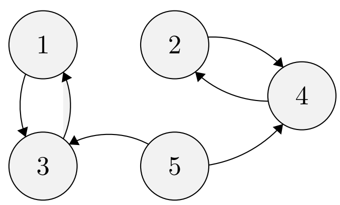
STAT 433: HOMEWORK 3

DANIEL PFEFFER

Problem 1. Consider the following transition matrices. Identify the transient and recurrent states, and the irreducible closed sets in the Markov chains.

(i) ???

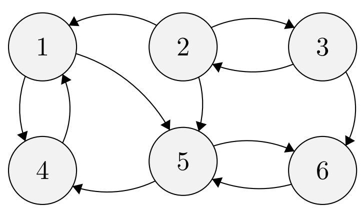
Solution. The directed graph representation is



Since $1 \to 2$ but $2 \not\to 1$, state 1 is transient. Since $5 \to 4$ but $4 \not\to 5$, state 5 is transient. Since $3 \to 2$ but $2 \not\to 3$, state 3 is transient. So, states 2 and 4 are recurrent. The closed irreducible set is $\{2, 4\}$. □

(ii) ???

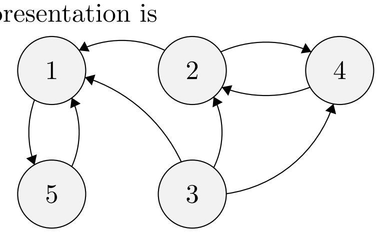
Solution. The directed graph representation is



Since $2 \to 1$ but $1 \not\to 2$, state 2 is transient. Since $3 \to 6$ but $6 \not\to 3$, state 3 is transient. So states 1, 4, 5, and 6 are recurrent. The closed irreducible set is $\{1, 4, 5, 6\}$. □

(iii) ???

Solution. The directed graph representation is

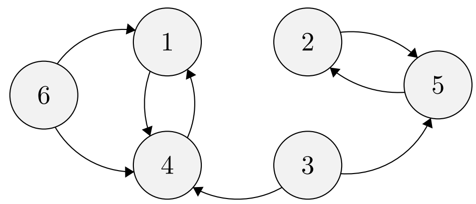


Since $3 \to 1$ but $1 \not\to 3$, state 3 is transient. So state 1, 2, 4, and 5 are recurrent. The closed irreducible sets are $\{1, 5\}$ and $\{2, 4\}$. □

Date: June 2, 2020.

(iv) ???

Solution. The directed graph representation is



Since $3 \to 4$ but $4 \not\to 3$, state 3 is transient, and since $6 \to 1$ but $1 \not\to 6$, state 6 is transient. So state 1, 2, 4, and 5 are recurrent. The closed irreducible sets are $\{1, 4\}$ and $\{2, 5\}$. $\square$

Problem 2. Let G be a connected graph. Let X_n be a simple random walk on G. Show that the Markov chain $\{X_n\}$ is irreducible. (Hint: for any two vertices x and y in G, consider a path of consecutive nodes from x to y.)

Solution. Let x and y be any two vertices in G, let $T_y = \min\{n \geq 0 : X_n = y\}$ be the first hitting time of y, and let K be the smallest number of steps it takes to get from x to y. Since G is connected, x and y are connected by with a path of edges by definition, which implies that $p^K(x, y) > 0$. So, for the the chain started in x, there exists a sequence $\{y_1, y_2, \ldots, y_{K-1}\}$ such that

$$p(x, y_1) > 0, p(y_1, y_2) > 0, \ldots, p(y_{K-1}, y_K) > 0,$$

where $y_i \neq x$ for $i = 1, 2, \ldots, K-1$. Then

$$P_x(T_y < \infty) = p(x, y_1)p(y_1, y_2) \cdots p(y_{K-1}, y_K)p(y_K, y) > 0.$$

By the same logic, for the chain started in y, we see that $\{X_n\}$ on G has exactly one communicating class, and hence is irreducible. $\square$

Problem 3. Let G be a graph with two disjoint components G_1 and G_2. Let X_n be a simple random walk on G.

(i) Prove $\{X_n\}$ is not an irreducible Markov chain.

Solution. Since $G_1 \cap G_2 = \emptyset$, there is no edge between any node in G_1 and any node in G_2, and hence G_1 and G_2 are disconnected. Then $\{X_n\}$ on G is not irreducible.

If it were the case that G_1 and G_2 are connected, then there would exists a path between any two nodes $x_1, y_1 \in G_1$, i.e., $x_1 \leftrightarrow y_1$, and $x_2, y_2 \in G_2$, i.e., $x_2 \leftrightarrow y_2$. However, G_1 and G_2 are disconnected as previously noted, $x_1, y_1 \not\leftrightarrow x_2, y_2$. But $x_1 \leftrightarrow y_1$ and $x_2 \leftrightarrow y_2$. So, there exists two communicating classes, and hence $\{X_n\}$ on G does not satisfy the definition of an irreducible Markov chain. $\square$

(ii) Let P be the transition matrix of X_n. Let P_1 and P_2 be the transition matrices for the SRW on G_1 and G_2, respectively. Let $V_1 = \{1, \ldots, k\}$ and $V_2 = \{k+1, \ldots, k+l\}$ be the set of vertices in G_1 and G_2. Show that P is of the following block diagonal form

$$P = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}. \tag{1}$$

Note a block diagonal matrix is a *reducible* matrix.

Solution. We have $P(i, j) = 0$ for all $[i] \neq [j]$ since $V_1 \cap V_2 = \emptyset$. If $[i] = [j]$ and $i \in V_1$, then X_n has transition matrix P_1, and if $i \in V_2$, X_n has transition matrix P_2. Hence, P is of the block diagonal form (1). $\square$

(iii) Show the the SRW $\{X_n\}$ on G has infinitely many stationary distributions.

Solution. Since the disjoint graphs G_1 and G_2 are finite, P_1 has a stationary distribution, call it π_1, and P_2 has stationarity distribution, call it π_2. Note that π_1 has dimension $1 \times k$, π_2 has dimension $1 \times l$. Let

$$\pi = \begin{cases} (\pi_1, 0) & \text{if } X_0 \in G_1 \\ (0, \pi_2) & \text{if } X_0 \in G_2 \end{cases}$$

be the stationary distribution for the operators P on G, and note that π has dimension $1 \times k + l$. Note also that P has dimension $(k+l) \times (k+l)$. Now, for the chain started on any vertex in G_1, let $\pi = (\pi_1, 0)$ where 0 is a $1 \times l$ dimensional vector. Then

$$\begin{pmatrix} \pi_1 & 0 \end{pmatrix} \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} = \begin{pmatrix} \pi_1 P_1 & 0 \end{pmatrix} = \begin{pmatrix} \pi_1 & 0 \end{pmatrix},$$

where the last equality follows since π_1 was defined to be the stationary distribution for P_1. For the chain started on any vertex in G_2, $\pi = (0, \pi_2)$ where here 0 is a $1 \times k$ vector. So

$$\begin{pmatrix} 0 & \pi_2 \end{pmatrix} \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} = \begin{pmatrix} 0 & \pi_2 P_2 \end{pmatrix} = \begin{pmatrix} 0 & \pi_2 \end{pmatrix}$$

For $0 \leq \lambda \leq 1$,

$$(\lambda \pi_1 + (1-\lambda)\pi_2) = (\lambda \pi_1 P + (1-\lambda)\pi_2 P)$$

is also stationary distribution for P since of the block diagonal form (1). Hence, there exists infinitely stationary for X_n on G. □

Problem 4. Consider a Markov chain with state space $S = \{1, 2\}$ and transition matrix

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix},$$

where $0 < a < 1$ and $0 < b < 1$.

(i) Find its stationary distribution π.

Solution. Let $\pi = (x, y)$. Then

$$\pi P = \pi \implies \begin{cases} (1-a)x + by = x \\ ax + (1-b)y = y \end{cases} \implies ax = by.$$

Recalling that $x + y = 1$ gives the unique stationary distribution

$$\pi = \left(\frac{b}{a+b}, \frac{a}{a+b} \right).$$

□

(ii) Use the Markov property to show that

$$P(X_{n+1} = 1) - \frac{b}{a+b} = (1-a-b)\left[P(X_n = 1) - \frac{b}{a+b} \right].$$

Solution. Start from the left hand side and condition on X_n to obtain

$$\begin{aligned}
P(X_{n+1} = 1) &= \sum_i P(X_{n+1} = 1 \mid X_n = i) P(X_n = i) \\
&= (1-a)P(X_n = 1) + b[1 - P(X_n = 1)] \\
&= P(X_n = 1) - aP(X_n = 1) + b - bP(X_n = 1) \\
&= (1-a-b)P(X_n = 1) + b \\
&= (1-a-b)P(X_n = 1) + \frac{b(a+b)}{a+b} \\
&= (1-a-b)P(X_n = 1) + \frac{ab}{a+b} + \frac{b^2}{a+b}.
\end{aligned}$$

Next, subtract $b/(a+b)$ from both sides

$$P(X_{n+1} = 1) - \frac{b}{a+b} = (1-a-b)P(X_n = 1) + \frac{ab}{a+b} + \frac{b^2}{a+b} - \frac{b}{a+b}.$$

Simplifying the above expression gives

$$P(X_{n+1} = 1) - \frac{b}{a+b} = (1-a-b)\left[P(X_n = 1) - \frac{b}{a+b}\right].$$

$\square$

(iii) Use mathematical induction to show that

$$P(X_n = 1) = \frac{b}{a+b} + (1-a-b)^n \left[P(X_0 = 1) - \frac{b}{a+b}\right]$$

Proof (by induction on n). The base case $n = 0$ holds since

$$\begin{aligned}
P(X_0 = 1) &= \frac{b}{a+b} + (1-a-b)^0 \left[P(X_0 = 1) - \frac{b}{a+b}\right] \\
&= \frac{b}{a+b} + P(X_0 = 1) - \frac{b}{a+b} = P(X_0 = 1),
\end{aligned}$$

and the base case $n = 1$ holds since

$$\begin{aligned}
P(X_1 = 1) &= \sum_i P(X_1 = 1 \mid X_0 = i) P(X_0 = i) \\
&= P(X_1 = 1 \mid X_0 = 1)P(X_0 = 1) + P(X_1 = 1 \mid X_0 = 2)P(X_0 = 2) \\
&= (1-a)P(X_0 = 1) + b(1 - P(X_0 = 1)) \\
&= \frac{b}{a+b} + (1-a-b)\left[P(X_0 = 1) - \frac{b}{a+b}\right]
\end{aligned}$$

where the last equality uses the same logic as in (ii).

Assume the claim is true for any n. We want to show that the result is also true for $n+1$. From (i), we know that

$$P(X_{n+1} = 1) = \frac{b}{a+b} + (1-a-b)\left[P(X_n = 1) - \frac{b}{a+b}\right].$$

Substituting in for $P(X_n = 1)$ gives

$$\begin{aligned} P(X_{n+1} = 1) &= \frac{b}{a+b} + (1-a-b)\left[\left[(1-a-b)^n \left[P(X_0 = 1) - \frac{b}{a+b}\right]\right] - \frac{b}{a+b}\right] \\ &= \frac{b}{a+b} + (1-a-b)^{n+1}\left[P(X_0 = 1) - \frac{b}{a+b}\right]. \end{aligned}$$

Shifting the indices back one period gives

$$P(X_n = 1) = \frac{b}{a+b} + (1-a-b)^n \left[P(X_0 = 1) - \frac{b}{a+b}\right].$$

□

(iv) Show that $P(X_n = 1)$ converges exponentially fast to $\pi(1)$ for the π you found in (i).

Solution. Since $0 < a < 1$ and $0 < b < 1$, the term $|1 - a - b| < 1$. Then, noting that $(1 - a - b)^n$ is the only term depending on n, we have

$$\begin{aligned} \lim_{n\to\infty} P(X_n = 1) &= \lim_{n\to\infty}\left[\frac{b}{a+b} + (1-a-b)^n \left[P(X_0 = 1) - \frac{b}{a+b}\right]\right] \\ &= \frac{b}{a+b} + \lim_{n\to\infty}(1-a-b)^n \left[P(X_0 = 1) - \frac{b}{a+b}\right] \\ &= \frac{b}{a+b} + O((1-a-b)^n). \end{aligned}$$

That is, convergence to the limiting distribution, $P(X_n = 1) \to \pi(1)$ as $n \to \infty$, is exponentially fast □