

一种无偏差的多通道多尺度样本熵算法*

李惟嘉¹⁾²⁾ 申晓红^{1)†} 李亚安²⁾

1) (西北工业大学航海学院, 西安 710072)

2) (西北工业大学, 海洋声学信息感知工业和信息化部重点实验室, 西安 710072)

(2023 年 7 月 14 日收到; 2024 年 4 月 7 日收到修改稿)

多通道数据采集技术的发展为复杂系统非线性动力学特性研究提供了更加丰富的先验信息. 然而传统的非线性特征量只能处理单通道数据, 无法直接提取多通道数据的非线性特征. 近年来, 有学者对多尺度样本熵算法进行了一般化处理, 提出了多通道多尺度样本熵算法. 该算法不仅可以对多通道数据整体的复杂度进行表征, 还可以有效提取通道内和通道间隐含的相关性信息. 但是, 该算法缺乏相应的理论支撑, 在实际应用中无法兼顾性能和稳定性. 针对以上问题, 本文提出了一种无偏差的多通道多尺度样本熵算法, 并利用概率论从理论上分析了多通道多尺度样本熵算法不稳定以及性能差的原因. 后续的仿真实验证明改进后的算法不但可以有效地提取通道内和通道间的相关性信息, 同时在处理复杂数据时表现出了良好的稳定性. 该算法为诸如模糊熵、排列熵等非线性特征量的算法一般化提供了思路和理论依据.

关键词: 非线性动力学, 多通道数据, 多通道多尺度样本熵

PACS: 05.45.-a, 89.70.Cf, 43.60.Cg

DOI: 10.7498/aps.73.20231133

1 引言

非线性动力学领域的研究工作一直致力于从有限数据中提取复杂非线性系统的特征. 精确地表征不同系统的特征对于后续的识别、分类和检测工作至关重要^[1-4]. 传统的研究方案是通过单通道的一维数据对系统隐含的非线性动力学特性进行表征^[5-7]. 例如, 针对一个 n 维的复杂非线性系统, 利用相空间重构将一维的数据映射到高维. 在高维空间中, 观察吸引子结构, 分析相空间中轨线的运动规律, 计算该系统诸如李雅普诺夫指数、分形维数、熵等非线性特征量^[8]. 从而揭示出该系统隐含的非线性动力学特性. 可以看出, 传统的非线性动力学研究都是利用单变量的观测数据提取复杂系统的非线性动力学特征. 这种研究思路的优点在于对先验信息的要求较少. 但这种研究方法的缺点是

无法完整、精确地描述系统的动力学特性.

相比于单通道数据, 利用多个传感器同时采集的多通道数据可以为非线性动力学系统的分析提供额外信息^[9,10]. 例如, 在水声信号处理领域, 阵列采集的多通道数据可提高特定方向上的信噪比^[11,12]. 从复杂系统的多通道数据中提取不同的特征向量, 可以有效地帮助神经网络、深度学习和支持向量机对不同的系统进行分类识别^[13,14]. 在非线性动力学研究领域中, 像李雅普诺夫指数、分形维数、熵等经典非线性特征量通常只适用于一维时间序列(单通道数据). 针对多通道数据, 也只能分别计算每一个通道对应的非线性特征量. 这样的处理方式忽略了通道之间的内在联系, 浪费了大量的有效信息.

作为一种经典的非线性特征量, 熵被广泛应用于机械故障诊断、水声信号分类识别、病理信号诊断等领域^[15-17]. 经过学者们的不断改进, 熵算法衍生出许多的分支. 如近似熵 (approximate entropy,

* 国家自然科学基金(批准号: 11874302, 62031021)资助的课题.

† 通信作者. E-mail: xhshen@nwpu.edu.cn

AE)^[18]、样本熵 (sample entropy, SE)^[19]、排列熵 (permutation entropy, PE)^[20]、多尺度样本熵 (multiscale sample entropy, MSE) 等^[21]。熵算法具有适用性强、计算量相对较小等优点。不同的熵算法从不同的角度对给定一维数据的复杂度进行了表征。在实际应用中,可以根据需要选取不同的熵算法提取目标信号的复杂度特征。然而,现有的熵算法都只能处理单通道数据,无法直接用来提取多通道数据的复杂度信息。

近年来, Ahmed 和 Mandic^[22] 将多尺度样本熵算法中的计算步骤进行了升维处理并提出了多通道多尺度样本熵算法 (multivariate multiscale sample entropy, MMSE)。该算法可以直接处理多通道数据,提取多通道数据的复杂度信息。然而,在将嵌入维数从 m 增加到 $m+1$ 并重构出对应的 $(m+1)$ 维相空间时, Ahmed 和 Mandic 发现 MMSE 算法无法在这一个计算步骤中得到一个确定的 $(m+1)$ 维相空间。为此, Ahmed 和 Mandic 提出了两种不同的多通道多尺度样本熵算法,分别是朴素多通道多尺度样本熵算法 (naive MMSE) 和严格多通道多尺度样本熵算法 (rigorous MMSE)。朴素多通道多尺度样本熵算法单独计算了每一种可能的 $(m+1)$ 维相空间中对应的概率,并将所有可能出现的 $(m+1)$ 维相空间对应的概率求和取平均得到最终需要的概率。而严格多通道多尺度样本熵算法将所有可能的 $(m+1)$ 维相空间合并,在合并后的相空间中计算出最终需要的概率。通过仿真研究, Ahmed 和 Mandic 认为严格算法是优于朴素算法的。因为严格算法可以提取不同通道之间的相关性信息。需要注意的是,采用严格多通道多尺度样本熵算法计算多通道数据的复杂度时,隶属于不同通道的样本点之间的相似性也会对结果产生影响。而不同通道间数据计算相似性对结果产生的影响是不确定的:一方面,在计算熵值时考虑不同通道的样本点之间的相似性可能给严格多通道多尺度样本熵算法带来了提取通道间的相关性信息的能力;另一方面,在实际应用中,不同通道之间数据的相关性是无法确定的。盲目地计算不同通道之间样本点的相似性可能使最终结果出现偏差。Ahmed 和 Mandic 直观地认为在算法中考虑不同通道之间样本点的相似性可以提升算法的性能,使算法可以提取通道间的相关性信息。

除此之外,多通道多尺度样本熵算法的理论证

明并不充分。在将多尺度样本熵升维处理时,事实上是有很多不同的升维策略的。现有的研究中没有对所有的策略进行充分完整的讨论,而是直接将多通道多尺度样本熵算法与其他诸如模态分解等技术结合并应用于故障诊断等领域^[23,24]。因此,有必要对可能的策略进行全面的分析和讨论,使多通道熵算法的理论更加可靠和完整。本文从单通道熵算法 MSE 的核心思想出发,分析了 MSE 算法的升维过程。通过理论研究,证明了严格多通道多尺度样本熵和朴素多通道多尺度样本熵都不是 MSE 算法在多通道情况下的精确推广。基于以上的理论分析,提出了一种无偏差的多通道多尺度样本熵算法 (UMMSE)。通过理论证明与仿真实验,证明了计算不同通道间样本点的相似性不能给算法带来提取通道间相关性信息的能力。也就是说,使用严格算法将得到一个有偏差的熵结果。

本文后续的结构如下:第 2 节介绍无偏差的多通道多尺度样本熵算法;第 3 节对无偏差的多通道多尺度样本熵算法、naive MMSE 以及 rigorous MMSE 算法升维过程中的策略进行分析讨论;第 4 节利用不同的仿真实验证明改进后算法优于原始多通道多尺度样本熵算法;最后第 5 节给出结论。

2 无偏差的多通道多尺度样本熵算法

Ahmed 和 Mandic 等^[22] 于 2011 年提出 MMSE 算法。该算法成功地将一种仅适用于一维数据的熵算法拓展到了多维情形。虽然仿真实验和实际数据处理的结果证明了该算法是一种有效的复杂度度量方法,并且可以有效地提取通道内和通道间的相关性信息。但是,至今没有学者对该算法进行严格的推导与证明。Ahmed 和 Mandic^[22] 没有从根本上解决 MMSE 算法存在的问题,而是提出了两种子算法 (严格算法以及朴素算法) 来弥补 MMSE 算法存在的缺陷。

针对 MMSE 算法存在的问题,本文进行了改进并提出无偏差的多通道多尺度样本熵算法。该算法由以下步骤组成。

1) 给定多通道数据 $\{x_{i,j}\}, 1 \leq i \leq N, 1 \leq j \leq p$ 。其中, N 表示每个通道内的样本点数, p 表示通道数。为了消除不同通道间的幅值差异对结果产生的影响,对各个通道进行归一化处理。

2) 在归一化处理后,利用下式将每一个通道分别进行粗粒化处理:

$$\{y_{i,j}^{(n)}\} = \frac{1}{n} \sum_{k=1}^n x_{ni-n+k,j}, \quad 1 \leq i \leq N_n, \quad 1 \leq j \leq p, \quad (1)$$

其中, n 表示尺度因子, $N_n = \left\lfloor \frac{N}{n} \right\rfloor$ 表示 $\frac{N}{n}$ 向下取整. 对于每一个通道来说, 粗粒化过程将原本的时间序列 $\{x_{i,j}\}$ 分割成了 N_n 个子序列, 并将每一段子序列求和取平均得到新的序列 $\{y_{i,j}^{(n)}\}$. 注意当尺度因子 $n = 1$ 时, $\{y_{i,j}^{(n)}\} = \{x_{i,j}\}$.

步骤 1) 和 2) 与后续的熵算法相互独立, 因此在下文介绍改进多通道样本熵算法时, 省略了尺度因子角标 n , 用 $\{y_{i,j}\}$ 表示经过粗粒化处理后的多通道数据.

3) 得到 $\{y_{i,j}\}$ 后, 进行相空间重构^[25-27]:

$$\begin{aligned} Y_M(i) = & [y_{i,1}, y_{i+\tau_1,1}, \cdots, y_{i+(m_1-1)\tau_1,1}, \\ & y_{i,2}, y_{i+\tau_2,2}, \cdots, y_{i+(m_2-1)\tau_2,2}, \cdots, \\ & y_{i,p}, y_{i+\tau_p,p}, \cdots, y_{i+(m_p-1)\tau_p,p}], \\ & 1 \leq i \leq q, \end{aligned} \quad (2)$$

式中, i 是 M 维相空间中各个向量的索引角标; $\mathbf{M} = [m_1, m_2, \cdots, m_p] \in \mathbb{R}^p$ 表示嵌入维数向量; $\boldsymbol{\tau} = [\tau_1, \tau_2, \cdots, \tau_p] \in \mathbb{R}^p$ 表示的是时间延迟向量; $M = \sum_{i=1}^p m_i$. 在重构出的 M 维相空间中共有 q 个向量.

4) 在利用 (2) 式重构出 M 维相空间后, 计算出相空间中任意两个不同向量之间的切比雪夫距离. 计算 $r \cdot \text{tr}(\mathbf{S})$ 作为阈值, 其中 r 是需要事先给定的系数, $\text{tr}(\mathbf{S})$ 是多通道数据协方差矩阵 \mathbf{S} 的迹. 在定义好距离并设置好阈值后, 针对 M 维相空间中的任意一个向量 $\mathbf{Y}_M(i)$, 就可以统计出在 M 维相空间中与其距离小于给定阈值的向量个数, 记为 $C_i^M(r)$. 当两个向量之间的距离小于阈值时, 称这两个向量相似.

5) 根据下式计算出相应的 $B_i^M(r)$:

$$B_i^M(r) = \frac{1}{q-1} C_i^M(r), \quad (3)$$

$B_i^M(r)$ 表示的是 M 维相空间中的其余向量与向量 $\mathbf{Y}_M(i)$ 之间的距离小于阈值的概率.

6) 遍历相空间中的所有向量, 将对应的 $B_i^M(r)$ 求和取平均, 求出

$$B^M(r) = \frac{1}{q} \sum_{i=1}^q B_i^M(r). \quad (4)$$

7) 得到 $B^M(r)$ 后, 增加嵌入维数使其从 $\mathbf{M} =$

$[m_1, m_2, \cdots, m_p]$ 变为 $\mathbf{M} + \mathbf{p} = [m_1 + 1, m_2 + 1, \cdots, m_p + 1]$. 根据新的嵌入维数, 重复步骤 3)—6), 重构出对应的 $(M + p)$ 维相空间并计算出对应的 $C_i^{M+p}(r)$, $B_i^{M+p}(r)$ 和 $B^{M+p}(r)$.

8) 最后, 求出无偏差的多通道多尺度样本熵结果:

$$\text{UMMSE}(\mathbf{M}, \boldsymbol{\tau}, r, N) = -\ln[B^{M+p}(r)/B^M(r)]. \quad (5)$$

无偏差的多通道多尺度样本熵算法实现了对多通道数据复杂度的估计. 一方面, 该算法是 MSE 算法在多维情况下的准确推广. 另一方面, 与 MSE 算法相比, 由于无偏差的多通道多尺度样本熵算法同时涉及了所有的通道, 因此也可以揭示不同通道之间的相关性信息.

2.1 多通道多尺度样本熵算法

为了更好地对不同算法进行对比, 本文在这里对 Ahmed 和 Mandic^[22] 提出的多通道多尺度样本熵算法进行介绍. 其中包含朴素和严格两种算法. 步骤 1)—6) 与无偏差的多通道多尺度样本熵算法相同. 为了与无偏差的多通道多尺度样本熵区分, MMSSE 的后续步骤编号会增加星号角标.

7*) 在得到 $B^M(r)$ 后, 增加嵌入维数使嵌入维数向量各元素之和从 M 变为 $M + 1$. 当通道个数为 p 时, 就会有 p 种可能的情况, 相应地可以重构出 p 种可能的 $(M + 1)$ 维相空间. 由于生成的 $(M + 1)$ 维相空间不唯一, 因此不能直接套用步骤 4)—6) 计算出相应的 $B^{M+1}(r)$. 下面将分别介绍朴素算法与严格算法.

朴素的多通道多尺度样本熵算法 针对每一个可能的 $(M + 1)$ 维相空间, 重复步骤 4)—6) 计算出对应的概率, 在这里用 $B^{M+1_i}(r)$ 表示第 i 个通道的嵌入维数增加 1 后重构的 $(M + 1)$ 维相空间对应的概率, 其中 $1 \leq i \leq p$. 对 p 个 $B^{M+1_i}(r)$ 求和取平均, 得到最终的 $B^{M+1}(r)$.

严格的多通道多尺度样本熵算法 现有 p 种可能的 $(M + 1)$ 维相空间, 每一个相空间中有 q' 个向量, 则共有 pq' 个向量. 针对所有的向量, 重复步骤 4)—6) 计算得到最终的 $B^{M+1}(r)$.

则根据 (6) 式即可计算出朴素的和严格的多通道多尺度样本熵结果:

$$\text{MMSE}(\mathbf{M}, \boldsymbol{\tau}, r, N) = -\ln[B^{M+1}(r)/B^M(r)]. \quad (6)$$

Ahmed 和 Mandic^[22] 提出的多通道多尺度样本熵算法可以直接处理多通道数据, 对多通道数据的复杂度进行表征. 但是在对算法进行升维时, 由于没有合理地增加嵌入维数导致无法重构出一个确定的相空间. 因此, Ahmed 和 Mandic^[22] 提出了朴素算法和严格算法. 经过分析发现, 不同通道之间数据点的相似性也会对严格算法的结果产生影响. 利用仿真实验对比不同方法的性能, Ahmed 和 Mandic 发现严格算法可以提取多通道数据通道间的相关性信息. 因此他们得出结论, 认为严格算法是优于朴素算法的.

事实上, 不论是严格算法还是朴素算法都缺乏相应的理论支撑. 不同通道的数据之间计算相似性会使熵结果不稳定. 而算法提取通道间相关性信息的能力并不是由不同通道之间数据计算相似性带来的. 本文将从概率论的角度出发, 对无偏差的多通道多尺度样本熵算法、朴素算法和严格算法进行分析.

3 算法分析与证明

无论是 Ahmed 和 Mandic^[22] 提出的多通道多尺度样本熵算法还是本文提出的无偏差的多通道多尺度样本熵算法, 其本质上都是多尺度样本熵算法在多通道数据情况下的推广. 因此, 本文将先对算法中涉及的不同事件进行定义, 再依次分析多尺度样本熵、朴素算法、严格算法、及无偏差的多通道多尺度样本熵算法. 证明无偏差的多通道多尺度样本熵算法的优越性. 在算法分析与证明这一部分中, 以两通道数据为例分析多通道情况下的多通道多尺度样本熵算法和无偏差的多通道多尺度样本熵算法. 在分析单通道情况下的多尺度样本熵算法时, 使用第一个通道的数据进行分析 (选择任意通道均不影响后续的分析与证明). 多尺度样本熵的其余计算参数也不需要专门进行设定, 使用多通道熵算法中参数的第一个分量即可. 由于粗粒化与归一化作为预处理步骤与后续的熵算法步骤独立, 因此在这里省略了步骤 1) 和 2), 等效为仅讨论在尺度为 1 时的情况.

3.1 事件定义

给定两通道数据 $\{x_{i,j}\}$, $i = 1, \dots, N; j = 1, 2$. 每个通道内包含 N 个样本点. 嵌入维数与时间延迟分别为 $M = [m_1, m_2]$ 和 $\tau = [\tau_1, \tau_2]$. 表 1 列举了算法涉及的事件以及对应的表示符号.

表 1 熵算法中涉及的事件及对应的符号
Table 1. Events involved in the entropy algorithm.

符号	事件
A	通道1中长度为 m_1 的子序列相似
B	通道2中长度为 m_2 的子序列相似
C	通道1中长度为 m_1 的子序列在第 $m_1 + 1$ 个样本点依然相似
D	通道2中长度为 m_2 的子序列在第 $m_2 + 1$ 个样本点依然相似

在单通道情况下, 多尺度样本熵算法根据给定的嵌入维数 m_1 和时间延迟 τ_1 可以重构出一个确定的 m_1 维相空间. 与无偏差的多通道多尺度样本熵算法类似, 在计算样本熵时, 需要先计算嵌入维数为 m_1 时对应的 $B^{m_1}(r)$, 再计算嵌入维数增加到 $m_1 + 1$ 时对应的 $B^{m_1+1}(r)$. 在得到 $B^{m_1}(r)$ 与 $B^{m_1+1}(r)$ 后, 计算 $B^{m_1+1}(r)/B^{m_1}(r)$ 的负自然对数. 推广到多通道的情况, 在改变嵌入维数时每一个通道的嵌入维数增量应该保证不大于 1. 根据最初给定的嵌入维数 $M = [m_1, m_2]$ 可以重构出一个确定的 M 维相空间 ($M = \sum_{i=1}^p m_i$), 并计算得到对应的 $B^M(r)$. 当增加嵌入维数时, 朴素算法与严格算法将嵌入维数之和从 M 增加到 $M + 1$. 而无偏差的多通道多尺度样本熵算法将嵌入维数之和从 M 增加到 $M + p$ (在这里, $p = 2$). 对于不同的嵌入维数增加策略, 重构出对应的相空间并计算对应的 $B^{M+i}(r)$, $i = 1, 2$. 表 2 列出了不同的嵌入维数与 $B^{M+i}(r)$, $i = 0, 1, 2$ 所对应的事件.

表 2 不同的嵌入维数所对应的事件
Table 2. Events corresponding to different embedding dimensions.

嵌入维数	可能的结果	$B^{M+i}(r)$
m_1	m_1	$P(A)$
$m_1 + 1$	$m_1 + 1$	$P(AC)$
M	$[m_1, m_2]$	$P(AB)$
$M + 1$	$[m_1 + 1, m_2]$	$P(ABC)$
	$[m_1, m_2 + 1]$	$P(ABD)$
$M + 2$	$[m_1 + 1, m_2 + 1]$	$P(ABCD)$

在对本文涉及的所有熵算法中包含的事件进行定义后, 下面利用上述事件及其概率分别分析多尺度样本熵算法、朴素算法、严格算法、以及无偏差的多通道多尺度样本熵算法.

3.2 多尺度样本熵

在多尺度样本熵算法中, 当嵌入维数为 m_1 时,

算法中的 $B^{m_1}(r)$ 可以等效为事件 A 的概率. 当嵌入维数增加到 $m_1 + 1$ 时, 对应的 $B^{m_1+1}(r)$ 可以等效为事件 A 与事件 C 的联合概率. 因此, 最终的结果即可以用概率表示为 $P(AC)/P(A) = P(C|A)$. 可以看出, 多尺度样本熵本质上计算了在特定时间尺度下, 当一段长度为 m_1 的子序列在增加一个样本点后依然保持相似的条件概率的负自然对数.

3.3 朴素的多通道多尺度样本熵

接下来对 Ahmed 和 Mandic^[22] 提出的多通道多尺度样本熵算法进行分析. 由于在算法中使用的策略是将嵌入维数增加到 $M + 1$, 从表 2 可以看到会有两种可能出现的嵌入维数结果. 这就是 Ahmed 和 Mandic 必须提出朴素算法与严格算法的原因. 因为没有有一个适用的标准来选择一个确定的嵌入维数, 导致后续重构相空间时也会出现两种可能的 $(M + 1)$ 维相空间.

朴素算法需要分别计算每一种可能的 $(M + 1)$ 维相空间对应的 $B^{(M+1)_i}$. 例如, 当嵌入维数为 $[m_1 + 1, m_2]$ 时, 对应的 $B^{(M+1)_1}$ 即为事件 ABC 的联合概率. 本文使用 $B^{(M+1)_i}$ 来表示第 i 个通道的嵌入维数增加 1 以后计算得到的概率. 那么, 经过计算得到的最终的 $B^{M+1}(r)$ 即可以用概率表示为

$$\begin{aligned} B^{M+1}(r) &= \frac{1}{2}[B^{(M+1)_1} + B^{(M+1)_2}] \\ &= \frac{1}{2}[P(ABC) + P(ABD)]. \end{aligned} \quad (7)$$

结合 (7) 式与表 2 中的结果, 可以得到朴素算法的概率表示:

$$\begin{aligned} \text{MMSE}(\mathbf{M}, \boldsymbol{\tau}, r, N) &= -\ln[B^{M+1}(r)/B^M(r)] \\ &= -\ln\left[\frac{B^{(M+1)_1} + B^{(M+1)_2}}{2B^M(r)}\right] \\ &= -\ln\left[\frac{P(ABC)}{2P(AB)} + \frac{P(ABD)}{2P(AB)}\right] \\ &= -\ln\left[\frac{1}{2}(P(C|AB) + P(D|AB))\right]. \end{aligned} \quad (8)$$

从概率论的角度对 (8) 式进行分析, 可以看出朴素算法本质上计算的依然是条件概率均值的负自然对数. 并且需要注意的是, (8) 式中的条件概率, 分别只考虑了某一个特定的通道. 并没有出现类似于 $P(CD|AB)$ 这样同时考虑所有通道的条件概率.

3.4 严格的多通道多尺度样本熵

对于严格算法, 在计算最终的 $B^{M+1}(r)$ 时同时考虑了所有可能出现的 $(M + 1)$ 维相空间. 为了便于理解, 这里对参数进行赋值. 给定每个通道的数据长度为 $N = 4$, $\mathbf{M} = [2, 2]$, $\boldsymbol{\tau} = [1, 1]$, $r = 0.15$, 先重构出对应的 M 维相空间. 重构出的相空间如 (9) 式所示:

$$\mathbb{V} = \begin{bmatrix} \mathbf{X}_M(1) \\ \mathbf{X}_M(2) \\ \mathbf{X}_M(3) \end{bmatrix} = \begin{bmatrix} [x_{1,1}, x_{2,1}, x_{1,2}, x_{2,2}] \\ [x_{2,1}, x_{3,1}, x_{2,2}, x_{3,2}] \\ [x_{3,1}, x_{4,1}, x_{3,2}, x_{4,2}] \end{bmatrix}. \quad (9)$$

根据 (9) 式中的向量 $\mathbf{X}_M(i)$, $1 \leq i \leq 3$ 和系数 r , 可以计算出对应的联合概率 $B^M(r) = P(AB)$. 当增加嵌入维数时, 会出现两种可能的结果. 分别是 $\mathbf{M}_1 = [3, 2]$ 和 $\mathbf{M}_2 = [2, 3]$. 相应地可以重构出两种不同的 $M + 1$ 维相空间. 这里使用 \mathbb{V}_1 和 \mathbb{V}_2 分别表示两种可能的 $(M + 1)$ 维相空间. 严格算法将两种可能的相空间合并, 结果如 (10) 式所示:

$$\begin{aligned} \mathbb{V}' &= \begin{bmatrix} \mathbb{V}_1 \\ \mathbb{V}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{M+1}^1(1) \\ \mathbf{X}_{M+1}^1(2) \\ \mathbf{X}_{M+1}^2(1) \\ \mathbf{X}_{M+1}^2(2) \end{bmatrix} \\ &= \begin{bmatrix} [x_{1,1}, x_{2,1}, x_{3,1}, x_{1,2}, x_{2,2}] \\ [x_{2,1}, x_{3,1}, x_{4,1}, x_{2,2}, x_{3,2}] \\ [x_{1,1}, x_{2,1}, x_{1,2}, x_{2,2}, x_{3,2}] \\ [x_{2,1}, x_{3,1}, x_{2,2}, x_{3,2}, x_{4,2}] \end{bmatrix}, \end{aligned} \quad (10)$$

其中, $\mathbf{X}_M^k(i)$ 表示相空间 \mathbb{V}_k 中的第 i 个向量, M 是各通道嵌入维数之和. 根据严格算法的定义, 由 \mathbb{V}' 计算得到的 $B^{M+1}(r)$ 可以写成 (11) 式中的概率形式:

$$B^{M+1}(r) = P(ABCDT), \quad (11)$$

式中, ABCD 所代表的事件在表 1 中已经进行了说明. 需要注意的是, (11) 式包含一个干扰事件 T. 这里首先给出事件 T 的定义. T: 隶属于不同通道的样本点之间的距离小于给定阈值, 即属于不同通道的样本点相似. (10) 式可以看到, 严格算法在增加嵌入维数后会出现计算不同通道样本点之间距离的情况. (12) 式即为一个典型例子.

$$\begin{aligned} d[\mathbf{X}_{M+1}^1(1), \mathbf{X}_{M+1}^2(2)] &= \max\{|x_{1,1} - x_{2,1}|, |x_{2,1} - x_{3,1}|, \\ &\quad |x_{3,1} - x_{2,2}|, |x_{1,2} - x_{3,2}|, |x_{2,2} - x_{4,2}|\}. \end{aligned} \quad (12)$$

需要注意的是, (11) 式中的事件 T 是事件 T 的子

集. 在 (12) 式中, 计算了 $|x_{3,1} - x_{2,2}|$ 这两个分属于不同通道之间样本点的距离. 距离在广义上可以理解为相似性. 也就是说在对比某一个通道内样本点的相似性时, 严格算法混入了小部分其他通道的样本点数据.

(13) 式给出了严格算法的概率表达:

$$\begin{aligned} \text{MMSE}(\mathbf{M}, \boldsymbol{\tau}, r, N) &= -\ln[B^{M+1}(r)/B^M(r)] \\ &= -\ln\left[\frac{P(\text{ABCDT})}{P(\text{AB})}\right] = -\ln[P(\text{CDT}|\text{AB})]. \end{aligned} \quad (13)$$

Ahmed 和 Mandic^[22] 认为是事件 T 使严格算法可以提取通道间的相关性信息. 然而对比 (8) 式和 (13) 式可以发现, 严格算法在条件概率中出现了不同通道之间的联合概率. 反观朴素算法, 则只有各个通道的条件概率 ($P(\text{C}|\text{AB}), P(\text{D}|\text{AB})$). 因此, 严格算法提取通道间相关性信息的能力可能并不是由事件 T 带来的. 相反, 这项能力可能是由不同通道之间联合的条件概率带来的.

3.5 无偏差的多通道多尺度样本熵

对于无偏差的多通道多尺度样本熵算法来说, 当嵌入维数从 M 增加到 $M+2$ 时. 由于约束条件的存在使得每一个通道的嵌入维数增量不能大于 1. 因此可以生成一个确定的 $(M+2)$ 维相空间. 根据生成的 $(M+2)$ 维相空间, 可以计算得到对应的 $B^{M+2}(r)$, 如下式所示:

$$B^{M+2}(r) = P(\text{ABCD}). \quad (14)$$

(14) 式表明当嵌入维数增加到 $M+2$ 时, $B^{M+2}(r)$ 是事件 ABCD 的联合概率. (15) 式给出了无偏差的多通道多尺度样本熵算法的概率表示方法:

$$\begin{aligned} \text{UMMSE}(\mathbf{M}, \boldsymbol{\tau}, r, N) &= -\ln[B^{M+2}(r)/B^M(r)] \\ &= -\ln\left[\frac{P(\text{ABCD})}{P(\text{AB})}\right] = -\ln[P(\text{CD}|\text{AB})]. \end{aligned} \quad (15)$$

对比 (8) 式, (13) 式和 (15) 式可发现仅有严格算法包含了事件 T 的子事件 T. 同时, 无偏差的多通道多尺度样本熵算法和严格算法均含有不同通道之间的联合条件概率. 本文认为, 事件 T 会使多通道熵算法的结果有偏. 而不同通道之间的联合条件概率使算法有能力提取通道间的相关性信息. 在后续的仿真实验中, 本文构造出特定的仿真信号, 对比不同算法之间性能的差异来证明本文的观点.

4 仿真实验

仿真实验主要分为以下几个部分: 首先, 研究数据长度、嵌入维数和阈值系数这 3 个参数对无偏差的多通道多尺度样本熵算法性能的影响, 并给出相应的参数选择范围. 在确定合适的参数后, 第一组仿真实验研究无偏差的多通道多尺度样本熵算法提取通道内相关性信息方面的能力; 第二组仿真实验则对比研究 3 种算法在提取通道间相关性信息方面的表现; 第三组仿真实验重点探究严格算法与无偏差算法之间的差异. 这里使用这两种算法计算了各通道间方差不同的高斯白噪声数据的熵值, 结果表明无偏差算法比严格算法更为稳定. 最后一组仿真实验比较了在相同参数条件下 3 种算法计算样本熵值所需的时间, 以研究不同多通道熵算法的计算效率. 需要说明的是, 除了第 4 组仿真实验外, 其余的仿真图均为误差条图. 绘制方法是: 独立重复多次进行仿真, 计算熵值的均值与方差. 根据计算得到的均值与方差, 绘制对应的误差条图.

在无偏差的多通道多尺度样本熵算法中, 主要涉及的参数分别为数据长度、嵌入维数以及阈值系数. 在计算过程中采用控制变量法, 控制其余参数不变, 单独研究某一个特定参数对结果产生的影响. 参数设置为: 数据长度 10000, 嵌入维数 $\mathbf{M} = [2, 2, 2]$, 阈值系数为 0.15. 结果如图 1 所示.

在图 1(a) 中, 计算熵值时改变数据长度, 范围为 $[600, 10500]$, 其余参数不变. 结果可以看出当数据长度过短时, 由于数据量不足, 在计算过程中可能会出现 $B^M(r)$ 或 $B^{M+p}(r)$ 过小甚至等于 0 的情况, 从而使熵结果不稳定. 当数据长度小于 5000 时, 数据长度的变化会对结果的稳定性产生较大影响. 当数据长度大于 5000 时, 数据长度的变化对稳定性产生的影响较小. 无偏差的多通道多尺度样本熵值随着数据长度的增加逐渐稳定. 考虑到实际应用中数据长度的获取成本以及计算效率的问题, 通常将数据长度的范围设置为 $[5000, 10000]$. 图 1(b) 展示了选取不同的嵌入维数时, 无偏差的多通道多尺度样本熵结果. 从结果可以看出通常每个通道的嵌入维数均应设置为 2. 最后, 图 1(c) 展示了阈值系数对无偏差的多通道多尺度样本熵结果的影响. 如果系数设置过小, 相空间中没有满足阈值条件的向量, $B^M(r)$ 或 $B^{M+p}(r)$ 会出现过小甚至等于 0

的情况. 如果系数设置过大, 熵值结果会趋于稳定. 但是过于宽松的阈值会导致算法无法准确地提取数据的复杂度特征并对不同类型的信号进行区分. 结合仿真实验的结果, 本文认为应在大于 0.1 的条件下选择尽可能小的阈值系数. 这样可以在保证算法稳定的条件下有效地对不同类型信号进行区分.

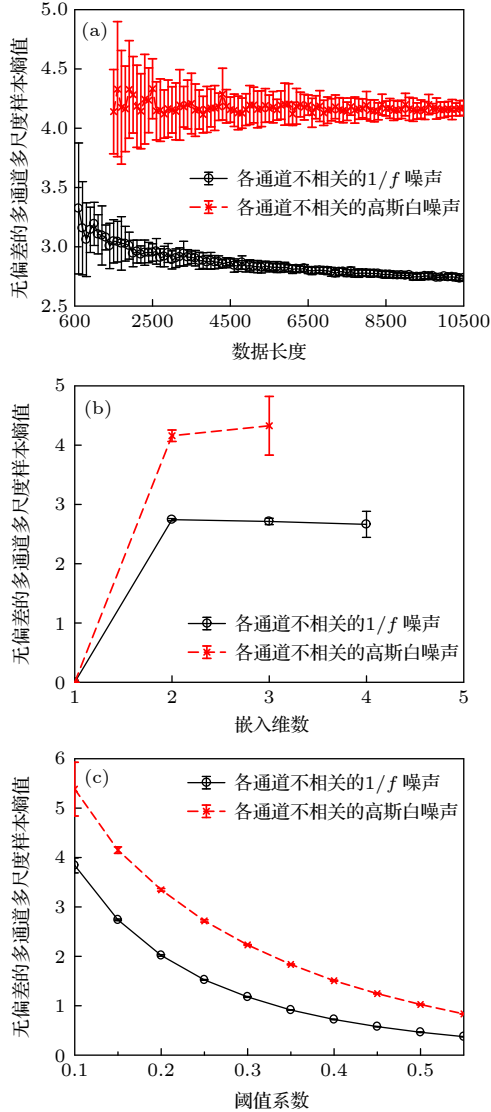


图 1 无偏差的多通道多尺度样本熵算法参数选择 (a) 数据长度; (b) 嵌入维数; (c) 阈值系数

Fig. 1. The parameter selection of the unbiased multivariate multiscale sample entropy algorithm: (a) Data length; (b) embedding dimension; (c) coefficient of the threshold.

第一组仿真实验, 构造一组 3 通道的高斯白噪声. 各通道之间的噪声相互独立. 进一步地, 逐步将 3 个通道的高斯白噪声替换为独立的 $1/f$ 噪声. 由于高斯白噪声和 $1/f$ 噪声的通道内相关性不同, 改变通道中的噪声类型会逐步改变数据通道内的相关性. 参数设置为 $M = [2, 2, 2]$, $\tau = [1, 1, 1]$ 和

$r = 0.15$. 每个通道的数据长度为 10000.

随着包含 $1/f$ 噪声的通道数增加, 多通道数据的通道内相关性增强. 从图 2 可以看出, 改进的多通道多尺度样本熵算法可以有效地提取多通道数据通道内相关性信息.

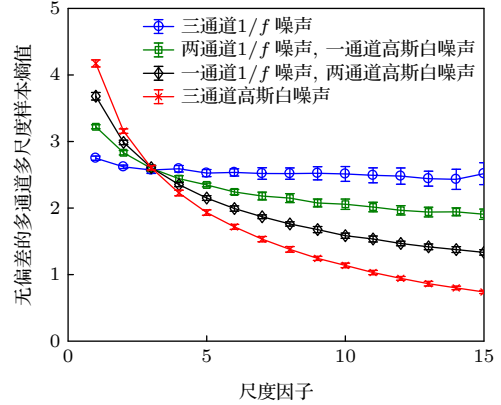


图 2 多通道高斯白噪声和 $1/f$ 噪声的无偏差的多通道多尺度样本熵结果

Fig. 2. Unbiased multivariate multiscale sample entropy results for multivariate white noise and $1/f$ noise.

第二组仿真实验, 构造两组 3 通道数据. 分别为通道间相关的高斯白噪声和通道间不相关的高斯白噪声. 分别计算其对应的朴素多通道多尺度样本熵结果、严格多通道多尺度样本熵结果和改进的多通道多尺度样本熵结果. 参数设置为 $M = [2, 2, 2]$, $\tau = [1, 1, 1]$ 和 $r = 0.12$. 每个通道的数据长度为 10000. 结果如图 3 所示.

从图 3 可以看出, 在不考虑稳定性的前提下严格算法和无偏差的多通道多尺度样本熵算法均可以区分相关和不相关的高斯白噪声. 但是考虑到严格算法的方差过大, 在实际应用中往往会使得对通道间相关性不同的信号出现误判. 而朴素多通道多尺度样本熵无法对两种不同的信号进行区分. 这证明了第 3.5 节(算法分析与证明)理论分析中的结论, 即事件 \mathbf{T} 不能给多通道熵算法带来提取通道间相关性信息的能力. 而不同通道间的联合条件概率可以有效地提取通道间相关性信息. 同时, 这也说明了严格多通道多尺度样本熵算法事实上是有偏的, 在处理实际信号时会使得结果不稳定.

第三组仿真实验, 构造两组 3 通道的高斯白噪声数据. 两组多通道噪声数据中各个通道的方差分别为 $[1, 1, 1]$ 和 $[1, 5, 10]$. 分别使用严格算法和无偏差的多通道多尺度样本熵算法计算其熵值. 结果如图 4 所示.

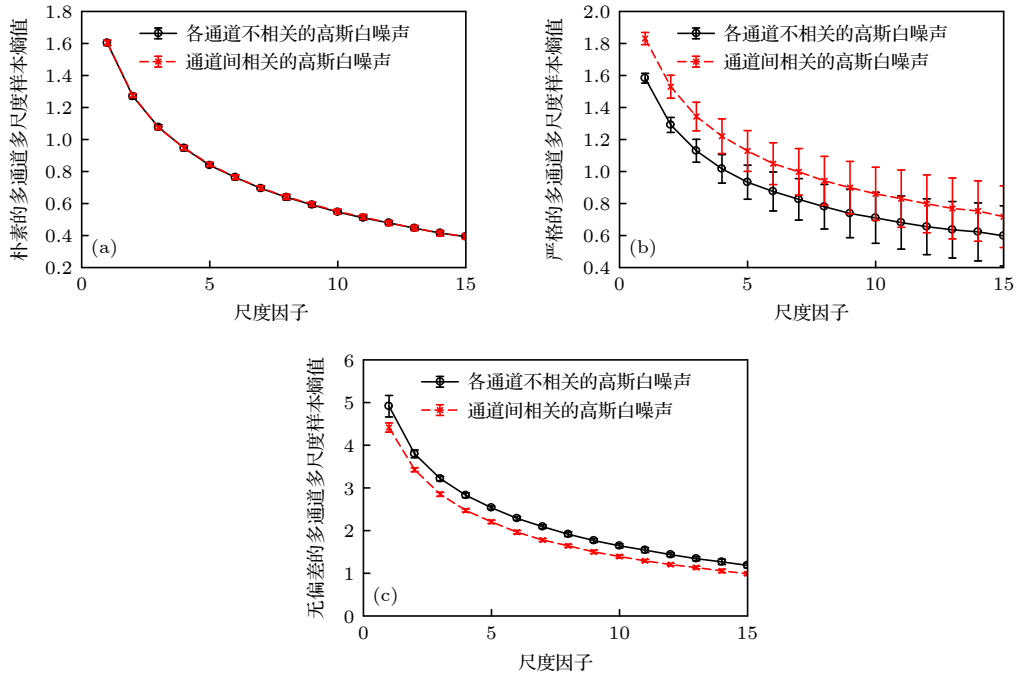


图 3 相关高斯白噪声与不相关高斯白噪声的多通道多尺度样本熵结果 (a) 朴素算法; (b) 严格算法; (c) 无偏差算法

Fig. 3. The entropy results of different MMSE algorithms processing correlated and uncorrelated white noise: (a) Naive MMSE result; (b) rigorous MMSE result; (c) unbiased multivariate multiscale sample entropy results.

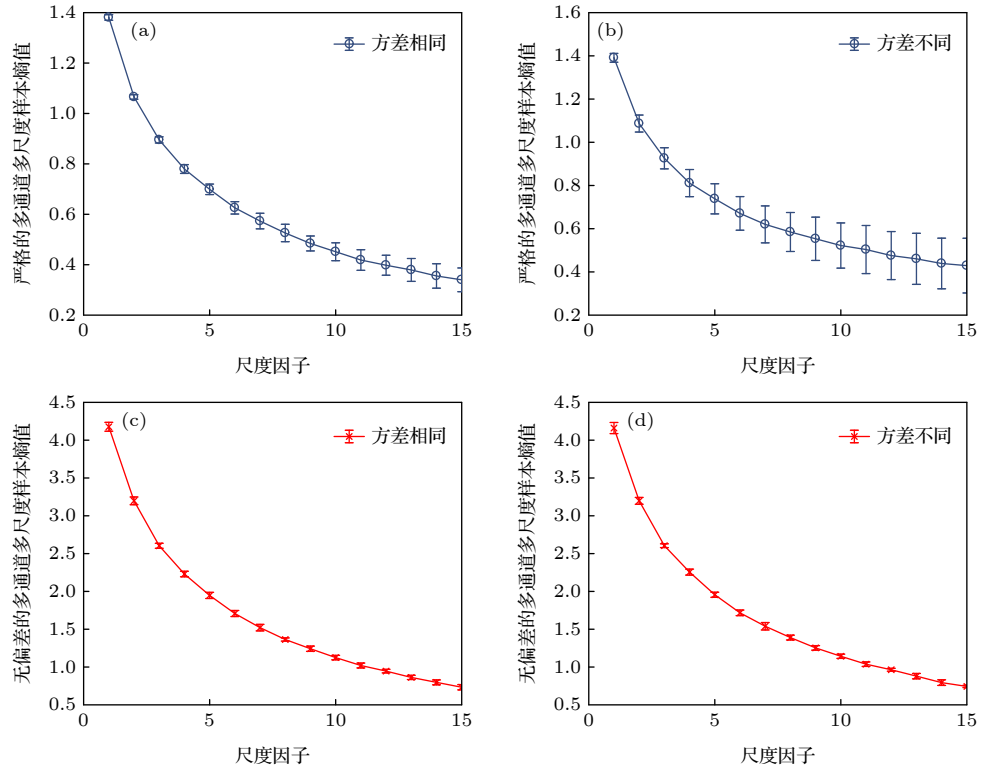


图 4 严格算法与无偏差的多通道多尺度样本熵算法稳定性分析 (a) 通道间方差相同时的严格算法结果; (b) 通道间方差不同时的严格算法结果; (c) 通道间方差相同时的无偏差算法结果; (d) 通道间方差不同时的无偏差算法结果

Fig. 4. The robustness analysis for the different multivariate entropy algorithms: (a) Rigorous MMSE result for the same variance of each channel; (b) rigorous MMSE result for the different variance of each channel; (c) unbiased multivariate multiscale sample entropy results for the same variance of each channel; (d) unbiased multivariate multiscale sample entropy result for the different variance of each channel.

由图4可以看出,从整体上看,严格算法的熵值方差比无偏差的多通道多尺度样本熵的熵值方差更大.并且,当通道间的方差区别较大时,严格算法的方差出现了明显的增大.这也说明了事件 T 在各通道之间方差较大时会使熵值结果不稳定.

在最后一组仿真实验中,本文对比研究了朴素算法、严格算法以及无偏差的多通道多尺度样本熵算法的计算效率.这一组仿真实验所使用的仿真信号为3通道的高斯白噪声.各通道之间相互独立.嵌入维数和时间延迟分别为 $M = [2, 2, 2]$ 以及 $\tau = [1, 1, 1]$.阈值系数为 $r = 0.15$.通过改变数据长度,得到不同的熵算法计算所需的时间,比较不同算法的计算效率.结果如图5所示.

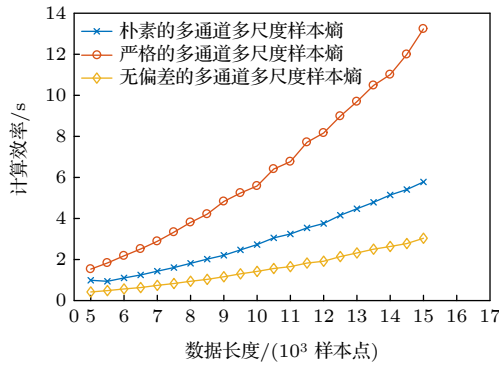


图5 不同多通道熵算法的计算效率对比

Fig. 5. Comparison of computational cost of different multi-channel entropy algorithms.

图5的横坐标是数据长度,纵坐标为计算时长.图5对比研究了数据长度不同时,不同的多通道熵算法计算效率的变化趋势.可以看出,当数据长度较短时3种多通道熵算法计算所需的时间比较接近.随着数据长度的增加,严格算法的计算时长增速要远高于其余两种算法.同时可以观察到,三种算法中无偏差的多通道多尺度样本熵算法的计算时长受数据长度影响最小,计算效率最高.

5 结 论

多通道多尺度样本熵(MMSE)作为近年来提出的一种新算法,能够有效地表征多通道数据的复杂度.然而,现有的朴素算法和严格算法未能兼顾算法性能和稳定性.为了解决这一问题,本文提出了一种无偏差的多通道多尺度样本熵算法.通过概率论的角度对多尺度样本熵、朴素算法、严格算法

以及无偏差的多通道多尺度样本熵算法进行了深入分析.理论分析结果显示,朴素算法未包含不同通道事件的联合条件概率,因此无法有效地提取多通道数据通道间的相关性信息.而严格算法由于考虑了干扰事件,导致计算结果有偏差且稳定性较差.相比之下,无偏差的多通道多尺度样本熵算法是多尺度样本熵在多通道情况下的准确推广.本文通过仿真实验确定了一般条件下无偏差的多通道多尺度样本熵算法的参数范围:计算无偏差多通道多尺度样本熵时的数据长度范围是 $[5000, 10000]$,每个通道的嵌入维数为2.阈值系数的范围应大于0.1.同时,利用仿真实验证明了无偏差多通道多尺度样本熵算法能够有效地提取通道内和通道间的相关性信息.此外,还证明了多通道的联合条件概率有助于熵算法提取通道间相关性信息,而干扰事件 T 则会导致结果有偏差且不稳定.最后,本文对比了不同多通道熵算法的计算效率,结果表明无偏差的多通道多尺度样本熵算法具有最小的计算成本和最高的效率.通过本文的理论推导与仿真实验分析,完善了经典非线性特征量——熵的一般化理论.证明了在增加嵌入维数时,将嵌入维数从 M 变为 $M + p$ 是最优的策略.这也为其他诸如排列熵、模糊熵等算法的一般化提供了理论支撑.

感谢卫卓君对本文的润色和语法检查.

参考文献

- [1] Fowler A C, Gibbon J D, McGuinness M J 1982 *Physica D* **4** 139
- [2] Gao Z, Jin N 2009 *Chaos* **19** 033137
- [3] López C, Naranjo á, Lu S, Moore K 2022 *J. Sound Vib.* **528** 116890
- [4] Wang Q B, Yang Y J, Zhang X 2020 *Chaos, Solitons Fractals* **137** 109832
- [5] Kunze M 2007 *Non-Smooth Dynamical Systems* (Heidelberg: Springer Berlin) pp63–140
- [6] Zarei A, Asl B M 2018 *IEEE J. Biomed. Health* **23** 1011
- [7] He H, Tan Y 2017 *Appl. Soft Comput.* **55** 238
- [8] Liu B Z, Peng J H 2004 *Nonlinear Dynamics* (Beijing: Higher Education Press) pp301–466 (in Chinese) [刘秉正, 彭建华 2004 非线性动力学 (北京: 高等教育出版社) 第301—466页]
- [9] Cranch G A, Nash P J, Kirkendall C K 2003 *IEEE Sens. J.* **3** 19
- [10] Zhu X, Murch R D 2002 *IEEE Trans. Commun.* **50** 187
- [11] Van Trees H L 2002 *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory* (New York: John Wiley and Sons) pp17–230
- [12] Eren F, Pe'eri S, Thein M W, Rzhaznov Y, Celikkol B, Swift M R 2017 *Sensors-Basel* **17** 1741
- [13] Shi Q Q, Li W, Tao R, Sun X, Gao L R 2019 *Remote Sens-Basel* **11** 419

- [14] Xing X W, Ji K F, Zou H X, Chen W T, Sun J X 2013 *IEEE Geosci. Remote Sens. Lett.* **10** 1562
- [15] Wang Z Y, Yao L G, Cai Y W 2020 *Measurement* **156** 107574
- [16] Thuraisingham R A, Gottwald G A 2006 *Physica A* **366** 323
- [17] Li W J, Shen X H, Li Y A 2019 *Entropy-Switz* **21** 793
- [18] Pincus S M 1991 *P. Natl. Acad. Sci. USA* **88** 2297
- [19] Richman J S, Moorman J R 2000 *Am. J. Physiol-Heart. C* **27** H2039
- [20] Bandt C, Pompe B 2002 *Phys. Rev. Lett* **88** 174102
- [21] Costa M, Goldberger A L, Peng C K 2005 *Phys. Rev. E* **71** 021906
- [22] Ahmed M U, Mandic D P 2011 *Phys. Rev. E* **84** 061918
- [23] Li Y, Tang B, Jiao S, Zhou Y 2024 *Chaos, Solitons Fractals* **179** 114436
- [24] Zhao C, Sun J, Lin S, Peng Y 2022 *Measurement* **195** 111190
- [25] Cao L Y, Mees A, Judd K 1998 *Physica D* **121** 75
- [26] Zhang Y C 1991 *J. Phys. I France* **1** 971
- [27] Takens F 1980 *Dynamical Systems and Turbulence* (Heidelberg: Springer Berlin) p366

Unbiased multivariate multiscale sample entropy^{*}

Li Wei-Jia¹⁾²⁾ Shen Xiao-Hong^{1)†} Li Ya-An²⁾

1) (School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China)

2) (Key Laboratory of Ocean Acoustics and Sensing, Ministry of Industry and Information Technology, Northwestern Polytechnical University, Xi'an 710072, China)

(Received 14 July 2023; revised manuscript received 7 April 2024)

Abstract

The development of multi-channel data acquisition techniques has provided richer prior information for studying the nonlinear dynamic characteristics of complex systems. However, conventional nonlinear feature extraction algorithms prove unsuitable in the context of multi-channel data. Previously, the multivariate multiscale sample entropy (MMSE) algorithm was introduced for multi-channel data analysis. Although the MMSE algorithm generalized the multiscale sample entropy algorithm, presenting a novel method for multidimensional data analysis, it remains deficient in theoretical underpinning and suffers from shortcomings, such as missing cross-channel correlation information and having biased estimation results. In this paper, unbiased multivariate multiscale sample entropy algorithm (UMMSE) is proposed. UMMSE increases the embedding dimension from M to $M + p$. This increasing strategy facilitates the reconstruction of a deterministic phase space. By virtue of theoretical scrutiny grounded in probability theory and subsequent experimental validation, this paper illustrates the algorithm's effectiveness in extracting inter-channel correlation information through the integration of cross-channel conditional probabilities. The computation of similarities between sample points across different channels is recognized as a potential source of bias and instability in algorithms. Through simulation experiments, this study delineates the parameter selection range for the UMMSE algorithm. Subsequently, diverse simulation signals are employed to showcase the UMMSE algorithm's efficacy in extracting both within-channel and cross-channel correlation information. Ultimately, this paper demonstrates that the new algorithm has the lowest computational cost compared with traditional MMSE algorithms.

Keywords: nonlinear dynamics, multi-channel signal, multivariate multiscale sample entropy

PACS: 05.45.-a, 89.70.Cf, 43.60.Cg

DOI: 10.7498/aps.73.20231133

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 11874302, 62031021).

[†] Corresponding author. E-mail: xhshen@nwpu.edu.cn

一种无偏差的多通道多尺度样本熵算法

李惟嘉 申晓红 李亚安

Unbiased multivariate multiscale sample entropy

Li Wei-Jia Shen Xiao-Hong Li Ya-An

引用信息 Citation: *Acta Physica Sinica*, 73, 110502 (2024) DOI: 10.7498/aps.73.20231133

在线阅读 View online: <https://doi.org/10.7498/aps.73.20231133>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

光纤探针测量及多尺度熵鉴别超临界类沸腾传热模式

Measurements and identification of supercritical pseudo-boiling heat transfer modes based on fiber optic probes and multiscale entropy

物理学报. 2023, 72(5): 057801 <https://doi.org/10.7498/aps.72.20222060>

基于多尺度熵的Duffing混沌系统阈值确定方法

Threshold determination method of Duffing chaotic system based on multi-scale entropy

物理学报. 2020, 69(16): 160501 <https://doi.org/10.7498/aps.69.20191642>

静息态功能磁共振成像评估健康老年人认知行为的多尺度熵模型研究

Study of multiscale entropy model to evaluate the cognitive behavior of healthy elderly people based on resting state functional magnetic resonance imaging

物理学报. 2020, 69(10): 108703 <https://doi.org/10.7498/aps.69.20200050>

用于多通道单分子定位的高精度图像配准方法

A high-precision image registration method for multi-channel single molecule localization

物理学报. 2024, 73(6): 068701 <https://doi.org/10.7498/aps.73.20231695>

双分子链中非线性多激子态的动力学研究

Theoretical study of nonlinear multi-exciton dynamics in coupled molecular chains

物理学报. 2020, 69(19): 197301 <https://doi.org/10.7498/aps.69.20200104>

开放式多通道多芯少模光纤表面等离子体共振生物传感器

Multi-channel few-mode multicore fiber based surface plasmon resonance biosensor with open air-hole

物理学报. 2020, 69(13): 137802 <https://doi.org/10.7498/aps.69.20200353>