1. 排列熵

给定单变量时间序列 $\{y_k\}_{k=1}^{M}$ ，引入嵌入维数 d ，时间延迟 τ 有 $Y_l^{d,\tau}=\left\{y_l,y_{l+\tau},\ldots,y_{l+(d-1)\tau}\right\}$ for $l=1,2,\ldots,\mathrm{M}-(d-1)\tau$.共 $L=M-(d-1)\tau$ 个向量，间隔为 τ ，长度为 d。 $L=M-(d-1)\tau$ 个向量一一映射到一个 ordinal pattern，一共有 d!可能的排列 $\left\{\pi_j^{d,\tau}\right\}_{j=1}^{d!}$ 可用 Π 表示.根据 Shannon entropy 的形式，排列熵的计算公式为：

$$H\left(p_{PE}\right)=-\sum_{\mathrm{j}:\pi_j^{d,\tau}\in\Pi}p\left(\pi_j^{d,\tau}\right)Inp\left(\pi_j^{d,\tau}\right)=-\sum p_{PE}Inp_{PE}$$

$p\left(\pi_j^{d,\tau}\right)$ 为

$$p\left(\pi_j^{d,\tau}\right)=p_{MPE}=\frac{\left\|l:l\le L,type\left(Y_l^{d,\tau}\right)=\pi_j^{d,\tau}\right\|}{L}$$

where $type(\cdot)$ denotes the map from pattern space to symbol space and $\|\cdot\|$ denotes the cardinality of a set.

例子:时间序列 x={3，2，5，8，9，6，1}，取 d=3,、τ=1。相当于 M=7，L=7-(3-1)*1=5 个向量一一映射到一个 ordinal pattern（motif）。
第一个(x1，x2，x3)=(3，2，5)映射到 ordinal pattern（213），因为 x2<x1<x3，第二个三维向量时(x1，x2，x3)=(2，5，8)映射到（123），第三个三维向量时(x1，x2，x3)=(5，8，9）映射到（123），第四个三维向量时(x1，x2，x3)=(8，9，6)映射到（231），第五个三维向量时(x1，x2，x3)=(9，6，1)映射到（321）。

然后一共有 6（即 3！）种可能的排列（123 的排列组合），然后可以计算出排列的概率分布:p(123)=2/5 、p(132)=0 、p(213)=1/5 、p(231)=1/5 、p(312)=0 、p(321)=1/5。
排列熵根据香农熵的形式求得：s[p]=-(2/5)log(2/5)-3(1/5)log(1/5)=1.3322

2.多尺度排列熵

给定时间序列 $\{x_i\}_i^N$ 和 尺度 s，粗粒化后的时间序列

$$y_k^{(s)}=\frac{1}{s}\sum_{i=(k-1)s+1}^{ks}x_i$$

其实就是每 s 项进行平均。对于粗粒化后时间序列再计算排列熵。步骤根据排列熵步骤即可。

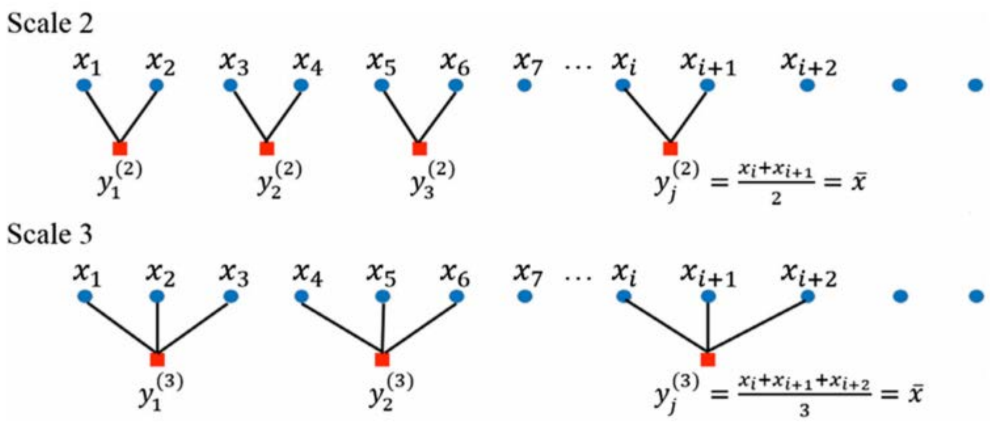

Fig. 1 Schematic diagram of the coarse-graining procedure when the scale are 2 and 3, respectively

例子： x={3,2,5,8,9,6,1,3}，粗粒化接上，取 s=2,y={(3+2)/2,(5+8)/2,(9+6)/3,(1+3)/2}={2.5,6.5,7.5,2}，取 d=2,τ=1。相当于 M=4,L=4-(2-1)*1=3 个向量一一映射到一个 ordinal pattern（motif）。第一个(x1,x2)=(2.5,6.5)映射到 ordinal pattern（12），第二个(x1,x2)=(6.5，7.5)映射到（12）第三个(x1,x2)=(7.5，2)映射到（21）

同样一共有 2（即 2！）种可能的排列：p(pi1)=p(12)=2/3，p(pi2)=p21)=1/3.

排列熵根据香农熵的形式求得：s[p]=-(1/3)log(1/3)- (2/3)log(2/3)=0.6365

3. 多尺度多变量排列熵

给定多变量（m 个变量）长度为 N 的时间序列 $X=\{x_{i,t}\}_{i=1,\dots,m}^{t=1,\dots,N}$，粗粒化后的 m 维时间序列 $Y_k^s=\left(y_{i,k}^s\right)_{i=1,\dots m}^{k=1,\dots,M}$ 由以下公式获得：

$$y_{i,k}^s=\frac{1}{s}\sum_{t=(k-1)s+1}^{ks}x_{i,t}$$

s 尺度因子 and $1\le k\le M=\dfrac{N}{s}$. 同样引入嵌入维数和时间延迟有 $Z_{i,l}^{d,\tau,s}=\left\{y_{i,l}^s,y_{i,l+\tau}^s,\dots,y_{i,l+(d-1)\tau}^s\right\}$ for $l=1,2,\dots,M-(d-1)\tau$，对每个变量映射，The relative frequencies $p_{i,j}^{s}$ are defined as:

$$p_{i,j}^{s}=\frac{\left\| l:l\le L,type\left(Z_{i,l}^{d,\tau,s}\right)=\pi_{j}^{d,\tau}\right\|}{L}=\frac{\sum_{l\le L}1_{u:type(u)=\pi_{j}^{d,\tau}}\left(Z_{i,l}^{d,\tau,s}\right)}{\sum_{l\le L}1_{u:type(u)=\Pi}\left(Z_{i,l}^{d,\tau,s}\right)}$$

其中 $L=\mathrm{M}-\left(d-1\right)\tau$ and $j=1,\ldots,n,n=d!$. $p_{i,j}^{s}$ are the entries of a matrix $P_{s(m,n)}=\left\{p_{i,j}^{s}\right\}$，它反映了粗粒度多元时间序列中 motifs 的分布，Here, $\sum_{i=i}^{m}\sum_{j=1}^{n}p_{i,j}^{s}=m$. 根据算法将原始的多元时间序列转化为时间相关矩阵，其中可以清楚地绘制相关统计量和熵。同时，描述 motif 分布的边缘相对频率可以计算为 $p_{mvMPE}=p_{\cdot,j}^{s}=\sum_{i=i}^{m}p_{i,j}^{s}$ 除以 m，for $j=1,\ldots,d!$, 多元多尺度排列熵计算公式：

$$H\left(p_{mvMPE}\right)=-\sum_{j=1}^{d!}p_{\cdot j}^{s}Inp_{\cdot j}^{s}=-\sum p_{mvMPE}Inp_{mvMPE}$$

注：添加多变量后其实也是按照每个变量求 motif 的概率，组成矩阵，然后多变量最后的 p 就是对矩阵按列求平均，列和除以变量个数。最后运用 Shannon 形式带入.

例子：给定 2 维时间序列：x={3,2,5,8,9,6,1;

2,4,6,8,5,7,3}

s=2 粗粒化后：y={(3+2)/2,(5+8)/2,(9+6)/2

(2+4)/2,(6+8)/2,(5+7)/2}=

{2.5,6.5,7.5;

3, 7, 6}

取 d=2,, τ=1 可得出每个 Z 的映射

Z11=(2.5,6.5) --(12) Z12=(6.5,7.5)---(12)

Z21(3,7)--(12) Z22=（7，6）--（21）

p11=p(12)=1 p12=p(21)=0

P21=p(12)=0.5 p22=p(21)=0.5 全部都是按变量求

P={1, 0;

0.5,0.5}

pmvmPE={(1+0.5)/2,(0+0.5)/2}={0.75,0.25}

s[p]= -(0.75)log(0.75)- (0.25)log(0.25)=0.5623

4.熵值标准化：

如果 S[P]=0，我们就能够确定地预测哪一种可能的结果，这是完全确定的，我的概率是由 p 给出的，在这种情况下，我们对概率分布描述的底层过程的知识是最大的。 当均匀分布时，所有结果出现可能性一样，我们无法根据概率预测出哪种结果出现更有可能，这是完全不确定的，此时是 s[p]最大为 Smax，我们的知识是最小的。

Smax=-(m!)(1/m!)log(1/m!)==log(m!)

标准化的熵 H=s[p]/Smax=s[p]/log(m!)

熵捕捉了时间序列的不确定性和无序性

对嵌入维数和时间延迟的选择问题，嵌入维数决定了可能的排列个数（d!个)和每个 motifi 的维数，因此太小或太大都不好，时间延迟为 1 是连续的取，当τ大于 1 时，组成排列的值被非连续地取，从而在不同的时间分辨率下映射系统的动力学。在 Permutation entropy: a natural complexity measure for time series 一文中 Bandt 和 Pompe 在[3,7]范围内考虑嵌入维数 d。至于其他参数的选项，他们特别建议τ=1 的时间延迟 参考这篇文章：Parameter selection for permutation entropy measurements。