在这种相关结构的条件下, 时间序列长度和时间序列数目对 Ns_{KLD} 的估计结果几乎没有影响, 而且与式 (14) 所示的解析结果也符合得很好. 于是, 多维序列中各时间序列之间的相关程度增加时, Ns_{KLD} 将会减小.

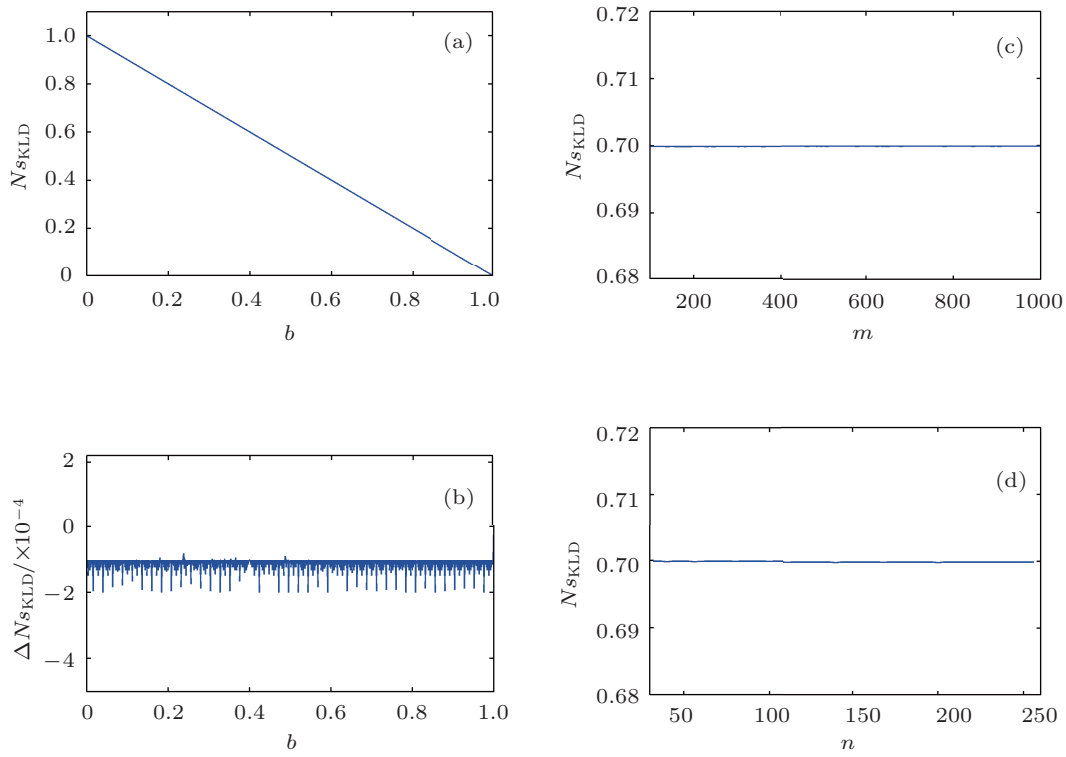

图 2 相关系数对 $Ns_{\rm KLD}$ 的影响 (a) b 变化时; (b) $\Delta Ns_{\rm KLD}$; (c) m 变化时; (d) n 变化时

3.2 N_f-tori 系统的非平稳多维序列

为了验证该方法的性能, 我们选择一系列标准的 N_f-tori 动力系统 [1], 即用 N_f 个等振幅正弦波振荡器的叠加来构成一个多维序列, 其中这 N_f 个振荡器的频率互相不可通约. 于是, N_f-tori 系统的构建如式 (16) 所示.

$$X_i(t) = \sum_{j=1}^{N_f} \sin(2\pi f_j t + \varphi_{ij}), \quad i = 1, \cdots, n. \tag{16}$$

其中频率 f_j 集合满足区间 $[0, f_{\max}]$ 内的均匀分布, 且 $f_{\max} = 20$. 所有通道的时间序列都选择同一个频率集合. 然而, 对于随机相位 φ_{ij}, 每一个通道的时间序列都是从区间 $[0, 2\pi]$ 内独立选取的.

对于 N_f-tori 系统的多维序列, 其相关结构是由相位 φ_{ij} 控制的; 如果相位 φ_{ij} 是 $[0, 2\pi]$ 区间内的均匀分布, 则 $X_i(t)$ 之间不相关, 任何偏离均匀分布的行为都将导致时间序列之间的相位相关 [1]. 于是, 为了产生相位相关的数据, 我们可以让 N_f-tori 系统的相位 φ_{ij} 符合给定均值和方差的高斯分布.

为了构建相关结构依赖于时间而变化的非平稳多维序列, 我们将时间序列分成时间上相继的三个子段, 每一段由 5120 个时间点组成. 在第一时间段中, 每一个通道时间序列的部分相位 φ_{ij} 从均匀分布的集合中独立地随机选取, 并在这个时间段内保持不变. 然后, 从一个给定均值和方差的高斯分布集合中选取一个新的随机部分相位子集 (对相位进行模 2π 运算). 这些高斯分布的部分相位任意指定到每一个 i 通道的 f_j. 在第二个时间段内, 部分相位从均匀分布的 φ_{ij} 线性地变化到相应高斯分布的 φ'_{ij}. 最后, 在第二时间段结束时, 部分相位已经达到相应高斯分布的相位值. 在第三时间段中, 所有的部分相位 φ'_{ij} 保持不变. 利用上述方法, 我们构建了 A, B, C, D 四种非平稳多维序列 ($n = 48$).

情形 A: 对所有通道, 25%最低频率的部分相位从均匀分布线性地变化到均值 $\mu_\varphi = \pi/2$, 方差 $\sigma_\varphi = \pi/4$ 的高斯分布, 其它频率的部分相位仍然保持均匀分布的值.

情形 B: 和 A 类似, 只是 50%最低频率的部分相位发生线性的变化.

情形 C: 和 A 类似, 只是 100%频率的部分相位发生线性的变化.

情形 D: 多维序列中, 其中两个通道中所有部分相位从均匀分布线性地变化到均值 $\mu_\varphi = \pi/2$,

方差 $\sigma_\varphi = \pi/8$ 的高斯分布, 而其它通道的部分相位仍然保持均匀分布的值.

利用滑动时间窗口技术, 将多维序列按 $m = 500$ 的时间窗口分段, 相邻两个窗口的重叠区域为 490, 对每一个窗口子多维序列分别求 Ns_{KLD}, 结果如图 3 和图 4 所示. 其中图 3 是非平稳多维序列情形 A, B, C 的结果, 图 4 是情形 D 的估计结果.

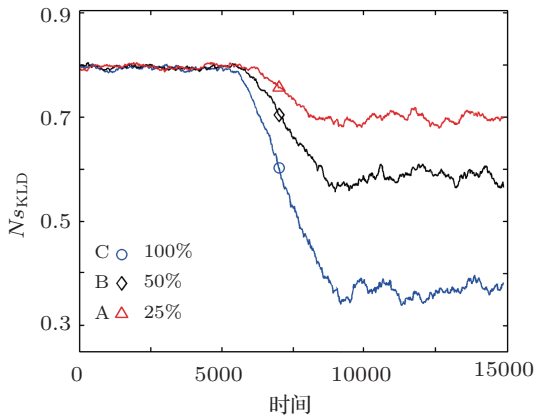

图 3 情形 A, B, C 的 Ns_{KLD} 随时间的演化

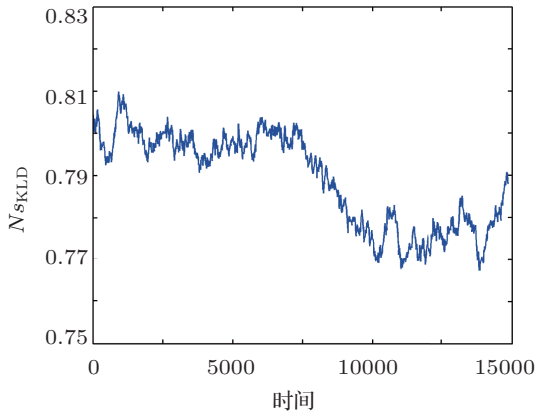

图 4 情形 D 的 Ns_{KLD} 随时间的演化

从图 3 不难看出, 在第一个时间段和第三个时间段, 部分相位分别为均匀分布和高斯分布, 而且在这两个时间段内部分相位均保持数值不变, 因此多维序列所有通道之间的相关结构变化很小, 于是 Ns_{KLD} 只有很小的波动. 而在第二个时间段中, 由于多维序列的部分相位从均匀分布线性地变化为高斯分布, Ns_{KLD} 也随之相应地变化, 有趣的是, 这种变化也几乎是线性的. 然而, 在时间段一和二以及时间段二和三的交接处, Ns_{KLD} 的变化稍显复杂, 这是因为在时间段的交接处, 多维序列通道之间的相关结构比其它时间复杂一些. 除此之外, 我们还发现, 第一时间段的 Ns_{KLD} 大于第三时间段的值. 正如前面所说的, 对于 N_f-tori 系统的多维时间序列, 如果部分相位是 0—2π 之间均匀分布, 则 $X_i(t)$ 之间不相关, 因此均匀分布的相位将导致较大的 Ns_{KLD}, 对均匀分布的任何偏离都会导致通道时间序列之间相关的增加, 于是 Ns_{KLD} 将会减小. 从图 4 中可以看出, 即使多维序列中相关结构的变化只是由两个通道引起, 在 Ns_{KLD} 中也会有所反映. 然而, 由于绝大部分通道的相关结构没有变化, 造成 Ns_{KLD} 的过渡期变大了, 于是线性变化的区域相对变小了.

3.3 耦合映象格子的非线性系统

为了验证该方法在时空非线性系统条件下的性能, 我们用 Ns_{KLD} 研究耦合映象格子 (CML) 的相关结构. CML 是复杂时空行为的一种简单模型, 但却展现出一系列有趣的特性 [23]. 我们讨论系统尺寸为 L 的一维最近邻耦合 CML, 模型表示如下:

$$x_{n+1}(i) = (1-\varepsilon) f(x_n(i)) + \frac{\varepsilon}{2}[f(x_n(i-1)) + f(x_n(i+1))], \tag{17}$$

其中 n 表示离散后的时间, ε 为耦合强度, 周期性边界条件由 $x_n(0) = x_n(L)$ 实现. CML 的局部动力学选择单峰映象, 即: $f(x) = 1 - \alpha x^2$. 众所周知, 随着参数 ε 和 α 的变化, CML 可以展现出非常广泛的复杂行为 [23].

我们利用随机初始条件和式 (17) 产生 CML 的多维序列, 并丢弃前面 8000 个时间点作为过渡过程. 这样 CML 多维序列由 $m = 500$, $n = L = 100$ 所组成. 其中参数 $\alpha = 2$, $\varepsilon \in [0, 1]$, 步长为 0.001. 估计 CML 多维时间序列的 Ns_{KLD}, 结果如图 5(b) 所示. 为了比较 Ns_{KLD} 的性能, 同时对该多维序列估计其最大 Lyapunov 指数 λ. 需要指出的是, 最大 Lyapunov 指数的估计采用基于动力学方程的 Wolf 方法 [24], 结果如图 5(a) 所示.

图 5 中可以看出, 当耦合强度 $\varepsilon \leqslant 0.120$ 和 $\varepsilon \geqslant 0.185$ 时, 最大 Lyapunov 指数和 Ns_{KLD} 都比较大, 各通道之间的相关程度比较低, 系统处于完全发展湍流模式. 当 $0.120 < \varepsilon \leqslant 0.164$ 时, 最大 Lyapunov 指数波动频繁, 而 Ns_{KLD} 虽然波动频繁, 但是其数值和波动的幅度都小得多, 这时系统处于缺陷扩散模式. 由于此时缺陷是在两个相互错了一个空间周期相位的区域之间的局部混沌, 是空间局域化的 [23], 因此有较小的 Ns_{KLD}.

当 $0.164 < \varepsilon < 0.185$ 时, 最大 Lyapunov 指数为负值, Ns_{KLD} 几乎为 0, 这时系统处于图案选择模式. 这是因为在图案选择模式中, 整个系统的混沌受到了抑制.

虽然最大 Lyapunov 指数和 Ns_{KLD} 具有类似的结果, 但是需要强调的是, 最大 Lyapunov 指数的估计需要用到其动力学方程. 而当一个系统的动力学方程未知时, 如分析实验中获得的数据, 正确估计最大 Lyapunov 指数仅用 500 个时间点是远远不够的, 这时 Ns_{KLD} 就显示出其特别的优越性.

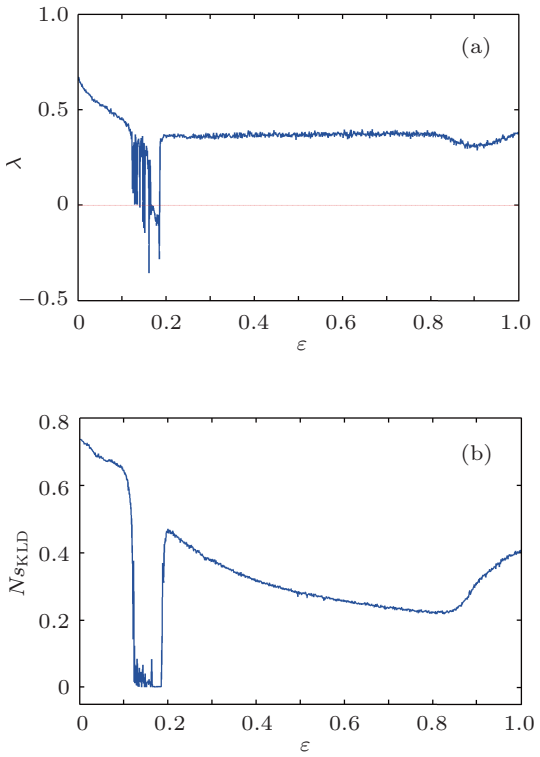

图 5 CML 多维序列 (a) 最大 Lyapunov 指数; (b) Ns_{KLD}

Ns_{KLD} 反映了多维序列通道之间的相关程度, 它很容易推广到多维序列通道之间同步的检测. 当多维序列中通道之间达到完全同步时, 则通道时间序列之间的相关系数均为 1, 相关矩阵为全 1 矩阵, 其 Ns_{KLD} 必定为 0. 反之, 当所有通道都完全不同步时, 其 Ns_{KLD} 必定为 1. 于是我们可以引入 KLD 同步系数 $\mathrm{Syn}_{\mathrm{KLD}} = 1 - Ns_{\mathrm{KLD}}$, 用 KLD 同步系数来刻画多维序列通道之间同步程度的大小. 很显然, 多维序列各通道之间的同步性越好, 则 $\mathrm{Syn}_{\mathrm{KLD}}$ 越大, 完全同步时达到 1. 反之, 多维序列各通道之间的同步性越差, 则 $\mathrm{Syn}_{\mathrm{KLD}}$ 越小, 完全不同步时 $\mathrm{Syn}_{\mathrm{KLD}}$ 达到 0. 利用 KLD 同步系数, 我们可以检测多维序列各通道之间的同步, 特别当系统的动力学方程未知时, 如多通道 EEG, fMRI 等.

4 结 论

KLD 维度和维密度的应用范围已经从线性领域拓展到刻画非线性的时空混沌系统, 从而广泛地应用于量化高维非线性系统的复杂度. 然而其存在固有的局限性, 特别是其忽略了相关矩阵中较小的特征值, 已经影响到 KLD 维度和维密度实际应用的进一步推广. KLD 系数的提出可以克服这些缺点. 利用完全不相关和完全相关的多维序列可以导出 KLD 维密度的上界和下界函数, 进而导出 KLD 系数上界和下界的解析结果, 于是 Ns_{KLD} 被提出. Ns_{KLD} 的取值范围为 [0,1], 与 KLD 维度和维密度不同, 它不仅反映了相关矩阵整个特征值谱的特性, 而且不依赖于某个任意的分数 f, 克服了 KLD 维度和维密度必须选择合适的分数 f 的难点.

解析结果和仿真的结果都证明, 多维序列中时间序列之间相关程度的变化可以引起 Ns_{KLD} 的线性变化. N_f-tori 系统的结果还表明, 即使多维序列中仅有其中的两个时间序列之间相关程度发生变化, Ns_{KLD} 也会产生相应的变化. 更为重要的是, Ns_{KLD} 还可以用来进行非平稳时间序列的分析.

CML 的结果表明, Ns_{KLD} 也可以有效地刻画非线性系统的相关结构, 进而可以推广到同步特性的研究, 特别是当系统的动力学方程未知时, 作为系统分析的一种数据驱动的方法, 有望推广到实际多维序列, 特别是短时间序列的分析中.

[1] Müller M, Baier G 2005 *Phys. Rev.* E **71** 046116

[2] He G G, Zhu P, Chen H P, Xie X P 2010 *Acta Phys. Sin.* **59** 5307 (in Chinese) [何国光, 朱萍, 陈宏平, 谢小平 2010 物理学报 **59** 5307]

[3] Quiroga R Q, Kraskov A, Kreuz T, Grassberger P 2002 *Phys. Rev.* E **65** 041903

[4] Pereda E, Quiroga R Q, Bhattacharya J 2005 *Progress in Neurobiology* **77** 1

[5] Stam C J 2005 *Clinical Neurophysiology* **116** 2266

[6] Bloomfield P 2000 *Fourier Analysis of Time Series* (New York: A Wiley-Interscience Publication, John Wiley & Sons, INC.) p1

[7] Xie X P, Cao Z T, Weng X C 2008 *NeuroImage* **40** 1672

[8] Zoldi S M, Greenside H S 1997 *Phys. Rev. Lett.* **78** 1687
[9] Mallat S 1989 *IEEE Pattern Anal. and Machine Intell.* **11** 674
[10] Raab C, Kurths J 2001 *Phys. Rev.* E **64** 016216
[11] Bauer M, Heng H, Martienssen W 1993 *Phys. Rev. Lett.* **71** 521
[12] Bünner M J, Hegger R 1999 *Physics Letters* A **258** 25
[13] Zeng X, Eykholt R, Pielke R A 1991 *Phys. Rev. Lett.* **66** 3229
[14] Sano M, Sawada Y 1985 *Phys. Rev. Lett.* **55** 1082
[15] Kramer M A 1991 *Amer. Inst. Chem. Engin. Journal* **37** 233
[16] Roweis S T, Saul L K 2000 *Science* **290** 2323
[17] Common P 1991 *Signal Processing* **24** 1
[18] Meixner M, Zoldi S M, Bose S, Schöll E 2000 *Phys. Rev.* E **61** 1382
[19] Xie X P, Zhao X H, Fang Y T, Cao Z T, He G G 2011 *Physics Letters* A **375** 1789
[20] Varela H, Beta C, Bonnefont A, Krischer K 2005 *Phys. Rev. Lett.* **94** 174104
[21] Plerou V, Gopikrishnan P, Rosenow B, Amaral L A N, Guhr T, Stanley H 2002 *Phys. Rev.* E **65** 066126
[22] Bazzani A, Castellani G C 2010 *Phys. Rev.* E **81** 051917
[23] Yang W M 1994 *Spatiotemporal Chaos and Coupled Map Lattices* (Shanghai: Shanghai Scientific and Technological Education Publishing House) p12 (in Chinese) [杨维明 1994 时空混沌和耦合映象格子 (上海: 上海科技教育出版社) 第 12 页]
[24] Wolf A, Swift J B, Swinney H L, Vastano J A 1985 *Physica* D **16** 285

The normalized KLD coefficient and its application in detection of correlation and synchronization from multivariable series*

Xie Xiao-Ping Chen Hong-Ping Cao Zhi-Tong He Guo-Guang†

(*Department of Physics, Zhejiang University, Hangzhou* 310027, *China*)

(Received 17 October 2011; revised manuscript received 28 November 2011)

Abstract

The KLD coefficient and the normalized KLD coefficient are proposed to characterize the correlation of multivariable series in order to overcome the intrinsic limitations of the KLD dimension density. Using the uncorrelated or perfectly correlated multivariable series, the upper and the lower bound functions of the KLD dimension density, and furthermore the upper and the lower bounds of the KLD coefficient are analytically deduced. Then, the normalized KLD coefficient is proposed in the paper. The analyses and numerical simulations prove that the changes of correlation of multivariable series can lead to linear variation of the normalized KLD coefficient. The simulations also prove that the normalized KLD coefficient can detect the changes of correlation sensitively, even if these are induced by only two channels of multivariable series. Furthermore, the normalized KLD coefficient can be used to analyze the nonstationary time series. The simulation results of coupled map lattice prove that the normalized KLD coefficient can also be used for the nonlinear system analysis.

Keywords: multivariable series analysis, correlation, detection, KLD

PACS: 05.45.−a, 05.45.Tp

* Project supported by the National Natural Science Foundation of China (Grant No. 61175060) and the Foundation of the Zhejiang Educational Committee of China (Grant No. Y200909251).

† E-mail: gghe@zju.edu.cn