基于贝叶斯估计的神经信息流

李岩 李小隼* 欧阳高翔 关新平

(燕山大学电气工程学院网络控制和生物信息中心, 秦皇岛 066004. *联系人, E-mail: xlli@ysu.edu.cn)

摘要 判断神经网络之间的相互影响是一个重要的神经科学问题. 目前已提出了多种成熟的方法计算神经网络之间的耦合强度, 但是对于神经网络之间耦合方向(信息流)的研究还属于起步阶段. 一般的香农熵计算方法仅仅利用了样本重复概率的统计信息, 而贝叶斯估计则充分利用了整体先验知识和样本重复概率. 基于最小平方误差的贝叶斯估计提出了一种新的基于信息论的相位耦合方向指数计算方法. 经过集总参数神经网络模型所产生的仿真信号检验表明, 提出的方法能够准确地判断两个系统间的耦合方向, 并且减少了对数据长度的依赖性, 使分析短时高噪的复杂生物信号成为可能. 应用该方法分析了癫痫病人临床信号, 结果表明该方法能判断出癫痫发作时各区域之间的影响方向, 并揭示了癫痫传播机制.

关键词 相位同步 耦合方向 条件互信息 贝叶斯估计 癫痫脑电信号

耦合系统之间的同步问题引起了各学科研究人员的广泛关注^[1]. 人们不仅在物理系统中观测到了大量的同步现象, 而且在生物系统中同步现象也广泛存在, 例如心肺之间的相互影响^[2]. 神经科学研究表明, 同步现象在大脑控制人体运动^[3]、睡眠^[4]、学习^[5]乃至意识形成^[6]的过程中有重要作用, 而过同步现象可能是癫痫发作的传播途径^[7]. 对这些系统人们不仅关注同步现象的存在, 更加关注的是各个子系统之间的驱动方向. Rosenblum等学者^[1,2]定义了归一化的方向指数, 基于傅里叶级数的模型映射法(EMA)和瞬时周期法(IPA). Rosenblum等应用EMA和IPA方法分析心肺之间的相互作用, 由于心肺信号相比脑电信号噪声较小, 信号的周期性也很明显, 所以得到了一些很好的结论. 由于EMA法对噪声比较敏感, 因此对噪音较少的心肺信号效果很好, 而脑电信号区别于心肺电信号的主要特点就是高度非平稳和富含噪声, 因此用来分析脑电信号效果并不理想. 针对EMA法统计上有偏的问题, Smirnov等人^[8,9]提出了改进方法和基于相空间的方法. Palus等^[10]提出了基于信息论的相位同步指数分析方法, 简称为IM方法. 该方法的优点在于适合分析非周期信号, 且其结果具有一定的统计意义, 但是该方法要求较长的信号, 才能得到可靠的计算结果. 因而用来分析变化频繁的脑电信号也不是一种理想的方法.

文献[10]中提到, IM方法往往需要比较大的数据量才能够得到比较准确和可靠的结果. 由于在计算的过程中受到信号噪声的影响, IM方法是有偏的估计方法, 而且该估计的方差较大, 随着计算的数据量增加, 误差和方差都随之减小, 这主要是由于该方法是定义在统计意义下的数学方法, 数据量越大, 所描述的过程越平稳, 得出的结果就越准确. 仿真结果表明, 该方法在数据达到 10^4 长度以上的情况下, 可以判断耦合混沌模型之间的相位耦合方向; 但是在数据长度低于这个数量级的时候误差较大. 为了克服IM方法这一固有的缺点, 本文提出了基于贝叶斯估计的改进算法, 因为贝叶斯估计是基于最小平方误差^[11], 所以可以减小在计算熵值过程中的误差, 从而提高方法的有效性. 通过仿真和实际临床信号证实, 该研究提供了切实可行的方法估计神经网络间的信息流.

1 方法

1.1 信号的相位获取

用相位同步方法来分析信号, 第一步是要从信号中获得准确的相位信息, 获得相位的方法一般有三种^[12], 分别适用于不同的情况. 本文利用希尔伯特变换来获得脑电的相位, $s_H(t)$ 是信号 $s(t)$ 的希尔伯特变换, 则信号的瞬时相位 $\phi(t)$ 和幅值 $A(t)$ 可以由下

式获得:

$$\zeta(t) = s(t) + is_H(t) = A(t)e^{i\phi(t)}. \quad (1)$$

由于希尔伯特变换不需要调整参数, 所以应用比较广泛. 当然这种方法也有一个限制, 就是仅适用于窄频带的信号. 对于宽频带的信号, 需要划分为几组较窄频带的信号, 再进行希尔伯特变换, 以得到可靠的相位. 由于癫痫病人的脑电信号是宽频信号, 因此本文参照文献[13]的方法首先对信号采用 bior 小波变换, 在选择所要分析频段时参考了文献[14]中的分析结果, 主要取 8~16 Hz 频段的信号进行分析.

1.2 基于信息论的相位同步方向指数

在计算相位同步方向指数之前, 应先计算相位同步强度, 因为方向指数是建立在强度指数的基础上. 首先两路信号必须存在同步, 才能考虑两路信号之间的影响方向. 通常用下式来估计同步强度:

$$Syn_{ab} = \left| \frac{1}{N} \sum_{j=1}^{N-1} \exp(i(\phi_a(j\Delta t) - \phi_b(j\Delta t))) \right|, \quad (2)$$

其中 N 是离散时间序列的长度, $1/\Delta t$ 是采样频率, Syn 值介于 0 和 1 之间. 当 Syn 值等于 0 时表示两路信号之间不存在同步; 当 Syn 值趋向于 1 时, 表示两路信号之间存在比较强的同步.

基于香农熵 [15,16] 的条件互信息理论, Palus 提出一种基于信息论的方法计算相位同步方向指数. 联合条件互信息定义如下:

$$I(X_1 \cap X_2 | X_3) = H(X_1 | X_3) + H(X_2 | X_3) - H(X_1, X_2 | X_3), \quad (3)$$

式中 X_1, X_2, X_3 是三个随机信号, $X_{1,2,3} = x_{1,2,3}(t_i)$, $t_i = i\Delta t$, $i = 1, \dots, N_x$, Δt 是采样时间. 考虑 X 和 Y 两个过程, 通过 Hilbert 变换得到 $[0, 2\pi]$ 或者 $[-\pi, \pi]$ 之间的相位 $\{\phi_{x,y}(t_i)\}_{i=1}^{N_\phi}$, 通过计算离散随机变量的联合事件概率和条件事件概率, 利用(3)式计算条件互信息:

$$\begin{aligned} I_{xy} &= I(\phi_x(t) \cap \Delta\phi_y(t+\tau) | \phi_y(t)), \\ I_{yx} &= I(\phi_y(t) \cap \Delta\phi_x(t+\tau) | \phi_x(t)), \end{aligned} \quad (4)$$

式中 I_{xy} 表示在已知随机过程 $\{Y(t)\}$ 的概率分布的条件下, $\{X(t)\}$ 在时间间隔 τ 内与 $\{Y(t)\}$ 的净信息, 可以看成在时间间隔 τ 内, 随机过程 $\{X(t)\}$ 对 $\{Y(t)\}$ 的影响; 相应的 I_{yx} 是已知随机过程 $\{X(t)\}$ 的概率分布的条件下, 过程 $\{Y(t)\}$ 在时间间隔 τ 内对过程 $\{X(t)\}$ 的净信息. X 和 Y 两过程的相位同步方向指数定义如下:

$$IM(x, y) = \frac{I_{xy} - I_{yx}}{I_{xy} + I_{yx}} \quad (5)$$

这里由于条件互信息的性质, I_{xy} 和 I_{yx} 是大于等于零的(当随机过程 $\{X(t)\}$ 和 $\{Y(t)\}$ 完全不相相关时, 两随机过程之间的条件互信息 I_{xy} 和 I_{yx} 等于零), 所以方向指数 $IM(x, y)$ 的值域为 $[-1, 1]$. 即当 $IM(x, y) > 0$ 时, 表示随机过程 $\{X(t)\}$ 驱动 $\{Y(t)\}$, 而当 $IM(x, y) < 0$ 时, 表示随机过程 $\{X(t)\}$ 被 $\{Y(t)\}$ 驱动, $IM(x, y) = 0$ 时表示两个随机过程的相互影响对称相等.

1.3 贝叶斯条件互信息方法

基于信息论的方向指数方法需要相当长的数据量, 如果数据长度较短或是存在较多噪声, 该方法的计算结果就不稳定. 如(3)式所示, 联合条件互信息可以转化为计算两个一维条件熵和一个二维联合条件熵. 而一维条件熵又可以通过一维信息熵和二维联合熵来进行计算; 而二维联合条件熵可以通过二维信息熵和三维联合熵来进行计算. 文献[11]提出了一维贝叶斯熵的改进算法并获得了比传统香农熵更好的计算结果. 本文首先将二维联合条件展开为二维和三维联合信息熵, 参照文献[11]中一维贝叶斯熵的改进方法, 并将二维联合信息熵和三维联合信息熵拓展为贝叶斯熵意义下的联合信息熵(具体推导见网络版附录), 最后在此基础上定义贝叶斯条件互信息耦合方向指数:

$$IM_{bys}(x, y) = \frac{\hat{I}_{xy} - \hat{I}_{yx}}{\hat{I}_{xy} + \hat{I}_{yx}}. \quad (6)$$

由于条件互信息需要计算二维和三维随机变量的联合信息熵, 因此计算贝叶斯熵过程中的分类情况较多, 计算量大, 本文应用泰勒级数展开的快速算法, 在保证精度的情况下, 提高算法的运算速度, 从而能够满足实时分析和处理脑电信号.

2 集总参数神经网络模型仿真结果

为了验证基于贝叶斯估计的耦合方向指数方法的有效性和精度, 仿真分析集总参数神经网络模型输出的癫痫样脑电信号, 并与 IM 方法进行比较.

2.1 集总参数神经网络数据模型

本文所采用的集总参数神经网络模型可以生成癫痫样信号模拟脑电数据 [17~20]. 该模型最重要的三个集总参数分别是神经网络兴奋指数 A 、抑制指数 B 和各神经网络之间的连接强度 K , 用来模拟神经网络中兴奋和抑制的平衡, 以及各神经网络之间的耦合方向和同步强度. 在每个神经网络模型中, 动态线性传递函数

$h_e(t) = u(t)Aate^{-at}$ 和 $h_i(t) = u(t)Bbte^{-bt}$ 分别用来模拟神经元的兴奋和抑制状态^[21], 该模型中各参数的详细意义可参考文献[20]. 当神经元被激发时, 将平均膜电位转化为平均脉冲密度电位, 这一电位沿着轴突将兴奋传导至其他神经网络. 三个串连的神经网络产生脑电信号. 开始的时候令三个神经网络的参数处于平衡状态, 不产生癫痫现象, 兴奋参数 $A = 3.25$, 抑制参数 $B = 22$; 持续 80 s 之后, 令第一个神经网络的兴奋参数 A 增加到 3.585, 其他两个神经网络的兴奋参数和抑制参数都保持不变, 此时神经网络 1 开始产生连续不断的尖波, 持续到 150 s 后, 兴奋参数 A 恢复为 3.25, 250 s 时再次增加参数 A 到 3.585, 神经网络 1 再次输出尖波, 如图 1(a) 所示. 为了验证所提出方法的有效性, 本文分别采用了不同的连接强度 K_{12} 和 K_{23} , 令 K_{12} 等于常数 200, K_{23} 采用均值为 130、方差为 50 的高斯分布随机信号. 这样设置使神经网络 1 和 2 之间的耦合较强, 可以保证神经网络 1 的所有兴奋都可以传到神经网络 2 上去(如图 1(b) 所示). 由于 K_{23} 的取值是一个随机变化的高斯分布信号, 神经网络 2 上面的兴奋会随机地传导到神经网络 3, 当连接强度比较小的时候, 神经网络 3 的兴奋有可能没有被神经网络 2 激发, 也可能被激发的尖波幅值较低(如图 1(c) 所示). 前面两个神经网络之间的连接属于理想状态; 后面两个神经网络之间属于耦合比较弱的情况, 尽可能接近大脑中的实际情况.

图1 三个串联的集总参数神经元群模型产生的模拟类癫痫信号

(a) 第一路信号; (b) 第二路信号; (c) 第三路信号

2.2 实验检验

在计算耦合方向的过程中, 时间延迟 τ 取为信号的平均周期, 经过计算本文在分析仿真信号时 τ 选为 100. 为了分析本文提出方法的有效性, 采用移动窗口方法逐段计算神经网络输出信号间的同步关系, 为了比较数据长度对方向指数计算结果的影响, 本文分三种情况进行讨论.

每次计算方向指数取 39 s 长度的神经网络信号 (10000 个点); 为了克服边缘效应, 每次选取的移动窗口有 50% 的重叠. 如图 2 所示, IM 指数在判断神经网络 1 和 2 之间较强耦合的时候, 效果尚可; 但是对于判断弱耦合的情况, 即分析神经网络 2 和 3 之间的耦合方向时, 出现误判, 把没有耦合情况判断为前者影响后者. 而本文提出的改进算法 IM_{bys} 很好地判断出三个神经网络之间的耦合方向, 尤其是对于相互之间没有影响的阶段(图中没有阴影的部分), 判断非常准确, 所有的 IM_{bys} 指数都等于零. 如图 2(c) 所示, 相位同步强度指数在发生癫痫时下降到 0.8, 如文献[22]中所提到的, 人们往往利用这一去同步过程进行癫痫预测, 而其他方法, 例如相似性^[23]、分形谱^[24]和 RQA^[25]等也发现了癫痫发作前期的特征. 本文 LPCC 模型很好地反映了这一点.

移动窗口取 19 s 的神经网络信号 (5000 个点), 每次有 50% 的重叠. 如图 3(a) 所示, 当数据长度下降到 10000 以下时, IM 方法的可靠性大大地下降了, 对神经网络 1 和 2 之间较强的耦合, 还有一部分指数可以

图2 移动窗口的长度为 39 s 时, 耦合方向和同步强度的模型脑电信号分析结果

(a) IM 指数; (b) IM_{bys} 指数; (c) 同步强度指数

图 3 移动窗口的长度为 19 s 时, 耦合方向和同步强度的模型脑电信号分析结果

(a) IM 指数; (b) IM_{bys} 指数; (c) 同步强度指数

判断出来, 对神经网络 2 和 3 之间的弱耦合情况, 方法出现了误判, 这主要是由于方法的灵敏度不够造成的. 而本文提出的改进方法对较强的神经网络耦合做出了正确的判断, 在癫痫发作期所有的贝叶斯方向指数都大于零, 在癫痫发作间歇期所有的贝叶斯方向指数都等于零.

移动窗口取 4 s 的神经网络信号(1000 个点), 每次有 50% 的重叠(图 4). 由图 4(a)可知, 原信息论方向指数无法在移动窗口长度较小的情况下判断神经网络之间的相互作用, 本文提出的贝叶斯方向指数在极短的数据窗口中仍然可以很好地判断模型间的相

图 4 移动窗口的长度为 4 s 时, 耦合方向和同步强度的模型脑电信号分析结果

(a) IM 指数; (b) IM_{bys} 指数; (c) 同步强度指数

互作用(图 4(b)), 大大地提高了算法的实时性. 结果表明, IM 方法在这样短的数据长度下基本失效了, 因为数据量太少, 结果不具有统计意义. 本文提出的方法依然可以判断出正确的耦合方向, 由于数据长度的影响, 方向指数的数值都比前两个实验中的方向指数数值要小.

3 癫痫发作期两路脑电信号之间的信息流

3.1 脑电信号的采集

本文分析的脑电信号由德国 Freiburg 大学的医学院癫痫中心提供, 采集癫痫病人的颅内脑电信号, 采样频率为 256 Hz, 16 位模数转换. 为了获得较高的信噪比, 减少虚假动作电位, 成网格状或条状在发作区域直接放置电极或者深度电极. 本文所分析的数据是在大脑中海马区深度电极采集的脑电信号, 由一排三个电极采集的三路脑电信号. 癫痫发作时的信号如图 5 所示.

图 5 大脑中海马区深度电极采集的脑电信号, 电极为条状分布的三个并排电极

(a) 1 号电极采集的脑电信号; (b) 2 号电极采集的脑电信号; (c) 3 号电极采集的脑电信号. 阴影部分标出的是癫痫病发作的时期

3.2 耦合方向指数的比较

首先用 bior 小波对脑电信号进行分解, 取 8 ~ 16 Hz 的频段进行分析. 通过希尔伯特变换获得脑电信号这一频道的相位信息后, 然后分别用两种方法, IM 指数和 IM_{bys} 指数对三路脑电信号之间的耦合方向进行分析. 为了保证结果的有效性和可信度, 两种方法分别取 390 s 的数据(10 万个数据点), 包括癫痫发作前 160 s, 发作时 150 s, 发作后 80 s 的数据. 每次计算移动窗口取 10 s 的数据(2500 个点)来计算相位方

向指数.

如图 6(c) 中同步强度指数所示, 在癫痫发作期 (图中阴影部分), 同步指数有一个明显的下降. 在大脑中靠近病灶的区域, 其脑电波在开始发作时的震荡频率和强度高于病灶周围的区域. 表现在采集的脑电信号上就是一路信号局部震荡的强度和频率增高, 而非病灶区域的脑电信号的震荡达不到这种强度, 从震荡理论的角度来说, 就是两路信号去同步的过程. 癫痫发作之后, 同步指数又恢复到了一个比较正常的水平. 由图 6~8 的 IM 指数和 IM_{bys} 指数所示, 在癫痫发作以前, 两路信号之间的同步属于一种平衡状态, 即两个区域之间的相互作用基本是平衡的, 图 6~8(a) 中 IM 指数偏离了期望的零轴, 出现了误判; 而图 6~8(b) 中 IM_{bys} 同步方向指数的数值误差非常小, 没有像 IM 指数图中出现大的波动. 而在癫痫发作期, IM 指数没有明确地判断出两路脑电信号之间相位信息的同步方向, 结果在 $[-1, 1]$ 之间振荡, 而本文的新方法辨别出了两路脑电信号之间的相位同步方向. 另外如图 6(b) 和 8(b) 中的癫痫发作部分所示, 第一路脑电信号对第二路脑电信号的影响强度要稍大于第一路对第三路的影响强度, 而从图 7 可以看出, 第二路信号对第三路信号的影响在癫痫发作期间更强. 从这三组图总地来看, IM 指数无论在发作前还是在发作中, 指数变化的方差都较大, 而 IM_{bys} 指数在发作前指数没有大的波动; 在癫痫发作期间, IM_{bys} 指数明显地判断出来两路信号之间的影响方向, 可信度要明显高于 IM 指数.

文献[2]采用 EMA 方法对三个单向耦合的 Van

图 6 癫痫病人第一路和第二路脑电信号耦合方向和同步强度的分析结果

(a) IM 指数; (b) IM_{bys} 指数; (c) 同步强度指数

图 7 癫痫病人第二路和第三路脑电信号耦合方向和同步强度的分析结果

(a) IM 指数; (b) IM_{bys} 指数; (c) 同步强度指数

图 8 癫痫病人第一路和第三路脑电信号耦合方向和同步强度的分析结果

(a) IM 指数; (b) IM_{bys} 指数; (c) 同步强度指数

Der Pol 混沌系统进行了拓扑结构研究. 文献[26]基于频率分解的条件因果检验方法对多变量三维自回归过程的拓扑结构进行了分析. 文献[27]采用 PDC 方法 (partial directed coherence) 分别对五维自回归过程、耦合的四个 Van Der Pol 混沌系统以及特发性震颤 (essential tremor) 进行了拓扑结构研究. 为了更加简明地说明在癫痫发作过程前后大脑中各区域之间的相互影响方向, 根据图 6~8 中的耦合方向的计算结果, 本文给出图 9 中的拓扑结构图. 如图 9(a) 所示, 取第 100 s 时刻的方向指数作为癫痫发作以前的代表值, 三个区域之间没有明显的相互驱动关系, 所有的耦合方向值都约等于零; 如图 9(b) 所示, 取癫痫发作时

的最大方向指数, 即第 280 s 时刻的方向指数作为癫痫发作时的代表值, 第一路信号对另外两路信号都产生了明显的影响, 因而可以推断第一路信号所在的区域更接近于癫痫发作的病灶区域, 而癫痫发作的过程是一个由病灶区域向外扩散的过程, 特别需要说明的是, 本分析结果仅表明这个病人在此次发作过程中三个区域的影响变化, 而不同的病人每次癫痫发作时的传播路线和传播方式可能是不同的。

图 9 第一路信号对其他两路信号的影响方向

(a) 癫痫发作间歇期; (b) 癫痫发作时

4 结论

针对 IM 方法对数据长度过于依赖的问题, 提出了基于贝叶斯估计的新方法。Palus 等 [10] 提出的 IM 方法只有在计算较长的数据时, 结果才具有比较好的统计特性, 能够比较准确地辨别两个模型之间的耦合方向。本文的仿真结果也证明了这一点, IM 方法对数据的长度比较敏感。本文采用基于贝叶斯估计的新方法, 充分地利用了数据中的各种信息, 采用比较短的数据得到了可靠的结果。在所取数据均为 39 s 的情况下, 两种方法都能够分辨出两个神经网络模型之间的相互作用方向, 但是 IM 方法的波动较大, 较大的数值波动使该方法的可靠性大大下降。而数据长度较短的情况下, 本文分别就 19 s 和 4 s 两种情况对两种方法进行了分析, 结果表明, IM 方向指数在数据长度较短的情况下已经完全失去了其方法的统计意义, 无法对两个神经网络之间的耦合方向做出正确判断; 而本文提出的方法甚至在非常短的数据(4 s) 情况下, 也能够较好地分辨出两个神经网络之间的同步方向。另外本文采用三个神经网络串联的方法, 尤其对神经网络之间连接较强和较弱两种情况分别进行讨论, 结果表明本文的方法对这两种情况都做

出了比较准确的判断。通过对癫痫神经网络模型模拟信号的分析, 本文提出的基于贝叶斯估计的条件互信息耦合方向指数方法, 在灵敏度和准确性等方面明显优于原有的基于信息论的方向指数方法。

最后对癫痫病人的颅内脑电信号进行了分析。分析结果表明, 在计算复杂的实际脑电信号时, 由于各种噪声的影响, IM 方法无法得出可靠的结论, 而本文提出的新方法准确地判断出了在癫痫发作时脑电信号之间的同步方向, 因此提出的贝叶斯方向指数的鲁棒性要优于原有的信息论方向指数。经过对大量脑电数据的分析, 通过对比相位信息方向指数的分析结果, 发现距离较近的两路信号之间的同步强度要大于距离较远的两路信号之间的同步强度。这说明癫痫发作时距离对癫痫波在大脑内传播的影响, 距离癫痫病灶越远的区域所受的病灶区的影响就越小, 相位同步强度随着电极距离的增加而减小, 而模型间的作用方向也是从病灶附近的神经元群指向外围的神经元群, 从另外一个角度也证明了同步现象是癫痫传播的重要途径。

参 考 文 献

- 1 Rosenblum G M, Pikovsky A. Detecting direction of coupling in interacting oscillators. *Phys Rev E*, 2001, 64: 045202[DOI]
- 2 Rosenblum G M. Identification of coupling direction: Application to cardio respiratory interaction. *Phys Rev E*, 2002, 65: 041909 [DOI]
- 3 Roelfsema P R, Engel A K, König P, et al. Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature*, 1997, 385: 157—161[DOI]
- 4 Steriade M, McCormick D A. Thalamocortical oscillations in the sleeping and aroused brain. *Science*, 1993, 262: 679—685[DOI]
- 5 Miltner W H, Braun C, Arnold M, et al. Coherence of gamma-band EEG activity as a basis for associative learning. *Nature*, 1999, 397: 434—436[DOI]
- 6 Rodriguez E, George N, Lachaux J P, et al. Perception's shadow: Long-distance synchronization of human brain activity. *Nature*, 1999, 397: 430—433
- 7 Niedermeyer E, Fernando L S. *Electroencephalography: Basic Principles, Clinical Applications and Related Fields*. Baltimore: Williams & Wilkins, 1993. 1097
- 8 Smirnov D A, Andrzejak R G. Detection of weak directional coupling: Phase-dynamics approach versus state-space approach. *Phys Rev E*, 2005, 71: 036207[DOI]
- 9 Smirnov D A, Bezruchko B P. Estimation of interaction strength and direction from short and noisy time series. *Phys Rev E*, 2003, 68: 046209[DOI]
- 10 Palus M, Stefanovska A. Direction of coupling from phases of interacting oscillators: An information-theoretic approach. *Phys Rev*

- E, 2003, 67: 055201[DOI]
- 11 Yokota Y. An approximate method for Bayesian entropy estimation for a discrete random variable. In: Proceedings of the 26th Annual International Conference of the IEEE EMBS, San Francisco, USA, 2004. 99—102
 - 12 Rosenblum G M, Pikovsky A. Phase Synchronization: From Theory to Data Analysis. Handbook of Biological Physics. Vol 4, Neuro-informatics. Elsevier Science, 2001. 279—321
 - 13 Li X L, Yao X, John R G, et al. Interaction dynamics of neuronal oscillations with wavelet. J Neurosci Meth, 2007, 160(1): 178—185 [DOI]
 - 14 Li X L. Temporal structure of neuronal population oscillations with empirical model decomposition. Phys Lett A, 2006, 356(3): 237—241 [DOI]
 - 15 Palus M. Synchronization as adjustment of information rates: Detection from bivariate time series. Phys Rev E, 2001, 63: 046211 [DOI]
 - 16 Cover T M, Thomas J A. Elements of Information Theory. New York: John Wiley & Sons, 1991
 - 17 Lopes da Silva F H, Hoek A, Smith H, et al. Model of brain rhythmic activity. Kybernetik, 1974, 15: 27—37
 - 18 Jansen B, Rit V G. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biol Cybern, 1995, 73(4): 357—366
 - 19 Jansen B H, Zouridakis G, Brandt M E. A neurophysiologically-based mathematical model of flash visual evoked potentials. Biol Cybern, 1993, 68(3): 275—283[DOI]
 - 20 Wendling F, Bellanger J J, Bartolomei F, et al. Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals. Biol Cybern, 2000, 83(4): 367—378[DOI]
 - 21 Li Y, Li X L, Ouyang G X, et al. Strength and direction of phase synchronization of neural networks. In: International Symposium on Neural Networks, Chongqing, China, 2005. 314—319
 - 22 Mormann F, Andrzejak R G, Kreuz T. Automated detection of a pre-seizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients. Phys Rev E, 2003, 67: 021912[DOI]
 - 23 Li X L, Ouyang G X. Nonlinear similarity analysis for epileptic seizures prediction. Nonlinear Anal-Theor, 2006, 64(8): 1666—1678[DOI]
 - 24 Li X L, Kapiris P G, Polygiannakis J, et al. Fractal spectral analysis of pre-epileptic seizures phase: In terms of criticality. J Neural Eng, 2005, 2: 11—15[DOI]
 - 25 Li X L, Ouyang G X, Yao X, et al. Dynamical characteristics of pre-epileptic seizures in rats with recurrence quantification analysis. Phys Lett A, 2004, 333: 164—171[DOI]
 - 26 Chen Y H, Bressler S L, Ding M Z. Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data. J Neurosci Meth, 2006, 150: 228—237 [DOI]
 - 27 Schelter B, Winterhalder M, Eichler M, et al. Testing for directed influences among neural signals using partial directed coherence. J Neurosci Meth, 2005, 152: 210—219[DOI]

附录

本文以二维条件熵为例, 给出得到贝叶斯熵的过程, 三维的联合条件熵可以由二维的条件熵推广得到.

一般的二元随机变量条件熵定义为

$$H(X|Y) = \text{ent}(\mathbf{r}|s) = \text{ent}(\mathbf{r}_{x,y}) - \text{ent}(s_y), \quad (\text{A1})$$

式中 X, Y 表示两个随机过程, $\mathbf{r}$ 为二维联合概率密度, s 为一维边缘概率密度.

贝叶斯意义下的条件熵定义 [11] 为

$$H_{\text{bys}}(X|Y) = \frac{E}{s|\mathbf{n}}(\text{ent}(\mathbf{r}|s)) = \iint [\text{ent}(\mathbf{r})p(\mathbf{r}|\mathbf{n}) - \text{ent}(s)p(s|\mathbf{n})]d\mathbf{r}ds, \quad (\text{A2})$$

式中 $E_{r|\mathbf{n}}[\bullet]$ 表示已知后验联合概率密度 $p(\mathbf{r}|\mathbf{n})$ 和边缘概率密度 $p(s|\mathbf{n})$ 的情况下对 $\bullet$ 的贝叶斯估计, $p(\mathbf{r}|\mathbf{n})$ 和 $p(s|\mathbf{n})$ 可以由下式获得:

$$p(\mathbf{r}|\mathbf{n}) = \frac{p(\mathbf{n}|\mathbf{r})p(\mathbf{r})}{p(\mathbf{n})} = \frac{p(\mathbf{n}|\mathbf{r})p(\mathbf{r})}{\int p(\mathbf{n}|\mathbf{r})p(\mathbf{r})d\mathbf{r}},$$

$$p(s|\mathbf{n}) = \frac{p(\mathbf{n}|s)p(s)}{p(\mathbf{n})} = \frac{p(\mathbf{n}|s)p(s)}{\int p(\mathbf{n}|s)p(s)ds}, \quad (\text{A3})$$

如果不知道先验概率 $\mathbf{r}$ 和 s 的概率分布的先验知识, 则可以假设先验概率 $\mathbf{r}$ 和 s 的先验概率密度分布 $p(\mathbf{r})$ 和 $p(s)$ 为一个常数 [11].

$$p(\mathbf{r}) = V, \forall \mathbf{r} \in R_r \equiv \left\{ \mathbf{r} \mid \sum_{i=1}^p \sum_{j=1}^q r(i, j) = 1, r(i, j) \geq 0 \right\},$$

$$p(s) = C, \forall s \in R_s \equiv \left\{ s \mid \sum_{j=1}^q s(j) = 1, s(j) \geq 0 \right\}, \quad (\text{A4})$$

将(A4)式引入(A3)式得

$$p(\mathbf{n}|\mathbf{r}) = \frac{p(\mathbf{n}|\mathbf{r})}{\int p(\mathbf{n}|\mathbf{r})d\mathbf{r}}, \quad p(s|\mathbf{n}) = \frac{p(\mathbf{n}|s)}{\int p(\mathbf{n}|s)ds}. \quad (\text{A5})$$

则概率事件发生次数 $\mathbf{n}$ 服从下面的多项式概率分布:

$$p(\mathbf{n}|\mathbf{r}) = P_{N,\mathbf{r}}(\mathbf{n}) = \frac{N!}{(n_{1,1}!) \cdots (n_{p,q}!)} r_{1,1}^{n_{1,1}} \cdots r_{p,q}^{n_{p,q}},$$

$$p(\mathbf{n}|s) = P_{N,s}(\mathbf{n}) = \frac{N!}{(n_1!) \cdots (n_q!)} s_1^{n_1} \cdots s_q^{n_q}. \quad (\text{A6})$$

根据 Dirichlet 分布的性质, 得到

$$\int r_{1,1}^{n_{1,1}} \cdots r_{p,q}^{n_{p,q}} d\mathbf{r} = \frac{(n_{1,1}!) \cdots (n_{p,q}!)}{(pq-1 + \sum_{i=1}^p \sum_{j=1}^q n_{i,j})!},$$

$$\int s_1^{n_1} \cdots s_q^{n_q} ds = \frac{(n_1!) \cdots (n_q!)}{\left(q-1 + \sum_{j=1}^q n_j \right)!}. \quad (\text{A7})$$

由(A4)和(A6)式可得, (A3)式的分母可以表示为

$$\int p(\mathbf{n}|\mathbf{r})d\mathbf{r} = \frac{N!}{(n_{1,1}!) \cdots (n_{p,q}!)} \int r_{1,1}^{n_{1,1}} \cdots r_{p,q}^{n_{p,q}} d\mathbf{r} = \frac{N!}{(pq-1+N)!},$$

$$\int p(\mathbf{n}|s)ds = \int \frac{N!}{(n_1!) \cdots (n_q!)} s_1^{n_1} \cdots s_q^{n_q} ds = \frac{N!}{(q-1+N)!}. \quad (\text{A8})$$

将(A6)和(A8)代入(A5)式, 后验联合概率密度 $p(\mathbf{r}|\mathbf{n})$ 和边缘概率密度 $p(s|\mathbf{n})$ 可以表示为

$$p(\mathbf{r}|\mathbf{n}) = \frac{(pq-1+N)!}{N!} p(\mathbf{n}|\mathbf{r}) = \frac{(pq-1+N)!}{(n_{1,1}!) \cdots (n_{p,q}!)} r_{1,1}^{n_{1,1}} \cdots r_{p,q}^{n_{p,q}},$$

$$p(s|\mathbf{n}) = \frac{(q-1+N)!}{(n_1!) \cdots (n_q!)} s_1^{n_1} \cdots s_q^{n_q}. \quad (\text{A9})$$

因此贝叶斯条件熵可以定义为

$$\hat{H}_{\text{bys}}(X|Y) = \iint \left[\text{ent}(\mathbf{r}) \frac{(pq-1+N)!}{(n_{1,1}!) \cdots (n_{p,q}!)} r_{1,1}^{n_{1,1}} \cdots r_{p,q}^{n_{p,q}} - \text{ent}(s) \frac{(q-1+N)!}{(n_1!) \cdots (n_q!)} s_1^{n_1} \cdots s_q^{n_q} \right] d\mathbf{r}ds. \quad (\text{A10})$$

计算(A10)式的多重定积分非常困难, (A6)式中幂级数的计算比较方便, 因此, 采用对熵函数 $\text{ent}(\mathbf{r})$ 和 $\text{ent}(s)$ 进行多元泰勒级数展开成幂级数的方法进行近似计算 [11], 具体如下:

$$\text{ent}(\mathbf{r}) = -\sum_{i=1}^p \sum_{j=1}^q r_{i,j} \log(r_{i,j})$$

$$= \text{ent}(\mathbf{r}^0) - \sum_{i=1}^p \sum_{j=1}^q \left[(1 + \log(r_{i,j}^0))(r_{i,j} - r_{i,j}^0) \right.$$

$$\left. + \sum_{k=2}^{\infty} \frac{(-1)^k}{k(k-1)} (r_{i,j}^0)^{-k+1} (r_{i,j} - r_{i,j}^0)^k \right],$$

$$\text{ent}(s) = \sum_{j=1}^q s_j \log(s_j) = \text{ent}(s^0)$$

$$- \sum_{j=1}^q \left[(1 + \log(s_j^0))(s_j - s_j^0) + \sum_{k=2}^{\infty} \frac{(-1)^k}{k(k-1)} (s_j^0)^{-k+1} (s_j - s_j^0)^k \right], \quad (\text{A11})$$

这里 $\mathbf{r}^0$ 和 s^0 表示向量 $\mathbf{r}$ 和 s 的泰勒级数在 $\mathbf{r}^0$ 和 s^0 附近进行展开, 推导表明, 为了使熵函数的泰勒展开式收敛, 应该使 $\mathbf{r}^0 = 1/2$ 和 $s^0 = 1/2$. 这里令 $r_{i,j}^0 = 1/2, i=1 \cdots p, j=1 \cdots q$ 和 $s_j^0 = 1/2, j=1 \cdots q$, 将(A11)代入(A10)式, 得到贝叶斯条件熵的近似计算公式如下:

$$\hat{H}_{\text{bys}}(X|Y)$$

$$= \sum_{i=1}^p \sum_{j=1}^q r_{i,j}^0 - \sum_{j=1}^q s_j^0 - \frac{1}{N+pq} \sum_{i=1}^p \sum_{j=1}^q (1 + \log(r_{i,j}^0))(n_{i,j} + 1)$$

$$+ \frac{1}{N+q} \sum_{j=1}^q (1 + \log(s_j^0))(n_j + 1)$$

$$+ \sum_{k=2}^K \frac{1}{k(k-1)} \sum_{l=0}^k \binom{k}{l} \frac{(N+pq-l)!}{(N+pq+l-1)!}$$

$$\begin{aligned} & \times (-1)^l \sum_{i=1}^p \sum_{j=1}^q (t_{i,j}^o)^{-i-j+1} \frac{(n_{i,j} + j)!}{n_{i,j}!} \\ & + \sum_{k=2}^K \frac{1}{k(k-1)} \sum_{l=0}^k \binom{k}{l} \frac{(N+q-1)!}{(N+q+l-1)!} (-1)^l \sum_{j=1}^q (s_{i,j}^0)^{-j+1} \frac{(n_j + j)!}{n_{i,j}!}. \end{aligned} \quad (\text{A12})$$

本文 K 取为 4 表示泰勒级数展开为 4 项多项式, 经过检验, 计算精度已经足够保证结果的稳定性和准确性. 为了计算条件联合互信息, 本文也给出了贝叶斯条件联合熵的近似计算公式:

$$\begin{aligned} \hat{H}_{\text{bys}}(X \cap Y | Z) &= \sum_{i=1}^p \sum_{j=1}^q \sum_{\lambda=1}^v t_{i,j,\lambda}^0 \\ &- \sum_{j=1}^q s_j^0 - \frac{1}{N + pqv} \sum_{i=1}^p \sum_{j=1}^q \sum_{\lambda=1}^v (1 + \log(t_{i,j,\lambda}^0))(n_{i,j,\lambda} + 1) \\ &+ \frac{1}{N + q} \sum_{j=1}^q (1 + \log(s_j^0))(n_j + 1) \\ &- \sum_{k=2}^K \frac{1}{k(k-1)} \sum_{l=0}^k \binom{k}{l} \frac{(N + pqv - 1)!}{(N + pqv + l - 1)!} \\ &\times (-1)^l \sum_{i=1}^p \sum_{j=1}^q \sum_{\lambda=1}^v (t_{i,j,\lambda}^o)^{-i-j-v+1} \frac{(n_{i,j,\lambda} + j)!}{n_{i,j,\lambda}!} \end{aligned}$$

$$+ \sum_{k=2}^K \frac{1}{k(k-1)} \sum_{l=0}^k \binom{k}{l} \frac{(N+q-1)!}{(N+q+l-1)!} (-1)^l \sum_{j=1}^q (s_{i,j}^0)^{-j+1} \frac{(n_j + j)!}{n_{i,j}!}, \quad (\text{A13})$$

式中 t 表示 $\{X\}, \{Y\}, \{Z\}$ 三个随机过程所产生的随机变量在样本空间中的三维联合概率密度, 公式(A13)中的 s 表示由随机过程 $\{Z\}$ 所产生的随机变量在样本空间中的边缘概率密度. 将 (A12)和(A13)代入(3)式中, 我们得到贝叶斯条件互信息:

$$\begin{aligned} \hat{I}_{xy} &= \hat{I}_{\text{bys}}(\phi_x(t) \cap \Delta\phi_y(t+\tau) | \phi_y(t)) \\ &= \hat{H}_{\text{bys}}(\phi_x(t) | \phi_y(t)) + \hat{H}_{\text{bys}}(\Delta\phi_y(t+\tau) | \phi_y(t)) \\ &\quad - \hat{H}_{\text{bys}}(\phi_x(t) \cap \Delta\phi_y(t+\tau) | \phi_y(t)), \\ \hat{I}_{yx} &= \hat{I}_{\text{bys}}(\phi_y(t) \cap \Delta\phi_x(t+\tau) | \phi_x(t)) \\ &= \hat{H}_{\text{bys}}(\phi_y(t) | \phi_x(t)) + \hat{H}_{\text{bys}}(\Delta\phi_x(t+\tau) | \phi_x(t)) \\ &\quad - \hat{H}_{\text{bys}}(\phi_y(t) \cap \Delta\phi_x(t+\tau) | \phi_x(t)). \end{aligned} \quad (\text{A14})$$

将(A14)代入(5)式, 定义贝叶斯条件互信息耦合方向指数:

$$IM_{\text{bys}}(x, y) = \frac{\hat{I}_{xy} - \hat{I}_{yx}}{\hat{I}_{xy} + \hat{I}_{yx}}. \quad (\text{A15})$$