Homework 6 (2023) — SOLUTIONS

JAIST — School of Information Science — I232 Information Theory

1. A Markov chain has two states $\{0, 1\}$. The probability of going from state 0 to 1 is p. The probability of going from state 1 to 0 is $(1-p)/2$, for $0 \le p \le 1$. Find the entropy rate of this Markov chain.

Solution: With the probability transition matrix,

$$\mathbf{P} = \begin{bmatrix} 1-p & p \\ \frac{1-p}{2} & \frac{1+p}{2} \end{bmatrix},$$

calculate the stationary distribution $\mathbf{z}$.

$$\mathbf{z} = \mathbf{z}\mathbf{P}$$
$$\mathbf{z}(\mathbf{P} - \mathbf{I}_2) = \mathbf{0}$$

Replace the first column of $(\mathbf{P} - \mathbf{I}_2)$ to all-one column denoted as $\tilde{\mathbf{Q}}$,

$$\tilde{\mathbf{Q}} = \begin{bmatrix} 1 & p \\ 1 & \frac{p-1}{2}, \end{bmatrix}$$

then

$$\mathbf{z} = (1, 0) \cdot \tilde{\mathbf{Q}}^{-1}$$
$$\mathbf{z} = (\frac{1-p}{1+p}, \frac{2p}{1+p}).$$

Then the entropy rate is given as

$$\begin{aligned} H(\mathcal{X}) &= -\sum_{i=1}^{2} z_i \sum_{j=1}^{2} P_{i,j} \log P_{i,j} \\ &= \frac{1-p}{1+p} h(p) + \frac{2p}{1+p} h(\frac{1-p}{2}), \end{aligned}$$

where $h(p) = -(p \log_2(p) + (1-p)\log_2(1-p))$ is the binary entropy function.

2. DNA, which encodes genetic instructions in all living things, uses a code over a quaternary (4-ary) alphabet. This code is abbreviated $\mathcal{X} = \{$ A, C, G, T $\}$. DNA consists of long strings from this alphabet: $\mathsf{X}_1, \mathsf{X}_2, \mathsf{X}_3, \ldots$. A possible first-order Markov model is given by (`https://bit.ly/2yxMzCf`):

	X_{n+1}			
X_n	A	C	G	T
A	0.180	0.274	0.426	0.120
C	0.171	0.367	0.274	0.188
G	0.161	0.339	0.375	0.125
T	0.079	0.355	0.384	0.182

   1. Find the steady-state distribution of this Markov chain.

Solution: The system of equations can be written as:

$$\mathbf{z} = \mathbf{z}\mathbf{P}$$
$$\mathbf{z}(\mathbf{P} - \mathbf{I}_4) = 0$$
$$(z_1, z_2, z_3, z_4)\begin{bmatrix} 0.180-1 & 0.274 & 0.426 & 0.120 \\ 0.171 & 0.367-1 & 0.274 & 0.188 \\ 0.161 & 0.339 & 0.375-1 & 0.125 \\ 0.079 & 0.355 & 0.384 & 0.182-1 \end{bmatrix} = 0$$

$z_i = 0$ is a trivial solution. Instead, add the constraint $\sum z_i = 1$ by *replacing* one column with 1's

$$(z_1, z_2, z_3, z_4)\begin{bmatrix} 0.180-1 & 0.274 & 0.426 & 1 \\ 0.171 & 0.367-1 & 0.274 & 1 \\ 0.161 & 0.339 & 0.375-1 & 1 \\ 0.079 & 0.355 & 0.384 & 1 \end{bmatrix} = (0, 0, 0, 1)$$
$$(0, 0, 0, 1) \cdot \begin{bmatrix} 0.180-1 & 0.274 & 0.426 & 1 \\ 0.171 & 0.367-1 & 0.274 & 1 \\ 0.161 & 0.339 & 0.375-1 & 1 \\ 0.079 & 0.355 & 0.384 & 1 \end{bmatrix}^{-1} = (z_1, z_2, z_3, z_4)$$
$$(z_1, z_2, z_3, z_4) = (0.155, 0.341, 0.350, 0.154)$$

which is the solution. Alternatively, solve the eigenvector problem $\mathbf{z} = \mathbf{z}P$ using Matlab `[v, d] = eig(P') ; z = v(:,1) / sum( v(:,1));`

2. Find the entropy of X.

Solution:

$$\begin{aligned} H(X) &= -\sum_{i_1}^{4} z_i \log z_i \\ &= -0.155 \log 0.155 - 0.341 \log 0.341 - 0.350 \log 0.350 - 0.154 \log 0.154 \\ &= 1.892 \text{ bits} \end{aligned}$$

3. Find the entropy rate $H(\mathcal{X})$ of this process.

Solution:

$$\begin{aligned} H(\mathcal{X}) &= -\sum_{i=1}^{4} z_i \sum_{j=1}^{4} P_{i,j} \log P_{i,j} \\ &= 0.4105 + 0.5272 + 0.5232 + 0.4117 \\ &= 1.8725 \ \text{ bits} \end{aligned}$$

In Matlab, `-sum( z' * (P .* log2(P)) )`. Here $\mathbf{z}$ is assumed as column vector.

3. Under what conditions is a Markov chain a stationary process?

Solution: A Markov chain is a stationary process if (1) it is time-invariant, (2) the initial distribution is equal to the steady-state distribution. That is,

$$\Pr[\mathsf{X}_1 = i] = \Pr[\mathsf{X}_{1+t} = i],$$
$$\Pr[\mathsf{X}_2 = j | \mathsf{X}_1 = i] = \Pr[\mathsf{X}_{2+t} = j | \mathsf{X}_{1+t} = i].$$

for all i, j, and t.