第 1 章 相空间重构

1.1 引言

一般时间序列主要是在时间域或变换域中进行研究，而在混沌时间序列处理中，无论是混沌不变量的计算、混沌模型的建立和预测都是在相空间中进行，因此相空间重构是混沌时间序列处理中非常重要的第一步。

为了从时间序列中提取更多有用信息，1980 年 Packard 等人提出了用时间序列重构相空间的两种方法：导数重构法和坐标延迟重构法[1]。从原理上讲，导数重构和坐标延迟重构都可以用来进行相空间重构，但就实际应用而言，由于我们通常不知道混沌时间序列的任何先验信息，而且从数值计算的角度看，数值微分是一个对误差很敏感的计算问题，因此混沌时间序列的相空间重构普遍采用坐标延迟的相空间重构方法[2]。坐标延迟法的本质是通过一维时间序列$\{x(n)\}$的不同时间延迟来构造m维相空间矢量：

$$\mathbf{x(i)} = \{x(i), x(i+\tau), \cdots, x(i+(m-1)\tau)\} \tag{1.1}$$

1981 年 Takens 等提出嵌入定理：对于无限长、无噪声的d维混沌吸引子的标量时间序列$\{x(n)\}$，总可以在拓扑不变的意义上找到一个m维的嵌入相空间，只要维数$m \geq 2d+1$[3]。Takens 定理保证了我们可以从一维混沌时间序列中重构一个与原动力系统在拓扑意义下等价的相空间，混沌时间序列的判定、分析与预测是在这个重构的相空间中进行的，因此相空间的重构是混沌时间序列研究的关键[2]。

1985 年 Grassberger 和 Procaccia 基于坐标延迟法，提出了关联积分的概念和计算公式，该方法适合从实际时间序列来计算混沌吸引子的维数，被称作 G-P 算法[4]。G-P 算法是混沌时间序列研究中的一个重要突破，从此对混沌时间序列的研究不仅仅局限于已知的混沌系统，而且也扩展到实测混沌时间序列，从而为混沌时间序列的研究进入实际应用开辟了一条道路[2]。

坐标延迟相空间重构技术有两个关键参数：即嵌入维m和时间延迟τ的确定。在 Takens 定理中，对于理想的无限长和无噪声的一维时间序列，嵌入维m和时间延迟τ可以取任意值，但实际应用最后等时间序列都是含有噪声的有限长序列，嵌入维数和时间延迟是不能任意取值，否则会严重影响重构的相空间质量。

有关时间延迟与嵌入维的选取方法，目前主要有两种观点。一种观点认为两者是互

不相关的，先求出时间延迟后再求出选择合适的嵌入维。求时间延迟τ比较常用的方法有自相关法[5]、平均位移法[5]、复自相关法[6]和互信息法[7, 8]等，目的是使原时间序列经过时间延迟后可以作为独立坐标使用。一个好的重构相空间是使重构后的吸引子和系统真正的吸引子尽可能做到拓扑等价，目前寻找最小嵌入维的方法主要是几何不变量法[9]、虚假最临近点法[10](FNN)和它的改进形式 Cao 方法[11]。另一种观点认为时间延迟和嵌入维是相关的，1996 年 Kugiumtzis 提出的时间窗长度是综合考虑两者的重要参数[12]。1999 年，Kim 等人基于嵌入窗法的思想提出了 C-C 方法，该方法使用关联积分同时估计出时延与嵌入窗[13]。C-C 方法也是实际时间序列中比较常用的方法，针对该方法的缺陷，国内学者作了相应的改进[14, 15]。

以上讨论的主要是求混沌不变量如关联维、Lyapunov 指数、Kolmogorov 熵或复杂度的常用相空间重构方法，重构的目标是重构吸引子和真正吸引子的近似程度达到全局最优。但由于无法得到混沌时间序列关于相空间重构的先验知识，因此上面提到的方法都具有一定的主观性[16]。目前并没有一种适合各种混沌时间序列的通用相空间重构方法，各种新的重构方法也不断被提出[17]，甚至有学者提出不需要进行重构直接描述系统混沌特征的新方法[18, 19]，但他们的可靠性有待于进一步验证[20, 21]。另外虽然均匀嵌入这种全局最优算法能保证求得的混沌不变量能体现系统全局特征，但不能达到很好的局部预测效果。从混沌时间序列建模和预测的角度，Judd 和 Small 提出非均匀嵌入和可变嵌入等将非线性建模和相空间重构相结合的方法[22-25]，可变嵌入方法也在不断发展中[26]。

本章中各节主要内容如下：1.2 节介绍确定时间延迟τ的常用方法，1.3 节介绍确定嵌入维m的常用方法，1.4 节介绍同时确定τ和m的时间窗长度方法，1.5 节介绍如何将相空间重构和非线性建模相结合。1.6 节为本章的重点内容，首先介绍海杂波的特点以及各种相空间重构方法处理后的结果，接着指出海杂波在相空间中表现出不完全混沌也不完全随机的特性，最后讨论噪声和采样间隔对相空间重构的影响。1.7 节为本章小结。

1.2 延迟时间τ的确定

时间延迟τ如果太小，则相空矢量$\mathbf{x(i)} = \{x(i), x(i-\tau), \cdots, x(i-m\tau)\}$中的任意两个分量$x(i-j\tau)$和$x(i-(j+1)\tau)$在数值上非常接近，以至于无法相互区分，从而无法提供两个独立的坐标分量；但如果时间延迟τ太大的话，则两坐标在统计意义上又是完全独立

的，混沌吸引子的轨迹在两个方向上的投影毫无相关性可言。因此需要用一定的方法来确定一个合适的τ值，目前确定τ的主要方法为序列相关法和相空间几何法。

1.1.1 自相关函数法

自相关法为序列相关法的一种，利用自相关函数选取延迟时间后，使得重构后的时间序列元素之间的相关性降低，同时尽可能使原序列动力学特征不丢失。

对于连续变量$x(t)$，其自相关函数(Autocorrelation function)定义为

$$C(\tau)=\lim_{T\to\infty}\int_{-\frac{T}{2}}^{\frac{T}{2}}x(t)x(t+\tau)dt \tag{1.2}$$

其中τ是时间的移动值，表示两时刻(t和$t+\tau$)的相互关联程度。当$x(t)$的幅值一定时，$C(\tau)$越大，则意味着$x(t)$和$x(t+\tau)$越相似。因此一般$x(t)$和$x(t+\tau)$的自相关函数$C(\tau)$随着τ的增加而逐渐变小，直至趋于零。

自相关函数是非常成熟的求时间延迟τ的方法，它主要是提取序列间的线性相关性。对于混沌时间序列$x_1,x_2,\cdots,x_n,\cdots$，可以先写出其自相关函数，我们设总点数为$N$，则序列$\{x_n\}$在时间跨度为$\tau$时的自相关函数为

$$R_{xx}(\tau)=\frac{1}{N}\sum_{i=0}^{N-1}x_i x_{i+\tau} \tag{1.3}$$

我们使用实际观测数据做出自相关函数随延迟时间τ变化的函数图像，当自相关函数下降到初始值的$1-1/e$时，所得的时间τ就是重构相空间的时间延迟τ[5]。

虽然自相关函数是是一种简便有效的计算延迟时间的方法，但它仅能提取时间序列间的线性相关性。如根据自相关函数法得到的τ可分别让x_i和$x_{i+\tau}$以及$x_{i+\tau}$和$x_{i+2\tau}$之间不相关，而x_i和$x_{i+2\tau}$之间的相关性却可能很强[12]。这一点也意味着由自相关法得到的时间延迟一般不可能推广到高维，因此在总结自相关法和平均位移法的基础上，林嘉宇等提出了一种更好求时间延迟的复自相关法[6]。

1.1.2 平均位移法

平均位移法(Average Displacement, AD)属于相空间重构几何法，可联系相关性准则，具有较强的物理意义[5]。与自相关法求序列时间上的相关性不同，Rosenstein 等提出的 AD 法是从几何的角度来确定时间延迟的。若时间延迟τ较小时，重构后的混沌吸引子

会被压缩在主对角线一带，而随着延迟时间的不断增大，吸引子的形状会逐渐展开。最后求得的τ应保证重构的相空间轨道从相空间的主对角线尽可能向外扩展，而又不发生折叠现象，最好的显示吸引子特征。

对于时间序列$\{x_n\}$按时间延迟τ进行相空间重构后，相空间的相邻两相点$\mathbf{x(i)}$和$\mathbf{x(i+\tau)}$之间的平均距离$S_m(\tau)$可以定义如下：

$$S_m(\tau)=\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{x(i+\tau)}-\mathbf{x(i)}\right\| \tag{1.4}$$

若嵌入维数m已经确定，则有：

$$S_m(\tau)=\frac{1}{N}\sum_{i=1}^{N}\sqrt{\sum_{j=1}^{m-1}(x_{i+j\tau}-x_i)^2} \tag{1.5}$$

$S_m(\tau)$随着时间延迟τ的增加，会逐渐从线性增加趋于饱和，其线性区的末端所对应的τ值即为最佳时间延迟，可选择 AD 法度量来$S_m(\tau)$曲线波形斜率，第一次降为其波形初始斜率的 40%以下对应的时间即为所求的时间延迟τ [5]。AD 法在总体变化趋势中可能夹有较强的抖动，具有一定的随机性，而且仅根据试验结果得到，其理论性不强[6]。

1.1.3 复自相关法

在分析求时间延迟的自相关法和平均位移法的基础上，林嘉宇等推导出一种较好的求时间延迟的方法,即复自相关法[6]。复相关法是将上面两种准则进行折衷的新方法，其计算复杂度不大，具有很好的抗噪能力。

由 AD 法可知，时间序列$\{x_n\}$的m维相空间重构下的平均位移$<S_m^2(\tau)>$如下

$$<S_m^2(\tau)>=\frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=1}^{m-1}(x_{i+j\tau}-x_i)^2 \tag{1.6}$$

其中，N为观测序列$\{x_n\}$的点数。平均位移$<S_m^2(\tau)>$体现了相空间矢量X_k离开相空间主对角线的距离。忽略边缘点带来的差别，可认为$\sum_{i=0}^{N-1}x_{t+j\tau}^2$对$0\le j\le m-1$为常数，统一记为$E=\frac{1}{N}\sum_{i=0}^{N-1}x_i^2$，展开(1.6)式，则有

$$<S_m^2(\tau)>=2(m-1)E-2\sum_{j=1}^{m-1}R_{xx}(j\tau) \tag{1.7}$$

其中，$R_{xx}(j\tau)$是序列$\{x_n\}$在时间跨度为$j\tau$的自相关函数，即：

$$R_{xx}(j\tau)=\frac{1}{N}\sum_{i=0}^{N-1}x_i x_{i+j\tau} \tag{1.8}$$

式(1.7)体现了 AD 法和自相关法之间的关系，事实上，二维平均位移法和自相关法具有大致相反的波形。由(1.7)，定义复自相关函数为：

$$R_{xx}^{m}(\tau)=\sum_{j=1}^{m-1}R_{xx}(j\tau) \tag{1.9}$$

可设 m 维相空间重构的复自相关法为：选取 $R_{xx}^{m}(\tau)$ 的第一个零点为时间延迟 τ 。

为了适应一般系统，计算中一般采用偏复自相关法，定义 m 维去偏自相关法为：

$$C_{xx}^{m}(\tau)=\frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=1}^{m-1}(x_t-\overline{x})(x_{i+j\tau}-\overline{x}) \tag{1.10}$$

显然，复自相关法是由 AD 法蜕化而得，继承了 AD 法在相空间重构几何学中的意义。同时它又可看成是自相关法的高维扩展，克服了自相关法的缺点。复自相关法不但能让 x_i 和 $x_{i+\tau}$ 以及 $x_{i+\tau}$ 和 $x_{i+2\tau}$ 之间不相关，也保证了 x_i 和 $x_{i+2\tau}$ 之间不相关。因此复自相关法具有较明确的理论依据，可作为时间延迟选择的序列相关准则和相空间几何准则之间的桥梁。

1.1.4 互信息法

自相关法本质上一个线性的概念，适合于判断线性相关性，而混沌系统是一个非线性系统，因此 Fraser 和 Swinney 提出了用互信息来判断系统的非线性相关性[7]。

考虑两离散信息系统 $\{s_1,s_2,\cdots,s_n\}$ 和 $\{q_1,q_2,\cdots,q_n\}$ 构成的系统 S 和 Q 。则由信息论，从两系统测量中所获得的平均信息量，即信息熵分别为：

$$\begin{aligned}H(S)&=-\sum_{i=1}^{n}P_i(s_i)\log_2 P_s(s_i)\\ H(Q)&=-\sum_{i=1}^{m}P_q(q_j)\log_2 P_q(q_j)\end{aligned} \tag{1.11}$$

其中， $P_s(s_i)$ 和 $P_q(q_j)$ 分别为 S 和 Q 中事件 s_i 和 q_j 的概率。

在给定 S 的情况下，我们得到的关于系统 Q 的信息，称为 S 和 Q 的互信息，如下：

$$I(Q,S)=H(Q)-H(Q\,|\,S) \tag{1.12}$$

其中
$$H(Q\,|\,s_i)=-\sum_{j}\left[P_{sq}(s_i,q_j)/P_s(s_i)\right]\log\left[P_{sq}(s_i,q_j)/P_s(s_i)\right] \tag{1.13}$$

因此有 $$I(Q,S)=\sum_i\sum_j P_{sq}(s_i,q_j)\log_2\left[\frac{P_{sq}(s_i,q_j)}{P_s(s_i)P_q(q_j)}\right] \tag{1.14}$$

其中，$P_{sq}(s_i,q_j)$ 为事件 s_i 和事件 q_j 的联合分布概率。

定义 $[s,q]=[x(t),x(t+\tau)]$，即 s 代表时间序列 $x(t)$，q 为其延迟时间为 τ 的时间序列 $x(t+\tau)$，则 $I(Q,S)$ 显然是与时间延迟有关的函数，记为 $I(\tau)$。$I(\tau)$ 的大小代表了在已知系统 S 即 $x(t)$ 的情况下，系统 Q 也就是 $x(t+\tau)$ 的确定性大小。$I(\tau)=0$，表示 $x(t+\tau)$ 完全不可预测，即 $x(t)$ 和 $x(t+\tau)$ 完全不相关；而 $I(\tau)$ 的极小值，则表示了 $x(t)$ 和 $x(t+\tau)$ 是最大可能的不相关，重构时使用 $I(\tau)$ 的第一个极小值作为最优延迟时间。

互信息的关键问题是联合分布概率 $P_{sq}(s_i,q_j)$ 的计算，Fraser 和 Swinney 采用的是等概率递推的方法，其划分与计算都很复杂，杨志安等提出了等间距格子法，计算相对更简便[8]。互信息法提出后，尽管其划分和计算复杂，但由于可以度量时间序列的非线性相关性，因此不但在计算最优时间延迟中被广泛采用，也大量应用于判定两非线性序列的相关性方面。

1.2 嵌入维数 m 的确定

选择嵌入维数的目的是使原始吸引子和重构吸引子拓扑等价。如果 m 选得过小，吸引子会发生折叠甚至某些地方会出现自相交，在某些区域的较小邻域内会包含吸引子不同轨道上的点，重构吸引子的形状和原始吸引子完全不同。如果 m 选得过大，虽然此时理论上讲是可行的，吸引子的几何结构被完全打开，但这样作一方面增大了计算量，另一方面噪声的影响也被进一步放大。因此我们需要选择合适的嵌入维数，既保证能准确计算各种混沌不变量，又尽量降低计算量和噪声的影响。

1.2.1 几何不变量法

关于最小嵌入相空间维数，Taken 等从理论上证明了当 $m\geq 2d+1$ 时可获得一个吸引子的嵌入，其中 d 是吸引子的分形维数，但这只是一个充分条件，对实测时间序列选择嵌入维 m 没有帮助。在实际应用中通常的方法是计算吸引子的某些几何不变量(如关联维 D、Lyapunov 指数等)，选择好延迟时间 τ 后逐渐增加 m，不断计算不变量直到他们停止变化为止。从理论上讲，由于这些不变量具有吸引子的几何性质，当 m 大于最小嵌

入维数时，几何结构被完全打开，此时这些不变量与嵌入维数无关，于是选择吸引子的几何不变量停止变化时的嵌入维 m 为重构的相空间维数[9]。

1.2.2 虚假最近邻点法

从几何的观点看，混沌时间序列是高维相空间混沌运动的轨迹在一维空间上的投影，在这个投影的过程中，混沌运动的轨迹会被扭曲。高维相空间中并不相邻的两点投影在一维空间轴上时有可能会称为相邻的两点，即虚假邻点，这就是混沌时间序列呈现出无规律的原因所在。重构相空间，实际上就是从混沌时间序列中恢复混沌运动的轨迹，随着嵌入维数 m 的增大，混沌运动的轨道就会逐渐打开，虚假邻点也会逐步被剔除，从而混沌运动的轨迹得到恢复，这个思想就是虚假最临近点法(False Neatest Neighbors, FNN)的出发点[10]。

在 d 维相空间中，每一个相点矢量为 $\mathbf{x(i)} = \{x(i), x(i+\tau), \cdots, x(i+(d-1)\tau)\}$，都有一个某距离内的最近临点 $\mathbf{x}^{\mathbf{NN}}\mathbf{(i)}$，其距离为

$$R_d(i) = \left\| \mathbf{x(i)} - \mathbf{x}^{\mathbf{NN}}\mathbf{(i)} \right\| \tag{1.15}$$

当相空间的维数从 d 维增加到 $d+1$ 维时，这两个相点的距离就会发生变化，两者的距离成为 $R_{d+1}(i)$ 且有

$$R_{d+1}^2(i) = R_d^2(i) + \left\| x(i+\tau d) - x^{NN}(i+\tau d) \right\| \tag{1.16}$$

如果 $R_{d+1}(i)$ 比 $R_d(i)$ 大很多，可以认为这是由于高维混沌吸引子中两个不相邻的点投影到低维轨道上时变成相邻的两点造成的，因此这样的临点是虚假的，令

$$a_1(i,d) = \frac{\left\| x(i+\tau d) - x^{NN}(i+\tau d) \right\|}{R_d(i)} \tag{1.17}$$

若 $a_1(i,d) > R_\tau$，则 $\mathbf{x}^{\mathbf{NN}}\mathbf{(i)}$ 是 $\mathbf{x(i)}$ 的虚假最近临点，阈值 R_τ 可在[10,50]之间选取。

对于含噪的有限长度数据，也可加入如下法则判断，若

$$R_{d+1}(i)/R_A \geq 2 \tag{1.18}$$

则 $\mathbf{x}^{\mathbf{NN}}\mathbf{(i)}$ 是 $\mathbf{x(i)}$ 的虚假最近临点，其中

$$R_A = \frac{1}{N}\sum_{i=1}^{N}\left[x(i) - \bar{x}\right], \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x(i) \tag{1.19}$$

对于实测时间序列，从嵌入维数的最小起始值开始，计算虚假最近临点的比例，然

后增加d，直到虚假最近临点的比例小于5%或者虚假最近临点不再随着d的增加而减少时，可以认为混沌吸引子已经完全打开，此时的d为嵌入维数，在相空间嵌入维数的确定方面，FNN 被认为是计算嵌入维很有效的方法之一。

1.2.2 伪最近邻点的改进方法－Cao 方法

FNN 算法对信号中含有的噪声敏感，虚假临近点的数目会随着噪声的影响而起伏，不进行单调变化。另外该方法在实际操作中需要选择R_τ和R_A，具有很强的主观性。Cao Liangyue 提出改进的 FNN 算法(Cao 方法)，该方法主要有以下优点：计算时只需要延迟时间τ一个参数；能够有效区分随机信号和确定性信号；使用较小的数据量就可以求得嵌入维[11]。

公式(1.17)代入$R_d(i)$后为如下形式：

$$a_1(i,d)=\frac{\left\|x(i+\tau d)-x^{NN}(i+\tau d)\right\|}{\left\|\mathbf{x_d(i)}-\mathbf{x_d}^{\mathbf{NN}}\mathbf{(i)}\right\|} \tag{1.20}$$

Cao 将上式改写为

$$a_2(i,d)=\frac{\left\|\mathbf{x_{d+1}(i)}-\mathbf{x_{d+1}}^{\mathbf{NN}}\mathbf{(i)}\right\|}{\left\|\mathbf{x_d(i)}-\mathbf{x_d}^{\mathbf{NN}}\mathbf{(i)}\right\|} \tag{1.21}$$

其中$\mathbf{x_d(i)}$和$\mathbf{x_d}^{\mathbf{NN}}\mathbf{(i)}$为$d$维空间的第$i$个向量和它的最临近点，$\mathbf{x_{d+1}(i)}$和$\mathbf{x_{d+1}}^{\mathbf{NN}}\mathbf{(i)}$是$d+1$维相空间的第$i$个向量和它的最临近点。

定义

$$\begin{aligned} E(m)&=\frac{1}{N-m\tau}\sum_{i=1}^{N-m\tau}a_2(i,m) \\ E1(m)&=E(m+1)/E(m) \end{aligned} \tag{1.22}$$

如果时间序列是确定的，则嵌入维是存在的，即$E1(m)$将在m大于某一特定值m_0后将不再变化。若时间序列是随机信号，则$E1(m)$应逐渐增加，但在实际应用中对一有限长序列$E1(m)$究竟是在缓慢变化还是已经稳定不容易判断，因此补充一个判断准则为

$$\begin{aligned} E^*(m)&=\frac{1}{N-m\tau}\sum_{i=1}^{N-m\tau}\left|x(i+m\tau)-x^{NN}(i+m\tau)\right| \\ E2(m)&=E^*(m+1)/E^*(m) \end{aligned} \tag{1.23}$$

对于随机序列，数据间没有相关性，即不具备可预测性，$E2(m)$将始终为 1；对于确定性序列，数据点之间的相关关系是依赖于嵌入维m值变化的，总存在一些m值使得$E2(m)$不等于 1。

1.3 同时确定嵌入维和延迟时间

以上介绍的是分别选取嵌入维数m和延迟时间τ的方法，还有一种观点认为m和τ是相互依赖的，在进行相空间重构时只需保证时间窗$\tau_w = (m-1)\tau$不变即可。这种方法可以同时确定嵌入维和延迟时间，比较有代表性的算法是 Kugiumtzis 提出的时间窗长度和 Kim 等人提出的 C-C 方法[12, 13]，国内学者也对 C-C 方法作了进一步改进[14, 15]。

1.3.1 时间窗长度

1996 年 Kugiumtzis 提出延迟时间τ的选取不应该独立于嵌入维数m，而应该依赖延迟时间窗口[12]

$$\tau_w = (m-1)\tau \tag{1.24}$$

具体算法为首先根据原时间序列的波动求出平均轨道周期τ_p，如果时间序列没有明显的平均周期，则将原数据经过滤波以后再求平均周期τ_p。在保证嵌入维m大于序列本身关联维d的前提下，固定τ_w值后依据式(1.24)变换m和τ的值，使用关联维作为验证指标，逐渐改变τ_w大小来确定最优的时间窗长度。经过多次试验发现在一定时间窗长度下，大致为$\tau_w \geq \tau_p$，只要m和τ的值满足式(1.24)，最后求出的关联维就保持不变。时间窗长度方法的计算量较大，另外很难找到复杂时间序列平均周期τ_p，滤波方法求得的值只是一种主观上的近似。毫无疑问，m和τ的选取是有一定联系的，求m和τ的方法都可理解为：确定时间延迟τ确保时间序列各成分相互依赖，虽不依赖于m但影响m的选择；而时间窗口τ_w依赖于m，且τ_w随m而变化。时间窗口长度和 C-C 方法的优势是能同时确定m和τ，但它们都是实践方法，还缺乏必要的理论根据。

1.3.2 C-C 方法

1999 年 Kim 等提出了 C-C 方法，该方法延续了时间窗口的概念，先定义关联积分，再构造统计量$S_1(m,N,r,t)$，依据 BDS 统计结论确定m,N,r的合适取值范围，实际计算中利用$S_2(m,N,r,t) \sim t$的统计结论，实现最优时延τ_d与嵌入窗τ_w的估计。

我们定义：τ_s指时间序列的采样间隔，$\tau_d = t\tau_s$指时间序列的最优延时，$\tau_w = (m-1)\tau_d$指延迟时间窗口，τ_p是平均轨道周期($\tau_w \geq \tau_p$)，$\tau(\tau = t)$为时间延迟的数值，m是嵌入维数，N是数据组的大小，数据点数为$M = N-(m-1)\tau$，则重构相空间中嵌入时间序列$Y(i)$的关联积分定义为：

$$C(m,N,r,t)=\frac{2}{M(M-1)}\sum_{1\le i\le j\le M}\theta(r-d_{ij}) \tag{1.25}$$

其中 $$r>0,\quad d_{ij}=\left\|Y_i-Y_j\right\|_\infty,\quad \theta(z)=\begin{cases}0 & z<0\\ 1 & z\ge 0\end{cases} \tag{1.26}$$

关联积分是一个累积分布函数，表示相空间中任意两点之间距离小于半径r的概率，这里点与点之间的距离用矢量之差的无穷范数表示。定义检测统计量为

$$S_1(m,N,r,t)=C(m,N,r,t)-C^m(1,N,r,t) \tag{1.27}$$

上式在实际计算时，需要先把时间序列$\{x(n)\}$拆分为t个不相交的子序列，长度均为$N_s=N/t$，t为重构时延

$$\begin{aligned}
&\mathbf{x^1}=\{x_1,x_{t+1},\cdots,x_{N-t+1}\}\\
&\mathbf{x^2}=\{x_2,x_{t+2},\cdots,x_{N-t+2}\}\\
&\cdots\quad\cdots\quad\cdots\quad\cdots\\
&\mathbf{x^t}=\{x_t,x_{2t},\cdots,x_N\}
\end{aligned} \tag{1.28}$$

然后将式(1.27)采用分块平均策略，计算每个序列的$S(m,N,r,t)$为

$$S_2(m,N,r,t)=\frac{1}{t}\sum_{s=1}^{t}[C_s(m,\frac{N}{t},r,t)-C_s^m(1,\frac{N}{t},r,t)] \tag{1.29}$$

当$N\to\infty$时， $$S_2(m,r,t)=\frac{1}{t}\sum_{s=1}^{t}[C_s(m,r,t)-C_s^m(1,r,t)],(m=2,3,\cdots) \tag{1.30}$$

如果时间序列是独立同分布的，那么对固定的m和t，当$N\to\infty$时，对于所有的r均有$S_2(m,r,t)$恒等于零。但实际时间序列是有限长且元素时间存在相关性，实际得到的$S_2(m,r,t)$一般不为零。$S_2(m,r,t)\sim t$反映了时间序列的自相关特性，仿照求时延的自相关算法，最优时延τ_d可取$S_2(m,r,t)\sim t$的第一个零点，或者取$S_2(m,r,t)\sim t$对所有半径r相互差别最小的时间点，此时重建相空间中的点最接近均匀分布，重构吸引子的轨道在相空间完全展开。选择对应值最大和最小的两个半径r，定义差量为

$$\Delta S_2(m,t)=\max\{S_2(m,r_j,t)\}-\min\{S_2(m,r_j,t)\} \tag{1.31}$$

上式度量了$S_2(m,r,t)\sim t$对所有半径r的最大偏差。所以局部最大时间t应该是$S_2(m,r,t)$的零点和$\Delta S_2(m,t)$的最小值，最优延迟时间τ_d对应着这些局部最大时间t中的第一个。

根据几种重要渐近分布的数学统计结果表明：当$2 \le m \le 5$，$\frac{\sigma}{2} \le r \le 2\sigma$，$N \ge 500$时，渐近分布可以通过有限序列很好的近似，并且$S_2(m,N,r,1)$能代表序列的相关性。这里的$\sigma$指时间序列的均方差或标准差，一般如果仅仅判别序列是否为混沌时间序列，取$N = 500$，甚至 100 就够了；需要重构时间序列时一般取$N = 3000$比较好。因为$S(m,N,r,1)$研究的是时间序列的非线性独立性，所以增大N是没有必要的。使用 C-C 方法计算延迟时间和嵌入维时，取$N = 3000$，$m = 2,3,4,5$，$r_i = \frac{i\sigma}{2}$，$i = 1,2,3,4$。

$$\overline{S}_2(t) = \frac{1}{16}\sum_{m=2}^{5}\sum_{j=1}^{4} S_2(m, r_j, t) \tag{1.32}$$

$$\Delta\overline{S}_2(t) = \frac{1}{4}\sum_{m=2}^{5} \Delta S_2(m,t) \tag{1.33}$$

$$S_{2cor}(t) = \Delta\overline{S}_2(t) + \left|\overline{S}_2(t)\right| \tag{1.34}$$

在上面三式中，C-C 方法取$\overline{S}_2(t)$的第一个零点或$\Delta\overline{S}_2(t)$的第一个极小值作为时间序列独立的第一个局部最大值，即最优时间延迟τ_d。并综合$\overline{S}_2(t)$和$\Delta\overline{S}_2(t)$，取$S_{2cor}(t)$的全局最小值作为时间序列的时间窗口长度：$\tau_w = (m-1)\tau_d$。

1.3.3 改进的 C-C 方法

经过深入分析后，陆振波等指出C-C方法存在三点不足[15]。

第一，实际中$\overline{S}_2(t)$的第一个零点并不等于$\Delta\overline{S}_2(t)$的第一个局部极小点，因此文中将$\overline{S}_2(t)$的第一个零点和$\Delta\overline{S}_2(t)$的第一个极小值都视为最优时延τ_d是不合适的，只需考虑$\Delta\overline{S}_2(t)$的第一个局部极小点作为最优时延τ_d。

第二，原方法在统计量的计算时，采用式(1.29)的分块平均策略，当时$t = KT$时(k为大于零的整数) $\Delta\overline{S}_2(t)$等于零，而且$\Delta\overline{S}_2(t)$出现随增大而不断增长的高频起伏，当最优时延τ_d值较大时，这种高频起伏甚至影响到$\Delta\overline{S}_2(t)$的第一个局部极小点的选择。

第三，理想情况下的全局最小点即是最优嵌入窗$S_{2cor}(t)$，实际中存在若干个局部极小点与全局最小点在数值上相当接近，干扰了全局最小点的判读；甚至最优嵌入窗$S_{2cor}(t)$所对应的不是全局最小点，最终导致最优嵌入窗τ_w的错误估计。

基于以上C-C方法的不足，[15]中提出如下改进的C-C方法的相空间重构参数选择。

第一，比较$S_1(m,N,r,t)$与$S_2(m,N,r,t)$后，文中指出式(1.29)中当固定$m,r,\ N \to \infty$)时，出现随增大而不断增长的高频起伏；而式(1.27)中在相同前提下，$S_1(m,N,r,t) \sim t$总

体上与$S_2(m,N,r,t)\sim t$具有相同的起伏规律，但去除了中的高频起伏。因此参照前面用$S_2(m,N,r,t)\sim t$求得$\overline{S}_2(t)$和$\Delta\overline{S}_2(t)$的方法，这里用$S_1(m,N,r,t)\sim t$求得$\overline{S}_1(t)$与$\Delta\overline{S}_1(t)$，改进的C-C方法寻找$\Delta\overline{S}_1(t)$的第一个局部极小点作为最优时延τ_d。与原C-C方法选择$\Delta\overline{S}_2(t)$第一个局部极小点相比，新方法对最优时延τ_d的选择更为准确，这种优势特别是在最优时延τ_d取值比较大时更为明显。

另外，改进的C-C方法寻找$\left|\overline{S}_1(t)-\overline{S}_2(t)\right|$的周期点作为最优嵌入窗$\tau_d$。与原C-C方法选择的全局最小点相比，$\left|\overline{S}_1(t)-\overline{S}_2(t)\right|$在周期点位置存在比较明显的局部峰值，使对最优嵌入窗τ_w选取更为可靠。

最后，为加快计算速度，文中指出统计量定义式(1.27)也可采用分块平均的方法计算关联积分。考虑混沌时间序列，以时延t，嵌入维m，重构相空间$X=\{X_i\}$。为表述方便，将关联积分定义式(1.25)左边$C(m,N,r,t)$改写成$C(X,r)$形式，右边不变。这样，检验统计量的定义式(1.27)可改写成

$$S_1(m,N,r,t)=C(X,r)-C^m(x,r) \tag{1.35}$$

再令

$$\begin{aligned} X^{k,s}&=\{X_s,X_{s+k},\cdots\},s=1,2,\cdots,k \\ x^{k,s}&=\{x_s,x_{s+k},\cdots\},s=1,2,\cdots,k \end{aligned} \tag{1.36}$$

这里$X^{k,s}$与$x^{k,s}$分别是X与x中个互不相交子集，为独立于时延的常数。因此统计量定义式(1.35)的近似表达式为：

$$S_1(m,N,r,t)=\frac{1}{k}\sum_{s=1}^{k}C(X^{k,s},r)-\left[\frac{1}{k}\sum_{s=1}^{k}C(X^{k,s},r)\right]^m \tag{1.37}$$

这里k是权衡计算精度与速度的可调参数。当$k=1$时，式(1.37)就与式(1.35)等价，此时统计量的计算精度最高。当k取值较大，分块计算关联积分存在误差，但这种误差对不同时延t几乎只相差一个常数，几乎不影响局部极值点的判读。另一方面，由于关联积分计算的时间复杂度为$O((M/k)^2)$，增大k值，计算速度急剧加快。因此在实际中，在不严重损失估计精度的前提下，适当地增大k的取值，可大大加快的计算速度。

原C-C方法在计算延迟时间τ_d和时间窗τ_w的结果往往不太稳定，新方法解决了计算的稳定性，但新方法计算出的时间窗是否有效还需要进一步证明。徐自励等提出对时间序列采用分段处理的简便方法来消除原C-C方法的不稳定性，也取得一定的效果[14]。

1.3.4 微分熵比方法

C-C方法可以同时求出延迟时间 τ_d 和时间窗 τ_w，在实践中获得广泛应用，这在一定程度上促进了人们寻找同时确定延迟时间和嵌入维方法的兴趣。Gautama等提出了微分熵方法进行相空间重构[17]，设数据的概率密度函数 $p(\mathbf{x})$，微分熵定义为

$$H(\mathbf{x}) = -\int_{-\infty}^{+\infty} p(\mathbf{x}) \ln p(\mathbf{x}) d\mathbf{x} \tag{1.38}$$

为了方便使用 Kozachenko-Leonenko(K-L)估计微分熵为

$$H(x) = \sum_{j=1}^{N} \ln(N\rho_j) + \ln 2 + C_E \tag{1.39}$$

其中，N 是数据长度，ρ_j 为第 j 个点和它最临近点的距离，C_E 为常数。

对给定的一个嵌入维为 m，延迟时间为 τ 的时间序列 $\{x_n\}$，定义 $H(x,m,\tau)$ 为它的微分熵，$H(x_{s,i},m,\tau)$ 为原始数据第 i 次替代数据的微分熵，则微分熵比为

$$I(m,\tau) = \frac{H(x,m,\tau)}{\left\langle H(x_{s,i},m,\tau) \right\rangle_i} \tag{1.40}$$

其中 $\langle \bullet \rangle_i$ 表示 i 次平均值，一般 i 取 10 或 5。使用最小描述长度(MDL)，加入对嵌入维的惩罚因子后，新的熵比定义为：

$$R_{ent}(m,\tau) = I(m,\tau) + \frac{m \ln N}{N} \tag{1.41}$$

熵可以用来描述时间序列的信息特征，差分熵可以衡量在不同嵌入情况下的时间序列的无序特征，也是一种同时确定 m 和 τ 的方法，该方法要求信号在相空间中显示出很好的结构特征，因此比较适合生物信号重构[17]。

1.4 非线性建模与相空间重构

以上我们讨论的相空间重构，整个时间序列采用的都是相同的延迟时间和相同的嵌入维，这种重构方式叫做均匀嵌入方式，均匀相空间重构适合于不呈现明显伪周期性的混沌时间序列，但数据具有很强的伪周期性时，使用均匀相空间重构效果就不好，又如当时间序列既有长周期又有短周期，使用均匀嵌入方式显然不能兼顾这两种情况。均匀嵌入的精度一般能够满足求取混沌不变量的要求，但是如果进行非线性建模和预测，这时嵌入参数对预测效果影响很大，就需要采用不同延迟时间但相同嵌入维的非均匀嵌入，甚至是不同延迟时间和不同嵌入维的变化嵌入方式[22-25]。

设一维时间序列$\{x(n)\}$采用均匀时间延迟来构造m维相空间矢量为：

$$\mathbf{x(i)} = \{x(i), x(i+\tau), \cdots, x(i+(m-1)\tau)\} \tag{1.42}$$

则采用非均匀时间延迟来进行相空间重构为：

$$\mathbf{x'(i)} = \{x(i), x(i+l_1), x(i+l_2), \cdots, x(i+l_n)\} \tag{1.43}$$

其中

$$0 \le l_1 < l_2 < \cdots < l_n \le (m-1)\tau = \tau_w \tag{1.44}$$

设$\phi: M \to M$代表动力系统的演化算子，$h: M \to R$是一个C^2可微分观测函数。试验时间序列记为$\{h(X_1), h(X_2), \cdots, h(X_N)\}$，令$x_i \equiv h(X_i)$，根据 Takens 定理有如下映射$g$成立

$$x_i \overset{g}{\mapsto} (x_i, x_{i-1}, x_{i-2}, \cdots, x_{i-d-1}) \tag{1.45}$$

映射g是如$g(x_i) = (x_i, x_{i-1}, \cdots, x_{i-m-1})^3$这样的同胚演化算子，设嵌入映射为如下形式

$$x_i \overset{\hat{g}}{\mapsto} (a_1 x_i, a_2 x_{i-1}, \cdots, a_d x_{i-d-1}) \tag{1.46}$$

其中

$$a = [a_1, a_2, \cdots, a_d],\ a_i \in \{0, 1\} \tag{1.47}$$

使用最短描述长度(MDL)作为约束条件，经过复杂推导后如下[23, 25]：

$$\begin{aligned} DL(x) \approx \frac{d}{2}\left[\frac{1}{d}\sum_{i=1}^{d}(x_i - \overline{x})^2\right] + \frac{N-d}{2}\ln\left[\frac{1}{N-d}\sum_{i=d+1}^{N} e_i^2\right] \\ + \frac{N}{2}(1 + \ln 2\pi) + d + DL(d) + DL(\overline{x}) + DL(p) \end{aligned} \tag{1.48}$$

上式中$DL(\overline{x})$为常数，当使用局部常数模型时$DL(p) = 0$，这样可以简化描述长度的计算。因此对于给定的(a, d)，在公式(1.48)中只有$\sum e_i^2$不能求出。设x_s是x_t的最临近点，则有

$$x_{t+1} \simeq x_{s+1}, \quad e_{t+1} = x_{t+1} - x_{s+1} \tag{1.49}$$

即使用时间序列中两个最临近点的下一点的差异作为预测误差。

当模型形式一定时，公式(1.48)求得的最小$DL(x)$可以作为时间窗长度τ_w来使用，可以看到描述长度与噪声水平、数据和模型长度及模型预测误差等诸多因素有关。此时时间序列的建模、预测和相空间重构就变成一个需要综合考虑的问题。

1.5 海杂波的相空间重构

内容 (略)

1.6 本章小结

本章首先回顾了时间延迟与嵌入维分别选取的各种方法，然后讨论了两者相互依赖的选取方法。相空间重构中目前广泛的方法使用都是均匀嵌入方式，虽然这些方法不能普遍适用于所有数据，但基本上能满足求混沌不变量的要求。相空间重构的另一个目的是建模和预测，此时相空间重构参数和建模、预测方法有关，简单的均匀嵌入方式不再适用，非均匀嵌入和可变嵌入方式是今后发展的方向。

1.7 后记

本文是相空间重构这一章的第一版，第二版中要增加如下内容：

（1） 相空间重构和建模、预测方法和相互关系之间更深入、更透彻的讨论

（2） 智能方法的相空间重构参数选取：主要有神经网络和遗传算法两种方法

（3） 海杂波的信号特点和相空间重构

（4） 噪声和采样间隔对相空间重构的影响

第二版内容因为要写小论文，不会很快共享出来，但欢迎大家一起讨论，你的关注是我前进的最大动力☺。

任何意见、建议、批评和讨论我都热烈欢迎。

我的信箱：xuxkboy@newmail.dlmu.edu.cn

欢迎大家到研学论坛(http://bbs.matwav.com/)混沌分形版进行讨论

许小可.海杂波的非线性分析与建模:(博士学位论文).大连:大连海事大学,2007

参考文献

[1] Packard N H,Crutchfield J P,Farmer J D, et al.Geometry from a time series.Physical Review Letters.1980,45(9):712-716.

[2] 陈铿,韩伯棠.混沌时间序列分析中的相空间重构技术综述.计算机科学.2005,32(04):67-70.

[3] Takens F.Detecting strange attractors in turbulence.In:Dynamical Systems and Turbulence.Berlin:Springer-Verlag,1981:366-381.

[4] Grassberger P,Procaccia I.Measuring the strangeness of strange attractors.Physica D: Nonlinear Phenomena.1983,9(1-2):189-208.

[5] Rosenstein M T,Collins J J,De L C, et al.Reconstruction expansion as a geometry-based framework for choosing proper delay times.Physica D: Nonlinear Phenomena.1994, 73(1-2):82-98.

[6] 林嘉宇,王跃科,黄芝平, et al.语音信号相空间重构中时间延迟的选择—复自相关法.信号处理.1999,15(03):220-225.

[7] Fraser A M,Swinney H L.Independent coordinates for strange attractors from mutual information.Physical Review A.1986,33(2):1134-1140.

[8] 杨志安，王光瑞，陈式刚.用等间距分格子法计算互信息函数确定延迟时间.计算物理.1995, 12(04):442-447.

[9] 张雨,任成龙.确定重构相空间维数的方法.国防科技大学学报.2005,27(06):101-105.

[10] Kennel M B,Brown R,Abarbanel H D, et al.Determining embedding dimension for phase-space reconstruction using a geometrical construction.Physical Review A.1992,45(6):3403-4311.

[11] Cao L.Practical method for determining the minimum embedding dimension of a scalar time series.Physica D: Nonlinear Phenomena.1997,110(1-2):43-50.

[12] Kugiumtzis D.State space reconstruction parameters in the analysis of chaotic time series -- the role of the time window length.Physica D: Nonlinear Phenomena.1996, 95(1):13-28.

[13] Kim H S,Eykholt R,Salas J D, et al.Nonlinear dynamics, delay times, and embedding windows.Physica D: Nonlinear Phenomena.1999,127(1-2):48-60.

[14] 徐自励,王一扬,周激流.估计非线性时间序列嵌入延迟时间和延迟时间窗的C-C平均方法.四川大学学报(工程科学版).2007,39(01):151-155.

[15] 陆振波,蔡志明,姜可宇.基于改进的C-C方法的相空间重构参数选择.系统仿真学报.2007,

19(11):in press.

[16] Cellucci C J,Albano A M,Rapp P E, et al.Comparative study of embedding methods.Physical Review E.2003,67(6):066210-1-13.

[17] Gautama T,Mandic D P,Van H M, et al.A differential entropy based method for determining the optimal embedding parameters of a signal.Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on,2003:29-32.

[18] Gottwald G A,Melbourne I.Testing for chaos in deterministic systems with noise.Physica D: Nonlinear Phenomena.2005,212(1-2):100-110.

[19] Zhang J,Luo X,Small M, et al.Detecting chaos in pseudoperiodic time series without embedding.Physical Review E.2006,73(1):016216-1-5.

[20] Hu J,Tung W,Gao J, et al.Reliability of the 0-1 test for chaos.Physical Review E.2005, 72(5):056207-1-5.

[21] Zhang J,Luo X,Nakamura T, et al.Detecting temporal and spatial correlations in pseudoperiodic time series.Physical Review E.2007,75(1):016218-1-10.

[22] Judd K,Mees A.Embedding as a modeling problem.Physica D: Nonlinear Phenomena.1998, 120(3-4):273-286.

[23] Small M,Tse C K.Optimal embedding parameters: a modelling paradigm.Physica D: Nonlinear Phenomena.2004,194(3-4):283-296.

[24] http://arxiv.org/abs/nlin/0312011

[25] Small M.Applied Nonlinear Time Series Analysis: Applications in Physics, Physiology and Finance.Singapore:World Scientific,2005.

[26] Manabe Y,Chakraborty B.A novel approach for estimation of optimal embedding parameters of nonlinear time series by structural learning of neural network.Neurocomputing.2007, 70(7-9):1360-1371.