LECTURE 4

Last time:

- Types of convergence
- Weak Law of Large Numbers
- Strong Law of Large Numbers
- Asymptotic Equipartition Property

Lecture outline

- Stochastic processes
- Markov chains
- Entropy rate
- Random walks on graphs
- Hidden Markov models

Reading: Chapter 4.

Stochastic processes

A stochastic process is an indexed sequence or r.v.s $X_0, X_1, \ldots$ characterized by the joint PMF $P_{X_0,X_1,\ldots,X_n}(x_0, x_1, \ldots, x_n)$, $(x_0, x_1, \ldots, x_n) \in \mathcal{X}^n$ for $n = 0, 1, \ldots$.

A stochastic process is stationary if

$$P_{X_0,X_1,\ldots,X_n}(x_0, x_1, \ldots, x_n)$$

$$= P_{X_l,X_{l+1},\ldots,X_{l+n}}(x_0, x_1, \ldots, x_n)$$

for every shift l and all $(x_0, x_1, \ldots, x_n) \in \mathcal{X}^n$.

Stochastic processes

A discrete stochastic process is a Markov chain if

$$P_{X_n|X_0,\ldots,X_{n-1}}(x_n|x_0,\ldots,x_{n-1}) = P_{X_n|X_{n-1}}(x_n|x_{n-1})$$

for $n = 1, 2, \ldots$ and all $(x_0, x_1, \ldots, x_n) \in \mathcal{X}^n$.

We deal with time invariant Markov chains

X_n: state after n transitions

- belongs to a finite set, e.g., $\{1, \ldots, m\}$
- X_0 is either given or random

(given current state, the past does not matter)

$$\begin{aligned} p_{i,j} &= \mathbf{P}(X_{n+1} = j \mid X_n = i) \\ &= \mathbf{P}(X_{n+1} = j \mid X_n = i,\ X_{n-1}, \ldots, X_0) \end{aligned}$$

Markov chain is characterized by probability transition matrix $\underline{P} = [p_{i,j}]$

Review of Markov chains

State occupancy probabilities, given initial state i:

$$r_{i,j}(n) = \mathbf{P}(X_n = j \mid X_0 = i)$$

Key recursion:

$$r_{i,j}(n) = \sum_{k=1}^{m} r_{i,k}(n-1)p_{k,j}$$

With random initial state:

$$\mathbf{P}(X_n = j) = \sum_{i=1}^{m} \mathbf{P}(X_0 = i)r_{i,j}(n)$$

Does r_{ij} converge to something?

Does the limit depend on initial state?

Review of Markov chains

Recurrent and transient states.

State i is **recurrent** if: starting from i, and from wherever you can go, there is a way of returning to i. If not recurrent, called **transient**. Recurrent class collection of recurrent states that "communicate" to each other and to no other state.

A recurrent state is **periodic** if: there is an integer $d > 1$ such that $r_{i,i}(k) = 0$ when k is not an integer multiple of d

Assume a single class of recurrent states, aperiodic. Then,

$$\lim_{n \to \infty} r_{i,j}(n) = \pi_j$$

where π_j does not depend on the initial conditions

$$\lim_{n \to \infty} \mathbf{P}(X_n = j \mid X_0) = \pi_j$$

- $\pi_1, \ldots, \pi_m$ can be found as the unique solution of the balance equations

$$\pi_j = \sum_k \pi_k p_{k,j}$$

together with

$$\sum_j \pi_j = 1$$

Entropy rate

The entropy rate of a stochastic process is

$$\lim_{n\to\infty} \frac{1}{n} H(\underline{X}^n)$$

if it exists

For a stationary stochastic process, the entropy rate exists and is equal to

$$\lim_{n\to\infty} H(X_n|\underline{X}^{n-1})$$

since conditioning decreases entropy and by stationarity, it holds that

$$\begin{aligned} H(X_{n+1}|\underline{X}^n) &\leq H(X_{n+1}|\underline{X}_2^n) \\ &= H(X_n|\underline{X}^{n-1}) \end{aligned}$$

so it reaches a limit (decreasing non-negative sequence)

Chain rule

$$\frac{1}{n} H(\underline{X}^n) = \frac{1}{n} \sum_{i=1}^{n} H(X_i|\underline{X}^{i-1})$$

since the elements in the sum on the RHS reach a limit, that is the limit of the LHS

Entropy rate

Markov chain entropy rate:

$$
\begin{aligned}
& \lim_{n\to\infty} H(X_n|\underline{X}^{n-1}) \\
= & \lim_{n\to\infty} H(X_n|X_{n-1}) \\
= & H(X_2|X_1) \\
= & -\sum_{i,j} p_{i,j}\pi_i \log(p_{i,j})
\end{aligned}
$$

Random walk on graph

Consider undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E}, \mathcal{W})$ where $\mathcal{N}, \mathcal{E}, \mathcal{W}$ are the nodes, edges and weights. With each edge there is an associated edge weight $W_{i,j}$

$$\begin{aligned} W_{i,j} &= W_{j,i} \\ W_i &= \sum_j W_{i,j} \\ W &= \sum_{i,j:j>i} W_{i,j} \\ 2W &= \sum_i W_i \end{aligned}$$

Random walk on graph

We call a random walk the Markov chain in which the states are the nodes of the graph

$$p_{i,j} = \frac{W_{i,j}}{W_i}$$

$$\pi_i = \frac{W_i}{2W}$$

Check: $\sum_i \pi_i = 1$ and

$$\begin{aligned}
\sum_i \pi_i p_{i,j} &= \sum_i \frac{W_i}{2W}\frac{W_{i,j}}{W_i} \\
&= \sum_i \frac{W_{i,j}}{2W} \\
&= \frac{W_j}{2W} \\
&= \pi_j
\end{aligned}$$

Random walk on graph

$$
\begin{aligned}
& H(X_2|X_1) \\
= & -\sum_i \pi_i \sum_j p_{i,j} \log(p_{i,j}) \\
= & -\sum_i \frac{W_i}{2W} \sum_j \frac{W_{i,j}}{W_i} \log\left(\frac{W_{i,j}}{W_i}\right) \\
= & -\sum_{i,j} \frac{W_{i,j}}{2W} \log\left(\frac{W_{i,j}}{W_i}\right) \\
= & -\sum_{i,j} \frac{W_{i,j}}{2W} \log\left(\frac{W_{i,j}}{2W}\right) \\
+ & \sum_{i,j} \frac{W_{i,j}}{2W} \log\left(\frac{W_i}{2W}\right) \\
= & -\sum_{i,j} \frac{W_{i,j}}{2W} \log\left(\frac{W_{i,j}}{2W}\right) + \sum_i \frac{W_i}{2W} \log\left(\frac{W_i}{2W}\right)
\end{aligned}
$$

Entropy rate is difference of two entropies

Note: time reversibility for Markov chain that can be represented as random walk on undirected weighted graph

Hidden Markov models

Consider an ALOHA wireless model

$\mathcal{M}$ users sharing the same radio channel to transmit packets to a base station

During each time slot, a packet arrives to a user's queue with probability p, independently of the other $\mathcal{M} - 1$ users

Also, at the beginning of each time slot, if a user has at least one packet in its queue, it will transmit a packet with probability q, independently of all other users

If two packets collide at the receiver, they are not successfully transmitted and remain in their respective queues

Hidden Markov models

Let $X_i = (N[1]_i, N[2]_i, \ldots, N[\mathcal{M}]_{\rangle})$ denote the random vector at time i where $N[m]_i$ is the number of packets that are in user m's queue at time i. X_i is a Markov chain.

Consider the random vector $Y_i = (Z[1], Z[2], \ldots, Z[\mathcal{M}])$ where $Z[m]_i = 1$ if user m transmits during time slot i and $Z[i] = 0$ otherwise

Is Y_i Markov?

Hidden Markov processes

Let $X_i, X_2, \ldots$ be a stationary Markov chain and let $Y_i = \phi(X_i)$ be a process, each term of which is a function of the corresponding state in the Markov chain

$Y_1, Y_2, \ldots$ form a hidden Markov chain, which is not always a Markov chain, but is still stationary

What is its entropy rate?

Hidden Markov processes

We suspect that the effect of initial information should decay

$$H(Y_n|\underline{Y}^{n-1}) - H(Y_n|\underline{Y}^{n-1}, X_1) = I(X_1; Y_n|\underline{Y}^{n-1})$$

should go to 0

Indeed,

$$\begin{aligned}\lim_{n\to\infty} I\left(X_1; \underline{Y}^n\right) &= \lim_{n\to\infty} \sum_{i=1}^{n} I\left(X_1; \underline{Y}^i|\underline{Y}^{i-1}\right) \\ &= \sum_{i=1}^{\infty} I\left(X_1; \underline{Y}^i|\underline{Y}^{i-1}\right)\end{aligned}$$

since we have an infinite sum in which the terms are non-negative and which is upper bounded by $H(X_1)$, the terms must tend to 0

6.441 Information Theory
Spring 2010