

APMA1690: Homework # 5 (Due by 11pm Oct 26)

“A drunk man will eventually find his way home, but a drunk bird may get lost forever.”

— Shizuo Kakutani

1 Review

Please read the review section before delving into the problem set.

1.1 Notations

- \mathbb{Z} = the collection of all integers.
- $\mathbb{Z}^d = \underbrace{\mathbb{Z} \times \mathbb{Z} \times \cdots \times \mathbb{Z}}_{\text{Cartesian product, } d \text{ times}}$.
- For a Markov chain¹ (MC) $\{X_n\}_{n=0}^\infty$, the subscript n is conventionally referred to as “time.” All the components X_n of the MC are random variables defined on an underlying probability space (Ω, \mathbb{P}) , i.e., $X_n : \Omega \rightarrow \mathcal{X}$.
- Let \mathbf{A} be a matrix. The [transpose](#) of \mathbf{A} is denoted as \mathbf{A}^\top .

1.2 Simple Random Walks

The following is the definition of the d -dimensional simple random walk (SRW), where $d \in \{1, 2, 3, \dots\}$.

Definition 1.1. Let $\xi_1, \xi_2, \dots, \xi_n, \dots$ be \mathbb{Z}^d -valued random variables, i.e.,

$$\begin{aligned}\xi_i : \Omega &\rightarrow \mathbb{Z}^d, \\ \omega &\mapsto \xi_i(\omega),\end{aligned}$$

for all $i = 1, 2, \dots$. Suppose the random variables $\xi_1, \xi_2, \dots, \xi_n, \dots$ are iid and satisfy

$$(1.1) \quad \mathbb{P}(\xi_i = \mathbf{e}_k) = \mathbb{P}(\xi_i = -\mathbf{e}_k) = \frac{1}{2d}, \quad \text{for all } k = 1, \dots, d,$$

where \mathbf{e}_k is the k -th axis of the d -dimensional space, i.e.,

$$\mathbf{e}_k = (0, \dots, 0, \underset{\substack{\uparrow \\ k^{th}}}{1}, 0, \dots, 0).$$

¹All Markov chains utilized throughout this semester are assumed to be homogeneous Markov chains (HMCs). The “homogeneous” will sometimes be suppressed for simplicity.

Particularly, when $d = 1$, we have $e_1 = 1$ and $-e_1 = -1$.

We define the sequence $\{X_n\}_{n=0}^\infty$ of random variables taking values in \mathbb{Z}^d as the following

$$\begin{aligned} X_0(\omega) &= x_0 \in \mathbb{Z}^d \quad \text{for all } \omega \in \Omega, \\ X_n(\omega) &= x_0 + \sum_{i=1}^n \xi_i(\omega), \quad \text{for all } n = 1, 2, \dots, \end{aligned}$$

where $x_0 \in \mathbb{Z}^d$ is a fixed point as an initialization. The sequence $\{X_n\}_{n=0}^\infty$ is called the *d-dimensional simple random walk*.

Each random variable ξ_i can be viewed as one step of the drunk man ($d = 2$) or the drunk bird ($d = 3$), and the initialization x_0 can be viewed as the “home” of the man/bird. The cumulation of the steps is the walking/flying path of the man/bird. Since the man/bird is drunk, each step ξ_i is totally random; thus, we have the “drunk step distribution” in Eq. (1.1).

1.3 Recurrence and Transience

Let $\{X_n\}_{n=0}^\infty$ be a homogeneous Markov chain taking values in the discrete state space \mathcal{X} , which can be either finite or infinite. For this Markov chain and each element $y \in \mathcal{X}$, we define the following random variable

$$\begin{aligned} T_y(\omega) &\stackrel{\text{def}}{=} \min \{n > 0 \mid X_n(\omega) = y\} \\ &= \text{the time at which the sequence } \{X_n(\omega)\}_{n=1}^\infty \text{ first visits } y. \end{aligned}$$

Here, we explicitly write down ω to emphasize that T_y is a random variable. We denote the following probability, which will be used to define “recurrence/transience” and “irreducibility.”

$$\begin{aligned} \rho_{xy} &\stackrel{\text{def}}{=} \mathbb{P}(T_y < \infty \mid X_0 = x) \\ &= \text{the conditional probability that MC will visit } y \text{ at least once, given that it starts from } x. \end{aligned}$$

With the notations $\{\rho_{xy}\}_{x,y \in \mathcal{X}}$, we can define recurrence and transience as follows.

Definition 1.2. Suppose we have a homogeneous Markov chain $\{X_n\}_{n=0}^\infty$. We provide the following two versions of the same definition.

1. (Rigorous/mathematical version) A state $y \in \mathcal{X}$ is said to be **recurrent** for this Markov chain if $\rho_{yy} = 1$; otherwise (i.e., $\rho_{yy} < 1$), the state y is said to be **transient** for this Markov chain.
2. (Heuristic version) If the Markov chain starting from y (i.e., $X_0 = y$) will almost surely (i.e., with probability one) return to y , the state y is said to be recurrent for this Markov chain; otherwise, y is said to be transient.

The following theorem gives an approach to identifying recurrence.

Theorem 1.1. Let $\{X_n\}_{n=0}^\infty$ be a homogeneous Markov chain taking values in the discrete state space \mathcal{X} . For any $y \in \mathcal{X}$, we have the following

1. The state y is recurrent for the Markov chain if and only if

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n = y | X_0 = y) = \infty.$$

2. If y is not recurrent, we have the following for all $x \in \mathcal{X}$

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n = y | X_0 = x) < \infty.$$

The proof of Theorem 1.1 involves the so-called “strong Markov property,” which is beyond the scope of this course. Hence, the proof of Theorem 1.1 is omitted to avoid a cheating proof.

The following theorem shows that recurrence is “contagious.”

Theorem 1.2. Let $\{X_n\}_{n=0}^{\infty}$ be a homogeneous Markov chain taking values in the discrete state space \mathcal{X} , which can be either finite or infinite. If x is recurrent (i.e., $\rho_{xx} = 1$) and $\rho_{xy} > 0$, then we have

1. $\rho_{yy} = 1$, i.e., the state y is recurrent;

2. $\rho_{xy} = \rho_{yx} = 1$.

That is, $\rho_{xx} = 1$ and $\rho_{xy} > 0$ implies $\rho_{xx} = \rho_{yy} = \rho_{xy} = \rho_{yx} = 1$.

The proof of Theorem 1.2 is given in my lecture notes (proof of Theorem 3.2.3 therein).

1.4 Irreducibility

Definition 1.3. Let $\{X_n\}_{n=0}^{\infty}$ is a homogeneous Markov chain taking values in the discrete state space \mathcal{X} and having transition probability p .

1. The Markov chain $\{X_n\}_{n=0}^{\infty}$ or transition probability p is said to be **recurrent** if all states of \mathcal{X} are recurrent for this Markov chain.

2. The Markov chain $\{X_n\}_{n=0}^{\infty}$ or transition probability p is said to be **irreducible** if $\rho_{xy} > 0$ for all $x, y \in \mathcal{X}$, i.e., we have a positive probability of transiting between any two states.

Because of the following theorem, the scenario where the state space \mathcal{X} is finite is important in the Markov chain theory.

Theorem 1.3. Let $\{X_n\}_{n=0}^{\infty}$ be a Markov chain taking values in the state space \mathcal{X} . If \mathcal{X} is finite (i.e., $\#\mathcal{X} < \infty$), we have

1. \mathcal{X} has at least one state that is recurrent for $\{X_n\}_{n=0}^{\infty}$;

2. furthermore, if $\{X_n\}_{n=0}^{\infty}$ is irreducible, $\{X_n\}_{n=0}^{\infty}$ is recurrent, i.e., all states in \mathcal{X} are recurrent for $\{X_n\}_{n=0}^{\infty}$.

1.5 Transition Matrices

Let $\{X_n\}_{n=0}^\infty$ be a homogeneous Markov chain taking values in $\mathcal{X} = \{x_1, x_2, \dots, x_S\}$ and having the transition probability p . We define the following **transition matrix** of the Markov chain

$$(1.2) \quad \mathbf{P} = \begin{pmatrix} p(x_1, x_1) & p(x_1, x_2) & \cdots & p(x_1, x_S) \\ p(x_2, x_1) & p(x_2, x_2) & \cdots & p(x_2, x_S) \\ \vdots & \vdots & \ddots & \vdots \\ p(x_S, x_1) & p(x_S, x_2) & \cdots & p(x_S, x_S) \end{pmatrix},$$

which is an $S \times S$ matrix.

2 Problem Set

1. Prove that the 1-dimensional simple random walk is not only a Markov chain but also a homogeneous Markov chain. Please derive the transition probability of the 1-dimensional simple random walk.

The 1-dimensional simple random walk is defined by

$$X_n(\omega) = \begin{cases} x_0 & n = 0 \\ x_0 + \sum_{i=1}^n \xi_i(\omega) & n = 1, 2, \dots \end{cases}$$

where

$$\xi_1, \xi_2, \dots, \xi_n \stackrel{iid}{\sim} \text{Unif}(\{-1, 1\}) \implies \mathbb{P}(\xi_i = 1) = \mathbb{P}(\xi_i = -1) = \frac{1}{2}$$

Observe:

$$\begin{aligned} X_{n-1}(\omega) &= x_0 + \sum_{i=1}^{n-1} \xi_i(\omega) \\ X_n(\omega) &= x_0 + \sum_{i=1}^n \xi_i(\omega) = X_{n-1}(\omega) + \xi_n(\omega) \end{aligned}$$

Then for two particular states $x, y \in \mathfrak{X}$, we can say

$$\mathbb{P}(X_n = y \mid X_{n-1} = x) = \mathbb{P}(\xi_n = y - x)$$

But since ξ_i are drawn iid, the probability does not depend on n – just on x and y :

$$\mathbb{P}(\xi_n = y - x) = \mathbb{P}(y - x = 1 \mid y - x \in \{-1, 1\}) = \mathbb{P}(y - x = -1 \mid y - x \in \{-1, 1\}) = \frac{1}{2}$$

Heuristically, this makes sense: by the recursive formula above, the Markov chain changes by ± 1 every time step with equal probability. So the outcome of being in a particular state one time step after being in an earlier state represents half of the state space.

And in fact, the above calculation is exactly the way to calculate the transition probability:

$$p(x, y) = \mathbb{P}(X_n = y \mid X_{n-1} = x) = \frac{1}{2}$$

Since this function does not depend on n , the MC is homogenous. ■

2. Prove that the point $0 \in \mathbb{Z}$ is recurrent for the 1-dimensional simple random walk.

You may consider applying the following results: (You do not need to prove any of the following results.)

- Theorem 1.1.
- Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be two sequences of nonnegative numbers. Suppose $\lim_{n \rightarrow \infty} (a_n/b_n) = 1$. Then, we have $\sum_{n=1}^{\infty} a_n < \infty$ if and only if $\sum_{n=1}^{\infty} b_n < \infty$.
- (Stirling's formula).

$$\lim_{n \rightarrow \infty} \frac{n!}{n^n \cdot e^{-n} \cdot \sqrt{2\pi n}} = 1.$$

By Theorem 1.1 the state $0 \in \mathbb{Z}$ is recurrent for the 1-d simple random walk if and only if

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n = 0 \mid X_0 = 0) = \infty$$

By the definition of this 1-d SRW

$$X_n(\omega) = \sum_{i=1}^n \xi_i(\omega)$$

with $\xi_i(\omega) \stackrel{iid}{\sim} \text{Unif}(\{-1, 1\})$. So

$$\mathbb{P}(X_n = 0) = \mathbb{P}\left(\sum_{i=1}^n \xi_i(\omega) = 0\right)$$

So when does $\sum_{i=1}^n \xi_i(\omega) = 0$? Clearly, it must be after an even number of steps. So, we need to calculate the probability that out of the $2n$ steps of the chain, exactly n of them are in one particular direction (WLOG say $\xi = 1$). Thus,

$$\mathbb{P}\left(\sum_{i=1}^n \xi_i(\omega) = 0\right) = \frac{1}{2^{2n}} \binom{2n}{n} = \frac{1}{2^{2n}} \cdot \frac{(2n)!}{n!(2n-n)!} = \frac{1}{2^{2n}} \cdot \frac{(2n)!}{n! \cdot n!}$$

By Sterling's formula,

$$\frac{1}{2^{2n}} \cdot \frac{(2n)!}{n! \cdot n!} \approx \frac{1}{2^{2n}} \cdot \frac{(2n)^{2n} e^{-2n} \sqrt{4\pi n}}{[n^n e^{-n} \sqrt{2\pi n}]^2} \approx \frac{1}{\sqrt{\pi n}}$$

(for large n).

So to review, we have shown that

$$\mathbb{P}(X_n = 0 \mid X_0 = 0) = \mathbb{P}\left(\sum_{i=1}^n \xi_i(\omega) = 0\right) = \frac{1}{\sqrt{\pi n}}$$

So

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n = 0 \mid X_0 = 0) = \sum_{n=1}^{\infty} \frac{1}{\sqrt{\pi n}}$$

But by the integral convergence test,

$$\int_1^{\infty} \frac{1}{\sqrt{\pi n}} \, dn = \frac{1}{\sqrt{\pi}} \int_1^{\infty} n^{-\frac{1}{2}} \, dn = \frac{1}{\sqrt{\pi}} [2\sqrt{n}]_1^{\infty} = \infty$$

So the sum is ∞ . Thus by Theorem 1.1, $0 \in \mathbb{Z}$ is recurrent. ■

3. (1 point) Let $\{X_n\}_{n=0}^\infty$ be a homogeneous Markov chain taking values in the state space $\mathcal{X} = \{x_1, x_2, \dots, x_S\}$ and having transition probability p .

- Let \mathbf{P} denote the transition matrix of the Markov chain (see Eq. (1.2)).
- $\pi : \mathcal{X} \rightarrow [0, 1]$ is a probability mass function. Denote the column vector $(\pi(x_1), \pi(x_2), \dots, \pi(x_S))^T$ as $\boldsymbol{\pi}$.

Prove that the matrix equation $\boldsymbol{\pi} = \mathbf{P}^T \boldsymbol{\pi}$ is equivalent to the following

$$\pi(x) = \sum_{y \in \mathcal{X}} \pi(y) \cdot p(y, x), \quad \text{for all } x \in \mathcal{X}.$$

$$\begin{aligned} \vec{\pi} &= P^T \vec{\pi} \\ &= \begin{pmatrix} p(x_1, x_1) & p(x_2, x_1) & \cdots & p(x_S, x_1) \\ p(x_1, x_2) & p(x_2, x_2) & \cdots & p(x_S, x_2) \\ \vdots & \vdots & \ddots & \vdots \\ p(x_1, x_S) & p(x_2, x_S) & \cdots & p(x_S, x_S) \end{pmatrix} \begin{pmatrix} \pi(x_1) \\ \pi(x_2) \\ \vdots \\ \pi(x_S) \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^S p(x_i, x_1) \cdot \pi(x_i) \\ \sum_{i=1}^S p(x_i, x_2) \cdot \pi(x_i) \\ \vdots \\ \sum_{i=1}^S p(x_i, x_S) \cdot \pi(x_i) \end{pmatrix} \end{aligned}$$

Then since $\mathfrak{X} = \{x_1, x_2, \dots, x_S\}$,

$$\vec{\pi}_i = \pi(x_i) = \sum_{j=1}^S p(x_i, x_j) \cdot \pi(x_j) = \sum_{y \in \mathfrak{X}} \pi(y) \cdot p(y, x_i)$$

So $\forall x \in \mathfrak{X}$,

$$\pi(x) = \sum_{y \in \mathfrak{X}} \pi(y) \cdot p(y, x) \quad \blacksquare$$