EE5137 Stochastic Processes: Problem Set 9
Assigned: 18/03/22, Due: 25/03/22

There are six (6) non-optional problems in this problem set.

1. Exercise 4.12 (Gallager's book)

   (a) Let λ_k be an eigenvalue of a stochastic matrix $[P]$ and let $\pi^{(k)}$ be a left eigenvector for λ_k. Show that for each component $\pi_j^{(k)}$ of $\pi^{(k)}$ and each n that

   $$\lambda_k^n \pi_j^{(k)} = \sum_i \pi_i^{(k)} P_{ij}^n. \tag{1}$$

   Solution: By the definition of an eigenvector, $\pi^{(k)}[P] = \lambda_k \pi^{(k)}$. By iterating this, $\pi^{(k)}[P^2] = \lambda_k \pi^{(k)}[P] = \lambda_k^2 \pi^{(k)}$, and in general $\pi^{(k)}[P^n] = \lambda_k^n \pi^{(k)}$. The desired expression is the j-th component of this vector equation.

   (b) By taking magnitudes of each side and looking at the appropriate j, show that

   $$|\lambda_k|^n \leq M. \tag{2}$$

   Hint for Part (b): Choose j to maximize $|\pi_j^{(k)}|$ for the given k.

   Solution: Choose j to maximize $|\pi_j^{(k)}|$ for the given k. Then from (a) and the fact that $P_{ij}^n \leq 1$,

   $$|\lambda_k^n| \max_j |\pi_j^{(k)}| \leq \sum_i |\pi_i^{(k)}| P_{ij}^n \leq \sum_i [\max_j |\pi_j^{(k)}|] = M[\max_j |\pi_j^{(k)}|]. \tag{3}$$

   Canceling the common term, $|\lambda_k^n| \leq M$.

   (c) Show that $|\lambda_k| \leq 1$.
   Solution: Note that $|\lambda_k|^n$ is exponential in n. Since $|\lambda_k|$ is nonnegative and real, this is exponential and increasing in n without bound if $|\lambda_k| > 1$. Since $|\lambda_k|^n$ is bounded by M, $|\lambda_k| \leq 1$.

2. Exercise 4.16 (Gallager's book)

   (a) Let λ be an eigenvalue of a matrix $[A]$, and let ν and π be the right and left eigenvectors respectively of λ, normalized so that $\pi\nu = 1$. Show that

   $$[[A] - \lambda\nu\pi]^2 = [A^2] - \lambda^2\nu\pi. \tag{4}$$

   Solution: We simply multiply out the original square,

   $$\begin{aligned} [[A] - \lambda\nu\pi]^2 &= [A^2] - \lambda\nu\pi[A] - \lambda[A]\nu\pi + \lambda^2\nu\pi\nu\pi && (5) \\ &= [A^2] - \lambda^2\nu\pi - \lambda^2\nu\pi + \lambda^2\nu\pi && (6) \\ &= [A^2] - \lambda^2\nu\pi. && (7) \end{aligned}$$

(b) Show that $[[A^n] - \lambda^n \nu\pi][[A] - \lambda\nu\pi] = [A^{n+1}] - \lambda^{n+1}\nu\pi$.
Solution: This is essentially the same as (a)

$$[[A^n] - \lambda^n \nu\pi][[A] - \lambda\nu\pi] = [A^{n+1}] - \lambda^n \nu\pi[A] - \lambda[A^n]\nu\pi + \lambda^{n+1}\nu\pi \tag{8}$$
$$= [A^{n+1}] - \lambda^{n+1}\nu\pi. \tag{9}$$

(c) Use induction to show that $[[A] - \lambda\nu\pi]^n = [A^n] - \lambda^n\nu\pi$.
Solution: (a) gives the base of the induction and (b) gives the inductive step.

3. Exercise 4.17 (Gallager's book)

(a) If we imagine multiplying $[P][P]$ by straightforward block multiplication, we get

$$[P^2] = \begin{bmatrix} [P_\mathcal{T} P_\mathcal{T}] & [P_\mathcal{T} P_{\mathcal{T}\mathcal{R}} + P_{\mathcal{R}\mathcal{T}} P_\mathcal{R}] \\ [0] & [P_\mathcal{R} P_\mathcal{R}] \end{bmatrix}$$

In the same way, using induction, and not tracking the upper right term

$$[P^2] = \begin{bmatrix} [P_\mathcal{T}^n] & [P_x^n] \\ [0] & [P_\mathcal{R}^n] \end{bmatrix}$$

Another approach is to look at P_{ij}^n. If $i, j \in T$, then any n step walk from i to j stays in $\mathcal{T}$. Similarly, if $i, j \in \mathcal{R}$, any walk from i to j stays in $\mathcal{R}$. Finally, there are no walks from $\mathcal{R}$ to $\mathcal{T}$, giving rise to the matrix $[0]$ in the lower left.

(b) As specified in Figure 4.5, t is the number of transient states. As explained in Section 4.3.4, each transient state has a path of at most t steps to a recurrent state. All extensions of that path remain in recurrent states and thus if the original path has fewer than t steps, any extension of that path to a walk with t steps still ends at a recurrent state and has positive probability. Thus q_i is positive.

(c) Since $\sum_{k\in\mathcal{R}} P_{ik}^t \geq q$, it follows that $\sum_{j\in\mathcal{T}} P_{ij}^t \leq 1 - q$ for all $i \in \mathcal{T}$. Using the Chapman-Kolmogorov equations for $i \in \mathcal{T}$,

$$\sum_{j\in\mathcal{T}} P_{ij}^{2t} = \sum_{k,j\in\mathcal{T}} P_{ik}^t P_{kj}^t \leq \sum_{k\in\mathcal{T}} P_{ik}^t (1-q) \leq (1-q)^2$$

Iterating this equation, we get $\sum_{j\in\mathcal{T}} P_{ij}^{nt} \leq (1-q)^n$ and thus $P_{ij}^{nt} \leq (1-q)^n$ for all $j \in \mathcal{T}$.

(d) Note that $\boldsymbol{\pi}$ is also an eigenvector of $[P^n]$ for all n, so that $\boldsymbol{\pi}_\mathcal{T}[P_\mathcal{T}]^n = \boldsymbol{\pi}_\mathcal{T}$. Since $\lim_{n\to\infty}[P_\mathcal{T}]^n = 0$, this shows that $\boldsymbol{\pi}_\mathcal{T} = 0$. We then have $\boldsymbol{\pi}_\mathcal{R}[P_\mathcal{R}]^n = \boldsymbol{\pi}_\mathcal{R}$ which has a unique solution. To prove uniqueness, please see the proof of Theorem 4.3.5 in the textbook.

(e) $\boldsymbol{e}$ must be a right eigenvector of $[P]$ since $[P]$ is stochastic (i.e., its rows all sum to 1, which is what $[P]\boldsymbol{e} = \boldsymbol{e}$ means). We now prove uniqueness. From (4.30) of the book, we see that

$$[P^n] = \sum_{i=1}^{M} \lambda_i^n \boldsymbol{\nu}^{(i)} \boldsymbol{\pi}^{(i)} \tag{10}$$

where $\{\boldsymbol{\nu}^{(i)}\}$ and $\{\boldsymbol{\pi}^{(i)}\}$ are the right- and left-eigenvectors respectively. For ergodic unichains (Theorem 4.3.7 and see the discussion below Eqn. (4.30) of the textbook), we know that

$$\lim_{n\to\infty} [P^n] = \mathbf{e}\boldsymbol{\pi}. \tag{11}$$

Note that $\lambda_1 = 1$ and

$$\mathbf{e} = \boldsymbol{\nu}^{(1)}, \quad \text{and} \quad \boldsymbol{\pi} = \boldsymbol{\pi}^{(1)}. \tag{12}$$

By Exercise 4.12, we know that all eigenvalues λ_k are such that $|\lambda_k| \le 1$. Thus by (10) and (11), we know that all terms except that corresponding to $i = 1$ die out as $n \to \infty$. Hence, it has a single eigenvalue equal to 1 or $|\lambda_i| < 1$ for all $i > 1$. Thus, the right-eigenvector $\mathbf{e}$ must be unique. Why exactly? Let $\lambda_2 \ne \lambda_1 = 1$. Then

$$[P]\boldsymbol{\nu}^{(i)} = \lambda_i \boldsymbol{\nu}^{(i)}, \quad i = 1, 2. \tag{13}$$

Then $\boldsymbol{\nu}^{(1)}$ must be linearly independent of $\boldsymbol{\nu}^{(2)}$. Suppose, to the contrary, they were linearly dependent, i.e., $\boldsymbol{\nu}^{(2)} = c\boldsymbol{\nu}^{(1)}$ for some $c \ne 0$. Then

$$[P]\boldsymbol{\nu}^{(2)} = \lambda_2 \boldsymbol{\nu}^{(2)} \quad \Longrightarrow \quad [P]c\boldsymbol{\nu}^{(1)} = \lambda_2 c\boldsymbol{\nu}^{(1)} \quad \Longrightarrow \quad [P]\boldsymbol{\nu}^{(1)} = \lambda_2 \boldsymbol{\nu}^{(1)} \tag{14}$$

The last equation means that $\lambda_2 = \lambda_1$ (by (13)), which is a contradiction.

4. Consider the Markov chain whose transition probability matrix is

$$[P] = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 \end{bmatrix}.$$

The first row/column corresponds to state 0 and the last row/column corresponds to state 5 and so on.

(a) (4 marks) Classify the states $\{0, 1, 2, 3, 4, 5\}$ into classes.

(b) (4 marks) Identify the recurrent and transient states.

(c) (4 marks) Compute the period of each recurrent class.

(d) (3 marks) Identify the ergodic states.

(e) (5 marks) If the chain starts from state 1, find the steady state probabilities in each of the states $(\pi_0, \pi_1, \ldots, \pi_5)$.

(f) (5 marks) Assuming again we start from state 1, evaluate

$$\lim_{n\to\infty} -\frac{1}{n} \log |[P^n]_{11} - \pi_1|,$$

where $[M]_{ij}$ is the (i, j) element of the matrix $[M]$.

Solution:

(a) The three classes are $\mathcal{C}_1 = \{1, 5\}$, $\mathcal{C}_2 = \{0, 2, 4\}$ and $\mathcal{C}_3 = \{3\}$.

(b) The first two classes are recurrent while the last is transient.

(c) Period of $\mathcal{C}_1 = 1$ and $\mathcal{C}_2 = 3$.

(d) The ergodic states are those in $\mathcal{C}_1$.

(e) Solving $\pi P = \pi$, we get

$$\pi^{(1)} = [0, 1/3, 0, 0, 0, 2/3] \quad \text{and} \quad \pi^{(2)} = [1/3, 0, 1/3, 0, 1/3, 0].$$

If we start from state 1, then the steady state vector is $\pi^{(1)}$.

(f) Since we start from state 1, the effective transition matrix is

$$[P] = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$$

The second largest eigenvalue is 1/2 and so the required limit is $\log 2$.

5. A fly moves along a straight line in unit increments. At each time period, it moves one unit to the left with probabiliy 0.3, one unit to the right with probability 0.3 and stays in place with probability 0.4, independent of past movements. Two spiders are lurking at positions 1 and M; if a fly lands in positions 1 or M, it is captured by the spider and the process terminates. Let $j \in \{1, 2, \ldots, M\}$ be the position of the fly. The Markov chain $[P]$ is thus given by

$$p_{11} = 1 \quad p_{mm} = 1, \quad p_{ij} = \begin{cases} 0.3 & \text{if } |j-i| = 1 \\ 0.4 & \text{if } j = i \end{cases} \qquad \text{for} \quad i = 2, 3, \ldots, M-1.$$

Write down a system of linear equations to deduce the expected number of steps the fly takes given it starts from state j before being captured by one of the two spiders. This is denoted as v_j. For $M = 4$, solve your equations to find v_2 and v_3.

Solution: The system of linear equations is

$$\begin{aligned} v_i &= 1 + P_{i,i-1}v_{i-1} + P_{i,i+1}v_{i+1} + P_{i,i}v_i \qquad i = 2, \ldots, M-1 \\ &= 1 + 0.3(v_{i-1} + v_{i+1}) + 0.4v_i \end{aligned}$$

When $M = 4$, we have

$$\begin{aligned} v_2 &= 1 + 0.3v_3 + 0.4v_2 \\ v_3 &= 1 + 0.3v_2 + 0.4v_3 \end{aligned}$$

which yields

$$v_2 = v_3 = \frac{10}{3}.$$

6. Consider the Markov chain

$$[P] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1/3 & 2/3 & 0 \\ 0 & 0 & 0 & 1 \\ 3/5 & 0 & 0 & 2/5 \end{bmatrix}$$

The steady-state probabilities are known to be

$$\pi_1 = \frac{6}{31} \quad \pi_2 = \frac{9}{31} \quad \pi_3 = \frac{6}{31} \quad \pi_4 = \frac{10}{31}.$$

Assume that the process is in state 1 before the first transmission.

(a) What's the probability that the process will be in state 1 just after the sixth transition?

(b) Determine the expected value and variance of the number of transitions up to and including the next transition during which the process returns to state 1.

(c) What is (approximately) the probability that the state of the system resulting from transition 1000 is neither the same as the state resulting from transition 999 nor the same as the state resulting from transition 1001?

Solutions:

(a) There are 3 different paths that lead back to state 1 after 6 transitions. One path makes two self-transitions at state 2, one path makes two self-transitions at state 4, one path makes one self-transition at state 2 and one self-transition at state 4. By adding the probabilities of these three paths, we obtain

$$r_{11}(6) = \frac{2}{3} \cdot \frac{3}{5} \cdot \left(\frac{1}{3} \cdot \frac{2}{5} + \frac{1}{9} + \frac{4}{25} \right) = \frac{182}{1125}.$$

(b) The time T until the process returns to state 1 is equal to 2 (the time it takes for the transitions from 1 to 2 and from 3 to 4), plus the time it takes for the state to move from state 2 to state 3 (this is geometrically distributed with parameter $p = 2/3$), plus the time it takes for the state to move from state 4 to state 1 (this is geometrically distributed with parameter $p = 3/5$). Using the formulas $\mathbb{E}[X] = 1 - p$ and $\mathrm{var}(X) = (1-p)/p^2$ for the mean and variance of a geometric random variable, we find that

$$\mathbb{E}[T] = 2 + \frac{3}{2} + \frac{5}{3} = \frac{31}{6}$$

and

$$\mathrm{var}(T) = \left(1 - \frac{2}{3}\right) \frac{3^2}{2^2} + \left(1 - \frac{3}{5}\right) \frac{5^2}{3^2} = \frac{67}{36}.$$

(c) Let A be the event that X_{999}, X_{1000}, and X_{1001} are all different. Note that

$$\Pr(A \mid X_{999} = i) = \begin{cases} 2/3 & i = 1, 2 \\ 3/5 & i = 3, 4 \end{cases}.$$

Thus, using the total probability theorem, and assuming that the process is in steady-state at time 999, we obtain

$$\Pr(A) = \frac{2}{3}(\pi_1 + \pi_2) + \frac{3}{5}(\pi_3 + \pi_4) = \frac{98}{155}.$$

---

7. (Optional) Exercise 4.11 (Gallager's book)

(a) If we first premultiply $[P]$ by $\boldsymbol{\pi}^{(j)}$, we get

$$\boldsymbol{\pi}^{(j)}[P]\boldsymbol{\nu}^{(i)} = \lambda_j \boldsymbol{\pi}^{(j)}\boldsymbol{\nu}^{(i)}.$$

If we postmultiply $[P]$ by $\boldsymbol{\nu}^{(i)}$, we get

$$\boldsymbol{\pi}^{(j)}[P]\boldsymbol{\nu}^{(i)} = \lambda_i \boldsymbol{\pi}^{(j)}\boldsymbol{\nu}^{(i)}.$$

Thus $(\lambda_j - \lambda_i)\boldsymbol{\pi}^{(j)}\boldsymbol{\nu}^{(i)} = 0$. Since $\lambda_j \neq \lambda_i$, we must have $\boldsymbol{\pi}^{(j)}\boldsymbol{\nu}^{(i)} = 0$.

(b) Order the right eigenvectors so that the ith column of $[U]$ is $\boldsymbol{\nu}^{(i)}$ with eigenvalue λ_i. Let $[V]$ be a matrix whose rows are linearly independent left eigenvectors of $[P]$, ordered so that the ith row is $\boldsymbol{\pi}^{(i)}$ with eigenvalue λ_i.

Then $[VU]_{ji}$ is $\boldsymbol{\pi}^{(j)}\boldsymbol{\nu}^{(i)}$. From part (a), this is 0 for $i \neq j$, from which it follows that $[VU]$ is diagonal. We next show that the diagonal elements of $[VU]$ must be nonzero. Since $\boldsymbol{\pi}^{(i)}$ is an eigenvector, it is nonzero by definition. Since $[U]$ is nonsingular, $\boldsymbol{\pi}^{(i)}[U]$ is nonzero. The elements of $\boldsymbol{\pi}^{(i)}[U]$ are $\{\boldsymbol{\pi}^{(i)}\boldsymbol{\nu}^{(j)} : 1 \leq j \leq M\}$. These elements are 0 for all $j \neq i$, so $\boldsymbol{\pi}^{(i)}\boldsymbol{\nu}^{(j)}$ must be nonzero. Thus the ith element on the diagonal on $[VU]$ is nonzero for each i.

The eigenvector $\boldsymbol{\pi}^{(i)}$ of λ_i (assuming M distinct eigenvalues) is unique within a scale factor, but that scale factor can now be chosen so that $\boldsymbol{\pi}^{(i)}\boldsymbol{\nu}^{(i)} = 1$. With this scaling, $[VU] = [I]$ which means that $[V] = [U^{-1}]$. Thus, the rows of $[U^{-1}]$ are the left eigenvectors of $[P]$, scaled so that $\boldsymbol{\pi}^{(i)}\boldsymbol{\nu}^{(i)} = 1$ for each i.

(c) Since each column of $[A]$ is zero other than the ith, each column of $[UA]$ is zero other than the ith. Since the ith column of $[U]$ is $\boldsymbol{\nu}^{(i)}$, the ith column of $[UA]$ is $a\boldsymbol{\nu}^{(i)}$. Thus we have

$$[UAU^{-1}]_{kj} = \sum_l [UA]_{kl}[U^{-1}]_{lj} = [UA]_{ki}[U^{-1}]_{ij} = a\nu_k^{(i)}\pi_j^{(i)}.$$

Putting these terms together,

$$[UAU^{-1}] = a\boldsymbol{\nu}^{(i)}\boldsymbol{\pi}^{(i)}.$$

(d) Note that $[\Lambda^n]$ is a diagonal matrix with $[\Lambda^n]_{ii} = \lambda_i^n$. Let $[\Lambda_i^n]$ be a diagonal matrix with a single non-zero element, λ_i^n in position i. Then $[\Lambda] = \sum_{i=1}^M [\Lambda_i^n]$. It follows that

$$[U\Lambda^n U^{-1}] = [U]\left[\sum_{i=1}^M [\Lambda_i^n]\right][U^{-1}] = \sum_{i=1}^M [U\Lambda_i^n U^{-1}] = \sum_{i=1}^M \lambda_i^n \boldsymbol{\nu}^{(i)}\boldsymbol{\pi}^{(i)},$$

where we have used the result of (c) in each term.

8. (Optional) Exercise 4.18 (Gallager's book)

(a) Number the states so that the transient states come first, then the states in recurrent class 1, and so forth up to the recurrent states in class κ. Then $[P]$ has the following block structure

$$[P] = \begin{bmatrix} [P_{\mathcal{T}}] & [P_{\mathcal{T}\mathcal{R}_1}] & \dots & \dots & [P_{\mathcal{T}\mathcal{R}_k}] \\ [0] & [P_{\mathcal{R}_1}] & [0] & \dots & [0] \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ [0] & [0] & \dots & \dots & [P_{\mathcal{R}_k}] \end{bmatrix}$$

The lth recurrent class has an $|\mathcal{R}_l|$ by $|\mathcal{R}_l|$ transition matrix which, viewed alone, has an eigenvalue $\lambda = 1$ of multiplicity 1, a corresponding unique (within a scale factor) left eigenvector, say $\boldsymbol{\pi}(\mathcal{R}_l)$, and a corresponding unique (within a scale factor) right eigenvector, $\boldsymbol{\nu}(\mathcal{R}_l) = (1, 1, \dots, 1)^T$ (see Theorem 4.4.2).

Let $\boldsymbol{\pi}^{(l)}$ be an M dimensional row vector whose components are equal to those of $\boldsymbol{\pi}(\mathcal{R}_l)$ over the states of $\mathcal{R}_l$ and equal to 0 elsewhere. Then it can be seen by visualizing elementary row/matrix multiplication on the block structure of $[P]$ that $\boldsymbol{\pi}^{(l)}[P] = \boldsymbol{\pi}^{(l)}$. This gives us κ left eigenvectors of eigenvalue 1, one for each recurrent class $\mathcal{R}_l$.

If we look at the determinant equation, $\det[P - \lambda I]$, the upper triangular block structure of $[P]$ implies that $\det[P - \lambda I]$ breaks into the product of $\kappa + 1$ determinants, one for the transient states and one for each recurrent class. Thus we can identify the eigenvalues with the classes, $|\mathcal{T}|$ eigenvalues for $\mathcal{T}$ and, for each l, $|\mathcal{R}_l|$ eigenvalues for $\mathcal{R}_l$. Each recurrent class has one eigenvalue equal to 1 and the transient set has no eigenvalues equal to 1. To verify this latter statement, assume the contrary, i.e., assume there is a row vector $\boldsymbol{\pi}^{(T)}$ hat is nonzero only on the components of $\mathcal{T}$ and which satisfies $\boldsymbol{\pi}^{(T)} = \boldsymbol{\pi}^{(T)}[P]$. Then we must also have $\boldsymbol{\pi}^{(T)} = \boldsymbol{\pi}^{(T)}[P^n]$ for all $n > 1$. However, we have seen that $P_{ij}^n \to 0$ as $n \to \infty$ for all $i, j \in \mathcal{T}$. This yields the desired contradiction, since there can then be no nonzero solution to $\boldsymbol{\pi}^{(T)} = \boldsymbol{\pi}^{(T)}[P]$ for large n.

We have thus shown that $\lambda = 1$ has multiplicity κ for matrix $[P]$ and we have identified all the left eigenvectors of eigenvalue 1 (or more strictly, we have identified a basis for the vector space of left eigenvectors of eigenvalue 1).

(b) If there are no transient states, then a set of κ right eigenvectors can be chosen in the same way as the left eigenvectors. That is, for each $1 \le l \le k$ the components of $\boldsymbol{\nu}^{(l)}$ can be chosen to be 1 for each state in $\mathcal{R}_l$ and 0 for all other states. This doesn't satisfy the eigenvector equation if there are transient states, however. We now show, instead, that for each $1 \le l \le k$ there is a right eigenvector $\boldsymbol{\nu}^{(l)}$ of eigenvalue 1 such that $\nu_i^{(l)} = 0$ for all $i \in \mathcal{R}_m$, for each $m \ne l$. Finally we will show that these κ vectors are linearly independent and have the properties specified in the problem statement.

The right eigenvector equation that must be satisfied by $\boldsymbol{\nu}^{(l)}$ with the conditions above can be written out component by component, getting

$$\nu_i^{(l)} = \sum_{j \in \mathcal{R}_l} P_{ij} \nu_j^{(l)} \qquad i \in \mathcal{R}_l$$
$$\nu_i^{(l)} = \sum_{j \in \mathcal{T}} P_{ij} \nu_j^{(l)} + \sum_{j \in \mathcal{R}_l} P_{ij} \nu_j^{(l)} \qquad i \in \mathcal{T}$$

The first set of equations above are simply the usual right eigenvector equations for eigenvalue 1 over the recurrent submatrix $[P_{\mathcal{R}_l}]$. Thus $\nu_j^{(l)} = 1$ for $j \in \mathcal{R}_l$ and this solution (over $\mathcal{R}_l$) is unique within a scale factor. Substituting this solution into the second set of equations, we get

$$\nu_i^{(l)} = \sum_{j \in \mathcal{T}} P_{ij} \nu_j^{(l)} + \sum_{j \in \mathcal{R}_l} P_{ij} \qquad i \in \mathcal{T}$$

We can view this as a vector/matrix equation by letting $\boldsymbol{\nu}_{\mathcal{T}}^{(l)}$ to be the vector $\boldsymbol{\nu}^{(l)}$ restricted to the transient states. Thus, $\boldsymbol{\nu}_{\mathcal{T}}^{(l)} = [P_{\mathcal{T}}]\boldsymbol{\nu}_{\mathcal{T}}^{(l)} + \boldsymbol{x}$ where $x_i = \sum_{j \in \mathcal{R}_l} P_{ij}$. Since $\mathcal{T}$ is a set of transient states, $[P_{\mathcal{T}}]$ does not have 1 as an eigenvalue, so $[P_{\mathcal{T}} - I]$ is non-singular. Thus this set of equations has a unique solution (we have already specified the scale factor on $\boldsymbol{\nu}^{(l)}$ by setting $\nu_i^{(l)} = 1$ for $i \in \mathcal{R}_l$).

We have now established that for each $1 \le l \le k$, $[P]$ has a right eigenvector of eigenvalue 1 that is nonzero only on $\mathcal{R}_l$ and $\mathcal{T}$. Since the components of $\boldsymbol{\nu}^{(l)}$ are 1 on $\mathcal{R}_l$ and 0 on all other recurrent classes, these κ vectors are clearly linearly independent. Since the eigenvalue 1 has multiplicity κ, these eigenvectors form a basis for the eigenvectors of eigenvalue 1.

It remains only to show that $\nu_i^{(l)} = \lim_{n \to \infty} \Pr(X_n \in \mathcal{R}_l | X_0 = i)$. We have seen many times that if $\boldsymbol{\nu}^{(l)}$ is a right eigenvector of $[P]$ of eigenvalue 1, then it is also a right eigenvector of $[P^n]$ of eigenvalue 1. Since $[P^n]$ has the same block structure as $[P]$, we can repeat the argument leading to the above displayed equation. For each $n > 1$, then, we have

$$\nu_i^{(l)} = \sum_{j \in \mathcal{T}} P_{ij}^n \nu_j^{(l)} + \sum_{j \in \mathcal{R}_l} P_{ij}^n, \qquad i \in \mathcal{T}.$$

In the limit $n \to \infty$, each P_{ij}^n for $i, j \in \mathcal{T}$ goes to 0. Also $\sum_{j \in \mathcal{R}_l} P_{ij}^n$ is the probability that $X_n \in \mathcal{R}_l$ given $X_0 = i$. Since there is no exit from $\mathcal{R}_l$, this quantity is non-decreasing in n and bounded by 1, so it has a limit. This limit is the probability of ever going from i to $\mathcal{R}_l$, completing the derivation.

(c) Unfortunately, this result is not true in general. The additional restriction that each recurrent class is ergodic must be added. In this case, P_{ij}^n for $i \in \mathcal{T}$ and $j \in \mathcal{R}_m$ converges to the probability that class $\mathcal{R}_m$ is entered from i times the steady-state probability of j within class $\mathcal{R}_m$.

9. (Optional) Consider an m-state Markov chain where each state is either transient or absorbing (i.e., a state i is absorbing if $P_{ii} = 1$). Fix an absorbing state s. Show that the probabilities a_i of eventually reaching state s starting from a state i are the unique solutions to the equations

$$\begin{aligned} a_s &= 1 & & \\ a_i &= 0 & & \text{for all absorbing } i \ne s \\ a_i &= \sum_{j=1}^{m} P_{ij} a_j & & \text{for all transient } i \end{aligned}$$

Solution: The equations $a_s = 1$ and $a_i = 0$ for all absorbing $i \neq s$ are evident. Let i be a transient state and A be the event that state s is eventually reached. Then

$$\begin{aligned}
a_i &= \Pr(A \mid X_0 = i) \\
&= \sum_{j=1}^{m} \Pr(A \mid X_0 = i, X_1 = j) \Pr(X_1 = j \mid X_0 = i) \\
&= \sum_{j=1}^{m} \Pr(A \mid X_1 = j) P_{ij} \\
&= \sum_{j=1}^{m} a_j P_{ij}
\end{aligned}$$

We now only have to prove uniqueness. To show this, let $\bar{a}_i$ be another solution. Let $\delta_i = \bar{a}_i - a_i$. Denote B to be the set of absorbing states. By using the fact that $\delta_j = 0$ for all $j \in B$,

$$\delta_j = \sum_{i=1}^{m} P_{ij}\delta_j = \sum_{j \notin B} P_{ij}\delta_j \qquad \text{for all transient } i.$$

Applying this relation m successive times, we obtain

$$\delta_i = \sum_{j_1 \notin B} P_{ij_1} \sum_{j_2 \notin B} P_{j_1 j_2} \cdots \sum_{j_m \notin B} P_{j_{m-1} j_m} \delta_{j_m}.$$

Hence,

$$\begin{aligned}
|\delta_i| &\leq \sum_{j_1 \notin B} P_{ij_1} \sum_{j_2 \notin B} P_{j_1 j_2} \cdots \sum_{j_m \notin B} P_{j_{m-1} j_m} |\delta_{j_m}| \\
&= \Pr(X_1 \notin B, \ldots, X_m \notin B \mid X_0 = i) |\delta_{jm}| \\
&\leq \Pr(X_1 \notin B, \ldots, X_m \notin B \mid X_0 = i) | \max_{j \notin B} |\delta_{jm}|
\end{aligned}$$

The above relation holds for all transient i, so we obtain

$$\max_{j \notin B} |\delta_j| \leq \beta \max_{j \notin B} |\delta_{jm}|$$

where $\beta := \Pr(X_1 \notin B, \ldots, X_m \notin B \mid X_0 = i)$.

Note that $\beta < 1$, because there is a positive probability that X_m is absorbing, regardless of the initial state. It follows that $\max_{j \notin B} |\delta_j| = 0$ or $a_i = \bar{a}_i$ for all i that are not absorbing. We also have $a_j = \bar{a}_j$ for all absorbing j, so $a_i = \bar{a}_i$ for all i.