EE5137 : Stochastic Processes (Spring 2021)
Some Additional Notes on Convergence of Markov Chains

Vincent Y. F. Tan

March 13, 2021

In this document, we provide some supplementary material to Lectures 8 and 9 on 12th and 19th of March 2021. You need to know everything here.

1 Different Types of Markov chains

Here are some different types of Markov chains. Here is a 6-state example from Fig. 4.1(a) of Gallager's book.

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

Here, there are two recurrent classes $\{2,3\}$ and $\{5\}$. The rest are transient states. Clearly, this is not a unichain. We see that

$$\pi^{(1)} = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \pi^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \tag{2}$$

are stochastic row vectors that satisfy $\pi^{(1)} = \pi^{(1)}[P]$ and $\pi^{(2)} = \pi^{(2)}[P]$. Hence, we do not have a unique steady-state vector. Indeed, we may get trapped in either of the two recurrent classes if we start at an arbitrary state in the chain. Another way to check this is to perform the eigen-decomposition of $[P]$. We see that the set of eigenvalues is $\{1, 1, 0.5, 1/3, 0, -1\}$. There is a repeated eigenvalue at 1, which means that there are two linearly independent π satisfying $\pi = \pi[P]$. One can check that the two (normalized) left-eigenvectors corresponding to the two eigenvalues 1 are those in (2). So we see that for a non-unichain, we do not have unique probability vectors π satisfying $\pi = \pi[P]$.

Next, I want to get rid of the recurrent class at $\{5\}$. I modify state 5's outgoing edges so that with probability 1/2 it stays at state 5; with probability 1/2 it goes to state 1. Hence, the modified state transition matrix is

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

Now the set of eigenvalues is $\{1, 0.5, 0.5, 1/3, 0, -1\}$ and there is only one 1. The solitary left-eigenvector corresponding to the eigenvalue 1 is $\pi^{(1)}$ in (2). This makes sense because all the states apart from $\{2,3\}$ are transient and intuitively, the fraction of time we spend in 2 and 3 is half once get trapped in that recurrent class with period 2. However, $[P^n]$ does not converge to a rank-1 matrix. In fact,

$$[P^\infty] = \begin{bmatrix} 0 & 0.3333 & 0.6667 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.5278 & 0.4722 & 0 & 0 & 0 \\ 0 & 0.5556 & 0.4444 & 0 & 0 & 0 \\ 0 & 0.7222 & 0.2778 & 0 & 0 & 0 \end{bmatrix} \tag{4}$$

which clearly has rank 2.

I further modify the chain to show another behavior. Now, everything stays unchanged except that $P_{23} = 0.1$ and $P_{26} = 0.9$ (added an outgoing edge from 2 so that the class $\{2, 3\}$ is no longer periodic) so

$$[P] = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0 & 0.9 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 \end{bmatrix}. \tag{5}$$

Now, this is a unichain because there is a single recurrent class $\{1, 2, 3, 5, 6\}$ that is aperiodic; only state 4 is transient. We can check that the set of eigenvalues $\{-0.3779 \pm 0.5917i, 1, 0.3779 \pm 0.2923i, 1/3\}$ only contains one 1. Furthermore,

$$[P^\infty] = \begin{bmatrix} 0.2118 & 0.2353 & 0.1294 & 0 & 0.2118 & 0.2118 \\ 0.2118 & 0.2353 & 0.1294 & 0 & 0.2118 & 0.2118 \\ 0.2118 & 0.2353 & 0.1294 & 0 & 0.2118 & 0.2118 \\ 0.2118 & 0.2353 & 0.1294 & 0 & 0.2118 & 0.2118 \\ 0.2118 & 0.2353 & 0.1294 & 0 & 0.2118 & 0.2118 \\ 0.2118 & 0.2353 & 0.1294 & 0 & 0.2118 & 0.2118 \end{bmatrix} \tag{6}$$

Thus, there is a unique solution to $\pi = \pi[P]$ and $[P^n]$ converges to a rank-1 matrix. This π is any row of $[P^\infty]$, i.e.,

$$\pi = \begin{bmatrix} 0.2118 & 0.2353 & 0.1294 & 0 & 0.2118 & 0.2118 \end{bmatrix}. \tag{7}$$

Furthermore, we see that state 4 is transient, so $\pi_4 = 0$.

2 Convergence of Markov chains: Positive $[P]$ case

Consider the transition matrix

$$[P] = \begin{bmatrix} .2 & .5 & .2 & .1 \\ .1 & .7 & .1 & .1 \\ .5 & .1 & .2 & .2 \\ .3 & .3 & .3 & .1 \end{bmatrix} \tag{8}$$

This is clearly ergodic. The stationary distribution is solved using $\pi = \pi[P]$ and is given by

$$\pi = \begin{bmatrix} 0.2080 & 0.5160 & 0.1600 & 0.1160 \end{bmatrix} \tag{9}$$

In Fig. 1, we plot on the left $\max_l P_{l1}^n$ and $\min_l P_{l1}^n$ against n. We observe that they converge to the same value π_1. The eigenvalues of P are given in the diagonal matrix

$$\Lambda = \begin{bmatrix} 1.0000 & 0 & 0 & 0 \\ 0 & -0.1449 & 0 & 0 \\ 0 & 0 & -0.0000 & 0 \\ 0 & 0 & 0 & 0.3449 \end{bmatrix} \tag{10}$$

Thus the second largest eigenvalue in magnitude is 0.3449. On the right of Fig. 1, we plot $|P_{22}^n - \pi_2|$ and λ_2^n as functions of n. We observe that there is strong agreement in the slope. Note that the y-axis is plotted on a log scale. Since the decrease of both curves is linear, this means that the convergence is geometric as predicted by the theory. The code is provided in Section 6.

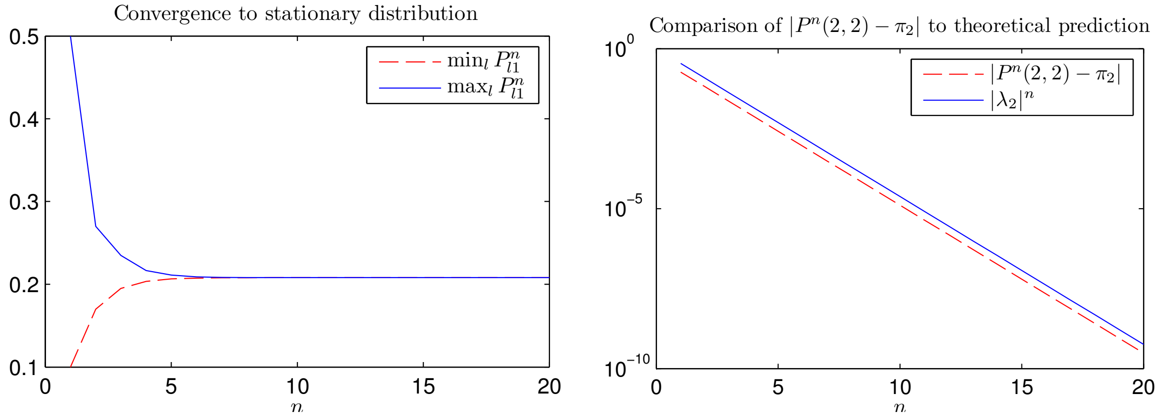

Figure 1: Convergence to stationary distribution

3 Convergence of Markov chains: Ergodic Unichain

Now we consider the transition matrix

$$[P] = \begin{bmatrix} .2 & .5 & .2 & .1 \\ .1 & .7 & .1 & .1 \\ 0 & 0 & .5 & .5 \\ 0 & 0 & .6 & .4 \end{bmatrix} \tag{11}$$

Note that states 1 and 2 are transient. If we start there, we will eventually get absorbed into the recurrent class $\{3, 4\}$. The recurrent class is ergodic and thus this chain is an ergodic unichain with 2 transient states. The stationary distribution is solved using $\pi = \pi P$ and is given by

$$\pi = \begin{bmatrix} 0 & 0 & 0.5455 & 0.4545 \end{bmatrix} \tag{12}$$

On the left plot of Fig. 2, we show the evolution of P_{1j}^n for $j \in \mathcal{S} = \{1, 2, 3, 4\}$. We see that two of the curves go to zero as expected. Next, we compute the eigenvalue matrix

$$\Lambda = \begin{bmatrix} 1.0000 & 0 & 0 & 0 \\ 0 & 0.7854 & 0 & 0 \\ 0 & 0 & 0.1146 & 0 \\ 0 & 0 & 0 & -0.1000 \end{bmatrix} \tag{13}$$

We see that the second largest eigenvalue in magnitude is $\lambda_2 = 0.7854$. We plot $|P_{1j}^n - \pi_j|$ for each $j \in \mathcal{S}$ together with λ_2^n against n in the right plot of Fig. 2. We again see that there is strong agreement in the slopes on a log scale y-axis. The code is provided in Section 6.

4 Markov Chain

A machine can be either working or broken down on a given day. If it is working, it will break down in the next day with probability b, and will continue working with probability $1 - b$. If it breaks down on a given

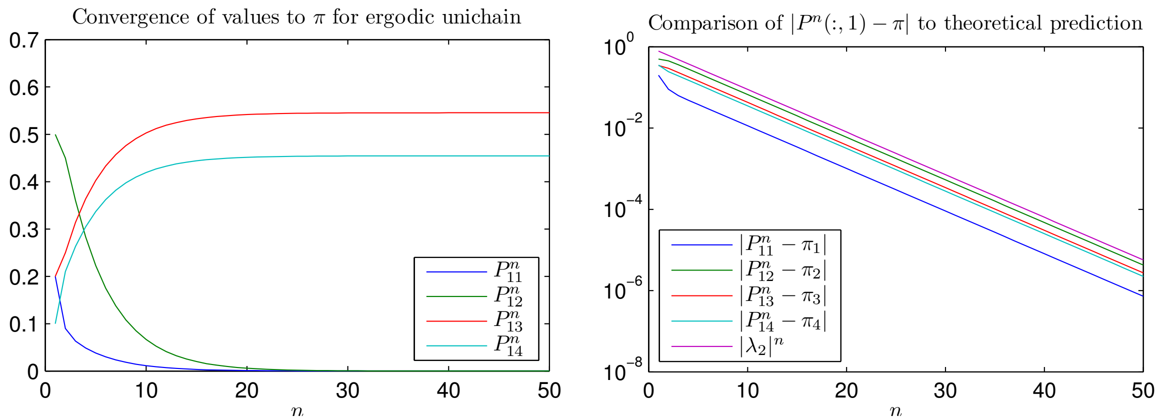

Figure 2: Convergence to stationary distribution for ergodic unichain with transient states

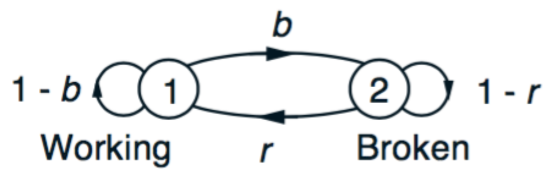

Figure 3: First Markov chain for Section 5

day, it will be repaired and be working in the next day with probability r, and will continue to be broken down with probability $1-r$. Draw the state transition diagram and find the steady-state probability that the machine is working on a given day?

Solution: For the state transition diagram, please see Fig. 5. The transition probability matrix is

$$P = \begin{bmatrix} 1-b & b \\ r & 1-r \end{bmatrix}.$$

This Markov chain has a single recurrent class that is aperiodic (assuming $0 < b < 1$ and $0 < r < 1$), and from the balance equations (i.e., $\pi = \pi P$), we obtain

$$\pi_1 = (1-b)\pi_1 + r\pi_2, \quad \pi_2 = b\pi_1 + (1-r)\pi_2.$$

or

$$b\pi_1 = r\pi_2.$$

This equation together with the normalization equation $\pi_1 + \pi_2 = 1$, yields the steady-state probabilities

$$(\pi_1, \pi_2) = \left(\frac{r}{b+r}, \frac{b}{b+r}\right).$$

Consider a variation of the previous example. If the machine remains broken for a given number of l days, despite the repair efforts, it is replaced by a new working machine. To model this as a Markov chain, we replace the single state 2, corresponding to a broken down machine, with several states that indicate the number of days that the machine is broken. These states are:

State $(2, i)$: The machine has been broken for i days, $i = 1, 2, \ldots, l$.

Draw the state transition diagram and find the steady-state probabilities.

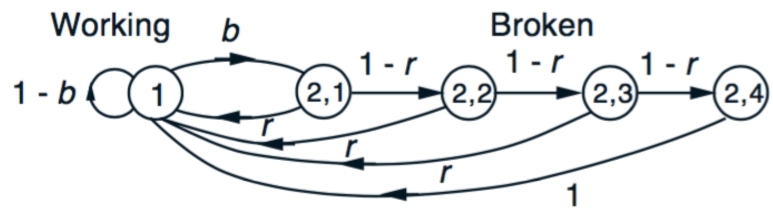

Figure 4: Second Markov chain for Section 5

Solution: For the state transition diagram with $l = 4$, please see Fig. 6. Again this Markov chain has a single recurrent class that is aperiodic. From the balance equations (i.e., $\pi = \pi P$), we have

$$\begin{aligned}
\pi_1 &= (1-b)\pi_1 + r(\pi_{(2,1)} + \ldots + \pi_{(2,l-1)}) + \pi_{(2,l)} \\
\pi_{(2,1)} &= b\pi_1 \\
\pi_{(2,i)} &= (1-r)\pi_{(2,i-1)}, \quad i = 2, \ldots, l
\end{aligned}$$

The last two equations can be used to express $\pi_{(2,i)}$ in terms of π_1 as follows:

$$\pi_{(2,i)} = (1-r)^{i-1} b\pi_1, \quad i = 1, \ldots, l.$$

Substituting into the normalization equation

$$\pi_1 + \sum_{i=1}^{l} \pi_{(2,i)} = 1,$$

we obtain

$$1 = \left(1 + b\sum_{i=1}^{l}(1-r)^{i-1}\right) = \left(1 + \frac{b(1-(1-r)^l)}{r}\right)\pi_1,$$

or

$$\pi_1 = \frac{r}{r + b(1-(1-r)^l)}.$$

Using the equation $\pi_{(2,i)} = (1-r)^{i-1} b\pi_1$, we can also obtain an explicit formula for $\pi_{(2,i)}$.

5 Markov Chain Example

A machine can be either working or broken down on a given day. If it is working, it will break down in the next day with probability b, and will continue working with probability $1-b$. If it breaks down on a given day, it will be repaired and be working in the next day with probability r, and will continue to be broken down with probability $1-r$. Draw the state transition diagram and find the steady-state probability that the machine is working on a given day?

Solution: For the state transition diagram, please see Fig. 5. The transition probability matrix is

$$P = \begin{bmatrix} 1-b & b \\ r & 1-r \end{bmatrix}.$$

This Markov chain has a single recurrent class that is aperiodic (assuming $0 < b < 1$ and $0 < r < 1$), and from the balance equations (i.e., $\pi = \pi P$), we obtain

$$\pi_1 = (1-b)\pi_1 + r\pi_2, \quad \pi_2 = b\pi_1 + (1-r)\pi_2.$$

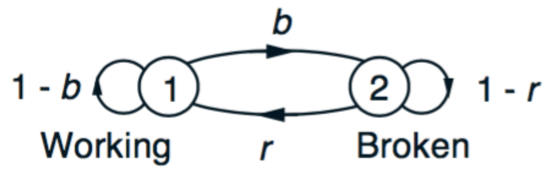

Figure 5: First Markov chain for Section 5

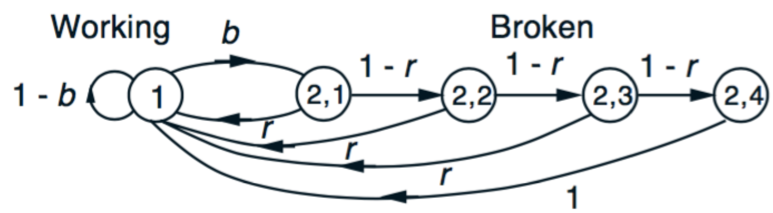

Figure 6: Second Markov chain for Section 5

or

$$b\pi_1 = r\pi_2.$$

This equation together with the normalization equation $\pi_1 + \pi_2 = 1$, yields the steady-state probabilities

$$(\pi_1, \pi_2) = \left(\frac{r}{b+r}, \frac{b}{b+r} \right).$$

Consider a variation of the previous example. If the machine remains broken for a given number of l days, despite the repair efforts, it is replaced by a new working machine. To model this as a Markov chain, we replace the single state 2, corresponding to a broken down machine, with several states that indicate the number of days that the machine is broken. These states are:

State $(2, i)$: The machine has been broken for i days, $i = 1, 2, \ldots, l$.

Draw the state transition diagram and find the steady-state probabilities.

Solution: For the state transition diagram with $l = 4$, please see Fig. 6. Again this Markov chain has a single recurrent class that is aperiodic. From the balance equations (i.e., $\pi = \pi P$), we have

$$\begin{aligned}
\pi_1 &= (1-b)\pi_1 + r(\pi_{(2,1)} + \ldots + \pi_{(2,l-1)}) + \pi_{(2,l)} \\
\pi_{(2,1)} &= b\pi_1 \\
\pi_{(2,i)} &= (1-r)\pi_{(2,i-1)}, \quad i = 2, \ldots, l
\end{aligned}$$

The last two equations can be used to express $\pi_{(2,i)}$ in terms of π_1 as follows:

$$\pi_{(2,i)} = (1-r)^{i-1} b\pi_1, \quad i = 1, \ldots, l.$$

Substituting into the normalization equation

$$\pi_1 + \sum_{i=1}^{l} \pi_{(2,i)} = 1,$$

we obtain

$$1 = \left(1 + b\sum_{i=1}^{l} (1-r)^{i-1} \right) = \left(1 + \frac{b(1-(1-r)^l)}{r} \right) \pi_1,$$

or

$$\pi_1 = \frac{r}{r + b(1 - (1-r)^l)}.$$

Using the equation $\pi_{(2,i)} = (1-r)^{i-1} b\pi_1$, we can also obtain an explicit formula for $\pi_{(2,i)}$.

6 Code

```
% ergodic Markov chain
P = [.2 .5 .2 .1; .1 .7 .1 .1; .5 .1 .2 .2; .3 .3 .3 .1];
% left eigenvalues
[evec,lambdas] = eig(P');
N = 20;
max_value = zeros(N,1); min_value = zeros(N,1);
value = zeros(N,1); diff_value = zeros(N,1);
p_limit = P^1000;
p_limit_theo = (abs(evec(:,1)))/sum(abs(evec(:,1)));

for i = 1:N
   Pi = P^i;
   max_value(i) = max(Pi(:,1));
   min_value(i) = min(Pi(:,1));
   value(i) = Pi(2,2);
   diff_value(i) = abs(value(i) - p_limit(2,2));
end

% ergodic unichain
P2 = [.2 .5 .2 .1; .1 .7 .1 .1; 0 0 .5 .5; 0 0 .6 .4];
p2_limit = P2^1000;
[evec2,lambdas2] = eig(P2');
p_limit_theo2 = (abs(evec2(:,1)))/sum(abs(evec2(:,1)));
N = 50;
value2 = zeros(N,4);
diff_value2 = zeros(N,4);

for i = 1:N
   Pi2 = P2^i;
   value2(i,:) = Pi2(1,:);
   diff_value2(i,:) = abs(value2(i,:) - p_limit_theo2');
end
```