2. Transience and Recurrence

The study of Markov chains, particularly the limiting behavior, depends critically on the random times between visits to a given state. The nature of this random times leads to a fundamental dichotomy of the states.

Basic Theory

As usual, our starting point is a (time homogeneous) Markov chain $X = (X_0, X_1, X_2, \ldots)$ with state space S and transition probability matrix P.

Hitting Times and Probabilities

Let A be a nonempty subset of S. Recall that the **hitting time** to A is the random variable that gives the first positive time that the chain is in A:

$$T_A = \min\{n \geq 1 : X_n \in A\}$$

Since the chain may *never* enter A, the random variable T_A takes values in $\mathbb{N}_+ \cup \{\infty\}$. When $A = \{y\}$, we will simplify the notation to T_y. This random variable gives the first positive time that the chain is in state y. Now for $x \in S$ and $n \in \mathbb{N}_+$, let

$$H_n(x, A) = \mathbb{P}(T_A = n | X_0 = x), \quad H(x, A) = \mathbb{P}(T_A < \infty | X_0 = x)$$

1. Note that $H(x, A) = \sum_{n=1}^{\infty} H_n(x, A)$

Again, when $A = \{y\}$, we will simplify the notation to $H_n(x, y)$ and $H(x, y)$, respectively. In particular, $H(x, x)$ is the probability, starting at x, that the chain eventually returns to x. If $x \neq y$, $H(x, y)$ is the probability, starting at x, that the chain eventually reaches y. Just knowing when $H(x, y)$ is 0, positive, and 1 will turn out to be of considerable importance in the overall structure and limiting behavior of the chain. As a function on S^2, we will refer to H as the **hitting matrix** of X.

2. Show that $H(x, y) > 0$ if and only if $P^n(x, y) > 0$ for some $n \in \mathbb{N}_+$.

    a. Show that $\{X_n = y\} \subseteq \{T_y < \infty\}$ for all $n \in \mathbb{N}_+$.

    b. Show that $\{T_y < \infty\} = \{X_k = y \text{ for some } k \in \mathbb{N}_+\}$.

    c. Use basic rules of probability and Boole's inequality to conclude that $P^n(x, y) \leq H(x, y) \leq \sum_{k=1}^{\infty} P^k(x, y)$ for all $n \in \mathbb{N}_+$.

The following exercise gives a basic relationship between the sequence of hitting probabilities and the sequence of transition probabilities.

3. Suppose that $(x, y) \in S^2$. Condition on T_y to show that

$$P^n(x, y) = \sum_{k=1}^{n} H_k(x, y) P^{n-k}(y, y), \quad n \in \mathbb{N}_+$$

Starting in state x, the chain is in state y at time n if and only if the chain hits y for the first time at some previous time k, and then returns to y in the remaining $n - k$ steps.

4. Suppose that $x \in S$ and $A \subseteq S$. Condition on X_1 to show that

   a. $H_{n+1}(x, A) = \sum_{z \notin A} P(x, z) H_n(z, A)$ for $n \in \mathbb{N}_+$
   b. $H(x, A) = P(x, A) + \sum_{z \notin A} P(x, z) H(z, A)$

Starting in state x, the chain first enters A at time $n + 1$ if and only if the chain goes to some state $z \notin A$ at time 1, and then from state z, first enters A in n steps. Similarly, starting in state x, the chain eventually enters A if and only if it either enters A at the first step, or moves to some other state $z \notin A$ at the first step, and then eventually enters A from z.

A state x is said to be **recurrent** if $H(x, x) = 1$ and is said to be **transient** if $H(x, x) < 1$. Thus, starting in a recurrent state, the chain will, with probability 1, eventually return to the state. As we will see, the chain will return to the state infinitely often with probability 1, and the times of the visits will form the arrival times of a renewal process. This will turn out to be the critical observation in the study of the limiting behavior of the chain. By contrast, if the chain starts in a transient state, then there is a positive probability that the chain will never return to the state.

Counting Variables and Potential

Again, suppose that A is a nonempty set of states. A natural complement to the hitting time to A is the **counting variable** that gives the number of visits to A (at positive times). Thus, let

$$N_A = \sum_{n=1}^{\infty} \mathbf{1}(X_n \in A)$$

We will mostly be interested in the special case $A = \{x\}$, and in this case, we will simplify the notation to N_x. Note that N_A takes value in $\mathbb{N} \cup \{\infty\}$.

5. Let $G(x, A) = \mathbb{E}(N_A | X_0 = x)$ for $x \in S$ and $A \subseteq S$. Show that
$$G(x, A) = \sum_{n=1}^{\infty} P^n(x, A)$$

6. Consistent with our notation in the Introduction, G is a transition measure on S (although, of course, not a transition *probability* measure). Show that
$$G(x, A) = \sum_{y \in A} G(x, y)$$

The matrix G is known as the **potential matrix**. (Some authors count the state at time 0 and thus the index n in the definition of the counting variable and in Exercise 5 starts at 0 rather than 1.)

The distribution of N_A has a simple representation in terms of the hitting probabilities. Note that because of the Markov property and time homogenous property, whenever the chain reaches state y, the future behavior is independent of the past and is stochastically the same as the chain starting in state y at time 0. This is the critical observation in the proof of the following exercise.

7. Suppose that x and y are states. Show that

$$\mathbb{P}\left(N_y = n \middle| X_0 = x\right) = \begin{cases} 1 - H(x, y), & n = 0 \\ H(x, y)\, H(y, y)^{n-1}\, (1 - H(y, y)), & n \in \mathbb{N}_+ \end{cases}$$
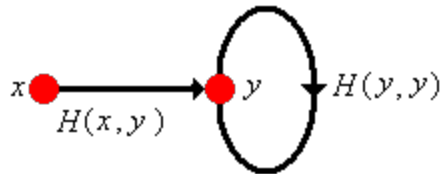
The essence of the proof is illustrated in the graphic above. The thick lines are intended as reminders that these are not one step transitions, but rather represent all paths between the given vertices. Note that in the special case that $x = y$ we have

$$\mathbb{P}\left(N_y = n \middle| X_0 = x\right) = H(y, y)^n\, (1 - H(y, y)) \quad n \in \mathbb{N}$$

In all cases, the counting variable N_y has essentially a geometric type distribution, but the distribution may well be **defective**, with some of the probability mass at ∞. The behavior is quite different depending on whether y is transient or recurrent.

8. Suppose that x and y are states and that y is transient. Show that

a. $\mathbb{P}\left(N_y < \infty \middle| X_0 = x\right) = 1$

b. $G(x, y) = \dfrac{H(x, y)}{1 - H(y, y)}$

c. $H(x, y) = \dfrac{G(x, y)}{1 + G(y, y)}$

9. Suppose that x and y are states and that y is recurrent. Show that

a. $\mathbb{P}\left(N_y = 0 \middle| X_0 = x\right) = 1 - H(x, y)$ and $\mathbb{P}\left(N_y = \infty \middle| X_0 = x\right) = H(x, y)$

b. $G(x, y) = 0$ if $H(x, y) = 0$ and $G(x, y) = \infty$ if $H(x, y) > 0$

c. $\mathbb{P}\left(N_y = \infty \middle| X_0 = y\right) = 1$ and $G(y, y) = \infty$

Note that there is an invertible relationship between the hitting probability matrix H and the potential matrix G; if we know one we can compute the other. In particular, we can characterize the transience or recurrence of a state in terms of G:

- State x is transient if and only if $H(x, x) < 1$ if and only if $G(x, x) < \infty$.
- State x is recurrent if and only if $H(x, x) = 1$ if and only if $G(x, x) = \infty$.

Relations

The hitting probabilities lead to an important relation on the state space S. For $(x, y) \in S^2$, we say that x **leads to** y and we write $x \to y$ if either $x = y$ or $H(x, y) > 0$. It follows immediately from Exercise 2 that $x \to y$ if and only if $P^n(x, y) > 0$ for some $n \in \mathbb{N}$. In terms of the graph of the chain, $x \to y$ if and only if there is a directed path from x to y. Note that the *leads to* relation is **reflexive** by definition: $x \to x$ for any $x \in S$.

⊞ 10. Show that the *leads to* relation is **transitive**: if $x \to y$ and $y \to z$ then $x \to z$.

   a. There exist $j \in \mathbb{N}$ and $k \in \mathbb{N}$ such that $P^j(x, y) > 0$ and $P^k(y, z) > 0$

   b. Show that $P^{j+k}(x, z) > 0$

A nonempty set of states A is **closed** if $x \in A$ and $x \to y$ implies $y \in A$. A closed set A is **irreducible** if A has no proper closed subset.

⊞ 11. Suppose that $A \subseteq S$ is closed.

   a. Show that P_A, the restriction of P to $A \times A$, is a transition probability matrix on A.

   b. Note that X restricted to A is a Markov chain with transition probability matrix P_A.

   c. Show that $P^n{}_A = P_A{}^n$ for $n \in \mathbb{N}$

Of course, the entire state space S is closed by definition. If it is also irreducible, we say the Markov chain X is **irreducible**.

⊞ 12. Suppose that A is a nonempty subset of S. Show that $\mathrm{cl}(A) = \{y \in S : x \to y \text{ for some } x \in A\}$ is the smallest closed set containing A, and is called the **closure** of A. That is, show that

   a. $\mathrm{cl}(A)$ is closed.

   b. $A \subseteq \mathrm{cl}(A)$

   c. If B is closed and $A \subseteq B$ then $\mathrm{cl}(A) \subseteq B$

Recall that for a fixed positive integer k, P^k is also a transition probability matrix, and in fact governs the k-**step** Markov chain. $(X_0, X_k, X_{2k}, \ldots)$. It follows that we could consider the *leads to* relation for this chain, and all of the results above would still hold (relative, of course to the k-step chain). Occasionally we will need to consider this relation, which we will denote by $\underset{k}{\to}$, particularly in our study of periodicity.

⊞ 13. Suppose that j and k are positive integers. Show that if $x \underset{k}{\to} y$ and $j|k$. then $x \underset{j}{\to} y$.

By combining the *leads to* relation $\to$ with its inverse, the **comes from** relation $\leftarrow$, we can obtain another very useful relation. For $(x, y) \in S^2$, we say that x **to and from** y and we write $x \leftrightarrow y$ if $x \to y$ and $y \to x$. By definition, this relation is **symmetric**: if $x \leftrightarrow y$ then $y \leftrightarrow x$. From our work above, it is also reflexive and transitive. Thus, the *to and from* relation is an equivalence relation. Like all equivalence relations, it partitions the space into mutually disjoint **equivalence classes**. We will denote the equivalence class of a state x by

$$[x] = \{y \in S : x \leftrightarrow y\}$$

Thus, for any two states x and y, either $[x] = [y]$ or $[x] \cap [y] = \varnothing$, and moreover, $\bigcup_{x \in S} [x] = S$
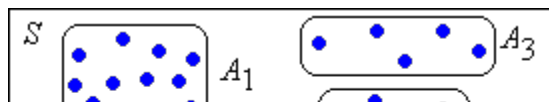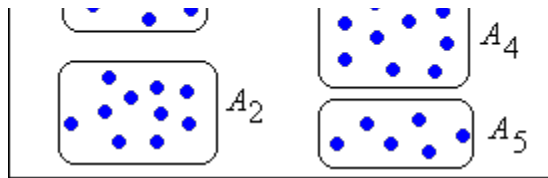
14. Draw state graphs to illustrate the following facts:

   a. A closed set is not necessarily an equivalence class.

   b. An equivalence class is not necessarily closed.

▨ 15. On the other hand, show that if A is a closed, irreducible set of states, then A is an equivalence class.

   a. Fix $x \in A$ and conclude from closure that $[x] \subseteq A$.

   b. For arbitrary $y \in A$, use irreducibility to argue that $\mathrm{cl}(x) = A$ and $\mathrm{cl}(y) = A$ and therefore $x \leftrightarrow y$.

   c. Conclude from (b) that $A \subseteq [x]$.

The *to and from* equivalence relation is very important because many interesting state properties turn out in fact to be **class properties**, shared by all states in a given equivalence class. In particular, the recurrence and transience properties are class properties.

Transient and Recurrent Classes

The following exercise gives the fundamental result of this section: a recurrent state can only lead to other recurrent states.

▨ 16. Suppose that x is a recurrent state and that $x \to y$. Show that y is recurrent and that $H(x, y) = H(y, x) = 1$.

   a. Show first that the result holds if $x = y$, so henceforth assume that $x \neq y$

   b. Let $\alpha(x, y)$ denote the probability, starting at x, that the chain reaches y without an intermediate return to x. Argue that $\alpha(x, y) > 0$ since $x \to y$. In terms of the graph of X, if there is a path from x to y, then there is a path from x to y without cycles.

   c. Starting at x, the chain could fail to return to x by first reaching y without an intermediate return to x, and then from y never reaching x. Use the Markov and time homogeneous properties to argue that
$$1 - H(x, x) \geq \alpha(x, y)(1 - H(y, x)) \geq 0.$$

   d. Conclude from (c) that $H(y, x) = 1$.

   e. Show that there exist positive integers j and k such that $P^j(x, y) > 0$ and $P^k(y, x) > 0$.

   f. Show that $P^{j+k+n}(y, y) \geq P^k(y, x) P^n(x, x) P^j(x, y)$ for any $n \in \mathbb{N}$.

   g. Sum over n in (f) to conclude that $G(y, y) = \infty$ and hence that y is recurrent.

   h. Finally, reverse the roles of x and y to conclude that $H(x, y) = 1$.

From the last exercise, note that if x is recurrent, then all states in the equivalence class of x are also recurrent. Thus, for each equivalence class, either all states are transient or all states are recurrent. We can therefore refer to **transient or recurrent classes** as well as states.

17. Suppose that A is a recurrent equivalence class. Show that A is closed and irreducible.

18. Suppose that A is finite and closed. Show that A has a recurrent state.

   a. Fix $x \in A$ and argue that $\mathbb{P}(N_A = \infty | X_0 = x) = 1$ since A is closed.

   b. Argue that $\mathbb{P}(N_y = \infty | X_0 = x) > 0$ for some $y \in A$ since A is finite.

   c. Conclude that y is recurrent.

19. Suppose that A is finite, closed, and irreducible. Show that A is a recurrent equivalence class.

   a. Note that A is an equivalence class by Exercise 15.

   b. Note that A has a recurrent state by Exercise 18.

   c. Conclude that all states in A are recurrent.

Thus, the Markov chain X will have a collection (possibly empty) of recurrent equivalence classes $\{A_j : j \in J\}$ where J is a countable index set. Each A_j is closed and irreducible. Let B denote the set of all transient states. The set B may be empty or may consist of a number of equivalence classes, but the class structure of B is not important to us. If the chain starts in A_j for some $j \in J$ then the chain remains in A_j forever, visiting each state infinitely often with probability 1. If the chain starts in B, then the chain may stay in B forever (but only if B is infinite) or may enter one of the recurrent classes A_j, never to escape. However, in either case, the chain will visit a given transient state only finitely many time with probability 1. This basic structure is known as the **canonical decomposition** of the chain, and is shown in graphical form below. The edges from B are in gray to indicate that these transitions may not exist.
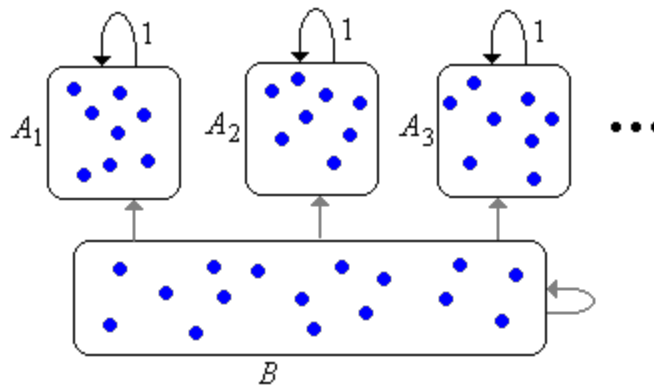
Staying Probabilities and a Classification Test

20. Suppose that A is a proper subset of S.

   a. Show that $(P_A{}^n \mathbf{1}_A)(x) = P_A{}^n(x, A) = \mathbb{P}(X_1 \in A, X_2 \in A, ..., X_n \in A | X_0 = x)$ for $x \in A$. This is the probability that the chain stays in A at least through time n, starting in state x.

   b. Show that $g_A(x) = \lim_{n \to \infty} (P_A{}^n \mathbf{1}_A)(x) = \mathbb{P}(X_1 \in A, X_2 \in A, ... | X_0 = x)$ for $x \in A$, using (a) and the continuity theorem for decreasing events. This is the probability that chain stays in A forever.

The **staying probability function** g_A is an interesting complement to the hitting matrix studied above. The following exercise characterizes this function and provides a method that can be used to compute it, at least in some cases.

⊞ 21. Show that g_A is the largest function on A that takes values in $[0, 1]$ and satisfies $g = P_A \, g$. Moreover, either $g_A = \mathbf{0}_A$ or $\sup\{g_A(x) : x \in A\} = 1$

   a. Note that $P_A{}^{n+1} \, \mathbf{1}_A = P_A \, P_A{}^n \, \mathbf{1}_A$ for $n \in \mathbb{N}$

   b. Take the limit in (a) as $n \to \infty$ and use the bounded convergence theorem to conclude that $g_A = P_A \, g_A$

   c. Suppose now that g is a function on A that takes values in $[0, 1]$ and satisfies $g = P_A \, g$. Argue that $g \le \mathbf{1}_A$ and hence that $g \le P_A{}^n \, \mathbf{1}_A$ for all $n \in \mathbb{N}$.

   d. Take the limit in (c) as $n \to \infty$ to conclude that $g \le g_A$.

   e. Let $c = \sup\{g_A(x) : x \in A\}$. Argue that $g_A \le c \, \mathbf{1}_A$ and hence $g_A \le c \, P_A{}^n \, \mathbf{1}_A$ for each $n \in \mathbb{N}$.

   f. Take the limit as $n \to \infty$ to conclude that $g_A \le c \, g_A$.

   g. Conclude that either $c = 0$ or $c = 1$

Note that the characterization in the last exercise includes a **zero-one law** of sorts: either the probability that the chain stays in A forever is 0 for every initial state x, or we can find states for which the probability is arbitrarily close to 1. The next two exercises explore the relationship between the staying function and recurrence.

⊞ 22. Suppose that X is an irreducible, recurrent chain with state space S. Show that $g_A = \mathbf{0}_A$ for any proper subset A of S.

   a. Fix $y \notin A$ and argue that $0 \le g_A(x) \le 1 - H(x, y)$ for every $x \in A$

   b. Argue that $g_A(x) = 0$ since the chain is irreducible and recurrent.

⊞ 23. Suppose that X is an irreducible Markov chain with state space S and transition probability matrix P. Show that if there exists a state x such that $g_A = \mathbf{0}_A$ where $A = S \smallsetminus \{x\}$ then X is recurrent.

   a. With A as defined above, show that $1 - H(x, x) = \sum_{y \in A} P(x, y) g_A(y)$.

   b. Conclude that x is recurrent and hence that the chain is recurrent.

More generally, suppose that X is a Markov chain with state space S and transition probability matrix P. The last two exercise can be used to test whether a closed, irreducible equivalence class C is recurrent or transient. We fix a state $x \in C$ and set $A = C \smallsetminus \{x\}$. We then try to solve the equation $g = P_A \, g$ on A. If the only solution taking values in $[0, 1]$ is $\mathbf{0}_A$, then the class C is recurrent by Exercise 23. If there are nontrivial solutions, then C is transient by Exercise 22. Often we try to choose x to make the computations easy.

Computing Hitting Probabilities and Potentials

We now know quite a bit about Markov chains, and we can often classify the states and compute quantities of interest. However, we do not yet know how to compute:

- $G(x, y)$ when x and y are transient
- $H(x, y)$ when x is transient and y is transient or recurrent.

These problems are related, because of the general inverse relationship between the hitting matrix and the potential matrix noted in our discussion above. As usual, suppose that X is a Markov chain with state space S, and let B denote the set of transient states. The next exercise shows how to compute G_B, the potential matrix restricted to the transient states. Recall that the values of this matrix are finite.

---

24. Show that G_B satisfies the equation $G_B = P_B + P_B G_B$ and is the smallest nonnegative solution. If B is finite, $G_B = (I_B - P_B)^{-1} P_B$.

   a. Argue that $P^n{}_B = P_B{}^n$ since a path between two transient states can only pass through other transient states.
   b. Conclude that $G_B = \sum_{n=1}^{\infty} P_B{}^n$
   c. Use the monotone convergence theorem to show that $P_B G_B = G_B - P_B$.
   d. Suppose that U is a nonnegative matrix on B satisfying $U = P_B + P_B U$. Show that
      $U = \sum_{k=1}^{n} P_B{}^k + P_B{}^{n+1} U$ for each positive integer n.
   e. Conclude that $U \geq \sum_{k=1}^{n} P_B{}^k$ for every positive integer n and hence $U \geq G_B$.
   f. Show that $(I_B - P_B)(I_B + G_B) = I_B$. Conclude that if B is finite, the matrix $I_B - P_B$ is invertible.

Now that we can compute G_B, we can also compute H_B using the result in Exercise 8. All that remains is for us to compute the hitting probability $H(x, y)$ when x is transient and y is recurrent. The first thing to notice is that the hitting probability is a class property.

---

25. Suppose that x is transient and that A is a recurrent class. Show that $H(x, y) = \mathbb{P}(T_A < \infty | X_0 = x)$ for $y \in A$. That is, the hitting probability to y is constant for $y \in A$, and is just the hitting probability to the class A.

As before, let B denote the set of transient states and suppose that A is a recurrent equivalence class. Let h_A denote the function on B that gives the hitting probability to class A, and let p_A denote the function on B that gives the probability of entering A on the first step:

$$h_A(x) = \mathbb{P}(T_A < \infty | X_0 = x), \quad p_A(x) = P(x, A); \quad x \in B$$

---

26. Show that $h_A = p_A + G_B p_A$.

   a. Show that $\mathbb{P}(T_A = n | X_0 = x) = \left(P_B{}^{n-1} p_A \right)(x)$ for $n \in \mathbb{N}_+$.
   b. Sum over n.

The result in the last exercise is adequate if we have already computed G_B (using the result in Exercise 24, for example). However, we might just want to compute h_A directly.

---

27. Show that h_A satisfies the equation $h_A = p_A + P_B h_A$ and is the smallest nonnegative solution. If B is finite, $h_A = (I_B - P_B)^{-1} p_A$.

   a. Show that $h_A = p_A + P_B h_A$ by conditioning on X_1.
   b. Suppose that h is nonnegative and satisfies $h = p_A + P_B h$. Show that $h = p_A + \sum_{k=1}^{n-1} P_B{}^k p_A + P_B{}^n h$ for each positive integer n.
   c. Conclude that $h \geq p_A + \sum_{k=1}^{n-1} P_B{}^k p_A$

d. Take the limit as $n \to \infty$ to conclude that $h \geq h_A$.

e. Note that the representation when B is finite follows from .

Examples and Applications

Finite Chains

⊞ 28. Consider a Markov chain with state space $S = \{a, b, c, d\}$ and transition matrix P given below:

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

a. Draw the state diagram.

b. Find the equivalent classes and classify each as transient or recurrent.

c. Compute the potential matrix G.

d. Compute the hitting matrix F.

⊞ 29. Consider a Markov chain with state space $S = \{1, 2, 3, 4, 5, 6\}$ and transition matrix P given below:

$$P = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{1}{4} & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \end{pmatrix}$$

a. Sketch the state graph.

b. Find the equivalence classes and classify each as recurrent or transient.

c. Compute the potential matrix G.

d. Compute the hitting matrix H.

⊞ 30. Consider a Markov chain with state space $S = \{1, 2, 3, 4, 5, 6\}$ and transition matrix P given below:

$$P = \begin{pmatrix} \dfrac{1}{2} & \dfrac{1}{2} & 0 & 0 & 0 & 0 \\[2mm] \dfrac{1}{4} & \dfrac{3}{4} & 0 & 0 & 0 & 0 \\[2mm] \dfrac{1}{4} & 0 & \dfrac{1}{2} & \dfrac{1}{4} & 0 & 0 \\[2mm] \dfrac{1}{4} & 0 & \dfrac{1}{4} & \dfrac{1}{4} & 0 & \dfrac{1}{4} \\[2mm] 0 & 0 & 0 & 0 & \dfrac{1}{2} & \dfrac{1}{2} \\[2mm] 0 & 0 & 0 & 0 & \dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix}$$

a. Sketch the state graph.
b. Find the equivalence classes and classify each as recurrent or transient.
c. Compute the potential matrix G.
d. Compute the hitting matrix H.

Special Models

31. Read again the definitions of the Ehrenfest chains and the Bernoulli-Laplace chains. Note that since these chains are irreducible and have finite state spaces, they are recurrent.

32. Read the discussion on recurrence in the section on the reliability chains and work the exercises.

33. Read the discussion on random walks on $\mathbb{Z}^k$ in the section on the random walks on graphs and work the exercises.

34. Read the discussion on extinction and explosion in the section on the branching chain and work the exercises.

35. Read the discussion on recurrence and transience in the section on queuing chains and work the exercises.

36. Read the discussion on recurrence and transience in the section on birth-death chains and work the exercises.