EECS 501
Discussion 10

1 Review

Random Process: an infinite (countable/uncountable) collection of random variables $\{X_n\}_{n\in I}$.

- Discrete-time: I is countable
- Continuous-time: I is uncountable

Markov Chains:

- Let $\{X_k\}$ be a discrete-time random process that takes on values in a countable set S called the state space. $\{X_k\}$ is said to be a Markov Chain if for all $n, x_0, x_1, x_2, \cdots, x_{n+1}$ we have,

$$P(X_{n+1} = x_{n+1}|X_n = x_n, \cdots, X_0 = x_0) = P(X_{n+1} = x_{n+1}|X_n = x_n)$$

- A Markov Chain is said to be time-homogeneous if $P(X_{n+1} = j|X_n = i)$ depends only on i and j. We define transition probability $p_{ij} = P(X_{n+1} = j|X_n = i)$.
- Let $P = [p_{i,j}]$ be the transition probability matrix. we call vector π a stationary distribution if

$$\pi = \pi P$$

  For finding the stationary distribution, we can use balance equations:

$$\pi_j = \sum_k \pi_k p_{kj} \ \forall j$$

$$1 = \sum_k \pi_k$$

- Reachable: state j is called **reachable** from state i if there exists time T such that

$$P(X_T = j|X_0 = i) > 0$$

  i.e. there exists a nonzero probability to go to state j from state i over a finite number of steps.
- Irreducible: a Markov chain is **irreducible** if j is reachable from $i \quad \forall i, j$.
- Period: state i is said to have a **period** k if the MC returns to state i in T steps only if T is a multiple of k:

$$k = gcd(\{T_1, T_2, \cdots\} \text{ such that } P(X_{t+T_j} = i|X_t = i) > 0,\ \forall T_j)$$

- Aperiodic: a Markov chain is **aperiodic** if all states have period 1.

- Every state in an irreducible Markov chain has the same period. Thus, in an irreducible Markov chain, if one state is aperiodic, the Markov chain is aperiodic.
- A **finite-state, irreducible** MC has a **unique** stationary distribution π such that

$$\pi = \pi P.$$

- A **finite-state, aperiodic, irreducible** MC has a **unique** stationary distribution π such that

$$\pi = \pi P.$$

  Furthermore,

$$\lim_{k\to\infty} p(k) = \pi$$

- The recurrence time T_i of state i of a Markov chain is defined as

$$T_i = \min\{n \geq 1 : X_n = i \text{ given } X_0 = i\}.$$

- A state i is said to be *recurrent* if $P(T_i < \infty) = 1$. Otherwise, it is called *transient.*
- The mean *recurrence time* M_i of state i is defined as $M_i = E[T_i]$.
- A recurrent state i is called *positive recurrent* if $M_i < \infty$. Otherwise, it is called *null recurrent.*
- A Markov chain is called *positive recurrent* if all of its states are positive recurrent.
- Suppose $\{X_k\}$ is irreducible and that one of its states is positive recurrent, then all of its states are positive recurrent. (The same statement holds if we replace positive recurrent by null recurrent or transient.)
- Consider a time-homogeneous Markov chain which is irreducible and aperiodic. Then, the following results hold.
  - If the Markov chain is positive recurrent, there exists a unique π such that $\pi = \pi P$ and $\lim_{k\to\infty} p(k) = \pi$. Further, $\pi_i = 1/M_i$.
  - If there exists a positive vector π such that $\pi = \pi P$ and $\sum_i \pi_i = 1$, it must be the stationary distribution and $\lim_{k\to\infty} p(k) = \pi$. (From the lemma above, this also means that the Markov chain is positive recurrent.)
  - If there exists a positive vector π such that $\pi = \pi P$, and $\sum_i \pi_i$ is infinite, a stationary distribution does not exist, and $\lim_{k\to\infty} p_i(k) = 0$ for all i.

2 Practice Problems

Problem 1 Let $P = \begin{bmatrix} 0.25 & 0.75 \\ 0.2 & 0.8 \end{bmatrix}$ be the transition probability matrix of a Markov Chain. Find the stationary distribution.

Solution:

Using balance equation ($\pi P = \pi$), we get $\pi = [\frac{4}{19} \ \frac{15}{19}]$.

Problem 2 A particle moves on the grid in Fig. 1. Time is discrete and at each time the particle can move to any of its neighboring grid points with equal probability. If $\{X_n\}_{n\geq 1}$ process denotes the position of the particle then note that it is a Markov Chain. Find its stationary distribution.

Solution:

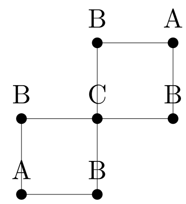

Figure 1: Particle on a grid

By symmetry, all the points of type A have the same steady state distribution π_A. Also, Points of type B have the same state distribution π_B. We have

$$2\pi_A + 4\pi_B + \pi_C = 1 \tag{1}$$

Now using the standard balancing relation for π we have

$$\pi_A = \frac{1}{2}\pi_B + \frac{1}{2}\pi_B \tag{2}$$
$$\pi_B = \frac{1}{2}\pi_A + \frac{1}{4}\pi_C \tag{3}$$

(2) and (3) together give $\pi_B = \pi_A$ and $\pi_C = 2\pi_A$. Substituting all these in (1) gives

$$2\pi_A + 4\pi_A + 2\pi_A = 1 \Rightarrow \quad \pi_A = \frac{1}{8} \quad \pi_B = \frac{1}{8} \quad \pi_C = \frac{1}{4} \tag{4}$$

Problem 3 For a discrete time homogeneous Markov chain $\{X_i\}_{i\geq 1}$, prove that the process $\{Y_i\}_{i\geq 1}$ is Markov in both the following cases,

1. $Y_i = X_{2i}$ for $i \geq 1$,
2. $Y_i = (X_i, X_{i+1})$ for $i \geq 1$.

Solution:

1.

$$\begin{aligned} P(Y_{i+1} = a \mid Y_i = b, Y_{i-1} = c) &= P(X_{2i+2} = a \mid X_{2i} = b, X_{2i-2} = c) && (5)\\ &= P(X_{2i+2} = a \mid X_{2i} = b) && (6)\\ &= \sum_j p_{bj}p_{ja} && (7)\end{aligned}$$

where the second line uses the Markov property for X. Clearly the above expression is independent of c.

2. For any a, b, c, d, e

$$\begin{aligned} &P(Y_{i+1} = (a,b) \mid Y_i = (d,c), Y_{i-1} = (e,d)) && (8)\\ &= P(X_{i+1} = a, X_{i+2} = b \mid X_{i+1} = c, X_i = d, X_{i-1} = e) && (9)\\ &= \mathbb{1}_c(a)P(X_{i+2} = b \mid X_{i+1} = c, X_i = d, X_{i-1} = e) && (10)\\ &= \mathbb{1}_c(a)p_{cb} && (11)\end{aligned}$$

where $\mathbb{1}_c(a)$ is equal to 1 if $a = c$ and is equal to 0 otherwise. Clearly the above expression is independent of e.

Problem 4 *Mean recurrence time (short term behavior of Markov Chains)*

Consider a discrete time homogeneous aperiodic recurrent Markov chain with state space S. Let $s \in S$ be a state. Denote by τ_i the expected time to reach the state s starting from the state i. Denote by M_s the mean recurrence time for state s, i.e. the expected time to return to state s, and finally let $\pi = \big(\pi(k)\big)_{k\geq 0}$ be the stationary distribution of the Markov chain. Prove the following relations:

$$\tau_i = \begin{cases} 1 + \sum_{j\in S} p_{ij}\tau_j & \text{if } i \neq s \\ 0 & \text{if } i = s \end{cases} \tag{12a}$$

$$M_s = 1 + \sum_{j\in S} p_{sj}\tau_j \tag{12b}$$

Then, derive the following relation between the stationary distribution and the mean recurrence time

$$\pi(s)M_s = 1 \tag{13}$$

Solution:

Define $T_i = min\{T \geq 0 : X_{t+T} = s | X_t = i\}$. For $i = s$, we clearly have $T_i = 0$. For $i \neq s$, and by using law of total expectation, we have

$$\tau_i = E[T_i] = 1 + \sum_{j\in S} E[T_i|\text{go to state j}]p_{ij} = 1 + \sum_{j\in S} E[T_j]p_{ij} = 1 + \sum_{j\in S} \tau_j p_{ij}$$

and the second equation follows by substituting $i = s$ and noticing the fact that we need to get out of state s and come back to it and therefore, $M_s > 0$. For the third result, we have

$$\pi(s)M_s + \sum_{i\neq s}\pi(i)\tau_i = \pi(s)M_s + \sum_{i\neq s}\pi(i)\Big(1 + \sum_j p_{ij}\tau_j\Big) \qquad \text{(by (12a))} \tag{14}$$

$$= \pi(s)M_s + \sum_{i\neq s}\pi(i) + \sum_{i\neq s}\pi(i)\sum_j p_{ij}\tau_j \tag{15}$$

$$= \pi(s)\Big(1 + \sum_j p_{js}\tau_j\Big) + \sum_{i\neq s}\pi(i) + \sum_{i\neq s}\pi(i)\sum_j p_{ji}\tau_j \quad \text{(by (12b))} \tag{16}$$

$$= \sum_i \pi(i) + \pi(s)\sum_j p_{sj}\tau_j + \sum_{i\neq s}\pi(i)\sum_j p_{ij}\tau_j \tag{17}$$

$$= \sum_i \pi(i) + \sum_i \pi(i)\sum_j p_{ij}\tau_j \tag{18}$$

$$= 1 + \sum_j \tau_j \sum_i \pi(i)p_{ij} \tag{19}$$

$$= 1 + \sum_j \tau_j\pi(j) \quad \text{(by balance equations} \tag{20}$$

$$= 1 + \sum_{j\neq s}\tau_j\pi(j) \quad \text{(since } \tau_s = 0 \text{ by (12a))} \tag{21}$$

Thus the above finally gives $\pi(s)M_s + \sum_{i\neq s}\pi(i)\tau_i = 1 + \sum_{j\neq s}\tau_j\pi(j)$ which gives $\pi(s)M_s = 1$.

Problem 5 *Random Walk on $\mathbb{Z}_+$ (Gambler's Ruin):* Assume a gambler starts with n units of money, where $0 \le n \le N$. He bets a single unit of money at every time step and wins with probability p or loses with probability $q = 1 - p$ (independently at every time step). The gambler's goal is to have a total of N units of money and he keeps betting until he gets there (and then stops gambling) or until he is ruined i.e. has 0 units of money and thus can't continue to gamble. Find the probability of winning. What happens as $N \to \infty$?

Solution:

Denote the total money after i time steps by X_i then we can write

$$X_i = X_0 + \sum_{k=1}^{i} Z_k \tag{Random Walk}$$

$$X_0 = n \tag{22}$$

where $Z_1, Z_2, \ldots$ are a sequence of i.i.d. random variables with distribution $\mathbb{P}(Z_i = 1) = p$ and $\mathbb{P}(Z_i = -1) = q$. Denote by τ_n the time that the gambling is stopped, i.e.

$$\tau_n = \min\big\{i \ge 1 : X_i \in \{0, N\} \mid X_0 = n\big\} \tag{23}$$

If $X_{\tau_n} = N$ then the gambler wins, if $X_{\tau_n} = 0$ then the gambler is ruined. Let $\alpha_n = \mathbb{P}(X_{\tau_n} = N)$ be the probability of winning when initial money is n units.

Due to X being Markov process, for any $1 \le m \le N - 1$ we can write

$$\alpha_m = p\alpha_{m+1} + q\alpha_{m-1} \tag{24}$$

Clearly we also have $\alpha_0 = 0$ and $\alpha_N = 1$. The above relation and the boundary conditions can now be used to evaluate α_n for any $1 \le n \le N - 1$.

(24) can be rewritten as $p\big(\alpha_{m+1} - \alpha_m\big) = q\big(\alpha_m - \alpha_{m-1}\big)$, which gives

$$\alpha_{m+1} - \alpha_m = \frac{q}{p}\big(\alpha_m - \alpha_{m-1}\big) \tag{25}$$

$$\Rightarrow \quad \alpha_{m+1} - \alpha_m = \left(\frac{q}{p}\right)^m (\alpha_1 - \alpha_0) \tag{26}$$

$$= \left(\frac{q}{p}\right)^m \alpha_1 \tag{27}$$

Thus we can write

$$\alpha_m = \sum_{j=1}^{m} (\alpha_j - \alpha_{j-1}) = \sum_{j=1}^{m} \left(\frac{q}{p}\right)^{j-1} \alpha_1 = \begin{cases} \frac{1-\left(\frac{q}{p}\right)^m}{1-\frac{q}{p}}\alpha_1 & \text{if } p \ne \frac{1}{2} \\ m\alpha_1 & \text{if } p = \frac{1}{2} \end{cases} \tag{28}$$

Finally the boundary condition $\alpha_N = 1$ gives

$$\alpha_1 = \begin{cases} \frac{1-\frac{q}{p}}{1-\left(\frac{q}{p}\right)^N} & \text{if } p \ne \frac{1}{2} \\ \frac{1}{N} & \text{if } p = \frac{1}{2} \end{cases} \tag{29}$$

Thus the solution for any $0 \le n \le N$ is

$$\boxed{\alpha_n = \begin{cases} \frac{1-\left(\frac{q}{p}\right)^n}{1-\left(\frac{q}{p}\right)^N} & \text{if } p \ne \frac{1}{2} \\ \frac{n}{N} & \text{if } p = \frac{1}{2} \end{cases}} \tag{30}$$

In the limit $N \to \infty$ there is a bifurcation in the behavior of the quantity α_n. Note that if $p > \frac{1}{2}$ then $\frac{q}{p} < 1$ and thus we have

$$\lim_{N\to\infty} \alpha_n = \begin{cases} 1 - \left(\frac{q}{p}\right)^n & \text{if } p > \frac{1}{2} \\ 0 & \text{if } p \leq \frac{1}{2} \end{cases} \tag{31}$$

Hence for any initial money $n \geq 1$ if the gambler continues to gamble, for $p \leq \frac{1}{2}$ he will lose all the money with probability 1 and for $p > \frac{1}{2}$ he can win infinite money with strictly positive probability.

Problem 6 *Imbedding a Markov Chain:* Consider a random process $X_1, X_2, X_3, \ldots$ where the odd numbered random variables $X_1, X_3, X_5 \ldots$ are i.i.d. Rademacher (i.e. take values in $\{-1, +1\}$ with equal probability). The even numbered random variables have the relation $X_{2m} = X_{2m-1}X_{2m+1}$ for $m \geq 1$. Show that

1. $\{X_i\}_{i\geq 1}$ is not a Markov process.
2. $\{Y_i\}_{i\geq 1}$ is a Markov process, where $Y_i \triangleq \big(X_i, X_{i+1}\big)$ for $i \geq 1$.

Solution:

1. Intuitively the process $\{X_i\}_{i\geq 1}$ is not Markov since the relation governing even numbered random variables has two-step dependency. More formally we can see from the relation $X_2 = X_1X_3$ that

$$\mathbb{P}(X_3 = 1 \mid X_2 = 1, X_1 = 1) = 1 \tag{32}$$
$$\mathbb{P}(X_3 = 1 \mid X_2 = 1, X_1 = -1) = 0 \tag{33}$$

which tells us that the distribution of X_3 given $X_2 = 1$ is NOT independent of the past.

2. Intuitively the process Y is Markov because it represents X with enough "memory" so that the two-step dependency $X_{2m} = X_{2m-1}X_{2m+1}$ takes on the one-step structure when written in terms of Y_i's. For $i = 2m$, $Y_{2m} = (X_{2m}, X_{2m+1})$ and $Y_{2m-1} = (X_{2m-1}, X_{2m})$ thus

$$Y_{2m}(1) = Y_{2m-1}(1)Y_{2m}(2) \quad \text{and} \quad Y_{2m-1}(2) = Y_{2m}(1) \tag{34}$$

where the indexes inside parenthesis refer to entry index for a vector.

More formally we have for any $a, b, c, d, e \in \{-1, +1\}$ and i odd,

$$\mathbb{P}(Y_{i+1} = (a, b) \mid Y_i = (c, d), Y_{i-1} = (e, c)) \tag{35}$$
$$= \mathbb{P}(X_{i+2} = b, X_{i+1} = a \mid X_{i+1} = d, X_i = c, X_{i-1} = e) \tag{36}$$
$$= \mathbb{1}_d(a)\mathbb{P}(X_{i+2} = b \mid X_{i+1} = d, X_i = c, X_{i-1} = e) \tag{37}$$

where $\mathbb{1}_d(a)$ is equal to 1 if $a = d$ and is equal to zero otherwise. All we want to show is that the above expression does not depend on e. But since i is odd then so is $i + 2$, which means X_{i+2} is generated independently of all past variables and thus the above expression is $\frac{1}{2}\mathbb{1}_d(a)$ (clearly does not depend on e).

For even i we consider any a, b, c, d, e that satisfy the relation $c = de$ (to ensure $X_i = X_{i-1}X_{i+1}$, else the conditioning event has zero probability) and again consider

$$\mathbb{P}(X_{i+2} = b, X_{i+1} = a \mid X_{i+1} = d, X_i = c, X_{i-1} = e) \tag{38}$$
$$= \mathbb{1}_d(a)\mathbb{P}(X_{i+2} = b \mid X_{i+1} = d, X_i = c, X_{i-1} = e) \tag{39}$$
$$= \mathbb{1}_d(a)\mathbb{P}(X_{i+3} = \frac{b}{d} \mid X_{i+1} = d, X_i = c, X_{i-1} = e) \tag{40}$$

and since $i+3$ is odd we know that X_{i+3} is generated independent of previous variables and hence the above expression is $\frac{1}{2}\mathbb{1}_d(a)$.

Problem 7 Let $\{X_n, n = 0, 1, 2, \dots\}$ be a random process given by

$$X_{n+1} = f(X_n, W_n)$$

where $f(x, w)$ is function of two variables and $\{W_n, n = 0, 1, 2, \dots\}$ is a sequence of independent random variables that are also independent of X_0. Show that $\{X_n, n = 1, 2, \dots\}$ is a Markov process.

Solution:

We first prove that X_n depends only on X_0 and $W_0, W_1, \dots, W_{n-1}$. That is, for any n there exists a function f_n such that $X_n = f_n(X_0, W_0, W_1, \dots, W_{n-1})$. This is done by induction.
We have $X_0 = X_0 = f_0(X_0)$. If $X_k = f_k(X_0, W_0, W_1, \dots, W_{k-1})$, for $n = k+1$ we have

$$X_{k+1} = f(X_k, W_k) = f(f_k(X_0, W_0, W_1, \dots, W_{k-1}), W_k) = f_{k+1}(X_0, W_0, W_1, \dots, W_k)$$

Using the above equation, since W_n is independent of $X_0, W_0, W_1, \dots, W_{n-1}$, W_n is also independent of $X_{n-1}, X_{n-2}, \dots, X_1, X_0$. Therefore, for any realization $x_0, x_1, \dots, x_n$ and any set A we have

$$\begin{aligned}
&P(X_{n+1} \in A | X_n = x_n, \dots, X_0 = x_0) \\
=&P(f(X_n, W_n) \in A | X_n = x_n, \dots, X_0 = x_0) \\
=&P(f(x_n, W_n) \in A | X_n = x_n, \dots, X_0 = x_0) \\
=&P(f(x_n, W_n) \in A) \text{ (Since } W_n \text{ is independent of } X_{n-1}, X_{n-2}, \dots, X_1, X_0 \text{)} \\
=&P(f(x_n, W_n) \in A | X_n = x_n) \\
=&P(f(X_n, W_n) \in A | X_n = x_n) \\
=&P(X_{n+1} \in A | X_n = x_n)
\end{aligned}$$

So $\{X_n, n = 0, 1, 2, \dots\}$ is Markov.