# $\boxed{\text{Topological Markov Chains and Parry measure}}$

Consider a signal or some $\underline{s} \in \Sigma_n$

$$\underline{s} = \ldots 35.7201 \ldots$$

Sometimes one just knows which symbol follows which and not what the probality is. This is coded in a transition matrix $M$ where

$$M_{ij} = 1 \quad \text{if } ij \text{ can occur}$$

$$M_{ij} = 0 \quad \text{if } ij \text{ cannot occur.}$$

For example $M = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ indicates that

00 is not allowable and 01, 10, 11 are

- Let $\Sigma_M = \{ \underline{s} \in \Sigma_n : A_{s_i s_{i+1}} = 1 \text{ for all } i \in \mathbb{Z} \}$

so $\Sigma_M$ is the collection of bi-infinite sequences which only contain allowable pairs.

- Recall the topology on $\Sigma_n$ is given by the metric
$$d(\underline{s}, \underline{t}) = \frac{1}{2^\ell} \quad \text{with } \ell = \min\{ |i| : s_i \neq t_i \}$$

- **FACTS**:

(1) $\Sigma_M$ is compact

(2) $\Sigma_M$ is completely invariant $\sigma(\Sigma_M) = \Sigma_M$ under the shift $\sigma$

(3) $\sigma$ is a homeomorphism

Proof:

(3), (2) is easy. For (1) recall that in $\Sigma_n$ cylinder sets are open (and closed) and $\sigma^k\,{}_0[ij] = {}_{-k}[ij]$. Let $X = \{{}_0[ij] : M_{ij} = 0\}$ so $X$ is the cylinders sets of length 2 disallowed words and $X$ is open so $\bigcup_{k\in\mathbb{Z}} \sigma^k X$ is open and

$$\Sigma_M = \Sigma_n - \bigcup_{k\in\mathbb{Z}} \sigma^k X$$

is thus closed and therefore compact. $\boxtimes$

Let $(p, A)$ by a pair defining a Markov measure $\mu$ so $\vec{p} > 0$, $\sum p_i = 1$, $\vec{p}A = \vec{p}$ and $A$ is stochastic.

As in the HW we have

<u>FACT</u> $spt(\mu) = \Sigma_M \iff (A_{ij} = 0 \iff M_{ij} = 0)$

So there are many Markov measures on $\Sigma_M$

So which do we chose? There is a special one which turns out to have the maximal entropy called the Parry measure or Shannon-Parry measure. Here is the construction.

---

Assume $M$ is a $\{0,1\}$ matrix that is irreducible $(\forall i,j\ \exists n \text{ with } (M^n)_{ij} > 0)$. By the Perron-Fröbenius Theorem there is a unique eigenvalue $\lambda > 0$ of maximal modulus and $\vec{l}, \vec{r} > 0$ with $\vec{l} M = \lambda \vec{l}$ and $M\vec{r} = \lambda \vec{r}$.

Since $\vec{l}, \vec{r} > 0$, $\vec{l} \cdot \vec{r} > 0$ so we rescale so that $\vec{l} \cdot \vec{r} = 1$

Notice $\vec{l}$ is a row vector, $\vec{r}$ a column one.

Define $\vec{p}$ via $p_i = r_i \ell_i$ $\quad i = 0, \dots, n-1$

$A$ via $A_{ij} = \dfrac{M_{ij} r_j}{\lambda r_i}$

<u>Fact for next HW</u>: $(\vec{p}, A)$ defines a Markov measure in $\Sigma_n$

so (1) $\sum p_i = 1$ (2) $\vec{p} > 0$ (3) $\vec{p}A = \vec{p}$ (4) $A$ is stochastic

<u>Theorem</u> If $M$ is as above and $\mu$ is its Parry measure then $h_\mu(\sigma) = \log \lambda$

<u>Proof</u>: We computed for a general Markov measure

$$h_\mu(\sigma) = -\sum_{i,j} p_i A_{ij} \log A_{ij}$$

Thus in the case of Parry measure

$$h_\mu(\sigma) = -\sum_{i,j} r_i l_i \frac{M_{ij} r_j}{\lambda r_i} \log\left( M_{ij} r_j (\lambda r_i)^{-1} \right)$$

$$= -\sum_{i,j} \frac{l_i M_{ij} r_j}{\lambda} \left( \log M_{ij} + \log r_j - \log \lambda - \log r_i \right)$$

We examine terms one by one

- Now $M_{ij} = 0$ or $1$ and so $M_{ij} \cdot \log M_{ij} = 0 \ \forall i,j$
- Since $\vec{l} M = \lambda \vec{l}$, $\sum_i l_i M_{ij} = \lambda l_j$, $\forall j$

so $\sum_{i,j} \frac{l_i M_{ij} r_j}{\lambda} \log r_j = \sum_j r_j \log r_j \left( \sum_i \frac{l_i M_{ij}}{\lambda} \right) = \sum_j l_j r_j \log r_j$

- Similary using $M\vec{r} = \lambda\vec{r}$ yields

$$\sum_{i,j} \frac{l_i M_{ij} r_j}{\lambda} \log r_i = \sum_i r_i l_i \log r_i$$

- Finally, $\displaystyle\sum_{i,j} \frac{l_i M_{ij} r_j}{\lambda} \log \lambda$

$$= \log\lambda \sum_{i,j} \frac{l_i M_{ij} r_j}{\lambda} = (\log\lambda) \sum_i r_i l_i = \log\lambda$$

using $M\vec{r} = \lambda\vec{r}$ again and the initial normallization $\sum r_i l_i = 1$.

So Finally, $h_\mu(\tau) = 0 - \sum_j l_j r_j \log r_j + \log\lambda$

$$+ \sum_i l_i r_i \log r_i = \log\lambda. \quad \blacksquare$$

<u>Example</u> Let $M = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ so all transitions are allowable so $\Sigma_M = \Sigma_2$. The eigenvalues of $M$ are 2 and zero

and $\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \end{bmatrix} = 2 \begin{bmatrix} 1 & 1 \end{bmatrix}$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Now $\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2$ so we normalize

$\vec{l} = \begin{bmatrix} \frac{1}{2} & 1/2 \end{bmatrix}$ $\vec{r} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (it doesn't matter which we normalize)

Now for the Markov measure

$p_i = r_i \ell_i$ so $\vec{p} = [1/2 \ 1/2]$

$$A_{ij} = \frac{M_{ij} r_j}{\lambda r_i} = 1/2 \quad \forall i j$$

so $A = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$

and $h_\mu(\sigma) = \log 2$

Notice that $\mu([s_0 s_1 .. s_k]) = p_{s_0} A_{s_0 s_1} \cdots A_{s_{k-1} s_k}$

$= 1/2 \, 1/2 \cdots 1/2$ so $\mu$ is just the $(1/2, 1/2)$

Bernoulli measure and using the entrop formula

from HW $h_\mu(\sigma) = -\left(\frac{1}{2}\log 1/2 + 1/2 \log 1/2\right) = \log 2$ again

(of course)

- Recall that we showed for any Bernoulli measure on $\Sigma_2$, $h_\mu(\sigma) \leq \log 2$. Thus the Bernoulli measure $(1/2, 1/2)$ which is also the Parry measure for $M = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ is maximal among those measures.

- We will eventually see that given a irreducible $\{0,1\}$-matrix with largest eigen value $\lambda$ that

$$\max\{h_\mu(\sigma) : \mu \text{ is an invariant probability Borel measure on } \Sigma_M\} = \log(\lambda)$$

which is the entropy of the Parry measure built from $M$. So the Parry measure is the <u>measure of maximal entropy.</u>