Handout 3

STAT5710 Applied Stochastic Processes

Last time we saw that a finite chain always has a stationary distribution, but need not have a limiting one. To characterize when limits exist we distinguish *recurrent* states, which are visited repeatedly, from *transient* states, which may never be visited.

Throughout, $\{X_n\}$ is a time-homogeneous Markov chain on a state space S with transition matrix P, and $p_{ij}^{(n)} = (P^n)_{ij}$. Three running examples on $S = \{1, 2, 3\}$:

$$P_A = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad P_B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \qquad P_C = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

Accessibility and communication

1. If $p_{ij}^{(n)} > 0$ for some $n \geq 0$, we say
   (a) state j is ____________________ from state i
   (b) states i and j ____________________ if i is accessible from j *and* j is accessible from i

   Note the condition is $n \geq 0$, not $n > 0$: the zeroth transition counts, since $P^0 = I$.

2. Communication is an **equivalence relation** — it is

   ____________________, ____________________, and ____________________

   The corresponding equivalence classes are the **communication classes** of the chain.

3. Draw the transition graph for each of the three chains above, then record its communication classes.

Hitting times and recurrence

4. Fix a state j. Two closely related random times:
   (a) the **first passage time** to j is $T_j = \inf\{\underline{\hspace{4cm}}\}$
   (b) the **hitting time** to j is $\tau_j = \inf\{\underline{\hspace{4cm}}\}$

Both are examples of **stopping times**.

A worked example (Dobrow 3.22)

Return to the two-state chain,

$$P = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} \end{array} \qquad \text{where} \qquad p, q \in (0,1)$$

5. Show that starting from state 1, the chain returns to state 1 with probability 1.
   (a) The event $\{T_1 \geq n\}$ says the chain has not yet come back. Write it as an event about $X_1, \ldots, X_{n-1}$.
   $$\{T_1 \geq n\} \;=\; \{\underline{\hspace{7cm}}\}$$
   (b) Factor the probability using the Markov property.

$$P(T_1 \geq n \mid X_0 = 1) \;=\; \underbrace{\prod_{k=1}^{n-2} P(X_{k+1} = 2 \mid X_k = 2)}_{A} \cdot \underbrace{P(X_1 = 2 \mid X_0 = 1)}_{B}$$

$$A \;=\; \underline{\hspace{4cm}} \qquad\qquad B \;=\; \underline{\hspace{4cm}}$$

$$= \; \underline{\hspace{4cm}}$$

$$\longrightarrow \; \underline{\hspace{2cm}} \qquad \text{as } n \to \infty$$

6. Now find the expected time to return, using the tail-sum formula $E(T) = \sum_{n\geq 1} P(T \geq n)$ for a nonnegative integer random variable.

$$E(T_1 \mid X_0 = 1) \;=\; \sum_{n=1}^{\infty} P(T_1 \geq n \mid X_0 = 1)$$

$$= \; 1 \;+\; \sum_{n=2}^{\infty} \underline{\hspace{4cm}}$$

$$= \; 1 \;+\; p \sum_{j=0}^{\infty} \underline{\hspace{4cm}}$$

$$= \; \underline{\hspace{4cm}}$$

7. Previously we found the stationary distribution of this chain.

$$\pi^T = \underline{\hspace{4cm}} \qquad \Longrightarrow \qquad E(T_1 \mid X_0 = 1) \;=\; \underline{\hspace{2cm}}$$

8. The expected *hitting* time from the other state takes a different argument: condition on X_1. It comes out shorter when the chain mixes fast, i.e. when $p + q > 1$.

$$E(\tau_1 \mid X_0 = 2) \;=\; 1 \cdot P(X_1 = 1 \mid X_0 = 2) \;+\; \left[1 + E(\tau_1 \mid X_1 = 2)\right] P(X_1 = 2 \mid X_0 = 2)$$

$$= \; \underline{\hspace{7cm}}$$

$$= \; 1 + (1-q)\, E(\tau_1 \mid X_0 = 2) \qquad \Longrightarrow \qquad E(\tau_1 \mid X_0 = 2) = \underline{\hspace{2cm}}$$

Recurrent and transient states

Denote the probability of ever returning to j and the probability of ever hitting j from i by

$$f_j = \underline{\hspace{4cm}} \qquad \text{and} \qquad h_{ij} = \underline{\hspace{4cm}}$$

respectively.

9. State j is
   (a) **recurrent** if ______________________
   (b) **transient** if ______________________

10. Let $S_1 < S_2 < \cdots$ denote the successive visit times to state j so that $S_1 = \tau_j$ and

$$S_{k+1} = \inf\{\underline{\hspace{4cm}}\} \quad \text{for} \quad k = 1, 2, \ldots$$

   (a) Then consider the total number of visits to state j, which we will denote by

$$N_j = \underline{\hspace{7cm}}$$

   (b) Now the visit count and visit times are related by $N_j \geq$ __________ $\iff$ $S_n <$ __________.

   (c) Build up $P(S_n < \infty \mid X_0 = i)$ one visit at a time.
      i. For $n = 1$:

$$P(S_1 < \infty \mid X_0 = i) = \underline{\hspace{4cm}}$$

      ii. For $n > 1$, use the Markov property at time S_{n-1}, where the chain is sitting in state j, together with time-homogeneity.

$$\begin{aligned} P(S_n < \infty \mid S_{n-1} < \infty) &= P(\underline{\hspace{4cm}}) \\ &= P(\underline{\hspace{4cm}}) \\ &= \underline{\hspace{1.7cm}} \end{aligned}$$

      iii. Chain the conditional probabilities together.

$$\begin{aligned} P(S_n < \infty \mid X_0 = i) &= \prod_{k=2}^{n} P(S_k < \infty \mid S_{k-1} < \infty) \cdot P(S_1 < \infty \mid X_0 = i) \\ &= \prod_{k=2}^{n} \underline{\hspace{4cm}} \\ &= \underline{\hspace{4cm}} \end{aligned}$$

(d) Sum the tail probabilities and recognize a geometric series.

$$E(N_j \mid X_0 = i) \;=\; \sum_{n=1}^{\infty} P(N_j \geq n \mid X_0 = i) \;=\; h_{ij} \sum_{n=1}^{\infty} \underline{\hspace{2cm}}$$

$$\implies \qquad E(N_j \mid X_0 = i) \;=\; \begin{cases} \underline{\hspace{2cm}}, & f_j < 1 \quad (j \text{ transient}) \\ \underline{\hspace{2cm}}, & f_j = 1,\; h_{ij} > 0 \quad (j \text{ recurrent}) \\ \underline{\hspace{2cm}}, & f_j = 1,\; h_{ij} = 0 \end{cases}$$

So if a recurrent state is visited at all, it is visited ________________ often.

(e) Specialize to $i = j$, where $h_{jj} = 1$.

$$E(N_j \mid X_0 = j) \;=\; \begin{cases} \underline{\hspace{2cm}}, & j \text{ transient} \\ \underline{\hspace{2cm}}, & j \text{ recurrent} \end{cases}$$

(f) Compute $E(N_j \mid X_0 = i)$ a second way, by exchanging expectation and sum.

$$E(N_j \mid X_0 = i) \;=\; E\Big[\sum_{n=0}^{\infty} \mathbf{1}\{X_n = j\} \;\Big|\; X_0 = i\Big] \;=\; \sum_{n=0}^{\infty} \underline{\hspace{4.5cm}}$$

$$= \sum_{n=0}^{\infty} \underline{\hspace{4.5cm}} \;=\; \sum_{n=0}^{\infty} \underline{\hspace{2cm}}$$

The above results yield an alternate characterization of recurrence and transience. State j is:

- transient just in case $\sum_{n=0}^{\infty} p_{jj}^{(n)} < \infty$
- recurrent just in case $\sum_{n=0}^{\infty} p_{jj}^{(n)} = \infty$.

11. Transience and recurrence are class properties. In other words, the states in a communication class are either all transient or all recurrent.

    Proof. Suppose i communicates with j, so $p_{ij}^{(r)} > 0$ and $p_{ji}^{(s)} > 0$ for some r, s.

    (a) Apply Chapman–Kolmogorov and keep a single term of the double sum.

    $$p_{ii}^{(r+n+s)} = \sum_k \sum_\ell p_{ik}^{(r)} p_{k\ell}^{(n)} p_{\ell i}^{(s)} \geq \underline{\hspace{4cm}}$$

    (b) Now sum over n and pull out the constants.

    $$\sum_{n=0}^{\infty} p_{ii}^{(n)} \geq \sum_{n=r+s}^{\infty} p_{ii}^{(n)} = \sum_{n=0}^{\infty} p_{ii}^{(n+r+s)} \geq \underline{\hspace{7cm}}$$

    (c) So if j is recurrent then ________________, and if i is transient then ________________.

12. A chain is **irreducible** if ____________________________

    If a finite Markov chain is irreducible, then all states are recurrent.

    Proof. Suppose instead that every state is transient.

    (a) By item 10(d), $E(N_j \mid X_0 = i) =$ ________ $< \infty$ for every i and j.

    (b) On the other hand, count visits a different way. Exchange the order of summation.

    $$\sum_j N_j = \sum_j \sum_n \mathbf{1}\{X_n = j\} = \sum_n \underbrace{\sum_j \mathbf{1}\{X_n = j\}}_{A} \qquad A = \underline{\hspace{2cm}}$$

    so $\sum_j N_j =$ ________ with probability 1.

    (c) Therefore $E[\sum_j N_j \mid X_0 = i] = \sum_j E(N_j \mid X_0 = i) =$ ________, contradicting (a).

Periodicity

Is recurrence enough to get a limiting distribution?

Consider the deterministic cycle $1 \to 2 \to 3 \to 1$. It is finite and irreducible, so by the lemma above every state is recurrent.

13. Fill in the table. Does $p_{12}^{(n)}$ converge?

n	0	1	2	3	4	5	6	7	8	9	10	11	12
$p_{12}^{(n)}$													

14. The **period** of state i is

$$d(i) = \begin{cases} \gcd\{n > 0 : p_{ii}^{(n)} > 0\}, & f_i > 0 \\ \infty, & f_i = 0 \end{cases}$$

and i is **aperiodic** if $d(i) =$ ________.

Informally, the possible return times to i are all ________________ of $d(i)$.

When $1 < d(i) < \infty$, passing through state i is periodic, which inhibits limiting behavior unless i is transient.

15. Periodicity is a class property: if i communicates with j, then $d(i) = d(j)$.

Proof. By hypothesis $p_{ij}^{(r)} > 0$ and $p_{ji}^{(s)} > 0$ for some $r, s > 0$.

(a) First show $d(i)$ divides $r + s$.

$$p_{ii}^{(r+s)} = \sum_k p_{ik}^{(r)} p_{ki}^{(s)} \geq \_\_\_\_\_\_\_\_\_\_\_\_ > 0 \quad \Longrightarrow \quad \_\_\_\_\_\_\_\_\_\_\_\_$$

(b) Now suppose $p_{jj}^{(n)} > 0$ for some $n > 0$, and run the excursion $i \to j \to j \to i$.

$$p_{ii}^{(r+s+n)} \geq \_\_\_\_\_\_\_\_\_\_\_\_ > 0 \quad \Longrightarrow \quad \_\_\_\_\_\_\_\_\_\_\_\_$$

It follows that $d(i)$ divides n and $d(i) \leq d(j)$.

(c) By the same argument with i and j switched, ________________

This completes the proof.

Limit theorem for a finite chain

Here we will develop necessary and sufficient conditions for the existence of limiting distributions. We first provide a sufficient condition.

16. **Lemma (L1).** For a finite Markov chain, a strong limit exists if the chain is irreducible and aperiodic.

    It can be shown (takes work) that a chain is aperiodic and irreducible just in case P is **primitive**: $(P^n)_{ij} > 0$ for every i, j at some n. We will take this as a given.

    For a primitive matrix:

    i. the Perron–Frobenius eigenvalue $\rho(P) = 1$ has algebraic multiplicity 1; and

    ii. $|\lambda| < 1$ for every other eigenvalue λ.

    Let π be the stationary distribution and denote

    $$L = \mathbf{1}\pi^T = \begin{bmatrix} \pi^T \\ \vdots \\ \pi^T \end{bmatrix}$$

    We claim that $P^n \to L$ as $n \to \infty$.

    Proof sketch.

    (a) L is the projection onto the one-dimensional right eigenspace spanned by $\mathbf{1}$, which is associated with $\lambda = 1$. Verify that it absorbs P from either side, using $P\mathbf{1} = \mathbf{1}$ and $\pi^T P = \pi^T$.

    $$PL \;=\; \underline{\hspace{4cm}} \qquad\qquad LP \;=\; \underline{\hspace{4cm}}$$

    (b) Now define $R = P - L$. Show that $LR = RL = 0$.

    (c) P and R have the same eigenvalues except for $\lambda = 1$, since L is a projection onto the associated eigenspace. So by (ii), it follows that

    $$\rho(R)\ \underline{\hspace{2cm}} \qquad \Longrightarrow \qquad R^n \longrightarrow \underline{\hspace{2cm}}$$

    (d) Finally expand the power, using $LR = RL = 0$.

    $$P^n \;=\; (L+R)^n \;=\; L^n + R^n \;=\; \underline{\hspace{4cm}} \longrightarrow \underline{\hspace{2cm}}$$

The canonical decomposition for reducible Markov chains

Irreducibility is not *necessary* for a strong limit. For example,

$$P = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad \Longrightarrow \qquad P^n \longrightarrow \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

All that is really needed is that the chain eventually transitions to one recurrent communication class with probability 1.

17. A communication class C is **closed** if no state outside C is accessible from within it:

$$p_{ij} = 0 \quad \text{for each } i \in \underline{\hspace{3em}} \text{ and } j \notin \underline{\hspace{3em}}$$

18. Recurrent classes are always closed.

    Proof. Let C be a recurrent class. Take $i \in C$ and $j \notin C$, and suppose $p_{ij} > 0$.

$$\begin{aligned} P(T_i < \infty \mid X_0 = i) &= \sum_k p_{ik}\, P(T_i < \infty \mid X_1 = k) \\ &= \sum_{k \neq j} p_{ik}\, P(T_i < \infty \mid X_1 = k) \\ &\leq \sum_{k \neq j} p_{ik} \;=\; \underline{\hspace{8em}} \\ &< 1 \end{aligned}$$

    But $f_i = 1$ by hypothesis, so this is a contradiction.

By the argument of item 12 with S replaced by C, it is immediate that for a finite chain, a communication class is closed if and only if it is recurrent.

19. It follows that for any finite Markov chain we may write the **canonical decomposition**

$$S \;=\; T \cup C_1 \cup \cdots \cup C_k$$

    where T is the set of ________________ and $C_1, \ldots, C_k$ are the ________________

Relabeling states — that is, permuting the rows and columns of P together — puts the transition matrix into **canonical form**:

$$P = \begin{array}{c} \\ T \\ C_1 \\ \vdots \\ C_k \end{array} \begin{array}{c} \begin{array}{cccc} T & C_1 & \cdots & C_k \end{array} \\ \begin{bmatrix} Q & R_1 & \cdots & R_k \\ 0 & P_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_k \end{bmatrix} \end{array} = \begin{bmatrix} Q & R \\ 0 & B \end{bmatrix}$$

Now under the canonical decomposition, we have for a reducible finite chain that:

$$P^n = \begin{bmatrix} Q^n & \sum_{m=0}^{n-1} Q^m R\, B^{n-m-1} \\ 0 & B^n \end{bmatrix}$$

20. **Lemma (L2).** P^n converges just in case every closed class is aperiodic.

    Proof sketch. Convergence of P^n reduces to convergence of the three blocks.

    (a) Owing to transience, $Q^n\mathbf{1} < \mathbf{1}$ entrywise for all $n > n^*$. Hence

    $$\rho(Q) \_\_\_\_\_\_\_\_ \quad \Longrightarrow \quad Q^n \longrightarrow \_\_\_\_\_\_\_\_ \quad \text{always}$$

    (b) B is block diagonal, so B^n converges just in case $P_1^n, \dots, P_k^n$ converge. We claim that

    $$P_i^n \text{ converges} \quad \Longleftrightarrow \quad C_i \text{ is aperiodic}$$

    We show this in two parts.

    i. ($\Leftarrow$) If each C_i is aperiodic, then each P_i is __________________, so by Lemma L1

    $$P_i^n \longrightarrow \_\_\_\_\_\_\_\_$$

    ii. ($\Rightarrow$) Suppose P_i^n converges but $d(j) > 1$ for some $j \in C_i$. Then

    $$(P_i^n)_{jj} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \quad \text{for } n \bmod d(j) \neq 0$$

    but

    $$(P_i^{nd(j)})_{jj} \_\_\_\_\_\_\_\_ \quad \text{for all } n > n^*$$

    which is a contradiction because ______________________________

    So $d(j) = 1$ and C_i is aperiodic.

    (c) Lastly, write $L = \lim_n B^n$ and split the off-diagonal block into an error term and a limit term.

    $$\sum_{m=0}^{n-1} Q^m R B^{n-m-1} = \underbrace{\sum_{m=0}^{n-1} Q^m R\big(B^{n-m-1} - L\big)}_{A} + \underbrace{\sum_{m=0}^{n-1} Q^m R L}_{B}$$

    $$A \longrightarrow \_\_\_\_\_\_\_\_ \qquad B \longrightarrow \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

So P^n converges if and only if $C_1, \dots, C_k$ are aperiodic, and the limit is

$$P^n \longrightarrow \begin{bmatrix} 0 & (I-Q)^{-1}R\Lambda \\ 0 & \Lambda \end{bmatrix}, \qquad \Lambda = \mathrm{diag}(\mathbf{1}\pi_1^T, \dots, \mathbf{1}\pi_k^T)$$

The matrix $F = (I - Q)^{-1}$ is called the **fundamental matrix** of the chain. We state the following without proof; they are true for any finite Markov chain.

quantity	interpretation
F_{ij}	expected number of visits to j starting from i
$(F\mathbf{1})_i$	expected time to absorption, $\tau_C = \inf\{n \geq 0 : X_n \notin T\}$
$(FR_i\mathbf{1})_\ell$	probability of absorption into C_i starting from ℓ
$E(T_j \mid X_0 = j) = 1/\pi_j$	expected return time, for $j \in C_1 \cup \cdots \cup C_k$

21. **Theorem (T1).** For a finite Markov chain $\{X_n\}$:

    i. a **weak** limiting distribution exists if and only if

    ii. a **strong** limiting distribution exists if and only if

    iii. a **stationary** distribution always exists.

    Proof. Statements (i)-(ii) follow from Lemma L2; statement (iii) was established previously from the Perron–Frobenius theorem.

22. **Algorithm.** To find the limiting distribution, when it exists:

    (a) write P in ________________ form;

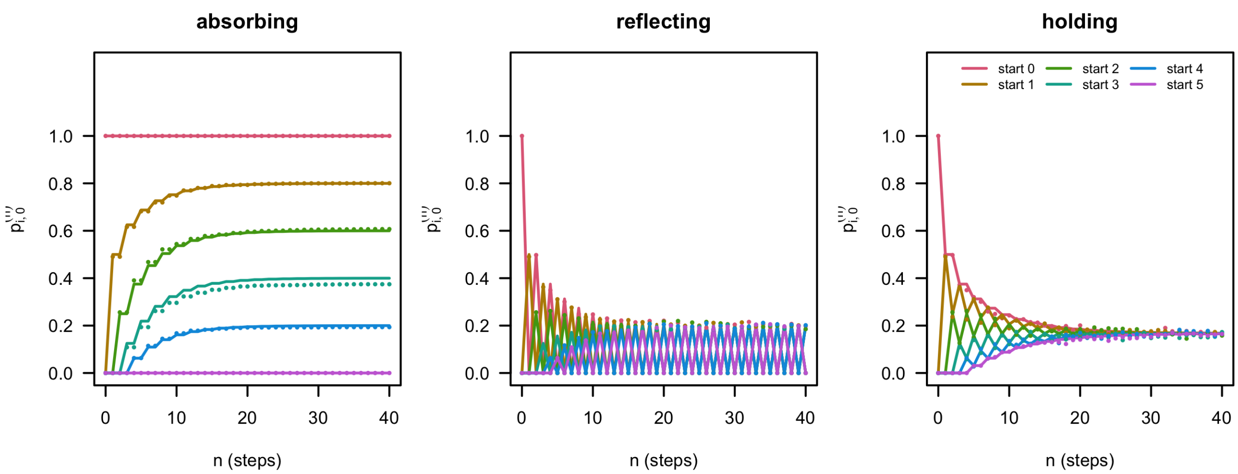
    (b) obtain the ________________ eigenvector π_i^T for each of $P_1, \ldots, P_k$;

    (c) compute ________________

Back to gambler's ruin

Here again are the three variants we simulated, on $S = \{0, 1, \dots, 5\}$ with $p = \frac{1}{2}$: **absorbing** barriers, **reflecting** barriers, and **holding** barriers.

$$\begin{array}{ccc}
\text{absorbing} & \text{reflecting} & \text{holding} \\
\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} &
\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} &
\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}
\end{array}$$



23. Work out whether a limit exists in each case, and if so what kind. Check your answer against the plots.

chain	communication classes	closed classes	periods	weak / strong / none
absorbing				
reflecting				
holding				